

**REVIEW**

Machine Learning-Based Intelligent Auscultation Techniques in Congenital Heart Disease: Application and Development

Yang Wang[#], Xun Yang[#], Mingtang Ye, Yuhang Zhao, Runsen Chen, Min Da, Zhiqi Wang, Xuming Mo and Jirong Qi^{*}

Department of Cardiothoracic Surgery, Children's Hospital of Nanjing Medical University, Nanjing, China

*Corresponding Author: Jirong Qi. Email: qjr7@163.com

[#]These two authors contributed equally to this work and should be considered co-first authors

Received: 04 December 2023 Accepted: 28 February 2024 Published: 16 May 2024

ABSTRACT

Congenital heart disease (CHD), the most prevalent congenital ailment, has seen advancements in the “dual indicator” screening program. This facilitates the early-stage diagnosis and treatment of children with CHD, subsequently enhancing their survival rates. While cardiac auscultation offers an objective reflection of cardiac abnormalities and function, its evaluation is significantly influenced by personal experience and external factors, rendering it susceptible to misdiagnosis and omission. In recent years, continuous progress in artificial intelligence (AI) has enabled the digital acquisition, storage, and analysis of heart sound signals, paving the way for intelligent CHD auscultation-assisted diagnostic technology. Although there has been a surge in studies based on machine learning (ML) within CHD auscultation and diagnostic technology, most remain in the algorithmic research phase, relying on the implementation of specific datasets that still await verification in the clinical environment. This paper provides an overview of the current stage of AI-assisted cardiac sounds (CS) auscultation technology, outlining the applications and limitations of AI auscultation technology in the CHD domain. The aim is to foster further development and refinement of AI auscultation technology for enhanced applications in CHD.

KEYWORDS

Congenital heart disease; heart sound auscultation; artificial intelligence; machine learning

1 Introduction

Congenital heart disease (CHD) is the most prevalent congenital ailment, affecting approximately 0.8% of live births [1]. Due to significant endeavors and technological strides, mending the heart has become a viable prospect, resulting in a noteworthy decrease in mortality rates [2]. In recent times, driven by the continuous progression of the “dual-indicator” screening program [3], early-stage diagnoses and interventions for children with CHD have become feasible, consequently elevating survival rates and significantly improving quality of life.

The information age has ushered in widespread adoption of non-invasive detection methods such as cardiac ultrasound and electrocardiogram, coupled with the integration of artificial intelligence (AI) technologies, encompassing machine learning (ML), artificial neural networks (ANN), and deep learning



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

(DL). This synergy has paved the way for early diagnosis and treatment modalities tailored for children with CHD. Notably, AI technology breakthroughs have facilitated the digital acquisition, storage, and analysis of heart sound signals [4], consequently giving rise to intelligent auscultation techniques for CHD sounds [5,6].

AI auscultation technology predominantly engages in the acquisition of heart sound signals, encompassing denoising, segmentation, feature extraction, and classification and recognition [5,7,8]. This manuscript provides a comprehensive overview of the recent evolution of intelligent heart sound auscultation technology. It meticulously outlines the present status and challenges associated with the application of AI auscultation-assisted decision-making technology in CHD screening. The goal is to promote the ongoing advancement and refinement of AI auscultation technology within the field of CHD applications.

2 CS and Auscultation

Cardiac sounds (CS), originating from the rhythmic dance between systole and diastole, manifest as distinct auditory patterns recorded as phonocardiograms (PCG). PCG are quasi-periodic, and the period of each cardiac signal is called the cardiac cycle. Depending on the order in which they occur in the cardiac cycle, CS can be categorized into four segments: the first heart sound (S₁), systolic, the second heart sound (S₂) and diastolic (Fig. 1) [9]. Beyond the symphony of S₁ and S₂, the auditory landscape may include cardiac murmurs, aberrant echoes arising from the vibrational interplay of ventricular walls, valves, or blood vessels. These murmurs, characterized by prolonged duration, diverse frequency, and amplitude variations, accompany CS during systole or diastole [4]. Analyzing raw PCG data provides valuable insights, serving as an ancillary diagnostic tool that contributes to the predictive assessment of cardiovascular diseases [4,10].

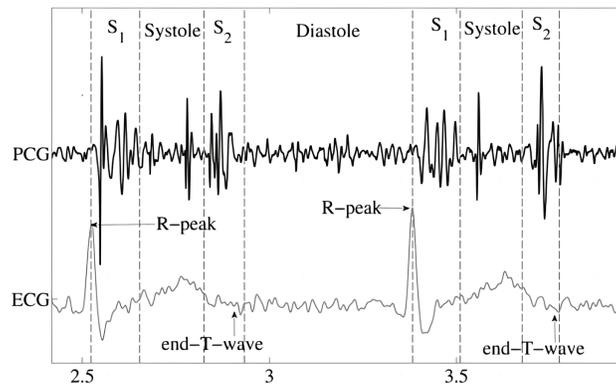


Figure 1: Phonocardiogram (PCG) [9]

The anatomical position of the heart valves relative to the chest wall determines the optimal auscultation position (Fig. 2). Therefore, the stethoscope should be placed in the following positions [11] for auscultation:

- Aortic valve I* (1): second intercostal space, right sternal border;
- Pulmonary valve* (2): second intercostal space, left sternal border;
- Aortic valve II* (3): third intercostal space, left sternal border;
- Tricuspid valve* (4): left lower sternal border;
- Mitral valve* (5): fifth intercostal space, midclavicular line(cardiac apex).

Additional CS is typically associated with murmurs, clicks, and snaps [4]. The genesis of cardiac murmurs holds paramount relevance to cardiovascular ailments [2], such as septal defects, unclosed

ductus arteriosus, pulmonary hypertension, or cardiac valve deficiencies [4,12]. Deciphering murmurs directly impacts a physician's diagnostic skills. Most cardiac murmurs correlate intimately with specific pathologies, and auscultation serves as the linchpin for recognizing and distinguishing these features. The ability to precisely characterize murmurs can ascertain whether a subject should be referred to a cardiac specialist [4], enhancing subsequent diagnostic examinations. Cardiac auscultation is a pivotal screening tool [9]; when used judiciously, it expedites treatment/referral processes, thereby improving patient prognosis and quality of life [10]. Insufficient auscultation expertise, non-digitized preservation of auscultation data, and subjective judgment criteria have emerged as pivotal factors hindering primary care physicians in conducting early CHD screenings [4,9,13].

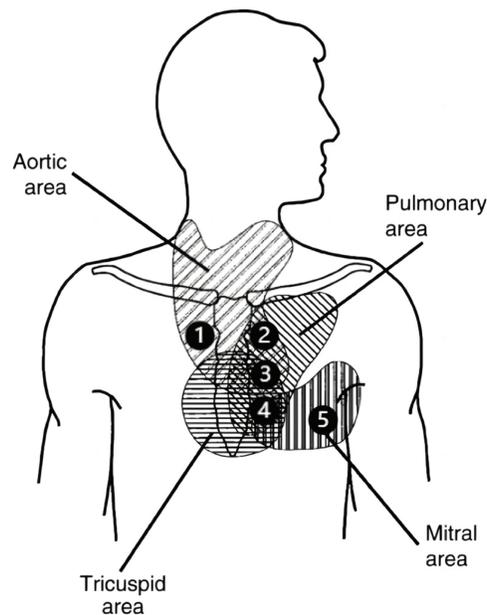


Figure 2: Schematic representation of the auscultation area in the prethoracic region (image adapted from [11]); 1 = right second intercostal space; 2 = left second intercostal space; 3 = left third intercostal space; 4 = left fourth/fifth intercostal space; 5 = fifth intercostal space (midclavicular line)

3 Research Status of AI Auscultation Techniques

AI-assisted detection technology is a swift, efficient, and cost-effective tool [10,14], applicable for the quantitative analysis of CS signals. By extracting key parameters from the PCG, it aims to yield more intuitive diagnostic results, facilitating the inference of potential cardiovascular ailments [14]. AI auscultation technology primarily encompasses denoising, segmentation, feature extraction, and classification recognition of PCG signals [5,7,8].

Influenced by the external environment, heart sound signals inevitably intertwine with electromagnetic interference, random noise, respiratory sounds, lung sounds, etc., during the acquisition process [15]. The diagnostic accuracy of CS diagnosis is directly affected by the quality of the signal and the subsequently extracted features. Therefore, noise reduction is necessary before signal analysis. Some widely studied and applied noise reduction methods for CS include discrete wavelet transform (DWT) [15,16], empirical mode decomposition (EMD) [17], and singular value decomposition (SVD) [15,18]. Noise reduction of CS signals using a combination of methods can achieve better results and help to improve the signal quality and detection accuracy. Mondal et al. [15] proposed a CS noise reduction method based on the combination of DWT and SVD, where the coefficients corresponding to the selected nodes are processed

by SVD in order to achieve the suppression of the noise component in the CS signal. Some interference signals (lung sounds, environmental sounds, etc.) may have more obvious frequency aliasing with the signal, which makes the noise reduction method based on the frequency decomposition of the signal difficult to apply, and greatly affects the extraction of CHD features. To address this, Shah et al. [19] decomposed the time-frequency spectra of single-channel mixed CS signals into non-negative matrix factorisation (NMF) parts related to heart or lung sounds. They then reconstructed the signal to obtain the source signals, achieving noise reduction and CS feature extraction.

Segmentation of CS is a prerequisite for feature extraction in CS classification, aiming to determine the positional intervals of S1, systole, S2, and diastole in the CS signal. This segmentation is advantageous for extracting the characteristics of heart murmurs from each CS state separately, thereby utilizing the rich pathological information within murmurs for CHD analysis and diagnosis. In response to this, numerous methods for heart sound segmentation have been proposed by experts and scholars [20–22], achieving commendable segmentation results. However, for some abnormal heart sounds affected by pathological murmurs and severe noise interference, the structure of the cardiac cycle is irregular, making it challenging to precisely delineate the cardiac cycle [21]. Segmentation methods lacking temporal structural information [22–24] often struggle to ensure that different fixed-duration segments of heart sound or raw heart sound sequences share the same starting point, consequently affecting recognition performance. Table 1 summarizes some of the recent literature on heart sound segmentation [9,20–23].

Table 1: Segmentation methods for acoustic cardiogram signals [9,20–23]

Year	Authors	Segmentation method	Heart sound data source	Results
2022	Guo et al. [20]	Proposed a novel method, C-LSTM-A (Convolution and Bidirectional Long-Short Term Memory neural network with Attention mechanism), which combines one-dimensional convolution with attention mechanism to accurately identify high-speed signals	Yunnan fuwai cardiovascular hospital heart sound dataset (700 normal signals and 450 pathological signals)/ PhysioNet CinC challenges dataset	Accuracy 97.00%, F1 96.8%
2020	Fernando et al. [21]	Developed a heart sound segmentation method using a bidirectional long short-term memory network (LSTM) with attention mechanism: focusing on the signal components of heart sound states through attention mechanism, combined with recurrent neural network (RNN) to learn features from irregular heart sounds, improving segmentation effectiveness	PhysioNet, M3dicine human heart sound database (M3-HU), and M3dicine animal heart sound database (M3-AN)	Sensitivity 97.20%, specificity 97.50%, F1 96.70%

(Continued)

Table 1 (continued)				
Year	Authors	Segmentation method	Heart sound data source	Results
2019	Renna et al. [22]	Utilized deep convolutional neural networks (CNN) for heart sound segmentation: combined with a sequential time modeling process to provide segmentation results compatible with the sequential nature of heart sounds; also incorporated into segmentation algorithms using HMM and HSMM as an effective emission distribution modeling method	PhysioNet CinC challenges dataset	Sensitivity 93.90%, positive predictive value (PPV) 94.00%
2017	Varghees et al. [23]	Applied Empirical Wavelet Transform (EWT) and Instantaneous Phase Waveform (IPW) to segment heart sound signals	Physionet, pascal, michigan, eGeneralMedical, and real-time PCG signals	Specificity 94.38%, Pp 97.25%, overall accuracy (OA) 91.92%
2016	Springer et al. [9]	Extracted features such as homomorphic envelope, power spectral density envelope, etc., after preprocessing of heart sound signals. Utilized logistic regression hidden semi-Markov model (LRHSMM) to learn the above features and segment S1 and S2 in noisy heart sounds	From massachusetts general hospital in the united states and home monitoring of mitral valve prolapse patients: a total of 123 cases (38 control group, 37 mitral valve prolapse patients, 36 benign murmurs, 5 aortic valve diseases, 7 other heart diseases)	Sensitivity 95.34%, positive predictive value (P+) 95.92%, F1 95.63

Before classifying CS, a few representative features from the CS category must be extracted from segmented signals to replace the high-dimensional original signals [8]. In general, classification models trained on features are more efficient and accurate than those trained on raw signals. Commonly used methods for CS feature extraction include DWT [25], short-time fourier transform (STFT) [26], mel-frequency cepstral coefficient (MFCC) [27,28] and variational mode decomposition (VMD) [29]. After extracting the features, the classification algorithm classifies the extracted features to achieve the screening and auxiliary diagnosis of CHD. CS classification problems come in various forms, and the choice of classification tasks can significantly impact the results [30]. Due to the rapid development of AI in medicine, ML-based classification algorithms have been widely researched and applied in the field of CHD intelligent auscultation [26,31].

4 Development and Application of ML in Intelligent Auscultation

Based on the principle of ML algorithms, CS classification algorithms can be divided into two categories: classification algorithms based on traditional ML and classification algorithms based on DL [32,33]. We provide a schematic of the principles of these two algorithms, as shown in Fig. 3. Among

them, traditional ML-based classification algorithms mainly involve manually designing and extracting effective CS features [31,32], followed by classification using classifiers. Classification algorithms based on DL refer to the algorithms that, with respect to the waveform map of the CS signals or extracted features in the time-frequency domain, use deep neural networks (DNN) to automatically extract more high-level abstract features and achieve classification [33,34].

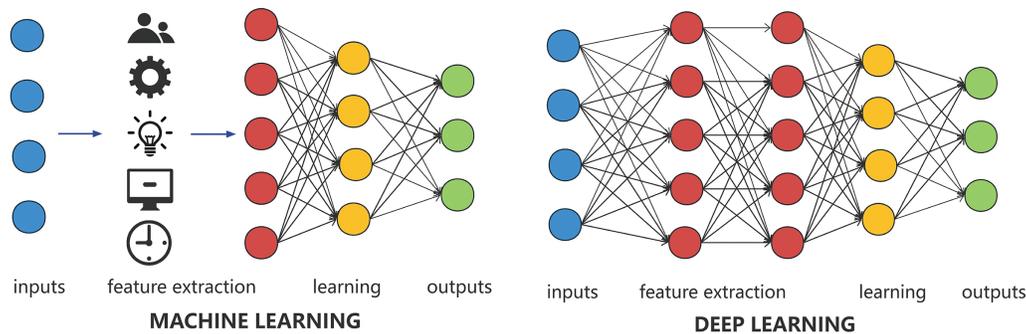


Figure 3: Schematic of machine learning (ML) and deep learning (DL)

4.1 Traditional ML Intelligent Auscultation

Traditional ML-based classification algorithms involve manually designing effective CS features. This enables computer systems to access and analyze data effectively, adjusting and enhancing functionalities based on patterns and experience, without the need for explicit programming. Commonly used classification algorithms include support vector machine (SVM) [35], decision trees [36], and MFCC [25] models. Deng et al. [35] utilized DWT to compute the envelope from sub-band coefficients of CS signals. They extracted autocorrelation features from the envelope, fused these features, resulting in a unified feature representation with a diffusion mapping, and eventually fed this unified feature into an SVM classifier to accomplish CS classification. Yaseen et al. [25] suggested using MFCC features combined with DWT features for CS classification. They compared SVM with the K-nearest neighbors (KNN) algorithm based on centroid displacement and DNN. The results indicated that SVM achieved optimal classification performance with an average accuracy of 97.9%.

In the domain of intelligent CHD auscultation-assisted diagnosis, Aziz et al. [32] aimed to achieve a three-class classification for atrial septal defect, ventricular septal defect, and normal CS. They proposed a method for extracting MFCC features from CS signals, combined with one-dimensional Local Ternary Patterns (1D-LTP) features to realize CHD classification. In the training phase, this model employed an SVM classifier, achieving a final average accuracy of 95.24%. Zhu et al. [37], using a backpropagation neural network as a classifier, analyzed and extracted features from CHD CS using both MFCC and Linear Predictive Cepstral Coefficients (LPCC). The final results indicated that MFCC features (specificity: 93.02%, sensitivity: 88.89%) outperformed LPCC features (specificity: 86.96%, sensitivity: 86.96%).

Traditional ML-based algorithms depend on manually designing effective CS features. Suboptimal feature selection can result in poor algorithm performance, underscoring the critical importance of proper feature selection. However, designing and selecting effective CS features for various pathological conditions in different environments poses exceptional challenges. This difficulty limits the further optimization of classification algorithms based on traditional ML [32]. With the increasing incidence of cardiovascular diseases, the volume of CS data to be processed is also on the rise [14]. To handle large datasets while ensuring classification accuracy, deep learning algorithms have emerged.

4.2 DL Intelligent Auscultation

DL is a collective term for ML techniques that leverage neural networks [34]. It involves constructing a multi-layered deep neural learning network by layering neural networks more deeply than traditional ML. This allows machines to autonomously recognize, extract, and express the more intrinsic features hidden within a database or dataset, thereby enhancing the accuracy and efficiency of classification. In 2006, Hinton et al. [38] first introduced the concept of DL, leveraging spatial relationships to combine low-level models into more complex high-level models, significantly improving system training performance. In recent years, DL algorithms have demonstrated practicality and reliability in various fields such as image recognition [39], biomedical data analysis [40,41], and signal processing [42]. DL models have been applied to the classification of CS signals, including DNN, convolutional neural networks (CNN), recurrent neural networks (RNN), and long short-term memory (LSTM) [10,32,37].

Among various artificial neural network models, CNN models excel in converting raw data into a two-dimensional matrix format and extracting image feature values, surpassing traditional ANN [43]. Thanks to operations like convolution and pooling, CNN models can construct deep CNN models by stacking multiple convolutional and pooling layers, achieving outstanding performance in computer vision tasks (images, video data, etc.) (Fig. 4) [40–42]. Kui et al. [44] combined log mel-frequency spectral coefficients (MFSC) with CNN for CS classification, achieving a binary classification accuracy of 93.89% and a multiclass accuracy of 86.25%. Krishnan et al. [45] used one-dimensional CNN and feedforward neural networks to classify unsegmented CS, achieving an average accuracy of 85.7%. Li et al. [46] extracted deep features of CS through denoising autoencoders and utilized one-dimensional CNN as a classifier, resulting in a classification accuracy of 99%. However, the study extended the original dataset, allowing segments of the same record to appear in both the training and test sets, leading to higher-than-realistic classification accuracy, which does not align with the clinical application environment. In research investigating the impact of different DL models on heart sound classification, Deng et al. [30] utilized MFCC features combined with first and second-order differential features as network inputs for CS classification. They compared various network models, and the results indicated that CNN has the optimal classification accuracy (98.34%).

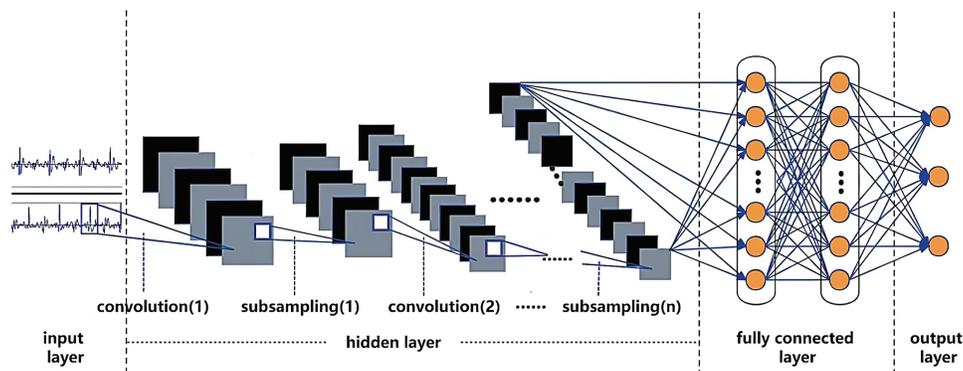


Figure 4: Schematic diagram of deep convolutional neural network (CNN)

5 Current Status of Intelligent Auscultation Technology in CHD

5.1 Automated Screening and Diagnosis

Intelligent cardiac auscultation technology enables the efficient automated analysis and identification of anomalies in cardiac audio signals, such as murmurs and arrhythmias, providing supplementary diagnostic information for physicians. This aids doctors in conducting timely, precise, and efficient screening and

diagnosis of CHD and facilitates prompt intervention when necessary. Xiao et al. [47] curated a pediatric CS dataset comprising 528 high-quality recordings from 137 subjects, including newborns and children, annotated by medical professionals. They employed two lightweight CNN models to develop a DL-based computer-aided diagnosis system for CHD. In a similar vein, Xu et al. [48] established a pediatric-CHD-CS database encompassing 941 CS signals. Utilizing classifiers based on Random Forest and Adboost, they achieved accurate classification of CHD. The outcomes demonstrated a sensitivity, specificity, and F1 score of 0.946, 0.961, and 0.953, respectively, indicating the effectiveness of this approach in CHD classification.

5.2 Remote Monitoring and Comprehensive Management

Intelligent cardiac auscultation devices, integrated with internet technology, facilitate the remote transmission and monitoring of cardiac audio signals. This capability proves highly advantageous for patients requiring long-term cardiac health monitoring, enabling real-time observation and recording of changes in cardiac audio signals to promptly detect and address any anomalies. Simultaneously, given the evolving demographic landscape, the concepts of prevention, control, and treatment for CHD need to extend into adulthood. It is imperative to implement tertiary prevention and comprehensive lifecycle management for individuals with CHD. In this regard, Chowdhury et al. [49] proposed a prototype model of an intelligent digital stethoscope system designed to monitor patients' CS and diagnose any abnormalities in real-time. The system captures patients' CS and wirelessly transmits them to a personal computer, visualizing the CS and achieving a binary classification of normal and abnormal CS. The accuracy of the optimized integrated algorithm, following adjustments, reached 97% for abnormal CS and 88% for normal CS.

5.3 Medical Education and Training

The intelligent cardiac auscultation technology proves instrumental in medical education and training, aiding medical students and practitioners in comprehending the intricacies of cardiac audio signals and diagnostic methodologies. While echocardiography currently stands as the primary modality for diagnosing CHD, cardiac auscultation remains a pivotal approach for both screening and diagnosis. Nevertheless, auscultation heavily relies on the subjective expertise of physicians, leading to substantial variations in diagnoses among different medical professionals. The absence of senior cardiac specialists in primary healthcare units, coupled with the inadequacy of auscultation skills among young interns [50], hinders the widespread adoption and application of auscultation screening for CHD. Through intelligent auscultation devices and related teaching software, medical students and doctors can engage in virtual auscultation and simulated diagnosis, enhancing their auscultation skills and diagnostic abilities. Therefore, the development of computer-aided auscultation algorithms and systems is an effective approach for CHD screening and diagnosis, and can also be utilized for retrospective teaching and clinical education [48].

6 Challenges and Issues in Intelligent Auscultation for CHD

6.1 Preprocessing of CS

During the data collection process, CS signals inevitably encounter interference from environmental noise, equipment artifacts, and other physiological organ noises. To mitigate these disturbances, currently prevalent methods for CS denoising include approaches based on Discrete Wavelet Transform (DWT) [15,16] and Empirical Mode Decomposition (EMD) [17]. Given the stochastic, non-stationary, and chaotic nature of interference noise in real-world collection environments, precise differentiation between CS signals and noise proves challenging for both DWT and EMD. The acquisition of an accurate adaptive threshold function becomes particularly elusive. Solely relying on spatial transformations and linear filtering struggles to entirely segregate noise from CS, leading to potential distortions in the CS. In

the majority of existing CS classification methods, segmenting CS serves as the fundamental step for feature extraction and classification. Researchers commonly adopt fixed or unfixed durations of raw CS sequences as study subjects [33,35,51], employing segmenting approaches lacking temporal structural information. Inevitably, this results in varying time shifts, mismatches with the cardiac cycle structure, and consequent impacts on recognition performance.

6.2 CS Data and Datasets

A fundamental step in developing a computer-aided decision system for cardiovascular disease screening through cardiac auscultation involves the collection of extensive annotated CS datasets [4]. These datasets should accurately represent and characterize the murmurs and anomalies of different cardiovascular disease patients. The detection of CS signals poses significant challenges due to the inherent characteristics of CS and the influence of environmental noise [4]. On one hand, the stochastic and variable nature of CHD symptoms results in the complexity and diversity of signal manifestations. On the other hand, CS signals are relatively weak, and the collection process of raw signals is susceptible to various noises and disturbances, thereby reducing the accuracy of relevant parameter extraction and increasing diagnostic uncertainty. The majority of CHD patients are minors, exhibiting significant differences in heart rate, body size, and cooperation levels [48]. Consequently, CS classification algorithms designed for adults cannot be directly applied to CHD screening and treatment. Challenges such as limited data volume, incomplete labels, and a lack of standardized recognition criteria have persistently impeded the improvement and testing of intelligent cardiac auscultation algorithms for CHD. The robustness and generalizability of these algorithms in clinical screening applications still require thoughtful consideration [52,53].

6.3 CS Algorithms and Classification Models

The predominant origin of the existing CS dataset from adults [4] results in a concentration of current algorithms and classification models on the screening and diagnosis of cardiac valve and major vascular diseases in adults. Differences in pediatric disease types, heart rate, and CS features mean that existing cardiac sound segmentation algorithms and classification models cannot be directly applied to pediatric CHD auxiliary auscultation decision systems [48]. Classification models based on compressed DL algorithms are more accurate compared to models based on traditional algorithms [47], and future research should focus on optimizing DL model structures, reducing feature parameter quantities, and improving model generalization performance. Despite DL models optimizing parameters through error backpropagation, their diagnostic results in clinical applications lack clear interpretability regarding the correlation or causation between output diagnoses and input data [48]. Therefore, seeking interpretability in models is also a significant challenge for current DL technology to achieve clinical application.

6.4 CS-Decision Systems and Clinical Applications

The CS-decision system requires an elevated level of robustness [52] in practical applications to ensure steadfast performance across diverse environments and conditions. Considering the susceptibility of cardiac sound data to surrounding noise interference [15], algorithms need to effectively counteract these disruptions to ensure reliability in practical clinical applications. The clinical application of the cardiac sound decision system faces issues related to practical operation and acceptance. System developers need to consider the system's integrative nature, user-friendliness, and the training requirements for healthcare personnel to ensure its effectiveness in real medical scenarios. The development and clinical application of cardiac sound decision systems involve collaboration across multiple disciplines, including medicine, engineering, computer science, requiring close cooperation between clinical physicians, researchers, and engineers to promote the broader service of CHD cardiac auscultation-assisted diagnostic systems in the medical field [14].

7 Conclusions

With the rapid evolution of AI, the realization of intelligent auscultation-assisted diagnostic technology for cardiovascular diseases has become achievable. In recent years, intelligent auscultation techniques based on ML models have made progress in the field of CHD. However, their validation in clinical settings is yet to be completed. Existing issues with practicality, consistency, and interpretability, coupled with a lack of mature commercial products, hamper their validation and dissemination in clinical CHD screening. Consequently, it is imperative for relevant medical institutions to convene experts for the collaborative formulation of a consensus and standardized procedure for auscultatory screening of congenital heart diseases. Building upon this foundation, the establishment of a comprehensively annotated pediatric CHD auscultation database is essential. Simultaneously, research institutions should dedicate efforts to further advance and enhance auscultation algorithms while ensuring the interpretability of the classification models. This endeavor aims to elevate the precision and diagnostic efficiency of healthcare professionals in the screening of congenital heart disease, providing comprehensive support for the lifecycle management of CHD-afflicted children. Simultaneously, the widespread application of this technology has the potential to expedite the establishment of dedicated networks for the screening, diagnosis, and monitoring of CHD.

Acknowledgement: Not applicable.

Funding Statement: This article was supported by Jiangsu Provincial Health Commission (Grant No. K2023036).

Author Contributions: Yang Wang, Xun Yang: responsible for drafting and chart to make; Mingtang Ye, Yuhang Zhao: draft modifies and charts; Runsen Chen, Min Da, Zhiqi Wang: responsible for the revision and polishing of the article; Xuming Mo, Jirong Qi: responsible for the proofreading and guidance of the article. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: Data availability is not applicable to this article as no new data were created or analyzed in this study.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. van der Linde D, Konings EEM, Slager MA, Witsenburg M, Helbing WA, Takkenberg JJM, et al. Birth prevalence of congenital heart disease worldwide. *J Am Coll Cardiol.* 2011;58(21):2241–7. doi:10.1016/j.jacc.2011.08.025.
2. Bouma BJ, Mulder BJM. Changing landscape of congenital heart disease. *Circ Res.* 2017;120(6):908–22. doi:10.1161/CIRCRESAHA.116.309302.
3. Zhao QM, Ma XJ, Ge XI, Liu F, Yan WI, Wu L, et al. Pulse oximetry with clinical assessment to screen for congenital heart disease in neonates in China: a prospective study. *Lancet.* 2014;384(9945):747–54. doi:10.1016/S0140-6736(14)60198-7.
4. Oliveira J, Renna F, Costa PD, Nogueira M, Oliveira C, Ferreira C, et al. The CirCor DigiScope dataset: from murmur detection to murmur classification. *IEEE J Biomed Health Inform.* 2022;26(6):2524–35. doi:10.1109/JBHI.2021.3137048.
5. Clifford G, Liu C, Moody B, Millet J, Schmidt S, Li Q, et al. Recent advances in heart sound analysis. *Physiol Meas.* 2017;38(8):E10–E25. doi:10.1088/1361-6579/aa7ec8.
6. Cheng J, Sun K. Heart sound classification network based on convolution and transformer. *Sens.* 2023;23(19): 8168. doi:10.3390/s23198168.

7. Bozkurt B, Germanakis I, Stylianou Y. A study of time-frequency features for CNN-based automatic heart sound classification for pathology detection. *Comput Biol Med.* 2018;100:132–43. doi:10.1016/j.compbimed.2018.06.026.
8. Chakir F, Jilbab A, Nacir C, Hammouch A. Phonocardiogram signals processing approach for PASCAL classifying heart sounds challenge. *Signal Image Video Process.* 2018;12(6):1149–55. doi:10.1007/s11760-018-1261-5.
9. Springer D, Tarassenko L, Clifford G. Logistic regression-HSMM-based heart sound segmentation. *IEEE T Bio-Med Eng.* 2016;63(4):822–32.
10. Huai X, Kitada S, Choi D, Siriaraya P, Kuwahara N, Ashihara T. Heart sound recognition technology based on convolutional neural network. *Inform Health Soc Care.* 2021;46(3):320–32. doi:10.1080/17538157.2021.1893736.
11. Pelech AN. The physiology of cardiac auscultation. *Pediatr Clin North Am.* 2004;51(6):1515–35. doi:10.1016/j.pcl.2004.08.004.
12. Riknagel D, Zimmermann H, Farlie R, Hammershoi D, Schmidt SE, Hedegaard M, et al. Separation and characterization of maternal cardiac and vascular sounds in the third trimester of pregnancy. *Int J Gynaecol Obstet.* 2017;137(3):253–9. doi:10.1002/ijgo.2017.137.issue-3.
13. Goethe Doualla FC, Bediang G, Nganou-Gnindjio C. Evaluation of a digitally enhanced cardiac auscultation learning method: a controlled study. *BMC Med Educ.* 2021;21(1):380. doi:10.1186/s12909-021-02807-4.
14. Li S, Li F, Tang S, Xiong W. A review of computer-aided heart sound detection techniques. *BioMed Res Int.* 2020;2020:5846191.
15. Mondal A, Saxena I, Tang H, Banerjee P. A noise reduction technique based on nonlinear kernel function for heart sound analysis. *IEEE J Biomed Health Inform.* 2018;22(3):775–84. doi:10.1109/JBHI.2017.2667685.
16. Gradolewski D, Magenes G, Johansson S, Kulesza W. A wavelet transform-based neural network denoising algorithm for mobile phonocardiography. *Sens.* 2019;19(4):957. doi:10.3390/s19040957.
17. Vican I, Kreković G, Jambrošić K. Can empirical mode decomposition improve heartbeat detection in fetal phonocardiography signals? *Comput Methods Programs Biomed.* 2021;203:106038. doi:10.1016/j.cmpb.2021.106038.
18. Zheng Y, Guo X, Jiang H, Zhou B. An innovative multi-level singular value decomposition and compressed sensing based framework for noise removal from heart sounds. *Biomed Signal Process Control.* 2017;38:34–43. doi:10.1016/j.bspc.2017.04.005.
19. Shah G, Koch P, Papadias CB. On the blind recovery of cardiac and respiratory sounds. *IEEE J Biomed Health Inform.* 2015;19(1):151–7. doi:10.1109/JBHI.2014.2349156.
20. Guo Y, Yang H, Guo T, Pan J, Wang W. A novel heart sound segmentation algorithm via multi-feature input and neural network with attention mechanism. *Biomed Phys Eng Express.* 2022;9(1):015012.
21. Fernando T, Ghaemmaghami H, Denman S, Sridharan S, Hussain N, Fookes C. Heart sound segmentation using bidirectional LSTMs with attention. *IEEE J Biomed Health Inform.* 2020;24(6):1601–9. doi:10.1109/JBHI.6221020.
22. Renna F, Oliveira J, Coimbra MT. Deep convolutional neural networks for heart sound segmentation. *IEEE J Biomed Health Inform.* 2019;23(6):2435–45. doi:10.1109/JBHI.6221020.
23. Nivitha Varghees V, Ramachandran KI. Effective heart sound segmentation and murmur classification using empirical wavelet transform and instantaneous phase for electronic stethoscope. *IEEE Sens J.* 2017;17(12):3861–72. doi:10.1109/JSEN.2017.2694970.
24. Papadaniil CD, Hadjileontiadis LJ. Efficient heart sound segmentation and extraction using ensemble empirical mode decomposition and kurtosis features. *IEEE J Biomed Health Inform.* 2014;18(4):1138–52. doi:10.1109/JBHI.2013.2294399.
25. Yaseen, Son GY, Kwon S. Classification of heart sound signal using multiple features. *Appl Sci.* 2018;8(12):106038.
26. Debbal SM, Bereksi-Reguig F. Computerized heart sounds analysis. *Comput Biol Med.* 2008;38(2):263–80. doi:10.1016/j.compbimed.2007.09.006.

27. Kay E, Agarwal A. DropConnected neural networks trained on time-frequency and inter-beat features for classifying heart sounds. *Physiol Meas*. 2017;38(8):1645–57. doi:10.1088/1361-6579/aa6a3d.
28. Nogueira DM, Ferreira CA, Gomes EF, Jorge AM. Classifying heart sounds using images of motifs, MFCC and temporal features. *J Med Syst*. 2019;43(6):168. doi:10.1007/s10916-019-1286-5.
29. Zeng W, Yuan J, Yuan C, Wang Q, Liu F, Wang Y. A new approach for the detection of abnormal heart sound signals using TQWT, VMD and neural networks. *Artif Intell Rev*. 2020;54(3):1613–47.
30. Deng M, Meng T, Cao J, Wang S, Zhang J, Fan H. Heart sound classification based on improved MFCC features and convolutional recurrent neural networks. *Neural Netw*. 2020;130:22–32. doi:10.1016/j.neunet.2020.06.015.
31. Fuadah YN, Pramudito MA, Lim KM. An optimal approach for heart sound classification using grid search in hyperparameter optimization of machine learning. *Bioeng*. 2022;10(1):45.
32. Aziz S, Khan MU, Alhaisoni M, Akram T, Altaf M. Phonocardiogram signal processing for automatic diagnosis of congenital heart disorders through fusion of temporal and cepstral features. *Sens*. 2020;20(13):3790. doi:10.3390/s20133790.
33. Potes C, Parvaneh S, Rahman A, Conroy B. Ensemble of feature: based and deep learning: based classifiers for detection of abnormal heart sounds. In: 2016 Computing in Cardiology Conference (CinC), 2016; Vancouver, British Columbia, Canada.
34. Chen W, Sun Q, Chen X, Xie G, Wu H, Xu C. Deep learning methods for heart sounds classification: a systematic review. *Entropy*. 2021;23(6):e0276264.
35. Deng SW, Han JQ. Towards heart sound classification without segmentation via autocorrelation feature and diffusion maps. *Future Gener Comput Syst*. 2016;60:13–21. doi:10.1016/j.future.2016.01.010.
36. Deo RC. Machine learning in medicine. *Circ*. 2015;132(20):1920–30. doi:10.1161/CIRCULATIONAHA.115.001593.
37. Zhu LL, Pan JH, Shi JH, Yang HB, Wang WL. Research on recognition of CHD heart sound using MFCC and LPCC. *J Phys Conf Ser*. 2019;1169:012011. doi:10.1088/1742-6596/1169/1/012011.
38. Hinton G, Osindero S, Welling M, Teh Y. Unsupervised discovery of nonlinear structure using contrastive backpropagation. *Cogn Sci*. 2006;30(4):725–31. doi:10.1207/s15516709cog0000_76.
39. Yu X, Pang W, Xu Q, Liang M. Mammographic image classification with deep fusion learning. *Sci Rep*. 2020;10(1):14361. doi:10.1038/s41598-020-71431-x.
40. Alipanahi B, DeLong A, Weirauch MT, Frey BJ. Predicting the sequence specificities of DNA- and RNA-binding proteins by deep learning. *Nat Biotechnol*. 2015;33(8):831–8. doi:10.1038/nbt.3300.
41. Shen D, Wu G, Suk HI. Deep learning in medical image analysis. *Annu Rev Biomed Eng*. 2017;19(1):221–48. doi:10.1146/bioeng.2017.19.issue-1.
42. Yu D, Deng L. Deep learning and its applications to signal and information processing [exploratory DSP]. *IEEE Signal Process Mag*. 2011;28(1):145–54. doi:10.1109/MSP.2010.939038.
43. Li F, Zhang Z, Wang L, Liu W. Heart sound classification based on improved mel-frequency spectral coefficients and deep residual learning. *Front Physiol*. 2022;13:102257.
44. Kui H, Pan J, Zong R, Yang H, Wang W. Heart sound classification based on log Mel-frequency spectral coefficients features and convolutional neural networks. *Biomed Signal Process Control*. 2021;69:102893. doi:10.1016/j.bspc.2021.102893.
45. Krishnan PT, Balasubramanian P, Umopathy S. Automated heart sound classification system from unsegmented phonocardiogram (PCG) using deep neural network. *Phys Eng Sci Med*. 2020;43(2):505–15. doi:10.1007/s13246-020-00851-w.
46. Li F, Liu M, Zhao Y, Kong L, Dong L, Liu X, et al. Feature extraction and classification of heart sound using 1D convolutional neural networks. *EURASIP J Adv Signal Process*. 2019;2019(1):102893.
47. Xiao B, Xu Y, Bi X, Li W, Ma Z, Zhang J, et al. Follow the sound of children’s heart: a deep-learning-based computer-aided pediatric CHDs diagnosis system. *IEEE Internet Things J*. 2020;7(3):1994–2004. doi:10.1109/JIoT.6488907.

48. Xu W, Yu K, Ye J, Li H, Chen J, Yin F, et al. Automatic pediatric congenital heart disease classification based on heart sound signal. *Artif Intell Med.* 2022;126:1084420.
49. Chowdhury MEH, Khandakar A, Alzoubi K, Mansoor S, Tahir M, Reaz A, et al. Real-time smart-digital stethoscope system for heart diseases monitoring. *Sens.* 2019;19(12):14361.
50. Leng S, Tan RS, Chai KTC, Wang C, Ghista D, Zhong L. The electronic stethoscope. *Biomed Eng Online.* 2015;14(1):012011.
51. Dominguez-Morales JP, Jimenez-Fernandez AF, Dominguez-Morales MJ, Jimenez-Moreno G. Deep neural networks for the recognition and classification of heart murmurs using neuromorphic auditory sensors. *IEEE Trans Biomed Circuits Syst.* 2018;12(1):24–34. doi:10.1109/TBCAS.2017.2751545.
52. Kumar A, Wang W, Yuan J, Wang B, Fang Y, Zheng Y, et al. Robust classification of heart valve sound based on adaptive EMD and feature fusion. *PLoS One.* 2022;17(12):e0276264. doi:10.1371/journal.pone.0276264.
53. Oliveira J, Nogueira D, Ferreira C, Jorge A, Coimbra M. The robustness of random forest and support vector machine algorithms to a faulty heart sound segmentation. *IEEE EMBC.* 2022;2022:1989–92.