Turkish Journal of Electrical Engineering & Computer Sciences

http://journals.tubitak.gov.tr/elektrik/

Research Article

Turk J Elec Eng & Comp Sci
(2020) 28: 875 – 889

doi:10.3906/elk-1905-165

Short unsegmented PCG classification based on ensemble classifier

Sinam Ajitkumar SINGH[1,*], **Swanirbhar MAJUMDER**[2]
[1]Faculty of Electronics & Communication Engineering, National Institute of Technology Sikkim, Sikkim, India
[2]Faculty of Information Technology, Tripura University, Tripura, India

Received: 29.05.2019 • **Accepted/Published Online:** 08.01.2020 • **Final Version:** 28.03.2020

Abstract: Diseases associated with the heart are one of the main reasons of death worldwide. Hence, early examination of the heart is important. For analysis of cardiac disorders, a study of heart sounds is a crucial and beneficial approach. Still, automated classification of heart sounds is a challenging task that mainly depends on segmentation of heart sounds and derivation of features using segmented samples. In the literature available for PCG classification provided by PhysioNet/CinC Challenge 2016, most of the research has focused on enhancing the accuracy of the classification model based on complicated segmentation processes and has failed to improve the sensitivity. In this paper, we present an automated heart sound classification by eliminating the segmentation steps using multidomain features, which results in enhanced sensitivity. The study is based on homomorphic envelogram, mel frequency cepstral coefficient (MFCC), power spectral density (PSD), and multidomain feature extraction. The extracted features are trained using the 5-fold cross-validation method based on an ensemble boosting algorithm over 100 independent iterations. Our proposed design is evaluated using public datasets published in PhysioNet/Computers in Cardiology Challenge 2016. Accuracy of 92.47% with improved sensitivity of 94.08% and specificity of 91.95% is achieved using our model. The output performance proves that our proposed model offers superior performance results.

Key words: Phonocardiogram, mel frequency cepstral coefficient, homomorphic filtering, ensemble classifier, feature extraction, machine learning

1. Introduction

Heart diseases are a primary cause of mortality in the world. Several cardiac anomalies are indicated by heart sound signals, which helps to identify cardiovascular diseases after carefully study of the heart sound signals. Auscultation is the commonly used method to analyze cardiac sounds by using a stethoscope in the clinical field. However, accurate auscultation requires an experienced cardiologist [1] and needs careful observation. As was summarized in [2], the auscultation accuracy when performed by an expert physician is approximately 80%. Hence, a computer-aided diagnosis (CAD) tool for analyzing cardiac signals is required to help in predicting cardiac diseases more accurately.

A cardiac sound recorded by a stethoscope that was generated due to the mechanical action of the heart is known as a phonocardiogram (PCG). PCG signal classification commonly consists of four steps: PCG acquisition and requisite preprocessing, segmentation, derivation of features, and classification using machine learning. The raw PCG signal requires preprocessing since the raw PCG signal is corrupted by noise due to breathing, talking, and environmental noise. PCG segmentation methods are mainly based on electrocardiogram

*Correspondence: ajit_sinam@yahoo.com

(ECG) [3, 4], envelogram energy [5], homomorphic filtering [6, 7], hidden Markov models [8, 9], and wavelet decomposition methods [10]. Extracted features are then fed to machine learning for further classification.

Previous work mostly focused on classifying the PCG signal based on the segmentation approach [11–21]. Despite increasing the classification performance, segmentation has the following disadvantages: 1) Segmentation of heart sounds is based on ECG as the reference required ECG recordings, but the collection of both ECG and heart sound signals is challenging in the case of a newborn baby. 2) The cost of collecting both ECG and PCG recordings is comparatively high. 3) A segmentation method based on envelope detection encounters two shortcomings: due to the unwanted noise and murmur, the first heart sound peaks are ignored along with the detection of a wrong peak, and due to the incorrect hypothesis employed during peak conditioning that the diastole period is longer than the systole period, the assumption fails in the case of cardiac patient and infants. 4) In statistical model-based segmentation, the first heart sound-related characteristics are analyzed using the approach model, but the characteristics of the first heart sound vary from newborn child to elderly people and from a healthy person to abnormal heart patients. Hence, it is troublesome to develop all the first heart sounds in a generalized model using a statistical approach.

Classification of a heart using unsegmented PCG signals was performed in limited papers [22–25]. Hamidi et al. [23] extracted the features from unsegmented PCG based on curve fitting and fractal dimension, which were then further classified using the K-nearest neighbors (KNN) classifier. Philip et al. [24] showed that unsegmented-based PCG classification was 11 times faster compared with segmented-based PCG classification. Deng et al. [22] illustrated that a high classification performance can be accomplished using a feature based on unsegmented heart sounds.

The literature on PCG classification using datasets provided by PhysioNet Challenge 2016 is explained in Table 1. Based on PhysioNet2016 datasets [26, 27], most researchers proposed segmentation algorithms for obtaining characteristics of heart sounds like the first heart sound (S_1), second heart sound (S_2), and correlated systolic and diastolic periods. Kamson et al. [18] segmented heart sounds based on a modified hidden semi-Markov model (HSMM) with an average F_1 score of 98.38%. Tang et al. [19] derived the multimodal features based on the HSMM segmentation method and predicted the abnormality of heart sounds using the SVM classifier. Bradley et al. [11] extracted the features using sparse coding and time domain followed by classification based on the SVM classifier. Messner et al. [16] segmented heart sounds based on a deep recurrent neural network with an average F_1 score of 96%. Wei et al. [21] extracted the feature after segmenting the heart sound using hidden Markov model (HMM) approach and mel frequency cepstral coefficient (MFCC) followed by classification based on a convolutional neural network (CNN). Masun et al. [14] extracted the feature based on time, frequency, and time-frequency domain. Zhang et al. [15] segmented the heart cycle using a combination of the wavelet decomposition method and Shannon energy followed by extraction of the feature based on tensor decomposition and finally predicted the model using the SVM classifier. After selecting useful features, the model was trained using a set of the ensemble classifier. Despite providing many features that may be valuable in predicting the abnormality of the heart sound, the segmentation-based feature extraction algorithm introduced complexity along with an increase in the computational load on the system.

The principal structure of this paper is shown in Figure 1 and it comprises the following three steps: preprocessing, extraction of features based on multidomain characteristics, and classification. In this paper, we focused on classifying heart sounds by elimination of complex segmentation steps using the time domain, frequency domain, entropy, high-order statistics, and cepstrum domain features that result in enhancing the

Table 1. A brief literature review of state-of-the-art methods using PhysioNet2016 PCG datsets.

First author	Preprocessing	Features	Classification	Samples	Results (%)
Tang [19]	-HPF[1] -Spike removal algorithm -Segmentation	Multidomain features	SVM[2]	Train: 2838 PCG Test: 315 PCG	Ac: 88±2 Sen: 88±4 Sp: 87±2
Dominguez [12]	-Segmentation -Sonogram image extraction using NAVIS[3]	CNN[4]	AlexNet	Train: 2345 PCG Valid: 469 PCG Test: 312 PCG	Ac: 97.00 Sen: 93.20 Sp: 95.12
Bradley [11]	-Segmentation -FFT[5]	SCV[11] based on sparse coding and time domain	SVM[2]	Train: 3153 PCG Test: 1277 PCG	Ac: 89.26 Sen: 90.07 Sp: 88.45
Mostafa [13]	-Downsample -Butterworth passband filter -Spike removal -Normalization -Segmentation	Time, time-frequency, and perceptual domain	FDA-ANN[6]	Train: 3153 PCG Test: 1277 PCG	Ac: 82.63 Sen: 76.96 Sp: 88.31
Masun [14]	-Downsample -Segmentation -CFS[5] algorithm	Multidomain	Ensemble	Train: 3153 PCG Test: 1277 PCG	Ac: 80.10 Sen: 79.60 Sp: 80.60
Plesinger [17]	-Segmentation -FFT[5] bandpass filter -Hilbert transformation	Time and statistical domain	Probability assessment	Train: 3153 PCG Test: 1277 PCG	Ac: 85.00 Sen: 89.0 Sp: 81.60
Vykintas [20]			Deep CNN[4] + MFSC[8]	Train: 3153 PCG Test: 1277 PCG	Ac: 84.15 Sen: 80.63 Sp: 87.66
Wei [21]	-Segmentation	MFCC[9]	CNN[4]	Train: 2916 PCG Test: 324 PCG	Ac: 91.50 Sen: 98.33 Sp: 84.67
Philip [24]	-Wavelet entropy	Spectral amplitude and wavelet entropy	Decision tree	Train: 2738 PCG Test: 300 PCG	Ac: 79.00 Sen: 77.00 Sp: 80.00
Hamidi [23]	-Downsample -Butterworth low pass filter -Wavelet -Entropy -Power spectral density	Time-frequency and cepstral frequency	KNN[10]	50% datasets for training, 50% for testing from 5 different datasets (a to e)	An overall accuracy of 81%, 92%, and 98% was achieved for three different datasets

[1]: High pass filter,
[2]: Support vector machine,
[3]: Neuromorphic Auditory VISualizer Tool,
[4]: Convolutional neural network,
[5]: Fast Fourier transform,
[6]: Fisher's Discriminant Analysis-Artificial Neural Network,
[7]: Correlation-based feature selection,
[8]: Mel frequency spectral coefficient,
[9]: Mel frequency cepstral coefficient,
[10]: K-nearest neighbors,
[11]: Spare Coefficient Vector.

high speed of the system. Hence, the derived features are more sensitive in classifying abnormalities of the heart, which further helps in diagnosis in clinical health centers.

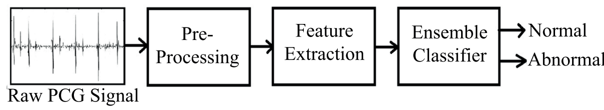


Figure 1. Algorithm of the proposed classification method.

2. Datasets and methods

2.1. Datasets

The PCG datasets used in this study consist of 6 different datasets of heart sound recordings provided by PhysioNet/CinC 2016 Challenge [26, 27]. The heart sound recordings were collected from cardiac patients and healthy subjects in different environments of clinical health centers or home care. The datasets contain a total of 3240 PCG recordings including both a training set and hidden set. The training and testing sets are mutually exclusive to each other to avoid overfitting. The PCG recordings were resampled at 2 kHz and the duration lasted from 5 s to 120 s. The datasets were stored in .wav format, collected from 764 subjects. Out of nine total locations, four different locations were selected, namely the mitral area, tricuspid area, pulmonic area, and aortic area, for collecting the PCG recordings. The PCG recordings were labeled either as –1 for normal or 1 for abnormal. The abnormal PCG recordings were from cardiac patients, typically with coronary artery diseases or aortic stenosis, and normal PCG recordings were from healthy subjects. Out of a total of 3240 PCG recordings, there are 2575 normal PCGs and the remaining 665 recordings are abnormal PCGs.

2.2. Extraction of mel frequency cepstral coefficients

The MFCC estimation using a filter bank is illustrated in Figure 2. The steps involved in the generation of MFCC coefficients are given below:

Algorithm for MFCC analysis.

1: First, the preprocessed PCG signals were divided into frames of 0.025 s with 0.010 s overlapping. With a sampling frequency of 2000 Hz, the frame size A was 50 and the overlapping size B was 20.

2: The amplitude spectrum of each frame was computed by applying a window. The Hamming window was employed to reduce the spectral deformity by decreasing the signal to zero at the start and end of each frame followed by applying DFT to convert the time domain to frequency domain for each frame.

3: Take the log of the above spectra.

4: Convert the above resultant spectrum into mel scale as shown in Equation 1.

5: Apply discrete cosine transform (DCT) to obtain our resultant coefficients.

The mel scale filter banks are estimated as follows:

$$M = 1127log_e(1 + \frac{f}{700}), \tag{1}$$

where M and f denote the resulting mel scale frequency and linear scale frequency, respectively. Using the equation as shown below, we have computed the MFCC from the spectrum of DCT as follows:

$$C_m = \sum_{n=1}^{N} F_n cos(\frac{n - 0.5}{N}\pi m), \tag{2}$$

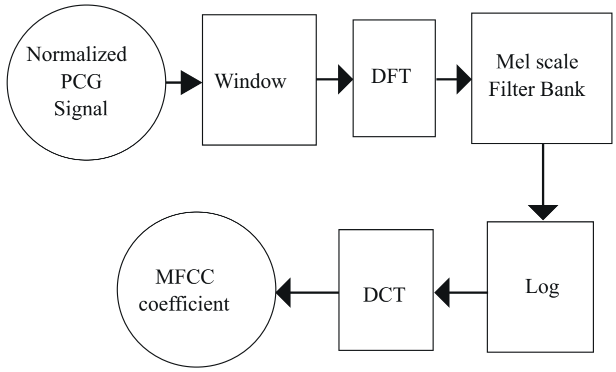


Figure 2. Steps for extracting the coefficient using MFCC approach.

where C_m denotes the mth coefficient of the MFCC, F_n represent the output of the kth filter bank channel, N represents the number of filter bank channels, and n = 1,2,3....N.

2.3. Homomorphic filtering

The PCG spectrum can be observed as slow and fast varying parts. Hence, homomorphic envelograms were extracted using a homomorphic filtering method by eliminating fast varying components. The steps for analyzing the envelope of the heart sound signal are given below.

Envelope detection using homomorphic filtering.

1: Let e(n) denote the energy of the PCG signal and p(n) be the PCG signal; then it can be expressed as:

$$e(n) = s(n)f(n), \tag{3}$$

where s(n) and f(n) denote amplitude (desired component) and oscillating components (unwanted high frequency component), respectively.

2: A logarithmic transformation was applied to convert multiplication to an additive operation:

$$y(n) = log[e(n)]. \tag{4}$$

Thus,

$$y(n) = log[s(n)] + log[f(n)]. \tag{5}$$

3: The unwanted high-frequency component was eliminated by applying a first-order Butterworth low-pass filter L with a cutoff frequency at 8 Hz:

$$y_1(n) = L[y(n)]. \tag{6}$$

We have:

$$y_1(n) = L[logs(n)] + L[logf(n)] \simeq logs(n). \tag{7}$$

Thus, by exponentiation:

$$Envelope(t) = exp[logs(n)] \simeq s(n). \tag{8}$$

The envelope of the PCG signal has been derived using Eq. 8.

2.4. Ensemble classifier

Several researchers used an ensemble learning method as a classifier for enhancing the classification performance of heart sound analysis [14, 28]. The ensemble method enhances the performance of machine learning by combining numerous machine learning methods like boosting, bagging, and stacking into one classification model. There are two groups of ensemble methods: 1) Sequential method: Derives a relationship between the base learners. This helps in enhancing the overall performance of the model by increasing the weight of previously mislabeled weights. 2) Parallel method: It helps in generating the base learners in parallel. Hence, it helps in reducing the error of the model by averaging the resultant value of the base learners.

Boosting enhances the classification performance of a model from the number of limited classifiers. The steps for improving classification performance using the boosting ensemble method are as follows: First, the model was created using the training data. Second, the errors were rectified by creating a second model. Finally, until the classification reached the best prediction performance, models were added. The first algorithm that was developed to boost the binary classification is also known as the AdaBoost algorithm [29]. In this study, we employed the AdaBoost algorithm as follows:

AdaBoost algorithm.

1: Initialize weight of the dataset as:

$$W(p_i, q_i) = \frac{1}{n}, \tag{9}$$

where n is the number of points in a dataset, $p_i \epsilon R^d$, $q_i \epsilon$[-1,1] (–1 represents negative class and 1 represents positive class), and i = 1,2,.....,n.

2: For m=1 to M proceed

i) By decreasing the weighted error function e_m, train a classifier $q_m(x)$:

$$e_m = \sum w_m[1_{q \neq f(x)}]. \tag{10}$$

ii) Compute mth weight of the weak classifier as:

$$\Theta_m = \frac{1}{2} ln(\frac{1 - e_m}{e_m}). \tag{11}$$

iii) Update the dataset weight as:

$$w_{m+1}(p_i, q_i) = \frac{w_m(p_i, q_i) e^{[-\Theta_m q_i f_m(p_i)]}}{N_m}, \tag{12}$$

where N_m is the normalization factor.

3: Using a final model, make a prediction using the equation as shown below:

$$F(x) = sign(\sum_{m=1}^{M} \Theta_m f_m(x)). \tag{13}$$

2.5. Base classifiers

2.5.1. Support vector machine (SVM)

A support vector machine is a machine learning approach generally used for binary classification. By creating hyperplanes in a multidimensional space, the SVM is used to analyze regression or predict the groups. Different types of SVM classifiers (linear or nonlinear) can be constructed based on kernel functions. In this study, we used a linear (radial basis function) kernel-based SVM classifier since nonlinear kernels cause overfitting in the model [11].

2.5.2. Classification tree

A decision tree is a flowchart-like formation in which every interior node indicates attribute testing, each branch indicates the test results, and each leaf node indicates a class. A decision tree comprises both a classification tree and regression tree (CART). A classification tree predicts outcomes by generating rules that can be easily interpreted. It includes splitting the data based on an attribute test value. Using a binary recursive partition, a model of the classification tree has been developed. According to the test attribute value, an algorithm has been assigned to every record. The purpose is to achieve a uniform set of classes. The method proceeds until no more valuable splits can be determined. The previous study demonstrated that the classification tree enhances the classification performance for heart anomaly prediction [24].

2.5.3. K-nearest neighbors (KNN)

The KNN algorithm is a supervised learning algorithm that predicts new outcomes depending on the distance function [30]. The training period requires the feature vector along with respective class labels. K-nearest neighbors have been determined by computing the minimum distance from the test point to the instance training points. Furthermore, by using a majority voting approach, the outcomes were predicted.

2.5.4. Linear discrimination analysis (LDA)

Linear discriminant analysis is an analytical technique traditionally used to predict binary class labels. In 1936, Fisher designed the linear discriminant analysis (LDA) technique to predict binary classification. The main intention of this approach is to optimize the ratio between the interclass alterations and class variations. Hence, this results in improving the maximum discriminant separation of unique classes.

3. Overall system design

PCG signals were preprocessed using a bandpass filter to remove unwanted noises that corrupt the signals. The features were derived from the signal and used as input to the ensemble classifier model. The envelope of the heart sound was analyzed since it helps in defining the event of interest. Hence, the time-domain features based on an envelope were derived. The spectral power of the PCG signal was analyzed using fast Fourier transform (FFT), which results in the extraction of frequency domain features. We extract cepstrum domain features based on the mel frequency cepstral coefficient (MFCC). The extracted features were fed as input to the model based on an ensemble classifier along with the class label for training the model.

3.1. Preprocessing

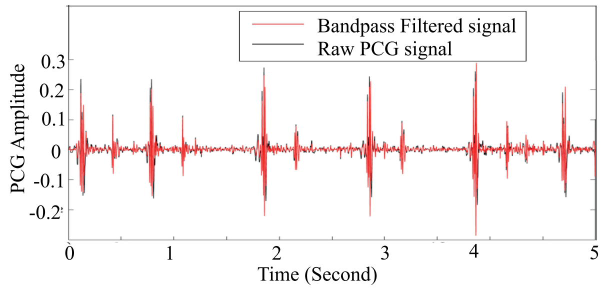
Before preprocessing steps, the recorded PCG signals were segmented using a window of 5 s in length to generate 13015 samples. Out of these 13015 samples, 3158 samples consist of abnormal heart sounds and the remaining 9857 samples are normal heart sounds. During the preprocessing step, a 4th order Butterworth passband filter was applied to remove unwanted noise and murmurs with a cut-off frequency at 25 Hz and 400 Hz. To analyze the increasing and decreasing amplitude of the raw PCG recording uniformly, the resultant signals are normalized using Equation 14 as shown below:

$$Normalized(signal) = \frac{signal - mean(signal)}{std(signal)}, \tag{14}$$

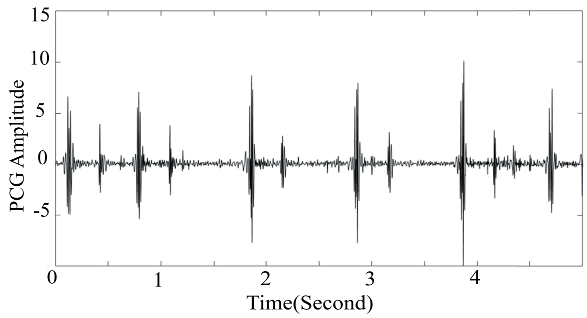
where "signal" represents the filtered PCG signal. The preprocessing of raw PCG signal is shown in Figure 3.

3.2. Feature extraction

Several researchers have been trying to choose efficient features since the classification performance of models varies directly with the selection of suitable features. Twenty-seven features were extracted from the prepro-



(a)



(b)

Figure 3. Preprocessing steps of PCG signal (PCG signal with annotation a0001 from 1 to 10000 samples, i.e. 5 s interval): (a) preprocessed PCG signal after applying bandpass filter, (b) normalization of filtered PCG signal.

cessed signal. They were calculated as time-domain, frequency-domain, cepstrum-domain, high-order statistics, and entropy features. Until recently only limited research has been employed for prediction of abnormality classification using features based on unsegmented PCG signals. To the authors' knowledge, feature extraction based on the combination of MFCC coefficients and homomorphic envelograms using unsegmented heart sounds is new for PCG classification. Besides, some of the recent studies used multidomain features, which were mostly based on segmented heart sounds. A summary of the proposed features is given in Table 2.

3.2.1. Time-domain features (7 features)

The envelope of the heart sound signal has been extracted using the homomorphic filtering approach. The envelope of the PCG recording helps in predicting the cardiac intervals as well as other relevant features. Thus, in general, it results in classifying the abnormality of the heart. The envelope detection procedure applied to the PCG signal using the homomorphic filtering approach is shown in Figure 4. Several researchers employed a homomorphic envelogram approach for extracting the amplitude envelopes of the heart sounds [7, 9]. After filtering the raw PCG signals, the envelope of the filtered PCG signals was analyzed using the homomorphic filtering method as stated above in Section 2.3 ("Envelope detection using homomorphic filtering). Seven time-domain features were extracted from the resultant envelogram of the PCG signal.

3.2.2. Frequency-domain features (3 features)

There are many approaches to achieve features based on the frequency domain. Three frequency-domain features were extracted after computing the power spectral density of the filtered PCG signals using Fourier analysis with a cutoff frequency of 256 Hz as employed by Arnott et al. [31].

Table 2. A brief description of the proposed features.

Domain	Total features	Feature names	Physical meaning	Motive
Time	7	mean_envelope	Mean value of the PCG envelope	To reflect the envelope of the signal.
		median_envelope	Median value of the PCG envelope	
		std_envelope	Standard deviation of the PCG envelope	
		mad_envelope	Mean absolute deviation of the PCG envelope	
		Q25_envelope	25th percentile value of the PCG envelope	
		Q75_envelope	75th percentile value of the PCG envelope	
		IQR_envelope	Inner quartile range value of the PCG envelope	
High-order statistics	2	skewness	Skewness of the PCG signal	To analyze the kurtosis and skewness of each PCG signal.
		kurtosis	Kurtosis of the PCG signal	
Entropy	2	sample_entropy	Entropy of the PCG signal	To compute the complexity of the PCG signal.
		spectral_entropy	Entropy of the PSD of the PCG signal	
Frequency	3	DFV	DFV of the PSD of the PCG signal	To reflect dominant frequency of the signal.
		DFM	DFM of the PSD of the PCG signal	
		DFR	Ratio of dominant frequency	
Cepstral	13	MFCC	MFCC coefficients	To reflect the acoustic property.

DFV: Dominant frequency value, DFM: dominant frequency magnitude, DFR: dominant frequency. ratio

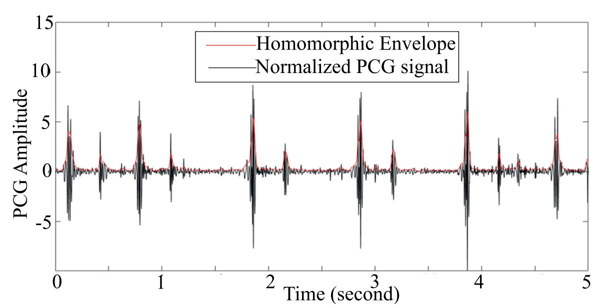


Figure 4. Envelope detection using homomorphic filtering method.

3.2.3. Cepstrum-domain features (13 features)

In most of the prediction models, signal processing functions operate after transforming the raw data into some parametric information. This information is then further analyzed and processed to convert some features. In this study, the collection of the raw PCG recordings is followed by preprocessing steps. The MFCC coefficients have been estimated to extract cepstrum domain features since MFCC denotes spectral information of a sound. The MFCC approach has been practiced by several researchers as the feature extraction approach to enhance classification accuracy [13, 19, 21]. Several researchers used the MFCC approach for extracting cepstrum domain features from heart sound signals since it is essential to analyze the acoustic properties that lead to enhancing the classification performance. Thirteen cepstral domain features were extracted from the filtered PCG signal by sliding over a 25 ms window with a step size of 10 ms using the coefficient based on the MFCC approach as explained in Section 2.2 ("Algorithm for MFCC analysis").

3.2.4. High-order statistic features (2 features)

In statistics, skewness is the measure of asymmetry from the normal distribution. It measures the lack of symmetry, defined using a positive and negative value. Hence, one feature has been calculated from the filtered PCG signal using the coefficient of skewness. Kurtosis is the measure of the degree of peakedness in the variable distribution curve. Using a coefficient of kurtosis, one feature has been calculated from the filtered PCG signal.

3.2.5. Entropy features (2 features)

Entropy and spectrum entropy of the PCG samples measure the complexity of a sequence. Spectrum entropy has been computed after analyzing the power spectral density using Fourier analysis. A detailed study for computing entropy can be observed in [32].

4. Classification and performance

In this study, we adopted an ensemble classifier based on the boosting method for predicting the abnormality of heart sounds. Based on the extracted features as explained in Section 3.2, we trained our proposed model using a 5-fold cross-validation method after carefully avoiding the overfitting of the training model. The algorithm for the ensemble classifier based on AdaBoost was explained in Section 2.4 ("AdaBoost algorithm"). To eliminate overfitting, the training samples and testing samples are mutually exclusive.

We compute two performance parameters, i.e. sensitivity (SN) and specificity (SP), for measuring the performance of our proposed ensemble classifier with other traditional classifier methods. Sensitivity measures the percentage of abnormal heart sounds correctly diagnosed in an abnormal class, whereas specificity measures the percentage of normal heart sounds correctly diagnosed in a healthy class:

$$Sensitivity = \frac{Tp}{Tp + Fn} 100\%, \tag{15}$$

$$Specificity = \frac{Tn}{Tn + Fp} 100\%, \tag{16}$$

$$Accuracy = \frac{Tp + Tn}{Tp + Tn + Fp + Fn} 100\%, \tag{17}$$

where Tp, Tn, Fp, and Fn denote true positive, true negative, false positive, and false negative, respectively.

5. Results

We randomly select 10% of the PCG recordings (10% abnormal recordings and 10% normal recordings) to train the ensemble model, and the remaining 90% of the PCG recordings (90% abnormal recordings and 90% normal recordings) are used to test the performance of the model. The training and testing data are mutually exclusive to each other. This study repeats 100 times to compute the stability. We evaluate the performance of the model by increasing 10% data for training and decreasing 10% data for testing until the training percentage reaches 70% to avoid overfitting the training model. The performance parameters here have been computed as mean $\pm$ standard deviation for the sake of indicating the classification performance.

The classification performance of the proposed method is illustrated in Table 3. It shows that efficient classification performance has been achieved by using 27 proposed features. In this study, we focus on increasing the sensitivity rather than the specificity since the cost for misclassifying abnormal heart sounds and healthy heart sounds are not the same. Hence, training a model using 70% data and testing with the remaining 30% achieved an accuracy of 92.47% with a sensitivity of 94.08%. Figure 5 illustrates the accuracy plot of the proposed ensemble model over 100 epochs for different training and testing samples.

Table 3. Classification performance of the proposed method.

Train%	Test%	Sensitivity%	Specificity%	Accuracy%
10%	90%	78.92±6	94.03±4	90.36±2
20%	80%	80.89±4	94.56±4	91.24±2
30%	70%	83.45±4	94.81±3	92.05±2
40%	60%	85.49±3	94.16±3	92.06±2
50%	50%	86.38±3	93.89±2	92.07±1
60%	40%	87.49±2	94..08±2	92.47±1
70%	30%	94.08±1	91.95±1	92.47±1

The training and testing samples are mutually exclusive to each other.
The performance parameters are presented in mean±SD.

6. Discussion

In this study, a total of 27 features are extracted from a single PCG recording. For evaluating the performance of the model, extracted features were trained based on the supervised learning method using the ensemble classifier. The classification of an abnormal or healthy heart sound is carried out by the ensemble model based on the AdaBoost algorithm as illustrated in Equation 13. The 27 extracted features were used as input to the ensemble model and class label as an output reference. The ensemble model was trained by part of the data and tested with the remaining data.

Based on the previous work for heart sound classification, few researchers employed feature extraction methods after eliminating a segmentation step using the PhysioNet 2016 PCG database [23–25, 33]. Singh and Majumder [33] used 11 features based on time-frequency characteristics, but their classification performance was unsatisfactory. However, in this study, we used multidomain features that capture complete information of the PCG signal based on the MFCC coefficient, power spectral density (PSD), and envelope extraction that eliminates complex segmentation steps. Our proposed ensemble-based classifier achieved an accuracy of 92.47% with sensitivity of 94.08% and specificity of 91.95%.

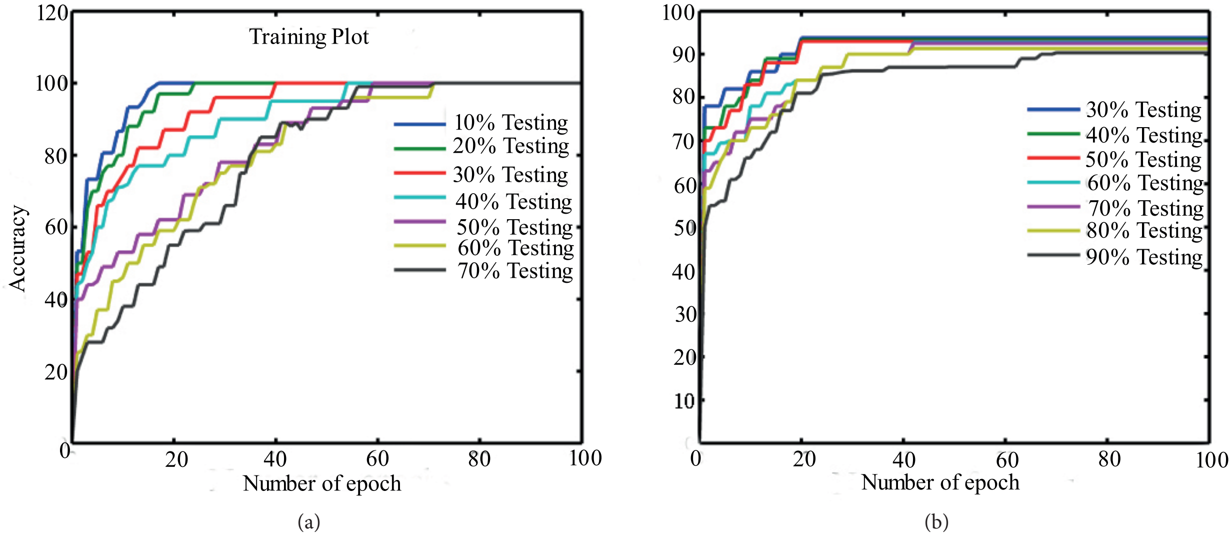


Figure 5. Accuracy plot for training and testing the proposed ensemble model: (a) training plot, (b) testing plot.

To estimate the review of our proposed model, we have computed the performance of several traditional classifiers using our proposed features. The comparison of our proposed model with other classifiers is explained in Table 4. The KNN-based model can significantly increase the performance as compared with other classifier models. The KNN classifier improves the classification accuracy from 86.22% to 94.03%. Although the specificity of the KNN model is higher than that of other models, we prefer to choose the model that can enhance the sensitivity. Hence, our proposed method based on the ensemble classifier achieved outstanding classification accuracy of 92.47% with improved sensitivity of 94.08%.

Table 4. Comparison of the proposed method with other existing methods.

Method	Sensitivity	Specificity	Accuracy
Proposed	94.08%	91.95%	92.47%
SVM	93.66%	83.83%	86.22%
KNN	87.64%	96.07%	94.03%
Classification tree	78.14%	93.33%	89.65%
LDA	73.07%	92.86%	88.06%

The comparison is based on the mean value of the performance parameter using the proposed model.

Comparison with the other state-of-the-art method that used the same datasets provided by PhysioNet/CinC 2016 Challenge is illustrated in Table 5. Our proposed method was trained using multimodel features with respective class labels as output. While reviewing previous studies, we observed that previous works focused on enhancing the accuracy but most of them neglected sensitivity. Sensitivity has been treated as an essential parameter for analyzing heart sound signals in the clinical field as the cost of misclassifying abnormal and normal signals is different. The variation of normal (9857) and abnormal samples (3158) results in degrading the performance of the model [34].

Table 5. Comparison of the proposed method with the existing methods.

Author	Sensitivity%	Specificity%	Accuracy%
Tang et al. [19]	88.00	87.00	88.00
Dominguez et al. [12]	93.20	95.12	97.00
Bradley et al. [11]	90.07	88.45	89.26
Mostafa et al. [13]	76.96	88.31	82.63
Masun et al. [14]	79.60	80.60	80.10
Plesinger et al. [17]	89.00	81.60	85.00
Vykintas et al. [20]	80.63	87.66	84.15
Wei et al. [21]	98.33	84.67	91.50
Philip et al. [24]	77.00	80.00	79.00
Singh et al. [33]	93.00	90.00	90.00
This study	94.08	91.95	92.47

In comparison with the state-of-art methods, the accuracy of the predicted models varies from 79% to 97% while the sensitivity ranges from 76.96% to 98.33%. The authors in [12] obtained overall accuracy of 97%, but they failed to improve the sensitivity by introducing a complex and time-consuming system as their approach was based on deep learning. Again based on unsegmented features, Philip et al. [24] achieved overall accuracy of 79% with 77% sensitivity, while our model-based classifier enhanced all the performance parameters with an accuracy of 92.47% with 94.08% sensitivity and 91.95% specificity. Another work [33] employed fewer features based on unsegmented heart sounds with classification accuracy of 90% along with sensitivity and accuracy of 93% and 90%, respectively. Hence, our proposed model using the time domain, high-order statistics, entropy, frequency, and cepstral coefficients is superior to the previous methods with improved sensitivity and overall accuracy. Furthermore, an ensemble-based classifier using relevant multidomain features is necessary to overcome the shortcomings suffered by the existing works. In the literature, most researchers focus on improving classification accuracy and specificity but have failed to enhance the sensitivity. Sensitivity is treated as an essential parameter concerning heart anomaly detection. Moreover, the cost of misclassification of an abnormal heart is considerably higher as compared to a normal one, so sensitivity is interpreted as a vital parameter in clinical treatment. Therefore, our proposed method based on unsegmented features overcomes all the hindrances confronted by the state-of-the-art method with a simple and low-cost model.

7. Conclusion

We have presented a model-based ensemble classifier for predicting anomalies of heart sound signals. We further prove that by eliminating a segmentation of the cardiac cycle, we can enhance the performance efficiency of the model with less complexity. The PCG samples were collected from the PhysioNet/CinC 2016 Challenge. Preprocessing steps remove the unwanted high-frequency noises and murmurs. The methods for extracting the features based on multidomain characteristics followed by classification using the ensemble boosting algorithm were presented. The results demonstrate accuracy of 92.47% based on 100 independent simulations. Furthermore, despite selecting features randomly while training, the ensemble classifier operates well with fewer features and high performance.

Conflict of interest statement

There are no conflicts of interest.

References

[1] Lloyd-Jones D, Adams RJ, Brown TM, Carnethon M, Dai S et al. Heart disease and stroke statistics–2010 update: a report from the American Heart Association. Circulation 2010; 121 (7): e46

[2] Etchells E, Bell C, Robb K. Does this patient have an abnormal systolic murmur? JAMA 1997; 277 (7): 564-571.

[3] Jabbari S, Ghassemian H. Modeling of heart systolic murmurs based on multivariate matching pursuit for diagnosis of valvular disorders. Computers in Biology and Medicine 2011; 41 (9): 802-811.

[4] Ahlstrom C, Hult P, Rask P, Karlsson JE, Nylander E et al. Feature extraction for systolic heart murmur classification. Annals of Biomedical Engineering 2006; 34 (11): 1666-1677.

[5] Maglogiannis I, Loukis E, Zafiropoulos E, Stasis A. Support vectors machine-based identification of heart valve diseases using heart sounds. Computer Methods and Programs in Biomedicine 2009; 95 (1): 47-61.

[6] Hassani K, Bajelani K, Navidbakhsh M, Doyle DJ, Taherian F. Heart sound segmentation based on homomorphic filtering. Perfusion 2014; 29 (4): 351-359.

[7] Gupta CN, Palaniappan R, Swaminathan S, Krishnan SM. Neural network classification of homomorphic segmented heart sounds. Applied Soft Computing 2007; 7 (1): 286-297.

[8] Springer DB, Tarassenko L, Clifford GD. Logistic regression-HSMM-based heart sound segmentation. IEEE Transactions on Biomedical Engineering 2015; 63 (4): 822-832.

[9] Schmidt SE, Holst-Hansen C, Graff C, Toft E, Struijk JJ. Segmentation of heart sound recordings by a duration-dependent hidden Markov model. Physiological Measurement 2010; 31 (4): 513.

[10] Golpaygani AT, Abolpour N, Hassani K, Bajelani K, Doyle DJ. Detection and identification of S1 and S2 heart sounds using wavelet decomposition method. International Journal of Biomathematics 2015; 8 (6): 1550078.

[11] Whitaker BM, Suresha PB, Liu C, Clifford GD, Anderson DV. Combining sparse coding and time-domain features for heart sound classification. Physiological Measurement 2017; 38 (8): 1701.

[12] Dominguez-Morales JP, Jimenez-Fernandez AF, Dominguez-Morales MJ, Jimenez-Moreno G. Deep neural networks for the recognition and classification of heart murmurs using neuromorphic auditory sensors. IEEE Transactions on Biomedical Circuits and Systems 2017; 12 (1): 24-34.

[13] Abdollahpur M, Ghaffari A, Ghiasi S, Mollakazemi MJ. Detection of pathological heart sounds. Physiological Measurement 2017; 38 (8): 1616.

[14] Homsi MN, Warrick P. Ensemble methods with outliers for phonocardiogram classification. Physiological Measurement 2017; 38 (8): 1631.

[15] Zhang W, Han J, Deng S. Heart sound classification based on scaled spectrogram and tensor decomposition. Expert Systems with Applications 2017; 84: 220-231.

[16] Messner E, Zöhrer M, Pernkopf F. Heart sound segmentation—an event detection approach using deep recurrent neural networks. IEEE Transactions on Biomedical Engineering 2018; 65 (9): 1964-1974.

[17] Plesinger F, Viscor I, Halamek J, Jurco J, Jurak P. Heart sounds analysis using probability assessment. Physiological Measurement 2017; 38 (8): 1685.

[18] Kamson AP, Sharma LN, Dandapat S. Multi-centroid diastolic duration distribution based HSMM for heart sound segmentation. Biomedical Signal Processing and Control 2019; 48: 265-272.

[19] Tang H, Dai Z, Jiang Y, Li T, Liu C. PCG classification using multidomain features and SVM classifier. BioMed Research International 2018; 1: 1-14. doi: 10.1155/2018/4205027

[20] Maknickas V, Maknickas A. Recognition of normal–abnormal phonocardiographic signals using deep convolutional neural networks and mel-frequency spectral coefficients. Physiological Measurement 2017; 38 (8): 1671.

[21] Han W, Yang Z, Lu J, Xie S. Supervised threshold-based heart sound classification algorithm. Physiological Measurement 2018; 39 (11): 115011.

[22] Deng SW, Han JQ. Towards heart sound classification without segmentation via autocorrelation feature and diffusion maps. Future Generation Computer Systems 2016; 60: 13-21.

[23] Hamidi M, Ghassemian H, Imani M. Classification of heart sound signal using curve fitting and fractal dimension. Biomedical Signal Processing and Control 2018; 39; 351-359.

[24] Langley P, Murray A. Heart sound classification from unsegmented phonocardiograms. Physiological Measurement 2017; 38 (8): 1658.

[25] Langley P, Murray A. Abnormal heart sounds detected from short duration unsegmented phonocardiograms by wavelet entropy. In: IEEE 2016 Computing in Cardiology Conference (CinC); Vancouver, Canada; 2016. pp. 545-548.

[26] Clifford GD, Liu C, Moody B, Springer D, Silva I et al. Classification of normal/abnormal heart sound recordings: the PhysioNet/Computing in Cardiology Challenge 2016. In: IEEE 2016 Computing in Cardiology Conference; Vancouver, Canada; 2016. pp. 609-612.

[27] Liu C, Springer D, Li Q, Moody B, Juan RA et al. An open access database for the evaluation of heart sound algorithms. Physiological Measurement 2016; 37 (12): 2181.

[28] Das R, Sengur A. Evaluation of ensemble methods for diagnosing of valvular heart disease. Expert Systems with Applications 2010; 37 (7): 5110-5115.

[29] Freund Y, Schapire RE. A decision-theoretic generalization of on-line learning and an application to boosting. Journal of Computer and System Sciences 1997; 55 (1): 119-139.

[30] Duba RO, Hart PE. Pattern Classification and Scene Analysis. New York, NY, USA: Wiley, 1973.

[31] Arnott PJ, Pfeiffer GW, Tavel ME. Spectral analysis of heart sounds: relationships between some physical characteristics and frequency spectra of first and second heart sounds in normals and hypertensives. Journal of Biomedical Engineering 1984; 6 (2): 121-128.

[32] Zhao L, Wei S, Zhang C, Zhang Y, Jiang X et al. Determination of sample entropy and fuzzy measure entropy parameters for distinguishing congestive heart failure from normal sinus rhythm subjects. Entropy 2015; 17 (9): 6270-6288.

[33] Singh SA, Majumder S. Classification of unsegmented heart sound recording using KNN classifier. Journal of Mechanics in Medicine and Biology 2019; 19 (4): 1950025.

[34] Sun Y, Kamel MS, Wong, AK, Wang Y. Cost-sensitive boosting for classification of imbalanced data. Pattern Recognition 2007; 40 (12): 3358-3378.