Caring is Sharing – Exploiting the Value in Data for Health and Innovation
M. Hägglund et al. (Eds.)

doi:10.3233/SHTI230198

Parralel Recurrent Convolutional Neural Network for Abnormal Heart Sound Classification

Arash GHAREHBAGHI[a,1], Elaheh PARTOVI[b] and Ankica BABIC[c,d]
[a] *School of Information Technology, Halmstad University, Halmstad, Sweden*
[b] *Department of Electrical Engineering, Amirkabir University, Tehran, Iran*
[c] *Department of Biomedical Engineering, Linköping University, Linköping, Sweden*
[d] *Department Information Science and Media Studies, University of Bergen, Bergen, Norway*

Abstract. This paper presents the results of a study performed on Parallel Convolutional Neural Network (PCNN) toward detecting heart abnormalities from the heart sound signals. The PCNN preserves dynamic contents of the signal in a parallel combination of the recurrent neural network and a Convolutional Neural Network (CNN). The performance of the PCNN is evaluated and compared to the one obtained from a Serial form of the Convolutional Neural Network (SCNN) as well as two other baseline studies: a Long- and Short-Term Memory (LSTM) neural network and a Conventional CNN (CCNN). We employed a well-known public dataset of heart sound signals: the Physionet heart sound. The accuracy of the PCNN, was estimated to be 87.2% which outperforms the rest of the three methods: the SCNN, the LSTM, and the CCNN by 12%, 7%, and 0.5%, respectively. The resulting method can be easily implemented in an Internet of Things platform to be employed as a decision support system for the screening heart abnormalities.

Keywords. Heart sound, deep learning, parallel convolutional neural network, convolutional neural networks, intelligent phonocardiography.

1. Introduction

Recent progress in development of different deep learning methods, created a leap towards intelligent decision making in various domains of healthcare including medical informatics and biomedical engineering. Extraction of medical information from time series of physiological activities has been traditionally regarded as an essential domain of research, sometimes with vital importance. Heart sound signal analysis, so called intelligent phonocardiography [1–4], is one of the highlighted domains of medical informatics and biomedical engineering which has been increasingly receiving attentions from the researchers, particularly after development of the powerful deep learning methods such as Convolutional Neural Network (CNN) and Long- and Short Term Memory (LSTM) [5–6]. The number of the publications addressing this domain has been doubled during the last three years with respect to the preceding half decade, mainly due to the high potential of deep learning methods in learning subtle details of the time series.

[1] Corresponding Author: Arash Gharehbaghi, E-mail: arash.ghareh-baghi@hh.se.

Among the various deep learning methods, spanning from the ones sophisticated for heart sound analysis, i.e. Time Growing Neural Network [7–9], to the image-based learning methods, CNN is by far the most popular learning method seen in different publications on heart sound analysis [5–7]. Although serious criticisms addressed reliability of the presented methods in terms of the structural risk [3,7,10], the study performed based on cascading a CNN and a recurrent neural network, which we named SCNN [5], and is still considered as the state-of-the art of this topic. It is evident that a reliable machine learning method can be incorporated into a digital stethoscope to serve as an easy-to-use and inexpensive screening tool for detecting heart abnormalities, which is, yet, considered as a priority for primary healthcare centers.

This paper introduces a Parallel combination of CNN and LSTM for heart sound signal analysis (PCNN), and compares accuracy of the presented method to the ones obtained by Deng et al. [5] in discriminating between normal and abnormal heart sounds. Performance of two other baseline methods were explored for the comparison: a Conventional CNN (CCNN) and a LSTM. Accuracy of the 4 methods is statistically validated using the widely known public dataset of heart sound, the PhysioNet/Computing in Cardiology Challenge 2016 (https://physionet.org). Results of the study published by Deng et al. are still considered as the state-of-the-art in terms of the performance and the methodology. This was the main motivation for selecting the Computing in Cardiology Challenge as the baseline for comparison.

2. Materials and Methods

The abovementioned repository of heart sound recordings contains 6 folders of data, named training-a to training-f, altogether comprising 3240 signals, from which 665 signals correspond to abnormal hearts. The subjects can have multiple recordings; however, the recordings have been allocated in the folders in a mutually exclusive manner. The recordings are all anonymous and have different lengths from 5 to 20 seconds. More details are found in (https://physionet.org/content/challenge-2016/1.0.0/).

2.1. Classification Methods

The classification methods are trained using the datasets exist in the mentioned repository, in which the number of the normal signals are by far higher than the abnormal ones, yielding a heavy class imbalance in the training data. In the PCNN case, an input signal is firstly divided into nonoverlapping segments of 5 seconds. In order to overcome the class imbalance, heavily seen in the repository, an augmentation method of SMOTE was employed. Details of the augmentation method can be found in [5]. The signal contents are mapped to 2-dimensional representation using the mel-frequencies representation [5]. Contents of the mel-frequency are employed by a CNN and a LSTM independently and the ultimate classification is performed using the two sets of the outcomes. In another attempt, a cascaded connection of CNN and LSTM, named CCNN, was employed in which the LSTM performs the ultimate classification. The rest of the processing including the mel-frequency representation remain identical to the PCNN. Figure 1 demonstrates the block diagram of the two methods. In both of the cases, the CNN and the LSTM are independently trained and optimized using Adam optimizer. The set of the hyper parameters is identically selected for the PCNN and the CCNN as listed in Table 1. The baselines for comparison are composed of a CCNN and a LSTM with the identical inputs of mel-frequency contents together with the similar set of the hyperparameters as the ones selected for the PCNN, Figure 1.

Table 1. The hyperparameters

Parameter	Value
Kernel size and stride size of convolution layer	3×3 and 1
Number of the convolutional layers and kernel at each layer	3 and (16, 32, 64)
Number and size of the Max Pool	3 and ((2×2), (4×4), (2×2))
Number of LSTM layers and units	1 and 74
Number of neurons in FC and the dropout rate	32 and 0.5
Activation function of the last layer	SoftMax
Activation function of the convolutional and FC layers	ReLu
Initial learning rate and exponential decay rates	0.01 and (0.9, 0.999)
Batch size and number of the epoch	512 and 50

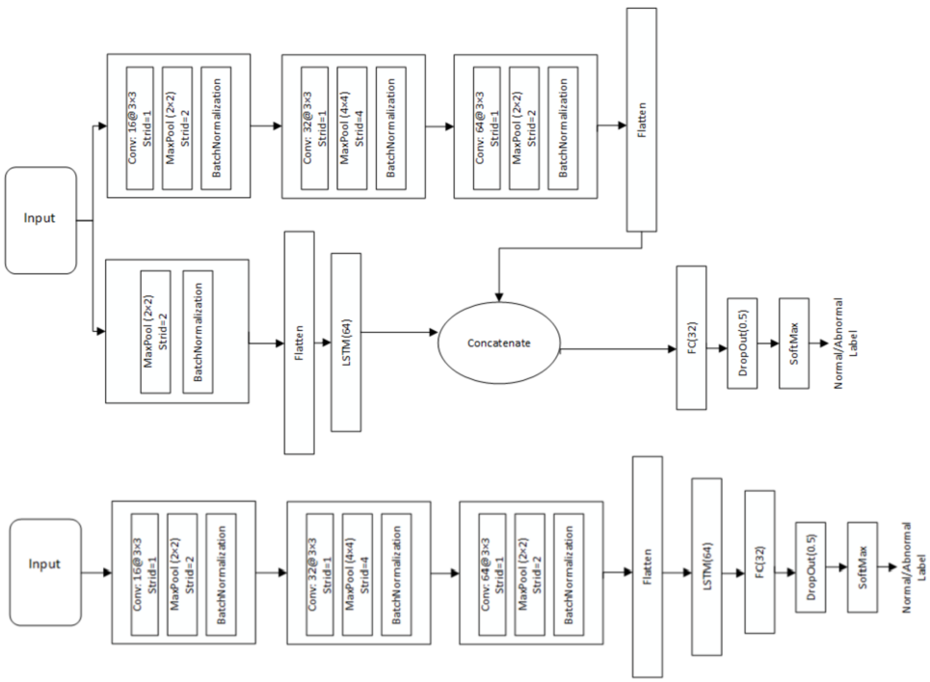

Figure 1. Block diagram of the PCNN (top) and CCNN (bottom)

3. Results

The recordings of each folder are employed for evaluation of the 4 methods, the PCNN, the SCNN, the CCNN, and the LSTM, with the training/validation/test of 75%/15%/10% selected from the Physionet datasets. The signals of each Physionet folder are firstly shuffled. Then, the training/validation/test recordings are selected according to the mentioned percentages. The three methods are independently evaluated 5 times and the descriptive statistics of two performance measures, the accuracy and the sensitivity, are calculated. Table 2 lists the estimated values for the average and standard deviation

(STD) of each performance measure, independently. As can be seen from the table, the PCNN offers the best performance in terms of the accuracy. The CCNN and LSTM both exhibit better accuracies comparing to the SCNN, however, this is not true with the sensitivities which are substantially lower with high standard deviations. This makes the CCNN and LSTM inappropriate to be employed as a screening tool. On the other hand, accuracy of the SCNN shows a high standard deviation against the evaluation data (the training/validation/test data), implying on its high structural risk. Nevertheless, for a certain selection of the evaluation data, an outperformance of SCNN might be observed, which could not be assumed objectively conclusive for the general population.

Table 2. The descriptive statistics of the performance measures for the three methods, PCNN, CCNN, and CNN for 5 runs of the algorithms with the train\validation\test of 75%\10%\15%

Method	Accuracy		Sensitivity	
	Average (%)	STD (%)	Average (%)	STD (%)
PCNN	87.2	2.7	76.5	14.9
SCNN	75.2	25.1	76.6	16.0
CCNN	86.7	2.6	64.5	21.6
LSTM	80.3	2.9	32.3	17.3

4. Discussion

The paper suggested a parallel combination of LSTM and CNN for discriminating between normal and abnormal heart sounds. This combination not only improves performance of the conventional CNN and LSTM, but also provides a better accuracy comparing to the cascaded combination which is considered as the state-of-the-art. The sensitivity is not improved by the parallel combination, though. The cascaded combination of CNN and LSTM was previously introduced by Deng et al. [5], with a very high accuracy. However, this is limited to a specific selection of the training/validation/test datasets, and thus very data-sensitive leading to a conclusion that cannot not be generalized as reflected by the high standard deviation of the performance measures. An important aspect of this study is the richness of dataset for training and testing of the methods. This richness is somewhat impaired by the class imbalance for the normal and abnormal heart sounds. In this study, an augmentation method, named SMOT was employed to overcome the class imbalance. Nonetheless, a large dataset with sufficient samples of the normal and abnormal samples, covering various pathological conditions, is crucial to arrive at an optimal selection for training the hyperparameters. The resulting method has the practical potential to improve the screening accuracy of cardiac auscultation at the primary healthcare centers. Studies showed that this screening accuracy is, yet, insufficient [11,12]. Integration of such artificial intelligence-based methods assigns a high level of sophistication to the electronic stethoscopes towards associating intelligence to the stethoscopes for various diagnostic objectives [13].

5. Conclusions

The parallel combination of LSTM and CNN improves performance of a cascaded LSTM and CNN for discrimination between normal and abnormal heart conditions using the heart sounds. This parallel combination offers significant enhancement of performance in terms of accuracy as compared to the conventional cases of CNN and LSTM, showing capability of such a combination of the deep learning methods as employed in this case study. However, the gain in accuracy could not guarantee an increase of the sensitivity or its reproducibility. The proposed machine learning method with high and stable accuracy can be integrated with an electronic stethoscope to be employed as a decision support system at primary healthcare centers.

References

[1] Gharehbaghi A, Sepehri AA, Lindén M, Babic A. Intelligent Phonocardiography for Screening Ventricular Septal Defect Using Time Growing Neural Network. Informatics Empowers Healthcare Transformation: IOS Press; 2017. 238:108-11.

[2] Gharehbaghi A, Lindén M. An Internet-Based Tool for Pediatric Cardiac Disease Diagnosis Using Intelligent Phonocardiography. In: Mandler B, Marquez-Barja J, Mitre Campista ME, Cagáňová D, Chaouchi H, Zeadally S, et al., editors. Internet of Things IoT Infrastructures: Second International Summit, IoT 360° 2015, Rome, Italy, October 27-29, 2015 Revised Selected Papers, Part I. Cham: Springer International Publishing; 2016. p. 443-7.

[3] Gharehbaghi A, Babic A. Structural Risk Evaluation of a Deep Neural Network and a Markov Model in Extracting Medical Information from Phonocardiography. Studies in health technology and informatics, 2018, 251:157-160.

[4] Gharehbaghi A, Sepehri AA, Lindén M, Babic A, A Hybrid Machine Learning Method for Detecting Cardiac Ejection Murmurs. EMBEC & NBC 2017; 787-790.

[5] Deng M, Meng T, Cao J, Wang S, Zhang J, Fan H. Heart sound classification based on improved MFCC features and convolutional recurrent neural networks. Neural Netw. 2020 Oct;130:22-32.

[6] Singh SA, Meitei TG, Majumder S. Short PCG classification based on deep learning. InDeep Learning Techniques for Biomedical and Health Informatics 2020 Jan 1 (pp. 141-164). Academic Press.

[7] Gharehbaghi A, Babic A. Deep Time Growing Neural Network vs Convolutional Neural Network for Intelligent Phonocardiography. Stud Health Technol Inform. 2022 Jun 29;295:491-494. doi: 10.3233/SHTI220772. PMID: 35773918.

[8] Gharehbaghi A, Lindén M, Babic A. An Artificial Intelligent-Based Model for Detecting Systolic Pathological Patterns of Phonocardiogram based on Time-Growing Neural Network. Applied Soft Computing, 2019. 83:105615.

[9] Gharehbaghi A, Lindén M. A Deep Machine Learning Method for Classifying Cyclic Time Series of Biological Signals Using Time-Growing Neural Network. IEEE Transactions on Neural Networks and Learning Systems. 2018; 29(9) 4102-4115.

[10] Gharehbaghi A. and Babic A., A-Test for Quantifying Structural Risk and Learning Capacity of Supervised Machine Learning Methods, IOS Press, 2022.

[11] Watrous RL, Thompson WR, Ackerman SJ. The impact of computer‐assisted auscultation on physician referrals of asymptomatic patients with heart murmurs. Clinical Cardiology: An International Indexed and Peer‐Reviewed Journal for Advances in the Treatment of Cardiovascular Disease. 2008 Feb;31(2):79-83. https://doi.org/10.1002/clc.20185

[12] DeGroff CG, Bhatikar S, Hertzberg J, Shandas R, Valdes-Cruz L, Mahajan RL. Artificial neural network–based method of screening heart murmurs in children. Circulation. 2001 Jun 5;103(22):2711-6..

[13] Gharehbaghi A, Ask P, Nylander E, Janerot-Sjoberg B, Ekman I, Lindén M, Babic A. A hybrid model for diagnosing sever aortic stenosis in asymptomatic patients using phonocardiogram. InWorld Congress on Medical Physics and Biomedical Engineering, June 7-12, 2015, Toronto, Canada 2015 (pp. 1006-1009). Springer International Publishing.