Research on Intelligent Analysis Model of Heart Sound based on Deep Learning

Hui Yu[1,2,3][a], Jing Zhao[1][b], Jinglai Sun[2,3][c] and Zhaoyu Qiu[2][d]
[1]*Academy of Medical Engineering and Translation Medicine, Tianjin University, Tianjin 300072, China*
[2]*School of Precision Instrument and Opto-Electronics Engineering, Tianjin University, Tianjin 300072, China*
[3]*Department of Biomedical Engineering, Tianjin Key Laboratory of Biomedical Detecting Techniques and Instruments, Tianjin University, Tianjin 300072, China*

Keywords: Heart Sound, Bispectrum Analysis, Convolutional Neural Networks.

Abstract: Heart sound auscultation is one of the most basic cardiac diagnosis techniques, but the traditional artificial auscultation method requires experienced clinicians and is limited by environmental factors. In this study, an intelligent analysis model of heart sound based on deep learning was designed to meet the daily public screening. Firstly, two public data sets and clinical self-collected data sets were fused, and pretreatments were carried out， such as normalization, denoising, overlapping cutting and subsampling. Then, the extraction and quantitative analysis of heart sound features were completed using bispectrum analysis technology. Finally, the features were input into the constructed improved convolutional neural network for classification. The results show that the accuracy, sensitivity, specificity and F1 score of normal and abnormal heart sounds were 85.5%, 85.7%, 85.3% and 85.9%, respectively, and the performance of pathological heart sounds classification was over 90%, reaching the highest level of this kind of research at present. This model provides a standardized evaluation with high classification performance and can quickly complete the intelligent analysis of heart sounds, which has important clinical significance.

1 INTRODUCTION

Heart sound is a mechanical wave phenomenon caused by the movement of the heart. The digital signals collected by sensors are called phonocardiogram (PCG). Heart sounds are clinically associated with many heart pathologies, common among which are aortic stenosis (AS), mitral stenosis (MS), mitral regurgitation (MR), and mitral valve prolapse (MVP) (Reed 2004).

Heart sound auscultation is of great significance in the diagnosis of cardiovascular diseases. It is one of the most commonly used cardiac diagnostic techniques because of its characteristics of non-invasive, fast and low cost. However, the traditional manual auscultation method requires experienced clinicians and is limited by environmental factors, making it highly subjective and easy to make mistakes. According to statistics, cardiologists' auscultation accuracy is about 80%, while that of primary care physicians is only in the range of 20% to 40%(Ma 2020). With the increasing demand for heart sound auscultation, clinical patients are eager to develop an accurate and rapid intelligent heart sound detection algorithm suitable for public screening.

At present, the automatic diagnosis of heart sound signal mainly has the following methods: Extracting features efficiently and using traditional pattern recognition methods such as support vector machine (SVM), empirical parameters and k-nearest neighbor (K-NN), the diagnosis is made, or the ability of the neural network itself to extract features and classify, such as convolutional neural network (CNN), deep neural network (DNN), recurrent neural network (RNN), deep confidence network (DBN) (Dominguez-Morales 2018, Abduh 2019, Chen 2018, Wu 2019).

Most of the current studies are based on foreign open data sets, and it remains to be seen whether the research results apply to Chinese clinical practice.

[a] https://orcid.org/0000-0002-8511-7296
[b] https://orcid.org/0000-0003-3231-5544
[c] https://orcid.org/0000-0003-3683-1968
[d] https://orcid.org/0000-0002-7728-7367

Yu, H., Zhao, J., Sun, J. and Qiu, Z.
Research on Intelligent Analysis Model of Heart Sound based on Deep Learning.
DOI: 10.5220/0011291400003444
In *Proceedings of the 2nd Conference on Artificial Intelligence and Healthcare (CAIH 2021)*, pages 244-248
ISBN: 978-989-758-594-4

Therefore, an intelligent heart sound analysis model based on deep learning was proposed to be suitable for clinical practice and higher classification accuracy. Firstly, two available data sets and clinical self-collected data sets were fused, and pretreatments such as normalization, denoising, overlapping cutting and sub-sampling were performed. Then, we used the bispectrum analysis technology to complete the heart sound feature extraction and quantitative analysis. Finally, the features were input into the constructed improved convolutional neural network for the dichotomies of healthy and abnormal heart sounds and four classifications of pathological heart sounds, and the performance of the model was evaluated by accuracy, sensitivity, specificity, accuracy rate and F1 score.

2 MATERIALS AND METHODS

2.1 Dataset

The first dataset used in this paper is the 2016 PhysioNet/CinC Challenge (Liu 2016), which includes 3,126 heart sound in two classes of abnormal and normal. The second used dataset is derived from GitHub, which contains four pathological and normal heart sound types uploaded by researcher Yaseen at Sejong University (Yaseen 2018). The third available dataset, including 29 kinds of the typical heart sound, is collected by our laboratory and Tianjin Chest Hospital of China.

2.2 Bispectrum Analysis

The feature engineering in this paper is based on the bispectrum analysis. Bispectrum analysis is one of the most commonly used higher order spectrum analysis methods. It can sufficiently suppress the signal's phase relationship and detect and quantify the phase coupling of non-Gaussian signals (Alqudah 2020). Its formula is as follows:

$$S_2^x(\omega_1,\omega_2) = \sum_{\tau_1=-\infty}^{\infty}\sum_{\tau_2=-\infty}^{\infty} C_3^x(\tau_1,\tau_2)e^{(-j(\omega_1\tau_1+\omega_2\tau_2))} \quad (1)$$

2.3 Improved Convolutional Neural Networks

In this study, a 19-layer improved convolutional neural network was constructed. A submodule contains a convolutional layer, BN layer, ReLU layer and max pooling layer. After being processed by three identical sub-modules, the pooling layer in the fourth sub-module is changed into a full connection layer, and then a layer of softmax layer is used to complete logistic regression to get our classification results. Figure 1 shows the overall network structure, and specific parameters of the model are shown in Table 1.

Table 1: Model parameter selection.

#	Layer	Information	#
1	Input layer	Size	256*256
		Number of filters	32
2	Conv_1	Kernel size	3*3
		Activation	ReLU
3	Batch_Norm_1	Number of channels	32
4	ReLU_1	Activation	ReLU
5	Maxpol_1	Kernel size	2*2
		Stride	2*2
		Number of filters	16
6	Conv_2	Kernel size	3*3
		Activation	ReLU
7	Batch_Norm_2	Number of channels	16
8	ReLU_2	Activation	ReLU
9	Maxpol_2	Kernel size	2*2
		Stride	2*2
		Number of filters	8
10	Conv_3	Kernel size	3*3
		Activation	ReLU
11	Batch_Norm_3	Number of channels	8
12	ReLU_3	Activation	ReLU
13	Maxpol_3	Kernel size	2*2
		Stride	2*2
		Number of filters	16
14	Conv_4	Kernel size	3*3
		Activation	ReLU
15	Batch_Norm_4	Number of channels	16
16	ReLU_4	Activation	ReLU
17	Fully connected layer	Number of channels	16
18	Softmax	Activation	Softmax
19	Output layer	Size	2*1

3 EXPERIMENTAL RESULTS AND DISCUSSIONS

3.1 Data Preprocessing

PhysioNet, Yaseen, and clinical self-collected data sets were used in this study. First, the heart sound signal was normalized to facilitate the subsequent study of the heart sound signal, and then the soft threshold denoising algorithm of the wavelet was used to process the heart sound to eliminate the influence of environmental noise. The db6 wavelet

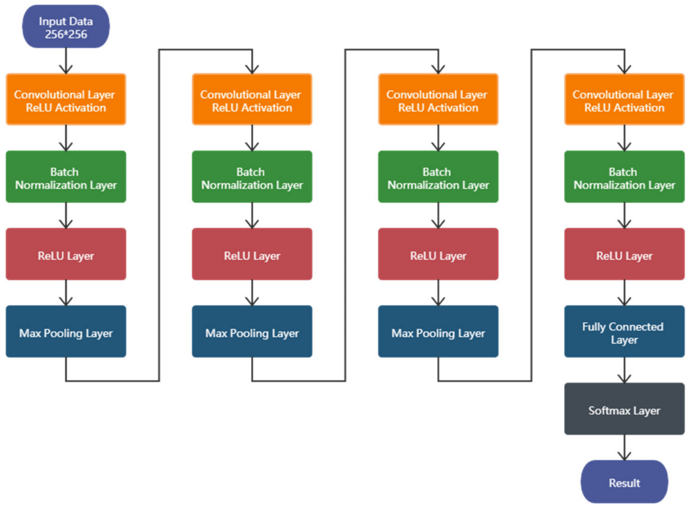
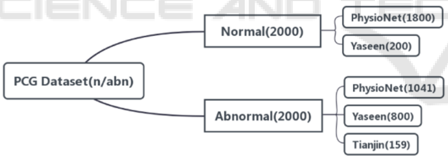
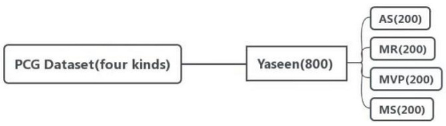


Figure 1: Diagram of the proposed network architecture.

was selected for four-layer wavelet decomposition, and the threshold function was 'sqtwolog'. Finally, the heart sound signal was sampled to 1000Hz, and the heart sound sample was cut by 2.5s, with 50% overlap each time, to realize samples were doubling while reducing the calculation pressure. The preprocessed heart sound samples were fused to build two fusion data sets, as shown in Figure 2, a normal/abnormal heart sound dataset was established, and an AS/MR/MVP/MS heart sound dataset was also established.



(a) Normal/Abnormal dataset

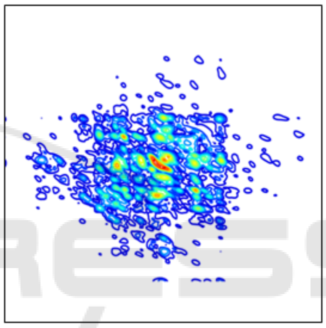


(b) AS/MR/MVP/MS dataset

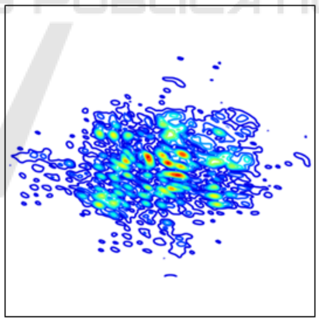
Figure 2: Construction of dataset.

3.2 Experimental Results

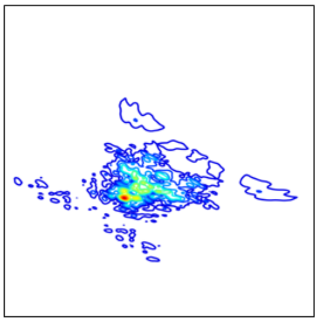
According to the bispectrum analysis technology, features were extracted from the heart sound signals in the data set. Figure 3 is the contour map of bispectrum plotted for two signals of aortic stenosis and one signal of mitral stenosis, and each image is the first quadrant image plotted. It can also be seen that the images of the two cases of aortic stenosis are similar but differ from the images of mitral stenosis.



(a) Aortic stenosis



(b) Aortic stenosis



(c) Mitral stenosis

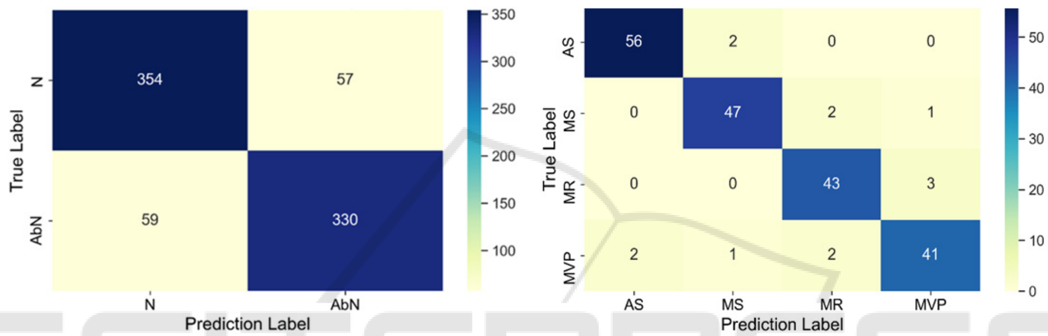
Figure 3: Contour images of bispectrum.

The extracted features were input into the 19-layer convolutional neural network previously established to distinguish between normal and abnormal heart sounds. The learning rate was set to 0.001, the loss function was set to the cross-entropy loss function, the optimizer was selected Adam, and the batch size was set to 16. A total of 4000 samples were divided into a training set and test set according to 4:1. The classification performance is shown in Table 2, and the confusion matrix is shown in Figure 4(a).

For the four categories of pathological heart sounds, the setting was the same as that for the two categories except batch-size 32. There were 800 samples for the four categories, and the number of the four diseases was the same, but the training set and test set were still allocated according to 4:1, and the model training converged around 300 rounds. The classification performance is shown in Table 3, and the confusion matrix is shown in Figure 4(b).

Table 2: Performance of the binary model.

Indicators	Score
Accuracy	0.980
Sensitivity	0.966
Specificity	0.986
F1 score	0.966



(a) Two categories (b) Four categories

Figure 4: Confusion Matrix

Table 3: Performance of the four classification models.

Category	Accuracy	Sensitivity	Specificity	F1 score
Aortic stenosis	0.980	0.966	0.986	0.966
Mitral stenosis	0.970	0.940	0.980	0.940
Mitral regurgitation	0.965	0.935	0.974	0.925
Mitral valve prolapse	0.955	0.891	0.974	0.901

3.3 Discussion

Although the model in this study has good classification performance in the four categories, there is still a particular gap compared with previous studies in the two categories. We suspect this is because the samples for the four categories come from the same database, while the samples for the two categories is a fusion of the three databases. Different databases have different collection methods and noise types, resulting in poor model performance in the fusion data sets. We will also increase the research on noise processing to narrow the differences between different databases in future research. In addition, sample collection will continue, especially the collection of pathological samples of the four categories, to expand the number of samples and improve the robustness and generalization ability of the model.

4 CONCLUSIONS

This article used three data sets: PhysioNet, Yaseen, and the clinical self-collected data sets. Firstly, normalization, wavelet denoising, subsampling and overlapping cutting were used to complete the data preprocessing, and then feature extraction was completed based on bispectrum. Finally, the features were put into the constructed neural network model of dichotomies and quadrotors for the corresponding classification, and the confusion matrix was calculated, and several evaluation parameters were obtained.

The results show that this model can provide a standardized evaluation with high classification performance and quickly complete the intelligent analysis of heart sounds, which has important clinical significance.

ACKNOWLEDGEMENTS

The authors express thanks to the doctors of Tianjin Chest Hospital of China for providing full help in the research process.

REFERENCES

Abduh, Z., Nehar, Y. E. A., Wahed, M. A., et al. (2019) Classification of heart sounds using Fractional Fourier Transform Based Mel-Frequency Spectral Coefficients and Stacked Autoencoder Deep Neural Network. J. Journal of Medical Imaging and Health Informatics, 9(1): 1-8.

Alqudah, A. M., Alquran, H. H., Abuqasmieh, I. (2020) Classification of heart sound short records using bispectrum analysis approach images and deep learning. J. Network Modeling Analysis in Health Informatics and Bioinformatics, 9(1):66.

Chen, L., Ren, J., Hao, Y., and Hu, X. (2018) The diagnosis for the extrasystole heart sound signals based on the deep learning. J. Journal of Medical Imaging and Health Informatics, 8(5):959-986.

Dominguez-Morales Juan, P., Jimenez-Fernandez Angel, F., Dominguez-Morales Manuel, J., Jimenez-Moreno Gabriel. (2018) Deep neural networks for the recognition and classification of heart murmurs using neuromorphic auditory sensors. J. IEEE transactions on biomedical circuits and systems, 12(1):1-11.

Liu, C., Springer, D., Li, Q., et al. (2016) An open access database for the evaluation of heart sound algorithms. J. Physiological Measurement, 37(12): 2181-2213.

Ma, L. Y., Chen, W. W., Gao, R. L., et al. (2020) China cardiovascular diseases report 2018: an updated summary. J. Journal of Geriatric Cardiology, 17(1): 1-8.

Reed, T. R., Reed, N. E., Fritzson, P. (2004) Heart sound analysis for symptom detection and computer-aided diagnosis. J. Simulation Modelling Practice and Theory, 12(2): 129-146.

Wu, M. T., Tsai, M. H., Huang, Y. Z., et al. (2019) Applying an ensemble convolutional neural network with Savitzky–Golay filter to construct a phonocardiogram prediction model. J. Applied Soft Computing Journal, 78(1):29-40.

Yaseen, Son, G. Y., Kwon, S. (2018) Classification of heart sound signal using multiple features. J. Applied Sciences, 8(12):2344.