

Design of Abnormal Heart Sound Recognition System Based on HSMM and Deep Neural Network

Hai Yin¹, Qiliang Ma², Junwei Zhuang¹, Wei Yu¹, Zhongyou Wang³

¹School of Biomedical Engineering and Medical Imaging, Xianning Medical College, Hubei University of Science and Technology, Xianning, 437100, People's Republic of China; ²School of Mathematics and Computer, Wuhan Textile University, Wuhan, 430200, People's Republic of China; ³School of Computer Science and Technology, Hubei University of Science and Technology, Xianning, 437100, People's Republic of China

Correspondence: Zhongyou Wang, Email wzy1963@163.com

Introduction: Heart sound signal is an important physiological signal of human body, and the identification and research of heart sound signal is of great significance.

Methods: For abnormal heart sound signal recognition, an abnormal heart sound recognition system, combining hidden semi-Markov models (HSMM) with deep neural networks, is proposed. Firstly, HSMM is used to build a heart sound segmentation model to accurately segment the heart sound signal, and then the segmented heart sound signal is subjected to feature extraction. Finally, the trained deep neural network model is used for recognition.

Results: Compared with other methods, this method has a relatively small amount of input feature data and high accuracy, fast recognition speed.

Discussion: HSMM combined with deep neural network is expected to be deployed on smart mobile devices for telemedicine detection.

Keywords: heart sound signal, recognition, hidden semi-Markov, neural network

Introduction

The heart sound signal is generated by the vibration of the whole heart structure caused by the opening and closing of the heart valves.¹ It is an important physiological signal in the human body and has an important role in the prevention and diagnosis of cardiovascular diseases.² At present, domestic and international hospitals still use manual auscultation in the diagnosis of heart sound, which is susceptible to both the interference of the external environment and personal experience. In addition, with the development of smart wear, a new requirement for remote real-time monitoring of heart sound signal has been put forward, and the traditional way of heart sound auscultation is unable to meet this requirement.

At present, the main research on heart sounds is extracting heart sound feature, and heart sound classification models mostly use convolutional neural network (CNN) algorithms. For example, Chaowen Tan et al^{3,4} proposed data acquisition of more than 1000 heart sound fragments in the time domain using Meier coefficients, combining the collected signal data into two-dimensional feature samples, and finally using convolutional neural networks to identify these heart sound signals; Noman et al⁵ combined Markov switching autoregressive (MSAR) model with switched linear dynamic system (SLDS) to build a heart sound segmentation model, which can accurately segment heart sounds and is convenient for subsequent heart sound recognition applications; Xingzhi Wang et al⁶ used stacked subband signal envelopes to construct time-frequency features and then used CNNs as classification models; Zeenat Tariq et al⁷ proposed a feature-based fusion network to transform audio data into image feature vectors to construct three optimal CNNfusions for heart sound classification; Meng Li-Nan et al⁸ used sliding window and Meier coefficients for feature extraction of heart sound signals to form a heart sound feature map, and then used CNNs as a classification model; Liu Wei-Wei et al⁹ made improvements on the classification model, but also used the combination between CNN and long short-term memory (LSTM) artificial neural network.

Although the heart tone feature extraction algorithms used in the above studies are different, they all convert the heart tone feature vectors into two-dimensional vectors and then put into the convolutional neural network. This is because CNN is originally used mainly for two-dimensional image recognition, thus, when processing one-dimensional data, it is often necessary to expand the dimension of the data, which leads to an increase in the amount of data input, thus increasing the resource consumption of the processing computer CPU and reducing the speed of model abnormal heart sound recognition, and this algorithm is not suitable for deployment on mobile smart devices with small computing resources. Therefore, it has good practical value to find a heart sound signal processing method suitable on mobile smart devices.

To solve the above problems, a dataset from the official “challenge heart sound” database was used.¹⁰ A minimal amount of heart sound feature input was performed using a deep neural network structure. An analytical model is obtained for fast training to identify abnormal heart sound signals.

Exemption from ethics was obtained from the Medical Ethics Committee of Hubei University of Science and Technology.

Overall System Design

As Figure 1 shows the structure of the above analysis model, the heart sound recognition system in this paper is mainly divided into two modules: data processing and neural network model. The first module is the data processing module, whose main functions are divided into data preprocessing, HSMM heart sound signal segmentation and heart sound signal feature value extraction. Among them, the data preprocessing is mainly to de-noise the input heart sound signal. The HSMM is responsible for the accurate segmentation of the heart sound signal into four segments: first heart sound, diastolic, second heart sound, and systolic. The feature extraction is to extract the feature data of the heart sound signal in terms of time and amplitude of the segmented four heart sound segments. Then, these feature data are sent to the trained neural network model. Once the neural network model receives the feature data, it will analyze the data to identify the abnormal heart sound signal and output the analysis results.

By using python language and with the help of TensorFlow 2.0 developed by Google in 2019, this paper can quickly build deep neural networks and train and test deep neural network models. Then, the data processing module uses MATLAB (2016 version) to process the data and packages the processing program as a dynamic link library that can be called directly in python.

Data Processing

Heart Sound Signal Dataset

The dataset used in this paper comes from the official heart sound database provided by the 2016 CHALLENGE HEART SOUND challenge. The database contains a total of 3240 heart sound clips from volunteers around the world. The quality of the heart sound clips varies due to the different health conditions of the volunteers and the different environments in which the heart sound was collected. Some of these heart sound clips are not even identifiable, and these unidentifiable heart sound clips are also classified as abnormal heart sound. These heart sound clips are usually between 5 and 120

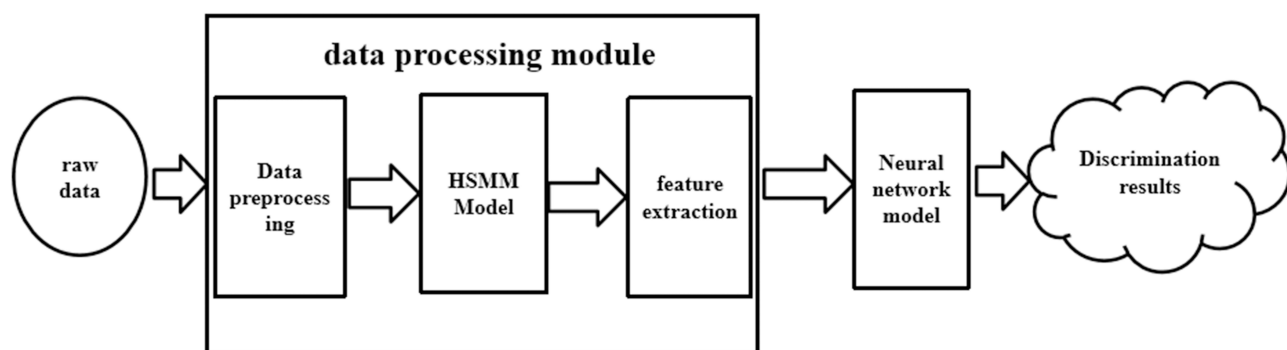


Figure 1 Overall system structure.

seconds in length, and each heart sound clip has been labeled by experts, so these heart sound data are authoritative. The details of this dataset are shown in Table 1.

Heart Sound Signal Preprocessing

The quality of heart sound in this dataset varies. Therefore, the heart sound signal needs to be de-noised before the heart sound segmentation. At present, the signal de-noise processing methods commonly used are low-pass filter, high-pass filter, band-pass filter, Fourier transform, short-time Fourier transform and other methods. These methods often filter out the useful signal when removing noise, while wavelet transform can greatly retain the detailed information of the signal while filtering out the noise, so it is widely used for signal de-noising processing. In this paper, according to the related research, db6 wavelet is used as wavelet base to perform 6-layer decomposition of heart sound signal, and soft thresholding is used to de-noise the decomposed signal.^{11,12}

Heart Sound Signal Segmentation

The heart sound signal contains important physiological information of human heart. The heart sound signal consists of one heartbeat cycle, and each heartbeat cycle is about 600ms – 1s. A heart cycle contains four parts: the first heart sound (S_1), the second heart sound (S_2), the third heart sound (S_3) and the fourth heart sound (S_4). Among these four parts, S_1 and S_2 are the most critical heart sound, also known as the basic heart sound. S_3 and S_4 , on the other hand, exist in specific populations and are not very relevant to the study. Within a cardiac cycle, the interval between S_1 and S_2 is the systolic period, and the interval between S_2 and the next cardiac cycle S_1 is the diastolic period.^{13,14}

Heart sound signal segmentation is a very important step in the heart sound abnormality recognition system. The effect of heart sound signal segmentation will directly affect the later accuracy of heart sound classification. The current heart sound segmentation methods include feature extraction analysis-based segmentation, machine learning-based segmentation, envelope-based segmentation and Hidden Markov model (HMM)-based segmentation.¹⁵ Among them, the heart sound signal segmentation method based on Hidden Markov model (HMM) requires very little computational effort in training and recognition, because of its ability to describe the smoothness and variability of short-time smooth signals well, and its dynamic modeling capability. Therefore, HMMs are very effective in dealing with speech-like signal analysis.¹⁶

HMM is a statistical model with implicit unknown parameters on time series. The HMM describes the relationship between the implied sequence and the corresponding observable sequence, a class of double discrete random sequences.¹⁷ As shown in Figure 2, $h_1, h_2 \dots h_{n-1}, h_n$ are the implied sequences of the HMM, which are unknowable. $v_1, v_2 \dots v_{n-1}, v_n$ are the observable sequences of HMM, which are knowable, and the rules between the implied sequences and the corresponding observed sequences are knowable, so we can then derive the implied sequences from the observed sequences.

The HMM is mainly composed of three parameters, namely, the initial state, the state transfer probability matrix and the observation probability matrix, and its model equation is

$$\theta = \{A, \pi, \varphi\} \quad (1)$$

In the above equation, A is the probability transfer matrix, π is the initial state, and φ is the observation probability matrix.

Table 1 Data Set

	Train	Valid	Test	Total
Normal heart sound	1558	503	514	2575
Abnormal heart sound	475	104	86	665
Total	2033	607	600	3240

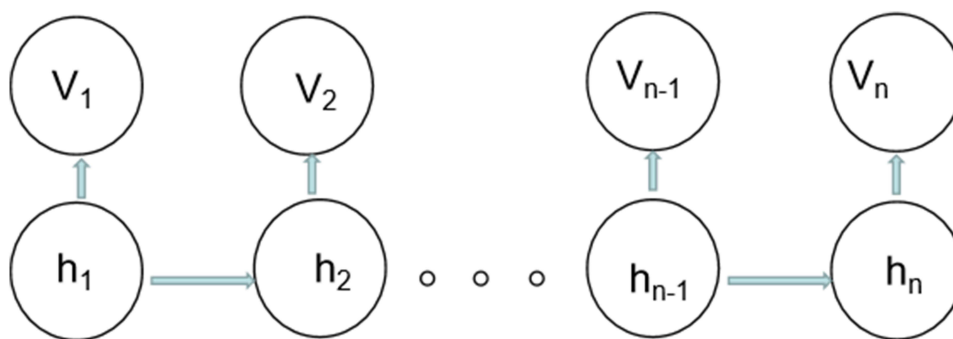


Figure 2 HMM process.

In the HMM, the state of the implied sequence transferred to the next moment is only related to the state of the implied sequence at this moment, but not to the residence time of the implied sequence with the observable sequence. This means that the probability of transitioning to the next state is the same no matter how long the model is in the same state, so there are some limitations of HMM in modeling.

In order to solve the limitations of HMM, hidden semi-Markov models (HSMM) have been proposed. This model makes up for the shortcomings of HMM by adding state residence time to the model. When judging the state transfer, the HSMM will only perform the state transfer when the joint probability distribution of the state jump time satisfies certain conditions, otherwise it still maintains the current state.¹⁸ Compared with HMM, HSMM adds a time parameter, which improves the accuracy and analysis ability of the model. The equation of the HSMM is

$$\theta = \{A, \pi, \varphi, D\} \tag{2}$$

In the above equation, A is the probability transfer matrix, π is the initial state, φ is the observation probability matrix, and D is the state residence time probability distribution.

A method based on HSMM for heart sound segmentation proposed by Springer David et al¹⁹ is used in this paper. As shown in Figure 3, the ECG signal R-wave position corresponds to the heart sound signal S1 position in the ECG signal and the heart sound signal T-wave end position corresponds to the heart sound signal S2 position in the ECG signal and the heart sound signal acquired simultaneously is used to segment the heart sound signal accurately. After that, the accurately segmented heart sound signal is used as the training set to train the HSMM. Finally, the trained HSMM is then used to segment the heart sound signal accurately.

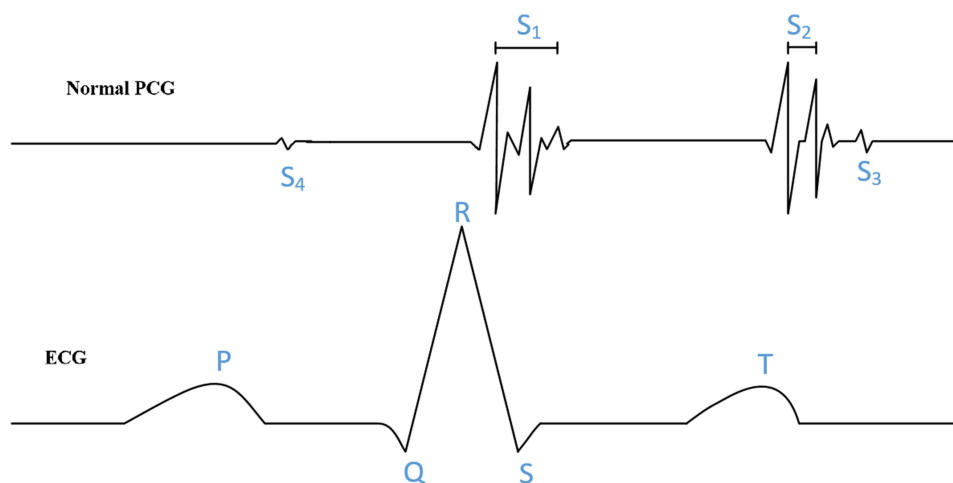


Figure 3 ECG signal and PCG relationship diagram.

Feature Extraction

After segmenting the heart sound signal, this paper obtains the feature values from the time of the heart sound signal divided into four segments and amplitude. In terms of time, the mean and standard deviation of each heart sound cardiac cycle, the mean and standard deviation of S_1 interval, the mean and standard deviation of S_2 interval, the mean and standard deviation of systole, the mean and standard deviation of diastole, the mean and standard deviation of systole to cardiac cycle ratio, the mean and standard deviation of diastole to cardiac cycle ratio, and the mean and standard deviation of systole to diastole ratio are obtained for each segment. In terms of amplitude, the mean and standard deviation of the absolute amplitude ratio between S_1 and systole and the mean and standard deviation of the absolute amplitude ratio between S_2 and diastole are selected in this paper.

Deep Neural Network Model Construction

In this paper, a number of neural networks are quickly built using TensorFlow2.0. The model is trained using the training set train and validation set valid in Table 1, and then the trained model is tested using the test set test. In this paper, the performance strengths and weaknesses of each neural network are evaluated by loss function and accuracy. These evaluations are then used to finally select the neural network structure. The experimental procedure is as follows:

Step 1: Find the optimal number of hidden layers of the neural network. In this paper, we borrow the selection method of neural network in the construction of heart sound intelligent analysis platform by Yurno Yang et al,²⁰ and construct neural networks with 1–5 hidden layers, respectively, so that the number of neurons in each hidden layer is the same, and then find the optimal number of layers by comparing with the loss function and accuracy. Considering the possible influence of the number of neurons on the loss function and accuracy, the number of neurons in each layer is set to 50 and 100 for 1–5 hidden layers, respectively. Experimenting separately on neural networks containing the above two numbers of neurons two data tables are obtained. By comparing the two tables, we can determine the most suitable number of hidden layers. The data obtained from the tests are shown in Tables 2 and 3.

From Tables 2 and 3, we can see that when the number of hidden layers is 1, the test accuracy is the highest, but the error is also the largest, so it is not suitable to be considered as the number of hidden layers. When the number of hidden

Table 2 Number of Neurons in Hidden Layer is 50 Test Data

Number of Hidden Layers	1	2	3	4	5
Train_loss	9.8775	3.7025	0.6715	0.5933	0.8389
Val_loss	14.1909	4.7949	0.5484	0.7584	0.5219
Val_accuracy	0.8194	0.8160	0.8194	0.7969	0.7865
Test_loss	6.5689	1.9301	0.3636	0.4143	0.4417
Test_accuracy	0.8854	0.8802	0.8819	0.8698	0.8490

Table 3 Test Data with 100 Hidden Neurons

Number of Hidden Layers	1	2	3	4	5
Train_loss	7.0267	0.6343	0.6611	0.6703	0.6641
Val_loss	8.6764	0.7915	0.6000	0.6109	0.5153
Val_accuracy	0.8090	0.7535	0.8160	0.8073	0.8299
Test_loss	4.7919	0.4325	0.3755	0.3888	0.4509
Test_accuracy	0.8837	0.8663	0.8819	0.8715	0.8715

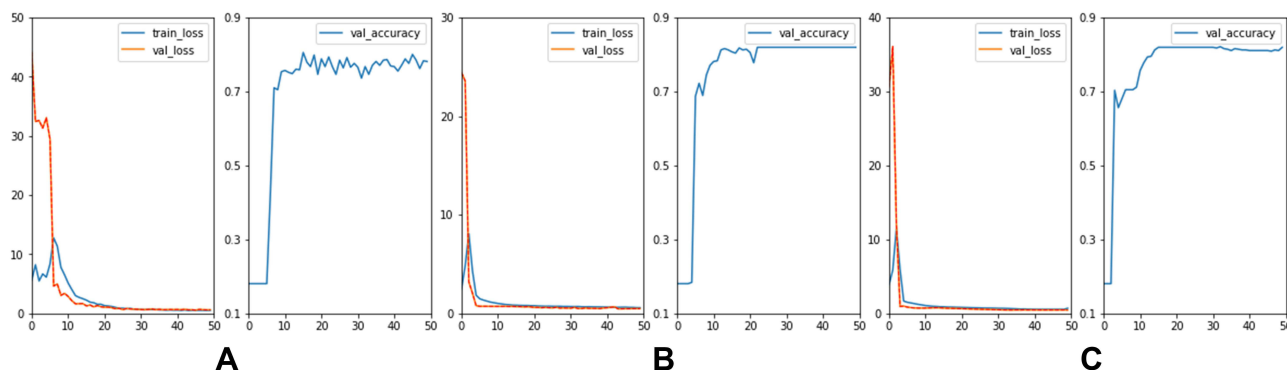


Figure 4 Loss function and test accuracy of different neural network structures. (A–C) denote the loss function and test accuracy of three neural network structures with designed hidden layers of 20–100–200, 100–200–20, and 200–100–20, respectively.

layers is 3, the loss is the smallest and the test accuracy is also higher, so the neural network with three hidden layers is chosen as the focus of the next study.

Step 2: According to the conclusion of the first step, three neural network structures with hidden layers of 20–100–200, 100–200–20, and 200–100–20 are designed, and these three neural network structures are set as A_1 , A_2 , A_3 . The same training set train_loss and validation set valid_loss are used to train the three neural networks, and the trained models are tested with the test set test. The training data and test data are shown in Figure 4A–C and Table 4. From Figure 4, it can be found that the result of A_2 shown in Figure 4B converges most quickly in the loss function and the loss is smallest, and the test accuracy is the highest compared with the results of A_1 and A_3 shown in Figure 4A and C. Therefore, A_2 is chosen to train the final anomalous heart sound recognition model in this paper.

Since the model recognition speed is related to the amount of data processed by the computer CPU, the less data the CPU needs to process, the faster the model recognition speed will be. Therefore, in this paper, the difference in recognition speed between the model designed in this paper and the mainstream CNNclassification model is verified to determine the difference between the two recognition algorithms in terms of CPU computing resources consumption and anomaly recognition speed. This paper calls the model trained by A_2 and the model trained by CNN to conduct 10 test experiments on the test set. The input feature vectors of the training set and the test set used in CNN will be expanded before the neural network is input.

The test results are shown in Tables 5 and 6. The maximum and minimum values are removed from the two tables, respectively, and the average value of the remaining 8 test results is found to be about 5.32 s for the average recognition

Table 4 Test Data of Three Neural Network Structures

Number of Hidden Layers	A_1	A_2	A_3
Train_loss	0.5274	0.5802	0.7056
Val_loss	0.6212	0.4749	0.4601
Val_accuracy	0.7812	0.8194	0.8194
Test_loss	0.3365	0.3881	0.3617
Test_accuracy	0.8767	0.8854	0.8785

Table 5 Recognition Speed of the Designed Model on the PCG Test Set

Number	1	2	3	4	5	6	7	8	9	10
Time/s	0.16	0.17	0.16	0.17	0.18	0.18	0.18	0.17	0.17	0.18

Table 6 Recognition Speed of CNN on the PCG Test Set

Number	1	2	3	4	5	6	7	8	9	10
Time/s	5.36	5.52	5.39	5.32	5.40	5.24	5.31	5.26	5.24	5.18

speed using CNN, and about 0.17 s for the model trained by the algorithm in this paper, and the algorithm in this paper is 31 times faster than the CNN algorithm in terms of recognition speed. Therefore, it can be concluded that the neural network model trained by the algorithm in this paper has faster recognition speed and occupies less CPU computing resources than the current mainstream method CNN abnormal heart sound recognition algorithm.

Conclusion

This system uses the powerful data processing function of MATLAB to process the raw heart sound signal, and packages the signal processing program into a MATLAB dynamic link library. The MATLAB dynamic link library can then be called directly in python to the trained deep neural network model for abnormal heart sound signal recognition. In the data processing function, the HSMM is used to accurately segment the heart sound signal, and 20 feature values are extracted for the segmented heart sound signal in terms of time and amplitude. This feature extraction method reduces the amount of data input to the deep neural network model and facilitates faster recognition and processing by the deep neural network model. In the neural network model, this paper conducts several experiments to find an optimal neural network structure for training, and obtains the anomaly heart sound recognition model. The model is also compared with the current mainstream abnormal heart sound classification model CNN classification model for experiments, and the results show that this classification model greatly outperforms the CNN classification model in recognition speed. The whole system is relatively streamlined and is expected to be deployed in embedded mobile devices, which is meaningful for telemedicine monitoring research. In future research, the abnormal heart sound recognition model can also be improved to further enhance the generalization ability and recognition accuracy of the system.

Acknowledgments

This work was partially funded by the Scientific Research Program of Hubei Education Department (No. B2020159). Thank you for sharing the official heart sound database provided by the 2016 CHALLENGE HEART SOUND challenge.

Disclosure

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Tang H, Li T, Park Y, Qiu T. Separation of heart sound signal from noise in joint cycle frequency-time-frequency domains based on fuzzy detection. *IEEE Trans Bio-Med Eng.* 2010;57(10):2438–2447. doi:10.1109/TBME.2010.2051225
2. Chen P, Zhang Q. Classification of heart sounds using discrete time-frequency energy feature based on S transform and the wavelet threshold denoising. *Biomed Signal Process Control.* 2020;57:1–11. doi:10.1016/j.bspc.2019.101684
3. Chaowen T, William W, Rong Z, Jiahua P, Lili Z. Convolutional neural network applied to heart sound signal classification study of precordial disease. *Comput Eng Appl.* 2019;55(12):174–180.
4. Tan CW, Wang WL, Zong R, Pan JH, Yang HB. Convolutional neural network-based heart sound signal classification algorithm for precordial disease. *J Biomed Eng.* 2019;36(05):728–736+744. doi:10.7507/1001-5515.201806031
5. Noman F, Salleh SH, Ting CM, et al. A Markov-switching model approach to heart sound segmentation and classification. *IEEE J Biomed Health.* 2020;24(3):705–716. doi:10.1109/JBHI.2019.2925036
6. Wang X, Yang H, Zong R, Pan J, Wang W. Heart sound classification algorithm based on subband envelope and convolution neural network. *J Biomed Eng.* 2021;38(05):969–978. doi:10.7507/1001-5515.202012024
7. Tariq Z, Shah SK, Lee Y. Feature-based fusion using CNN for lung and heart sound classification. *Sensors.* 2022;22:1521. doi:10.3390/s22041521
8. Linan M, Wei X, Chen N, Yang F. Heart sound classification algorithm based on deep convolution neural network. *Comp Meas Control.* 2021;29(08):211–217. doi:10.16526/j.cnki.11-4762/tp.2021.08.041
9. Liu WW, Sang SB, Zhang HP. Research on improved heart sound classification model based on CNN+LSTM. *Electron Des Eng.* 2022;30(02):38–42. doi:10.14022/j.issn1674-6236.2022.02.009.
10. Liu C, Springer D, Li Q, et al. An open access database for the evaluation of heart sound algorithms. *Physiol. Meas.* 2016;37(12):2181–2213.

11. Chen TH, Han LQ, Xing SH, Guo PY. Research on heart sound signal filtering method based on wavelet transform. *Comput Simul.* 2010;27(12):401–405.
12. Liu X, Sun J, Zhao. Y, Wang WL. Research on heart sound signal feature extraction and identification based on MFCC. *IET Sci Meas Technol.* 2018;41(02):1–5.
13. Berkaya SK, Uysal AK, Gunal ES, et al. A survey on ECG analysis. *Biomed Signal Process Control.* 2018;43:216–235.
14. Hassani K, Bajelani K, Navidbakhsh M, et al. Heart sound segmentation based on homomorphic filtering. *Perfusion.* 2014;29(4):351–359. doi:10.1177/0267659114523463
15. Xu CD, Zhou J, Ying DW, Hou LJ, Long QH. A DHMM-based segmentation method for low heart rate variability heart sound. *Data Acqui Proc.* 2019;34(04):605–614.
16. Guo XM, Duan Y, Zhong LS. Research on heart sound signal identification based on HMM and WNN. *Comput App Res.* 2010;27(12):4561–4564.
17. Bryan J, Levinson SE. Autoregressive hidden Markov models and the speech signal. *Procedia Comput Sci.* 2015;61(6):328–333. doi:10.1016/j.procs.2015.09.151
18. Schmidt SE, Holst-Hansen C, Graff C, et al. Segmentation of heart sound recordings by a duration-dependent hidden Markov model. *Physiol Meas.* 2010;31(4):513–529. doi:10.1088/0967-3334/31/4/004
19. Springer David B, Lionel T, Clifford Gari D. Logistic regression-HSMM-based heart sound segmentation. *IEEE Trans Biomed Eng.* 2016;63(4):822–832. doi:10.1109/TBME.2015.2475278
20. Yang YN, Zhang GL, Sun KX, Cheng XF. Construction of heart sound intelligent analysis platform based on deep learning network. *Inf Technol Dev.* 2019;29(07):130–134.

Medical Devices: Evidence and Research

Dovepress

Publish your work in this journal

Medical Devices: Evidence and Research is an international, peer-reviewed, open access journal that focuses on the evidence, technology, research, and expert opinion supporting the use and application of medical devices in the diagnosis, monitoring, treatment and management of clinical conditions and physiological processes. The identification of novel devices and optimal use of existing devices which will lead to improved clinical outcomes and more effective patient management and safety is a key feature of the journal. The manuscript management system is completely online and includes a very quick and fair peer-review system. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/medical-devices-evidence-and-research-journal>