Received 31 March 2024, accepted 25 April 2024, date of publication 30 April 2024, date of current version 4 June 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3395389

RESEARCH ARTICLE

Precision Diagnosis: An Automated Method for Detecting Congenital Heart Diseases in Children From Phonocardiogram Signals Employing Deep Neural Network

HEND KHALID ALKAHTANI[1], (Member, IEEE), IHTISHAM UL HAQ[2], YAZEED YASIN GHADI[3], NISREEN INNAB[4], MASOUD ALAJMI[5], (Member, IEEE), AND MEKEBAYEV NURBAPA[6]

[1]Department of Information Systems, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, Riyadh 11671, Saudi Arabia
[2]Department of Information and Communication Technology, University of Calabria, 87036 Rende, Italy
[3]Department of Computer Science and Software Engineering, Al Ain University, Abu Dhabi, United Arab Emirates
[4]Department of Computer Science and Information Systems, College of Applied Sciences, Almaarefa University, Diriyah, Riyadh 13713, Saudi Arabia
[5]Department of Computer Engineering, College of Computers and Information Technology, Taif University, Taif 21944, Saudi Arabia
[6]Kazakh National Women's Pedagogical University, 050000 Almaty, Kazakhstan

Corresponding authors: Hend Khalid Alkahtani (Hkalqahtani@pnu.edu.sa) and Ihtisham Ul Haq (ihtisham.haq@dimes.unical.it)

This work was supported by the Princess Nourah Bint Abdulrahman University Researchers Supporting Project, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia, under Grant PNURSP2023R384.

ABSTRACT According to the World Health Organization (WHO), congenital heart disorders (CHDs) impact a significant proportion of babies worldwide, with prevalence rates ranging from 0.8% to 1.2%. Phonocardiography (PCG) is the leading non-invasive method for detecting congenital heart defects (CHDs), offering important information about cardiac signals i.e. S1, S2, S3, and S4 and heartbeat patterns. This research study focuses on developing a strong binary classification system for congenital heart diseases (CHDs) utilizing deep neural networks. The system will be trained on a combined dataset that includes both local and publicly available repositories. The local dataset (LD) consists of 583 signals containing both normal and aberrant PCG recordings, while the public dataset (PD) obtained from Michigan University consists of 23 PCG recordings. In order to achieve consistency and compatibility, both datasets are subjected to down-sampling, resulting in a frequency of 8 kHz. A well-engineered band-pass filter efficiently removes signals outside the 20–650 Hz range, allowing for exclusive processing of the desired frequencies. The signals are divided into segments, each lasting exactly 4 seconds. The study utilizes data augmentation techniques, specifically pitch-shifting, to boost the model's robustness. This is accomplished by implementing a 1D convolutional neural network (CNN). The most notable outcomes are achieved in case C, exhibiting a sensitivity of 99.0%, specificity of 98.0%, F1 score of 98.56%, precision of 98.57%, and an accuracy of 98.56%. This study enhances the progress of automated techniques for detecting congenital heart disease (CHD), demonstrating the potential of deep neural networks in precision medicine for pediatric cardiology.

INDEX TERMS Digital stethoscope, congenital heart diseases, pitch-shifting, deep neural network, chunking, data augmentation.

I. INTRODUCTION

Congenital heart diseases (CHDs) encompass a broad spectrum of cardiac abnormalities that are present since birth. These conditions encompass tricuspid (TVD) and mitral

The associate editor coordinating the review of this manuscript and approving it for publication was Genoveffa Tortora.

valve disease (MVD), aortic valve disease (AVD) and pulmonic valve disease (PVD). Both juvenile and adult populations face significant obstacles due to these disorders. Notable cases include Atrial Septal Defect (ASD), which is a condition where there is an abnormality in the septum resulting in the mixing of deoxygenated and oxygenated blood; Patent Ductus Arteriosus (PDA), characterized by the persistence of an open passage between the heart's main blood vessels; Tetralogy of Fallot (TOF), a complex combination of ventricular septal defect (VSD), pulmonary stenosis, right ventricular hypertrophy, malposition aorta, and Ventricular Septal Defect (VSD), which involves a breach in the wall separating the heart's lower chamber. In the event of a sufficiently extensive breach, there is potential for enduring damage to lungs and heart, hence elevating the likelihood of a heart attack. These disorders show the significance of having a thorough understanding and employing creative techniques to confront the difficulties of CHDs in varied patient populations.Cardiac abnormalities are a significant contributor to mortality in individuals of all ages, including both pediatric and adult populations [1]. A study conducted by Michigan University indicates that congenital heart abnormalities are present in approximately 1% of babies globally [2]. The tetralogy of Fallot is the predominant congenital heart abnormality among cyanotic congenital heart diseases. Although medical research has made advancements, the mortality rate remains a significant worry, particularly with the inadequate treatment and preventative measures for infants born with cardiac abnormalities [3]. Traditional diagnostic methodologies can be costly and time-consuming. The objective of this research is to address these issues by suggesting the development of an advanced decision support system that is specifically tailored for the classification of CHD. We intend to revolutionize the existing diagnosis method by integrating cutting-edge technology. This would significantly enhance the accuracy, efficiency, and accessibility of cardiac anomaly diagnostics for pediatric patients. The objective of this project is to integrate traditional diagnostic methods with advanced technologies to improve results and provide timely intervention for people with CHDs. During our research at Lady Reading Hospital (LRH) and Rehman Medical Institute (RMI) in Peshawar, we observed a significant pattern. A substantial number of patients seeking cardiac care came not just from the surrounding regions but also from Afghanistan. Figure 1 demonstrates that our local dataset (LD) revealed a notable frequency-based distribution of cardiac illnesses. Ventricular septal defects (VSD) were the prevailing condition observed in the patients, affecting 30% of them.

Furthermore, 20% of the subjects were found to have pulmonic valve disease (PVD). Furthermore, there was a significant prevalence of various cardiac ailments. This study, conducted by an experienced researcher, aimed to identify the many heart problems present in the local community, with a specific emphasis on the common ailments affecting children. To effectively address the healthcare needs of vulnerable children, it is crucial to possess a thorough comprehension of the frequency of cardiac diseases, specifically in this demographic. This knowledge is necessary for the development of successful healthcare programs and interventions. The objective of our research was to determine and describe the frequency of heart disorders, particularly in the pediatric population. Our primary objective was to analyze the patterns of occurrence and distribution among patients residing in both local and cross-border regions.

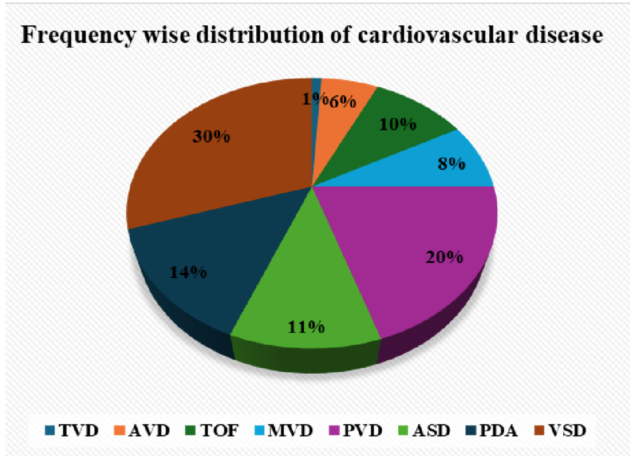


FIGURE 1. Distribution of cardiovascular disease incidence in local data by frequency.

II. RELATED WORK

Medical research has focused extensively on the study of CHDs and the development of diagnostic techniques for these conditions. Through an in-depth analysis of the available literature, we gain a thorough comprehension that reveals the intricate and subtle aspects related to these cardiac abnormalities. Previous research has thoroughly examined many diagnostic methods, including both conventional approaches and advanced technology, with the goal of improving the precision, effectiveness, and availability of identifying and categorizing CHDs. The purpose of this literature review is to consolidate and expand upon the current body of information, establishing a basis for our study that centers on creating an advanced decision support system for categorizing CHDs.

Alanazi et al. [4] evaluate the acoustic analysis of three stethoscopes—Starkey SLI-ST3, Cardionics E-Scope II (both electronic), and Littmann Classic SE II (conventional)—in evaluating simulated heart sounds. The investigation examined normal, aortic valvular stenosis (AVS), and pulmonic heart sound simulations at 85, 250, 400, 550, and 1050 Hz in a soundproof booth. Amplified stethoscopes, such the Starkey SLI-ST3, amplify normal and abnormal heart sounds across all frequencies more than traditional ones. Cardionics E-Scope II amplified aberrant heartbeats better. The study advises health practitioners and hearing-impaired people to be wary of product claims and suggests studying these effects using hearing aids.

A lightweight convolutional neural network (CNN) is presented for the screening of valvular heart disease utilizing phonocardiogram signals on mobile devices in a research article [5] by panelShichao. The model accomplishes a remarkable average accuracy of 99.4% in 10-fold cross-validation by utilizing self-supervised learning (SSL) with unlabeled data. On typical mobile devices, the suggested method achieves near-real-time performance, surpassing bigger CNN models in terms of speed and power economy. An effective and efficient method for screening for valvular heart disease using artificial intelligence in mobile health applications is proposed. The goal of this research [6] by panelBaris Bozkurt is to use digital phonocardiogram (PCG) signals to automate the diagnosis of structural cardiac anomalies in juvenile heart disease screening. This study investigates how well convolutional neural networks (CNNs) perform when taught to segmented PCG frames represented by time-frequency data. This study compares two high-quality databases and indicates that sub-band envelopes perform better than the commonly used MFCC and Mel-Spectrogram features. The study also found that sub-band envelopes are influenced by domain knowledge. Also, in CNN-based designs for heart sound classification, the research indicates that period synchronous windowing works better than asynchronous windowing. Researchers A. Zhang,et. al; set out to improve intelligent diagnosis systems' ability to identify children's heart sounds in their study published in [7]. Using wavelet soft threshold algorithm (WST) and variational modal decomposition (VMD), the article presents a new method for reducing noise. With a focus on reducing weeping sounds, the suggested method successfully denoised 103 phonocardiogram data. When compared to WST alone, the results showed a significant improvement in the signal-to-noise ratio, particularly when dealing with Gaussian noise. In comparison to WST alone, the intelligent classification system trained on post-denoising features outperformed WST in terms of accuracy (92.23%), sensitivity (92.42%), and specificity (91.89%) when it came to congenital heart disease. By demonstrating how well the suggested noise reduction method removes undesired heart sound noise, this study improves the intelligent screening for congenital heart disorders in children. In [8], Roy et al. investigates the use of deep learning and machine learning methods to develop effective classifiers for the identification of valvular heart disease. Upon comparison with existing deep learning models, the innovative Xception network model outperformed them all while requiring the least amount of time to test, achieving an astounding 99.45% accuracy. Considerations such as Root Mean Square, Energy, and Zero Crossing Rate were part of the feature extraction process. Based on our comparisons, Random Forest and support vector machine are the top two machine learning algorithms. The suggested Xception model exhibited promise for rapid and accurate diagnosis of valvular heart disease, and it outperformed the competition. Discrete wavelet transformations (DWTs) and Mel-frequency cepstral coefficients (MFCCs) are extracted in the investigation of cardiac sound signal classification in [9]. The study utilizes SVMs, deep neural networks, and k-nearest neighbors (kNN) for reliable classification. Researchers also made a significant addition to the field in [10], where they created an automated PCG diagnostic approach to forecast the outcome of children and teenagers diagnosed with Ventricular Septal Defect (VSD). The study employs kNN and features extracted from MFCC on a dataset comprising 55 children aged between 6 months and 2 years. Automated detection of normal and abnormal cardiac sounds is accomplished by Chen et al. [11] employing the Physionet Challenge 2016 dataset (CinC). They developed a new method by merging a 1D-CNN with an LSTM. Employing trained support vector machine (SVM), k-nearest neighbor (kNN), and multilayer perceptron (MLP) models, [12] performed a comprehensive evaluation in the field of cardiac sound signal classification. Wavelet transform (WT), maximum frequency clustering (MFCCs), power spectrum, and other factors are taken into account. In their study on PCG recording classification, Nassralla et al. [13] present a thorough method that includes segmentation, preprocessing, feature extraction, and random forest (RF) classification. They achieve an astounding 92% accuracy. Employing a combination of CNN and infinite impulse response filters, their suggested model for CVD detection achieved an astounding 97% accuracy on the PhysioNet dataset and 87% accuracy on the PASCAL dataset. In [19], authors use MobileNetv2 classifiers to detect PVCs in ECGs with 99.90% accuracy without feature extraction or cross-validation. The author discusses PVC recognition using image analysis and clustering to find ECG patterns in [20]. Preliminary data suggests four PVC layouts. Author in [21]. Deep neural networks like ResNet-50 and Inception V3 categorized QRS complexes to identify PVCs. Besides high accuracy (up to 99.8%), failure analysis to explain classifier flaws improves reliability. These studies provide a solid groundwork for future improvements in our planned study by providing significant insights into various approaches for heart sound classification.

III. MATERIALS & METHOD

Our complex method for classifying juvenile CHDs using phonocardiogram (PCG) data is summarized in Figure 2. When gathering Heart Sound (HS) signals from private and publicly available sources, it is critical to keep the sample rate constant at 8 kHz. This ensures consistency and helps reduce inconsistencies in the initial sample frequencies. Following widely-accepted methods described in prior research [15], [16], [17], we have partitioned our dataset into 4-second chunks to guarantee rapid identification of cardiac signal anomalies. By doing so, we can successfully detect any discrepancies within the allotted time frame of 3 to 6 seconds. By applying sample padding, signals that deviate from this segmentation can be corrected. This method improves the consistency of the dataset, which in turn can boost the efficiency of the real-time model. After segmentation and preprocessing, the HS signals are classified with a 1D-CNN classifier, which achieves an impressive 98.56% accuracy in differentiating normal from abnormal HS. A strong

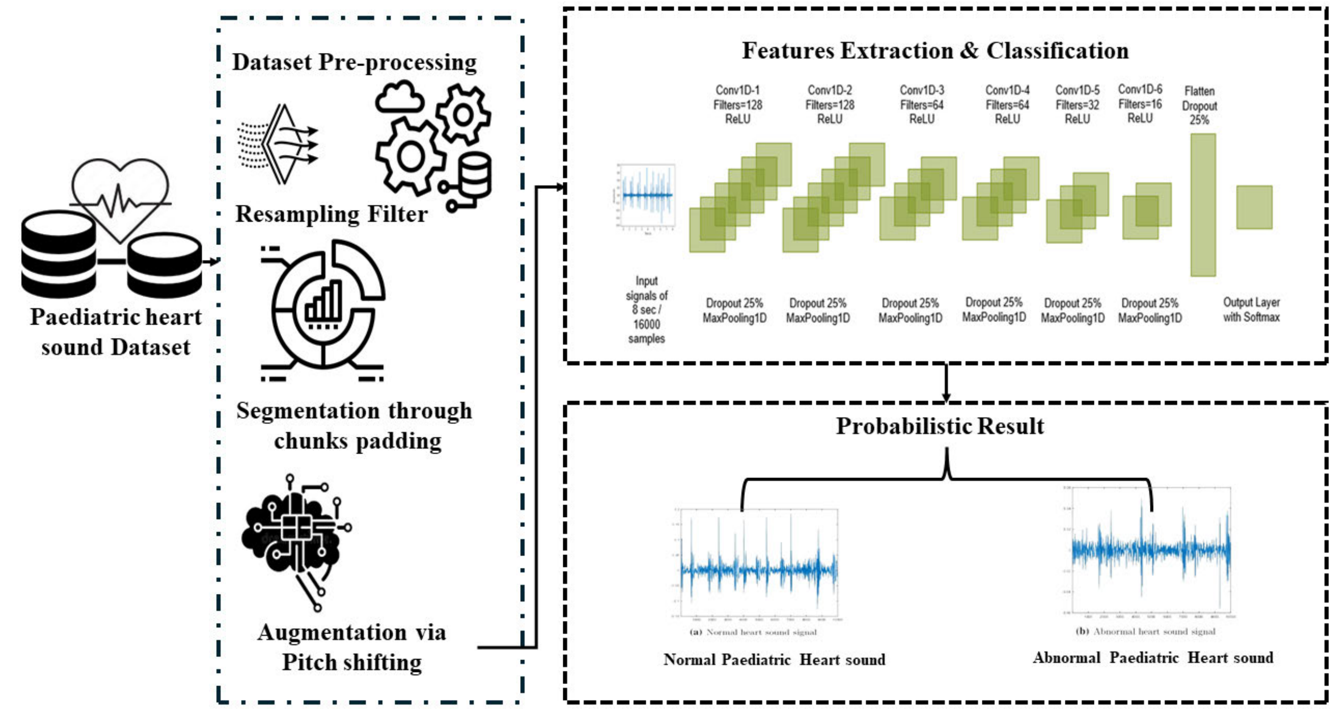


FIGURE 2. **Graphical visualization of developed methodology.**

foundation for categorizing hypertrophic cardiomyopathy (HS) in children has been developed, in large part due to this study's emphasis on pediatric congenital heart abnormalities (CHDs). We go above and beyond basic data analysis in our thorough evaluation of the demographic distribution of CHDs during our data collection approach. This study's findings should inform new policies in Pakistan that target the early detection and screening of CHDs, as they provide crucial information regarding the prevalence of these medical conditions in the country. In particular, our proposed method makes use of a one-dimensional Convolutional Neural Network (CNN) trained on a combination dataset, the majority of which consists of data collected locally. This design style makes it easier to spot congenital heart defects in children at an early age. We investigate the complexities of using local data to train deep neural networks and find solutions to the problems that come with it. Many factors must be taken into account, including the complexities of data annotation, the dynamics of dataset acquisition, the challenges of generalizing models, and the nuanced process of hyperparameter tuning. Our research has improved our understanding of the frequency of CHD in Pakistan and contributed significantly to the area of CHD classification. The results of our research can have a major impact on the creation and execution of policies concerning CHD screening and early detection, thanks to our extensive understanding of the subject. In the region, this could have a major impact on pediatric healthcare.

A. DATASET DESCRIPTION

The proposed methodology employs the amalgamation of two distinct datasets, each offering unique insights and augmenting the study's comprehensiveness.

1) LOCAL DATASET

The Local Dataset (LD) is meticulously curated from reputable local healthcare institutions, namely Lady Reading Hospital (LRH) and Rehman Medical Institute (RMI) in Peshawar. This dataset comprises a comprehensive compilation of 168 pediatric Heart Sound (HS) signals that demonstrate anomalies, showing various types of CHDs. In addition, it has 415 pediatric HS signals that are classified as normal. The duration of each signal recording varies from 10 to 12 seconds, providing a full understanding of cardiac anomalies in the presence of ambient and lung noises. Precise and trustworthy data were ensured through the utilization of advanced instruments such as electronic stethoscopes (uSteth), Audacity, MATLAB and Google Colab, software during the acquisition procedure.

2) PUBLIC DATASET

The Public Dataset (PD) is obtained from the prestigious University of Michigan Medical School [18]. This dataset includes heart sound signals from four specific locations: pulmonic area in a sitting position, aortic region in a sitting position, apex area in a left decubitus position, and apex area in a supine position. The signals were collected using an electronic stethoscope with a sample rate of 1 kHz. This dataset consists of 21 aberrant and 2 normal heart sound signals, making it a significant external resource. It offers further insights into the variances in heart sounds, which can complement local data. The duration of each signal in the PD dataset varies from one minute to one minute and fifteen seconds, providing a wider temporal viewpoint. By amalgamating these datasets, the suggested technique acquires a more comprehensive and diverse depiction of pediatric HS signals. The study's trustworthiness is greatly enhanced by employing a dual-sourced technique, which allows for broader application of the findings. The time domain plots for both datasets are visually depicted in Figure 3, offering insight into the distinctive characteristics of the collected data. The meticulous selection and organization of this dataset is vital to guarantee the study's triumph in precisely categorizing children CHDs using Phonocardiogram signals.

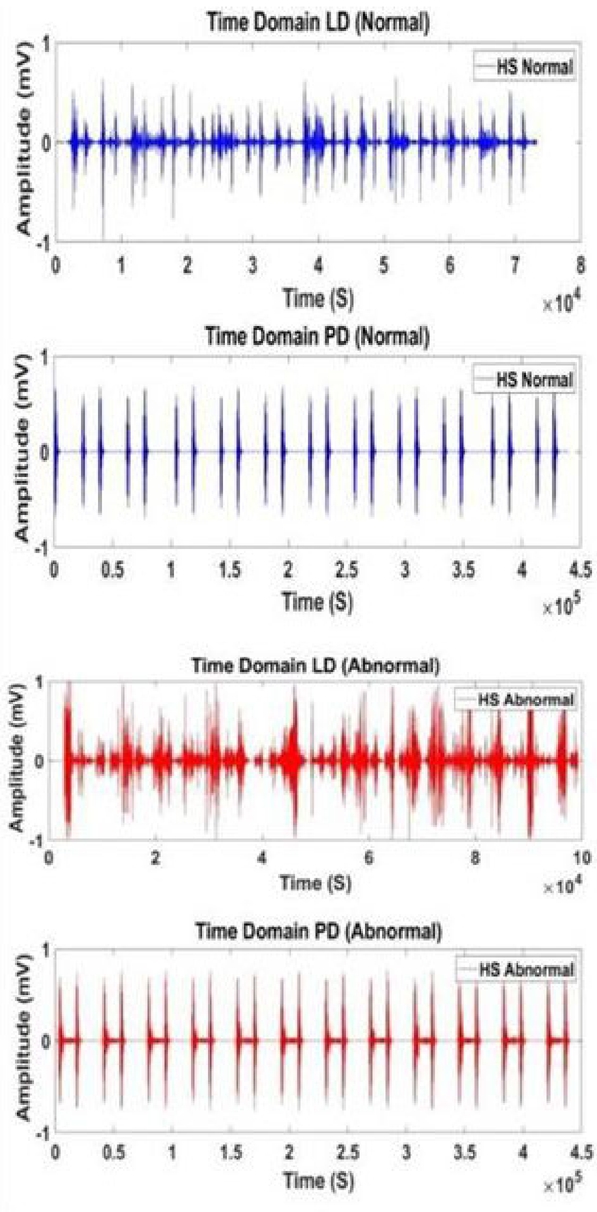


FIGURE 3. In raw format, the dataset displays four categories: public normal, local normal, public abnormal, and local abnormal.

B. DISTRIBUTION OF DATASET

The restricted amount of Heart Sound (HS) signals in the locally obtained dataset had an impact on its dispersion. To address this constraint and augment the dataset, a deliberate decision was taken to merge both datasets. The dataset comprises 189 anomalous HS signals with lengths ranging from 8 seconds to 15 second and from 15 seconds to 1 minute, together with 417 regular HS signals, each lasting 12 seconds. Subsequently, the signals were systematically partitioned into segments of 4 seconds each, yielding a dataset including 2074 samples, each with a signal duration of 4 seconds. Table 1 provides a detailed breakdown of both datasets, giving a full summary of the makeup and features of the dataset. The consolidation of the data resolved the limits of the local dataset and resulted in a more extensive dataset for the subsequent phases of the research investigation.

C. DATASET PREPROCESSING

During the pre-processing phase, the dataset we collected locally included a wide range of heart sounds along with different environmental noises, such as infants crying, conversations in hospital settings, and interference from stethoscope movements. For improved accuracy, specificity, and sensitivity in the subsequent analysis, we have incorporated filtering and pre-processing techniques. These techniques are described below:

a) Resampling: The dataset collected from the local source was sampled at 8 kHz, while the public dataset (PD) had an original sampling rate of 44.1 kHz. It is essential to maintain a balanced quality and consistent time duration for optimal effectiveness of deep learning models. In order to ensure consistency, we resampled the Phonocardiogram (PCG) Heart Sound (HS) signals from the PD at a rate of 8 kHz.

b) Filtering: Understanding the frequency range of heart murmurs [16], we developed and implemented a filter with the parameters specified in Table 2. By implementing a meticulous filtering process, the analysis is able to concentrate on the specific frequency range linked to heart sounds. This enhances the dataset, leading to more precise and significant outcomes.

TABLE 1. Distribution of dataset.

Dataset	Recording Sensor	Sampling Rate	Abnormal HS	Normal HS
Public Dataset	Electronic stethoscope	44.1 KHz	21	2
Local Dataset	U-Steth	8KHz	168	415

TABLE 2. Specification of filter.

ROLL-OFF DB PER OCTAVE	STOP BAND FREQUENCY	PASS BAND FREQUENCY	REQUIRED BAND
48	650	20	20-650

D. DATA AUGMENTATION

Synthetic data must be generated via techniques such as pitch-shifting, time-stretching, and time compression under the domain of Data Augmentation (DA). In cases where specific sections of the dataset are missing, these strategies become extremely beneficial. In this specific case, pitch-shifting has been employed as a targeted strategy to increase the dataset's diversity and, consequently, the results.

Pitch Shifting: Pitch-shifting (PS) is a method that specifically modifies the high-frequency component of an audio signal. This is important because Heart Sounds (HS) have unique characteristics, such as the sharp peaks related to S1 and S2. Pitch-shifting, as opposed to changing the original signal, adds subtle changes to the frequency components of the signal, resulting in a slightly enhanced dataset. The mathematical expression for pitch shifting is given below

$$\text{(s)pitch} - \text{shift} = s * \left(\frac{f}{\beta}\right), \text{ where } \beta = \frac{2^{k}}{12} \tag{1}$$

In equation (1), f represents the frequency tone, while β represents the pitch-shifting factor and k depicts semitones. When k is more than 0, it indicates upshifting. Conversely, when k is less than or equal to 0, it indicates downshifting. The frequency components of the augmented data are visually depicted in Figures 4 and 5 as spectrograms. These spectrograms exhibit the converted frequency components of public abnormal (PA), public normal (PN), local abnormal (LA), local normal (LN) data after pitch-shifting.

E. BINARY CLASSIFICATION

CNN are widely recognized for their effective utilization of subsampling and convolution techniques in the field of classification. The 1D-CNN, a modified version of these networks, has demonstrated remarkable efficacy in various domains, including precisely discerning and rapidly addressing individual biomedical data and detecting aberrations. In the context of this investigation, the 1D CNN serves as the classification model, adept at extracting features from time series data. Prior to being fed into the 1D-CNN for the categorization of normal and abnormal Heart Sound (HS) signals, both the local dataset (LD) and public dataset (PD) undergo comprehensive preprocessing, segmentation, and data augmentation. Figure 6 illustrates the architectural design of the 1D-CNN employed in this study, offering a graphical depiction of the network's configuration and the layers utilized for the classification procedure.

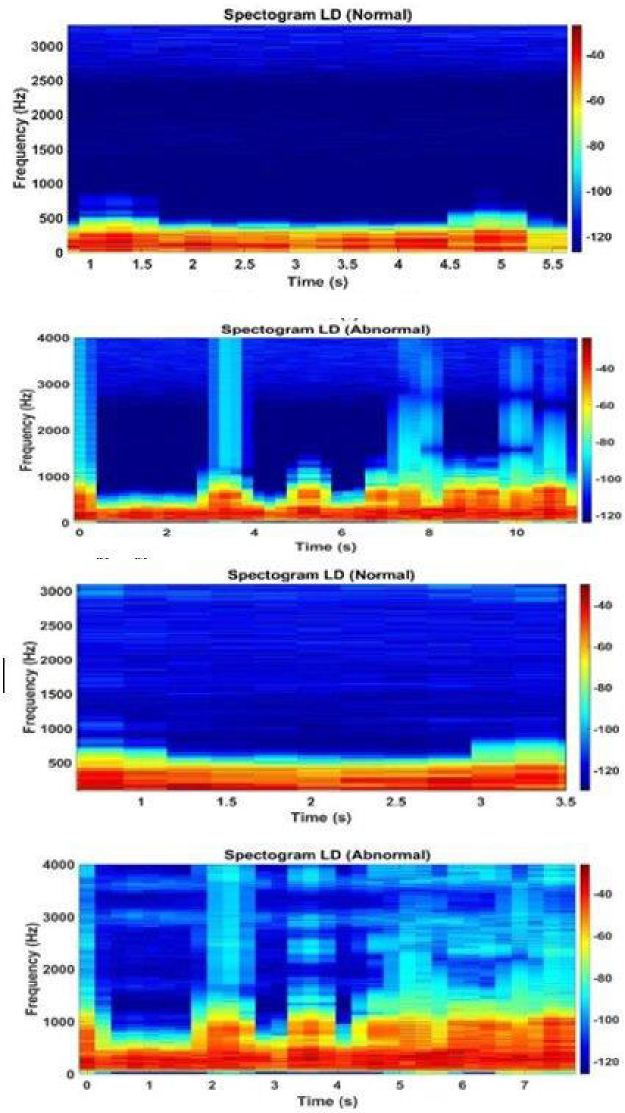


FIGURE 4. The PCG signal graphs for the local dataset show (a) the unpitched LN and LA and (b) the pitched LN and LA.

IV. RESULTS

The aforementioned research approach was assessed in three distinct examples, which will be elaborated upon in the subsequent sections. For the 1D-CNN, I used specific hyper-parameters to achieve optimal results.

These included a batch size of 4, training the model for 100 epochs, and experimenting with different learning rates i.e. 0.1, 0.001, and 0.0001. Additionally, I split the data into 85% for training, 10% for validation, and 5% for real-time local testing.

Scenario A: In scenario A, the 1D-CNN model was trained utilizing raw, unprocessed, unsegmented, and unfiltered data. The dataset comprised 894 anomalous Heart Sounds (HS) and 959 regular HS signals, each with varying durations. The results are given in Table 3. As an adept researcher, the primary objective of this study was to precisely detect and identify anomalous heart sounds with exceptional F1 scores,

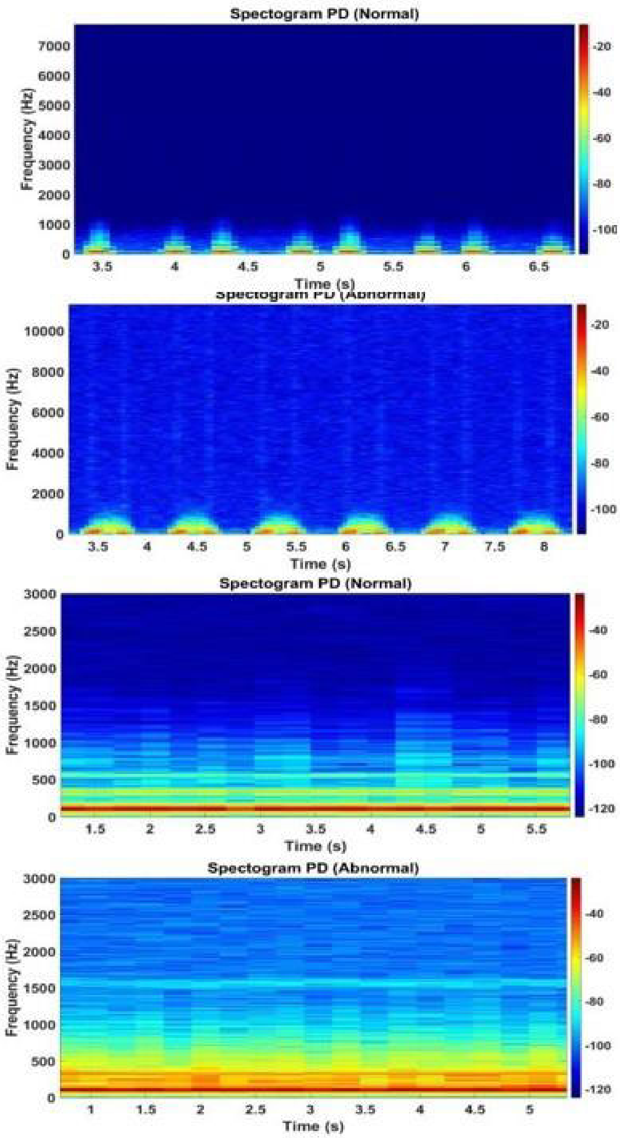


FIGURE 5. The PCG signal graphs for the public dataset show (a) the unpitched LN and LA and (b) the pitched LN and LA.

TABLE 3. Scenario A results.

Performance Metric	Testing Data	Validation Data
Sensitivity	99	95
Specificity	98	98
F1-Score	98.56	97.15
Precision	98.57	97.15
Accuracy	98.56	97.15

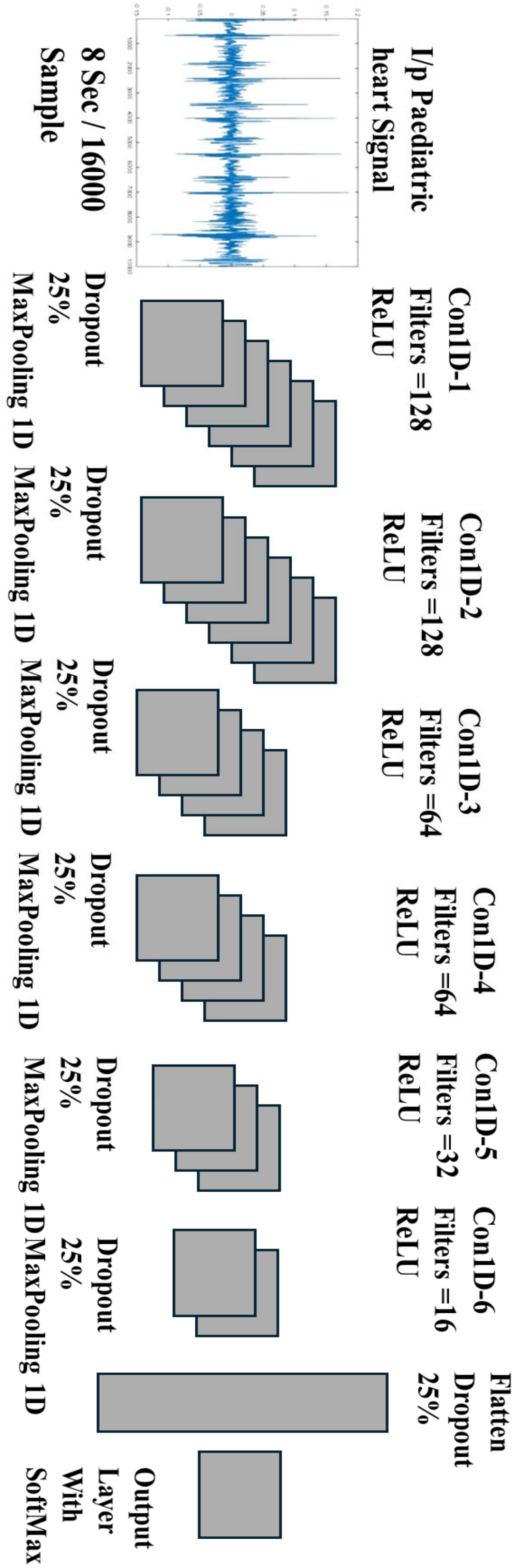


FIGURE 6. Customized 1D CNN architecture.

precision, accuracy and sensitivity. Both datasets included abnormal sounds i.e. murmurs, which are distinguished by heart sounds with a very low frequency. Case A had a sensitivity rating of 93%, indicating a significant level of precision. However, the model encountered difficulties during the

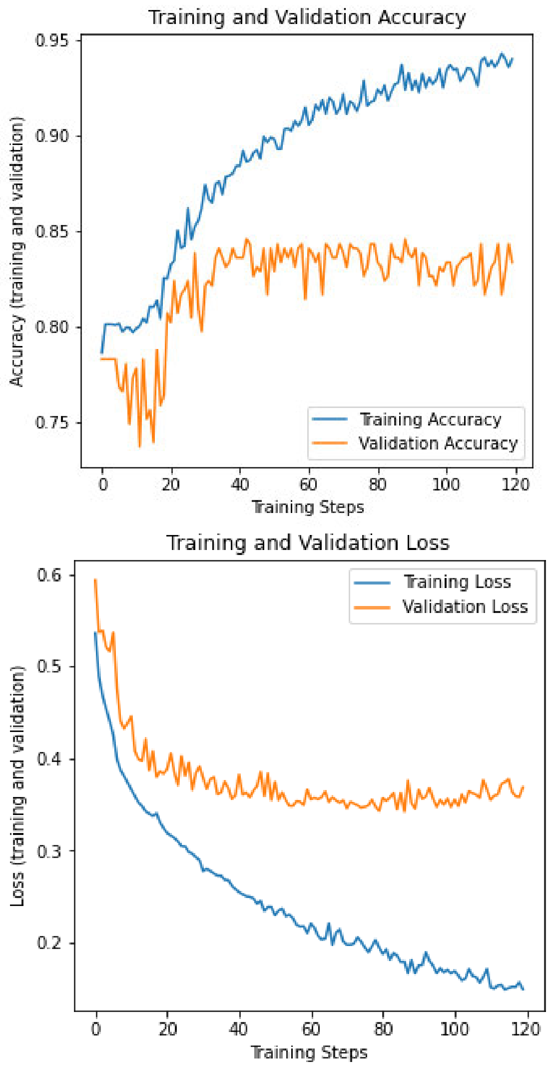


FIGURE 7. Accuracy scores for validation and training datasets, loss during validation and training.

training process, as depicted in Figure 7, and yielded suboptimal results. The curves demonstrate the discrepancies in the training dataset and the existence of substantial random variation.

Following a thorough assessment, we executed tactics to enhance outcomes. These strategies encompassed methods such as data augmentation (chunking, pitch shifting) and data filtering. Figure 7 illustrates the training and validation accuracy with loss, demonstrating the notable enhancements made by the deployed strategies.

Scenario B: The implementation of a thorough enhancement plan was carried out in Scenario B. The data were filtered with the help of the Audacity software, which utilised a cutoff frequency range that extended from sixty to six hundred and fifty hertz. Following the completion of pre-processing, filtration, and segmentation into 4-second chunks, the dataset was increased to include 1639 normal Heart Sounds (HS) and 1632 abnormal HS signals. A presentation of the findings can be found in Table 4.

TABLE 4. Scenario B results.

Performance Metric	Testing Data	Validation Data
Sensitivity	95	94
Specificity	99	96
F1-Score	97.25	95.17
Precision	97.35	96.23
Accuracy	97.25	96.16

The dataset that was being tested showed a significant increase in sensitivity, reaching a level of 95%, which is indicative of a positive influence. In spite of the fact that there was an improvement, there was still some unpredictability in the curves, as shown in Figure 8, which indicates that there is the potential for further improvement.

A visual representation of the training and validation accuracy, coupled with the loss, is presented in Figure 8. This figure illustrates the gradual improvement that was created by the strategies that were implemented.

Scenario C: During the implementation of Scenario C, a comprehensive enhancement strategy was carried out. This strategy included preprocessing, data augmentation, data segmentation and data filtration as methodologies. The segmented dataset underwent extension through the use of pitch-shifting, resulting in an increased count of normal Heart Sounds (HS) to 3230 and aberrant HS to 2302. A comprehensive breakdown of the results of this expanded dataset can be seen in Table 5. Particularly noteworthy is the fact that the trained model's sensitivity underwent a significant rise, ultimately reaching 98.0%. As shown in Table 6, this accomplishment stands out as having the highest sensitivity on the combined Local Dataset (LD) and Public Dataset (PD). This is the case when compared with references [4], [5], and [7]. A visual representation of the training and validation accuracy with loss is presented in Figure 9, which demonstrates the significant increase that was achieved through the combination of the strategies that were mentioned.

TABLE 5. Scenario C results.

Performance Metric	Testing Data	Validation Data
Sensitivity	99	95
Specificity	98	98
F1-Score	98.56	97.15
Precision	98.57	97.15
Accuracy	98.56	97.15

The author in [4] utilizes the CinC dataset to forecast heart disease. They utilized the KNN technique to extract time domain features and attained an accuracy of 97.82%. In [5], the author utilized a Fuzzy Inference technique on time domain characteristics using a 300-sample general dataset, achieving an accuracy of 97.82%. The author utilized self-collected data to extract features from time and frequency

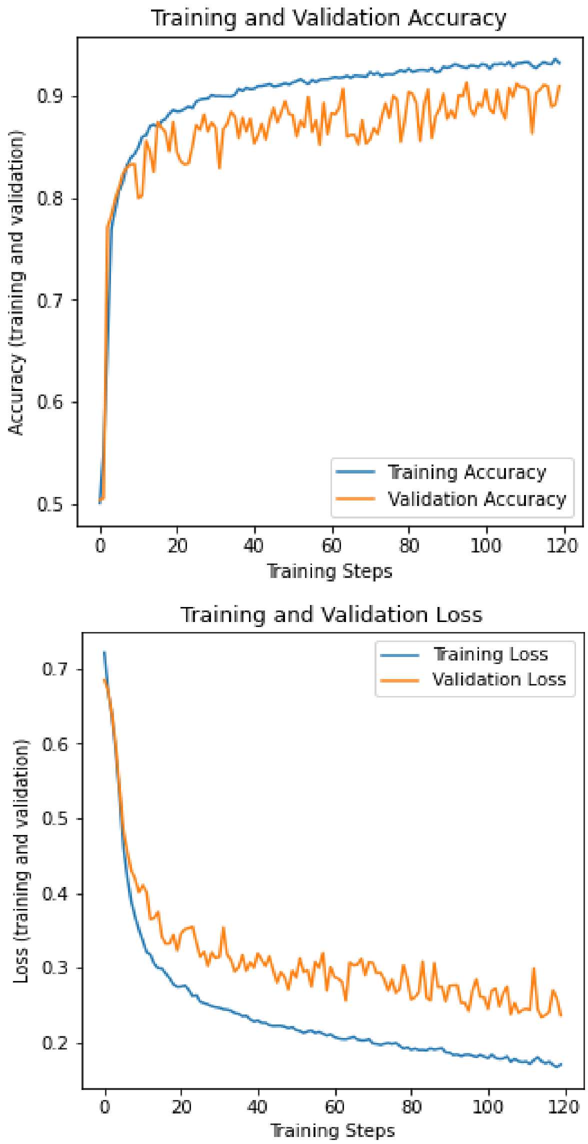


FIGURE 8. Accuracy scores for validation and training datasets, loss during validation and training.

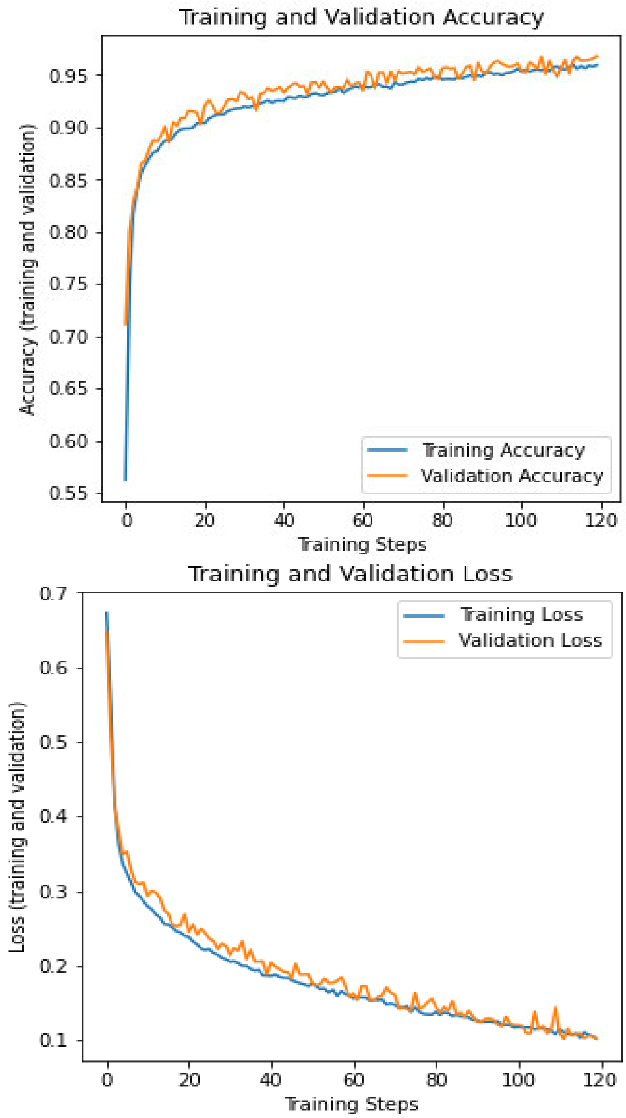


FIGURE 9. Accuracy scores for validation and training datasets, loss during validation and training.

domains employing WST and VMD approaches, resulting in an accuracy of 97.82%. Author in [7] Utilizing self-collected data, features were retrieved using Mel-frequency cepstral coefficients (MFCC) and one-dimensional local ternary patterns (1D-LTPs). An accuracy of 95.24% was achieved by employing the SVM algorithm. The author in [8] utilized both CinC and Pascal datasets to extract MFCC features and implemented a CNN algorithm, resulting in an accuracy of 95.24%. In a study utilizing the CinC dataset, the author in [9] utilized MFCC and Wavelet Entropy features, applied the RF algorithm, and obtained an accuracy of 92%. Author in [10] retrieved features such as RMSE, Skewness, and Kurtosis from a dataset of 1000 samples from CinC. RF, KNN, SVM, and Naïve Bayes algorithms were used, achieving accuracies of 99%, 76%, 99%, and 98% respectively. In [11], the author extracted MFCC and WT characteristics from self-collected data and applied SVM and KNN algorithms, resulting in an accuracy of 92.5%. The author utilized the CinC dataset to extract MFCC features and implemented a 1D CNN followed by LSTM, achieving an accuracy of 93.2% [12]. The author utilized self-collected data to extract MFCC features and implemented the KNN algorithm, achieving an accuracy of 93.2% [13]. The author in [14] utilized a mix of MFCC and DWT characteristics by making use of self-collected data. Algorithms KNN, SVM, and BNN were used, achieving accuracies of 97.2%, 92%, and 97.2% respectively. During the Proposed Methodology section, the author utilized WST, VMD, MFCC, and LTPs for feature extraction by utilizing self-collected data and the CinC dataset. A tailored 1D CNN algorithm was employed, achieving an accuracy of 98.56%.Table 7 shows the statistical justification for cardiovascular disease.

TABLE 6. Results comparison with existing techniques.

Ref	Dataset	Features	Deployed Algorithm	Accuracy
[4]	CinC	Time domain	KNN	97.82
[5]	300 Sample General Dataset	Time domain	Fuzzy Inference	97.82
[6]	Self-collected Data	Time & Frequency domain	WST, VMD	97.82
[7]	Self-collected Data	MFCC & 1D-LTPs	SVM	95.24
[8]	CinC & Pascal	MFCC	CNN	95.24
[9]	CinC	MFCC & Wavelet Entropy	RF	92
[10]	1000 Samples from CinC	RMSE, Skewness Kurtosis	RF, KNN, SVM Naïve Bayes	99, 76, 99, 98
[11]	Self-collected Data	MFCC WT	SVM KNN	92.5
[12]	CinC	MFCC	1D CNN LSTM	93.2
[13]	Self-collected Data	MFCC	KNN	93.2
[14]	Self-collected Data	MFCC+DWT DWT, MFCC	KNN, SVM BNN	97.2, 92, 97.2
Proposed Methodology	**Self-collected Data, CinC**	**WST, VMD, MFCC, LTPs**	**Customize 1D CNN**	**98.56**

TABLE 7. Cross validation/ statistical justification.

S. No	*Classes*	Dataset for Cardiovascular Surgeon test	Predicted True	Cardiova--scular expert opinion	Error Percent age
1	***Cardiovascular disease sound Signal***	30	29	1	5%
2	***Normal Heart signal***	30	28	2	10%

V. DISCUSSION

The study evaluated three specific instances in the categorization of paediatric CHDs utilizing Phonocardiogram (PCG) data. In Case A, the absence of preprocessing and segmentation led to less than optimum outcomes, which prompted the use of improved approaches in following cases. In Case B, the implementation of filtration resulted in an enhanced sensitivity of 95%, nevertheless, random variance continued to exist. Case C, which involved preprocessing, filtration, segmentation, and data augmentation, reached a remarkable sensitivity of 98.0%, surpassing previous studies. The application of pitch-shifting significantly improved the model's ability to detect aberrant heart sounds. The results highlight the effectiveness of thorough preprocessing and augmentation strategies in improving the accuracy of CHD classification.

Due to constraints in the dataset and the limited timeframe of the study, this analysis specifically focused on a small range of CHDs, which resulted in intrinsic limitations. Initially, the study is limited to binary classification, namely distinguishing between normal and abnormal cases. This calls for further investigation into multiclass classification in order to conduct a more thorough examination. In addition, the existence of murmurs, which are characterized by low-frequency noise, can potentially compromise the accuracy of results due to the influence of ambient noise. While it is currently impossible to completely eliminate environmental noise, continuous progress in sensor and filter technology provides exciting opportunities to tackle and reduce this issue in future research endeavors.

VI. CONCLUSION

Exploring the immense possibilities of artificial intelligence in the field of healthcare management, mainly in the field of accurate illness diagnosis, is highly promising, as demonstrated by its application in CHDs. Over the past few years, experts in academia and healthcare have made remarkable progress in this area, placing a greater focus on employing signal processing and deep learning techniques to classify heart sounds (HS). This study is unique since it introduces a dataset that focuses on paediatric patients, which is different from prior studies that mostly focused on diagnosing HS in adults. By employing a 1D-CNN-based intelligent model, we have attained exceptional performance metrics that outperform models trained on publicly accessible datasets for both adult and paediatric HS. Our proposed methodology exhibited outstanding precision and sensitivity when evaluated on three separate datasets. Case C stood out as the most sophisticated method, with an exceptional accuracy of 98.56% on real-time unseen test data. Additionally, it demonstrated remarkable precision, specificity, and sensitivity. This performance highlights the effectiveness of our algorithmic system in classifying paediatric CHD. Our work proposes that artificial intelligence algorithms can play a revolutionary role in overcoming the existing limits of digital stethoscopes, which are insufficient in analysing heart sounds and recognizing disorders. Our suggested approach enables a thorough evaluation and analysis of HS, which has the potential to be the fundamental component of an intelligent device. The purpose of this device is to provide cardiologists with advanced diagnostic capabilities and help improve the detection of anomalies in the population. Implementing our algorithm into current digital stethoscopes signifies a significant advancement in healthcare innovation, paving the way for enhanced diagnostic capabilities and better patient treatment.

Only a few congenital heart disorders (CHDs) were examined, limiting its thoroughness. Binary categorization of normal and abnormal instances is limited in the study, hence multiclass classification is needed. Murmurs, low-frequency noise, can also affect outcomes owing to ambient noise. Due to dataset constraints and research timeline, the paper admits its shortcomings. To reduce environmental noise in future research, sensor and filter technologies must advance.

REFERENCES

[1] *Congenital Heart Disease (CHD), CS Mott Children's Hospital, Michigan Medicine*. [Online]. Available: https://www.mottchildren.org/conditions-treatments/pedheart

[2] A. Chelu, S. G. Williams, B. D. Keavney, and D. Talavera, "Joint analysis of functionally related genes yields further candidates associated with tetralogy of fallot," *J. Hum. Genet.*, vol. 67, no. 10, pp. 613–615, Oct. 2022, doi: 10.1038/S10038-022-01051-Y.

[3] R. E. Gross, "Surgical management of the patent ductus arteriosus: With summary of four surgically treated cases," *Ann. Surgery*, vol. 110, no. 3, pp. 321–356, Sep. 1939, doi: 10.1097/00000658-193909000-00001.

[4] A. Alanazi, S. Atcherson, C. Franklin, and M. Bryan, "Frequency responses of conventional and amplified stethoscopes for measuring heart sounds," *Saudi J. Med. Med. Sci.*, vol. 8, no. 2, p. 112, 2020.

[5] S. Ma, J. Chen, and J. W. K. Ho, "An edge-device-compatible algorithm for valvular heart diseases screening using phonocardiogram signals with a lightweight convolutional neural network and self-supervised learning," *Comput. Methods Programs Biomed.*, vol. 243, Jan. 2024, Art. no. 107906.

[6] B. Bozkurt, I. Germanakis, and Y. Stylianou, "A study of time-frequency features for CNN-based automatic heart sound classification for pathology detection," *Comput. Biol. Med.*, vol. 100, pp. 132–143, Sep. 2018, doi: 10.1016/J.COMPBIOMED.2018.06.026.

[7] A. Zhang, J. Wang, F. Qu, and Z. He, "Classification of children's heart sounds with noise reduction based on variational modal decomposition," *Frontiers Med. Technol.*, vol. 4, May 2022, Art. no. 854382.

[8] T. S. Roy, J. K. Roy, and N. Mandal, "Classifier identification using deep learning and machine learning algorithms for the detection of valvular heart diseases," *Biomed. Eng. Adv.*, vol. 3, Jun. 2022, Art. no. 100035.

[9] D. Chen, W. Xuan, Y. Gu, F. Liu, J. Chen, S. Xia, H. Jin, S. Dong, and J. Luo, "Automatic classification of normal-abnormal heart sounds using convolution neural network and long-short term memory," *Electronics*, vol. 11, no. 2022, p. 1246, 2022.

[10] S. Akbari, H. Ghassemian, and Z. Akbari, "Systolic murmurs diagnosis improvement by feature fusion and decision fusion," in *Proc. IEEE Int. Conf. Signal Image Process. Appl. (ICSIPA)*, Sep. 2019, pp. 6–11, doi: 10.1109/ICSIPA45851.2019.8977784.

[11] M. Nassralla, Z. E. Zein, and H. Hajj, "Classification of normal and abnormal heart sounds," in *Proc. 4th Int. Conf. Adv. Biomed. Eng. (ICABME)*, Oct. 2017, pp. 1–4, doi: 10.1109/ICABME.2017.8167538.

[12] M. Boulares, R. Alotaibi, A. AlMansour, and A. Barnawi, "Cardiovascular disease recognition based on heartbeat segmentation and selection process," *Int. J. Environ. Res. Public Health*, vol. 18, no. 20, p. 10952, Oct. 2021, doi: 10.3390/IJERPH182010952.

[13] M. Ladrova, M. Sidikova, R. Martinek, R. Jaros, and P. Bilik, "Elimination of interference in phonocardiogram signal based on wavelet transform and empirical mode decomposition," *IFAC-PapersOnLine*, vol. 52, no. 27, pp. 440–445, 2019, doi: 10.1016/J.IFACOL.2019.12.703.

[14] (2020). *Murmur Library*. [Online]. Available: https://www.med.umich.edu/lrc/psb open/html/repo/primer

[15] *Professional Skill Builder—Heart Sound Murmur Library*. [Online]. Available: https://www.med.umich.edu

[16] P. T. Krishnan, P. Balasubramanian, and S. Umapathy, "Automated heart sound classification system from unsegmented phonocardiogram (PCG) using deep neural network," *Phys. Eng. Sci. Med.*, vol. 43, no. 2, pp. 505–515, Jun. 2020, doi: 10.1007/S13246-020-00851-W.

[17] C. Potes, S. Parvaneh, A. Rahman, and B. Conroy, "Ensemble of feature-based and deep learning-based classifiers for detection of abnormal heart sounds," in *Proc. Comput. Cardiology Conf. (CinC)*, Sep. 2016, pp. 621–624.

[18] Yaseen, G.-Y. Son, and S. Kwon, "Classification of heart sound signal using multiple features," *Appl. Sci.*, vol. 8, no. 12, p. 2344, Nov. 2018, doi: 10.3390/APP8122344.

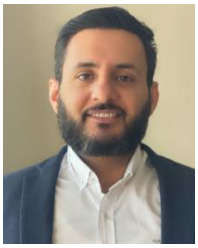
[19] F. De Marco, F. Ferrucci, M. Risi, and G. Tortora, "Classification of QRS complexes to detect premature ventricular contraction using machine learning techniques," *PLoS ONE*, vol. 17, no. 8, Aug. 2022, Art. no. e0268555.

[20] F. De Marco, L. Di Biasi, A. A. Citarella, M. Tucci, and G. Tortora, "Identification of morphological patterns for the detection of premature ventricular contractions," in *Proc. 26th Int. Conf. Inf. Visualisation (IV)*, Jul. 2022, pp. 393–398.

[21] F. De Marco, D. Finlay, and R. R. Bond, "Classification of premature ventricular contraction using deep learning," in *Proc. Comput. Cardiol.*, Sep. 2020, pp. 1–4.

HEND KHALID ALKAHTANI (Member, IEEE) received the Bachelor of Science degree in computer science from the School of Engineering and Applied Science, The George Washington University, in 1992, the Master of Science degree in information management from the Department of Engineering Management, The George Washington University, in 1993, and the Ph.D. degree in information security from the Department of Computer Science, Loughborough University, in 2018. She is currently pursuing the Ph.D. degree with the Information Systems Department, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University. She is an Associate Professor with the Information Systems Department, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University. She has 23 years of work experience as a Lecturer, the Computer Center President, and the Statistics Center President in faculty collages. She received an Award from the SIDF Academy: Leading Creative Transformation in Critical Time Program, Stanford University, Center for Professional Development.

IHTISHAM UL HAQ received the bachelor's degree in mechatronics engineering and the master's degree in automation and control from UET Peshawar, in 2022, showcasing his commitment to continuous learning and growth. He is currently pursuing the Ph.D. degree with the ICT Department, University of Calabria, Italy, where he is delving deeper into the captivating realms of clinical image processing, artificial intelligence, machine learning, deep learning, and explainable AI. His bachelor's degree laid the strong foundation for his academic pursuits. He is a passionate and dedicated Researcher. He is a Research Associate with NCAI UET, Peshawar. His journey in academia began as a Research Assistant with NCAI UET, from June 2021 to June 2022, where he honed his skills and nurtured his curiosity for cutting-edge technologies. Prior to that, he contributed his expertise as a Lab Engineer with the Mechatronics Department, UET Peshawar, for two impactful years, from 2019 to 2021. His passion for exploring the potential of these cutting-edge technologies shines through in his research, where he strives to push the boundaries of knowledge and make a positive impact on society. With an insatiable thirst for innovation and a dedication to scientific inquiry, he is undoubtedly an emerging force in the field of research, poised to make significant contributions to the advancement of technology and its applications in the medical domain.



YAZEED YASIN GHADI received the Ph.D. degree in electrical and computer engineering from Queensland University. He is currently an Assistant Professor of software engineering with Al Ain University. He was a Postdoctoral Researcher with Queensland University, before joining Al Ain University. He has published more than 80 peer-reviewed journals and conference papers. He holds three pending patents. His current research interests include developing novel electro-acoustic-optic neural interfaces for large-scale high-resolution electrophysiology and distributed optogenetic stimulation. He was a recipient of several awards. His dissertation on developing novel hybrid plasmonic photonic on-chip biochemical sensors received the Sigma Xi Best Ph.D. Thesis Award.

NISREEN INNAB received the Ph.D. degree in computer information systems, in 2008. She is currently an Associate Professor with the Department of Computer Science and Information Systems, College of Applied Sciences, Almaarefa University, Riyadh, Saudi Arabia. She is also the Program Director of the Information Systems Program. She supervised many theses for the Master of Information Security. Furthermore, she also supervised many graduation projects for computer science, information systems, and health information systems. She has published many scientific research in international high-prestigious journals indexed in Web of Science and Scopus. Her research interests include information security, cyber security, machine learning, artificial intelligence, and the IoT.

MASOUD ALAJMI (Member, IEEE) received the B.S. degree in electrical engineering from the King Fahad University of Petroleum and Minerals (KFUPM), in 2004, and the M.S. degree in electrical engineering and the Ph.D. degree in electrical and computer engineering from Western Michigan University, Kalamazoo, MI, USA, in 2010 and 2016, respectively. He has over four years of experience in the industry. He was with Zamel and Turbag Consulting Engineers, Al-Khobar, Saudi Arabia, as an Electrical Engineer, for three months. After that, he joined Saudi Electricity Company (SEC), Abha, Saudi Arabia, where he worked as a Pre-Commissioning Engineer, from 2004 to 2008. During that period, he completed many training programs in the technical and administrative fields at well-known institutes. Also, he was assigned to be a Commissioning Leader for many projects in Saudi Arabia. He was assigned to be the SEC Representative to supervise factory acceptance tests for Siemens Company, Berlin, Germany, in 2007, and Hyundai Heavy Industries Company Ltd., Ulsan, South Korea, in 2008. From 2012 to 2015, he was a Teaching Assistant with the Electrical and Computer Engineering Department, Western Michigan University. He is currently an Associate Professor with the Computer Engineering Department, Taif University, Taif, Saudi Arabia. He is involved in various technical committees. He is the coauthor of several papers in international journals and conference proceedings. His research interests include signal processing, biomedical image processing, image encryption, watermarking, steganography, data hiding, machine learning, smart grids, and renewable energy. He received the 2014–2015 Graduate Teaching Effectiveness Award from Western Michigan University for excellent teaching skills.

MEKEBAYEV NURBAPA is currently with Kazakh National Women's Pedagogical University, Almaty, Kazakhstan.