Research Article

A novel pulse plethysmograph signal analysis method for identification of myocardial infarction, dilated cardiomyopathy, and hypertension

Muhammad Umar KHAN[*] , **Sumair AZIZ**[*]
Department of Electronics Engineering, University of Engineering and Technology Taxila, Taxila, Pakistan

Abstract: Cardiac diseases (CDs) are one of the leading causes of the growing global mortality rate. Early detection of CDs is necessary to avoid a high increase in the mortality rate. Machine learning-based computer-aided diagnosis of CDs using various physiological signals has recently been used by researchers. Since pulse plethysmograph (PuPG) signal contains a wealth of information about cardiac pathologies, therefore, this paper presents an expert system design for the automatic diagnosis of cardiac disorders like hypertension, dilated cardiomyopathy and myocardial infarction using a novel fingertip PuPG signal analysis. The proposed system first performs signal denoising of raw PuPG sensor data using discrete wavelet transform (DWT). After signal segmentation, it extracts discriminant and simplest time-domain features, which are used to perform the detection of normal and abnormal subjects through a support vector machine (SVM) classifier. The proposed detection and classification systems are tested using 10-fold cross-validation which yielded an average accuracy of 98.90%, sensitivity of 100.00%, and specificity of 98.02% for detection (normal vs. abnormal) experiments with only four features and an average accuracy of 97.57% for the multiclass problem using five computationally inexpensive features. Comparative analysis with existing methods based on electrocardiogram, photoplethysmograph, and phonocardiogram revealed that the proposed system has high efficiency in terms of CD detection with very low computational complexity. The findings of this work provide insights into the contribution of PuPG signal analysis towards accurate detection of cardiac disorders through innovative, low cost, and noninvasive methods. Such a system with mobile cardiac health monitoring could be used as a counterpart or second opinion with clinical diagnoses and provide patients with additional but subtle indicators of varying heart dynamics.

Key words: Pulse plethysmograph (PuPG), biomedical signal analysis, support vector machine, feature extraction, discrete wavelet transform, cardiac disorders

1. Introduction

According to the World Health Organization reports, "Cardiovascular diseases (CVDs) are the number one cause of death globally: more people die annually from CVDs than from any other cause. An estimated 17.7 million people died from CVDs in 2015, representing 31% of all global deaths". Cardiac diseases (CDs) are triggered by disorders of the heart and blood vessels, abnormality in blood pressure, coronary heart disease, cerebrovascular disease, rheumatic heart fever, congenital heart disease, and heart failure [1]. It is estimated that 90% of CDs are preventable [2] with the timely usage of indicative tools for diagnosis of CDs such as electrocardiograph, photoplethysmograph (PPG) [3] and sphygmograph [4, 5]. Recently, pulse sensors have gained considerations due to their noninvasive nature and convenience in handling. The most commonly used technologies for pulse measurements are photoplethysmography that monitors the varying blood volume and pulse plethysmography

---

*Correspondence: sumair.aziz@uettaxila.edu.pk

(PuPG) that use piezoelectric material to sense the pressure changes in blood flow during each cardiac cycle.

An expert system for detection and classification of dilated cardiomyopathy, hypertension and myocardial infarction using PuPG signal analysis is presented in this article. The cardiac diseases addressed in this work include: (i) Dilated cardiomyopathy [6] is a disorder affecting heart muscles. In this disorder, the pumping ability of the affected heart is decreased due to the enlargement of the main pumping chamber. This disturbs the time of ventricular relaxation of the heart. Cardiomyopathies are the major reasons for heart failure and on average causes 10,000 deaths in the United States each year [7]. (ii) Hypertension is a significant risk factor for cardiac mortality and it is responsible for events like heart stroke, myocardial infarction, heart failure and chronic kidney disease [8]. Due to high blood pressure, the arteries experience extra stress. Hypertension is considered idiopathic and has no clear culprit for its cause [9, 10]. (iii) Myocardial infarction occurs when the flow of blood to the heart is reduced causing decreased oxygenation to the heart tissues. This is one of the major causes of worldwide deaths. Myocardial Infarction is often considered as a the first sign of coronary artery disease (CAD) [11].

2. Related works

Various techniques including electrocardiogram (ECG), phonocardiogram (PCG), and photoplethymograph (PPG) are currently being used to detect heart pathologies.

Myocardial infarction is detected from ECG waveforms [12] using fractal and entropy based features and K-nearest neighbor (KNN) classifier. The method achieved an accuracy of 99.62% and 99.74% for noisy and clean datasets respectively. ECG signals are also used to classify dilated cardiomyopathy, myocardial infarction, and hypertrophic cardiomyopathy [13] using nonlinear features. The method achieved 99.27% accuracy with K-nearest neighbors using 15 features. Detection of myocardial infarction using artificial neural network is discussed in [14]. ECG data from PhysioNet is optimized using hybrid firefly and particle swarm optimization. The algorithm yielded a 99.3% detection accuracy. Cardiac disorders are detected from multilead ECG using multiscale phase alteration features [15]. The system achieved sensitivities of 0.94, 0.78, and 0.80 and for myocardial infarction, bundle branch block, and heart muscle defect classes, respectively.

Diagnosis of MI from the fusion of ECG, clinical data, and PCG using KNN classification method yields in 99.0% of accuracy, 98.0% of sensitivity, and 100% specificity [16]. ECG signals were used to detect MI using multiscale energy and eigenspace [17]. The system was 96% accurate in diagnosing MI with the SVM classifier. In [18], the authors use 12-lead ECG beats to detect MI. The method consisted of a multilead attention technique with a convolutional neural network and the bidirectional gated recurrent unit framework achieved a mean accuracy of 96.50%.

A relatively new technique, called pulse plethysmograph (PuPG), is used to monitor blood flow or volume of blood during each heartbeat. Plethysmograph is a Greek word that has two parts, 'plethysmo' which means increase, and graph means to write [19]. A method to identify MI from PuPG signal analysis [20] achieved 98.5% accuracy using three features and an SVM classifier. DCM was detected [21] from PuPG time-series with an accuracy of 99.7% using discrete wavelet transform (DWT) and SVM. Diagnosis of acute coronary syndrome (ACS) from PuPG signal analysis by employing empirical mode decomposition (EMD) and SVM is presented in [22]. The system obtained a diagnostic accuracy of 99.42% on 5-fold cross-validation.

A similar sensing technique for measuring heart rate is called photoplethysmograph [23, 24]. The main difference between pulse and photo plethysmograph is that photoplethysmograph uses light to detect the volume

of blood flow while pulse plethysmograph has piezoelectric material that senses the pressure changes in blood flow during each cardiac cycle.

In another study, second derivative photoplethysmography features were employed for the diagnosis of myocardial infarction [25]. The binary logistic regression model for classification achieved a detection accuracy of 90.6%. Hypertension and diabetes were detected with an accuracy of 96.74% using isolation forest, synthetic minority oversampling technique Tomek link, and ensemble learning approach [26].

Fingertip PPG signals are used to classify episodes of intradialytic hypotension. The study used the genetic algorithm and AdaBoost to achieve a mean accuracy of 94.5%. Prediction of hypertension using heart rate variations (HRV) derived from PPG signals is proposed in [27]. The standard deviation of all RR intervals was used as an HRV feature to detect Hypertension in PPG with an accuracy of 85.47%. A system to detect hypertension using six morphological descriptors derived from PPG with 92.31% accuracy is discussed in [28]. Identification of hypertension patients from ballistocardiograms (BCG) is presented in [29]. The system achieved a mean accuracy of 84.4% using class association rules (CAR) classifier and morphological features.

Medical instrumentation in hospitals can easily be impacted by electromagnetic interference (EMI) in a complicated electromagnetic environment [30, 31]. ECG signals are most commonly affected by the EMI and preconditioning circuits, noise removal and shielded cables are required to minimize this interference. PCG requires highly-trained specialists to auscultate the PCG signal [32]. Variations in ambient light and temperature greatly influence the PPG signal. PPG signal acquisition is vulnerable to motion artifacts [33]. The frequency of the PPG signal is about 1–3 Hz [24] so it also requires a high order filter for signal processing.

3. Contribution

The objective of the current study was to develop an automated, reliable and noninvasive system for the diagnosis of hypertension, dilated cardiomyopathy, and myocardial infarction. Main contributions of this work are as follows:

- This is the first study that targeted the classification of hypertension (HYP), dilated cardiomyopathy (DCM) and myocardial infarction (MI) using a novel pulse plethysmograph signal analysis approach.

- The proposed method offers a significant accuracy of 98.90% for detection of cardiac disorders using only four features.

- The proposed approach provides an average accuracy of 97.57% for categorization of HYP, DCM, MI, and normal PuPG signals using only five features.

- The proposed method is feasible for implementing on embedded devices to monitor cardiac health and could be used to complement clinical diagnoses.

This study will lead to developing a real-time and low-cost system to assist physicians in diagnosing heart abnormalities using noninvasive modalities with higher accuracy and efficiency.

4. Materials

4.1. PuPG signals acquisition system

The PuPG signals were recorded using the PTN-104 Plethysmograph sensor attached to the subject's finger. The sensor was interfaced with the NI myDAQ device [34] for acquiring data on the computer using NI LabVIEW.
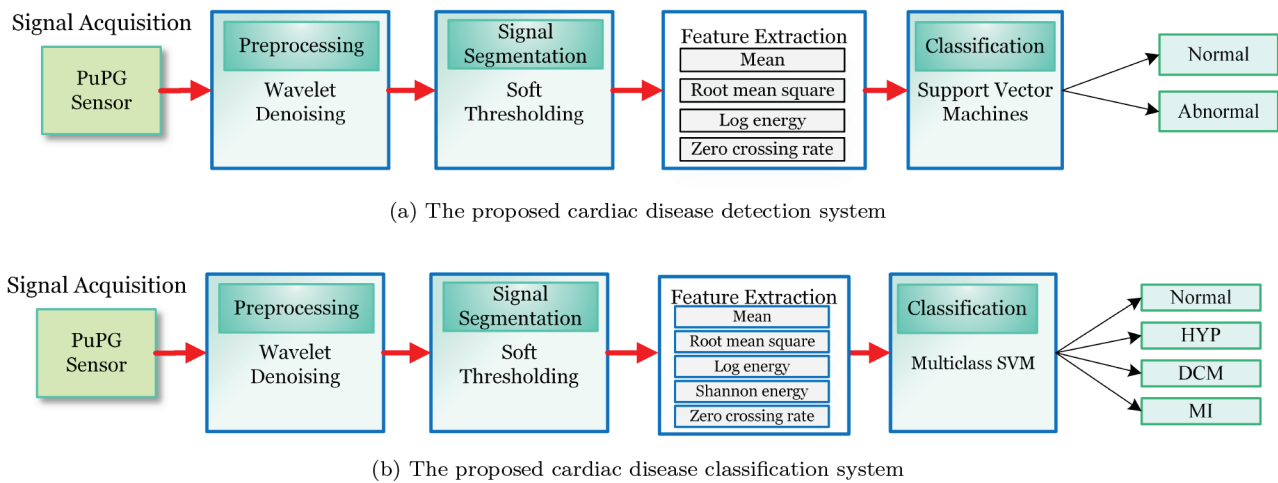
(a) The proposed cardiac disease detection system

(b) The proposed cardiac disease classification system

Figure 1. Block diagrams of the cardiac detection (binary class) and classification (4 class) systems using PuPGsignals

PuPG sensor senses variations in blood flow that generates voltage because of the piezoelectric material. This analog voltage was converted to digital by myDAQ at a sampling frequency of 1KHz.

4.2. Dataset description

In this study, we collected 1827 PuPG signals from healthy and abnormal subjects. Signals were acquired from patients admitted in the cardiology ward, Pakistan Institute of Medical Sciences, Islamabad, Pakistan. The abnormal class consists of a total of 438 signals collected from 13 patients of DCM, 448 PuPG signals gathered from 18 MI patients, and 443 signals recorded from 30 HYP subjects. A total of 498 PuPG signals were also collected from normal/healthy persons who were not having any cardiac abnormality. Each PuPG signal was recorded for 10 s. The total length of each recording was 10,000 samples.

4.3. Ethical approval

Since human subjects were involved in this study, due to ethical approval was taken from the ethics committee. Informed consent was obtained from all individual subjects included in this study.

5. Proposed method

5.1. Overview of proposed methodology

The proposed scheme of work is comprised of four main stages which include, (i) preprocessing of the raw PuPG sensor data, (ii) signal segmentation for extraction of the region of interest, (iii) feature extraction, and (iv) classification. In this study, we proposed separate frameworks for the detection and classification of cardiac disorders through PuPG signal analysis. Figure 1 illustrates the block diagrams of detection and classification systems.

Raw PuPG signals were denoised through wavelet decompositions. Region of interest was extracted from the denoised signal using soft thresholding. Preprocessing and signal segmentation stages are similar for both detection and multiclass system. Feature extraction for the detection system employs four features, namely mean, root mean square, log energy, and zero-crossing rate. Whereas, the multiclass system utilizes a total of
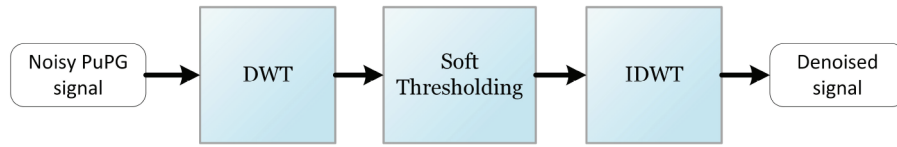
Figure 2. Preprocessing through DWT.

five features such as mean, root mean square, log energy, Shannon energy, and zero-crossing rate. Extracted features for the detection of cardiac abnormality were fed to SVM to distinguish normal and abnormal data classes. While multiclass SVM was employed to categorize HYP, DCM, MI, and normal data classes through extracted five features.

5.2. Preprocessing – wavelet denoising

To eradicate gain dependencies induced by the electronics of the PuPG front-end acquisition module, the acquired signal $p(t)$ was normalized. Figure 2 illustrates the steps of signal denoising through DWT. The noise/power line distortion incurred by the data acquisition module can be seen as a sinusoidal oscillatory signature in the PuPG Hat with a frequency around 50Hz. This distortion must be filtered out of the PuPG signal before the boundary allocation of the region of interest. To avoid the experimental tweaking of filter or filter banks, discrete wavelet transform (DWT) [35, 36] was used to remove this odd signature.

In comparison with methods based on fixed window Fourier transform, which have a dilemma in frequency and time resolution and localization, wavelet-based methods offer high time resolution for high-frequency components and high-frequency resolution for low-frequency components [37]. The DWT splits the signal into a patch of high and low-frequency components. These low frequency and high-frequency signal components are termed as an approximation *cApp* and detailed coefficients *cDet* respectively. The procedure was executed again on this approximation patch to break it up further into the next approximation and high-frequency components and it was iterated up to several levels. DWT follows reconstruction rule and it can also be used to reconstruct noise-free information from the combination of its approximation and detailed components [38].

The Symlet wavelet template 'Sym5' is chosen as a mother wavelet because of its morphological similarities to the PuPG signal. The 4th level approximation coefficients $cApp_4$ obtained after the noisy PuPG $p(t)$ is decomposed using Sym5 filter banks contains the clean PuPG signal. Initially, wavelet decomposition, up to four levels was performed over the noisy PuPG signal $p(t)$. Soft thresholding was performed on the detailed coefficients of each level ($cDet_1, cDet_2, cDet_3, cDet_4$). The denoised signal $v(t)$ was obtained by taking the inverse discrete wavelet transform of the decomposed components [39].

5.3. Signal segmentation – soft thresholding

At this step, the region of interest was extracted from the preprocessed signal $v(t)$ It was observed from extensive experimentation that discriminant information of healthy and abnormal subjects is more concentrated in the upper hat of the PuPG signal. Extraction of the region of interest from a signal reduces the redundant information present in different classes which provides a better feature extraction and classification. Signal

segmentation using soft thresholding was performed using Equation (1).

$$y(t) = \begin{cases} v(t) & v(t) \geq T \\ T & otherwise \end{cases} \tag{1}$$

$$T = min(v(t)) + 0.7 \times \{max(v(t)) - min(v(t))\}, \tag{2}$$

where $v(t)$ was normalized version of the preprocessed signal. T is the threshold parameter defined by Equation (2) and $y(t)$ is a segmented signal. Figures 3a–3c show the raw, preprocessed and segmented signals for normal, hypertension, DCM and MI classes, respectively.

5.4. Feature extraction

Feature extraction is the critical phase in biomedical signal investigation and analysis. This step targets the extraction of significant features from the segmented PuPG signal $y(t)$ that differentiate normal and diseased patients. Features were selected by analyzing the data following the physiology and morphology of the cardiac cycle in the case of normal and pathological PuPG as described earlier.

In this work, we carried out extensive experimentation via a pool of features from time, frequency, and textural domains. But time and statistical features [40, 41] such as Shannon energy (SE), zero-crossing rate (ZCR), mean (M), log energy (LE), standard deviation (SD), and root mean square (RMS) yielded better performance as compared to frequency and textural domain features. Therefore, these features were computed to extract the discriminant information of PuPG signals related to normal and abnormal classes. Further analysis of feature combination is presented in the results section.

5.5. Feature fusion

The final features for both detection (X_d) and 4-class (X_c) systems were constructed through a simple serial concatenation of several features. The features in these vectors were selected after extensive experimentation over a range of classification methods. Equations (3) and (4) represent the final feature vectors for both systems.

$$X_d = [M, RMS, LE, ZCR] \tag{3}$$

$$X_c = [M, SD, SE, LE, ZCR] \tag{4}$$

5.6. Classification – support vector machines (SVM)

After extracting significant features from the PuPG signal in the previous step, cardiac disorders were identified through binary and multiclass-SVM. SVM is one of the most extensively used classifiers for biomedical applications [42–44]. It works by constructing a separating hyper-plane in a feature space that splits the training data into two classes [45–48]. Nonlinear data distributions were separated using kernel functions which transform the original data to a higher dimensional feature space, where features might become linearly separable. The generalization ability of the SVM classifier is the best while dealing with a small sample problem. SVM was originally proposed for binary classification but later on extended for multiclass problems using one-vs-one (OVO) and one-vs-rest (OVR) approaches. Linear SVM is given by the following equation.

$$y = W^T X - b, \tag{5}$$

(a) Raw PuPG signal

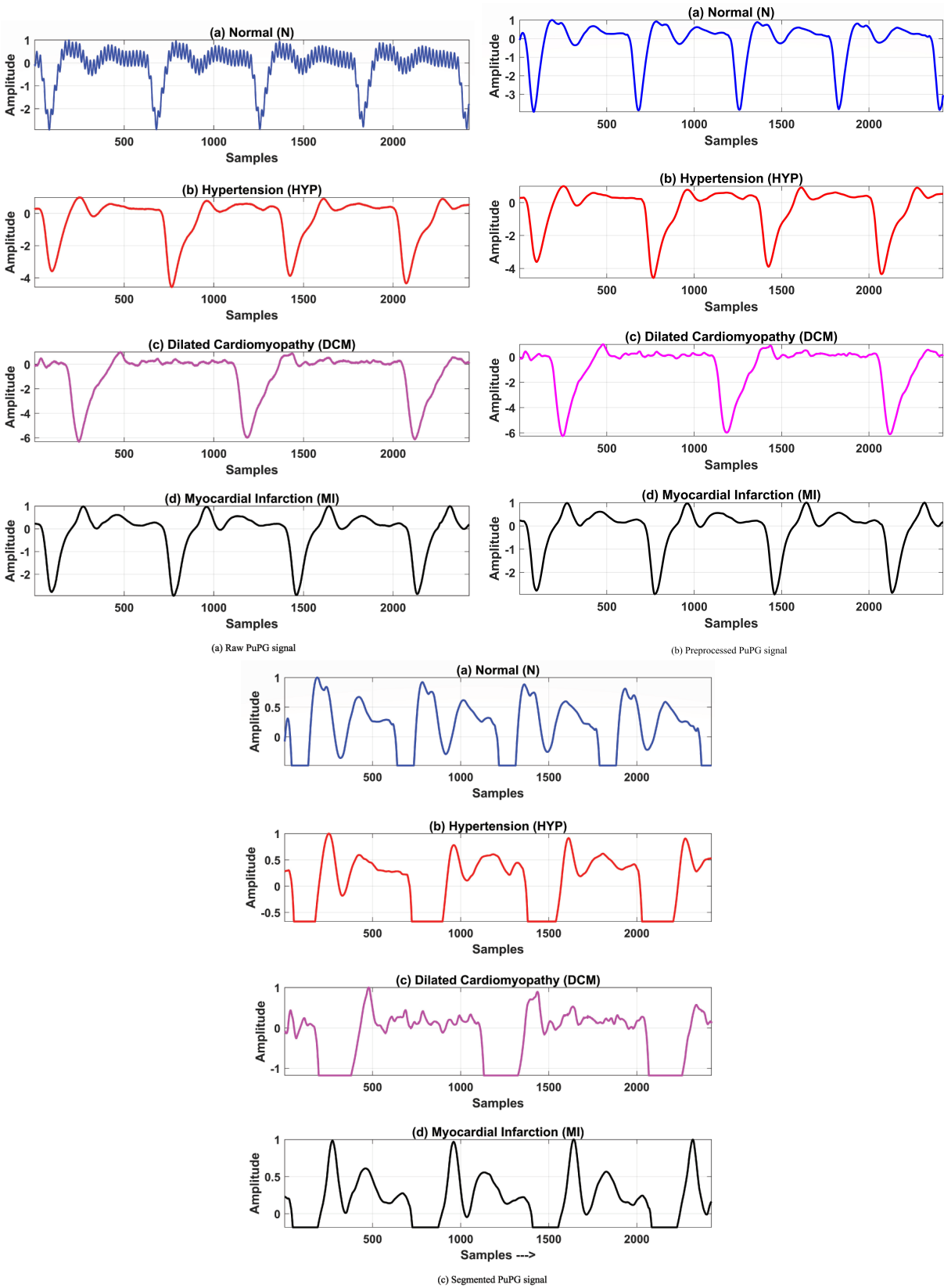(b) Preprocessed PuPG signal

(c) Segmented PuPG signal

Figure 3. Raw, preprocessed and segmented version of PuPG signals for normal, HYP, DCM and MI classes.

where X is the extracted feature vector from preprocessed PuPG signal, W is the weight vector perpendicular to the decision hyper-plane and b represents offset of the hyper-plane. In this work, binary and multiclass SVMs were trained using quadratic, Gaussian and cubic kernels for detection and classification tasks. Two experiments were performed with different datasets and classifier settings. i) Binary class experiments, where features extracted from each PuPG signal were labeled as normal and abnormal and a binary SVM is trained. ii) Multiclass experiments, where training data were labeled as normal and according to the disease type (hypertension, dilated cardiomyopathy, myocardial infarction) and finally multiclass SVM was trained for the four-class problem.

6. Results

Results of the proposed expert system for cardiac disease classification using pulse plethysmography (PuPG) signals are analyzed in this section. After signal preprocessing through DWT, time domain and statistical features were computed to extract the characteristic information contained by different classes of PuPG signal. A pool of features consisting of time, spectral, and textural domains were tested with a range of classification methods. The results with best performing features are shown in this work, for instance, mean (M), standard deviation (SD), skewness (S), kurtosis (K), root mean square (RMS), Shannon energy (SE), log energy (LE), and zero-crossing rate (ZCR). The features were evaluated through comprehensive experimentation over a range of classification methods (SVM, KNN, decision tree).

In this article, all results are stated using the 10-fold cross-validation strategy, and final results were averaged over 100 iterations. The quantities measured here include;

1. True positive (TP): PuPG signal belongs to an abnormal class and is correctly predicted as an abnormal signal by the classifier.

2. False positive (FP): PuPG signal belonged to a normal subject and falsely predicted as an abnormal subject.

3. True negative (TN): PuPG signal belonged to a normal subject and correctly predicted as a normal subject.

4. False negative (FN): PuPG signal belonged to an abnormal subject and falsely predicted as a normal subject.

Based on the values of the abovementioned parameters, we calculated accuracy (Acc), specificity (Sp), sensitivity (Sen), positive predictive value (PPV), negative predictive value (NPV) and error rate (Err). All experiments were performed in MATLAB 2018a on personal computer Core i7 with 16GB of RAM. Experiments were conducted for two different problems; (i) detection of cardiac disorders (normal vs abnormal), (ii) classification of cardiac disorders (normal, hypertension, dilated cardiomyopathy, myocardial infarction).

6.1. Detection of cardiac disorders

In detection experiments, datasets were divided into two classes; normal and abnormal. All PuPG signals belonging to HYP, DCM and MI classes were labeled as "abnormal". Representation of the PuPG signal of different classes through extracted features was validated through comprehensive experimentation. Classification method with different basis functions such as SVM-linear (SVM-L), SVM-quadratic (SVM-Q), SVM-cubic

(SVM-C), SVM-fine Gaussian (SVM-FG), and SVM-medium Gaussian (SVM-MG) was applied on different feature combinations. The results of the best performing feature combinations are summarized in Table 1.

Table 1. Performance comparison of SVM with different feature combinations for detection (binary) experiment. Significant results are highlighted.

Classifier	Features								Performance indicators					
	M	SD	S	K	RMS	SE	LE	ZCR	Acc(%)	Sen(%)	Sp(%)	PPV(%)	NPV(%)	Err(%)
SVM-linear	*	*	*	*	*	*	*	*	91.17	94	88.93	87.04	94.94	8.83
	*		*	*	*	*		*	90.73	93	88.93	86.92	94.14	9.27
	*	*	*	*	*	*		*	90.73	93.5	88.54	86.57	94.51	9.27
	*	*	*	*		*	*	*	90.51	90.5	90.51	88.29	92.34	9.49
	*		*		*	*	*	*	90.51	88	92.49	90.26	90.7	9.49
SVM-quadratic	*				*	*	*	*	96.69	96	97.23	96.48	96.85	3.31
	*		*	*	*	*	*	*	96.47	97	96.05	95.1	97.59	3.53
	*	*		*	*	*	*	*	96.47	95.5	97.23	96.46	96.47	3.53
	*	*	*			*	*	*	96.25	96	96.44	95.52	96.83	3.75
	*	*	*	*		*	*	*	96.25	95.5	96.84	95.98	96.46	3.75
SVM-cubic	*		*		*			*	97.57	97	98.02	97.49	97.64	2.43
	*		*	*	*	*		*	97.13	96.5	97.63	96.98	97.24	2.87
	*	*			*		*	*	97.13	95.5	98.42	97.95	96.51	2.87
	*	*	*					*	97.35	96.5	98.02	97.47	97.25	2.65
	*	*	*			*		*	97.57	96.5	98.42	97.97	97.27	2.43
SVM-medium Gaussian	*		*		*			*	95.81	95	96.44	95.48	96.06	4.19
	*		*		*	*	*	*	95.36	94	96.44	95.43	95.31	4.64
	*	*	*					*	95.14	95	95.26	94.06	96.02	4.86
	*	*	*		*			*	95.36	94.5	96.05	94.97	95.67	4.64
	*	*	*	*				*	95.36	94.5	96.05	94.97	95.67	4.64
SVM-fine Gaussian	*****				*****		*****	*****	**98.9**	**100**	**98.02**	**97.56**	**100**	**1.1**
	*	*			*	*		*	98.68	99	98.42	98.02	99.2	1.32
	*	*	*			*		*	98.68	99.5	98.02	97.55	99.6	1.32
	*				*	*	*	*	98.45	99.5	97.63	97.07	99.6	1.55
	*	*				*		*	98.45	99.5	97.63	97.07	99.6	1.55

Table 2. Comparison of different classifiers for binary class experiments.

Method	Accuracy (%)	Sensitivity(%)	Specificity (%)	PPV(%)	NPV(%)	Error rate(%)
SVM (linear)	91.17	94.00	88.93	87.04	94.94	8.83
SVM (quadratic)	96.69	96.00	97.23	96.48	96.85	3.31
SVM (cubic)	97.13	96.50	97.63	96.98	97.24	2.87
SVM (fine Gaussian)	**98.90**	**100.00**	**98.02**	**97.56**	**100.00**	**1.10**
KNN medium	92.53	90.91	95.24	96.97	86.21	7.47
KNN cubic	92.53	90.91	95.24	96.97	86.21	7.47
KNN weighted	97.51	97.73	97.14	98.29	96.23	2.49
DT	95.73	96.59	94.29	96.59	94.29	4.27

SVM-L classifier achieved a maximum 91.17% accuracy on the combination of all eight features. SVM-Q achieved 96% mean accuracy with all five feature set combinations and the best results (96.69%) for this classifier were achieved on feature combination of root mean square, Shannon energy, log energy, and zero-crossing rate. The average score of 97% was achieved through the SVM-C classifier on five the best feature sets. Overall,
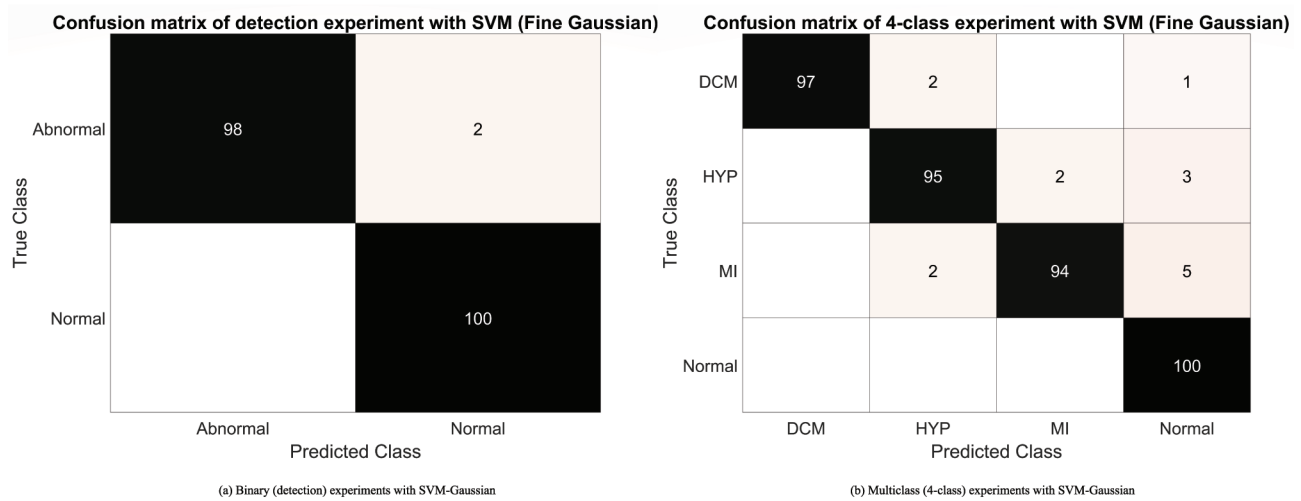
(a) Binary (detection) experiments with SVM-Gaussian

(b) Multiclass (4-class) experiments with SVM-Gaussian

Figure 4. Confusion matrix of binary and multiclass (4-class). All values are presented in percentage.

the best results for the detection of cardiac disorder were achieved through the SVM-FG classifier through a combination of only four features namely mean, root mean square, log energy and zero-crossing rate. Where, SVM-FG yielded 98.90% accuracy, 100% sensitivity and 98.02% specificity (Table 2).

Figure 4a shows class-wise accuracy in the form of a confusion matrix with SVM-FG. It can be seen that all instances of the normal class were correctly predicted while only 2% of instances of the abnormal class were misclassified. To further validate the proposed methodology, we performed classification with a different type of learners i.e. K nearest neighbor (KNN) with K=5 (KNN-medium), KNN with the cubic kernel (KNN-cubic), weighted KNN (KNN-W) and decision tree (DT). It is observed that accuracy and sensitivity performance measures were well above 90% for all classification methods and 98% accuracy were achieved through SVM-FG (Table 2).

6.2. Evaluation of multiple cardiac disorders

To unveil the efficiency of the proposed methodology as a reliable cardiac disorder classification system, we broadened the evaluation scope by performing the validation over the multiclass problems. This included training and testing on features that belonged to normal subjects as well as subjects belonging to multiple cardiac disorders. In these experiments, the whole dataset was divided into normal, hypertension (HYP), dilated cardiomyopathy (DCM) and myocardial infarction (MI) classes. In multiclass problems, the high correlation between the different classes makes the classification task even more challenging.

To unveil the features having the best representation of various classes of PuPG signal, comprehensive experimentation was performed by testing all possible feature combinations through multiclass SVM classifiers with several kernel functions. Performance of SVM-L, SVM-C, SVM-Q, SVM-FG and SVM-MG classifiers in terms of statistical measures are shown in Table 3. The SVM-C achieved a classification accuracy of 96.69% with only four features. Classification performance was further improved by the inclusion of mean feature with SVM-FG classifier. Overall, the best accuracy of 97.57%, sensitivity of 100% and specificity of 97.23% is achieved through SVM-FG with only five features. It was shown that mean and zero-crossing rate features have a critical role in achieving good performance for all classifiers.

Table 3. Performance comparison of SVM with different feature combinations for multiclass (4-class) experiment. Significant results are highlighted.

Classifier	Features								Performance indicators					
	M	SD	S	K	RMS	SE	LE	ZCR	Acc	Sen	Sp	PPV	NPV	Err
SVM-linear	*		*		*	*	*	*	90.51	98.5	91.3	89.95	98.72	9.49
	*		*	*	*	*	*	*	90.07	97.5	91.7	90.28	97.89	9.93
	*	*	*		*	*	*	*	87.86	93.5	90.12	88.21	94.61	12.14
	*		*	*	*	*	*		87.64	95	88.14	86.36	95.71	12.36
	*	*	*	*	*	*	*	*	88.3	92	93.28	91.54	93.65	11.7
SVM-quadratic	*	*	*		*	*		*	95.58	96.5	96.84	96.02	97.22	4.42
	*		*	*	*	*	*	*	95.14	96.5	96.05	95.07	97.2	4.86
	*		*		*	*		*	95.58	97.5	96.05	95.12	97.98	4.42
	*	*		*		*	*	*	94.7	95	95.65	94.53	96.03	5.3
	*		*	*	*	*		*	94.04	95.5	94.47	93.17	96.37	5.96
SVM-cubic	*	*					*	*	96.69	96	99.21	98.97	96.91	3.31
	*		*	*	*		*	*	96.47	97.5	98.81	98.48	98.04	3.53
	*			*	*		*	*	96.25	96	99.21	98.97	96.91	3.75
	*	*				*	*	*	96.69	97.5	97.63	97.01	98.02	3.31
	*				*		*	*	96.69	96.5	98.81	98.47	97.28	3.31
SVM-medium Gaussian	*	*	*		*		*	*	95.58	100	92.49	91.32	100	4.415
	*	*	*		*	*	*	*	95.36	100	92.09	90.91	100	4.636
	*		*		*	*	*	*	94.92	99.5	92.49	91.28	99.574	5.077
	*			*		*	*	*	95.36	99.5	92.49	91.28	99.574	4.636
	*		*		*		*	*	94.48	100	90.12	88.89	100	5.519
SVM-fine Gaussian	*				*	*	*	*	**97.57**	**100**	**97.23**	**96.62**	**100**	**2.428**
	*	*				*	*	*	97.35	100	97.23	96.62	100	2.649
	*	*		*	*	*	*	*	97.57	100	97.23	96.62	100	2.428
	*	*					*		97.35	100	97.23	96.62	100	2.649
	*		*	*	*	*		*	97.35	99.5	97.23	96.6	99.595	2.649

To further investigate the proposed methodology, KNN and DT classification methods were also employed and compared with baseline SVM with various kernels. The reliability of SVM-FG for multiclass problems against other classification methods was also confirmed (Table 4). KNN-W, SVM-C, and SVM-Q provided evidence for the high discriminative characteristics and efficiency of selected features for the multiclass classification task.

Regarding the performance of individual classes, it is observed (Figure 4b) that features extracted from the normal class of PuPG signals contained high discrimination information that resulted in an increased separation from features associated with the patients. This yielded accuracy as high as 100% for a normal class. Abnormal data classes DCM, HYP, and MI showed accuracies by 97%, 95%, and 94%, respectively. Based on these evaluations, it can be concluded that the proposed method in 4-class problems is effective in terms of representing PuPG signals for normal and abnormal subjects.

7. Discussion
Detection of cardiac disorders using noninvasive sensing techniques has recently attracted attention. One to one comparison is not possible since all previous research utilized photo plethysmograph (PPG) signals for the detection of cardiac disorders [23, 25, 49, 50] while this work targeted detection through pulse plethysmograph

Table 4. Performance comparison of different classification methods for multiclass (4-class) experiments.

Method	Accuracy (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Error rate (%)
SVM (linear)	90.51	98.50	91.30	89.95	98.72	9.49
SVM (quadratic)	95.58	96.50	96.84	96.02	97.22	4.42
SVM (cubic)	96.69	96.00	99.21	98.97	96.91	3.31
SVM (fine Gaussian)	**97.57**	**100.00**	**97.23**	**96.62**	**100.00**	**2.43**
KNN medium	88.26	90.34	94.29	96.36	85.34	11.74
KNN cubic	80.43	94.89	61.90	80.68	87.84	19.57
KNN weighted	94.66	98.86	96.19	97.75	98.06	5.34
DT	88.97	93.75	90.48	94.29	89.62	11.03

(PuPG) signals. However, we compared the performance of the proposed methodology with recent studies. The proposed method in this study was compared with the recent state-of-the-art methods (Table 5). Our proposed method utilized PuPG sensor data for cardiac disorder detection and classification has shown to be better in all performance measures. Previously, systems have been designed by using SVM [23, 49] and logistic regression classification [25, 50] where existing literature only focused on two-class problems (normal vs. abnormal). Instead, the current study presented binary and multiclass disease identification in two separate experimental settings.

Table 5. Performance comparison with other studies.

Ref.	Modality	Classes	Method	Sensitivity	Specificity	Accuracy
[17]	ECG	2 (Healthy/MI)	Multiscale energy, eigenspace and SVM	85%	78%	
[25]	PPG	2 (Normal/MI)	Amplitude and interval derived SDPPG features and binary logistic regression	93.90%	87.50%	
[27]	PPG	2 (Normal/Hyp)	HRV features	–	–	85.47%
[28]	PPG	2 (Normal/Hyp)	Morphological features	–	–	92.31%
[29]	BCG	2 (Normal/Hyp)	Morphological features and CAR	–	–	84.4%
[50]	PPG	2 (Normal/MI)	SDPPG features and logistic regression	90.60%	87.50%	
[20]	PuPG	2 (Normal/MI)	Lowpass filter with Lyapunov exponent, mean frequency and spurious free dynamic range features and SVM	100%	95.1%	98.5%
[21]	PuPG	2 (Normal/DCM)	DWT with kurtosis, peak to peak and root mean square features and SVM	100%	99%	99.7%
[22]	PuPG	2 (Normal/ACS)	EMD with statistical and fractal features and SVM	99.43%	99.41%	99.42%
This paper	PuPG	2-class (Normal /Abnormal) and 4-class (Normal, HYP, DCM, MI)	DWT, time and statistical features and multiclass SVM	Binary: 100% Multiclass: 100%	Binary: 98.02% Multiclass: 97.23%	Binary: 98.90% Multiclass: 97.57%

Previously, only three studies [20–22] utilized the same PuPG signatures to identify MI, DCM, and ACS

with 98.5%, 99.7%, and 99.42% accuracies respectively. All of those works were based on a two-class problem. But this study proposed a multiclass approach to detect four classes and on a comparatively larger dataset as compared to that used in [20–22].

Table 6 shows the mean processing times of each step involved in detection and 4-class systems. The detection time of the proposed method was 103 ms, while the total execution time of the multiclass system was 107.29 ms. Timing and performance analysis of the proposed method ensures its feasibility for its real-time implementation on embedded devices.

Table 6. Step-wise execution times for detection and 4-class systems.

Stage-wise execution times (ms)		
Processing stage	Binary class	4-class
Signal normalization time	0.074	0.074
Preprocessing time	14.519	14.519
Segmentation time	0.121	0.121
Feature extraction time	4.925	5.673
Feature fusion time	0.053	0.087
Prediction time	83.316	86.818
Total time	103.008	107.292

8. Conclusion

In this paper, we proposed a novel PuPG signal-based cardiac disease identification system for DCM, HYP, and MI. PuPG signals have embedded information about cardiac health and signal characteristics change significantly with change in the pathological condition of the heart. PuPG signals were first denoised using DWT and afterward, segmentation using thresholding was performed to extract the region of interest. Time domain and statistical features were extracted to represent the PuPG signal of different classes and finally, the SVM classifier was trained to perform effective detection of cardiac diseases. Here, two independent experiments were performed for detection and classification. In the first experiment, the detection of heart abnormalities using only two classes (normal vs. abnormal) was performed. The proposed method achieved 100% sensitivity and 98% specificity which are higher when compared to previously proposed systems. In the second experiment, the dataset was divided among normal, HYP, DCM, and MI classes. This is the first study that addressed the classification of multiple cardiac disorders through only PuPG signals with an accuracy of 97.57% for the multiclass problem.

Literature reveals that there are many pieces of research with various promising technologies using various biosignals to detect multiple cardiac diseases, but they are not mature enough to create an influential and statistically significant contribution towards the validation purpose with clinical diagnoses. However, to ensure reliability, more data will be incorporated in the future for each disease. The efficiency of the proposed method will be examined on large or multiple datasets. Future studies should also focus on the physiological implications of PuPG signatures, such as the relation with gender, age, and other constraints, as well as comorbid clinical circumstances. This study will lead to developing a real-time and low-cost system to assist physicians in diagnosing heart abnormalities using noninvasive modalities with higher accuracy and efficiency.

References

[1] Banerjee R, Bhattacharya S, Bandyopadhyay S, Pal A, Mandana K. Non-invasive detection of coronary artery disease based on clinical information and cardiovascular signals: a two-stage classification approach. In: IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS); Karlstad, Sweden; 2018. pp. 205-210.

[2] McGill Jr HC, McMahan CA, Gidding SS. Preventing heart disease in the 21st century: implications of the Pathobiological Determinants of Atherosclerosis in Youth (PDAY) study. Circulation 2008; 117 (9): 1216-1227.

[3] Rubins U, Grabovskis A, Grube J, Kukulis I. Photoplethysmography analysis of artery properties in patients with cardiovascular diseases. In: 14th Nordic-Baltic Conference on Biomedical Engineering and Medical Physics; Riga, Latvia; 2008. pp. 319-322.

[4] Gunarathne A, Patel JV, Hughes EA, Lip GY. Measurement of stiffness index by digital volume pulse analysis technique: clinical utility in cardiovascular disease risk stratification. American Journal of Hypertension 2008; 21 (8): 866-872.

[5] Bedford DE. The ancient art of feeling the pulse. British Heart Journal 1951; 13 (4): 423.

[6] Caragnano A, Aleksova A, Bulfoni M, Cervellin C, Rolle IG et al. Autophagy and inflammasome activation in dilated cardiomyopathy. Journal of Clinical Medicine 2019; 8 (10): 1519.

[7] Towbin JA, Bowles NE. The failing heart. Nature 2002; 415 (6868): 227-233. doi: 10.1038/415227a

[8] Calvillo L, Gironacci MM, Crotti L, Meroni PL, Parati G. Neuroimmune crosstalk in the pathophysiology of hypertension. Nature Reviews Cardiology 2019. doi: 10.1038/s41569-019-0178-1

[9] Solak Y, Afsar B, Vaziri ND, Aslan G, Yalcin CE et al. Hypertension as an autoimmune and inflammatory disease. Hypertension Research 2016; 39: 567. doi: 10.1038/hr.2016.35

[10] Chan G, Cooper R, Hosanee M, Welykholowa K, Kyriacou PA et al. Multi-site photoplethysmography technology for blood pressure assessment: challenges and recommendations. Journal of Clinical Medicine 2019; 8 (11): 1827.

[11] Thygesen K, Alpert JS, Jaffe AS, Simoons ML, Chaitman BR et al. Third universal definition of myocardial infarction. Nature Reviews Cardiology 2012; 9: 620. doi: 10.1038/nrcardio.2012.122

[12] Sharma M, San Tan R, Acharya UR. A novel automated diagnostic system for classification of myocardial infarction ECG signals using an optimal biorthogonal filter bank. Computers in Biology and Medicine 2018; 102: 341-356.

[13] Adam M, Oh SL, Sudarshan VK, Koh JE, Hagiwara Y et al. Automated characterization of cardiovascular diseases using relative wavelet nonlinear features extracted from ECG signals. Computer Methods and Programs in Biomedicine 2018; 161: 133-143.

[14] Kora P. ECG based myocardial infarction detection using hybrid firefly algorithm. Computer Methods and Programs in Biomedicine 2017; 152: 141-148.

[15] Tripathy R, Dandapat S. Detection of cardiac abnormalities from multilead ECG using multiscale phase alternation features. Journal of Medical Systems 2016; 40 (6): 143.

[16] Zarrabi M, Parsaei H, Boostani R, Zare A, Dorfeshan Z et al. A system for accurately predicting the risk of myocardial infarction using PCG, ECG and clinical features. Biomedical Engineering: Applications, Basis and Communications 2017; 29 (03): 1750023.

[17] Sharma L, Tripathy R, Dandapat S. Multiscale energy and eigenspace approach to detection and localization of myocardial infarction. IEEE Transactions on Biomedical Engineering 2015; 62 (7): 1827-1837.

[18] Fu L, Lu B, Nie B, Peng Z, Liu H et al. Hybrid network with attention mechanism for detection and location of Myocardial Infarction based on 12-lead electrocardiogram signals. Sensors 2020; 20 (4): 1020.

[19] Kamshilin AA, Margaryants NB. Origin of photoplethysmographic waveform at green light. Physics Procedia 2017; 86: 72-80.

[20] Khan MU, Aziz S, Malik A, Imtiaz MA. Detection of myocardial infarction using pulse plethysmograph signals. In: 2019 International Conference on Frontiers of Information Technology (FIT); Islamabad, Pakistan; 2019. pp. 95-955.

[21] Khan MU, Aziz S, Amjad F, Mohsin M. Detection of dilated cardiomyopathy using pulse plethysmographic signal analysis. In: 22nd International Multitopic Conference (INMIC); Islamabad, Pakistan; 2019. pp. 1-5.

[22] Khan MU, Aziz S, Iqtidar K, Zainab A, Saud A. Prediction of acute coronary syndrome using pulse Plethysmograph. In: 4th International Conference on Emerging Trends in Engineering, Sciences and Technology (ICEEST); Karachi, Pakistan; 2019. pp. 1-6.

[23] Paradkar N, Chowdhury SR. Coronary artery disease detection using photoplethysmography. In: 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); Bali, Indonesia; 2017. pp. 100-103.

[24] Ave A, Fauzan H, Adhitya SR, Zakaria H. Early detection of cardiovascular disease with photoplethysmogram (PPG) sensor. In: International Conference on Electrical Engineering and Informatics (ICEEI); Bali, Indonesia; 2015. pp. 676-681.

[25] Mahri N, Gan KB, Meswari R, Jaafar MH, Mohd Ali MA. Utilization of second derivative photoplethysmographic features for myocardial infarction classification. Journal of Medical Engineering & Technology 2017; 41 (4): 298-308.

[26] Fitriyani NL, Syafrudin M, Alfian G, Rhee J. Development of disease prediction model based on ensemble learning approach for diabetes and hypertension. IEEE Access 2019; 7: 144777-144789.

[27] Lan Kc, Raknim P, Kao WF, Huang JH. Toward hypertension prediction based on PPG-derived HRV signals: a feasibility study. Journal of Medical Systems 2018; 42 (6): 103.

[28] Liang Y, Chen Z, Ward R, Elgendi M. Hypertension assessment using photoplethysmography: a risk stratification approach. Journal of Clinical Medicine 2019; 8 (1): 12.

[29] Liu F, Zhou X, Wang Z, Cao J, Wang H et al. Unobtrusive mattress-based identification of hypertension by integrating classification and association rule mining. Sensors 2019; 19 (7): 1489.

[30] Baranchuk A, Kang J, Shaw C, Campbell D, Ribas S et al. Electromagnetic interference of communication devices on ECG machines. Clinical Cardiology: An International Indexed and Peer-Reviewed Journal for Advances in the Treatment of Cardiovascular Disease 2009; 32 (10): 588-592.

[31] Klein A, Djaiani G. Mobile phones in the hospital–past, present and future. Anaesthesia 2003; 58 (4): 353-357.

[32] Sujadevi V, Soman K, Vinayakumar R, Sankar AP. Anomaly detection in phonocardiogram employing deep learning. In: Computational Intelligence in Data Mining; Xiamen, China; 2019. pp. 525-534.

[33] Kranjec J, Beguš S, Geršak G, Drnovšek J. Non-contact heart rate and heart rate variability measurements: a review. Biomedical Signal Processing and Control 2014; 13: 102-112.

[34] Wang D, Khuon L. Using myDAQ and LabView to develop a single-channel EEG for a multi-modality epileptic seizure detection platform. In: 39th Annual Northeast Bioengineering Conference; New York, NY, USA; 2013. pp. 161-162.

[35] Başpınar U, Şenyürek VY, Doğan B, Varol HS. A comparative study of denoising sEMG signals. Turkish Journal of Electrical Engineering & Computer Sciences 2015; 23 (4): 931-944.

[36] Hekim M. The classification of EEG signals using discretization-based entropy and the adaptive neuro-fuzzy inference system. Turkish Journal of Electrical Engineering & Computer Sciences 2016; 24 (1): pp. 285-297.

[37] Bhowmik T, Dey J, Tiwari VN. A novel method for accurate estimation of HRV from smartwatch PPG signals. In: 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); Honolulu, HI, USA; 2017. pp. 109-112.

[38] Arvanaghi R, Daneshvar S, Seyedarabi H, Goshvarpour A. Fusion of ECG and ABP signals based on wavelet transform for cardiac arrhythmias classification. Computer Methods and Programs in Biomedicine 2017. 151: 71-78.

[39] Joseph G, Joseph A, Titus G, Thomas RM, Jose D. Photoplethysmogram (PPG) signal analysis and wavelet de-noising. In: Annual International Conference on Emerging Research Areas: Magnetics, Machines and Drives (AICERA/iCMMD); India; 2014. pp. 1-5.

[40] Ghazali F, Hacine-Gharbi A, Ravier P, Mohamadi T. Extraction and selection of statistical harmonics features for electrical appliances identification using k-NN classifier combined with voting rules method. Turkish Journal of Electrical Engineering & Computer Sciences 2019; 27 (4): 2980-2997. doi: 10.3906/elk-1812-80

[41] Khan MU, Aziz S, Ibraheem S, Butt A, Shahid H. Characterization of term and preterm deliveries using electro-hysterograms signatures. In: IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON); New York, NY, USA; 2019. pp. 899-905.

[42] Kabaoğlu RO. A fault detection, diagnosis, and reconfiguration method via support vector machines. Turkish Journal of Electrical Engineering & Computer Sciences 2015; 23 (2): 589-601. doi: 10.3906/elk-1301-96

[43] Işık Ş, Özkan K, Ergin S. Biometric person authentication framework using polynomial curve fitting-based ECG feature extraction. Turkish Journal of Electrical Engineering & Computer Sciences 2019; 27 (5): 3682-3698.

[44] Aziz S, Khan MU, Alhaisoni M, Akram T, Altaf M. Phonocardiogram Signal Processing for Automatic Diagnosis of Congenital Heart Disorders through Fusion of Temporal and Cepstral Features. Sensors 2020; 20 (13): 3790. doi: 10.3390/s20133790

[45] Weston J, Watkins C. Multi-class support vector machines. Technical Report CSD-TR-980-4. London, UK: Royal Holloway University of London, 1998.

[46] Aziz S, Awais M, Akram T, Khan U, Alhussein M et al. Automatic scene recognition through acoustic classification for behavioral robotics. Electronics 2019; 8 (5): 483. doi: 10.3390/electronics8050483

[47] Khan MU, Aziz S, Zainab A, Tanveer H, Iqtidar K et al. Biometric system using PCG signal analysis: a new method of person identification. In: International Conference on Electrical, Communication, and Computer Engineering (ICECCE); İstanbul, Turkey; 2020. pp. 1-6.

[48] Khan MU, Aziz S, Ch JM, Shahjehan A, Imtiaz A et al. Detection of acute coronary syndrome using electrocardiogram signal analysis. In: International Conference on Electrical, Communication, and Computer Engineering (ICECCE); İstanbul, Turkey; 2020. pp. 1-6.

[49] Banerjee R, Vempada R, Mandana K, Choudhury AD, Pal A. Identifying coronary artery disease from photoplethysmogram. In: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct; USA; 2016. pp. 1084-1088.

[50] Mahri N, Gan KB, Mohd Ali MA, Jaafar MH, Meswari R. Analysis of myocardial infarction signals using optical technique. Journal of Medical Engineering & Technology 2016; 40 (4): 155-161.