

Heart Murmur Detection and Clinical Outcome Prediction with a Hybrid Deep Learning Model using Phonocardiogram Recordings

Ehsan Karegar Foroogh
Department of Mechatronics, School of Intelligent
Systems, College of Interdisciplinary Science and
Technology
University of Tehran
Tehran, Iran
ehsan.karegarfo@ut.ac.ir

Farshid Hajati*
School of Science and Technology, Faculty of Science,
Agriculture, Business and Law
University of New England
Armidale, Australia
fhajati@une.edu.au

Amirhosein Zobeiri
Department of Mechatronics, School of Intelligent
Systems, College of Interdisciplinary Science and
Technology
University of Tehran
Tehran, Iran
ah.zobeiri@ut.ac.ir

Alireza Rezaee
Department of Mechatronics, School of Intelligent
Systems, College of Interdisciplinary Science and
Technology
University of Tehran
Tehran, Iran
arrezae@ut.ac.ir

Abstract

Cardiovascular diseases (CVDs) are a leading cause of global mortality, with millions of deaths annually, particularly in low- and middle-income countries. Early diagnosis is crucial for effective management and improved outcomes. Heart murmurs, caused by turbulent blood flow in the heart, are key indicators of certain CVDs, such as valve defects and congenital heart disorders. We developed a hybrid model with a combination of CNN and LSTM layers for murmur detection and clinical outcome prediction on a dataset including PCG recordings and demographic information of 942 patients from George B. Moody PhysioNet Challenge 2022 dataset. Our murmur detection model uses only PCG recordings, but the model trained for clinical outcome prediction, uses demographic information as well. We used focal categorical cross entropy for training phase of murmur detection, while trained clinical outcome prediction model using categorical cross entropy. Our hybrid CNN-LSTM model could reach weighted accuracy of 0.82 for murmur detection and clinical cost metric of 10786 for outcome prediction, outperforming leading models developed in Physionet challenge.

CCS Concepts

• **Applied computing**; • **Health care information systems**; • **Information systems applications**;

Keywords

Murmur Detection, Deep Learning, Phonocardiogram

*Corresponding Author.



This work is licensed under a Creative Commons Attribution 4.0 International License. *ICMHI 2025, Kyoto, Japan*

© 2025 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-1514-3/2025/05

<https://doi.org/10.1145/3761712.3761724>

ACM Reference Format:

Ehsan Karegar Foroogh, Amirhosein Zobeiri, Farshid Hajati, and Alireza Rezaee. 2025. Heart Murmur Detection and Clinical Outcome Prediction with a Hybrid Deep Learning Model using Phonocardiogram Recordings. In *2025 9th International Conference on Medical and Health Informatics (ICMHI 2025)*, May 16–18, 2025, Kyoto, Japan. ACM, New York, NY, USA, 7 pages. <https://doi.org/10.1145/3761712.3761724>

1 Introduction

Cardiovascular diseases (CVDs) encompass a diverse spectrum of disorders that disrupt the integrity and functionality of the heart and its blood vessels, ranging from coronary heart disease (CHD) to rheumatic heart disease (RHD) [1-3]. These conditions not only contribute significantly to morbidity but also serve as the leading global cause of mortality, responsible for nearly one-third of all deaths worldwide [4].

Heart murmurs are critical auscultatory phenomena that serve as potential markers of cardiovascular abnormalities [5]. Turbulent blood flow, commonly caused by structural abnormalities or functional impairments in the heart or its vessels, is a key factor in the development of heart murmurs [5, 6]. The two main categories of these audio signals are innocent or physiological murmurs and abnormal or pathologic murmurs [7]. Common, especially in children, innocent murmurs are usually harmless and do not present any serious health risks [8]. Conversely, abnormal murmurs suggest the presence of underlying cardiovascular disorders and necessitate immediate medical intervention. They may occur as a result of different factors, such as congenital heart disease, valvular disorders, or as complications of other systemic illnesses [9]. For example, abnormal murmurs can result from stenosis which is the narrowing of the heart valves or from regurgitation which is the backflow of blood through valves that do not close properly [10].

Aside from identifying heart diseases, the type, location, and timing of the murmur gives almost all the needed information to determine the severity and nature of the condition as well as guide in determining the treatment strategies. Therefore, the timely

identification and classification of heart murmurs are essential for management of problems involving the cardiovascular system. In a clinical setup, one of the important cardiovascular examination procedures is to listen to the heart sounds using a stethoscope and identify any unexpected anomalies like murmurs. Because of being simple, inexpensive, and non-invasive, this technique continues to be used for actual practice even after many decades [11]. Using this technique is very demanding with regard to training and expertise due to the overly intricate nature of heart sounds, the overlapping frequency and the background sounds, which make diagnosis particularly difficult with faint murmurs [12, 13]. Additionally, the accuracy of diagnostic outcomes is highly dependent on the subjective evaluation of different heart sounds and the practitioner's expertise performing the examination. A physician is required to be present during the processes of auscultation which makes this method impractical for mass screening programs over remote locations. Because of these limitations, there has been a growing focus on the development of automated techniques to improve the accuracy and alternates of the heart murmur diagnosis in a non-invasive manner [14, 15].

The development of artificial intelligence (AI) technology like neural networks (NNs) has significantly increased the complexity of medical diagnostics processes, and has changed the entire spectrum of the healthcare industry [16]. NNs are proficient in the processing of complex and large datasets, and with the application of their self-learning capabilities, can discover and apply complex patterns in clinical data. This capability, in turn, has enabled the prognosis and diagnosis of several categories of diseases like cancer, cardiovascular diseases, and serious neurological disorders [17-19].

Phonocardiograms (PCGs) hold significant value in the automated diagnosis of heart issues. They serve as raw data for testing different AI algorithms designed to detect murmurs in heartbeat sounds. Although there are a number of public PCG datasets, most are limited due to small sample sizes, excessive localization, and poor quality of labelling for both normal and abnormal murmurs [20, 21]. For example, the PASCAL CHSC 2011 [20] and PhysioNet 2016 [21] datasets have been extensively used in research; however, neither includes labelled murmurs, which restricts their utility for specific diagnostic applications. The PhysioNet 2022 challenge dataset [22, 23] attempts to solve this problem by adding labelled murmur data and clinical outcomes. This data encompasses recordings of heart sounds from 1,568 people, together with their age, sex and significant medical history.

Our aim in this work is to devise a strong deep learning model that can identify heart murmurs and estimate clinical outcomes simultaneously. For this reason, we designed a hybrid model that includes both CNNs and LSTMs and uses their complementary strengths together. Capturing spatial features such as frequency patterns and amplitude in sounds, CNNs are strong, whereas LSTMs are particularly useful for seeing temporal connections in heart sound sequences and tracking changes over time.

To address the challenges of class imbalances in the dataset, we employ a modified focal loss function that prioritizes harder-to-classify samples. Additionally, we explore how using demographic information into the model alters the results. Our analysis of heart sound recordings compares the effects of patient-specific factors in diagnostic measures, whether we have their information or not.

Not only did our models exceed the challenge benchmarks, but they also achieved higher accuracy, F1-score and cost for clinical outcomes. These results point out the usefulness of these models in cases where reaching specialists is a challenge. The objective is to automate cardiovascular diagnostics by inventing advanced approaches for heart murmur detection and outcome prediction, to fill the gap and bring such automated services to more people.

2 Related Works

Extensive research has been done regarding the automation of diagnosing heart diseases using PCG signals and AI-driven techniques, highlighting considerable studies with varying classification methods.

Choi and Jiang [24] proposed a new methodology in cardiac sound analysis where feature extraction was performed using Normalized Auto Regressive Power Spectral Density (NARPSD) curve and classification was done via Support Vector Machine (SVM). By normalizing the power spectral density with auto-regressive determination, their method reduced noise and enhanced feature reliability.

Yaseen et al. [25] studied heart sound classification utilizing features extracted from PCG recordings. They included Mel-Frequency Cepstral Coefficients (MFCCs) and Discrete Wavelet Transform (DWT) feature extraction. SVM, DNN, and centroid-based KNN classifiers were used to test the performance evaluation. The first attempts performed with MFCCs achieved some promising results, however, the best results were obtained when combining MFCCs with DWT features. By employing SVM with fused features, maximum accuracy achieved was 97.9% and F1 score of 99.7%.

Alkhodari and Fraiwan [26] performed heart sound classification using WT for feature extraction in combination with a hybrid deep learning model to classify heart sounds as normal or abnormal. The features were subsequently processed through a 1D-CNN, which autonomously discovered and extracted indicative spatial patterns. To improve temporal analysis, the output of 1D-CNN was input into a Bidirectional Long Short-Term Memory (Bi-LSTM) network.

Chakraborty et al. [27] trained a 3-layered 2D-CNN to work with the obtained spectrograms and created a deep learning architecture to enhance the prediction process. They achieved high accuracy in distinguishing the three classes, demonstrating the potential of spectrograms for representing heart sounds while retaining their defining features.

Lubis and Gondawijaya [28] analyzed Bark Frequency Cepstrum Coefficients (BFCCs) and MFCCs for feature extraction in heart sound analysis using backpropagation neural network and made comparisons. Both sets of features were useful, but MFCCs were shown to provide better results for overall accuracy.

To find heart murmurs and predict their effects, a model of stacked Bi-LSTM was suggested by Monteiro et al. [29]. his model was designed for PCG recordings simply because it made the most of Bi-LSTMs which excel at capturing series trends over time. Because sound signals have intricacy temporal relationships, this bidirectional processing improved the model's knowledge of these details.

McDonald et al. [30] used Hidden Semi-Markov Models (HSMMs) as approaches of heart murmur detection. The goal was

to make the models capture how the duration of these sounds changes through each part of the cycle, that is, systolic and diastolic. The method modeled both normal and murmur heart sounds together using parallel HSMs, resulting in better ability to accurately set differences by using probability. The results displayed strong discriminatory ability to distinguish between the presence and absence of a murmur, relying on high sensitivity and specificity in both subgroups. A CatBoost gradient-boosted tree was applied by the authors to forecast clinical outcomes based on the characteristics of heart sounds and information about the patient such as age and pregnancy status.

3 Materials and Methods

3.1 Dataset

We conducted our study using the dataset provided by the PhysioNet 2022 Challenge [22, 23]. The training set has 942 open-source heart sound samples plus corresponding demographic information about each patient. The files included with the data have complete information on the study’s annotations and metadata. In the dataset, age, sex, height, weight and pregnancy are all considered part of demographic characteristics. Besides explaining the heart sound recordings and the location of the murmur, the report marks the murmur and lists what clinical results were found. A frequency of 4000 Hz is used to sample the heart sound recordings, ensuring high-resolution audio data. By using the dataset’s detailed tags and structure, researchers are able to investigate the relationships between age, sex, heart sound features and the results people have experienced after treatment. Additionally, the diversity among the patients and how well murmur events are described in the records means the dataset can be used for different machine learning based methodologies and support more progress in cardiovascular research.

3.2 Preprocessing

The initial sampling rate of 4000 Hz put a large strain on the system and reduced its ability to learn and apply models in different scenarios. To solve this problem, we pulled the sampling rate down to 800 Hz, so that the heart sounds were still useful for the detection task and would reduce the amount of time the software requires to process the data [31]. After down-sampling, we scale each heart sound recording, ensuring that the data has a mean of zero and a standard deviation of one.

Next, we employed a sliding window technique to segment each heart sound recording into fixed-length 4-second segments. This strategy not only helps train the model on smaller, more manageable chunks of data, but also increases the amount of training data available, improving the robustness of the model. By sliding the window across the heart sound signal with some overlap, we generate multiple data points from each recording, allowing the model to capture more nuanced patterns in the heart sounds.

Since the murmur classes are significantly imbalanced, we used different overlaps for the sliding windows corresponding to the present, unknown, and absent murmur labels. For patients with a present murmur, we applied an 80% overlap, which allowed the model to capture more relevant features from these recordings. For patients with an unknown murmur, we used a 92.5% overlap

to provide additional context and maximize the model’s ability to identify subtle patterns, given the ambiguity of the label. Finally, for patients with an absent murmur, we set no overlap in the sliding window, as these recordings were already considered less complex and needed less augmentation. These overlapping strategies were specifically tailored for the training phase to enhance model learning, while during the testing and prediction phases, a 50% overlap was applied to ensure consistent segmentation of heart sound signals. After applying the sliding windows, the distribution of the training set was as follows: 34.29% of the data belonged to the “present” murmur class, 29.57% to the “unknown” class, and 36.13% to the “absent” class. It is also important to note that heart sound recordings from patients with a present murmur, but recorded in locations where murmurs were absent, were excluded from the training phase. These recordings did not contain relevant murmur information and could confuse the model by introducing misleading patterns. Therefore, we removed them to improve the model’s accuracy and prevent overfitting to irrelevant data.

On the other hand, the clinical outcome classes, which categorize the patients as either normal or abnormal based on their heart condition, were balanced at the beginning, ensuring an equitable representation of both classes. This balance was maintained even after the sliding window technique was applied. After segmenting the heart sound recordings, 59.14% of the training set belonged to the abnormal class, while 40.86% belonged to the normal class. Despite this slight imbalance, the ratio between the normal and abnormal classes was still considered reasonable for effective training, allowing the model to learn the patterns associated with both classes without significant bias.

To incorporate demographic information into the model, we first addressed any missing values using the “SimpleImputer”, which automatically fills in the missing entries based on the mean, median, or most frequent value of the respective feature. Additionally, we applied the “StandardScaler” to age, height, and weight to scale these numerical features to have a mean of zero and a standard deviation of one. This preprocessing stage helps enhance model performance and stability by making all the demographic variables attend equally to the prediction task.

3.3 Label Processing

We labeled each clip of 4-second duration with the same patient label, both for murmur and clinical outcome labels. This process of labeling ensured that the labels would remain aligned with the patient-level nature of our training model approach. By keeping patient-level assignments within each segment, we provided the necessary contextual consistency for the model to predict correctly at the recording level. This kept the correlation between heart sound recordings and the corresponding clinical outcomes intact while training the model.

Following the labeling procedure, the segmented recordings were randomized to avoid any biases related to their initial position in the dataset. After randomization, we performed the split of the data into training and validation sets. The validation set was meant to be 20% of the entire dataset, thus ensuring a representative distribution for the model performance evaluation.

3.4 Loss Function

Focal Loss is simply a more flexible version of the categorical cross-entropy which is designed to mitigate class imbalance in training datasets. By introducing a class weighting coefficient (α) to the loss function, it dynamically modifies the function and illuminates model performance on imbalanced data. The weighting coefficient α allows the loss to be contoured or scaled based on the imbalance of classes, addressing the imbalance in data and constructing an appropriate loss function to avoid favoring the dominant class. The representation of focal loss is seen in (1) - α specifying the scaling of the loss, which invariably changes the contribution of previously underrepresented classes to the scaling of the total loss.

$$\text{Focal Loss} = - \sum_{i=0}^C \alpha_i \log(p_i) \quad (1)$$

During our testing, we made a clear choice to use the weighting factor (α) so that we could accurately weight our dataset. This really mattered to us based on the differences in class sizes within our training dataset. We gave each class values for α .

α value for murmur, unknown, and absent classes was set to 1.06, 1.2, and 1, respectively, and the reasoning behind our selection of these values related to the scarcity or abundance of the respective labels in the training set, so we could also give all labels equal rank in importance during training. α values for the abnormal and normal classes were set to 0.41 days. and 0.59 days. Again, we reasoned these numbers to help balance the class differences, so that the model would have a balanced loss space to learn from.

What we had is a very powerful focal loss function that was tuned in such a way to balance classes and mitigate the consequences of class imbalance in the dataset. Because we were able to tune α , we ensured that the model initially understood the minority classes prior to those majority classes significantly impacting or dominating learning through the training cycle. This approach empowered the model to better learn more subtle differences in the data and to generalize better, ultimately learning about all classes equally at evaluation.

3.5 Models

We combined 1D CNNs with LSTMs to process the sequential time steps in the PCG recordings, followed by dense layers to help us better learn complex patterns within the data. This hybrid model was set up to handle murmur detection and clinical outcome prediction by using relations between points in both spatial patterns and temporal passage of time.

To predict patient's outcome, we applied both PCG recordings and demographic details, doing so in separate steps to avoid confusion when blending the different kinds of information. Initially, we sent the PCG recording into the model that used the same initial processing as the murmur detection model. For this, a CNN layer is used to get the heart sound features, MaxPooling reduces the size of these features, and two LSTM layers are applied subsequently to detect how the heart sound changes over time.

Following the PCG recording, we included the patient's age, sex, height and weight in our analysis. After that, the output from the PCG recording layers was connected with this information, so the model would consider both sound data and patient-specific factors.

Following this, we had a dense layer with 5 nodes to refine the mix of features and then a dense layer with 2 nodes to improve the predictions of the model. This last layer of the network relied on a softmax activation function to return the chances of each clinical outcome class. With two nodes in the output layer, the model is able to classify recordings either as normal or abnormal by using both PCG signals and demographic information.

Used in this way, the model can take advantage of the timing of the PCG scans and the details in the patient's profile, leading to an improved prediction of possible results. Figure 1 explains how the model has been built. We have employed both categorical cross entropy and categorical focal cross entropy to train this model.

To detect murmurs, the model was trained using the PCG recordings as input. In the first part of the model a 1D CNN layer was used, which excels at identifying spatial characteristics within raw audio signals. After that, MaxPooling was used to lower the number of features while making sure the main spatial data does not get lost. Then, two LSTM layers were added to spot temporal relationships and sequential changes in the heart sound recordings. Also, adding two more thick layers allowed the model to better pick up on the intricacies and complex representations in the data. The output layer is a dense layer made up of just three nodes, one for every class (present, absent and unknown murmur). In the last dense layer, softmax is used to convert the output into a probability distribution for the three classes.

This design means the model learns important features both from the location and time stamps in the PCG and the dense layers then refine the outcome of these readings. Since softmax returns probabilities, it is suitable for categorizing murmurs from different types. Figure 2 presents the architecture and the process of this model.

3.6 Prediction

Since this is a health-related problem, we made sure our work was done with great accuracy. We split every PCG recording of the patients into 4s segments. After splitting the signals, we further segmented by a sliding window with a 50% overlap to capture subtler patterns that appeared across these overlapping parts.

Next, we took the mean of all probability values which were predicted for each 4-second section. Because of this approach, the prediction for the whole recording remained stable and accurate, minimizing the effect of any single misclassified segment.

To get the final prediction for the PCG recording, we chose the one with the biggest probability for the "present" or "abnormal" class. As a result, the longest prediction for each recording became the final answer, as it is especially vital for health applications that errors can harm patients. By following this method, both the accuracy of the results and the attention to patients' interests were kept at their best.

4 Results and Discussion

The models were evaluated through different measures suggested by the PhysioNet Challenge 2022. For the murmur detection (Task

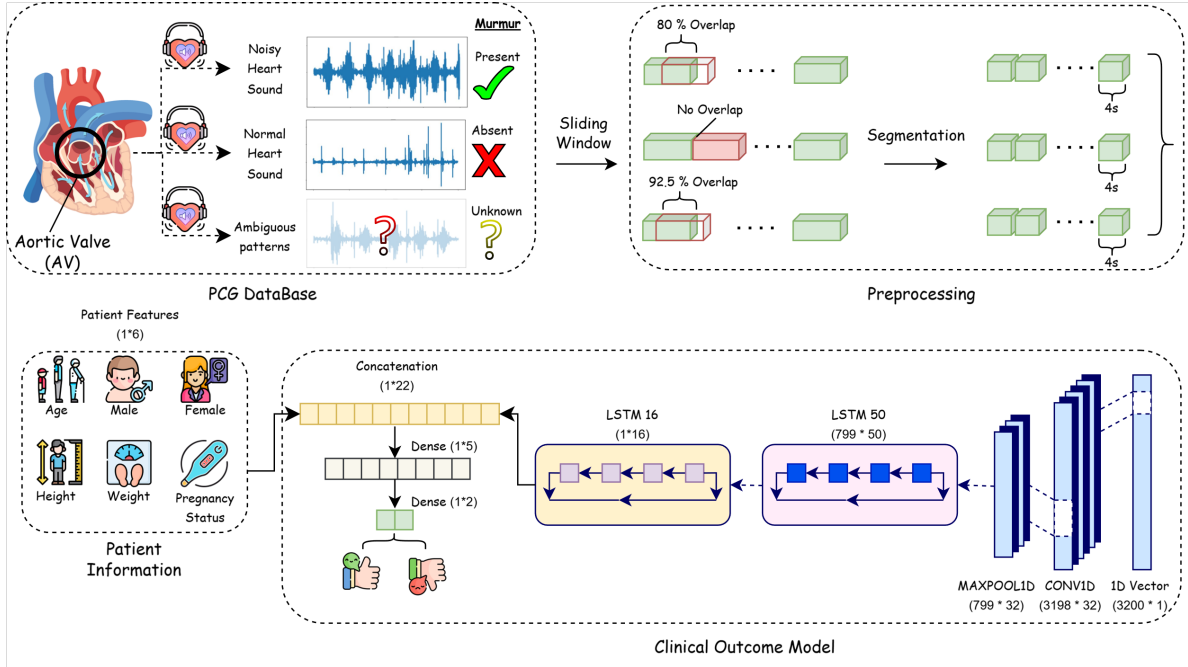


Figure 1: The Hybrid Model for Clinical Outcome

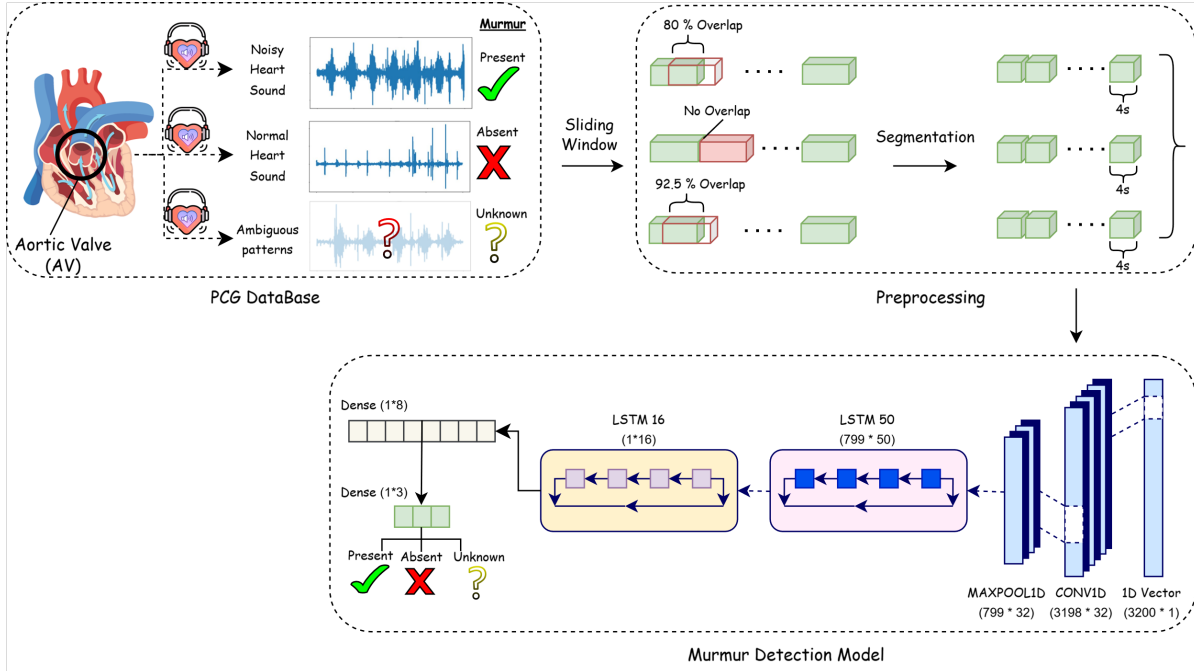


Figure 2: The Hybrid Model for Murmur Detection

I), weighted accuracy (WAcc) was considered as the proposed evaluation measure and is defined as below:

$$S_{outcome} = \frac{5\eta_{TP} + \eta_{TN}}{5(\eta_{TP} + \eta_{FN}) + (\eta_{FP} + \eta_{TN})} \quad (2)$$

As you can see in the Figure 3, The model did well in classifying the Present and Unknown murmur classes. The sensitivity for Present was 0.93, while the sensitivity for Unknown murmurs was

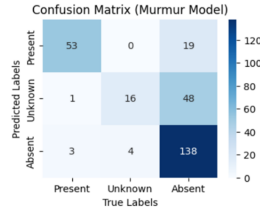


Figure 3: The test set Confusion Matrix for murmur detection model

0.80. These high sensitivities indicate the model usefulness for detecting both present and unknown murmurs diagnosis skills.

Additionally, a specificity of 0.67 indicates the model can detect “Absent” murmur class correctly, reflecting its reasonable ability to correctly identify recordings where no murmur is present, minimizing false positives.

The weighted accuracy of this model stands at 0.82 which surpasses the weighted accuracy performance of every other model tested in the challenge. By obtaining such excellent results the model proves its ability to correctly classify heart murmurs within different classes while demonstrating robust performance to handle data imbalances and maintain high accuracy across all categories.

Our model was trained for clinical outcome prediction based on both categorical cross-entropy (CCE) and focal categorical cross-entropy (FCCE). As you can see from Figure 4, the confusion matrix in FCCE (b) displays better class balance between different classes. This is largely due to the class weights we used during training, which helped address the class imbalance and improved the model’s ability to differentiate between the different murmur categories. The model trained with this approach reached a mean cost of 13862, which indicates a balanced performance across all classes.

On the other hand, when we used categorical cross-entropy after segmentation with the sliding window technique, the model achieved higher sensitivity but lower specificity. This approach resulted in a mean cost of 10786. The increase in sensitivity is beneficial, especially for detecting the presence of murmurs, but the lower specificity indicates a higher rate of false positives, where the model incorrectly predicts the presence of a murmur when there is none.

Although the model trained with categorical cross-entropy is less balanced and has lower specificity, it performs well in terms of the cost metric. This is because the abnormal class has a larger influence on the cost metric, and the model’s ability to detect abnormal cases contributes significantly to lowering the cost. With a sensitivity of 0.92 and a specificity of 0.34, this model can outperform other leading models in the challenge, particularly when minimizing the cost associated with detecting abnormal murmurs.

Table 1 presents a comparison of our study’s results with those of other articles in terms of weighted accuracy and clinical cost.

5 Conclusions

It was clear from our study that making use of Hybrid CNNs and LSTMs layers gave us effective results for both murmur detection and clinical outcome predictions using the PCG recordings. This approach achieved the highest scores in the PhysioNet 2022 Challenge compared to all other approaches. General feature extraction through CNN layers occurs effectively because these layers detect spatial patterns in PCG signals which include both frequency components and localized signal features. The presence of these abilities is necessary to find specific patterns in heart signals that may mean there is a murmur. Using two LSTM layers, the model can now recognize complex changes in data over time period. Doing this allows medical professionals to notice small changes in heart murmurs and cardiac outcomes as the heartbeat progresses. LSTM layers in the model help monitor how heart sounds vary from one time to another, often pointing to different heart illnesses.

When age, sex, height, weight and pregnancy status were added in addition to the PCG signals, the model performed more accurately in clinical studies. Thanks to demographic information, the model can deduce which clinical conditions could be reflected in heart sounds. For instance, during pregnancy or the first few months of life, there are usually changes in what the heart sounds like. A better view of patient conditions is achieved when both PCG signals and demographic information are considered in the model.

This system indicates that feature extraction along with temporal modeling are both key to handle complex biomedical signal analysis tasks. Additionally, when we merge heart sound data with patient demographics, it seems that whole-scale machine learning can play an important role in clinical work. Not only do such algorithms improve diagnostic accuracy, but they provide hope for wider use in personalized healthcare, where patient-specific factors are very

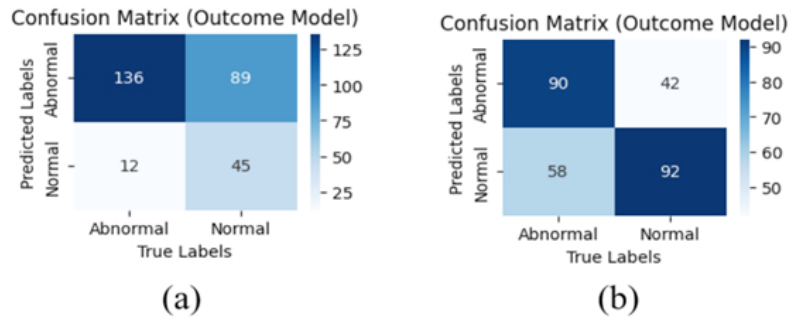


Figure 4: Confusion Matrixes for clinical outcome prediction using (a) CCE and (b) focal CCE

Table 1: Performance Comparison of our Proposed Method with other models in the challenge

Authors	Model	Weighted accuracy	Cost
Lu et al. [32]	CNN	0.793	11731
Gemke et al. [33]	Bi-LSTM	0.64	11386
Monteiro et al. [29]	Bi-LSTM	0.652	12434
Singstad et al. [31]	CNN	0.497	12315
McDonald et al. [30]	RNN	0.817	10565
Our Study	CNN-LSTM	0.82	10786

important in treating and managing diseases. Applying this detailed approach makes it possible to improve and better tailor healthcare programs which can potentially help in treating heart murmurs as well as different systemic heart problems.

Acknowledgments

This work was supported by the University of New England, Australia, under grant number A24/2676.

Author Statement

This paper was written entirely by the authors without the use of any artificial intelligence (AI) tools.

References

- [1] R. B. P. Libby, D. Mann, and D. Zipes, *Braunwald's Heart Disease: A Textbook of Cardiovascular Medicine*, 8 ed. Elsevier Science, 2007.
- [2] E. J. Benjamin et al., "Heart Disease and Stroke Statistics—2019 Update: A Report From the American Heart Association," *Circulation*, vol. 139, no. 10, pp. e56-e528, 2019.
- [3] E. Marijon, M. Mirabel, D. S. Celermajer, and X. Jouven, "Rheumatic heart disease," *The Lancet*, vol. 379, no. 9819, pp. 953-964, 2012, doi: 10.1016/S0140-6736(11)61171-9.
- [4] M. Di Cesare et al., "The Heart of the World," *Global Heart*, 2024, doi: 10.5334/gh.1288.
- [5] W. C. Members et al., "2020 ACC/AHA guideline for the management of patients with valvular heart disease: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines," *Journal of the American College of Cardiology*, vol. 77, no. 4, pp. e25-e197, 2021.
- [6] B. Erickson, *Heart sounds and murmurs across the lifespan*, 4 ed. Elsevier, 2003.
- [7] J. Pedrosa, A. Castro, and T. T. Vinhoza, "Automatic heart sound segmentation and murmur detection in pediatric phonocardiograms," in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2014: IEEE, pp. 2294-2297.
- [8] L. Mesrobian and Y. Sanil, "Innocent heart murmurs from the perspective of the pediatrician," *Pediatric Annals*, vol. 45, no. 8, pp. e306-e309, 2016.
- [9] M. S. Ahmad, J. Mir, M. O. Ullah, M. L. U. R. Shahid, and M. A. Syed, "An efficient heart murmur recognition and cardiovascular disorders classification system," *Australasian physical & engineering sciences in medicine*, vol. 42, pp. 733-743, 2019.
- [10] "Corrigendum to: 2017 ESC/EACTS Guidelines for the management of valvular heart disease," *European Heart Journal*, vol. 39, no. 21, pp. 1980-1980, 2017, doi: 10.1093/eurheartj/ehx636.
- [11] S. Mangione, "Cardiac auscultatory skills of physicians-in-training: a comparison of three English-speaking countries," *The American journal of medicine*, vol. 110, no. 3, pp. 210-216, 2001.
- [12] D. S. Panah, A. Hines, and S. McKeever, "Exploring the impact of noise and degradations on heart sound classification models," *Biomedical Signal Processing and Control*, vol. 85, p. 104932, 2023.
- [13] H. Pasterkamp and H. Melbye, "Machines Are Learning Chest Auscultation. Will They Also Become Our Teachers?," *CHEST Pulmonary*, vol. 2, no. 4, 2024, doi: 10.1016/j.chpulm.2024.100079.
- [14] R. u. S. Ahmad, M. S. Khan, M. E. Hilal, B. Khan, Y. Zhang, and B. L. Khoo, "Advancements in wearable heart sounds devices for the monitoring of cardiovascular diseases," *SmartMat*, vol. 6, no. 1, p. e1311, 2025.
- [15] J. Prince et al., "Deep Learning Algorithms to Detect Murmurs Associated With Structural Heart Disease," *Journal of the American Heart Association*, vol. 12, no. 20, p. e030377, 2023, doi: doi:10.1161/JAHA.123.030377.
- [16] E. J. Topol, "High-performance medicine: the convergence of human and artificial intelligence," *Nature Medicine*, vol. 25, no. 1, pp. 44-56, 2019/01/01 2019, doi: 10.1038/s41591-018-0300-7.
- [17] G. Petmezias et al., "Recent advancements and applications of deep learning in heart failure: A systematic review," *Computers in Biology and Medicine*, p. 108557, 2024.
- [18] H. M. Rai, "Cancer detection and segmentation using machine learning and deep learning techniques: a review," *Multimedia Tools and Applications*, vol. 83, no. 9, pp. 27001-27035, 2024/03/01 2024, doi: 10.1007/s11042-023-16520-5.
- [19] A. Zobeiri, A. Rezaee, F. Hajati, A. Argha, and H. Alinejad-Rokny, "Post-Cardiac arrest outcome prediction using machine learning: A systematic review and meta-analysis," *International Journal of Medical Informatics*, p. 105659, 2024.
- [20] N. G. Bentley P. Coimbra M, and Mannor S. *The PASCAL Classifying Heart Sounds Challenge (CHSC2011)*. [Online]. Available: <http://www.peterjbentley.com/heartchallenge/index.html>
- [21] Physionet Challenge [Online] Available: <https://physionet.org/content/challenge-2016/1.0.0/>
- [22] 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.
- [23] M. A. Reyna et al., "Heart murmur detection from phonocardiogram recordings: The george b. moody physionet challenge 2022," *PLOS Digital Health*, vol. 2, no. 9, p. e0000324, 2023.
- [24] S. Choi and Z. Jiang, "Cardiac sound murmurs classification with autoregressive spectral analysis and multi-support vector machine technique," *Computers in biology and medicine*, vol. 40, no. 1, pp. 8-20, 2010.
- [25] Yaseen, G.-Y. Son, and S. Kwon, "Classification of heart sound signal using multiple features," *Applied Sciences*, vol. 8, no. 12, p. 2344, 2018.
- [26] M. Alkhodari and L. Fraiwan, "Convolutional and recurrent neural networks for the detection of valvular heart diseases in phonocardiogram recordings," *Computer Methods and Programs in Biomedicine*, vol. 200, p. 105940, 2021.
- [27] D. Chakraborty, S. Bhattacharya, A. Thakur, A. R. Gosthipaty, and C. Datta, "Feature extraction and classification of phonocardiograms using convolutional neural networks," in *2020 IEEE 1st International Conference for Convergence in Engineering (ICCE)*, 2020: IEEE, pp. 275-279.
- [28] C. Lubis and F. Gondawijaya, "Heart sound diagnose system with BFCC, MFCC, and backpropagation neural network," in *IOP conference series: materials science and engineering*, 2019, vol. 508, no. 1: IOP Publishing, p. 012119.
- [29] S. Monteiro, A. Fred, and H. P. da Silva, "Detection of heart sound murmurs and clinical outcome with bidirectional long short-term memory networks," in *2022 Computing in Cardiology (CinC)*, 2022, vol. 498: IEEE, pp. 1-4.
- [30] A. McDonald, M. J. Gales, and A. Agarwal, "Detection of heart murmurs in phonocardiograms with parallel hidden semi-markov models," in *2022 Computing in Cardiology (CinC)*, 2022, vol. 498: IEEE, pp. 1-4.
- [31] B.-J. Singstad, A. M. Gitau, M. K. Johnsen, J. Ravn, L. A. Bongo, and H. Schirmer, "Phonocardiogram classification using 1-dimensional inception time convolutional neural networks," in *2022 Computing in Cardiology (CinC)*, 2022, vol. 498: IEEE, pp. 1-4.
- [32] H. Lu et al., "A lightweight robust approach for automatic heart murmurs and clinical outcomes classification from phonocardiogram recordings," in *2022 Computing in Cardiology (CinC)*, 2022, vol. 498: IEEE, pp. 1-4.
- [33] P. Gemke, N. Spicher, and T. Kacprowski, "An LSTM-based Listener for Early Detection of Heart Disease," in *2022 Computing in Cardiology (CinC)*, 2022, vol. 498: IEEE, pp. 1-4.