



Research article

A multimodal parallel method for left ventricular dysfunction identification based on phonocardiogram and electrocardiogram signals synchronous analysis

Yajing Zeng¹, Siyu Yang², Xionikai Yu¹, Wenting Lin¹, Wei Wang¹, Jijun Tong^{2,*} and Shudong Xia^{1,*}

1 The Fourth Affiliated Hospital Zhejiang University School of Medicine, Jinhua 321000, China

2 School of Information and Technology, Zhejiang Sci-Tech University, Hangzhou 310000, China

* **Correspondence:** Email: jijuntong@zstu.edu.cn, shystone@zju.edu.cn; Tel: +8613116789476, +8618867961072.

Abstract: Heart failure (HF) is widely acknowledged as the terminal stage of cardiac disease and represents a global clinical and public health problem. Left ventricular ejection fraction (LVEF) measured by echocardiography is an important indicator of HF diagnosis and treatment. Early identification of LVEF reduction and early treatment is of great significance to improve LVEF and the prognosis of HF. This research aims to introduce a new method for left ventricular dysfunction (LVD) identification based on phonocardiogram (PCG) and electrocardiogram (ECG) signals synchronous analysis. In the present study, we established a database called Synchronized ECG and PCG Database for Patients with Left Ventricular Dysfunction (SEP-LVDb) consisting of 1046 synchronous ECG and PCG recordings from patients with reduced ($n = 107$) and normal ($n = 699$) LVEF. 173 and 873 recordings were available from the reduced and normal LVEF group, respectively. Then, we proposed a parallel multimodal method for LVD identification based on synchronous analysis of PCG and ECG signals. Two-layer bidirectional gate recurrent unit (Bi-GRU) was used to extract features in the time domain, and the data were classified using residual network 18 (ResNet-18). This research confirmed that fused ECG and PCG signals yielded better performance than ECG or PCG signals alone, with an accuracy of 93.27%, precision of 93.34%, recall of 93.27%, and F1-score of 93.27%. Verification of the model's performance with an independent dataset achieved an accuracy of 80.00%, precision of 79.38%, recall of 80.00% and F1-score of 78.67%. The Bi-GRU model outperformed Bi-directional long short-term memory (Bi-LSTM) and recurrent neural network (RNN) models with a best selection

frame length of 3.2 s. The Saliency Maps showed that SEP-LVDPN could effectively learn features from the data.

Keywords: multi-modal; phonocardiogram; electrocardiogram; neural network; left ventricular dysfunction

1. Introduction

Heart failure (HF) is a terminal stage of cardiac disease associated with a poor prognosis, high mortality, and expensive medical costs [1] and has become a serious clinical and public health problem [2,3]. Although HF incidence rates in developed countries have stabilized or even decreased, HF incidence rates in low-income areas and the total number of HF patients worldwide continue to increase [3]. Indeed, HF represents a major economic and medical burden, with approximately 64.34 million patients affected worldwide [4].

At present, echocardiography, B-type and N-terminal pro-B-type natriuretic peptide analysis are the mainstays of diagnosis of HF. Left ventricular ejection fraction (LVEF) measured by echocardiography is not only an indicator of HF diagnosis but is also applied in treatment [5,6]. Based on LVEF measurements, HF can be classified into three types: HF with preserved EF (HFpEF; $LVEF \geq 50\%$), HF with mid-range EF (HFmrEF; $40\% \leq LVEF < 50\%$), and HF with reduced LVEF (HFrEF; $LVEF < 40\%$) [7]. Current evidence suggests that the 1-year mortality rates are 14.1 and 12.1% in HFrEF and HFpEF patients and in the middle for HFmrEF patients [8,9]. Previous studies suggested that HF with restored LVEF or HF with improved LVEF had a better prognosis than HF with persistently decreased LVEF [1]. Changes in LVEF in HF patients are more likely to occur in earlier stages of the disease [10]. Therefore, early identification of reduced LVEF is significant for diagnosis and treatment. However, LVEF is mainly measured by echocardiography, which is highly dependent on examiners' skills, image quality, and modality [11]. Therefore, there is an urgent need for a simple and accurate method for left ventricular dysfunction (LVD) in clinical practice.

It is widely acknowledged that electrocardiogram (ECG) and phonocardiography (PCG) signals reflect mechanical and electronic movements, respectively. Both are non-invasive, low-cost, and easily available for medical examinations. Over the years, researchers have extensively analyzed PCG or ECG using deep learning methods to detect cardiovascular diseases. Wu et al. built an ensemble CNN model for heart sounds classification using the PhysioNet/Computing database in the Cardiology Challenge of 2016 [12], which achieved a sensitivity of 86.46% and specificity of 85.63% in hold-out testing. Using the same database, Deng et al. proposed a new Mel-frequency cepstrum calculation method, and their model based on a deep convolutional and recurrent neural network achieved a classification accuracy of 98% [13]. Li et al. built a fusion framework based on multi-domain features and deep learning features of PCG for coronary artery disease detection [14]. They confirmed that the fusion framework performed better than the multi-domain or deep learning features alone. Interestingly, a previous study also used a deep learning network to recognize cardiac murmurs [15]. Clinically, ECGs are mainly used to detect arrhythmia (ARR), myocardial infarction, ventricular hypertrophy, and electrolyte disturbances. Dami and Yahaghizadeh proposed a long short-term memory-deep belief network (LSTM-DBN) model to predict arterial events a few weeks or months before the event by

analyzing ECG with a mean accuracy of 88.42% [16]. Last but not least, Gumpfer et al. proposed an artificial intelligence model based on a CNN to detect myocardial scars with a sensitivity of 70%, specificity of 84.30%, and accuracy of 78% [17].

Computer-aided detection technology has been used to analyze PCG or ECG signals for HF identification in recent years. Liu et al. used an extreme learning machine classifier to identify HFpEF by analyzing PCG features extracted by multifractal detrended fluctuation analysis (MF-DFA) with an accuracy, sensitivity, and specificity of 96.32, 95.48 and 97.10%, respectively [18]. However, the features were manually extracted in this research and may have omitted other important features. Gao et al. proposed a gated recurrent unit (GRU) model that distinguished healthy people, HFpEF patients, and HFrEF patients with an average accuracy of 98.82% [19]. This research showed that the GRU model performed better than the long short-term memory (LSTM) model, FCN model and support vector machine (SVM). However, HS databases for HFrEF and HFpEF are lacking; therefore, generalized tests cannot be performed on other public databases. Furthermore, HF recordings in this study were collected from 42 HFrEF and 66 HFpEF patients. Gjoreski et al. combined the traditional method with a deep learning method for chronic heart failure identification based on HF data obtained from only 51 CHF patients [20]

Cho et al. developed a 12-lead ECG analysis artificial intelligence algorithm for HFrEF identification based on a deep learning network which yielded an area under the curve for internal and external verification of 0.913 and 0.961, respectively [21]. Their study achieved a sensitivity, specificity, and accuracy of 90.5, 75.6 and 77.5% during internal validation and 91.5, 91.1 and 91.1%, respectively, during external validation. Li et al. proposed a deep convolutional neural network-recurrent neural network (CNN-RNN) model for different stages of HF recognition [22]. This research showed that ECG signals between normal subjects and HF patients were significantly different, based on a combination of ECG features with many other clinical features for classification, such as gender, age, coronary heart disease, hypertension, history of diabetes, and percutaneous coronary intervention, which may be inconvenient for clinical practice. Eltrass et al. proposed a new ECG diagnosis algorithm that combined CNN with the Constant-Q non-stationary Gabor transform for congestive heart failure (CHF) and ARR identification with an accuracy, sensitivity, specificity and precision of 98.82, 98.87, 99.21 and 99.20% [23]. This study adopted the BIDMC Congestive Heart Failure Database, containing only 30 ECG recordings. Previous studies that used PCG or ECG signals for HF identification are shown in Table 1.

It is well-established that acoustic cardiography combines ECG and PCG to evaluate cardiac function. The major cardiac acoustic biomarkers associated with HF include electromechanical activation time (EMAT) [24], systolic dysfunction index (SDI) [25], EMAT/RR interval (%EMAT), and left ventricular systolic time (LVST) [26]. One of the most critical biomarkers is EMAT, defined as the period from the onset of the Q wave to the first peak of the first heart sound (S1). This reflects the time delay of the electrical excitation and mechanical movement. In a recent study, Li et al. demonstrated that an $EMAT \geq 104$ ms diagnosed $LVEF < 50\%$ with a sensitivity of 92.1% and specificity of 92% [24]. A previous study showed that $\%EMAT \geq 0.15$ diagnosed $LVEF < 40\%$ with a sensitivity of 54%, specificity of 92%, and accuracy of 72% [27]. Moyers et al. confirmed that EMAT/LVST performed better than EMAT in detecting left ventricular dysfunction (defined as the presence of both $LVEDP > 15$ mmHg and $LVEF < 50\%$) [26]. Acoustic cardiography comprehensively evaluates the mechanical and electronic functions of the heart [28]. Li et al. proposed a multi-modal machine learning method to predict cardiovascular diseases by integrating ECG and PCG [29]. This

study showed that the performance of multi-modal method outperformed the other cases with single model based on ECG or PCG. Integrating ECG and PCG features may also play an essential role in assessing HF. However, there is no research on the simultaneous analysis of PCG and ECG based on deep learning networks.

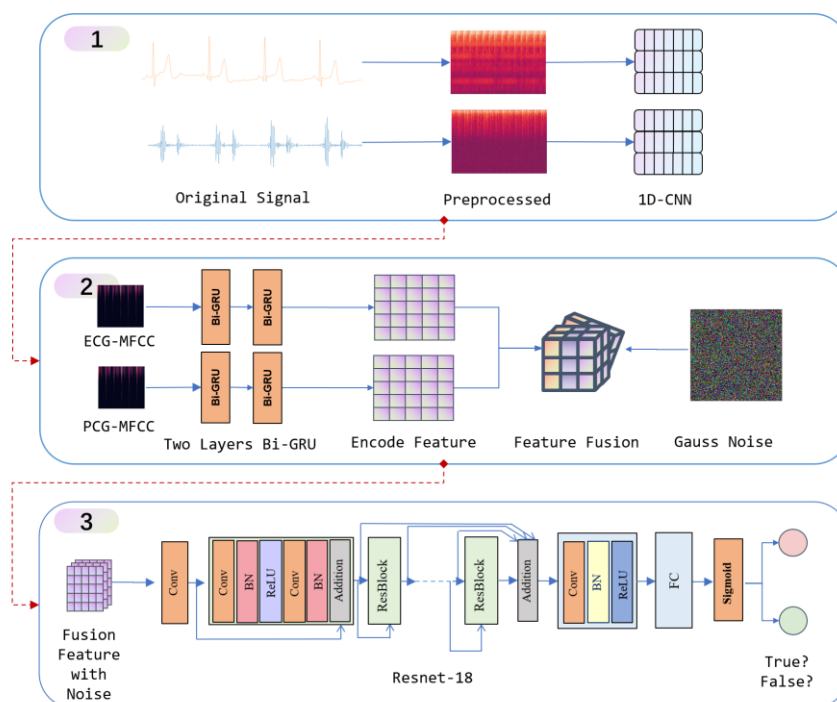


Figure 1. Comprehensive SEP-LVDPN framework. This figure describes how the SEP-LVDPN model works with data flow as a medium. The input consists of synchronized but independent PCG and ECG signals sent to the neural network for prediction after preprocessing. The input data are first independently extracted using two-layer Bi-GRU for high-dimensional features in the neural network. Then, the two independent feature matrices are fused per channel, and Gaussian noise is mixed in an additional channel. Finally, the fused feature blocks are learned and classified by the residual network, and the prediction results are output.

In the present study, we first established a dataset called “Synchronized ECG and PCG Database for Patients with Left Ventricular Dysfunction (SEP-LVDb)” a medium-scale ontology of cardiac physiological signals, including 1046 recordings, which is, to our knowledge, the first deep learning dataset containing synchronized ECG and PCG signals using LVEF as a binary label. Signals in this database include PCG and ECG synchronized in the time dimension, and the basic information contains sex, age, systolic pressure, diastolic pressure, and LVEF. Based on this dataset, we proposed a deep neural network called “Synchronous ECG and PCG Left Ventricular Dysfunction Prediction Network (SEP-LVDPN)” as a performance benchmark (Figure 1). SEP-LVDPN is a two-stage multimodal fusion neural network consisting of a two-layer bidirectional gate recurrent unit (Bi-GRU) and residual network 18 (ResNet-18). This model was designed for left ventricular dysfunction

screening by simultaneous analysis of PCG and ECG. The input signals of our model are one-dimensional PCG and ECG signals.

Table 1. Summary of other research using PCG or ECG signals for HF classification.

Authors	Purposes	Recordings(subjects)	Methods	Results
Liu et.al (2019) [18]	HFpEF vs. normal	401 normal 441 HFpEF	PCG features were extracted by MF-DFA and classified by ELM	Acc = 96.32% Sen = 95.48% Spe = 97.10%
Gao et.al (2020) [19]	normal vs. HFpEF vs HFrEF	Unknown (42 HFrEF) Unknown(66HFpEF) 1286 normal	PCG features were learned and classified by the GRU model	Acc = 98.82%
Gjoreski et.al (2020) [20]	Normal vs. Chronic HF Recomp. Vs. Decomp	159 Healthy (110) 22 Recomp. (22) 52 Decomp. (51)	PCG features were extracted by an ML and an end-to-end DL, then classified by a recording-based ML	Acc = 84.2% Sen = 66.3% Spe = 93.5% Acc = 93.2% Sen = 90.9% Spe = 95.5%
Li et.al (2019) [22]	Normal vs. Stage A vs. Stage B vs. Stage C vs. Stage D	172 normal 84 Stage A 156 Stage B 105 Stage C 56 Stage D	ECG features were extracted by CNN and classified by RNN combined with other clinical features	Acc = 97.6% Sen = 96.3% Spe = 97.4%
Eltrass et.al (2021) [23]	CHF vs. arrhythmia vs. normal	576 ARR (47) 180 CHF (15) 216 NSR (18)	ECG features were extracted by the AlexNet and discriminated by MLP	Acc = 98.82% Sen = 98.87% Spe = 99.21% Pre = 99.20%
Cho et.al (2021) [21]	HFrEF vs. non-HFrEF	Hospital A 20,882 Non-HFrEF (19693) 2,021 HFrEF (342) Hospital B 4173 Non-HFrEF (4020) 189 HFrEF (156)	ECG features were learned and classified by CNN	internal validation Sen = 90.5% Spe = 75.6% Acc = 77.5% external validation Sen = 91.5% Spe = 91.1% Acc = 91.1% Acc=73.9%
Sun et.al (2021) [30]	LVD vs. non-LVD	1262 LVD 25530 non-LVD	ECG features were extracted by LeNet-5 architecture	Sen= 69.2% Spe= 70.5% PPV=70.1% NPV= 69.9%

*Note: Abbreviation:(1) Sen: sensitivity; Spe: specificity; (2) MF-DFA: multifractal detrended fluctuation analysis; (3) ELM: extreme learning machine; (4) ML: machine learning; (5) DL: deep learning; (6) MLP: Multi-Layer perceptron; (7) ARR: arrhythmia; (8) CHF: congestive heart failure; (9) NSR: normal sinus rhythm; (10) PPV: positive predictive value; (11) NPV: negative predictive value

2. Materials and methods

2.1. Database

Herein, we aimed to establish a deep learning network model to analyze PCG and ECG simultaneously to identify patients with LVD. To the best of our knowledge, no database containing synchronized ECG and PCG signals using LVEF as a binary label has hitherto been documented in the literature. A new database called “SEP-LVDb” is introduced in this section. All recordings were collected from inpatients in the Fourth Affiliated Hospital of Zhejiang University School of Medicine from March 2021 to August 2021.

2.1.1. Ethics approval of research

This research was approved by the Human Research Ethics Committee of the Fourth Affiliated Hospital of the Zhejiang University School of Medicine. Informed consent was obtained from all the patients before collection. The adverse reactions and risks to the subjects were minimal and written informed consent may pose a threat to the subjects' privacy. Thus, the Human Research Ethics Committee of the Fourth Affiliated Hospital of the Zhejiang University School of Medicine approved the use of oral consent.

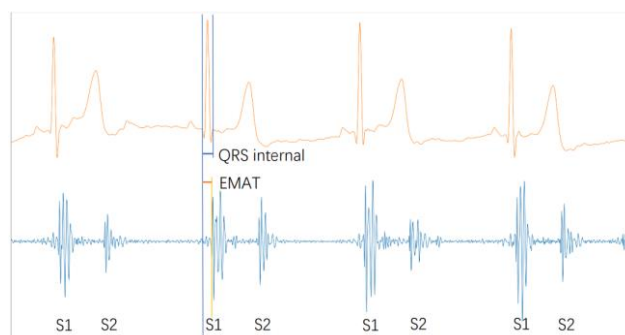
2.1.2. Acquisition process

This research included patients aged from 18 to 90 years. Most patients came from the Department of Cardiology, and a small part came from the Department of Endocrinology and Nephrology. Patients with any of the following conditions were excluded: (1) ventricular paced rhythm, (2) sick sinus syndrome, (3) third-degree atrioventricular block, (4) pre-excitation syndrome, (5) onset ventricular tachycardia or reentrant tachycardia, (6) after valve surgery and (7) dextrocardia. The causes of LVD included myocardial infarction, valvular diseases, ischemic cardiomyopathy, and non-ischemic dilated cardiomyopathy. All patients completed echocardiography, and the results were interpreted by experts.

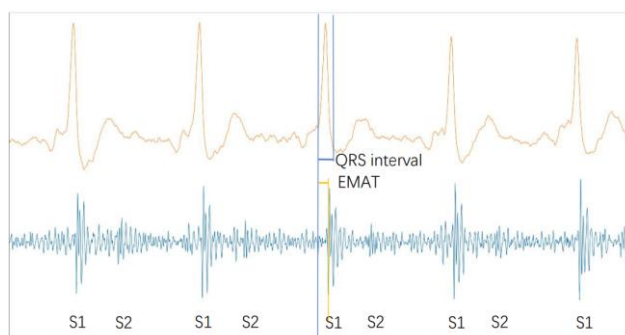
In this study, the recordings were collected using the DUO101 DUO ECG + digital stethoscope (Diglo, United States), which could record PCG and ECG signals synchronously. This stethoscope has four types of audio filters: diaphragm (100–500 Hz), bell mode (20–200 Hz), midrange (50–500 Hz), and extended (20–2000 Hz). The PCG frequency was 10–200 Hz, but the frequency of some murmurs could be extended to 600 Hz [31]. Recordings were collected in a real ward environment, with significant surrounding noise. Therefore, we chose a midrange audio filter during acquisition in this study. The recording length was 15 s.

Patients were placed in a supine position during data collection. Due to the severe condition of some patients who could not hold their breath, we only required the patients to breathe lightly. We collected the recordings from the precordial area. The stethoscope probe was placed in the 3rd to 4th intercostal spaces to the left side of the sternum at an angle of 30° with the sternum. A software called Eko was downloaded. The stethoscope was connected to the mobile phone through a spike connection. Signals were collected using a stethoscope probe. The recordings were automatically saved on the cloud platform. The recordings were then downloaded from the platform. ECG and PCG recordings were all saved in the wav format. If the patient's LVEF was less than 50% for the first time, we

completed the recording collection within 48 h before and after echocardiography. The recordings of the normal LVEF and reduced LVEF groups are shown in Figure 2. The ECG signals collected in this database were from single-lead ECG devices. The study flow chart is shown in Figure 3.



(a) Normal LVEF



(b) Reduced LVEF

Figure 2. (a) One recording of the normal LVEF group; (b) One recording of the reduced LVEF group.

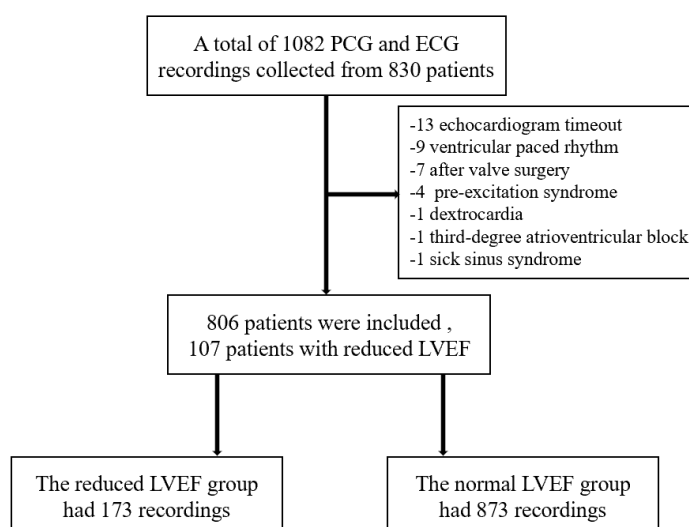


Figure 3. Study flowchart.

2.1.3. Database description

SEP-LVDb is a medium-scale ontology of cardiac physiological signals. This database contains a total of 1046 recordings from 107 patients with reduced LVEF and 699 patients with normal LVEF. Patients with reduced LVEF included 75 men and 32 women. Patients with normal LVEF included 397 men and 302 women. One to five recordings were collected per patient. In this study, according to the LVEF, the data were divided into two groups: reduced LVEF group with 173 recordings and the normal LVEF group with 873 recordings. Patients with heart failure usually have many cardiac and non-cardiac complications, some of which may influence ECG and PCG signals. These complications also occur in normal patients. It has been established that the most common non-cardiac complication is chronic obstructive pulmonary disease (COPD)/bronchiectasis (26%) [32]. Causes of cardiac complications are hypertension (55% in elderly patients), coronary artery disease, atrial fibrillation, bundle branch block, and valvular heart disease [33,34]. In this database, participants were not excluded because of the above situations in both the reduced and normal LVEF groups. The details of SEP-LVDb are shown in Table 2 and Figure 4.

Table 2. Details of SEP-LVDb and demographic information of subjects. Every recording is considered to be collected from an independent individual.

Parameters	Reduced LVEF Group (173)	Normal LVEF Group (873)
LVEF (Mean \pm SD, %)	37.30 \pm 8.36	64.90 \pm 5.27
Age (Mean \pm SD)	68.30 \pm 13.2	62.92 \pm 13.15
Male (%)	121 (69.9%)	495 (56.7%)
Blood Pressure (mmHg)	116.37/66.55	125.48/69.15
Hypertension	91 (52.6%)	542 (62.1%)
COPD	27 (15.6%)	45 (5.2%)
Atrial Fibrillation	26 (15.0%)	45 (5.2%)
Complete Left or Right Bundle Branch Block	31 (17.9%)	46 (5.7%)
Moderate or Severe Valve Regurgitation	58 (33.5%)	49 (5.6%)
Moderate or Severe Valvular Stenosis	6 (3.4%)	8 (0.9%)

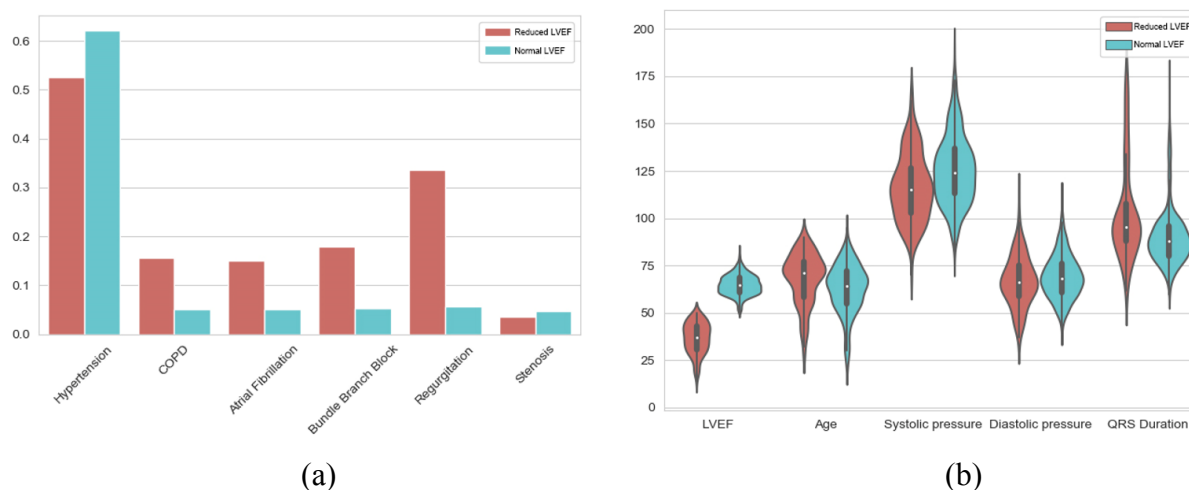


Figure 4. (a) Percentage of complications in reduced and normal LVEF groups. There is a significant difference in some indexes between the two groups. (b) Violin plot of basic patient information with LV dysfunction and non-LV dysfunction featuring a kernel density estimation of the multiple underlying distributions at once.

As the recordings were collected in a ward environment, the PCG recordings contained significant noises, such as conversations, alarm sounds from medical instruments, television noises, footsteps, rubbing of the stethoscope and chest wall, breathing, and intestinal peristalsis. The ECG recordings were mainly affected by poor contact with the electrode and the myoelectric activity of the respiratory muscles. Given that the electrodes need to be in close contact with the skin, it was challenging to collect recordings from underweight patients.

2.1.4. SEP-LVDb relevance

1) To the best of our knowledge, this is the first documented database containing PCG and ECG synchronous signals using LVEF as a binary label. Each recording was provided with the corresponding clinical information, such as the patient's gender, age, blood pressure, echocardiography results, and comorbidities, which facilitate further research.

2) This database was designed to develop computer-aided technology for LV dysfunction detection. As patients with LV dysfunction usually have many complications, participants were not excluded because of these complications, increasing the difficulty of PCG and ECG analysis, but in line with reality.

3) The PCG signals compiled in this database can be used for medical education on cardiac auscultation for medical students.

2.2. Methods

This model uses a multimodal parallel method to construct a dual-mode, dual-input, and multimodal deep neural network. During the preprocessing process, multimodal input signals were converted into a synchronized data frame and transformed into an input suitable for the neural network due to the heterogeneity of sampling rates. SEP-LVDPN is a two-stage model consisting of Bi-GRU and Resnet-18. The preprocessed data were extracted and classified by this model.

2.2.1. Data preprocessing

ECG signals are well-recognized as low-frequency signals with an effective frequency range of 0.05–100 Hz. PCG signals are also low-frequency signals, and the signal components are mainly concentrated in the range of 10–200 Hz. The Nyquist theorem was used for data preprocessing, using the following formula.

$$F_s > 2 * F_n$$

It can be concluded that the sampling frequency of the ECG signal is at least 200 Hz and the sampling frequency of the PCG signal is not less than 400 Hz. Since the purpose of the algorithm is to detect anomalies in the signal, the sampling rate for collecting ECG was set to 500 Hz, and the sampling rate for PCG to 4000 Hz.

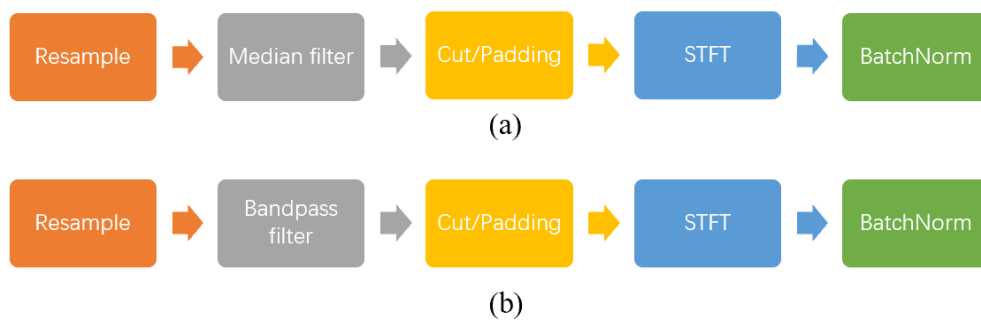


Figure 5. (a) ECG preprocessing process. (b) PCG preprocessing process.

In this study, the ECG recordings were collected from the hospital at a frequency of 500 Hz, which conformed to the sampling theorem. The ECG preprocessing is shown in Figure 5(a). The ECG signal was extremely weak; accordingly, the baseline was easily affected by external interference (i.e., poor electrode contact and myoelectric activity of respiratory muscles). In this study, low-frequency interference in the ECG signal was eliminated through a median filter. The median filtering method has a good filtering effect on the impulse noise, and the baseline drift phenomenon of the ECG signal can be eliminated. While filtering the noise, the median filter can protect the edge of the ECG signal and prevent it from being blurred.

$$g(x, y) = \text{med}\{f(x - k, y - l), (k, l \in W)\}$$

During acquisition, ECG signals were easily affected by high-frequency signals (such as electromyogram signals).

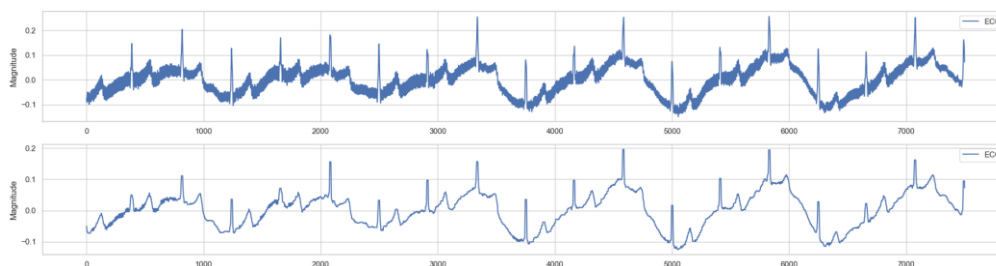


Figure 6. ECG preprocessing based on median filtering and wavelet transformation. The original signal is displayed above and the signal after median filtering below.

For ECG signals that underwent the above processing, every segment with 5000 sampling points (the time length of each segment is 10 s) was intercepted at a random starting position. The intercept position ensures that the distance from the end of the signal is greater than or equal to 5000. If the total length of the signal is less than 5000, it will be filled as 0. Finally, this signal segment is converted into a spectrogram after a short-time Fourier transform (STFT). The STFT formula is as follows:

$$STFT(t, f) = \int_{-\infty}^{+\infty} x(\tau)h(\tau - t)e^{-j2\pi f\tau} d\tau$$

Where $h(\tau - t)$ is the analysis window function, the window length is 50 (0.1 s, 10% of sample rate), corresponding to the signal's spectrogram. The ECG signal transformation involved using Fourier transform, with time on the horizontal axis, frequency on the vertical axis, and color representing the amplitude. The amplitude is the time-frequency distribution of the signal and can be understood as the color representing the energy distribution. The data format of the spectrogram is an $n \times m$ matrix. An ECG processed by STFT is shown in Figure 7.

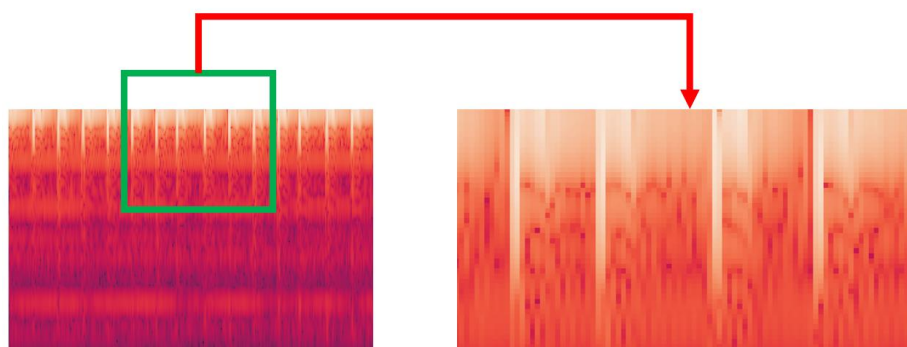


Figure 7. Spectrogram based on STFT.

The PCG signals were handled similarly to the ECG signals, and the preprocessing is shown in Figure 5(b). They were also collected from the hospital at a frequency of 4000 Hz. The PCG signal was also weak and susceptible to external interference.

We filtered out high- and low-frequency signals through a 10–2000 Hz bandpass filter during this research. Then, the processed PCG signal was intercepted with 40,000 sampling points (10 s) for analysis and converted to a spectrogram through STFT change.

2.2.2. Synchronous ECG and PCG left ventricular dysfunction prediction network (SEP-LVDPN)

SEP-LVDPN is a two-stage multimodal fusion neural network consisting of two-layer Bi-GRU and ResNet-18. The details of the model are shown in Figure 1. The two-layer Bi-GRU was used to extract features in the time domain. Next, the two input modalities were dimensionally spliced to obtain a two-dimensional feature matrix, and the data were classified using Resnet-18 and Sigmoid activation function.

The SEP-LVDPN model is an essential innovation in this study; it harnesses the powerful feature extraction capabilities of Resnet-18 and Bi-GRU to directly classify normal and abnormal heart signals from the original data. This model omitted the intricate segmentation and manual feature extraction steps of PCG and ECG and thoroughly used their global features and extracted frequency and time domain information simultaneously. Therefore, the SEP-LVDPN model achieved an excellent performance in classifying LVD patients using synchronized PCG and ECG.

Stage 1: Extraction and encoding of time-series features

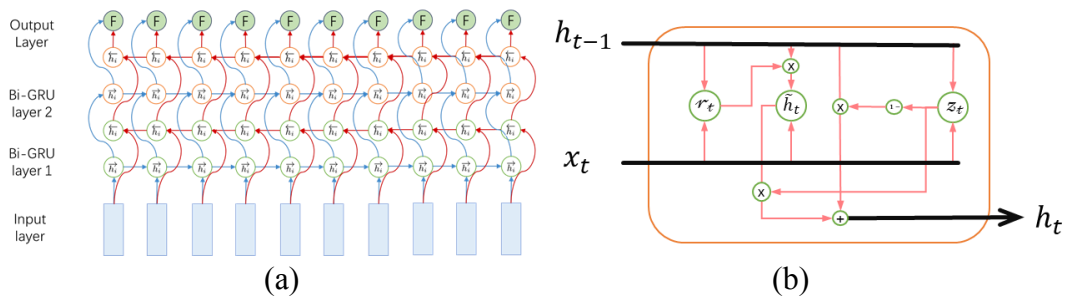


Figure 8. (a) Framework of two-layers Bi-GRU. (b) Cell of Bi-GRU.

ECG and PCG signals are highly relevant data in the time domain; accordingly, an RNN is suitable for this scenario. However, RNNs are subject to limitations such as vanishing gradients. RNN cells gradually forget what they have learned before or find it difficult to obtain new knowledge when the length of the input signal increases [35]. Bi-GRU is a time-series neural network that can extract the time series features contained in the data well. Bi means “bidirectional”; indeed, a neural network experiences positive and negative time flow cycles. First, the compressed spectrogram progresses forward with time and in the reverse direction when time is reversed [36–38]. The GRU is a variant of the LSTM network with a more straightforward structure than the LSTM network, yielding a relatively better effect in some scenarios [37,38].

The two-layer Bi-GRU model and the cell of the Bi-GRU are shown in Figure 8. The calculation formula for each GRU unit is as follows:

$$z_t = \sigma(W_z * [h_{t-1}, x_t])$$

$$r_t = \sigma(W_r * [h_{t-1}, x_t])$$

$$\tilde{h}_t = \tanh(W * [r_t * h_{t-1}, x_t])$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$

h_{t-1} is the output at the previous moment, while x_t and h_t represent the input and output at the current moment, respectively. It has an update gate and a reset gate. The update gate (z_t) is used to control the percentage of the previous state information brought into the current state. The larger the

value of the update gate, the more the information on the state from the previous moment. The reset gate (r_t) is used to control the degree to which the state information is ignored at the previous moment. The smaller the reset gate value, the more information is ignored. Every gated unit adaptively learns how much new information should be remembered and how much old information should be forgotten during the training process. A dropout layer was added between each Bi-GRU layer to alleviate the over-fitting phenomenon, and the rate was set to 0.6. Moreover, they can reduce the excessive dependence on the feature learning process and improve the generalization ability of the model [39].

Stage 2: Feature fusion and classification

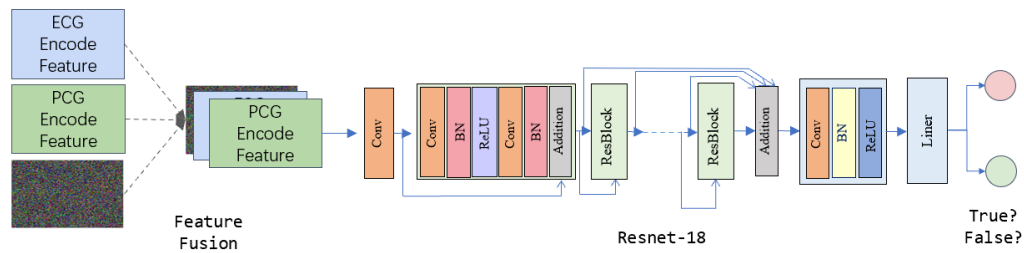


Figure 9. Feature fusion and classification using residual neural network.

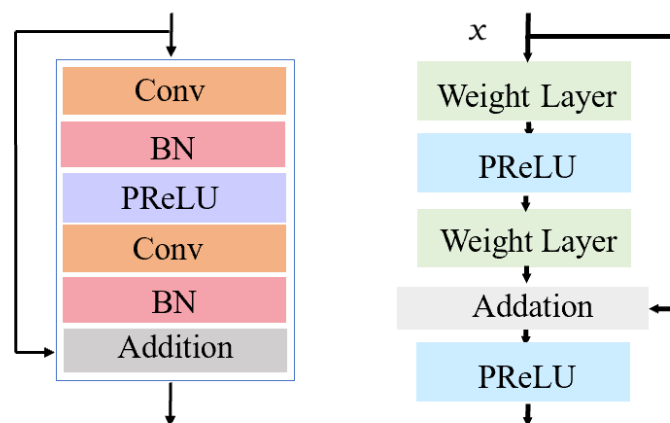


Figure 10. Convolution operation block (left) and remote jump connection (right).

In the third segment of the network, the PCG and ECG features were independently extracted by the Bi-GRU model and fused, and Gaussian noise was mixed to improve the model's generalization ability. Gaussian noise is a type of noise whose probability density function obeys a Gaussian distribution.

As shown in Figure 9, the network input is an $m \times n \times 3$ feature matrix, where m and n are the dimensions of the matrix, and the number of channels is three. The fused features with three channels will be input into the residual network. Each operation block of the residual network is composed of a convolutional layer, pooling layer, batch normalization (BN) layer, and PRelu layer. BN normalizes the input of specific layers through a mini-batch, thereby fixing the mean and variance of the input signal of each layer. This is a more effective local response normalization method to prevent gradient dispersion [40]. A long-distance jump connection is built between every two convolutional layers. At the end of the network, there is a linear layer and a sigmoid activation function. The output of the

activation function is the classification result. The presence of LV dysfunction was determined according to the value of the output neuron and confidence. The convolution operation block and remote jump connection are shown in Figure 10.

$$y = H(x, W_h) + x$$

The output y is a linear superposition of $H(x, W_h)$; x , is the identity map channel to the gradient; x is the input data and $H(x, W_h)$ is the output of the weight layer. The operation of each layer is given by the formula above. The residual network solves the vanishing gradient problem of deep convolutional networks through remote jump connections. The remote jump connection directly maps the shallow features to the deep network, thus simplifying the learning process and enhancing gradient propagation [41,42]. The residual network reduces degradation and increases expressive ability by breaking the asymmetry.

2.2.3. Experiment

All data are collected from frontline clinics by professional medical staff. All data were cleaned, and the missing and abnormal values were corrected. All codes were run in the Python 3.6 environment, and the neural network part was mainly constructed by the Pytorch deep learning framework. All experiments were performed on the workstation configured as follows: (1) CPU: Ryzen R9-5900X with 12 cores and 24 threads, 4.8 GHz; (2) GPU: Nvidia RTX 3090 with 24 G memory; (3) Memory: Multi-channel 3200 frequency 64 G memory; (4) Operating system: Ubuntu 20.04.

During the model training process, owing to the excellent video memory of the GPU, we set the batch size to 256. A larger batch size could provide a specific regularization effect for the deep learning model, increasing its generalization ability. For each model training, 800 iterations were conducted. Based on a large amount of data in our dataset, the optimizer chose Adam [43]. The initial learning rate was set to 0.0005, and as the training progressed, the ReduceLRonPlateau algorithm was used for adaptive adjustment. After finding that the loss no longer decreased or the accuracy value no longer increased, we reduced the learning rate [44,45]. Therefore, the matching degree of the learning rate and the learning process were maintained to the greatest extent to ensure that the model could fully absorb knowledge from the dataset [44–46]. The learning rate change process and loss value decrease are shown in Figure 11. The model successfully converged through learning and was close to the optimal value.

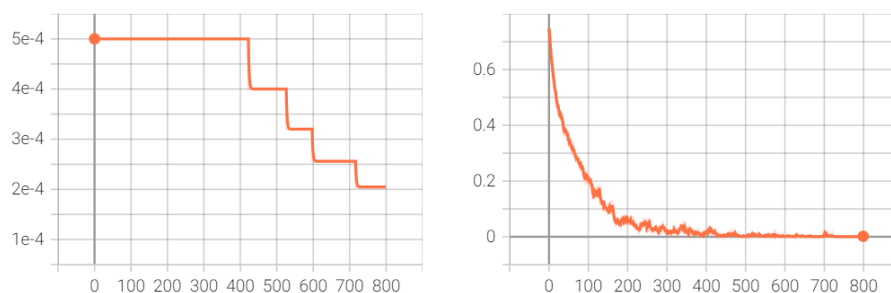


Figure 11. Adaptive correction of learning rate and convergence of loss value.

2.2.4. Performance evaluation

Five-fold cross-validation was conducted in this study. First, the PCG and ECG recordings of all subjects were divided into five subsets by stratified sampling, and the recordings in each subset were chopped into 3.2 s segments. Four subsets were used as the training set and the remaining as the validation set. Then, five iterations were performed, and the final classification result was the average of the cross-validations. As data of the reduced and normal LVEF groups were imbalanced, we replicated a set of reduced LVEF group data before stratified sampling. T, F, N and P were used to define true, false, negative and positive. Four evaluation indicators (EI) were introduced, including accuracy (Acc), precision (Pre), recall (Rec) and F-Score for assessing classification performance.

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

$$Pre = \frac{TP}{TP + FP}$$

$$Rec = \frac{TP}{TP + FN}$$

$$F - Score = 2 * \frac{Pre * Rec}{Pre + Rec}$$

Furthermore, to evaluate the performance of imbalanced data, the weighted avg was also used to consider the balanced performance with Pre, Rec and F-Score.

$$Weight\ Ave = EI_P * \frac{P}{P + N} + EI_N * \frac{N}{P + N}$$

3. Results

3.1. Separation and fusion of features

This study used the fusion features of ECG and PCG for LVD prediction. We compared the performances of the fusion, single ECG, and single PCG features for the LVD classification based on the Bi-GRU model with a 3.2 s slice length. Every feature was trained ten times independently, and the optimal performance obtained in the iteration process was recorded. Finally, the average accuracy of each model was calculated and compared. As the data was imbalanced, the weighted avg was used to consider the balanced performance.

Table 3. Experimental results of fusion feature, ECG, and PCG (Mean \pm STD).

Input	Accuracy	Precision	Recall	F-Score
ECG + PCG	93.27 \pm 1.36	93.34 \pm 1.45	93.27 \pm 1.36	93.27 \pm 1.39
ECG	90.43 \pm 1.35	91.23 \pm 1.31	90.43 \pm 1.35	90.62 \pm 1.31
PCG	90.32 \pm 1.40	90.45 \pm 1.53	90.31 \pm 1.43	90.32 \pm 1.45

As shown in Table 3 and Figure 12, the fusion feature yielded the best performance compared to ECG and PCG signals alone in terms of average accuracy (93.27 vs. 90.43 and 90.32%), precision (93.34 vs. 91.23 and 90.45%), recall (93.27 vs. 90.43 and 90.31%) and F-Score (93.27 vs. 90.62 and 90.32%, respectively).

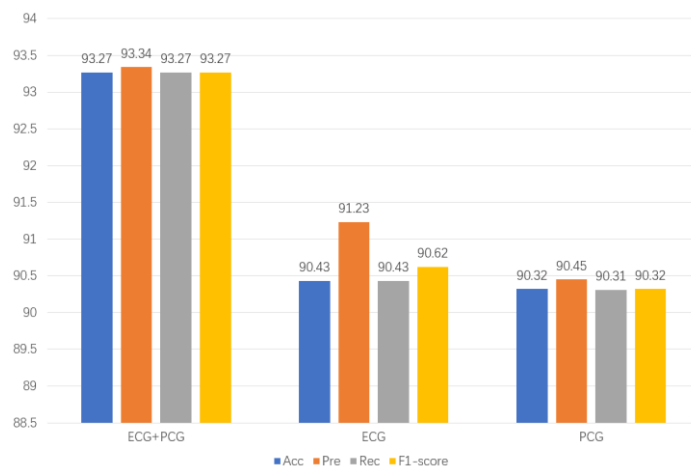


Figure 12. Comparison of different physiological signals based on the Bi-GRU model. Notably, fusion features yielded the best performance.

3.2. Different time series models

The high-dimensional fusion features obtained from a time-series network to classify LVD were the core of SEP-LVDPN. Therefore, it was necessary to carefully design the timing network structure to extract features from the time and frequency domains simultaneously. We compared three mainstream timing models of neural networks: RNN, Bi-GRU and Bi-LSTM.

Table 4 shows the experimental results of the Bi-GRU, RNN, and Bi-LSTM models. The Bi-GRU model achieved an accuracy, precision, recall and F-score of 93.98, 94.09, 93.99 and 93.98%, respectively. The RNN model achieved an accuracy, precision, recall and F-score of 90.94%, 91.13, 90.94 and 91.00%, respectively. The Bi-LSTM model achieved an accuracy, precision, recall and F-score of 88.46, 89.22, 88.46 and 88.67%, respectively. As shown in Figure 13, the Bi-GRU model yielded the best performance and was relatively stable.

Table 4. Experimental results of Bi-GRU, RNN and Bi-LSTM models (Mean \pm STD).

Models	Accuracy	Precision	Recall	F-Score
Bi-GRU	93.98 \pm 0.76	94.09 \pm 0.67	93.99 \pm 0.75	93.98 \pm 0.73
RNN	90.94 \pm 0.51	91.13 \pm 0.64	90.94 \pm 0.51	91.00 \pm 0.55
Bi-LSTM	88.46 \pm 2.19	89.22 \pm 1.75	88.46 \pm 2.19	88.67 \pm 1.84

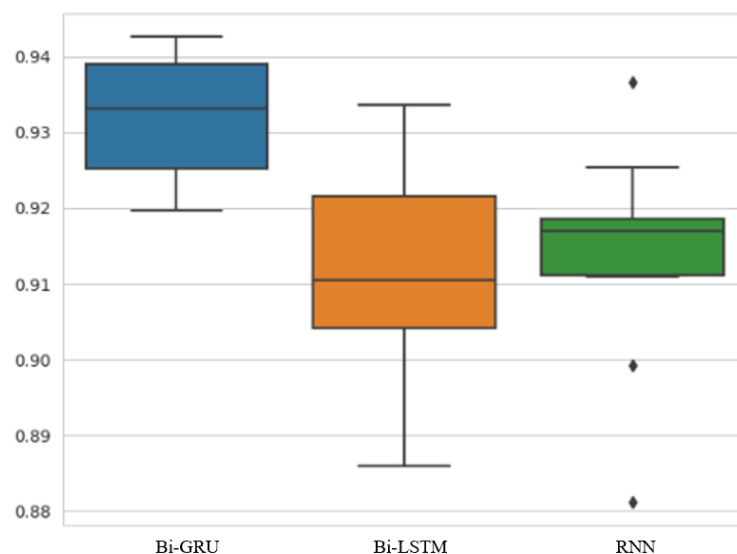


Figure 13. Accuracy comparison of the Bi-GRU, Bi-LSTM and RNN models. Notably, the Bi-GRU model performed the best.

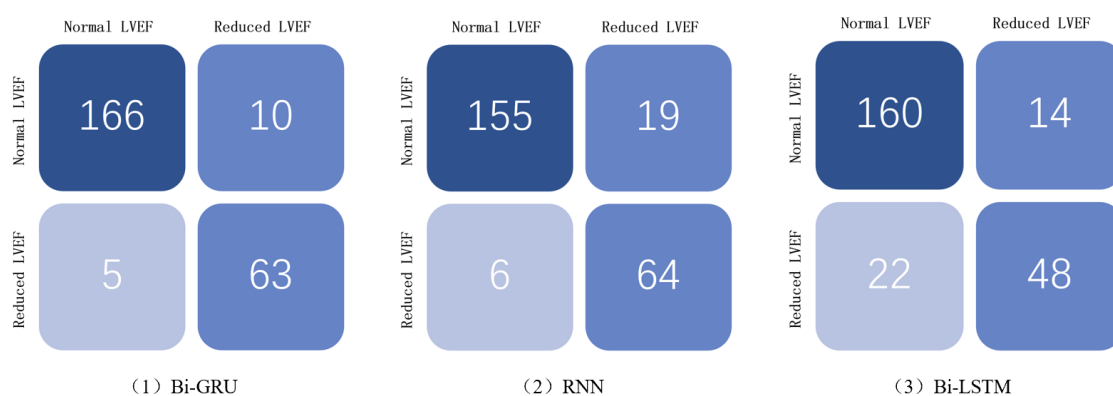


Figure 14. Confusion matrix of Bi-GRU, RNN and Bi-LSTM. The data predicted by the model and the actual data were largely consistent.

3.3. Frame length

We found that the length of the slice window significantly affected the model's performance. Different time slice lengths also significantly affected the calculation load. Given that the heartbeat cycle was approximately 0.8 s, we used a time interval of 1.6 s to search for the best slice length. The time interval of the search start position was 1.6 s, and the end position was 11.2 s. We added 1.6 s each time and retrained the neural network multiple times during the search process, and took the average value to comprehensively evaluate the model's performance. As shown in Figure 15, an optimal selection frame length of 3.2 s yielded the best performance and conducted the correct amount of calculation.

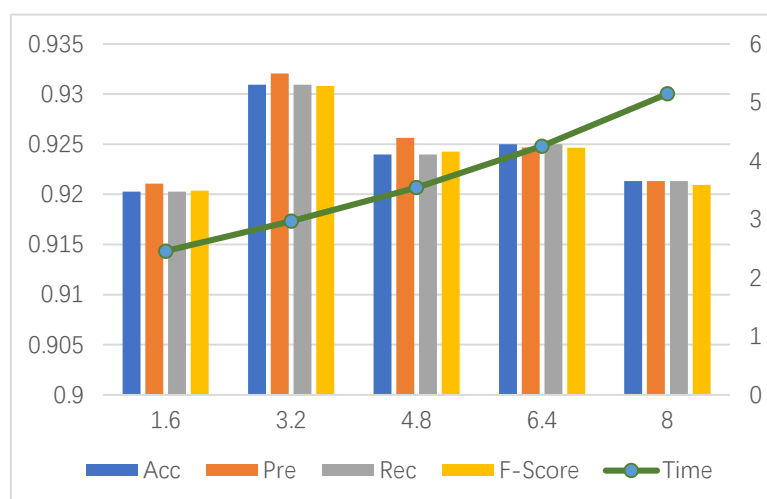


Figure 15. Comparison of the impact of different time-slice lengths on model performance. The model was trained for each time slice length, and the average values of Acc, Pre, Rec and F-score were counted. At the same time, the prediction time for 256 data pieces was calculated. The left axis is model performance, and the right axis is time.

3.4. Validation by an independent dataset

Table 5. Details of the independent dataset.

Parameters	Reduced LEVF Group (12)	Normal LVEF Group (28)
LVEF (Mean \pm SD, %)	40.06 \pm 7.41	63.04 \pm 4.32
Age (Mean \pm SD)	61.92 \pm 14.76	58.25 \pm 13.31
Male (%)	9 (75.0%)	17 (60.7%)
Hypertension	7 (58.3%)	19 (67.9%)
COPD	0 (0.0%)	1 (3.6%)
Atrial Fibrillation	1 (8.3%)	2 (7.1%)
Complete Left or Right Bundle Branch Block	1 (8.3%)	0 (0.0%)
Moderate or Severe Valve Regurgitation	6 (50.0%)	1 (3.6%)
Moderate or Severe Valvular Stenosis	0 (0.0%)	0 (0.0%)

Due to the lack of available databases, we collected 40 synchronous ECG and PCG recordings from 39 inpatients in different periods to establish a separate dataset to verify the model's performance.

The reduced LEVF group had 12 recordings collected from 11 inpatients, while the normal LEVF group had 28 recordings collected from 28 inpatients. Based on the Bi-GRU model with a time-slice length of 3.2 s, this model achieved an accuracy, precision, recall and F-score of 80.00, 79.38, 80.00 and 78.67%, respectively. Details of the independent dataset are shown in Table 5, and the confusion matrix of the results is shown in Figure 16.

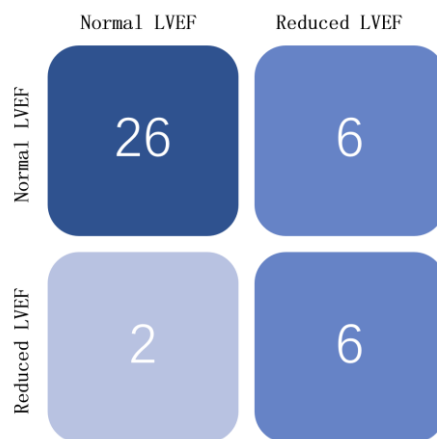


Figure 16. Confusion matrix of the results of validation by an independent dataset.

3.5. Interpretable visualization experiments

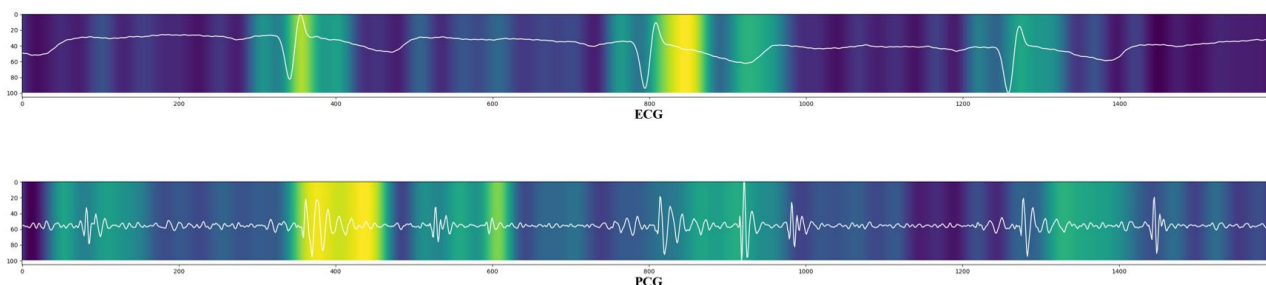


Figure 17. The Saliency Maps of SEP-LVDPN, it highlighted the primacy of the regions around the QRS wave, ST segment, T wave, the first heart sound (S1), the second heart sound (S2) and LVST.

The interpretability of deep learning has always been one of the world's trickiest problems with sources of its unreliability. Saliency Maps are adopted for model interpretation, which can be employed to figure out the important variables for the model. It can also express the importance of each feature by calculating the gradient while the value of the derivative reflects the influence of changes in the input data on the final results. We applied the Saliency Maps, combine the calculated derivative matrix with the original signal to analyze its attention distribution in the original data. The Saliency Maps of SEP-LVDPN were shown in Figure 17, with highlights of the primacy of the regions around the QRS wave, ST segment, T wave, the first heart sound (S1), the second heart sound (S2), and LVST (interval between S1 and S2).

4. Discussion

An important feature of this study was that LVD was detected by using synchronized ECG and PCG signals based on a neural network. In this study, we found that fusion features were significantly better than ECG or PCG alone for LVD classification. Over the years, ECG or PCG signals have been extensively used for HF classification [18–23]. Synchronized ECG and PCG signals analysis imply integrating ECG features with PCG features, such as QRS wave, ST segment, T wave, S1, S2 and LVST. Furthermore, an increasing body of evidence suggests that the time delay of cardiac electrical excitation and cardiac mechanical movement was extended in HF patients [24,26,27]. Moreover, SEP-LVDPN was designed for LVD screening by simultaneous analysis of PCG and ECG. It could recognize features in the frequency domain and learn the time phase features between them in the time domain. These may account for the good performance of the PCG and ECG synchronous analyses. The Bi-GRU model outperformed the Bi-LSTM and RNN models in this study. Consistently, in a study conducted by Gao et al., the performance of the GRU model was better than the LSTM model [19]. Although the LSTM network structure was more complex than GRU, its model performance was unstable and subject to large variations in this study. The possible reason was that the LSTM model might be too large and easily lead to overfitting. A simple RNN network might not be able to solve the gradient explosion problem; accordingly, its performance was significantly weaker than the Bi-GRU model. This study also demonstrated that 3.2 s was the optimal frame length option. A possible reason was that the field of view of the neural network and the continuity between features might be destroyed if the frame length was too small. In contrast, a large frame length might give the neural network extremely messy information, making it difficult for the network to focus its attention on the feature set effectively.

5. Conclusions

This research led to the development of SEP-LVDb, a medium-scale ontology of cardiac physiological signals, including 1046 recordings. The reduced and normal LVEF groups consisted of 173 and 873 recordings. Patients with or without LVD may have many complications that influence ECG or PCG signals. An essential feature of this database is that these subjects were not excluded, which makes this dataset more broadly representative. Furthermore, detailed clinical information is available for every recording, providing the foothold for further research.

This research proposed a multimodal parallel method for LVD identification based on PCG and ECG synchronous analysis. PCG and ECG signals were converted to spectrograms by STFT. Then, a two-layer Bi-GRU was used to extract the time domain and frequency domain features from PCG and ECG, respectively, and the features were fused and mixed into Gaussian noise. Finally, the multi-features used the Resnet-18 neural network for feature learning and classification.

We conducted experiments and comparisons of fusion features with PCG or ECG alone to validate that hybrid feature learning is effective. The results showed that fusion features were significantly better than the single feature. To better extract the features to be fused from the time-frequency domain, three different time-series neural networks were compared: RNN, Bi-GRU and Bi-LSTM. Bi-GRU achieved an optimal score owing to its appropriate model capacity and strong feature learning capabilities. The time slice length affects the calculation time and affects the model's performance. Interestingly, when a time slice length of 3.2 s is selected, the model can obtain better performance while ensuring that the amount of calculation does not increase significantly. We also conducted

interpretable visualization experiments in this study. The Saliency Maps showed that SEP-LVDPN could effectively learn the features from the data.

A larger database with high-quality data from multiple centers will be used for analysis in our future works. To increase the generalizability of our model, a hospital field noise signal model can be added to the input terminal, whereby branch reduction technology is applied to reduce redundant units and the amount of calculation, and improve the overall performance.

Acknowledgments

The authors acknowledge financial support from the National Natural Science Foundation of China (No. 81971688) and Medical and public health projects in Zhejiang Province (No. 2020386297).

Conflict of interest

The authors declare no conflict of interest.

References

1. G. Savarese, D. Stolfo, G. Sinagra, L. H. Lund, Heart failure with mid-range or mildly reduced ejection fraction, *Nat. Rev. Cardiol.*, **19** (2021), 100–116. <https://doi.org/10.1038/s41569-021-00605-5>
2. V. L. Roger, Epidemiology of Heart Failure: A Contemporary Perspective, *Circ. Res.*, **128** (2021), 1421–1434. <https://doi.org/10.1161/CIRCRESAHA.121.318172>
3. A. Groenewegen, F. H. Rutten, A. Mosterd, A. W. Hoes, Epidemiology of heart failure, *Eur. J. Heart. Fail.*, **22** (2020), 1342–1356. <https://doi.org/10.1002/ejhf.1858>
4. S. L. James, D. Abate, K. H. Abate, S. M. Abay, C. Abbafati, N. Abbasi, et al., Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017, *Lancet*, **392** (2018), 1789–1858. [https://doi.org/10.1016/s0140-6736\(18\)32279-7](https://doi.org/10.1016/s0140-6736(18)32279-7)
5. X. Zhang, Y. Sun, Y. Zhang, F. Chen, S. Zhang, H. He, et al., Heart failure with midrange ejection fraction: Prior left ventricular ejection fraction and prognosis, *Front. Cardiovasc. Med.*, **8** (2021), 697221. <https://doi.org/10.3389/fcvm.2021.697221>
6. J. N. Njoroge, J. R. Teerlink, Systolic time intervals in patients with heart failure: time to teach new dogs old tricks, *Eur. J. Heart Fail.*, **22** (2020): 1183–1185. <https://doi.org/10.1002/ejhf.1725>
7. P. Ponikowski, A. A. Voors, S. D. Anker, H. Bueno, J. G. Cleland, A. J. Coats, et al., 2016 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC). Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. *Eur. J. Heart Fail.*, **18** (2016), 891–975. <https://doi.org/10.1002/ejhf.592>

8. O. Chioncel, M. Lainscak, P. M. Seferovic, S. D. Anker, M. G. Crespo-Leiro, V. P. Harjola, et al., Epidemiology and one-year outcomes in patients with chronic heart failure and preserved, mid-range and reduced ejection fraction: an analysis of the ESC Heart Failure Long-Term Registry, *Eur. J. Heart Fail.*, **19** (2017), 1574–1585. <https://doi.org/10.1002/ejhf.813>
9. Meta-analysis Global Group in Chronic Heart, The survival of patients with heart failure with preserved or reduced left ventricular ejection fraction: an individual patient data meta-analysis, *Eur. Heart J.*, **33** (2012), 1750–1757. <https://doi.org/10.1093/eurheartj/ehr254>
10. J. Lupón, G. Gavidia-Bovadilla, E. Ferrer, M. de Antonio, A. Perera-Lluna, J. López-Ayerbe, et al., Dynamic trajectories of left ventricular ejection fraction in heart failure, *J. Am. Coll. Cardiol.*, **72** (2018), 591–601. <https://doi.org/10.1016/j.jacc.2018.05.042>
11. J. Butler, S. D. Anker, M. Packer, Redefining heart failure with a reduced ejection fraction, *JAMA*, **322** (2019), 1761–1762. <https://doi.org/10.1001/jama.2019.15600>
12. J. M. T. Wu, M. H. Tsai, Y. Z. Huang, S. K. H. Islam, M. M. Hassan, A. Alelaiwi, et al., Applying an ensemble convolutional neural network with Savitzky–Golay filter to construct a phonocardiogram prediction model, *Appl. Soft Comput.*, **78** (2019), 29–40. <https://doi.org/10.1016/j.asoc.2019.01.019>
13. M. Deng, T. Meng, J. Cao, S. Wang, J. Zhang, H. Fan, Heart sound classification based on improved MFCC features and convolutional recurrent neural networks, *Neural Netw.*, **130** (2020), 22–32. <https://doi.org/10.1016/j.neunet.2020.06.015>
14. H. Li, X. Wang, C. Liu, Q. Zeng, Y. Zheng, X. Chu, et al., A fusion framework based on multi-domain features and deep learning features of phonocardiogram for coronary artery disease detection, *Comput. Biol. Med.*, **120** (2020), 103733. <https://doi.org/10.1016/j.combiomed.2020.103733>
15. J. S. Chorba, A. M. Shapiro, L. Le, J. Maidens, J. Prince, S. Pham, et al., Deep learning algorithm for automated cardiac murmur detection via a digital stethoscope platform, *J. Am. Heart Assoc.*, **10** (2021), e019905. <https://doi.org/10.1161/JAHA.120.019905>
16. S. Dami, M. Yahaghizadeh, Predicting cardiovascular events with deep learning approach in the context of the internet of things, *Neural Comput. Appl.*, **33** (2021), 7979–7996. <https://doi.org/10.1007/s00521-020-05542-x>
17. N. Gumpfer, D. Grun, J. Hannig, T. Keller, M. Guckert, Detecting myocardial scar using electrocardiogram data and deep neural networks, *Biol. Chem.*, **402** (2021), 911–923. <https://doi.org/10.1515/hsz-2020-0169>
18. Y. Liu, X. Guo, Y. Zheng, An automatic approach using ELM classifier for HFpEF identification based on heart sound characteristics, *J. Med. Syst.*, **43** (2019), 285. <https://doi.org/10.1007/s10916-019-1415-1>
19. S. Gao, Y. Zheng, X. Guo, Gated recurrent unit-based heart sound analysis for heart failure screening, *Biomed. Eng. Online.*, **19** (2020), 3. <https://doi.org/10.1186/s12938-020-0747-x>
20. M. Gjoreski, A. Gradisek, B. Budna, M. Gams, G. Poglajen, Machine learning and end-to-end deep learning for the detection of chronic heart failure from heart sounds. *IEEE Access.*, **8** (2020), 20313–20324. <https://doi.org/10.1109/access.2020.2968900>
21. J. Cho, B. Lee, J. M. Kwon, Y. Lee, H. Park, B. H. Oh, et al., Artificial intelligence algorithm for screening heart failure with reduced ejection fraction using electrocardiography, *ASAIO J.*, **67** (2021), 314–321. <https://doi.org/10.1097/MAT.0000000000001218>

22. D. Li, X. Li, J. Zhao, X. Bai Automatic staging model of heart failure based on deep learning. *Biomed. Signal Process. Control*, **52** (2019), 77–83. <https://doi.org/10.1016/j.bspc.2019.03.009>
23. A. S. Eltrass, M. B. Tayel, A. I. Ammar, A new automated CNN deep learning approach for identification of ECG congestive heart failure and arrhythmia using constant-Q non-stationary Gabor transform, *Biomed. Signal Process. Control*, **65** (2021), 102326. <https://doi.org/10.1016/j.bspc.2020.102326>
24. X. C. Li, X. H. Liu, L. B. Liu, S. M. Li, Y. Q. Wang, R. H. Mead, Evaluation of left ventricular systolic function using synchronized analysis of heart sounds and the electrocardiogram, *Heart Rhythm*, **17** (2020), 876–880. <https://doi.org/10.1016/j.hrthm.2020.01.025>
25. S. Wang, F. Fang, M. Liu, Y. Y. Lam, J. Wang, Q. Shang, et al., Rapid bedside identification of high-risk population in heart failure with reduced ejection fraction by acoustic cardiography, *Int. J. Cardiol.*, **168** (2013), 1881–1886. <https://doi.org/10.1016/j.ijcard.2012.12.064>
26. B. Moyers, M. Shapiro, G. M. Marcus, I. L. Gerber, B. H. McKeown, J. C. Vessey, et al., Performance of phonoelectrocardiographic left ventricular systolic time intervals and B-type natriuretic peptide levels in the diagnosis of left ventricular dysfunction, *Ann. Noninvasive Electrocardiol.*, **12** (2007), 89–97. <https://doi.org/10.1111/j.1542-474X.2007.00146.x>
27. S. Efstratiadis, A. D. Michaels, Computerized acoustic cardiographic electromechanical activation time correlates with invasive and echocardiographic parameters of left ventricular contractility, *J. Card. Fail.*, **14** (2008), 577–582. <https://doi.org/10.1016/j.cardfail.2008.03.011>
28. Y. N. Wen, A. P. Lee, F. Fang, C. N. Jin, C. M. Yu, Beyond auscultation: acoustic cardiography in clinical practice, *Int. J. Cardiol.*, **172** (2014), 548–560. <https://doi.org/10.1016/j.ijcard.2013.12.298>
29. P. Li, Y. Hu, Z. P. Liu, Prediction of cardiovascular diseases by integrating multi-modal features with machine learning methods, *Biomed. Signal Process. Control*, **66** (2021), 102474. <https://doi.org/10.1016/j.bspc.2021.102474>
30. J. Y. Sun, Y. Qiu, H. C. Guo, Y. Hua, B. Shao, Y. C. Qiao, et al., A method to screen left ventricular dysfunction through ECG based on convolutional neural network, *J. Cardiovasc. Electrophysiol.*, **32** (2021), 1095–1102. <https://doi.org/10.1111/jce.14936>
31. C. Liu, D. Springer, Q. Li, B. Moody, R. A. Juan, F. J. Chorro, et al., An open access database for the evaluation of heart sound algorithms, *Physiol. Meas.*, **37** (2016), 2181–2213. <https://doi.org/10.1088/0967-3334/37/12/2181>
32. C. C. Lang, D. M. Mancini, Non-cardiac comorbidities in chronic heart failure, *Heart*, **93** (2007), 665–671. <https://doi.org/10.1136/hrt.2005.068296>
33. M. Metra, V. Zaca, G. Parati, P. Agostoni, M. Bonadies, M. Ciccone, et al., Cardiovascular and noncardiovascular comorbidities in patients with chronic heart failure, *J. Cardiovasc. Med.*, **12** (2011), 76–84. <https://doi.org/10.2459/JCM.0b013e32834058d1>
34. E. Sze, J. P. Daubert, Left bundle branch block-induced left ventricular remodeling and its potential for reverse remodeling, *J. Interv. Card. Electrophysiol.*, **52** (2018), 343–352. <https://doi.org/10.1007/s10840-018-0407-2>
35. W. Zaremba, I. Sutskever, O. Vinyals, Recurrent neural network regularization, preprint, arXiv:1409.2329.
36. S. Hochreiter, J. Schmidhuber, Long short-term memory, *Neural Comput.*, **9** (1997), 1735–1780. <https://doi.org/DOI 10.1162/neco.1997.9.8.1735>

37. J. Chung, Ç. Gülçehre, K. Cho, Y. Bengio, Empirical evaluation of gated recurrent neural networks on sequence modeling, preprint, arXiv:1412.3555.
38. K. Cho, B. v. Merriënboer, D. Bahdanau, Y. Bengio, On the properties of neural machine translation: Encoder-decoder approaches, preprint, arXiv:1409.1259.
39. N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, Dropout: a simple way to prevent neural networks from overfitting, *J. Machine Learn. Res.*, **15** (2014), 1929–1958.
40. S. Ioffe, C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift, preprint, arXiv:1502.03167
41. K. M. He, X. Y. Zhang, S. Q. Ren, J. Sun, Deep residual learning for image recognition, in *2016 Ieee Conference on Computer Vision and Pattern Recognition (Cvpr)*, (2016), 770–778. <https://doi.org/10.1109/Cvpr.2016.90>
42. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. E. Reed, D. Anguelov, et al., Going deeper with convolutions, in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (2015), 1–9. <https://doi.org/10.1109/CVPR.2015.7298594>
43. D. P. Kingma, J. Ba, Adam: A method for stochastic optimization, preprint, arXiv:1412.6980.
44. L. N. Smith, Cyclical Learning Rates for Training Neural Networks, in *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)*, (2017), 464–472. <https://doi.org/10.1109/WACV.2017.58>
45. I. Loshchilov, F. Hutter, SGDR: Stochastic gradient descent with restarts, preprint, arXiv:1608.03983.
46. Y. Lecun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, in *Proceedings of the IEEE*, **86** (1998), 2278–2324. <https://doi.org/10.1109/5.726791>



AIMS Press

©2022 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)