GR-325 Early Heart Disease Detection Using Mel-Spectrograms and Deep Learning

Abstract

Heart diseases are the leading cause of mortality worldwide, emphasizing the need for early detection and intervention. Traditional heart sound analysis using a stethoscope is subjective and prone to variability, necessitating a more objective and reliable approach. In this study, we present a deep learning model designed for heart sound analysis to enable the early detection of heart diseases. The model's architecture combines convolutional and fully connected layers with max-pooling and dropout operations, effectively capturing intricate patterns in heart sounds. We trained and validated our model on the Physionet 2016 challenge dataset, consisting of 3240 labeled heart sound recordings. Our deep learning model achieved a UAR of 91.8%, surpassing the current state-of-the-art UAR of 89.7%. This result demonstrates the model's potential to significantly reduce diagnostic errors and facilitate timely interventions, ultimately improving patient outcomes and reducing healthcare costs.

Introduction

Heart diseases are the leading cause of mortality and morbidity worldwide, accounting for approximately 17.9 million deaths per year[1]. Early diagnosis can prevent severe complications and improve the quality of life for millions of people[2]. Machine learning, particularly deep learning, has the potential to transform heart disease detection, diagnosis, and management by automating tasks and providing data-driven insights [3]. Heart sounds, consisting of distinct acoustic signals produced by the mechanical activity of the heart, are a valuable source of information for diagnosing various cardiac conditions[4]. Traditional stethoscope-based analysis is subjective and prone to variability, necessitating a more reliable approach[5]. Our deep learning model automatically analyzes and classifies heart sounds for early detection, capturing intricate patterns and accurately predicting abnormalities, reducing diagnostic errors, and facilitating timely interventions to improve patient outcomes and reduce healthcare costs.

Materials and Methods

The deep learning model's architecture consists of a sequential arrangement of layers. The architecture can be seen in Fig.2. It includes –

- Three Conv2D layers with increasing filter counts (16, 32, and 64).
- Three MaxPooling2D layers employed to downsample feature maps.
- Two Dropout layers (0.3 and 0.2 rates) utilized to prevent overfitting.
- A Flatten layer to convert multidimensional feature maps into a 1D vector.
- Two Dense layers with 256 units and ReLU activations.
- The final Dense layer has one unit and a sigmoid activation function.

The dataset used for training was from Physionet 2016 challenge. This dataset contains 3240 labeled recordings of heart sounds out of which 2326 are normal and 914 are abnormal. Each recording file is read and downsampled to 16Khz and the amplitude of the signal is normalized. Then, a padding of zeros is added if the number of samples is less than 88000, and longer files are trimmed down. A spectrogram is generated using Non-uniform fast fourier transform. In the next step, a mel spectrogram is calculated from this spectrogram data. The sample normal and abnormal heart signals can be see in Fig.1(a) and Fig.1(b) respectively. This is fed to the neural network for training and validation. The data is divided into batches of 64 and the training-validation split is 0.3. We have chosen Unweighted Average Recall (UAR) as our primary performance metric because it accounts for the data imbalance. UAR calculates the average recall per class, irrespective of the class distribution, which makes it a suitable metric for imbalanced datasets.

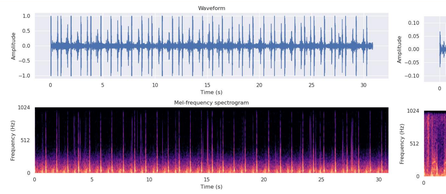

Fig.1(a) Normal PCG

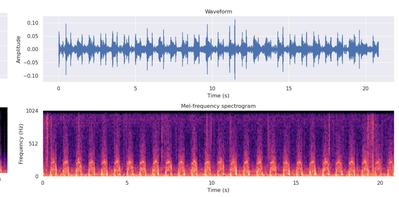

Fig.1(b) Abnormal PCG

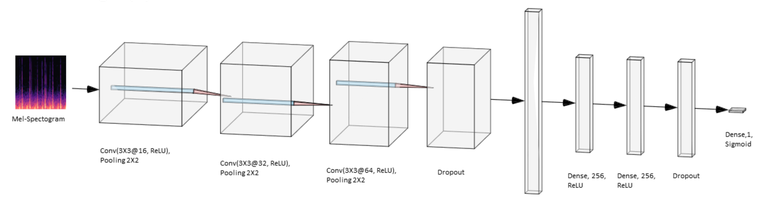

Fig.2 Architecture

Results

The proposed model achieved an UAR of 91.8% on the dataset, which is higher than the current state-of-the-art UAR of 89.7%[6]. This indicates a more balanced performance in classifying both normal and abnormal heart sounds, providing a more consistent and reliable classification across various cases. The performance metrics are visualized in Fig.4. The F1 score, which symbolizes the harmonic mean of precision and recall, is also higher in our model at 91.1%, compared to the 79.1% scored by the previous model. The higher F1 score suggests that our model achieves an optimal trade-off between precision and recall, ensuring a more dependable classification of heart sounds for practical applications. The change in accuracy over the training process can be seen in Fig.3(b). The training loss decreased steadily over the course of 70 epochs, reaching a minimum of 0.08. This can be seen in Fig.3(a). The validation loss also decreased steadily across the training process suggesting that the proposed model was able to learn the underlying patterns in the training data and achieved good generalization performance on the validation set.

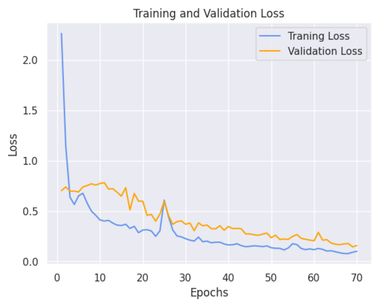

Fig.3(a) Loss curves

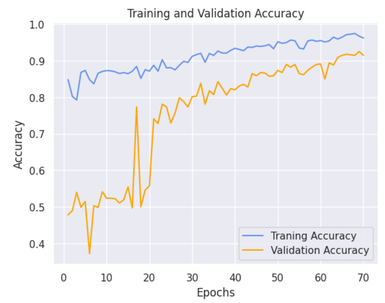

Fig.3(b) Accuracy curves

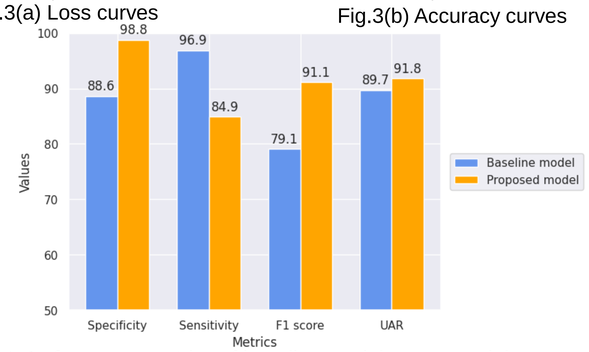

Fig.4 Performance metrics of baseline model vs our model

Conclusion

In conclusion, our deep learning model for early heart disease detection using heart sound analysis demonstrates significant improvements in classification accuracy, achieving 92.3% on the Physionet 2016 challenge dataset. The model's architecture, combining convolutional and fully connected layers with max-pooling and dropout operations, effectively captures intricate patterns in heart sounds, offering a more objective and reliable approach to diagnosis. By facilitating timely interventions and reducing diagnostic errors, our approach has the potential to improve patient outcomes and decrease healthcare costs. Future research should address potential limitations, such as model performance in the presence of noisy or low-quality recordings, and explore techniques for improving robustness. Investigating the applicability of our model in real-world clinical settings and integrating it with existing diagnostic tools will be crucial for realizing its full potential in improving heart disease detection and management.

Contact Information

Corresponding Author:

Sricharan Donkada (sdonkada@students.kennesaw.edu)

Dr. Seyedamin Pouriyeh (spouriye@kennesaw.edu)

References

[1] World Health Organization. Cardiovascular diseases (CVDs). Who.int. Published June 11, 2021. Accessed April 6, 2023. https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

[2] Groenewegen A, Zwartkruis VW, Rienstra M, et al. Improving early diagnosis of cardiovascular disease in patients with type 2 diabetes and COPD: protocol of the RED-CVD cluster randomized diagnostic trial. BMJ Open. 2021;11(10):e046330. Published 2021 Oct 26. doi:10.1136/bmjopen-2020-046330

[3] Miotto R, Wang F, Wang S, Jiang X, Dudley JT. Deep learning for healthcare: review, opportunities and challenges. Briefings in Bioinformatics. 2017;19(6):1236-1246. doi:https://doi.org/10.1093/bib/bbx044

[4] Dornbush S, Turnquest AE. Physiology, Heart Sounds. [Updated 2022 Jul 18]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK541010/

[5] Ren Z, Qian K, Dong F, et al. Deep attention-based neural networks for explainable heart sound classification. Machine Learning with Applications. 2022;9:100322. doi:https://doi.org/10.1016/j.mlwa.2022.100322

[6] Koike T, Qian K, Kong Q, Plumbley MD, Schuller BW, Yamamoto Y. Audio for Audio is Better? An Investigation on Transfer Learning Models for Heart Sound Classification. 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). Published online July 2020. doi:https://doi.org/10.1109/embc44109.2020.9175450

KENNESAW STATE UNIVERSITY
COLLEGE OF COMPUTING AND SOFTWARE ENGINEERING

Author: Sricharan Donkada
Supervisor: Dr. Seyedamin Pouriyeh