Utilization of Deep Learning and Expert Feature Classifier for Detection of Heart Murmurs

Petr Nejedly[1], Jan Pavlus[1], Radovan Smisek[1], Eniko Vargova[1], Zuzana Koscova[1], Ivo Viscor[1], Pavel Jurak[1], Filip Plesinger[1]

[1]Institute of Scientific Instruments of the Czech Academy of Sciences, Brno, Czech Republic

Abstract

This paper introduces our solution (ISIBrno-AIMT team) to the Physionet Challenge 2022. The main goal of the challenge was a classification of heart murmurs from phonocardiographic recordings into three mutually exclusive classes (i.e., present, unknown, and not present) and to determine whether the patient's overall status is Normal or Abnormal. We propose a deep learning method that classifies whether there is a heart murmur in the phonocardiographic recording and also provides heart sound segmentation. Furthermore, the expert feature classifier assesses whether the patient's status is normal or abnormal. Our approach achieved a hidden test challenge score of 0.755 in the murmur classification task and a score of 12313 in the patient outcome classification task. Our team was ranked as 9th and 12th out of 40 teams in the official ranking for murmur and outcome classification, respectively.

1. Introduction

Auscultation, i.e., analysis of hearable sounds produced in the chest, is a fast technique to screen the heart; more specifically, heart valves. When heart valves do not work correctly, abnormally circulating blood produce murmurs. Murmur severity is then considered by murmur occurrence and patient age; if needed, further screening is done by more expensive devices such as ECG or echocardiography. Although auscultation devices (phonendoscopes) are inexpensive in terms of medical devices, correct analysis outcome requires well-experienced personnel. Therefore, automated murmur analysis could cross this gap and help in scenarios with a limited budget for more complex methods, seasoned personnel, and population screenings

The Physionet Challenge 2022[1] is similar to Physionet Challenge 2016 [2] in that both challenges ask participants to process heart sound recordings. In 2016, most approaches mainly focused on feature-extracting methods that were classified with simple classifiers. For example, our team used a custom-made feature extraction tool and subsequent probability assessment [3, 4], ranked 7th in the following phase. The winner of the 2016 challenge was already using convolutional neural networks[5]. Since then, deep learning techniques with outstanding results have emerged in various engineering and scientific fields, e.g., image recognition and captioning, machine text translation, speech-to-text, and biological processing [6]. The deep-learning methods use end-to-end learning, which means manual feature engineering is omitted, and features are automatically extracted from data during the training process. The disadvantage of supervised deep-learning methods is that large-scale datasets with standard gold labels must be provided.

This paper presents our solution to Physionet Challenge 2022, which consists of two classifiers. The first is a deep learning based, and the second works with expert features. Each classifier solves one of the Physionet Challenge 2022 tasks. The first task aims to classify heart murmurs from phonocardiographic recordings into three mutually exclusive classes (i.e., present, unknown, and not present).

2. Methods

2.1. Murmur classification

The heart murmur detector is based on a deep-learning method that processes phonocardiogram data (fs = 4 kHz) converted into spectrograms by the Short Time Fourier Transform (STFT) with a window size of 256 and overlap of 128 samples. Each row of the spectrogram was normalized using a z-score. The neural network architecture consists of 5 convolutional layers (Figure 1), accompanied by batch normalization [7], ReLU activation, and dropout. The first convolutional layer has kernel size dependent on the spectrogram height and temporal dimension of 1. Subsequent convolutional layers use kernel 1xN, where N is temporal domain kernel size 3, 5, 7, and 9, respectively. Subsequently, the gated recurrent unit layer (GRU) is utilized to process the temporal domain of the data. The output from the GRU layer is split into two output heads. The

 ISSN: 2325-887X DOI: 10.22489/CinC.2022.041

first output head provides heart sound segmentation, and the second output head provides information on whether there is a murmur present in the recording or not. We observed that the GRU layer was improving accuracy for the segmentation task. However, there was not any effect on the classification task. The model was trained with an Adam optimizer [8] for 10 epochs, with a learning rate of 1e-4. For the training purpose, the model is trained using a combination of two loss functions, i.e., cross-entropy loss for murmur classification and second segmentation cross-entropy loss for heart sound segmentation. To provide robust classification, we have trained the model 15 times, each with a different subset of training data. Then we selected the top 5 models to form a classification ensemble (Figure 2).

2.2. Outcome classification

For the outcome classification task, we present an approach based on the extraction of expert features and subsequent classification using a shallow neural network with two fully connected layers (82 and 128 neurons). The features were extracted from each auscultation point, and the features from possible missing auscultation points were filled with zeros. The following two lists represent extracted features from the patient metadata and phonocardiographic recordings, respectively.

List of extracted features (patient metadata): Age, sex, pregnancy status, height, weight, number of auscultation points measured

List of extracted features (from each auscultation location) time-domain mean, std, skewness, kurtosis, length in seconds, power, power of bandpass envelopes in bands (15-90Hz, 55-150Hz, 100-250Hz, 200-450Hz, 400-800Hz), and correlation coefficients of former bandpass envelopes. The bandpass envelope frequency bands were taken from [3]. The band envelopes were calculated as the absolute value of the Hilbert transformed bandpass signal. The bandpass signals were computed with Butterworth 3rd order acausal filter.

The shallow neural network was trained using Adam (lr=1e-3, batch size=64) for 30 epochs, and validation scores (challenge outcome metric) were monitored after each epoch. The model with the lowest score (the best performing model) was saved for the subsequent inference phase. This approach was repeated five times to form an ensemble model.

3. Results

The performance of the models on the first task is assessed based on a weighted accuracy score [1]. The second task aims to classify whether the patient's overall status is considered normal or abnormal. The performance of the models on the second task is assessed according to the custom challenge metrics representing the cost of treatment. Thus the lower value indicates the better classifier. This metric is designed to strongly penalize the false negative detections, which forces the classifiers to be very sensitive. A detailed description of the scoring metrics is described in [1].

Methods proposed for this challenge evaluated on a hidden validation set achieved the murmur score of 0.709 while requiring two hours of model training. The feature extraction approach for outcome classification achieved an outcome score of 9657. Based on our experiments, we noticed that local cross-validation results for murmur score are approaching 0.80, and outcome score is approaching 8000. This indicates that the proposed methods are prone to overtraining.

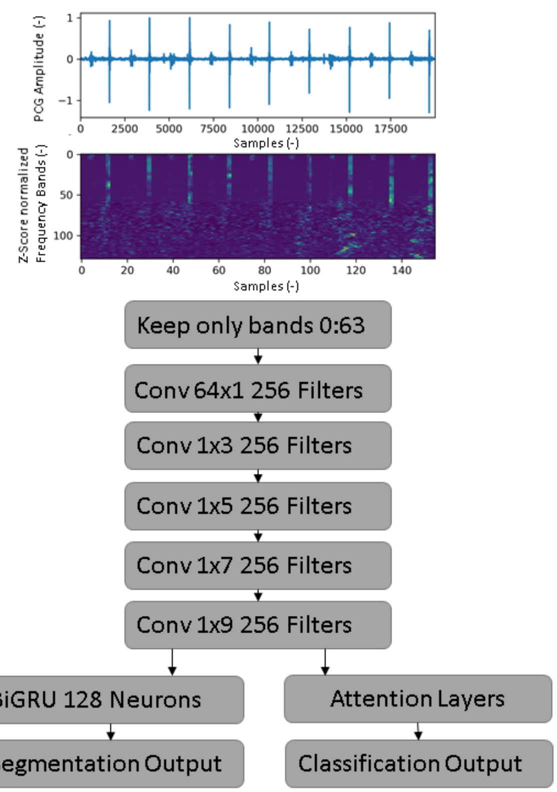

Figure 1. The proposed Neural Network architecture. The input phonocardiographic signal is transformed into a spectrogram by STFT and z-score normalized in each row. Subsequently, the STFT is processed by five convolutional layers that keep the same dimension as the input (paddings are not shown in the figure). The model has two output heads. One classifies each STFT timepoint into four classes of heart sounds, i.e., S1, Systole, S2, Diastole. The second output is used to classify whether the murmur is present, absent, or unknown.

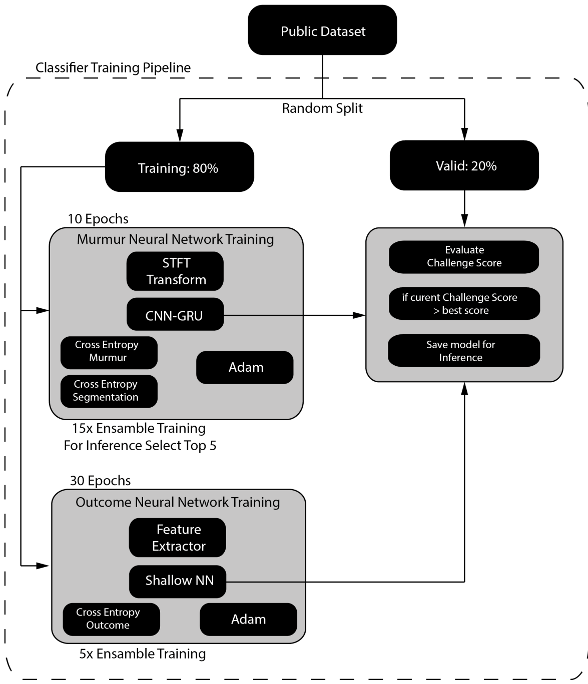

Figure 2. Training pipeline for murmur and outcome classifiers.

Training	Validation	Test	Ranking
0.827	0.709	0.755	9/40

Table 1. Weighted accuracy metric scores (official Challenge score) for our final selected entry (team ISIBrno-AIMT) for the murmur detection task, including the ranking of our team on the hidden test set. The results are obtained from official scores published by challenge organization team.

4. Discussion

During the official round of the challenge, we prepared several variants of our approach. Firstly, we tried to perform murmur classification and outcome classification using one model that takes the STFT data as an input and has three outputs, i.e., murmur, segmentation, and outcome. Nevertheless, this model was not able to train the outcome class. Considering that, we investigated several combinations of different loss functions and their weighting factors, but this approach was not successful.

Then we designed our outcome classifier based on the original challenge example code, which extracts expert features. We modified this approach to use a shallow neural network and added additional features that were successful for phonocardiogram classification during Physionet Challenge 2016. This approach achieved a better score than a purely deep-learning-based approach (outcome score 9657 vs. 13826 on validation set).

We also tried to train architecture for sound classifications to murmur tasks. This approach utilized only raw data in the temporal domain (without using STFT). Nevertheless, this approach was highly ineffective murmur validation score 0.543, with 10 hours of training.

Training	Validation	Test	Ranking
11089	9657	12313	12/40

Table 2. Cost metric scores (official Challenge score) for our final selected entry (ISIBrno-AIMT) for the clinical outcome identification task, including the ranking of our team on the hidden test set. The results are obtained from official scores published by challenge organization team.

5. Conclusion

In this paper, we presented our approach to the Physionet Challenge 2022. Our team was ranked 9th and 12th for murmur and outcome classification tasks, respectively. The murmur classification approach based on deep learning achieved the murmur score of 0.755 on the hidden test set. The outcome classification method based on expert feature extraction achieved the outcome score of 12313 on the hidden test set.

Acknowledgments

This work was supported by the Czech Technological Agency grant number FW01010305 and the CAS project RVO:68081731.

References

[1] Reyna MA, Kiarashi Y, Elola A, Oliveira J, Renna F, Gu A, et al. Heart murmur detection from phonocardiogram recordings: The George B. Moody PhysioNet Challenge 2022. medRxiv 2022;URL `https://doi.org/10.1101/2022.08.11.22278688`.

[2] Clifford G, Liu C, Springer D, Moody B, Li Q, Abad R, et al. Classification of normal/abnormal heart sound recordings: the PhysioNet/Computing in Cardiology Challenge 2016. 2016 Computing in Cardiology Conference CinC ;.

[3] Plesinger F, Viscor I, Halamek J, Jurco J, Jurak P. Heart sounds analysis using probability assessment. Physiological Measurement jul 31 2017;38(8):1685–1700.

[4] Plesinger F, Jurco J, Halamek J, Jurak P. Discrimination of normal and abnormal heart sounds using probability assessment. Computing in Cardiology Conference CinC 2016;.

[5] Potes C, Parvaneh S, Rahman A, Conroy B. Ensemble of Feature:based And Deep learning:based Classifiers for De-

tection of Abnormal Heart Sounds. 2016 Computing in Cardiology Conference CinC ;.

[6] Goodfellow I, Bengio Y, Courville A. Deep learning. MIT Press, nov 18 2016.

[7] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. https://arxiv.org/abs/1502.03167, feb 11 2015.

[8] Kingma DP, Ba J. Adam: A method for stochastic optimization. https://arxiv.org/abs/1412.6980, dec 22 2014.

Address for correspondence:

Petr Nejedly
nejedly@isibrno.cz
Institute of Scientific Instruments of the CAS, v. v. i.
Královopolská 147
612 64 Brno
Czech Republic