Intra-Patient and Inter-Patient Multi-Classification of Severe Cardiovascular Diseases Based on CResFormer

Dengao Li*, Changcheng Shi, Jumin Zhao, Yi Liu, and Chunxia Li

Abstract: Severe cardiovascular diseases can rapidly lead to death. At present, most studies in the deep learning field using electrocardiogram (ECG) are performed on intra-patient experiments for the classification of coronary artery disease (CAD), myocardial infarction, and congestive heart failure (CHF). By contrast, actual conditions are inter-patient experiments. In this study, we proposed a deep learning network, namely, CResFormer, with dual feature extraction to improve accuracy in classifying such diseases. First, fixed segmentation of dual-lead ECG signals without preprocessing was used as input data. Second, one-dimensional convolutional layers performed moderate dimensionality reduction to accommodate subsequent feature extraction. Then, ResNet residual network block layers and transformer encoder layers sequentially performed feature extraction to obtain key associated abstract features. Finally, the Softmax function was used for classifications. Notably, the focal loss function is used when dealing with unbalanced datasets. The average accuracy, sensitivity, positive predictive value, and specificity of four classifications of severe cardiovascular diseases are 99.84%, 99.68%, 99.71%, and 99.90% in intra-patient experiments, respectively, and 97.48%, 93.54%, 96.30%, and 97.89% in inter-patient experiments, respectively. In addition, the model performs well in unbalanced datasets and shows good noise robustness. Therefore, the model has great application potential in diagnosing CAD, MI, and CHF in the actual clinical environment.

Key words: dual-lead ECG signals; coronary artery disease; myocardial infarction; congestive heart failure; inter-patient experiments; intra-patient experiments; CResFormer

1 Introduction

In general, cardiovascular diseases refer to ischemic or hemorrhagic diseases that are due to hyperlipidemia, thick blood, atherosclerosis, and hypertension, which seriously threaten human life and health. Such diseases include asymptomatic myocardial ischemia, coronary artery disease (CAD), angina pectoris, myocardial infarction (MI), and ischemic heart disease (IHD). Based on the data from the World Health Organization, nearly 17 million deaths are due to cardiovascular diseases[1], which ranks first among all diseases. By 2035, the number of adults with cardiovascular diseases will exceed 130 million[2]. Among cardiovascular diseases, CAD is the occurrence of atherosclerotic lesions in coronary arteries. Fibrous plaques begin to form thick areas on the inner wall of the arteries, which slows down the blood flow to the heart and hinders normal blood circulation[3]. In severe cases, the disease may lead to fatal heart diseases, such as congestive heart failure (CHF), MI, and IHD[4]. In 2014, approximately 26 million adults with CHF were reported worldwide[5, 6]. In 2015, approximately 110 000 Americans died of MI in the United States. The annual incidence of MI

- Dengao Li and Changcheng Shi are with the College of Data Science, Taiyuan University of Technology, Jinzhong 030600, China. E-mail: lidengao@tyut.edu.cn; 2814389394@qq.com
- Jumin Zhao and Yi Liu are with the College of Information and Computer, Taiyuan University of Technology, Jinzhong 030600, China. E-mail: zhaojumin@tyut.edu.cn; yi.liuvip@foxmail.com.
- Chunxia Li is with Shanxi Bethune Hospital, Taiyuan 030032, China. E-mail: 87211232@qq.com.
- *To whom correspondence should be addressed.
  Manuscript received: 2022-03-01; accepted: 2022-03-30

is about 605 000 cases[7]. The characteristics of these typical cardiovascular diseases are as follows: Fast onset, susceptibility to severe, and widespread distribution. Therefore, accurate identification and early treatment of these diseases are necessary.

Electrocardiogram (ECG) is an effective and common tool for the diagnosis of serious cardiovascular diseases. It is non-invasive and intuitive. In the current clinical environment, 12-lead ECG can be used for comprehensive diagnosis. The ECG manifestations of CAD are as follows: Low and flat T waves, ST-segment shifting down or up, and decreasing or disappearing changes of R waves. In general, MI is manifested as ST-segment elevation and arching upward. T waves result in the occurrence of leads facing the ischemic myocardium surrounding the injured area. The Q wave deepens and widens in the infarct zone[8]. Severe MI is marked by ST-segment depression in leads I, II, V4, V5, and V6 and ST-segment elevation in lead aVR. MI can lead to myocardial necrosis, abnormal pumping of the heart, and systemic dysfunction. It can also trigger heart failure that manifests as a period of low ST-segment or QRS wave groups that exhibit low amplitude in ECG[9].

However, the specific interpretation of ECG is generally performed by a specialist physician[10]. The precise interpretation of ECG depends on the training, experience, and maturity of knowledge of the physician[11]. Given the complexity of the ECG signal information, even experts in the field cannot obtain enough information from the ECG signal to ensure an accurate diagnosis[12]. In recent years, many researchers have applied deep learning to processing signals. For example, the use of eight digital signal modulation recognition techniques fused with deep learning applications is more accurate and flexible[13]. However, deep learning algorithms require large batches of datasets for training to form stable and accurate models. For example, digital twins facilitate the expansion of real and valid data[14]. Therefore, a large volume of datasets and deep learning algorithms should be used for computer-aided diagnosis to build effective models and ensure a fast and accurate diagnosis.

2   Related Work

To date, machine-learning-based algorithmic models and deep-learning-based network models have been widely used in the field of autonomous driving, image recognition, machine translation, and signal processing, which have achieved notable results. Tables 1–3 shows

algorithms for the automatic detection and classification of ECG signals, which summarizes studies on the diagnosis of CAD, CHF, and MI. The first stage is the preprocessing of the signal. The signal is preprocessed by denoising to remove interference to the signal, detecting R-peaks, and extracting time-frequency domain features[15, 16, 18, 25, 26, 28, 29, 33, 44, 45], statistical features[20, 24, 33], and morphological features[30]. Second, key features are extracted and classified by machine learning algorithms, deep learning network algorithms, artificial neural networks[27, 34, 35], or various types of novel networks based on basic algorithms, such as typical machine learning algorithms: Support vector machines (SVM)[15–17, 20, 26, 27, 33], K-nearest neighbor algorithms (KNN)[18, 28, 37, 44, 45], random forest algorithm[27]. In addition, typical deep learning network models, deep belief network[17], convolutional neural network (CNN)[21, 22, 30, 32, 36, 41], long short-term memory network (LSTM)[21], and hybrid model of convolutional neural network[38, 46] are used for the extraction of such key features. Synthesizing the literature collected over the last six years in Tables 1–3, the total number of literature collected on diagnostic models for CAD and its complications with a classification accuracy of more than 95% account for nearly 85% of the literature collected. As shown in Table 1, according to Patidar et al.[15], the accuracy, specificity, and sensitivity of normal and CAD models are 99.72%, 99.81%, and 99.63%, respectively. In addition, according to Sudarshan et al.[28], the accuracy, specificity, and sensitivity of normal and CHF models are 99.86%, 99.94% and 99.78%. Moreover, according to Liu et al.[40], the accuracy, specificity, and sensitivity of normal and MI models are 99.90%, 99.54%, and 99.57%, respectively. This result indicates that stable and reliable results have been achieved in two-classification diagnostic models for serious cardiovascular diseases such as CAD. The recent literature on three-classification models for severe cardiovascular diseases is shown in Table 2. The models used by Yang et al.[42, 43] on the three-classification problem of cardiovascular diseases were outstanding; the accuracy, specificity, and sensitivity of their models are 99.89%, 99.93%, 99.82%, and 99.96%, 99.98%, 99.93%, respectively. Such models also had a shorter classification time, allowing for rapid diagnosis. At present, research on the issue of triple classification of severe cardiovascular diseases complicating CAD is also maturing. The literature on the study of four-classification problems, namely, normal, CAD, CHF,

Table 1　Literature studies on two-classifications of normal, CAD, CHF, and MI.

Disease classification	Author	Year	Approach	Accuracy (%)	Specificity (%)	Sensitivity (%)
Normal and CAD	Patidar et al.[15]	2015	Tunable-Q wavelet transform (TQWT) Principal component analysis (PCA) Least squares support vector machine (LS-SVM)	99.72	99.81	99.63
	Kumar et al.[16]	2017	Flexible analytic wavelet transform (FAWT) Cross information potential (CIP), Parameter Least squares support vector machine (LS-SVM)	99.60		
	Altan et al.[17]	2017	Deep belief model and hilbert transform	98.05	96.02	98.88
	Acharya et al.[18]	2017	HOS bispectrum and cumulants K-nearest neighbor (KNN)	98.99	98.50	99.70
	Caliskan and Yuksel[19]	2017	A stacked autoencoder (SAE) network with a Softmax classifier	92.20		
	Dolatabadi et al.[20]	2017	Principal component analysis (PCA), SVM	99.20	100	98.43
	Tan et al.[21]	2018	Long short-term memory (LSTM) Convolutional neural network (CNN)	99.85		
	Acharya et al.[22]	2017	CNN	95.11	95.88	91.13
	Abdar et al.[23]	2019	A new genetic training N2Genetic-nuSVM	93.08		
Normal and CHF	Chandrakar[24]	2015	Detrended fluctuation analysis (DFA) short-term and intermediate-term fractal scaling exponents	98.20		
	Chandrakar[25]	2015	Approximate entropy, sample entropy, permutation entropy, energy entropy	99.90		
	Acharya et al.[26]	2017	Empirical mode decomposition (EMD) Probabilistic neural network (PNN), SVM	97.64	98.24	97.01
	Masetic and Subasi[27]	2016	Random forest, SVM, artificial neural networks (ANN), KNN, C4.5 decision tree classifiers	100.00		
	Sudarshan et al.[28]	2017	Dual tree complex wavelet transform KNN and decision tree (DT) classifiers	99.86	99.94	99.78
	Kumar et al.[29]	2017	FAWT fuzzy entropy, accumulated permutation entropy	98.21		
	Acharya et al.[30]	2019	CNN	98.97	98.87	99.01
	Tripathy et al.[31]	2019	Stockwell (S)-transform and frequency division Sparse representation classifier	98.78	99.09	98.48
	Khade et al.[32]	2019	CNN	88.30		
	Hussain et al.[33]	2020	Linear features (time domain, frequency domain) Statistical and nonlinear features, entropy-based complexity features, wavelet entropy features SVM	93.10		
Normal and MI	Safdarian et al.[34]	2014	T-wave and total integral features Artificial neural network	94.74		
	Kora and Kalva[35]	2015	Bat algorithm Levenberg-marquardt neural network	98.90	92.20	93.34
	Acharya et al.[36]	2017	CNN	95.22	95.49	94.19
	Sharma et al.[37]	2018	Optimal biorthogonal filter bank KNN	99.74		
	Feng et al.[38]	2019	CNN and LSTM model	95.54	86.50	98.20
	Han and Shi[39]	2020	Feature fusion and neural network	99.92		
	Liu et al.[40]	2020	MFB-CBRNN	99.90	99.54	99.97
	Baloglu et al.[41]	2019	CNN	99.78		

and MI, is shown in Table 3. The accuracy, specificity, and sensitivity of the particle swarm optimization-based KNN algorithm studied by Acharya et al.[45] are 99.55%, 99.24%, and 99.93%, respectively. All three metrics exceeded 99%. Zhang et al.[47] used their own multilevel discrete wavelet transform dense network (MDD-Net) to automatically diagnose cardiovascular diseases; the accuracy, specificity, and sensitivity of this network

Table 2   Literature studies on three-classifications of normal, CAD, CHF, and MI.

Disease classification	Author	Year	Approach	Accuacy (%)	Specificity (%)	Sensitivity (%)
Normal, CAD, and MI	Acharya et al.[18]	2017	Discrete cosine transform (DCT) Continuous wavelet transform (CWT) EMD KNN	98.50	99.70	98.50
Normal, CAD, and CHF	Yang et al.[42]	2020	ECG fragment alignment (EFA)-PCA convolutional network (EFAP-Net) Canonical correlation analysis (CCA) PCA	99.89	99.93	99.82
Normal, CAD, and CHF	Yang et al.[43]	2021	Independent component analysis (ICA)-PCA convolutional network A dempster-shafer (D-S) theory-based linear SVM	99.96	99.98	99.93

Table 3   Literature studies on four classifications of normal, CAD, CHF, and MI.

Author	Year	Approach	Accuacy (%)	Specificity (%)	Sensitivity (%)
Fujita et al.[44]	2017	Wavelet packet decomposition Relief KNN	97.98	94.84	99.61
Acharya et al.[45]	2017	Contourlet transform particle swarm optimization KNN	99.55	99.24	99.93
Lih et al.[46]	2020	CNN coupled with LSTM	98.51	97.89	99.30
Zhang et al.[47]	2020	Multilevel discrete wavelet transform densely network (MDD-Net)	99.74	99.83	98.67

are 99.74%, 99.83%, and 98.67%, respectively. In addition, this network handles unbalanced data well, and it is more flexible in handling the study of four-classification problems of severe cardiovascular diseases. Therefore, technical algorithms for the diagnosis of severe cardiovascular diseases can allow for the rapid and comprehensive diagnosis of multiple diseases.

The above-mentioned results show that the classification of severe cardiovascular disease diagnoses has shown a stable effect that can be applied in practice. However, a number of issues must be addressed. First, severe cardiovascular diseases are easily confused in individual presentations, resulting in misdiagnosis of diseases. Second, the vast majority of the above-mentioned studies have been conducted on the basis of intra-patient experiments. Intra-patient experiments, which train and test models using the same patient's heartbeats, can achieve excellent performance by learning the characteristics of each patient through multiple training phases[48]. Under realistic conditions, the model must predict and handle patients in a real environment. Patient heartbeats cannot be used as training data. The inter-patient paradigm suggests that the training and test datasets have heartbeats from different patients[49]. Third, the vast majority of researchers use open data without evident noise and other interfering signals for their experiments. The

ECGs are rarely distorted or blurred, which can be tested directly to achieve better results. However, the actual ECGs collected from patients are subject to varying degrees of noise interference, and the ECG data content is diverse. Fourth, the number of people with severe cardiovascular diseases in the real-world setting is still small, and the reality is complex. The degree of unbalance in the data required remains unpredictable. Fifth, the ECG data used in models of severe cardiovascular diseases are single-lead data. However, the single-lead ECG data measured do not indicate the patient's condition because such data are susceptible to interference.

In addressing the above-mentioned issues, we set a validated four-classification model of severe cardiovascular diseases for the effective diagnosis of CAD, MI, and CHF to reduce misdiagnosis of patients. We trained and tested models using the ECG information obtained from multiple leads to enhance the comprehensiveness and certainty of the information. Here we performed a combined input operation of the two ECG information leads. Based on the literature[47], inter-patient diagnosis on the issue of four classifications of severe cardiovascular diseases still needs to be improved. This paper proposes a CResFormer (CR-Former) deep learning hybrid network structure, which utilizes a dual extraction of features. Network's residual

extraction and full-domain intercorrelation of features can extract and effectively bridge key features, thereby reducing the loss of key features, enhancing the mutual information among features, and improving the accurate discrimination of data. In particular, the network was structured in two phases: A feature extraction phase and a classification phase.

First, we built a simple convolutional pooling operation using the unique advantages of CNNs in processing high-dimensional data to perform the corresponding feature extraction and some degree of dimensionality reduction in the feature extraction phase. Then, the network constructed multiple connections of residual blocks to obtain a richer set of key features based on the unique feature of the Resnet residual network, retain the original information of the extracted features, and fused it with the extracted feature information. Finally, the transformer encoder feature extraction unit was constructed to enhance the tightness of the mutual information association among key features[50]. The extracted feature phase of this paper shows a fusion of a convolutional neural network, ResNet residual network blocks, and transformer encoders.

Second, we apply the Softmax function as the final classification function to obtain four precise decision values in the classification phase. The advantages of this network structure are as follows:

(1) The CR-Former network utilizes a dual extraction of features, which has achieved good results for inter-patient tests.

(2) The network uses multi-lead data input to respond to all aspects of data, further improving the diagnosis of diseases among patients.

(3) The network can ignore upfront data preprocessing operations, and it has excellent noise robustness in processing data with different noise disturbances without noise processing.

(4) The network applies the focal loss function to reduce the internal sample weights and achieve excellent results when dealing with unbalanced datasets.

3 Materials

3.1 Database used

The ECG signal data used in this study were obtained from a public database (PhysioBank)[51]. In the database storage network, we downloaded four public databases of ECGs on normal, CAD, MI, and CHF as our experimental data, namely, MIT-BIH Normal Sinus Rhythm Database (NSRDB)[51], St Peters-burg INCART 12-lead Arrhythmia Database (INCARTDB)[51], PTB Diagnostic ECG Database (PTBDB)[51], and Beth Israel Deaconess Medical Centre (BIDMC) Congestive Heart Failure Database[51]. Their ECG signal map segments are shown in Fig. 1, and the specific information of the ECG recordings in the database is shown in Table 4.

3.2 Datasets division

The anthropometric sites of the ECG1 and ECG2 leads of the MIT-BIH NSRDB and BIDMC CHFDB are consistent with the II and V1 leads of the St INCARTDB
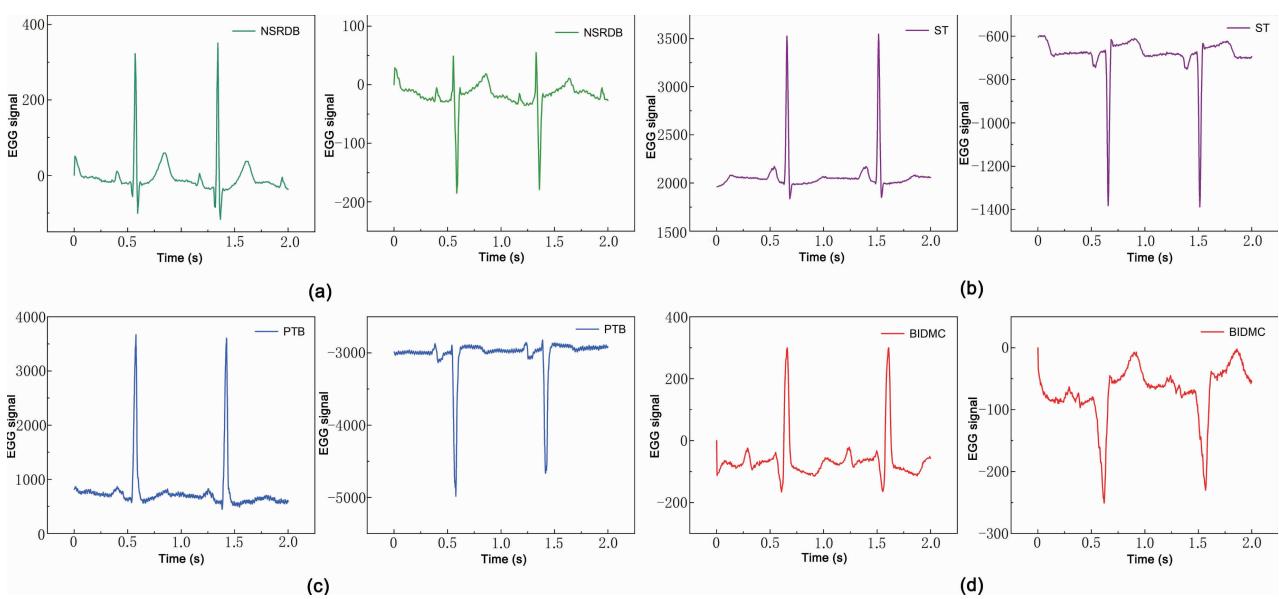

Fig. 1 ECG signal waveform obtained from two leads. (a) NSRDB; (b) St INCARTDB; (c) PTB; (d) BIDMC.

Table 4   Information of four databases of ECGs used.

Database	Diagnostic type	Lead	Sampling frequency (Hz)	Number	Duration of each case	Specific case
MIT-BIH NSRDB	Normal	ECG1, ECG2	128	18	24 h	16265, 16272, 16273, 16420, 16483, 16539, 16773, 16786, 16795, 17052, 17453, 18177, 18184, 19088, 19090, 19093, 19140, 19830
St INCARTDB	CAD	I, II, III, aVR, aVL, aVF V1, V2, V3, V4, V5, V6	257	7	30 min	I01, I02, I20, I21, I22, I35, I36, I37, I38, I39, I44, I45, I46, I57, I58, I72, I73
PTBDB	MI	I, II, III, AVR, AVL, AVF V1, V2, V3, V4, V5, V6 VX, VY, VZ	1000	148	2 min	001–103, 108, 111, 120, 128, 135, 138–142, 145, 148, 149, 152, 158, 160, 163, 183, 189, 193, 195, 197, 205, 207, 211, 223, 226, 230, 231, 259, 261, 265, 268, 270, 273, 274, 280, 282, 283, 287, 290–294
BIDMC CHFDB	CHF	ECG1, ECG2	250	15	20 h	chf01, chf02, chf03, chf04, chf05, chf06, chf07, chf08, chf09, chf10, chf11, chf12, chf13, chf14, chf15

and PTBDB. Therefore, their lead data were used as input data for experiments. The paper was divided into several datasets, including division of two-classification experimental datasets, three-classification experimental datasets, and four-classification experimental datasets, to verify the effectiveness of the network model in dealing with multi-classification problems. Nine datasets (A–I) were divided for the demonstration of the experiments. The details are shown in Tables 5–8.

(1) Datasets A–F and G–I were used for intra-patient and inter-patient classification experiments for normal, CAD, MI, and CHF, respectively.

(2) We selected the same data points from each ECG signal to construct balanced Datasets A–C. In addition,

Table 5   Information of balanced datasets for intra-patient experiments.

Diagnostic type	Number of datasets		
	A	B	C
Normal	20 000	20 000	20 000
CAD	–	20 000	20 000
MI	–	–	20 000
CHF	20 000	20 000	20 000

Table 6   Information of unbalanced datasets for intra-patient experiments.

Diagnostic type	Number of datasets		
	D	E	F
Normal	35 000	35 000	35 000
CAD	–	$35\,000/N_1$	$35\,000/N_1$
MI	–	–	$35\,000/N_2$
CHF	$35\,000/N_3$	$35\,000/N_3$	$35\,000/N_3$

Table 7   Information of training datasets for inter-patient experiments.

Diagnostic type	Training ID	Number of datasets		
		G	H	I
Normal	16265, 16272, 16420, 16773, 16795, 17052, 18177, 19088, 19093, 19830	25 000	25 000	25 000
CAD	I01, I20, I35, I39, I44, I45, I72	–	17 500	17 500
MI	033, 140, 152, 193, 211, 268	–	–	15 000
CHF	chf01, chf02, chf04, chf09, chf10, chf12	15 000	15 000	15 000

Table 8   Information of test datasets for inter-patient experiments.

Diagnostic type	Training ID	Number of datasets		
		G	H	I
Normal	16273, 16483, 16539, 16786, 17453, 18184, 19140	17 500	17 500	17 500
CAD	I21, I22, I38, I58, I73	–	15 000	15000
MI	56, 149, 205, 231, 283	–	–	15 000
CHF	chf03, chf06, chf08, chf14	10 000	10 000	10 000

we must compose unbalanced Datasets D–F to validate the performance of the model in handling unbalanced datasets. In particular, retaining the normal ECG points unchanged in Datasets D–F, we scaled the representation of ECG data points for CAD, MI, and CHF. Proportion N (N_1, N_2, and N_3) denotes the proportion of imbalances,

and *N* is an integer greater than 1. Datasets G–I have the same number of ECG datasets for each normal and patient ECG signal. This results are shown in Tables 5–8.

4 Method

4.1 Data input

In this paper, the ECG signals from two leads are used as input to verify the diagnostic rate of multi-lead data for diseases. The input data are presented as follows:

$$X = [x_{1,1}, x_{1,2}, \ldots, x_{1,M}, x_{2,1}, x_{2,2}, \ldots, x_{2,N}] \quad (1)$$

4.2 CR-Former model

In this paper, the required processing method is shown in Fig. 2. We do not need to perform signal processing operations such as denoising of ECG signals, but we must segment ECG signals from two leads and input them into the model. The details are shown as follows.

4.2.1 Feature extraction

(1) CResnet feature extractor

In addressing the large feature dimensionality of the input data, this paper used CNN to process high-dimensional data to construct a one-dimensional convolutional pooling layer. The data were dimensionally reduced to ensure that most of the key features are present. The number of layers in the convolutional network was too deep, which leads to low processing results. Appropriate addition of residual structure can deepen the network and improve the classification results when the convolutional layers were effectively expanded.

The original ResNet network was proposed by He et al.[52] to address deep layers of convolutional neural networks leading to gradient explosion. This approach caused the network not to converge easily, making it prone to network computation errors. The basic structure of Resnet is shown in Fig. 3, noting that its overall structure consists of a convolutional and pooling layer, four residual blocks, and a fully connected layer. First, the data were fed through a convolutional and pooling layer, which applies convolutional operations for feature extraction and pooling operations for feature reduction. Then, in the ResNet residual block structure, each residual block had the same size convolutional kernels, and it consisted of a different number of convolutional layers, with the same number of channels of convolutional kernels in each residual block. In the structure, the input features in each residual block were halved, whereas the number of filter channels in
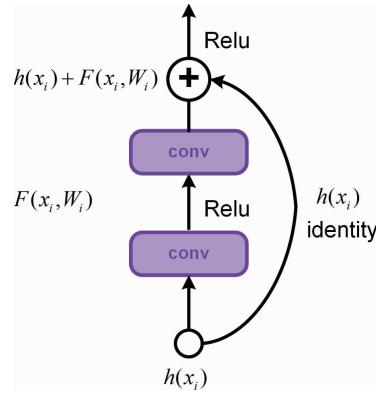

**Fig. 3   Diagram of basic structure of ResNet.**
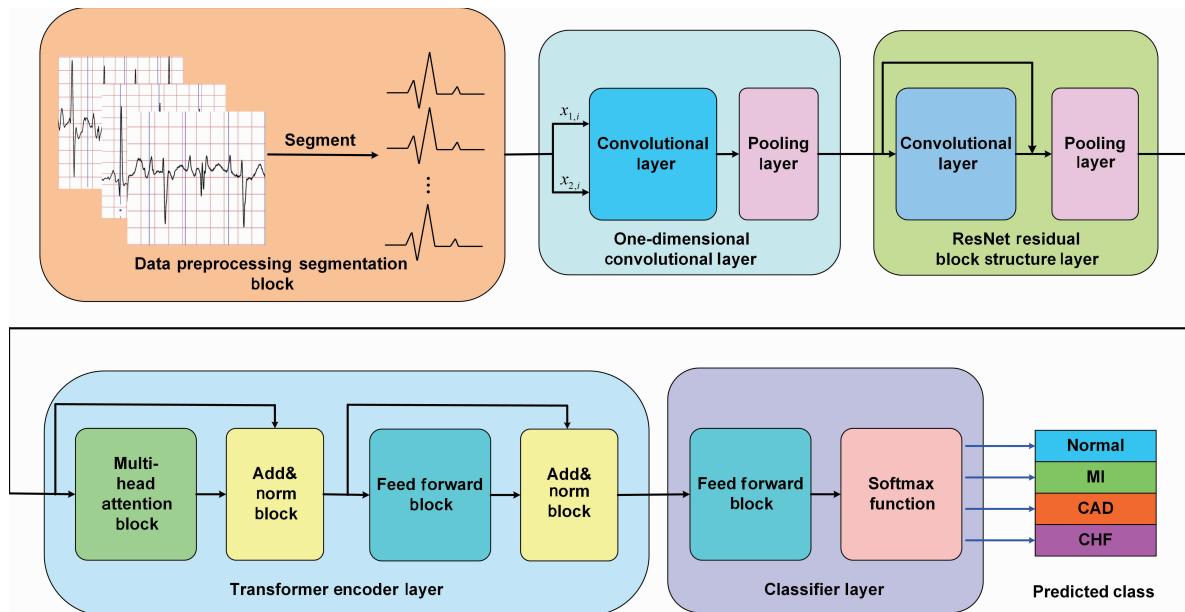
Fig. 2   Flow chart of severe cardiovascular diseases classification.
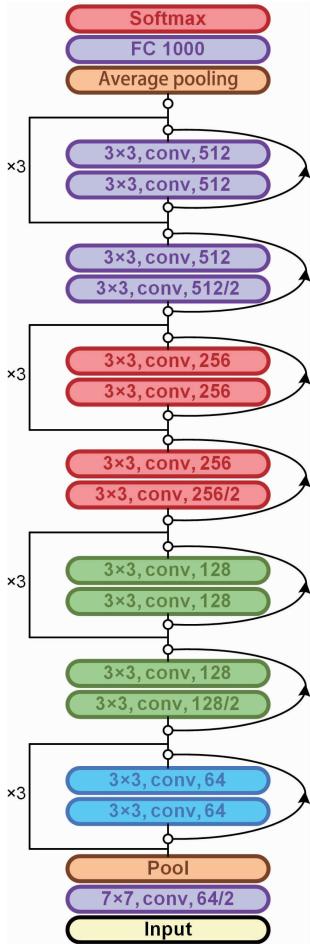
Fig. 4   Diagram of basic structure of ResNet residual block. Conv means convolution.

the convolutional layer was doubled. Finally, a global averaging pooling operation was performed, and the fully connected layer was concatenated to produce results.

The structure of each residual block is shown in Fig. 4. This structure is the core idea proposed by He et al.[52]. Equations (2) and (3) are mathematical representations of the residual units.

$$y_i = h(x_l) + F(x_i, W_i) \qquad (2)$$
$$x_{i+1} = f(y_i) \qquad (3)$$

where x_i and x_{i+1} denote the input and output of the i-th residual unit, respectively; f denotes the residual function, which represents the learned residuals. Each residual unit contains multiple convolutional layers, batch normalization layers[53], and the ReLU function. Notably, the batch normalization layers can prevent gradient disappearance or explosion problems because of the excessive number of layers. $h(X_i) = X_i$ denotes constant mapping, and f denotes the ReLU activation function after the final summation of residuals. Based on

Eq. (4), we can derive the learning characteristics from a shallow layer i to a deep layer l.

$$x_l = x_i + \sum_{k=1}^{l-1} F(x_k, W_k) \qquad (4)$$

The great advantage of the Resnet network structure is the residual structure. The residual connectivity of this structure allows the input of the previous layer to be added to the residual features, enriching the input features of the next layer. This paper combines the advantages of dimensionality reduction of convolutional neural networks and feature fusion of Resnet residual networks. A CResnet feature extractor is constructed to enrich the key features.

(2) Transformer feature extractor

After the primary feature extraction using the CResnet feature extractor, we proceeded with the secondary feature extraction using the Transformer feature extractor. The Transformer model has parallelizability and global dependency for processing sequence data. In addition, the model can focus on the mutual part of clinical problems[54], particularly for the processing of ECG signal features. In the construction of the transformer feature extractor, the encoder part of the transformer model must be considered because of the direct predictive nature of ECG signals for the specific unit structure (Fig. 5).

(i) Input embedding: Based on the transformer
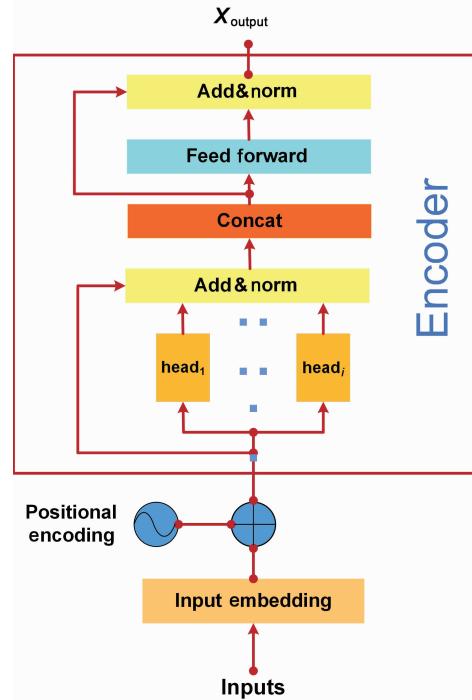
Fig. 5   Diagram of structure of transformer model encoder unit.

model's reference in the numerical domain of each word mapping in machine translation, we used the ECG feature values extracted by the CResnet feature extractor as the size d_{model} of the input dimension after tiling. We must obtain the position encoding of each point and overlay it with the embedding information to obtain the specific information of each point with location information. Position embedding (PE) can be characterized by the periodicity of the sine and cosine functions to obtain the sequence length relative position information, which is encoded by the following equations:

$$PE(\text{pos}, 2i) = \sin\left(\text{pos}/10\,000^{2i/d_{\text{model}}}\right) \quad (5)$$

$$PE(\text{pos}, 2i+1) = \cos\left(\text{pos}/10\,000^{2i/d_{\text{model}}}\right) \quad (6)$$

As shown in the above-mentioned Eqs. (5) and (6), the location function embedding period varies from 2π to $10\,000 \times 2\pi$ orders of magnitude, and each location can obtain different combinations of values of sin and cos functions with different periods in the embedding dimension, yielding unique location texture information. Such combinations result in position dependencies and temporal information.

(ii) Multi-headed attention: Multi-headed attention can describe the complete relationship of multiple dependencies, focusing on representational information at different locations. Assuming that the input matrix for the transformer encoder unit is $Y \in \mathbf{R}^{d_{\text{model}} \times d_{\text{model}}}$, we can represent the query matrix, key matrix, and value matrix as Q, K, and V, respectively, which are equivalent to Y. By performing the partitioning operation on each matrix, we obtained $Q_i = QW_i^Q$, $K_i = KW_i^K$, and $V_i = VW_i^V$, where $W_i^Q \in \mathbf{R}^{d_{\text{model}} \times d_k}$ (the parameter matrix of query matrix Q), $W_i^K \in \mathbf{R}^{d_{\text{model}} \times d_k}$ (the parameter matrix of key matrix K), $W_i^V \in \mathbf{R}^{d_{\text{model}} \times d_k}$ (the parameter matrix of value matrix V). Each head was calculated in accordance with Eq. (7).

$$\text{Head}_i = \text{soft max}\left(\frac{Q_i K_i^{\text{T}}}{\sqrt{d_k}}\right) V_i \quad (7)$$

The attention query was calculated and mapped to its own key value matrix to obtain the similarity among the points. Then, the Softmax function was applied to calculate the weights, and the corresponding weights were updated for each point to obtain the attention matrix for each head. Multiple heads performed different linear mappings separately to further control the attention weights of multiple messages, and each calculated head was concatenated to map the corresponding results. This process is shown in Eq. (8):

$$Z = \text{Concat}(\text{head}_1, \text{head}_2, \ldots, \text{head}_l)W^{\circ} \quad (8)$$

where $W^{\circ} \in \mathbf{R}^{d_{\text{model}} \times d_{\text{model}}}$ (the parameter matrix for concating of multi-headed attention matrices).

(iii) Residual linking: The original input location embedding was added to the specific information of the location weights after the multi-headed attention operation to obtain more accurate information about the interdependent attention mechanism. It also prevented over-fitting caused by an excessive number of layers, as shown in Eq. (9).

$$X = Y + Z \quad (9)$$

(iv) Forward propagation: Two fully connected layers were used. The middle layer was a linear correction unit, which reduced the original feature information.

4.2.2 Model composition

The basic architecture of our proposed CR-Former model is illustrated in Fig. 6. In this paper, the input of the model is ECG signal segments. The ECG signals were moderately reduced in dimensionality by passing through two convolutional layers, two batch normalization layers, two ReLU activation layers, and two maximum pooling layers. Then, we constructed residual block structures based on the residual structure of the ResNet model. We also applied four residual blocks. After each residual operation, the dimension size remained the same, but the number of channels changed. A maximum pooling operation was applied after each residual operation, which allowed the construction of a deeper model to extract more critical features. After the last residual operation, we applied a global average pooling operation and perform tiling of data features. We fed the features extracted by the CResnet feature extractor into the transformer encoder unit and performed the same transformer encoder operation three times. Finally, a fully connected layer and a softmax function layer were used to obtain the corresponding classification results for heart failure-related diseases. The specific model parameters are shown in Table 9.

5 Model Evaluation

5.1 10-fold cross-validation

In this paper, we use 10-fold cross-validation for intra-patient experimental validation[55]. In each epoch, we randomly selected 10% of the data as the test data and 90% of the data as the training validation data. Within this 90% training validation set, we used 80% of the

Input
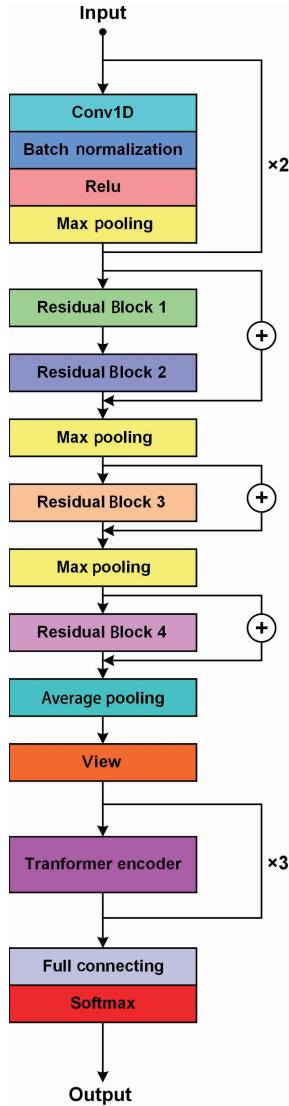
Fig. 6 Diagram of basic architecture of proposed CR-Former model.

data as the training dataset and the remaining 20% as the validation set. The specific 10-fold cross-validation scheme is shown in Fig. 7.

5.2 Evaluation indicators

We applied the confusion matrix to obtain the results. Then, widely prevalent evaluation metrics (accuracy A_{cc}, specificity S_{pe}, sensitivity S_{en}, precision P_p, F1 score, overall accuracy (OA) were calculated on the basis of the confusion matrix results to assess the performance of classifications, as follows:

$$A_{cc} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \times 100\% \quad (10)$$

$$S_{pe} = \frac{\text{TN}}{\text{TN} + \text{FP}} \times 100\% \quad (11)$$

$$S_{en} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100\% \quad (12)$$

$$P_p = \frac{\text{TP}}{\text{TP} + \text{FP}} \times 100\% \quad (13)$$

$$\text{F1} = 2 \times \frac{P_p \times S_{en}}{P_p + S_{en}} \times 100\% \quad (14)$$

$$\text{OA} = \frac{\text{Correctly classified instances}}{\text{Total number of instances}} \times 100\% \quad (15)$$

where the confusion matrix records the number of true positives, false positives, true negatives, and false negatives, and OA is the proportion of correct classifications to all classifications, true positive (TP), false positive (FP), true negative (TN), and false negative (FN) are defined as follows: True positive (TP) is a class where the prediction is positive and the actual is also positive, false positive (FP) is a class where the prediction is positive and the actual is negative, true negative (TN) is a class where the prediction is negative and the actual is positive, and false negative (FN) is a class where the prediction is negative and the actual is also negative. Accuracy (A_{cc}) is the proportion of correct classifications of the current class to the current class. Specificity (S_{pe}) is the proportion of samples that are judged to be in the negative category out of those that are actually in the negative category. Sensitivity (S_{en}) is the proportion of the sample that is judged to be in the positive category out of those that are actually in the positive category. F1 score is the weighted average of the precision and recall rates. The maximum value of this score is 1 and the minimum value is 0.

6 Results

6.1 Experimental deployment

We conducted the experiment on a computer server with a 9th generation Core i7-9800X CPU, 16 GB of running memory, and four GeForce 20 RTX 2070 GPUs from NVIDIA. In this study, we described other details of some of the parameters of the experiments. In this study, Xavier initialization was used to initialize the weights of the model[56]. The cross-entropy function was used as the loss function of the network on the balanced dataset, and the focal loss function was used as the loss function on the unbalanced dataset. Using the Adam optimization algorithm, the learning rate was adjusted to 2×10^{-5} to achieve a fast data convergence process. During the iterative process of the model, 50 batches were used for each iteration of the operation. These parameter functions were adjusted to achieve the best results.

6.2 Results of input data length comparison

A small amount of ECG signals processed in a single

Table 9   Parameters of CR-Former model.

Layer	Operation	Filter size	Number of filters	Stride length	Numbe of heads	d_{inner}	Output size
0	Input	–	–	–	–	–	1000×1
1	Conv1D	16	4	1	–	–	985×4
2	Batch normalization	–	–	–	–	–	985×4
3	Max pooling	2	4	2	–	–	492×4
4	Conv1D	8	6	1	–	–	485×6
5	Batch normalization	–	–	–	–	–	485×6
6	Max pooling	2	6	2	–	–	242×6
7	ResNet1	8	8	1	–	–	242×8
8	ResNet2	16	10	1	–	–	242×10
9	Max pooling	2	10	2	–	–	121×10
10	ResNet3	16	12	1	–	–	121×12
11	Max pooling	2	12	2	–	–	60×12
12	ResNet4	4	14	1	–	–	30×14
13	Average pooling	2	14	2	–	–	30×14
	View	–	–	–	–	–	420
14	Transformer encoding	–	–	–	5	248	420
15	Transformer encoding	–	–	–	5	248	420
16	Transformer encoding	–	–	–	5	248	420
17	Full connecting	–	–	–	–	–	248
18	Softmax	–	–	–	–	–	4

Fig. 7   Training and test methods for 10-fold cross-validation.

session may result in the model capturing fewer key features, and a larger amount of ECG signals processed in a single session will result in many distracting factors. Thus, inaccurate features may be captured. In this paper, we used different initial signal lengths to systematically show their importance to the performance of the model. If we select a large input feature, then the hyperparameters in the Transformer feature extractor in this model will increase dramatically and cause the model to run slowly. Therefore, a length of 500 to 3000 was selected for our experiments. Tables 10 and 11 show the experimental results with different signal lengths for the four classifications of severe cardiovascular diseases. The highest values of accuracy and F1 scores were obtained for each classification with an initial signal length of 1000, indicating that a single signal with an initial length of 1000 is suitable for the model.

6.3   Results of intra-patient experiments

6.3.1   Results of intra-patient balanced datasets

The experimental results for two-, three-, and four-classification diagnoses for the intra-patient balanced

Table 10   Experimental results (accuracy) of different signal lengths for the four disease classifications.

Disease classification	Accuracy (%)				
	500	1000	1500	2000	3000
Normal	99.80	99.88	99.76	99.78	99.71
MI	99.72	99.82	99.68	99.70	99.66
CAD	99.74	99.82	99.70	99.68	99.66
CHF	99.78	99.86	99.72	99.70	99.68

Table 11   Experimental results (F1 score) of different signal lengths for the four disease classifications.

Disease classification	F1 score (%)				
	500	1000	1500	2000	3000
Normal	99.68	99.76	99.64	99.65	99.58
MI	99.54	99.66	99.52	99.54	99.50
CAD	99.51	99.64	99.47	99.49	99.44
CHF	99.64	99.71	99.62	99.58	99.54

datasets are shown in Fig. 8. The accuracy of each classification exceeds 99.8%. The accuracy, sensitivity, specificity, and F1 score for identifying CHF in the two-classification diagnosis (Fig. 8a) are 99.94%, 99.92%, 99.95%, and 99.93%, respectively, with all metrics exceeding 99.9%. The sensitivity, precision, specificity, and F1 scores for identifying CAD (Fig. 8b) are 99.79%, 99.75%, 99.88%, and 99.77%, respectively. As for identifying CHF, the sensitivity, precision, specificity and F1 scores reaches 99.85%, 99.82%, 99.91%, and 99.83%. Figures 8c and 8d show the confusion matrix and experimental results among normal, CAD, MI, and CHF. In Fig. 8d, the lowest value

(a) Two-classification indicators for normal and CHF on intra-patient experiments

(b) Three-classification indicators for Normal, CAD, and CHF on intra-patient experiments

(c) Four-classification confusion matrix for Normal, MI, CAD, and CHF on intra-patient experiments

(d) Four-classification indicators for Normal, MI, CAD, and CHF on intra-patient experiments
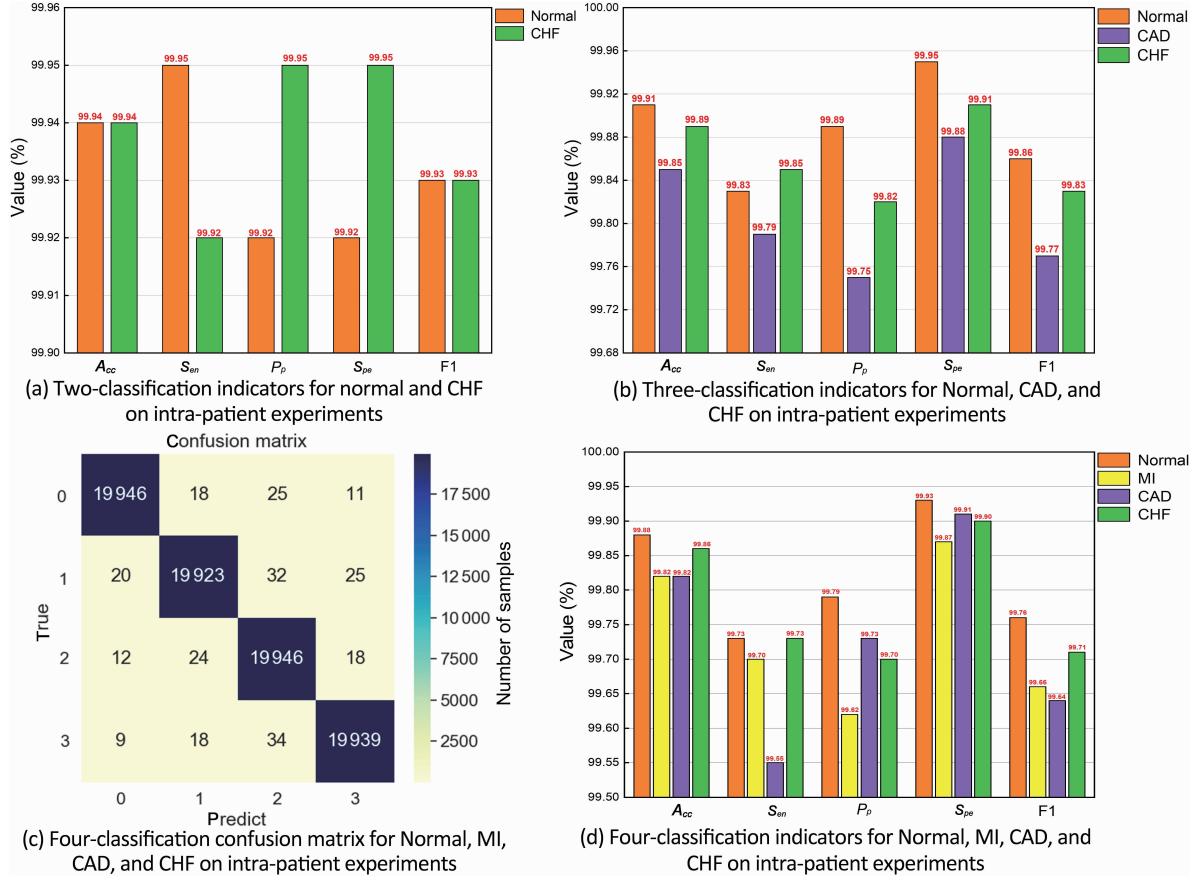
Fig. 8 Results of the multi-class diagnosis experiments on intra-patient balanced datasets.

of 99.55% and the highest value of 99.93% are found for all evaluation indicators, with an overall indicator value of around 99.6%. Thus, the model proposed in this paper shows excellent performance in differentiating among normal, CAD, MI, and CHF for two-, three-, and four-classification diagnoses in intra-patient experiments.

6.3.2 Results of intra-patient unbalanced datasets

In this paper, the specific N_1, N_2, and N_3 assigned experimental groups are shown in Table 12. In addition, the experimental results of the intra-patient unbalanced dataset are shown in Table 13. In Table 13, the values for all metrics exceed 99%, and they do not fluctuate by more than 1% in the experiments for different balanced indices. The maximum and minimum values of the F1 score are 99.93% (Group 4) and 99.83%

Table 12 Proportional distribution of intra-patient unbalanced datasets of the five groups for experiments.

N_i	Value of N_i				
	Group 1	Group 2	Group 3	Group 4	Group 5
N_1	2	5	7	10	14
N_2	5	7	10	14	20
N_3	10	14	5	2	7

(Groups 1 and 2) for the two-classification diagnostic Dataset D. The maximum and minimum value of the F1 score is 99.75% (Group 3) and 99.70% (Group 2) for the three-classification diagnostic Dataset E, respectively. In particular, in the four-classification diagnostic Dataset F for severe cardiovascular diseases, the diagnostic CAD results have a maximum and minimum sensitivity of 99.52% (Group 1) and 99.23%

Table 13 Experimental results of intra-patient unbalanced datasets of the five groups for experiments.

(%)

		Group 1				Group 2				Group 3				Group 4				Group 5			
		S_{en}	P_p	S_{pe}	F1	S_{en}	P_p	S_{pe}	F1	S_{en}	P_p	S_{pe}	F1	S_{en}	P_p	S_{pe}	F1	S_{en}	P_p	S_{pe}	F1
D	CHF	99.74	99.91	99.93	99.83	99.72	99.89	99.94	99.83	99.87	99.92	99.94	99.91	99.92	99.94	99.94	99.93	99.84	99.93	99.92	99.88
E	CAD	99.71	99.84	99.78	99.47	99.56	99.77	99.84	99.70	99.74	99.78	99.76	99.75	99.65	99.76	99.82	99.73	99.54	99.72	99.90	99.72
F	CAD	99.52	99.68	99.54	99.53	99.46	99.73	99.57	99.51	99.48	99.75	99.55	99.51	99.34	99.78	99.70	99.52	99.23	99.72	99.68	99.45
	MI	99.66	99.64	99.48	99.57	99.61	99.56	99.52	99.56	99.59	99.66	99.54	99.56	99.46	99.60	99.54	99.49	99.32	99.62	99.72	99.52

(Group 5) and a maximum and minimum F1 score of 99.53% (Group 1) and 99.45% (Group 5), respectively. In addition, the diagnostic MI results have a maximum and minimum sensitivity of 99.66% (Group 1) and 99.32% (Group 5) and a maximum and minimum F1 score of 99.57% (Group 1) and 99.49% (Group 4), respectively. The above-mentioned results indicate that the model proposed in this paper can handle unbalanced data well, and it has an accurate and stable performance in distinguishing normal, CAD, MI, and CHF in two-, three-, and four-classification diagnoses.

6.4 Results of inter-patient experiments

Figure 9 shows the inter-patient experimental results of two-, three-, and four-classification diagnoses of severe cardiovascular diseases. Figure 9a shows that all indicators exceed 99.5%, 94.2%, and 90% for two-, three-, and four-classification diagnoses, respectively. In particular, the sensitivity, specificity, and F1 scores of the method are 99.74%, 99.84%, and 99.79% for the two-classification diagnosis of CHF; 99.12%, 96.81%, and 97.95% for the three-classification diagnosis of CHF; and 99.12%, 96.81%, and 97.95% for the four-

classification diagnosis of CHF, respectively. In addition, the sensitivity, specificity, and F1 scores of the method are 97.48%, 97.89%, and 93.96%, respectively. By contrast, the sensitivity, specificity, and F1 scores of the method for the diagnosis of CAD are 96.03%, 97.72%, and 94.74%, respectively. Furthermore, the overall accuracy (OA) in the four-classification diagnostic experiment is 94.97%, with only 5.03% of heartbeats being incorrectly classified. Therefore, based on the above-mentioned results, the model can obtain excellent inter-patient experimental results for two-, three-, and four-classification diagnoses of severe cardiovascular diseases.

6.5 Results of multi-level noise interference signals

We added multilevel noise interference to the ECG signal to verify the noise robustness of the model proposed in this paper. Figure 10 shows the waveforms of the normal, CAD, MI, and CHF ECG signals with added noise used in this paper. The figure shows the ECG signal waveforms acquired from two different leads, and all signal noise is added by using MATLAB. As can be seen from the signal curves in Fig. 10, noise has a


(a) Two-classification indicators for normal and CHF on inter-patient experiments

(b) Three-classification indicators for normal, CAD, and CHF on inter-patient experiments

(c) Four-classification confusion matrix for Normal, MI, CAD, and CHF on inter-patient experiments

(d) Four-classification indicators for Normal, MI, CAD, and CHF on inter-patient experiments
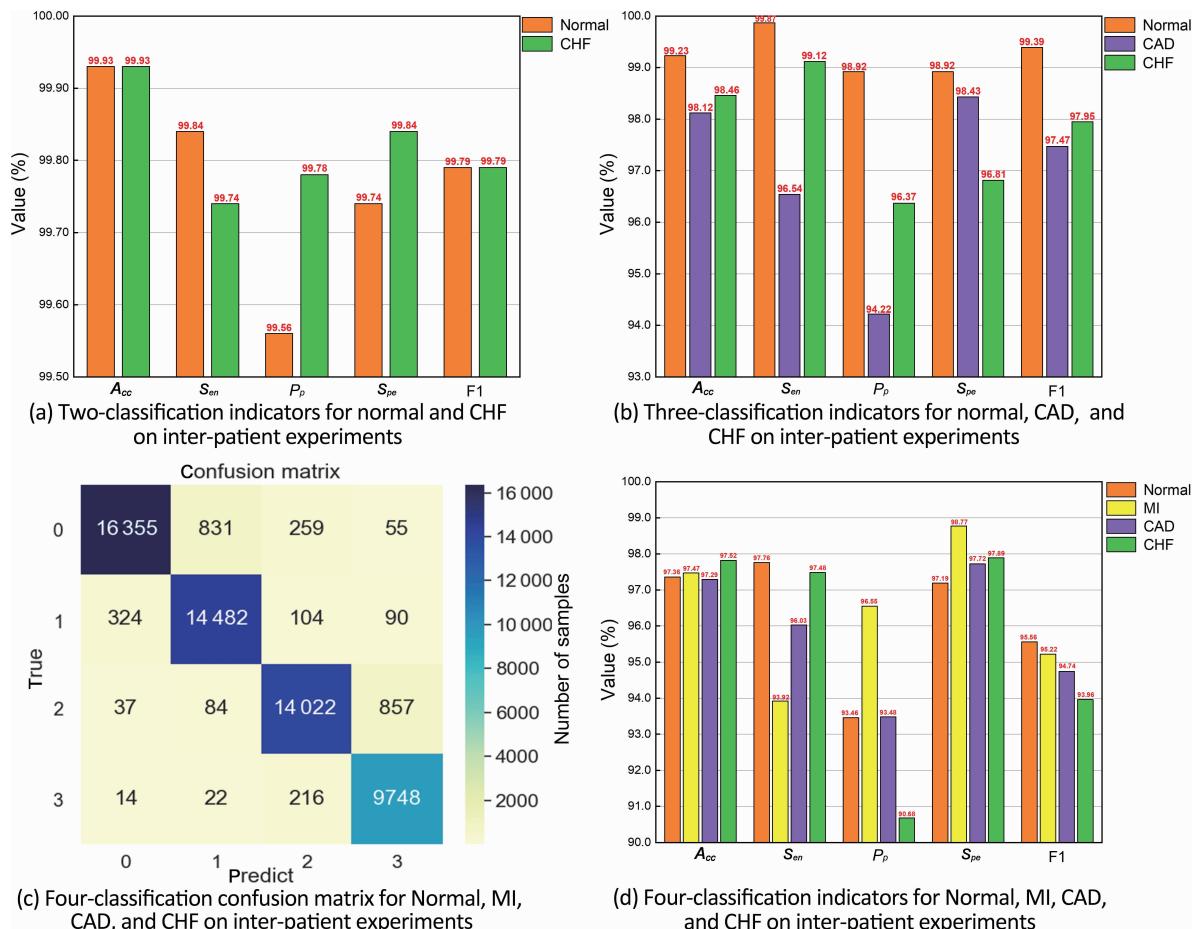
Fig. 9   Experimental results of the inter-patient datasets of two-, three-, and four-classification diagonses of diseases.

large effect on signal interference. In particular, the ECG signal waveform is severely disturbed at signal-to-noise ratios (SNR) below 18 dB. Figure 11 and Table 14 show the experimental results for Sets A, B, C, G, H, and I with different SNRs of noise interference. Based on the results shown in Fig. 11 and Table 14, the OA obtained

decreases with the increase of the SNR for the same categories. A large difference in the OA values for an SNR of 6 dB and 0 dB is found, with a minimum and maximum difference of 5.15% and 9.9%, respectively. This data difference indicates that the noise has a large impact on the ECG signal waveform at SNR of 0 dB,
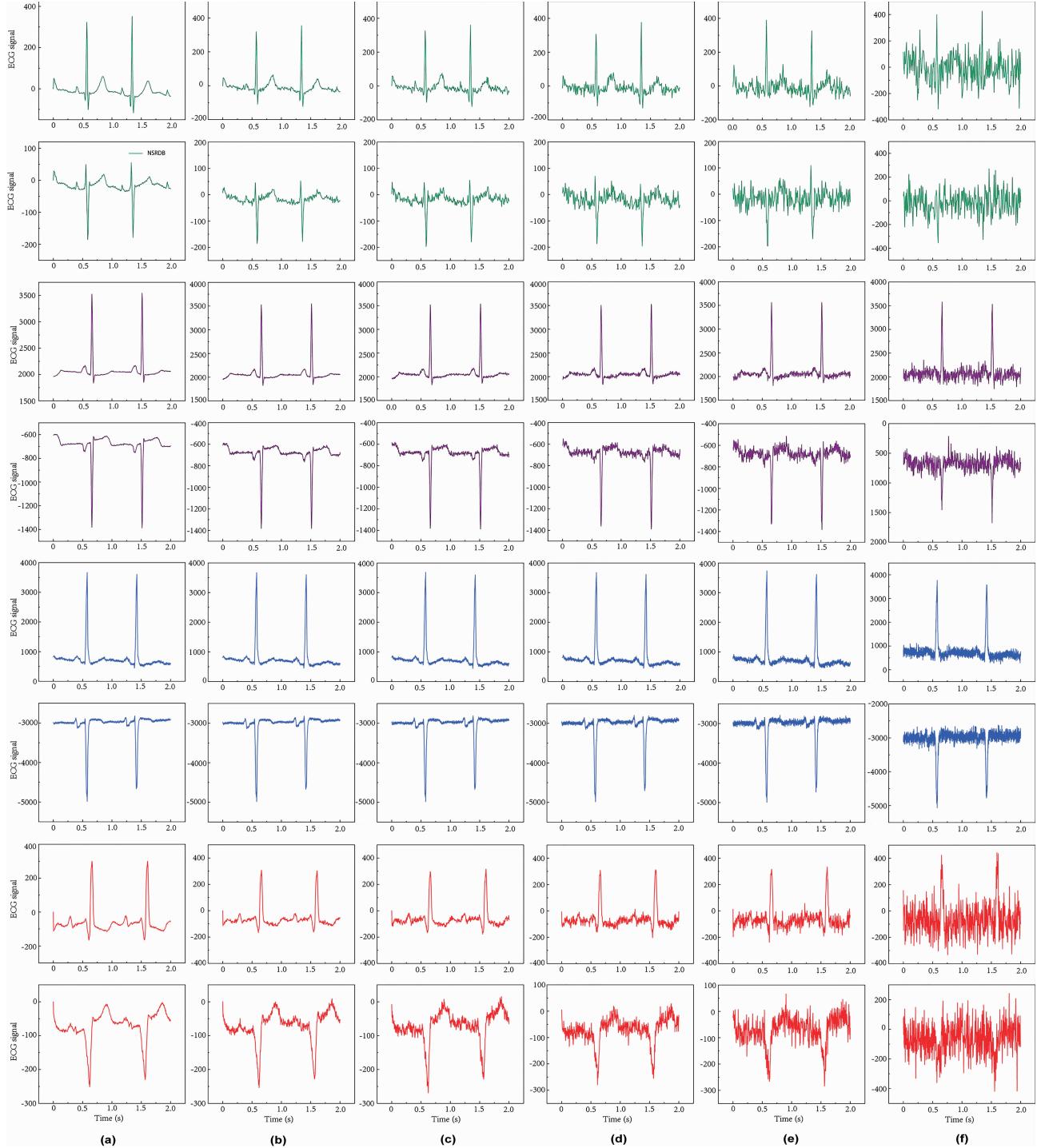
Fig. 10   **ECG signals under different signal-to-noise ratios (a) ∞ dB; (b) 24 dB; (c) 18 dB; (d) 12 dB; (e) 6 dB; (f) 0 dB. Each row is different from top to bottom. Green, purple, blue, and red curves indicate normal heartbeat, CAD heartbeat, MI heartbeat, and CHF heartbeat, respectively.**
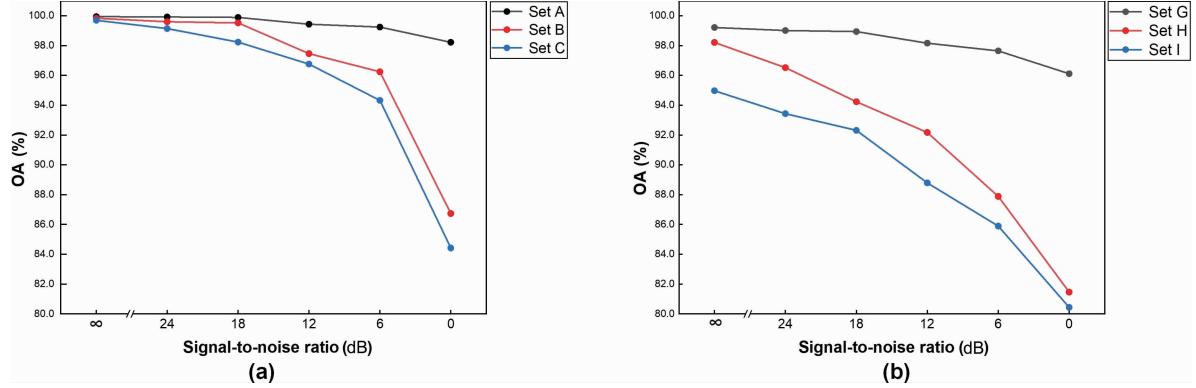
Fig. 11   Results (OA) of ECG signals under noise interference with different signal-to-noise ratios. (a) Balanced datasets (A, B, and C); (b) unbalanced datasets (G, H, and I).

Table 14   Results (OA) of ECG signal processing under noise interference with different SNR for the six datasets.

(%)

SNR (dB)	OA					
	A	B	C	G	H	I
∞	99.94	99.83	99.69	99.21	98.21	94.97
24	99.92	99.60	99.14	99.01	96.52	93.43
18	99.89	99.52	98.23	98.94	94.23	92.31
12	99.43	97.46	96.76	98.16	92.17	88.79
6	99.24	96.24	94.32	97.64	87.88	85.89
0	98.22	86.73	84.42	96.11	81.46	80.44

which seriously affects judgment performance. In all experiments where the SNR is greater than or equal to 6 dB, the experimental result of OA is greater than 90% in 90% of the cases. However, the experimental results on Datasets H and I with an SNR of 0 dB exceed 80%. The above-mentioned results indicate that the model has good noise robustness in experiments with two-, three-, and four-classification diagnoses in distinguishing normal, CAD, MI, and CHF.

6.6   Results of traditional models

In the paper, we use datasets C and I. We followed the input size of the model proposed in the paper. The performance of several traditional models and current popular models was also evaluated. The other model architectures used are consistent with the model architectures shown in the original paper. The experimental results are shown in Tables 15 and 16. The evaluation metrics include the average metrics diagnosed for each specific classification. All models achieved classification metric results exceeding 97% on intra-patient experimental datasets. In addition, the classification metrics on inter-patient experimental datasets achieved excellent performance for the currently available models, although the performance varied

Table 15   Experimental results compared with traditional models (intra-patient datasets).

(%)

Model	A_{cc}	S_{en}	P_p	S_{pe}	F1
VGG_16[57]	97.76	98.32	97.84	97.78	98.05
ResNet_18[52]	98.72	98.23	98.46	98.64	98.43
ResNet_34[52]	98.74	98.26	99.48	98.60	98.42
ResNet_50[52]	98.70	98.17	98.24	98.56	98.36
CNN-LSTM[46]	98.81	99.54	98.97	98.21	98.87
Transformer[50]	99.82	99.74	99.76	99.82	99.72
Proposed	99.84	99.68	99.71	99.90	99.69

Table 16   Experimental results compared with traditional models (inter-patient datasets).

(%)

Model	A_{cc}	S_{en}	P_p	S_{pe}	F1
VGG_16[57]	80.03	62.48	70.21	87.21	78.94
ResNet_18[52]	91.46	78.43	82.54	92.96	84.77
ResNet_34[52]	92.34	79.46	84.68	94.38	86.34
ResNet_50[52]	90.21	77.85	78.37	91.79	82.98
CNN-LSTM[46]	93.14	89.37	92.34	94.63	92.46
Transformer[50]	94.57	91.64	94.79	95.93	93.75
Proposed	97.48	93.54	96.30	97.89	94.87

relatively and widely. In the CNN-LSTM model[46], transformer model[50], and the model proposed in the paper, almost all inter-patient classification metrics exceeded 90%. In particular, the overall performance metrics of the current CNN-LSTM model[46] and the transformer model[50] are higher than those of the traditional models based on the experimental results of intra-patient datasets. The accuracy, sensitivity, positive predictive value, specificity, and F1 scores of the model presented in this paper were 99.84%, 99.68%, 99.71%, 99.90%, and 99.69%, respectively, of which the accuracy and specificity values are the highest, followed by the F1 score and sensitivity and positive prediction values. The model proposed in this paper

still outperformed other traditional and popular models in intra-patient experiments. Based on the experimental results of inter-patient datasets, all evaluation metrics of the proposed model are the highest, followed by those of the transformer model[50]. Therefore, our proposed model can utilize the CNN, Resnet residual network, and transformer model to obtain excellent results.

7   Discussion

In this paper, a great breakthrough is achieved by our proposed model with regard to intra-patient four-classification experiments and inter-patient four-classification experiments in severe cardiovascular diseases. The comparative results are shown in Tables 17 and 18.

The results of these experiments are compared with the latest results in the relevant literature. We found that all performance evaluation metrics of the model proposed in this paper are the highest in intra-patient and inter-patient experiments. In addition, the model proposed in this paper can accurately detect severe cardiovascular diseases using actual ECG signals.

The construction of the CR-Former model uses the information retention benefits of the residual structure of the ResNet network, allowing for the construction of high-level networks and the expansion of a certain amount of key information features. The network also utilizes the multi-headed attention mechanism in the transformer encoder, which allows all information to be interconnected and noticed. Despite its excellent performance, the model has some drawbacks. First, the model is built with several layers, which can result in excessive computational power and high computational complexity of the model, thereby accelerating computer

Table 17   Comparison results of multi-diagnostic studies (intra-patient datasets).

					(%)
Author	Year	A_{cc}	S_{en}	P_p	S_{pe}
Fujita et al.[44]	2017	97.98	99.61	–	94.84
Acharya et al.[45]	2017	99.55	**99.93**	–	99.24
Lih et al.[46]	2020	98.51	99.30	–	97.89
Gong et al.[47]	2020	99.74	98.67	99.09	99.83
Proposed	–	**99.84**	99.68	**99.71**	**99.90**

Table 18   Comparison results of multi-diagnostic studies (inter-patient datasets).

					(%)
Author	Year	A_{cc}	S_{en}	P_p	S_{pe}
Zhang et al.[47]	2020	96.92	89.18	92.17	97.77
Proposed	–	**97.48**	**93.54**	**96.30**	**97.89**

wear and tear over long periods of time. Second, the model takes a long time to diagnose. There are situations where we need to diagnose the condition quickly to buy the doctor's time for resuscitation.

8   Conclusion

The development of an integrated multi-classification diagnostic system for severe cardiovascular diseases is essential to rapidly save the lives of critically ill patients. It is also necessary for artificial intelligence in medicine. In this paper, we propose a new model that consists of a short-layered 1D CNN, a multilayer Resnet-network-structured block layer constructed using the structure and properties of the ResNet network, a transformer feature extraction layer consisting of three transformer model encoders, a softmax classifier, and a focal loss function. We applied a 10-fold cross-validation approach. In within-patient experiments, we applied Sets A, B, and C. The mean accuracies obtained are 99.94%, 99.88%, and 99.84%. In experiments on multilevel imbalanced datasets D (two-classification CHF diagnosis), E (three-classification CAD diagnosis), and F (four-classification CAD and MI diagnosis), the mean accuracies are 99.82%, 99.64%, 99.41%, and 99.53%, respectively. In inter-patient experiments, the average accuracies obtained for experiments conducted on datasets G, H, and I are 99.93%, 98.60%, and 97.48%, respectively. Therefore, the model produces excellent results for two-, three-, and four-classification diagnoses in inter-patient experiments. Good results are achieved for experimental classification diagnosis in mixed multilevel noise data, thereby demonstrating the model's extreme noise robustness. The model presented in this paper shows great application potential in the multi-classification diagnosis of cardiovascular diseases in a practical clinical environment.

In future work, we must work toward creating more concise and sensitive systems for the classification of severe cardiovascular diseases. In addressing the shortcomings of the model mentioned in this paper, the complexity of the algorithm must be optimized to save computer memory. In addition, we must perform feature extraction of dual-lead information and enhance the comprehensiveness of the information by using more leads to extract more relevant features. We can build fewer network layers to save training and classification time. Moreover, we must find more actual ECG signal data from hospitals to validate the

generalization capability of the model. Hence, we can determine whether the model can still have more accurate results when using ECG signals obtained in different environments for diagnosis.

Acknowledgment

References

[1] O. Sayadi and M. B. Shamsollahi, Multiadaptive bionic wavelet transform: Application to ECG denoising and baseline wandering reduction, *EURASIP J. Adv. Signal Process.*, vol. 2007, no. 1, p. 041274, 2007.

[2] E. J. Benjamin, S. S. Virani, C. W. Callaway, A. M. Chamberlain, A. R. Chang, S. S. Cheng, S. E. Chiuve, M. Cushman, F. N. Delling, R. Deo, et al., Heart disease and stroke statistics-2018 update: A report from the American Heart Association, *Circulation*, vol. 137, no. 12, pp. e67–e492, 2018.

[3] L. M. Buja and J. T. Willerson, The role of coronary artery lesions in ischemic heart disease: Insights from recent clinicopathologic, coronary arteriographic, and experimental studies, *Hum. Pathol.*, vol. 18, no. 5, pp. 451–461, 1987.

[4] E. D. Grech, *ABC of Interventional Cardiology*. 2nd ed. New York, NY, USA: John Wiley Sons, 2011.

[5] A. L. Bui, T. B. Horwich, and G. C. Fonarow, Epidemiology and risk profile of heart failure, *Nat. Rev. Cardiol.*, vol. 8, no. 1, pp. 30–41, 2011.

[6] A. P. Ambrosy, G. C. Fonarow, J. Butler, O. Chioncel, S. J. Greene, M. Vaduganathan, S. Nodari, C. S. P. Lam, N. Sato, A. N. Shah, et al., The global health and economic burden of hospitalizations for heart failure: Lessons learned from hospitalized heart failure registries, *J. Am. Coll. Cardiol.*, vol. 63, no. 12, pp. 1123–1133, 2014.

[7] E. J. Benjamin, P. Muntner, A. Alonso, M. S. Bittencourt, C. W. Callaway, A. P. Carson, A. M. Chamberlain, A. R. Chang, S. S. Cheng, S. R. Das, et al., Heart disease and stroke statistics-2019 update: A report from the American Heart Association, *Circulation*, vol. 139, no. 10, pp. 56–528, 2019.

[8] K. Thygesen, J. S. Alpert, A. S. Jaffe, M. L. Simoons, B. R. Chaitman, H. D. White, Writing Group on the Joint ESC/ACCF/AHA/WHF Task Force for the Universal Definition of Myocardial Infarction, Third universal definition of myocardial infarction, *Eur. Heart J.*, vol. 33, no. 20, pp. 2551–2567, 2012.

[9] J. E. Madias, ECG changes in response to diuresis in an ambulatory patient with congestive heart failure, *Congest. Heart Fail.*, vol. 12, no. 5, pp. 277–283, 2006.

[10] R. C. Schlant, R. J. Adolph, J. P. DiMarco, L. S. Dreifus, M. I. Dunn, C. Fisch, A. Garson Jr, L. J. Haywood, H. J. Levine, J. A. Murray, et al., Guidelines for electrocardiography. A report of the American College of Cardiology/American Heart Association Task Force on assessment of diagnostic and therapeutic cardiovascular procedures (Committee on Electrocardiography), *J. Am. Coll. Cardiol.*, vol. 19, no. 3, pp. 473–481, 1992.

[11] R. B. Schnabel, X. Y. Yin, P. Gona, M. G. Larson, A. S. Beiser, D. D. Mcmanus, C. Newton-Cheh, S. A. Lubitz, J. W. Magnani, P. T. Ellinor, et al., 50 year trends in atrial fibrillation prevalence, incidence, risk factors, and mortality in the Framingham Heart Study: A cohort study, *Lancet*, vol. 386, no. 9989, pp. 154–162, 2015.

[12] W. G. Morrison and I. J. Swann, Electrocardiograph interpretation by junior doctors, *Arch. Emerg. Med.*, vol. 7, no. 2, pp. 108–110, 1990.

[13] F. G. Liu, Z. W. Zhang, and R. L. Zhou, Automatic modulation recognition based on CNN and GRU, *Tsinghua Science and Technology*, vol. 27, no. 2, pp. 422–431, 2022.

[14] J. J. Pang, Y. Huang, Z. Z. Xie, J. B. Li, and Z. P. Cai, Collaborative city digital twin for the COVID-19 pandemic: A federated learning solution, *Tsinghua Science and Technology*, vol. 26, no. 5, pp. 759–771, 2021.

[15] S. Patidar, R. B. Pachori, and U. R. Acharya, Automated diagnosis of coronary artery disease using tunable-Q wavelet transform applied on heart rate signals, *Knowl.-Based Syst.*, vol. 82, pp. 1–10, 2015.

[16] M. Kumar, R. B. Pachori, and U. R. Acharya, Characterization of coronary artery disease using flexible analytic wavelet transform applied on ECG signals, *Biomed. Signal Process. Control*, vol. 31, pp. 301–308, 2017.

[17] G. Altan, N. Allahverd, and Y. Kutlu, Diagnosis of coronary artery disease using deep belief networks, *Eur. J. Eng. Nat. Sci.*, vol. 2, no. 1, pp. 29–36, 2017.

[18] U. R. Acharya, H. Fujita, M. Adam, O. S. Lih, V. K. Sudarshan, T. J. Hong, J. E. Koh, Y. Hagiwara, C. K. Chua, C. K. Poo, et al., Automated characterization and classification of coronary artery disease and myocardial infarction by decomposition of ECG signals: A comparative study, *Information Sciences*, vol. 377, pp. 17–29, 2017.

[19] A. Caliskan and M. E. Yuksel, Classification of coronary artery disease data sets by using a deep neural network, *The EuroBiotech Journal*, vol. 1, no. 4, pp. 271–277, 2017.

[20] A. D. Dolatabadi, S. E. Z. Khadem, and B. M. Asl, Automated diagnosis of coronary artery disease (CAD) patients using optimized SVM, *Comput. Methods Programs Biomed.*, vol. 138, pp. 117–126, 2017.

[21] J. H. Tan, Y. Hagiwara, W. Pang, I. Lim, S. L. Oh, M. Adam, R. S. Tan, M. Chen, and U. R. Acharya, Application of stacked convolutional and long short-term memory network for accurate identification of CAD ECG signals, *Comput. Biol. Med.*, vol. 84, pp. 19–26, 2018.

[22] U. R. Acharya, H. Fujita, O. S. Lih, M. Adam, J. H. Tan, and C. K. Chua, Automated detection of coronary artery disease using different durations of ECG segments with convolutional neural network, *Knowl.-Based Syst.*, vol. 132, pp. 62–71, 2017.

[23] M. Abdar, W. Książek, U. R. Acharya, R. S. Tan, V. Makarenkov, and P. Pławiak, A new machine learning

technique for an accurate diagnosis of coronary artery disease, *Comput. Methods Programs Biomed.*, vol. 179, p. 104992, 2019.

[24] K. Chandrakar, A new approach to detect congestive heart failure using detrended fluctuation analysis of electrocardiogram signals, *J. Eng. Sci. Technol.*, vol. 10, no. 2, pp. 145–159, 2015.

[25] K. Chandrakar, Entropy measures of irregularity and complexity for surface electrocardiogram time series in patients with congestive heart failure, *J. Adv. Comput. Res.*, vol. 6, no. 4, pp. 1–11, 2015.

[26] U. R. Acharya, H. Fujita, V. K. Sudarshan, S. L. Oh, A. Muhammad, J. E. W. Koh, J. H. Tan, C. K. Chua, K. P. Chua, and R. S. Tan, Application of empirical mode decomposition (EMD) for automated identification of congestive heart failure using heart rate signals, *Neural Comput. Appl.*, vol. 28, no. 10, pp. 3073–3094, 2017.

[27] Z. Masetic and A. Subasi, Congestive heart failure detection using random forest classifier, *Comput. Methods Programs Biomed.*, vol. 130, pp. 54–64, 2016.

[28] V. K. Sudarshan, U. R. Acharya, S. L. Oh, M. Adam, J. H. Tan, C. K. Chua, K. P. Chua, and R. S. Tan, Automated diagnosis of congestive heart failure using dual tree complex wavelet transform and statistical features extracted from 2 s of ECG signals, *Comput. Biol. Med.*, vol. 83, pp. 48–58, 2017.

[29] M. Kumar, R. B. Pachori, and U. R. Acharya, Use of accumulated entropies for automated detection of congestive heart failure in flexible analytic wavelet transform framework based on short-term HRV signals, *Entropy*, vol. 19, no. 3, p. 92, 2017.

[30] U. R. Acharya, H. Fujita, S. L. Oh, Y. Hagiwara, J. H. Tan, M. Adam, and R. S. Tan, Deep convolutional neural network for the automated diagnosis of congestive heart failure using ECG signals, *Appl. Intell.*, vol. 49, no. 1, pp. 16–27, 2019.

[31] R. K. Tripathy, M. R. A. Paternina, J. G. Arrieta, A. Zamora-Méndez, and G. R. Naik, Automated detection of congestive heart failure from electrocardiogram signal using Stockwell transform and hybrid classification scheme, *Comput. Methods Programs Biomed.*, vol. 173, pp. 53–65, 2019.

[32] S. Khade, A. Subhedar, K. Choudhary, T. Deshpande, and U. Kulkarni, A system to detect heart failure using deep learning techniques, *Int. Res. J. Eng. Technol.*, vol. 6, no. 6, pp. 384–387, 2019.

[33] L. Hussain, I. A. Awan, W. Aziz, S. Saeed, A. Ali, F. Zeeshan, and K. S. Kwak, Detecting congestive heart failure by extracting multimodal features and employing machine learning techniques, *BioMed Res. Int.*, vol. 2020, p. 4281243, 2020.

[34] N. Safdarian, N. J. Dabanloo, and G. Attarodi, A new pattern recognition method for detection and localization of myocardial infarction using T-wave integral and total integral as extracted features from one cycle of ECG signal, *J. Biomed. Sci. Eng.*, vol. 7, no. 10, pp. 818–824, 2014.

[35] P. Kora and S. R. Kalva, Improved Bat algorithm for the detection of myocardial infarction, *SpringerPlus*, vol. 4, no. 1, p. 666, 2015.

[36] U. R. Acharya, H. Fujita, S. L. Oh, Y. Hagiwara, J. H. Tan, and M. Adam, Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals, *Information Sciences*, vol. 415–416, pp. 190–198, 2017.

[37] M. Sharma, R. S. Tan, and U. R. Acharya, A novel automated diagnostic system for classification of myocardial infarction ECG signals using an optimal biorthogonal filter bank, *Comput. Biol. Med.*, vol. 102, pp. 341–356, 2018.

[38] K. Feng, X. T. Pi, H. Y. Liu, and K. Sun, Myocardial infarction classification based on convolutional neural network and recurrent neural network, *Appl. Sci.*, vol. 9, no. 9, p. 1879, 2019.

[39] C. Han and L. Shi, ML-ResNet: A novel network to detect and locate myocardial infarction using 12 leads ECG, *Comput. Methods Programs Biomed.*, vol. 185, p. 105138, 2020.

[40] W. H. Liu, F. Wang, Q. J. Huang, S. Chang, H. Wang, and J. He, MFB-CBRNN: A hybrid network for MI detection using 12-lead ECGs, *IEEE J. Biomed. Health Inform.*, vol. 24, no. 2, pp. 503–514, 2020.

[41] U. B. Baloglu, M. Talo, O. Yildirim, R. S. Tan, and U. R. Acharya, Classification of myocardial infarction with multi-lead ECG signals and deep CNN, *Pattern Recognit. Lett.*, vol. 122, pp. 23–30, 2019.

[42] W. Y. Yang, Y. J. Si, D. Wang, G. Zhang, X. Liu, and L. L. Li, Automated intra-patient and inter-patient coronary artery disease and congestive heart failure detection using EFAP-Net, *Knowl.-Based Syst.*, vol. 201–202, p. 106083, 2020.

[43] W. Y. Yang, Y. J. Si, G. Zhang, D. Wang, M. Q. Sun, W. Fan, X. Liu, and L. L. Li, A novel method for automated congestive heart failure and coronary artery disease recognition using THC-Net, *Information Sciences*, vol. 568, pp. 427–447, 2021.

[44] H. Fujita, V. K. Sudarshan, M. Adam, S. L. Oh, J. H. Tan, Y. Hagiwara, K. C. Chua, K. P. Chua, and U. R. Acharya, Characterization of cardiovascular diseases using wavelet packet decomposition and nonlinear measures of electrocardiogram signal, in *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*, S. Benferhat, K. Tabia, and M. Ali, eds. Arras, France: Springer, 2017, pp. 259–266.

[45] U. R. Acharya, H. Fujita, V. K. Sudarshan, S. L. Oh, M. Adam, J. H. Tan, J. H. Koo, A. Jain, C. M. Lim, and K. C. Chua, Automated characterization of coronary artery disease, myocardial infarction, and congestive heart failure using contourlet and shearlet transforms of electrocardiogram signal, *Knowl.-Based Syst.*, vol. 132, pp. 156–166, 2017.

[46] O. S. Lih, V. Jahmunah, T. R. San, E. J. Ciaccio, T. Yamakawa, M. Tanabe, M. Kobayashi, O. Faust, and U. R. Acharya, Comprehensive electrocardiographic diagnosis based on deep learning, *Artif. Intell. Med.*, vol. 103, p. 101789, 2020.

[47] G. Zhang, Y. J. Si, W. Y. Yang, and D. Wang, A robust multilevel DWT densely network for cardiovascular disease classification, *Sensors*, vol. 20, no. 17, p. 4777, 2020.

　　　　　　　　　　　　　　　　　　　　Tsinghua Science and Technology, April 2023, 28(2): 386–404

[48] P. De Chazal, M. O'Dwyer, and R. B. Reilly, Automatic classification of heartbeats using ECG morphology and heartbeat interval features, *IEEE Trans. Biomed. Eng.*, vol. 51, no. 7, pp. 1196–1206, 2004.

[49] H. F. Huang, J. Liu, Q. Zhu, R. P. Wang, and G. S. Hu, A new hierarchical method for inter-patient heartbeat classification using random projections and RR intervals, *Biomed. Eng. Online*, vol. 13, p. 90, 2014.

[50] G. S. Yan, S. Liang, Y. C. Zhang, and F. Liu, Fusing transformer model with temporal features for ECG heartbeat classification, presented at the 2019 IEEE Int. Conf. Bioinformatics and Biomedicine (BIBM), San Diego, CA, USA, 2019, pp. 898–905.

[51] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, and H. E. Stanley, PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals, *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[52] K. M. He, X. Y. Zhang, S. Q. Ren, and J. Sun, Deep residual learning for image recognition, presented at the 2016 IEEE Conf. Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 2016, pp. 770–778.

[53] S. Ioffe and C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift, in *Proc. 32nd Int. Conf. Machine Learning*, Lille, France, 2015, pp. 448–456.

[54] H. Song, D. Rajan, J. Thiagarajan, and A. Spanias, Attend and diagnose: Clinical time series analysis using attention models, *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, no. 1, pp. 4091–4098, 2018.

[55] R. P. Lippmann, Pattern classification using neural networks, *IEEE Commun. Mag.*, vol. 27, no. 11, pp. 47–50, 1989.

[56] X. Glorot and Y. Bengio, Understanding the difficulty of training deep feedforward neural networks, in *Proc. 13th Int. Conf. Artificial Intelligence and Statistics*, Sardinia, Italy, 2010, pp. 249–256.

[57] K. Simonyan and A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv: 1409.1556, 2015.

Dengao Li received the PhD degree from Taiyuan University of Technology, Jinzhong, China, in 2010. He is a professor at the College of Data Science, Taiyuan University of Technology, Jinzhong, China. He is the chairman of the Taiyuan Branch of the International Computer Society (ACM) and a member of the China Computer Federation (CCF) Internet of Things Committee. His main research interests include air-space-ground integration and navigation technology and big data analysis technology and application.

Jumin Zhao received the PhD degree from Taiyuan University of Technology, Jinzhong, China, in 2008. She is a professor at the College of Information and Computer Science, Taiyuan University of Technology, Jinzhong, China. She is the deputy secretary-general of the Taiyuan Branch of the ACM and a member of the CCF Internet of Things Committee. Her main research interests include intelligent perception, internet of things technology, and big data analysis technology and application.

Changcheng Shi received the BE degree from Taiyuan University of Technology, Jinzhong, China, in 2018. He is currently pursuing the MS degree at Taiyuan University of Technology, Jinzhong, China. His research interest is medical data processing.

Chunxia Li received the MS degree from Shanxi Medical University in 2011. She is pursuing the PhD degree at the Department of Dardiology, Shanxi Bethune Hospital, Taiyuan, China. Her main research interests include diagnosis and treatment of myocardial infarction and heart failure.

Yi Liu received the MS degree from Ningxia University, Yinchuan, China, in 2018. He is currently pursuing the PhD degree at Taiyuan University of Technology, Taiyuan, China. His research interest is medical data processing.