Development of a Classification Model for Pediatric Cardiac Abnormality Detection Through the Integration of Child Physiological Data and Phonocardiogram Features

Yuyu Yen, Chun Chang, Jui-Hung Kao and Woei-Chyn Chu

November 16, 2024

Development of a Classification Model for Pediatric Cardiac Abnormality Detection Through the Integration of Child Physiological Data and Phonocardiogram Features

Yu-Yu, Yen[1], Chun Chang [2], Jui-Hung Kao[3] and Woei-Chyn Chu[1]*

[1] Department of Biomedical Engineering,
National Yang Ming Chiao Tung University, Taipei 112304, Taiwan R.O.C.
sheepkelly19.be11@nycu.edu.tw (Y.Y.Y.); wchu@nycu.edu.tw (W.C.C.) *
[2] Department of Biological Science and Technology,
National Yang Ming Chiao Tung University, Hsinchu
[3] Department of Information Management,
Shih Hsin University, Taipei, Taiwan

Abstract. Heart murmurs play an irreplaceable role in the assessment of cardiac conditions. The type of heart murmur is an important characteristic of cardiovascular diseases (CVDs). This study aims to classify the types of heart murmurs in patients using Support Vector Machine (SVM) from phonocardiogram (PCG) signals and to determine whether the patient's heart is normal. The study used heart sound recordings from the CirCor DigiScope dataset, which provides labels for murmur classification, pitch, quality, timing, and more. To incorporate physiological information into the model's judgment and consider the impact of age on heart development, this study focused on subjects labeled "Child." Additionally, due to the limited number of subjects with diastolic murmurs, the research focused solely on the evaluation of systolic murmurs. Since heart murmurs are a type of noise, we further calculate the first and second derivatives of the Mel spectrogram features as our characteristics. Finally, we use support vector machines for classification to determine whether the subject is a patient with heart murmurs. Ultimately, the precision of the diagnosis for patients reaches 82.7%.

Keywords: Physiological data, Phonocardiogram (PCG), Cardiac Abnormaliy, Child.

1 Introduction

Cardiovascular diseases (CVDs) are the leading cause of death worldwide. According to World Health Organization statistics, approximately 17.9 million people died from cardiovascular diseases in 2019, accounting for 32% of all global deaths. Phonocardiograms (PCGs) are signals produced by the mechanical activity of the heart, containing important information about heart condition [1]. Through cardiac auscultation, it is possible to assess whether a subject requires more expensive follow-up examinations in a cost-effective and efficient manner [2]. Therefore, in middle- and low-income countries where cardiovascular diseases (CVDs) are prevalent and pose a higher risk of death, auscultation can help reduce delays in diagnosis and treatment, thus lowering the risk of mortality [3].

Although cardiac auscultation has its advantages in diagnosis, mastering this skill requires extensive training and clinical experience [3], making it difficult to popularize cardiac auscultation in areas of middle and low-income countries where shortage of cardiology specialists is a problem [4]. To effectively address this issue, the fundamental approach needs to tackle the cost of training doctors in cardiac auscultation, computer-assisted auscultation can effectively solve this problem [5]. Most studies have only focused on whether murmurs are present or if heart sounds are normal, without incorporating the physiological characteristics of the subjects. Therefore, the combination of grading heart sounds, pitch, timing, quality, and physiological information can provide a new perspective for cardiac assessment [6].

Heart murmurs include systolic and diastolic murmurs, which occur during the contraction and relaxation phases of the heart, respectively. Since this study excludes diastolic heart murmurs, we will only discuss systolic heart murmurs. According to Leathem's classification, systolic heart murmurs are mainly divided into systolic ejection murmurs and systolic regurgitant murmurs [7]. The quality of heart murmurs described as "harsh" is most commonly associated with systolic ejection murmurs, which are usually related to benign systolic flow murmurs as well as valve and vascular obstructions [8]. The qualities "blow" and "harsh" may appear in systolic regurgitant murmurs, which are associated with mitral or tricuspid regurgitation and ventricular septal defects. This characteristic can help determine whether a patient's heart sounds are normal [9].

2 Related Works

So far, there have been many algorithms and models applied to heart sound classification and detection, but most research has focused on determining whether heart sounds contain noise or whether the heart sounds are normal. Currently, only [10] focuses on the classification and detection of heart murmur quality. To our knowledge, there is currently no comprehensive research on heart murmur quality, shape, grading, and occurrence time, among other related information. The use of Discrete Wavelet Transform (DWT) and Mel-Frequency Cepstral Coefficients for feature extraction, followed by classification using machine learning algorithms, is a common research method. Utilizing deep learning algorithms for heart sound research is currently a trend. Some studies directly use neural networks to extract features and classify the heart sound waveform (PCG), for example[11, 12] ; another approach is to first extract time-frequency feature maps (two-dimensional matrices) and then classify the extracted features, for example [13]. Self-supervised learning and unsupervised learning, compared to supervised learning, do not require a large amount of labeled training data, which can reduce research costs.

After determining whether there are heart murmurs in the heart sounds, uncertain that the heart sounds of the subject are normal is still a problem. From the dataset, we can see that there are a total of 495 subjects with heart murmurs, but only 321 subjects are labeled as abnormal. This indicates that not all subjects with heart murmurs have abnormal heart sounds. Additionally, it is currently known that there is no machine learning model that combines multiple labels and physiological information. The purpose of this study is to detect the types, grades, pitch, timing, and quality of heart murmurs, and to explore how the characteristics of these murmurs contribute to the assessment of abnormal heart sounds. This will aid in the diagnosis of heart diseases and in establishing standards for the quality assessment of heart murmurs. Additionally, we will combine physiological information at the end to determine whether the heart sounds of patients with murmurs are normal. Specifically, the contributions of this study are as follows:

- Design a machine learning model to extract features (first derivative, second derivative) from the log-mel spectrogram of PCG segments, and to classify the types, grades, pitch, occurrence time, and quality of systolic heart sounds. Finally, combine this with physiological information to determine whether the heart sounds are normal.
- Evaluate the advantages and disadvantages of deep learning methods in the classification of heart murmur quality.
- Using deep learning models to explore the relationship between physiological information and heart sound features.

3 Dataset

The dataset used in this study is publicly available and was initiated by the George B. Moody PhysioNet challenge, with data from 2022, named the CirCor DigiScope phonocardiogram dataset. This dataset contains 3,136 phonocardiogram (PCG) recordings from 942 subjects, collected in northeastern Brazil between July and August 2014 and June and July 2015, with a sampling rate of 4000 Hz. Phonocardiogram recordings from the same subject can come from up to five different auscultation sites, but most recordings are concentrated in four main locations: aortic valve (AV), pulmonary valve (PV), tricuspid valve (TV), and mitral valve (MV), with only a few from other auscultation sites. As shown in Table 1, the primary subjects in this dataset are children, with a total of 664 relevant recordings, so this paper will focus on children as the main research subjects. Additionally, we can see from Table 2 that the gender ratio in this dataset is very balanced.

In the .tsv file of the dataset, the segment labels for the phonocardiogram records are as follows: 0 for unmarked signal segments, 1 for the first heart sound, 2 for systole, 3 for the second heart sound, and 4 for diastole. In the subjects' .txt files, the locations and characteristics of the heart murmurs are recorded, including the loudest location, timing, shape, pitch, grading, and quality, all of which are manually labeled by cardiologists. Specifically, the grading of murmurs is based on the Levine grading system, divided into three levels: "I/VI," "II/VI," and "III/VI"; the pitch of the murmur can be categorized as "high," "medium," and "low"; the timing of the murmur is classified as "early systolic," "mid-systolic," and "late systolic"; the shape of the murmur can be described as "crescendo," "decrescendo," "diamond-shaped," and "plateau-shaped"; and in terms of quality, it is labeled as "blowing," "rough," and "musical.

During the data organization process, as shown in Table 3, not all subjects diagnosed with heart murmurs (Abnormal) exhibited obvious murmurs (Present), which was contrary to our expectations. Therefore, in subsequent tests, we will include the data from Table 2 where murmurs are classified as Unknown, as well as those where murmurs are absent but the outcome is Abnormal, into the model testing set. From Table 4, it can be seen that the new murmurs mostly occur during systole and almost never appear during diastole. The only data during diastole comes from the subject with Patient ID 85119, so we will exclude this data. The resulting model will focus on the assessment of systolic heart murmurs.

Table 1. Age Distribution of original dataset.

Age	Count	Percentage
Neonate	6	0.6%
Child	664	70.5%
Infant	126	13.4%
Adolescent	72	7.6%
Unknown	74	7.9%
Total	942	100%

Table 2. Child gender ratio.

Gender	Count	Percentage
Male	65	49.2%
Child	67	50.8%
Total	132	100%

Table 3. Condition of murmur present and outcome.

Condition murmur/outcome	Count	Percentage
Absent/Abnormal	190	28.6%
Absent/Normal	305	45.9%
Present/Abnormal	110	16.6%
Present/Normal	22	3.3%
Unknown/Abnormal	21	3.2%
Unknown/Normal	16	2.4%
Total	664	100%

Table 4. Murmur conditions in Child.

Conditions	Count	Percentage
Which phase		
Systolic	131	99.2%
Diastolic	1	0.8%
Total	132	100%
Systolic murmur timing		
Early-systolic	44	33.6%
Holosystolic	70	53.4%
Late-systolic	1	0.8%
Mid-systolic	16	12.2%
Total	131	100%
Systolic murmur shape		
Crescendo	2	1.5%
Decrescendo	24	18.3%
Diamond	28	21.4%
Plateau	77	58.8%
Total	131	100%
Systolic murmur grading		
I/VI	68	51.9%
II/VI	24	18.3%
III/VI	39	29.8%
Total	131	100%
Systolic murmur pitch		
High	31	23.7%
Low	64	48.9%
Medium	36	27.4%
Total	131	100%
Systolic murmur quality		
Blowing	58	44.3%
Harsh	69	52.7%
Musical	4	3.0%
Total	131	100%

4 Method

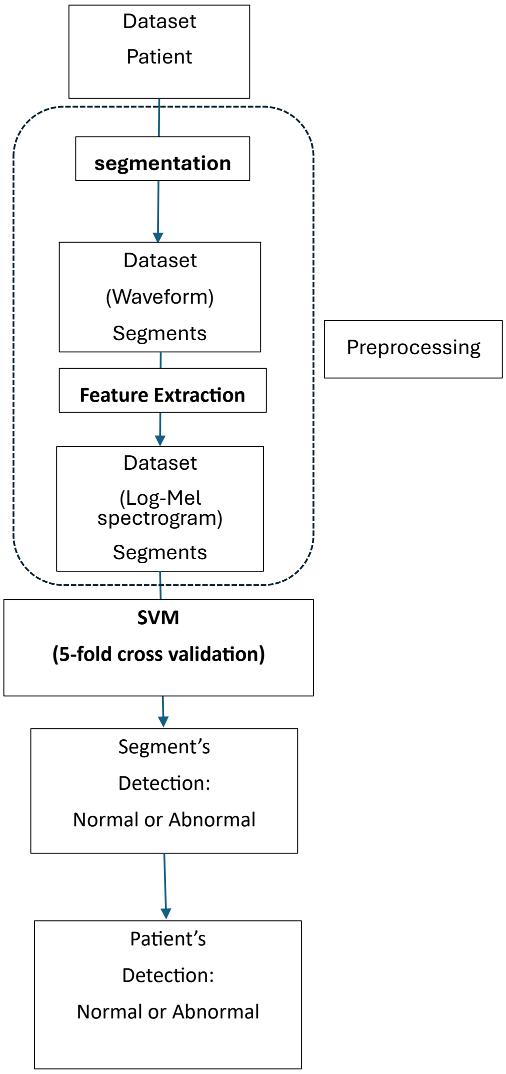

Fig. 1. Steps for detecting whether a subject's heart sound is normal using a Support Vector Machine.

In this study, the heart sound signals of the subjects are segmented into systolic segments, allowing for a more detailed analysis of the heart sound characteristics. After segmentation, feature extraction is performed on these systolic segments, where representative features are derived, such as Mel-frequency cepstral coefficients (MFCC), pitch, spectral energy, and other relevant features that aid in the subsequent classifier training and analysis. Following feature extraction, a Support Vector Machine (SVM) is used to classify these heart sound segments. Finally, the classification results of individual segments are integrated to determine whether the subject's overall heart sound can be classified as normal or abnormal. This approach improves accuracy by focusing on the most critical part of the cardiac cycle and combining segment-level classifications for a holistic assessment.

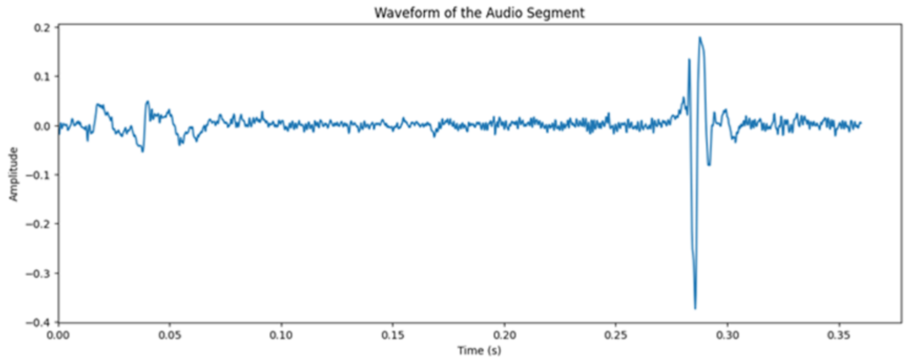

Fig. 2. PCG segment in waveform.

4.1 Feature Extraction

For each PCG segment figure 2., a 2D logarithmic Mel spectrogram is extracted, a method frequently applied in deep learning and machine learning to replicate logarithmic Mel spectrogram will undergo differentiation to obtain the first and second derivatives, capturing dynamic changes in the sound figure 3. These derivative features will then be combined in various configurations to create enriched feature sets. Ultimately, these feature combinations will serve as input for the Support Vector Machine (SVM) classifier, enhancing its ability to distinguish between different heart sound characteristics.

Short-Time Fourier Transform (STFT):

$$X(\omega, t) = \sum_{n=-\infty}^{\infty} x[n]w[n-t]e^{-j\omega n} \tag{1}$$

Power Spectrum:

$$P(\omega, t) = |X(\omega, t)|^2 \quad (2)$$

Mel Scale Conversion:

$$\boldsymbol{m(f) = 2595 \cdot log_{10}\left(1 + \frac{f}{700}\right)} \quad (3)$$

Mel Spectrogram Calculation:

$$\boldsymbol{S_{mel}(m, t) = \sum_{k=1}^{K} P(\omega_k, t) \cdot H_m(\omega_k)} \quad (4)$$

Logarithm of Mel Spectrogram:

$$\boldsymbol{Log - Mel\ Spectrum(m, t) = log(S_{mel}(m, t) + \epsilon)} \quad (5)$$

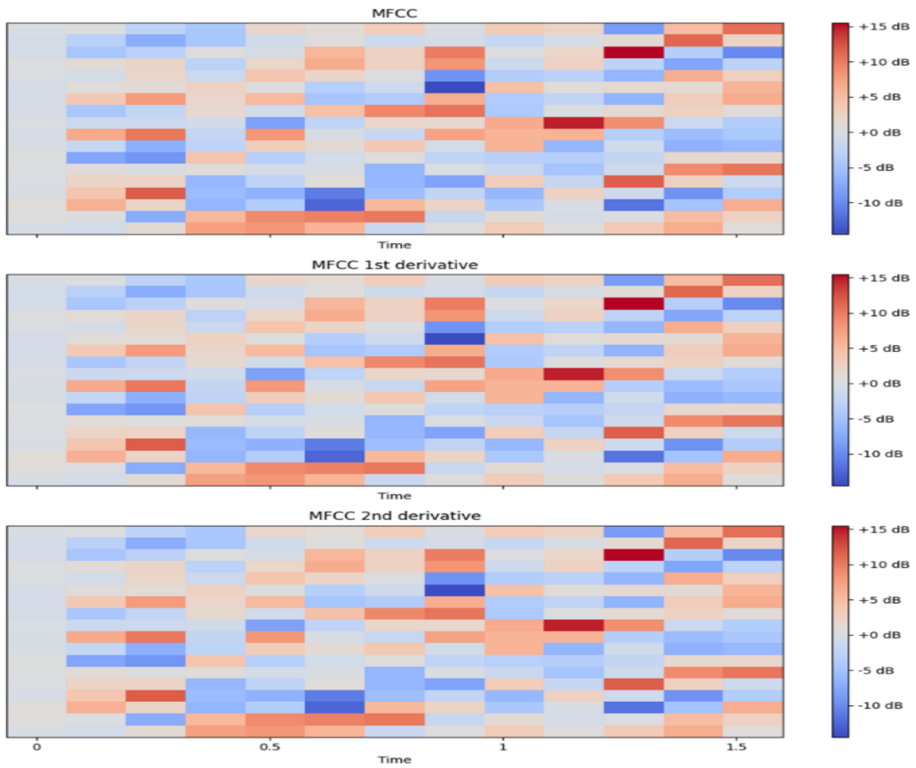

Fig. 3. From top to bottom logarithmic Mel spectrogram, 1st derivative, 2nd derivative.

4.2 SVM Classification

Based on the presence of heart murmurs and the classification of heart sounds as normal or abnormal, the subjects are divided into five groups:

1. Those with heart murmurs and normal heart sounds
2. Those with heart murmurs and abnormal heart sounds
3. Those without heart murmurs and with normal heart sounds
4. Those without heart murmurs and with abnormal heart sounds
5. Others (this group is excluded during classification)

The classification features include physiological information (such as height, weight, and gender) and audio features. The audio features consist of the logarithmic Mel spectrogram, along with its first and second derivatives. These audio features are standardized to a fixed length to ensure consistency and facilitate model training. Finally, the model is trained using five-fold cross-validation, which helps assess its performance and generalizability by partitioning the dataset into five subsets, training the model on four subsets and validating it on the remaining subset for each fold.

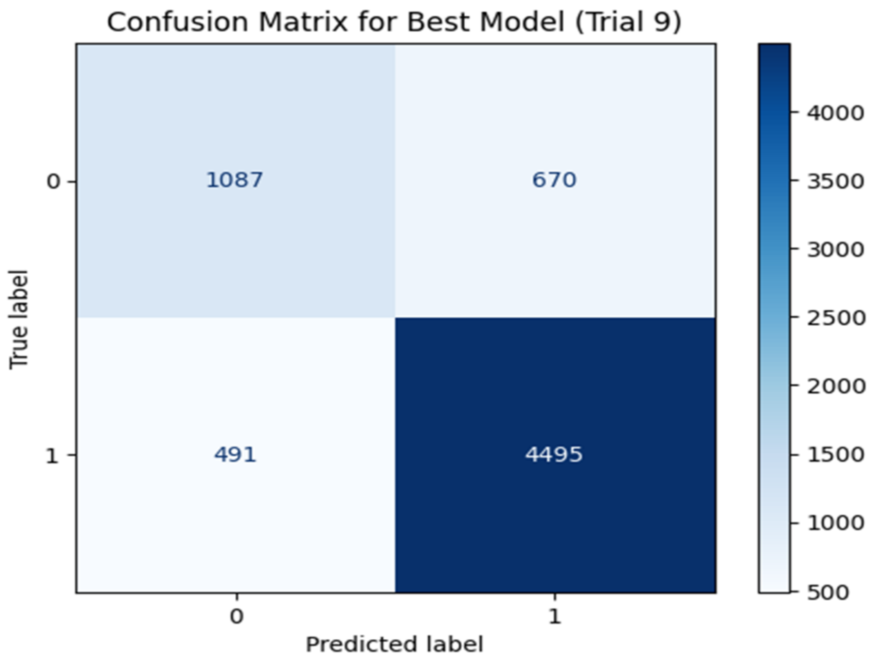

Fig. 4. Confusion matrix of SVM.

5 Conclusion

This study successfully developed a classification model for heart sound analysis by combining both physiological information and dynamic audio features, specifically the second derivative of the Mel spectrogram. The model achieved an impressive 82% accuracy in determining whether a subject's heart sound is normal figure 4. This improvement highlights the crucial role that dynamic changes in heart sound features, such as the first and second derivatives of the Mel spectrogram, play in enhancing the model's classification capability. The inclusion of physiological data, such as height, weight, and gender, provided valuable context to the audio features, allowing the model to better understand the subject's overall condition. By focusing on systolic heart murmurs and incorporating physiological factors alongside the audio features, the model was able to make more informed decisions, particularly in distinguishing between normal and abnormal heart sounds.

Furthermore, the model’s use of five-fold cross-validation ensured robust performance, allowing for generalization across different subsets of the dataset. This cross-validation process helped in assessing the model’s ability to handle diverse heart sound characteristics and physiological variations, which is essential for real-world applications. Incorporating the second derivative of the Mel spectrogram allowed the model to capture dynamic temporal changes in the heart sounds, further refining its ability to distinguish subtle nuances in murmur characteristics. The differentiation of the Mel spectrogram into its first and second derivatives proved to be an effective approach for improving classification accuracy, as it captures changes in the frequency and intensity of heart sounds, which are critical in diagnosing heart conditions.

Overall, this study demonstrates the potential of using a hybrid approach that combines both audio and physiological features to achieve a high level of accuracy in heart sound classification. By leveraging machine learning techniques like Support Vector Machine (SVM) and feature engineering, this approach opens new possibilities for cost-effective and efficient cardiac diagnosis, especially in regions with limited access to specialized healthcare professionals. The findings of this research offer significant insights into the potential of AI-powered auscultation tools for improving early detection and intervention for cardiovascular diseases.

References

1. A. K. Roy, A. Misal, and G. Sinha, "Classification of PCG signals: a survey," Int. J. Comput. Appl, vol. 975, p.8887, (2014).
2. I. R. Hanna and M. E. Silverman, "A history of cardiac auscultation and some of its contributors," The American journal of cardiology, vol. 90, no. 3, pp. 259-267, (2002).
3. J. Oliveira et al., "The CirCor DigiScope dataset: from murmur detection to murmur classification," IEEE journal of biomedical and health informatics, vol. 26, no. 6, pp. 2524-2535, (2021).
4. A. Elola et al., "Beyond heart murmur detection: automatic murmur grading from phonocardiogram," IEEE Journal of Biomedical and Health Informatics, vol. 27, no. 8, pp. 3856-3866, (2023).

5. L. Zühlke, "Computer-assisted auscultation as a screening tool for cardiovascular disease: a cross-sectional study," (2011).
6. J. K. Roy, T. S. Roy, and S. C. Mukhopadhyay, "Heart sound: Detection and analytical approach towards diseases," Modern Sensing Technologies, pp. 103-145, (2019).
7. F. Mokeddem, F. Meziani, and S. Debbal, "Study of murmurs and their impact on the heart variability," International Journal of Medical Engineering and Informatics, vol. 12, no. 3, pp. 291-301, (2020).
8. F. Mokeddem, F. Meziani, and S. Debbal, "Study of murmurs and their impact on the heart variability," International Journal of Medical Engineering and Informatics, vol. 12, no. 3, pp. 291-301, (2020).
9. F. Li, H. Tang, S. Shang, K. Mathiak, and F. Cong, "Classification of heart sounds using convolutional neural network," *Applied Sciences,* vol. 10, no. 11, p. 3956, (2020).
10. A. K. Dwivedi, S. A. Imtiaz, and E. Rodriguez-Villegas, "Algorithms for automatic analysis and classification of heart sounds–a systematic review," IEEE Access, vol. 7, pp. 8316-8345, (2018).
11. C. Turner and A. Joseph, "A wavelet packet and mel-frequency cepstral coefficients-based feature extraction method for speaker identification," *Procedia Computer Science,* vol. 61, pp. 416-421, (2015).
12. . K. Abdul and A. K. Al-Talabani, "Mel frequency cepstral coefficient and its applications: A review," *IEEE Access,* vol. 10, pp. 122136-122158, (2022).
13. D. M. Nogueira, C. A. Ferreira, E. F. Gomes, and A. M. Jorge, "Classifying heart sounds using images of motifs, MFCC and temporal features," *Journal of medical systems,* vol. 43, no. 6, p. 168, (2019).