Systematic Review for Phonocardiography Classification Based on Machine Learning

Abdullah Altaf[1], Hairulnizam Mahdin[2], Awais Mahmood[3], Mohd Izuan Hafez Ninggal[4], Abdulrehman Altaf[5], Irfan Javid[6]

Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia, Parit Raja, Batu Pahat, Johor, Malaysia[1, 2, 5]

Computer Engineering Dept, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia[3]

Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, Selangor, Malaysia[4]

Department of Computer Science & IT, University of Poonch Rawalakot, AJK Pakistan[6]

***Abstract*—Phonocardiography, the recording and analysis of heart sounds, has become an essential tool in diagnosing cardiovascular diseases (CVDs). In recent years, machine learning and deep learning techniques have dramatically improved the automation of phonocardiogram classification, making it possible to delve deeper into intricate patterns that were previously difficult to discern. Deep learning, in particular, leverages layered neural networks to process data in complex ways, mimicking how the human brain works. This has contributed to more accurate and efficient diagnoses. This systematic review aims to examine the existing literature on phonocardiography classification based on machine learning, focusing on algorithms, datasets, feature extraction methods, and classification models utilized. The materials and methods used in the study involve a comprehensive search of relevant literature and a critical evaluation of the selected studies. The review also discusses the challenges encountered in this field, especially when incorporating deep learning techniques, and suggests future research directions. Key findings indicate the potential of machine and deep learning in enhancing the accuracy of phonocardiography classification, thereby improving cardiovascular disease diagnosis and patient care. The study concludes by summarizing the overall implications and recommendations for further advancements in this area.**

Keywords—Heart sounds classification; Phonocardiogram (PCG); CVDs; deep learning

I. Introduction

Phonocardiography (PCG) is one of the basic techniques used to understand the heart's state and assess whether the heart is in a natural state or has some abnormal pattern. PCG is a diagnostic procedure that allows a visual record of the sounds and murmur created by the contracting heart, including its valves and connected large vessels. In the absence of diagnosis equipment, the stethoscope is only the tool available to general physicians to examine a patient's heart sounds [1]. Cardio-specialist can understand the heartbeat as a specialist and recommend further medical procedures to the patients according to their heart condition but usually, in the unavailability of a cardio-specialist, the general physicians cannot detect, if the heart is functioning properly or if there is any type of exception due to the closure of heart walls. Environmental interferences, such as those caused by friction between the device and a human's skin, Electromagnetic Interference (EI), and unrelated noises like breath, lung, and ambient sounds, can readily interfere with the process of PCG because signals in the form of sound generated by the human heart are frequently paired with EI, out-of-band noise must be removed [2], [3].

To cope with the limitations in traditional phonocardiography techniques, machine learning (ML) based methods can be a good solution for phonocardiography for several reasons [4]. ML models can automate the process of analyzing PCG recordings, which can be time-consuming and subject to human error when done manually. ML models can be trained on large datasets of labeled PCG recordings, which can improve their accuracy in detecting and diagnosing heart conditions. ML-based methods can handle large amounts of data, which is important in PCG as it requires analyzing audio signals over time and adapting to new data, and improving their performance over time, which can be useful in handling diverse populations and detecting new conditions [5].

Using ML-based approaches to achieve real-time heart disease detection from audio signals is challenging because heart sounds can vary significantly depending on factors such as the person's age, sex, and underlying medical conditions [6]. This makes it difficult to develop an ML model that can accurately detect and classify heart sounds in a wide range of individuals. Heart sounds can be difficult to distinguish from other sounds in the body, such as breathing and blood flow [7]. Additionally, external noise such as background noise or equipment noise can also interfere with the recording and analysis of heart sounds [5]. Collecting a large and diverse dataset of heart sounds for training machine learning models can be difficult and time-consuming.

Heart sounds are complex and have a lot of variations, which can make it difficult for machine learning algorithms to accurately classify them. Overfitting is a common problem when building machine learning models, and can occur when a model is trained on a limited dataset and then performs poorly on new, unseen data [8]. The interpretation of phonocardiography signals requires expertise and knowledge of human anatomy and physiology, this is a challenge when using machine learning algorithms to interpret the signals. Designing a phonocardiography system using machine learning is a big challenge because it requires overcoming several technical and logistical hurdles in accurately and reliably detecting a classifying heart sound [9]. The core phases involved in the phonocardiography process i.e., segmentation, feature extraction, and final classification results, are all considerably impacted by denoising. Preprocessing is required in phonocardiography systems to clean and enhance the quality of the raw phonocardiography signal before it is further analyzed.

It is necessary to create these components, incorporating the applicable criteria that follow.

II. An Overview of Phonocardiography

Machine Learning based phonocardiography systems models used in real-time to analyze Phonocardiography (PCG) recordings and provide diagnostic information, which can be useful in critical care settings. Overall, the combination of machine learning and PCG data can provide more accurate and objective results and helps to improve the diagnosis and monitoring of heart conditions [10]. The generic ML-based phonocardiography systems are depicted in Fig. 1.

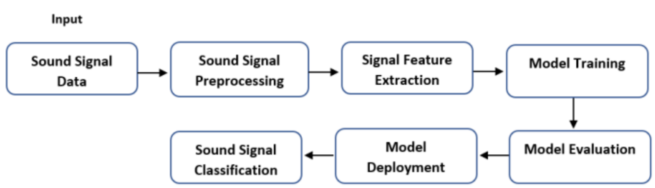


Fig. 1. ML-based phonocardiography general framework.

A. *Heart Sound Signal*

Heart sound signals, also known as phonocardiograms or PCG signals, are acoustic signals generated by the mechanical activities of the heart. These signals provide important information about the structure, function, and abnormalities of the cardiovascular system. Heart sound signals typically consist of two prominent components: the first heart sound (S1) and the second heart sound (S2). S1 is produced by the closure of the mitral and tricuspid valves during the systolic phase of the cardiac cycle, while S2 is generated by the closure of the aortic and pulmonary valves during the diastolic phase. These components are accompanied by additional sounds such as the third heart sound (S3) and fourth heart sound (S4), which can indicate specific cardiac conditions.

Heart sound signals are characterized by their duration, intensity, frequency content, and temporal relationships between different components. They contain valuable information about heart rate, heart rhythm, valve function, and the presence of murmurs, stenosis, or other cardiovascular abnormalities. Traditionally, heart sound signals were recorded using specialized electronic stethoscopes or phonocardiography equipment. However, with advancements in technology, heart sound signals can now be captured using digital stethoscopes, wearable devices, or even smartphone applications equipped with appropriate sensors.

Heart sounds (signals) are produced from a specific cardiac event such as the closure of a valve or tensing of a chordae tendineae. Most normal heart sound (Lub, Dub) signal rates at rest are between about 60 and 100 beats per minute. Sound is the pressure of air propagating to our ears. The digital audio file is gotten from a sound sensor that can detect sound waves and convert them to electrical signals.

B. *Normal Heart Sound Signal*

Normal heart sounds, also known as physiological heart sounds, are the characteristic sounds produced by a healthy heart during its regular functioning. These sounds are a result of the synchronized mechanical activities of the heart's valves and chambers. The normal heart sound consists of two primary components: the first heart sound (S1) and the second heart sound (S2). S1 is a low-frequency sound that occurs at the beginning of each cardiac cycle and is caused by the closure of the mitral and tricuspid valves. S2 is a higher-pitched sound that occurs at the end of the cardiac cycle and is produced by the closure of the aortic and pulmonary valves. These two sounds create the familiar "lub-dub" rhythm associated with a normal heartbeat. The normal heart sound signifies the proper functioning of the heart's valves and chambers, reflecting a healthy cardiovascular system. Understanding the characteristics and timing of normal heart sounds is crucial in differentiating them from abnormal sounds and diagnosing various cardiac conditions [11].

C. *Artifact Heart Sound Signal*

Artifact heart sounds refer to extraneous or spurious sounds that may be inadvertently recorded or introduced during the process of capturing heart sound signals. These sounds are not of physiological origin and do not reflect the actual functioning of the heart. Artifact heart sounds can arise from various sources, such as environmental noise, patient movement, electrical interference, or improper placement of the recording device. They can manifest as random noise, clicking sounds, buzzing, or other irregular patterns that may obscure or distort the true heart sounds. Artifact heart sounds pose a challenge in the accurate analysis and interpretation of heart sound signals, as they can interfere with the detection of abnormal cardiac conditions or mask important diagnostic information. Efforts are made to minimize artifacts during data collection by ensuring proper recording techniques, reducing environmental noise and employing noise cancellation methods. Additionally, careful signal processing and expert interpretation are essential to distinguish artifact heart sounds from genuine physiological sounds and ensure the reliability and accuracy of heart sound analysis in clinical practice and research [12].

D. *Extrastole Heart Sound Signal*

Extrastole, also known as an extra heart sound or premature beat, refers to an abnormal additional sound that occurs in the cardiac cycle, occurring either before or after the normal heart sounds. It is typically characterized by a distinctive "gallop" or "clicking" sound. Extrastole is caused by premature contractions of the heart's ventricles, atria, or both. These premature contractions disrupt the normal rhythm and timing of the cardiac cycle. Common types of Extrastole include atrial premature complexes (APCs) and ventricular premature complexes (VPCs). Extrastole can be indicative of underlying heart conditions such as arrhythmias, valvular disorders, or heart muscle abnormalities. Detecting and analyzing Extrastole in heart sound signals is crucial for diagnosing and monitoring cardiac abnormalities. Advanced signal processing techniques and machine learning algorithms are employed to identify and classify Extrastole patterns accurately. Understanding the presence and characteristics of Extrastole heart sounds aids in the comprehensive evaluation and management of cardiovascular health [13].

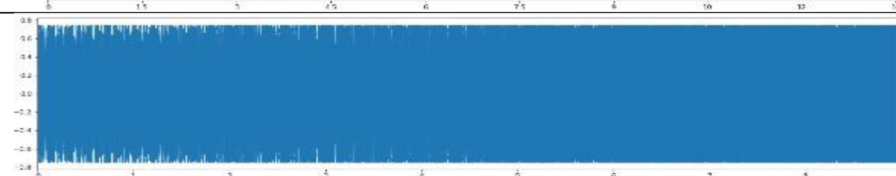
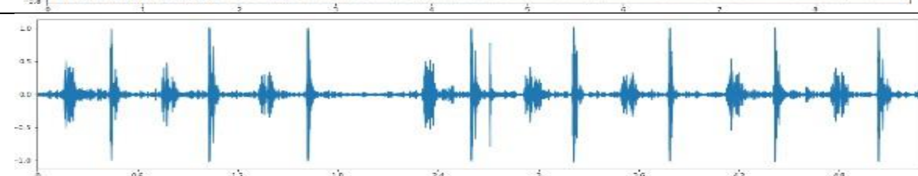
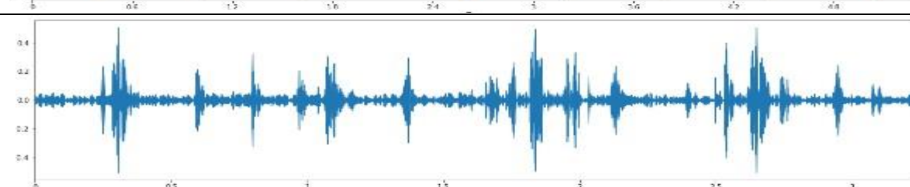
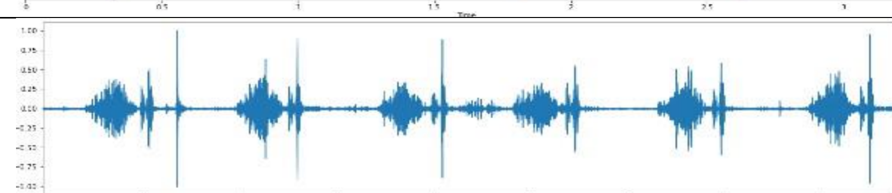
E. Extrahls Heart Sound Signal

Extrahls, also known as extraneous heart sounds or adventitious heart sounds, refer to abnormal sounds that are superimposed on normal heart sound signals. These sounds are not typically associated with the regular functioning of the heart and can arise from various pathological conditions or abnormalities within the cardiovascular system. Extrahls can manifest as additional clicks, murmurs, or abnormal sounds that occur in addition to the first and second heart sounds. They can be indicative of structural heart defects, valve abnormalities, turbulent blood flow, or other cardiac disorders. Analyzing Extrahls in heart sound signals is crucial for diagnosing and monitoring cardiovascular conditions, as they can provide valuable insights into the presence and severity of cardiac abnormalities. Advanced signal processing techniques, such as spectrogram analysis, wavelet analysis, or machine learning algorithms, are employed to identify and characterize Extrahls accurately. By detecting and analyzing Extrahls in heart sound signals, clinicians and researchers can improve their understanding of cardiac pathologies and make informed decisions regarding patient care and treatment strategies [14].

F. Murmur Heart Sound Signal

Murmur heart sound signals refer to abnormal or atypical sounds that are heard during auscultation of the heart. Murmurs are characterized by a prolonged, swishing, or whooshing sound that occurs between the normal heart sounds (S1 and S2). These sounds are caused by turbulent blood flow within the heart or blood vessels, typically due to structural abnormalities such as valve defects, stenosis, regurgitation, or abnormal blood flow patterns. Murmurs can be classified based on their timing, intensity, pitch, and location within the cardiac cycle. They are often graded on a scale from 1 to 6, with higher grades indicating more pronounced murmurs. Accurate identification and characterization of murmurs in heart sound signals are crucial for diagnosing and managing various cardiovascular conditions. Advanced signal processing techniques, such as spectral analysis, time-frequency analysis, or machine learning algorithms, can aid in the automated detection and classification of murmur patterns. By analyzing murmurs in heart sound signals, healthcare professionals can assess the severity of underlying cardiac abnormalities, determine appropriate treatment strategies, and monitor the effectiveness of interventions for improved patient care [12]. Table I depicts a human's heart-generated sound Wave-form signal.

TABLE I. HEART SOUND WAVE-FORM SIGNAL, X-AXIS REPRESENTS TIME AND Y-AXIS MEASURES AMPLITUDE (GENERATED BY USING PYTHON LIBRARY: MATPLOTLIB)

S.No	Heart State	Human's Heart Generated Sound Wave-form Signal
1.	Normal	
2.	Artifact	
3.	Extrastole	
4.	Extrahls	
5.	Murmur	

G. Signal Denoising

Signal denoising refers to the process of removing unwanted noise or interference from a signal while preserving the underlying information of interest. It is a fundamental technique in signal processing used to improve the quality and reliability of signals in various domains, such as audio, image, and biomedical signals. The presence of noise in a signal can distort or mask important features, making it challenging to extract meaningful information or make accurate measurements. Signal denoising methods aim to reduce or eliminate this noise, enhancing the signal's clarity and fidelity. These methods employ various techniques, such as filtering, statistical analysis, wavelet transforms, or machine learning algorithms, to suppress or attenuate the noise components while preserving the desired signal components. Signal denoising is widely used in applications where the accuracy and reliability of signal analysis, interpretation, or decision-making are critical, allowing researchers, engineers, and practitioners to obtain cleaner and more accurate signals for further analysis or processing.

Signal denoising is highly important in various applications due to its ability to enhance signal quality, improve data analysis accuracy, facilitate signal interpretation, increase measurement precision, optimize signal processing techniques, enable better signal visualization, and enhance communication and signal transmission reliability. By removing unwanted noise and interference from signals, signal denoising improves the accuracy, clarity, and reliability of the underlying information, leading to more meaningful analysis, interpretation, and decision-making. It is a crucial step in fields such as biomedical signal processing, image and audio processing, communication systems, and scientific research, where accurate and reliable signal analysis is essential for successful outcomes.

Environmental interferences, such as those caused by friction between the device and a human's skin, Electromagnetic Interference (EI), and unrelated noises like breath, lung, and ambient sounds, can readily interfere with the process of recording heart sounds [15]. Because signals in the form of sound generated by the human heart are frequently paired with EI, out-of-band noise must be removed. The segmentation, feature extraction, and final classification results are all considerably impacted by denoising. Wavelet denoising, variational mode deconstruction denoising, and Digital Filter Denoising (DFD) are the three most often used denoising techniques [16]. A new line of study in the field of heart sound feature extraction is the creation of a wavelet function for the human heart's signals based on the past understanding of heart sound data [17].

A sound spectrum refers to the distribution of the different frequencies present in a sound signal. It represents the energy or intensity of each frequency component within the signal. The spectrum provides valuable information about the composition and characteristics of the sound, allowing for the identification and analysis of specific frequency components. In the context of heart sound signals, a sound spectrum can reveal the presence and intensity of different sound frequencies associated with normal or abnormal heart sounds. By analyzing the spectrum, healthcare professionals and researchers can gain insights into the underlying physiological conditions and abnormalities of the heart.

Spectral analysis techniques, such as Fourier transform or wavelet transform, are commonly employed to compute the sound spectrum and visualize the frequency content of the signal. This information can assist in the diagnosis, monitoring, and treatment of various cardiovascular disorders, providing valuable insights into the acoustic properties of the heart. Table II depicts different Heart Sound spectrums after denoising of heart's sound wave signal.

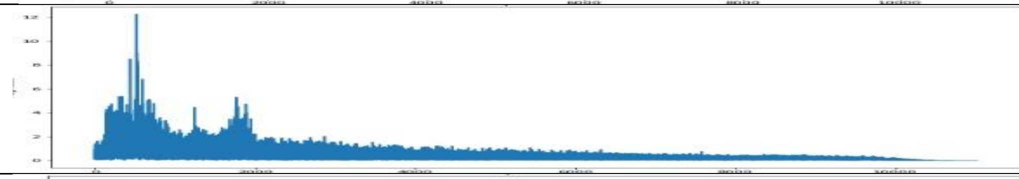
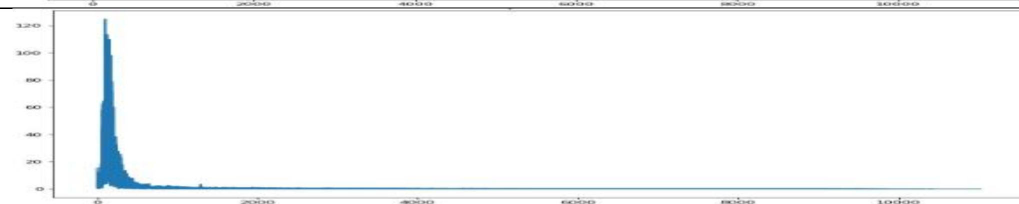
H. Signal Segmentation

Signal segmentation is a fundamental process in signal processing that involves dividing a continuous signal into shorter segments or intervals. This technique is essential for isolating specific regions of interest within a signal, allowing for targeted analysis and processing. Signal segmentation is commonly used in various fields, such as speech recognition, audio processing, image analysis, and biomedical signal analysis. By segmenting a signal, researchers can focus on specific time intervals or frequency components for further analysis, enabling the extraction of meaningful features and patterns. This approach facilitates tasks such as event detection, signal classification, anomaly detection, and time-frequency analysis. Additionally, signal segmentation helps in dealing with non-stationary signals by breaking them down into smaller, more manageable segments. Overall, signal segmentation plays a vital role in signal processing applications, allowing researchers and practitioners to effectively analyze and understand complex signals by examining their constituent segments individually.

As part of the segmentation process, the heart sounds of the first human (S1), the second human (S2), and the diastole are split into four parts or segments. Each section has useful components that help distinguish between the different types of heart sounds. However, individual differences in the length of the human heartbeat beat cycle, the number of human heart sounds, and the kinds of heart murmurs result in erroneous PCG signal segmentation. Thus, segmenting the FHS is a crucial step in the automated PCG analysis process.

In recent years, envelope-based techniques have been among the most frequently utilized techniques for segmenting heart sounds [18], [19]. Electrocardiogram (ECG) [20], feature-based methods [21], time-frequency analysis methods [22], and probabilistic model methods **[**23], [24], [25], [26] are some important segmentation methods. The underlying premise of the employed algorithms is that the diastolic interval is more prolonged than the systolic time. In actuality, especially in newborns and cardiac patients, this supposition is not always accurate for an aberrant heart sound **[27].** Based on the similarity between the ECG and the human heart signals, it has been discovered that algorithms that combine the cardiac cycle with an ECG signal perform better at segmenting data. They do have higher hardware and software requirements, though.

TABLE II. HEART SOUND SPECTRUM, X-AXIS REPRESENTS FREQUENCY AND Y-AXIS REPRESENTS MAGNITUDE (GENERATED BY USING PYTHON LIBRARY: MATPLOTLIB)

S.No	Heart State	Human's Heart Sound Signal Spectrum
1.	**Normal**	
2.	**Artifact**	
3.	**Extrastole**	
4.	**Extrahls**	
5.	**Murmur**	

Humans can hear sound not only at a particular time by its intensity but also by its pitch. The pitch is the frequency of the sound, a higher pitch corresponds to a higher frequency, and vice versa. So, to have a representation that is closer to the human brain, another dimension which is frequency is added to the representation, which is the spectrogram.

III. REVIEW

This review will provides a compendious review of Phonocardiography, Machine Learning (ML) literature including experimental, empirical, and theoretical studies. The modern advancements and contributions from recent studies related to major and compelling administration depict this research.

Previously proposed phonocardiography, Machine learning techniques such as Deep Learning (DL), Extreme Learning Machines (ELMs), Deep Extreme Learning Machines (DELMs), and previous technologies and techniques are comprehensively reviewed. The literature mentioned is either the most benchmark research contributions, most relevant, highly cited, or published in well-reputed research journals.

A. Phonocardiography

Phonocardiography is a non-invasive diagnostic technique used to capture and analyze the sounds produced by the heart during its normal functioning. It involves recording the sounds made by the heart using a sensitive microphone or electronic sensor and then analyzing the recordings to identify any abnormal sounds or patterns that may indicate the presence of heart disease. Fig. 2 shows the Phonocardiography Signals of Normal Heart. In the phonocardiograph, the process of determining the state of the heart is done using the waves that come from the heart, and the process of classification of the heartbeats is done as normal and abnormal [28].

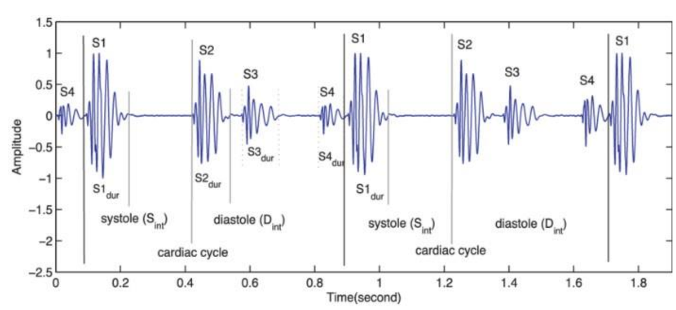


Fig. 2. Phonocardiography signals of normal heart.

The normal is defined as the state that an adult's blood pressure at rest varied from 60 to 100 beats per minute. In general, improved cardiac function and improved cardiovascular health are demonstrated by a lower resting heart rate. As for the abnormal state, the heart is in a state of disorder, unstable, and known any erratic heartbeat is known to be an abnormal heart rate or an increased heart rate. An irregular heart rhythm (rapid heartbeat, called tachyarrhythmia, or slow pulse, called slow arrhythmia-bradyarrhythmia) can accompany arrhythmia [29]. Fig. 3 shows the Phonocardiography Wave Signals of Arrhythmia-Bradyarrhythmia.

One of the most prominent application features of phonocardiography is the detection and monitoring of the evaluate the effectiveness of treatment of the key heart abnormalities symptoms like ventricular dysfunction, aortic regurgitation, mitral regurgitation, aortic stenosis, heart failure, hypertrophic cardiomyopathy, tricuspid regurgitation, coronary artery disease, pulmonic stenosis, atrial fibrillation, ventricular tachycardia, heart murmurs, aortic aneurysms, systolic heart murmurs, ventricular hypertrophy, ventricular septal defects and ductus arteriosus [30].

Left ventricular dysfunction refers to a condition where the heart's left ventricle, which is responsible for pumping oxygenated blood to the rest of the body, is not functioning properly. This can occur due to a variety of reasons, including damage to the heart muscle, high blood pressure, heart valve problems, or coronary artery disease. Left ventricular dysfunction can lead to a range of symptoms, including shortness of breath, fatigue, swelling in the legs and feet, and chest pain. It can also increase the risk of heart failure, heart attack, and other cardiovascular complications. Phonocardiography-based detection of left ventricular dysfunction was reviewed in detail by Jagannath [31].

Some notable contributions along with the feature extraction techniques, datasets, modeling techniques, and results are depicted in Table III.

From Table III, it can be noted that most of the research work has been done using different feature extraction techniques including MFCC. Also, different researchers used different machine learning techniques like KNN, ANN, and SVM.

All the results have differences due to the use of different datasets, feature extraction techniques, and machine learning techniques. Most of the datasets used were private and the researchers did not share their data set on the web, or in other words, we can say that those data sets are not publicly available. We have found that Ziaee Hospital recorded Ardakan dataset. The dataset contains 148 PCG signals from 22 subjects (8 males, 14 females; between 3 and 85 years old), unfortunately, the data is not publicly available from the literature.

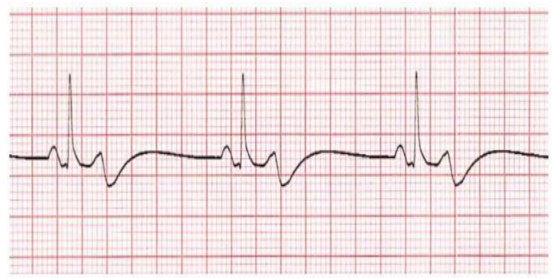


Fig. 3. Phonocardiography wave signals of arrhythmia-bradyarrhythmia.

TABLE III. DETAILS OF FEATURE EXTRACTION AND ACCURACY

Article	Feature Extraction Technique	Machine Learning Techniques	Dataset	Results Accuracy in %
[32]	MFCC	CNN	PASCAL	87.65
[33]	SEE, HT, MFCC, WT	SVM, KNN	hospital in Ardakan	98.78
[34]	WPT, SVD	CNN, RNN, LSTM	PhysioNet	79.8
[35]	MFCC, Time & Freq	HMM, SVM, CNN	PhysioNet	89.22
[36]	MFCC	DFT, CNN	PhysioNet	86.02
[37]	MFCC	CNN, Springer's algorithm	PASCAL	90.4
[38]	DL	RNN, LSTM , GRU	PASCAL	76.9
[15]	DWT, CWT, STFT, MFCC	SVM, ANN, KNN	PhysioNet	99.7
[39]	MFCC	SVM, RF, MLP	PhysioNet	84.88
[40]	FT, WT, FLP	LVQ, PNN, LS-SVM	PhysioNet	98
[41]	MN, SD, RPAB, SC	SVM	Not available	71.13
[70]	LDA	K-mean	Peter Bentley heart sound	84.39

Another dataset that some researchers have used is GitHub. GitHub dataset they have obtained the PCG signals from a public database but the best result was from GitHub. There are 1,000 PCG recordings present in the database for different subjects. Out of these 1,000 recordings each class contains 200 recordings. There are five classes of PCG signals given in the database namely the healthy control (HC), AS, MS, MR, and mitral valve collapse (MVP), Most of this dataset it's not clear and even the data set is not labeled.

Considered the most popular among the research and the most used in the experiment, the dataset (PASCAL challenge dataset) consists of 312 auscultations collected in the Real Hospital Portugués (RHP) Maternal and Fetal Cardiology Unit in Recife, Brazil, using the DigiScope. Each inspection is reported for six to ten seconds. The normal count is 200. And the abnormal count of 112. In reference [37] the collected dataset is divided into two parts: Dataset A contains 31 normal heart sounds and 34 abnormal heart sounds. For the training set, 15 normal and 17 abnormal sounds are selected, and the rest is for the test. Dataset B contains 200 normal and 66 murmurs, 100 normal and 33 abnormal for the training, and the rest is left for the test.

After we collected a large number of databases and searched for them, we found that what is publicly available is PASCAL and PhiysoNet. Thus our research based on them in addition to the availability of a new database called (heartbeat sound) through the site "Kaggle" which is the largest site for contests and review of databases.

B. Mel-Frequency Cepstral Coefficients

Mel Frequency Cepstral Coefficients (MFCC) is a feature extraction technique used in speech processing and audio signal analysis. The MFCC algorithm was first proposed by Davis and Mermelstein in 1980 [42]. It has since become one of the most widely used techniques in speech and audio signal processing. The basic intuition behind MFCC is to extract a compact representation of the spectral envelope of an audio signal, which can then be used for further analysis. The spectral envelope is essentially a smoothed version of the power spectrum of the signal, and it contains information about the distribution of energy across different frequency bands [43]. Fig. 4 shows the signal in the Time Domain.

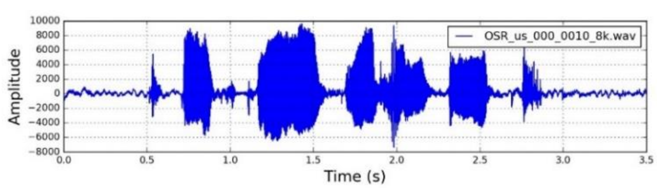


Fig. 4. Signal in the time domain.

To extract the spectral envelope, the MFCC algorithm first divides the signal into short overlapping frames, typically 20-40 MS in duration. Each frame is then transformed into the frequency domain using a Fourier Transform. The resulting frequency spectrum is then passed through a bank of Mel-scale filter banks, which are spaced uniformly on the Mel-scale, a perceptual scale that is more closely related to the way humans perceive pitch than the linear frequency scale [44]. The output of each filter bank is then logarithmically scaled, and the resulting values are transformed using the Discrete Cosine Transform (DCT) to obtain the MFCCs.

The DCT is used to decorrelate the filter bank outputs and to obtain a set of décor-related coefficients that are more suitable for further analysis [45]. The number of MFCCs extracted depends on the specific application, but typically, 12-13 coefficients are used for speech recognition tasks. The resulting MFCCs can be used as features for machine learning algorithms such as Hidden Markov Models (HMMs), which are commonly used in speech recognition [2].

A Comparative Study on MFCC and MEL spectrogram features for automatic speech recognition presented by J.M. Azevedo, et al. [46]. This study compared the performance of MFCC and Mel spectrogram features for automatic speech recognition using deep learning models. The results showed that Mel spectrogram features outperformed MFCCs in terms of recognition accuracy [46].

Exploring the impact of MFCC and its derivatives on EEG-based emotion recognition was proposed by S. Almogbel, et al. [47]. In this study, the authors investigated the use of MFCC and its derivatives for EEG-based emotion recognition. The results showed that incorporating MFCC and its derivatives improved the accuracy of emotion recognition compared to using raw EEG signals alone [47].

MFCC-based speech enhancement using Deep Neural Networks (DNN) was investigated by Mohamed, et al. [48]. This study proposed a deep neural network-based approach for speech enhancement using MFCC features. The results showed that the proposed approach improved speech quality and intelligibility compared to traditional MFCC-based speech enhancement methods [48]. MFCC and its derivatives-based features extraction for automated speech recognition of spontaneous Tamil language was put forward by S. Sivaprakasam, et al. [49]. This study investigated the use of MFCC and its derivatives for automated speech recognition of the spontaneous Tamil language. The results showed that using higher-order derivatives of MFCCs improved the recognition accuracy compared to using only MFCCs [49].

Comparison of MFCC and Gamma tone filter bank features for speech emotion recognition presented by S. Mohapatra and S. Lenka [50]. This study compared the performance of MFCC and gamma tone filter bank features for speech emotion recognition using machine learning models. The results showed that both features performed similarly in terms of recognition accuracy [50]. An Improved MFCC algorithm for speech recognition lodge by J. Wu, et al. [51]. This study proposed an improved MFCC algorithm for speech recognition by adding a Gaussian filter to the Mel filter bank. The results showed that the proposed algorithm outperformed traditional MFCCs in terms of recognition accuracy [51].

A Comparative Study of MFCC and Gammatone Filter Bank features for phoneme recognition advanced by Y. Sun, et al. [52]. This study compared the performance of MFCC and Gamma tone filter bank features for phoneme recognition using deep learning models. The results showed that Gammatone filter bank features outperformed MFCCs in terms of recognition accuracy [52].

Speaker recognition based on MFCC and XGBoost initiated by Z. Yang, et al. [53]. This study proposed a speaker recognition system based on MFCC features and XGBoost, a gradient-boosting algorithm. The results showed that the proposed system achieved high accuracy in speaker recognition tasks [53]. A Comparative Study of MFCC and DL features for speech emotion recognition argued by X. Liu, et al. [54]. This study compared the performance of MFCC and deep learning features for speech emotion recognition. The results showed that deep learning features outperformed MFCCs in terms of recognition accuracy [54].

MFCC-based speech separation using Convolutional Recurrent Neural Networks tender by M. Hossain, et al. [55]. This study proposed a convolutional recurrent neural network-based approach for speech separation using MFCC features. The results showed that the proposed approach outperformed traditional MFCC-based speech separation methods [55].

C. Machine Learning and Medical Diagnosis

Machine Learning (ML) continues to hold immense significance in medical diagnosis, especially in this digital and AI age. One of its key contributions lies in the increased accuracy it offers. ML algorithms possess the capability to analyze vast amounts of medical data, encompassing patient records, medical images, and genomic information. By identifying complex patterns and relationships within this data, ML models can provide more accurate and reliable diagnoses, often surpassing human capabilities. ML algorithms excel at the early detection and prevention of diseases by analyzing patient data over time, these models can identify subtle indicators and early signs of conditions such as cancer, heart disease, or neurological disorders. Early detection enables healthcare professionals to intervene sooner, improving patient outcomes and potentially saving lives. ML also enables personalized medicine, as it allows for the development of tailored treatment plans based on an individual's unique characteristics, such as genetics, lifestyle, and medical history. By considering these factors, ML algorithms can predict the effectiveness of different treatment options, helping physicians make informed decisions that are specifically catered to each patient's needs [56].

In the field of medical imaging, and medicine phonography ML algorithms have made significant strides. They can automatically analyze and interpret medical images like X-rays, MRIs, and CT scans, detecting patterns and anomalies that may be challenging for human observers to identify from the image, and sound data. This capability enhances radiologists' efficiency and accuracy, leading to more precise diagnoses and reducing the chances of misinterpretation. ML's integration into electronic health records (EHRs) enables the analysis of large-scale patient data, allowing for data integration and decision support. ML algorithms can uncover hidden correlations between symptoms, risk factors, and treatment outcomes, providing healthcare professionals with valuable insights to make informed decisions. ML-powered decision support systems can suggest diagnoses, treatment plans, and medication recommendations, ultimately improving the overall quality of patient care [57].

Furthermore, ML contributes to increased efficiency and cost savings in healthcare. By automating time-consuming tasks like data entry, documentation, and administrative processes, ML frees healthcare providers to focus more on direct patient care, reducing the burden of paperwork and enhancing operational efficiency. Additionally, ML helps optimize resource allocation and treatment pathways, potentially reducing healthcare costs.

ML process passes through many phases like pre-processing, learning, and evaluation. Data is mostly in the form of unstructured, inconsistent, incomplete, redundant, heterogeneous, as well as noisy. Using techniques like data cleaning, extraction, fusion, transformation, etc. data pre-processing helps prepare raw data usable and consistent for the subsequent phases. The learning phase selects algorithms of learning and refines the parameters of the model to get the required results through the usage of the pre-processed input data. The evaluation phase follows to ensure the attained performance of the learning model, i.e. performance evaluation of the learning algorithm included the selection of dataset, error estimation, measuring performance, and the different tests of statistics. The results of the evaluation phase may lead to the parameters of adjusting for the selected learning algorithm or selecting various algorithms and classifiers. Through the multiple facets like nature of learning, learning target, type of input data, data availability timing, users (stakeholders), and domain factors. Fig. 5, illustrates the Machine Learning framework.

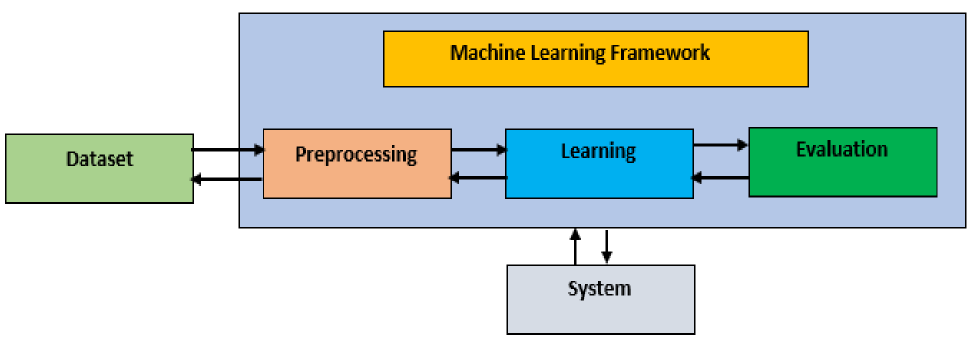


Fig. 5. Machine learning framework (Rizvi, 2019).

D. Supervised Learning

There are three major administrations of conventional machine learning which are supervised, unsupervised, and semi-supervised learning. Supervised learning (SL) is claimed as the uttermost decisive and intrusive annex of machine learning (ML) and pattern recognition. In supervised learning, the learning system is commenced with examples of input and output braces and the prime objective is to learn a function that decorously maps inputs and outputs [58].

Supervised machine learning which is also known as a classification learning advent is employed for analyzing training or labeled data to draft concealed and unseen instances of data for future and imminent classification. To train a classifier that learns to differentiate between different pattern classes, extracted features from recognition units are used [59]. The supervised learning approach carves an acceptable amount of training data or labeled data for the classification of unseen test data [1].

SL, which is also called sometimes a classification learning advent, is implemented for data labeling or training analysis for drafting unseen and concealed data instances for both the imminent and future classification. In training, a classifier for learning to differentiate between the distinct pattern's classes and extracted features from the units of recognition are utilized. The SL approach carves an acceptable amount of training data or labeled data for the classification of unseen test data [59]. The SL model is shown in Fig. 6.

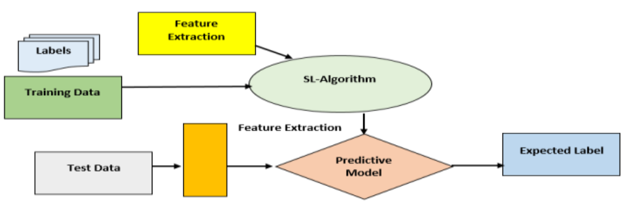


Fig. 6. Supervised machine learning model [59].

E. Artificial Neural Network

Being an SL-based information processing discipline, Artificial Neural Network (ANN) is used for solving circuitous problems. Besides, ANN facilitates understanding the conduct of complex systems through the usage of computer simulations. The final objective of the ANN paradigm is to solve computational problems in the way, the human brain would do. A very simple ANN comprises abounding elementary nodes/processors, often dubbed neurons. These neurons cause a catenation of real-valued activation. Apart from that, sensory input is evoked by the input neurons through the perception of the environment. In the same way, other neurons receive the input of activation from weighted connections of the past antecedent neurons. In this way, only a few neurons, output neurons, for instance, may affect the environment through various prompting actions [60].

The discipline of ANN can be thought of as a set of massively parallel computing systems comprising many interconnected processors. ANN models try to utilize organizational principles like generalization, learning, and computation in a network of weighted graphs. The nodes being used in this network are artificial neurons. Further, the edges with the weights are joined with the input and the output neurons. In ANN, the process of learning is the determination of weights that force the ANN to show the expected behavior or output. Comprehending the conduct of neurons may entail an array of computational stages and phases in a non-linear manner. Further, each stage converts the network's accumulated activation. Specifically, the techniques of ANN deputize the credit beyond many such stages. Besides, this paradigm is normally adopted due to its vibrant adequacy of the mapping for nonlinear systems [61].

F. Extreme Learning Machines

Both the DL and the classical ML techniques proved a springboard for yet another efficient approach to learning which was introduced during the last decade. This new approach to learning is termed Extreme Learning Machines (ELM). ELM based on the SL approach has been introduced by Huang et al. [62]. Both the bias of ELM and the spawning of input weights in a random fashion are the source of major anomalies between classical DL and ELMs, which are prone to fast learning speed [63]. On the other hand, ELM is a very straightforward, naive, and efficient algorithm, which stands on the principle of Single Layer Feed-Forward Neural Networks (SLF-FNN).

ELM is a special architecture of a Multi-Layer Neural Network, consisting of a Single-Layer Feed-Forward Neural Network (SLFFNN) equipped with hidden neurons, and designated input weights. Further, it has random bias values in the hidden layer, whereas the output is computed by utilizing single multiplication of the weights of the vector matrix. ELM utilizes SLFFNN due to its sufficiency to advance any continuous function and to assert any discontinuous area. If we compare the efficiency of ELM and traditional Back Propagation Neural Networks (BP-NN), the ELM has far better learning time for N sample data. In ELM settings, parameters being used in the hidden layers are absolute i.e., independent of the data. Further, hidden level parameters for the function of activation are haphazardly generated before the perception of training data [64], [65].

G. Deep Extreme Learning Machines

DELM is normally employed for regression and classification purposes in diverse settings since its rate of learning is very rapid. Further, it is very effective as far as the rate of computational convolution is concerned. Classical ANN algorithms require sophisticated measurements and the learning times should be very slow. Further, they can override the learning model given by Khan [66]. Deep Extreme Learning Machines (DELMs) take advantage and benefit from both ELMs and DL techniques. In the phenomenon of DELMs, hidden layers get enhanced in the network structure of ELM and the random initialization process for the weights of input-layer and initial hidden-layer weights along with the initial hidden layer's bias.

Besides, DELMs utilize a new way to calculate the parameters to be used in all the hidden layers except the first using the Least Square Method (LSM) for calculating the output network weight [67]. DELM consists of the architecture with the multi-layer network which has been distributed in two halves. The first half of these two halves learns the original data in depth to get the most representative novel data through the usage of ELM-AE. As far as the second half is concerned, it calculates the parameters of the network by utilizing the kernel ELM algorithm [68].

There are two main steps of the Deep Extreme Learning Machines DELMs based classification of image-set. In the first step, the global domain-specific DELM model is found by utilizing training images. In the second step, an initialization building of class-specific DELM is carried out through the usage of global representation. It is to be noted that the encoding of both domain-level, as well as class-specific properties of data, is essential in the performance of both steps [69].

IV. Discussion and Conclusions

This review paper explored various studies, research papers, and scholarly works related to the integration of phonocardiography and machine learning. This paper provides a detailed introduction that highlights the importance of phonocardiography and machine learning in the field of healthcare, providing background information on phonocardiography by discussing the components of the phonocardiogram and the challenges associated with signal acquisition and processing. Transitioning to machine learning, the algorithms used in analyzing phonocardiograms and the steps involved in developing a machine learning model.

In the view of the research that we have collected that most of the researchers use Mel-frequency cepstral coefficients (MFCC), which is one of the conventional feature extraction techniques. Looking at the machine learning techniques that researchers have tried, we found that two techniques are the most used and the highest performing, which are support vector machine and Artificial neural network (SVM, ANN).

Review outcomes provide a roadmap for future research by pinpointing underexplored areas, automated detection and classification of heart sounds for the early detection of cardiac abnormalities. Studies in arrhythmia detection, heart sound segmentation, and abnormality classification. The paper concludes with a discussion on the limitations and future directions of phonocardiography and machine learning, addressing challenges and suggesting future research possibilities including deep learning models. The presented literature review provides a comprehensive overview of phonocardiography and machine learning, covering fundamental concepts, applications, and potential advancements in the field.

Acknowledgment

I want to express my profound appreciation and gratitude to the supervisor for their unwavering support and encouragement throughout the development of this paper. Their active participation and valuable insights have played a pivotal role in shaping the direction and content of our research. Additionally, I extend heartfelt thanks to the reviewers for their diligent assessment and constructive feedback, which has significantly enhanced the quality and credibility of this work. This paper would not have been possible without the valuable contributions of both the supervisor and the reviewers, and I am immensely thankful for their involvement in this research journey.

References

[1] Li, J., Guo, Y., Wang, C., and Li, Z. (2020). "A Robust Speech Recognition System Based on MFCC and CNN." 2020 IEEE 3rd International Conference on Information and Computer Technologies (ICICT), pp. 62-66.

[2] Li, J. P., Haq, A. U., Din, S. U., Khan, J., Khan, A., & Saboor, A. (2020). Heart disease identification method using machine learning classification in e-healthcare. IEEE Access, 8, 107562-107582.

[3] Chakir, F.; Jilbab, A.; Nacir, C.; Hammouch, A. Phonocardiogram signals processing approach for PASCAL Classifying Heart Sounds Challenge. Signal Image Video Process. 2018, 12, 1149–1155.

[4] Jamil, S., & Roy, A. M. (2023). An efficient and robust Phonocardiography (PCG)-based Valvular Heart Diseases (VHD) detection framework using Vision Transformer (ViT). Computers in Biology and Medicine, 106734.

[5] Chen, W., Sun, Q., Chen, X., Xie, G., Wu, H., & Xu, C. (2021). Deep learning methods for heart sound classification: a systematic review. Entropy, 23(6), 667.

[6] Chang, V., Bhavani, V. R., Xu, A. Q., & Hossain, M. A. (2022). An artificial intelligence model for heart disease detection using machine learning algorithms. Healthcare Analytics, 2, 100016.

[7] Gahane, A., & Kotadi, C. (2022, February). An Analytical Review of Heart Failure Detection based on IoT and Machine Learning. In 2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS) (pp. 1308-1314). IEEE.

[8] Kernbach, J. M., & Staartjes, V. E. (2022). Foundations of machine learning-based clinical prediction modeling: Part II—Generalization and overfitting. Machine Learning in Clinical Neuroscience: Foundations and Applications, 15-21.

[9] Yildirim, M. (2022). Diagnosis of Heart Diseases Using Heart Sound Signals with the Developed Interpolation, CNN, and Relief Based Model. Traitement du Signal, 39(3).

[10] Athreya, A. M., Paramesha, K., Avani, H. S., & Madhu, S. (2022). Neural Networks for Detecting Cardiac Arrhythmia from PCG Signals. In Intelligent Vision in Healthcare (pp. 103-115). Springer, Singapore.

[11] Li, S, Li, F, Tang, S, Xiong, W "A Review of Computer-Aided Heart Sound Detection Techniques", BioMed Res. pp: 1–10, 2020.

[12] Raza, A., Mehmood, A., Ullah, S., Ahmad, M., Choi, G. S., & On, B. W. (2019). Heartbeat sound signal classification using deep learning. Sensors, 19(21), 4819.

[13] Deperlioglu, O. (2021). Heart sound classification with signal instant energy and stacked autoencoder network. Biomedical Signal Processing and Control, 64, 102211.

[14] Khan, M. U., Ali, S. Z. E. Z., Ishtiaq, A., Habib, K., Gul, T., & Samer, A. (2021, July). Classification of Multi-Class Cardiovascular Disorders using Ensemble Classifier and Impulsive Domain Analysis. In 2021 Mohammad Ali Jinnah University International Conference on Computing (MAJICC) (pp. 1-8). IEEE.

[15] Li, S, Li, F, Tang, S, Xiong, W "A Review of Computer-Aided Heart Sound Detection Techniques", BioMed Res. pp: 1–10, 2020.

[16] Thalmayer, A, Zeising, S, Fischer, G, Kirchner. J. A. "Robust and Real-Time Capable Envelope-Based Algorithm for Heart Sound Classification: Validation under Different Physiological Conditions", Sensors 2020.

[17] Kapen, P.T.; Youssoufa, M.; Kouam, S.U.K.; Foutse, M.; Tchamda, A.R.; Tchuen, G. Phonocardiogram: A robust algorithm for generating synthetic signals and comparison with real life ones. Biomed. Signal Process. Control 2020, 60, 101983.

[18] Giordano, N.; Knaflitz, M. A Novel Method for Measuring the Timing of Heart Sound Components through Digital Phonocardiography. Sensors 2019, 19, 1868.

[19] Wei, W.; Zhan, G.; Wang, X.; Zhang, P.; Yan, Y. A Novel Method for Automatic Heart Murmur Diagnosis Using Phonocardiogram. In Proceedings of the 2019 International Conference on Artificial

Intelligence and Advanced Manufacturing, AIAM, Dublin, Ireland, 16–18 October 2019; Volume 37, pp. 1–6.

[20] Malarvili, M.; Kamarulafizam, I.; Hussain, S.; Helmi, D. Heart sound segmentation algorithm based on instantaneous energy of electrocardiogram. Comput. Cardiol. 2003, 2003, 327–330.

[21] Chen, T.E.; Yang, S.I.; Ho, L.T.; Tsai, K.H.; Chen, Y.H.; Chang, Y.F.; Wu, C.C. S1 and S2 heart sound recognition using deep neural networks. IEEE Trans. Biomed. Eng. 2017, 64, 372–380.

[22] Liu, Q.; Wu, X.; Ma, X. An automatic segmentation method for heart sounds. Biomed. Eng. Online 2018, 17, 22-29.

[23] Dixon, P., Pavan, A., & Vinodchandran, N. V. (2018). On pseudodeterministic approximation algorithms. In 43rd International Symposium on Mathematical Foundations of Computer Science (MFCS 2018). Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.

[24] Kamson, A.P.; Sharma, L.; Dandapat, S. Multi-centroid diastolic duration distribution based HSMM for heart sound segmentation. Biomed. Signal Process. Control. 2019, 48, 265–272.

[25] Renna, F.; Oliveira, J.H.; Coimbra, M.T. Deep Convolutional Neural Networks for Heart Sound Segmentation. IEEE J. Biomed. Health Inform. 2019, 23, 2435–2445.

[26] Liu, C.; Springer, D.; Clifford, G.D. Performance of an open-source heart sound segmentation algorithm on eight independent databases. Physiol. Meas. 2017, 38, 1730–1745.

[27] Deng, S. W., & Han, J. Q. (2016). Towards heart sound classification without segmentation via autocorrelation feature and diffusion maps. Future Generation Computer Systems, 60, 13-21.

[28] Aziz, S., Khan, M. U., Alhaisoni, M., Akram, T., & Altaf, M. (2020). Phonocardiogram signal processing for automatic diagnosis of congenital heart disorders through the fusion of temporal and cepstral features. Sensors, 20(13), 3790.

[29] Kumar, V., Arora, A., Suri, P., & Arora, V. (2022). Managing arrhythmias-A guide to Physicians. J Med Sci Res, 10(1), 30-38.

[30] Pugliese, M., Biondi, V., La Maestra, R., & Passantino, A. (2021). Identification and clinical significance of heart murmurs in puppies involved in puppy trade. Veterinary Sciences, 8(8), 139.

[31] Jagannath, D. J., Dolly, D. R. J., & Peter, J. D. (2020). Composite Deep Belief Network approach for enhanced Antepartum foetal electrocardiogram signal. Cognitive Systems Research, 59, 198-203.

[32] Almanifi, O. R. A., Ab Nasir, A. F., Razman, M. A. M., Musa, R. M., & Majeed, A. P. A. (2022). Heartbeat murmurs detection in phonocardiogram recordings via transfer learning. Alexandria Engineering Journal, 61(12), 10995-11002.

[33] Lubaib, P., & Muneer, K. A. (2016). The heart defect analysis based on PCG signals using pattern recognition techniques. Procedia Technology, 24, 1024-1031.

[34] Sujadevi, V. G., Mohan, N., Sachin Kumar, S., Akshay, S., & Soman, K. P. (2019). A hybrid method for fundamental heart sound segmentation using group-sparsity denoising and variational mode decomposition. Biomedical Engineering Letters, 9(4), 413-424.

[35] Noman, F., Ting, C. M., Salleh, S. H., & Ombao, H. (2019, May). Short-segment heart sound classification using an ensemble of deep convolutional neural networks. In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 1318-1322). IEEE.

[36] Krishnan, P. T., Balasubramanian, P., & Umapathy, S. (2020). Automated heart sound classification system from unsegmented phonocardiogram (PCG) using deep neural network. Physical and Engineering Sciences in Medicine, 43(2), 505-515.

[37] El Badlaoui, O., Benba, A., & Hammouch, A. (2020). Novel PCG analysis method for discriminating between abnormal and normal heart sounds. Irbm, 41(4), 223-228.

[38] Li, H., Wang, X., Liu, C., Zeng, Q., Zheng, Y., Chu, X., ... & Karmakar, C. (2020). A fusion framework based on multi-domain features and deep learning features of phonocardiogram for coronary artery disease detection. Computers in biology and medicine, 120, 103733.

[39] Karan, B., Thakur, G., Rath, A., & Sahu, S. S. (2021, September). Heart Sound Abnormality Detection using Wavelet Packet Features and Machine Learning. In 2021 International Symposium of Asian Control Association on Intelligent Robotics and Industrial Automation (IRIA) (pp. 310-314). IEEE.

[40] Dhar, P., Dutta, S., & Mukherjee, V. (2021). Cross-wavelet assisted convolution neural network (AlexNet) approach for phonocardiogram signals classification. Biomedical Signal Processing and Control, 63, 102142.

[41] Chen, T. E., Yang, S. I., Ho, L. T., Tsai, K. H., Chen, Y. H., Chang, Y. F., ... & Wu, C. C. (2016). S1 and S2 heart sound recognition using deep neural networks. IEEE Transactions on Biomedical Engineering, 64(2), 372-380.

[42] Davis, S.B., and Mermelstein, P. (1980). "Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences." IEEE Trans. Acoust. Speech Signal Process., vol. 28, no. 4, pp. 357-366.

[43] Ghosh, S., Tripathy, S., Saha, G., and Paul, S. (2021). "An Improved Feature Extraction Method for Speech Recognition using Deep Learning." 2021 IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), pp. 728-733.

[44] Wang, Y., Huang, W., and Zhao, Z. (2020). "MFCC and CQCC-based Automatic Speaker Recognition Using Convolutional Neural Network." 2020 9th International Conference on Systems and Control (ICSC), pp. 49-53.

[45] Nguyen, T.T., Wang, R., and Nguyen, T.T. (2020). "A Comparative Study of Feature Extraction Techniques for Vietnamese Speech Recognition." 2020 IEEE 17th International Conference on Mobile Ad Hoc and Sensor Systems (MASS), pp. 188-193.

[46] Azevedo, J. M., Costa, M. J., & Neto, J. P. (2020). A comparative study on MFCC and MEL spectrogram features for automatic speech recognition. Proceedings of the 2020 IEEE International Conference on Signal and Image Processing Applications (ICSIPA), 407-412.

[47] Almogbel, S., Alshammari, R., & Alshamrani, M. (2020). Exploring the impact of MFCC and its derivatives on EEG-based emotion recognition. Proceedings of the 2020 7th International Conference on Signal Processing and Integrated Networks (SPIN), 88-93.

[48] Mohamed, M. A., Abo-Elsoud, M. A., & El-Said, S. A. (2020). MFCC-based speech enhancement using deep neural networks. Proceedings of the 2020 International Conference on Computer and Information Sciences (ICCIS), 1-6.

[49] Sivaprakasam, S., Alagappan, K., & Sundararajan, K. (2020). MFCC and its derivatives based features extraction for automated speech recognition of spontaneous Tamil. Proceedings of the 2020 4th International Conference on Computing Methodologies and Communication (ICCMC), 410-416.

[50] Mohapatra, S., & Lenka, S. K. (2020). Comparison of MFCC and Gammatone filterbank features for speech emotion recognition. Proceedings of the 2020 International Conference on Advances in Computing, Communication, & Control (ICAC3), 1-6.

[51] Wu, J., Gao, H., & Gao, L. (2021). An improved MFCC algorithm for speech recognition. Proceedings of the 2021 11th International Conference on Information Science and Technology (ICIST), 418-423.

[52] Sun, Y., Wang, X., & Yu, L. (2021). A comparative study of MFCC and Gammatone filterbank features for phoneme recognition. Proceedings of the 2021 IEEE 14th International Conference on Advanced Infocomm Technology (ICAIT), 178-183.

[53] Yang, Z., Chen, J., & Yang, Y. (2021). Speaker recognition based on MFCC and XGBoost. Proceedings of the 2021 IEEE 4th International Conference on Signal Processing Systems (ICSPS), 138-142.

[54] Liu, X., Chen, M., & Fu, Y. (2021). A comparative study of MFCC and deep learning features for speech emotion recognition. Proceedings of the 2021 IEEE 23rd International Conference on High Performance Computing and Communications (HPCC), 367-372.

[55] Hossain, M., Islam, M. R., & Razzak, M. A. (2021). MFCC-based speech separation using convolutional recurrent neural networks. Proceedings of the 2021 IEEE 7th International Conference on Computer and Communication Systems (ICCCS), 356-361.

[56] Ahsan, M. M., Luna, S. A., & Siddique, Z. (2022, March). Machine-learning-based disease diagnosis: A comprehensive review. In Healthcare (Vol. 10, No. 3, p. 541). MDPI.

[57] Amjad, M., Raza, H., Muneer, S., & Aslam, M. A. (2022). A Systematic Review on Brain Tumor Detection Using Machine Learning. Journal of NCBAE, 1(4), 11-16.

[58] Masoumian, A., Rashwan, H. A., Cristiano, J., Asif, M. S., & Puig, D. (2022). Monocular depth estimation using deep learning: A review. Sensors, 22(14), 5353.

[59] Rizvi, S. S. R., Khan, M. A., Abbas, S., Asadullah, M., Anwer, N., & Fatima, A. (2022). Deep extreme learning machine-based optical character recognition system for Nastalique Urdu-like script languages. The Computer Journal, 65(2), 331-344.

[60] Kurani, A., Doshi, P., Vakharia, A., & Shah, M. (2023). A comprehensive comparative study of artificial neural network (ANN) and support vector machines (SVM) on stock forecasting. Annals of Data Science, 10(1), 183-208.

[61] Patel, L., Shukla, T., Huang, X., Ussery, D. W., & Wang, S. (2020). Machine learning methods in drug discovery. Molecules, 25(22), 5277.

[62] Huang.G.-B, Zhu.Q.-Y, Siew.C.K "Extreme Learning Machine: Theory and Applications", Neurocomputing, pp:1-13, 2005.

[63] Gu.Y, Chen.Y, Liu.J, Jian.X "Semi-Supervised Deep Extreme Learning Machine for Wi-Fi based Localization", Neurocomputing, pp: 282-293, 2015.

[64] Demertiz.K, Lliadis.L.S, Anezakis.V "Extreme Deep Learning in Biosecurity: The Case of Machine Hearing for Marine Species Identification", Journal of Information and Telecommunication, pp: 492-510, 2018.

[65] Yao.L, Ge.Z "Deep Learning of Semi-supervised Process Data with Hierarchical Extreme Learning Machine and Soft Sensor Application", IEEE Transactions on Industrial Electronics, pp: 1-8, 2017.DOI: 10.1109/TIE.2017.2733448.

[66] Khan, M. U., Samer, S., Alshehri, M. D., Baloch, N. K., Khan, H., Hussain, F., ... & Zikria, Y. B. (2022). Artificial neural network-based cardiovascular disease prediction using spectral features. Computers and Electrical Engineering, 101, 108094.

[67] Fayaz.M, Kim.D "A Prediction Methodology of Energy Consumption Based on Deep Extreme Learning Machine and Comparative Analysis in Residential Buildings" Electronics, 1-22, 2017.

[68] Xiao.D, Li.B, Mao.Y "A Multiple Hidden Layers Extreme Learning Machine Method and Its Application", Mathematical Problems in Engineering, pp. 1-10, 2017.

[69] Uzair.M, Shafait.F, Ghanem.B, Mian.A "Representation of Learning with Deep Extreme Learning Machines for Efficient Image Set classification", Neural Comput & Applic, pp. 1-13, 2016. DOI: https://doi.org/10.1007/s00521-016-2758-x.

[70] Sujadevi, V. G., Soman, K. P., Vinayakumar, R., & Sankar, A. P. (2017, December). Deep models for phonocardiography (PCG) classification. In 2017 international conference on intelligent communication and computational techniques (ICCT) (pp. 211-216). IEEE.