Classification of Murmurs in PCG Using Combined Frequency Domain and Physician Inspired Features

Julia Ding[1], Jing-Jing Li[2], Max Xu[3]

[1] Emory University, Atlanta, GA, USA
[2] University of California at Berkeley, Berkeley, CA, USA
[3] Georgia Institute of Technology, Atlanta, GA, USA

Abstract

Physiological machine learning methods have a unique opportunity to augment deep-learning engineered features with additional features derived from prior pathological knowledge. We propose an phonocardiogram (PCG) classifier that combines raw spectrogram features with crafted, physician-inspired features with an end-to-end neural network architecture. Learned spectrogram features were obtained by training a convolutional neural network (CNN) directly on the raw mel-spectrogram representation of the PCG time-series. Crafted features were based on the four stages of the cardiac cycle (S1, systole, S2, and diastole).

The spectrogram features have the advantage of introducing flexibility for the model to learn abstract, low-level information that captures a variety of different rhythmic abnormalities and the latter has the advantage of using segmentation to elucidate specific, high-level, human-interpretable information. Combined features are fed into a fully connected neural network which is able to learn the relationship between the two feature types. In the George B. Moody PhysioNet Challenge 2022 test set, our team (”lubdub”) received a weighted accuracy score of 0.835 with a cost of 14905 in the clinical outcome task (ranked 31/39). For the murmur prediction task, our model received a weighted accuracy score of 0.525 and a cost of 15083 (ranked 33/40).

1. Introduction

Auscultation of heart sounds for murmurs is vital in identifying cardiovascular disorders in children. In regions with lack of infrastructure or access to cardiology specialists, a non-invasive assessment of heart sounds via phonocardiogram (PCG) can provide life-saving information for pediatric patients with congenital or acquired heart disease. The analyzed data, provided by the Moody Challenge 2022, consisted of PCG recordings from 1568 pediatric patients collected in northeastern Brazil [1–3].

Ensemble models that combine information across raw and curated feature types have found broad success, such as in the winners 2016 and 2020 George B. Moody PhysioNet Challenges [4, 5]. We hypothesize that the neural network will be able to learn to identify heart murmur patterns within the time-frequency domain paired with our physician-inspired, human-interpretable information.

2. Methods

2.1. Preprocessing

For each patient, there were at most four PCG recordings that were recorded at the four classic auscultation locations on the chest corresponding to the aortic, pulmonic, tricuspid, and mitral valves of the heart. The samples are downsampled from 4000 to 1000 Hz, as we found that 1000 hz was a dense enough sampling rate to preserve PCG morphology and frequency information, while not being too dense for machine learning models. We then followed up with standard signal preprocessing techniques with a 25Hz highpass to remove low frequency interference and 400Hz lowpass 2nd order Butterworth filter in order to remove high frequency noise [6].

We followed the procedure outlined by [7] to remove outlier spikes by first dividing the recording into 500ms windows, finding the maximum absolute amplitude (MAA) in each window, and identifying samples where at least one MAA was greater than three times the median value of the MAAs in the entire sample. In these identified samples, the max noise spike was found to be the max point in the outlier MAA. The beginning and end of the spike were found by identifying the last zero-crossing point before the max point and the first zero-crossing after the max point, respectively.

2.2. CNN on Spectrogram Features

These time-series are then transformed into mel-spectrogram and concatenated to be used as inputs into a

 Page 1 ISSN: 2325-887X DOI: 10.22489/CinC.2022.065

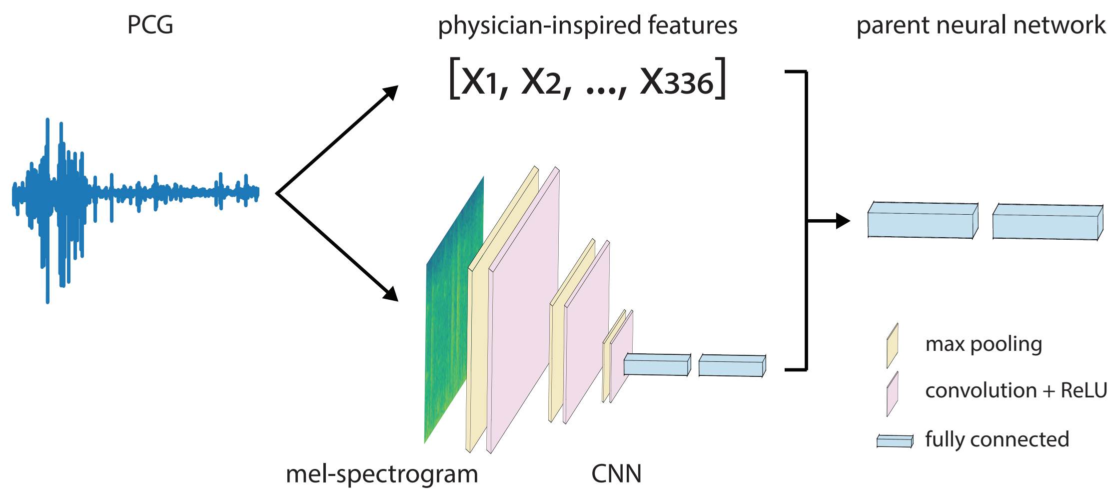

Figure 1. Network architecture. The PCG recording is used to extract both physician-inspired features and a mel-spectrogram. The spectrogram is input to a CNN containing three convolutional layers and two fully connected layers. Both the output of the CNN and the physician-inspired feature vector are concatenated and input into a fully connected neural network with a final softmax layer to predict the label of the recording.

CNN used for embedding the time-series. Specifically, this block is composed of three sequential blocks of a convolution layer, leaky relu, max-pooling, and dropout. The intuition of this is that unlike Recurrent Neural Network-based architectures, such as LSTMs, which seek to model information from each time-point individually in a memory, the convolutional nature allows us to capture and model morphological information more readily without any potential vanishing gradient issues or information lost across time [8]. Following this, we take the CNN-embed time-series and synthesize information from long-term dependencies with two fully-connected layers to model any potential long-term temporal dependencies. Batch normalization is applied after each block of layers to make training faster and more stable.

2.3. Physician-inspired Features

Recordings were segmented into key components of the cardiac cycle (S1, systolic interval, S2, and diastolic intervals) using a pre-trained logistic regression, hidden semi-Markov model with a modified Viterbi decoding algorithm [6]. [6] proposes a segmentation method that is robust to in-band noise, stemming from speech interference, motion artifacts, physiological sounds, etc., without the use of a separate reference signal, such as ECG, as the ground truth. A nonergodic hidden semi-markov model was used in order to incorporate prior information about how state transitions are allowed to occur and expected duration of each state, which each state represents one of the four heart sounds. A logistic regression issue was used to model emission probabilities, allowing for a greater discrimination between being in a given state.

The first and second heart sounds (S1 and S2) heard through a stethoscope are commonly used by physicians to identify and classify pathological murmurs before further medical intervention [9]. In the cardiac cycle, the first heart sound corresponds to the closing of the mitral and tricuspid valves while the second heart sound corresponds to the closing of the aortic and pulmonic valves. The heart actively pumps blood into the circulation during the time interval between S1 and S2 (systole) and is filled again by returning blood in the interval between S2 and S1 (diastole). Increased blood flow, either forwards or backwards, or a defective valve often results in murmurs [10]. Murmurs are traditionally classified based on their amplitude and timing in relation to S1 and S2. Despite the rising popularity of point-of-care ultrasound in identifying cardiac defects, the stethoscope remains the fastest, cheapest, and most universally accessible tool in recognizing cardiac pathology [11]. Automatic interpretation of auscultation results is especially vital in resource-poor health systems that may lack more advanced technology or the trained ear of a cardiologist.

The mean, median, and standard deviation of these cardiac cycle intervals as well as between these intervals were then calculated as summary statistics. Separately, we used discrete Fourier transform to calculate frequency domain features for each segmented window. For 10 Hz frequency bands between 30 and 800, the spectra of S1, systole, S2, and diastole were calculated [12]. Because murmurs are often due to an underlying structural issue persistent within the cardiac cycle and will continuously affect a given heart-

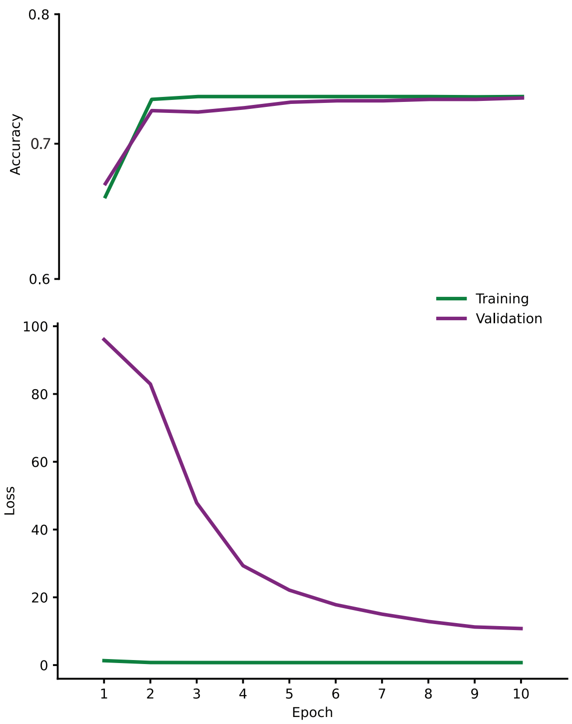

Figure 2. 6-fold cross-validation results of our method on the George B. Moody PhysioNet 2022 Challenge dataset. Within 10 epochs, accuracy converged to an unweighted training accuracy of 73.8% and validation accuracy of 73.6%.

sound region, summary statistics were used here to consolidate information about each of the intervals. [13]

These crafted features based on the cardiac cycle were concatenated with the spectrogram learned features and fed into a fully connected network with a softmax layer with cross-entropy to predict the label of the recording for both the murmur detection challenge and the clinical outcome prediction challenge. The full network architecture can be found in Figure 1. The murmur detection challenge included three classes (”Present”, ”Absent”, and ”Unknown”) while the clinical outcome challenge included two (”Abnormal” and ”Normal”).

3. Results

3.1. Cross-validation

Here we report the 6-fold cross-validation performance of our method on the George B. Moody PhysioNet 2022 Challenge dataset (Fig. 2). Both loss and accuracy converged within 10 epochs, resulting in unweighted training accuracy of 73.8% and validation accuracy of 73.6%. Although accuracy plateaued after a few epochs, validation loss continued to decrease towards the end of training.

	Training	Validation	Test	Ranking
Weighted Accuracy	0.501	0.467	0.525	33/40
Cost	14851	13836	15083	–

Table 1. Murmur detection task’s weighted accuracy score (official Challenge score) and cost scores for our final selected entry (team lubdub), including the ranking of our team on the hidden test set.

	Training	Validation	Test	Ranking
Weighted Accuracy	0.825	0.756	0.835	–
Cost	14584	13836	14905	31/39

Table 2. Clinical outcome task’s weight accuracy score and cost score (official Challenge score) for our final selected entry (team lubdub), including the ranking of our team on the hidden test set.

3.2. Challenge score

In the evaluation of the George B. Moody PhysioNet 2022 Challenge, our team’s (”lubdub”) model achieved a weighted accuracy of 0.525 and a cost of 15083 on the murmur detection task. Then on the outcome prediction task, we achieved an accuracy of 0.835 and a cost of 14905. Please see results in Tables 1 and 2. Our model was developed with Tensorflow-CPU version 2.4.1, and with our entry’s submission, our training took 7 hours 46 minutes with the model evaluation being completed in 4 hour and 12 minutes.

4. Discussion

In conclusion, our model demonstrates a proof-of-concept end-to-end architecture that flexibly combines features developed from physician knowledge with medical historical basis with the raw spectrogram representation of an input heart murmur audio time-series. Additionally, we provide the first public Python implementation of the widely-used PCG cardiac cycle segmentation algorithm [6] in our Github repository, which can be utilized by researchers in future works.

To further improve our algorithm, future research could optimize the hyper-parameters used in the model, which have not been carefully tuned. Extensive feature analysis could reveal the salience of each feature in classifying murmurs and highlight informative dimensions to assist diagnosis. Most importantly, we hope our work will facilitate and inspires the development of interpretable heart murmur classification algorithms that are able to utilize physician information within the model design and implementation.

References

[1] Reyna MA, Kiarashi Y, Elola A, Oliveira J, Renna F, Gu A, Perez-Alday EA, Sadr N, Sharma A, Mattos S, et al. Heart murmur detection from phonocardiogram recordings: The george b. moody physionet challenge 2022. medRxiv 2022; n.p.

[2] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng C, Stanley HE. Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals. Circulation June 2000;101(23):215–220.

[3] Oliveira J, Renna F, Costa PD, Nogueira M, Oliveira C, Ferreira C, Jorge A, Mattos S, Hatem T, Tavares T, Elola A, Rad AB, Sameni R, Clifford GD, Coimbra MT. The circor digiscope dataset: From murmur detection to murmur classification. IEEE Journal of Biomedical and Health Informatics 2022;26(6):2524–2535.

[4] Potes C, Parvaneh S, Rahman A, Conroy B. Ensemble of feature-based and deep learning-based classifiers for detection of abnormal heart sounds. In 2016 Computing in Cardiology Conference (CinC). 2016; 621–624.

[5] Alkhodari M, Hadjileontiadis LJ, Khandoker AH. Identification of cardiac arrhythmias from 12-lead ecg using beatwise analysis and a combination of cnn and lstm. In 2020 Computing in Cardiology. 2020; 1–4.

[6] Springer DB, Tarassenko L, Clifford GD. Logistic regression-hsmm-based heart sound segmentation. IEEE Trans Biomed Eng Apr. 2016;63(4):822–832.

[7] Schmidt SE, Holst-Hansen C, Graff C, Toft E, Struijk JJ. Segmentation of heart sound recordings by a duration-dependent hidden markov model. Physiol Meas Apr. 2010; 31(4):513–29.

[8] Yin W, Kann K, Yu M, Schütze H. Comparative study of cnn and rnn for natural language processing. arXiv preprint arXiv170201923 2017;.

[9] Pelech AN. The cardiac murmur: when to refer? Pediatric Clinics of North America 1998;45(1):107–122.

[10] Dornbush S, Turnquest AE. Physiology, heart sounds 2019; n.p.

[11] Thompson WR. In defence of auscultation: a glorious future? Heart Asia 2017;9(1):44–47.

[12] Tang H, Dai Z, Jiang Y, Li T, Liu C. Pcg classification using multidomain features and svm classifier. BioMed Research International 2018;2018(Article ID 4205027).

[13] Thomas SL, Heaton J, Makaryus AN. Physiology, cardiovascular murmurs. In StatPearls [Internet]. StatPearls Publishing, 2021; n.p.

Address for correspondence:

Julia Ding
100 Woodruff Circle
Atlanta, GA 30322 USA
julia.ding@emory.edu