A Deep Learning Model for Recognizing Pediatric Congenital Heart Diseases Using Phonocardiogram Signals

Md Hassanuzzaman[1], Nurul Akhtar Hasan[2], Mohammad Abdullah Al Mamun[3], Khawza I. Ahmed[1], Ahsan H. Khandoker[4], Raqibul Mostafa[1]

[1]Department of Electrical and Electronic Engineering, United International University, Dhaka, BD
[2]Children & Women Cardiac Unit, National Heart Foundation Hospital & Research Institute, Dhaka, BD
[3]Department of Pediatric Cardiology, Bangladesh Shishu (Children) Hospital and Institute, Dhaka, BD
[4]Healthcare Engineering Innovation Center, Department of Biomedical Engineering, Khalifa University, Abu Dhabi, UAE

Abstract

Diagnosing congenital heart disease (CHD)in children through heart sound auscultation requires extensive medical training and understanding. However, the quality of PCG data may be compromised due to the sensor location, a child's developing heart, and the complex and changeable cardiac acoustic environment. This study proposes a one-dimensional Convolution Neural Network (1D-CNN) with a residual block that classifies PCG signals to predict heart abnormalities in 751 patients with PCG signals aged five months to twenty years. After assessing the signal quality, only good-quality signals are used as input features of the neural network. The study's results showed the accuracy of 0.93 accuracy and 0.98 sensitivity. The Receiver Operating Characteristic (ROC) plot yielded an Area Under Curve (AUC) value of 0.98, and the F1-score was 0.94. The proposed model required only 15 sec of the PCG signals to predict CHD cases (4.2 ms processing time). Thus, it can be implemented as a primary screening tool for remote-end pediatricians by providing cheaper and faster interpretations of PCG signals before referring the cases to specialists.

1. Introduction

Pediatric CHD refers to a group of structural defects in the heart that are present at birth. These conditions can range from mild abnormalities to severe, life-threatening diseases. CHD is the most common congenital disability, affecting approximately 1 in every 100 live births worldwide [1]. In the United States alone, around 40,000 babies are estimated to be born with CHD yearly [1]. However, it is recognized as a global health concern, affecting children in developed and developing countries. In low-income countries, access to specialized cardiac care and timely interventions may be limited, increasing morbidity and mortality rates for children with CHD. Full recovery and return to normal functioning are possible if cardiac disease in children is identified early and treated.

On the other hand, if the issue is discovered too late, irreparable alterations in the heart muscle structure develop, necessitating urgent surgery [2]. Many CHD cases can be diagnosed prenatally through advanced fetal imaging, allowing for early intervention. However, despite these advancements, there are still challenges in detecting pediatric CHD effectively due to the limited access to specialized care in remote areas, and the high cost of treatment can pose significant barriers to optimal care for many children with CHD, particularly in resource-constrained settings. Automatic CHD diagnosis has several advantages, such as quicker diagnostic processes that enable faster treatments. It lowers the possibility of an incorrect diagnosis, mainly when CHD symptoms are not immediately apparent or are confused with symptoms of other illnesses. Also, it lessens the strain on medical professionals by allowing them to focus on patient care rather than scheduling diagnostic tests and assessments.

Machine learning techniques have recently been used to generate models identifying neonatal mortality linked to this disease. According to Hussain L. et al. [3], the ANN prediction model was assessed for accuracy in predicting the outcome of neonates with CHD after being trained on medical data sets. Dritsas, E. et al. [4], risk categories for neonates with CHD can be identified using screening techniques and prediction models. Balakrishnan, M. et al. [5], data gathered from maternal laboratory testing, clinical laboratory data, and other research predicting CHD may be analyzed using machine learning techniques like logistic regression and decision trees. A deep-learning-based computer-aided approach for pediatric CHD diagnosis was developed by Xiao et al. [6] using two lightweight convolution neural networks. Chen et al. [7] used 1D-CNN + LSTM with preprocessing techniques to distinguish between normal and abnormal heart sounds. A segment-based heart sound segmentation method was used by Weize et al. [8] to identify abnormal heart sounds. Using a continuous wavelet transform-based method, Abbas et al. [9] extracted representative features from PCG data. Convolutional Vision Transformers architecture was also developed to classify heart sounds. Most automatic heart-sound classification methods now concentrate on several features, including time interval, state frequency spectrum, state amplitude, energy, record frequency spectrum, high-order statistics, and entropy. Sometimes, the manually obtained features are highly complex, which might lead to subjectivity and variability biases.

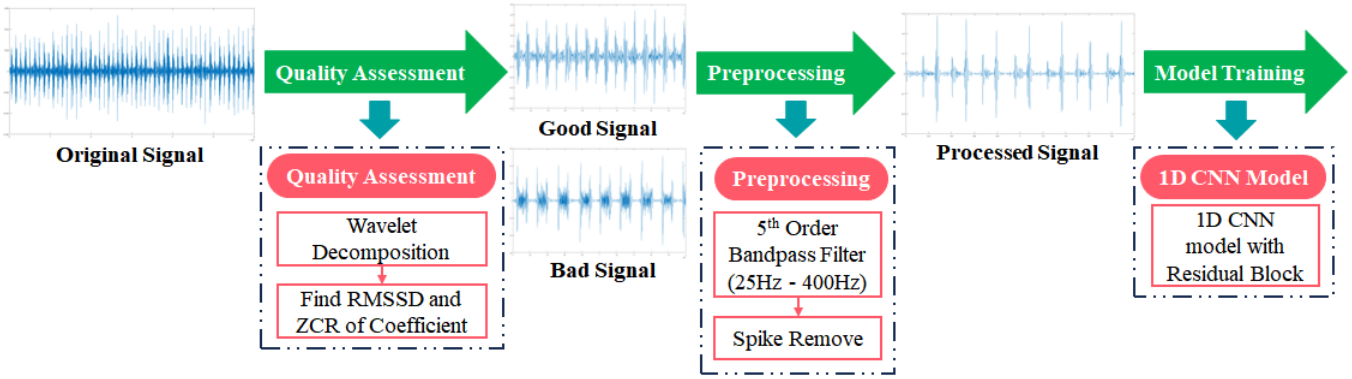

Figure 1. A systematic flow diagram of the proposed system for classifying pediatric congenital heart diseases.

First, using an electronic stethoscope on 751 distinct patients, a dataset of pediatric heart sounds is created for this work. The quality of the cardiac sound is then evaluated. A 1D CNN with a residual block is presented to classify processed heart sounds. This study also demonstrates the model's efficacy with patient demographic data, such as age, gender, and mode of delivery (MOD). The study also examines the model's performance in terms of each patient individually. A deep learning network's classification decisions are also examined using the gradient-weighted class activation mapping (Grad-CAM) approach. It offers insights into the particular regions of the PCG signal that improve interpretability and confidence in the model's predictions. The model has no dependency on the data acquisition position to predict heart abnormalities.

2. Data and methods

Our proposed CHD diagnosing method's flowchart is shown in Fig. 1.

2.1. Data Acquisition and Description

The data acquisition method has been explained in detail in [10]. Each signal was collected at a sample rate of 4000Hz for 15s. The datasets contained PCG signals from 751 distinct patients (total recordings: 3435), ranging in age from 5 months to 17 years, with a male-to-female ratio of 60:40 and a CHD-to-Non-CHD ratio of 61:39. The number of signals in this instance is the total number of signals gathered from patients.

2.2. Data Preprocessing

The recording and labeling of pediatric heart sounds is a difficult task because, in comparison to adult subjects, children are less able to fully cooperate with the doctors' instructions during the recording process, which results in the addition of extra noise like cries, screams, coughs, talking, moving, and breathing sounds. They are consequently unsuitable for pediatric heart sound analysis and direct algorithmic research. Signal quality assessment methods are applied [11] to address this issue. A three-level discrete wavelet transform was used to decompose heart sound data using the Daubechies 2 as the mother wavelet. The root mean square of successive differences (RMSSD) and the zero crossings rate (ZCR) of estimated coefficients of wavelet decomposition were utilized as assessment criteria [Eq - 1].

$$RMSSD < 0.2 \ \& \ ZCR < 0.4 \qquad \text{(Eq 1)}$$

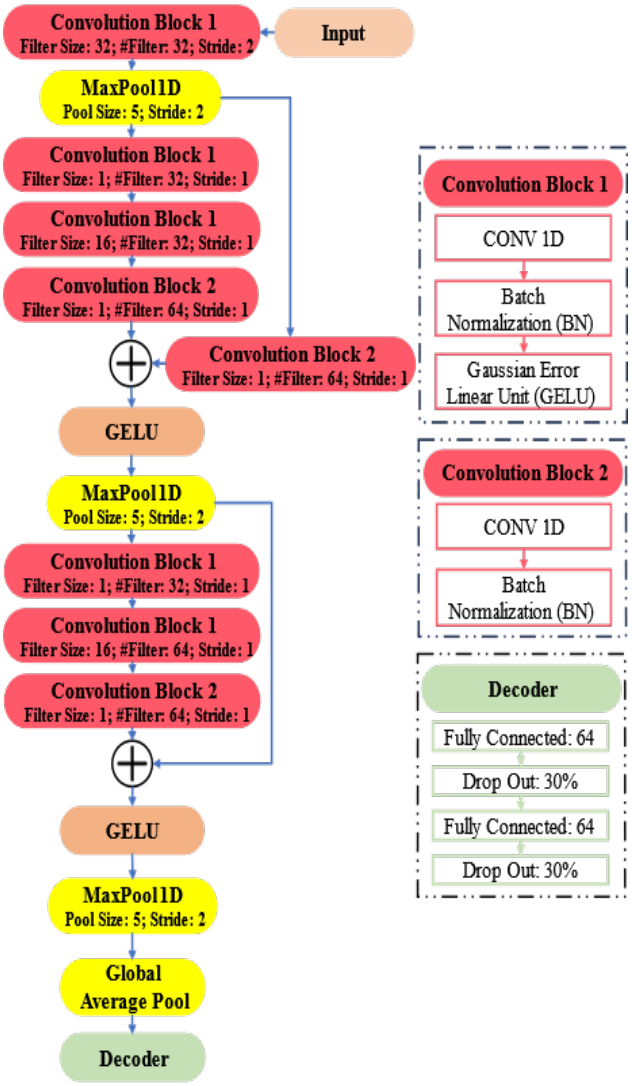

Figure 2. The proposed 1D convolutional neural network with residual block architecture.

The two preset indicators of the two-level wavelet decomposition approximation coefficients of the heart sound signal were calculated using the quality evaluation

procedure based on those mentioned above two preset indicators. Here, the RMSSD and ZCR threshold values 0.2 and 0.4 are used. The signal was regarded as good if the threshold value was satisfied. If the prerequisite was not satisfied, the sample was considered invalid. Following quality analysis, the primary distribution frequency band of the heart sound information was retained using the Butterworth 5th order band-pass filter, with the cut-off frequency in the 25–400 Hz range. Then, Schmidt's [12] technique was applied to locate and remove spikes generated by the acquisition process.

2.3. Model Architecture

Following the successful completion of preprocessing, the classification of congenital heart disease into CHD and non-CHD categories using a 1D convolutional neural network with residual block method. The Z-score approach was used to standardize each effective PCG signal before it was fed into the suggested deep-learning model. As illustrated in Fig. 2, we used residual blocks to build the model. Convolution Block 1 process uses a batch normalization (BN) and a Gaussian Error Linear Unit (GELU) after the 1D convolution layer. To avoid gradient disappearance and explosion, BN and GELU are utilized. Convolution Block 2 and Convolution Block 1 are comparable. It has no activation layer. Two fully connected layers comprise the decoder layer, and two dropout layers have a 30% drop rate. There are 64 and 32 neurons in each of the fully connected layers, respectively.

3. Results

For training the model, the training dataset is split into 95/5 for training/validation based on the patients in the experiment. The model that had achieved a high accuracy of 0.93, a sensitivity of 0.98, and a precision of 0.91 in the test dataset was next tested using patient-wise, entirely unbiased test cases. We used the Inverse-Frequency approach to apply class weights in the loss function during the network's training phase to solve the class imbalance issue. The receiver operating characteristic curves led to the AUC score of 0.98. The model's lowest False Negative Rate (FNR) was 0.022, which is essential. FNR is crucial in the medical domain since making erroneous negative predictions might severely impact patients. The F1 score is essential in the medical industry to guarantee that a diagnostic model can accurately identify instances of a specific condition while reducing both false positive and false negative outcomes. The experiment's F1- score is 0.94. The model's effectiveness is also analyzed from the patient's demographic data perspective, such as age, gender, and MOD. Our study shows that model performance declines as age increases, as seen in Table 1. Regarding gender, the model's performance in female patients is better than in male patients, as shown in Table 2. Additionally, we noticed that model performance varies based on the MOD of the patient. Lower Uterine segment Cesarean Section (LUCS) delivery patients perform better than Normal Vaginal Delivery (NVD) delivery patients in the proposed model. The data analysis revealed that LUCS patients exhibited more consistent physiological markers and presented with fewer confounding factors, such as variability in labor progression or fetal distress. The proposed model performs well for particular patients. The proposed model can be used to accurately classify a PCG signal obtained from any of the valve positions, as shown in Fig. 3. This demonstrates the potential of the proposed model to effectively diagnose and monitor heart conditions related to these valve positions. The model's reasoning for making a classification decision regarding a test signal is depicted in Fig. 4. The proposed model has appropriately learned to detect S1, S2, systole, and diastole peaks at the signal. Evidently, the model correctly emphasizes the S1, S2, systole, and diastole peaks on the signal that convey useful information. The models can use these valuable features to predict disease when they are correlated with disease prevalence. Moreover, it improves accountability and transparency in the decision-making process. This explainability can also help improve doctors' confidence in and adoption of AI systems.

Table 1. Model performance to age.

Age (Year)	#Signals	Accuracy	Sensitivity	F1 Score
0 - 1	20	0.87	1	0.80
1 - 5	81	1	1	1
5 - 10	52	0.88	0.95	0.92
> 10	11	0.87	1	0.90

Table 2. Model performance to gender.

Gender	#Signals	Accuracy	Sensitivity	F1 Score
Male	77	0.91	0.95	0.92
Female	87	0.95	1	0.96

Table 3. Model performance to the MOD.

MOD	#Signals	Accuracy	Sensitivity	F1 Score
NVD	69	0.89	0.95	0.92
LUCS	95	0.96	1	0.96

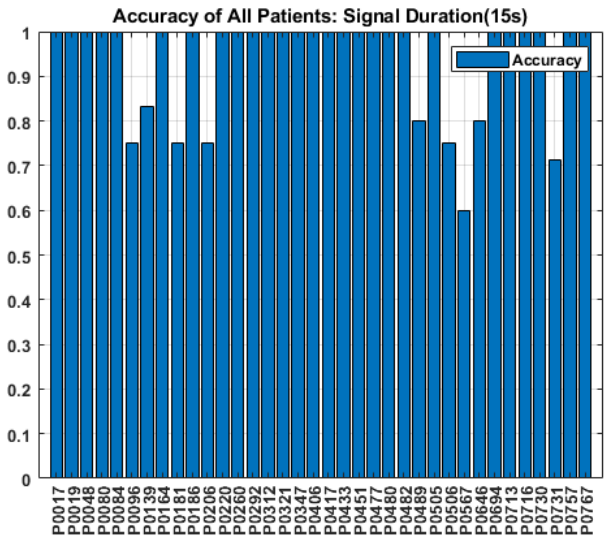

Figure 3. Model performance with respect to the patient.

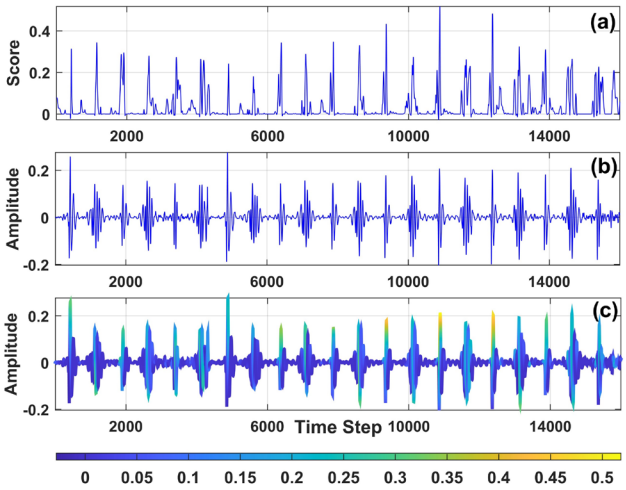

Figure 4. Explanation of model predictions using Grad-CAM (a) Importance score, (b) PCG signal, and (c) Importance score in the PCG signal.

4. Discussion and future work

A deep learning model based on a one-dimensional convolution neural network (1D-CNN) with a residual block is developed to classify pediatric heart sounds. The model has a high AUC, along with good accuracy and sensitivity. The study also examines how demographic elements like age, gender, and delivery method affect the model's effectiveness. Our findings emphasize the need to consider demographic aspects when creating and applying Diagnostic models to ensure their applicability across various patient groups. Data quality evaluation and preprocessing methods are essential to increase the PCG signals' dependability. Explainability can also assist in highlighting model and data biases and make the model predictions more transparent. Information about the model's decision-making process helps the doctor understand how and why specific predictions are made. This knowledge can be helpful in high-stakes applications like healthcare, where inaccurate predictions might have disastrous outcomes. The proposed deep learning model can serve as a primary screening tool for identifying heart abnormalities in pediatric patients with its interpretability, especially in remote areas with limited access to specialized care.

However, the study acknowledges limitations, such as the specific dataset used and the need for further validation on more extensive and diverse datasets. Overall, the study highlights the potential of deep learning techniques in diagnosing and monitoring heart conditions and calls for additional research and validation to evaluate the model's effectiveness in real-world clinical settings.

Acknowledgments

The Institute for Advanced Research (IAR) of the United International University in Bangladesh has funded the study. Under Award: UIU/IAR/02/2019-20/SE/10. The Healthcare Engineering Innovation Center (HEIC), located at Khalifa University in the United Arab Emirates, partially contributed to its funding. Under Award: 8474000132.

References

[1] Congenital Heart Defects (CHD) | Children's Hospital Colorado. (n.d.). https://www.childrenscolorado.org/ conditions-and-advice/conditions-and-symptoms/conditions/ congenital-heart-defects-chd/

[2] Bukhari, F.; Idrees, M.; Iqbal, W. Predictive Analysis of Congenital Heart Defects Prior to Birth. In Proceedings of the 2021 International Conference on Robotics and Automation in Industry (ICRAI), Xi'an, China, 30 May–5 June 2021; IEEE: New York, NY, USA, 2021; pp. 1–6.

[3] Hussain, L.; Awan, I.A.; Aziz, W.; Saeed, S.; Ali, A.; Zeeshan, F.; Kwak, K.S. Detecting congestive heart failure by extracting multimodal features and employing machine learning techniques. BioMed Res. Int. 2020, 123–138. [CrossRef]

[4] Dritsas, E.; Alexiou, S.; Moustakas, K. Cardiovascular Disease Risk Prediction with Supervised Machine Learning Techniques. ICT4AWE 2022, 315–321.

[5] Balakrishnan, M.; Christopher, A.A.; Ramprakash, P.; Logeswari, A. Prediction of Cardiovascular Disease using Machine Learning. In Journal of Physics: Conference Series; IOP Publishing: Bristol, UK, 2021; Volume 1767, p. 012013.

[6] Xiao B, Xu Y, Bi X, Li W, Ma Z, Zhang J, Ma X. 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.https://doi.org/10.1109/ jiot.2019.2961132. PubMed PMID: WOS:000522265900035.

[7] 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. 8, p. 1246, 2022.

[8] Xu, Weize, et al. "Automatic pediatric congenital heart disease classification based on heart sound signal." Artificial Intelligence in Medicine 126 (2022): 102257.

[9] Abbas, Q.; Hussain, A.; Baig, A.R. Automatic Detection and Classification of Cardiovascular Disorders Using Phonocardiogram and Convolutional Vision Transformers. Diagnostics 2022, 12, 3109. https://doi.org/10.3390/diagnostics12123109.

[10] Hassanuzzaman, Md, et al. "Recognition of Pediatric Congenital Heart Diseases by Using Phonocardiogram Signals and Transformer-Based Neural Networks." 2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE, 2023.

[11] Mei, Na, et al. "Classification of heart sounds based on quality assessment and wavelet scattering transform." Computers in Biology and Medicine 137 (2021): 104814.

[12] S.E. Schmidt, C. Holst-Hansen, C. Graff, et al., Segmentation of heart sound recordings by a duration-dependent hidden Markov model, Physiol. Meas. 31 (4) (2010) 513–529, https://doi.org/10.1088/0967-3334/31/4/004.

Address for correspondence:

Md Hassanuzzaman
Department of Electrical and Electronic Engineering
United International University, Dhaka, Bangladesh (BD)
hassanuzzaman1503@gmail.com
https://sites.google.com/view/mdhassan