Diffusion-Based Conditional Phonocardiogram Generation: Multichannel Audio from a Single Channel Source

Introduction

The leading cause of death globally each year is cardiovascular disease [1], with prompt diagnosis being required to manage the disease. Auscultation, the most common non-invasive pre-screening method, lacks accuracy [2,3], emphasising a need for reliable alternatives. Multichannel computer-aided pre-screening has found greater success over a single channel [4]; however, limited multichannel phonocardiogram data makes it hard to train robust classifiers. Recent innovations in audio synthesis using diffusion models such as DiffWave [5] make it possible to generate multichannel audio from a single channel source.

Methods

Recordings were obtained from subjects using a multichannel device with six phonocardiogram sensors recording at different auscultation sites [6]. Ninety-six patients were diagnosed with coronary artery disease through angiography, and 58 were control patients. Using 232 minutes of audio taken from these patients, a 60%-20%-20% train-validation-test split was used.

Pre-processing consisted of a bandpass filter between 25-400Hz. Additional data augmentation, such as time stretching, rearranging of heart cycles and addition of filtered noise, was also applied to the signals during training to increase data distribution.

Based on the DiffWave architecture, as shown in Figure 1, a diffusion model was trained on each channel as the reference and each other channel as the conditioner to allow for all channels to be generated from one source signal.

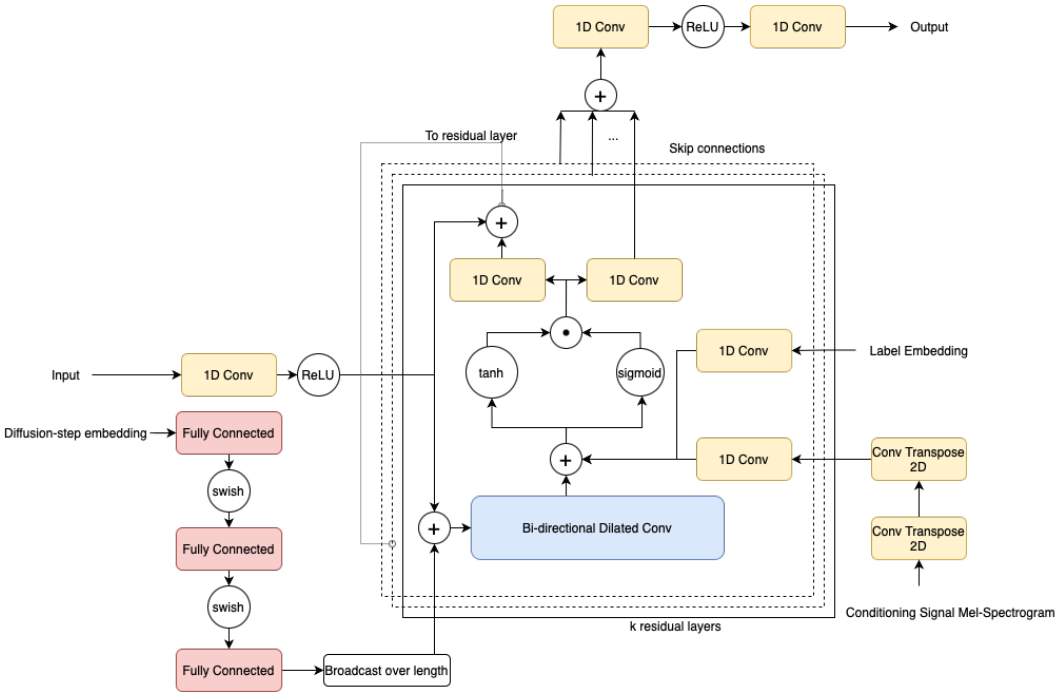

Figure 1. Diffusion Model Architecture.

Results

Figure 2 below compares the generated, reference and conditioning signals where the model tries to generate the reference signal from the conditioning.

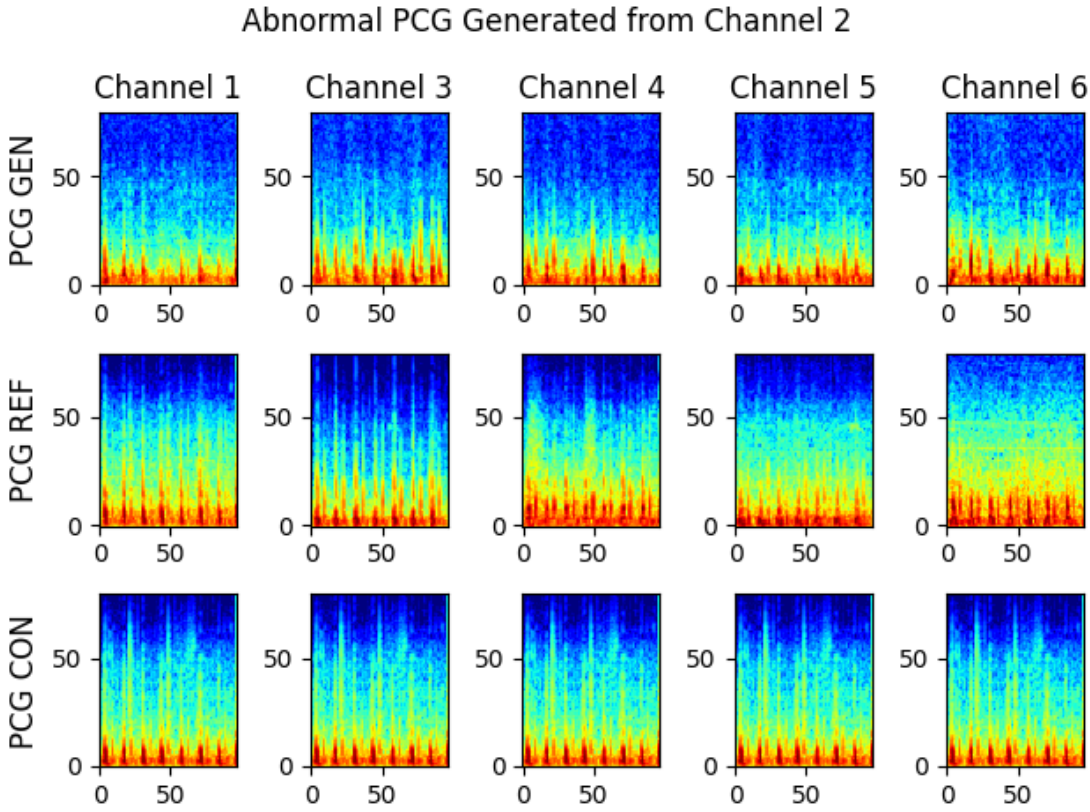

Figure 2. Comparison of generated (first row), reference (second row) and conditioning (third row) signals

Conclusion

The results show that more multichannel data can be created from existing single-channel datasets, the first work in this area, allowing easier training of deep-learning-based classification models.

References

1. WHO (2021) Cardiovascular Diseases (CVDs). Geneva, Switzerland: WHO.
2. Chizner M. A (2008) Cardiac Auscultation: Rediscovering the Lost Art. https://doi.org/10.1016/j.cpcardiol.2008.03.003
3. Feddock C. A (2007) The Lost Art of Clinical Skills. https://doi.org/10.1016/j.amjmed.2007.01.023
4. Samanta P, Pathak A, Mandana K, Saha G (2019) Classification of coronary artery disease and normal subjects using multi-channel phonocardiogram signal. https://doi.org/10.1016/j.bbe.2019.02.003

5. Kong Z, Ping W, Huang J, Zhao K, Catanzaro B (2021) DiffWave: A Versatile Diffusion Model for Audio Synthesis. https://doi.org/10.48550/arXiv.2009.09761
6. Rong Y, Fynn M, Nordholm S (2023) A pre-screening technique for coronary artery disease with multi-channel phonocardiography and electrocardiography. https://doi.org/10.1201/9781003346678