IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 29 January 2024, accepted 23 February 2024, date of publication 7 March 2024, date of current version 13 March 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3373191

RESEARCH ARTICLE

Multi-Scale and Multi-Channel Information Fusion for Exercise Electrocardiogram Feature Extraction and Classification

JUTAO WANG[1], FUCHUN ZHANG[2], MENG LI[2], AND BAIYANG WANG[2]
[1]School of Physical Education and Health, Linyi University, Linyi, Shandong 276005, China
[2]School of Information Science and Engineering, Linyi University, Linyi, Shandong 276005, China

Corresponding author: Fuchun Zhang (fufcfd97@163.com)

This work was supported by the Research on the Promotion of Adolescents' Physical Health in Shandong Province Based on Big Data Analysis, Shandong Social Science Planning Research Project, under Project 21CTYJ03.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Massachusetts Institute of Technology.

ABSTRACT Increased physical activity can help reduce the occurrence of cardiovascular disease. However, cardiovascular disease during strenuous exercise also brings certain risks, so a convenient and effective method is needed to accurately identify heart rate. Due to the low amplitude characteristics of ECG signals, automatic classification of the imperceptibility and irregularities of ECG signals remains a great challenge. To address this issue, we propose an automatic heart rate detection method using a two-dimensional convolutional neural network. First, the ECG data is preprocessed to convert the multi-lead single-channel ECG data into dual-channel ECG data. Then, the ECG image is input into the convolutional neural network. To enhance the diagnostic accuracy of an abnormal heart rate diagnosis model, this study integrates a multi-scale pyramid module into the network to fully extract image contextual information. Experimental analysis utilizes MIT-BIH and Sudden Cardiac Death Holter public datasets for training and testing. The results show that the final classification accuracy reaches 99.12% and 98.40%, respectively. The method requires no manual pre-processing of converted ECG images, has high accuracy, can effectively monitor abnormal heart rate, and minimize sudden death caused by abnormal heart rate.

INDEX TERMS ECG signal, multi-scale, deep learning, convolutional neural network, information fusion.

I. INTRODUCTION

According to the World Health Organization, cardiovascular disease has become one of the most important causes of death worldwide [1]. In 2016, over 17 million deaths were attributed to cardiac arrhythmias, and the number is expected to increase to 24 million by 2030 [2]. Abnormal heart rhythm is a common condition in the early stages of many cardiovascular diseases, with symptoms of tachycardia, bradycardia, and arrhythmia [3]. Vigorous exercise is one of the important causes of abnormal heart rate. Abnormal heart rate during exercise can cause adverse reactions in the body, and in people with poor physical fitness, may even lead to sudden death. Accordingly, monitoring abnormal heart rate during exercise is crucially important. In traditional methods, ECG collection and diagnosis are not synchronized. Diagnosis requires collecting ECG signal during exercise in advance. The doctor then diagnoses the abnormal heart rate from the ECG post-exercise. They can provide appropriate treatment according to different heart rate types [4].

An electrocardiogram (ECG) is a waveform that records potential changes during human heart activity [5]. An abnormal heartbeat frequency will make the ECG signal we obtain abnormal, and the duration of this abnormal ECG waveform is short, making it very difficult for doctors to diagnose early abnormal heartbeats [6]. Additionally, heart rate analysis is

The associate editor coordinating the review of this manuscript and approving it for publication was Yiqi Liu.

a long-term monitoring process. Manual diagnosis requires considerable time and effort. Even experienced doctors may exhibit subjective uncertainties and human errors during analysis. Thus, there's an urgent need for computer-aided methods to reduce errors from fatigue and differences and improve abnormal heart rate detection.

An ECG of a cardiac cycle contains three waves, two intervals, and two segments. Three significant waves represent three different electrical phenomena of the heart during a heartbeat cycle: P wave (atrial depolarization), QRS complex (ventral depolarization), and T wave (repolarization) [7]. The length of the two intervals (PR interval and QT interval) represents the time the heart needs to complete corresponding electrical changes. Most heart diseases can be diagnosed based on abnormalities in the ECG signal. Typical heart problems include myocardial infarction, coronary artery disease, and various types of arrhythmias. Arrhythmias can be classified into 16 types including tachycardia, bradycardia, supraventricular arrhythmia, ventricular arrhythmia, and atrial fibrillation. However, most studies do not divide them into so many categories. In the case of multiple leads, the ECG can clarify the specific location of myocardial infarction.

ECG signals are often impacted by various interference components like muscle artifacts, EEG, and white noise, which can cause abnormal changes in the received signal in terms of frequency and amplitude [8]. The reason is that during an ECG recording, electrodes not only record the signal from the patient's heart, but also receive electrical signals from many different sources. The recording is performed through electrodes on the skin, which not only record the electrical activity of the heart but are also affected by electrical signals from other body parts and the external environment. Currently, researchers utilize various methods like canonical correlation analysis (CCA), principal component analysis (PCA) [9], maximum noise fraction (MNF), independent component analysis (ICA) [10], singular value decomposition (SVD)/ICA [11], [12], periodic component analysis, and parallel linear predictor (PLP) filter [13] to identify and separate mixed signals. Building on ECG data processing and effective signal separation, we leverage deep learning technology for ECG signal classification. Compared to some existing pattern recognition research, deep learning-based methods don't require additional feature extraction and selection steps [14]. A neural network can extract representative features from input data and classify them with the help of an underlying softmax [15]. For example, Zhao et al. [16] used the wavelet function for preprocessing experimental data. They divided ECG signals into 9 sub-signals, then used wavelet reconstruction to address noise issues. Finally, they used a convolutional neural network for data classification. The accuracy rate reaches 86.46%. Wang et al. [17] considered the R peak interval beneficial for classification, so they extracted features from it and achieved 98.74% accuracy. Wu et al. [18] proposed a convolutional neural network (CNN) based algorithm for real-time arrhythmia detection using a binarized model, reducing computational power and memory required for model operations. Yao et al. [19] fused attention into a convolutional neural network to achieve the fusion of ECG signal time and space, reducing parameters and achieving 81.2% accuracy. Du et al. [20] proposed a fine-grained multi-label electrocardiogram (FM-ECG) framework using a weakly supervised fine-grained classification mechanism, achieving good detection results. Based on the ResNet model, Zhao et al. [21] applied a low-pass filter to remove noise in ECG signals, achieving high classification accuracy across 5 ECG types. LUO et al. [22] proposed a Hybrid Convolutional Recurrent Neural Network (HCRNet) on ECG time-series signals, solving the problem of imbalanced ECG data and achieving effective results.

Existing studies have achieved good signal recognition accuracy, but signal feature extraction remains somewhat cumbersome. Therefore, building on prior research and referencing excellent image algorithms, we propose a classification method for exercise electrocardiograms based on multi-scale feature extraction and multi element information fusion. Considering deep network gradient disappearance during training, we employ an improved multi-scale network based on VGG16 for ECG classification. Our contributions include:

(1) Convert the one-dimensional ECG signal into a two-dimensional image, and fuse the ECG signals of two channels into one image as the input of the network.

(2) We added a multi-scale feature fusion module on the basis of VGG. This module is added to the deep layer of the network, so that when the image features reach the deep layer of the network, the high-level and low-level image information can be combined through this module. When the network learns image features, it makes full use of the context information of the image at different scales, so as to combine more image information to learn more content, so as to alleviate the gradient disappearance of the deep network and complete a higher accuracy ECG classification work.

The rest of this paper is organized as follows. Materials and methods are included in Section II. Then, Section III gives the experimental study and evaluates the advantages of the method and the rationality of the proposed method. Finally, we conclude the paper in Section IV and discuss possible future work.

II. MATERIALS AND METHODS

A. DATASET PRE-PROCESSING

Data pre-processing is a critical technology for classification tasks. In this study, we used the matplotlib method in the python language to fuse two groups of single-channel ECG data. After data fusion and processing, the data was input into a network model. We took 400 data points as a sample. Data 1400 was the first sample generated by cutting. The starting point of the next sample was 401800 until the end of the ECG data. If the last data had less than 400 data points, they were discarded. The definition of each sample is shown in

formula (1), where p is the current data point and n is the interval length of the selected data point.

$$V(p-n) \leqslant V(p) \leqslant V(p+n), p \in (n, 2\,n, 3\,n, \ldots, p\,n) \tag{1}$$

The divided data points are shown in Figure 1, and every 400 data points is a sample, and this cycle is repeated.

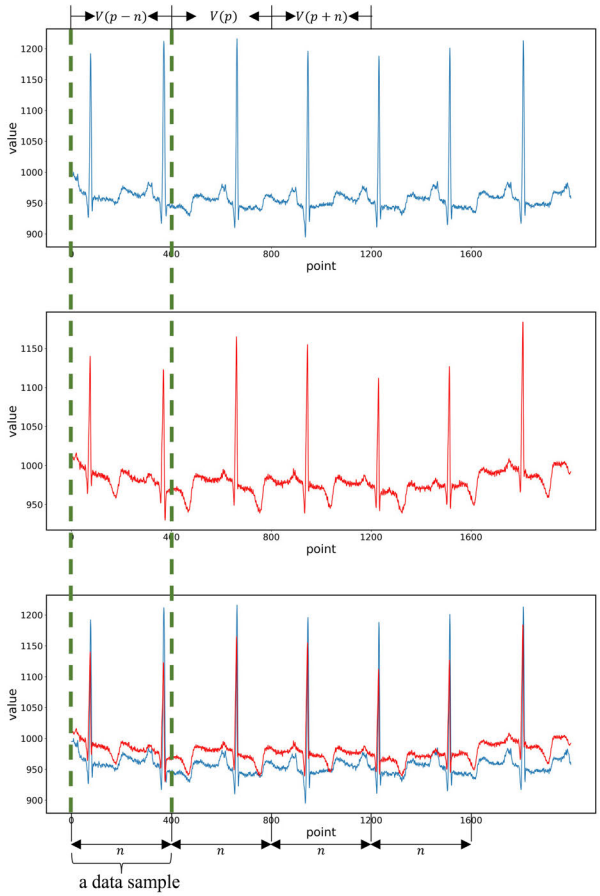

FIGURE 1. Data point sampling. Blue is the ECG data of the MLII channel, and red is the ECG data of the V5 channel. The intersection of red and blue is dual-channel fused ECG data.

B. VGG NETWORK OVERVIEW

VGG [23], the full name of visual geometry group, is a series of convolutional neural network models starting with VGG published by the department of science and engineering of Oxford university, which can be applied to face recognition [24], [25], [26], image classification [27], [28] and other aspects. The original purpose of studying the depth of convolution network is to find out how the depth of convolution network affects the accuracy and accuracy of large-scale image classification and recognition. It was originally VGG16, known as very deep convolution network. This paper introduces an im-proved multiscale ECG classification algorithm. Its network architecture is based on VGG16 network structure, and the extraction of feature information at different scales is considered.

C. OVERVIEW OF DILATED CONVOLUTION

In previous studies, the size of the receptive field of the convolution kernel is limited by its size. Generally, in order to reduce the number of parameters, the size of the convolution kernel needs to be limited to a smaller size. Therefore, the receptive field can be expanded by the down-sampling operation or increasing the depth and width of the deep learning network. However, repeated de-sampling operations will lose the resolution of the feature map and a lot of detailed feature information, which cannot be effectively restored. At the same time, the method of increasing the network width may produce more parameters, resulting in over-fitting of the network. In addition, the method of increasing the network depth may cause the gradient of the network to disappear when the network depth is too deep. In order to solve the above problems, dilated convolution came into being [29].

We add dilated convolution to VGG method and reduce the number of down-sampling layers, so that the network can have a large receptive field when extracting signal characteristics. For the dilated convolution of a two-dimensional image, w is the dilated convolution core, and w is each pixel on the feature map. Then the relationship between the input feature map x and the output feature map y is

$$f\,(i) = \sum\nolimits_k x\,(i + r \cdot k) w(k) \tag{2}$$

where r is the dilation rate, and k is the sampling interval of the dilated convolution. are other pixels except the center pixel in the dilated convolution process.

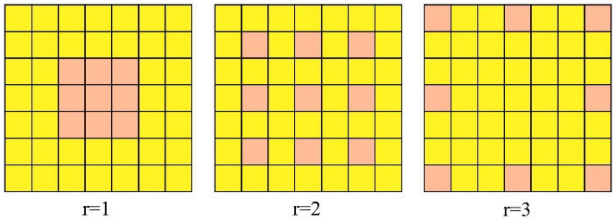

FIGURE 2. Schematic image of dilated convolution with different dilation rates.

In dilated convolution, the size of the convolution kernel is fixed at 3 × 3. When the expansion rate increases, the sampling interval of the convolution kernel increases, and the receptive field also increases accordingly. Therefore, dilated convolution can achieve the purpose of expanding the receptive field by setting different dilated rates. Dilated convolution is applied to the convolution layer at the bottom of the VGG16 network to obtain an appropriate receptive field by adjusting the expansion rate, thus avoiding irreversible detail feature loss caused by excessive down-sampling operations. Figure 2 shows the schematic diagram of cavity convolution with different cavity rates. r = 1 represents standard convolution, which is a special form of hollow convolution. It covers 3 × 3 subregions in the image through convolution operation. r = 2 represents the hollow convolution with a cavity rate of 2, and the 5 × 5 subregions in the

image are covered by the convolution operation. r = 3 represents a cavity convolution with a cavity rate of 3, covering 7 × 7 subregions in the image.

D. IMPROVED VGG NETWORK

The VGG16 network structure does not take into account the extraction of image feature information at different scales. In order to make full use of these feature information, we added a multi-scale module in the network structure to better extract the details of the image.

The overall network framework proposed in this paper is shown in Figure 3, from which modules of the network architecture can be clearly seen. Firstly, a 3-channel image with a resolution of 224 × 224 is input into the convolutional layer. Once the convolution operation is completed, the data enters the pooling layer again following nonlinear function mapping. Local feature information is gradually ignored. After more than 4 cycles of operation, the ECG characteristic information enters the fully connected layer following all convolution operations, with all feature graphs containing local information mapped to 2048 dimensions. The convolution kernel size of the convolutional layer of VGG16 network structure is all 3 × 3, the step size is 1, and the filling method is "same". Compared with other filling methods, this filling method can make the size of the convolution result obtained after each convolution unchanged. The pooling layer uses a 2 × 2 sized pooling kernel, which is filled in the same way as the convolutional layer, and the excitation function is ReLU.

When the network reaches a certain depth, we add a multi-scale module to perform dilated convolution operation on the acquired features at different dilated rates. This operation will make the network consider the feature information at different scales. After the completion of the multiscale module, the output of the multiscale module is fully connected, and finally the category of the output ECG is activated with softmax.

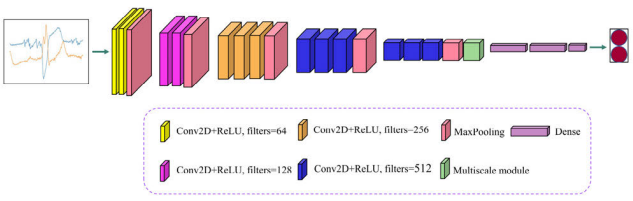

FIGURE 3. The overall architecture of the network.

We combine VGG16 with multi-scale strategy to design a multi-scale module using dilated convolution for extracting global features. The module comprises dilated convolution, batch normalization, and ReLU activation layers. It initially performs convolution and batch normalization on input features. By including batch normalization, the model's generalizability is enhanced. The resulting feature information is then convolved with dilated ratios of r = 1, r = 3, and r = 5. After each dilated convolution, the features are output in turn. Finally, all output features are concatenated using a connection layer. This module considers information at different scales and combines them for more accurate classification. See Figure 4 for details.

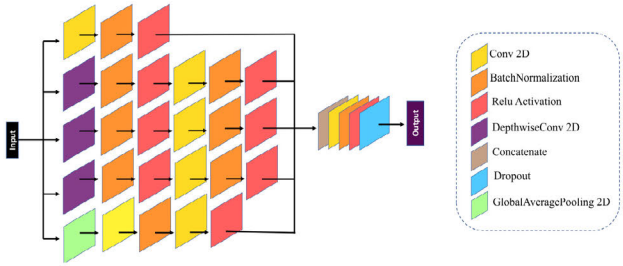

FIGURE 4. The overall architecture of multi-scale modules.

E. RECOGNITION METHODS

The method flow mainly includes: data pre-processing, training model and ECG image classification. Figure 5 shows the flow of the method proposed in this paper.

Step1. Collect ECG data during exercise by placing sensors on the human body. These ECG data are one-dimensional time series data.

Step 2. Preprocess the collected one-dimensional ECG data. Select fixed data points to create data samples. Use the matplotlib method to fuse single-channel data into a two-dimensional multi-channel ECG image.

Step 3. The ECG dataset is divided into training and validation sets in a 9:1 ratio.

Step 4. To fully extract features from multi-channel ECG images, a multi-scale network module is designed to enhance recognition accuracy.

Step 5. Train the designed model to obtain a network model for ECG image diagnosis.

Step 6. Validate the trained model on the test set, and calculate the accuracy of ECG image recognition.

III. EXPERIMENTAL AND DISCUSSION

A. EXPERIMENTAL SETTINGS

This paper uses the framework of python and tensorflow to complete the experiment. The program uses GPU NVIDIA GeForce GTX 2080 Ti for training and testing. During the training process, in order to reduce the memory usage, the batch size is set to 32, the learning rate is set to 0.0001, and the number of trainings is 50 times.

B. DATASET INTRODUCTION

The experiment was completed using the MIT-BIH arrhythmia dataset [30], [31]. This dataset was integrated by 47 researchers from 1975 to 1979. These mixed data were collected from the ECG signal data of inpatients and outpatients in Boston Hospital. These data include 48 30-minute dual-channel (i.e. lead MLII and V5) dynamic ECG recording fragments, 23 of which are randomly selected from 4000 24-hour dynamic ECG signal records. The records are sampled 360 times per second, with 11-bit resolution, representing a

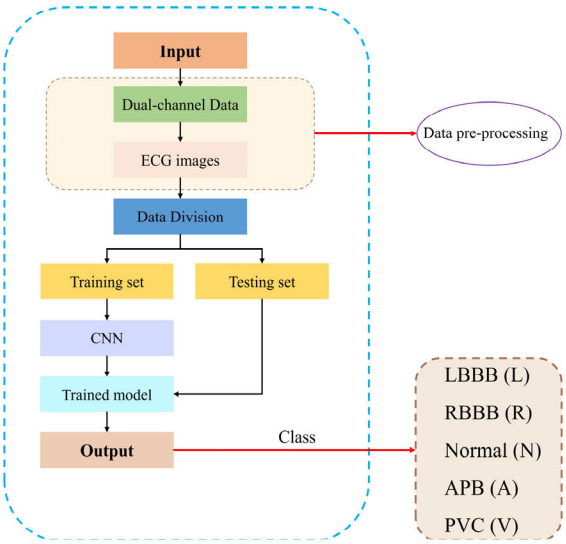

FIGURE 5. Flowchart of the proposed automatic classification algorithm. Including ECG signal acquisition, data processing and model training.

range of 10mv. The remaining 25 records are small sample data that are not common but have clinical significance, and can't well represent the symptoms of arrhythmia.

TABLE 1. MIT-BIH database number of ECG images.

Classes	ECG images
LBBB(L)	8048
RBBB(R)	5379
Normal(N)	10279
APB(A)	2545
PVC(V)	6090
Total	32341

Due to the unequal number of ECG images of each type of arrhythmia in the MIT-BIH dataset, the data is unbalanced. The data imbalance will make the image in-formation learned by the depth learning algorithm more inclined to the category with large data samples, which will reduce the accuracy of the depth learning model. Therefore, we train the model by using the arrhythmia category with a small difference in the number of data samples to reduce the impact of dataset imbalance. We selected five different types of arrhythmias from the MIT-BIH dataset, as follows: left bundle branch block (LBBB), right bundle branch block (RBBB), normal (N), atrial premature beat (APB) and premature ventricular contractions (PVC). Table 1 shows the number of two-dimensional ECG dataset samples used in this paper. Among them, the number of ECG for each arrhythmia category is also relatively close, which can reduce the impact of data imbalance during training.

In order to further verify the validity of this experiment, this article uses the Sudden Cardiac Death Holter Database [32] to further verify the experimental results of this article. The database currently includes 18 patients with underlying sinus rhythm (4 with intermittent pacing), 1 with continuous pacing, and 4 with atrial fibrillation. All patients had sustained ventricular tachyarrhythmias, and most had actual cardiac arrest. The data set is divided into four categories, namely normal, Premature ventricular contraction (PVC) and Isolated QRS-like artifact (I). The number of ECGs used in this dataset is shown in Table 2.

TABLE 2. Sudden cardiac death holter database number of ECG images.

Classes	ECG images
Normal(N)	10932
PVC(V)	10695
I	10914
Total	32541

C. DATASET DIVISION

The dataset is divided into training set and test set. In our experiments, 90% of the dataset is used for training and the remaining 10% is used for testing. The number of ECG images in the specific training dataset and test dataset is shown in Table 3. During training, 7244 images were used for the LBBB category, 4842 images were used for the RBBB category, 9252 images were used for the normal category, 2291 images were used for the APB category, and PVC category used 5481 images. During the test, 804 images were used for the LBBB category, 537 images were used for the RBBB category, 1027 images were used for the normal category, 254 images were used for the APB category, 609 images were used for the PVC category.

TABLE 3. The training set and test set of the MIT-BIH dataset.

Classes	Training set	Testing set
LBBB(L)	7244	804
RBBB(R)	4842	537
Normal(N)	9252	1027
APB(A)	2291	254
PVC(V)	5481	609

Table 4 shows the number of training and testing data for the Sudden Cardiac Death Holter Database.

TABLE 4. Number of ECG images in sudden cardiac death holter database during training and testing.

Classes	Training set	Testing set
Normal(N)	9839	1093
PVC(V)	9626	1069
I	9526	1091

D. EVALUATION INDICATORS

The network model is optimized by using a loss function. When training the model, the loss function is used to calculate

the difference between the neural predicted value and the true value. According to the derivative of the loss function, the neural network transmits the error back along the direction with the smallest gradient, and corrects each weight value in the forward calculation formula until the loss function reaches a satisfactory value, and the iteration stops. The network model is continuously trained, and its loss function is continuously reduced, the performance of the model will be better, and the classification accuracy will be higher. The cross-entropy loss function is a non-negative function and it is used for multi-classification tasks. Given N training samples $\left\{\left(\boldsymbol{x}^{(n)},\boldsymbol{y}^{(n)}\right)\right\}_{n=1}^{N}$,softmax regression uses the cross-entropy loss function to learn the optimal parameter matrix W. C represents the class label, and for a small class c, its vector is expressed as:

$$y = [(I(1 = c), I(2 = c), \ldots, I(C = c))]^T \tag{3}$$

where I $(\cdot)$ is the indicator function.

The cross-entropy loss function is defined as equation (4):

$$\mathcal{L}(W) = -\frac{1}{n}\sum_{n=1}^{N}\sum_{c=1}^{C} y_c^{(n)} log\hat{y}_c^{(n)} \tag{4}$$

The accuracy metric is used to express the proportion of correct ECG beat classification to all ECG beats. It is defined as:

$$ACC_i = \frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i} \tag{5}$$

TP_i (True Positive) and FN_i (False Negative) represent the number of classes i correctly predicted and the number of classes i assigned to other classes, respectively. TN_i (True Negative) and FP_i (False Positive) are the number of other classes not classified as class i and the prediction of other classes as class i, respectively.

E. EXPERIMENTAL RESULTS

Using the plt function in the Matplotlib package in python, the single-channel data (V5 channel and MLII channel) of one-dimensional time series are sequentially con-verted into 2D ECG images. The single-channel data and multi-channel data are trained using the same VGG network, and the training results are shown in Figure 6 and Table 5. In Figure 6, the first row represents the training set and validation set loss function values of the V5 channel, and the training set and validation set accuracy values. The second row represents the training set and validation set loss function values of the MLII channel, and the training set and validation set accuracy values. The third row represents the training set and validation set loss function values, and the training set and validation set accuracy values for the dual-channel data. Table 5 shows the loss function values and accuracy values during training and validation.

The accuracy rates of V5 channel, MLII channel and dual channel data are 98.15%, 98.65% and 99.09% on the validation set. In the training set, the recognition accuracy of the two-channel data is not significant. However, in the validation set, the two-channel data achieve the lowest loss function value and the highest accuracy, and it achieves better results than the single-channel data. It can be seen that the dual-channel data is more conducive to im-proving the recognition accuracy of ECG images.

TABLE 5. Loss value and accuracy for single-Channel Datasets.

Methods	Train_loss	Val_loss	Train_acc(%)	Val_acc(%)
V5	0.008	0.08	99.74	98.15
MLII	**0.007**	0.07	**99.77**	98.65
Dual-channel	0.01	**0.06**	99.70	**99.09**

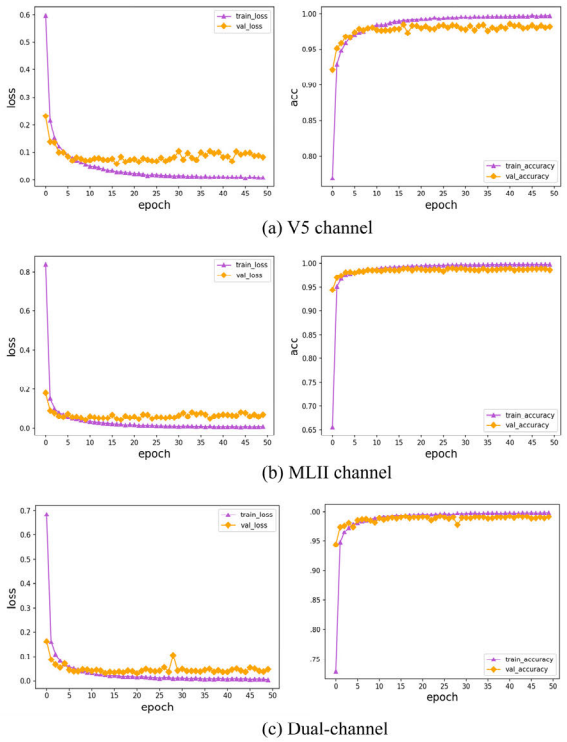

FIGURE 6. Experimental results of single-channel data: (a) loss value and accuracy of V5 channel, (b) loss value and accuracy of MLII channel, (c) loss value and accuracy of dual-channel.

In order to verify the effectiveness of the method of combining dual-channel data with multi-scale modules (DCMS), this paper uses dual-channel data to train the VGG network, and also uses dual-channel data to train the VGG network integrating multi-scale modules. Finally, the training results of the two methods are compared. The number of samples for training and testing is the same as that for single channel. The training results using dual-channel data are shown in Figure 7. In Figure 7, (a) shows the loss function value of the training set and verification set of the dual-channel data, as well as its accuracy value of the training set and verification set. (b) shows the loss function values of the training set and verification set of the DCMS method, as well as its accuracy values of the training set and verification set. The orange

TABLE 6. The loss value and acc of the dual-channel MIT-BIH dataset.

Methods	Train_loss	Val_loss	Train_acc(%)	Val_acc(%)
Dual-channel	0.01	0.06	**99.70**	99.09
DCMS	**0.01**	**0.03**	99.30	**99.12**

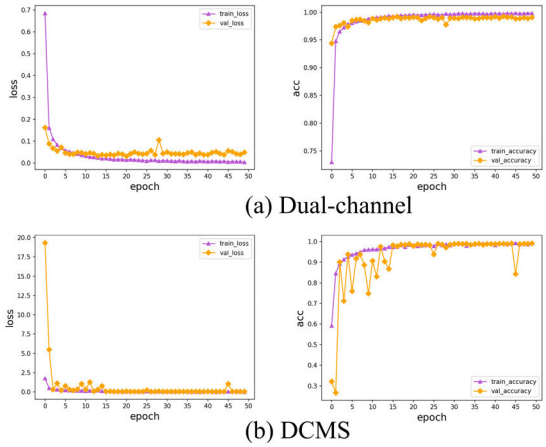

FIGURE 7. Experimental results of dual-channel data: (a) Experimental results of dual-channel (b) experimental results of dual-channel combined with multi-scale modules.

curve represents the result of the verification set, and the purple curve represents the result of the training set. It can be seen from the figure that the convergence speed of the two methods is equivalent, but the DCMS method has a high accuracy. From Table 6, the accuracy of the verification set of the DCMS method reached 99.12%, and its loss function value of the verification set also reached the lowest. It shows that DCMS method can make full use of dual-channel image information and achieve higher recognition accuracy in VGG network.

For the Sudden Cardiac Death Holter Database, this paper also uses dual-channel data to train the VGG network and the VGG network that integrates multi-scale mod-ules, and finally compares the experimental results. As shown in Table 7· and Figure 8. Table 7 shows the loss function value and accuracy value of training and verification. Under the condition of the same dual-channel data, the accuracy of the fusion multi-scale module is 98.94% lower than that of the comparison experiment, but the difference between the experimental results is very small. The experimental result is lower than the comparison experiment, which may be a problem with the experimental equipment. During the training process, it is limited by the experimental equipment, and the batch size of the training data cannot be adjusted.

Figure 8 shows the stability and convergence of the Sudden Cardiac Death Holter Database training process. During the training process of these two methods, the value and accuracy of the loss function are relatively stable, and the loss function and accuracy fluctuate greatly during verification.

TABLE 7. The loss value and acc of the dual-channel sudden cardiac death holter database.

Methods	Train_loss	Val_loss	Train_acc(%)	Val_acc(%)
Dual-channel	**0.007**	0.09	**99.76**	**98.94**
DCMS	0.22	**0.09**	98.02	98.40

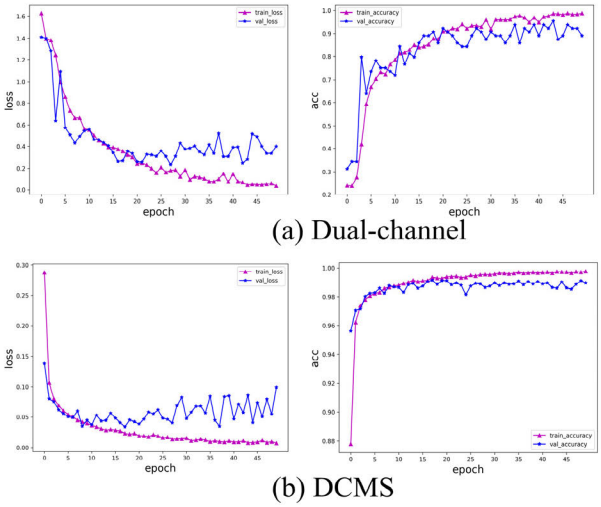

FIGURE 8. The experimental results of the two-channel Sudden Cardiac Death Holter Database: (a) the experimental results of the two-channel (b) the experimental results of the two-channel combined with the multi-scale module.

In contrast, the network validation accuracy is more stable when integrating multi-scale modules. Although the accuracy of the comparative experiments in Table 7 is higher than that of the method proposed in this paper, the method proposed in this paper is more stable in terms of validation accuracy.

F. DISCUSSION

The proposed method can achieve 5-class classification of ECG images with 99.12% accuracy. To further verify the proposed method's applicability, a confusion matrix is utilized to display recognition results. The confusion matrix is a summary of the prediction results for classification problems. By calculating the correct and incorrect number of predictions for each classification category and subdividing them, it shows which categories will be confused by the classification model during prediction. Not only can we identify which classification model is erroneous in predicting which classification through the confusion matrix, but more importantly, we can also discern the types of misclassifications. The confusion matrix compensates for limitations in classification accuracy. Therefore, we will use the confusion matrix to demonstrate advantages of this method.

As shown in Figure 9 (b), the DCMS method achieves recognition accuracy of 99% for LBBB, 100% for RBBB, 98% for Normal, 98% for APB and 98% for PVC. The proposed method is capable of clearly distinguishing various

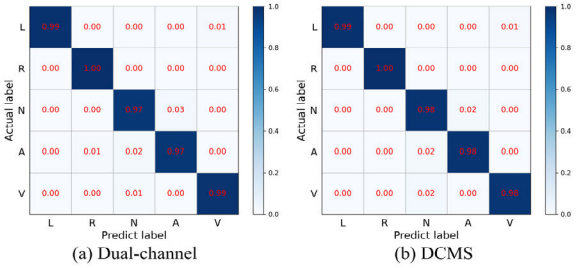

FIGURE 9. Confusion matrix plots for Dual-channel and DCMS methods.

types of arrhythmias. The results of the confusion matrix for dual-channel data are shown in Figure 9 (a). The recognition accuracy is 99% for LBBB, 100% for RBBB, 97% for Normal, 97% for APB, and 99% for PVC. Although PVC accuracy is higher than DCMS, DCMS is 1% more accurate for Normal and APB prediction than the dual-channel method. Comprehensively, DCMS method can effectively extract information from multi-channel ECG images. It significantly improves the accuracy of electrocardiogram classification and has better performance.

The Sudden Cardiac Death Holter Database is further utilized to visualize the experimental results, as shown in Figure 10. Figure 10 (a) shows the verification results of the comparative experiment using a confusion matrix. The verification results for all three categories reached 1.00. Figure 10 (b) shows the verification results for the fusion multi-scale module. The verification accuracy for the normal category reaches 99%, and the verification accuracy for other categories reaches 1.00.

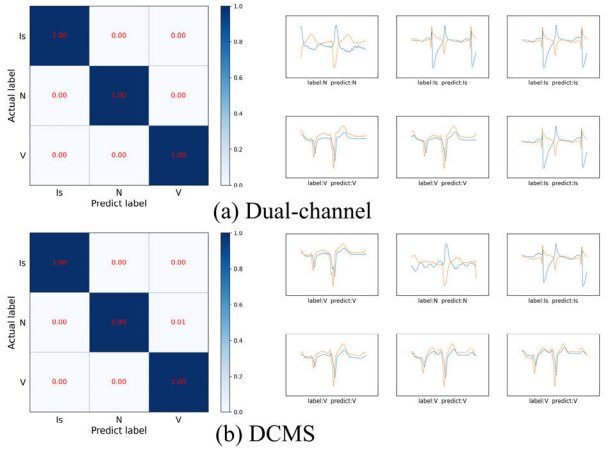

FIGURE 10. Confusion matrix plots of sudden cardiac death holter database dual-channel and DCMS methods.

We utilize a method based on convolutional neural networks to enable automatic classification of ECG signals. To verify the classification performance of the convolutional neural network, we also utilized the t-Disturbed Stochastic Nervous Embedding (t-SNE) [33] for visualizing classification results. T-SNE is a dimensionality reduction technology well-suited for visualizing high-dimensional data, and is widely used in deep learning methods.

Figure 11 shows a t-SNE diagram representing classification performance for dual-channel and DCMS methods. To better showcase classification performance and visualization results, we reduced the number of ECG images. Figure 9 shows that dual channels have some abnormal values mixed in with different arrhythmia types. The DCMS method may also have outliers, but fewer than the dual-channel method. This means the DCMS method can better complete the task of classifying arrhythmias.

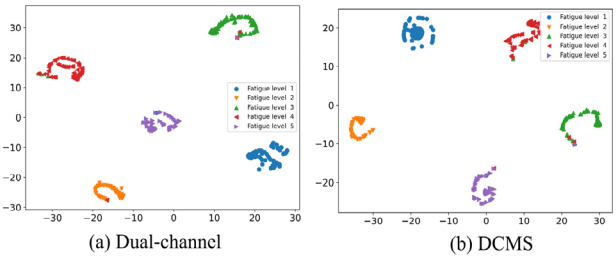

FIGURE 11. t-SNE images for Dual-channel and DCMS methods.

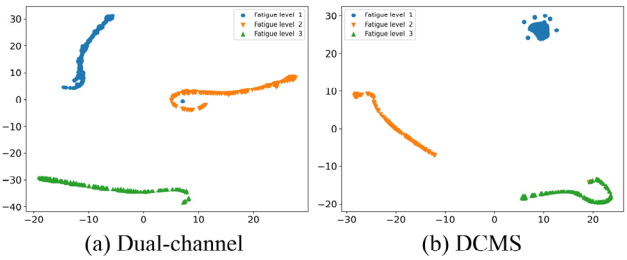

FIGURE 12. t-SNE images of sudden cardiac death holter dilated database two-channel and DCMS method.

Figure 12 uses t-SNE to show verification results for both methods. Their classification results are comparable. The dual-channel data is inaccurate for verifying the Isolated QRS-like artifact category, and the method of integrating multi-scale modules is not accurate for verifying the normal category. From the verification results, the proposed method is not accurate enough for verifying the normal category, and recognition accuracy needs further improvement.

We also compared ECG signal classification using convolutional neural networks under the same dataset in recent years. From Table 8, we can see that the ac-curacy of this method is 0.12% higher than Li et al., 0.10%, 0.38% and 0.79% higher than Ullah, Wang and Khan et al., 5.93%, 1.12% and 3.22% higher than Atal, Shaker and Xu et al., 0.52% and 1.70% higher than Zhao and Izci et al., respectively. The above results prove that the proposed method has good performance. Finally, the results obtained by this method and other methods are visualized, and the visualized results are shown in Figure 13. It can be clearly seen from Figure 13 that the accuracy of our proposed method exceeds the results obtained by other convolutional neural networks,

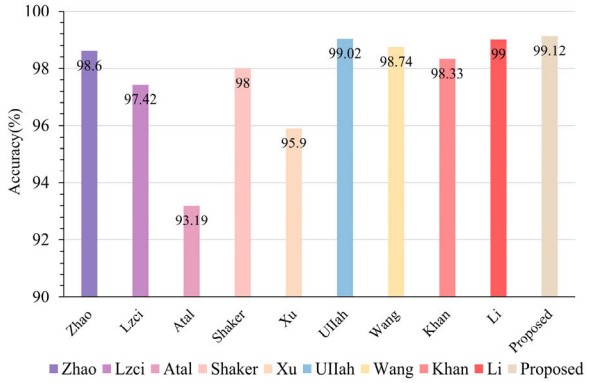

FIGURE 13. Comparison between DCMS method and other methods.

TABLE 8. Comparison between DCMS method and other methods.

Previous Methods	Year	Accuracy(%)
Zhao et al.[21]	2019	98.60%
Izci et al. [15]	2019	97.42%
Atal et al. [6]	2020	93.19%
Shaker et al. [34]	2020	98.00%
Xu et al. [35]	2020	95.90%
Ullah et al. [36]	2021	99.02%
Wang et al. [17]	2021	98.74%
Khan et al. [37]	2021	98.33%
Li et al. [38]	2022	99.00%
DCMS	2023	99.12%

and has excellent performance. Accurate classification of ECG data can effectively reduce misdiagnosis rates of various heart diseases, enabling timely detection and treatment. Convolutional neural networks can automatically extract features without manual intervention, avoiding impacts of lacking domain expertise on model classification results.

IV. CONCLUSION

This paper presents an abnormal ECG classification method that integrates multi-scale and multi-channel information. To obtain multi-channel information, the single-channel ECG data from MLII and V5 are fused to convert the one-dimensional ECG data into a two-dimensional ECG image. The processed data is then input into the neural network. To fully extract ECG signal features, a multi-scale feature extraction module is integrated into the model. Finally, the neural network is trained to obtain an abnormal heart rate diagnosis model based on multi-scale feature extraction and multi-lead information fusion. Through experiments and parameter optimization, the accuracy of MIT-BIH algorithm database and the Sudden Cardiac Death Holter Database are 99.12% and 98.40%, respectively. This method is compared with other ECG signal classification methods from recent years. This method has lower complexity and higher diagnostic accuracy. In future work, we will continue to develop more portable and comfortable wearable devices, optimize the diagnostic model, and diagnose more real-time collected samples to enhance its generalization ability.

REFERENCES

[1] A. Rath, D. Mishra, G. Panda, and S. C. Satapathy, "Heart disease detection using deep learning methods from imbalanced ECG samples," *Biomed. Signal Process. Control*, vol. 68, Jul. 2021, Art. no. 102820.
[2] X. Liu, H. Wang, Z. Li, and L. Qin, "Deep learning in ECG diagnosis: A review," *Knowl.-Based Syst.*, vol. 227, Sep. 2021, Art. no. 107187.
[3] S. Hadiyoso, K. Usman, and A. Rizal, "Arrhythmia detection based on ECG signal using Android mobile for athlete and patient," in *Proc. 3rd Int. Conf. Inf. Commun. Technol. (ICoICT)*, May 2015, pp. 166–171.
[4] J. Zhuang, J. Sun, and G. Yuan, "Arrhythmia diagnosis of young martial arts athletes based on deep learning for smart medical care," *Neural Comput. Appl.*, vol. 35, no. 20, pp. 14641–14652, Jul. 2023.
[5] W. Fan, Y. Si, W. Yang, and G. Zhang, "Active broad learning system for ECG arrhythmia classification," *Measurement*, vol. 185, Nov. 2021, Art. no. 110040.
[6] D. K. Atal and M. Singh, "Arrhythmia classification with ECG signals based on the optimization-enabled deep convolutional neural network," *Comput. Methods Programs Biomed.*, vol. 196, Nov. 2020, Art. no. 105607.
[7] Neha, H. K. Sardana, R. Kanwade, and S. Tewary, "Arrhythmia detection and classification using ECG and PPG techniques: A review," *Phys. Eng. Sci. Med.*, vol. 44, no. 4, pp. 1027–1048, Dec. 2021.
[8] F. Buendía-Fuentes, M. A. Arnau-Vives, A. Arnau-Vives, Y. Jiménez-Jiménez, J. Rueda-Soriano, E. Zorio-Grima, A. Osa-Sáez, L. V. Martínez-Dolz, L. Almenar-Bonet, and M. A. Palencia-Pérez, "High-bandpass filters in electrocar-diography: Source of error in the interpretation of the ST segment," *Int. Scholarly Res. Notices*, vol. 2012, Jun. 2012, Art. no. 706217, doi: 10.5402/2012/706217.
[9] L. Shoker, S. Sanei, and J. Chambers, "Artifact removal from electroencephalograms using a hybrid BSS-SVM algorithm," *IEEE Signal Process. Lett.*, vol. 12, no. 10, pp. 721–724, Oct. 2005.
[10] Z. Zhang, H. Li, and D. Mandic, "Blind source separation and artefact cancellation for single channel bioelectrical signal," in *Proc. IEEE 13th Int. Conf. Wearable Implant. Body Sensor Netw. (BSN)*, Jun. 2016, pp. 177–182.
[11] S. Ziani, A. Jbari, L. Bellarbi, and Y. Farhaoui, "Blind maternal-fetal ECG separation based on the time-scale image TSI and SVD—ICA methods," *Proc. Comput. Sci.*, vol. 134, pp. 322–327, Jan. 2018.
[12] L. Taha and E. Abdel-Raheem, "A null space-based blind source separation for fetal electrocardiogram signals," *Sensors*, vol. 20, no. 12, p. 3536, Jun. 2020.
[13] W. Zheng, L. Hongxing, and C. Jianchun, "An adaptive filtering in phase space for fetal ECG estimation from an abdominal ECG signal and a thoracic ECG signal," *IET Signal Process*, vol. 6, no. 3, pp. 171–177, 2012.
[14] X. Xu and H. Liu, "ECG heartbeat classification using convolutional neural networks," *IEEE Access*, vol. 8, pp. 8614–8619, 2020.
[15] E. Izci, M. A. Ozdemir, M. Degirmenci, and A. Akan, "Cardiac arrhythmia detection from 2D ECG images by using deep learning technique," in *Proc. Med. Technol. Congr. (TIPTEKNO)*, Oct. 2019, pp. 1–4.
[16] Y. Zhao, J. Cheng, P. Zhan, and X. Peng, "ECG classification using deep CNN improved by wavelet transform," *Comput., Mater. Continua*, vol. 64, no. 3, pp. 1615–1628, 2020.
[17] T. Wang, C. Lu, Y. Sun, M. Yang, C. Liu, and C. Ou, "Automatic ECG classification using continuous wavelet transform and convolutional neural network," *Entropy*, vol. 23, no. 1, p. 119, Jan. 2021.
[18] Q. Wu, Y. Sun, H. Yan, and X. Wu, "ECG signal classification with binarized convolutional neural network," *Comput. Biol. Med.*, vol. 121, Jun. 2020, Art. no. 103800.
[19] Q. Yao, R. Wang, X. Fan, J. Liu, and Y. Li, "Multi-class arrhythmia detection from 12-lead varied-length ECG using attention-based time-incremental convolutional neural network," *Inf. Fusion*, vol. 53, pp. 174–182, Jan. 2020.
[20] N. Du, Q. Cao, L. Yu, N. Liu, E. Zhong, Z. Liu, Y. Shen, and K. Chen, "FM-ECG: A fine-grained multi-label framework for ECG image classification," *Inf. Sci.*, vol. 549, pp. 164–177, Mar. 2021.
[21] W. Zhao, J. Hu, D. Jia, H. Wang, Z. Li, C. Yan, and T. You, "Deep learning based patient-specific classification of arrhythmia on ECG signal," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2019, pp. 1500–1503.

[22] X. Luo, L. Yang, H. Cai, R. Tang, Y. Chen, and W. Li, "Multi-classification of arrhythmias using a HCRNet on imbalanced ECG datasets," *Comput. Methods Programs Biomed.*, vol. 208, Sep. 2021, Art. no. 106258.
[23] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.
[24] S. A. Dar, "Neural networks (CNNs) and VGG on real time face recognition system," *Turkish J. Comput. Math. Educ.*, vol. 12, no. 9, pp. 1809–1822, 2021.
[25] M. Wang, S. Wang, and P. Kong, "Simplified VGG based super resolution restoration for face recognition," in *Proc. 8th Int. Conf. Comput. Pattern Recognit.*, Oct. 2019, pp. 385–389.
[26] D. Su, Y. Li, Y. Zhao, R. Xu, B. Yuan, and W. Wu, "A face recognition algorithm based on dual-channel images and VGG-cut model," *J. Phys., Conf. Ser.*, vol. 1693, no. 1, Dec. 2020, Art. no. 012151.
[27] K. H. Cheah, H. Nisar, V. V. Yap, C.-Y. Lee, and G. R. Sinha, "Optimizing residual networks and VGG for classification of EEG signals: Identifying ideal channels for emotion recognition," *J. Healthcare Eng.*, vol. 2021, pp. 1–14, Mar. 2021.
[28] T. Kaur and T. K. Gandhi, "Automated brain image classification based on VGG-16 and transfer learning," in *Proc. Int. Conf. Inf. Technol. (ICIT)*, Dec. 2019, pp. 94–98.
[29] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," 2015, *arXiv:1511.07122*.
[30] A. L. Goldberger, L. A. N. 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, Jun. 2000.
[31] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Eng. Med. Biol. Mag.*, vol. 20, no. 3, pp. 45–50, May/Jun. 2001.
[32] S. D. Greenwald, "Development and analysis of a ventricular fibrillation detector," M.S. thesis, Dept. Elect. Eng. Comput. Sci., MIT, Cambridge, MA, USA, 1986.
[33] L. van der Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, no. 11, pp. 2579–2605, 2008.
[34] A. M. Shaker, M. Tantawi, H. A. Shedeed, and M. F. Tolba, "Generalization of convolutional neural networks for ECG classification using generative adversarial networks," *IEEE Access*, vol. 8, pp. 35592–35605, 2020.
[35] X. Xu, S. Jeong, and J. Li, "Interpretation of electrocardiogram (ECG) rhythm by combined CNN and BiLSTM," *IEEE Access*, vol. 8, pp. 125380–125388, 2020.
[36] A. Ullah, S. U. Rehman, S. Tu, R. M. Mehmood, Fawad, and M. Ehatisham-Ul-Haq, "A hybrid deep CNN model for abnormal arrhythmia detection based on cardiac ECG signal," *Sensors*, vol. 21, no. 3, p. 951, Feb. 2021.
[37] A. H. Khan, M. Hussain, and M. K. Malik, "Cardiac disorder classification by electrocardiogram sensing using deep neural network," *Complexity*, vol. 2021, pp. 1–8, Mar. 2021.
[38] H. Li, Z. Lin, Z. An, S. Zuo, W. Zhu, Z. Zhang, Y. Mu, L. Cao, and J. D. P. García, "Automatic electrocardiogram detection and classification using bidirectional long short-term memory network improved by Bayesian optimization," *Biomed. Signal Process. Control*, vol. 73, Mar. 2022, Art. no. 103424.

JUTAO WANG was born in Linyi, Shandong, China, Han nationality. He received the Ph.D. degree in education.

In December 2003, he joined the Communist Party of China. He is currently an Associate Professor and a Master's Tutor with the School of Physical Education and Health, Linyi University. Since 2009, he has chaired or undertaken nine scientific research projects and published four academic monographs. He is mainly engaged in teaching and research in sports humanities and sociology. His research interests include data analysis, sports health promotion, and sports industry management.

FUCHUN ZHANG was born in Liaocheng, Shandong, China, in 1997. He received the bachelor's degree from Taishan University, in 2020, and the master's degree from Linyi University, in 2023.

He has published four SCI articles, one patent, and three Chinese core papers. He is mainly engaged in electrocardiogram signal detection and medical image processing. His research interests include signal processing, data analysis, and artificial intelligence.

MENG LI was born in Jining, Shandong, China, Han nationality. She received the bachelor's and master's degrees from Linyi University, in 2020 and 2023, respectively.

She has published three SCI articles, one patent, three Chinese core papers, and participated in two scientific research projects. She is mainly engaged in image processing. Her research interests include ECG signal processing and artificial intelligence.

BAIYANG WANG was born in Jinan, Shandong, China, Han nationality. He received the master's degree from Linyi University, in 2023.

He has published four SCI articles and two patents. He is mainly engaged in signal processing. His research interests include signal processing and artificial intelligence.