International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 16 (2024) pp. 131–147
© Cerebration Science Publishing

Article

Heart Sound Analysis with Machine Learning Using Audio Features for Detecting Heart Diseases

Sathyanarayanan Swaminathan *, Srikanta Murthy Krishnamurthy, Chandrashekar Gudada, Satish Kumar Mallappa and Neeraj Ail

Department of Mathematical and Computational Sciences Sri Sathya Sai University for Human Excellence, Gulbarga, Karnataka 585313, India; srikantamurthy.k@sssuhe.ac.in (S.M.K.); chandrugudada@gmail.com (C.G.); satishkumar697@gmail.com (S.K.M.); neerajail98@gmail.com (N.A.)
* Correspondence author: sathyanarayanan.s@sssuhe.ac.in or sathyanarayanan.brn@gmail.com

Received date: 5 March 2024; Accepted date: 21 March 2024; Published online: 24 May 2024

Abstract: Society has been impacted by the influence of artificial intelligence (AI) across various aspects of daily life over the last few years. AI can positively impact healthcare by making it cheaper, quicker, more effective, and more accessible. AI has impacted the detection and treatment of cardiovascular diseases (CVD). The analysis of heart sound recordings using AI has been studied during the last decade in the study of noninvasive diagnosis of CVD. This study aims to construct a machine learning model that requires the least computational resources and computation time for the classification of heart sounds using a novel set of time-domain, frequency-domain, and statistical-domain features extracted from heart sound recordings. A public dataset of heart sound recordings comprising five classes including one normal category and four different categories of valvular diseases was used in this study. Two combinations of data were used in the experiments. The first combination dataset consisted of normal and abnormal heart sounds. The second combination consisted of heart sound recordings belonging to one normal category and four different valvular diseases of heart sounds. The model's performance can be deemed excellent, with the first combination giving an accuracy of 99.89% and the second combination giving an overall accuracy of 99.26%.

Keywords: phonocardiogram; cardiovascular disease; machine learning; auscultation; MFCC

1. Introduction

Artificial Intelligence (AI) has impacted society in several ways in various domains. Several domains of healthcare have been affected by AI, including radiology, cancer detection, and diabetic retinopathy. AI can reduce healthcare costs, make it more accessible, and improve the speed of response by using it for screening, detection, diagnosis and prognosis [1]. A few examples are provided here to demonstrate the application of AI in diverse areas of healthcare. Heli Shah et al. [2] used machine learning (ML) to generate a treatment plan for oral cancer. Chandrika R et al. [3] have proposed a hybrid segmentation for the detection and classification of lung cancer. Maalej et al. [4] used transfer learning and data augmentation techniques [4] to classify breast cancers. A hybrid deep learning (DL) network was proposed for skin lesion segmentation by Rout et al. [5].

According to a report by the World Heart Federation (WHF) released in 2023, 20.5 million deaths due to cardiovascular diseases (CVDs) occurred in 2021. Timely detection and treatment can prevent 80% of stroke and heart attacks [6]. One of the major factors is congenital heart disease (CHD), which affects one out of every 100 children born [7]. Only approximately 10% of children with CHD have access to proper healthcare. Developing countries have significantly higher mortality rates due to CHD than developed countries [8]. Communities with lower socioeconomic status are at a higher risk of

CVDs [9].

Advancements in treatment techniques have occurred over the last few decades. However, the sections of society that need it do not have access to quality treatment as the development is concentrated in richer countries. This healthcare gap must be addressed by integrating AI into healthcare to make it more accessible and cheaper. One way to enhance access to cardiovascular care for weaker sections of society is to integrate AI into cardiovascular care to make it more accessible, quicker, cheaper, and more effective. AI is being applied in CVD detection and treatment to make the treatment more effective and to help in early detection.

Early detection: Early detection of cardiovascular disease can be beneficial in both treating and preventing it for a variety of reasons. First, it enables timely intervention through lifestyle adjustments, medications, or minimally invasive procedures, possibly halting the progression of the disease and its complications. The long-term health and quality of life of patients can be significantly improved using this approach [10]. Second, early detection often involves identifying and managing underlying risk factors such as high cholesterol, high blood pressure, and diabetes. By addressing these factors through lifestyle modifications or pharmacotherapy, individuals can reduce the risk of developing further complications. Third, it serves to prevent serious complications, such as heart attacks, strokes, and heart failure, which can be dangerous to life and place a burden on healthcare systems. Proactive management can reduce the healthcare costs associated with the treatment of advanced stages of the disease. Finally, early diagnosis promotes increased awareness among individuals, motivating them to adopt healthier lifestyle habits, undergo regular screenings, and follow treatment plans, contributing to improved health outcomes and potentially reducing the burden on healthcare systems [11].

Researchers began using digital signal processing techniques more than three decades ago to automate the detection of heart diseases by analysing heart sounds. The researchers then began using machine learning (ML) techniques to analyse phonocardiograms. However, a very high accuracy could not be achieved. DL techniques have been applied during the last decade to increase the accuracy and make the system more effective. However, building DL models requires more data samples and a large amount of computational power. However, they have not reached the desired levels of accuracy and effectiveness, where they can be integrated into an electronic stethoscope. There is a need to find methods which require minimal computing power, reach nearly 100% accuracy and sensitivity, and are sufficiently small to be embedded in an electronic stethoscope or an equivalent portable device.

This study aims to fill this research gap by finding a method that requires the least computing resources and provides high values for all important performance metrics. This study aimed to develop a classification model that uses a novel combination of audio features as input, is less computationally intensive, and uses low-end hardware. ML techniques were used in this study to classify heart sounds as diseased or normal using features extracted from heart sounds. Furthermore, abnormal heart sound recordings were classified into one of the four valvular diseases. This method is a low-cost, simple, and noninvasive technique for detecting CVDs.

Several studies have been conducted in this area, using audio features [12] and spectrograms generated from audio data samples [13]. Some researchers have used time-domain and spectral-domain features of audio datasets. This study examined the efficiency of various ML algorithms for classifying heart sounds based on audio features. Broadly, these features fall into four categories: time-domain, frequency-domain, time-frequency domain, and statistical domain features. Researchers have used ML and deep learning (DL) algorithms, some of which are computationally intensive and require high-end hardware.

The list of features included mel-frequency cepstral coefficients (MFCCs), energy (RMS), spectral flux, spectral rolloff, spectral bandwidth, spectral contrast, energy, mel-spectrogram, FFT values, degree of periodicity, spectral entropy, short-time energy, spectral centroid, flatness, power mean value, normalised signal, temporal crest factor, discrete wavelet transform, zero-crossing rate, skewness, kurtosis, peak frequency, mean frequency, median frequency, LPC_AVG, and high-frequency distortion.

ML algorithms are used to ensure accurate classification while processing large amounts of data. As a result, the accuracy and efficiency are enhanced in classifying heart sounds and detecting cardiac events. Five classification methods have been applied: multilayer perceptron (MLP), support vector machines (SVM), k-nearest neighbours (kNN), naive Bayes (NB), and random forests (RF).

2. Background

2.1. Heart Murmurs

The heart generates extra sounds called murmurs, which are generated by the turbulent blood flow within the heart. Murmurs can be caused by various factors, including valve defects and congenital heart defects (CHD). Murmurs can be classified according to their timing, duration, and intensity to provide clues regarding the basic cause.

Listening to heart sounds to understand the details of heart sound and murmurs and interpreting heart sounds for detecting heart diseases is called auscultation. This is the traditional technique used by the physicians. However, it is highly subjective, requires training, and has the possibility of a wrong diagnosis.

2.2. Automated Analysis of Heart Sounds

Automated analysis of heart sounds using ML and DL methods has been studied over the last several years. Cardiovascular disease (CVD) can be detected early using this method [14]. Different heart sound patterns can be identified and categorised using it, which can offer an important understanding of a patient's cardiovascular health [15]. The accurate heart sound classification can also reduce the burden on healthcare systems by enabling more efficient allocation of resources and reducing healthcare costs [14].

Audio recordings of heart sounds, called phonocardiograms, have been analysed using ML techniques to detect heart disease. One approach is the segmentation of heart sounds after identifying the fundamental heart sounds and subsequently extracting features from the audio which are then channelled as input to the ML algorithms. Some researchers have extracted audio features from unsegmented audio data to analyse heart sounds. Several researchers have provided raw audio data as input directly to DL algorithms for classification.

2.3. Audio Features

For this training, 27 audio features were selected for this training. These audio features are classified into four broad categories: frequency, time, time-frequency domain, and statistical domain features. These features help in studying the condition of the heart, as any abnormality in the heart can affect some of the features. A weakened heart may lead to a lower amplitude (energy). Irregular heart valve functioning may alter the heart rate and hence affect the frequency spectrum. The murmurs and additional sounds may be of different frequencies than a normal heart sound. The time-domain features can be affected by valvular diseases as the duration of various phases within the heart cycle is affected. The selected features are discussed below.

Discrete wavelet transforms: Discrete wavelet transforms can be used to extract different frequency bands. This allows the extraction of both time and frequency information. The DWT can capture transient changes and variations in sounds.

$$\mathrm{DWT}(x, y) = \int_{-\infty}^{\infty} Z(t) \frac{1}{\sqrt{|2^x|}} \varphi\left(\frac{t - 2^x y}{2^x}\right) dt \tag{1}$$

MFCCs: MFCCs, which are acronyms for Mel Frequency Cepstral Coefficients, are utilised to record the spectral envelope of an audio signal, thereby depicting the dispersion of sound energy at various frequencies. MFCCs mimic the human auditory system's perception of pitch and loudness [16]. The Mel-scale frequency was derived using (2):

$$\mathrm{Mel}\ (f) = 2595 \log (1 + f/100) \tag{2}$$

Extracting MFCCs from heart sound recordings helps to obtain the coefficients of the dominant frequencies and their relative strengths in the sample. This information can be analysed to characterise normal heart sounds. S1 and S2, the main heart sounds, have characteristic MFCC patterns based on the valve-closure frequencies. Abnormal sounds caused by turbulent blood flow can manifest as distinct changes in the MFCC spectrum, highlighting additional frequency components or altering the relative strengths of existing ones. Specific MFCC features can help to distinguish murmurs arising from different valve problems (e.g., mitral regurgitation vs. aortic stenosis).

The MFCCs are robust to noise. They are relatively resistant to background noise and environmental factors, making them reliable for analysing real-world recordings. MFCCs reduce high-dimensional frequency information into a set of lower-dimensional coefficients, thereby reducing the computational complexity.

Calculating and extracting MFCCs from an audio signal involves the following steps [17].

(Optional) Pre-emphasize high frequencies.

Divide the signal into short, overlapping frames.

Apply the windowing function to each frame.

Each frame is converted using a Fast Fourier Transform (FFT) to the frequency domain.

Apply mel-scale filtering and logarithmic compression to mimic human hearing.

Use the Discrete Cosine Transform (DCT) to extract a low-dimensional feature vector.

The formula for the MFCCs can be summarised as follows:

$$\text{MFCC} = \text{DCT}\ (\log\ (\ \text{MelFilter}\ (\text{FFT}(\text{Window}\ (\ \text{Pre-emphasis}\ (\ \text{signal}))))))\qquad(3)$$

The resulting MFCC features capture the spectral characteristics of audio in a manner relevant to human perception, making them valuable in applications such as speech recognition and speaker identification.

Spectral rolloff: Spectral rolloff is the frequency below which a certain percentage of the total signal energy is present. Data on the frequency content of heart sound recording can be indicative of certain cardiovascular conditions [18].

$$Spectral\ Rolloff = 0.9 * \sum_{n=1}^{N} |x(n)| \qquad (4)$$

where x(n) is the frequency-component amplitude in bin *n*.

Spectral entropy: The spectral entropy of a signal measures the randomness of its frequency distribution. Normal heart sounds can be distinguished from abnormal sounds using this technique [18] because noise has a higher entropy value. Mathematically, entropy is given by (5).

$$Entropy = \sum_{n=1}^{N} -x(n) * log2x(n) \qquad (5)$$

Temporal crest factor: The peak-to-peak amplitude of an audio sample divided by the root mean square of the signal over a given time interval is called the temporal crest factor, and is given by (6). It helps in the differentiation of various types of heart sounds and identification of distinct patterns associated with different cardiovascular conditions [18].

$$CF_t = \frac{Peak\ amplitude}{RMS\ amplitude} \qquad (6)$$

Spectral flatness: A signal's spectral flatness is determined by how evenly it is distributed across different frequencies and is the ratio of the geometric mean to the arithmetic mean of the power spectrum. Higher values indicate more noise, whereas lower values indicate tonal or harmonic characteristics.

$$SF = \frac{Geometric\ mean\ of\ the\ power\ spectrum}{Arithmetic\ mean\ of\ the\ power\ spectrum} \qquad (7)$$

Short-time energy: Short-time energy measures the energy of a signal within a short period and is given by (8). This provides information regarding the intensity or amplitude variations in the heart sound signal, which can be indicative of certain cardiac abnormalities [15,18].

$$STE = \frac{1}{N} \sum_{n=1}^{N} [x(n) * w(m-n)]^2 \qquad (8)$$

Degree of periodicity: The frequency with which a signal repeats is called periodicity. Normal heart sounds can exhibit some degree of periodicity. Pathological heart sounds could be less periodic owing to irregularities in valve function.

Spectral energy: The dispersal of the signal energy across the frequencies of the audio sample is given by the spectral energy. This feature can help to identify abnormal patterns or characteristics associated with different heart conditions.

$$E(f) = |X(f)|^2 \qquad (9)$$

where E(f) is the spectral energy at frequency f, X(f) is the Fourier transform of the audio signal x(t) evaluated at frequency f, and |X(f)| represents the magnitude of X(f), which captures the amplitude of the frequency component of audio data at frequency f. This formula calculates the energy content of a signal at a specific frequency by squaring the magnitude of the Fourier transform coefficient at that frequency, which provides a measure of the energy distribution of the signal across different frequency components.

Spectral contrast: Spectral contrast is a feature used in audio signal processing to quantify the difference in energy between the peaks and valleys in the spectrum of an audio signal. Spectral contrast can help to identify abnormalities or changes in the frequency content associated with specific cardiovascular conditions.

$$C_b = \frac{\max(\mu_b) - GM}{GM} \qquad (10)$$

where μ_b is the average energy in frequency band b, GM is the geometric mean of the average energies over all the bands, and C_b is the spectral contrast for each band.

Spectral bandwidth: The frequency range in which most of the signal energy resides is called spectrum bandwidth. This feature provides insights into the distribution and intensity of frequencies present in heart sounds [15].

$$BW = \sqrt{\sum_{k=1}^{N} P(f_k)\ .\ (f_k - f_c)^2} \qquad (11)$$

Fast Fourier Transform values: FFT values are commonly used in heart sound analysis to determine the signal's frequency content. Using FFT, the frequency components linked to CVDs can be identified from the time domain of the heart sound signal [18]. The FFT coefficients are determined by applying discrete Fourier transforms to each signal, as given in (12), to obtain the spectrum of the signals.

$$F(w) = \sum_{n=1}^{N} (f(n). e^{-} j2\,\pi wn/N) \tag{12}$$

Zero crossing rate: This rate measures the rate at which a signal crosses the horizontal axis (zero amplitude). This provides information about the periodicity of heart sounds [18].

$$ZCR = \frac{1}{N} \sum_{n=1}^{N} |(sign(x(n)) - \,sign\,(x(n-1))| \tag{13}$$

Spectral flux: The spectral flux measures the spectral energy changes in the spectral energy distribution over time. By analysing the spread of spectral flux values within different classes (normal vs. abnormal or specific valve diseases), the model can learn to associate changes in the spectrum with specific heart conditions [18].

$$SpectralFlux = \sum_{n=1}^{N} (f[x(n)] - f[x(n-1)]^2) \tag{14}$$

Spectral centroid: The spectral centroid identifies the main frequency existing in the heart sound by mapping the centre of mass of the power spectrum. Different heart sound patterns can be distinguished in this way [14,18]. It is calculated using the power spectra of the preceding and current frames, as given in (15).

$$SpectralCentroid = \frac{\sum_{n=1}^{N} n.f(x(n))}{\sum_{n=1}^{N} f(x(n))} \tag{15}$$

Skewness: The skewness of a probability distribution is a statistical measure of asymmetry. It provides information regarding the shape of the distribution of a feature [15]. Normal heart sounds have skewness values closer to 0. Certain types of abnormal sounds, such as murmurs, can exhibit more pronounced skewness [17].

$$Skewness = \frac{\sum_{k=b_1}^{b_2} (f_k - \mu_1)^3 s_k}{(\mu_2)^3 \sum_{k=b_1}^{b_2} s_k} \tag{16}$$

Kurtosis: Kurtosis is a statistical measure that quantifies the peakedness or flatness of a probability distribution. It assesses the extent to which the distribution deviates from a bell-shaped curve, by focusing on the presence of extreme values (outliers). Positive kurtosis can be observed in murmurs or extra heart sounds [17].

$$Kurtosis = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})}{\left(\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2\right)^2} - 3 \tag{17}$$

Median frequency: Median frequency refers to the frequency that divides the power distribution in a signal into two equal parts. Various valvular diseases can result in distinct median frequency patterns.

Mean frequency: Recorded heart sounds are represented by their average or mean frequency. Heart sounds with murmurs often exhibit dissimilar mean frequencies compared with normal sounds.

$$Mean(\mu) = \frac{\sum_r \sum_c \{J(a,b)\}}{A \times B} \tag{18}$$

Peak frequency: This is the frequency at which the signal reaches its maximum amplitude.

High-frequency distortion: The amount of noise or distortion present at high frequencies in the heart sound is called high-frequency distortion.

Linear Predictive Coding average: The Linear Predictive Coding average (LPC_AVG) calculates the average energy distribution across all frequencies in the heart sound signal. Normal versus abnormal sounds can be distinguished from LPC_AVG patterns and can assist in identifying specific valve diseases based on their characteristic frequency content. The formula for the LPC is given in Equation (19), the time index is denoted by n, the LPC coefficient index is denoted by k, and the residual prediction error is denoted by e(n). Signal redundancy is removed by predicting the next value as a linear combination of previous values, as shown in (19).

$$s(n) = \sum_{k=1}^{P} a_k . s\,(n-k) + e\,(n)) \tag{19}$$

Energy (RMS): Audio signals are measured by their energy (RMS) or Root Mean Square (RMS) energy and the average power (or intensity) over time, and are indicative of the amount of sound energy present in the signal. This is the average magnitude of a changing signal computed by taking the square root of the average of the squared values of the signal over a period. This method quantifies the overall “loudness” of the signal by considering its variations in amplitude. The term “energy” represents the amount of sound energy present in the signal.

$$E = \sqrt{\frac{1}{N}\sum_{n=0}^{N-1}\left[x[n]\right]^2} \tag{20}$$

where N is the total count of samples in the audio sample and x[n] is a sequence of samples of discrete-time signals.

Mel-spectrograms: Mel-spectrograms use the mel-scale, which reflects the non-linear response of the human auditory system to frequency, in contrast to the traditional spectrograms (A visual representation of the distribution of frequencies of an audio signal over a period through the application of the short-time Fourier transform (STFT) method on an audio signal) that employ a linear frequency scale [19]. By compressing the frequency axis at lower frequencies and expanding it at higher frequencies, mel-spectrograms mimic human perception of loudness, while the logarithmic intensity scale highlights changes in perceived loudness rather than absolute amplitude, effectively capturing temporal sound variations and exhibiting resilience to noise, with the mel-scale frequency being defined by Equation (2).

Normalised signals: This is the signal with the mean adjusted to zero and standard deviation adjusted to one. This is done to reduce the differences in the frequency ranges of the different types of sounds [20].

$$\mathbf{X} = \frac{\mathbf{X} - \min(\mathbf{X})}{\max(\mathbf{X}) - \min(\mathbf{X})} \tag{21}$$

Heart sounds with murmurs could have higher RMS values than normal sounds, owing to the increased energy in specific frequency bands. Cases such as weak heart muscle contraction or valvular insufficiency result in lower energy (RMS) values.

Among the several audio features that can be extracted from heart sound recordings for being given as input for ML models, the features that have added impact include spectral bandwidth, spectral centroid, zero crossing rate, FFT values, and spectral flux. These features capture a variety of aspects of an audio signal, such as the frequency distribution, rhythmic patterns, spectral energy changes, and dominant frequencies. A concise description of the features is given in Table 1.

Table 1. A list of audio features extracted and their description.

Feature	Description
Frequency domain Features	
MFCCs	The spectral envelope of the signal is represented on a mel-scale, like human perception.
Spectral Flux	Measure how quickly the spectrum changes over time.
Spectral Rolloff	Frequency above which 90% of the spectral energy lies.
Spectral Bandwidth	Difference between the frequencies at 75% and 25% of the spectral energy.
Spectral Contrast	Measure of spectral shape differences between adjacent frequency bands.
Spectral Energy	Distribution of energy across the frequency spectrum.
FFT Values	Magnitudes and phases of the signal's frequency components.
Spectral Entropy	A measure quantifying the spectral distribution's randomness
Spectral Centroid	Spectral energy distribution's centre of mass
Spectral Flatness	The ratio of total spectral energy to the energy in the highest frequency band.
Peak Frequency	The frequency with the maximum amplitude
Mean Frequency	The signal's average frequency
Median Frequency	Frequency that divides the signal's energy in half.
LPC_AVG	Average of the LPC coefficients, representing the formant frequencies.
High-Frequency Distortion	Measure the amount of high-frequency content lost due to processing or transmission.
Time domain features	

Energy	Signal’s overall loudness in a short time window.
ZCR	Number of times the signal crosses zero amplitude in a window.
Degree of Periodicity	Measure how periodic the signal is.
Short Time Energy	Energy within a short time window; is used for dynamic analysis.
Power Mean Value	Measure the overall level of the signal, capturing both high and low amplitudes.
Temporal Crest Factor	The ratio of the peak amplitude to the RMS amplitude.
Normalized Signal	Signal with mean and standard deviation adjusted to zero and one, respectively.
Time-Frequency doman features	
DWT	Multi-scale depiction of the signal using wavelets.
Mel Spectrogram	Visual illustration of the mel-filtered spectrogram over time.
Statistical domain features	
Skewness	Measure of the asymmetry of the signal's amplitude distribution.
Kurtosis	Measure of the "tailedness" or peakedness of the signal's amplitude distribution.

3. Related Work

Several researchers have attempted to develop automated algorithms for heart sound classification using ML techniques, and significant progress has been made. Sathyanarayanan et al. [21] have surveyed the work done during the last few years in heart sound classification using ML techniques. Some studies have focused on extracting relevant features from heart sound recordings and feeding them into ML algorithms to develop accurate classification models [22]. The models produced satisfactory results. to accurately categorise heart sounds and detect various cardiovascular conditions [14].

A significant contribution was made by Clifford et al. (2016), who organised the PhysioNet/Computing in Cardiology Challenge, focusing on the classification of abnormal and normal heart sound recordings. This challenge highlights the diversity of methodologies employed in the field, and offers a comparison of different algorithmic approaches. The results from this challenge indicated that ML, especially DL models, was reasonably effective in handling the complexity and variability inherent in heart sound data [23].

Mahesh Kumar et al. [24] experimented with a swift spectral analysis-driven statistical feature extraction technique for detecting the phases of Aortic Stenosis (AS) using heart sounds by identifying the statistical difference between AS heart sounds and normal heart sounds. Spectral statistical features can identify the normal/unhealthy condition of the heart, serving as a rapid predictor of AS.

In a study conducted by Li et al. [15], the authors utilised a set of audio features of heart sounds for classification purposes. Researchers have determined that MFCCs are among the most important features. The authors also claimed that spectral entropy, short-time energy, spectral centroid, and spectral flatness are important for the classification of heart sounds. An automated heart sounds assessment was conducted by Tang et al. [14]. They selected ten features without segmentation, including kurtosis, energy ratio, and sound periodicity, and achieved 90.4% accuracy for binary classification and 85.7% accuracy for triple classification.

Soto-Murillo et al. [25] calculated fifty-two audio features from three categories of analysed heart sounds, including Linear Predictive Coding (LPC) coefficients, statistical features, and MFCCs, and used them for further analysis and classification. They evaluated six classifiers: k-nearest neighbours, SVM, naive Bayes, logistic regression, decision trees, and Artificial Neural Networks (ANNs) and reported the highest accuracy of 70.73%.

Khan et al. [18] used frequency and time domain features for training the different models in automated heart sound classification without performing segmentation on the input signals. The extracted features include the spectral roll-off frequency, root mean square value, peak value, crest factor, LPC, MFCCs, entropy-based features, wavelet transform-based features, and features extracted from the power spectral density. Automated heart sound classification was performed using various algorithms including SVM, ANN, and Cartesian Genetic Programming Evolved ANNs (CGPANN). They reported an accuracy of 73.64% using SVM. The SVM outperformed the other algorithms.

Around ten features of PCGs were extracted from PCGs by Yadav et al. [12] and channelled into four different ML algorithms to build a classification model. Furthermore, four prominent features were selected and given as inputs for building ML models which resulted in higher accuracy. Upretee et al. [26] implemented a classification algorithm using only the spectral centroid feature and reported a 96% accuracy for multiclass classification using the Yaseen dataset.

Three heart sound classes were considered by Zeinali et al. [27], that is, normal, S3 and S4 sounds for multiclass classification. Statistical, signal, wavelet, and information theory features extracted from the audio dataset were processed using a feature selection algorithm to select the relevant features. ML classifiers, including gradient boosting, support vectors, and random forests were used for classification. They reported accuracies of 87.5% for multiclass classification and 98% for binary classification.

A study involving the extraction of texture recognition features HOG and LBP from spectrograms generated from an audio dataset which was then channelled to various ML algorithms as input, was conducted by Sathyanarayanan et al., with excellent results [13].

Several ML algorithms have been tested for classifying heart sounds [14]. These algorithms utilise features derived from the dataset as inputs to develop a classification model [22].

Gap in research: This study aims to build an ML model that uses a novel set of acoustic features derived from each training sample as input and classifies heart sounds. The ML model should not be computationally intensive in terms of both time and computing power. The model could then be embedded in an electronic stethoscope to detect heart disease. Therefore, screening can be performed by using this device.

4. Dataset Details

The heart sounds utilised for this research were related to valvular diseases and is obtained from the dataset curated by Yaseen et al. [16]. It comprises 1000 audio samples with an equal distribution of 200 samples across each of the five classes. The audio samples were filtered and converted into a monochannel format. These samples are characterised by a single-channel audio format, 128 kbps bit rate, 8 KHz sampling rate, and a sample size of 16 bits. The duration of the recordings ranged from 1 to 3 seconds, with the majority being 2 seconds in duration. To ensure uniformity in the training data, audio samples shorter than 2 s were excluded and longer samples were trimmed to 2 s. Because of this step, potential errors owing to uneven training data were avoided. The sample waveform for each category is shown in Figure 1.

According to Bao et al. [28], heart sound recordings lasting 2 s were deemed optimal for several reasons. Longer samples do not necessarily result in higher accuracy and can lead to unnecessary resource consumption. Moreover, audio data with a duration shorter than 2 s might not provide adequate information for the identification of certain patterns and the reduction of random errors. Consequently, this study employed 957 audio samples for analysis, as listed in Table 2.

Table 2. Details of the Yaseen dataset.

Class	Type	Number of samples
Normal (N)	Normal	200
Aortic stenosis (AS)	Abnormal	200
Mitral regurtitation (MR)		184
Mitral stenosis (MS)		186
Mitral valve prolapse (MVP)		187
Total		957

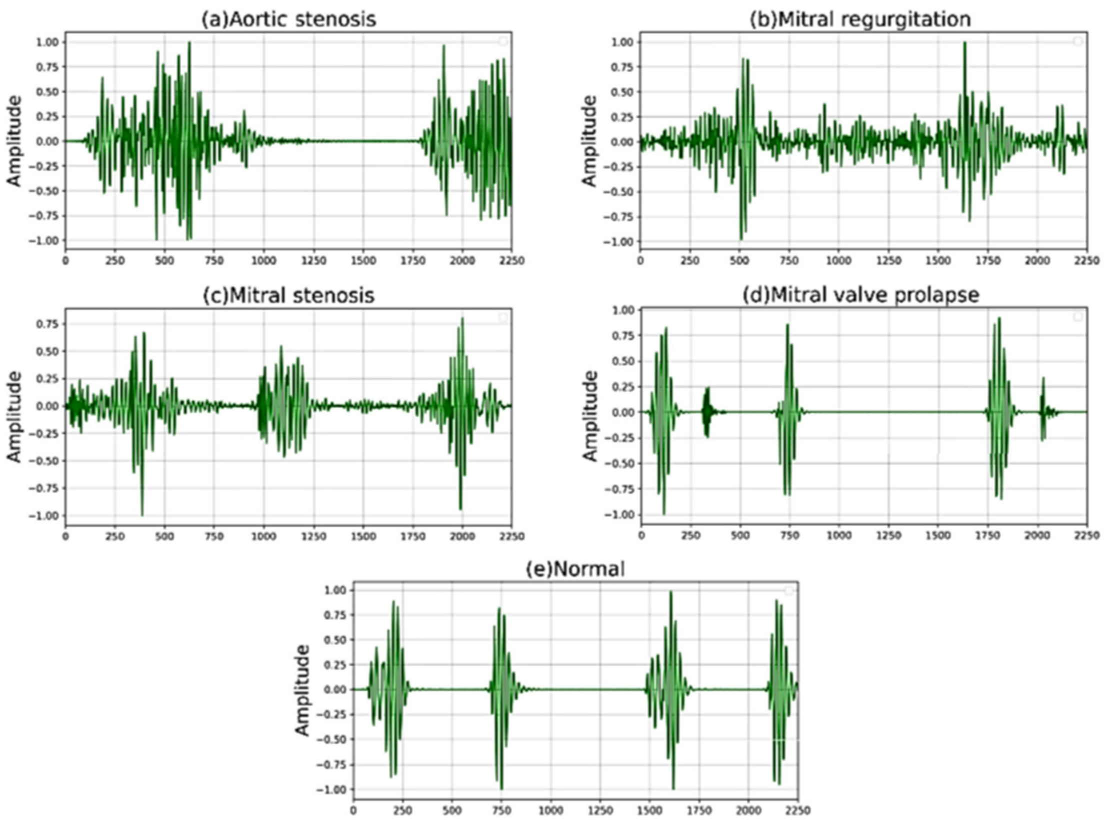

Figure 1. Waveforms of existing CVD classes in the PCG dataset [29].

5. Methodology

The Yaseen dataset was used in this study. The dataset consisting of audio recordings of cardiac sounds was preprocessed to ensure standardisation. Two variations of the dataset were used in the experiments. The first combination comprised both the abnormal and normal categories. The second combination of datasets encompassed a single normal category of cardiac sound recordings and four distinct categories of abnormal cardiac sound recordings, each of which included audio samples of a valvular ailment. In total, 957 spectrograms were obtained from this dataset.

This work was performed in five steps, as depicted in Figure 2, for both combinations of datasets.

(1) Acquire the heart sounds dataset
(2) Pre-process the audio dataset
(3) Extract the relevant features from the heart sounds
(4) Channel the features to an ML model and perform model training
(5) Performance evaluation was based on various ML metrics, including precision, recall, accuracy, specificity, and F1-score.

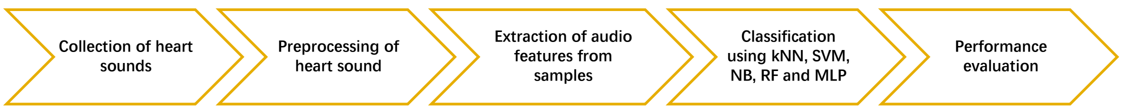

Figure 2. Steps in the methodology to classify the heart sound signals.

The features include MFCCs, energy, spectral rolloff, spectral bandwidth, spectral flux, spectral contrast, spectral energy, mel-spectrogram, FFT values, zero-crossing rate, degree of periodicity, spectral entropy, spectral centroid, short-time energy, spectral flatness, power mean value, normalised signal, temporal crest factor, DWT, peak frequency, median frequency, skewness, kurtosis, mean frequency, LPC_AVG, and high-frequency distortion.

6. Results and Discussion

6.1. Performance Metrics

A brief description is provided for the metrics used to evaluate the model's performance in Equations (22)–(26).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{20}$$

$$Precision = \frac{TP}{TP + FP} \tag{22}$$

$$Sensitivity = Recall = True\ Positive\ Rate = \frac{TP}{TP + FN} \tag{23}$$

$$Specificity = \frac{TN}{TN + FP} \tag{24}$$

$$F1 - score = \frac{2\ \times\ (Precision\ \times\ Recall)}{Precision + Recall} \tag{25}$$

where TP, FP, TN, and FN represent the number of true positives, false positives, true negatives, and false negatives,respectively.

Matthews Correlation Coefficient: The Matthews Correlation Coefficient (MCC) is a statistical metric used to evaluate the performance of the binary classification model. It considers true positives, false positives, false negatives, and true negatives from a confusion matrix to provide a single value that summarises the performance of the classifier [30]. The advantages of MCC over other metrics, such as accuracy, AUC, and F1 score, include its ability to handle imbalanced datasets and provide more informative results. It is a reliable measure for evaluating binary classifications in various scientific fields because it considers both types of errors, and is given by (27).

$$MCC = \frac{(TP * TN) - (FP * FN)}{\sqrt{(TP + FP) * (TP + FN) * (TN + FP) * (TN + FN)}} \tag{27}$$

PRC Area: The PRC area, also known as the Average Precision (AP), refers to a performance metric that measures the balance between precision and recall across all possible thresholds for classifying positive cases. The PRC area is calculated by plotting the Precision-Recall Curve (PRC), which shows precision on the y-axis and recall on the x-axis for different thresholds, and calculating the AUC of PRC. A higher PRC area indicates a better model, meaning that it consistently maintains good precision and recall across varying thresholds. This is useful for imbalanced data sets. It focuses on the quality of positive predictions, making it suitable for problems in which the identification of true positives is crucial.

Mean Absolute Error: Mean Absolute Error (MAE) is used for evaluating the regression model's *performance* in ML. It provides the average magnitude of the errors between the actual and predicted values, without considering the direction of the errors, and is given by (28).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y_i}) \tag{28}$$

where the number of observations is denoted by n, the actual value of the i-th observation is denoted by y_i and $\hat{y}_i$ is the value predicted for the i-th observation.

Root Mean Squared Error: Root Mean Squared Error (RMSE) measures the average magnitude of the errors between the predicted and actual values in the dataset. This is given by (29).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y_i})} \tag{29}$$

where n is the count of observations, y_i is the real value of the i-th observation and $\hat{y}_i$ is the predicted value of the i-th observation.

6.2. Discussion of Results

Overall performance: The study yielded outstanding results with both the 2-class and 5-class classification tasks. Random forests consistently outperformed the other ML algorithms across all metrics, demonstrating their effectiveness in discriminating between abnormal and normal heart sounds.

5-Class Classification: All ML models yielded high accuracy, with random forests again yielding 99.27%, as displayed in Table 3. There was a reduction in accuracy compared to the 2-class results, indicating the increased difficulty of multiclass classification. Precision and specificity were observed to be very high, with random forests maintaining high values for both the metrics. The F-measure and recall metrics slightly decreased for all models in comparison to the 2-class case, suggesting some difficulty in distinguishing between specific valvular diseases. Although random forests achieved near-perfect performance for the ROC Area, kNN and SVM dropped significantly, indicating potential challenges in separating certain disease classes. Random forests performed well against the PRC Area metric, with other models showing a wider range of performance, suggesting that different models might be better suited for specific disease classification tasks.

The *MAE* values for random forests (0.0296 for five classes and 0.0127 for two classes) were lower than those obtained by the other models, signifying that its predictions were better on average. The RMSE values for random forests (0.0816 for five classes and 0.0456 for two classes) were lower than those obtained by other models, indicating more accurate predictions with minimal errors in both cases.

Table 3. Performance metrics for 5 classes.

Metrics\Classifier	**Random Forests**	**kNN**	**SVM**	**Naive Bayes**	**MLP**
Correctly Classified Instances	950	922	844	693	928
Incorrectly Classified Instances	7	35	113	264	29
Mean absolute error	0.0296	0.0217	0.2469	0.1121	0.0169
Root mean squared error	0.0816	0.1069	0.3267	0.3097	0.1005
Specificity	0.002	0.009	0.03	0.069	0.008
Precision	0.993	0.964	0.883	0.722	0.97
Recall	0.993	0.963	0.882	0.724	0.97
F-Measure	0.993	0.963	0.879	0.712	0.97
MCC	0.993	0.955	0.852	0.653	0.962
ROC Area	1	0.996	0.944	0.922	0.991
PRC Area	0.999	0.987	0.825	0.827	0.982
Accuracy	**99.2685**	96.6573	88.1923	72.4138	96.9697

The *accuracy* metrics of all the classifiers for both combinations are shown in Figure 3. Tables 4–8 display the confusion matrix for the 5-class combination of the five ML techniques. Only seven instances were incorrectly classified, further proving the superiority of the random forests technique.

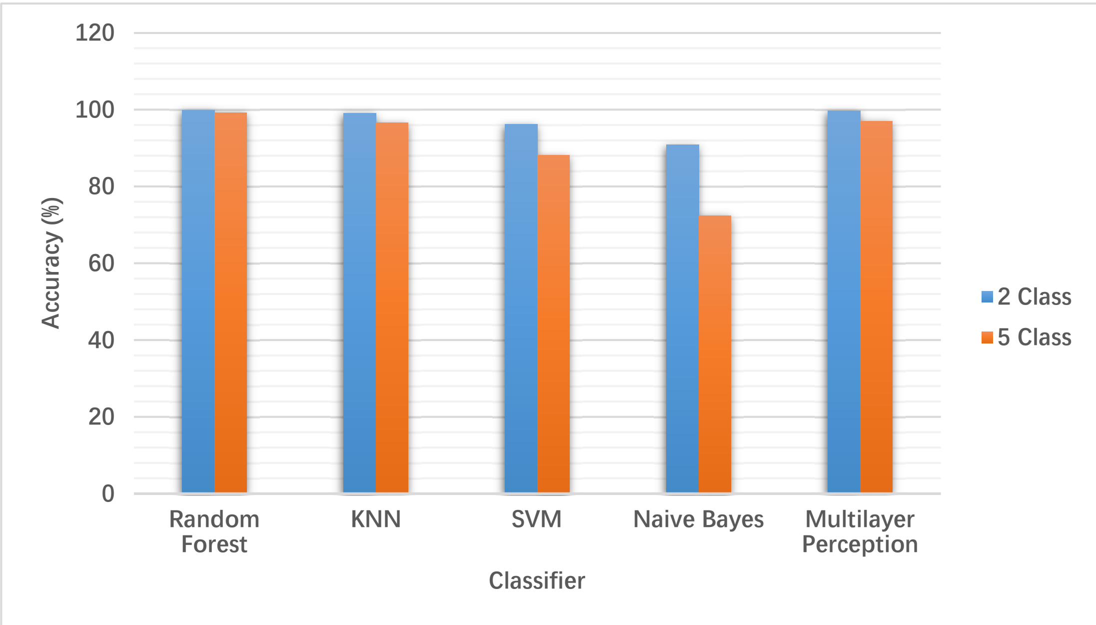

Figure 3. Accuracy of different classifiers.

The proposed models performed exceptionally which can be attributed to the fact that the audio features which give details about the patterns in the audio, changes in frequency and energy, and minor variations due to different types of murmurs were considered for the proposed model.

Table 4. Confusion matrix for 5 classes and Random forests.

AS	MR	MS	MVP	N
200	0	0	0	0
0	182	1	1	0
0	0	186	0	0
2	1	1	183	0
1	0	0	0	199

Table 5. Confusion matrix for 5 classes and kNN.

AS	MR	MS	MVP	N
195	3	2	0	0
0	179	4	0	1
0	2	177	3	4
2	3	7	173	2
0	0	1	1	198

Table 6. Confusion matrix for 5 classes and SVM.

AS	MR	MS	MVP	N
195	0	1	3	1
18	152	3	9	2
4	1	167	11	3
3	14	10	134	26
4	0	0	0	196

***Table** 7*. Confusion matrix for 5 classes and Naïve Bayes.

AS	MR	MS	MVP	N
81	22	67	30	0
24	143	7	10	0
7	0	165	14	0
4	10	13	116	44
4	0	1	7	188

Table 8. *Confusion* matrix for 5 classes and MLP (NN).

AS	MR	MS	MVP	N
198	0	1	0	1
5	174	2	3	0
0	2	181	3	0
3	5	4	175	0
0	0	0	0	200

2-Class Classification: All models attained an accuracy exceeding 90%, with random forests achieving a near-perfect 99.89% accuracy, as listed in Table 9. All models yielded high specificity and correctly identified normal sounds with only a few false positive rates. Random forests maintained high precision and recall for both normal and abnormal sounds, indicating accurate identification of both classes. Both the ROC Area and PRC Area values reached 1 for random forests, kNN, and MLP, indicating perfect discrimination between normal and abnormal sounds. SVM, naive Bayes, and MLP performed slightly lower, suggesting that the extracted features and task complexity might be better suited for ensemble methods such as random forests.

Table 9. Performance metrics for 2 classes.

Metrics\Classifier	Random Forests	kNN	SVM	Naive Bayes	MLP
Correctly Classified Instances	956	948	921	870	955
Incorrectly Classified Instances	1	9	36	87	2
Mean absolute error	0.0127	0.013	0.0376	0.0894	0.0052
Root mean squared error	0.0456	0.0805	0.194	0.2865	0.0395
Specificity	0.004	0.006	0.039	0.064	0.008
Recall	0.999	0.991	0.963	0.909	0.998
Precision	0.999	0.991	0.965	0.927	0.998
F-Measure	0.999	0.972	0.892	0.913	0.998
MCC	0.997	0.972	0.962	0.767	0.994
PRC Area	1	1	0.953	0.96	1
ROC Area	1	1	0.962	0.956	1
Accuracy	**99.8955**	99.0596	96.2382	90.9091	99.791

Tables 10–14 show the confusion matrix for the 2-class combination using the five ML techniques. Only one instance was incorrectly classified, indicating the superiority of the random forests technique. Except for the naïve Bayes model which performed reasonably well, the other models performed exceptionally well.

Table 10. Confusion Matrix for 2-classes and random forests.

Normal	Abnormal
199	1
0	757

Table 11. Confusion Matrix for 2-classes and kNN.

Normal	Abnormal
199	1
8	749

Table 12. Confusion Matrix for2-classes and SVM.

Normal	Abnormal
192	8
28	729

Table 13. Confusion Matrix for 2-classes and naïve Bayes.

Normal	Abnormal
189	11
76	681

Table 14. Confusion Matrix for 2-classes and MLP (NN).

Normal	Abnormal
198	2
0	757

Table 15 lists the outcome of this study and compares it against the results obtained by others.

Table 15. Comparison of results.

Reference	Author(s)	Methodology	Dataset Used	Accuracy (%)
5-class classification				
[16]	Yaseen et al.	MFCC + Discrete wavelet transform features combined with SVM, DNN and centroid displacement kNN	Yaseen dataset	97.9
[31]	Chowdhury et al.	DNN	Yaseen dataset	97.77
[32]	Khan et al.	CNN and power spectrogram	Yaseen dataset	98.87
[12]	Yadav et al.	Statistical features	Private dataset	97.78
[26]	Upretee et al.	Spectral centroid frequency with kNN and SVM	Yaseen dataset	96.50
[13]	Sathyanarayann et al.	ML techniques with HOG and LBP texture feature as input	Yaseen dataset	98.62
Proposed method		Audio features and ML	Yaseen dataset	99.26
Binary classification				
[26]	Upretee et al.	Spectral centroid frequency with kNN and SVM	Yaseen dataset	99.60
[33]	Taneja et al.	LBP + chromagram	PhysioNet 2016	94.87
[34]	Ibrahim et al.	Temporal, spectral and geometric audio features and cubic SVM	Private dataset	97
[35]	Milani et al.	Time-domain features and ANN	PhysioNet 2016	93.33
[36]	T. Li, &Yin et al.	Freguency domain features and 2D-CNN	PhysioNet 2016	86
[37]	Vinay et al.	XgBoost; No segmentation	PhysioNet 2016	92.85
[15]	F. Li &Zhang et al.	Improved MFCC and ResNet	PhysioNet 2016	94.43
[13]	Sathyanarayanan et al.	ML techniques with HOG and LBP textural features as input	Yaseen dataset	99.79
Proposed method		Audio features and ML	Yaseen dataset	99.89

7. Conclusions

This study investigated the possibility of building an ML model for the classification of heart sounds using a unique combination of audio features extracted from an audio dataset for training. The algorithm was not computationally intensive, and was built using a low-end system with minimal RAM. These results suggest the potential of the model as a valuable tool for accurate heart sound diagnosis in clinical settings. Further validation and refinement can significantly improve early disease detection and patient care. This study demonstrated the potential of using ML models for accurate heart sound classification.

The extensive set of extracted features likely contributed to the success of the model. Exploring the features that had the most significant impact on each model's performance could further optimise the classification process. The model can be further optimised by experimenting with a subset of these features and studying the impact of each feature on the model's performance.

This study indicates the potential of ML models, particularly random forests, for classifying normal and abnormal heart sounds with very high accuracy. In both the binary and five-class settings, the extracted audio features combined with the chosen algorithms achieved near-perfect discrimination in many instances. These results show a promising future for AI-powered heart sound diagnosis, paving the way for earlier and more precise detection of cardiovascular diseases.

The clinical implications of these findings are significant. Early detection of valvular diseases and other cardiac abnormalities is crucial to increasing the chances of patients being treated effectively. Ultimately, this could lead to timely intervention and optimised treatment. Cardiac diagnostic equipment integrated with AI, which is cheaper, quicker, reliable, and accurate, is set to revolutionise the first level of screening for heart diseases at the grassroots level.

Author Contributions

Conceptualisation—S.S. and S.M.K.; Methodology—S.S., C.G. and S.M.K.; Software—N.A. and C.G.; Formal analysis—S.S. and S.M.K.; Investigation—S.S. and C.G.; Resources—S.S., C.G. and S.K.M.; Data curation—S.S.; Writing-original draft preparation—S.S.; Writing-review and editing—S.S., S.M.K., C.G. and S.K.M.; Visualization—S.S. and C.G.; Supervision—S.M.K.; Project administration—S.S.; Funding acquisition—S.S. All authors have read and agreed to the published version of the manuscript.

Funding

The funding for this project was done by Sri Sathya Sai University for Human Excellence, Kalaburagi, Karnataka.

Conflict of Interest Statement

The authors have no financial or nonfinancial affiliations that may be interpreted as a source of conflict.

Data Availability Statement

No new dataset was created. The dataset created by Yaseen et al. was used for this work.

Acknowledgements

I thankful to Sri Sathya Sai University for Human Excellence, Gulbarga, Karnataka, India for supporting my research. I thank Shobha Sathyanarayanan for her assistance.

References

1. S. Sathyanarayanan and S. Chitnis, 'A Survey of Machine Learning in Healthcare', in *Artificial Intelligence Applications for Health Care*, I., M. K. Ahirwal, N. D. Londhe, and A. Kumar, Eds. Boca Raton: Taylor & Francis Ltd, 2022, pp. 1–22.
2. H. Shah, Y. Patil, and R. Patel, 'Prediction of Oral Cancer Treatment Plan using Machine Learning', *International Journal of Computer Information Systems and Industrial Management Applications*, vol. 15, no. 2023, pp. 408–417, 2023.
3. R. R. Chandrika, V. Anitha, N. R. Behera, P. V. Krishna, R. N. Jogekar, and K. Singh, 'Detection of Lung Cancer Using Optimal Hybrid Segmentation and Classification', *International Journal of Computer Information Systems and Industrial Management Applications*, vol. 15, no. 2023, pp. 360–370, 2023.
4. R. Maalej, A. Mezghani, M. Elleuch, I. M. Ben Ahmed, and M. Kherallah, 'Transfer Learning and Data Augmentation for Improved Breast Cancer Histopathological Images Classifier', *International Journal of Computer Information Systems and Industrial Management Applications*, vol. 15, no. 2023, pp. 268–277, 2023.
5. R. Rout, P. Parida, and S. Dash, 'Automatic Skin Lesion Segmentation using a Hybrid Deep Learning Network', *International Journal of Computer Information Systems and Industrial Management Applications*, vol. 15, no. 2023, pp. 238–249, 2023.
6. World Heart Federation, 'World Heart Report 2023: Confronting the World's Number One Killer', Geneva, 2023.
7. K. W. Giang *et al.*, 'Congenital heart disease: changes in recorded birth prevalence and cardiac interventions over the past half-century in Sweden', *European Journal of Preventive Cardiology*, vol. 30, no. 2, pp. 169–176, 2023, doi: 10.1093/eurjpc/zwac227.

8. Z. Su *et al.*, 'Global, regional, and national time trends in mortality for congenital heart disease, 1990–2019: An age-period-cohort analysis for the Global Burden of Disease 2019 study', *eClinicalMedicine*, vol. 43, no. January, p. 101249, 2022, doi: 10.1016/j.eclinm.2021.101249.
9. O. Hahad *et al.*, 'Cumulative social disadvantage and cardiovascular disease burden and mortality', *European Journal of Preventive Cardiology*, vol. 31, no. 1, pp. 40–48, 2023, doi: 10.1093/eurjpc/zwad264.
10. Y. Wang, K. Ng, and B. R. J., 'Early Detection of Heart Failure with Varying Prediction Windows by Structured and Unstructured Data in Electronic Health Records', 2017, doi: 10.1109/EMBC.2015.7318907.
11. Y. Muhammad, M. Tahir, M. Hayat, and K. T. Chong, 'Early and accurate detection and diagnosis of heart disease using intelligent computational model', *Scientific Reports*, vol. 10, no. 1, pp. 1–17, 2020, doi: 10.1038/s41598-020-76635-9.
12. A. Yadav, A. Singh, M. K. Dutta, and C. Travieso, 'Machine learning-based classification of cardiac diseases from PCG recorded heart sounds', *Neural Computing and Applications*, vol. 32, pp. 17843–17856, Oct. 2020, doi: 10.1007/s00521-019-04547-5.
13. S. Swaminathan, S. Murthy, S. K. Mallappa, and C. Gudada, 'Machine Learning Approach Using HOG and LBP Features of Spectrograms-based Heart Sounds Analysis for the Detection of Heart Diseases', in *15th Soft Computing and Pattern Recognition (SoCPaR'23)*, 2023, pp. 1–11.
14. H. Tang, M. Wang, Y. Hu, B. Guo, and T. Li, 'Automated Signal Quality Assessment for Heart Sound Signal by Novel Features and Evaluation in Open Public Datasets', *BioMed Research International*, vol. 2021, p. 7565398, 2021, doi: 10.1155/2021/7565398.
15. F. Li, Z. Zhang, L. Wang, and W. Liu, 'Heart sound classification based on improved mel-frequency spectral coefficients and deep residual learning', *Frontiers in Physiology*, vol. 13, no. December, pp. 1–16, 2022, doi: 10.3389/fphys.2022.1084420.
16. Yaseen, G. Y. Son, and S. Kwon, 'Classification of heart sound signal using multiple features', *Applied Sciences (Switzerland)*, vol. 8, no. 12, 2018, doi: 10.3390/app8122344.
17. G. Sharma, K. Umapathy, and S. Krishnan, 'Trends in audio signal feature extraction methods', *Applied Acoustics*, vol. 158, p. 107020, 2020, doi: 10.1016/j.apacoust.2019.107020.
18. N. M. K. M. S. K. G. M. K. Khan, 'Automated Heart Sound Classification from Unsegmented Phonocardiogram Signals Using Time Frequency Features', *International Scholarly and Scientific Research & Innovation*, vol. 12, no. 8, pp. 598–603, 2018.
19. G. Zhou, Y. Chen, and C. Chien, 'On the analysis of data augmentation methods for spectral imaged based heart sound classification using convolutional neural networks', *BMC medical informatics and decision making*, vol. 22, no. 1, p. 226, 2022, doi: 10.1186/s12911-022-01942-2.
20. K. Bai, Y. Zhou, Z. Cui, W. Bao, N. Zhang, and Y. Zhai, 'HOG-SVM-Based Image Feature Classification Method for Sound Recognition of Power Equipments', *Energies*, vol. 15, no. 12, p. 4449, 2022, doi: 10.3390/en15124449.
21. S. Sathyanarayanan, Srikanta Murthy K., and S. Chitnis, 'A Comprehensive Survey of Analysis of Heart Sounds using Machine Learning Techniques to Detect Heart Diseases', *The Journal of Population Therapeutics and Clinical Pharmacology*, vol. 30, no. May, pp. 375–384, 2023, doi: 10.47750/jptcp.2023.30.11.038.
22. S. Li, F. Li, S. Tang, and W. Xiong, 'A Review of Computer-Aided Heart Sound Detection Techniques', *BioMed Research International*, vol. 2020, 2020, doi: 10.1155/2020/5846191.
23. G. D. Clifford *et al.*, 'Classification of normal/abnormal heart sound recordings: The PhysioNet/Computing in Cardiology Challenge 2016', in *2016 Computing in Cardiology Conference (CinC)*, 2016, vol. 43, pp. 609–612, doi: 10.22489/cinc.2016.179-154.
24. S. V. Mahesh Kumar, P. Dhinakar, and R. Nishanth, 'Aortic Stenosis Detection Using Spectral Statistical Features of Heart Sound Signals', *EAI Endorsed Transactions on Pervasive Health and Technology*, vol. 8, no. 4, pp. 1–6, 2022, doi: 10.4108/EETPHT.V8I4.3168.
25. M. A. Soto-Murillo *et al.*, 'Automatic evaluation of heart condition according to the sounds emitted and implementing six classification methods', *Healthcare (Switzerland)*, vol. 9, no. 3, pp. 1–15, 2021, doi: 10.3390/healthcare9030317.
26. P. Upretee and M. E. Yüksel, 'Accurate classification of heart sounds for disease diagnosis by a single time-varying spectral feature: Preliminary results', *2019 Scientific Meeting on Electrical-Electronics and Biomedical Engineering and Computer Science, EBBT 2019*, pp. 1–4, 2019, doi: 10.1109/EBBT.2019.8741730.
27. Y. Zeinali and S. T. A. Niaki, 'Heart sound classification using signal processing and machine learning algorithms', *Machine Learning with Applications*, vol. 7, p. 100206, Mar. 2022, doi: 10.1016/j.mlwa.2021.100206.
28. X. Bao, Y. Xu, and E. N. Kamavuako, 'The Effect of Signal Duration on the Classification of Heart Sounds : A Deep Learning Approach', *sensors*, 2022, doi: 10.3390/s22062261.
29. S. B. Shuvo, S. N. Ali, S. I. Swapnil, M. S. Al-Rakhami, and A. Gumaei, 'CardioXNet: A Novel Lightweight Deep Learning Framework for Cardiovascular Disease Classification Using Heart Sound Recordings', *IEEE Access*, vol. 9, pp. 36955–36967, 2021, doi: 10.1109/ACCESS.2021.3063129.
30. M. Aslan and S. Sen, 'A dynamic trust management model for vehicular ad hoc networks', *Vehicular Communications*, vol. 41, no. April, 2023, doi: 10.1016/j.vehcom.2023.100608.

31. S. Chowdhury, M. Morshed, and S. A. Fattah, 'SpectroCardioNet: An Attention-Based Deep Learning Network Using Triple-Spectrograms of PCG Signal for Heart Valve Disease Detection', *IEEE Sensors Journal*, vol. 22, no. 23, pp. 22799–22807, 2022, doi: 10.1109/JSEN.2022.3196263.
32. J. S. Khan, M. Kaushik, A. Chaurasia, M. K. Dutta, and R. Burget, 'Computer Methods and Programs in Biomedicine Cardi-Net : A deep neural network for classification of cardiac disease using phonocardiogram signal', *Computer Methods and Programs in Biomedicine*, vol. 219, p. 106727, 2022, doi: 10.1016/j.cmpb.2022.106727.
33. K. Taneja, V. Arora, and K. Verma, 'Classifying the heart sound signals using textural-based features for an efficient decision support system', *Expert Systems*, vol. 40, no. 6, p. e13246, Feb. 2023, doi: 10.1111/exsy.13246.
34. N. Ibrahim, N. Jamal, M. N. A. H. Sha'abani, and L. F. Mahadi, 'A Comparative Study of Heart Sound Signal Classification Based on Temporal, Spectral and Geometric Features', in *Proceedings - 2020 IEEE EMBS Conference on Biomedical Engineering and Sciences, IECBES 2020*, 2021, pp. 24–29, doi: 10.1109/IECBES48179.2021.9398810.
35. M. G. M. Milani, P. E. Abas, L. C. De Silva, and N. D. Nanayakkara, 'Abnormal heart sound classification using phonocardiography signals', *Smart Health*, vol. 21, no. February, p. 100194, 2021, doi: 10.1016/j.smhl.2021.100194.
36. T. Li, Y. Yin, K. Ma, S. Zhang, and M. Liu, 'Lightweight end-to-end neural network model for automatic heart sound classification', *Information (Switzerland)*, vol. 12, no. 2, pp. 1–11, 2021, doi: 10.3390/info12020054.
37. V. Arora, R. Leekha, R. Singh, and I. Chana, 'Heart sound classification using machine learning and phonocardiogram', *Modern Physics Letters B*, vol. 33, no. 26, pp. 1–24, 2019, doi: 10.1142/S0217984919503214.