**Progress in Medical Devices**
**eISSN: 2957-5478**

Review Article

# Research process on deep learning methods for heart sounds classification

Weifeng Wu, Yongqian Zhang, Qianfeng Xu, Jiuzhou Zhao, Rongguo Yan

School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China

**Declaration of conflict of interest:** None.

Received February 7, 2023; Accepted August 25, 2023; Published September 30, 2023

## Highlights

- Denoising, segmentation, and feature extraction of heart sounds as well as its classification process are reviewed.
- A detailed exposition of diverse deep learning methods for heart sounds classification is presented.

## Abstract

Cardiovascular diseases are still the primary threats to people's health around the world. Automatic heart sound classification technology, as a fast and efficient means for diagnosis and treatment, is of great clinical significance. With the rapid development of artificial intelligence technology, deep learning algorithms are widely used in automatic heart sound classification. This paper reviewed the key technologies related to the automatic classification of heart sounds in recent years, including heart sound denoising, segmentation, feature extraction, and classification recognition. The classification and recognition technologies related to deep learning are presented in detail, with a focus on the application and development of convolutional neural network and recurrent neural network, as well as various combination models for heart sound classification in the past five years.

**Keywords**: Cardiovascular disease, deep learning, heart sounds classification, convolutional neural network, recurrent neural network

## Introduction

Cardiovascular diseases (CVDs) are associated with a high fatality rate and posing an escalating threat to public health [1]. In 2019, approximately 17.9 million global deaths were attributed to CVDs, with heart disease and stroke constituting 85% of it. In low- and middle-income countries, approximately 75% of the financial burden endured by individuals with cardiovascular conditions can be attributed to a lack of adequate medical resources [1].

CVDs, such as hypertension, coronary heart disease, and heart failure, are featured by rapid onset and substantial damage. Thus, early detection of heart diseases is of vital importance. A significant physiological signal directly associated with CVDs is the heart sound signal, a primary tool for diagnosing CVDs. Through analysis of cardiac sound waves, physicians can identify abnormal characteristics. However, heart sound auscultation currently presents two key limitations. Firstly, human auditory sensitivity constraints may complicate the discernment of faint physiological sounds from human internal organs [2]. Secondly, the accuracy of heart sound auscultation may be critically constrained when physicians lack comprehensive knowledge and clinical experience, as the diagnostic results are limited by doctors' subjective expertise and subjectivity [3]. As a consequence, the use of computer software-based heart sound analysis in the precise diagnosis of cardiovascular illnesses is in line with the trend.

Phonocardiogram (PCG) refers to audio data captured via the electronic stethoscope, and the heart sound signals can be classified by using computer-assisted techniques [4]. The parameters of these heart sound signals can vary based on the condition of the heart. Notably, there is a substantial divergence between normal and pathological cardiac sounds, as their corresponding PCG signals differ in character-

**Address correspondence to: Rongguo Yan**, School of Health Science and Engineering, University of Shanghai for Science and Technology, NO.516, Jungong Road, Shanghai 200093, China. Email: yanrongguo@usst.edu.cn.

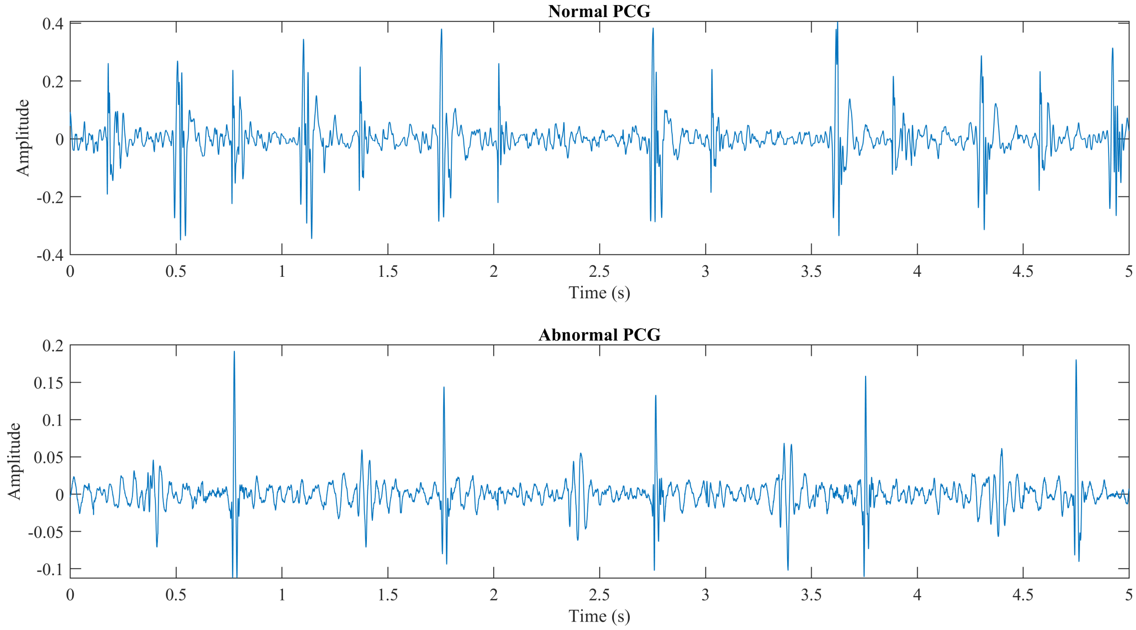


**Figure 1. Samples of PCG records from PhysioNet /Computing in Cardiology Challenge 2016 dataset.** PCG, Phonocardiogram.

istics such as magnitude, duration, intensity, spectrum, and uniformity [5].

With the precipitous advancement in digital signal processing and artificial intelligence technologies, deep learning has been extensively explored in the field of automatic heart sound analysis. It has the potential to reduce the labor costs associated with manual extracting features by employing data-driven feature learning. This approach is anticipated to bring about significant advances in the field of heart sound auscultation, and it is expected to be implemented in mobile devices.

## Heart sounds and datasets

Heart sound signals are mechanical vibrations generated by the contraction and relaxation of the myocardium, as well as the interactions of blood flow with the valves, the atrioventricular walls, and vascular systems. These vibrations are transmitted through the surrounding tissues to the skin, generating sound signals on body surface. Heart sound signals are quasi-periodic. A typical cardiac cycle signal consists of four components: the first heart sound, the systolic phase, the second heart sound, and the diastolic phase [6]. In certain populations, the third and fourth heart sounds may also exist.

Heart murmur is an abnormal heart sound signal in the cardiac cycle, which contains pathological information of various CVDs. The most common heart murmurs are caused by valvular diseases, such as aortic stenosis, mitral regurgitation, mitral stenosis, mitral valve prolapse, and tricuspid regurgitation [7].

Among the existing studies, PhysioNet/Computing in Cardiology Challenge 2016 dataset is the most commonly used dataset to verify the performance of the heart sound classification algorithms, which contains only normal and abnormal heart sounds [8]. **Figure 1** shows examples of these two heart sounds. There is another dataset that is widely used in the multi-classification task of heart sounds [9]. This dataset contains five categories of sounds, namely normal, aortic stenosis, mitral regurgitation, mitral stenosis, and mitral valve prolapse. **Figure 2** shows the examples of these five types of heart sounds.

## Process of heart sound classification

The automatic heart sound classification processes include denoising, segmentation, feature extraction, and classification.

### *Denoising*

Due to the susceptibility to external factors, heart sound signals are commonly disrupted by disturbances, such as electromagnetic interference, random noise, and respiratory sounds. Therefore, the preliminary step in automatic heart sound classification involves the denoising of these signals. Two commonly employed techniques for heart sound signal denoising are discrete wavelet transform (DWT) and empirical mode decomposition (EMD).

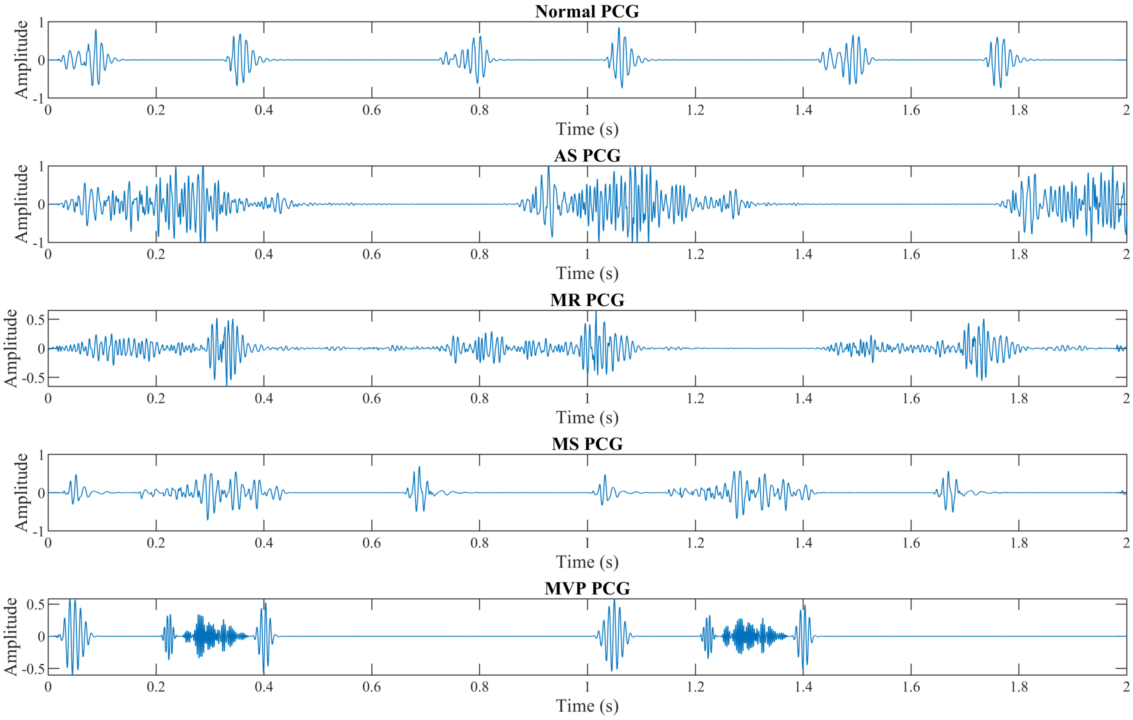


**Figure 2. Samples of PCG records of five categories from the dataset.** PCG, phonocardiogram; AS, aortic stenosis; MR, mitral regurgitation; MS, mitral stenosis; MVP, mitral valve prolapse.

DWT is a pivotal tool used for the analysis in the time-frequency domain. The core of this technique involves calculating the wavelet coefficients of both the noise and the heart sound signals. This is achieved by adjusting the scaling and shifting parameters of the so-called mother wavelet. Following this step, an appropriate threshold function can be identified and utilized to reconstruct the denoised heart sound signals [10]. Ali et al. conducted an examination on the influences of the wavelet basis function type, the threshold function type, and the signal decomposition level on the performance of the discrete wavelet denoising methodology [11]. In a distinct approach, Zhang et al. proposed an adjustable threshold function which can be fine-tuned using two parameters [12]. These adjustments alter the soft and hard threshold properties of the function, thus enhancing the integrity of the denoised heart sound signals. Despite the effectiveness of the DWT methodology, one significant drawback is that the denoised heart sounds may suffer from localized distortion, especially in the presence of anomalous heart sound signals.

Contrarily, EMD technique separates heart sound signals into several intrinsic mode functions, each comprising both heart sound components and noise components. By selectively eliminating the noise components, heart sound denoising can be achieved [13]. Furthermore, there have been experiments that combine the two methods. For instance, Dong et al. proposed a wavelet transform denoising technique grounded in complementary population EMD [14]. This method partitions the cardiac sound signals into intrinsic mode function components of varying frequencies. Then, the autocorrelation coefficient is deployed to distinguish between the pertinent signals and the noise within the range of the modal components. Following this, the heart sound signals are filtered and reconstructed employing the wavelet transform, resulting in the acquisition of the denoised heart sound signal.

The two methods have shown broad applicability in the field of heart sound denoising, although each of them faces some challenges. For DWT-based methods, their denoising effectiveness largely depends on thresholds, wavelet basis functions, and decomposition levels. The selection of these variables greatly affects the results, thereby leading to difficulties in parameter selection [15]. In contrast, EMD-based methods are highly sensitive to noise and suffer from issues such as the modal confusion [16]. Therefore, despite their wide application, these two strategies still have a large amount of research potential and room for optimization.

### *Segmentation*

Heart sound segmentation is the division of the heart sound signals into multiple cardiac cycles. In order to improve the accuracy of heart sound classification, automatic heart sound analysis focuses on segmentation of heart sounds. Currently, there are many successful heart sound segmentation methods, which are roughly classified into two categories. One is envelope-based segmentation method, and the other is based on statistical modelling.

The envelope-based segmentation is commonly employed in segmentation algorithms for heart sounds. The cardiac cycle is determined by detecting the peak value of the first or second heart sound following the extraction of the heart sound envelope. In the heart sound segmentation methods based on envelope, there are techniques, such as normalized average Shannon energy, Hilbert transform, homomorphic filtering, heart sound feature waveform extraction, Hilbert-Huang transform, and short-time corrected Hilbert transform [17-22]. In an in-depth research, Choi et al. examined and analyzed the envelope extraction techniques of the normalized Shannon envelope, the Hilbert transform envelope, and the heart sound characteristic waveform [23]. They found that the heart sound characteristic waveform envelope extraction method has uniform first heart sound and second heart sound information compared to the normalized Shannon envelope and Hilbert transform envelope curve.

In addition to using envelope analysis for heart sound segmentation, methods based on statistical models are also widely used, such as the hidden Markov model, the ensemble EMD method, the K-means clustering algorithm, and dynamic clustering [24-27]. These methods segment heart sound signals by utilizing different characteristics of the heart sound signals, such as the distribution of time-frequency energy, systole, and time-relatedness.

Both the two types of methods are widely used in the field of heart sound segmentation, but each has its own drawbacks. Envelope-based heart sound segmentation methods are susceptible to environmental noise interference, often missing the true peaks of the heart sounds but detecting noise peaks instead. Heart sound segmentation methods based on statistical models are constrained by the high degree of signal specificity in heart sounds across different individuals, making it challenging to model all heart sound signals with a unified model.

### *Feature extraction*

The main purpose of feature extraction of heart sounds is to find some effective features to replace the high-dimensional original information. Presently, the characteristics employed for classifying heart sounds can be categorized as time domain, frequency domain, statistical domain, and time-frequency domain features. The methods including DWT, continuous wavelet transform, short-time Fourier transform, and Mel frequency cepstral coefficients (MFCC) are widely  feature extraction [28-32]. Mel frequency spectral coefficient, as a MFCC without the discrete cosine transform, retains more of the original data [33-39]. MFCC and Mel frequency spectral coefficient can extract comparable characteristics of human perception of loudness and tone from audio data, hence they are commonly utilized in speech signal processing.
In the field of heart sound feature extraction, several improved methods have been proposed. For instance, Deng et al. proposed an improved MFCC feature extraction method, which reduces the computational complexity and can dynamically represent the heart sound signals [40]. Likewise, Cheng et al. proposed an improved approach for classifying heart sounds by integrating MFCC with Gammatone Frequency Cepstral Coefficients [41].

In conclusion, due to the unique nature of heart sound signals, feature extraction is a critical step in the automatic classification of heart sounds, playing a significant role in the accuracy of subsequent heart sound classification results. Therefore, it is essential to continuously conduct research and improve methodology to optimize heart sound feature extraction.

### *Classification*

Classification of heart sounds primarily provides qualitative data of heart sound detection and often separates heart sound signals into normal and pathological forms. Currently, machine learning techniques, such as support vector machine, hidden Markov model, artificial neural network, K-nearest neighbor method, and Euclidean distance, are utilized extensively in heart sound classification. By extracting the P value of features, Yadav et al. examined the performance of support vector machines, Naive Bayes model, random forest, and K-nearest neighbor in classifying heart sounds [42]. Deep learning techniques, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), which are detailed in depth in Section 4, are the most popular approaches for classifying heart sounds.

## Application of deep learning in heart sound classification

**Table 1. Literature on heart sound classification using deep learning**

| No. | Year | Reference | Method | Input Features | Results | | |
|---|---|---|---|---|---|---|---|
| 1 | 2018 | Meintjes et al. [29] | CNN | Scalograms | Sensitivity 87.4% | Specificity 86.7% | Accuracy 93.8% |
| 2 | 2018 | Bozkurt et al. [43] | CNN | Sub-band Envelopes | Sensitivity 84.5% | Specificity 78.5% | MAcc 81.5% |
| 3 | 2019 | Tan et al. [33] | CNN | MFSC | | Accuracy 89.6% | |
| 4 | 2019 | Tan et al. [34] | CNN | MFSC | Sensitivity 91% | Specificity 88% | Accuracy 89.5% |
| 5 | 2021 | Xu et al. [44] | CNN | PSD | Sensitivity 77.6% | Specificity 94.6% | Accuracy 84.7% |
| 6 | 2021 | Boulares et al. [31] | CNN | MFCC | Sensitivity 94.6% | Specificity 94.6% | Accuracy 97% |
| 7 | 2022 | Li et al. [32] | CNN | MFCC | Sensitivity 89.5% | Specificity 89.7% | |
| 8 | 2021 | Shen et al. [45] | CNN | Spectrograms | Sensitivity 91.2% | Specificity 92.1% | Accuracy 91.1% |
| 9 | 2021 | Wang et al. [46] | CNN | Sub-band Envelopes | | Accuracy 85.7% | |
| 10 | 2022 | Azam et al. [47] | CNN | MFCC&Fbank | Accuracy 86.2% | F1 84% | MAcc 85.1% |
| 11 | 2022 | Chen et al. [41] | CNN | MFCC&GFCC | Sensitivity 93.8% | Specificity 88.8% | Accuracy 91.3% |
| 12 | 2021 | Chen et al. [35] | CNN | MFSC | | Accuracy 99.5% | |
| 13 | 2021 | Meng et al. [36] | CNN | MFSC | Sensitivity 98.5% | Specificity 97.2% | Accuracy 98.2% |
| 14 | 2017 | Maknickas et al. [37] | CNN | MFSC | Sensitivity 80.6% | Specificity 86.66% | Accuracy 93.7% |
| 15 | 2022 | Zhang et al. [38] | CNN | MFSC | Sensitivity 92.5% | Specificity 98.6% | Accuracy 96.6% |
| 16 | 2022 | Zhu et al. [39] | CNN-LSTM | MFSC | Sensitivity 84.3% | Specificity 85.6% | Accuracy 84.4% |
| 17 | 2020 | Li et al. [48] | CNN-LSTM | Spectrograms | | Accuracy 85.7% | |
| 18 | 2022 | Liu et al. [49] | CNN-LSTM | Second order spectrum | Sensitivity 96.6% | Specificity 95.6% | Accuracy 95.3% |
| 19 | 2022 | Al-Issa et al. [50] | CNN-LSTM | FFT | Accuracy 93.8% | F1 85.6% | AUC 95.1% |
| 20 | 2022 | Chen et al. [51] | CNN-LSTM | Sequence | Sensitivity 87.0% | Specificity 89.0% | Accuracy 86.0% |

Note: CNN, convolutional neural network; MFSC, Mel frequency spectral coefficient; PSD, power spectral density; MFCC, Mel frequency cepstral coefficients; GFCC, Gammatone Frequency Cepstral Coefficients; LSTM, long short-term memory networks; FFT, fast fourier transform; AUC, area under the curve; MAcc, mean accuracy.

This part reviews the applications of CNNs, RNNs, and Hybrid methods in heart sound classification. The research literature in recent years is shown in **Table 1**.

*CNN methods for heart sound classification*

CNN is a traditional model of deep learning. The LeNet is a typical CNN model proposed by Lecun et al. , as illustrated in **Figure 3** [52]. It consists of an input layer, two convolutional layers, two pooling layers, two fully connected layers, and an output layer. CNNs extract data features through the convolution layer based on local perception and decrease the parameter scale of the model via a weight-sharing technique, and the convolution process has local perception and weight sharing characteristics. The pooling layer, also known as the downsampling layer, efficiently reduces network parameters to minimize overfitting. The completely linked layer resides at the end of the CNNs. Each neuron of the fully connected layer is connected to the neurons of the previous layer, and the multi-dimensional feature vectors calculated by the previous layer are mapped into one-dimensional vectors, which reduces the influence of the feature position arrangement on the classification results and increases the network's robustness.

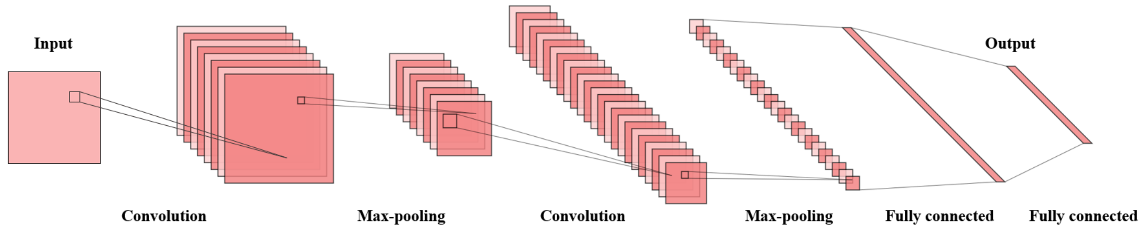


**Figure 3. Network structure of LeNet.**

Several optimization algorithms based on LeNet have been explored by researchers, with representative algorithms including AlexNet, GoogleNet, VGGNet, ResNet, and DenseNet. These algorithms have shown exemplary performance in heart sound classification tasks. For instance, Maknickas et al. proposed a heart sound classification model based on AlexNet, and obtained an accuracy of 93.7% [37]. Distinct from the most fundamental CNN model, AlexNet employs ReLU instead of Sigmoid as the activation function to avoid gradient vanishing when the network is deep and utilizes dropout to reduce the complexity of the model to prevent overfitting.

The performance of CNNs is affected by depth, width, and the size of the convolution kernel. The deepening of the number of network layers will have more linear transformations, which will help improve network performance. However, as the depth reaches a certain degree, model training becomes more difficult, culminating in the disappearance of gradients. In 2015, He et al. presented ResNet based on residual learning module, with network parameters of up to 152 layers, which can more effectively address the training challenges caused by network depth [53]. ResNet is currently more common in the classification of heart sounds.

Subsequent studies have built upon this foundation. For example, Zhang et al. presented a ResNet50 based on bilinear homologous transformation [38]. The model is comprised of multi-layer combinations of bottleneck structures with cross-layer connections and incorporates convolutional attention modules, better cross-entropy loss functions, and label smoothing techniques extensively. The greatest accuracy score reached 96.6%. Not resting on these laurels, further research pushed the boundaries of accuracy. Chen et al. proposed a heart sound classification model based on Inception-ResNet, using ResNet to increase the network's depth and convergence speed, and to improve the Inception network's accuracy [41]. The maximum possible score could reach 99.5% correctness.

However, the superiority of ResNet over other network models is not always clear-cut. As illustrated by Li et al., when comparing the performance of VGG and ResNet, the lack of a more complex network model can lead to an equivalent performance of the two [32]. In fact, the VGG model employing Focal loss and Log Mel features achieved an Unweighted Average Recall of 89.6% across all test groups.

Efforts to enhance CNN performance didn't stop at ResNet. Meng et al. compared the performance of CNN based on single connection (SingleCNN), skip connection (SkipCNN), and dense connection (DenseCNN) in the classification of heart sounds [36]. They investigated the effect of several convolutional layer connection methods on the classification of heart sounds. DenseCNN includes dense connections that enable the reuse of heart sound information, allowing the network to extract more heart sound features with fewer layers. DenseCNN had superior performance compared to the other two models, achieving 98.2% accuracy, 98.5% sensitivity, 97.2% specificity, and 0.0557 loss.

The application of CNNs in heart sound classification has witnessed considerable progression over the years, with successive studies contributing fresh perspectives and avenues for future exploration. These advancements underscore the extensive potential of CNNs within the realm of biomedical engineering. However, they also remind us of the crucial role that continuous inquiry and research play in further propelling the advancement within this discipline.

### *RNN methods and hybrid methods for heart sound classification*

Compared with CNNs, RNNs have excellent time modeling ability and is widely used in natural language processing, speech recognition, and other fields. RNNs are feed-forward neural network variant with internal memory. With the

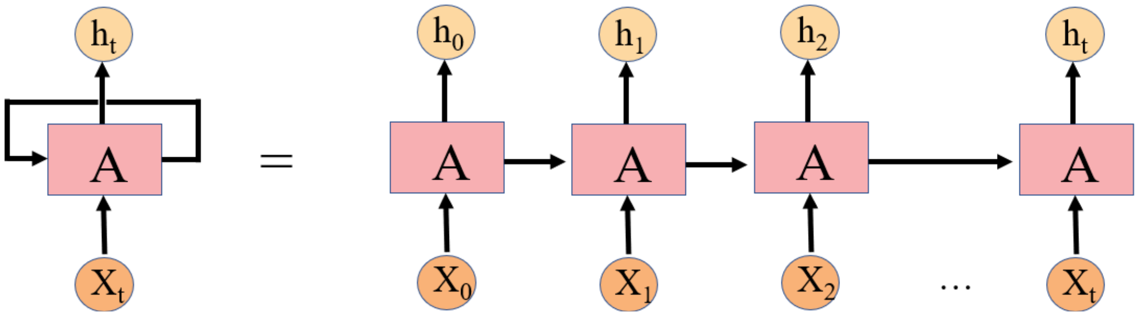


**Figure 4. Network structure of RNN.** RNN, recurrent neural network.

help of the internal memory structure, RNNs can memorize historical data and effectively predict future data. The typical structure of the RNN model is shown in **Figure 4**. The Xt is the input of the current module, and the ht is calculated as the output and participates in the input of the next module. Although RNNs have great advantages in time series modeling, it also has problems such as vanishing and exploding gradient. In order to solve these problems, researchers have proposed improved RNN models, including long short-term memory networks (LSTM) and gated recurrent units (GRU). LSTM is composed of forgetting gate, input gate and output gate, which make up for the deficiency of ordinary RNNs by introducing linear self-circulation unit. GRU is an important variant of LSTM. It uses update gate instead of forgetting gate and output gate, which makes GRU have fewer parameters and is easier to converge.

Further research by Bao et al. delved into the effects of different feature inputs on these network models [54]. In their study, the MFCC static features, and the extended MFCC with dynamic features, were used respectively as inputs to test their impacts on CNN and RNN models. Their findings suggest that the extended MFCC did not significantly improve the performance of RNN, although the accuracy of CNN using MFCC with dynamic features did show a slight improvement. Interestingly, when the static features of MFCC were used as input, RNN models outperformed CNN. However, among different RNN models, such as LSTM, bidirectional LSTM, GRU, and bidirectional GRU, no significant differences in performance were observed.

In audio processing, the CNN is more suitable for identifying specific portions of sounds, while the RNN is more excellent at handling sequences that vary over time. Each has its respective advantages, hence, the integration of CNN and RNN has become a prevalent direction in research.

A variety of studies have examined this approach. Chen et al. proposed the use of one-dimensional sequences as feature inputs, and the performance of a network combination of CNN and LSTM outperforms single networks, achieving an accuracy of 86.0% [36]. While using one-dimensional sequences as feature inputs does not have the best performance, it is beneficial for applications in real-time detection. On a similar note, Li et al. utilized CNN to extract the frequency domain features of data, complemented by the use of RNN to draw out the time domain features. This method yielded a classification accuracy rate of 85.7% [48]. In another study, Liu et al. explored the effectiveness of different CNN-LSTM structures in heart sound classification, demonstrating that the convolution kernel's size within the CNN can yield better outcomes within a certain range [49]. Additionally, they found that the single-layer unidirectional and single-layer bidirectional LSTM models were the most effective. However, they also noted that increasing the number of LSTM layers could impact the model's accuracy negatively. Despite this, their approach achieved a noteworthy classification accuracy of 95.2%.

Moreover, Al-Issa et al. designed a lightweight CNN-LSTM model to distinguish between five categories of heart valvular disorders [50]. They highlighted that using frequency domain features post-FFT transformation as input could enhance heart sound classification performance. They achieved an accuracy of 93.8%, a F1 score of 85.6%, and an AUC of 95.1% on the binary classification dataset. Additionally, on the enhanced five-class classification dataset, they attained an accuracy of 99.87%, a F1 score of 99.87%, and an AUC of 99.85%.

The integration of CNNs and RNNs, among other networks, is a well-adopted strategy, serving to unite the individual strengths of these architectures whilst offsetting their respective weaknesses. CNNs are particularly proficient in handling tasks related to image processing or spatial pattern recognition. On the other hand, RNNs are highly effective in dealing with sequence data and temporal pattern recognition. By combining these two, the stability and generalization capability of the model can be greatly enhanced. In real-world applications, this hybrid approach has been proven to be successful in numerous instances, demonstrating the flexibility offered by deep learning techniques and highlighting the potential of such combined models.

## Summary and prospective

In recent years, as the frequency of CVDs has increased, automatic heart sound diagnosis technologies have enabled clinicians to diagnose patients more quickly and precisely, contributing to the healthcare system. This review summarizes the most recent research in the field of heart sound automatic classification during the past five years, the research status of denoising, segmentation, and feature extraction, and the use of a deep learning model in heart sound classification.

With the ongoing advancement of deep learning, numerous effective deep learning models, including CNN and RNN algorithms, have been presented and used in heart sound categorization. The algorithm model that combines the properties of the CNN and RNN is currently the study focus, and examination of datasets reveals that these algorithms have excellent performance and strong classification results. However, the complexity of the present heart sound categorization model needs to be reduced. The deployment of the lightweight model in mobile devices is anticipated to significantly improve the application of heart sound automatic categorization technologies.

Due to the lack of heart sound datasets, the majority of existing heart sound classification algorithms utilize both normal and abnormal heart sound datasets. However, various kinds of cardiac murmurs exist. Thus, to further assist physicians in the diagnosis of cardiovascular disorders, it is still important to increase and improve the heart sound datasets and investigate the multi-classification model in order to more precisely identify the types of heart murmurs.

## References

[1] World Health Orgnization. Cardiovascular diseases (CVDs): Fact sheets. 2021.

[2] Xu WZ, Yu K, Ye JJ, et al. Automatic pediatric congenital heart disease classification based on heart sound signal. Artif Intell Med 2022;126:102257.

[3] Germanakis I, Petridou ET, Varlamis G, et al. Skills of primary healthcare physicians in paediatric cardiac auscultation. Acta Paediatr 2013;102(2):e74-e8.

[4] Fernando T, Ghaemmaghami H, Denman S, et al. Heart Sound Segmentation Using Bidirectional LSTMs With Attention. IEEE J Biomed Health Inform 2020;24(6):1601-1609.

[5] Wang M, Tang H, Feng TF, et al. Continuous and Noninvasive Estimation of Right Ventricle Systolic Blood Pressure Using Heart Sound Signal by Deep Bidirectional LSTM Network. Appl Sci (Basel) 2020;10(16):5466.

[6] Debbal SM, Bereksi-Reguig F. Time-frequency analysis of the second cardiac sound in phonocardiogram signals. Medical Phys 2005;32(9):2911-2917.

[7] Itagaki S, Adams DH, Anyanwu AC. Triggers for Surgical Referral in Degenerative Mitral Valve Regurgitation. Circu J 2013;77(1):28-34.

[8] Liu CY, Springer D, Li Q, et al. An open access database for the evaluation of heart sound algorithms. Physiol Meas 2016;37(12):2181-213.

[9] Yaseen, Son GY, Kwon S. Classification of Heart Sound Signal Using Multiple Features. Appl Sci (Basel) 2018;8(12):2344.

[10] Emmanouilidou D, McCollum ED, Park DE, et al. Adaptive Noise Suppression of Pediatric Lung Auscultations With Real Applications to Noisy Clinical Settings in Developing Countries. IEEE Trans Biomed Eng 2015;62(9):2279-2288.

[11] Ali MN, El-Dahshan ELSA, Yahia AH. Denoising of Heart Sound Signals Using Discrete Wavelet Transform. Circuits Sys Signal Process 2017;36(11):4482-4497.

[12] Zhang NN. The Application of an Improved Wavelet Threshold Function in De-noising of Heart Sound Signal. 2020 32nd Chinese Control And Decision Conference (CCDC) 2020;1768-1772.

[13] JiaRY, Jiann SS, Huang NE. Complementary ensemble empirical mode decomposition: a novel noise enhanced data analysis method. Adv Adapt Data Anal Theory Appl (Singapore) 2010;2(2):135-156.

[14] Dong L, Guo X, Zheng Y. Wavelet packet de-noising algorithm for heart sound signals based on CEEMD. J Vib Shock 2019;38(9):192-198,222.

[15] Li R, Bao C, Dou H. Speech enhancement algorithm based on wavelet transform. J Data Acquis Process (China) 2009;24(3):362-368.

[16] Gu J, Peng YX. An improved complementary ensemble empirical mode decomposition method and its application in rolling bearing fault diagnosis. Digital Signal Process2021;113:103050.

[17] Maglogiannis I, Loukis E, Zafiropoulos E, et al. Support Vectors Machine-based identification of heart valve diseases using heart sounds. Comput Methods Programs Biomed 2009;95(1):47-61.

[18] Huang NE, Shen Z, Long SR, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society a-Mathematical Physical and Engineering Sciences. 1998;454(1971):903-995.

[19] Kumar D, Carvalho P, Antunes M, et al. Heart Murmur Recognition and Segmentation by Complexity Signatures.Annu Int Conf IEEE Eng Med Biol 2008; 2128-2132.

[20] Jiang ZW, Choi SJ. A cardiac sound characteristic waveform method for in-home heart disorder monitoring with electric stethoscope. Expert Syst Appl 2006;31(2):286-298.

[21] Tseng YL, Ko PY, Jaw FS. Detection of the third and fourth heart sounds using Hilbert-Huang transform. BioMed Eng OnLine 2012;11(1):8.

[22] Sun SP. An innovative intelligent system based on automatic diagnostic feature extraction for diagnosing heart diseases. Knowledge-Based Systems. 2015;75:224-238.

[23] Choi SJ, Jiang ZW. Comparison of envelope extraction algorithms for cardiac sound signal segmentation. Expert Syst Appl 2008;34(2):1056-1069.

[24] Gill D, Gavrieli N, Intrator N, et al. Detection and identification of heart sounds using homomorphic envelogram and self-organizing probabilistic model. Comput Cardiol 2005;957-960.

[25] Papadaniil CD, Hadjileontiadis LJ. Efficient Heart Sound Segmentation and Extraction Using Ensemble Empirical Mode Decomposition and Kurtosis Features. IEEE J Biomed Health Inform 2014;18(4):1138-1152.

[26] Gupta CN, Palaniappan R, Swaminathan S, et al. Neural network classification of homomorphic segmented heart sounds. Appl Soft Comput 2007;7(1):286-297.

[27] Tang H, Li T, Qiu TS, et al. Segmentation of heart sounds based on dynamic clustering. Biomed Signal Proce Control 2012;7(5):509-516.

[28] Coskun H, Deperlioglu O, Yigit T, et al. Implementation of Wavelet Transform Extrasystole Heart Sound With Convolution Method for Feature Extraction. 25th Signal Processing and Communications Applications Conference (SIU) 2017;1-4.

[29] Meintjes A, Lowe A, Legget M, et al. Fundamental Heart Sound Classification using the Continuous Wavelet Transform and Convolutional Neural Networks. 40th Annual International Conference of the IEEE-Engineering-in-Medicine-and-Biology-Society (EMBC) 2018;409-412.

[30] Chen QY, Zhang WB, Tian X, et al. Automatic Heart and Lung Sounds Classification using Convolutional Neural Networks. Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA) 2016;1-4.

[31] Boulares M, Alotaibi R, AlMansour A, et al. Cardiovascular Disease Recognition Based on Heartbeat Segmentation and Selection Process. Int J Environ Res Public Health. 2021;18(20).

[32] Li Z, iChang Y, Schuller BW. CNN-Based Heart Sound Classification with an Imbalance-Compensating Weighted Loss Function. Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual International Conference 2022;4934-4937.

[33] Tan ZW, Wang WL, Zong R, et al. Research on Classification of Congenital Heart Disease Heart Sound Signal Using Convolutional Neural Network. Compu Eng Appl 2019;55(12):174-180.

[34] Tan ZW, Wang WL, Zong R, et al. Classification of heart sound signals in congenital heart disease based on convolutional neural network. J Biomedl Eng 2019;36(5):726-738.

[35] Chen W, Sun Q, Qi YY, et al. Deep Convolutional Neural Networks for Heart Sound Classification. Compu Eng Appl 2021;57(16):182-189.

[36] Meng LN, Xie HW, Ning C, et al. Heart Sound Classification Algorithm Based on Deep Convolutional Neural Network. Compu Meas Control. 2021;29(8):211-217,222.

[37] Maknickas V, Maknickas A. Recognition of normal-abnormal phonocardiographic signals using deep convolutional neural networks and mel-frequency spectral coefficients. Phys Meas 2017;38(8):1671-1684.

[38] Zhang JF, Zhang GY. Heart sound classification based on homologous bilinear residual network. Inf Technol 2022(5):36-42,47.

[39] Zhu LX, Qian K, Wang ZH, et al. Heart Sound Classification based on Residual Shrinkage Networks. 2022 44th Annual International Conference of the IEEE Engineering

in Medicine & Biology Society (EMBC) 2022;4469-4472.
[40] Deng MQ, Meng TT, Cao JW, et al. Heart sound classification based on improved MFCC features and convolutional recurrent neural networks. Neural Networks 2020;130:22-32.
[41] Chen C, Pan J, Sun J. Research on Heart Sound Classification of CHD Based on MFCC and GFCC Mixed Features. Compu Simul 2022;39(7):263-269.
[42] Yadav A, Singh A, Dutta MK, et al. Machine learning-based classification of cardiac diseases from PCG recorded heart sounds. Neural CompuAppl 2020;32(24):17843-17856.
[43] Bozkurt B, Germanakis I, Stylianou Y. A study of time-frequency features for CNN-based automatic heart sound classification for pathology detection. Compu Biol Med 2018;100:132-143.
[44] Xu CD, Xin PL, Zhou J, et al. Classification of Heart Sounds Using Power Spectral Density and Convolutional Neural Networks. Compu Eng App 2021;57(10):125-132.
[45] Shen YJ, Wang X, Tang M, et al. Recognition of S1 and S2 heart sounds with two-stream convolutional neural networks J Biomed Eng 2021;38(1):138-144.
[46] Wang XZ, Yang HB, Zong R, et al. Heart sound classification based on sub-band envelope and convolution neural network. J Biomed Eng 2021;38(5):969-978.
[47] Azam FB, Ansari MI, Nuhash SISK, et al. Cardiac anomaly detection considering an additive noise and convolutional distortion model of heart sound recordings. Artif Intell Med 2022;133:102417.
[48] Li W, Yang XD, Chen K. Heart sound classification based on CNN and RNN. Compu Eng Des 2020;40(1):46-51.
[49] Liu W, Sang S, Zhang H. Study on improved heart sound classification model based on CNN+LSTM. Electron Des Eng 2022;30(2):38-42.
[50] Al-Issa Y, Alqudah AM. A lightweight hybrid deep learning system for cardiac valvular disease classification. Sci Rep 2022;12(1):14297.
[51] Chen D, Xuan WP, Gu YX, et al. Automatic Classification of Normal-Abnormal Heart Sounds Using Convolution Neural Network and Long-Short Term Memory. Electronics 2022;11(8).
[52] Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition. Proc IEEE 1998;86(11):2278-2324.
[53] He KM, Zhang XY, Ren SQ, et al. Deep Residual Learning for Image Recognition. IEEE Conference on Computer Vision and Pattern Recognition 2016;770-778.
[54] Bao XQ, Xu YJ, Kamavuako EN. The Effect of Signal Duration on the Classification of Heart Sounds: A Deep Learning Approach. Sensors 2022;22(6):2261.