World Academy of Science, Engineering and Technology
International Journal of Biomedical and Biological Engineering
Vol:18, No:1, 2024

An Auxiliary Technique for Coronary Heart Disease Prediction by Analyzing ECG Based on ResNet and Bi-LSTM

Yang Zhang, Jian He

***Abstract*—**Heart disease is one of the leading causes of death in the world, and coronary heart disease (CHD) is one of the major heart diseases. Electrocardiogram (ECG) is widely used in the detection of heart diseases, but the traditional manual method for CHD prediction by analyzing ECG requires lots of professional knowledge for doctors. This paper presents sliding window and continuous wavelet transform (CWT) to transform ECG signals into images, and then ResNet and Bi-LSTM are introduced to build the ECG feature extraction network (namely ECGNet). At last, an auxiliary system for CHD prediction was developed based on modified ResNet18 and Bi-LSTM, and the public ECG dataset of CHD from MIMIC-3 was used to train and test the system. The experimental results show that the accuracy of the method is 83%, and the F1-score is 83%. Compared with the available methods for CHD prediction based on ECG, such as kNN, decision tree, VGGNet, etc., this method not only improves the prediction accuracy but also could avoid the degradation phenomenon of the deep learning network.

***Keywords*—**Bi-LSTM, CHD, coronary heart disease, ECG, electrocardiogram, ResNet, sliding window.

I. Introduction

HEART disease is the main cause of death nowadays, and CHD is the most common form of cardiovascular disease, accounting for about 13% of deaths in the United States [1]. Timely diagnosis of CHD is crucial to reduce the health risks caused by CHD such as cardiac arrest, so researchers began to study auxiliary diagnostic techniques for CHD.

In medicine, the auxiliary diagnostic techniques of CHD mainly include the auxiliary diagnosis based on physiological indicators, based on cardiac medical imaging and based on ECG. Among them, studies on auxiliary diagnosis based on physiological indicators are as follows: Kannel et al. found that major risk factors such as hypertension, high cholesterol and diabetes are related to CHD [2]. Irie et al.'s study on the general population showed that high levels of creatinine in blood can increase the risk of CHD [3]. In addition, blood cholesterol and glycoprotein levels in patients with CHD have been found to be consistently and significantly increased [4]. However, these measurements and analyses of many physiological indicators of CHD will increase the complexity of auxiliary diagnosis. In contrast, the auxiliary diagnosis of CHD based on medical imaging has the characteristics of accuracy and efficiency. For example, Madani et al. used a deep learning model [5] on echocardiogram images to judge CHD. Shi and Gu applied Coronary Arteriography (CAG) technique for the identification and diagnosis of CHD in Chinese medicine and achieved good results [6]. However, imaging test for CHD are expensive and the use of imaging techniques can be physically damaging to patients. ECG monitors patients' real-time ECG signals to assist doctors in disease diagnosis, and is a commonly used auxiliary diagnostic technique for CHD [7]. For example, Jin et al. used ECG signals to predict clinically important parameters related to patients with CHD (such as heart rate and axial migration) [8]. Wang et al. analyzed the nonlinear dynamic characteristics of ECG to diagnose CHD [9].

With the development of IT technology, researchers apply machine learning techniques to the study of CHD auxiliary diagnostic technology. Among them, the techniques using traditional machine learning algorithms are based on statistical analysis, decision tree and artificial neural network. For example, Cross et al. proposed a risk scoring system for clinical risk factors of CHD based on Cox regression model using physiological indicators such as serum protein, etc. to prevent the occurrence of CHD [10]. Meghan et al. studied the relationship between serum ferritin and the risk of CHD by using logistic regression and found that the risk of CHD increased by 5.1% for every 10 units of serum ferritin increase [11]. However, the simple prediction effect of these regression analysis methods is not significant. Therefore, the following methods are introduced, among which the relevant research based on decision tree is as follows: Maryam et al. established a prediction model of CHD based on 12 physiological indicators by using decision tree algorithm [12]. Karaolis et al. used C4.5 decision tree to predict the occurrence of CHD based on three groups of physiological indicators, with an accuracy of 75% [13]. However, the method based on decision tree ignores the correlation between data and is prone to overfitting. In terms of research based on artificial neural network: Rajeswari et al. proposed to use artificial neural network technology to mine knowledge from medical data to identify the risk level of CHD [14]. However, artificial neural networks are less efficient to train against large architectures. Most of the above machine learning methods for CHD diagnosis are based on numerous physiological indicators. With the success of deep learning technology in natural language processing (NLP), computer vision and other aspects, researchers have tried to apply deep

Yang Zhang is a Master's Student, Faculty of Information Technology, Beijing University of Technology, China (corresponding author, phone: (+86)18851966351; e-mail: 1271157595@qq.com).

Jian He is An Associate Professor, Faculty of Information Technology, Beijing University of Technology.

learning technology in the auxiliary diagnosis of CHD in recent years. For example, Han and Liang found in a comparative study that VGGNet19 convolutional network could effectively improve the efficiency of left ventricular segmentation of echocardiography and play an important role in the diagnosis of CHD [15]. Li et al. proposed a Deep Neural Network (DNN) based model named craftNet, which is used to accurately identify manual features to detect CHD and achieved a good accuracy [16]. Although ECG is an important technology in the auxiliary diagnosis of CHD, there is currently a lack of application of deep learning in the analysis of ECG data. In this paper, ECG scalogram is constructed by sliding window and CWT, and ECG time-frequency feature extraction network is constructed by combining ResNet and Bi-LSTM techniques, an ECG auxiliary diagnosis network with ResNet and Bi-LSTM is implemented named ECGNet. In addition, the ECG datasets exposed by MIMIC-3 was used for network training and testing, and the effectiveness of the proposed method was verified by experiments.

The remainder of this paper is organized as follows. Section II analyzes the ECG differences between CHD patients and normal people, and introduces the technology of Butterworth high-pass filter to denoise the original ECG signals, and the technology of sliding window and CWT to extract the time-frequency features of ECG. Section III introduces the network architecture of ECGNet and the principles of ResNet and Bi-LSTM respectively. Section IV introduces the implementation method of ECGNet. Section V uses MIMIC-3 ECG datasets for network training and verification, and analyzes the experimental results. Section VI presents the conclusion, and the suggestion for future works.

II. Methodology

In this section, based on the analysis of ECG features of CHD patients, noise in ECG data is filtered by Butterworth high-pass filter, secondly, sliding window and CWT techniques are introduced to realize time-frequency feature conversion of ECG signals, which provide the basis for the subsequent construction of deep learning-based auxiliary diagnosis of CHD.

A. ECG Feature Analysis for CHD

Since myocardial ischemia caused by CHD can cause specific changes in ECG, it is an important tool to monitor heart status and is commonly used in clinical diagnosis of CHD. Compared with the ECG of a normal person, the ECG of patients with CHD has three characteristics: (1) the T-wave is low; (2) the ST segment has downward movement; (3) The R-wave has decrease. Fig. 1 compares the ECG images of normal person and patients with CHD. I, II and III in the figure are the data sampled from the three lead positions of the ECG, it can be seen that data from CHD patient in all three lead positions may have all three of these characteristics, so it is only necessary to select one of the lead positions to study the characteristics of coronary patients, and in this paper the lead III portion of the ECG is selected.

Generally, ECG data collected by instruments and devices are susceptible to effects such as respiration, which can generate ultra-low frequency signal noise, thus leading to baseline roaming effect [17]. To this end, a 0.5 Hz Butterworth high-pass filter is used to remove baseline drift caused by motion, sensor impedance and respiration. The ideal high-pass filter cannot be realized by electronic components and has ringing phenomenon. Therefore, Butterworth high-pass filter is the most commonly used high-pass filter in practice. The transfer function of the filter can be calculated as (1):

$$H(u,v) = 1/(1+(D_0 / D(u,v))^{2n}) \quad (1)$$

$D(u,v)$ represents the distance from the midpoint of the frequency to the frequency plane and is the cut-off frequency. When $D(u,v)$ is greater than the D_0, $H(u,v)$ gradually close to 1, the high frequency part can pass through the filter; when $D(u,v)$ is less than the D_0, $H(u,v)$ gradually close to 0, the low frequency part cannot pass through the filter. In this paper, the low-frequency signal noise in ECG data is filtered by Butterworth high-pass filter.

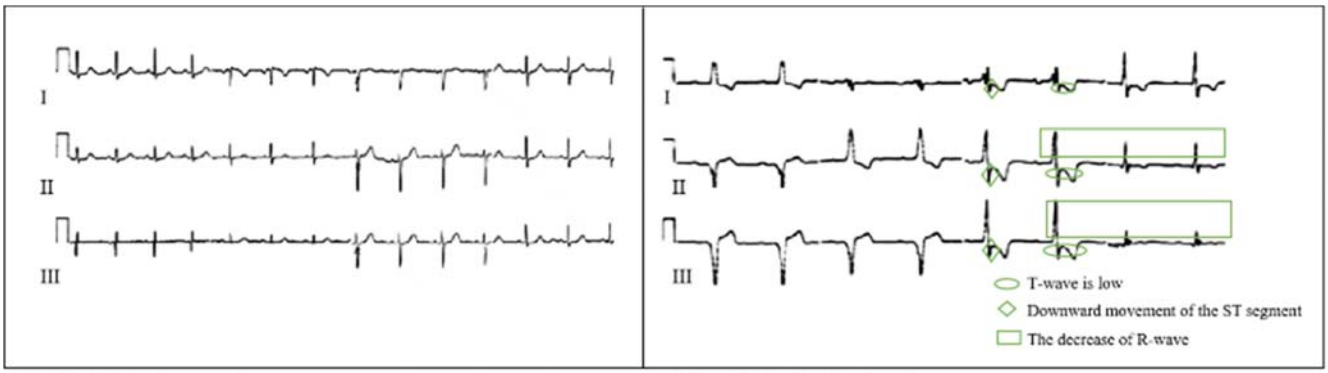

Fig. 1 Comparison of ECG between normal and CHD patients

B. Sliding Window for ECG

Since the human ECG data collected by ECG are continuous stream data, they cannot be directly applied to machine learning algorithms. In this regard, the sliding window technique is

introduced in this paper to segment the ECG data. In this paper, a timestamp is set for each ECG data, as new ECG data are continuously generated, the system continuously updates the old data in the window according to the first-in-first-out principle based on the timestamp. The sliding window maintains the integrity and temporal order of ECG data, which provides the basis for subsequent feature extraction and analysis based on deep learning models.

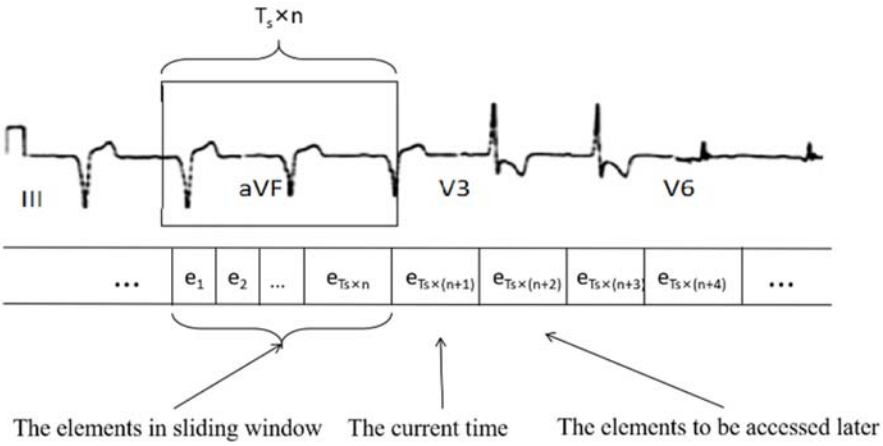

Fig. 2 Sliding window of ECG data

Fig. 2 shows an example of a sliding window of ECG data. In this case, the newly sensed data elements come from the right side of the window, and the elements on the left side of the window are moved out of the window according to the first-in-first-out principle. The sliding window contains a total of $Ts \times n$ time periods of sensed data (i.e., the window size). Ts is the sampling period, n is the sampling frequency, each element e_j in the window is the denoised ECG data sampled at moment j, and T is the total duration of ECG data sampling for a patient. With the sliding window, the continuous ECG data of a patient can be divided into N segments, (2).

$$N = \frac{T - T_s \times n}{s} + 1 \tag{2}$$

C. Frequency Feature Analysis Based on Wavelet

For the ECG data after denoising and sliding window processing, this paper uses the CWT to convert the ECG data into a scalogram which contains the time-frequency domain features of the ECG to provide a basis for developing ECG-based deep learning algorithms. The scalogram is defined as the absolute value of the CWT of the signal, which as a function of time and frequency [18] can identify the low-frequency and fast-changing frequency components of the ECG signal. The original ECG data are a one-dimensional vector signal that can be converted into a three-channel RGB image by the CWT. Compared with the short-time Fourier transform, the CWT can provide better temporal localization for short-time, high-frequency events, and better frequency localization for low-frequency, long-time events.

ECG signals have rich information in the time-frequency domain. The wavelet coefficients of the CWT can be used to locate the different frequency components. The CWT is defined as (3):

$$CWT_x^{\psi}(\tau, s) = \psi_x^{\psi}(\tau, s) = \frac{1}{\sqrt{|s|}} \int x(t) \psi\left(\frac{t-\tau}{s}\right) dt \tag{3}$$

$x(t)$ is primary time domain signal, $\Psi(t)$ is wavelet basis, τ and s are translation and scale transformation of the wavelet basis.

Fig. 3 shows the ECG scalogram generated by CWT of a segment of ECG data, which contain the time-domain and frequency-domain features of the ECG signal of the patient over a period of time, which provides a basis for the study of deep learning-based auxiliary diagnosis techniques for CHD.

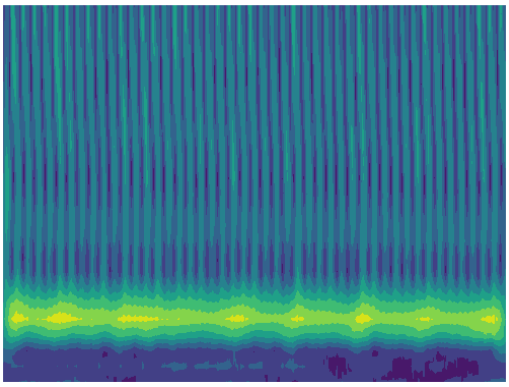

Fig. 3 Scalogram image of an ECG data

III. CHD Prediction Network Based on ECG Signal

In this section, ResNet and Bi-LSTM are introduced to construct the network architecture for CHD auxiliary diagnosis, and the principles and important parameters of ResNet and Bi-LSTM are introduced.

A. Architecture of the ECGNet

The denoised original ECG is converted into ECG scalogram by sliding window and CWT, and the features of these

scalograms can be extracted using convolutional networks. Based on the analysis and comparison of classical convolutional neural networks such as LeNet-5 [19], AlexNet [20], VGGNet [21], ResNet, and combining the advantages of RNN recurrent neural networks in processing continuous serialized data, we propose a network architecture oriented to ECG feature extraction, namely ECGNet. As shown in Fig. 4, ECGNet consists of input layer, ResNet layer, Bi-LSTM layer, Fully connected layer, and Softmax layer.

The input layer denoises the patient's original ECG data and generates a multi-segment ECG scalograms through sliding window and CWT processing. The ResNet layer consists of several convolutional and residual blocks and pooling layers, which extracts time-frequency features from multi-segment ECG scalograms. After that, the Bi-LSTM layer extracts the N segment feature data from the ResNet layer through the bidirectional long and short-term memory network. Finally, the feature data extracted by ResNet layer and Bi-LSTM layer are classified through the Fully connected layer and Softmax layer, and the category with the highest probability is output as the prediction result.

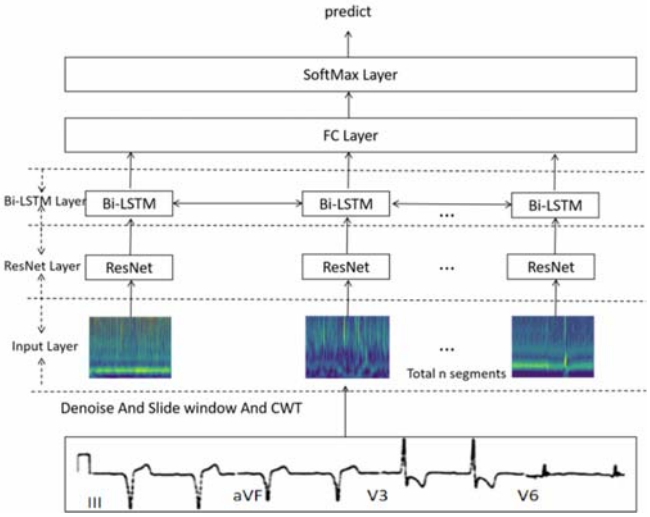

Fig. 4 ECGNet Architecture

B. ECG Feature Extraction Based on ResNet

Residual network is a kind of DNN, which is composed of multiple residual units. As shown in Fig. 5, each residual unit is implemented through a forward neural network and a shortcut connection. The core idea is the introduction of residual edge, that is, an edge connected directly from the input to the operation "⊕", on which h transformation is applied to the original input X_i. The residual unit can realize the fusion of features at different scales, and the residual network formed by it can effectively solve the gradient dispersion and degradation problems due to the increasing depth of deep learning network [22].

The i-th residual unit is shown in Fig. 5, whose input is X_i, F is network mapping transformation of input X_i, which can be a multilayer perceptron network, can also be a convolutional neural network and so on. X_{i+1} is the output of the "⊕" operation after the F and h transformations of the input X_i, and then the activation function f. The residual unit can be calculated as (4) and (5):

$$x_{i+1} = f(h(x_i) + F(x_i, w_i)) \tag{4}$$

If the equation h is the identity function, i.e., $h(X_i) = X_i$; According to (4), the features learned from shallow layer l to deep layer L can be calculated as (5):

$$x_L = x_l + \sum_{i=1}^{L-1} F(x_i, w_i) \tag{5}$$

Experimental analysis proves that when h is the identity function, F is the convolutional network, and the activation function f is ReLU, the residual network effect is optimal [22]. At the same time, a deep network containing hundreds or thousands of layers can be formed based on the residual network, which can effectively solve the gradient dispersion and degradation problems caused by the increasing number of layers in the deep learning network and can extract multi-scale feature fusion. Therefore, this paper adopts the fusion of convolutional network and constant shortcut connection of residual units to construct the feature extraction network of ECG scalogram.

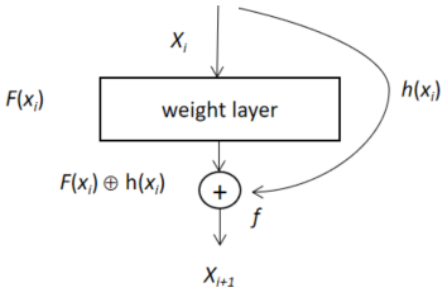

Fig. 5 Residual block

C. ECG Feature Extraction Based on Bi-LSTM

ResNet can generate n ECG feature vectors based on the sliding window, and there are temporal dependencies between these neighboring feature vectors. In this paper, Bi-LSTM is introduced to extract these temporal dependencies. As an improved recurrent neural network (RNN), Bi-LSTM can solve the problem of long distant dependence between data that RNN cannot process.

As shown in Fig. 6, the Bi-LSTM structural model can be divided into two independent LSTMS. X_1, X_2... X_n is the input, $\overrightarrow{h_t}$ and $\overleftarrow{h_t}$ represents the output of forward LSTM and reverse LSTM, that is, the output of Bi-LSTM network H_i is the stack of forward LSTM and reverse LSTM. The input sequences were respectively input into two LSTM neural networks in forward and reverse directions for feature extraction, whose expressions are shown as (6) and (7):

$$\vec{h}_t = f(\vec{b} + [x_t, h_{t-1}] \times \vec{w}) \tag{6}$$

$$\overleftarrow{h_t} = f(\overleftarrow{b} + [x_t, h_{t+1}] \times \overleftarrow{w}) \quad (7)$$

f is the activation function; the output of forward LSTM is $\overrightarrow{h_t}$; forward bias and weight are the $\overrightarrow{b}$ and $\overrightarrow{w}$; the output of the reverse LSTM neural network is $\overleftarrow{h_t}$; and the reverse bias and weight are $\overleftarrow{b}$ and $\overleftarrow{w}$. By splicing forward LSTM neural network and reverse LSTM neural network together, the output H_t of Bi-LSTM neural network can be obtained as (8):

$$H_t = c + g([\overrightarrow{h_t}, \overleftarrow{h_t}] \times U) \quad (8)$$

g is the activation function; H_t is the output of Bi-LSTM; the biases and weights are c and U. The idea of Bi-LSTM model design is to make the characteristic data obtained at time t contain the information between the past and the future. Experimental results show that this neural network model is superior to a single LSTM model in feature extraction efficiency and performance.

In this paper, ECGNet adopts a single layer Bi-LSTM network architecture to extract the ECG feature vectors fused with ECG temporal correlations. These ECG feature vectors pass through the subsequent FC layer and Softmax layer to classify whether the patient has coronary heart disease.

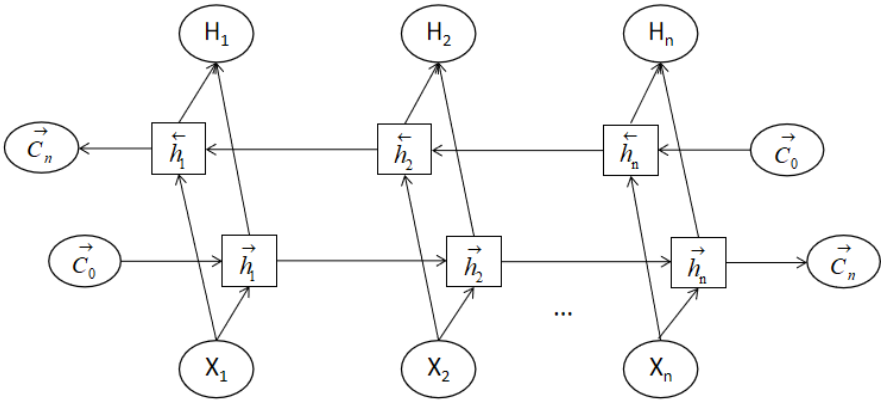

Fig. 6 Structure of Bi-LSTM

IV. Implementation

In this section, based on the above ECGNet architecture, we optimize the ResNet18 residual network and combine Bi-LSTM to design and implement ECGNet.

A. Implementation of ResNet

This paper optimizes the network model of ResNet18 and generates the residual network model shown in Fig. 7. This model includes 8 residual units, which are composed of 17 convolution layers and two pooling layers. For the 640x480x3 ECG scalogram generated by CWT, it was first uniformly reduced to 320x240x3 size images as the input of the residual network.

C1 is the first convolution block, the block contains only one layer, the kernel size is 7 x7, stride is 2, padding is 3, output channel is 64, after passing through C1 layer, the feature vector is transformed into dimensions of 64x160 x120.

S1 is the Max pooling layer, the kernel size is 3x3, stride is 2, padding is 1, after passing through S1 layer, the feature vector is transformed into dimensions of 64x80x60.

C2 to C5 are convolution blocks with residual structure. Every convolution block contains four convolution layers and two residual units, every two convolution layers have one shortcut links, each convolution layer's kernel size is 3x3, and the padding is 1. C2 convolution block's stride is 1, output channel is 64, after the C2 layer feature vector is still 64x80x60. C3 convolution in addition to the first layer's stride is 2, the rest of the stride is 1, the output of the channel is 128, after passing through C3 layer, the feature vector is transformed into dimensions of 128x40x30. C4 convolution block layer in addition to the first layer's stride is 2, the rest of the stride is 1, the output of the channel is 256, after passing through C4 layer, the feature vector is transformed into dimensions of 256x20x15. C5 in addition to the first layer's stride is 2, the rest of the stride is 1, the output channel number is 512, after passing through C5 layer, the feature vector is transformed into dimensions of 512x10x7.

S2 is an Average pooling layer, after passing through S2 layer, the feature vector is transformed into dimensions of 512x1x1.

Through the above residual network, the input ECG scalogram with the size of 320x240x3 is converted into 512x1x1 feature vectors for subsequent Bi-LSTM network processing.

B. Implementation of Bi-LSTM

Since the sliding window divides a patient's ECG data into N segments, a patient's ECG data will generate N dimensions of 512x1x1 feature vectors after passing through ResNet. These vectors are arranged in the order X_1~X_N. After a single layer of Bi-LSTM, a feature vector v containing the entire ECG data of a patient generated. The hidden layer of Bi-LSTM is set to 256, and other parameters are default values. After that, classification results are output through the Fully connected layer and Softmax layer, as in Fig. 8.

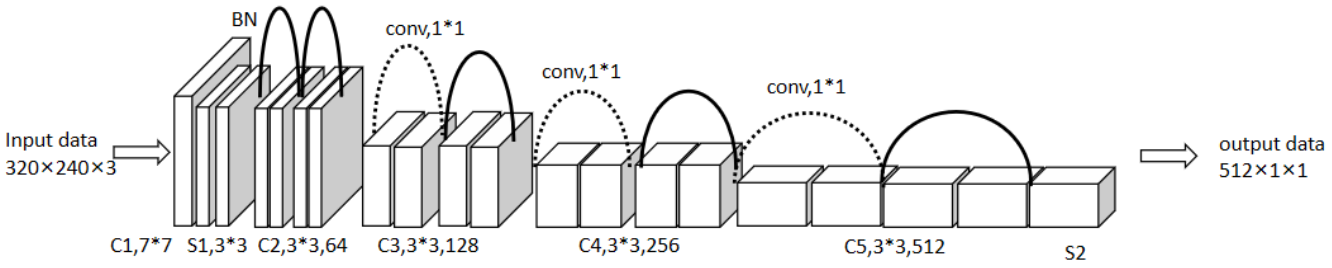

Fig. 7 Implementation of ResNet

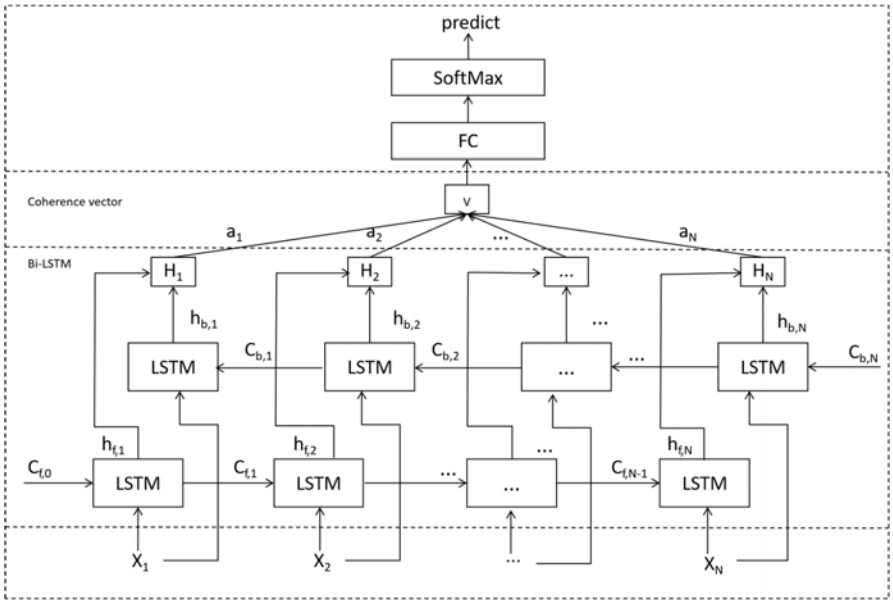

Fig. 8 Implementation of Bi-LSTM

C. Loss Function

The loss function adopted by ECGNet is cross-entropy loss function, which is often used in the loss calculation of classification problems. It can capture the difference of the relative size of prediction probability and further measure the performance of different classifiers in a more detailed way The cross-entropy loss function can be calculated as (9):

$$\text{Loss} = \frac{1}{N}\sum_{i=1}^{N} y_{ij} \log(p_{ij}), j = 1,2,...C \tag{9}$$

N is the sample number; C is classification number; y_{ij} expresses the case of a sample i belongs to the sample j, which only has two values, 0 or 1; and p_{ij} expresses the probability of a sample i forecast for the sample j, with value range of [0, 1].

V. Experiment and Its Analysis

In this section, we use MIMIC-3 ECG datasets for network training and testing, and analyze the experimental results.

A. Dataset Preparation

In this paper, we use PhysioNet's MIMIC-3 (Marketplace for Medical Information in Intensive Care) database [23], which is a publicly available multiparametric intensive care database provided by the Massachusetts Institute of Technology. The dataset contains physiological data such as ECG, photovolumetric pulse wave signals (PPG), arterial blood pressure signals (ABP) and respiratory signals (RESP) collected from patients in ICU wards, which has been successfully used in several research areas after more than 10 years of multidisciplinary construction.

This paper uses ECG data from the matched subset of the MIMIC-3 waveform database [24] to predict CHD. This dataset uses ICD-9 codes to code for CHD and assigns ICD-9 codes to each CHD patient in the MIMIC-3 database. This paper randomly selected 1230 patients from the large MIMIC-3 dataset for the study, of which 406 patients were diagnosed with CHD. Data from 904 of the 1230 patients were randomly selected for training (which contained 302 patients with CHD), 100 patients were used for validating (which contained 35 patients with CHD) and 226 patients were used for testing (which contained 69 patients with CHD).

For the ECG data of 1230 patients, we sequentially performed denoising, sliding window, and CWT to generate multi-segment ECG scalograms.

B. Model Evaluation Criteria and Systematic Experiment

In this paper, accuracy rate, recall rate and F1-score which take both accuracy rate and recall rate are used as evaluation

indexes of the model, F1-score is shown as (10):

$$F_1 = \frac{2 \times \mathrm{Pr}\,ecision \times \mathrm{Re}\,call}{\mathrm{Pr}\,ecision + \mathrm{Re}\,call} \tag{10}$$

Precision describes how many of the positive cases predicted by the bi-classifier are accurate, which can be calculated as (11). And recall describes how many of the true positive cases in the test set are selected by the bi-classifier, which can be calculated as (12). *TP* in (11) and (12) indicates that the real and predicted results are positive examples; *FP* indicates that the true result is a negative case and the predicted result is a positive case; and *FN* indicates that the true result is a positive case and the predicted result is a negative case.

$$\mathrm{Pr}\,ecision = \frac{TP}{TP + FP} \tag{11}$$

$$\mathrm{Re}\,call = \frac{TP}{TP + FN} \tag{12}$$

This paper selected a server equipped with Ubuntu operating system, which was specifically configured with E5-2620 CPU, 125 GB memory and TESLA M40 graphics acceleration card. This paper used Pytorch and Python to implement ECGNet.

The training parameters of ECGNet network are as follows: 60 batches are used, batch size is 16, 70 epoches are run, learning rate is 0.001, Adam optimizer is used, cross entropy loss function is adopted. The above parameter sets are used as inputs to train the model.

C. Experimental Analysis

First of all, in order to verify the effect of Butterworth high-pass filter on denoising, this paper compares the results of the ECGNet with and without denoising in the early stage, as shown in Table I. The results show that it is necessary to introduce Butterworth high-pass filter to filter low frequency waveforms.

TABLE I
COMPARISON RESULTS OF DENOISING OR NOT

Denoise or not	Accuracy	Recall	F1-score
Denoise	0.83	0.85	0.83
Not Denoise	0.72	0.73	0.72

In order to verify the validity of the ECGNet network model proposed in this paper, the research group compared the model with two traditional machine learning algorithms, K-NN [25] (k takes 3) and decision tree, as well as the popular deep learning algorithm for medical image classification: VGGNet (9 convolutional layers, 3 pooling layers and 1 global average pooling layer, all of the kernel size is 3x3) and ResNet18, ResNet34 and ResNet50, as shown in Table II. The results of table show that ECGNet model is significantly superior to these models. This may be due to the limited ability of traditional machine learning methods such as K-NN and decision tree to extract time-frequency features, while deep learning models such as VGGNet only rely on convolution, so they cannot extract and remember sequential features in ECG data. Therefore, the effectiveness of the proposed model combining sliding window and Bi-LSTM is proved.

TABLE II
COMPARISON RESULTS OF DIFFERENT MODELS

Model	Accuracy	Recall	F1-score
VGGNet	0.63	0.66	0.65
ResNet18	0.7	0.71	0.7
ResNet34	0.7	0.68	0.69
ResNet50	0.66	0.65	0.66
ECGNet	0.83	0.85	0.83
K-NN	0.61	0.59	0.6
Decision Tree	0.58	0.6	0.59

This paper also conducted experimental analysis on the sliding window length and step size used for data segmentation, and the results are shown in Fig. 9. It can be seen that as the window length increases from 20 s to 45 s, the F1-score predicted by CHD generally increases first and then decreases, among which, the window length reaches the maximum peak at 30 s. This indicates that the F1-score of CHD prediction is not higher with the longer sliding window length, but has an optimal state is at 30 s. Too small or too large window length is not conducive to feature extraction. After determining the size of the sliding window length, this paper analyzed the influence of the window step size. The window step size is usually smaller than the window length, which causes a partial overlap of active data between two adjacent windows. As can be seen from Fig. 9, with the window step size increasing from 1 s to 25 s, F1-score of CHD prediction generally increased first and then decreased, reaching the peak value at 15 s. Through experiment, this paper found that the prediction accuracy of CHD is the highest when the window step is half of the window length.

At the same time, this paper also compared the performance of increasing the number of Bi-LSTM layers with the same parameter settings as above, as shown in Table III. This comparison test is based on the ECGNet, just simply increasing the number of Bi-LSTM layers to compare and observe the effect of the model with different layers of Bi-LSTM. The table shows that the Bi-LSTM model with single layer has better results than that with two or three layers. Perhaps because single-layer Bi-LSTM is sufficient to memorize ECG data temporal dependencies. Therefore, a single-layer Bi-LSTM is used in this paper.

TABLE III
COMPARISON RESULTS OF DIFFERENT BI-LSTM LAYERS

Layer number	Accuracy	Recall	F1-score
1	0.83	0.85	0.83
2	0.79	0.80	0.79
3	0.75	0.72	0.73

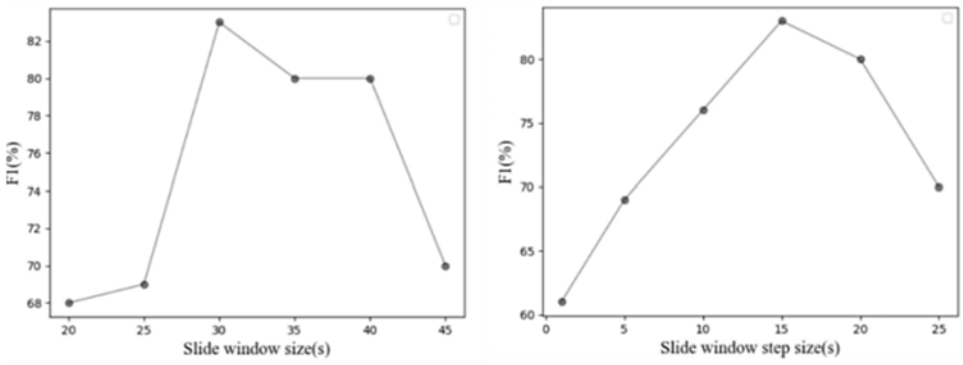

Fig. 9 Comparison of different sliding window size and step size

VI. Conclusion

In this paper, we use sliding window, CWT technique to extract the ECG time-frequency domain features of CHD patients. Then we propose a CHD auxiliary diagnosis model ECGNet that fuses ResNet and Bi-LSTM network, through this model, the ECG signals of patients can be used for CHD auxiliary diagnosis. And this paper conducts experimental comparison based on the publicly available MIMIC-3 ECG database, the experimental results show that the method is better than some traditional image classification methods K-NN, decision tree, VGGNet, ResNet18, ResNet34 and ResNet50. By comparing different sliding window length and step size, it is found that setting the sliding window length as 30 s and step size as 15 s is optimal. And after the comparison of increasing the number of Bi-LSTM layers, it was found that ECGNet is a better model for CHD auxiliary diagnosis.

The deep learning CHD auxiliary diagnosis network designed in this paper based on ECG data analysis can be further improved and refined in subsequent studies. For example, although ECG data are an effective mean to CHD, in practice, doctors often combine physiological indicators such as blood pressure, cholesterol and blood glucose to auxiliary diagnosis. Therefore, in the future, the group will study the auxiliary diagnostic technique of CHD by integrating ECG with other physiological indicators.

References

[1] Benjamin, E. J. et al. (2019). Heart Disease and Stroke Statistics—2019 Update: A Report from the American Heart Association. American Heart Association, 139, 56–528.

[2] Kannel, W. B., Castelli, W. P., Gordon, T. & McNamara, P. M. (1971). Serum cholesterol, lipoproteins, and the risk of coronary heart disease. The Framingham study. Ann Intern Med, 74(1), 1-12.

[3] Irie, F., Iso, H., Sairenchi, T., Fukasawa, N., Yamagishi, K., Ikehara, S., Kanashiki, M. (2006). The relationships of proteinuria, serum creatinine, glomerular filtration rate with cardiovascular disease mortality in Japanese general population. Kidney Int., 69(7), 1264-71.

[4] Burchfiel, C. M., Tracy, R. E., Chyou, P. & Strong, J. P. (1997). Cardiovascular Risk Factors and Hyalinization of Renal Arterioles at Autopsy. Arteriosclerosis, Thrombosis, and Vascular Biology, 17(4), 760–768.

[5] Madani, A., Arnaout, R., Mofrad, M., Arnaout, R. (2018) Fast and accurate view classification of echocardiograms using deep learning. npj Digital Medicine, 1, 6.

[6] Shi Z X, Gu W L. Exploration of TCM syndrome differentiation of coronary heart disease and coronary arteriography (J). Chinese Journal of Integrated Traditional & Western Medicine, 2007, 27(1):76.

[7] Yan Z, Jiang S, Jiao N, et al. The clinical diagnosis effect analysis of electrocardiogram (ECG) and ultrasonic cardiogram (UCG) for coronary atherosclerotic heart disease (CHD) (J). China Modern Doctor, 2015.

[8] Jin, Z., Sun, Y., Cheng, A. C. (2009) Predicting cardiovascular disease from real-time electrocardiographic monitoring: An adaptive machine learning approach on a cell phone. Conf Proc IEEE Eng Med Biol Soc., 6889-92.

[9] Wang Z, Ning X, Du G, et al. Nonlinear Dynamical Characteristics of ECG Signals of CHD Patients (J). Journal of Naijing University (Natural Sciences), 2001.

[10] Cross D S, Mccarty C A, Hytopoulos E, et al. Coronary risk assessment among intermediate risk patients using a clinical and biomarker based algorithm developed and validated in two population cohorts (J). Current Medical Research & Opinion, 2012, 28(11):1819.

[11] Meghan, E, Olesnevich, et al. Serum ferritin levels associated with increased risk for developing CHD in a low-income urban population. (J). Public Health Nutrition, 2012.

[12] Maryam, Tayefi, Mohammad, et al. hs-CRP is strongly associated with coronary heart disease (CHD): A data mining approach using decision tree algorithm (J). Computer Methods & Programs in Biomedicine, 2017.

[13] Karaolis M A, Moutiris J A, Hadjipanayi D, et al. Assessment of the Risk Factors of Coronary Heart Events Based on Data Mining With Decision Trees(J). IEEE Trans Inf Technol Biomed, 2010, 14(3):559-566.

[14] Rajeswari K, Vaithiyanathan D V, Amirtharaj D P. A Novel Risk Level Classification of Ischemic Heart Disease using Artificial Neural Network Technique - An Indian Case Study(J). 2011.

[15] Han X, Liang G. Echocardiographic Features of Patients with Coronary Heart Disease and Angina Pectoris under Deep Learning Algorithms(J). Hindawi Limited, 2021.

[16] Li Yong, He Zihang, Wang Heng, Li Bohan, Li Fengnan, Gao Ying, Ye Xiang. Craftnet: a deep learning ensemble to diagnose cardiovascular diseases. Biomed Signal Process Control 2020;62:102091.

[17] A. E. Awodeyi, S. R. Alty, and M. Ghavami, “Median based method for baseline wander removal in photoplethysmogram signals,” in 2014 IEEE International Conference on Bioinformatics and Bioengineering. IEEE, 2014, pp. 311–314.

[18] Y.-H. Byeon, S.-B. Pan, and K.-C. Kwak, “Intelligent deep models based on scalograms of electrocardiogram signals for biometrics,” Sensors, vol. 19, no. 4, p. 935, 2019.

[19] Wang, Changhong, Low-Power Fall Detector Using Triaxial Accelerometry and Barometric Pressure Sensing (C). IEEE Transactions on Industrial Informatics (2016):1-1. DOI:10.1109/TII.2016.2587761.

[20] Gjoreski, Hristijan, RAReFall — Real-time activity recognition and fall detection system (C). IEEE International Conference on Pervasive Computing & Communications Workshops IEEE, 2014. 7395664, pp.139-145 DOI:10.1109/PerComW.2014.6815182.

[21] Koniusz P, Cherian A, Porikli F. Tensor Representations via Kernel Linearization for Action Recognition from 3D Skeletons (Extended Version) (C). 14th European Conference, ECCV 2016. Vol.9908, pp.37-53 DOI:10.1007/978-3-319-46493-0_3.

[22] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition(C)// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016.

[23] A. E. Johnson, T. J. Pollard, L. Shen, L.-W. H Lehman, M. Feng, M. Ghassemi et al., “Mimic-iii, a freely accessible critical care database,” Scientifific data, vol. 3, no. 1, pp. 1–9, 2016.

[24] B. Moody, G. Moody, M. Villarroel, G. Clifford, I. Silva, "Mimic-iii waveform database matched subset (version 1.0)," 2020. (Online). Available: https://physionet.org/content/mimic3wdb-matched/1.0/
[25] Simonyan K, Zisserman A.Very Deep Convolutional Networks for Large-Scale Image Recognition(J). Computer ence, 2014.