This is the accepted manuscript published in Computer Science Review vol 41, August 2021, 100420.
https://doi.org/10.1016/j.cosrev.2021.100420

PCG Signals for Biometric Authentication Systems: An In-depth Review

El-Sayed. A. El-Dahshan[1, 2] Mahmoud. M.Bassiouni[2], Septavera S. Sharvia[3], and Abdelbadeeh M. Salem[4]
[1]Department of Physics, Faculty of Science, Ain Shams University, Postal Code: 11566, Cairo, Egypt
[2]Egyptian E-Learning University (EELU), 33 El-messah Street, Eldoki, Postal Code: 11261, El-Giza, Egypt
[3]School of Engineering and Computer Science, University of Hull, Hull HU6 7RX, U.K
[4]Faculty of Computer and Information Science, Ain Shams University, Abbassia, Postal Code: 11566, Cairo, Egypt
e_eldahshan@yahoo.com, mbassiouni@eelu.edu.eg, s.sharvia@hull.ac.uk, abmsalem@yahoo.com

Highlights

- A review presented based on research studies that used heart sound signals as a biometric from inception to current status.
- A survey explaining the main common phases used in the PCG biometric systems
- Illustrating the main advantages, disadvantages, classification of the techniques used in each phase.
- The limitations, challenges, future work of the PCG biometric system are demonstrated in detail.

Abstract

This review aims to present a survey of the technologies and methodologies used in the phonocardiogram (PCG) biometric systems. The phases used in the PCG which are explored in this paper include data acquisition, de-noising, extracting PCG peaks, feature extraction, feature reduction, classification, and evaluation. As part of this study, we performed a systematic review that summarizes the well-known approaches used in PCG since its inception to the current status. Out of 157 manuscripts available in the academic databases from 2006 until 2020, 35 primary studies focused on "heart sounds" like a biometric and this is related to the objective of this research. Out of those studies, 11 matched the inclusion criteria. The estimation performance of these systems is close to an acceptable level in the consideration of PCG as a biometric. The use of PCG signals is a promising field. Finally, the limitations and some future works are discussed.

Keywords: Cardiac sound signals, Biometrics, Authentication, Machine learning.

1. Introduction

The process of identifying the correct person before the release of the secure resources is known as authentication. To achieve this a counter configuring unique information is obtained by the person. These types of information can be divided into three main types which are knowledge, by using username or password; tokens, focusing on the personal identification number (PIN), and based verification [1]. Biometric authentication [2] is an automatic method that identifies or verifies the user depending on the measurement of his or her unique physiological traits such as face [3], palm[4], iris[5], etc or behavioral traits such as keystrokes dynamics [6], voice [7], signature [8]. Physiological biometrics related to natural biological features and behavioral biometrics are those traits that are naturally grown to and behavioral biometrics are those which are learned.

The process of recording cardiac vibrations is known as phonocardiography. The control that can be made on the flow of the blood is managed by two sets of valves. These sets are the atrioventricular and semilunar valves, and the first valve opens to let the blood flow in the heart while the second value opens to let the blood flow out of the heart. Those known as the sounds of the heart are sometimes called heart sounds or phonocardiogram (PCG). As often noted in the literature, PCG or heart sounds biometric is attractive because of the following advantages which are not shared with another type of biometrics.

In Fig. 1 the radar plot presents a comparison between different biometrics using the seven parameters of measurements. Those seven measurements are universality, uniqueness, permanence, measurability, performance, acceptability, and circumvention. The total score for appropriateness is represented by the outermost right of the radar plots. The values of each parameter range from 0 (lowest) to 10 (highest). The area of each radar plot can be computed by dividing the shape into a set of triangles with the same angles [9]. The area of the plot indicates the strength of the biometric techniques. For example, for ECG, DNA, Face, and Iris, score the highest in four out of seven measurement categories. In the case of ECG biometric, it has high universality, uniqueness, permanence, performance, and Face as a biometric has high universality, acceptability, collectability and circumvention, and so on other biometrics. This is followed by the Ear, Gait, Fingerprint biometric as they have a maximum of three measurements out of seven. After that, Signature and palmprint have a maximum of two measurements out of seven. Voice, PCG, and keystrokes have the lowest scores, as voice and PCG have a maximum of one favorable measurement out of seven, and keystrokes are the weakest biometric as it scores poorly on almost all areas. In comparison, ECG studies the electrical mechanism of the human heart which consists of a set of peaks which are P, Q, R, S, and T [10]. Fiducial dependent and independent (non-fiducial) are two types of ECG recognition approaches that can be highlighted [11].

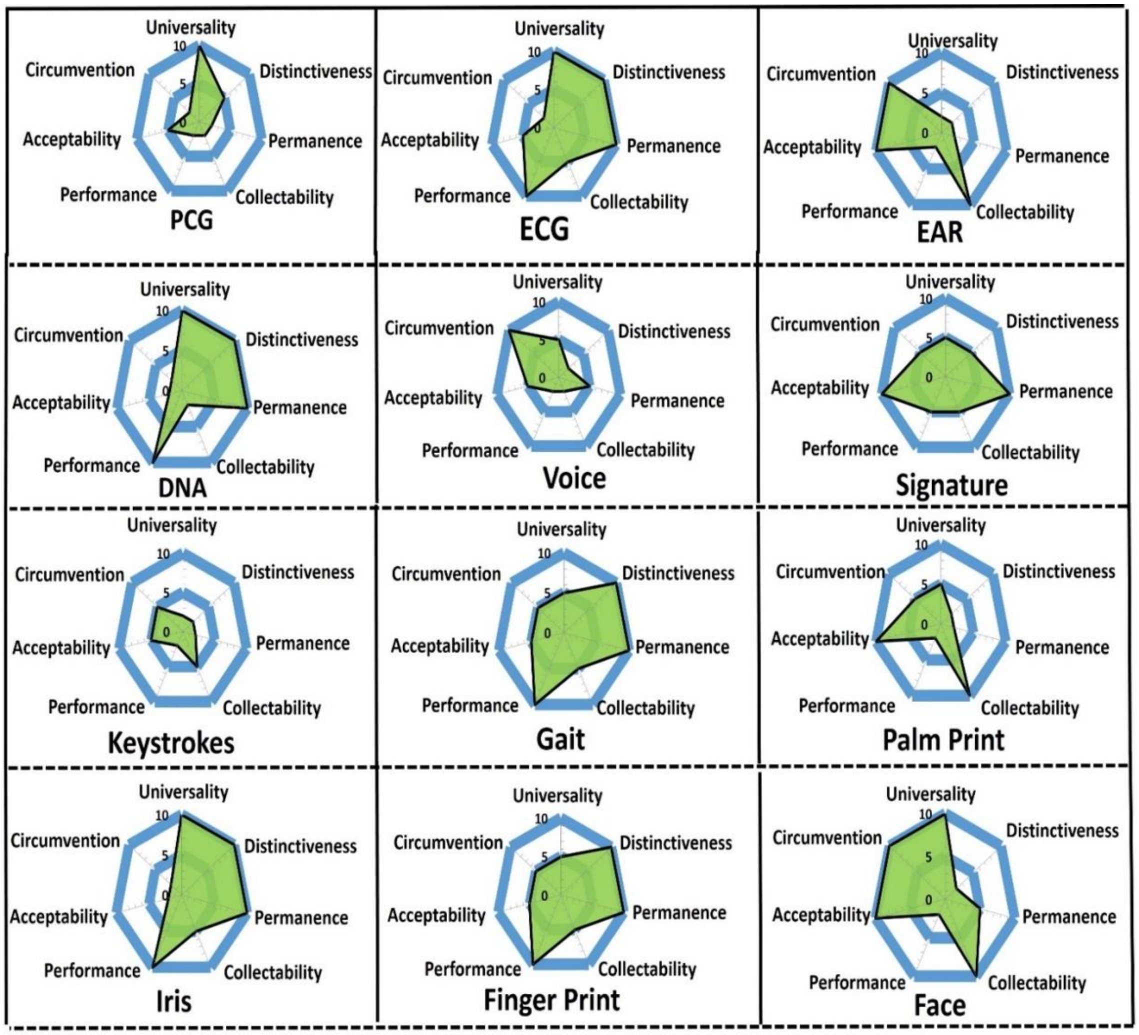

Fig. 1. Radar plots comparison between physiological and behavioral biometrics on seven variables.

While the fiducial dependent methods utilize temporal or spatial features within some points in the ECG heartbeat, the fiducial independent methods utilize the signal holistically. On the other hand, heart sounds are a set of bursts and variations of varying intensity and frequency, and it consists of two main sounds which are S1 and S2 when the heart contracts and pumps blood.

The PCG signal can be formed through the contribution of blood flow and the opening or closing of the valves. Finally, PCG has been lately qualified for heart diagnostics and biometrics. PCG and ECG are the main signals that are captured from the heart, and they have their biometric modality. In other words, ECG and PCG have high universality, and they have a medium value for acceptability. and collectability. ECG has high percentages of distinctiveness, permanence, and performance with a medium level of collectability, but it has a low level of circumvention. On the other hand, PCG has an average level of distinctiveness and a low level in all the remaining measurements. PCG can be deemed to be accepted as a biometric more than other types such as key strokes and Voice traits. We found that data acquisition is based on multiple online datasets for PCG as a biometric.

The preprocessing depends on six types of filtrations methods down-sampling, low and high pass filtering, energy thresholding, normalization; total variation de-noising, and wavelet-based de-noising. While the segmentation is based on four types, for example, framing and windowing, autocorrelation, Shannon energy envelop, and zero-crossing rate with short-term amplitude. The feature extraction is categorized into five types of domain methods such as time, frequency, time-frequency, time scale, and fusion between different features. Finally, the classification depends on five types of classifiers defined by statistical, similarity, pattern recognition, neural network, and other approaches. The structure of the paper is as follows. Section 1 presents the introduction and outlines the background on biometric techniques.

The contributions, search criteria, and the background on the heart and PCG signals are discussed in sections 2, 3, and 4 respectively. Section 5 examines the different approaches employed in PCG Signals which include data acquisition, pre-processing, segmentation, feature extraction, and classification. Section 6 presents the performance measurements, section 7 shows the discussion, open issues, limitations, and future directions, and section 8 illustrates the applications and finally, section 9 presents the conclusion discrimination.

2. Heart and its Sound Signal

Four main chambers compose the human heart [12]. These are the right and left atrium and right and left ventricle as shown in Fig.2. The heart's right part is much smaller with less myocardium in their heart wall. Based on the two circulatory loops' size, there is a variation and difference in the main function of the left and right sides. Blood is pumped to the extremities of the human body by the heart's left side, while the human heart's right side operates as a pulmonary circulation to the lungs [12]. The functionality mechanism generates acoustic signals and vibrations that can be gained and obtained over the wall chest. The normal heart sounds are defined as "lubb" and "dupp", and they are caused by pushing blood on the valves of the human heart as shown in Fig. 3.

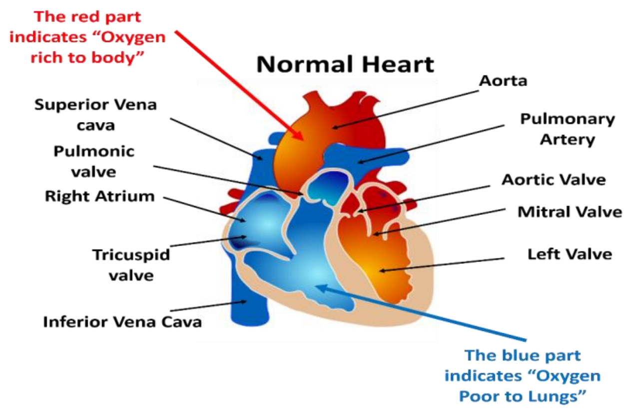

Fig. 2. Normal Heart Structure.

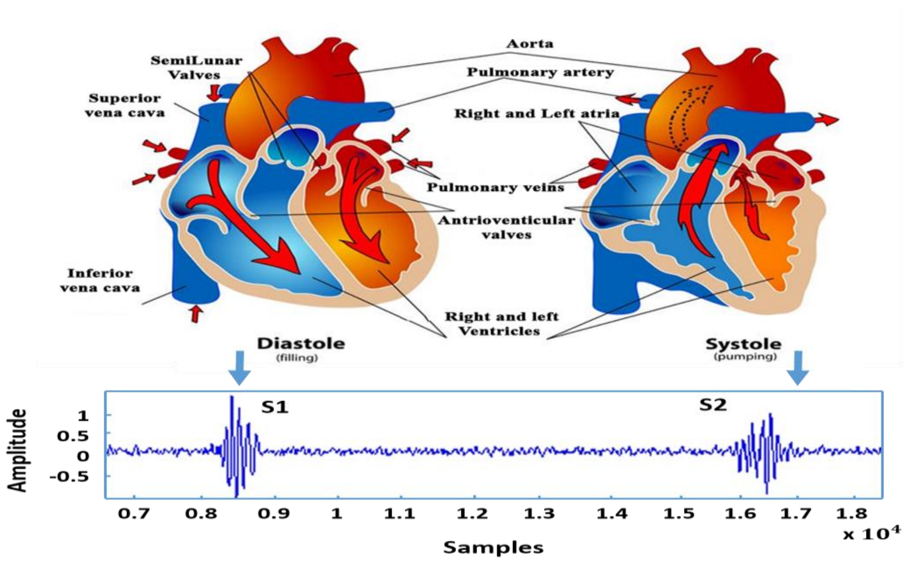

Fig 3. (a) The diastole (S1) and systole (S2) cycle in the heart [14].

The "lubb" sound is known by S1 and it comes first in the human heartbeat and it is longer than that of the two heart sounds. Lubb is generated by closing the AV valves that are located at the beginning of the ventricular systole. The "dupp" sound is also known as S2 and it comes next in the human heartbeat, and it is a shorter and sharper sound that is resulted from the closing of the semilunar valves at the end of the ventricular systole. The pattern of S1-S2 or lubb-dupp is repeated in the heart. Some problems or different sounds such as gurgling or liquid rushing in the heart may indicate problems in the heart that cause defects in the ventricular or atrial or leakage in the valves. Heart sounds can be used in a lot of applications for diagnosis or biometrics. In this survey, we will explore the role of heart sounds as a biometric technique, studying its role in the identification and verification processes [13]. Heart sounds can be used as a biometric-based on a lot of advantages. The most important merit is that the heart sound can be captured only from a living human body, therefore, it is difficult to forge or steal someone's PCG [15, 16]. Heart sounds are measured non-invasively in their entirety while being socially acknowledged. It is quantifiable and has uniqueness and vulnerability [17]. The heart sounds biometric system's main advantage, however, is the high security it provides because the sounds of a person's heart cannot be forged. PCG signals have several characteristics which support this. For example, heart sounds are one-dimensional signals, and heart sounds are easy to process and have low-frequency characteristics.

3. History of PCG Biometric

Robert Hooke (1635-1703), is deemed to be the first English polymath to determine the power potential and the diagnostic of cardiac Auscultation. Afterward, **Ren Laennec** (1781 -1826) discovered the stethoscope in 1816, and it has become a fundamental tool for clinics and other usages as it remains to this day.

In 2006, Phua et al., [18] introduced the heart acoustic signals as a biometric method in a workshop event on multimodal user authentication. In 2007, heart sounds were first used for human recognition by Beritelli et al., [19] proposed an intelligent system for 20 participants based on z-chirp transform (CZT) for feature extraction, and the Euclidian distance (ED) was used for classification as a matching technique. The result of the implemented system showed a rejection rate of 5% and a false acceptance of 2.2%.

In 2008 Phua et al., [20] proposed an experiment based on linear frequency banks cepstral (LFBC) to identify individuals based on PCG, and they built an authentication system using two classifiers which are GMM “Gaussian mixture models” and VQ “vector quantization”. They worked on 10 user's achieving an accuracy of 96%. Beritelli et al., [21] developed an approach based on sub-band aggregation for verification. The PCG signal is segmented to obtain S1 and S2 from the PCG signal. S1 and S2 are passed to CZT transform separately obtaining sub-band selection from S1 and S2. A matching algorithm based on ED was used to compare the test and the template and this produced an equal error rate (EER) of 10 % on 70 subjects.

For the improvement of the accuracy of the human recognition system, Beritelli et al., [22] 2009 resumed their work in PCG as a biometric. A larger number of subjects consisting of 50 people were the basis of their work. Additionally, 13 Mel frequency cepstral coefficients extracted (MFCC) from the basic PCG heart sounds, S1 and S2, were employed. Finally, the First-to-Second ratio (FSR) was applied for reaching a 9% EER rate. The previous algorithms presented in [22] were applied by Beritelli et al., [23] on 40 people. The system revealed a 5% EER, and they offered their explanations for the impact of the increase of the test set to 80. However, no negative impact because of this was perceived on EER.

In 2010 Fateman et al., [24, 25] showed a verification and identification system for heart sounds. These two main systems worked on 21 subjects, and they used wavelets for filtration of the PCG. Moreover, for feature extraction and classification, they employed short-time Fourier transform (STFT) and linear discernment analysis (LDA) respectively. The results revealed that the identification system reached 100% and 33% EER for the verification system. Again, Beritelli et al., [26] utilized a method that is focused on Gaussian Mixture Models (GMM) for classification, and features are extracted from spectral and time domain for improving the performance. The results were 13.70% on a dataset of 165 people that outperforms other similar approaches. Beritelli et al., [27] implemented two approaches statistical and non-statistical for PCG signals as a biometric. The latter approach yields an EER of 29.08% and the former approach yields 15.53% based on GMM on 147 subjects.

An application of a proposed system based on obtaining the cycle-power-frequency-drawing combined with the D-S information was given by Tao et al., [28] for the determination of the identity verification system depending on PCG signals working on a dataset of 5 – 100 people to reach 99% accuracy. Huy et al., [29] Proposed a feature extraction based on 8 feature sets. These sets are based on spectral, temporal, harmonic, cepstral, cardiac, rhythmic, and GMM features. Then, from these features sets two features, sets were selected, and then they are passed to the support vector machine (SVM) for classification. The results of the two feature sets were applied to 52 users. The results of the identification of the first experiment reached an accuracy of 80% and the second one achieved higher than 90%.

The linear prediction cepstral coefficient (LPCC) was applied as a feature extraction method by Guo et al., [30]. Then, the hidden markov model (HMM) with wavelet neural network (WNN) was used as a classification method for the PCG signals for the verification of the identity of 80 people. Based on wavelet transform (WT), Jasper et al., [31] gave an analysis of the PCG signals in the time-frequency domain. Then, there was an extraction of the Shannon energy envelogram

(SEE) as the feature set. Depending on a database of 10 people, the performance was deemed acceptable reaching 98.67%.

In 2011 Cheng Xie Feng [32] developed a PCG identification system applying improved circle convolution (ICC) as feature extraction, with the combination of independent sub-band function (ISF). The recognition of the PCG depended on S1 and S2 sounds. Each sound was entered separately through the feature extraction used to ensure validity. They used the similarity distance to verify different human heart sound features. This method was verified using 10 records of heart sounds and the performance results showed an accuracy of 85.7% in both modes. An identification system based on MFCC and VQ was introduced by Zhao et al., [33] for both feature extraction and classification respectively. The numbers of subjects used were 30 subjects achieving 100% accuracy.

In 2012, a PCG verification system was developed by Cheng Xie Feng [34] which used the linear band frequency cepstra (LBFC) as feature extraction and the similarity distance for classification. Their proposed system worked on 12 PCG signals reaching an error acceptance rate of 1% - 8%, verification rate of 95%, and error rejection rate below 3%. Rasha Wahid et al., [35] proposed a verification system based on two main feature extraction methods, the first one reached an accuracy of 100%, and the second one reached 85% final accuracy. The classification was done using GMM for both feature extraction techniques. Chen W et al., [36] applied wavelet and MFCC for the implementation of a PCG biometric system for filtering and feature extraction. Afterward, a feature reduction step was taken based on principle component analysis (PCA) and a 90% recognition rate was verified by the results. Karmaker et al., [37] used an approach based on time and frequency domain. The de-noising phase was done using Butterworth low pass filter, and then the segmentation was performed for S1 and S2. In the feature extraction phase, the db2 wavelet was applied to extract the detail coefficients. The average energy was determined from S1, and the same was done for S2. Finally, the MLP was used for classification achieving an accuracy of 96.178%, In 2013, there was a proposed biometric system by Zhong L et al., [38] based on PCG signal and with the dependence on GMM as a classification combined with cepstral coefficients. MFCC and linear predication cepstral coefficients (LPCC) were the kinds of cepstral coefficients used. The system was based on 100 heart sounds from 50 people to classify the PCG signals. Spadaccini et al., [39] applied an identification system based on PCG signals. In the de-noising phase, there was a detection of the

S1 and S2 from the endpoints of the PCG signal, and a computing of the (FSR) of the S1 and S2 sound signal. The structural and statistical systems were the two biometric systems implemented and tested. Based on linear frequency cepstral coefficient (LFCC) combined with FSR and the GMM, the statistical system was used for classification, while MFCC and FSR were used by the structural system for feature extraction and Euclidean distance (ED) for matching. A database of 206 individuals was used for applying the results which reached 13.66% for the statistical system.

Zhao et al., [40] showed a PCG biometric system working on 40 subjects using marginal spectrum analysis algorithm as feature extraction and VQ for classification. The results showed an identification rate of 94.4%, and when the number of subjects increased and reached 80 the accuracy down sampled to 92%. Girish et al., [41] implemented a method for the identification of the PCG signals. The method was based on a low pass filter for de-noising. The segmentation of the heart sound was based on thresholding and framing. While the feature extraction was done based on the wavelet decomposition and the classification was done using the MLP ANN. The results were verified on 10 volunteers reaching an EER of 9.48% and accuracy of 90.52%

An approach for PCG as a biometric tool was proposed by Tan et al., [42] in 2014 with a set of steps. Preprocessing is the first step based on low-pass filtering. The second step is the segmentation of the heart sounds S1 and S2 based on the zero crossing rate (ZRC) technique followed by short-term amplitude (STA). MFCC was used for extracting features from S1 and S2, and the classification was done using a sparse representation classifier (SRC). The approach was applied on 15 subjects selected randomly with an accuracy of 85.45% achieved. A system was proposed by Abo el zahad et al., [43] for PCG signal identification depending on LFCC, MFCC, DWT, and Bark frequency cepstral coefficient (BFCC). Then, conical correlation analysis (CCA) was used for applying the fusion between the features, while Bayes rule and GMM were used for classification. The work was implemented on 17 participates achieving a performance of 99%. The first one who applied the fusion concept in PCG signals was Abo el zahad et al, while Swati et al., [44] presented a biometric PCG system with the use of MFCC for obtaining features and support vector machine (SVM) as a classifier. The system was applied on 30 subjects reaching an accuracy of 96%.

In 2015, S. Bindu et al., [45] worked on an identification system based on pre-processing and extracting the heart sounds which are S1 and S2 as features and template matching for classification. Abo el zahad et al., [46, 47] developed a new approach for human verification based on heart sounds. The approach is based on wavelet packet cepstral coefficients (WPCC) reaching identification accuracy 91.05% and verifying them with 3.2% EER. Discrete wavelet decomposition was used for pre-processing and wavelet packet cepstral coefficients for feature extraction. LDA and Bayes rule are used for classification. Abo el zahad et al., [48] showed a human identification system based on PCG signals using MFCC, LFCC, WPCC, and Non-linear filter cepstral coefficient (NLFCC). The classification was done using LDA and Bayes achieving an EER of 2.88% and 2.13% on 206 and 21 subjects respectively.

In 2016, Abo el zahad et al., [49] proposed a new methodology for the individual identification based on PCG signal. The basis of this approach was MFCC, linear frequency cepstral coefficient (LFCC), modified Mel frequency coefficient (M-MFCC), and WPCC coefficients to reach 91.05% identification accuracy. Moreover, they were verified with 3.2% EER on 206 subjects and 2.68% EER on 21 subjects. Discrete wavelet decomposition was used for pre-processing and wavelet packet cepstral coefficients for feature extraction. LDA and Bayes rule was used for classification.

In 2017, T. E Chen et al., [50] presented a recognition system for the PCG signals based on some acoustic features. Their main work focused on two sets which are the training and the test sets. The number of subjects was 16 with 626 heart sounds for the training set, and 6 subjects with 120 heart sound like a test set. Their feature extraction was based on MFCC and k-mean to enhance the features. Finally, the features are fed into five classifiers which are LR, GMM, KNN, SVM, and DNN. The experiment resulted that DNN has the highest accuracy of 91.12%.

In 2018, TG Meitei et al., [51] concentrated on presenting PCG as a biometric as a few sources are available in this area deeming it to be nascent, and showed a PCG biometric system. The preprocessing stage was based on wavelet and the features were extracted from previous researcher's techniques. In the matching phase, ED, GMM, FSR, and VQ methods were examined. In 2019, Imran et al., [52] proposed a method for PCG identification for biometric-based on autoregressive modeling. They applied their methods on 50 subjects taken from a publicly available dataset. They used wavelets based on DWT for de-noising, and then they used the Hilbert envelope for segmentation. For each PCG beat, AR Burg modeling is applied to compute reflection coefficients, and finally, they used bagged decision trees for classification. They achieved an accuracy of 86.7%. El-Sayed et al., [53] worked on 60 and 50 subjects from two datasets. Multi-resolution decomposition and multi-resolution reconstruction (MRD-MRR) were the basis of their pre-processing. The feature extraction was based on Shannon Energy Envelope (SEE) and Multi-Scale wavelet transforms for time-scale. For time-frequency, it was based on framing, windowing, and (MFCC, BFCC, and LFCC). The classification was based on Random Forest (RF), SVM, ANN, and KNN. RF proved the highest accuracy in the time-frequency domain analysis, and SVM proved the highest accuracy in the time-scale domain analysis. Xiefeng et al., [54] performed their work on 80 heart sounds from 40 subjects. They based their proposed model on the decomposition of every single pulse of heart sound into a set of intrinsic model functions (IMFs). Then, after segmenting the IMFS, frames are produced. These frames are then used for obtaining the multiscale dispersion entropy as the main representation of the PCG signal. A combination of logistic regression (LR) and hidden semi-Markov model (HSMM) was used for extracting the features, and Fisher ratio (FR) was used for feature selection.

Finally, the proposed method was tested using an (ED) reaching an accuracy of 96.08%. The chronological development of PCG as biometric techniques is presented in Fig. 4.

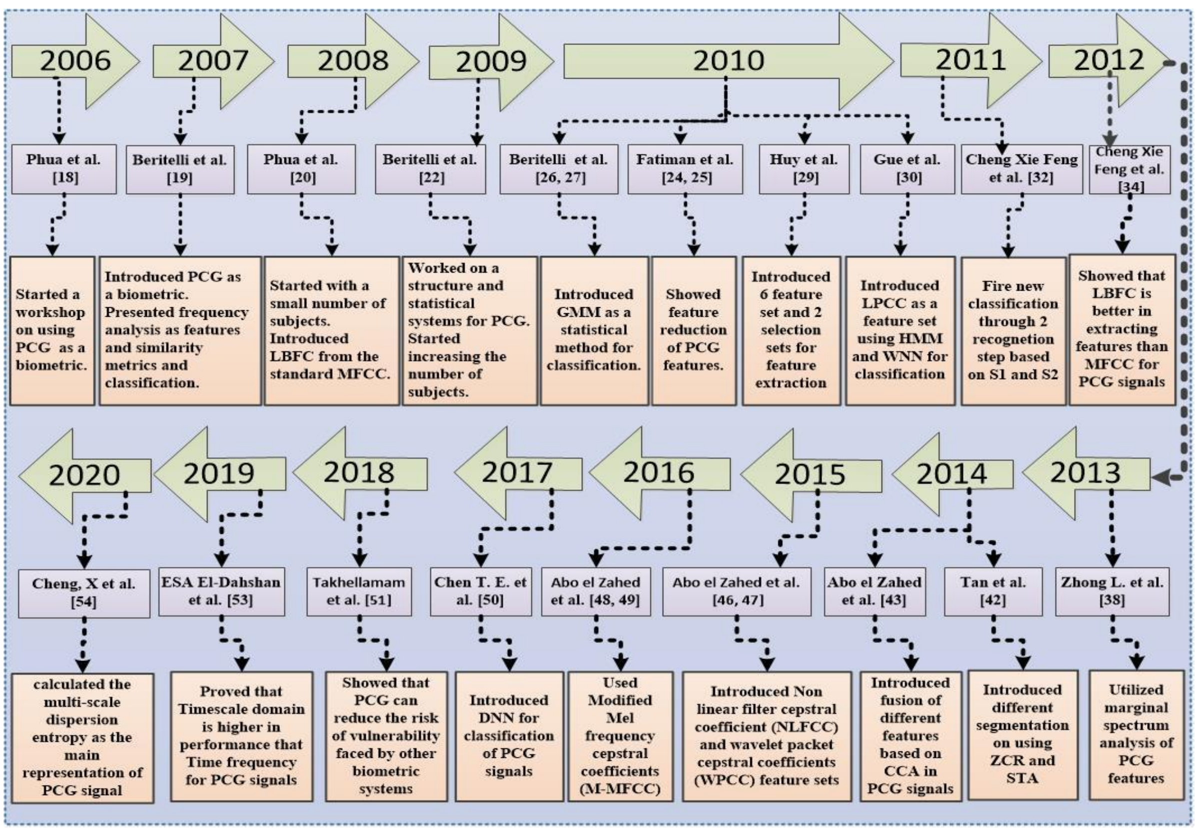

Fig.4. The chronological evolution of PCG as a biometric with a series of milestones that have been covered in this field so far from 2006 to the current status.

4. Relevant Work and Main Contribution

In comparison to other previously published works of literature, our main focus in this review is PCG as a biometric. In this paper, we explained the data acquisition, illustrated what is available online with several subjects and recordings, and what is not available. Moreover, the de-noising techniques in the heart sound as a biometric are explained in detail. The segmentation stage, feature extraction, and various segmentation techniques used in PCG as a biometric are studied and compared. In addition to this, the advantages and the disadvantages of most of the de-noising, segmentation, and classification stages used in PCG as a biometric are determined to help the researchers to choose suitable techniques. The number of articles and journals studied in our review is considerably more than the number of articles used in other review papers on PCG as a biometric. The authors that targeted a review for the PCG as a biometric used about 5 papers for their survey, while our survey considered 35 papers published in different journals and conferences for PCG as biometric, and we used them for discussion and review. Other papers **[55]** have reviewed the use of PCG and ECG together as a biometric, but it did not concentrate on a specific biometric trait or the concerns in both PCG and ECG. The data acquisition and the de-noising stages, and the segmentation step were not discussed. Furthermore, the merits and the demerits of those methods are not explained in their survey. The main contributions of our review are discussed in the following lines:

1. Presents a survey investigating relevant recent research papers which cover a total of 35 publications from journals and conference proceeding, from at least seven scientific libraries
2. Complements neglected area and gap information in the earlier reviews, such as acquisition, de-noising, segmentation, extracting the most discriminate features, and classification.
3. Thoroughly discusses the advantages and the disadvantages of techniques and methods applied in segmentation and preprocessing over others to determine the robust techniques.
4. Illustrates the important measurements used for identification and verification systems for PCG biometric.
5. Provides a strong reference and lowers the barriers of entry to the field of PCG biometrics.
6. Offers a great range of views and comparison in terms of experimental evaluation, classification, and categorization.
7. Recommends a potential opportunity for enhancement and exploitation.
8. Demonstrates the applications in which the PCG signals can be applied as a biometric technique and how it can be combined with other biometrics techniques to increase performance accuracy.

5. Processing steps of PCG Biometric System

The PCG biometric system processes are divided into 3 steps. The first step in which the database is created is known as the enrollment stage, where storing the PCG feature set of the individual is performed. The authentication stage is the second step in which there is an extraction of PCG features, and then a comparison between them and the feature templates that exist in the database is conducted to find a similarity known as the identification mode. Then, they can be compared with template features of the requested identity which is known as verification mode. The authentication and the enrolment modes are shown and explained clearly in Fig.5. The green lines show the enrolment phase and the black lines show the authentication phase.

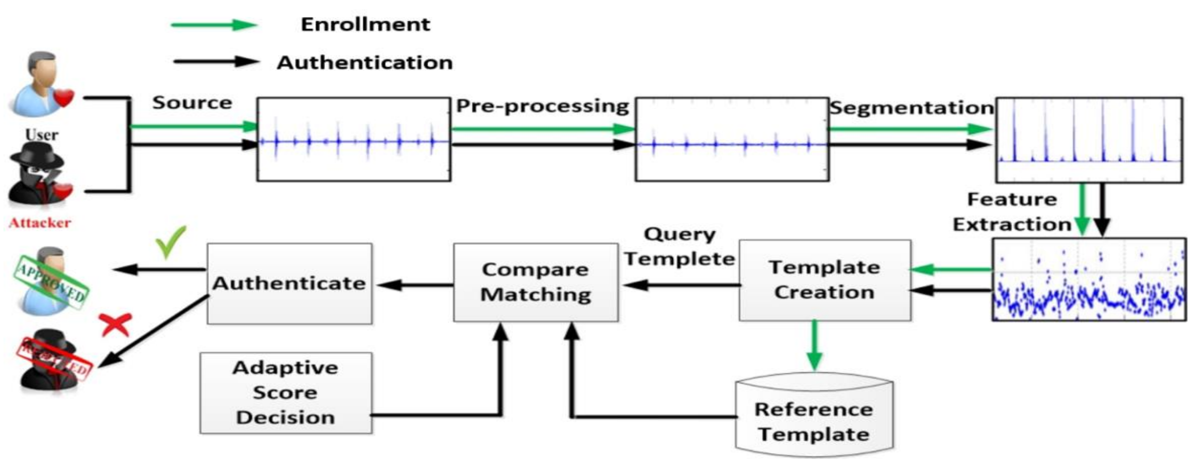

Fig. 5. Shows the schematic diagram of a generic biometric system.

5.1 Data Acquisition

Data Acquisition is one of the most important steps in PCG biometric authentication. Heart sounds can be captured from many devices, including mobile phones. One of the main devices in capturing the PCG signal is the stethoscope. There are some sample datasets available online containing PCG signals. These datasets can be used in the applied studies to evaluate the methods proposed in the research studies discussed for PCG signals as a biometric, and they are divided into five categories as shown in Fig. 6.

In Table 1. Data acquisition in the PCG biometric systems are presented according to the dataset, devices used to capture it, records, subjects, Fs, digitization, and studies that have used them.

The remaining studies discussed in this manuscript built their dataset and used it for identification and verification.

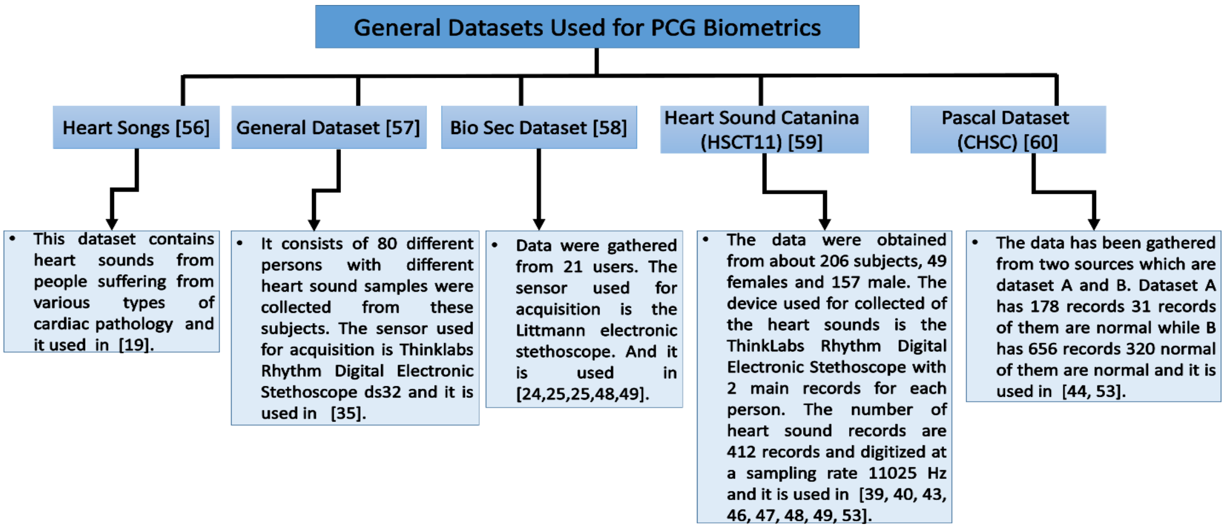

Fig. 6. The datasets used in PCG biometric.

Table 1. Datasets available for PCG classification

Datasets	Devices	Description	S/R	FS/DR
Heart Songs [56]	N/A	• It includes PCG signals that suffer from different types of cardiac pathology [19]. • It includes different types of pathologies such as mitral regurgitation variation, mitral stenosis, mitral regurgitation, innocent systolic murmur, and the third second [19]	S: N/A R: N/A	Fs: 11025 DR: 16
General Dataset [57]	Thinklabs ds32 (Rhythm Digital Electronic Stethoscope)	• It includes various people with various age ranges, and it incorporates pregnant ladies, people enduring distinctive heart abnormalities, and healthy individuals [35].	S: N/A R: 80	Fs: N/A DR: N/A
Bio Sec 2010 [58]	Littmann electronic stethoscope	• For each user, six recordings were taken, each under the rest condition [24, 25, and 26]. • These datasets are saved as an audio file with the extension ".au" and aren't accessible for free [48, 49].	S: 126 R: 21	Fs: 8000 DR: 16
Heart Sounds Catania (HSCT11) 2011 [59]	ThinkLabs ds 32 (Rhythm Digital Electronic Stethoscope)	• These databases were gotten from the clients during the acquisition stage the individual was in a sitting situation and resting state. • The stethoscope was situated close to the pulmonary valve. It is considered the largest dataset used for PCG as a biometric [39, 40, 43, 46, 47, 48, 49, and 53].	S: 412 R:206 157 male 49 female	Fs: 11025 DR: N/A
PASCAL (CHSC 2011)	iStethoscope Pro iPhone app digital	• The data are obtained from datasets "A" and "B" [44, 53]. ➢ Dataset "A" is accumulated utilizing the	S: Dataset A 178 Dataset B	Fs: N/A DR: N/A

[60]	stethoscope DigiScope	iStethoscope Pro iPhone application. ➢ Dataset "B" is collected from patients in the hospitals using the stethoscope device. Dataset A consists of four categories normal, murmur, extra heart sound, and artifacts. While dataset B has three main different classes which are murmur, extra systole, and the normal beats.	656 R: N/A	

N/A: Not Defined **S:** Subjects **R:** Records **Fs**: Sampling frequency **DR**: Digitization Resolution

5.2 Pre-processing

Heart sounds are considered monodimensional signals which have the possibility of being processed to some extent. There are many techniques used to work on mono-dimensional signals taking into account the structure and the components of the signals. PCG signals are subject to noise and artifacts that the external effects introduce. The pre-processing stage has the objective of producing a signal with reduced noise. The filtering method maps the input signal into an output signal, and this facilitates extracting noise from the input signal. It is possible to further divide the noise heart sound signals into many variant types as clear in Fig.7 depending on the signal features based on frequency, time, and spectral features.

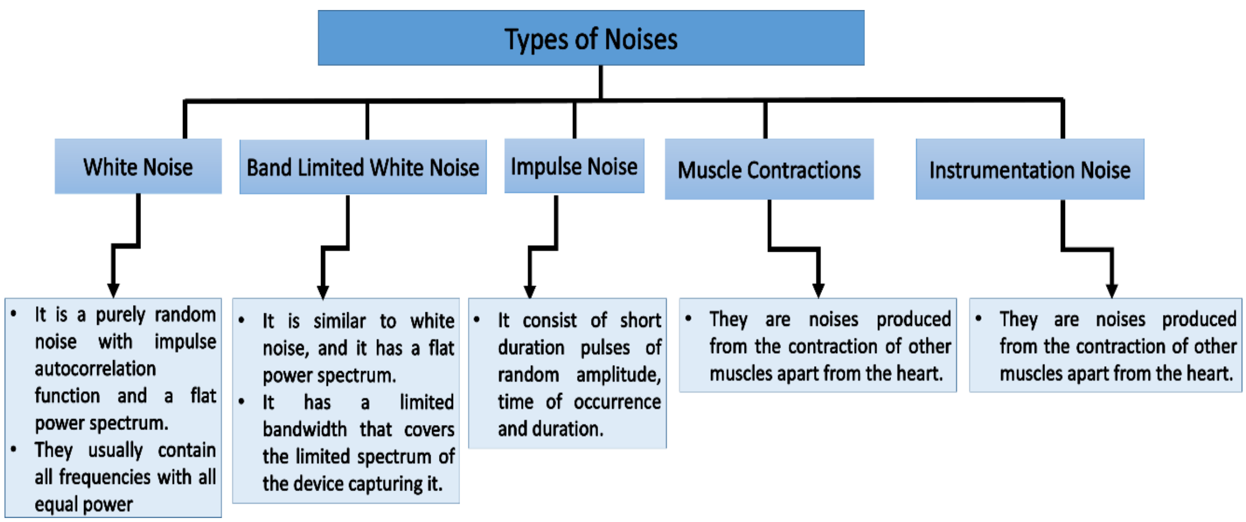

Fig. 7. The most known types of Noise in the heart sound signals.

Several de-noising approaches can be used to reduce these noises and improve the quality of the PCG signal. They can be divided into five main categories. Those categories are based on down--sampling, low pass filtering, energy thresholding, normalization, total variation de-noising (TV). Since frequency components of PCG signals are concentrated below 250 Hz. Some approaches of down-sampling are applied to obtain a sampling frequency of approximately

1KHz. The down-sampling factor used is 11 [46, 47, 48]. Other approaches used analog low pass filtering it has a great effect on the heart sounds, and a good low pass filter can filter the noise by leaving a large number of samples for further processing. It is designed to eliminate the high-frequency components in the PCG signal. A sixth-order low butter worth filter with a cut-off frequency of 400 Hz and 800 Hz is used in [22,23] respectively. A low butter worth filter with a frequency range from 0 to 300 Hz to remove noise components beyond 300 Hz is used by [41]. An eight-order elliptical low pass filter is applied to remove noise in the PCG signals [30]. A 5^{th}-order Chebyshev Type I low pass filter with a cut-off frequency of 880 Hz was used by [40]. Another approach uses the concept of energy thresholding for de-noising [20, 24, 25, 43, 46, 47, 48, 49]. It is considered to be an optimal process for de-noising. Normalization is a pre-processing phase to achieve accurate heart sound identification. The maximum amplitude taken in [30, 39] is 1. An approach is based on a high pass filter It recompenses the part of high frequency that were blocked during the production of heart sounds [44]. PCG signals can also be filtered using (TVD) as it is considered to be a new technique that proved its performance in the PCG signal de-noising [45].

Other approaches used DWT [47, 48, 52, 53] on the PCG signals for a clear filtration of the signal. This approach is conducted by passing the input signal through a series of high and low pass filter banks. The output of this filtering is the detailed and the approximation coefficients. The de-noising process is executed by applying a thresholding schema for wavelet coefficients. The thresholding selection rules are based on rigrsure, sqtwolog, Heursure, and Minimaxi. Rigrsure is known as the adaptive threshold selection using the principle of Stein's unbiased risk estimate, while sqrtwolog is the fixed form threshold and it is the square root of twice log of the length of the signal. Heursure is a variant of rigrsure and sqrtwolog, and minimax is the minimax thresholding. The thresholding rescaling values are based on (one, sln, mln). Those rescaling values represent the thresholding values, for example, one means no rescaling, sln for rescaling based on a single estimation of the noise level using the 1^{st}-level coefficients of the decomposition structure, and mln is for rescaling done based on level-dependent estimation of the noise level. A 5^{th}-order DWT is decomposed to five scales and reconstructed using coefficient only from the third, fourth, fifth scales with two energy thresholding rules used in [24, 25, 43, 48] for de-noising PCG signal. Others are set to zero and reconstructed using Inverse DWT (IDWT). Decomposition using DWT 5th level four thresholding estimation rules and three threshold rescaling values are applied for wavelet coefficients. The rigrsure is considered the best thresholding selection rule and the sln is considered the best threshold rescaling [46, 47, 48].

Fig. 8 shows a summary of the most commonly used preprocessing techniques with definition and effects on the PCG signal in terms of de-noising. It is shown that some techniques have advantages over other techniques.

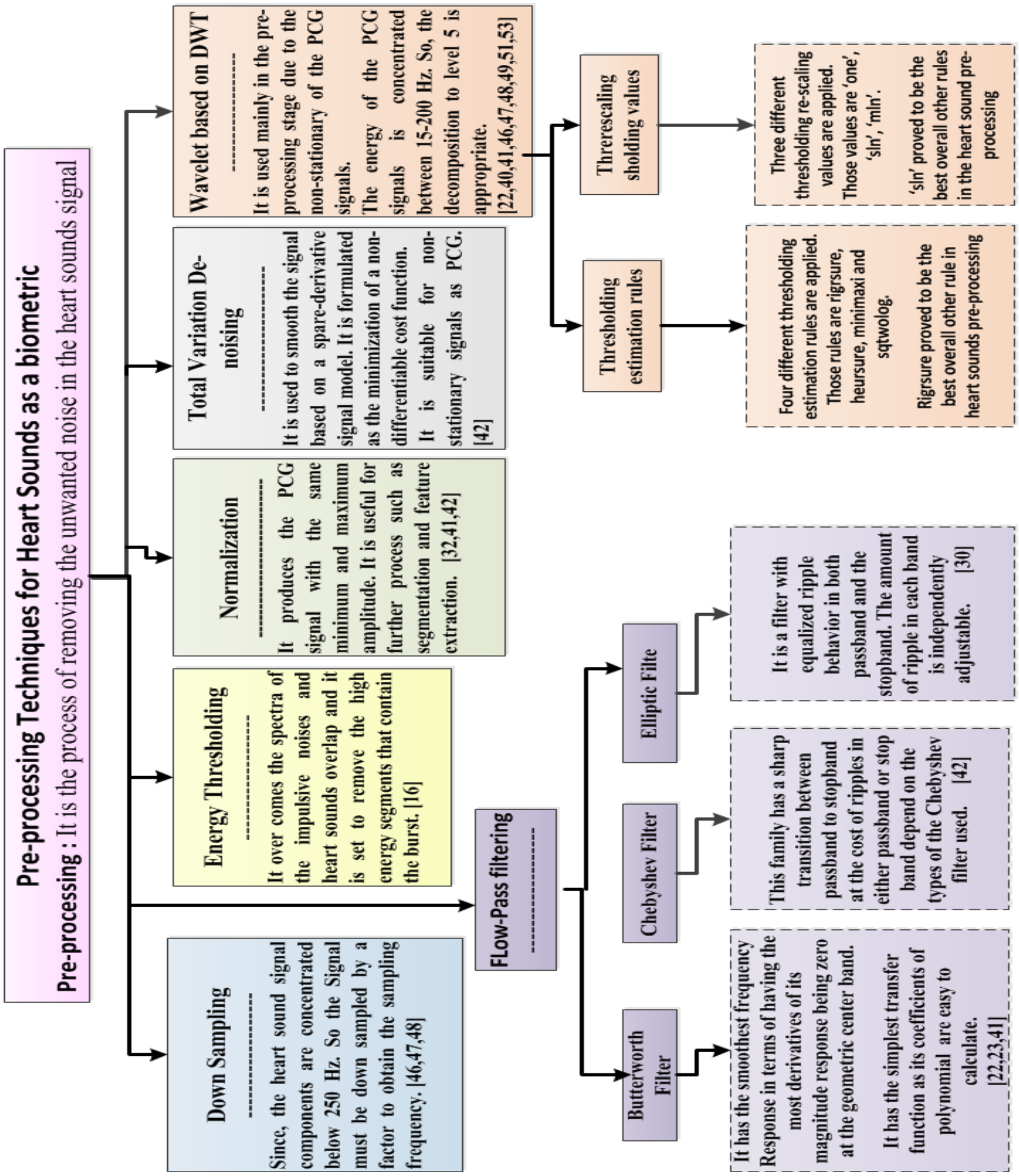

Fig. 8. Overview of the most commonly used Pre-processing Techniques for Heart Sounds as a biometric

Wavelets-based techniques have an irregular shape so they can perfectly reconstruct functions with linear and higher-order polynomial while the FFT fails to do this. Therefore wavelet-based techniques can denoise particular signals far better than the conventional filters that are based on FFT, for example (low pass filters). Wavelets have more advantages in de-noising the heart sound signal mainly because they can deconstruct complex parts in the PCG signals into basics signals of finite bandwidth and then reconstruct them again with little information loss. The result is a little or no heart sound signal leakage or phase-shifting of the original heart sounds signal. There are generally problems in conventional filters about heart sound signal leakage or phase-shifting. These problems must be addressed, or at least acknowledged in the de-noised signal. Furthermore, the conventional filters work most efficiently in removing out-of-band signals. If applied to in-band signals, it will remove the signal of interest. To de-noise the PCG signal, an algorithm was implemented based on a multi-resolution decomposition and multi-resolution reconstruction (MRD-MRR) scheme. The introduction of this algorithm serves to separate discontinuous sounds of the lung. It is regarded as a wavelet filtering technique that depends on the most important peaks in the time domain having large components over different wavelet scales [61]. When the noise and signal spectra overlap, the wavelet functions well in removing noise. TV is a de-nosing method in which the output can be obtained using a numerical algorithm, unlike the conventional method. TV is the most appropriate method for piece constant signals, and it can be modified and extended to be effective for more general signals. Table 2 shows the pre-processing techniques used for PCG. It shows the parameters used in the pre-processing approach based on order, cut-off frequency, maximum amplitude, down-sampling factor, and others.

Table 2 Pre-processing Techniques used for PCG biometric.

Pre-processing Approach	**Functionality**	**Parameters**
Energy Thresholding	• De-noising spike noise as the nature of the heart sounds experiences impedances caused by movement of the stethoscope, spectra of rash noises and heart sounds overlap [20]. • It is used to eliminate the overlapping noise components and to detect the spikes [24, 25, 26] • It can be divided into high energy thresholding and low energy thresholding. The high energy thresholding suppresses the spike, whereas the low energy suppresses the noise component [40, 43, 46, 47, 48, 49].	N/A
Low pass Filter	• **Butterworth**: It is used to remove the high-frequency components and to obtain the heart sounds that can be used in the next phase which are obtaining the features and to extract two different sets of biometric indexes [22, 23, 37, 41].	6^{th} order Butterworth filter [22,23] The cutoff frequency of 400 Hz [22] The cutoff frequency of 800 Hz [23]

		Butterworth filter from 0 – 300 Hz [37,41]
	• **Elliptic filter**: It filters the PCG signals focusing on the low-frequency parts, and it removes the interference that appears in the shape of spikes that falls in the higher frequency part. [30].	8^{th} order elliptical filter [30]
	• **Chebyshev filter**: It can be used to remove the noise located in the background of the heart sounds, and it also eliminates the sounds that have a frequency greater than the cut-off one[42].	Chebyshev filter-type I- 5^{th} order with Cutoff-frequency at 880 Hz [42]
High Pass Filter	• It leads to a higher frequency which increases the energy of the signal [44].	-
Normalization	• It confines the PCG signal inside a fixed range (1 to - 1) [38, 39].	Maximum amplitude is 1 [39]
Down-sampling	• It can be applied to decrease the sampling frequency and eventually band limit the content of the frequency in the heart sound signal [46, 47, and 48].	Down-sampling factor = 11
TVD	• It preserves the edges of the signal used for pre-processing and it is very flexible with Spike-like, piecewise-constant signals [45].	-
DWT	• The primary target utilizing the DWT for de-noising is to cancel high and low-frequency noise components. • It also removes other interferences that affect the heart sounds, for example, lung sounds, body movement, and other kinds of surrounding sounds. • DWT can be utilized to hold a large portion of the PCG signal energy as it is concentrated inside its scales. Additionally, it is beneficial for unwanted signal removal due to the non-stationary feature of the signal and the grouping of the PCG signal frequency in 15-200Hz.	5^{th} order DWT into 5 scales reconstructed from 3^{rd},4^{th} and 5^{th} scales + Two
		5^{th} order DWT into 6 scales obtained from 3^{rd},4^{th}, 5^{th} and 6^{th} scales and retained based on energy-threshold other scales are set to zero and reconstructed using Inverse DWT (IDWT) [46,47,49]
		Decomposition using DWT 5^{th} level 4 Thresholding estimation rules 3 Threshold rescaling values [40]

5.3 Segmentation

In this study, we have surveyed the segmentation techniques employed for heart sounds, especially normal heart sounds for biometric systems. In this section, we present a comparison including many published segmentation techniques for S1 and S2. Table 3 reveals the segmentation algorithms and techniques developed for the segmentation of S1 and S2 from the

PCG signal. It is shown that most of the segmentation techniques used are based on framing and then windowing with different values for the frameshift (fs) and frame length (fl). The framing process is done by extracting the signal energy information to obtain a more detailed profile of the energy tend. The windowing step can be performed using different functions such as hamming, hanning, and rectangular. Some approaches [20, 24, 25, 26, 35, 43] depend on framing, and implemented short-time discrete Fourier transform (STDFT) with fl of 256 ms or 512samples at a sampling rate of 2 kHz). It is also deduced that the non-over-lapping windowing should be the most optimal. However, the samples are very short over-lapping should be used. Some approaches perform framing and windowing depending on a hamming window with a length of 256m for frame [46, 47, 48]. These two steps aim for avoiding difficulties in the transactions of the heart sound signals and for increasing the signal smoothing. Moreover, the PCG signal is fairly stationary and this means that it is possible to examine the heart sound signals over a short time. As a result, the signals can be analyzed in small time segments.

Some authors applied framing with an fl = 256ms and fs = 64ms as an optimal framing [38]. They tried different types of windowing and showed that the hamming window gave the highest CCR. Other approaches were based on dividing the PCG signal into frames. Some authors applied framing with an fl = 256ms and fs = 64ms as an optimal framing [38]. They tried different types of windowing and showed that the hamming window gave the highest CCR. Other approaches were based on dividing the PCG signal into frames. Then an autocorrelation function was applied for ordering and determining the variant periodic participles of the er signal [19, 21, 22, 28, 29, 39]. Multiplication of the frame values by a hamming window performed for minimizing the discontinuation disruption at the start and the end of each fi This leads to the estimation of the power spectral density and the determination of the sounc has the greater energy. Furthermore, the bounds of these sounds are identified in tern samples in an audio track from cardiac auscultation. Some segmentation approaches were t on the Shannon energy operator [23,50] with 0.02s segments and 0.01s segment overlap This approach was used to identify the quality of the sound lobes in the heart sounds ai extract the signal envelope.

Wavelets were used for the segmentation of the PCG signal sounds [36]. They found that s s2 consist of three main kinds of sub-wavelets which are four peaks, twin-peak, four- wavelets. S1 consists of seven twin-peak, 1 four-peak, 5 three-peak subwavelets. Whil involves 9 twin-peaks, 1 three-peak, and 3 four-peak subwavelet. The main decomposition depends on DWT. The main aim of the DWT is to obtain the coefficients in detail of the PCG sounds, and it can represent the personal features of the PCG signals as a whole.

Finally, a segmentation approach was based on framing with an fl of 5 m and applying ZCR and STA algorithms to detect S1 and S2 from the PCG signal [42]. Fig.9 shows the advantages and the disadvantages of the segmentation techniques used in the PCG as a biometric for framing and windowing, autocorrelation, Shannon energy envelope, and zero-crossing and short-term amplitude.

Table 3 Segmentation Techniques used for PCG biometric.

Segmentation Approach	Functionality	Parameters
Framing and windowing based on an Auto Correlation function	• To estimate the power spectral density • To determine which sound has a greater energy • It identifies the boundaries of the heart sound signal, in terms of some samples in the PCG signal trace obtained from the heart auscultation. • The main aim of this framing is to create independence between samples and to perform an accurate estimate. • It helps in finding the variance of each segment and to increase the effect of the segment parts S1 and S2 and enhance them [19, 21, 22, 28, 29, 39],	Frame length = 15 ms Frame shift = 5 ms
Framing based on STDFT	• To make independence between samples, it is required to precisely estimate the framing characteristics [20, 24, 25, 26, 36, and 43].	Frame length = 256 ms
Shannon energy operator	• To identify the boundaries of all sound lobes in the heart sounds [23]. • To extract the signal envelope [47]. • This technique includes the intensity of the heart sound signal, and it also reduces the impact of low intensity of the PCG signals [53].	0.02s segments 0.01s segment overlapping
DWT Wavelet	• To synthesize a set of simulated heart sounds [36]	-
Framing and windowing based on hamming window	• For avoiding the problems with the transactions of the signal and aides in the smoothing of the signal [46, 47, 48, 49].	Frame length = 256 ms
Framing and windowing based on Hamming window	• The consequent feature extraction depends on each frame formed [38, 52].	The length of the frame is “256 ms” And the frame is shifted by “64 ms”.
Framing and applying ZCR and STA algorithms	• To increase the effectiveness of the segment part S1 and S2 • To detect S1 and S2 from the PCG signal. • This method can be applied to determine the silent parts in the audio signals, and it is very useful in detecting the audio from the background noise during the start and endpoints. [42].	Frame length = 5ms

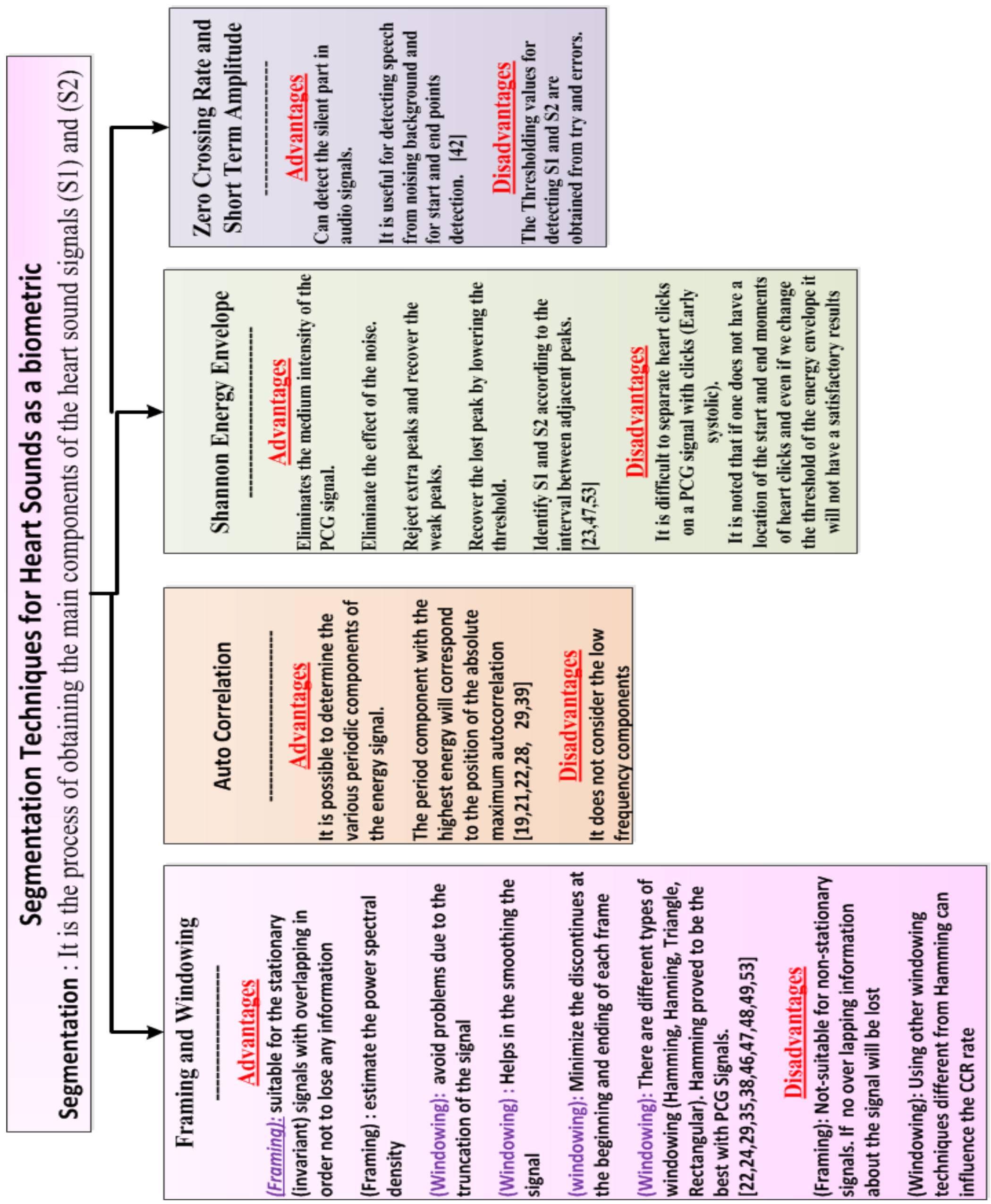

Fig. 9. Illustrates the advantages and disadvantages of methods utilized in PCG as a biometric

5.4 Feature Extraction and Reduction

The process of characterizing the attributes similar to all the features belonging to a certain type is called feature extraction. Feature extraction mainly aims to search for features that can reach optimal or near-optimal results regarding certain performance measures [62]. Table 4 provides a detailed explanation of major feature extraction methods.

There are different feature extraction methods used in heart sounds as a biometric. These approaches can be based on time, frequency, time, and frequency feature as shown in Fig.10.

Time-domain features refer to the analysis of the PCG signal with respect to time. In the time domain, the signal samples are known for all real numbers. Time domain features are based on signal energy envelope, first second ratio (FSR), log-attack time, temporal increase, signal strength, zero crossing rate (ZCR).

- **Signal energy envelope**: This method extract features based on the normalization of PCG signal after filtering. The first step is computing the magnitude of the normalized signal and computing the Shannon entropy of the positively valued signal. This is followed by adaptive thresholding of the Shannon entropy sequence, and finally, the signal is smoothed for the production of the time and domain features [30, 45].
- **First second ratio (FSR):** It is a feature based on which obtains the ratio between S1 and S2. It is used in addition to features produced to determine whether two heart sounds sequence belong to the same person. In [26] FSR was introduced as a time-domain feature based on the ratio of the power S1 over S2. This achieves a higher performance as discussed in [22, 23, 24, 25, 26, and 27].
- **Log-Attack Time (LAT):** LAT is an algorithm of the time duration between the point where the PCG starts to the point it reaches a stable part [29].
- **Temporal Increase:** It is a feature that represents the change with respect to time. The temporal increase means a large number of unique frames are repeated after each other [29].
- **Signal Strength:** It is the physical energy or generated energy from the PCG signal [29].
- **Zero-Crossing rate:** It is the rate of significant change along with the PCG signal. It is the rate at which the signal changes from positive to negative or back [29].

Frequency domain features can be classified into two groups Chirp Z-transform (CZT), discrete cosine transform (DCT).

- **CZT:** One of the fast Fourier transform (FFT) techniques that aim to compute the discrete Fourier transform (DFT) of different arbitrary sizes is known as chirp Z-transform (CZT). The main advantage of this algorithm is that it does not require a power of two numbers of samples. This can be achieved by representing the DFT as a convolution. It was also used due to the energy of S1 and S2 is vitally concentrated around frequencies less than 300 Hz.

- **DCT:** It is a frequency domain algorithm used in most of the studies for feature reduction and generates features with lower dimensionality for classification. It was used as a reduction tool in the following studies [20, 35, 40, 46, 47, 48, and 49].

Time-frequency domain features consist of several different types of characteristics such as harmonic, spectrum, and cepstral features [63]. The Cepstral features are divided into 7 groups discussed as follows:

- **Mel Frequency Cepstral Coefficients (MFCC):** They are features based on customary filters that have a triangular shape that is represented equally on the Mel scale. [22] applied MFCC as the first feature and a filter bank of 7 filters, [23, 33, and 44] used 24 filters built from bandpass filters ranging from 0 to 750 Hz [22] and from 0 to 4000 Hz [23]. Using a filter can achieve higher performance by widening the band and it is related to a low-pass filter. 13 coefficients are produced and for each heart sound 182 feature vectors are produced. [44] used MFCC and DCT for reduction.[50] used MFCC and K-means to cluster the acoustic features of the PCG segment. The magnitude of the spectral coefficients is passed to MFCC. MFCC coefficients are based on LDA for dimensionality reduction [24, 25]. [36] Implemented MFCC for feature extraction and PCA for dimensionality reduction. [38] used MFCC to extract representative features. [39] showed a structural approach based on MFCC and FSR for feature extraction. Other studies also used MFCC [29, 42, 45, 46, 48, 49, 53] in comparison with other feature extraction methods.
- **Modified Mel Frequency Cepstral Coefficients (M-MFCC):** It is based on MFCC and it aims to increase the non-linearity of the triangular filter banks. It introduces a parameter called $\alpha < 700$ to increase the non-linearity of the triangular filters in the frequency below 250 Hz in the range of the heart sounds frequency. M-MFCC showed the highest performance over different feature extraction Techniques such as MFCC, LFCC, WPCC, and it was used by [49].
- **Bark Frequency Cepstral Coefficients (BFCC):** They are features that are based on a set of triangular features that are equally spaced on the bark scale to generate bark

coefficients. It was used by [42, 53] to test its performance as a feature extraction to others.

- **Linear Frequency Cepstral coefficients (LFCC):** They are features based on a set of triangular features that are equally spaced on the linear scale to generate linear coefficients. It was used by the studies presented in [29, 43, 46, 48, 49, and 53] and its performance was not better than MFCC and Bark. It was also introduced as a statistical approach based on LFCC and FSR and it was better in performance compared to MFCC with FSR [39].
- **Non-Linear Frequency Cepstral Coefficients (NLFCC):** They are features based on LFCC as it focuses on the increase of the non-linearity in the linear scale from 0 to 250 Hz as most of the energy is concentrated from 0 to 250 Hz. NLFCC is the nonlinearity of the Mel-frequency scale, and it rises for frequencies exceeds 1000 Hz, The Mel-scale is considered linear under this value. This modification was made by [48] and it was higher than MFCC, LFCC, and WPCC in performance.
- **Linear Frequency Band Cepstral Coefficients (LFBC):** The concentration of the heart sound spectrum is within the range of 20–150 Hz. To capture more information of the signal spectrum, full resolution is necessary for such a narrow band. To distinguish the heart sound feature from the standard MFCC, it was known as the LFBC feature set. This set was used by [20] and DCT for dimensionality reduction of the feature produced from LFBC.
- **Linear Predictive Cepstral Coefficients (LPCC)**: They are features produced by estimating the n+ 1 sample using a linear combination of its previous n samples. These are known as predictor coefficients. They are from the smoothed Auto-Regressive power spectrum instead of the period gram estimate of the power spectrum [30, 38].
- **Heart Sounds Linear Band Frequency Coefficients (HS-LBFC)**: In the PCG signals, we need to include more data in the spectrum of the signal, and the signals that are narrowband need to be concentrated on. In the process of obtaining the MFCC coefficients, the energy of the Mel filter group is required. According to the natural properties of the logarithm, it has a rapidly increasing large slope in the low-frequency sections. So, a piece-wise function is used to replace the log function. In this case, the signal attenuation in the low-frequency sections will be more appropriate [35].

- Other features are based on harmonic features such as Harmonicity, Harmonic attenuation, Harmonic Spectral Deviation, Harmonic Energy Ratio shown in [29] study. Applied in [29] study, there are some rhythmic features based on the bass, max, aggressiveness, gravity, addition and low-frequency domination, and modulation spectrogram, while other features are based on spectrum features. The bases of such spectral features are spectral attenuation, spectral roll-off, spectral centroid, spectral flatness measurement (SFM), spectral divergence, spectral kurtosis, spectral skewness, tonality, spectral range, s-transform, and Hilbert Spectrum [64].

One of the feature extraction techniques that was used is the Hilbert spectrum in [40]. It uses EEMD for each frame in the heart sounds to get the intrinsic mode functions (IMFs). Hilbert spectrum used hung-transform (HT) to each IMF as it determines the instantaneous frequency. It also shows the amplitude in a 3D plot with respect to time. It can reveal the heart sound physical properties in a time-frequency plane. A marginal spectrum is produced based on applying a Hilbert spectrum in a three-dimensional form for each frame, and it is obtained in the integral form of the time domain.

The reduction of the features is done using DCT and amplitude normalization.

Time-Scale domain features can be classified into three types. Those types are discrete wavelet transform (DWT), improved circle convolution (ICC), and wavelet packet cepstral coefficient (WPCC).

- **DWT**: The study [31] used **DWT** [65- 69] to extract features from the PCG signal and decompose it to level 6 using db5. The coefficients produced are all the details from level 1 to 6 and the 4th and 5th levels approximation and then finally they are introduced to SEE as features. Another study was based on the DWT to generate features. The extracted S1 and S2 are sent to wavelet db2 level 2 decomposition. The detail coefficients D2 are divided into N = 20 windows and the energy is calculated for each window**.** The final feature-length is 40. Where the first 20 is feature extracted from S1 and next 20 is feature extracted from S2 [36].
- **MT-DWT**: The study [53] used DWT to obtain coefficients. Those coefficients are introduced to multi-Scale features such as Energy, Entropy, Standard deviation, and wavelength. The decomposition was made to level 11 producing 44 feature samples.

- **ICC:** ICC was used as a feature extraction method in the study [32]. The ICC works by using a different resolution to perform successive approximations about the heart sound signal based on wavelet packets to divide the PCG into layers by the (ICC). Then an independent sub-band function is used to analyze the independent components of these layers.
- **Wavelet Packet Cepstral Coefficient (WPCC**): A study introduced WPCC [46, 47, 48, 49] as a feature extraction method based on the time scale domain. Unlike the DWT that decomposes the signal's lower frequency part, WPD decomposes both the details and the approximation coefficients to obtain a good resolution and representation of the signal. In those studies that used WPD different wavelet functions and different linear and non-linear filter banks were tested. After PCG decomposition using WPD, the energy of each filter output is calculated. Finally, it takes the logarithm and applies the DCT to filters energy to obtain WPCC features. It proved to produce a better performance in MFCC, BFCC, and LFCC but not better than M-MFCC and N-LFCC.

Fusion between features: Some features were combined to generate features based on the fusion and they are divided into two groups.

- **Discrete Wavelet Mel Frequency Cepstral Coefficient (DW-MFCC):** A fusion between the features of DW and MFCC and it proved to have a better performance compared to each of MFCC, BFCC, and LFCC alone [43].
- **Conical Correlation Analysis (CCA):** A fusion between features was introduced using CCA to produce different features based on serial and parallel strategies. In serial and parallel strategy MFCC+DW-MFCC was the highest performance than any combination using CCA. Overall the parallel strategy achieved higher performance than the serial strategy [43].

Some features were produced from tools used in the feature extraction of the speaker recognition but handled with parameters to generate features from the heart sounds. A tool called Sfbcep and is a part of the Sprout suite provided by an ALIZE/ SpkDet as an open source for speaker identification. It has a lot of features and facilities as it can perform cepstral analysis using filter bank on the signal, and it allows changing several parameters of the filter bank by obtaining the difference between heart sounds and speech features. It produces 3 feature sets (A, B, C) with the difference in the number of cepstra and the existence of first and second-order derivative features. This tool was used by [26, 27] and they showed that set C was the highest performance based on set C.

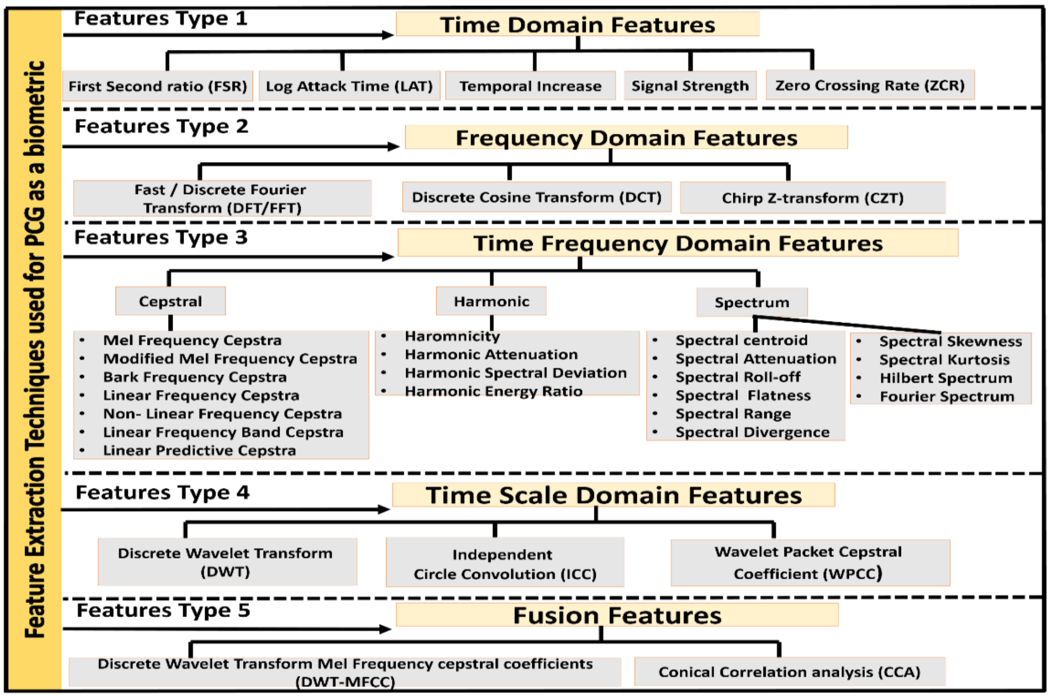

Fig. 10. An overview of the feature extraction approaches utilized in the PCG as a biometric.

Table 4 Feature extraction approaches utilized for PCG biometric

Authors	Types of Features	Approach
Beritelli et al. [19] Beritelli et al. [21]	Frequency analysis	CZT
Phua et al. [20]	Spectrum	Linear frequency bands cepstral (LFBC), (DCT)
Beritelli et al. [22,23] Zhao et al., [33] Swati et al., [44] T.E Chen et al. [50]	Time-Frequency based on Cepstral features	MFCC + FSR
Fateman et al. [24,25]	Time Frequency based on Cepstral features	MFCC + LDA
Beritelli et al., [26] Beritelli et al., [27]	Time Frequency based on Cepstral features	Feature extracted using MFCC + FSR, BFCC + FSR and LFCC + FSR
Tao et al., [28]	Mixed features	Cycle - power-frequency
Huy et al., [29]	Time domain, Time-Frequency, Frequency domain features	6 feature sets Temporal, Spectral, Cepstral, Harmonic, Rhythmic, GMM
Guo et al., [30]	Time-Frequency based on Cepstral features	linear prediction cepstrum coefficient (LPCC)
Jasper et al., [31]	Time Scale domain features based on DWT	Discrete wavelet transform (Db5), Shannon Energy Envelope (SEE)

Cheng Xie Feng et al. [32]	Time-Scale domain based on ICC	Improved circle convolution (ICC) Independent sub-Band function
Cheng Xie Fenget al. [34]	Time-Frequency domain based on Cepstral features	HS-LBFC
Rasha Wahid et al. [35]	Time-Frequency domain based on Cepstral and spectral features	2 feature extraction algorithms MFCC + DCT Spectral Magnitude + DCT
Chen W et al. [36]	Time-Frequency domain based on Cepstral	MFCC, PCA
Zhong L et al. [38]	Time-Frequency domain based on Cepstral	MFCC LFCC
Spadaccini et al., [39]	Time-Frequency domain based on Cepstral	MFCC/FSR LFCC/FSR
Zhao et al., [40]	Time-Frequency domain based on Spectrum	Ensemble Empirical Mode Decomposition (EEMD) based on Hilbert Spectrum
Girish et al., [41]	Time Scale features	Wavelet decomposition Using db2 level = 2
Abo el zahad et al., [43]	Time and Frequency Domain + Fusion between different domains	MFCC, LFCC, BFCC, DW-MFCC Discrete wavelet- (MFCC), CCA for fusion.
S. Bindu et al., [45]	Time Domain	Shannon energy envelope
Abo el zahad et al., [46, 47]	Time-frequency domain Time Scale domain	Wavelet packet cepstral coefficient (WPCC) LDA
Abo el zahad et al., [48]	Time-frequency domain Time Scale domain	MFCC, LFCC, WPCC, NLFCC, and LDA
Abo el zahad et al., [49]	Time-frequency domain Time Scale domain	MFCC, LFCC, WPCC, MFCC Modified Mel-scaled filter banks, LDA
El-Sayed et al., [53]	Time-frequency domain Time Scale domain	MFCC, BFCC, LFCC and MS_DWT

5.5 Classification

Classification is a process that finds the optimal class that is nearest to the classified pattern. The classifiers used in heart sounds as a biometric is classified into 4 types. Type one is based on similarity methods like Euclidian, nearest, and similarity distances. Some studies applied Euclidean distance to measure the spectra of two signals. Considering the spectra as N-dimensional vectors, the distances concerning segmented heart sound spectra extracted from the heart sounds of different people yield higher values than those with spectra alone [23,26,27,28,29,31,39]. Others used the nearest distance to compare anonymous PCG input samples against the whole templates stored in the database. Once the entered template matches the input heart sound signal based on the nearest distance, it will be preserved to be the user identified [27]. Other authors classified the PCG signals using similarity distance as it requires a signal to compare with which as the minimum accumulative distance with the verified signal in the database of the heart sounds. They used dynamic time wrapping directly [28, 30].

Some approaches were based on statistical classifiers. One of them is VQ with the main idea of wrapping up a huge number of cepstral data into a group of code data. For the reduction of quantization error, a training of VQ is usually conducted with LBG [70]. This method has the main advantage of achieving.

A relatively high identification rate with very short test signals. Sometimes, LVG is used for training VQ for a direct definition of the classification borders between classes by the nearest-neighbor rule. Performance of LVG was higher than LBG but, more costly. VQ gives the highest performance and lowest computational time [20, 33, and 40]. Some studies used GMM as a classification method [20, 24, 25, 27, 35, and 39]. GMM is considered the generalization of the VQ. Contrary to LBQ-VQ, the GMM technique is a stochastic model with a probabilistic matching pattern, since the results are in conditional probability observation taking into consideration the model or measure of the likelihood. The use of the GMM is for the classification stage and it constitutes a compromise between the performance and computational time. Furthermore, it offers the data as a weighted sum of Gaussian distributions with variance, mean, and weight. K-means are used for the calculation of the initial means and weights are calculated. The classification decision is based on Bayes theory. GMM is utilized in most of the PCG as biometric studies. Other studies used LDA, which is a special case of discriminant analysis that assumes that all the classes have the same covariance. LDA gives the poorest performance and highest computational time. The classifier decision is performed using the optimum Bayes decision rule which maximizes the posterior probability or its logarithm. Other approaches used multi-class SVM [71- 73] as a classifier to separate the PCG features. SVM is a discriminative classifier formally defined by separating a hyperplane. SVM works using different kernel functions (linear, quadratic, polynomial…..etc). SVM didn't show a good performance in some studies [48] but give a great reduction of error and a good accuracy in other studies [33, 48].

Table 5 Classification Techniques used for PCG biometric

Approach	Classifier	Authors	Results
Similarity	Euclidean distance	Beritelli and Serrano [19]	FRR = 5.2% FAR = 2.2%
		F. Beritelli [21]	EER= 9%

		Beritelli and Spadaccini [22]	EER= 9%
		Beritelli and Spadaccini [23]	EER= 5%
		Fateman et al.[24, 25]	A = 100%
		Tao et al. [28]	A = 99%
		Spadaccini et al.[39]	EER = 36.86%
		El-Sayed et al. [53]	KNN and showed the lowest performance accuracy in the classifiers used
	Nearest Distance	Beritelli et al., [26]	A = 98.67%
	Similarity Distance	Tao et al.[28]	A= 85.7%, EAR < 7%, RER < 10%
		Guo et al. [30]	A= 95%, EAR= 1%-8%, RER < 3%
Statistical	VQ	Phua et al. [18]	A= 93.58%
		X. Cheng [34]	A = 100%
		Z. Zhao, et al. [40]	CRR = 94.15%
			CRR = 84.93%
	GMM	Phua et al. [18]	A= 96.01%
		Fateman et al. [24, 25]	EER= 13.70%
		Beritelli et al. [26]	EER= 15.53%
		Rasha Wahid et al.[35]	A = 85%
		Spadaccini et al.[39]	
	LDA + Bayes	Abo el zahad et al. [43]	A = 99.5%
		Abo el zahad et al. [46]	A= 91.5% EER = 3.2%
		Abo el zahad et al. [47]	A = 90.79% ERR =2.88%
		Abo el zahad et al. [48]	A = 92.71% EER = 2.13%
		Abo el zahad et al. [49]	A = 90.05% EER = 3.2% A = 92.82% ERR = 2.66 % A = 98.57% EER = 1.83 %
Machine Learning	Multi-Class SVM	Huy et al. [29]	Reduction of 4% for all EER
		Swati et al. [40]	A = 96% SVM
		Abo el zahad et al. [48]	SVM and it didn't show a good Performance in verification
		El-Sayed et al. [53]	A = 100%
	Random Forest	El-Sayed et al. [53]	RF proved to have a high accuracy performance in Time-frequency domain features
Neural Network	HMM, WNN	Guo et al.[30]	HMM is higher than GMM and the hybrid has a higher rate.
		Zhong L et al. [38]	HMM higher than GMM
	MLP-ANN	Chen W et al. [37]	A= 90.53% EER = 9.48%
		El-Sayed et al. [53]	A = 100%
	DNN	T.E Chen et al. [50]	A= 91%
Other Approaches	KSRC	Tan et al., [42]	A = 85.45%

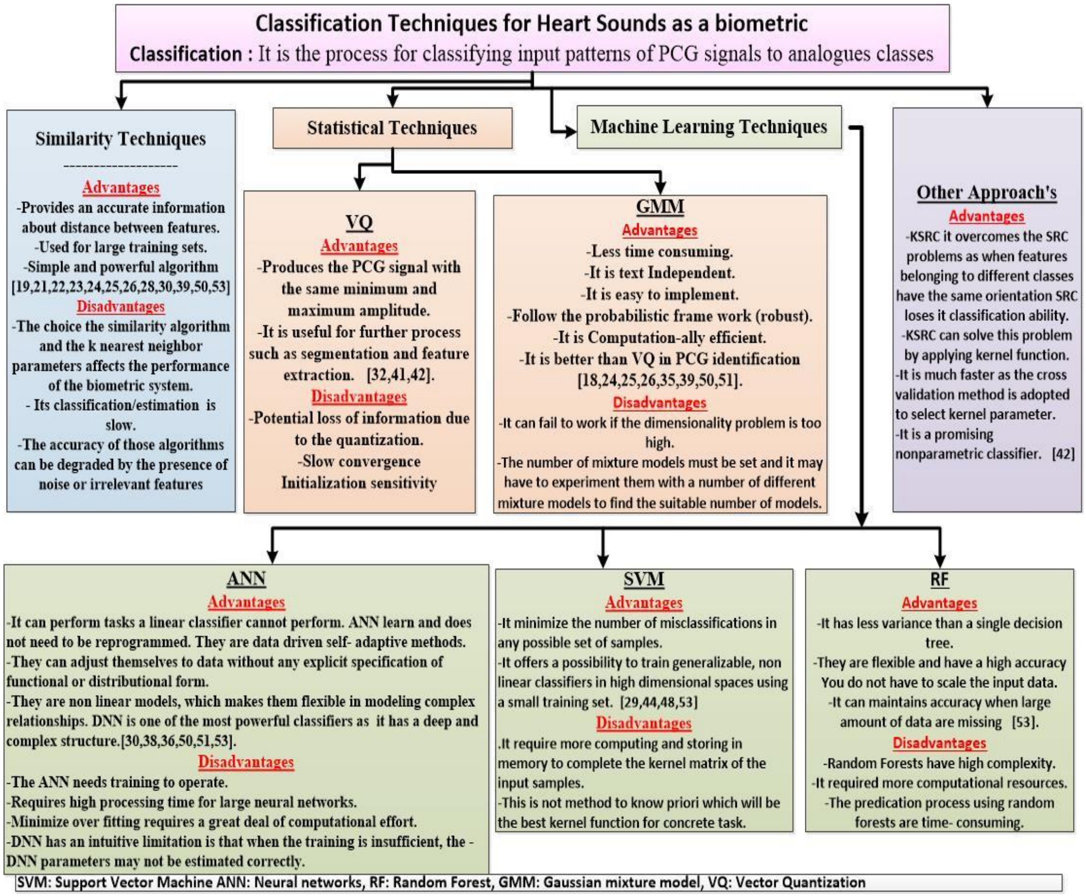

Fig. 11. and the disadvantages of the classification techniques used in PCG as a biometric

Some other techniques are used to form decision boundaries to obtain a certain error criterion. These techniques depend on neural network classifiers [74, 76]. Some studies used HMM to train the time sequence of the heart sounds and to compute the output score. Then, the score was used as an input and it made n nonlinear mapping by WNN to acquire the classification information [30, 38]. They proved that HMM-WNN is better in performance than GMM. Due to its simplicity and quick processing, MLP-ANN was considered a very popular classifier and was used by other studies. In addition, for the classification of speech and heart sounds recognition, MLP ANN is considered the preferred choice [36].

DNN is a strong trend recently used in classification. A study applied DNN with some hidden layers to strengthen the classification or the regression capability, and the standard back-propagation was applied to compute the parameters in DNN Model. It proved to be more efficient than other classifiers [50].

Aside from the methods of classification mentioned earlier, KSRC was applied as a non-parametric learning method in one of the studies. This method can perform a direct assignment of a class label to a test sample depending on a dictionary that is composed of training samples. Then, for changing the sample distribution, there is an application of the kernel tricks to the classifier. Furthermore, an application of mapping in the kernel feature space of high dimension into linear separable is conducted for a variation in the linear non-separable samples [42]. Table 5 presents the studies that investigated the classifiers which used in PCG as a biometric and the accuracy achieved from each classifier. Fig.11 Illustrates the most used classification techniques based on similarity, statistical, machine learning, and other approaches. It shows their advantages and their disadvantages and the reference of the paper used the classifier as well as its pros and cons based on the accuracy achieved.

6 Performance Measures

Testing the performance of any biometric system includes an open-set or a closed-set type. In the closet-set testing, the individuals enrolled are expected to have the ability to access the system, although this can hardly be guaranteed practically. The open-set testing has a focus on the presence of unknown subjects.

Its application is possible through plotting the probability distributions of consistent matching scores to the allowed individual and the impostor. Heart sounds biometric identity can be determined based on two main systems: identification and verification. A biometric identity identification model can be viewed by a set of feature vectors which are used to verify the nearest match in the database templates if the obtained distance to the nearest template is low.

A biometric identification system generates an error in the identification if the assigned class vector is not the true one. The biometric system that is used for identification can also use the Cumulative Match curve (CMC), and it draws the cumulative recognition rate as a mathematical function for ranking the recognition. When the closed-set testing is applied, there are no distributions of the score found to be evaluated. In the heart sound biometric system Correct Recognition rate (CRR) is the most common metric used for identification. A biometric identity verification system functions as a binary classifier. The systems that work using binary classification compare the matched score with a given threshold. The threshold is determined based on the context of the specified system. According to the chosen threshold, the accuracy is nearly linked.

Two main likely errors can be made using the binary classifier. The first error is the False Match error and it is a kind of error that happens when a method accepts an entity claims a match if the template matches with the template stored in the model. The second type of error is false non-match which is an error that rejects the entity claim even if the template matches with the template stored in the model. Depending on how the biometric systems operate, the importance of the error in the operational context varies. For instance in environments that depend on high security. False match error can be critical, while false non-match error could be tolerated. A threshold-independent approach is also needed to measure the performance of the heart sound identification models. A major problem with this is that we cannot know the applications in advance. The commonly used error performance in verification is Equal Error Rate (EER), False Match rate (FMR), and False Non-Match Rate (FNMR) as shown in Fig. 12.

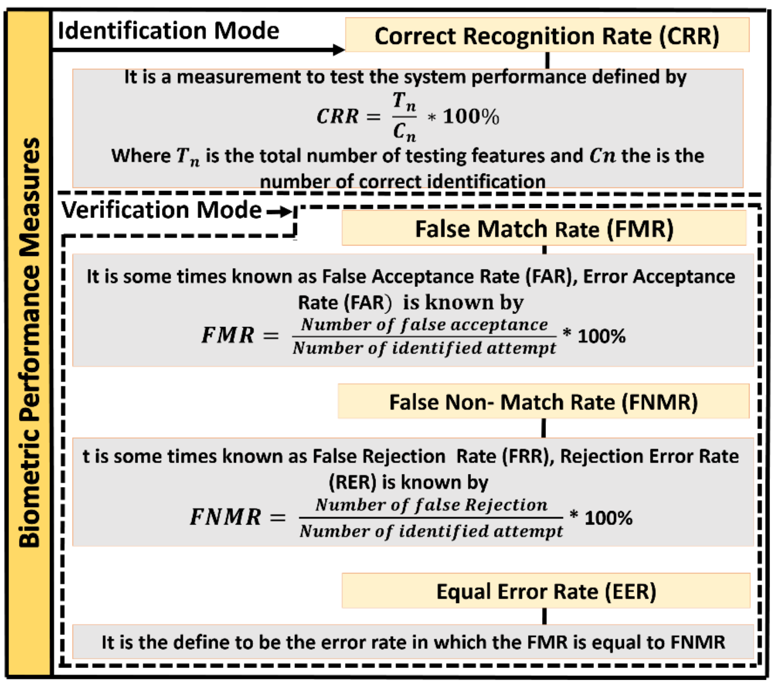

Fig. 12. The most commonly used performance measure in heart sounds as a biometric.

One of the ways used to evaluate the performance of the PCG biometric system is to plot the detection error tradeoff (DEF) curve. This curve is a relation between the FMR against FNMR. DET curves study the performance of the low FMR or FNMR introduced to the PCG biometric system. This curve is considered to be a relation between security and usability. On one hand, when the FMR is low for a specific system it means that it is highly secure, therefore it can result in a great number of non-matches. This kind of system may ask the user to try more than one authentication step. On the other hand, when the FNMR is low for the specific biometric system it means that the system will be more permissive and tolerant. This will result in a lot of false match errors and more unauthorized users to be accepted. To determine the correct choice between the two measures and the level of security that is intermediate is deemed to be application dependent. Table 6 presents the results of each study in the last ten decades. The table presents the year and the author of the paper, data sets used, pre-processing, segmentation techniques, feature extraction, classification methods, and the results. It presents the work display in the last ten years in the PCG signal as a biometric.

Fig. 13. Shows the effective parameters that affect the heart sounds biometric performance. It is characterized by 6 main causes which are: (data capturing, filtering, segmentation, feature gathering classification, and measurements evaluations) and 24 secondary causes divided into (2 for data acquisition), (5 for pre-processing), (6 for segmentation), (5 for feature extraction), (4 for classification) and (2 for measurements).

The process of placing the main causes in the upper zone of the lower zone of the fishbone was made according to some conditioning, meaning that each phase in the fishbone depends on the previous one for example; feature extraction or segmentation depends on its previous phase which is pre-processing and data acquisition and so on. The same principal tried to be respected for the secondary causes. For example for the segmentation to be performed it depends on its 6 secondary causes and the same for the main causes. The result of those causes will lead to achieving our goal or using PCG as a biometric.

Table 6. Survey of the most common techniques used in the PCG Identification

Year & Authors	DataSet	Pre-processing	Segmentation	Feature extraction	Classification	Results
2007 Beritelli et al., [19]	Heart songs 20 people	-	Autocorelation and hamming window	z-chirp CZT	ED	FRR = 5.0% and FAR = 2.2%
2008 Phua et al., [20]	10 people 1000 HS	Energy thresholding	Framing based on STFT FL = 256, FS = 256	LFBC	VQ GMM	GMM was higher than VQ with 60 reachs 96%
2008 Beritelli et al., [21]	70 people	-	Autocorelation and hamming window	z-chirp CZT applied on each sub-bands (S1-S2)	ED	EER = 9%
2009 Beritelli et al., [22]	50 people	Low pass filter	Autocorelation and hamming window	13 cofficients from MFCC + FSR	ED	EER < 9% ERR = 8.70%
2009 Beritelli et al., [23]	40 people	Low pass filter	Autocorrelation and hamming window	13 coefficients from MFCC + FSR	ED	EER = 5%
2010	21	Wavelets	Framing based	MFCC +	ED The	A = 100%

Fateman et al. [24, 25]	subjects	using db5	onSTFT FL = 250,500,1000ms	LDA	Distance Threshold = 6,8,10	
2010 Beritelli et al., [26]	165 People	-	-	A Tool called Sfbcep that performs filter-bank cepstral analysis	GMM	EER = 13.70%
2010 Beritelli et al., [27]	147 People	-	-	A Tool called Sfbcep that performs filter-bank cepstral analysis	GMM	ERR = 15.53%
2010 Tao et al., [28]	5-100 people	-	Autocorrelation and hamming window	Fusion between cycle, power, frequency, and drawing features	Similarity distance	A Close to 99%
2010 Huy et al., [29]	52 users	-	Autocorrelation and hamming window	8 feature sets + RFE for feature selection +	First experiment: using 8 feature sets + SVM without selection. Second experiment: using 8 feature sets + SVM with RFE selection	Two experiments were applied 1st experiment A was over 80% for GMM and LFCC features 2nd experiment A was over 90% for GMM features
2010 Guo et al., [30]	160 heart sounds from 80 subjects	-	-	LPCC	WNN+ HMM	Better than GMM
2010 Jasper et al., [31]	10	Low pass filter	-	DWT decomposition + selecting appropriate bands + Shannon energy	Template matching	98.67% with Shannon energy 77.33% without Shannon energy
2011 Cheng Xie et al., [32]	10	-	-	ICC + Independed sub-band function	Similiarty distance	A = 85.7%, EAR < 7%, RER < 10%
2011 Zhao et al., [33]	30	Normalization	hamming window	MFCC	VQ	A= 100%
2012 Cheng Xie [34]	300 heart sounds	-	Wavelet Family	LBFC	Similarity distance	Verification: 12 Heart sounds signal for train, 12 Heartsounds signals for test A = 100% Identificaiton: EAR < 1-8%, ERR < 3%, A = 99%
2012 Rasha Wahid et	80 heart sound samples	-	FFT using hamming window FL = 256 ms	FEAL1: MFCC + DCT FEAL2 : Spectral magnitude + DCT	GMM	FEAL1 : A = 100% for 7 samples FEAL2 :

al., [35]			STFT FL = 256 ms			A = 100% for 6 samples
2012 Chen W et al. [36]	-	Wavelet transform	-	MFCC PCA	-	A could reach 90%
2012 Karmakar et al. [37]	-	Low pass filter	-	Wavelet and windowed 2nd level coefficients	MLP	96.178%
2013 Zhong L et al., [38]	100 heart sounds From 50 people	Wavelet transform	-	LPCC MFCC	GMM	LFCC is more suitable than MFCC
2013 Spadacci ni et al., [39]	HSCT-11 206 people	Low pass filter	Cross-correlation windowing Computing the variance of each segment	Structural system : MFCC + FSR Statistical System: LFCC + FSR	Structural system: Template matching Statistical System: GMM	Structural system ERR = 36.86% Statistical system ERR = 13.66%
2013 Zhao et al., [40]	40 Subjects 280 heart sounds	DWT using db5 family	Hamming Hanning Rectangular Hamming was the best	FS (Fourier Spectrum) MS (Marginal Spectrum)	VQ	For FS A= 84.93% For MS A = 94.16%
2013 Girish et al., [41]	10 4000 PCG Samples	Normalization Low pass filter	Autocorrelation + segmentation using thresholding	LFBCC Wavelet decomposition using db2	MLP-ANN	LFBCC with A = 89.68% Wavelet with A = 90.52%
2014 Tan et al., [42]	15 Subjects	Low pass filter	ZCR+STA “zero-crossing rate and short-term amplitude”.	MFCC DCT	KSRC SVM KNN SRC	A = 85.45% A = 84.87% A= 84.57% A = 78.78%
2014 Abo el zahad et al., [43]	HSCT-11 17 Subjects	DWT using db5 + Thresholding of wavelet coefficients	Hamming window	MFCC, BFCC, LFCC, DW-MFCC + Fusion between them using CCA	LDA + GMM with kmeans with a decision-based Bayes theory	A= 94.4%,A=94.325 %,A=93.7% A=95.12%, A of parallel fusion was the best between MFCC+DW-MFCC features A = 99.5%
2014 Swati et al., [44]	30 subjects	High pass filter	Framing using Hamming window	MFCC DCT	SVM	A = 96%
2015 S. Bindu et al., [45]	-	TVD	-	Signal Energy envelope	Template matching	-
2015 Abo el zahad et al., [46,47]	HSCT-11 206 people	DWT using db5 and 4 thresholding techniques	Framing using hamming window FS = 1000ms, FS =500ms	WPCC using linear and non-linear filters + LDA Based on different wavelets the best was a demy	Bayes	The best accuracy achieved using WPCC with non-linear filtering reaching A= 91.05%

2016 Abo el zahad et al., [48]	HSCT-11 206 people Bio-Sec 21 people	DWT using db5 and 4 thresholding techniques	Framing using hamming window FS = 1000ms, FS =500ms	MFCC,LFCC,NLFCC,WPCC + LDA Based on different wavelets the best was demy	Bayes	HSCT-11 database A = 91.61%, A=91.15, A=92.51 and 90.26% Bio-Sec database A=97.31%, A= 96.94%,A = 97.02% andA=98.05%
2016 Abo el zahad et al., [49]	HSCT-11 206 people Bio-Sec 21 people	DWT using db5 and 4 thresholding techniques	Framing using hamming window FS = 1000ms, FS =500ms	MFCC,LFCC,MMFCC, WPCC + LDA Based on different wavelets the best was demy	Bayes	HSCT-11 database A = 91.15%, A=91.61, A=92.82 and 90.26% Bio-Sec database A=96.94%, A= 97.31%,A = 98.57% and A=97.02%
2017 T.E Chen et al. [50]	16 people Total of 616 HS	-	Heart sound activity detection based on SEE	MFCC + kmeans	DNN, KNN, LR SVM GMM	A= 91.12% A= 78.11% A= 87.57% A= 90.53% A=86.98%
2018 TG Meitei et al., [51]	-	Wavelets	-	-	ED, GMM, FSR, and VQ	-
2019 Fahad et al. [52]	50	DWT	Hilbert modeling	AR burg modeling	Bagged decision Tree	A= 86.7%
2019 El-dahshan et al. [53]	60 from HSCTI 50 from PASCAL	MRD-MRR	Framing and windowing, Shannon energy envelope	MS_DWT	RF ANN SVM KNN	A = 100% using SVM with (Db9) on 60 Subjects A = 100% using ANN with (Db10) on 50 Subjects
2020 Cheng, X et al. [54]	80 HS from 40 subjects	-	IMF + multiscale dispersion entropy	LR HSMM + FR for reduction	ED	96.08%

7 Discussion, Limitations, and Challenges

There have been contributions from diverse research papers to the field of automated verification and identification of the PCG signals as a biometric recorded in heart sound form. In this examination, 29 publications that focus on heart sounds as a biometric are taken into consideration. This literature was published during the years from 2006 until 2020. The yearly distribution of 35 articles (13 journals, 20 conferences, and 2 theses) together with the repetition of the keywords (biometric and heart sound) in the title of the paper is shown in table 7.

Table 7. shows the number of manuscripts used to review PCG as a biometric

Year	No. of Paper Journals	No. of Paper Conferences	No. of Thesis	Total No. of publications	Keyword	
					Biometric	Heart Sound
2006	-	1	-	1	-	1
2007	1	-	-	1	1	-
2008	1	1	-	2	1	1
2009	-	2	-	2	-	1
2010	-	7	1	8	5	6
2011	1	2	-	3	1	3
2012	2	-	1	3	-	1
2013	1	2	-	3	3	3
2014	1	1	-	2	3	1
2015	1	2	-	3	4	2
2016	2	1	-	3	3	1
2017	1	-	-	1	-	-
2018	1	-	-	1	2	1
2019	1	1	-	2	-	-
Total	**13**	**20**	**2**	**35**	**23**	**21**

The data acquisition phase is considered an important step in the phases of PCG authentication. From the survey, not many datasets were presented due to the limitations in the number of them in the applications of PCG biometric. We recommend the researchers publish their datasets online to encourage the development of different techniques.

The pre-processing phase is a critical step in the process of the PCG signal as a biometric. Therefore, almost all of the filtering methods mentioned in the study and their combinations can be performed to increase the performances. Studies about the filtering techniques focused on down-sampling, low-pass filtering, energy-thresholding, normalization, total variation de-noising, and wavelets. The filter type can change depending on the type of noise in the PCG signal. Studies showed several techniques used in the segmentation phase. Those techniques are based on framing and windowing, auto-correlation, Shannon energy envelope, and zero-crossing rate with short-term amplitude. The studies showed that this step is very important for detecting the heart sounds in the PCG signal before extracting any features. The segmentation technique should be chosen based on the appropriate context and with the relevant advantages and disadvantages. The focus should also be placed on the concentration of S1 and S2 and without removing any important information from the PCG signal.

Although researchers have proposed different types of feature extraction for PCG signals, there is no one size fits all. The feature extraction in the studies was based on the time, frequency, time, and frequency domain features. We recommend using time-frequency features, frequency, and then the time domain respectively in the efficiency. Some studies fused between frequency and time-frequency features and achieved high accuracy on a large number of subjects.

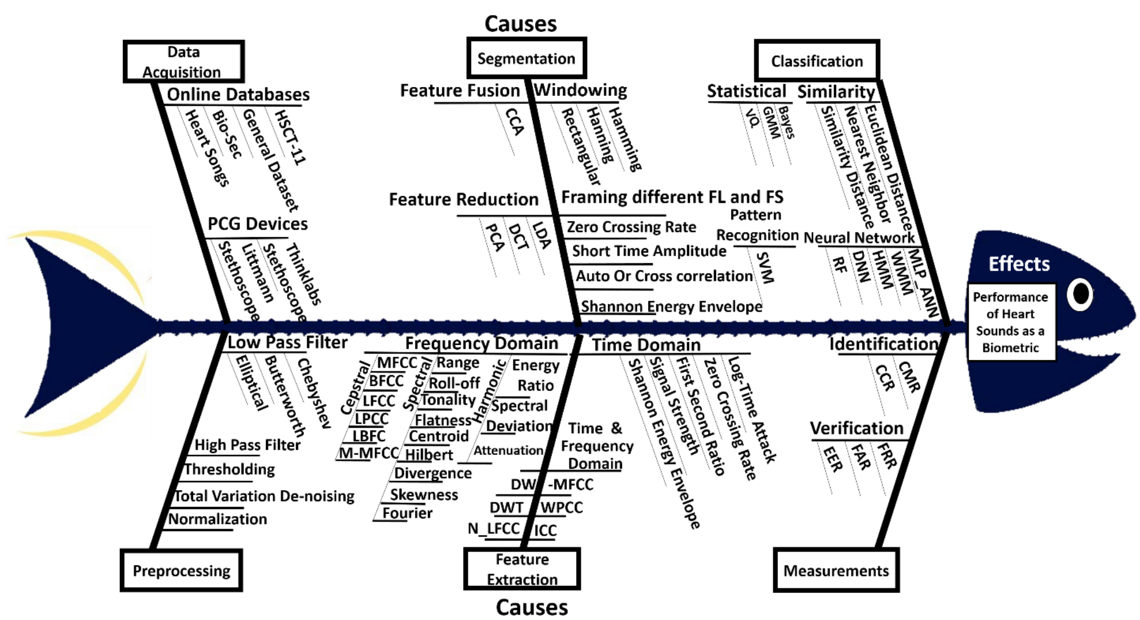

Fig. 13. Fishbone diagram with the parameters that affect the heart sounds as a biometric

There are many different conditions to decide the feature extraction techniques to be used, for example, the de-noising, processing time expectations, classification, and the dimensionality of the feature space. So, one should choose what features to be used by applying all these factors.

To overcome this problem feature transformation or selection should be applied as it is deemed to be a salient step. The process transforms the original features space into a lower-dimensionality sub-space some feature selection methods focus on producing the optimal solution, while others might give suboptimal results. So there is a tradeoff between the processing time and the optimality. Different success measures for PCG as a biometric and classification based on identification and verification. For identification, CRR can be used, and for verification, common methods include EER, FMR, FAR, EAR, FNMR, FRR, and RER. EER is considered to be the most often used measurement among all these measures in verification. However, most of the researchers prefer to present their results with more than one measurement due to the non-stationary nature of the PCG signal.

There are different challenges and limitations in the use of heart sounds as a biometric. Some are discussed as below:

1- The efficient segmentation of the essential PCG S1 and S2 sounds is one of the main tasks of phonocardiography. The detection and classification of the heart sound signal automatically constitute the challenge in the heart sound signal. The techniques used in the segmentation process of PCG signals do not give an accurate accuracy of extracting S1 and S2 from heart sound signals [77]. The classification of the heart sounds for the individuals as a biometric in the identification or verification phase did not reach satisfying accuracy.

2- It is possible to evaluate and compare different feature extraction techniques. To increase the system's performance, especially in the case of fusion with ECG, the techniques of feature selection and fusion can be adopted [10-78]. Furthermore, multimodal biometrics can be used and combined with PCG to improve the strength of the biometric system [79]. There are merits and downsides for each type of biometric. Hence, multi-model biometric fusion is introduced to improve identification performance. It is deemed that this model is a new application of information fusion. Information fusion can be defined as a mixture between different data sources for the generation of one format or taking an accurate decision. This model is considered to be a new application of information fusion. Three variant levels can be used for carrying out the fusion between ECG and PCG: Data-level, feature-level, and decision-level fusion. A study has investigated the multimodal biometric system's performance in the identification mode [24]. The evaluation is tested using data collected from ECG and PCG signals of 21 subjects. The results showed a recognition rate of 98.4%. Hence, it is proved that fusion between them has gained higher accuracy than each biometric alone.

3- In the feature extraction phase, most of the research committees applied wavelet-based methods and cepstral coefficient over other features. Wavelet proved its ability to provide accurate results in extraction during the PCG analysis as a biometric [80-83]. A valid example is the wavelets can define the local change in the heart sound cycle. Moreover, the high-frequency components need to be examined in the PCG signal, and wavelets proved to be more localized at the higher parts of frequency, therefore, wavelets proved to be better than FFT [84, 85]. In contrast, it is difficult for wavelets to experience a local event that occurs at the local frequency range. On the contrary, Cepstral coefficients are not efficient in the case of the existence of additive noise, therefore, it is common to perform a normalization operation to the values of the PCG signals to decrease the presence of this noise.

4- The ability of EMD and empirical wavelet transform (EWT) is proven in overcoming the non-adaptability of the wavelets. Their performance in feature extraction leads to better results. Some further studies work on obtaining a group of well-defined and accepted features for the heart sound based on their clinical and statistical significance.

5- ANN is a traditional and popular classifier that is used for the classification of many different PCG signals as a biometric, but the ANN suffers from a lot of difficulties such as the problem of over-fitting, the increase in the complexity because of the redundancy in the features and the hidden noises. Research communities should explore the performance of the hybridization and fusion between classifiers and adding some meta-heuristic techniques for optimal decisions [86].

6- This study also discusses that three main variant classification concepts. The first concept is based on similarity. The second concept is based on a probabilistic or statistical approach. The third concept is based on the construction of decision boundaries by optimizing certain error criteria. There is a need to apply other classification approaches to the heart sound features and show their performances [55].

7- A challenging research topic in the field of biometrics is using PCG signals for biometry. Several works have been presented so far on this topic by the academic community. However, the problem with most of them is that the evaluation is conducted on small databases. Thus, the results obtained are difficult to be generalized. Only one large database is used for PCG as a biometric [58]. Upon the availability of larger databases of heart sounds for the scientific research communities, future research will need to address a lot of issues and challenges [87– 90].

8- For more conventional biometrics (fingerprint, face), heart sounds disorders cannot be considered a damaging factor in some cases. They can only cause limitations in the PCG biometrics methods. A range of disorders can emerge, from isolated irregularity to severe conditions where immediate medical assistance is required [91-93].

   A move toward deep learning techniques is evident in the incline in the pattern recognition fields. There is a lot of specialties involved. Handcrafting feature extraction method is specifically avoided by locating discriminating regularities in the raw data. It has a good performance with large and diverse datasets. Furthermore, the performance is also good in practical settings because a lot of data with large variations are produced by the clinical routine. Another main point to be considered is that our survey yielded only 29 papers. To be specific, one paper discussed deep learning as a future work and its effect on the accuracy of authentication of the individual using heart sounds. In several pattern recognition domains, deep learning has become the state of art. Likewise, may also be useful in PCG recognition [94-96].

9- The identification performance is very low for larger datasets. This will entail matching algorithms that will be fine-adjusted with a suitable feature set that can be properly identified via the combination of elements from both frequency and time domains. Then, there will be an assessment of the mid-term and long-term reliability of heart sound signals, where the variations in the heart acoustic signal as biometric as time progresses will be analyzed. It is expected from the community to exert effort in developing systems and algorithms for the heart- sounds biometry. Furthermore, common databases need to be created for evaluating different research approaches over a shared dataset. These databases will allow their performances to be compared to refine them over time. Methods and approaches that might be deployed in real-life scenarios also need to be developed [97].

10- Some changes must be made on the databases that are collected for the PCG biometric because most of them are gathered at the same time, therefore, the results will not be accurate for the stability of the heart sound signals over time. This poses a challenge in obtaining the characteristic of permanence as it is considered to be one of the most important features of a biometric trait. That is why a new different dataset should be contrasted with a great time duration between the testing and training heart sounds recordings. In addition to this, the datasets are gathered under rest conditions, and the research communities in the future should focus on capturing heart sound under variant physical activities. Many other factors can affect the performance of the heart sound biometric verification systems such as age, gender, and heart rate, and it needs more methods and study to match these factors [98].

11- Some practical issues like ad-hoc sensors with embedded matching algorithms and computational efficiency must be addressed. This can be achieved using algorithms more sophisticated and different self-determining feedback as this will give a positive evaluation of the notion. This will lead to the advancement of PCG biometry and it can act as an alternative or supportive solution to many biometrics methods [99 – 100].
12- ECG signals can be captured when single session and multi-session data, and most of the ECG data obtained are based on single-session from public repositories, clinical sources, or using custom free-living protocols. The databases available for heart sounds are obtained at the same session, and there are no datasets available for PCG signals captured at multi-session data. More studies need to be performed to determine whether if a single or multi-session is better for authentication [101].
13- Lastly, one main limitation started to be apparent in the last two years. The available datasets for PCG as a biometric online started to be not accessible on the servers. Also, the research papers about PCG as a biometric in the last two years started to become sparse.

8 Application Fields for PCG as a Biometric

PCG signals as a biometric can be used in several applications. It is significant to correctly and efficiently identify the user in many applications used within the defense, security, finance, airport, hospitals, and personal identity industries. The features of the PCG signals allow the improvement of different and motivating applications, where nonstop authentication is a critical factor. There are valid examples of that such as electronic trading platforms in which more security and permanent authentication are required, the gaming industry in which the PCG sensor can be used to verify the players in a multi-player mode, the auto industry for car-sharing programs and fleet management applications. Moreover, PCG signals can be used to detect abnormality and heart defects [102-106], heartbeat detection [107-110], and emotion recognition [111].

9 Conclusions

With the advancement of computer intelligence and machine learning, biometrics attracts more attention as a means of authentication as it is used in many different applications. It has become one of the major research areas in security traits and verification. There are different types of biometrics and we focused on a behavior biometric namely the use of heart sounds due to its advantages. In this paper, we provide an intensive review and survey of research on heart sounds described in the work. The steps of processing the heart sound signal as a biometric are organized into data acquisition, filtering, segmentation, feature extraction, and classification and evaluation. In the data acquisition, we discussed the most commonly used datasets in our survey and it is shown there are limited datasets available for the heart sounds. Also, the presented datasets do not discern between the ages, gender, and size. The techniques used in pre-processing based on the survey were categorized into down-sampling, low-pass filtering, energy-thresholding, normalization, total variation de-noising, and wavelets. It is concluded from the survey that the wavelets and total variation de-noising are much better than the conventional filters in de-noising the heart sound signal. This is because it is considered to be non-stationary signals especially total variation de-noising that can be generalized for general signals.

From the survey, the techniques used in the segmentation were based on framing and windowing, auto-correlation, Shannon energy envelope, and zero-crossing rate with short-term amplitude. Each technique has its advantages and disadvantages. Therefore the selection of the segmentation method depends on what types of features are needed after segmentation. In the feature extraction stage, we reviewed all the techniques used in heart sounds as a biometric. Those techniques were divided into three domains, time-domain features, frequency domain features, and both time and frequency domain features. It is observed from the survey that fusion between those feature domains with a large number of subjects achieves the lowest EER and FAR.

The final stage is a classification based on verification and identification of the individual using a heart sounds signal. From the survey, most of the classification techniques were based on similarity, statistical, pattern recognition approaches and we have also considered the traditional classifiers. There are some improved classifiers such as the DNN approach and SRC that are discussed in two papers in the survey. Those classifiers can improve the performance and increase the accuracy of the heart sound biometric system. The challenge remains that there is a need to provide a generalized heart sound biometric system regardless of data size and quality.

References

1. A. K. Jain, A. Ross, S. Prabhakar, An introduction to biometric recognition, IEEE. Trans. Circuits. Syst. Video. Technol. 14 (1) (2004) 4-20, **doi:** 10.1109/TCSVT.2003.818349.
2. S. Nanavati, M. Thieme, R Nanavati, Biometric’s Identity Verification in a Networked World, John Wiley and Sons Inc./Wiley Computer Publication, 2003.

3. P. Karczmarek, A. Kiersztyn, W. Pedrycz, M.Dolecki, An application of chain code-based local descriptor and its extension to face recognition, Pattern Recognit. 65 (2017) 26-34, **doi**:10.1016/j.patcog.2016.12.008.
4. A. Morales, A. Kumar, M. A Ferrer, Interdigital palm region for biometric identification. Compute. Vis. Image. Und. 142 (2016) 125-133, **doi**: 10.1016/j.cviu.2015.09.004.
5. K. W. Bowyer, M. J. Burge, 2016. Handbook of iris recognition. Springer London, **doi**: 10.1007/978-1-4471-4402-1.
6. J. Kim, H. Kim, P. Kang, Keystroke dynamics-based user authentication using freely typed text based on user-adaptive feature extraction and novelty detection, Appl. Soft Comput. 62 (2017) 1077-1087, **doi**: 10.1016/j.ins.2014.08.070.
7. J. Yang, J. Chen, Y. Su, Q. Jing, Z. Li, F. Yi, Z. L. Wang, Eardrum-Inspired Active Sensors for Self-Powered Cardiovascular System Characterization and Throat-Attached Anti-Interference Voice Recognition. J. Adv. Mater. 27(8) (2015) 1316-1326, **doi**:10.1002/adma.201404794.
8. J. Galbally, M. Diaz-Cabrera, M. A. Ferrer, M. Gomez-Barrero, A. Morales, J. Fierrez, On-line signature recognition through the combination of real dynamic data and synthetically generated static data. Pattern Recognit. 48(9) (2015) 2921-2934, **doi**: 10.1016/j.patcog.2015.03.019.
9. Y. Zhang, Z. Dong, G. Ji, S. Wang, Effect of spider-web-plot in MR brain image classification, Pattern Recogn Lett. *62* (2015) 14-16, **doi**: 10.1016/j.patrec.2015.04.016.
10. S. K. Berkaya, A. K. Uysal, E. S. Gunal, S. Ergin, S. Gunal, M. B. Gulmezoglu, A survey on ECG analysis. Biomed Signal Process. 43 (2018) 216-235.
11. M. Hejazi, S. A. R. Al-Haddad, Y. P. Singh, S. J. Hashim, and A. F. A. Aziz. ECG biometric authentication based on non-fiducial approach using kernel methods. Digit. Signal Process. 52 (2016) 72-86.
12. J. Herzig, A. Bickel, A. Eitan, N. Intrator, Monitoring cardiac stress using features extracted from S1 heart sounds. IEEE. Trans. Biomed. Eng. 62(4) (2015) 1169-1178.
13. S. M. Debbal, F. Bereksi, Computerized heart sounds analysis. Comput. Biol. Med. 38, (2008), 263–280.
14. Difference between systolic and diastolic Available from: http://www.differencebetween.net/science/health/difference-between-systolic-and diastolic/.
15. V. S. Chourasia, A. K. Tiwari, and R. Gangopadhyay, A novel approach for phonocardiographic signals processing to make possible fetal heart rate evaluations. Digit. Signal Process. , 30 (2014) 165-183.
16. S. Geetha, A.V. Phamila, Combating Security Breaches and Criminal Activity in the Digital Sphere, **doi**: 10.4018/978-1-5225-0193-0. IGI Global, 2016.
17. E. J. Kindt, Privacy and data protection issues of biometric applications (Vol. 1). Springer, 2016. **doi**: 10.1007/978-94-007-7522-0.
18. K. Phua, T. H. Dat, J. Chen, L. Shue, (2006, May). Human identification using heart sound. In Second International Workshop on Multimodal User Authentication, Toulouse, France 227(10), **doi**: 10.1.1.94.3041.
19. F. Beritelli, S. Serrano, Biometric identification based on frequency analysis of cardiac sounds. IEEE T Inf Foren Sec, 2(3) (2007) 596-604, **doi**:10.1109/TIFS.2007.902922.
20. K. Phua, J. Chen, T. H. Dat, L. Shue, Heart sound as a biometric. Pattern Recognit. 41(3) (2008) 906-919, **doi**:10.1016/j.patcog.2007.07.018.
21. F. Beritelli, A multiband approach to human identity verification based on phonocardiogram signal analysis. In *Biometrics Symposium, 2008. BSYM'08* (pp. 71-76). IEEE, **doi**: 10.1109/BSYM.2008.4655525.

22. F. Beritelli, A. Spadaccini, Human identity verification based on mel frequency analysis of digital heart sounds. In *Digital Signal Processing, 2009 16th International Conference on* (pp. 1-5). IEEE, **doi**: 10.1109/ICDSP.2009.5201109.
23. F. Beritelli, A. Spadaccini, (2009, December). Heart sounds quality analysis for automatic cardiac biometry applications. In *Information Forensics and Security, 2009. WIFS 2009. First IEEE International Workshop on* (pp. 61-65). IEEE, **doi**: 10.1109/WIFS.2009.5386481.
24. S. Z. Fatemian, (2009). A wavelet-based approach to electrocardiogram (ECG) and phonocardiogram (PCG) subject recognition. University of Toronto.
25. S. Z. Fatemian, F. Agrafioti, D. Hatzinakos, (2010, September). HeartID: Cardiac biometric recognition. In *Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on* (pp. 1-5). IEEE, **doi**: 10.1109/BTAS.2010.5634493.
26. F. Beritelli, A. Spadaccini, (2010, September). An improved biometric identification system based on heart sounds and gaussian mixture models. In *Biometric Measurements and Systems for Security and Medical Applications (BIOMS), 2010 IEEE Workshop on* (pp. 31-35). IEEE **doi**:10.1109/BIOMS.2010.561044.
27. F. Beritelli, A. Spadaccini, (2010, July). A statistical approach to biometric identity verification based on heart sounds. In *2010 Fourth International Conference on Emerging Security Information, Systems and Technologies* (pp. 93-96). IEEE, **doi**: 10.1109/SECURWARE.2010.23.
28. T. Ye-wei, S. Xia, Z. Hui-xiang, W. Wei, (2010, May). A biometric identification system based on heart sound signal. In *Human System Interactions (HSI), 2010 3rd Conference on* (pp. 67-75). IEEE, **doi**: 10.1109/HSI.2010.5514589.
29. D. H. Tran, Y. R. Leng, H. Li, (2010, March). Feature integration for heart sound biometrics. In *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on* (pp. 1714-1717). IEEE, **doi**:10.1109/ICASSP.2010.5495476 .
30. X. M. Guo, Y. Duan, L. S. Zhong, (2010). Study of human identity recognition based on HMM and WNN. *Jisuanji Yingyong Yanjiu*, *27*(12), 4561-4564, **doi**:10.3923/jse.2016.408.415.
31. J. Jasper, K. R. Othman, (2010, August). Feature extraction for human identification based on envelogram signal analysis of cardiac sounds in time-frequency domain. In *Electronics and Information Engineering (ICEIE), 2010 International Conference On* (Vol. 2, pp. V2-228). IEEE, **doi**: 10.1109/ICEIE.2010.5559770.
32. X. F. Cheng, Y. W. Tao, Z. J. Huang, Heart sound recognition-A prospective candidate for biometric identification. J. Adv. Mater. Res. 225 (2011) 433-436.Z.
33. Zhao, J. Wang, (2011, September). Heart sound identification system. In *Electronics, Communications and Control (ICECC), 2011 International Conference on* (pp. 2079-2082). IEEE, doi:10.1109/ICECC.2011.6067675.
34. X. Cheng, Y. Ma, C. Liu, X. Zhang, Y. Guo, Research on heart sound identification technology. Sci. China Technol Sci. 55(2) (2012) 281-292, **doi**: 10.1007/s11432-011-4456-8.
35. R. Wahid, N. I. Ghali, H. S. Own, T. H. Kim, A. E Hassanien (2012). A Gaussian mixture models approach to human heart signal verification using different feature extraction algorithms. In *Computer Applications for Bio-technology, Multimedia, and Ubiquitous City* (pp. 16-24). Springer, Berlin, Heidelberg, **doi**: 10.1007/978-3-642-35521-9_3.
36. W. Chen, Y. Zhao, S. Lei, Z. Zhao, M. Pan, Study of biometric identification of heart sound base on Mel-Frequency cepstrum coefficient. J. Biomed. Eng 29(6), (2012) 1015-1020, **doi**: 10.1109/ICDSP.2009.5201109.
37. A. Karmakar, S. Ari, Biometric identification and verification based on time-frequency analysis of phonocardiogram signal. M.S. thesis, Department of Electronics and Communication Engineering, National Institute of Technology, Rourkela, India, June (2012).

38. Zhong, L., Wan, J., Huang, Z., Guo, X., Duan, Y. (2013). Research on biometric method of heart sound signal based on GMM. Chin Med. J. 37(2), 92-5.
39. A. Spadaccini, F. Beritelli, (2013, July). Performance evaluation of heart sounds biometric systems on an open dataset. In *Digital Signal Processing (DSP), 2013 18th International Conference on* (pp. 1-5). IEEE, **doi**: 10.1109/ICDSP.2013.6622835.
40. Z. Zhao, Q. Shen, F. Ren, Heart sound biometric system based on marginal spectrum analysis. Sensors (Basel), 13(2) (2013) 2530-2551, **doi**: 10.3390/s130202530.
41. G. Gautam, D. Kumar (2013, April). Biometric system from heart sound using wavelet based feature set. In *Communications and Signal Processing (ICCSP), 2013 International Conference on* (pp. 551-555). IEEE, **doi**:10.1109/iccsp.2013.6577115.
42. W. C. Tan, H. M. Yeap, K. J. Chee, D. A. Ramli, (2014).Towards real time implementation of sparse representation classifier (SRC) based heartbeat biometric system. In *Computational problems in engineering* (pp. 189-202). Springer, Cham, **doi**: 10.1007/978-3-319-03967-1_15.
43. M. Abo-Zahhad, S. M. Ahmed, S. N. Abbas, (2014, May). PCG biometric identification system based on feature level fusion using canonical correlation analysis. In *Electrical and Computer Engineering (CCECE), 2014 IEEE 27th Canadian Conference on* (pp. 1-6). IEEE, **doi**: 10.1109/CCECE.2014.6901068.
44. S. Verma, T. Kashyap, Analysis of heart sound as biometric using mfcc & linear SVM classifier. IJAREEIE. 3(1) (2014) 6626 - 6633.
45. S. B. B. Kumar, M. Jagannath, (2015, January). Analysis of phonocardiogram signal for biometric identification system. In *Pervasive Computing (ICPC), 2015 International Conference on* (pp. 1-4). IEEE **doi**: 10.1109/PERVASIVE.2015.7086994.
46. M. Abo-Zahhad, S. M. Ahmed, S. N. Abbas, (2015, December). A new biometric authentication system using heart sounds based on wavelet packet features. In *Electronics, Circuits, and Systems (ICECS), 2015 IEEE International Conference on* (pp. 17-20). IEEE, **doi**: 10.1109/ICECS.2015.7440238.
47. M. Abo-Zahhad, S. M. Ahmed, S. N. Abbas, Biometrics from heart sounds: Evaluation of a new approach based on wavelet packet cepstral features using HSCT-11 database. Comput Electr Eng, 53 (2016) 346-358 **doi**: 10.1109/ICECS.2015.7440238.
48. M. Abo-Zahhad, M. Farrag, S. N. Abbas, S. M. Ahmed, A comparative approach between cepstral features for human authentication using heart sounds. Signal Image Video P. 10(5), (2016) 843-85, **doi**: 10.1007/s11760-015-0826-9.
49. S. N. Abbas, M. Abo–Zahhad, S. M. Ahmed, M. Farrag, Heart-ID: human identity recognition using heart sounds based on modifying mel-frequency cepstral features. IET Biometrics. 5(4) (2016) 284-296, **doi**: 10.1049/iet-bmt.2015.0033.
50. T. E. Chen, S. I. Yang, L. T. Ho, K. H. Tsai, Y. H. Chen, Y. F. Chang, C. C. Wu, S1 and S2 heart sound recognition using deep neural networks. IEEE T Bio- Med Eng 64(2) (2017) 372-380, **doi**: 10.1109/TBME.2016.2559800.
51. T. G. Meitei, S.A. Singh, S. Majumder, (2018). PCG-Based Biometric. *In Handbook of Research on Information Security in Biomedical Signal Processing* (pp. 1-25). IGI Globel.
52. I. Fahad, M. A. R Apu, A., Ghosh, and S. A. Fattah, (2019, October). Phonocardiogram Heartbeat Segmentation and Autoregressive Modeling for Person Identification. In *TENCON 2019-2019 IEEE Region 10 Conference (TENCON)* (pp. 1942-1946). IEEE.
53. El-Dahshan, E. S. A., & Bassiouni, M. M. (2020). Intelligent methodologies for cardiac sound signals analysis and characterization in cepstrum and time-scale domains. *Comput. Intell.*, *36*(2), 427-458.
54. Cheng, X., Wang, P., & She, C. (2020). Biometric identification method for heart sound based on multimodal multiscale dispersion entropy. *Entropy*, *22*(2), 238.

55. M. Abo-Zahhad, S. M. Ahmed, S. N. Abbas, Biometric authentication based on PCG and ECG signals: present status and future directions. Signal Image Video P, 8(4) (2014), 739-751.
56. Heart Song database Available: http://www.cardiosource.com/heartsounds/index.asp .
57. PCG Dataset Thinklabs Rhythm Digital Electronic Stethoscope (ds32a), http://www.thinklsbamedical.com/.
58. BioSec. PCG database Available at http://www.comm.utoronto.ca/~biometrics/databases.html. accessed 1-12-2017.
59. HSCT11 database: http://www.diit.unict.it/hsct11/ last accessed on 3-7-2021.
60. Pascal database: http://www.peterjbentley.com/heartchallenge/ last accessed on 3-7-2021.
61. I. Daubechies, The wavelet transform, time-frequency localization and signal analysis. IEEE Trans, 36(5) (1990) 961-1005, **doi**: 10.1109/18.57199.
62. K. Revett, F. Gorunescu, A. B. Salem, and E. S. El-Dahshan, Evaluation of the feature space of an erythematosquamous dataset using rough sets. Annals of the University of Craiova-Mathematics and Computer Science Series, 36(2) (2009), 123-130.
63. A. Moukadem, Z. Bouguila, D. O. Abdeslam, and A. Dieterlen. A new optimized Stockwell transform applied on synthetic and real non-stationary signals. Digit. Signal Process, 46 (2015) 226-238.
64. B. Boashash, N.A. Khan and T. Ben-Jabeur. Time–frequency features for pattern recognition using high-resolution TFDs: A tutorial review. Digit. Signal Process, 40 (2015) 1-30.
65. I. Daubechies, Ten lectures on wavelets. 61 (1992) Siam. , https://doi.org/10.1137/1.9781611970104.
66. F. D Lee, Conceptual Wavelets in Digital Signal Processing. Space & Signals Technologies 200 2009, 1-30.
67. S. Mallat, A wavelet tour of signal processing. Academic press 1999.
68. M. Vetterli, C. Herley, Wavelets and filter banks: Theory and design. IEEE Trans. Signal Process. 40 (9), (1992) 2207-2232, doi: 10.1109/78.157221
69. S. G. Mallat, A theory for multiresolution signal decomposition: the wavelet representation. Pattern Anal. Mach. Intell. 11 (7) (1989) 674-693, **doi**:10.1109/34.192463.
70. Y. Linde, A. Buzo, R. Gray, An algorithm for vector quantizer design. IEEE T Commun, 28(1) (1980) 84-95, **doi**: 10.1109/TCOM.1980.1094577.
71. V. Vapnik, The Nature of Statistical Learning Theory, Springer-Verlag, NewYork, 1995, **doi**: 10.1007/978-1-4757-2440-0.
72. T. M. Mitchell, Machine Learning (McGraw-Hill International Editions Computer Science Series), 1997 doi: 10.1002/(SICI)1099-1689(199909)9:3<191::AID-STVR184>3.0.CO;2-E .
73. D.W Aha, D. Kibler, M. K. Albert., Instance-based learning algorithms, Mach learn. 6(1), (1991) 37-66.
74. R. S. Michalski, J. G. Carbonell, T. M. Mitchell, Machine learning: An artificial intelligence approach. Springer Science & Business Media, 2013 **doi**: 10.1007/978-3-662-12405-5.
75. S. Haykin, Neural networks and learning machines (Vol. 3). Upper Saddle River, NJ, USA Pearson, 2009.
76. M. T. Hagan, H. B. Demuth, M. H. Beale, O. De Jesús, Neural network design (Vol. 20). Boston: Pub, 1996.
77. Y. Kagaya, M. Tabata, Y. Arata, J. Kameoka, S. Ishii, Variation in effectiveness of a cardiac auscultation training class with a cardiology patient simulator among heart sounds and murmurs. Am J Cardiol. 70(2) (2017) 192-198.
78. A. Ross, A. Jain, Information fusion in biometrics. Pattern Recognit. Let. 24(13) (2003) 2115-2125, **doi**: 10.1016/S0167-8655(03)00079-5.

79. O. Al-Hamdani, A. Chekima, J. Dargham, S. H. Salleh, F. Noman, H. Hussain, A. M. Noor, Multimodal biometrics based on identification and verification system. J Biomet Biostat. 4(2) (2013), 1-8.
80. F Meziani, S. M. Debbal, A. Atbi, Analysis of phonocardiogram signals using wavelet transform, J. Med. Eng. Technol. 36(6) (2012) 283–302, **doi**: 10.3109/03091902.2012.684830.
81. L.R. Rabiner, R.W. Schafer, Introduction to digital speech processing. (Now Publishers Inc., 70–73, 2007, **doi:** http://dx.doi.org/10.1561/2000000001.
82. S. M. Debbal, F. Bereksi-Reguig, Computerized heart sounds analysis. Comput Biol Med, 38(2) (2008) 263-280, **doi**: 10.1016/j.compbiomed.2007.09.006.
83. M. Unser, A. Aldroubi, (1996). A review of wavelets in biomedical applications. Proceedings of the IEEE, 84(4), 626-638, **doi**: 10.1109/5.488704.
84. M. N. Ali, E. S. A. El-Dahshan, E. S. A., and A. H. Yahia (2017). De-noising of heart sound signals using discrete wavelet transform. *Circuits, Syst. Signal Process.* *36*(11), 4482-4497.
85. L. H. Cherif, S. M. Debbal, F. Bereksi-Reguig, Choice of the wavelet analyzing in the phonocardiogram signal analysis using the discrete and the packet wavelet transform. Expert Syst Appl, 37(2) (2010) 913-918. **doi**: 10.1016/j.eswa.2009.09.0.
86. M. M. Bassiouni, E. S. A. El-Dahshan, W. Khalefa, & A. M. Salem, (2018). Intelligent hybrid approaches for human ECG signals identification. *Signal Image Video Process.* 12(5), 941-949.
87. L. Khadra, M. Matalgah, B. El-Asir, S. Mawagdeh, The wavelet transform and its applications to phonocardiogram signal analysis. Inform. 16(3) (1991) 271-277.
88. M. S. Obaidat, Phonocardiogram signal analysis: techniques and performance comparison. J. Med. Eng. Technol. 17(6) (1993) 221-227.
89. L. G. Durand, P. Pibarot, Digital signal processing of the phonocardiogram: review of the most recent advancements. Crit. Rev. Biomed. Eng., 23 (1995) 3-4.
90. S. Ismail, I. Siddiqi, U. Akram, Localization and classification of heart beats in phonocardiography signals—a comprehensive review. Eurasip J Adv Sig Pr. 1 (2018) 26. https://**doi**.org/10.1186/s13634-018-0545-9 .
91. M. Nabih-Ali, E. S. A. El-Dahshan, A. S. Yahia, A review of intelligent systems for heart sound signal analysis. J. Med. Eng. Technol. 41(7) (2017) 553-563, **doi**: 10.1080/03091902.2017.1382584.
92. A. K. Dwivedi, S. A. Imtiaz, E. Rodriguez-Villegas. Algorithms for Automatic Analysis and Classification of Heart Sounds–A Systematic Review. IEEE Access. 7, (2019) 8316-8345.
93. J. R. Pinto, J. S. Cardoso, A. Lourenço, Evolution, current challenges, and future possibilities in ECG biometrics. IEEE Access, 6, (2018) 34746-34776.
94. M. M. Al Rahhal, Y. Bazi, H. AlHichri, N. Alajlan, F. Melgani, R. R. Yager, Deep learning approach for active classification of electrocardiogram signals. Inf. Sci. 345 (2016) 340-354.
95. T. Chakrabarti, S. Saha, S. Roy, I. Chel, (2015, October). Phonocardiogram signal analysis-practices, trends and challenges: A critical review. In 2015 International conference and workshop on computing and communication (IEMCON) (pp. 1-4). IEEE.
96. S. B. B. Kumar, M. Jagannath, Analysis of Phonocardiogram signal for biometric identification system. In 2015 International Conference on Pervasive Computing (ICPC), pp. 1-4. IEEE, 2015.
97. S. K. Randhawa, Singh, M., Survey of Different Methodologies Used in Phonocardiogram Signal Analysis. Int. J. Comput. Appl., 117(9) (2017) 18-21.
98. L. A. H. Clevenger, Classification of Heart Sound Biometrics for Active User Authentication and Clinical Cardiac Applications 2017 (Doctoral dissertation, Pace University).

99. S. Gupta, L. Singh, (2017, October). A study on new biometric approaches. In 2017 International Conference on Computing and Communication Technologies for Smart Nation (IC3TSN) (pp. 306-310). IEEE.
100. Bhanu, B. and Govindaraju, V. Multibiometrics for Human Identification. Cambridge University Press (2011).
101. Da Silva, H. P., Fred, A., Lourenço, A., Jain, A. K. (2013, September). Finger ECG signal for user authentication: Usability and performance. In *2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS)* (pp. 1-8). IEEE.
102. V. Arora, R. Leekha, R., Singh, I. Chana, Heart sound classification using machine learning and phonocardiogram. *Modern Physics Letters B*, 33(26) (2019), 1950321.
103. Gjoreski, M., Gradišek, A., Budna, B., Gams, M., & Poglajen, G. (2020). Machine learning and end-to-end deep learning for the detection of chronic heart failure from heart sounds. *IEEE Access*, 8, 20313-20324.
104. Z. Abduh, E. A. Nehary, M. A. Wahed, Y. M. Kadah, Classification of heart sounds using fractional fourier transform based mel-frequency spectral coefficients and traditional classifiers. Biomed Signal Process, 57 (2020), 101788.
105. S. K. Ghosh, R. N. Ponnalagu, R. K. Tripathy & U. R. Acharya, (2020). Automated detection of heart valve diseases using chirplet transform and multiclass composite classifier with PCG signals. *Comput. Biol. Med.* 118, 103632.
106. .Humayun, A. I., Ghaffarzadegan, S., Ansari, M. I., Feng, Z., & Hasan, T. (2020). Towards domain invariant heart sound abnormality detection using learnable filterbanks. *IEEE J BIOMED HEALTH,* 24(8), 2189-2198.
107. M. S. Manikandan, and K. P. Soman, Robust heart sound activity detection in noisy environments. Electron. Lett, *46*(16) (2010), 1100-1102.
108. K. A. Babu, B. Ramkumar, and M. S. Manikandan, Automatic identification of S1 and S2 heart sounds using simultaneous PCG and PPG recordings. IEEE Sens. J., *18*(22) (2018) 9430-9440.
109. S. Shukla, S. K. Singh, and D. Mitra, An efficient heart sound segmentation approach using kurtosis and zero frequency filter features. Biomed Signal Process *57* (2020), 101762.
110. S. Li, F. Li, S. Tang, W. Xiong, A review of computer-aided heart sound detection techniques, *Biomed Res. J,* (2020) http://dx.doi.org/10.1155/2020/5846191, Article ID 5846191.
111. L. Shu, J. Xie, M. Yang, Z. Li, D. Liao, X. Yang, A review of emotion recognition using physiological signals. *Sensors, 18*(7) (2018) 2074.

Abbreviations

ANN	Artificial Neural Network
BFCC	Bark Frequency Cepstral Coefficient
CCA	Conical Correlation Analysis
CCR	Correct Recognition rate
CZT	Z-Chirp transforms
DCT	Discrete Cosine Transform
DWT	Discrete Wavelet Transform
DW-MFCC	Discrete Wavelet-Mel Frequency Cepstral coefficients
ED	Euclidian Distance
EEMD	Ensemble Empirical Mode Decomposition
EER	Equal Error Rate

FFT	Fast Fourier Transform
FS	Fourier Spectrum
FSR	First-to-Second Ratio
GMM	Gaussian Mixture Models
HMM	Hidden Markov Model
HS	Hilbert Spectrum
HS-LBFC	Heart Sound-Linear Band Frequency Cepstra
HT	Hung-Transform
ICC	Improved Circle Convolution
IDWT	Inverse Discrete Wavelet Transform
IMF	Intrinsic Mode Functions
ISF	Independent Sub-band Function
KNN	K-Nearest Neighbor
KSRC	Kernel Sparse Representation
LBQ	Linde Buzo & Gray algorithm
LDA	Linear Discernment Analysis
LFBC	Linear Frequency banks cepstral
LFCC	Linear Frequency Cepstral Coefficient
LPCC	Linear Predication Cepstral Coefficient
LR	Linear Regression
LVQ	Linear Vector Quantization
MFCC	Mel Frequency Cepstral Coefficient
MLP	Multi-layer Perception
M-MFCC	Modified Mel Frequency Cepstral Coefficients
N-LFCC	Nonlinear Frequency Cepstral Coefficients
PCA	Principle Component Analysis
PCG	Phonocardiogram
SEE	Shannon Energy Envelogram
SFM	Spectral flatness Measurement
SRC	Sparse Representation Classifier
STA	Short Term Amplitude
STFT	Short Time Fourier Transform
SVM	Support Vector Machine
TVD	Total Variation De-noising
VQ	Vector Quantization
WNN	Wavelet Neural Network
WPCC	Wavelet Packet Cepstral Coefficients
WT	Wavelet Transform
ZCR	Zero Crossing Rate