

Chapter 9

Peak Spectrogram and Convolutional Neural Network-Based Segmentation and Classification for Phonocardiogram Signals



Anam Abid and Zo-Afshan

Abstract Heart diseases are one of the contributing reasons for the human loss in the world. The Phonocardiogram (PCG) provides information for the exact and timely detection of Cardiovascular Diseases (CVDs). Digital stethoscopes record the heart sound and store it as a PCG signal for the identification of abnormal sounds. This chapter discusses heart sound segmentation and classification algorithms to diagnose the abnormal symptoms of the heart. Firstly, the breakdown of heart signal into “S1”, “systole”, “S2”, and “diastole” states is performed using multi-level threshold values for peak detection. Phonocardiograms are the non-stationary signals for which the identification of the exact location of the peaks is difficult. Using the multi-level threshold method for peak detection and with the peak spectrogram generation, the identification of the peak locations is improved to 91.2%. Subsequently, features of the generated peak spectrograms are extracted from these four states to perform the naming of PCG as “normal” or “abnormal” using a traditional “Support Vector Machine” (SVM) and “Convolutional Neural Network”. The developed algorithms are tested on the PhysioNet2016 heart sound challenge dataset. The results show the suitability of the developed methods for the identification of “normal” and “abnormal” PCG for CVDs identification and scored an accuracy of 93.3%.

A. Abid (✉)

Department of Mechatronics Engineering, University of Engineering and Technology,
Peshawar, Pakistan

e-mail: anam.abid@uetpeshawar.edu.pk

Zo-Afshan

Department of Electrical and Computer Engineering, Air University, Islamabad, Pakistan

e-mail: zoafshan@mail.au.edu.pk

9.1 Introduction

Medical Signal Processing, an emerging field, has come into the limelight and shown a vital role in the identification of solutions for various diseases. Along with the recent advancement, it is now realized that various medical issues require intelligent systems and assistance. Particularly, “Deep Learning” algorithms i.e., “Convolutional Neural Networks” (CNN), “Long-Short Term Memory”, and “Recursive Neural Networks”, have had a great impact on biomedical fields. These systems are supposed to work like a human, think like a human, possess the decision-making ability, and also explain how to take action. The quintessence of designing an intelligent system of this kind is Deep Learning Convolutional Neural Network: convolution that extracts dominant features and neural networks that are designed to recognize patterns and adapt the environmental changes to cope with real-time scenarios.

An intelligent system has made nearly impossible tasks into reality. It is frequently advantageous to use multi-disciplinary computing techniques and methodologies in cooperation rather than exclusively. One such synergistic construction of an intelligent system is the prelude of this chapter. In particular, the integration of two complementary approaches: Signal Processing and Machine Learning, results in an innovative approach for the analysis and early detection of cardiovascular disease through phonocardiogram (PCG) signals.

Agreeing to the “World Health Organization” (WHO), CVDs are one of the most important reasons for death in the world. From the statistics reported by WHO, 17.9 million people lost their lives in 2017 alone due to CVDs [1]. Considering the severity of the situation, there is a dire need for an automated system that can be used for the evaluation of heart conditions and timely detection of any abnormality.

9.1.1 Auscultation

The heart system comprises of “heart”, “blood vessels”, and “blood”, whereas the heart is its most important organ. Therefore, the functioning and monitoring of the heart become crucial to avoid a heart attack, angina, or stroke. A heart abnormality originates due to the constriction or complete blockage of the blood vessels and may result in serious complications. Auscultation is a medical screening process in which a medical expert/physician listens to the sounds/murmurs of the body organs by using a stethoscope. During heart sound auscultation, heartbeats present very useful information for the recognition of abnormalities. However, only a handful of skilled medical experts can correctly identify heart abnormality by listening to the heart sounds. Biomedical signal processing and artificial intelligence techniques have been facilitating doctors and cardiologists in heart abnormality diagnosis over the last few decades.

9.1.2 Phonocardiogram Signal

Stethoscopes are considered essential screening instruments for the diagnosis of heart and lung pathologies. With the development of digital stethoscopes, the process of auscultation of heart sounds becomes easier and more convenient. Digital stethoscopes record the heart sound and store them for the identification of heart abnormalities. A Phonocardiogram (PCG) signal (refer to Fig. 9.1) consists of various events like “S1” sound, “S2” sound, and “murmurs” [2, 3]. “S1” and “S2” correspond to the primary PCG components (i.e., systolic and diastolic activities of the heart respectively) which are important for heart sound segmentation. In particular, each heart sound segment contains different characteristics. In segmented data analysis, the information from each heart sound segment is retrieved for an in-depth analysis and heart disease detection [4]. Consequently, in unsegmented data analysis methods, the entire PCG signal is given as input. A discussion on segmented and unsegmented PCG data analysis is given in the following section.

The heart pumping process represents a synergetic combination of mechanical and electrical activities which form a certain set of activities over the entire cycle. The blood is circulated with efficient coordination of atria and ventricles with each other. The cardiac cycle includes two phases: Systole (the contraction phase) and Diastole (the relaxation phase). During systole, contraction of atria or ventricles occurs and the blood is pushed to the arteries whereas during diastole, relaxation of heart muscles occurs and the blood is supplied to the heart. The systole period represents the contraction of the right and left ventricles and discharge of blood into the aorta and pulmonary artery which is allowed through the opening of the aortic and pulmonic valves while the atrioventricular valves remain closed during the systole period to prevent the blood flow into the ventricles. Diastole represents the relaxation of left and right ventricles. The blood runs through the mitral and tricuspid valves. Left and right atria contract at the end of the diastole period pushing an extra

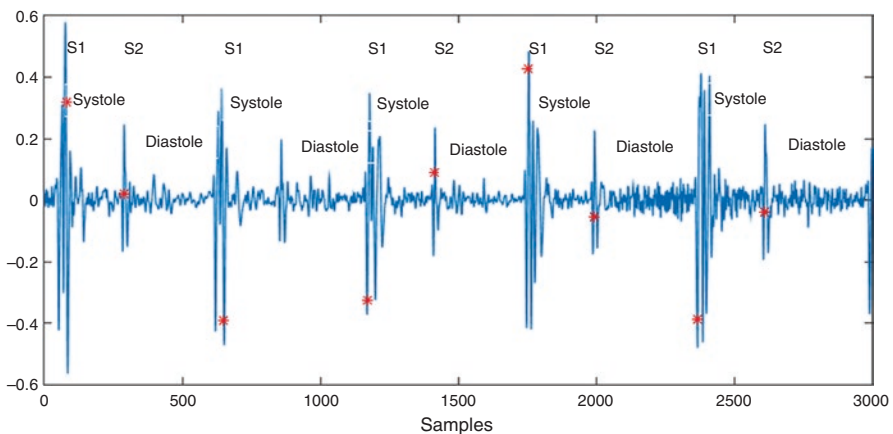


Fig. 9.1 Visualization of a phonocardiogram signal

amount of blood into the ventricles. The generated electrical signal during the heart pumping process force the blood flow between heart chambers and throughout the body. The heart produces sounds as a result of the heart beating and blood flow through it during the cardiac cycle. Also, the vibrations are produced with the closure of heart valves creating turbulence, which is audible and can be listened to through a stethoscope during cardiac auscultation by the examiner. The heart sounds are distinct and unique which gives valuable acoustic information about the heart condition. Normally in adults, there are two heart sounds, i.e., “Lub” and “Dub”, generated due to the closing of the semilunar and atrioventricular valve.

There are two normal primary heart sounds, i.e., “S1” and “S2”, associated with heart valves closing. S1 is also called first heart sounds or “Lub”. S1 is generated by the closing of tricuspid and bicuspid valves during the start of the systole period. The vibrations are produced as a result of turbulence during ventricles contraction in systole and they could be easily heard with a stethoscope placed at the heart vertex. It has two parts: M1 which is caused by mitral valve closing and T1 which is caused by tricuspid valve closure. M1 occurs before T1 with an approximate 25–45 cycles per second whereas it elapses for an interval of around 0.14–0.15 s [4].

S2 is also known as second heart sound or “Dub”. S2 is produced by semilunar valve closure during the end of the systole or early diastole period. S2 is best heard with the stethoscope placed in the aortic area. It has two components: A2 caused by the aortic valve closure and P2 caused by the closure of the pulmonary valve. Generally, S2 sound is louder and high-pitched as compared to S1 sound with a frequency falling in the range of 40–70 Hz. In addition, its duration is relatively longer which elapses for an interval of around 0.11–0.12 s [4].

9.1.3 PCG Signal Acquisition

PCG signals are correlated with the mechanical activity of the heart and provide a means of visualization for better analysis. PCG signals provide the most valuable qualitative and quantitative heart-related attributes. PCG signal acquisition process is categorized as one-channel acquisition and multiple-channel acquisition. In the one-channel case, the PCG signal is fragmented using the actual signal without any prior knowledge. In the multiple-channel scenario [5], certain signals for example an electrocardiogram, photoplethysmogram, and carotid pulse are simultaneously obtained along with PCG. As a result, the performance of the multiple-channel acquisition setup is more effective than its one-channel counterpart. Nonetheless, simultaneous acquisition of multiple signals (modalities) becomes expensive and unmanageable, especially when conditions are ambulatory. Hence, field experts prefer one-mode segmentation methods over multiple-channel counterparts [6–8].

The rest of the chapter is organized as follows. Section 9.2 discusses the recent work carried out for PCG segmentation and classification. Section 9.3 discusses the quality assessment and pre-processing of the PCG signal and Sect. 9.4 presents a threshold-based peak detection method. Section 9.5 details the proposed

segmentation techniques i.e., identification of “S1” and “S2” states by (i) calculating statistical features, and (ii) converting 1D PCG signals into their 2D spectrogram respectively. Section 9.6 discusses the final labeling of PCG signal into “S1”, “systole”, “S2”, and “diastole” states, followed by PCG classification. In Sect. 9.7 experiment and results are presented based on the developed methods. Sections 9.8 and 9.9 present a comparison study, discussions, and conclusions of the chapter respectively.

9.2 Related Work

Automated analysis of a PCG signal can be classified into various steps. The typical approach involves (i) pre-processing, (ii) segmentation, and (iii) classification. In pre-processing PCG signal is filtered and extra noise spikes are removed to make the signal more appropriate for starting the analysis process. Most of the heartbeat segmentation methods follow a similar preprocessing approach, starting with noise reduction by applying filters and the normalization of the signal using the absolute maximum. This is followed by the application of envelope detection methods, such as the Hilbert Transform, Homomorphic Envelope, Wavelet Transform, Shannon Energy, and Power Spectral Density [9]. In the unsegmented PCG signal processing approach, deep features are obtained from the unsegmented PCG chunks, and deep learning algorithms are employed for classification.

9.2.1 Segmentation

In earlier works, the rule of thumbs and facts-based distinction was used to differentiate between “S1” and “S2”, (i.e., interval lengths). Other techniques [10, 11] involve PCG energy calculation. Gomes [11] designed a system that changes the phonocardiogram signal into individual segmentation fragments. In [10], the Shannon envelope of the signal is extracted from the overlapping fragments of the entire signal. Threshold-based peak detection is performed in each window/fragment. Heart sound signals are mainly known as non-stationary signals. Hence, by applying energy-based calculation methods only, better results cannot be obtained. Considering this issue, researchers combined these methods with some transforms, for instance, Short Time Fourier Transform [12], Wavelet Transform [13, 14]. In [14], a wavelet-transform-based segmentation algorithm was employed to extract temporal, time, and frequency domain attributes of PCG. In [15], the authors used an advanced mode of decomposition method for PCG segmentation using various modes of the decomposed PCGs and variational mode decomposition. Other than these techniques, neural networks have also been applied for heart sound segmentation [16]. Specifically, a neural network algorithm was proposed for PCG segmentation using a Hidden Markov Model. Features extracted for the PCG classification

were the time and frequency domains representing the underlying characteristics of the phonocardiogram signal. The inter-patient differences challenge was addressed in [17] where the main focus is to compare the heart sounds within and across the patient's PCG dataset using "Dynamic Time Warping" (DTW). The combination of DTW and "Mel Frequency Cepstral Coefficients" (MFCC) features was given to an SVM classifier. As compared to MFCC, DTW-based features computed in an unbiased dataset condition performed well. In [18] heart sound recordings are classified by first performing the heart sound segmentation, followed by 1D waveforms transformation into 2D time-frequency heat maps using MFCC, and finally, classification was performed using CNN.

9.2.2 *Extracted Features and Classifiers*

After the segmentation of PCG signals as "S1" and "S2", these segments are passed to the feature (distinctive attribute of an item) extraction stage, followed by the classification stage. For PCG classification features are extracted to analyze the changes in the signal over time and frequency contents within the signal. In the segmentation-based heart sound analysis approach, different features are calculated at two different stages. At first, the features are extracted to find the S1 and S2, and in the second stage (i.e., classification) features are extracted to classify the signal as normal and abnormal. Some common features are mean, median, kurtosis, energies, entropy, spectral edges, etc.

Some classification methods are based on clustering like K-Means [19]; others use statistical analysis like Hidden Markov Model [20], K-Nearest Neighbor [21], etc. Machine learning (ML) models are applied to PCG databases with several feature extraction approaches [22–24]. In [9], Springer used a modified form of the Hidden Markov Model (the Hidden Semi Markov Model) for the classification of PCGs. Similarly in [25], Kaur used fuzzy K-NN, Bayesian, and Gaussian mixer Model-based KNN for the classification. In [26], two stages were employed, the first stage performed segmentation using SVM, and "Artificial Neural Networks" (ANN) were used for the final classification of PCGs. Most of the recent studies employed classification techniques like "multi-layer perceptron" (MLP) [27], SVM [28], CNN [18], etc. The above-mentioned methods use different preprocessing techniques to segment PCG signals and extract suitable features from the PCG segments using techniques such as Short Time Fourier Transform [12], Wavelet Transform [13, 14], DTW [17], MFCC [29], etc. These ML methods are subjective and time-consuming due to the handcrafted feature selection process. In addition, deep features using deep neural networks [30, 31], spectrograms of heart sounds [32, 33], and a continuous wavelet transform-based scalogram [34] were also used for PCG analysis.

The existing research work conducted on CVD identification using ML and deep learning on different medical databases has contributed to the detection of heart

sound abnormalities and most of them achieved significant results. The proposed study is focused on spectrogram-based segmentation of PCG using CNN.

9.2.3 Unsegmented PCG Classification

The second approach for the PCG classification does not involve segmentation. Researchers have opted for this approach to directly classify the PCG signals into normal and abnormal classes skipping the intermediate segmentation stage. In [35] authors performed the classification of heart sound without segmentation using 5 categories of features that include “Linear Predictive Coefficient”, Entropy-based features, MFCC, wavelet transform, power spectral density. From the set of 40 features, 18 features were chosen using one of the search algorithms (called wrapper-based) for the feature selection. In this method, a sequential forward selection search was used. A total of 20, 2-layer feed-forward ANNs were used for the classification (25 neuron nodes per hidden layer). In the output layer, 4 neurons were used for two classification tasks at the same time, two for normal vs. abnormal and two for good vs. bad. In [36] authors classify the heart sound recording as normal or abnormal by extracting the morphological features of the PCG signals. Several features are extracted from both temporal and spectral domains, and the classification is performed using an SVM. In another approach [37], wavelet entropy at a wavelet scale of 1.7 and with a threshold of 7.8 was employed. The heart signal was recorded for 5 seconds, and then wavelet coefficients were calculated. Afterward, wavelet energy and entropy were calculated and it was passed to the criterion function which used a threshold for signal classification. Another CNN-based PCG classification approach [38] employed Power Spectral Density (PSD) features with a window of 150 ms. These spectrograms were fed to the network for the classification task. SVM, logistic regression, and random forest were also applied for PCG classification and their results were compared. Another method using temporal dynamics of the signal using Markov features along with other statistical and frequency domain features was presented in [39]. These features were trained over the ensemble of artificial neural networks and gradient boosting trees.

9.3 Quality Assessment and Pre-processing of PCG Signals

In a real-world environment, during auscultation, the recorded PCG signal is often contaminated with noise. It is always necessary to check the suitability of the PCG signal before carrying out any kind of processing. For this purpose, firstly quality assessment of the PCG signal is carried out [16] in which the suitability of the signal is tested based on evaluation criteria. If the PCG signal fails the criteria, it is declared as “unsure” and no further processing is carried out for that PCG signal. The details of quality assessment and pre-processing stages are discussed in this section.

9.3.1 *Evaluation Criteria*

The classification task requires the determination of a heart sound recording as normal, abnormal, or unsure (due to the high content of noise). For this purpose, three measures of quality assessment are taken [16]. If any of the criteria does not meet, the signal is not called suitable. These criteria are; (i) root mean square of successive differences, (ii) number of counted normal peaks in the specified size window, and (iii) number of zero crossings in the whole PCG signal. PCG signal is tested over these criteria, and the suitable signal is passed on to the next step. On the other hand, if any criterion fails, the PCG signal will be declared as “unsure” and no further processing will be carried out for that PCG signal.

9.3.2 *Filtering and Spike Removal*

Heart sound recording and analysis are generally employed as an effective and low-cost alternative for heart abnormality screening. Nonetheless, there are a few challenges involved in this process. Firstly, the accurate localization of primary heart sounds (i.e., “S1” and “S2”) is very important for the detection of any heart abnormality as it provides the basis for the upcoming classification stage. Another challenge is the vulnerability of heart sounds from different noise sources. In particular, external noises present in the nearby setting of the signal acquisition arrangement (e.g., human speech, noise generated by appliances and devices), measurement noise due to the involvement of sensors, and other components data acquisition system. In contrast, internal noises (coming from the patient body), for instance, sounds originating from lungs and other body parts, speech, etc. may also deteriorate the desired PCG signal. Consequently, it is the foremost task to remove the undesired noises from the acquired signal using appropriate noise filtering methods.

Filters play an essential role in the field of signal processing. A filter is a special type of process which is used to remove the unwanted part of the signal, suppress the effect of unwanted or unnecessary signal, or restore the original signal from corruption. Normal PCG signals have low frequencies, ranging from 40 Hz to 200 Hz. Murmurs and extra heart sounds have frequency ranges up to 400 Hz. In literature the usage of Butterworth and Chebyshev filters [17, 28, 40] and their variants are found quite often. These linear filters are used for the separation of noise from the signal using different cutoff frequencies provided that the signal does not overlap in the frequency domain. In this study, the PCG signal is resampled to 1 kHz and the resampled signal is filtered with a Butterworth low-pass filter with a cut-off frequency of 400 Hz and order 4. After that output of the first filter is passed to a Butterworth high-pass filter with a cut-off of 25 Hz.

Noisy spikes are then removed from the filtered signal to make the PCG signal clean from extra spikes other than the actual peaks of the heartbeat. Some common spike removal methods are Nonlinear Median filters, Schmidt spike removal

function [20], etc. After that envelope of the PCG using the Hilbert transform is extracted and it is normalized using the mean and standard deviation values of the extracted PCG envelope. Researchers also suggest calculating the Shannon Energy [10] for PCG envelope calculation.

9.4 Single and Multi-Level Threshold-Based Peak Detection Methods

During PCG signal segmentation, a signal is segmented into its fundamental heart sounds, i.e., “S1” and “S2”. Therefore, the localization is threefold; the identification of all peaks present in the normalized PCG, extraction of true peaks, and removal of false peaks. Firstly, the local maxima function is employed to identify the location of all peaks, called candidate peaks, from the normalized envelope of the PCG signal. The second step is the determination of true peaks which becomes difficult due to the sub-par quality of the PCG acquisition process. Generally, clinical settings or ambulatory conditions affect the recorded signal quality due to the presence of external and observational noise. In such situations, a specific single-threshold is not a suitable measure for extracting true peaks for the reason that signals usually have a “Signal to Noise ratio” (SNR) of a different range. Due to the degraded performance of single-threshold techniques, employment of one perfect, global threshold value for the determination of true peaks i.e., “S1” and “S2”, is not possible. Another associated challenge of a one-specific threshold is the determination of the threshold level. For instance, by selecting a low threshold value, peaks in systole/diastole intervals are also selected along with true peaks. In contrast, by selecting a high threshold value sometimes the misdetection of “S1” peaks from the PCG signal increases considerably, however, the peaks in systole and diastole intervals are not detected anymore. This is because the signal with a high SNR performs well with a low threshold value and a signal with low SNR requires a high threshold value. In literature, a multi-threshold algorithm [28] is suggested to find out the candidate peaks for S1 and S2 from the pool of all detected peaks. The summary of the multi-threshold algorithm is shown in Fig. 9.2.

The proposed method employs multiple threshold levels for true peak selection namely, “moderate-level (MLT)”, “high-level (HLT)”, and “low-level (LLT)”. These levels are applied in sequence and the criterion for the selection of a threshold is a count of candidate peaks, ‘count’, occurring in a pre-defined window size (0.2 s in our case). To begin with, the ‘count’ is computed with MLT (0.1 in our case). When ‘count’ is above the specified upper limit (sp1), HLT is incorporated. Similarly, when ‘count’ is below the specified lower limit (sp2), LLT is employed. Afterwards, true peaks fulfilling the requirements i.e., equal to or above the updated threshold are selected using the freshly updated threshold. Over each candidate peak, a window of 1 ms with overlapping of 0.5 ms is placed to segment that particular portion of the signal.

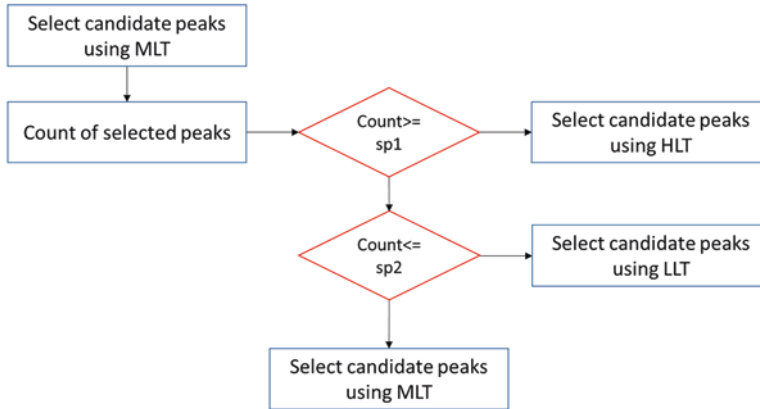


Fig. 9.2 Multi-Level threshold algorithm

9.5 Segmentation Methods of PCG Signals

In the proposed methodology, two approaches are employed for the classification of “S1” and “S2”:

- Segmentation based on statistical features and Support Vector Machine (SVM)
- Segmentation based on Peak Spectrogram and Convolutional Neural Networks (CNN)

Both approaches are discussed in detail in this section.

9.5.1 Segmentation Based on Statistical Features and Support Vector Machine

In this approach, a total of 11 features obtained from both time and frequency domains are extracted from the windowed PCG segments (wS) and complete heart sound (HS). A list of these features is mentioned in Table 9.1. These features provide statistical values for the classification of “S1” and “S2”. Different classifiers are trained using these 11 features namely K-Nearest-Neighbor (KNN), ANN, and SVM. Accuracy results obtained from all mentioned classifiers are given in Sect. 9.8.

Table 9.1 Proposed features for heart sound peak classification

1. SD of 'wS'/SD of heart 'HS'
2. μ of 'wS'/ μ of 'HS'
3. S of 1st-level approximation coefficient of 'HS'
4. S of 1st-level detail coefficients of 'HS'
5. S of 2nd-level approximation coefficient of 'HS'
6. Spectral edge frequency of 'HS'
7. Spectral S for dyadic bands of 'HS'
8. Fractional dimensions of 'HS'
9. Hjorth parameters of 'HS'
10. Skewness of 'HS'
11. Kurtosis of 'HS'

Note: SD, Standard deviation; μ , Mean; S, Entropy

9.5.2 Segmentation Based on Spectrograms and Convolutional Neural Network

Convolutional Neural Network (CNN) works better on 2D data. On the other hand, the signal acquired during auscultation is 1D. To make it useful for CNN, in this approach, Short Time Fourier Transform of PCG segment determined after which 1D signal is converted into a 2D peak spectrogram. The time-domain representation of "S1" and "S2" and their respective spectrograms are shown in Fig. 9.3.

These peak spectrograms are fed to a CNN which classifies them into "S1" and "S2". The architecture used for the CNN model is shown in Fig. 9.4. At the end of this step, all candidate peaks are assigned with their respective labels, like "S1" or "S2". This information is used in the next phase of the proposed methodology to get fully labeled cardiac cycles in the PCG signal.

9.6 Post-processing and Classification of PCG Signals

9.6.1 Post-processing and PCG Labeling

In this step, the marked positions of "S1" and "S2" along with the duration distribution provided by Schmidt et.al [20]. are utilized to label the systole and diastole regions in the PCG signal. A fully labeled PCG signal with states "S1", systole, "S2", and diastole is obtained after the post-processing. The example of a labeled PCG signal obtained after segmentation is presented in Sect. 9.7.

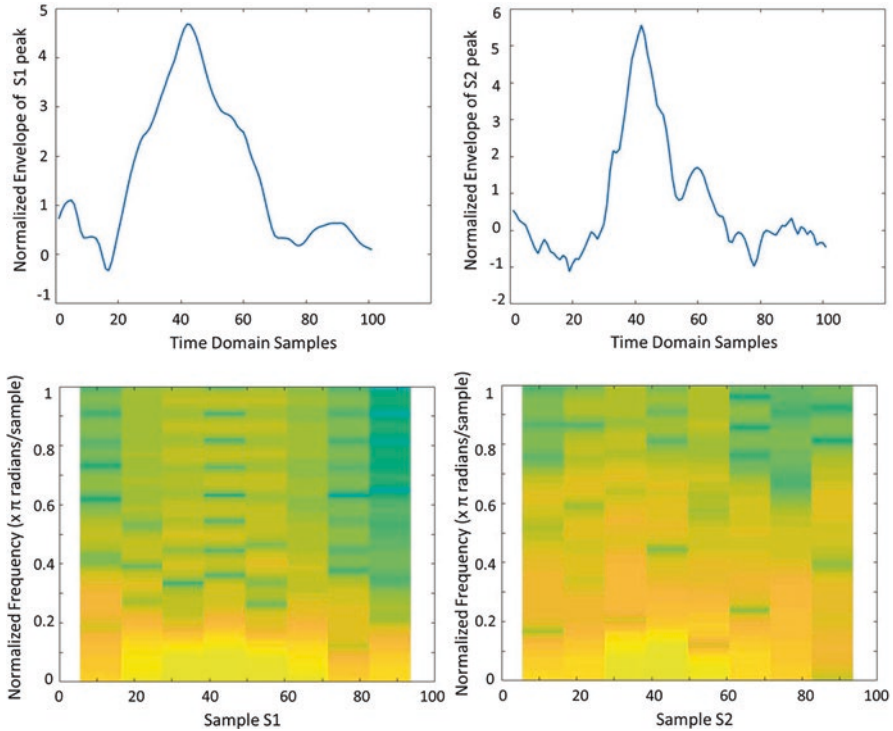


Fig. 9.3 Peak Spectrogram generation using short-time-Fourier-transform

9.6.2 PCG Classification

In this step, features are extracted utilizing post-processed state labels for PCG signals that are used to train the classifier. A total of 50 features are extracted (20 time-domain, 30 frequency-domain). A list of these features is given in Table 9.2. These features are used to train SVM. As mentioned earlier, for segmentation, two approaches were proposed. Both of them follow the same classification step separately and their accuracies are reported in Sect. 9.7.

9.7 Experimentation on the PhysioNet2016 Challenge Dataset

This section discusses the experiments and the obtained results of the proposed method implementation. A detailed description of the results obtained in each step is given below.

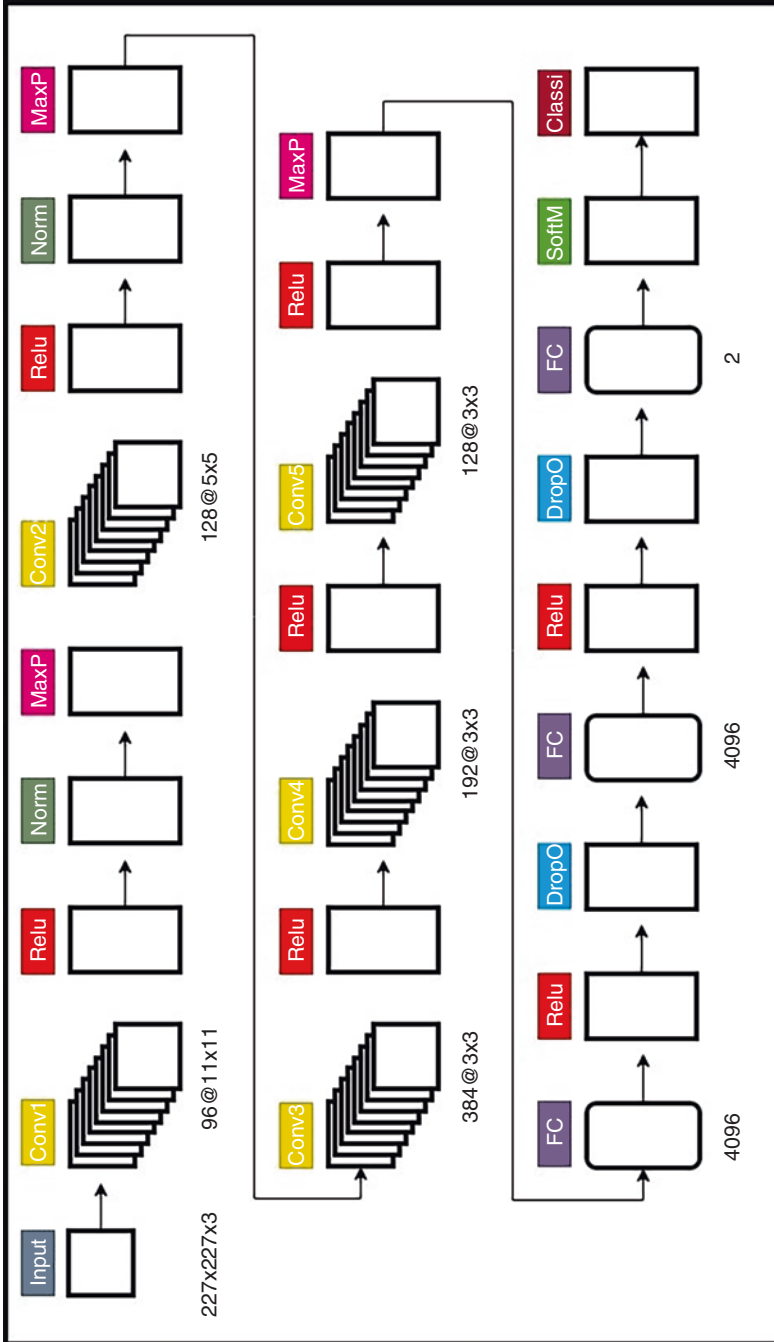


Fig. 9.4 Proposed CNN Architecture

Table 9.2 Proposed features for SVM Classifier Training

<i>Time-Domain Features</i>					
Interval/ Interval Ratio	Features				
	Mean			Standard Deviation	
R-R	X			X	
S1	X			X	
S2	X			X	
Systole	X			X	
Diastole	X			X	
Systole, R–R ratio	X			X	
Diastole, R–R ratio	X			X	
Systole, diastole ratio	X			X	
Systole period, S1 period ratio	Mean absolute amplitude ratios			Mean absolute amplitude ratios	
Diastole period, S2 period ratio	Mean absolute amplitude ratios			Mean absolute amplitude ratios	
<i>Frequency-Domain Features</i>					
Segments/ Segments Ratio	Frequency Band	Features			
		Power	Band Power	Amplitude	Q-factor
Cardiac Cycle	150–350 Hz 200–400 Hz	Mean	Mean	–	–
Systole	150–350 Hz 200–400 Hz	Mean	Mean	–	–
Diastole	150–350 Hz 200–400 Hz	Mean	Mean	–	–
S1		Mean	Mean	–	X
S2		Mean	Mean	–	X
Systole		Mean	Mean	Mean	X
Diastole		Mean	Mean	Mean	X
S1		Mean	Mean	Mean	X
S2		Mean	Mean	Mean	X
Cardiac Cycle, systole ratio	100–300 Hz 200–400 Hz	Average	–	–	–
Cardiac Cycle, diastole ratio	100–300 Hz 200–400 Hz	Average	–	–	–
Diastole, systole ratio	100–300 Hz 200–400 Hz	Average	–	–	–

9.7.1 Dataset

The proposed methodology is implemented on the PhysioNet2016 challenge dataset [41]. In this dataset, a total of 3226 data samples of PCG recording both from healthy and pathological patients were collected. 1610 instances of the total data samples are used for training and 1616 samples are used for testing. The physioNet dataset consists of multiple sub-datasets including A, B, C, D, E, and F. The audio recordings present in each sub-dataset are 409, 488, 29, 53, 2137, and 110 respectively. A 50-50% split training-testing strategy is incorporated in this study. The actual number of recordings used for training are 204, 244, 14, 26, 1067, and 55 and for testing 205, 244, 15, 27, 1070, and 55 respectively. For segmentation, “S1” and “S2” labels for each PCG signal sample are obtained using the Springer Algorithm [9] which are compared with the segmentation results of the developed method.

9.7.2 Results of Pre-processing

After obtaining the suitable (classifiable) signals the preprocessing techniques are performed. First, the signal is passed through the low pass and high pass filters. Furthermore, the removal of spikes is carried out using a Schmidt spike removal function. Followed by the extraction of the envelope using the Hilbert transform. Finally, the resultant signal is normalized using simple mean and standard deviation formula. The results of different preprocessing operations are illustrated in Fig. 9.5. Figure 9.5a shows a smaller chunk of the original (classifiable) signal. Figure 9.5b shows the signal obtained after filtering and spike removal operations. It is clear that after preprocessing the signal became smooth and spikes were removed. Furthermore,

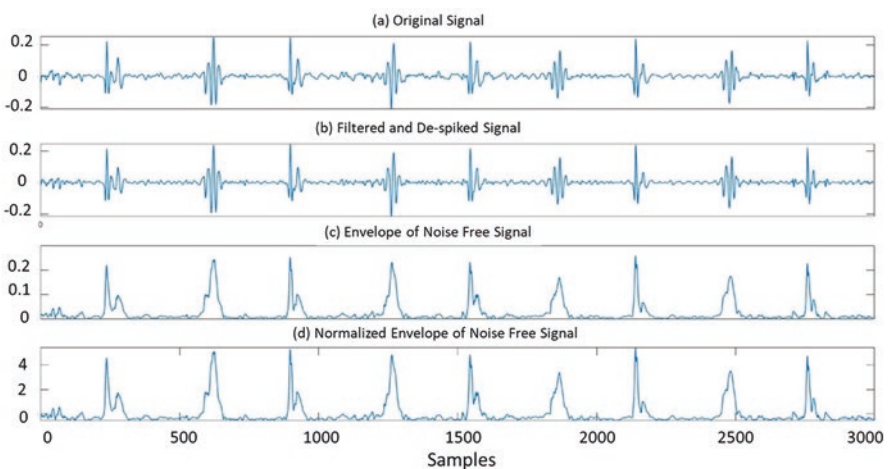


Fig. 9.5 Pre-processing results

Fig. 9.5c shows the obtained envelope of the preprocessed signal which contains useful information regarding the “S1” and “S2” activities. Finally, a normalized envelope of the processed signal is obtained as shown in Fig. 9.5d.

9.7.3 Results of Segmentation

In this stage, true peaks are extracted from the preprocessed signal. Firstly all peaks are identified using peak finders which are called the candidate peaks for the “S1” and “S2”. Afterward, a multi-level threshold is employed for the identification of the true peaks. Consequently, true peaks are selected from the pool of peaks which are sent to the feature extractor to classify them as “S1” and “S2” based on their features. The results of the peak finder stage illustrating the candidate peaks are given in Fig. 9.6a. Subsequently, the candidate obtained using the starting threshold is given in Fig. 9.6b. Afterward, true peaks are detected and false peaks are eliminated using the developed multi-level threshold, as shown in Fig. 9.6c.

Classification of peaks is performed using a spectrogram of obtained true peaks. The windowing procedure is applied to obtain the true peak, and its spectrogram is obtained using Short Time Fourier Transform. These spectrograms are fed as input to the convolutional neural network which learns features in 100 iterations and trains its model. Final testing of the model is performed on unseen samples of spectrogram which classify them as “S1” and “S2”. The spectrogram and CNN combination gives an overall segmentation accuracy of 91.20% as compared to SVM and ANN classifiers. Figure 9.7 illustrates the peak identification results where the true peaks are labeled based on the peak spectrograms.

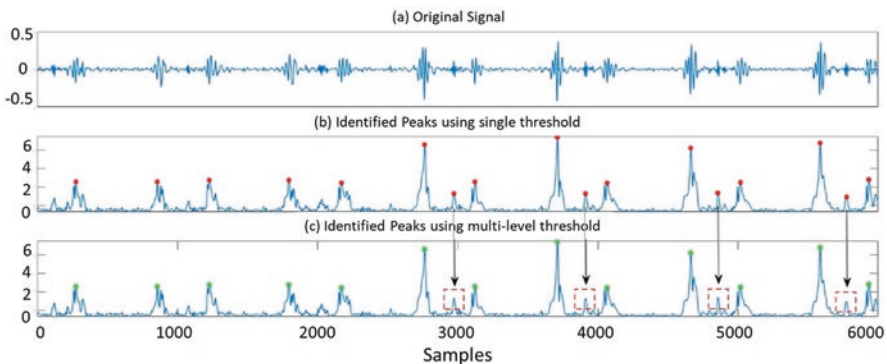


Fig. 9.6 True peak detection results based on multi-level threshold

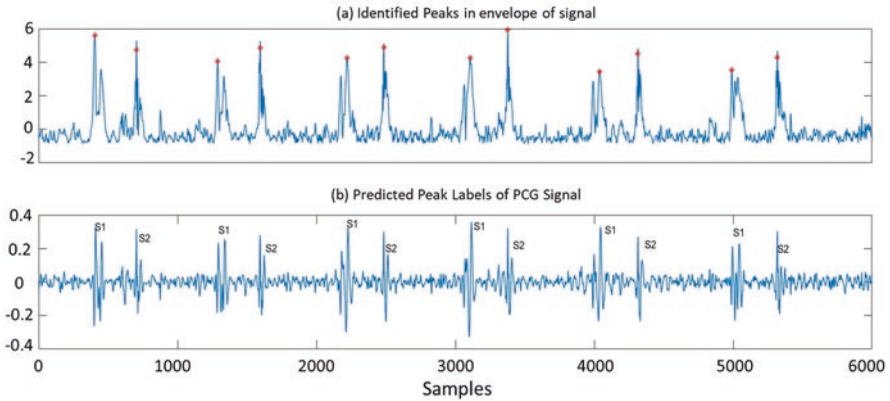


Fig. 9.7 Predicted labels using peak spectrogram and convolution neural network

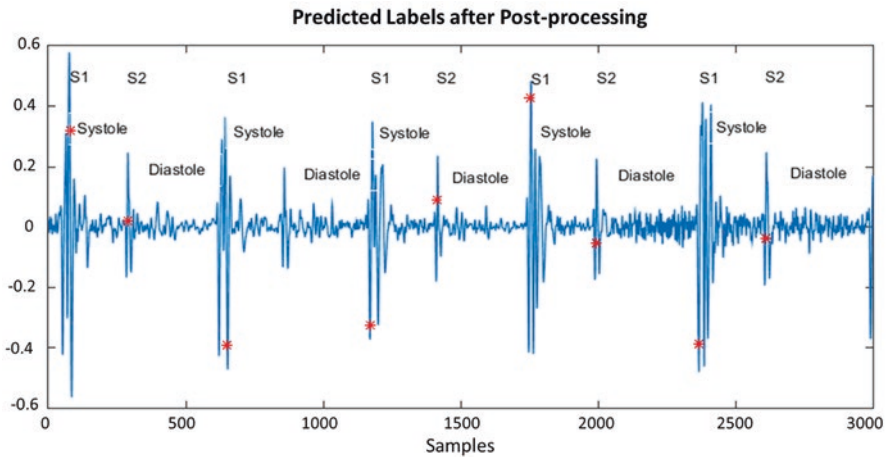


Fig. 9.8 Predicted label after post-processing

9.7.4 Results of Post-processing

After the assignment of “S1” and “S2” peaks by the classifier, a post-processing step is required. This completes one cycle as “S1”, “systole”, “S2”, and “diastole” as shown in Fig. 9.8.

9.7.5 Results of PCG Segmentation

The state sequences (i.e., “S1”, systole, “S2”, diastole) obtained after post-processing are forwarded to the final stage in which features are extracted based on the intervals between the states, their ratios, mean, standard deviation, amplitudes,

Table 9.3 Segmentation results on the PhysioNet 2016 challenge dataset

Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)
CNN (Proposed)	91.20	94.05	88.53
SVM	87.57	92.08	83.08
ANN	82.01	87.33	76.93

Table 9.4 Comparison Analysis on the PhysioNet 2016 challenge dataset

Approach	Technique	Score/ Accuracy
Proposed Methodology I	SVM without spectrogram	86.89
Proposed Methodology II	SVM with spectrogram	93.33
Classifying Heart Sound Recording Using Deep Convolutional Neural Network [18]	CNN	88%
Morphological Determination of Pathological PCG Signal by Time and Frequency Domain Analysis [36]	SVM	81%
PCG Classification using Neural Network Approach [16]	ANN	79%

and other power-energy features. The classifier employed for this purpose is the SVM which performs the binary classification. In another set of experiments, PCG signal segmentation was performed using the developed spectrogram and CNN-based segmentation approach. Afterward, PCG signal classification is performed using an SVM classifier and the classification result are shown in Table 9.3 which shows that with the incorporation of spectrogram and CNN combination the PCG classification accuracy improves.

9.8 Comparison Analysis and Discussions

In this section, a comparison study is presented between the proposed approach and the state-of-the-art counterparts on the same datasets i.e., PhysioNet 2016, as discussed in Sect. 9.7. Comparison results with other classifiers (i.e., Artificial Neural Networks, Convolutional Neural networks, and Support Vector Machine) are also shown in Table 9.4. In [16] PCG signal is classified using Hidden Markov Model (HMM) for extracting features and ANN focusing only on the statistical features which results in an accuracy of 79%. Support Vector Machine with Time and Frequency domain features is employed in [38] for analysis of PCG signals and it obtained an accuracy of 81%. In [18] the emphasis is on the state-of-the-art CNN, and this approach leads to an accuracy of 88%. In this chapter, two different classification approaches were presented. In the Proposed Methodology I, PCG signals were analyzed under the category of segmented approach using a multi-level threshold, extracting the time and frequency domain features along with the Support

Vector Machine for final classification with the accuracy of 86.89%. In the Proposed Methodology II, the same strategy was followed except that segmentation was carried out using a CNN along with the SVM as a final classifier, resulting in the highest accuracy of 93.33%.

9.9 Conclusions

Phonocardiograms are the non-stationary signals which make the task of identifying the exact location of the peaks difficult. Using the multi-threshold method for peak detection and with the peak spectrogram, the identification of the peak locations was improved. Also, PCG signal classification into normal and abnormal was improved to 93.33% with our developed method. The segmentation and classification results reported for our developed approach using peak spectrogram and state-of-the-art convolutional neural network have an accuracy of 91% and 93.33% respectively.

Classification of PCGs can be bettered by calculating advanced features which extract the information of signal in more depth or by using deep learning models. In this paper, separate methods are used for PCG segmentation and classification. Nonetheless, there can be a possibility to use a unified framework for segmenting and classifying PCG both. In addition, instead of binary classification non-binary classification can be performed to find out the exact CVD in future endeavors.

References

1. H. Mehrtash, R. Laing, V.J. Wirtz, Comparative analysis of essential medicines for cardiovascular diseases in countries of the WHO Eastern Mediterranean region. *East. Mediterr. Health. J.* **24**(5), 427–434 (2018)
2. M.W. Arnaudin, J.J. Mintzes, Students' alternative conceptions of the human circulatory system: A cross-age study. *Sci. Educ.* **69**(5), 721–733 (1985)
3. A.J. Weinhaus, K.P. Roberts, Anatomy of the human heart, in *Handbook of Cardiac Anatomy, Physiology, and Devices*, 2nd edn., (2005), pp. 59–85
4. Q. Mubarak, M.U. Akram, A. Shaukat, A. Ramazan, Quality Assessment and Classification of Heart Sounds Using PCG Signals. no. February, (2018), pp. 1–11
5. N.K. Al-Qazzaz, I.F. Abdulazez, S.A. Ridha, Simulation recording of an ECG, PCG, and PPG for feature extractions. *J. Al-Khwarizmi Eng. J.* **10**(4), 81–91 (2014)
6. S.I. Malik, M.U. Akram, I. Siddiqi, Localization and classification of heartbeats using robust adaptive algorithm. *Biomed. Signal Process. Control* **49**, 57–77 (2019)
7. S. Ari, K. Hembram, G. Saha, Detection of cardiac abnormality from PCG signal using LMS based least square SVM classifier. *Expert Syst. Appl.* **37**(12), 8019–8026 (2010)
8. C.D. Papadaniil, L.J. Hadjileontiadis, Efficient heart sound segmentation and extraction using ensemble empirical mode decomposition and kurtosis features. *IEEE J. Biomed. Heal. Informatics* **18**(4), 1138–1152 (2014)
9. D.B. Springer, L. Tarassenko, G.D. Clifford, Logistic regression-HSMM-based heart sound segmentation. *I.E.E.E. Trans. Biomed. Eng.* **63**(4), 822–832 (2016)

10. F. Chakir, A. Jilbab, C. Nacir, A. Hammouch, A. Hajjam, Detection and identification algorithm of the S1 and S2 heart sounds, in *International Conference on Electrical and Information Technologies (ICEIT)*, (2016), pp. 418–420
11. E.F. Gomes, P.J. Bentley, M. Coimbra, E. Pereira, Y. Deng, Classifying heart sounds: Approaches to the PASCAL challenge, in *Heal. 2013 – Proceedings of the International Conference Heal. Informatics*, (2013), pp. 337–340
12. K.N. Khan, F.A. Khan, A. Abid, T. Olmez, Z. Dokur, Deep learning based classification of unsegmented phonocardiogram spectrograms leveraging transfer learning. *Physiol. Meas.* **42**, 095003 (2021)
13. Z. Dokur, T. Ölmez, Feature determination for heart sounds based on divergence analysis, in *19th Digital Signal Processing*, (2009), pp. 521–531
14. B. Bozkurt, I. Germanakis, Y. Stylianou, A study of time-frequency features for CNN-based automatic heart sound classification for pathology detection. *Comput. Biol. Med.* **100**, 132–143 (2018)
15. K.A. Babu, B. Ramkumar, M.S. Manikandan, S1 and S2 heart sound segmentation using variational mode decomposition, in *IEEE Region 10 Conference (TENCON)*, (2017), pp. 1629–1634
16. I. Grzegorzczak, M. Soliński, M. Łepeck, A. Perka, J. Rosiński, J. Rymko, K. Stępień, J. Gierałowski, PCG Classification Using a Neural Network Approach, in *Conference on Computing in Cardiology*, (2016), pp. 1–4
17. J. Gonz, C.P. Phoo, J. Wiens, Heart Sound classification based on temporal alignment techniques, in *Computing in Cardiology Conference (CinC)*, (2016), pp. 1–4
18. J. Rubin, R. Abreu, A. Ganguli, S. Nelaturi, I. Matei, K. Sricharan, Classifying heart sound recordings using deep convolutional neural networks and mel-frequency cepstral coefficients, in *Computing in Cardiology Conference (CinC)*, (2016), pp. 1–4
19. Z. Xiong, T. Liu, G. Tse, M. Gong, P.A. Gladding, B.H. Smaill, M.K. Stiles, A.M. Gillis, J. Zhao, A machine learning aided systematic review and meta-analysis of the relative risk of atrial fibrillation in patients with diabetes mellitus. *Front. Physiol.* **9**, 835 (2018)
20. S.E. Schmidt, C. Holst-Hansen, C. Graff, E. Toft, J.J. Struijk, Segmentation of heart sound recordings by a duration-dependent hidden Markov model. *Physiol. Meas.* **31**(4), 513–529 (2010)
21. J. Park, K. Lee, K. Kang, Arrhythmia detection from heartbeat using k-nearest neighbor classifier, in *IEEE International Conference on Bioinformatics and Biomedicine*, (2013), pp. 15–22
22. F. Safara, S. Doraisamy, A. Azman, A. Jantan, A.R.A. Ramaiah, Multi-level basis selection of wavelet packet decomposition tree for heart sound classification. *Comput. Biol. Med.* **43**(10), 1407–1414 (2013)
23. J. Li, L. Ke, Q. Du, Classification of heart sounds based on the wavelet. *Entropy* **21**(5), 472 (2019)
24. S. Patidar, R.B. Pachori, Classification of cardiac sound signals using constrained tunable-Q wavelet transform. *Expert Syst. Appl.* **41**(16), 7161–7170 (2014)
25. J.M. Keller, M.R. Gray, J.A. Givens, A fuzzy k-NN neighbor algorithm. *IEEE Trans. Syst. Man Cybern.* **SMC-15**(4), 580–585 (1985)
26. S.K. Randhawa, M. Singh, Classification of heart sound signals using multi-modal features. *Procedia Comput. Sci.* **58**, 165–171 (2015)
27. H.M. Hadi, M.Y. Mashor, M.Z. Suboh, M.S. Mohamed, Classification of heart sound based on s-transform and neural network, in *10th International Conference on Information Science, Signal Processing and their Applications*, (2010), pp. 189–192
28. A.A. Zo-Afshan, F. Hussain, Localization of phonocardiogram signals using multi-level threshold and support vector machine, in *2019 IEEE 19th International Symposium on Signal Processing and Information Technology, ISSPIT 2019*, (2019), pp. 1–5
29. F.A. Khan, A. Abid, M.S. Khan, Automatic heart sound classification from segmented/unsegmented phonocardiogram signals using time and frequency features. *Physiol. Meas. Pap.* **41**, 055006 (2020)

30. F. Demir, A. Şengür, V. Bajaj, K. Polat, Towards the classification of heart sounds based on convolutional deep neural network. *Health Inf. Sci. Syst.* **7**(16) (2019)
31. H. Alaskar, N. Alzhrani, A. Hussain, F. Almarshed, The implementation of pretrained AlexNet on PCGClassification, in *International Conference on Intelligent Computing*, (2019), pp. 784–794
32. K. Wołk, A. Wołk, Early and remote detection of possible heartbeat problems with convolutional neural networks and multipart interactive training. *IEEE Access* **7**, 145921–145927 (2019)
33. R. Yamashita, M. Nishio, T.K. Do RKG, Convolutional neural networks: an overview and application in radiology. *Insights Imaging* **9**, 611–629 (2018)
34. S.A. Singh, S. Majumder, Classification of unsegmented heart sound recording using KNN classifier. *J. Mech. Med. Biol* **19**, 1950025 (2019)
35. C. Potes, S. Parvaneh, A. Rahman, B. Conroy, Ensemble of feature: Based and deep learning: Based classifiers for detection of abnormal heart sounds, in *2016 Computing in Cardiology Conference (CinC)*, vol. 43, (2017)
36. M.A. Goda, P. Hajas, Morphological determination of pathological PCG signals by time and frequency domain analysis. *Comput. Cardiol.* **2010**(43), 1133–1136 (2016)
37. P. Langley, A. Murray, Abnormal heart sounds detected from short duration unsegmented phonocardiograms by wavelet entropy. *Comput. Cardiol.* **2010**(43), 545–548 (2016)
38. G. Clifford, C. Liu, D. Springer, B. Moody, Q. Li, R. Abad, J. Millet, I. Silva, A. Johnson, R. Mark, Classification of normal/abnormal heart sound recordings: The PhysioNet/computing in cardiology challenge 2016, in *2016 Computing in Cardiology Conference (CinC)*, vol. 43, (2017), pp. 3–6
39. S. Vernekar, S. Nair, D. Vijaysenan, R. Ranjan, A novel approach for classification of normal/abnormal phonocardiogram recordings using temporal signal analysis and machine learning. *Comput. Cardiol.* **2010**(43), 1141–1144 (2016)
40. A. Bouril, D. Aleinikava, M.S. Guillem, G.M. Mirsky, S. Llc, U.P. De València, Automated classification of Normal and abnormal heart sounds using support vector machines. *Comput. Cardiol.* **43**, 549–552 (2016)
41. G.D. Clifford, I. Silva, B. Moody, Q. Li, D. Kella, A. Shahin, T. Kooistra, D. Perry, R. Mark, The PhysioNet/computing in cardiology challenge 2015: Reducing false arrhythmia alarms in the ICU. *Comput. Cardiol.* **2010**(42), 273–276 (2015)