Received: 3 October 2023 | Revised: 20 November 2023 | Accepted: 21 December 2023

DOI: 10.1111/exsy.13535

REVIEW

Expert Systems WILEY

Artificial intelligence for heart sound classification: A review

Junxin Chen[1] | Zhihuan Guo[2] | Xu Xu[3] | Gwanggil Jeon[4] | David Camacho[5]

[1]School of Software, Dalian University of Technology, Dalian, China

[2]College of Medicine and Biological Information Engineering, Northeastern University, Shenyang, China

[3]School of Computer Science and Engineering, Northeastern University, Shenyang, China

[4]College of Information Technology, Incheon National University, Incheon, Korea

[5]Department of Computer Systems Engineering, Universidad Politécnica de Madrid, Madrid, Spain

Correspondence
Gwanggil Jeon, College of Information Technology, Incheon National University, Incheon, 22012, Korea.
Email: gjeon@inu.ac.kr

Funding information
National Natural Science Foundation of China, Grant/Award Number: 62171114; Fundamental Research Funds for the Central Universities, Grant/Award Number: DUT22RC(3)099; MCIN/AEI/CHISTERA IV Cofund 2021 program, Grant/Award Number: MCIN/AEI/10.13039/501100011033-PCI2022-134990-2(MARTINI); MCIN/Spanish Ministry of Science and Innovation, Grant/Award Number: PID2020-117263GB-I00(FIGHTDIS)

Abstract

Heart sound signal analysis is very important for the early identification and treatment of cardiovascular illness. With rapid advancements in science and technology, artificial intelligence technologies are providing tremendous opportunities to enhance diagnosis and clinical decision-making. Instruments can now perform clinical diagnoses that previously could only be handled by human experts more conveniently and efficiently. Despite multiple works on automatic heart sound analysis, there are few summarization and review works. This article attempts to give a thorough overview of various heart sound analysis subtasks and examine the improvements made in each subtask by both machine learning techniques and deep learning algorithms. It goals to highlight the potential of AI to revolutionize cardiovascular healthcare by enabling accurate and automated analysis of heart sounds. The findings of this review are beneficial for researchers, clinicians, and engineers in the development and application of AI-based solutions for improved heart sound classification and diagnosis.

KEYWORDS
cardiovascular diseases, deep learning, heart sound analysis, machine learning

1 | INTRODUCTION

Cardiovascular diseases (CVDs) are seriously threatening the safety of human life (Chakrabarti et al., 2015; Chen, Sun, et al., 2022), which are demonstrated to affect 48% of the population (≥ 20 years old). Cardiac auscultation with a stethoscope has a long history in the diagnosis and treatment of CVDs, and is also popular nowadays. It is based on the features extracted from the acquired heart sounds. Heart sound is a signal that carries physiological information coming from the opening and closing of heart valves and the turbulent blood flow in arteries (Reyna et al., 2022), its graphical representation recorded by electronic instruments is denoted as phonocardiogram (PCG). A normal PCG cycle typically refers to the first heart sounds (S1) produced at the start of systole and the second heart sounds (S2) produced at the start of diastole and can be described accordingly (Chizner, 2008). Manual auscultation remains the most commonly used method of cardiac examination, but its diagnostic sensitivity and accuracy are limited because the accuracy of this technique depends on extensive clinical experience. Automatic computer recognition of cardiac sound pathology information can overcome the subjectivity of manual auscultation and is still an important area of exploration.

In 1956, John MacCarthy first proposed artificial intelligence at the Dartmouth College Conference (Long et al., 2020). Since then, artificial intelligence (AI) technology has entered the public's field of vision and become popular. The computer science discipline of AI aims to create machines capable of carrying out tasks that traditionally require human intelligence (Chen, Wang, et al., 2023). Echocardiography is an effective

David Camacho is the co-author of this manuscript.

diagnostic tool for detecting diseases at an early stage. In a recent collaboration, cardiologists and radiologists worked with AI scientists to create a completely computerized echocardiographic assessment application for use in medical care (Zhang et al., 2018). The above system demonstrates that AI has great potential to transform the current clinical practice model and popularize high-quality medical care. The 2022 George B. Moody PhysioNet Challenge, previously known as the 2016 PhysioNet Challenge, offers a chance to explore the use of AI in PCG diagnosis. Teams are asked to create entirely automated methods for finding anomalies in children's PCG records as part of the competition (Reyna et al., 2022). This challenge demonstrates the limitless potential of AI in the field of automated diagnosis of PCG recordings.

Classical machine learning (ML) methods, such as hidden Markov models (HMMs) (Rabiner, 1989) or classifiers like support vector machines (SVM) and decision trees (Whitaker et al., 2017), have been applied to heart sound signal classification with promising results. However, these methods often require manual feature extraction and their performance on large datasets has been limited. In contrast, deep learning (DL) has lately become a superior approach for examining heartbeats. A time-frequency representation of the 1D cardiac sound signal or the raw audio stream are frequently used as inputs for the DL models (Chowdhury et al., 2020; Xiao et al., 2020). Their strong learning ability enables them to perform well on large-scale datasets.

The approaches and models covered in the selected papers are described using ML and DL. The ML stands for traditional machine learning algorithms. The DL is a subclass of ML that involves convolutional neural learning networks. In the past, several review studies on heart sound analysis have been carried out (Brites et al., 2021; Chen, Sun, Chen, et al., 2021; Dwivedi et al., 2018; Ghosh, Nagarajan, & Tripathy, 2020; Kumar et al., 2022; Nabih-Ali et al., 2017), and the latest review study found in Google Scholar is Ren et al. (2023). Dwivedi et al. (2018) presented a systematic review and critical analysis of existing approaches for automated identification and classification of heart sounds. Kumar et al. (2022) introduced a survey on the use of AI techniques, including ML and DL, in diagnosing various diseases such as Alzheimer's, cancer, diabetes, and chronic heart disease. Nabih-Ali et al. (2017) reviewed recent advancements in pre-processing, feature extraction, and classification techniques for PCG signal analysis in intelligent computer-aided diagnosis systems. Ghosh, Nagarajan, and Tripathy (2020) provided a review of state-of-the-art approaches for heart sound acquisition and pre-processing techniques, which were crucial for diagnosing heart valve-related disorders and other applications. Brites et al. (2021) presented a systematic mapping study that reviews the literature on the application of ML and IoT in disease identification through heart sounds. Some of these studies have not followed the latest AI techniques for heart sound analysis, and others have not focused on the area of heart sound.

Different from the above studies, this article fully introduces the summary and discussion of the methods involved in each heart sound signal analysis subtask. We will review recent ML and DL techniques favoured by researchers for heart sound analysis. For a more comprehensive search of papers, we first search the papers based on Google Scholar, which indexes a wide range of academic sources and offers advanced search features that enable users to refine their search queries. Second, we consider using a combination of keywords and Boolean operators to create a search query. Next, adjusting keywords from the initial results to retrieve relevant papers. Finally, these papers are adopted or excluded based on three factors, that is, publication date, citations, and journal level. Furthermore, this review summarizes traditional ML and DL approaches in denoising, segmentation, obtaining characteristics, and identifying cardiac sound signals. Compared to Ren et al. (2023), we have added a discussion of DL denoising methods. We will also discuss issues requiring staff attention and future research directions in this survey, helping to advance research in heart sound analysis.

The overall structure of this article flows as follow. Section 2 provides a background introduction and theoretical basis of heart sound analysis. Section 3 presents the heart sound database and algorithmic evaluation metrics. Section 4 reviews the ML methods that require manual feature extraction. Section 5 discusses the application of DL algorithms to various subtasks of heart sound analysis. Sections 6 and 7, respectively, give the study's discussion and conclusion.

2 | RELATED WORK

2.1 | Generation of heart sounds

The heart's primary role is to circulate oxygen throughout the body and use oxygenated blood to provide energy for various physiological activities. Blood circulation is controlled by heart valves (Potdar et al., 2022). The entire heart structure vibrates as the heart valves open and close (Ahlstrom et al., 2006). Heart sounds, which are heard at the chest wall and can reveal details about the condition of the heart, are these vibrations. The rhythmic contractions and expansions of the heart generate pressure that controls the process of opening and closing the heart valves, commonly referred to as systole and diastole, respectively. Typically, systole lasts about 0.35 s and diastole lasts about 0.45 s, resulting in a total of approximately 0.8 s of cardiac cycle time (Li, Li, et al., 2020).

2.2 | Heart sound auscultation

To non-invasively monitor and diagnose cardiovascular disease, doctors frequently apply a device called a stethoscope to listen to the heart rhythms. Heart sounds can be divided into four components based on their order in the cardiac cycle, the first sound of the heart (S1), the 2nd heart sound (S2), the 3rd sound of the heart (S3), and the 4th heart sound (S4). Figure 1, which depicts the cardiac cycle encompassing systole and diastole, demonstrates that S1 and S2 are the primary heart sounds. The generations of S1 and S2 are illustrated in Figure 2. For the S1, it is generated by the closeness of the tricuspid and mitral valves. Furthermore, the S2 is generated by the closeness of the aortic and mitral valves. Some young, healthy individuals have S3 in their cardiac cycles, although S4 does not happen in a typical cardiac cycle (Li, Li, et al., 2020). The presence of S3 may indicate heart failure (Ren et al., 2023), and the frequencies of S1 and S2 range from 20 to 200 Hz, while the frequencies of both S3 and S4 range from 15 to 65 Hz (Naseri & Homaeinezhad, 2013). Murmurs, which are additional sounds given by disturbed blood circulation, are significant in the diagnosis of cardiovascular disease. A heart structure condition, particularly valvular heart disease, which can prevent valves from fully opening and closing, is the most frequent reason for a murmur. Clinically, there are two types of murmurs: systolic and diastolic. Aortic stenosis, mitral regurgitation, and tricuspid regurgitation produce systolic murmurs, and mitral and tricuspid stenosis produce diastolic murmurs (Noor & Shadi, 2013), that is, PCG. Heart sounds refer to the audible vibrations produced by the heart. In addition, PCG is a diagnostic tool that records and analyzes these heart sounds along with murmurs, extra sounds, and other cardiac abnormalities. The PCG provides a visual representation of the heart sounds, aiding in the interpretation and diagnosis of various cardiac conditions.

The diagnostic results of cardiac auscultation may determine whether further, more expensive tests should be done. Since the auscultation of heartbeats is highly subjective, it relies on the physician's skill, experience, and expertise to make an accurate diagnosis. However, the overall sensitivity of cardiac auscultation is poor (i.e., between 0.21 and 1.00) due to the interference of various uncertainties (Alam et al., 2010). In order to avoid the above-mentioned problems of cardiac auscultation, it has become popular to record heart sounds with electronic instruments.

Phonocardiography, developed in the 1930s and 1940s, is a valuable tool for studying murmurs, particularly in mitral valve disease and congenital heart disease cases (Sprague, 1957). In recent decades, the rapid development of digital technology has made computer-aided PCG signal auscultation play an essential role in diagnosing and treating diseases (Quiceno-Manrique et al., 2010). The electronic equipment is used to record heart sounds, which are amplified to overcome the low sound level of traditional stethoscopes (Nowak & Nowak, 2018). This gives the electronic stethoscope an advantage over conventional stethoscopes regarding heart sound collection, processing, and analysis. The PCG signal captured with a digital device is typically uploaded to a computer for viewing and further research. Recently, researchers have developed new portable electronic stethoscopes, which can be connected to mobile devices or Bluetooth for remote processing, such as real-time remote monitoring and diagnosis (Hadiyoso et al., 2020; Shin et al., 2013; Xu, Li, Fan, et al., 2023). The COVID-19 epidemic has accelerated the development of a non-

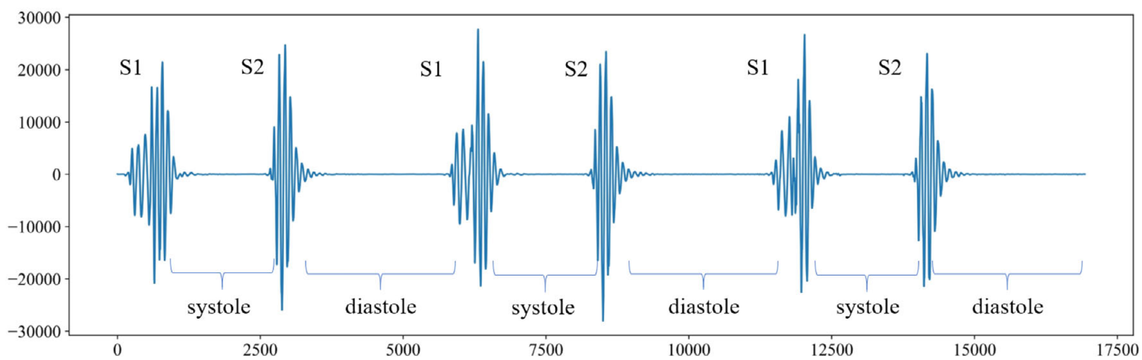

FIGURE 1 Diagram of the heart sound cycle.

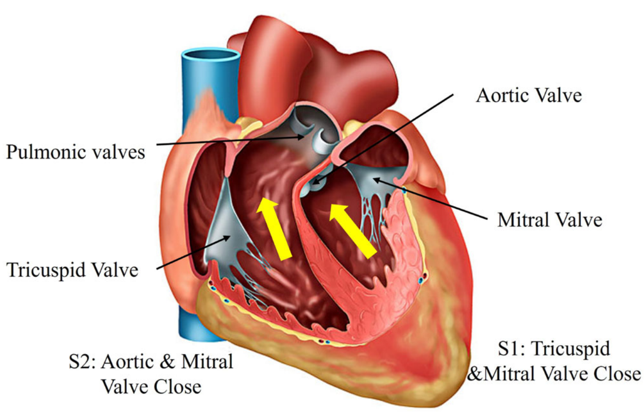

FIGURE 2 Generation of the heart sounds.

contact electronic stethoscope, and in 2021, a non-contact electronic stethoscope was designed for the auscultation of patients with COVID-19 (Yang et al., 2021).

2.3 | Diagnosis of cardiovascular diseases

There are several diagnostic methods for cardiovascular disease. A typical screening method is an electrocardiography (ECG), which uses P-QRS-T waves to show the electrical impulses of the heart (Martis et al., 2014). However, it is difficult to detect structural abnormalities in heart valves (Molcer et al., 2010). Additionally, several medical imaging tools capable of visualizing the cardiovascular system have become increasingly popular. Cardiovascular computed tomography, also called CT, and cardiac magnetic resonance imaging (CMRI) can offer additional details on the heart's condition and provide information about the size, form, building, and function of the cardiovascular system. However, due to the high cost of the inspections, as mentioned earlier, and the need for professional medical personnel to operate, the conditions mentioned above are unavailable in remote and poor areas, clinics, and small and medium-sized hospitals.

3 | DATABASES AND PERFORMANCE METRICS

In this section, we briefly introduce the currently existing heart sound datasets. Provide some references for future related research. The datasets commonly mentioned in previous studies is summarized and listed in Table 1.

3.1 | Open databases

The lack of large-scale and high-quality datasets will affect the development of heart sound analysis algorithms. In past studies, several databases were mentioned. Table 1 briefly introduce the following accessible database.

Currently, the most famous heart sound database is PhysioNet[1], which consists of datasets recorded from multiple sets of different devices from all over the world. The public training set contains 3240 records (totalling 20.216 h) of subjects from 764 individuals. The details of PhysioNet are shown in Figure 3. Severe data imbalance in this database is visualized.

The open-source heart sound database[2] released by Spadaccini and Beritelli (2013). It aims to develop systems and methods for biometric identification using heart sounds as physiological features. This data set contains heart-sound information of 157 men and 49 women, a total of 206 people. The researchers used the ThinkLabs Rhythm Digital Electronic Stethoscope to collect two recordings of each person on the same day. The signal sampling rate was 11,025 Hz and stored in WAVE format.

With durations ranging from 1 to 30 s, the PASCAL database[3] contains 656 recordings for heart sound categorizing and 176 recordings for the heart sound segmentation process. However, due to the application of a low-pass filter during recording, the frequency band of a heartbeat signal is limited to below 195 Hz, which may remove some valuable components for clinical diagnosis.

GitHub also provides an open-source heart sound database, collecting 1000 audio files (0.679 h in total).[4] Recordings of five different types of heart sounds are included in this dataset: normal, aortic stenosis, mitral regurgitation, mitral stenosis, and mitral valve prolapse. Each category has 200 pieces of data.

TABLE 1 Common heart sound datasets.

Database	Sampling frequency	Subjects	Samples
PhysioNet/CinC challenge (Liu et al., 2016)	2000 Hz	764	3240
HSCT-11 (Spadaccini & Beritelli, 2013)	11,025 Hz	206	412
PASCAL (Gomes et al., 2013)	4000 Hz	Unknown	176/656
Data on Github (Son & Kwon, 2018)	8000 Hz	Unknown	1000
Michigan	44.1 KHz	Unknown	23
DigiScope (Oliveira et al., 2018)	4000 Hz	29	29
CirCor DigiScope (Oliveira et al., 2021)	Unknown	1568	5282
Heart Sounds Shenzhen corpus	4000 Hz	170	845
Fetal (Cesarelli et al., 2012)	333 Hz	26	26
Simultaneous electrocardiogram and phonocardiogram (Kazemnejad et al., 2021)	8000 Hz	24	69

The Open Michigan Heart Sound and Murmur Library[5] records heart sound data from the apical region when the subject is supine, the apical region when lying on the left side, the aortic region when sitting, and the lung region when lying supine. A total of 23 records were provided with a total duration of 0.413 h.

The literature (Oliveira et al., 2018) publishes a paediatric dataset containing samples from 29 individuals of various health levels. Two cardiopulmonologists noted each heart sound event's start and termination. This dataset's heart sounds were captured using a stethoscope equipped with DigiScope Collector technology, the Littmann 3200.

The CirCor DigiScope database (Oliveira et al., 2021) was collected from a paediatric population aged 21 years or younger for the George B. Moody PhysioNet Challenge 2022. A total of 5272 PCG records at one or more locations (pulmonary, aortic, mitral, tricuspid) were collected from 1568 young patients in rural areas of Brazil using electronic auscultation equipment (Littmann 3200 stethoscope) (Reyna et al., 2022). This database's publicly available training set consists of 3163 audio samples from 942 participants, totalling 20.094 h.

The Shenzhen University General Hospital gathered the Heart Sounds Shenzhen corpus, which was initially made public during the INTERSPEECH 2018 COMPARE Heart Sound Sub-challenge. It contained 845 records from 170 individuals (male: 55, female: 115), each of which each audio sample is annotated into one of three categories: normal, mild, moderate/severe.

Cesarelli et al. (2012) recorded fetal heartbeats from 35 pregnant women in the final stages of a single fetal physiological pregnancy using a collector. The signals were digitized with an 8-bit ADC and a sampling frequency of 333 Hz.

In a recent study (Kazemnejad et al., 2021), a dataset of 69 simultaneous electrocardiograms (ECG) and PCG recordings was provided, including 61 recordings taken in an indoor sports centre and 24 recordings obtained from healthy adults aged between 23 and 29 years (mean: 25.4 $\pm$ 1.9 years). The dataset[6] also contains 8 sample recordings of the 30 s obtained at rest.

In addition to the above, there are many other heart sound data sets, such as the heart murmurs database (eGeneralMedical)[7], the heart sound library of Thinklabs,[8] Texas,[9] Biosciences Database of heart Sound signals.[10] Furthermore, D. Mason's book (Mason, 2000) includes some heart sound recordings. Additionally, some review papers have utilized heart sound signals from the audiovisual presentation by Tavel et al. (1973) as a database.

Above mentioned datasets play a crucial role in the field of heart sound pattern recognition. They can provide researchers, developers, and the wider community with valuable resources for studying and advancing the understanding of cardiac health. First, researchers can use these open databases to study heart sound patterns and develop algorithms for automatic heart sound analysis. This includes tasks such as heartbeat segmentation, extraction of acoustic features, and classification of abnormal heart sounds. Furthermore, these databases enable the development and evaluation of diagnostic models for various cardiac conditions. By analysing heart sound recordings from patients with different cardiac pathologies, researchers can develop algorithms to aid in the diagnosis of diseases such as valvular disorders, heart murmurs, and heart failure.

3.2 | Performance metrics

Before discussing evaluation measures for classification tasks, it is essential to describe the idea of a confusion matrix. A confusion matrix summarizes the predictions made by a classification model. It is a valuable tool for summarizing the proportion of accurate and unreliable predictions made for every group and presents this information in an intuitive and concise numerical format. The confusion matrix shows which data the classification model predicts is ambiguous, allowing us to understand the types of errors the classification model makes. This detailed breakdown of the results overcomes the limitations of using accuracy alone. All evaluation indicators are calculated based on the confusion matrix.

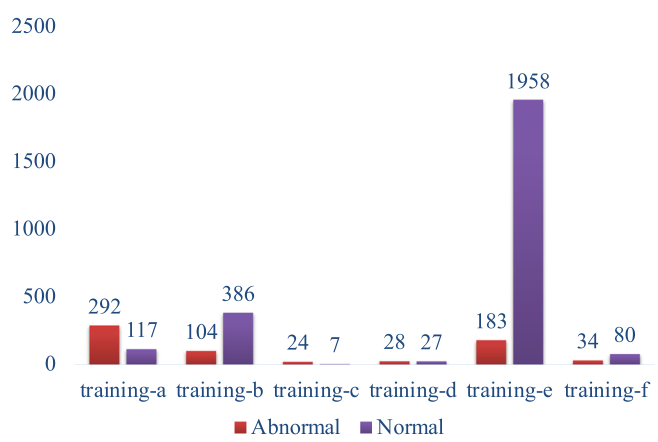

FIGURE 3 Detail of PhysioNet/CinC challenge.

In binary classification tasks, the more common evaluation indicators are sensitivity, specificity and accuracy (Ren et al., 2023). Precision and recall are frequently used in multi-class classification problems to assess how well a model performs for each class. Recall is the percentage of real positives among all real positives, whereas precision measures the percentage of real positives among all accurate positive predictions. Usually, the trend of recall and precision are opposite. From this, the F1-score is derived, which is the harmonic mean of the precision and recall. Both metrics are useful for evaluating the performance of a model on a specific class. On the other hand, accuracy quantifies the percentage of correctly classified samples to the total samples, regardless of class. However, accuracy does not perform well on unbalanced sample sets, more detailed metrics are needed to objectively describe the quality of classification models. Sensitivity refers to the proportion of correctly classified positive samples; similarly, specificity refers to the proportion of negative samples that are correctly classified. The connection between sensitivity and specificity for a classification model is graphically represented by the ROC curve (receiver operating characteristic curve). It offers a thorough index to assess a model's effectiveness concerning continuous variables. The area under the curve (AUC) is a standard metric used to quantify the overall diagnostic accuracy of a model, with a larger AUC indicating higher accuracy. Table 2 below summarizes the evaluation indicators of some common heart sound classification models, where the TP stands for the percentage of the PCG signals that are correctly identified as normal, the TN for the percentage of abnormal PCG signals that are correctly detected, the FP for the percentage of normal PCG signals that are incorrectly identified as abnormal, and the FN for the percentage of abnormal PCG signals that are misclassified as normal. These indicators are commonly used to evaluate the performance of model by measuring the difference between the predicted results and the ground truth.

4 | MACHINE LEARNING FOR HEART SOUND CLASSIFICATION

Early numerical simulations are used in a various fields such as image encryption (Chen, Chen, & Zhou, 2020), and tunnel construction (He et al., 2023; He et al., 2024). With the advent of the ML, the computation of various tasks is further accelerated. This section summarizes the classical ML techniques in the automatic sound heart classification task. Noteworthy is the fact that the machine-learning techniques referred to in this chapter are limited to studies that require manual feature extraction. Figure 4 summarizes the automatic heart sound classification process.

Specifically, data pre-processing methods such as denoising and segmentation are essential components of many ML heart sound classification algorithms. The effectiveness of the ML algorithms relies heavily on data pre-processing. This section summarizes the common types of manual feature extraction and data pre-processing methods in different studies. Generally speaking, there are three types of ML: supervised, unsupervised, and reinforcement learning. Supervised learning algorithms learn the mapping between inputs and labels to predict the output based on information. Common supervised learning approaches in heart sound analysis include artificial neural networks (ANN), SVM, K-nearest neighbours (KNN), Naive Bayesian models, HMMs, and decision tree models. Unsupervised learning automatically identifies and classifies the potential classification rules of unlabelled data through algorithms. Unsupervised learning can be used to discover patterns and relationships in data without needing labelled examples. For example, unsupervised learning has been used to identify three distinct cardiac phenotypes in type 2 diabetic individuals (Vickers, 2017). Principal component analysis, K-means clustering, and hierarchical clustering are common unsupervised techniques used for this purpose (Xu, Li, Lan, et al., 2023). Reinforcement learning is the strategy in which a machine learns by interacting with outcomes to maximize rewards (Long et al., 2020). Different from supervised or unsupervised learning, its learning process is dynamic. Clinical trials have employed Q-learning as an illustration of reinforcement learning to find the most effective, customized lung cancer treatment (Zhao et al., 2011).

TABLE 2 Metrics for evaluating model results.

Metric	Definition	Formula
Accuracy	Proportion of correctly classified samples to the total samples	$\frac{TP+TN}{TP+TN+FP+FN}$
Sensitivity	Proportion of correctly classified positive samples (TPR)	$\frac{TP}{TP+FN}$
Specificity	Proportion of negative samples that are correctly classified (TNR)	$\frac{TN}{TN+FP}$
FPR	Proportion of individuals who tested positive in a sample that were actually negative	$\frac{FP}{FP+TN}$
FNR	Proportion of samples that tested negative for the true label as positive	$\frac{FN}{TP+FN}$
Precision	Proportion of true positives among all samples predicted to be positive	$\frac{TP}{TP+FP}$
Recall	Proportion of true negatives among all samples predicted negative	$\frac{TP}{TP+FN}$
F1-score	Harmonic mean of the precision and recall	$\frac{2\times Sen\times Pre}{Sen+Pre}$
UAR	Averaged recall for each class of heart sounds	$\frac{\sum_{i=1}^{n} Recall_i}{n}$
ROC	Plotting the area under the probability curve for TPR and FPR at different probability	-

4.1 | Pre-processing

4.1.1 | Denoising

It is actually difficult for us to obtain pure heart sound signals. The main reason is that we will be affected by the following noise sources during the acquisition process, such as lung sounds caused by breathing, background sounds, intestinal activities, and stethoscope movements, sensor changes (Liu et al., 2016). The influence of noise is a problem that cannot be ignored for computer-assisted heart sound analysis task. This section discusses several common heart sound denoising methods for future reference.

Signal filters are a common method of removing noise. Band-pass filters can simultaneously remove high-frequency and low-frequency noise and retain signal data within their cut-off frequency range, often used for denoising heart sounds (Gao et al., 2023; Ghosh et al., 2019; Yadav et al., 2018). In several heart sound categorization investigations, Butterworth band-pass filters with various orders and cut-off frequencies have been employed (Ahmad et al., 2021; Noman et al., 2019; Singh et al., 2020). For instance, a 4th-order Butterworth filter with a 25–400 Hz cut-off frequency was applied in Singh et al. (2019), while a 5th-order Butterworth filter with a 25–500 Hz cut-off frequency was used in Li et al. (2021).

In addition, common heart sound denoising methods also include tunable-Q wavelet transform, adaptive noise canceller, discrete wavelet transform (DWT), variational mode decomposition, and combination of singular value decomposition (SVD) and compressed sensing, non-negative matrix factorization (NMF) with adaptive contour representation computation (ACRC) from the corresponding spectrogram. Table 3 lists common research on heart sound denoise algorithms in recent years.

4.1.2 | Segmentation

Segmentation usually refers to the identification and segmentation of the boundaries between heart sounds and murmurs. Heart sound segmentation is a crucial stage in heart sound analysis, according to several research. Numerous algorithms have been proposed and are often used as a foundation for developing more advanced segmentation techniques. Methods for segmenting heart sounds that are often utilized include Hilbert envelopes, Shannon Energy Envelograms, wavelet transforms, and fractal decomposition. Additionally, several segmentation techniques have been developed by combining these methods. For example, the combination of peak spacing patterns with Hilbert envelopes was proposed in Milani et al. (2019), Teague Energy Operator and wavelet transform were proposed together in Ceyhan et al. (2017). Other methods include hidden semi-Markov model (HSMM) based CNN methods (HSMM-CNN) and GMM (Gaussian mixture model) based HSMM (HSMM-GMM). This section will focus on four popular segmentation algorithms: the wavelet transform, the fractal decomposition, the Hilbert envelope, and the Shannon energy envelope.

The study (Ren et al., 2023) states that to obtain better classification performance for a classifier that is not very powerful, segmentation may need to be used as an additional supplementary or as an internal process of a robust classifier. Experiments in Dissanayake et al. (2020) show that segmentation does not significantly improve the model's performance. The reason may be that the model is already compelling and robust, and the middle layer automatically does the segmentation. The author Dissanayake et al. (2020) proved the above point through the SHapley Additive Explanation (SHAP) model (Lundberg & Lee, 2017).

The literature (Renna et al., 2019) summarizes three categories of solutions for performing PCG segmentation: envelope-based segmentation methods, feature extraction and classifier-based methods, and statistical model-based methods. Table 4 lists common research on heart sound segmentation algorithms in recent years.

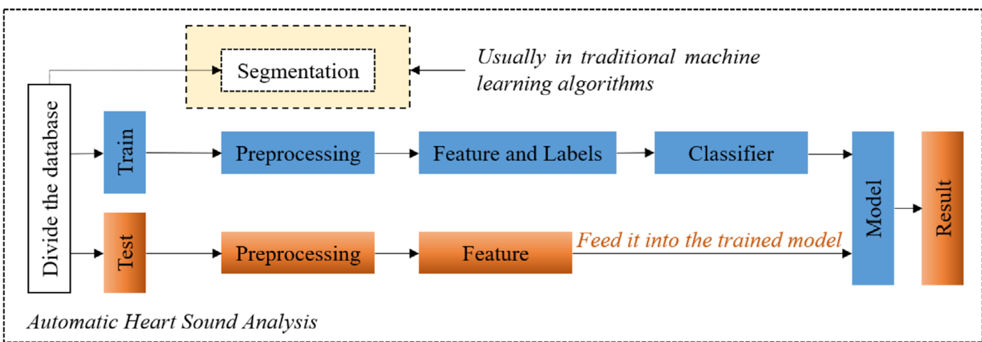

FIGURE 4 Block of algorithm.

TABLE 3 Denoising in machine learning heart sound classification algorithm.

Method	Reference	Description
A bandpass filter	Singh et al. (2019)	A 4th-order Butterworth filter with a 25–400 Hz cutoff frequency.
Adaptive noise canceller	Potdar et al. (2015)	It needs to include the main input of the damaged signal and the reference input containing noise.
Discrete wavelet transform (DWT)	Ghosh, Tripathy, and Ponnalagu (2020)	Based on its ability to compact energy and localize signals in both temporal and frequency domains, it can provide better performance in noisy environments.
Tunable-Q wavelet transform (TQWT)	Zeng et al. (2021)	The discrete wavelet transform known as TQWT is effective at analysing oscillating signals.
Empirical mode decomposition (EMD)	Gavrovska et al. (2013)	EMD, suited for non-stationary signal analysis, is based on empirical basis functions.
Variational mode decomposition	Lahmiri and Boukadoum (2014)	Due to the ability of VMD to capture relevant centre frequencies, it helps to achieve frequency band separation better and identify various discontinuities in non-stationary signals.
Combination of compressed sensing and singular value decomposition (SVD)	Zheng et al. (2017)	SVD is a multi-level decomposition and reconstruction of the original heartbeats, which separates the fundamental heart sounds from the signal mixed with noise as much as possible. Compressed sensing based sparse reconstruction is then performed for further denoising.
Non-negative matrix factorization with adaptive contour representation computation	Pham et al. (2018)	After the signal has been subjected to NMF to remove high-energy noise, ACRC extracts the feature. Then, these features are used for signal reconstruction.

4.2 | Feature engineering

Feature extraction and selection are critical tasks in heart sound signal analysis. After denoising and segmentation, features are usually extracted manually from the raw heart sound data, which are transformed into temporal, spectral, and features containing both the temporal and frequency domain information.

Temporal features are commonly used because they are easy to extract and quantify. Statistical indicators such as signal energy (Ibrahim et al., 2021; Khan et al., 2018), amplitude (Eslamizadeh & Barati, 2017), envelope (Singh et al., 2019; Varghees & Ramachandran, 2017), and kurtosis (Xu et al., 2022) can be calculated to analyse heart sound signals. The most commonly used time-domain features include the locations of S1 and S2, systolic and diastolic intervals (Liang et al., 1997), and amplitude information such as the mean and standard deviation of time duration.

While time-domain analysis may miss some physiological and pathological information, the signal characteristics can be visualized in spectral analysis. For example, power spectral density (Ibrahim et al., 2021) demonstrates the frequency domain power distribution of a signal, and the location of the maximum power distribution can also be estimated from it. Mel-frequency cepstral coefficients (MFCCs) have also been widely used in recent years. Figure 5a shows the feature map of healthy heart sounds, and Figure 5b shows the graphical representation of MFCCs for heartbeats in patients with mitral valve prolapse.

Time-frequency representation of a signal is usually better than separate time-domain or frequency-domain graphics in capturing the components and changing trends of the instantaneous signal. Heart sound signals have been represented using a variety of transforms, including time-frequency representation using S-transform (Moukadem et al., 2013), short-time Fourier transforms (STFT) (Chen, Guo, et al., 2023), and wavelet transform. Although the STFT is widely used for heart sound analysis, obtaining appropriate resolution feature representation can be challenging using STFT (Zhang et al., 2017a).

Wavelet analysis has become a popular alternative method. Wavelet transform has been extensively described in the literature and is favoured by researchers because it may provide high temporal and frequency resolution and better depict S1 and S2 parts (Deng et al., 2020). It adopts and expands the concept of the STFT while improving the shortcomings of window size, which does not vary with frequency. It is an ideal tool for frequency analysis and processing. Standard features in heart sound analysis tasks include DWT (Yuenyong et al., 2011), continuous wavelet transform (CWT) (Debbal & Tani, 2016), and mel-scale wavelet transform (Lubaib & Muneer, 2016).

Some existing feature sets can be extracted using the open-source toolkit openSMILE for heart sound classification, including the COMPARE feature set (Schuller et al., 2018) and the eGeMAPS feature set (Eyben et al., 2015). Table 5 mainly lists the features used in traditional ML algorithms for heart sound classification.

4.3 | Models

The goal of classification is to distinguish abnormal PCG signals from normal ones. ANNs, GMMs, random forests, SVM, and The HMM are examples of conventional ML techniques, typically combine hand-crafted features to complete heart sound classification tasks. The DL techniques,

TABLE 4 Segmentation methods of heart sound signals.

Year	Authors	Methods	Dataset	Results
2019	Renna et al. (2019)	CNNs+HMMs/HSMMs	PhysioNet	Average sensitivity of 93.9% on the PhysioNet dataset
2019	Giordano and Knaflitz (2019)	The PCG envelope	24 healthy volunteers	99.2%
2018	Jain and Tiwari (2018)	TQWT	The Texas Heart Institute's heart sound series at St. Luke's Episcopal Hospital	Enhanced classification efficiency in noisy cases
2018	Alexander et al. (2018)	HMM	2016 PhysioNet computing in cardiology challenge database	90.3% sensitivity and 89.9% specificity
2018	Oliveira et al. (2018)	HSMM-GMM	PhysioNet, Pascal datasets and a paediatric dataset	F-score 92%
2017	Gavrovska et al. (2017)	Fractal decomposition	Over a thousand sequences belonging to paediatric subjects	Accuracy (higher than 95%)
2017	Othman and Khaleel (2017)	Shannon energy	Three abnormal MR cases three abnormal MS cases three normal cases	The objective is to distinguish normal heart sounds from those that are classified as MR or MS
2017	Babu et al. (2017)	Variational mode decomposition	The study utilized various standard databases such as Physionet, Pascal, Michigan, and eGeneralMedical, as well as recorded signals	Greater than 95% and a detection error rate of less than 6%
2017	Varghees and Ramachandran (2017)	Empirical wavelet transform	PhysioNet/CinC challenge database, PASCAL database and real-time PCG signals	Average accuracy for PCG signals with a 20 dB signal-to-noise ratio (SNR) was 95.5%
2016	Springer et al. (2015)	HSMM	A total of 123 adult patients underwent concurrent PCG and ECG recordings, resulting in 405 recordings that lasted between 30 and 40 s each	F1-score average of 95.63 ± 0.85%
2014	Sun et al. (2014)	Hilbert transform	The Michigan database and sounds from clinical heart diseases	97.37%

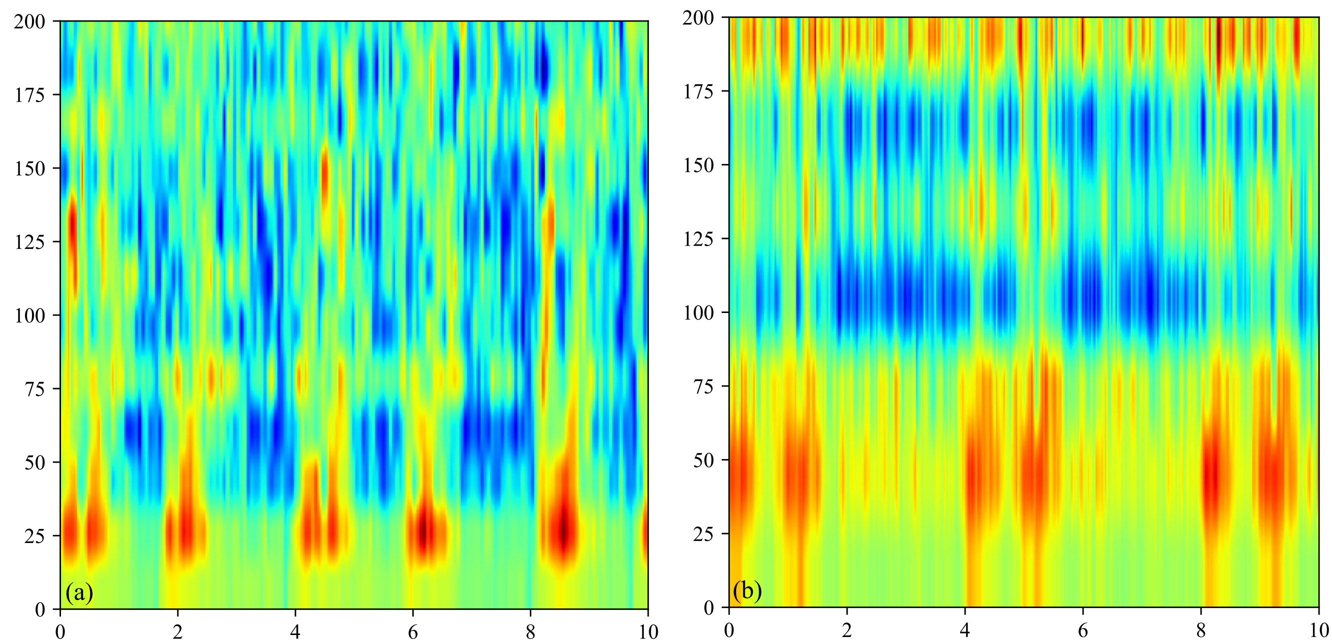

FIGURE 5 Heart sound wave MFCCs of healthy and unhealthy hearts in the HVDB: (a) MFCCs of normal; (b) MFCCs of MVP.

including deep convolutional neural networks, have recently become popular due to their ability to automatically perform feature extraction. For more details, please refer to Section 5.

SVM is a classical supervised learning method used for classification tasks. To increase the effectiveness of training, the kernel function's parameters might be changed for classification tasks, which may require the help of an additional optimizer (Redlarski et al., 2014). In general, to

TABLE 5 Common manually extracted features.

Features	Reference	Description
Temporal	Liang et al. (1997)	The mean and standard deviation of time duration
	Potes et al. (2016)	Skewness
	Xu et al. (2022)	Kurtosis and short-term energy of each stage
	Varghees and Ramachandran (2017) and Singh et al. (2019)	Envelope
	Eslamizadeh and Barati (2017)	Amplitude
	Khan et al. (2018) and Ibrahim et al. (2021)	Energy
	Khan et al. (2018) and Soares et al. (2020)	Entropy
Spectral	Langley and Murray (2017) and Khan et al. (2018)	Spectral amplitude
	Khaled et al. (2022) and Arora et al. (2019)	Dominant frequency ratio and value
	Noman et al. (2019)	Spectral energy
	Ibrahim et al. (2021)	Power spectral density
	Alqudah (2019)	Instantaneous frequency
	Yadav et al. (2018)	Cepstrum
Mel-frequency	Maknickas and Maknickas (2017)	Mel-frequency
	Khan et al. (2020), Chen and Dang (2019), Chen, Dang, and Li (2022), Rahmandani et al. (2018), and Adiban et al. (2019)	Mel-frequency cepstral coefficients (MFCCs)
	Abduh et al. (2020)	Fractional Fourier transform-based mel-frequency
Wavelet analysis	Baydoun et al. (2020)	Wavelet transform
	Li et al. (2019) and Mei et al. (2021)	Wavelet scattering transform
	Ghosh et al. (2019)	Wavelet synchrosqueezing transform
	Sawant et al. (2021)	Tunable quality wavelet transform
	Langley and Murray (2017)	Wavelet entropy

improve the performance of SVM classifiers, most studies modify previous SVM-based classification methods or update features. SVM classifiers are often used in the classification of heart valve diseases in recent studies. Studies have pointed out that SVM classifiers are more suitable for identifying innocent noise than ANN (Kang et al., 2016). For high-dimensional classification problems, SVM classifiers are applicable even with small sample sizes (Tseng et al., 2016). Studies have shown that the performance of SVM classifiers is not necessarily dependent on the dimensionality of the input vectors (Ari et al., 2010).

The KNN algorithm is based on the idea that if most of the K nearest neighbouring samples of a given sample in the feature space belong to a specific class, then it can be concluded that this sample not only belongs to that class, but also shares the characteristics of the samples in that class. The KNN algorithm has been successfully applied in heart sound classification and noise detection (Delgado-Trejos et al., 2009). It is found that the k parameters used in the algorithm will determine the classification performance, and k can be applied in combination with various features to improve the classification performance. But the algorithm requires a relatively large memory space (Wang et al., 2017).

In addition, some studies have used the HMM classifier to distinguish between normal and abnormal heart sounds. HMM is a type of Markov chain, which is a stochastic process characterized by the Markov property. During the investigation, it was found that the HMM model is mainly used for segmentation, and it has the disadvantage of slow interpretation in the process of heart sound classification. HMM classifiers usually contain a large number of parameters, and their classification accuracy depends on the HMM parameters chosen in the model (Fahad et al., 2018). Table 6 lists the common traditional ML algorithms for heart sound classification.

5 | DEEP LEARNING FOR HEART SOUND CLASSIFICATION

The DL has achieved remarkable success in various fields (He et al., 2023), such as image recognition, it has also been applied to heart sound signal analysis. We present and discuss the application of DL methods to heart sound classification in this section.

TABLE 6 Traditional machine learning classifiers commonly used in heart sound classification tasks.

Model	Reference	Remark
KNN	Sengur and Turkoglu (2008)	A fuzzy KNN algorithm based on artificial immune system (AIS) for identifying heart valve diseases, was proposed by Sengur and Turkoglu
	Quiceno-Manrique et al. (2010)	This study presents a method for detecting systolic murmurs in phonocardiographic recordings by calculating spectral power over time
	Avendano-Valencia et al. (2010)	The extracted features from this time-frequency representation (TFR) were utilized to train a straightforward k-NN classifier
SVM	Zhang et al. (2017b)	Zhang and colleagues proposed a heart sound classification method that combines support vector machine (SVM) with proportional spectrogram and partial least squares regression (PLSR)
	Deng and Han (2016)	To obtain a feature representation, autocorrelated features are extracted from the envelope of the discrete wavelet decomposition (DWT). The resulting features are then used to train a support vector machine (SVM) learner for heartbeat classification
	Zheng et al. (2015)	After pre-processing, the heart sound signal is reconstructed using wavelet decomposition, and features such as energy fraction and sample entropy are chosen. A support vector machine (SVM) is then utilized to detect abnormalities in the heart sound
HMM	Fahad et al. (2018)	Fahad et al. proposed a scheme to fuse adaptive neuro-fuzzy inference system (ANFIS) and HMM to identify heart murmurs
	Uğuz et al. (2007)	The wavelet and short-time Fourier transform to extract features and feed them into the hidden Markov model (HMM)
	Chauhan et al. (2008)	In this study, the authors developed a hidden Markov model (HMM) based on the mel frequency cepstrum coefficient feature for the classification task
ANN	Beritelli et al. (2018)	Beritelli et al. proposed a method for diagnosing cardiac activity based on Gram polynomials and probabilistic neural networks
	Abdollahpur et al. (2017)	This article extracts four sets of features in the temporal, time-frequency and perceptual domains. Then a voting system based on three feedforward neural networks and between classifiers is used to accomplish the main classification process
	Kay and Agarwal (2017)	Following segmentation, the original signal is processed using features such as continuous wavelet transform (CWT) and mel frequency cepstrum coefficients (MFCC), which are then fed into two fully connected hidden layer neural networks to detect the presence of murmurs

5.1 | Deep learning for pre-processing

5.1.1 | Deep learning for denoising

According to the literature (Ali et al., 2023), heart sound denoising has seen limited use of DL architectures, and further exploration of its application in this field is necessary. Our research found only two studies that utilized DL-based architectures for denoising heart sound signals. One study evaluated the effectiveness of 2D U-Net and denoising convolutional neural network (DnCNN) in denoising heart sound signals (Sharan et al., 2020), while another proposed an end-to-end DL model for real-time denoising of heart sounds (Ali et al., 2023). Considering the impact of practical denoising algorithms on classification performance is essential, as their inclusion may have a negative effect. However, the effect of denoising on PCG classification performance was not evaluated in the denoising method described in Sharan et al. (2020). In Ali et al. (2023), the authors address this issue using an FCN-based denoising autoencoder and a 1D U-Net as baseline systems. These methods were inspired by a similar denoising approach used for ECG signals (Chiang et al., 2019). The experimental evaluation metrics demonstrate that the proposed LU-Net outperforms the FCN and U-Net. These results highlight the importance of carefully selecting denoising methods for optimal classification performance.

5.1.2 | Deep learning for segmentation

The recent advancements in DL's success in image segmentation have motivated researchers to develop heart sound segmentation methods (Wang et al., 2023). Current advanced solutions for heart sound segmentation utilize statistical models like HMM and HSMM. However, the application of CNNs to the heart sound signal segmentation task has also been explored. Renna et al. (2019) were among the first to attempt this by developing a temporal modelling solution. The estimated segmentation sequence is produced by combining the CNN outputs corresponding to different PCG parts. Recurrent neural networks (RNNs) have also been used to localize heart sound states. For example, bidirectional gated recurrent units (GRU-RNNs) were developed in Fan et al. (2018) to directly segment heart sounds. The literature (Chen, Lv, et al., 2020) proposed a duration-long-short-term memory (LSTM) to address the problem of poor utilization of cardiac cycle duration information due to the inability of envelope features to effectively model intrinsic sequence features. In order to overcome the difficulties brought on by erratic and noisy PCG recordings, Fernando et al. (2019) proved the usefulness of a mix of RNN-based temporal modelling and attention-based salient feature extraction strategies. Lastly, Chen, Sun, Lv, et al. (2021) proposed an end-to-end approach that employs Convolutional CLSTM architecture. This method uses convolutional layers to extract relevant features and downsample them, and LSTM layers for sequence recognition. These methods showcase the potential of DL for heart sound segmentation, emphasizing the significance of considering diverse neural network architectures to design effective segmentation methods.

5.2 | Deep learning for classification

This subsection outlines the latest progress in DL research concerning heart sound signal analysis in recent years. Hinton et al. (2006) first proposed the idea of DL in 2006. Recently, DL has demonstrated its effectiveness in various fields, including biomedical data analysis (Shen et al., 2017) and signal processing (Yu & Deng, 2010). DL models have been explored and utilized for intelligent heart sound signal classification. These include transfer learning, mechanisms for attention, end-to-end learning strategies, CNNs, RNNs, combination of CNNs and RNNs.

The 2D CNN approach for classifying heart sounds is inspired by the successes achieved in computer vision. Pre-processing is used to create 2D feature maps that reflect both the frequency and time domains from the initial 1D heart sound input. These maps are then used as inputs to the 2D CNN for classification. Commonly used 2D representations of heartbeat signals include Mel spectrograms and short-time Fourier transform time-frequency maps. In the literature (Nguyen et al., 2023), two DL models, the LSTM and the 2D CNN, are suggested for categorizing heartbeat sounds based on extracted features. The log-mel spectrum characteristics are then retrieved after the cardiac signal has been initially cropped to a constant length. These features are then fed into the LSTM and 2D CNN models for training and testing. The suggested technique achieves an accuracy of 99.67% on the GitHub dataset. In a published study (Xiang et al., 2023), the effect of different two-dimensional features on heart sound classification models utilizing the 2D CNN architecture was examined. The study evaluated features such as mel log spectrograms and signal waveform maps and concluded that log-mel and log power features outperformed envelope and waveform features in heart sound classification. The literature (Zhang et al., 2019) utilized the short-time Fourier transform (STFT) to extract the spectrogram of the heartbeat signal. However, the resolution of the signal can be affected by the STFT window size, leading some researchers to suggest that wavelet transform-based features may be more effective. Heartbeat classification tasks have been successfully completed using wavelet analysis. In the study (Qian et al., 2019), a system founded on wavelet representation and deep recurrent neural networks was proposed to identify three types of heartbeats: normal, mild, and severe. The proposed approach demonstrated promising results in heart sound classification tasks. Although 2D CNN has been demonstrated to perform effectively in tasks requiring the classification of heart sounds, they need further feature transformations. Several CNN architectures have been proposed for 1D CNN-based methods to identify abnormal heart sounds, including in studies such as Fakhry and Brery (2022), Krishnan et al. (2020), and Li, Yao, et al. (2020). In a typical example, as demonstrated in Zeng et al. (2023), raw data is utilized as input to a 1D CNN without additional transformations. RNNs are well-suited for processing sequential data, including heart sound signals, due to their solid temporal correlation (Xu et al., 2020; Zhang et al., 2019). RNN architectures, including gated recurrent units (GRU) and LSTM, are commonly employed in various fields due to their ability to model sequential data with temporal dependencies.

Various integration methods for heart sound classification have been proposed. One such method is the fusion of CNN and RNN models, which can extract both frequency and time domain information, fully leveraging the global characteristics of heart sound data and achieving better performance than individual models (Li et al., 2021; Shuvo et al., 2021). A table, as shown in Table 7, offers a summary of notable literature that discusses various heart sound classification techniques.

6 | DISCUSSION

The use of computers to assist in heart sound analysis has gained popularity recently. This section delves into the specifics of heart sound classification methods. Additionally, the section addresses future trends and challenges in the development of heart sound signal recognition methods.

Both ECG analysis and heart sound analysis are valuable techniques for predicting CVDs, but they provide different types of information about the heart's functioning. Here are the advantages and differences of heart sound analysis compared to ECG for predicting CVDs as follow.

1. Different information. The ECG primarily measures the electrical activity of the heart, providing information about the heart's rhythm, conduction abnormalities, and other electrical parameters. On the other hand, heart sound analysis focuses on capturing the acoustic signals produced by the heart, namely the S1 and S2 sounds, as well as additional abnormal sounds like murmurs or gallops. These sounds reflect the mechanical events occurring during the cardiac cycle, including the closing of the heart valves.

TABLE 7 Deep learning for classification.

Year	Reference	Dataset	Pre-processing	Input features	Model	Segment	Crop	Results
2019	Qian et al. (2019)	The Heart Sounds Shenzhen corpus	Unknown	Wavelet energy features	DRNN	No	No	UAR: 43.0%, Chance Level UAR: 33.3%
2019	Zhang et al. (2019)	PhysioNet 2016	The bilinear interpolation	The temporal features obtained from the Short-Time Fourier Transform spectrogram	LSTM	No	Yes	Overall scores of 0.9484%
2019	Koike et al. (2020)	PhysioNet 2016	Resampled with 2 KHz	Spectrogram and log mel spectrogram	Transfer learning models (PANNs)	No	Yes	Unweighted average recall at 89.7%
2020	Gao et al. (2020)	Acquired from University-T own Hospital of Chongqing Medical University+PhysioNet 2016	Down-sampled at 600 Hz S1 marking and segmentation Normalization	Without denoising and hand-crafted feature extraction	Gated recurrent unit (GRU) model	Yes	Yes	Average accuracy of 98.82%
2021	Shuvo et al. (2021)	Github+ PhysioNet 2016	Resampled at 2 kHz and amplitude normalized	Three parallel CNN	CRNN	No	Yes	99.60% accuracy and an accuracy of 86.57%
2021	Li et al. (2021)	PhysioNet 2016	Normalized + a fifth-order Butterworth bandpass filter with cutoff frequency of 25–500 Hz.	CRNN	CRNN	No	Yes	95.00%
2022	Nguyen et al. (2023)	Github	The signal samples were cropped	Log-mel spectrogram	2DCNN/LSTM	No	Yes	Overall accuracy of about 99.67%
2022	Xiang et al. (2023)	A collection of four open datasets	Wavelet transform, denoising, down-sample	Log-mel spectrogram, logarithmic power spectrogram, waveform diagram, the envelope of signals	2DCNN	No	Yes	Average accuracy around 94%
2022	Ismail et al. (2023)	Pascal	Decimation and windowing	Spectrograms	CNN + SVM	No	Yes	Precision, recall, and accuracy values exceeding 95%
2022	Zeng et al. (2023)	Github	Butterworth bandpass filter (25–900 Hz)	Teager-Kaiser energy operator (TKEO) + rational dilation wavelet transform	1DCNN	No	No	Overall average accuracy 99.00%
2022	Fakhry and Brery (2022)	PhysioNet 2016 (select 150 healthy and 150 unhealthy)	A low-pass filter with a 15th-order Butterworth design and a cut-off frequency of 250 Hz	Mean, mode, variance, skewness, kurtosis, Shannon energy, Shannon entropy, zero-crossing rate	The bidirectional long short-term memory	No	Yes	Accuracy of 89.10%

2. Complementary insights. Heart sound analysis offers complementary insights to ECG analysis. While ECG is well-established and widely used, it may not always capture certain structural abnormalities or subtle changes in cardiac function. By incorporating heart sound analysis, clinicians can gain additional information about the heart's mechanical function, valvular abnormalities, and potential structural defects that may not be evident solely through ECG.
3. Early detection of certain conditions. Heart sound analysis has shown promise in the early detection of specific cardiac conditions. For example, certain murmurs or abnormal heart sounds can indicate valvular disorders such as aortic stenosis. Identifying these conditions at an early stage can facilitate timely intervention and prevent further progression.
4. Non-invasive and portable heart sound analysis can be performed using non-invasive methods with portable devices, making it accessible and convenient for screening purposes. This advantage is particularly relevant in resource-constrained settings where access to advanced diagnostic tools like ECG machines may be limited.

Referring to a survey (Dwivedi et al., 2018), heart sound classification utilizing DL models generally has higher accuracy compared to traditional ML approaches. Classical ML techniques and DL techniques result in distinct models when used for heartbeats classification. Traditional ML models are trained using small batches of data and often involve complex data pre-processing methods. The effect of conventional heart sound classification ML methods depends on whether the collected features can clearly distinguish pathological information. The purpose of data pre-processing is to retain more original heart sound signals to remove noise and obtain practical features, which is a crucial part of the traditional heart sound classification ML model training. Generally, the segmentation algorithm is used to locate the position of the primary heart sound to extract the temporal domain, domain of frequency, and statistical features of the primary heart sound. The introduction of the segmentation algorithm will first increase the computational complexity. Second, the effectiveness of manual feature extraction depends on the effectiveness of the segmentation algorithm for locating heart sounds. We cannot guarantee the 100% accuracy of the segmentation algorithm, which will inevitably add a little bit uncertainty to the initial classification task. Small batches of training data will reduce the generalization and robustness of the model. The model's training time will be greatly lengthened by huge training data sets. The balance between the two is required the researcher's concerns.

The DL can effectively and automatically learn features through large batches of data training. Usually, researchers only need to iteratively optimize model parameters, but the learning process of the model is like a black box to us. At present, some researchers pay attention to the explainable model, and they also pay attention to why the model has good results while pursuing the efficiency of the model. Building interpretable models based on DL can increase the credibility. A past study (Ren et al., 2022) used an attention mechanism to visualize each feature unit's contribution to the final heart sound prediction. The SHAP is similarly used in Lundberg and Lee (2017) to compute the contribution of each feature unit. The end-to-end network simplifies the prediction process. A single architecture of DL methods can simultaneously achieve feature extraction and classification. It is also the reason why most work chooses DL architecture. Segmentation algorithms are often rarely chosen for DL architectures on heart sound classification tasks. Some literature (Dissanayake et al., 2020) thinks that whether the segmentation algorithm has a positive effect on the classification effect may be related to the robustness of the model itself. For some algorithms, segmentation may not lead to a significant improvement in the model's performance. The reason may be that the model is already potent and robust (Ren et al., 2023). The author Dissanayake et al. (2020) also demonstrated this view via the SHAP model (Lundberg & Lee, 2017). It can be seen from Table 4 that many methods are proposed to cut heart sounds into short segments of the same length, thereby reducing the complexity of automatic auscultation.

Validation techniques are typically used to evaluate the effectiveness of the research methodology. Here are some common validation techniques that are summarized from previous heart sound research.

1. Cross-validation. It is a technique commonly used in statistical modelling and ML to assess the performance and generalizability of predictive models. Cross-validation involves splitting the available data into subsets for training and testing the model. By evaluating the model's performance on multiple independent subsets, researchers can determine its ability to generalize to new data and detect potential overfitting.
2. External validation. It involves comparing the results or predictions of a model or study with independent external data or established benchmarks. This technique helps assess the generalizability and applicability of research findings beyond the original dataset or context.
3. Expert evaluation. It involves seeking the opinion and feedback of subject matter experts or professionals in the field to validate the research methodology, findings, or conclusions. Expert input can provide valuable insights and perspectives, especially in complex or specialized domains.

There are numerous approaches to classifying heart sound signals. However, this research field still faces many technical challenges. For example, it is difficult to obtain clean data because the recording environment may be disturbed by environmental noise and speech. Denoising is thus an essential pre-processing step to improve audio quality. Moreover, evaluating the effectiveness of existing algorithms is challenging due to the diverse range of test datasets utilized, making it difficult to draw direct comparisons between them. The current theoretical research needs to include a combination of practical applications. In future work, the algorithm architecture should consider the effect and the clinical practicality.

7 | CONCLUSION

This survey aims to assist researchers in reviewing recent advancements in automated algorithms for heartbeats analysis and diagnosis. It lists common heart sound databases, assessment metrics, segmentation, categorization algorithms, and denoising, segmentation, and feature extraction techniques. The survey focuses on implementing conventional ML and DL techniques in various heart sound analysis sub-tasks. Additionally, it discusses current challenges in heart sound analysis. The survey's findings serve as as a guide for future research, facilitating the clinical application of heart sound recognition in disease diagnosis.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

ORCID

Xu Xu https://orcid.org/0000-0002-3934-9096
Gwanggil Jeon https://orcid.org/0000-0002-0651-4278
David Camacho https://orcid.org/0000-0002-5051-3475

ENDNOTES

[1] https://www.physionet.org/content/challenge-2016/1.0.0/.

[2] http://www.diit.unict.it/hsct11.

[3] www.peterjbentley.com/heartchallenge.

[4] https://github.com/yaseen21khan/Classification-of-Heart-Sound-Signal-Using-Multiple-Features–/.

[5] www.med.umich.edu/lrc/psb/heartsounds/index.htm.

[6] https://physionet.org/content/ephnogram/1.0.0/.

[7] http://www.egeneralmedical.com/listohearmur.html.

[8] https://www.thinklabs.com/heart-sounds-old.

[9] http://feeds.texasheart.org/HeartSoundsPodcastSeries.

[10] http://int-prop.lf2.cuni.cz/heart_sounds/h14/sound.htm.

REFERENCES

Abdollahpur, M., Ghaffari, A., Ghiasi, S., & Mollakazemi, M. J. (2017). Detection of pathological heart sounds. *Physiological Measurement*, *38*(8), 1616.

Abduh, Z., Nehary, E. A., Wahed, M. A., & Kadah, Y. M. (2020). Classification of heart sounds using fractional fourier transform based mel-frequency spectral coefficients and traditional classifiers. *Biomedical Signal Processing and Control*, *57*, 101788.

Adiban, M., BabaAli, B., & Shehnepoor, S. (2019). Statistical feature embedding for heart sound classification. *Journal of Electrical Engineering*, *70*(4), 259–272.

Ahlstrom, C., Hult, P., Rask, P., Ahlstrom, C., Hult, P., Rask, P., Karlsson, J.-E., Nylander, E., Dahlstrom, U., & Ask, P. (2006). Feature extraction for systolic heart murmur classification. *Annals of Biomedical Engineering*, *34*, 1666–1677.

Ahmad, B., Khan, F. A., Khan, K. N., & Khan, M. S. (2021). Automatic classification of heart sounds using long short-term memory. In *2021 15th International Conference on Open Source Systems and Technologies (ICOSST)* (pp. 1–6). IEEE.

Alam, U., Asghar, O., Khan, S. Q., Hayat, S., & Malik, R. A. (2010). Cardiac auscultation: an essential clinical skill in decline. *British Journal of Cardiology*, *17*(1), 8–10.

Alexander, B., Nallathambi, G., & Selvaraj, N. (2018). Screening of heart sounds using hidden Markov and Gammatone filterbank models. In *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)* (pp. 1460–1465). IEEE.

Ali, S. N., Shuvo, S. B., Al-Manzo, M. I. S., Hasan, A., & Hasan, T. (2023). An end-to-end deep learning framework for real-time denoising of heart sounds for cardiac disease detection in unseen noise. *IEEE Access*, *11*, 87887–87901.

Alqudah, A. M. (2019). Towards classifying non-segmented heart sound records using instantaneous frequency based features. *Journal of Medical Engineering & Technology*, *43*(7), 418–430.

Ari, S., Hembram, K., & Saha, G. (2010). Detection of cardiac abnormality from PCG signal using LMS based least square SVM classifier. *Expert Systems with Applications*, *37*(12), 8019–8026.

Arora, V., Leekha, R., Singh, R., & Chana, I. (2019). Heart sound classification using machine learning and phonocardiogram. *Modern Physics Letters B*, *33*(26), 1950321.

Avendano-Valencia, L., Godino-Llorente, J., Blanco-Velasco, M., & Castellanos-Dominguez, G. (2010). Feature extraction from parametric time–frequency representations for heart murmur detection. *Annals of Biomedical Engineering*, *38*, 2716–2732.

Babu, K. A., Ramkumar, B., & Manikandan, M. S. (2017). S1 and S2 heart sound segmentation using variational mode decomposition. In *TENCON 2017-2017 IEEE Region 10 Conference* (pp. 1629–1634). IEEE.

Baydoun, M., Safatly, L., Ghaziri, H., & El Hajj, A. (2020). Analysis of heart sound anomalies using ensemble learning. *Biomedical Signal Processing and Control*, *62*, 102019.

Beritelli, F., Capizzi, G., Lo Sciuto, G., Napoli, C., & Scaglione, F. (2018). Automatic heart activity diagnosis based on gram polynomials and probabilistic neural networks. *Biomedical Engineering Letters*, *8*, 77–85.

Brites, I. S. G., Silva, L. M., Barbosa, J. L. V., Rigo, S. J., Correia, S. D., & Leithardt, V. R. Q. (2021). Machine learning and IoT applied to cardiovascular diseases identification through heart sounds: A literature review. *Informatics*, *8*, 73.

Cesarelli, M., Ruffo, M., Romano, M., & Bifulco, P. (2012). Simulation of foetal phonocardiographic recordings for testing of FHR extraction algorithms. *Computer Methods and Programs in Biomedicine*, *107*(3), 513–523.

Ceyhan, M., Orhan, Z., & Domnori, E. (2017). Health service quality measurement from patient reviews in Turkish by opinion mining. In *Proceedings of the International Conference on Medical and Biological Engineering 2017* (pp. 649–653). Springer.

Chakrabarti, T., Saha, S., Roy, S., & Chel, I. (2015). Phonocardiogram signal analysis-practices, trends and challenges: A critical review. In *2015 International Conference and Workshop on Computing and Communication (IEMCON)* (pp. 1–4). IEEE.

Chauhan, S., Wang, P., Lim, C. S., & Anantharaman, V. (2008). A computer-aided MFCC-based HMM system for automatic auscultation. *Computers in Biology and Medicine*, *38*(2), 221–233.

Chen, J., Chen, L., & Zhou, Y. (2020). Cryptanalysis of a DNA-based image encryption scheme. *Information Sciences*, *520*, 130–141.

Chen, J., Dang, X., & Li, M. (2022). Heart sound classification method based on ensemble learning. In *2022 7th International Conference on Intelligent Computing and Signal Processing (ICSP)* (pp. 8–13). IEEE.

Chen, J., Guo, Z., Xu, X., Zhang, L.-B., Teng, Y., Chen, Y., Wozniak, M., & Wang, W. (2023). A robust deep learning framework based on spectrograms for heart sound classification. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*.

Chen, J., Sun, S., Zhang, L. B., Yang, B., & Wang, W. (2022). Compressed sensing framework for heart sound acquisition in Internet of Medical Things. *IEEE Transactions on Industrial Informatics*, *18*(3), 2000–2009.

Chen, J., Wang, W., Fang, B., Liu, Y., Yu, K., Leung, V. C. M., & Hu, X. (2023). Digital twin empowered wireless healthcare monitoring for smart home. *IEEE Journal on Selected Areas in Communications*, *41*(11), 3662–3676.

Chen, J. F., & Dang, X. (2019). Heart sound analysis based on extended features and related factors. In *2019 IEEE Symposium Series on Computational Intelligence (SSCI)* (pp. 2189–2194). IEEE.

Chen, W., Sun, Q., Chen, X., Xie, G., Wu, H., & Xu, C. (2021). Deep learning methods for heart sounds classification: A systematic review. *Entropy*, *23*(6), 667.

Chen, Y., Lv, J., Sun, Y., & Jia, B. (2020). Heart sound segmentation via duration long–short term memory neural network. *Applied Soft Computing*, *95*, 106540.

Chen, Y., Sun, Y., Lv, J., Jia, B., & Huang, X. (2021). End-to-end heart sound segmentation using deep convolutional recurrent network. *Complex & Intelligent Systems*, *7*, 2103–2117.

Chiang, H. T., Hsieh, Y. Y., Fu, S. W., Hung, K. H., Tsao, Y., & Chien, S. Y. (2019). Noise reduction in ECG signals using fully convolutional denoising autoencoders. *IEEE Access*, *7*, 60806–60813.

Chizner, M. A. (2008). Cardiac auscultation: Rediscovering the lost art. *Current Problems in Cardiology*, *33*(7), 326–408.

Chowdhury, T. H., Poudel, K. N., & Hu, Y. (2020). Time-frequency analysis, denoising, compression, segmentation, and classification of PCG signals. *IEEE Access*, *8*, 160882–160890.

Debbal, S., & Tani, A. M. (2016). Heart sounds analysis and murmurs. *International Journal of Medical Engineering and Informatics*, *8*(1), 49–62.

Delgado-Trejos, E., Quiceno-Manrique, A., Godino-Llorente, J., Blanco-Velasco, M., & Castellanos-Dominguez, G. (2009). Digital auscultation analysis for heart murmur detection. *Annals of Biomedical Engineering*, *37*, 337–353.

Deng, M., Meng, T., Cao, J., Wang, S., Zhang, J., & Fan, H. (2020). Heart sound classification based on improved MFCC features and convolutional recurrent neural networks. *Neural Networks*, *130*, 22–32.

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.

Dissanayake, T., Fernando, T., Denman, S., Sridharan, S., Ghaemmaghami, H., & Fookes, C. (2020). A robust interpretable deep learning classifier for heart anomaly detection without segmentation. *IEEE Journal of Biomedical and Health Informatics*, *25*(6), 2162–2171.

Dwivedi, A. K., Imtiaz, S. A., & Rodriguez-Villegas, E. (2018). Algorithms for automatic analysis and classification of heart sounds—A systematic review. *IEEE Access*, *7*, 8316–8345.

Eslamizadeh, G., & Barati, R. (2017). Heart murmur detection based on wavelet transformation and a synergy between artificial neural network and modified neighbor annealing methods. *Artificial Intelligence in Medicine*, *78*, 23–40.

Eyben, F., Scherer, K. R., Schuller, B. W., Sundberg, J., Andre, E., Busso, C., Devillers, L. Y., Epps, J., Laukka, P., Narayanan, S. S., & Truong, K. P. (2015). The Geneva minimalistic acoustic parameter set (GeMAPS) for voice research and affective computing. *IEEE Transactions on Affective Computing*, *7*(2), 190–202.

Fahad, H., Ghani Khan, M. U., Saba, T., Rehman, A., & Iqbal, S. (2018). Microscopic abnormality classification of cardiac murmurs using ANFIS and HMM. *Microscopy Research and Technique*, *81*(5), 449–457.

Fakhry, M., & Brery, A. F. (2022). A comparison study on training optimization algorithms in the biLSTM neural network for classification of PCG signals. In *2022 2nd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)* (pp. 1–6). IEEE.

Fan, T., Zhu, J., Cheng, Y., Li, Q., Xue, D., & Munnoch, R. (2018). A new direct heart sound segmentation approach using bi-directional GRU. In *2018 24th International Conference on Automation and Computing (ICAC)* (pp. 1–5). IEEE.

Fernando, T., Ghaemmaghami, H., Denman, S., Sridharan, S., Hussain, N., & Fookes, C. (2019). Heart sound segmentation using bidirectional LSTMs with attention. *IEEE Journal of Biomedical and Health Informatics*, *24*(6), 1601–1609.

Gao, S., Zheng, Y., & Guo, X. (2020). Gated recurrent unit-based heart sound analysis for heart failure screening. *Biomedical Engineering Online*, *19*, 1–17.

Gao, Y., Xu, X., Li, C., Liu, J., & Wu, T. (2023). Semi-automatic framework for voxel human deformation modeling. *Current Medical Imaging*.

Gavrovska, A., Slavkovic, M., Reljin, I., & Reljin, B. (2013). Application of wavelet and EMD-based denoising to phonocardiograms. In *International Symposium on Signals, Circuits and Systems ISSCS2013* (pp. 1–4). IEEE.

Gavrovska, A., Zajic, G., Bogdanovic, V., Reljin, I., & Reljin, B. (2017). Identification of S1 and S2 heart sound patterns based on fractal theory and shape context. *Complexity*, *2017*, 1580414.

Ghosh, S. K., Nagarajan, P. R., & Tripathy, R. K. (2020). Heart sound data acquisition and preprocessing techniques: A review. In *Handbook of research on advancements of artificial intelligence in healthcare engineering* (pp. 244–264). IGI Global.

Ghosh, S. K., Tripathy, R. K., & Ponnalagu, R. (2020). Evaluation of performance metrics and denoising of PCG signal using wavelet based decomposition. In *2020 IEEE 17th India Council International Conference (INDICON)* (pp. 1–6). IEEE.

Ghosh, S. K., Tripathy, R. K., Ponnalagu, R., & Pachori, R. B. (2019). Automated detection of heart valve disorders from the PCG signal using time-frequency magnitude and phase features. *IEEE Sensors Letters*, *3*(12), 1–4.

Giordano, N., & Knaflitz, M. (2019). A novel method for measuring the timing of heart sound components through digital phonocardiography. *Sensors*, *19*(8), 1868.

Gomes, E. F., Bentley, P. J., Pereira, E., Coimbra, M. T., & Deng, Y. (2013). Classifying heart sounds-approaches to the PASCAL challenge. In *HEALTHINF* (pp. 337–340). SciTePress.

Hadiyoso, S., Mardiyah, D. R., Ramadan, D. N., & Ibrahim, A. (2020). Implementation of electronic stethoscope for online remote monitoring with mobile application. *Bulletin of Electrical Engineering and Informatics*, *9*(4), 1595–1603.

He, B. G., Lin, B., Li, H. P., & Zhu, S. Q. (2024). Suggested method of utilizing soil arching for optimizing the design of strutted excavations. *Tunnelling and Underground Space Technology*, *143*, 105450.

He, B. G., Wang, L., Feng, X. T., & Zhen, R. L. (2023). Failure modes of jointed granite subjected to weak dynamic disturbance under true-triaxial compression. *Rock Mechanics and Rock Engineering*, *56*, 7939–7957.

Hinton, G., Osindero, S., Welling, M., & Teh, Y. W. (2006). Unsupervised discovery of nonlinear structure using contrastive backpropagation. *Cognitive Science*, *30*(4), 725–731.

Ibrahim, N., Jamal, N., Ibrahim, N., Jamal, N., Sha'abani, M. N. A.-H., & Mahadi, L. F. (2021). A comparative study of heart sound signal classification based on temporal, spectral and geometric features. In *2020 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)* (pp. 24–29). IEEE.

Ismail, S., Ismail, B., Siddiqi, I., & Akram, U. (2023). PCG classification through spectrogram using transfer learning. *Biomedical Signal Processing and Control*, *79*, 104075.

Jain, P. K., & Tiwari, A. K. (2018). A robust algorithm for segmentation of phonocardiography signal using tunable quality wavelet transform. *Journal of Medical and Biological Engineering*, *38*, 396–410.

Kang, S., Doroshow, R., McConnaughey, J., & Shekhar, R. (2016). Automated identification of innocent still's murmur in children. *IEEE Transactions on Biomedical Engineering*, *64*(6), 1326–1334.

Kay, E., & Agarwal, A. (2017). DropConnected neural networks trained on time-frequency and inter-beat features for classifying heart sounds. *Physiological Measurement*, *38*(8), 1645.

Kazemnejad, A., Gordany, P., & Sameni, R. (2021). An open–access simultaneous electrocardiogram and phonocardiogram database. *bioRxiv*.

Khaled, S., Fakhry, M., Esmail, H., Ezzat, A., & Hamad, E. (2022). Analysis of training optimization algorithms in the NARX neural network for classification of heart sound signals. *International Journal of Scientific and Engineering Research*, *13*(2), 382–390.

Khan, F. A., Abid, A., & Khan, M. S. (2020). Automatic heart sound classification from segmented/unsegmented phonocardiogram signals using time and frequency features. *Physiological Measurement*, *41*(5), 055006.

Khan, N. M., Khan, M. S., & Khan, G. M. (2018). Automated heart sound classification from unsegmented phonocardiogram signals using time frequency features. *International Journal of Computer and Information Engineering*, *12*(8), 598–603.

Koike, T., Qian, K., Kong, Q., Plumbley, M. D., Schuller, B. W., & Yamamoto, Y. (2020). Audio for audio is better? An investigation on transfer learning models for heart sound classification. In *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* (pp. 74–77). IEEE.

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*, 505–515.

Kumar, Y., Koul, A., Singla, R., & Ijaz, M. F. (2022). Artificial intelligence in disease diagnosis: A systematic literature review, synthesizing framework and future research agenda. *Journal of Ambient Intelligence and Humanized Computing*, *14*, 8459–8486.

Lahmiri, S., & Boukadoum, M. (2014). Biomedical image denoising using variational mode decomposition. In *2014 IEEE Biomedical Circuits and Systems Conference (BioCAS) Proceedings* (pp. 340–343). IEEE.

Langley, P., & Murray, A. (2017). Heart sound classification from unsegmented phonocardiograms. *Physiological Measurement*, *38*(8), 1658.

Li, C., Yao, G., Xu, X., Yang, L., Zhang, Y., Wu, T., & Sun, J. (2020). DCSegNet: Deep learning framework based on divide-and-conquer method for liver segmentation. *IEEE Access*, *8*, 146838–146846.

Li, J., Ke, L., Du, Q., Ding, X., Chen, X., & Wang, D. (2019). Heart sound signal classification algorithm: A combination of wavelet scattering transform and twin support vector machine. *IEEE Access*, *7*, 179339–179348.

Li, S., Li, F., Tang, S., & Luo, F. (2021). Heart sounds classification based on feature fusion using lightweight neural networks. *IEEE Transactions on Instrumentation and Measurement*, *70*, 1–9.

Li, S., Li, F., Tang, S., & Xiong, W. (2020). A review of computer-aided heart sound detection techniques. *BioMed Research International*, *2020*, 5846191.

Liang, H., Lukkarinen, S., & Hartimo, I. (1997). Heart sound segmentation algorithm based on heart sound envelogram. In *Computers in Cardiology 1997* (pp. 105–108). IEEE.

Liu, C., Springer, D., Li, Q., Moody, B., Juan, R. A., Chorro, F. J., Castells, F., Roig, J. M., Silva, I., Johnson, A. E. W., Syed, Z., Schmidt, S. E., Papadaniil, C. D., Hadjileontiadis, L., Naseri, H., Moukadem, A., Dieterlen, A., Brandt, C., Tang, H., … Clifford, G. D. (2016). An open access database for the evaluation of heart sound algorithms. *Physiological Measurement*, *37*(12), 2181–2213.

Long, Q., Ye, X., & Zhao, Q. (2020). Artificial intelligence and automation in valvular heart diseases. *Cardiology Journal*, *27*(4), 404–420.

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

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, *30*, 4768–4777.

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

Martis, R. J., Acharya, U. R., & Adeli, H. (2014). Current methods in electrocardiogram characterization. *Computers in Biology and Medicine*, *48*, 133–149.

Mason, D. (2000). *Listening to the heart: A comprehensive collection of heart sounds and murmurs*. FA Davis Company.

Mei, N., Wang, H., Zhang, Y., Liu, F., Jiang, X., & Wei, S. (2021). Classification of heart sounds based on quality assessment and wavelet scattering transform. *Computers in Biology and Medicine*, *137*, 104814.

Milani, M., Abas, P. E., & De Silva, L. C. (2019). Identification of normal and abnormal heart sounds by prominent peak analysis. In *Proceedings of the 2019 International Symposium on Signal Processing Systems* (pp. 31–35). ACM.

Molcer, P. S., Kecskes, I., Delić, V., Domijan, E., & Domijan, M. (2010). Examination of formant frequencies for further classification of heart murmurs. In *IEEE 8th International Symposium on Intelligent Systems and Informatics* (pp. 575–578). IEEE.

Moukadem, A., Dieterlen, A., Hueber, N., & Brandt, C. (2013). A robust heart sounds segmentation module based on S-transform. *Biomedical Signal Processing and Control*, *8*(3), 273–281.

Nabih-Ali, M., El-Dahshan, E. S. A., & Yahia, A. S. (2017). A review of intelligent systems for heart sound signal analysis. *Journal of Medical Engineering & Technology*, *41*(7), 553–563.

Naseri, H., & Homaeinezhad, M. (2013). Detection and boundary identification of phonocardiogram sounds using an expert frequency-energy based metric. *Annals of Biomedical Engineering*, *41*, 279–292.

Nguyen, M. T., Lin, W. W., & Huang, J. H. (2023). Heart sound classification using deep learning techniques based on log-mel spectrogram. *Circuits, Systems, and Signal Processing*, *42*(1), 344–360.

Noman, F., Salleh, S. H., Ting, C. M., Samdin, S. B., Ombao, H., & Hussain, H. (2019). A Markov-switching model approach to heart sound segmentation and classification. *IEEE Journal of Biomedical and Health Informatics*, *24*(3), 705–716.

Noor, A. M., & Shadi, M. F. (2013). The heart auscultation. From sound to graphical. *Journal of Engineering and Technology (JET)*, *4*(2), 73–84.

Nowak, L. J., & Nowak, K. M. (2018). Sound differences between electronic and acoustic stethoscopes. *Biomedical Engineering Online*, *17*(1), 1–11.

Oliveira, J., Renna, F., Costa, P. D., Nogueira, M., Oliveira, C., Ferreira, C., Jorge, A., Mattos, S., Hatem, T., Tavares, T., Elola, A., Rad, A. B., Sameni, R., Clifford, G. D., & Coimbra, M. T. (2021). The CirCor DigiScope dataset: From murmur detection to murmur classification. *IEEE Journal of Biomedical and Health Informatics*, *26*(6), 2524–2535.

Oliveira, J., Renna, F., Mantadelis, T., & Coimbra, M. (2018). Adaptive sojourn time HSMM for heart sound segmentation. *IEEE Journal of Biomedical and Health Informatics*, *23*(2), 642–649.

Othman, M. Z., & Khaleel, A. N. (2017). Phonocardiogram signal analysis for murmur diagnosing using Shannon energy envelop and sequenced DWT decomposition. *Journal of Engineering Science and Technology*, *12*(9), 2393–2402.

Pham, D. H., Meignen, S., Dia, N., Fontecave-Jallon, J., & Rivet, B. (2018). Phonocardiogram signal denoising based on nonnegative matrix factorization and adaptive contour representation computation. *IEEE Signal Processing Letters*, *25*(10), 1475–1479.

Potdar, M. R., Meshram, M., Dewangan, N., & Kumar, R. (2015). Implementation of adaptive algorithm for PCG signal denoising. *International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering*, *3*(4), 33–42.

Potdar, R. M., Kumar, A., & Kumar, R. (2022). Performance analysis of DWT based denoising of vibration signals: A case study of heart sound signals. *International Journal of Mechanical Engineering*, *7*, 2810–2822.

Potes, C., Parvaneh, S., Rahman, A., & Conroy, B. (2016). Ensemble of feature-based and deep learning-based classifiers for detection of abnormal heart sounds. In *2016 Computing in Cardiology Conference (CinC)* (pp. 621–624). IEEE.

Qian, K., Ren, Z., Dong, F., Lai, W. H., Schuller, B. W., & Yamamoto, Y. (2019). Deep wavelets for heart sound classification. In *2019 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)* (pp. 1–2). IEEE.

Quiceno-Manrique, A., Godino-Llorente, J., Blanco-Velasco, M., & Castellanos-Dominguez, G. (2010). Selection of dynamic features based on time–frequency representations for heart murmur detection from phonocardiographic signals. *Annals of Biomedical Engineering*, *38*, 118–137.

Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, *77*(2), 257–286.

Rahmandani, M., Nugroho, H. A., & Setiawan, N. A. (2018). Cardiac sound classification using mel-frequency cepstral coefficients (MFCC) and artificial neural network (ANN). In *2018 3rd International Conference on Information Technology, Information System and Electrical Engineering (ICITISEE)* (pp. 22–26). IEEE.

Redlarski, G., Gradolewski, D., & Palkowski, A. (2014). A system for heart sounds classification. *PLoS One*, *9*(11), e112673.

Ren, Z., Chang, Y., Nguyen, T. T., Tan, Y., Qian, K., & Schuller, B. W. (2023). A comprehensive survey on heart sound analysis in the deep learning era. arXiv preprint arXiv:2301.09362.

Ren, Z., Qian, K., Dong, F., Ren, Z., Qian, K., Dong, F., Dai, Z., Nejdl, W., Yamamoto, Y., & Schuller, B. W. (2022). Deep attention-based neural networks for explainable heart sound classification. *Machine Learning with Applications*, *9*, 100322.

Renna, F., Oliveira, J., & Coimbra, M. T. (2019). Deep convolutional neural networks for heart sound segmentation. *IEEE Journal of Biomedical and Health Informatics*, *23*(6), 2435–2445.

Reyna, M. A., Kiarashi, Y., Elola, A., Oliveira, J., Renna, F., Gu, A., Perez Alday, E. A., Sadr, N., Sharma, A., Kpodonu, J., Mattos, S., Coimbra, M. T., Sameni, R., Rad, A. B., & Clifford, G. D. (2022). Heart murmur detection from phonocardiogram recordings: The George B. Moody Physionet Challenge 2022. *PLOS Digital Health*, *2*, e0000324.

Sawant, N. K., Patidar, S., Nesaragi, N., & Acharya, U. R. (2021). Automated detection of abnormal heart sound signals using Fano-factor constrained tunable quality wavelet transform. *Biocybernetics and Biomedical Engineering*, *41*(1), 111–126.

Schuller, B., Steidl, S., Batliner, A., Marschik, P. B., Baumeister, H., Dong, F., Hantke, S., Pokorny, F. B., Rathner, E.-M., Bartl-Pokorny, K. D., Einspieler, C., Zhang, D., Baird, A., Amiriparian, S., Qian, K., Ren, Z., Schmitt, M., Tzirakis, P., & Zafeiriou, S. (2018). The INTERSPEECH 2018 computational paralinguistics challenge: Atypical and self-assessed affect, crying and heart beats. In *Proceedings of Interspeech 2018* (pp. 122–126). Universität Augsburg.

Sengur, A., & Turkoglu, I. (2008). A hybrid method based on artificial immune system and fuzzy k-NN algorithm for diagnosis of heart valve diseases. *Expert Systems with Applications*, *35*(3), 1011–1020.

Sharan, T. S., Bhattacharjee, R., Sharma, S., & Sharma, N. (2020). Evaluation of deep learning methods (DnCNN and U-net) for denoising of heart auscultation signals. In *2020 3rd International Conference on Communication System, Computing and IT Applications (CSCITA)* (pp. 151–155). IEEE.

Shen, D., Wu, G., & Suk, H. I. (2017). Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, *19*, 221–248.

Shin, J. Y., L'Yi, S., Jo, D. H., Bae, J. H., & Lee, T. S. (2013). Development of smartphone-based stethoscope system. In *2013 13th International Conference on Control, Automation and Systems (ICCAS 2013)* (pp. 1288–1291). IEEE.

Shuvo, S. B., Ali, S. N., Swapnil, S. I., Al-Rakhami, M. S., & Gumaei, A. (2021). CardioXNet: A novel lightweight deep learning framework for cardiovascular disease classification using heart sound recordings. *IEEE Access*, *9*, 36955–36967.

Singh, S. A., Majumder, S., & Mishra, M. (2019). Classification of short unsegmented heart sound based on deep learning. In *2019 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)* (pp. 1–6). IEEE.

Singh, S. A., Meitei, T. G., & Majumder, S. (2020). Short PCG classification based on deep learning. In *Deep learning techniques for biomedical and health informatics* (pp. 141–164). Elsevier.

Soares, E., Angelov, P., & Gu, X. (2020). Autonomous learning multiple-model zero-order classifier for heart sound classification. *Applied Soft Computing*, *94*, 106449.

Son, G. Y., & Kwon, S. (2018). Classification of heart sound signal using multiple features. *Applied Sciences*, *8*(12), 2344.

Spadaccini, A., & Beritelli, F. (2013). Performance evaluation of heart sounds biometric systems on an open dataset. In *2013 18th International Conference on Digital Signal Processing (DSP)* (pp. 1–5). IEEE.

Sprague, H. B. (1957). History and present status of phonocardiography. *IRE Transactions on Medical Electronics*, *PGME-9*, 2–3.

Springer, D. B., Tarassenko, L., & Clifford, G. D. (2015). Logistic regression-HSMM-based heart sound segmentation. *IEEE Transactions on Biomedical Engineering*, *63*(4), 822–832.

Sun, S., Jiang, Z., Wang, H., & Fang, Y. (2014). Automatic moment segmentation and peak detection analysis of heart sound pattern via short-time modified Hilbert transform. *Computer Methods and Programs in Biomedicine*, *114*(3), 219–230.

Tavel, M. E., Campbell, R. W., & Gibson, M. E. (1973). *Heart sounds and murmurs: An audio visual presentation*. Yearbook Medical Publisher.

Tseng, Y. L., Lin, K. S., & Jaw, F. S. (2016). Comparison of support-vector machine and sparse representation using a modified rule-based method for automated myocardial ischemia detection. *Computational and Mathematical Methods in Medicine*, *2016*, 9460375.

Uğuz, H., Arslan, A., & Türkoğlu, İ. (2007). A biomedical system based on hidden Markov model for diagnosis of the heart valve diseases. *Pattern Recognition Letters*, *28*(4), 395–404.

Varghees, V. N., & Ramachandran, K. (2017). Effective heart sound segmentation and murmur classification using empirical wavelet transform and instantaneous phase for electronic stethoscope. *IEEE Sensors Journal*, *17*(12), 3861–3872.

Vickers, N. J. (2017). Animal communication: When i'm calling you, will you answer too? *Current Biology*, *27*(14), R713–R715.

Wang, W., Yu, X., Fang, B., Zhao, Y., Chen, Y., Wei, W., & Chen, J. (2023). Cross-modality LGE-CMR segmentation using image-to-image translation based data augmentation. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *20*(4), 2367–2375.

Wang, Z., Yang, Z., & Dong, T. (2017). A review of wearable technologies for elderly care that can accurately track indoor position, recognize physical activities and monitor vital signs in real time. *Sensors*, *17*(2), 341.

Whitaker, B. M., Suresha, P. B., Liu, C., Clifford, G. D., & Anderson, D. V. (2017). Combining sparse coding and time-domain features for heart sound classification. *Physiological Measurement*, *38*(8), 1701.

Xiang, M., Zang, J., Wang, J., Wang, H., Zhou, C., Bi, R., Zhang, Z., & Xue, C. (2023). Research of heart sound classification using two-dimensional features. *Biomedical Signal Processing and Control*, *79*, 104190.

Xiao, B., Xu, Y., Bi, X., Zhang, J., & Ma, X. (2020). Heart sounds classification using a novel 1-D convolutional neural network with extremely low parameter consumption. *Neurocomputing*, *392*, 153–159.

Xu, W., Yu, K., Ye, J., Li, H., Chen, J., Yin, F., Xu, J., Zhu, J., Li, D., & Shu, Q. (2022). Automatic pediatric congenital heart disease classification based on heart sound signal. *Artificial Intelligence in Medicine*, *126*, 102257.

Xu, X., Li, C., Fan, X., Lan, X., Lu, X., Ye, X., & Wu, T. (2023). Attention mask R-CNN with edge refinement algorithm for identifying circulating genetically abnormal cells. *Cytometry Part A*, *103*(3), 227–239.

Xu, X., Li, C., Lan, X., Fan, X., Lv, X., Ye, X., & Wu, T. (2023). A lightweight and robust framework for circulating genetically abnormal cells (CACs) identification using 4-color fluorescence in situ hybridization (FISH) image and deep refined learning. *Journal of Digital Imaging*, *36*, 1687–1700.

Xu, X., Ruan, Z., & Yang, L. (2020). Facial expression recognition based on graph neural network. In *2020 IEEE 5th International Conference on Image, Vision and Computing (ICIVC)* (pp. 211–214). IEEE.

Yadav, A., Dutta, M. K., Travieso, C. M., & Alonso, J. B. (2018). Automatic classification of normal and abnormal PCG recording heart sound recording using Fourier transform. In *2018 IEEE International Work Conference on Bioinspired Intelligence (IWOBI)* (pp. 1–9). IEEE.

Yang, C., Zhang, W., Pang, Z., Zhang, J., Zou, D., Zhang, X., Guo, S., Wan, J., Wang, K., & Pang, W. (2021). A low-cost, ear-contactless electronic stethoscope powered by raspberry pi for auscultation of patients with COVID-19: Prototype development and feasibility study. *JMIR Medical Informatics*, *9*(1), e22753.

Yu, D., & Deng, L. (2010). Deep learning and its applications to signal and information processing [exploratory DSP]. *IEEE Signal Processing Magazine*, *28*(1), 145–154.

Yuenyong, S., Nishihara, A., Kongprawechnon, W., & Tungpimolrut, K. (2011). A framework for automatic heart sound analysis without segmentation. *Biomedical Engineering Online*, *10*, 1–23.

Zeng, W., Su, B., Yuan, C., & Chen, Y. (2023). Automatic detection of heart valve disorders using Teager–Kaiser energy operator, rational-dilation wavelet transform and convolutional neural networks with PCG signals. *Artificial Intelligence Review*, *56*(1), 781–806.

Zeng, W., Yuan, J., Yuan, C., Wang, Q., Liu, F., & Wang, Y. (2021). A new approach for the detection of abnormal heart sound signals using TQWT, VMD and neural networks. *Artificial Intelligence Review*, *54*(3), 1613–1647.

Zhang, J., Gajjala, S., Agrawal, P., Tison, G. H., Hallock, L. A., Beussink-Nelson, L., Lassen, M. H., Fan, E., Aras, M. A., Jordan, C. R., Fleischmann, K. E., Melisko, M., Qasim, A., Shah, S. J., Bajcsy, R., & Deo, R. C. (2018). Fully automated echocardiogram interpretation in clinical practice: Feasibility and diagnostic accuracy. *Circulation*, *138*(16), 1623–1635.

Zhang, W., Han, J., & Deng, S. (2017a). Heart sound classification based on scaled spectrogram and partial least squares regression. *Biomedical Signal Processing and Control*, *32*, 20–28.

Zhang, W., Han, J., & Deng, S. (2017b). Heart sound classification based on scaled spectrogram and tensor decomposition. *Expert Systems with Applications*, *84*, 220–231.

Zhang, W., Han, J., & Deng, S. (2019). Abnormal heart sound detection using temporal quasi-periodic features and long short-term memory without segmentation. *Biomedical Signal Processing and Control*, *53*, 101560.

Zhao, Y., Zeng, D., Socinski, M. A., & Kosorok, M. R. (2011). Reinforcement learning strategies for clinical trials in nonsmall cell lung cancer. *Biometrics*, *67*(4), 1422–1433.

Zheng, Y., Guo, X., & Ding, X. (2015). A novel hybrid energy fraction and entropy-based approach for systolic heart murmurs identification. *Expert Systems with Applications*, *42*(5), 2710–2721.

Zheng, Y., Guo, X., Jiang, H., & Zhou, B. (2017). An innovative multi-level singular value decomposition and compressed sensing based framework for noise removal from heart sounds. *Biomedical Signal Processing and Control*, *38*, 34–43.

AUTHOR BIOGRAPHIES

Junxin Chen is currently a Full Professor at the School of Software, Dalian University of Technology, Dalian 116621, China. He received the B. Sc. Degree in Communications Engineering and M.Sc. and Ph.D degrees in Communications and Information System, all from Northeastern University in 2007, 2009, and 2016 respectively. From 2019 to 2020, he was a Postdoc Research Fellow in Department of Computer and Information Science, under the UM Macau Talent Programme (Class A), University of Macau. He worked as an Assistant Professor and Associate Professor in the College of Medicine and Biological Information Engineering, Northeastern University, Shenyang, China. His research interests include Internet of Medical Things, smart health care, artificial intelligence, and information security. He has authored/co-authored over 80+ scientific papers in international peer-reviewed journals and conferences. He has an H-index of 33 and a total of 3200 citations.

Zhihuan Guo was born in Fuxin, China. She received the bachelor's degree in computer science and technology from Shenyang Aerospace University. She received the master's degree in electronic information with Northeastern University, Shenyang, China. Her research interest includes heart sound classification based on artificial intelligence.

Xu Xu was born in Luan, China. He received the B.S. degree in computer science and technology from Anhui Science and Technology University and M.S. degree in information and communication engineering with China Academy of Information and Communications Technology. He is currently pursuing the Ph.D. degree in information and communication engineering with Northeastern University, Shenyang, CN. His research interests include image processing, artificial intelligence, and computer vision.

Gwanggil Jeon received the B.S., M.S., and Ph.D. (summa cum laude) degrees from the Department of Electronics and Computer Engineering, Hanyang University, Seoul, Korea, in 2003, 2005, and 2008, respectively. From 2009.09 to 2011.08, he was with the School of Information Technology and Engineering, University of Ottawa, Ottawa, ON, Canada, as a Post-Doctoral Fellow. From 2011.09 to 2012.02, he was with the Graduate School of Science and Technology, Niigata University, Niigata, Japan, as an Assistant Professor. From 2014.12 to 2015.02 and 2015.06 to 2015.07, he was a Visiting Scholar at Centre de Mathématiques et Leurs Applications (CMLA), École Normale Supérieure Paris-Saclay (ENS-Cachan), France. From 2019 to 2020, he was a Prestigious Visiting Professor at Dipartimento di Informatica, Università degli Studi di Milano Statale, Italy. He is currently a Full Professor at Xidian University, Xi'an, China and at Incheon National University, Incheon, Korea. Hewas a Visiting Professor at Sichuan University, China, Universitat Pompeu Fabra, Barcelona, Spain, Xinjiang University, China, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand, and University of Burgundy, Dijon, France.

David Camacho is full professor at Computer Systems Engineering Department of Universidad Politécnica de Madrid (UPM), and the head of the Applied Intelligence and Data Analysis research group (AIDA: https://aida.etsisi.uam.es) at UPM. He holds a Ph.D.in Computer Science from Universidad Carlos III de Madrid in 2001 with honors (best thesis award in Computer Science). He has published more than 300 journals, books, and conference papers (https://scholar.google.es/citations?hl=en&user=fpf6EDAAAAA). His research interests include Machine Learning (Clustering/Deep Learning), Computational Intelligence (Evolutionary Computation, Swarm Intelligence), Social Network Analysis, Fake News and Disinformation Analysis. He has participated/led more than 50 research projects (Spanish and European: H2020, DG Justice, ISFP, and Erasmus+), related to the design and application of artificial intelligence methods for datamining and optimization for problems emerging in industrial scenarios, aeronautics, aerospace engineering, cybercrime/cyber intelligence, social networks applications, or video games among others.

How to cite this article: Chen, J., Guo, Z., Xu, X., Jeon, G., & Camacho, D. (2024). Artificial intelligence for heart sound classification: A review. *Expert Systems*, *41*(4), e13535. https://doi.org/10.1111/exsy.13535