Highlights

Rapid diagnosis technology for acute heart failure based on auscultation

Hui Yu,Zhaoyu Qiu,Zhigang Li,Jinglai Sun,Guangpu Wang,Jing Zhao,Shuo Wang

- An auscultation dataset has been established, containing 2999 recordings from heart failure patients, each with rich annotations, and are publicly accessible on https://github.com/qiuzhaoyu/AHF-Rapid-Diagnosis/Database.

Rapid diagnosis technology for acute heart failure based on auscultation

Hui Yu[a], Zhaoyu Qiu[a], Zhigang Li[b], Jinglai Sun[a], Guangpu Wang[a], Jing Zhao[a,*] and Shuo Wang[a,*]

[a]Department of Biomedical Engineering, Tianjin University, Tianjin, 300072, China
[b]Department of Emergency Medicine, Tianjin 4TH Centre Hospital, Tianjin, 300142, China

ARTICLE INFO

Keywords:
Heart Sound Signal
Acute Heart Failure
Mel Frequency Cepstrum Coefficient
Lightweight Deep Learning Model

ABSTRACT

Background and objectives: Acute Heart Failure (AHF) leads to over 26 million hospital admissions worldwide annually and is now a major global health concern. Currently, AHF diagnosis relies on biochemical markers and echocardiography, which takes more than 20 minutes. Auscultation, a quick and non-invasive clinical practice, is used alongside the gold standard. Recognizing the need for rapid clinical AHF diagnosis, this paper presents a model for feature extraction and diagnosis using short heart sound signals.

Methods: In this paper, discrete wavelet transform is applied for heart sound denoising, and the Mel Frequency Cepstrum Coefficient is applied for feature extraction. A new DenseHF-Net is proposed for the diagnosis of heart failure. A feature fusion method is proposed for multi-region fusion auscultation, including mitral, aortic, and pulmonic valves. An ensemble method is proposed for long auscultation in the mitral valve region.

Results: An auscultation dataset containing 2999 recordings has been established, each with rich annotations. A proposed wavelet denoising algorithm achieves a signal-to-noise ratio of 7.8 dB. For multi-region fusion auscultation, using DenseHF-Net, the average accuracy is 99.35%. For mitral valve ensemble auscultation, using DenseHF-Net, the average accuracy is 94.41%.

Conclusions: The above method enables rapid auscultation of AHF, providing accurate results based on a 3-second auscultation recording. Multi-region fusion auscultation achieves good auscultation accuracy, but mitral valve ensemble auscultation provides a good balance between efficiency and accuracy. The above research has the potential to be used for cardiac auscultation that can be used on mobile phones, cloud, or electronic gloves.

1. Introduction

1.1. Background

Cardiovascular disease (CVD), including heart failure (HF) and stroke, has become the leading cause of death and disability worldwide [1–4]. The upward trend in the age-standardized rate of CVD is occurring in almost all non-high-income countries [2]. Heart failure is an end-stage in the development of CVD characterized by dysfunction of the contractile/stretching function of the heart. Acute heart failure is a leading cause of emergency hospital admissions, particularly in the elderly population [5, 6]. Acute heart failure (AHF) accounts for more than 26 million hospital admissions each year, with a mortality rate of 20-30% [7]. Emergency procedures for AHF include ambulance response, door-to-balloon, emergency, and ward transfer. Shortening the first two procedures can significantly reduce patient mortality and morbidity [8, 9]. Taking myocardial infarction as an example, which is one of the causes of AHF: the control group reduced ambulance response time by 15.3 minutes and door-to-balloon time by 36 minutes, decreased the length of hospital stay by 6.3 days, lowered the mortality rate by 12.33%, and reduced the rehospitalization rate by 19.69% [9]. Therefore, shortening the door-to-balloon time in the emergency department for patients with AHF, and even completing the basic disease diagnosis during the ambulance response, can significantly improve the survival rate of patients. However, the traditional process of AHF diagnosis needs to be optimized. While some rapid methods, such as auscultation, are part of clinical examinations, their effectiveness needs further quantitative evaluation.

Traditional AHF diagnosis relies heavily on clinical biochemical trials, and there is a lack of convenient and accurate diagnostic methods, especially for 20% of patients with no history of chronic heart failure (CHF). According to the European Society of Cardiology's AHF First Aid Guidelines, the diagnostic criteria for AHF include clinical evaluation, electrocardiogram (ECG), chest x-ray and imaging techniques, laboratory tests, and echocardiography [10]. The diagnosis of AHF is usually based on the history and physical examination as well as physical tests to assist in the examining of ECG, echocardiography, and biochemical tests of brain natriuretic peptide (BNP) and NT-proBNP. At present, according to ten mainstream NT-proBNP and BNP biochemical detection equipment on the market, the waiting time for BNP results is 9-16min, and the waiting time for NT-proBNP results is 11-21min [11]. The ejection fraction obtained by echocardiography is also an important indicator in the preoperative evaluation of AHF [12], which takes about 20-30 minutes for a single examination, including five minutes for equipment location.

*Corresponding author
zhaojing_zj@tju.edu.cn (J. Zhao); ws111@tju.edu.cn (S. Wang)
ORCID(s): 0000-0002-7728-7367 (Z. Qiu)

Auscultation, as part of the clinical evaluation, is an effective means of diagnosing heart failure and is recommended by the European Society of Cardiology as class-1 [10]. Heart sounds are produced by the mechanical motion of the heart's dynamic system. They are the sum of various mechanical vibrations caused by the blood movement within the cardiovascular system. Considering the mechanism by which heart sounds are produced, they contain a wealth of information regarding the physiology of the cardiovascular system and are regarded as a visual description of the contractile and stretch function of the heart [4, 13, 14]. Normal heart sounds consist of 4 tones, called first heart sound (S1), second heart sound (S2), third heart sound (S3), and fourth heart sound (S4) in the order during the cardiac cycle. Mitral valve, aortic valve, and pulmonary valve are the three most commonly used auscultation areas. Different from ECG, heart sounds are a manifest of the ventricle's ability to pump blood, which can reflect pathological information better than ECG in a single cardiac cycle. In general, abnormal heart sounds have more background noise, greater S1 amplitude fluctuations and more high-frequency details than normal heart sounds.

1.2. Related works

The current research on digital auscultation technology encompasses aspects such as datasets, signal processing, and diagnostic models.

In terms of datasets, the most commonly utilized dataset is the one established by PhysioNet for the Heart Sound Classification Challenge in 2016 [15], including 665 abnormal heart sounds and 2575 normal heart sounds. Yaseen et al. [16] provided a five-classification dataset of Aortic Stenosis (AS), Mitral Regurgitation (MR), Mitral Stenosis (MS), Mitral Valve Prolapse (MVP) and Normal (N), with 200 cases of each type. Nonetheless, for the development of algorithms aimed at diagnosing a particular medical condition, the current dataset may exhibit limitations in terms of annotation richness. Therefore, to develop rapid diagnostic models for AHF, additional efforts are needed to refine the data and standardize inclusion criteria.

In terms of pre-processing and feature extraction for heart sounds, in Vepa's research [17], heart sound signals were processed based on Short-Time Fourier Transform (STFT) and Discrete Wavelet Transform (DWT). Wu et al. [18] extracted the MFCC components of heart sound signals based on the hidden Markov model (HMM) model. Existing techniques, such as feature extraction based on MFCC, have been highly successful and widely utilized in the processing of heart sounds. However, current heart sound processing techniques are primarily designed for long-duration signals, with some even exceeding 60 seconds, rendering them unsuitable for AHF rapid diagnosis.

In terms of diagnostic models, Rubin [19] used a convolutional neural network (CNN) for heart sound signal classification by time-frequency characteristics. Arora et al. [20] performed transfer learning of heart sound signals based on VIZ, MobileNet, Xception, VGG, ResNet, DenseNet and Inception. Scholars such as Li and Shuvo [21, 22] have developed end-to-end lightweight neural networks for clinical mobile devices. The established models are too large in size for rapid diagnosis of AHF. Therefore, we have developed lightweight models and introduced different auscultation strategies for various clinical scenarios.

To address the challenges of signal processing and diagnostic model in AHF auscultation, our main contributions are:

- An auscultation dataset has been established, containing 2999 recordings from heart failure patients, aiming at addressing the challenges in heart failure auscultation. This dataset encompasses comprehensive information, including diagnostic results, collection area annotations, medical history records, and annotations related to BNP and NT-proBNP. Additionally, this dataset has undergone a rigorous ethical review process and is publicly accessible on https://github.com/jimmytju/AHF-Rapid-Diagnosis/Database.
- A wavelet denoising algorithm and a lightweight DenseHF-Net have been developed for short-duration auscultation signals, aiming at rapid diagnosis of acute heart failure. The denoising algorithm proposed in this paper has improved the average signal-to-noise ratio to 7.8 dB. DenseHF-Net has only 3.82M parameters and is easily ported to low-computation scenarios such as mobile terminals.
- We have introduced two auscultation strategies: multi-region fusion auscultation and mitral valve auscultation. Multi-region fusion auscultation is designed for scenarios where long-time auscultation is possible, such as monitoring situations. It utilizes the three most crucial regions, namely, the mitral valve region, aortic valve region, and pulmonic valve region. On the other hand, mitral valve auscultation is specifically designed for rapid diagnosis in cases of heart failure emergencies, focusing solely on the mitral valve auscultation region. It completes the diagnosis within 15 seconds with an accuracy of 94.41%, representing only a 4.94% reduction compared to the multi-region fusion strategy.

2. Materials and Methods

Fig. 1 illustrates the methodology employed in this paper. Firstly, in order to address the data scarcity issue in the development of AHF diagnostic models, we have created several datasets. We create a multi-region fusion auscultation dataset, containing 540 healthy cases and 389 heart failure cases. Each case includes three audio recordings corresponding to the mitral valve, aortic valve, and pulmonic valve. Additionally, a mitral valve auscultation dataset is established, containing 1620 healthy cases and 1379 heart failure cases, with each recording lasting between three to five seconds.

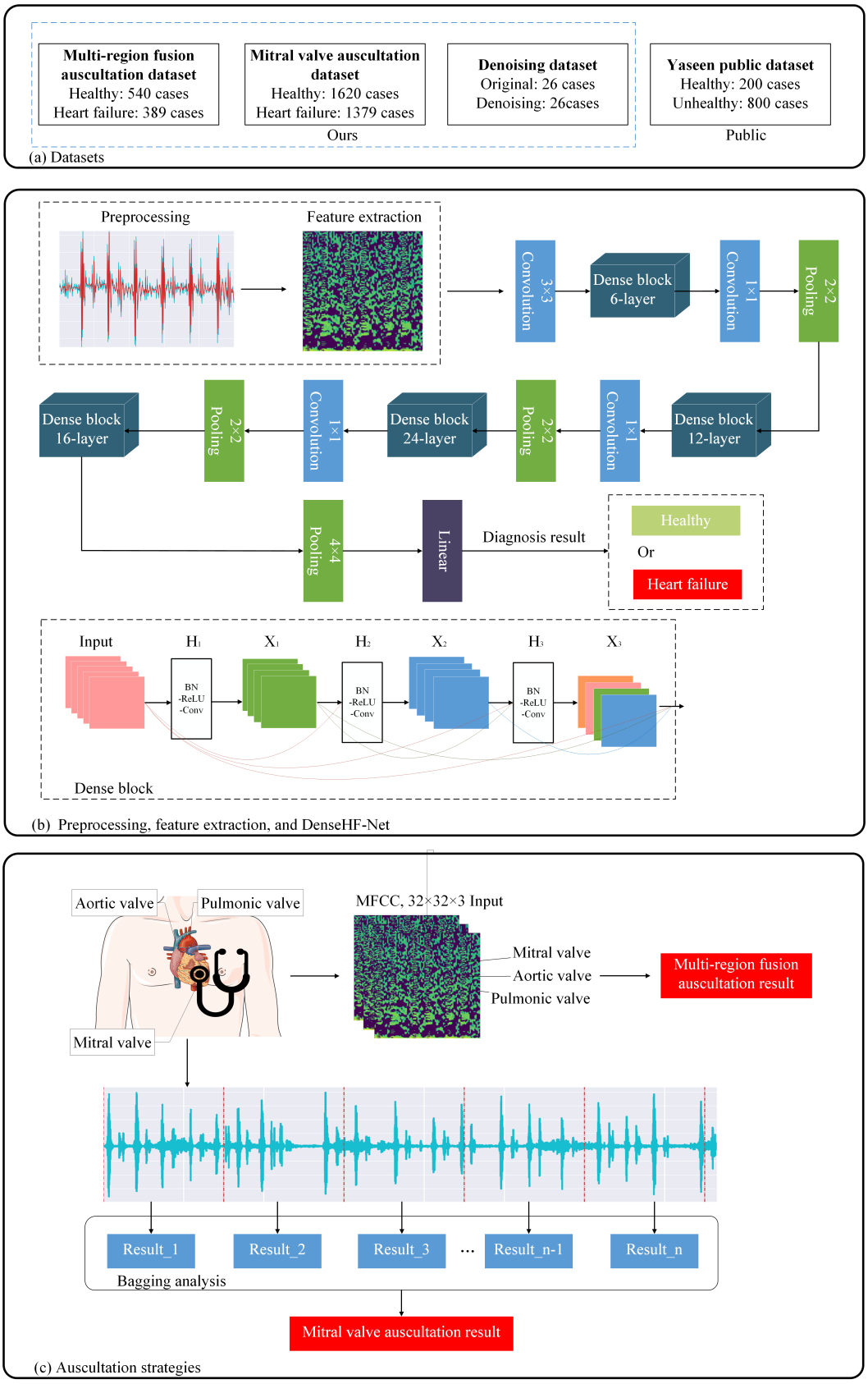


Figure 1: Methodology of the proposed work. (**a**) A dataset was built in this paper, including two HF diagnostic subsets, a denoising subset, and a public subset. (**b**) The layered architecture of the proposed DenseHF-Net contains four Dense blocks. (**c**) Multi-region fusion auscultation and mitral valve auscultation.

Next, in response to the signal processing and lightweight requirements for rapid AHF diagnosis, we have developed a wavelet denoising algorithm and a lightweight DenseHF-Net. We explore a wavelet denoising algorithm specifically designed for short-duration heart sound signals to reduce noise effectively. Subsequently, we employ the MFCC algorithm to extract one-dimensional sound signals into two-dimensional image features. Finally, a DenseHF-Net is used to train on the MFCC features.

Two distinct auscultation strategies are introduced in this paper. 1. Multi-region fusion auscultation is designed for scenarios where long-time auscultation is possible, using the fusion characteristics of the mitral valve, aortic valve, and pulmonic valve as input features. 2. Mitral valve auscultation is specifically designed for AHF rapid diagnosis in cases of ambulances. In the case of mitral valve auscultation lasting more than 10 seconds, an ensemble method is proposed.

2.1. Datasets

2.1.1. HF auscultation datasets

The heart sound databases were acquired at Tianjin 4th Center Hospital of China between 2021 and 2022. The data were recorded using a 3M™ Littmann® electronic stethoscope 3200, with a sampling rate set at 22kHz. This project was approved by the medical ethics committee of Tianjin 4th Center Hospital of China (No. 2022-T050). All volunteers have signed an informed consent.

Under the guidance of two chief physicians, we collected heart sounds from heart failure patients in various departments to create a pathological dataset. We recruited volunteers among patients diagnosed with heart failure, recorded auscultation at three different regions, and simultaneously documented their gender, age, hospitalization information, medical history, and the latest biochemical markers. Afterward, the chief physicians reviewed the data to exclude samples that had already recovered from heart failure or had unclear signs of heart failure features. Finally, heart sounds from a healthy population were collected in the same manner to serve as a comparison group. As shown in Tab. 2, the HF auscultation dataset consists of a total of 71.6 minutes of heart failure auscultation and 81 minutes of the comparison group auscultation.

The multi-region fusion auscultation dataset comprises 540 healthy cases and 389 cases of heart failure. Each case includes three audio recordings of the mitral valve, aortic valve, and pulmonic valve. To ensure robustness, we established a 10-fold cross-validation database after shuffling.

The mitral valve auscultation dataset consists of 1620 healthy cases and 1379 cases of heart failure, with each case featuring audio recordings lasting three to five seconds. Similar to the previous dataset, we created a 10-fold cross-validation database after shuffling.

2.1.2. Denoising dataset

As shown in Fig. 1a, we have created a comparative dataset before and after denoising, encompassing 26 different pathological descriptions. The initial 26 recordings are characterized as noisy signals, encompassing various common abnormal heart sounds. In contrast, the remaining 26 control recordings were subsequently reviewed and verified by two medical professionals to eliminate background noise while preserving all relevant pathological information.

2.1.3. Yaseen public dataset

As shown in Fig. 1a, we also use a publicly available Yaseen dataset [16]. This dataset includes Aortic Stenosis (AS), Mitral Regurgitation (MR), Mitral Stenosis (MS), Mitral Valve Prolapse (MVP), and Normal (N). The main purpose is to evaluate the model's generalization ability in diagnosing normal and abnormal heart sounds. We employ 80% training data and 20% testing data. Each case is recorded for a duration of 2 seconds.

2.2. Preprocessing

The preprocessing of heart sound signals aims to depress the noisy background in clinical environments. Wavelet transform is used to decrease the noise of heart sounds based on mother wavelets, such as Haar, db, Coif, Sym, and Biorthogonal (bior). Chen et al. [23] achieved optimal denoising results using the db6 wavelet basis. Zhao et al. [24] reported the best outcomes with the bior5.5 basis. Cheng et al. [25] utilized wavelet-based adaptive algorithms to enhance the denoising of heart signals, resulting in a signal improvement of 12.4 dB compared to the pre-denoising state.

In this paper, we consider three wavelet functions: db6, sym8, and coif5. These wavelet bases are used for the discrete wavelet decomposition of heart sound recordings. Following the decomposition, discrete wavelet reconstruction is performed, with a coefficient shrinkage function applied at a threshold of 20% modulo the maximum hard threshold.

Secondly, in order to determine the coefficient contraction strategy, various coefficient contraction functions are applied during both the discrete wavelet decomposition and reconstruction processes.

We introduce a novel self-adaptive threshold function, as depicted in Eq.(1).

$$f_{self}(x,T) = \begin{cases} e^{\frac{x+T}{2}} - e^{\frac{-x-T}{2}} & x \le -T \\ 0 & -T \le x \le T \\ e^{\frac{x-T}{2}} - e^{\frac{-x+T}{2}} & x \ge T \end{cases} \tag{1}$$

f_{self} has the following three advantages:

- f_{self} satisfies:

$$\lim_{x \to -T^-} f_{self}(x) = \lim_{x \to -T^+} f_{self}(x) = 0$$

$$\lim_{x \to T^-} f_{self}(x) = \lim_{x \to T^+} f_{self}(x) = 0$$

  f_{self} is differentiable at $x = \pm T$

- f_{self} is odd, with smooth and monotonically increasing curves.

- f_{self} can overcome the problem of discontinuities in the hard threshold function so that the reconstructed signal retains more detailed information after the reconstruction.

We collect statistical data on the average signal-to-noise ratio (SNR) under different coefficient contraction functions and thresholds.

2.3. Feature extraction

Feature extraction for heart sound signals aims to reduce the dimensionality of the data, highlight key information, and thereby improve the subsequent data processing and analysis. Mel Spectrum and MFCC have widely employed feature extraction methods in speech recognition. Human perception of frequency is non-linear, with greater sensitivity to low-frequency signals compared to high-frequency ones. Consequently, frequency conversion is performed according to the equation Eq.(2).

$$Mel(f) = 2595 \ln\left(1 + \frac{f}{700}\right) \tag{2}$$

To reproduce the Meier scale in the processing of discrete digital signals, the power spectral estimates of the resulting periodic plot are filtered using a Mehr filter bank (usually 26 V-Band Pass filter banks), as shown in Eq.(3).

$$H_m(k) = \begin{cases} \frac{k - f(m-1)}{f(m) - f(m-1)} & f(m-1) \le k \le f(m) \\ \frac{f(m+1) - k}{f(m+1) - f(m)} & f(m) \le k \le f(m+1) \\ 0 & others \end{cases} \tag{3}$$

Multiply with FFT to get the Mel spectrum:Eq. (4).

$$MelSpec(m) = \sum_{k=f(m-1)}^{f(m+1)} H_m(k) * |X(k)|^2 \tag{4}$$

Calculate the logarithmic energy output of the V-Band Pass filter bank: Eq.(5). Discrete cosine transform (DCT): Eq. (6) to obtain the MFCC coefficient.

$$S(m) = \ln\left(\sum_{k=0}^{N-1}) \left|X_a(k)\right|^2 H_m(k)\right), 0 \le m \le M \tag{5}$$

$$C(n) = \sum_{m=0}^{N-1} S(m) \cos\left(\frac{\pi n(m - 0.5)}{M}\right), n = 1, 2, ..., L \tag{6}$$

The L order refers to the order of the MFCC coefficient, usually 12-16. M is the number of triangular filters.

The MFCC feature design in this paper is defined as Wu et al. [18] and has achieved the same time-frequency extraction effect as Vepa [17], as shown in Fig. 1a.

2.4. DenseHF-Net

The common deep learning model architecture for heart sound diagnosis includes Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Long Short-Term Memory networks (LSTM), and Convolutional Neural Network-Long Short-Term Memory network hybrids (CNN-LSTM) [19–22]. This paper focuses on model design specifically tailored for AHF rapid diagnosis, aiming to develop a model that balances lightweight characteristics with accuracy.

We introduce DenseHF-Net, which is based on the CVPR 2017 Best Paper, Dense-Net [28]. To achieve model lightweight, DenseHF-Net employs only four Dense Blocks: DenseBlock-1, DenseBlock-2, DenseBlock-3, and DenseBlock-4. This choice reduces the model's parameter count and computational load while still maintaining a certain level of depth and feature extraction capability. Three transition layers with small compression rates are employed simultaneously to reduce the output channel numbers of the 1x1 convolutional layers, thereby reducing the size of feature maps. Ultimately, the diagnostic results are obtained after passing through a linear layer. The parameter settings for the three models are detailed in Tab. 1.

The classic ResNet [26] and MobileNet [27] models are also trained for comparison with DenseHF-Net. All input features are resized to 32×32 to make it more convenient for mobile terminals. The parameter numbers of the three models are 3.82M, 42.61M, and 12.24M. The Memory Access Cost (MAC) of the three models are 130.89M, 556.97M, and 46.28M.

Experimental environment system: Ubuntu 18.04; CPU: Intel(R) Core(TM) i7-9700K CPU @ 3.60GHz; GPU: NVIDIA GeForce RTX 3090 with 24G VRAM; CUDA Version: 11.4; Pytorch Version:1.10.

2.5. Auscultation strategies

The multi-region fusion auscultation strategy is designed for AHF diagnosis in hospitals, with the aim of reducing the door-to-balloon time. Then to reduce the mortality rate and rehospitalization rate of HF patients [9]. Heart sound signals are synchronously collected from three regions, including the mitral valve region, aortic valve region, and pulmonic valve region.

The processing pipeline includes three main steps:

1. Applying wavelet transform to the three channels of heart sound to reduce noise. 2. Feature extraction using the MFCC. 3. Fusion of the three MFCC feature sets to produce input of the same dimension as that of the mitral valve auscultation.

The mitral valve auscultation strategy is designed for emergency medical services (EMS) or general screening scenarios, with a stronger emphasis on convenience, aiming to complete the diagnosis within 15 seconds.

As shown in Fig. 1c, for mitral valve auscultation durations exceeding 10 seconds, there arises the need to address the challenge of reducing false positive diagnoses. To tackle

Table 1
Models parameter setting

DenseHF-Net (ours) **Number of parameters:** 3.82M **Memory access cost:** 130.89M **Forward/backward pass size:** 43.49M			**ResNet-18** [26] **Number of parameters:** 42.61M **Memory access cost:** 556.97M **Forward/backward pass size:** 13.63M			**MobileNetV1-28** [27] **Number of parameters:** 12.24M **Memory access cost:** 46.28M **Forward/backward pass size:** 10.31M		
Layer	Filters	Output	Layer	Filters	Output	Layer	Filters	Output
Feature Map	$\begin{bmatrix}3\times3, 24\end{bmatrix}\times1$	32×32	Feature Map	$\begin{bmatrix}3\times3, 64\end{bmatrix}\times1$	32×32	Feature Map	$\begin{bmatrix}3\times3, 32\end{bmatrix}\times1$	30×30
Dense Block	$\begin{bmatrix}1\times1, 48\\3\times3, 12\end{bmatrix}\times6$	32×32	Conv2_x	$\begin{bmatrix}3\times3, 64\\3\times3, 64\end{bmatrix}\times2$	32×32	Conv_dw,pw	$\begin{bmatrix}3\times3, 32\\1\times1, 64\end{bmatrix}\times1$	30×30
Transition	$\begin{bmatrix}1\times1, conv\\2\times2, pool\end{bmatrix}$	16×16	Conv3_x	$\begin{bmatrix}3\times3, 128\\3\times3, 128\end{bmatrix}\times2$	16×16	Conv_dw,pw	$\begin{bmatrix}3\times3, 64\\1\times1, 128\end{bmatrix}\times1$	15×15
Dense Block	$\begin{bmatrix}1\times1, 48\\3\times3, 12\end{bmatrix}\times12$	16×16	Conv4_x	$\begin{bmatrix}3\times3, 256\\3\times3, 256\end{bmatrix}\times2$	8×8	Conv_dw,pw	$\begin{bmatrix}3\times3, 128\\1\times1, 128\end{bmatrix}\times1$	15×15
Transition	$\begin{bmatrix}1\times1, conv\\2\times2, pool\end{bmatrix}$	8×8	Conv5_x	$\begin{bmatrix}3\times3, 512\\3\times3, 512\end{bmatrix}\times2$	4×4	Conv_dw,pw	$\begin{bmatrix}3\times3, 128\\1\times1, 256\end{bmatrix}\times1$	8×8
Dense Block	$\begin{bmatrix}1\times1, 48\\3\times3, 12\end{bmatrix}\times24$	8×8	Pooling	4×4	$1\times1\times512$	Conv_dw,pw	$\begin{bmatrix}3\times3, 256\\1\times1, 256\end{bmatrix}\times1$	8×8
Transition	$\begin{bmatrix}1\times1, conv\\2\times2, pool\end{bmatrix}$	4×4	Linear		1×2	Conv_dw,pw	$\begin{bmatrix}3\times3, 256\\1\times1, 512\end{bmatrix}\times1$	4×4
Dense Block	$\begin{bmatrix}1\times1, 48\\3\times3, 12\end{bmatrix}\times16$	4×4				Conv_dw,pw	$\begin{bmatrix}3\times3, 512\\1\times1, 512\end{bmatrix}\times5$	4×4
Pooling	4×4	$1\times1\times384$				Conv_dw,pw	$\begin{bmatrix}3\times3, 512\\1\times1, 1024\end{bmatrix}\times1$	2×2
Linear		1×2				Conv_dw,pw	$\begin{bmatrix}3\times3, 1024\\1\times1, 1024\end{bmatrix}\times1$	2×2
						Pooling	2×2	$1\times1\times1024$
						Linear		1×2

this issue, this research paper introduces an ensemble learning method as a strategic approach.

$$\begin{aligned} Result\_i &= \max_{\alpha_i} Softmax(\alpha_i) \\ &= \max_{\alpha_i} \frac{\exp(\alpha_i)}{\sum_{i=1}^{2}\exp(\alpha_i)} \end{aligned} \tag{7}$$

Eq.7 is used to calculate the diagnostic results for a single fragment. α_i represents the model output. $Result_i$ is 0 for healthy and 1 for heart failure.

$$Output = Result\_1 \vee Result\_2 \vee \cdots \vee Result\_n \tag{8}$$

Eq.8 represents the result of a one-to-one OR operation applied to auscultation segments, designed to minimize the occurrence of false positives. Here, 'n' denotes the number of fragments, typically set to 3.

3. Results

3.1. Datasets

Tab. 2 shows the details of the dataset in this paper. The details include the age and gender composition of auscultation volunteers, and case descriptions for the denoising dataset. Auscultation dataset contains 2999 recordings from heart failure patients, each with rich annotations, and are publicly accessible on https://github.com/jimmytju/AHF-Rapid-Diagnosis/Database.

3.2. Preprocessing

Firstly, we conduct an analysis to determine the average SNR under various combinations of wavelet bases and decomposition layers, using a coefficient shrinkage function based on the 20% modulo maximum hard threshold. The results of this analysis are presented in Fig. 2, which clearly indicates that the Sym8 base at the 7-layer decomposition level yields the most effective denoising results among the tested methods.

Secondly, we further investigate the average SNR across different shrinkage functions and threshold values while maintaining the sym8 base at the 7-layer decomposition level. Our findings indicate that the use of the 20% modulo maximum with the f_{self} threshold function emerged as the optimal denoising method.

Finally, the average SNR of the denoising algorithm proposed in this paper is 7.8 dB.

3.3. Multi-region fusion auscultation

The quality of models refers to Eq.9: Accuracy (Acc), Eq.10: Sensitivity (Se), Eq.11: Specificity (Sp) and Eq.12: F1-Score.

$$Acc = \frac{TP + TN}{TP + FP + TN + FN} \tag{9}$$

$$Se = \frac{TP}{TP + FN} \tag{10}$$

Table 2
The details of the dataset information.

HF datasets			
Group	Age($\bar{x} \pm sd$)	Sex	Length(min)
Healthy	24	Male	27.0
Healthy	26 ± 2	Female	54.0
HF	72.6 ± 12.0	Male	34.8
HF	77.9 ± 11.1	Female	36.8

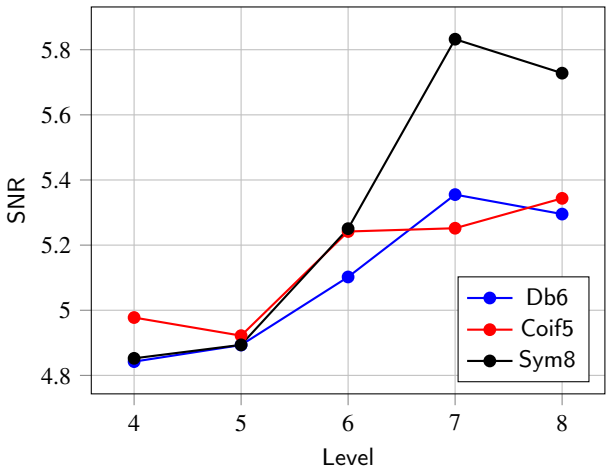
Denoising dataset	
Description	Length(s)
Systolic murmur	68
Functional aortic stenosis	4
Musical noise	23
Organic mitral regurgitation	59
Functional mitral stenosis	9
Organic mitral stenosis	9
Diplogue	39
Paradoxically divided	13
Varied S1	23
Weak S1	3
Split S1	4
Strong S1	6
Weak S2	29
Split S2	12
Ventricular septal defect	9
Open flap sound	10
Late contraction	9
Relative mitral regurgitation	4
Sinus bradycardia	7
Sinus tachycardia	4
Functional pulmonary regurgitation	3
Pulmonary stenosis	67
Diastolic tetratone	15
Continuous murmur	15
Overlapping galloping sound	9
Pendulum sound	6

Yaseen dataset	
Description	Number
Aortic Stenosis	200
Mitral Regurgitation	200
Mitral Stenosis	200
Mitral Valve Prolapse	200
Normal	200

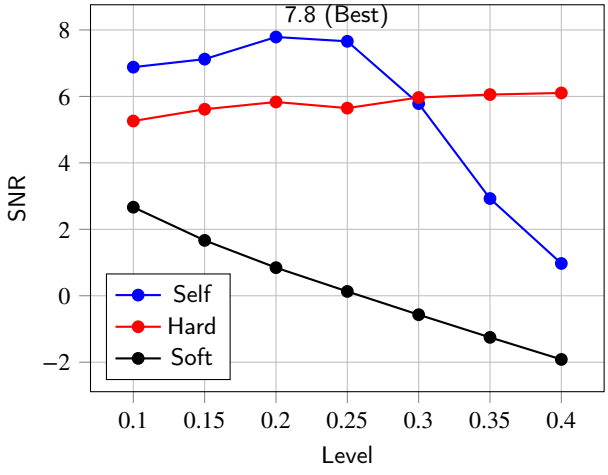
$$Sp = \frac{TN}{TN + FP} \tag{11}$$

$$F1 - Score = 2 \times \frac{Se \times Sp}{Se + Sp} \tag{12}$$

For multi-region fusion auscultation, the average accuracies are 99.35%, 98.71%, and 99.14%, respectively. In terms of average sensitivity, these models achieve 99.08%, 99.09%, and 100.00%, respectively. Furthermore, the average specificity for these models is measured at 99.75%, 98.10%, and 97.95%, respectively. The average F1-Scores are 99.42%, 98.59%, and 98.97%, respectively.



(a) Choice of base and levels



(b) Choice of threshold parameters

Figure 2: Wavelet denoising of short signal. (**a**) Best: wavelet denoising with sym8 base at 7-layer decomposition. (**b**) Best: wavelet denoising with 20% modulo maximum f_{self}.

Tab.3 also presents the results obtained from the Yaseen dataset. DenseHF-Net, ResNet-18, and MobileNetV1-28 typically reached convergence around 50 epochs. The AS diagnosis exhibits the best performance, with sensitivity and specificity exceeding 82% across all three models. For MS and MR, all three models achieve correct diagnoses, with sensitivity and specificity exceeding 87% in both ResNet-18 and MobileNetV1-28. However, MVP diagnosis by DenseHF-Net yields suboptimal results, with a specificity of only 30% and an F1-Score of only 45.88%.

3.4. Mitral valve auscultation

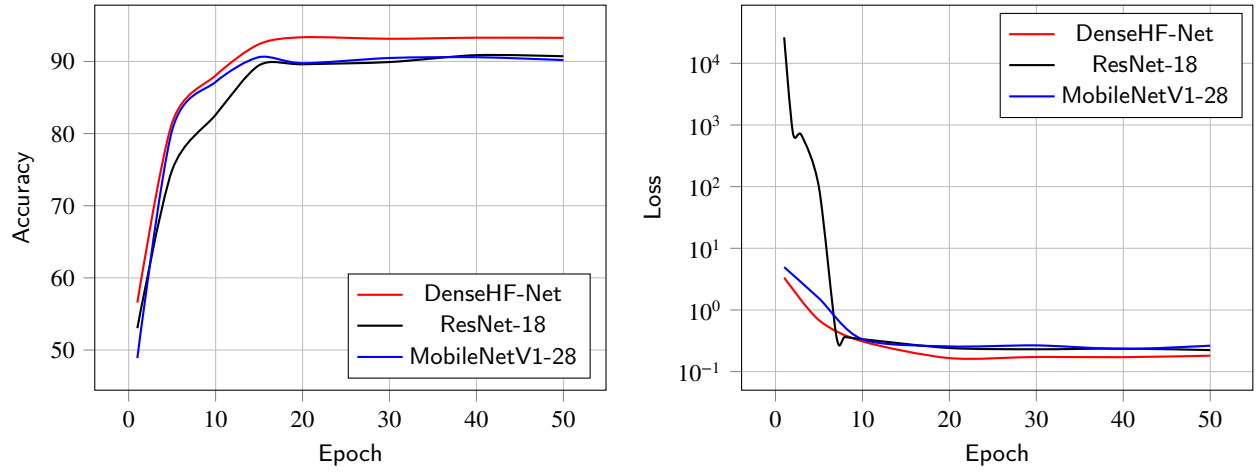
Fig.3 provides a comprehensive overview of the average performance across the mitral valve auscultation dataset. Notably, ResNet-18 and MobileNetV1-28 exhibit comparable performance, while DenseHF-Net exhibits the most rapid rate of improvement. Importantly, all three models

Table 3
Classification results of HF-Diagnosis dataset and public Yassen Dataset.

Multi-region fusion auscultation results					
DenseHF-Net(ours)	$Average \pm sd$	ResNet-18	$Average \pm sd$	MobileNetV1-28	$Average \pm sd$
$Acc(\%)$	99.35 ± 0.71	$Acc(\%)$	98.71 ± 1.26	$Acc(\%)$	99.14 ± 0.65
$Se(\%)$	99.08 ± 1.22	$Se(\%)$	99.09 ± 0.92	$Se(\%)$	100.00 ± 0.00
$Sp(\%)$	99.75 ± 0.75	$Sp(\%)$	98.10 ± 2.68	$Sp(\%)$	97.95 ± 1.54
$F1-Score(\%)$	99.42 ± 0.64	$F1-Score(\%)$	98.59 ± 1.49	$F1-Score(\%)$	98.97 ± 0.78
Mitral valve auscultation results					
DenseHF-Net(ours)	$Average \pm sd$	ResNet-18	$Average \pm sd$	MobileNetV1-28	$Average \pm sd$
$Acc(\%)$	94.41 ± 1.78	$Acc(\%)$	91.33 ± 2.48	$Acc(\%)$	90.73 ± 1.80
$Se(\%)$	96.11 ± 1.88	$Se(\%)$	92.32 ± 2.59	$Se(\%)$	90.68 ± 2.49
$Sp(\%)$	92.00 ± 4.08	$Sp(\%)$	90.20 ± 4.49	$Sp(\%)$	90.80 ± 2.97
$F1-Score(\%)$	93.95 ± 2.15	$F1-Score(\%)$	91.17 ± 2.58	$F1-Score(\%)$	90.69 ± 1.75

Yaseen dataset results

	DenseHF-Net(ours)				ResNet-18				MobileNetV1-28			
	AS	MR	MS	MVP	AS	MR	MS	MVP	AS	MR	MS	MVP
$Acc(\%)$	91.25	58.75	75.00	63.75	92.50	88.75	95.00	83.75	82.50	92.50	85.00	88.75
$Se(\%)$	90.00	97.50	75.00	97.50	95.00	80.00	90.00	80.00	87.50	87.50	87.50	87.50
$Sp(\%)$	92.50	20.00	75.00	30.00	90.00	97.50	100.00	87.50	77.50	97.50	82.50	90.00
$F1-Score(\%)$	91.23	33.19	75.00	45.88	92.43	87.89	94.74	83.58	82.20	92.23	84.93	88.73



(a) Mitral valve auscultation testing accuracy.

(b) Mitral valve auscultation testing loss value.

Figure 3: Average 10-fold CV history. (**a**) Average testing accuracy of three models. (**b**) Average testing loss of three models (without augmentation).

exhibit effective convergence of the loss function. It is worth highlighting that both ResNet-18 and MobileNetV1-28 demonstrate similar performance trends, with DenseHF-Net demonstrating the fastest convergence among them.

Tab.3 lists the 10-fold results of HF diagnosis. DenseHF-Net, ResNet-18, and MobileNetV1-28 basically converge around 50 epochs.

For mitral valve auscultation, the average accuracies achieved by the individual models are 93.63%, 91.33%, and 90.73%, respectively. In terms of average sensitivity, these models reach 94.28%, 92.32%, and 90.68%, respectively. Furthermore, the average specificity for these models is measured at 92.95%, 90.20%, and 90.80%, respectively. The average F1-Scores for these models are 93.57%, 91.17%, and 90.69%, respectively.

When combined with DenseHF-Net and the ensemble method, the overall performance improves significantly. The resulting average accuracy, sensitivity, specificity, and F1-score are enhanced to 94.41%, 96.11%, 92.00%, and 93.95%, respectively.

3.5. Ablation Study

3.5.1. Effects of the Model Architecture

aaaaa

3.5.2. Effects of Preprocessing and Feature Extraction

aaa

4. Discussion

4.1. How auscultation works?

Heart sounds carry essential physiological information related to the heart's ability to pump blood. This paper analyzes the time-frequency characteristics of heart sounds by utilising short-term Fourier transforms. The primary auscultation area of interest is the mitral valve, situated at the point of the strongest apical beat.

In Fig.4a, the stable S1 amplitude is primarily concentrated within the 100 Hz frequency range. The aortic valve, located in the second intercostal space of the right sternal border, is further from the heart compared to the mitral valve. It is significantly influenced by lung sounds and exhibits lower amplitude. On the other hand, the pulmonic valve, situated in the second intercostal space of the left sternal border, demonstrates signal characteristics falling between those of the mitral and aortic valves.

This study involved the collection of mitral valve heart sounds from the same patient before and after receiving initial medical intervention, as depicted in Fig. 4d and Fig.4e. In comparison to healthy subjects, the amplitudes of S1 are found to be unstable, accompanied by noisy lung sounds. However, following the initial treatment, which took place two days later, there was a noticeable improvement. The heart sound cycle becomes clearer, and the energy amplitude of S1 is more pronounced.

4.2. How long should auscultate for?

The optimal input length for heart sounds in AI models is a subject of debate. Cardiac auscultation offers significant predictive value for cardiac diagnosis, characterized by its rapid and cost-effective nature [29]. In the PhysioNet 2016 dataset, which includes 665 abnormal heart sounds and 2575 normal heart sounds, the lengths of recordings range from 3 seconds to 60 seconds. Meanwhile, the Yaseen dataset consists of 1000 abnormal heart sounds and 200 normal heart sounds, all standardized at a length of 2 seconds.

As illustrated in Fig.5, the choice of input length for heart sounds directly impacts the number of cardiac events included, such as S1, diastole, S2, and systole. Fig.5a demonstrates that a 1.5-second input can theoretically encompass two diastolic and two systolic events. Fig.5b shows that a 2-second input can include at least two complete cardiac cycles. Fig.5c reveals that a 3-second input can accommodate three or four full cardiac cycles.

For the purposes of this paper, the input length is standardized at 3 seconds.

4.3. Results discussion

4.3.1. Preprocessing and feature extraction

This paper conducts an in-depth investigation into the optimal decomposition layers, wavelet bases, and threshold shrinkage functions for the denoising of short heart sounds. Fig.2 presents our findings, indicating that a 7-layer decomposition using the f_{self} threshold based on the sym8 wavelet yields the most effective denoising results. It's worth noting that Db6 and sym8 wavelets exhibit similar properties, but sym8, due to its shorter support length and superior energy concentration, is more morphologically aligned with heart sounds.

In contrast to previous studies by Chen [23] and Zhao [24], this paper employs Signal-to-Noise Ratio (SNR) as the primary metric for evaluating noise reduction, avoiding the limitations of waveform comparison. Additionally, our denoising experiments encompass 26 pathological heart sounds, enhancing the clinical relevance of our findings. Furthermore, in comparison to Cheng [25], we propose a new f_{self} construction method that significantly improves SNR (7.8 dB).

As depicted in Fig.6, this study delves into the intricacies of feature extraction, particularly focusing on Mel-spectrum and MFCC. The MFCC feature extraction method for heart sounds possesses the advantage of preserving not only temporal waveform features but also capturing frequency energy distribution. This approach retains valuable information aligned with the Mel scale, which corresponds to human auditory perception.

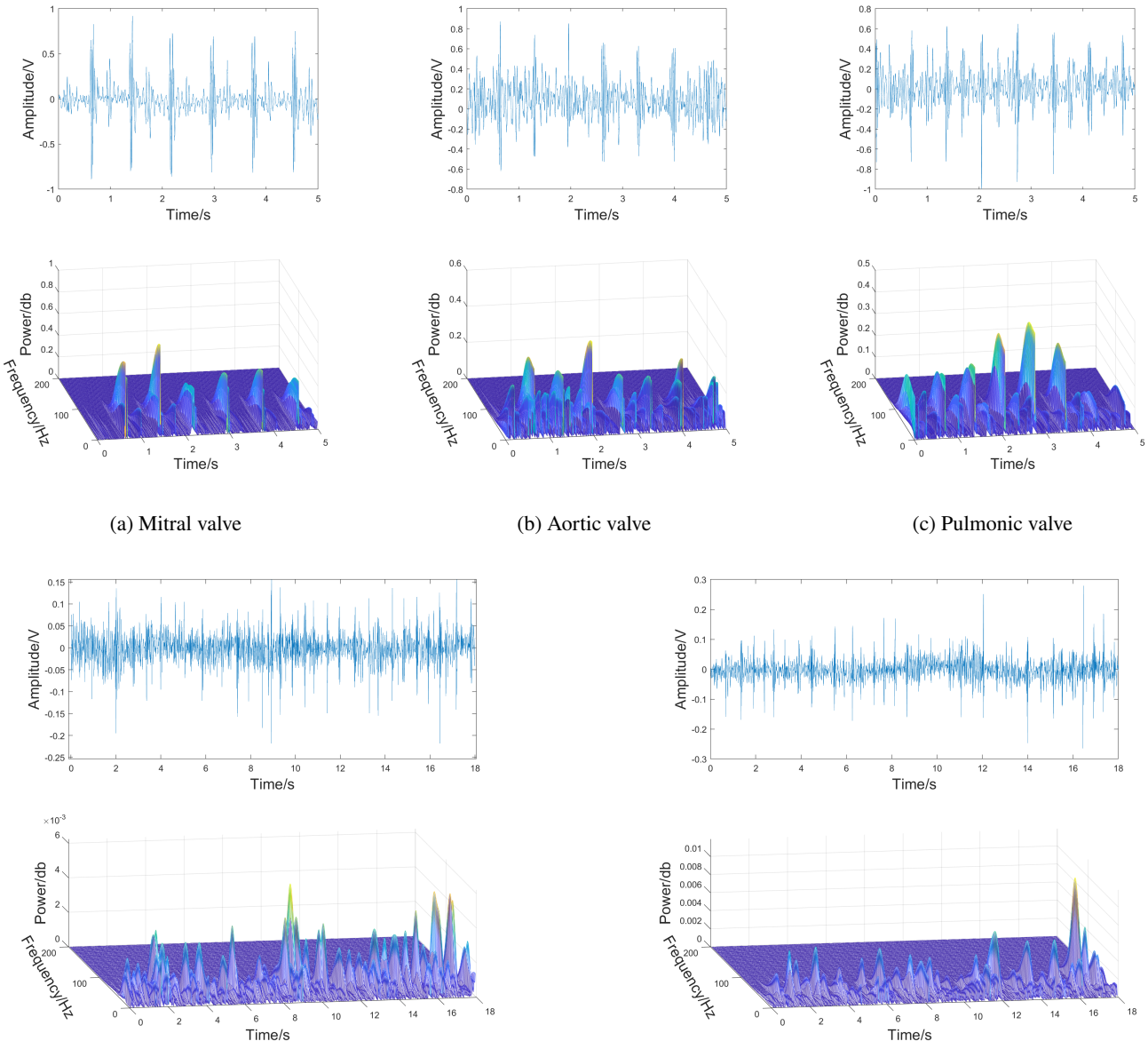
Fig.6a illustrates a representation of normal heart sounds, while Fig.6b showcases an instance of heart sounds from a patient with Mitral Valve Prolapse. Both Mel spectrum and MFCC representations encompass a comprehensive range of waveform characteristics, effectively capturing the intricacies of high-frequency signal features in the form of heat maps.

4.3.2. Diagnostic performance

Tab. 3 presents the diagnostic outcomes for two heart failure diagnostic datasets and the publicly available Yaseen dataset.

For multi-region fusion auscultation, DenseHF-Net, ResNet-18, and MobileNetV1-28 undergo 10-fold cross-training to calculate average accuracy, sensitivity, specificity, and F1-Score. Among these three models, DenseHF-Net achieves the highest performance, followed by MobileNetV1-28 and ResNet-18. DenseHF-Net attains an accuracy of 99.35%, 0.64% higher than ResNet-18 and 0.21% higher than MobileNetV1-28. In terms of sensitivity, DenseHF-Net records 99.08%, which is 0.01% lower than ResNet-18 and 0.92% lower than MobileNetV1-28. Regarding specificity, DenseHF-Net demonstrates 99.75%, 1.65% lower than ResNet-18 and 1.80% lower than MobileNetV1-28. The F1-Score for DenseHF-Net stands at 99.42%, 0.83% higher than ResNet-18 and 0.45% higher than MobileNetV1-28.

For mitral valve auscultation, DenseHF-Net exhibits an accuracy of 94.41%, which is 3.08% higher than ResNet-18 and 3.68% higher than MobileNetV1-28. The sensitivity of DenseHF-Net is 96.11%, 3.79% higher than ResNet-18, and 5.43% higher than MobileNetV1-28. The specificity for DenseHF-Net reaches 92.00%, 1.80% higher than ResNet-18, and 1.20% higher than MobileNetV1-28. The F1-Score for DenseHF-Net stands at 93.95%, 2.78% higher than ResNet-18 and 3.26% higher than MobileNetV1-28.



(a) Mitral valve (b) Aortic valve (c) Pulmonic valve

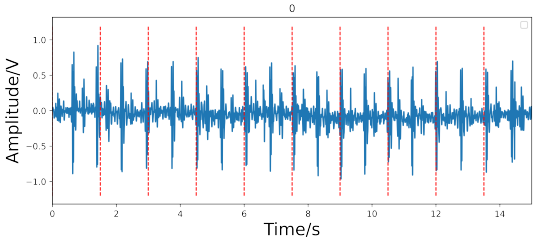
(d) Before first-aid (e) After first-aid

Figure 4: Heart sounds in five different situations. (**a**) Heart sounds from the mitral valve. (**b**) Heart sounds from the aortic valve. (**c**) Heart sounds from the pulmonic valve. (**d**) Heart sounds before first-aid. (**e**) Heart sounds after first-aid.

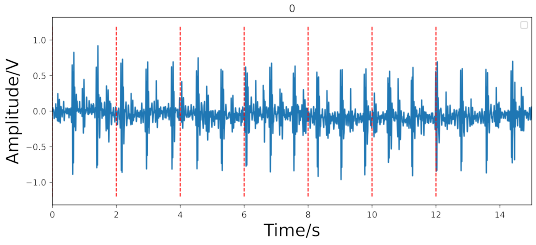
Regarding the public Yaseen dataset [16], the models exhibit robust generalization abilities for short heart sounds. For the diagnosis of aortic stenosis, mitral regurgitation, mitral stenosis, and mitral valve prolapse, both ResNet-18 and MobileNetV1-28 perform well, achieving accuracies exceeding 82.50%. However, DenseHF-Net does not perform as well in the diagnosis of mitral regurgitation, mitral stenosis, and mitral valve prolapse, with an accuracy of 63.75% and an F1-Score of 45.88%.

Previous researchers have also explored various intelligent algorithms for heart failure diagnosis or prediction, as shown in Tab. 4. Khade [30] and Rao [31] predicted heart failure incidence rates based on physiological parameters or medication records. Acharya [32] and Matsumoto [33] examined the potential of diagnosing heart failure using ECG or X-ray images. In comparison to Khade [30] and Rao [31], this paper offers enhanced medical diagnostic value and interpretability, along with Acharya [32] and Matsumoto [33]. When compared to Acharya [32] and Matsumoto [33], this study theoretically boasts superior diagnostic speed, reduced device size, and more suitable application in ambulance settings.

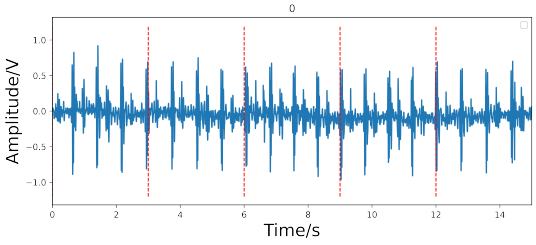
Compared with previous classification research efforts [17–22], this study incorporates several well-established and successful methods, including wavelet denoising, MFCC feature extraction, and the use of CNNs. We address the issue



(a) Cut with a length of 1.5s



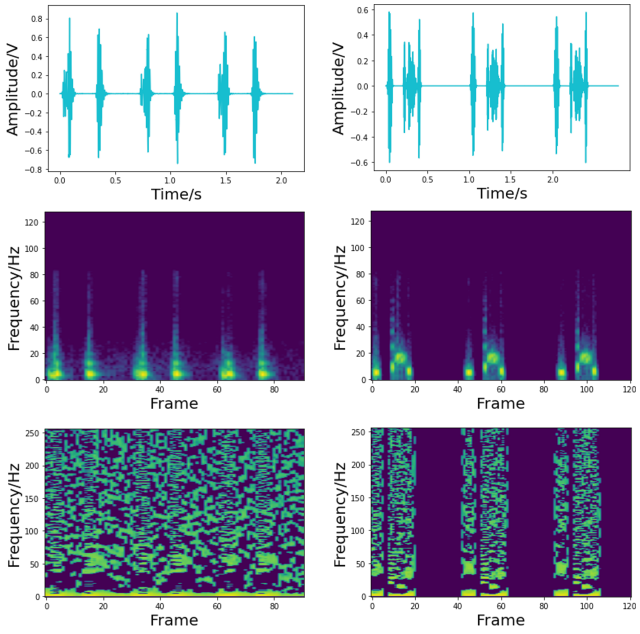
(b) Cut with a length of 2s



(c) Cut with a length of 3s

Figure 5: Effect of different cutting lengths. (**a**) A heart sound is cut with lengths of 1.5s. (**b**) The same heart sound is cut with lengths of 2s. (**c**) The same heart sound is cut with lengths of 3s.

of signal processing and diagnosis in rapid short-duration auscultation. However, it lacks strong generalization capabilities in diagnosing mitral regurgitation, indicating that the model still suffers from overfitting problems. Another limitation is the insufficient discussion of subtypes of AHF.



(a) Normal heart sound (b) Unhealthy heart sound

Figure 6: Feature extraction result. (**a**) Waveform, Mel-spectrum, MFCC of normal heart sound. (**b**) Waveform, Mel-spectrum, MFCC of unhealthy MVP heart sound.

Table 4
Comparison with other researches

Comparison with other heart failure researches		
Authors	Database	Aim
Ours	HF-Diagnosis dataset	HF diagnosis
Khade [30]	Physiological parameters	CHF prediction
Rao [31]	Medication records	CHF prediction
Acharya [32]	ECG	HF diagnosis
Matsumoto [33]	X-Ray Images	HF diagnosis

Comparison with other heart sound researches				
Authors	Feature extraction	Classification method	Databases	Acc(%)
Ours	MFCC	DenseHF-Net	Mitral valve auscultation dataset	94.41
Vepa [17]	STFT, DWT	kNN, MLP, SVM	Normal, systolic, diastolic records	95.2
Wu [18]	MFCC	HMM	325 cycles of 10-type	95.08
Rubin [19]	MFCC	CNN	PhysioNet 2016	95.2
Arora [20]	STFT	CNN	PhysioNet 2016	89.04
Li [21]	STFT	CNN	PhysioNet 2016	85
Shuvo [22]	Time-invariant features	CNN	Yaseen database	99.6

5. Conclusion

Cardiac auscultation is a crucial clinical method for detecting heart failure and holds significant promise for rapid heart failure detection. This paper discusses signal processing and diagnostic model techniques for the rapid diagnosis of AHF. The average SNR of heart sounds reached 7.8 dB. The models successfully achieve effective feature learning and effective loss convergence. DenseHF-Net stands out as the top-performing model in heart failure diagnosis, achieving an accuracy of 94.41%. Both two auscultation strategies can meet the requirements for emergency vehicles and hospital rooms. Future research will focus on collecting multi-center AHF data for model training, aiming to achieve better generalization performance. Simultaneously, we will actively promote clinical testing and assessment to validate its practicality and explore its clinical emergency application value.

6. Data and code availability

Data collection has been approved by the medical ethics committee of Tianjin 4th Center Hospital of China (No. 2022-T050). The data used in this study is available at https://github.com/qiuzhaoyu/AHF-Rapid-Diagnosis/Database. The code used in this study is available at https://github.com/qiuzhaoyu/AHF-Rapid-Diagnosis.

7. Funding

This research was supported by The National Natural Science Foundation of China (No. 72174138). 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] S. S. Virani, A. Alonso, E. J. Benjamin, M. S. Bittencourt, C. W. Callaway, A. P. Carson, A. M. Chamberlain, A. R. Chang, S. Cheng, F. N. Delling, et al., Heart disease and stroke statistics—2020 update: a report from the american heart association, Circulation 141 (2020) e139–e596.

[2] G. A. Roth, G. A. Mensah, C. O. Johnson, G. Addolorato, E. Ammirati, L. M. Baddour, N. C. Barengo, A. Z. Beaton, E. J. Benjamin, C. P. Benziger, et al., Global burden of cardiovascular diseases and risk factors, 1990–2019: update from the gbd 2019 study, Journal of the American College of Cardiology 76 (2020) 2982–3021.

[3] G. A. Mensah, G. A. Roth, V. Fuster, The global burden of cardiovascular diseases and risk factors: 2020 and beyond, 2019.

[4] E. M. Boorsma, J. M. Ter Maaten, K. Damman, W. Dinh, F. Gustafsson, S. Goldsmith, D. Burkhoff, F. Zannad, J. E. Udelson, A. A. Voors, Congestion in heart failure: a contemporary look at physiology, diagnosis and treatment, Nature Reviews Cardiology 17 (2020) 641–655.

[5] L. Sinnenberg, M. M. Givertz, Acute heart failure, Trends in Cardiovascular Medicine 30 (2020) 104–112.

[6] M. Arrigo, M. Jessup, W. Mullens, N. Reza, A. M. Shah, K. Sliwa, A. Mebazaa, Acute heart failure, Nature Reviews Disease Primers 6 (2020) 1–15.

[7] B. Chapman, A. D. DeVore, R. J. Mentz, M. Metra, Clinical profiles in acute heart failure: an urgent need for a new approach, ESC heart failure 6 (2019) 464–474.

[8] S. M. Victor, A. Gnanaraj, S. Vijayakumar, S. Pattabiram, A. S. Mullasari, Door-to-balloon: where do we lose time? single centre experience in india, indian heart journal 64 (2012) 582–587.

[9] Z. Fan, F. Zhang, Effects of an emergency nursing pathway on the complications and clinical prognosis of patients with acute myocardial infarction, Int J Clin Exp Med 14 (2021) 661–668.

[10] M. Nieminen, M. Böhm, M. Cowie, H. Drexler, G. Filippatos, G. Jondeau, Y. Hasin, et al., Task force on acute heart failure. executive summary of the guidelines on the diagnosis and treatment of acute heart failure: the task force on acute heart failure of the european society of cardiology, Eur Heart J 26 (2005) 384–416.

[11] R. A. Lewis, C. Durrington, R. Condliffe, D. G. Kiely, Bnp/nt-probnp in pulmonary arterial hypertension: time for point-of-care testing?, European Respiratory Review 29 (2020).

[12] P. Menon, P. M. Kapoor, M. Choudhury, Echocardiography for left ventricular assist device patients, Journal of Cardiac Critical Care TSS 6 (2022) 155–161.

[13] M. Johnston, S. P. Collins, A. B. Storrow, The third heart sound for diagnosis of acute heart failure, Current Heart Failure Reports 4 (2007) 164–169.

[14] J. Wynne, The clinical meaning of the third heart sound, The American Journal of Medicine 111 (2001) 157–158.

[15] G. D. Clifford, C. Liu, B. Moody, D. Springer, I. Silva, Q. Li, R. G. Mark, Classification of normal/abnormal heart sound recordings: The physionet/computing in cardiology challenge 2016, in: 2016 Computing in cardiology conference (CinC), IEEE, 2016, pp. 609–612.

[16] G.-Y. Son, S. Kwon, Classification of heart sound signal using multiple features, Applied Sciences 8 (2018) 2344.

[17] J. Vepa, Classification of heart murmurs using cepstral features and support vector machines, in: 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, 2009, pp. 2539–2542.

[18] H. Wu, S. Kim, K. Bae, Hidden markov model with heart sound signals for identification of heart diseases, in: Proceedings of 20th International Congress on Acoustics (ICA), Sydney, Australia, 2010, pp. 23–27.

[19] J. Rubin, R. Abreu, A. Ganguli, S. Nelaturi, I. Matei, K. Sricharan, Classifying heart sound recordings using deep convolutional neural networks and mel-frequency cepstral coefficients, in: 2016 Computing in cardiology conference (CinC), IEEE, 2016, pp. 813–816.

[20] V. Arora, K. Verma, R. S. Leekha, K. Lee, C. Choi, T. Gupta, K. Bhatia, Transfer learning model to indicate heart health status using phonocardiogram (2021).

[21] T. Li, Y. Yin, K. Ma, S. Zhang, M. Liu, Lightweight end-to-end neural network model for automatic heart sound classification, Information 12 (2021) 54.

[22] S. B. Shuvo, S. N. Ali, S. I. Swapnil, M. S. Al-Rakhami, A. Gumaei, Cardioxnet: A novel lightweight deep learning framework for cardiovascular disease classification using heart sound recordings, IEEE Access 9 (2021) 36955–36967.

[23] J. Y. Zhao, H. Y. Liu, H. S. Ma, H. D. Zhou, Research of the approach for the fetal heart sound signal's extracting based on coif5 wavelet transform, Chinese Journal of Biomedical Engineering (2006).

[24] T. Chen, L. Han, S. Xing, Research of de-noising method of heart sound signals based on wavelet transform, Computer Simulation 27 (2010) 401–405.

[25] X. Cheng, Z. Zhang, Denoising method of heart sound signals based on self-construct heart sound wavelet, Aip Advances 4 (2014) 087108.

[26] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.

[27] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, H. Adam, Mobilenets: Efficient convolutional neural networks for mobile vision applications, arXiv preprint arXiv:1704.04861 (2017).

[28] G. Huang, Z. Liu, L. Van Der Maaten, K. Q. Weinberger, Densely connected convolutional networks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 4700–4708.

[29] A. J. Taylor, Learning Cardiac Auscultation: From Essentials to Expert Clinical Interpretation, Springer, 2015.

[30] S. Khade, A. Subhedar, K. Choudhary, T. Deshpande, U. Kulkarni, A system to detect heart failure using deep learning techniques, Int. Res. J. Eng. Technol.(IRJET) 6 (2019) 384–387.

[31] S. Rao, Y. Li, R. Ramakrishnan, A. Hassaine, D. Canoy, J. Cleland, T. Lukasiewicz, G. Salimi-Khorshidi, K. Rahimi, An explainable transformer-based deep learning model for the prediction of incident heart failure, IEEE Journal of Biomedical and Health Informatics 26 (2022) 3362–3372.

[32] U. R. Acharya, H. Fujita, S. L. Oh, Y. Hagiwara, J. H. Tan, M. Adam, R. S. Tan, Deep convolutional neural network for the automated diagnosis of congestive heart failure using ecg signals, Applied Intelligence 49 (2019) 16–27.

[33] T. Matsumoto, S. Kodera, H. Shinohara, H. Ieki, T. Yamaguchi, Y. Higashikuni, A. Kiyosue, K. Ito, J. Ando, E. Takimoto, et al., Diagnosing heart failure from chest x-ray images using deep learning, International Heart Journal 61 (2020) 781–786.