SMGroup

SM Journal of Biomedical Engineering

Research Article

Noise-Invariant Component Analysis for Wearable Sensor based Electrocardiogram Monitoring System

Kemeng Chen[1], Linda S Powers[1,2,3]* and Janet M Roveda[1,3]

[1]*Department of Electrical and Computer Engineering, University of Arizona, USA*
[2]*Department of Biomedical Engineering, University of Arizona, USA*
[3]*BIO5 Institute, University of Arizona, USA*

Article Information

Received date: Apr 27, 2018
Accepted date: May 25, 2018
Published date: May 29, 2018

***Corresponding author**

Powers LS, Department of Electrical and Computer Engineering, The University of Arizona, 1230 E. Speedway Blvd, Tucson, Arizona, 85721, USA,
Tel: 520-621-7634;
Email: lspowers@email.arizona.edu

Keywords Noise invariance; Noise tolerance; Wearable sensors; Electrocardiogram; Automatic diagnosis

Abbreviations DSP: Digital Signal Processing; ECG: Electrocardiogram; EEG: Electroencephalography; EMG: Electromyography; PPG: Photoplethysmogram; QRS: Q: Q-wave, R: R peak, S: S-wave; AI: Artificial intelligence; ADC: analog to digital converter

Abstract

Sensors have been widely used in various data acquisition systems, especially for medical applications. However, once developed for wearable use, these have suffered from various types of noise which greatly degrade data quality and consequently, the reliability. Low data quality is a major obstacle for computer-based diagnosis. Thus, the noise tolerance ability plays a crucial role in wearable sensor based data acquisition and analysis. This work proposes a novel method: noise-invariant component analysis (N-ICA), to expose the influence of noise on this data and provides noise removal and dimensionality reduction. The proposed N-ICA based approach extracts information from data which undergoes minimal change with noise and directly shows the extent to which the true information has been corrupted by noise. This work also implements a simulated wearable sensor based ECG automatic diagnosis system together with a noise generator to validate N-ICA noise tolerant enhancement. Test data is selected from the MIT-BIH Arrhythmia Database. The simulated ECG monitoring system achieves 99.42% accuracy in classifying eight types of heartbeats. Experimental results demonstrate that the signal-to-noise ratio is improved by applying N-ICA based on ECG data contaminated by five noise sources. QRS detection accuracy is also improved to above 95% under the highest noise level tested. Dimensionality reduction reduces the data to 6.5% of the original size. Finally, diagnostic accuracy of four different classifiers is significantly improved when applied in our simulated ECG automatic monitoring system.

Introduction

Wearable sensors have been widely used in a clinical context for disorder detection, treatment efficiency assessment, home rehabilitation, and healthcare research [1]. Wearable sensors such as ECG, EEG and EMG are designed to process bio-potential signals observed at various locations on the human body. The collected data from these sensors can be transmitted to other devices or the cloud for remote home monitoring since modern AI, machine learning, and signal processing techniques enable computer based automatic analysis to assist diagnosis. Wearable sensors are used in uncontrolled conditions and have a simplified hardware design at the expense of signal quality. Since wearable sensors provide more flexibility with fewer physical limits to users, they are often used in homes and for remote monitoring where no medical professionals are involved. Consequently, the activity of the wearer results in more noise than hospital-based measurements. In addition, fewer ADC bits, a lower sampling rate due to transmission speed, power constraints, and physical size of the sensors are other concerns. In addition, muscle movement and sweat can cause noise. Thus, measured data from wearable sensor systems is a combination of various types of noise which are difficult for computer-based automatic diagnosis systems to process accurately.

A number of investigations have examined the sources of noise in most wearable sensors [2-5]. For instance, it has been shown that tissue/body fluid movement during measurement introduced motion artifacts. In addition, electrode-body interface noise (e.g., offset voltages occurring in the electrodes for ECG and EEG), flicker noise and white noise in circuits also exist in wearable sensors measurements [2-5]. De-noising techniques in the sensor hardware design, differential pairs, and common-mode designs have been adopted in the past to reduce the noise [6-8]. However, none of these remove noise completely: they only reduce the amplitude. It is important that none of these approaches can effectively remove motion artifacts which generally introduce large amplitude noise. Several different filtering techniques have been developed to remove or attenuate noise. For example, low pass, high pass, and band pass filters remove components within a particular frequency range. Thus, noise which exists in that frequency range will be filtered. However, useful information may also be removed. Adaptive filters [9] utilize a reference signal which must resemble the desired input to remove the noise via iterative computation that minimizes error using Least Mean Square or Recursive Least Square algorithms. Adaptive filtering needs a good reference signal which may not always be available or with high penalty in a wearable sensor system. For example, to adaptively cancel motion artifacts from ECG or PPG, an extra accelerometer sensor is

usually added to provide the reference signal [8,10]. Savitzky-Golay filtering [11] smooths a signal by approximating it to a polynomial function. Ideally, the polynomial would only fit the signal and not the noise. However, some particular noise is also a smoothed signal and thus may be preserved after polynomial fitting. Newly developed methods, e.g., [12], utilized alexander polynomials and R-L differential equations of fractional calculus to minimize noise. In summary, a method is needed to study and analyze the characteristics of noise in wearable sensors and provide the capability of strong noise tolerance for the wearable sensor data acquisition system. This paper proposed a method to extract the invariant information of feature to differentiate true information from noise. The rest of the paper is organized as follows: section two presents method of noise-invariant component analysis. Section three introduces ECG, noise and the monitoring system architecture used in this work including R peak detection, feature extraction, and heartbeat classification. Section four introduces the experiments design, data preparation, and evaluation metrics for verification. Finally, the conclusion is section five.

Methods

Signals collected from wearable sensors in medical systems are usually corrupted by at least one noise source. These noise sources stem from the subject being measured and/or the data acquisition system and are finally reflected on the acquired signal. To deal with this problem in wearable sensor systems, we propose a method termed N-ICA. The goal is to differentiate between true information and noise by utilizing the invariant feature of true information in the signal. As the true information is generated by a biological mechanisms (in medical applications), randomness is relatively low and the true information collected over time is often self-correlated. For example, electrocardiogram signals generally repeat themselves. As noise and the true information are generated by different mechanisms, they could be uncorrelated under some projections, thus showing different characteristics. Based on the above assumptions, we isolate the uncorrupted portion of the measured information.

We define the noise-corrupted signal collected from wearable sensors as a time series $x=\{x_1..., x_n\}$ where each x_i is digitized data collected at time stamp i. Φ is a function that maps x to another domain (e.g., wavelet transform) and represents x as α coefficients. We also define a function δ which evaluates the variation of α. Finally, N-ICA is defined as a subset of α which exhibits minimal changes with noise. This is based on an assumption that noise damages only a portion of the true information; otherwise, N-ICA would be an empty set. Equation (1) mathematically defines N-ICA where $\alpha=\Phi(x)$ and K is a subset of S:

$$C_{N-IC} = \underset{j\in K}{\alpha_j} = \underset{i\in S}{\arg\min}\left(\delta\left(\alpha_i\right)\right) \quad (1)$$

From equation (1), coefficients in subset K preserve the uncorrupted information as much as possible and we define it as the Noise-Invariant Components. However, when the true information is completely destroyed by noise, there could be very few or no Noise-Invariant Components (K is an empty set). N-ICA directly shows the extent to which true information has been corrupted by noise and aids in separation of corrupted and uncorrupted components. Thus, N-ICA aids in separation of noise affected coefficients from those less affected or not affected at all. As noise usually exhibits certain statistical characteristics, coefficients affected by noise (e.g., Noise-Variant Components) may also have similar properties, assuming Φ is a linear mapping. Thus, these statistical characteristics can be used to enhance noise removal. Assuming Φ has an inverse projection or transform Φ_{inv} such that $x=\Phi_{inv}(\alpha)$, L is a function used to attenuate the effect of noise on noise-variant coefficients and θ is the parameter learned from the noise. The de-noising process can be expressed as

$$\hat{x} = \varphi_{inv}\left(\underset{j\in K}{\alpha_j} \cup \underset{i\in(S-K)}{L(\alpha_i;\theta)}\right) \quad (2)$$

Where $\hat{x}$ is the de-noised data, α_j are noise-invariant coefficients, and α_i belongs to the noise variant coefficients. The de-noised data is obtained by inverse transform of both noise-invariant coefficients and processed noise-variant coefficients.

In terms of classification based automatic diagnosis, information used in a particular pattern for diagnosis is only the portion of the data which makes it an invariant identity even when corrupted by noise. Thus, extracted N-ICs are also good indicators of both classification accuracy and reduced data dimensionality when data is corrupted by noise.

Wearable Sensor Based ECG Monitoring System

ECG measures heart movements through electrical bio-potential changes via electrodes attached to skin. Bio-potentials from electrodes are then amplified, filtered and converted to digital format for further processing. Generally, there are five kinds of major noise sources in an ECG acquisition system: powerline interference, instrument noise, baseline wander (due to respiration), motion artifacts, and muscle contraction noise [13]. Powerline interference is from the electromagnetic field generated mainly by AC power used for artificial light which persistently exists during monitoring. It can be modeled as a combination of sinusoid signals centered around 60 Hz (e.g., U.S.) with random phases. Instrument noise typically from electronic circuits is usually modeled as white noise. Baseline wander mainly comes from respiration which results in roughly slow periodic noise (based on the assumption that ECG electrodes are placed on the chest) that can be modeled as a sinusoid waveform with a random period. Motion artifacts are a consequence of changes in electrode-skin impedance due to electrode motion from human movement such as talking, jumping, running, etc. Since people wear wearable sensors in their daily life, motion noise results from a combination of multiple random motions. Thus, measured motion noise shows non-repeatable patterns and exhibits randomness in both amplitude and shape (pattern). Muscle contraction noise is generated by bio-potential changes caused by muscles other than the cardiac muscle. Since ECG is the reflection of bio-potential changes created by the cardiac muscle, other muscles such as a chest muscle which is even closer to electrode can generate strong bio-potential changes. This kind of noise consists of sharp spikes reflected in the ECG waveform which sometimes resembles the R peaks that can be modeled as spikes at random locations whose amplitude is also randomized within a range. Usually people using wearable ECG monitoring devices are more active, muscle contraction noise is expected to be stronger and more frequent than for in-hospital measurements. This noise greatly degrades the acquired ECG signal quality and negatively affects the

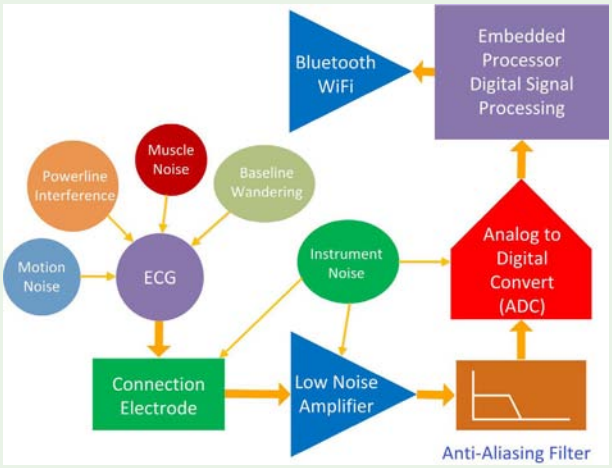

Figure 1: General architecture of a wearable sensor system for ECG monitoring.

performance of subsequent computer automatic processing analysis. Thus, a strong noise tolerant ability is the key for the design of a wearable sensor. Figure 1 shows the general architecture of ECG monitoring system.

According to American Health Association, useful information in ECG mainly occurs from 0.67 to 40 Hz [14]. In particular, P-T waves concentrate at low frequency bands while the QRS complex occupies the high frequency band. Figure 2 displays six pairs of Time-Frequency Spectra calculated by continuous wavelet transform and corresponding ECG measurements corrupted by different kinds of noise. Each pair contains two figures: ECG plot (top) and the Time-Frequency spectrum (bottom). The ECG segment is eleven seconds long and corrupted by noise. The top left pair displays the effect of power line interference which is mainly concentrated around 60 Hz. The top middle pair is the instrument noise. The top right and bottom left pairs are the baseline wander and motion artifacts, respectively. These generally affect the coefficients at lower frequencies. The bottom middle figure is the Time-Frequency spectrum for muscle contraction noise and the bottom right is the mixed noise composed of all kinds of the noise in the figure. As stated in the above section, certain kinds of noise only consistently affect a portion of coefficients which are a subset of all coefficients. Thus, characteristics of the effect of noise in the Time-Frequency domain can be captured and then used to remove noise. Figure 3 displays the de-noising example on an ECG signal with mixed noise. The black line is the original uncorrupted ECG signal and the blue plot represents ECG corrupted by mixed noise. Red plot is the de-noised ECG signal. As observed in Figure 3, noise has been greatly attenuated and critical information of the original ECG signal has been preserved after de-noising. Note that the red plot does not perfectly overlap the black line. The reason is that a subset of coefficients results from noise (Noise-Variant Components) also has overlap with coefficients that represent the true ECG information. Thus, when removing noise by manipulating noise variant components, the true information of ECG may also be altered. Detailed results of de-noising using real ECG data corrupted by five noises with quantitatively evaluation of this method is presented in results and discussion section.

Four step processing flow for automatic ECG processing and analysis

The goal of wearable sensor data acquisition system is not only to collect data with good quality, but also to understand the data, e.g., clustering or classification based preliminary diagnosis. To perform these tasks, we extracted noise-invariant components as the most informative part of data (the unique identification of features) and also reduce data dimensionality. In the ECG data acquisition system, a four-step processing flow is designed to perform automatic heartbeat classification based diagnosis. Input for this flow is the raw

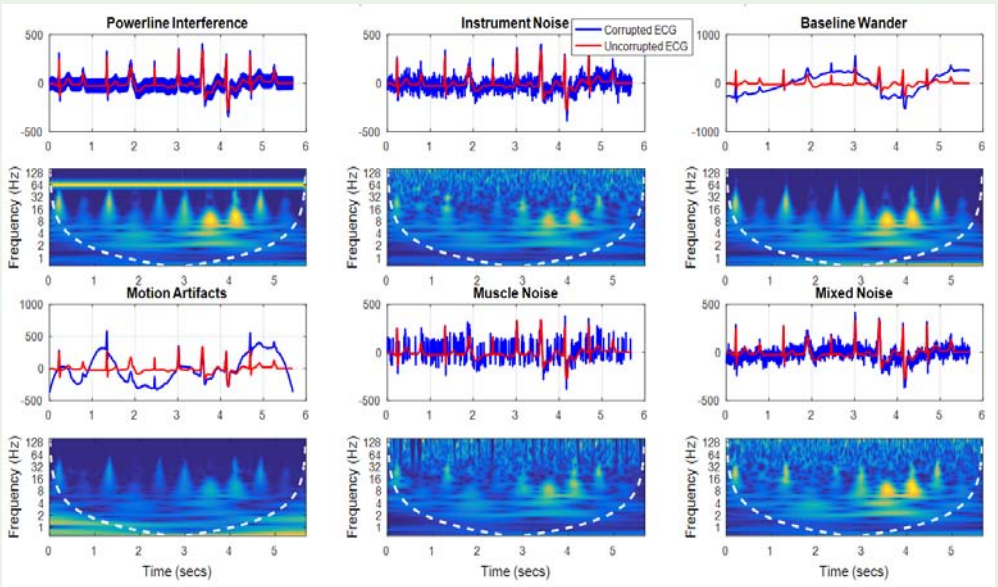

Figure 2: ECG signals corrupted by synthetic noise and the corresponding Time-Frequency spectrum. Below each noise label are the ECG and corresponding Time-Frequency spectrum. In the ECG plot, the vertical axis is the voltage signal amplitude. The frequency (vertical) axis is algorithmic scale as indicated by white dashed lines. The color indicates the amplitude of the time-frequency component with lighter color indicating larger amplitude.

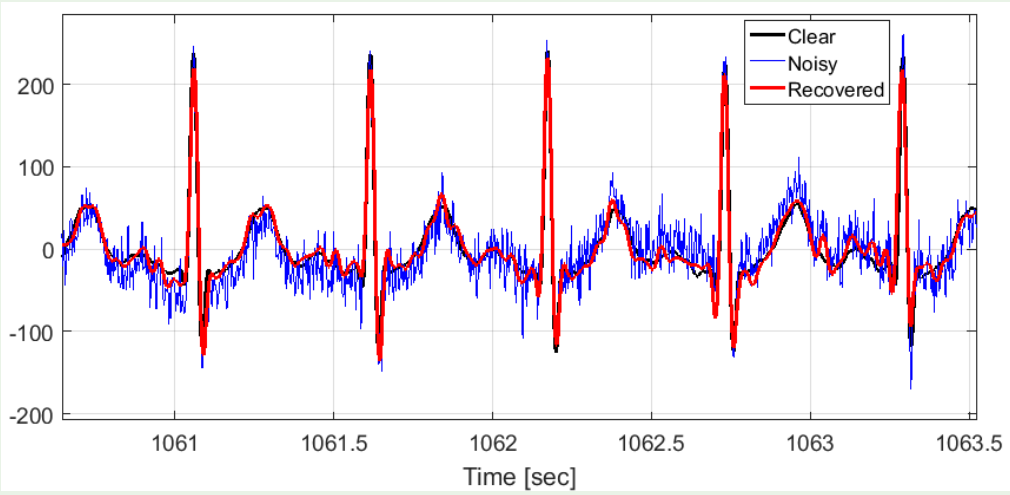

Figure 3: ECG corrupted by mixed noise after de-noising procedure.

ECG signal and the output is the associated heartbeat classification based diagnostic results. The four-step flow comprises de-noising to remove potential noise, detecting R peaks to isolate heartbeat, extracting features of the heartbeat to capture critical information and reduce dimensionality, and finally, classification of different types of heartbeats for automatic diagnosis. Figure 4 displays the block diagram of the proposed flow. In addition, the proposed flow is also flexible to employ new algorithms at each stage to enhance performance.

R peak detection

R peak detection is the first step in calculating other ECG metrics such as heart rate, heart rate variability, etc. which determines the performance of many applications such as stress management [15], emotion tracking [16], etc. Thus, R peak detection accuracy plays a critical role in the performance of the ECG automatic monitoring system. R peaks are detected by the time stamp in the input ECG time series in order to locate each potential heartbeat. While several methods to extract R peaks have been proposed in the literature, our R peak extraction algorithm is based on the classic Pan-Tompkins QRS extraction algorithm [17]. The Pan-Tompkins QRS extraction algorithm demonstrated very high accuracy (99.3% when tested on the MIT-BIH Arrhythmia database [17]). However, when noise is present in the signal, performance of R peak detection can be degraded. Thus, the goal of noise removal is not only to recover the general shape of each heartbeat, but also maintain the R peak information. Detailed experimental results of QRS detection with noise will be discussed in section four.

Feature extraction

The feature extraction step extracts the most important information from the data to provide a more concise form of the original information. The main goals include: (1) reduce dimensionality of raw data; (2) exclude influence from noise; (3) distinguish the difference between different types of heartbeats; and (4) enhance the similarity between heartbeats of the same type. Dimensionality reduction determines the feature vector size for the classification step, which in turn reduces the amount of computation performed by the classifier [18]. The feature extraction step attenuates the effect of noise since the Noise-Invariant Component or Coefficients (N-ICs)

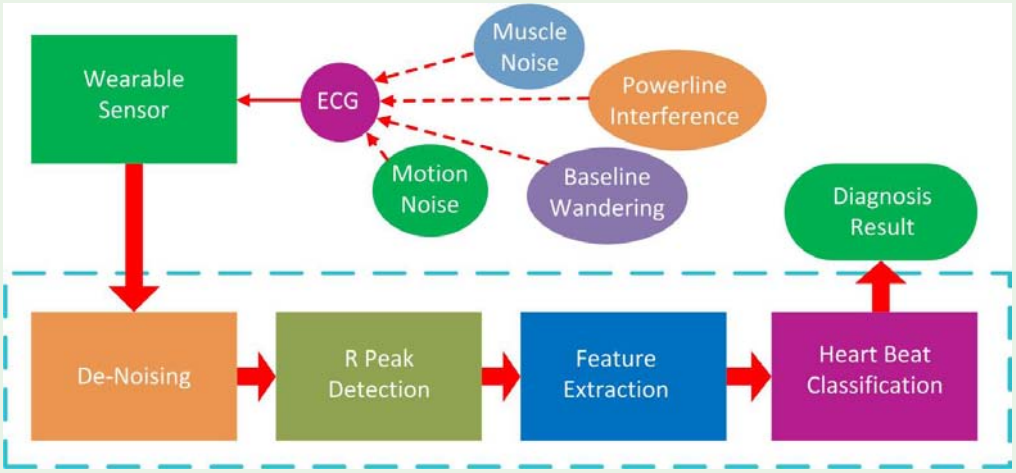

Figure 4: Block diagram of the proposed four-step ECG processing flow.

are used as the extracted feature vectors. Thus, N-ICs preserve the heartbeat information but exclude most of the influence of the noise. Since we have shown that noise is observed in the Time-Frequency domain, we compute N-ICs in the Time-Frequency domain. We also adopt the discrete wavelet transform [19] instead of the continuous wavelet transform to reduce computation. The Discrete Wavelet Transform uses a pair of orthogonal filters to decompose the signal into two sub-bands of different frequencies (approximate signal at low frequency and detailed signal at high frequency). Each sub-band of the signal is half in size of the original signal. By repeating the decomposition onto the low frequency band signal obtained from the previous decomposition, the discrete wavelet transform provides a comprehensive view of the signal from a frequency perspective. Figure 5 (top 2 rows) displays eight types of different heartbeats. These heartbeats are calculated by averaging multiple heartbeats of the same type. Figure 5 (bottom 2 rows) displays the N-ICs as the extracted feature for each type of heartbeat.

Heartbeat classification

After de-noising, R peak detection, and feature extraction step, ECG data is transformed into feature vectors of each beat. A classifier is used as the last step to produce automatic heartbeat diagnosis. Given training data in the format of heartbeats and corresponding diagnosis, the goal of the classifier is to predict the diagnosis of heartbeats through a training process which minimizes the error between correct diagnosis and predicted diagnosis. Our analysis adopts four widely used classifiers to classify each heartbeat (Neural Network, Support Vector Machine (SVM), Naive Bayes Classifier, and Nearest Neighbor). Mathematical details of the four classifiers can be found in supplementary material.

Results and Discussion

To validate performance of the proposed N-ICA with de-noising, R peak detection, and classification based diagnosis, we implemented a simulated wearable sensor based ECG monitoring system with our proposed four step processing flow using MATLAB 2016b. The simulated system includes a noise generator which can generate five types of noise as described in the third section, N-ICA based de-noising block, R peak detection block, feature extraction block, and classification block containing four different classifiers. By using a noise generator, it is possible to quantitatively control noise and preserves ground truth signal for performance evaluation. We also compare the performance with other related work. Ground truth data used in this work are from the MIT-BIH Arrhythmia Database [20-23]. ECG from this database is Holter recordings digitalized at 360 Hz used as ground truth wearable sensor measured ECG. In this work, we selected normal beats together with seven kinds of abnormal beats including atrial premature beats, aberrated atrial premature beats, ventricular escape beats, nodal (junctional) premature beats, nodal (junctional) escape beats, left bundle branch block beats, and right bundle branch block beats from files 209, 201, 207, 234, 222, 109, 118, respectively. This selected testing data contains a total of 8118 heartbeats.

To quantitatively measure the effectiveness of de-noising, we use the Signal to Noise Ratio (SNR in dB) as a numeric metric to compare corrupted ECG and de-noised ECG with ground truth ECG. We also adopted the Percent of Root Mean Square Error (PRMSE) as the metric to quantitatively control noise. The PRMSE is the percent of Root Mean Square Error (RMSE) as defined in equation (3):

$$PRMSE = 100\% \times \frac{\sqrt{\sum_{i=1}^{N}(x_i - \hat{x}_i)^2}}{\sqrt{\sum_{i=1}^{N}(x_i)^2}} \quad (3)$$

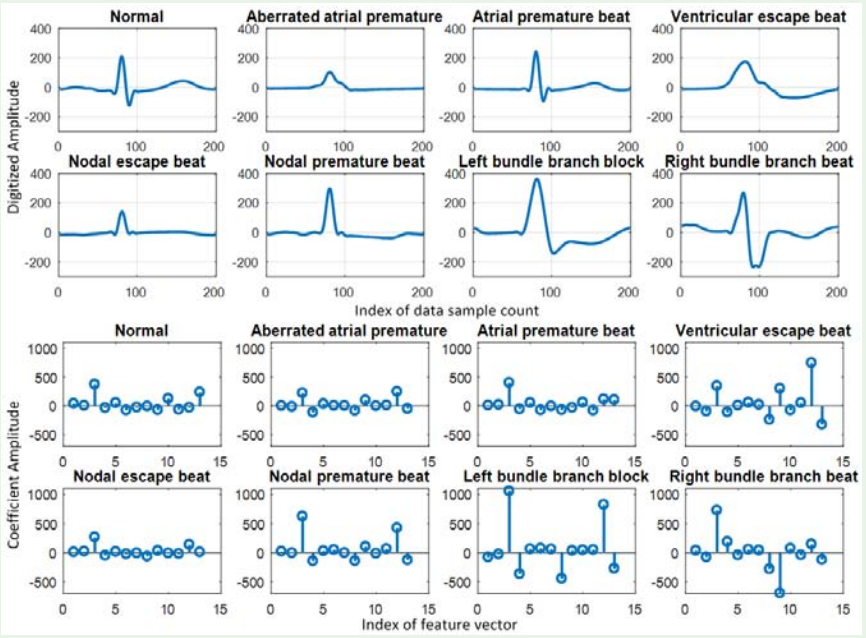

Figure 5: Typical heartbeats (row one and two) calculated by averaging heartbeats of the same type and noise-invariant components based feature extraction (row three and four) where top 15 noise-invariant components are selected.

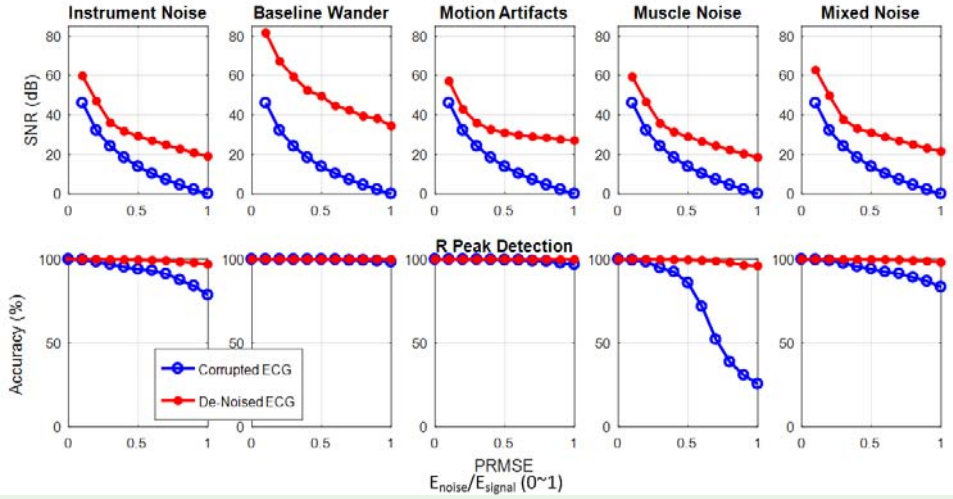

Figure 6: De-noise performance (upper row) and R peak detection accuracy (lower row). The horizontal axis is the PRMSE from zero to one with increments of ten percent. The vertical axis is the SNR in dB. Blue circles indicate the noise-corrupted ECG signals and red dots represent de-noised signals.

The numerator is the root square error and denominator is the square root of the signal energy. Thus, PRMSE can also be interpreted as square root of noise energy over signal energy. PRMSE equals zero means no noise. When PRMSE is one, the energy of the noise is equal to the energy of the signal (SNR is zero (dB)). Figure 6 (top row) displays the result de-noising of four types of noise and their combination. Our results do not include powerline interference since it could be filtered out with a band stop or notch filter at around 60Hz.

Generally, among the four types of noise and their combination, de-noised ECG shows much higher SNR than the noise-corrupted data. In particular, since baseline wander mainly exists in the low frequency range, the frequencies of the rest data are all noise-invariant components which preserve more information compared with other noise. This demonstrates that the de-noising block shows better performance on baseline wander noise than other noise.

On average, the SNR of the instrument noise, motion artifacts, muscle noise, and mixed noise is around 60 dB when PRMSE is 10% (46.05 dB) and decreases to around 35 dB when PRMSE is 30% (24.08 dB) as shown in Table 1. SNR changes within the above range are roughly a linear decrease as shown in Figure 6 (top plot). The SNR decrease rate slows down after PRMSE equals 30% and exhibits approximately a linear decrease from PRMSE equals 30% to 100%. Table 1 shows the numerical results selected at critical points.

Noise has a strong impact on the accuracy of QRS detection which identifies the R peak location of each heartbeat. To evaluate the accuracy of R peak detection, accuracy in this work is defined as one minus the number of false positives and the number of false negatives as shown in equation (4):

$$accuracy = 100\% \times \left(1 - \frac{FP + FN}{N}\right) \quad (4)$$

Where N is the total number of beats in test data. FP is false positive which indicate detecting a wrong location as an R peak and FN is false negative which means that the R peak location is not detected. Figure 6 (bottom) shows R peak detection accuracy on both noise-corrupted ECG data and de-noised ECG data.

Figure 6 bottom row displays the R peak detection results. For each plot, the horizontal axis is PRMSE from zero to one. The vertical axis is the R peak detection accuracy defined using equation (10). The red dot is the accuracy of de-noised ECG data and the blue circle is that for the noise-corrupted ECG data. As observed from Figure 6, different noise has a very different impact on the R peak detection. Instrument noise and muscle contract noise contains many sharp spikes which resemble the R peaks. These are very likely to confuse the R peak detection algorithm and significantly degrade detection accuracy. As shown in Table 2, accuracy decreased to 78.53% for instrument noise and 25.39% for muscle noise when the SNR is zero (dB). Particularly, accuracy tested on muscle noise decreased much faster than that of other noise. In contrast, the baseline wander artifact is generally slow curves that change the amplitude of ECG

Table 1: De-Noise result for each type of noise in (dB).

PRMSE (%)	SNR (dB)	Instrument Noise	Baseline Wander	Motion Artifacts	Muscle Noise	Mixed Noise
10%	46.05	59.69	81.64	57.08	59.24	62.7
30%	24.08	35.96	59.3	35.84	35.65	37.67
60%	10.22	26.94	44.56	29.85	26.52	28.77
100%	0	18.98	34.4	27.05	18.41	21.46

Numerical results of de-noising using five kinds of noise. Left column is PRMSE and SNR (horizontal axis of Figure 6 upper plot). Corresponding SNR values after de-noising for five kinds of noise are listed under noise names in the remaining columns.

Table 2: R peak detection accuracy.

PRMSE (%)	SNR (dB)	Instrument Noise (%)		Baseline Wander (%)		Motion Artifacts (%)		Muscle Noise (%)		Mixed Noise (%)	
0%	N/A	99.8	**99.8**	99.8	**99.8**	99.8	**99.8**	99.8	**99.8**	99.8	**99.8**
50%	13.86	93.93	**99.54**	99.73	**99.78**	99.47	**99.75**	85.65	**99.51**	94.12	**99.63**
100%	0	78.53	**96.86**	98.26	**99.7**	96.82	**99.63**	25.39	**95.98**	83.26	**98.07**

R peak detection performance. First two columns: PRMSE and SNR (horizontal axis of Figure 6 bottom plot). R peak detection accuracy using equation (7) of each kind of noise is listed under corresponding noise name. For each noise name, left column is accuracy on corrupted ECG and right column is on de-noised ECG (**bold**).

baseline which has very small negative impact on R peak detection accuracy. As shown in Table 2, the R peak detection accuracy of these two different kinds of noise remains higher than 95%, even when SNR is zero. In general, de-noising maintains the R peak detection accuracy to above 95% when PRMSE is 100% (SNR equals zero) which demonstrated a very strong noise tolerant ability. Table 2 contains numerical results selected at critical PRMSE values.

In the last step, we verify the performance of heartbeat classification based diagnosis taking R peak detection accuracy into consideration. Since R peak detection is the first step in ECG data analysis, it makes no sense to classify a heartbeat which could not be detected especially when ECG is corrupted by noise. Therefore, performance at this step is the overall performance of the entire system. Since we proposed our own ECG processing flow as a standard testing platform, we first examine the performance of this platform and compare its performance with other related work in terms of classification accuracy. Accuracy in this work is defined as the percentage of beats correctly classified according to annotated files divided by the total number of beats in the data. Table 3 provides a detailed comparison.

Overall, our proposed platform achieves 99.13%, 99.42%, 98.26%, and 93.42% accuracy in classifying eight different types of beats tested on the ECG data containing 8118 beats from MIT-BIH Arrhythmia Database [24-27]. Feature extraction in the proposed platform reduces dimensionality to 6.5% of the raw data size. Thus, the computational complex of classification algorithms would reduce to 4.23% of that without feature extraction assuming $O(n^2)$ complexity of the classifier. Since noise also has a strong effect on classification accuracy, we also studied the influence of different noise and our de-noising strategy on the performance of classification. Four classifiers behave differently in terms of noise corrupted data due to their mathematical approach. This work does not intend to compare the performance of those classifiers. Instead, we focus on the noise tolerant ability and feature extraction and their enhancement to system performance. Figure 7 shows the result of the diagnostic accuracy. Each column represents one type of noise and each row represents one type of classifier. Each plot contains three accuracy curves on corrupted (green), de-noised (blue), and de-noised plus feature extracted (red) ECG data. The horizontal axis is the PRMSE from zero to one indicating the extent the ECG data is corrupted by noise and the vertical axis is the diagnostic accuracy. In general, the accuracy of de-noised ECG and feature extracted ECG perform better than corrupted ECG. Table 4 displays numerical values of accuracy using five types of noise for four classifiers selected from PRMSE values of 0%, 30%, 70% and 100%. For each classifier, three values represent the accuracy using corrupted, de-noised, and extracted features data, respectively [28-30].

Noise generally decreases slowly with classification accuracy as PRMSE increases as shown in Figure 7. Both de-noising and feature extraction improve the accuracy compared with corrupted data, especially when PRMSE is high. This indicates an enhancement of noise tolerance. In particular, when PRMSE is zero, feature extraction makes accuracy slightly lower than without extraction. Since we would like to achieve high dimensionality reduction, noise-invariant components for feature extraction was selected when PRMSE is one. Thus, at a low PRMSE level, part of the information will not be preserved which reduces accuracy. Feature extraction in most conditions slightly lowers the accuracy more than purely de-noising except for the Naive Bayes classifier which favors feature extraction [31-33].

Overall, neural networks and SVM behave similarly with noise tolerance in terms of accuracy. The Naive Bayes classifier and nearest neighbor classifier have weaker noise tolerance compared with other classifiers. As discussed in above section, Naive Bayes classifier assumes that each cluster is independent of all other clusters. However, adding the same kind of noise to all clusters may create dependency between different clusters. Therefore, equation (6) becomes inaccurate which results in low accuracy of the entire classifier. Since feature extraction only preserves coefficients unaffected by noise, features extracted data of each cluster is more independent than raw data. Thus, feature extracted data accuracy is higher than purely de-noised data for the Naive Bayes Classifier. Nearest neighbor classifies data by comparing

Table 3: Comparison of heartbeat classification performance.

Comparable related works	*Top Accuracy*	*#Types*
Proposed flow using Neural Network	**99.13%**	**8**
Proposed flow using SVM	**99.42%**	**8**
Proposed flow using Naïve Bayes Classifier	**98.39%**	**8**
Proposed flow using Nearest Neighbor	**95.57%**	**8**
X. Dong, C. Wang, W. Si	97.78%	5
M. Thomas, M. K. Das, S. Ari	94.64%	5
A. F. Khalaf, M. I. Owis, I. A. Yassine	98.60%	5
S.-N. Yu, Y.-H. Chen	99.65%	6
S. Osowski, T. H. Linh	96.06%	7
I. Guler, E. D. Ubeyli	97.78%	5
M. Engin	98.00%	4
Y. H. Hu, S. Palreddy, W. J. Tompkins	92.20%	4
J. Kim, H. S. Shin, K, Shin, M. Lee	98.72%	6
J. J. Oresko, A. C. Cheng	>90%	5

*Comparison of ECG automatic diagnosis accuracy performance comparison with literature reports.

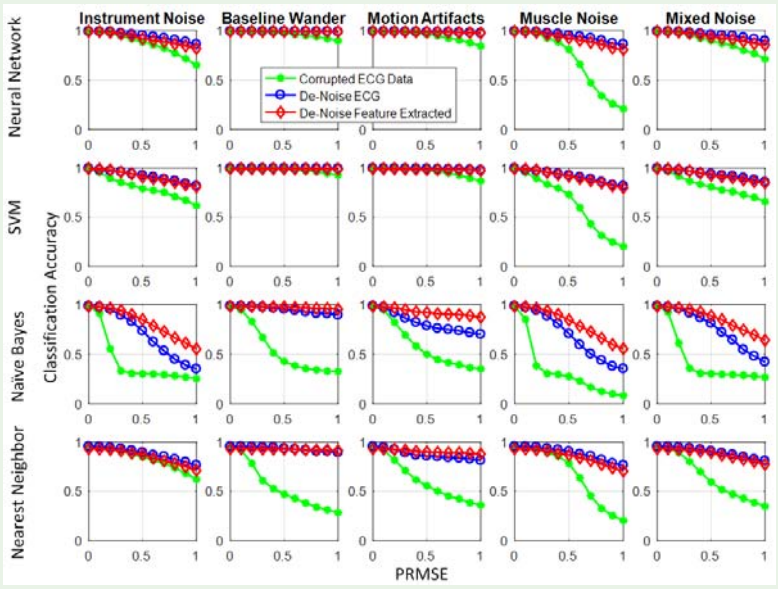

Figure 7: Performance of four classifiers on ECG data corrupted by five kinds of noise with PRMSE from 0 to 1 (10% increment intervals). Each plot represents one classifier accuracy under one type of noise with PRMSE from 0 to 1. Green is corrupted ECG data, blue is de-noised ECG data and red is de-noised plus feature extracted data.

Table 4: Selected heartbeat automatic diagnostic accuracy of four classifiers with different noise at different PRMSEs.

PRMSE	Neural Network			SVM			Naïve Bayes			Nearest Neighbor		
	+	*	-	+	*	-	+	*	-	+	*	-
Instrument Noise												
0%	99.13	99.13	98.93	99.62	99.62	99.23	98.39	98.39	98.07	95.57	95.57	93.24
30%	94.39	97	96.19	85.26	96.03	95.94	33.2	89.26	94.03	90.84	93.31	91.18
70%	82.27	92.08	88.37	75.43	88.68	87.54	29.4	53.43	72.93	78.78	85.29	81.03
100%	64.87	85.58	81.8	61.62	81.73	80.73	25.33	35.02	55.31	66.24	76.46	71.35
Baseline Wander												
30%	98.54	97.95	98.89	99.18	99.58	99.19	66.85	97.81	97.81	60.95	94.8	92.92
70%	94.68	92.87	98.79	97.33	99.58	99.1	35.37	96.69	96.69	38.2	91.92	92.21
100%	89.32	98.64	98.54	92.83	99.36	98.84	32.5	95.3	95.3	28.24	89.28	91.3
Motion Artifacts												
30%	98.32	98.61	98.53	99.1	99.35	98.83	69.18	86.4	94.72	71.16	89.08	91.86
70%	92.07	98.17	98.11	94.75	98.54	98.28	41.57	74.92	90.27	45.15	84.25	89.28
100%	84.11	97.44	97.2	86.66	97.76	97.27	35.22	70.05	87.28	35.97	81.24	87.87
Muscle contraction noise												
30%	92.31	97.06	95.81	83.44	95.8	95.52	30.61	88.42	93.85	89.32	93.55	91.14
70%	47.09	92.1	88	42.95	88.71	87.06	16.77	50.34	72.72	45.37	85.8	81.23
100%	21.1	85.95	80.65	19.99	81.78	79.77	8.19	35.29	55.52	20.21	76.43	70.76
Mixed noise												
30%	95.43	97.54	96.83	86.54	99.62	99.23	98.39	98.39	98.07	95.57	95.57	93.24
70%	84.81	94.38	90.98	75.88	96.81	96.94	35.69	91.56	95.05	80.26	93.83	92.03
100%	71.11	89.38	85.19	65.86	91.68	89.94	29.46	64.36	78.92	46.95	87.25	84.63

Note: +Corrupted ECG, *De-Noised ECG, and -feature extraction in (%). *Selected diagnostic accuracy (%) of +corrupted, *de-noised, and –feature extracted ECG. In particular, when PRMSE is zeros, accuracy of each noise is the same.

the number of samples of each cluster within a sphere centered at the input data. Ideal input data stays in the central area of the cluster and thus, the sphere would include mostly data sampled from the same cluster. The effect of noise on the input data results in a shift of data location. If noise deviation is evenly distributed to each dimension, corrupted input data has a better chance to stay closer to the original location in terms of Euclidean distance. However, when the noise is not evenly distributed to each dimension, the above assumption would not be valid and accuracy decrease quickly as more noise is added to data. Since instrument noise has random amplitude on each dimension, nearest neighbor algorithms show strong noise tolerant ability for instrument noise than other noise types.

Instrument noise shows a roughly linear decrease of accuracy for neural networks, SVM, and nearest neighbor methods. However, Naive Bayes classifier accuracy decreases very quickly from PRMSE equals 10% to 30% (accuracy changes from 98.39% to 33.2%) and accuracy maintains stable after PRMSE is greater than 30% (accuracy changes from 33.2% to 25.33%). In contrast, de-noising and feature extraction lead to a more stable decrease. Baseline wanders and motion artifacts have a small effect on the accuracy of neural network and SVM, but strong impact on Naive Bayes classifier and nearest neighbor. From Table 4, accuracy of neural networks and SVM remained around 90% maximal noise (PRMSE equals one) while Naive Bayer and nearest neighbor accuracy decreased quickly to 32% and 28%, respectively. Proposed de-noising and feature extraction bring accuracy back to around 95% for the Naive Bayes classifier and to 90% for nearest neighbor. Muscle noise created a sharp decrease of accuracy for four classifiers. As the performance metric takes R peak detection accuracy into account, the sharp accuracy decrease may due to the R peak detection accuracy as shown in Figure 6. Both de-noising and feature extraction greatly boost diagnosis accuracy. Mixed noise contains all four kinds of noise, leading to a decrease in accuracy for all four classifiers to 71.11%, 65.86%, 28.68%, and 34.85% respectively. The proposed de-noising and feature extraction improve diagnostic accuracy to above 85% for neural networks and SVM.

From our analytical methods, noise significantly degrades signal quality and creates a big challenge for ECG monitoring and automatic diagnosis systems since it reduces R peak detection accuracy and further decreases diagnosis accuracy. The proposed N-ICA based de-noising improves SNR, thus enhancing system noise tolerance. N-ICA based dimensionality reduces the data to only 6.5% of its original size which significantly reduces computation of the classifier. As SNR improved, R peak detection accuracy together with classification based automatic diagnosis accuracy is improved.

Conclusion

This paper proposes a novel approach to analyze noise and attenuate the negative effects of noise in order to enhance the performance of wearable sensor based data acquisition and automatic analysis systems. The proposed method was applied in data de-noising and dimensionality reduction. To validate its effectiveness, we designed and implemented a four-step processing flow for a simulated ECG monitoring and automatic diagnosis system together with a noise generator. ECG data from the MIT-BIH Arrhythmia database containing 8118 heartbeats including normal beats and seven types of abnormal beats was used as testing ground truth data. Using ground truth data, the ECG monitoring system performance in terms of diagnosis accuracy outperformed most related methods and is able to classify more types of heart diseases. Experimental results using noise corrupted data demonstrated significant improvement of noise tolerance as measured using PRMSE and SNR as well as system performance improvement in both R peak detection and automatic diagnosis accuracy.

Acknowledgement

The authors would like to thank Mr. Ao Li for help convert data to plaintext format.

Supplementary Information

The neural network is a fully connected network. The general architecture of the fully connected neural network consists of an input layer, output layer, and several hidden layers. The number of neurons in the input layers is determined by the size or dimensionality of the input feature vector. The number of neurons in the output layers is determined by the number of heartbeat types. Each neuron in a particular neural layer receives input from all neurons of the previous layer. These inputs are first weighted and summed. Then, an activation function determines the output. Equation (5) mathematically shows the calculation of a single neuron [34]:

$$y_i = f\left(\sum_{j=1}^{n}(w_{i,j}\cdot x_j + b_i)\right) \tag{5}$$

Where $w_{i,j}$ is the weight connecting current neuron to the j^{th} neuron in the previous layer, x_j is the output of the j^{th} neuron in the previous layer, bi is a bias of the current neuron, n is the number of neurons in the previous layer, and $f(\cdot)$ is the activation function. The learning process takes in a set of input data together with their corresponding labels and iteratively computes the weights and bias of the entire network via a back-propagation algorithm [35]. The goal of learning is to minimize the difference between neural network predictive output and the correct prediction result measured using a cost function. After the learning process, the network is able to make predictions given input data. Detailed information about neural networks can be found in [36].

Support Vector Machine (SVM) [37] training process calculates an optimal separating hyperplane [38] which stays between two different clusters of data that maximize the margin between two different groups of data. The classification process then determines to which cluster the current input data belongs using the hyper-plane. A hyper-plane based classification of two groups of data can be mathematically represented as equations (6) and (7) [36].

$$(w\cdot x)-b=0,\ \ |w|=1 \tag{6}$$

$$y_i=\begin{cases}1 & \text{if } (w\cdot x_i)-b\geq\Delta\\ -1 & \text{if } (w\cdot x_i)-b\leq-\Delta\end{cases} \tag{7}$$

Equation (6) is the hyper-plane with parameters w and b. Equation (7) is the classification process with Δ as the margin. To classify multiple classes, 'one' versus the 'rest' strategy is usually adopted.

Assuming there are N clusters, N hyper-plane (namely N classifiers) will be learned, each classifying one cluster versus the remaining N-1 clusters. These classifiers produce a real-value instead of a binary value. The classifier with largest output is the final classification result. More details of SVM one-versus-rest strategy can be found in [39].

Naïve Bayes classifier calculates the posterior probability of each cluster c_i given the current input x represented in equation (8):

$$p(c_i \mid x) = p(c_i) \cdot \frac{p(x \mid c_i)}{p(x)} \propto p(c_i, x) = p(c_i) \cdot \prod_{k=1}^{N} p(x_k \mid c_i) \quad (8)$$

$$y = \arg\max\ p(c_i) \cdot \prod_{k=1}^{N} p(x_k \mid c_i) \quad (9)$$

$p(x)$ is evidence of data which is a constant, $p(c_i)$ is a prior probability of each c_i among all clusters, and $p(x|c_i)$ is the conditional probability (likelihood) of data given cluster c_i. The Naive Bayes classifier assumes that each feature is independent of other features. Thus, the conditional probability can be calculated as the product of each conditional probability of x given c_i as shown in equation (8). Finally, the cluster with the highest probability, $p(c_i|x)$ will be selected as the classification result as shown in equation (9). More details can be found in [40].

Nearest Neighbor is a density estimation based classification algorithm. Assume there are K clusters. For each input data x, there could be a sphere of volume V centered at x. Within this sphere, there are Ni training samples belonging to the i^{th} cluster. The estimated density of each cluster within the sphere centered at x can be represented as a probability $p(c_i|x)$ and the cluster with maximal density will be selected as the result as shown in equation (10).

$$y = \underset{i \in \{1,\ldots K\}}{\arg\max}\ p(c_i \mid x) = \frac{N_i}{\sum_{j=1}^{K} N_j} \quad (10)$$

References

1. Patel S, Park H, Bonato P, Chan L, Rodgers M. A review of wearable sensors and system with application in rehabilitation. J Neuroeng Rehabil. 2012; 9: 21.
2. Ha S, Kim C, Chi YM, Cauwenberghs G. Low-Power Integrated Circuit Design for Wearable Bio potential Sensing. Wearable Sensors: Fundamentals, Implementation and Applications. 2014; 323-352.
3. Yousefi R, Nourani M, Panahi I. Adaptive cancellation of motion artifact in wearable biosensors. Annual International Conference of IEEE, Engineering in medicine and Biology Society (EMBC). 2012.
4. Huigen E, Peper A, Grimbergen C. Investigation into the origin of the noise of surface electrode. Medical & Biological Engineering & Computing. 2002; 40: 332-338.
5. Ha S, Kim C, Chi YM, Akinin A, Maier C, Ueno A, et al. Integrated circuits and electrode interfaces for noninvasive physiological monitoring. IEEE Transaction Biomedical Engineering. 2014; 61: 1522-1537.
6. Asada HH, Shaltis P, Reisner A, Rhe S, Hutchinson RC. Mobile monitoring with wearable photolethymographic Biosensors. IEEE Engineering in Medicine and Biology Magazine. 2003; 22: 28-40.
7. Wood LB, Asada HH. Low variance adaptive filters for cancelling motion artifact in wearable sensor photoplethysmogram sensor signals. Conf Proc IEEE Eng Med Biol Soc. 2007; 652-655.
8. Kim SH, Ryoo DW, Bae C. Adaptive noise cancellation using accelerometers for the PPG signal from forehead. Conf Proc IEEE Eng Med Biol Soc. 2007; 2564-2567.
9. Vijay E, Madisetti K and Williams DB. Digital Signal Processing Handbook. CPC Press. 1997.
10. Liu SH. Motion Artifact Reduction in Electrocardiogram Using Adaptive Filtering. Journal of Medical and Biological Engineering. 2009; 31: 67-72.
11. Savitzky A, Golay JE. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Analytical Chemistry. 1964; 36: 1627-1639.
12. Verma AK, Saini I and Saini BS. Alexander fractional differential window filter for ECG denoising. Australas Phys Eng Sci Med. 2018; 1-21.
13. Friesen GE, Jannett TC, Jadallah MA, Yates SL, Quint SR, Nagle HT. A Comparison of the Noise Sensitivity of Nine QRS Detection Algorithms. IEEE Transactions on Biomedical Engineering. 1990; 37.
14. Michael, Crawford H. ACC/AHA Guidelines for Ambulatory Electrocardiography. Journal of the American College of Cardiology. 1999; 34: 3.
15. Chen K, Fink W, Roveda JM, Lane RD, Allen J, Vanuk J. Wearable sensor based stress management using integrated respiratory and ECG waveform. IEEE 12th International Conference on Wearable and Implementable Body Sensor Networks (BSN). 2015.
16. Roveda JM, Fink W, Chen W, Wu WT. Psychological Health Monitoring for Pilots and Astronauts by Tracking Sleep-Stress-Emotion Changes. IEEE Aerospace Conference. 2016.
17. Tompkins WJ, Pan J. A real-time QRS detection algorithm. IEEE Transactions on Biomedical Engineering. 1985; 32: 230-236.
18. Yan J, Zhang J, Liu J, Yan J, Cheng Q, Fan W, et al. Effective and efficient dimensionality reduction for large-scale and streaming data preprocessing. IEEE Transactions on Knowledge and Data Engineering. 2006; 18: 320-333.
19. Mallat SG. A theory for multiresolution signal decomposition: the wavelet representation. IEEE Transaction on Pattern Analysis and Machine Intelligence. 1989; 11: 674-693.
20. Moody GB, Mark RG. The impact of the MIT-BIH Arrhythmia Databased. IEEE Engineering of Biology Magazine. 2001; 20: 45-50.
21. Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, et al. PhysioNet: Research Resource for Complex Physiological Signals. Circulation. 2000; 101: 215-220.
22. Mark RG, Schluter P, Moody GB, Devlin P, Chernoff D. An annotated ECG database for evaluating arrhythmia detectors. IEEE Transaction Biomedical Engineering. 1982; 29: 600.
23. Moody GB, Mark RG. The MIT-BIH Arrhythmia Database on CD-ROM and software for use with it. Proceedings of Computers in Cardiology. 1990.
24. Dong X, Wang C, Si W. ECG beat classification via deterministic learning. Neurocomputing. 2017; 240: 1-12.
25. Thomas M, Das MK, Ari S. Automatic ECG arrhythmia classification using dual tree complex wavelet based features. AEU-International Journal of Electronics and Communication. 2015; 69: 715-721.
26. Khalaf AF, Owis MI, Yassine IA. A novel technique for cardiac arrhythmia classification using spectral correlation and support vector machines. Expert Systems with Applications. 2015; 42: 8361-8368.
27. Yu SN, Chen YH. Electrocardiogram beat classification based on wavelet transform and probabilistic neural network. Pattern Recognition Letters. 2007; 28: 1142-1152.
28. Osowski S, Linh TH. ECG beat recognition using fuzzy hybrid neural network. IEEE Transactions on Biomedical Engineering. 2001; 48: 1265-1271.

29. Guler I, Ubeyli ED. A modified mixture of experts network structure for ECG beats classification with diverse features. Engineering Applications of Artificial Intelligence. 2005; 18: 845-856.

30. Engin M. ECG beat classification using neuro-fuzzy network. Pattern Recogition Letters. 2004; 25: 1715-1722.

31. Hu YH, Palreddy S, Tompkins WJ. A patient adaptable ECG beat classifier using a mixture of experts approach. IEEE Transactions on Biomedical Engineering. 1997; 44: 891-900.

32. Kim J, Shin HS, Shin K, Lee M. Robust algorithm for arrhythmia classification in ECG using extreme learning machine. BioMedical Engineering OnLine. 2009; 8: 31.

33. Oresko JJ, Cheng AC. A wearable Smartphone-Based Platform for Real-Time Cardiovascular Disease Detection via Electrocardiogram Processing. IEEE Transcations on Information technology in Biomedicine. 2010; 14: 734-740.

34. McCulloch WS, Pitts W. A logical calculus of ideas immanent in nerves activity. The bulletin of mathematical biophysics. 1943; 5: 115-133.

35. Rumerlhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. Nature. 1986; 323: 533-536.

36. Vapnik VN. The Nature of Statistical Learning Theory. Springer-Verlag. New York: Berlin Heidelberg. 2000.

37. Boser BE, Guyon IM, Vapnik VN. A training algorithm for optimal margin classifiers. COLT'92 Proceedings of the fifth annual workshop on computational learning theory. 1992; 144-152.

38. Vapnik VN, Chervonenkis AY, Moskva N. Pattern theory, Journal of Statistical Learning Problem. 1974.

39. Bishop CM. Pattern Recognition and Machine Learning. Springer. 2006.

40. Manning CD, Raghavan P, Schutze H. Introduction to Information Retrieval. NY: Cambridge University Press. 2008.