Automated atrial fibrillation and ventricular fibrillation recognition using a multi-angle dual-channel fusion network

Weiyi Yang [a,*], Di Wang [b], Wei Fan [c], Gong Zhang [d], Chunying Li [e], Tong Liu [f]

[a] School of Artificial Intelligence, Nanjing University of Information Science and Technology, Nanjing 210044, China
[b] School of Electronics & Information Engineering, Tiangong University, Tianjin 300387, China
[c] College of Communication Engineering, Jilin University, Changchun 130012, China
[d] School of Computer Science, Hubei University of Technology, Wuhan 430068, China
[e] The Cancer Hospital of the University of Chinese Academy of Sciences (Zhejiang Cancer Hospital), Hangzhou Institute of Medicine (HIM), Chinese Academy of Sciences, Hangzhou, Zhejiang 310022, China
[f] School of Information and Electrical Engineering, Ludong University, Yantai 264025, China

ARTICLE INFO

ABSTRACT

Atrial fibrillation (AFIB) and ventricular fibrillation (VFIB) are two common cardiovascular diseases that cause numerous deaths worldwide. Medical staff usually adopt long-term ECGs as a tool to diagnose AFIB and VFIB. However, since ECG changes are occasionally subtle and similar, visual observation of ECG changes is challenging. To address this issue, we proposed a multi-angle dual-channel fusion network (MDF-Net) to automatically recognize AFIB and VFIB heartbeats in this work. MDF-Net can be seen as the fusion of a task-related component analysis (TRCA)-principal component analysis (PCA) network (TRPC-Net), a canonical correlation analysis (CCA)-PCA network (CPC-Net), and the linear support vector machine-weighted softmax with average (LS-WSA) method. TRPC-Net and CPC-Net are employed to extract deep task-related and correlation features, respectively, from two-lead ECGs, by which multi-angle feature-level information fusion is realized. Since the convolution kernels of the above methods can be directly extracted through TRCA, CCA and PCA technologies, their training time is faster than that of convolutional neural networks. Finally, LS-WSA is employed to fuse the above features at the decision level, by which the classification results are obtained. In distinguishing AFIB and VFIB heartbeats, the proposed method achieved accuracies of 99.39 % and 97.17 % in intra- and inter-patient experiments, respectively. In addition, this method performed well on noisy data and extremely imbalanced data, in which abnormal heatbeats are much less than normal heartbeats. Our proposed method has the potential to be used as a diagnostic tool in the clinic.

1. Introduction

In recent years, the number of cardiovascular disease (CVD) patients has been increasing worldwide [1]. According to the World Health Organization (WHO) [2], CVDs cause nearly 18 million deaths every year, accounting for the highest proportion of deaths due to various non-communicable diseases, such as diabetes, malignant tumors, and tuberculosis. One of the most serious CVDs is arrhythmia, which is a general term for a heterogeneous set of types of abnormal cardiac electrical activity. Among various arrhythmias, atrial fibrillation (AFIB) and ventricular fibrillation (VFIB) [3–6] are two types of important diseases. They are characterized by rapid and irregular contraction of the atria and ventricles at an abnormal rate. In general, the occurrence of

AFIB and VFIB may lead to the decline or even complete disruption of cardiac blood pumping function, which may eventually cause death. Hence, it is essential to identify potential AFIB and VFIB patients at an early stage.

One of the most common diagnostic tools for AFIB and VFIB is electrocardiograms (ECGs). ECG is a kind of periodic electrical signal and can be collected conveniently through noninvasive methods. In general, a one-cycle ECG signal, namely, a heartbeat, is composed of a P wave, a QRS complex, and a T wave [7]. Among them, the P wave captures the potential changes during depolarization of the right and left atria, while the QRS complex and T wave reflect the changes in the two ventricles during depolarization and repolarization, respectively. Hence, ECGs contain rich information reflecting the state of the heart and are

Table 1
The relevant studies for classifying AFIB and VFIB.

Author	Source	Method	Accuracy
AFIB			
Martis et al. (2013) [15]	MIT-BIH arrhythmia database (MITdb), MIT-BIH atrial fibrillation database (AFdb)	Independent component analysis, Naive Bayes (NB) classifier	99.33 % (NB)
Jiang et al. (2019) [16]	A real Paroxysmal Atrial Fibrillation dataset	17-Layer one-dimensional convolution neural networks (CNN)	95.8 %
Radhakrishnan et al. (2021) [17]	Physionet Computing in Cardiology Challenge 2017, MITdb, AFdb, MIT-BIH Normal Sinus Rhythm database (NSRdb)	Time-frequency representation (TFR), deep convolutional bidirectional long short-term memory (BLSTM) network	99.18 %
Maghawry et al. (2021) [18]	A real Paroxysmal Atrial Fibrillation dataset	QRS complex duration, RR interval, sparse representation, Extreme Learning Machine classification technique	97 %
Wei et al. (2022) [19]	PhysioNet Computing in Cardiology Challenge 2017	Spectrogram, Fine-tuned EfficientNet B0	97.3 %
Rahul et al. (2022) [20]	NSRdb, the MIT-BIH atrial fibrillation collected from Boston's Beth Israel Hospital	Time-frequency representation, Bidirectional long short-term memory (BLSTM) network	98.5 %
Andersen et al. (2019) [21]	AFdb, MITdb, and NSRdb	RR intervals, Convolutional- and Recurrent-Neural Networks (CNN and RNN)	98.96 %
Parsi et al. (2021) [22]	Atrial fibrillation prediction database (AFPdb)	RR intervals, Poincaré representation, Linear kernel SVM	98 %
Lee et al. (2020) [23]	Cardiology Center of Soonchunhyang Bucheon Hospital (SBH)	Root-mean square of the successive differences (RMSSD) and the Shannon entropy (ShE), ensemble methods of neural networks (NNs), k-nearest neighbors (kNN), and decision trees (DT)	97.94 %
Lee et al. (2021) [24]	Real data collected by themselves	RMSSD and the Shannon entropy of the RR interval (ECG features), kNN, DT, and NNs	98.9 % (KNN) 97.8 % (DT), 98.9 % (NN)
VFIB			
Panigrahy et al. (2021) [25]	MIT-BIH malignant ventricular arrhythmia database, arrhythmia database, and CUDB database	Support vector machine (SVM), adaptive boosting (AdaBoost) and differential evolution (DE) algorithms	98.2 %
AFIB and VFIB			
Fujita et al. (2019) [26]	MIT-BIH malignant ventricular arrhythmia database (VFdb), AFdb, MITdb.	6-Layer deep CNN	97.78 %
Fujita et al. (2019) [27]	VFdb, AFdb and MITdb	8-Layer deep CNN	98.45 %

Table 1 (*continued*)

Author	Source	Method	Accuracy
Tripathi et al. (2022) [28]	MITdb, VFdb, Fantasia database, AF termination challenge database, CU ventricular tachyarrhythmia database	Superlet transform (SLT), DenseNet-201	96.2 %

often used by medical staff to diagnose AFIB and VFIB [8,9]. However, the changes in ECG waveforms are occasionally subtle and similar, making it difficult for medical staff to visually identify ECG abnormalities [10]. Hence, it is important to develop an accurate and fast computer-aided system to detect cardiac abnormalities [11–14].

To realize automatic detection of AFIB and VFIB, relevant researchers usually develop ECG classification systems by using signal processing and machine learning technologies. In general, the common ECG classification system is composed of three stages: preprocessing, feature extraction, and classification. At the preprocessing stage, the noise in raw ECGs is first removed by the signal processing method. Then, the noise-free ECGs are divided and normalized to obtain a series of mean-removed heartbeats. Next, the feature extraction stage is employed to mine key discriminatory information from heartbeats that is conducive to identifying cardiac abnormalities. In previous studies, common features of heartbeats mainly included time-domain features [18,21,22], frequency-domain features [17,19,20,28], statistical features [15,23,24], and deep features [17,19–21,26–28]. At the stage of classification, support vector machine (SVM) [22,25], K-nearest neighbor (KNN) [23,24], and different kinds of neural networks [16,17,20,21,23,24,26,27] have been employed as classifiers to process features of heartbeats to obtain the final labels. Table 1 shows the relevant studies for classifying heartbeats of AFIB and VFIB patients.

Although the above studies achieved excellent performance in identifying AFIB and VFIB, some issues have not been addressed [15–28], by which these methods may not be used in real scenarios. 1) First, most of the studies failed to verify the applicability of their methods in recognizing new patients. Specifically, these studies adopted the data of the same patients as the training and testing data (intra-patient experimental data) to verify the performance of their methods. For example, in 2019, Fujita et al. [27] proposed an 8-layer deep CNN based ECG classification method for recognizing cardiac diseases including AFIB and VFIB. This method achieved an accuracy of 98.43 % in the intra-patient experiment using 10-fold cross validation. However, the performance of this method has not been validated in inter-patient experiments. In 2022, Rahul et al. [20] proposed a time-frequency representation and BLSTM based ECG classification method. They only tested the performance of their method in the intra-patient experiment and not in the inter-patient experiment. However, in real scenarios, it is essential to adopt the proposed method to recognize the ECG data of new patients, for whom the heartbeats have not been used as training data. Due to the differences in the ECGs among individuals, the models developed based on intra-patient experiments are likely to cause overfitting in recognizing the ECGs of new patients. Hence, it is particularly important to ensure the excellent performance of the proposed method in inter-patient experiments. 2) Second, the ability of the proposed methods to process noisy data has not been ensured. Specifically, many researchers have used signal processing algorithms to remove noise in ECGs at the preprocessing stage. However, due to the difference in collection equipment and environments, the levels of noise between ECGs are usually different, causing the process of completely removing noise from ECGs to inevitably lead to the loss of information useful for recognizing cardiac abnormalities. Hence, to avoid excessive dependence on an accurate denoising algorithm, it is necessary to ensure that the method has certain noise robustness when developing an ECG recognition method; that is, it can effectively identify ECG data containing weak noise.

2

Table 2
Databases used.

Database	MIT-BIH Normal sinus rhythm (NSRDB)	MIT-BIH Atrial Fibrillation Database (AFDB)	MIT-BIH Malignant Ventricular Ectopy Database (VFDB)
Class	Normal	AFIB	VFIB
ECG	ECG1 andECG2	ECG1 and ECG2	ECG1 and ECG2
Sampling rate (Hz)	128	250	250
Number of records	18	23	4
ID	16265, 16272, 16273, 16420, 16483, 16539, 16773, 16786, 16795, 17052, 17453, 18177, 18184, 19088, 19090, 19093, 19140, 19830	04015, 04043, 04048, 04126, 04746, 04908, 04936, 05091, 05121, 05261, 06426, 06453, 06995, 07162, 07859, 07879, 07910, 08215, 08219, 08378, 08405, 08434, 08455	422, 424, 426, 430

Although some studies did not adopt a denoising algorithm at the pre-processing stage, most of the noise in the raw ECGs was basically filtered out by the database creators. Hence, the results in these studies are not sufficient to prove the noise robustness of the proposed method. 3) Third, researchers should ensure that the proposed method can effectively process multi-level extremely imbalanced data, in which the number of abnormal ECGs is far less than that of normal ECGs. Specifically, the number of patients is usually less than that of normal individuals in real data settings, causing the numbers of ECG data between categories to be extremely imbalanced. Since machine learning methods tend to assign testing data to the category with the highest proportion of data used for training classification models, the small proportion of data is usually difficult to classify accurately. Hence, it is particularly important to verify the applicability of the proposed method to extremely imbalanced data. Although most researchers have not adopted fully balanced data, they have usually constructed experimental datasets using more abnormal ECGs than normal ECGs, which is obviously different from the data distribution in real scenarios. In addition, the degrees of imbalance adopted in these studies are usually less than an order of magnitude, indicating that there is only a small difference in the number of heartbeats between different categories. Hence, it is essential to verify the performance of the proposed method on extremely imbalanced experimental datasets, in which there are far fewer abnormal ECGs than normal ECGs. 4) Fourth, the proposed methods in the above studies can only process single-lead ECGs, resulting in the inability to utilize diversified information in multi-lead ECGs. Specifically, ECGs can be collected from the limb or chest leads. Among them, limb- and chest-lead ECGs record the state on the cardiac horizontal plane and sagittal plane, respectively, by which the single-lead ECG used in the above studies can only provide limited discriminant features. Hence, it is important to make full use of multi-lead ECGs when diagnosing cardiac abnormalities.

To address the above issues, we propose a novel ECG classification method using a multi-angle dual-channel fusion network (MDF-Net). MDF-Net can be seen as the combination of the TRCA-PCA network (TRPC-Net) and the CCA-PCA network (CPC-Net), which are employed to extract deep features of multi-lead ECGs from different angles. Hence, MDF-Net can achieve two stages of feature-level information fusion. TRPC-Net is mainly composed of a TRCA convolutional layer, a PCA convolutional layer, and a fully connected layer. The aim of the TRCA convolutional layer is to extract the task-related components from multi-lead ECGs by maximizing the reproducibility of data under the same task [29]. Then, the PCA convolutional layer further weakens the correlation of ECGs between categories. Finally, the fully connected layer processes the outputs of the PCA convolutional layer to yield the feature vectors,

called inter-lead task-related features. CPC-Net consists of a CCA convolutional layer, a PCA convolutional layer, and a fully connected layer. Among them, the CCA convolutional layer is employed to extract the correlation information from the multi-lead ECGs, while the functions of the PCA convolutional layer and the fully connected layer are similar to those in TRPC-Net. Hence, we regard the CPC-Net-based features as the inter-lead correlation features. Similar to convolutional neural networks (CNNs), MDF-Net can extract the deep discriminatory features from ECGs layer by layer. However, since the kernels of MDF-Net can be directly extracted by the TRCA, CCA, and PCA methods, MDF-Net has a faster training speed than CNNs. Next, a linear support vector machine (LSVM) is adopted to process the inter-lead task-related features and correlation features to yield two sets of primary results. These primary results are then fused by weighted softmax with average (WSA) to obtain the final labels. Since TRCA, CCA and PCA [30–32] can remove certain noise, MDF-Net, which combines the above methods using cascaded and convolutional structures, has significant noise robustness. In addition, since TRPC-Net and CPC-Net can extract features of multi-lead ECGs from different angles, including enhancing correlation and reproducibility under the same task, the complete MDF-Net has excellent feature mining ability, by which the key information conducive to recognizing cardiac abnormalities can be effectively extracted from extremely imbalanced and inter-patient ECG data. In this study, we adopted the ECGs of normal, AFIB, and VFIB patients in NSRdb, AFdb, and VFdb as experimental data to verify the performance of the proposed method.

The main contributions of this work are as follows:

1) An MDF-Net-based classification method, consisting of the TRPC-Net, CPC-Net, and LS-WSA fusion methods, was developed to automatically diagnose AFIB and VFIB heartbeats.
2) Inter-lead task-related feature extraction: We developed a TRPC-Net consisting of a TRCA convolutional layer, a PCA convolutional layer, and a fully connected layer. It is used to conduct the first stage of feature-level information fusion for multi-lead ECGs.
3) Inter-lead correlation feature extraction: CPC-Net is constructed by a CCA convolutional layer, a PCA convolutional layer, and a fully connected layer. It can realize the second stage of feature-level information fusion for multi-lead ECGs.
4) Decision-level information fusion: We designed an LSVM-weighted softmax with average (LS-WSA) method, including the LSVM, weight assignment, softmax function and average methods. It is employed to make decisions by conducting a stage of decision-level information fusion.

2. Materials used

2.1. Databases

We adopt the ECGs of normal, atrial fibrillation (AFIB), and ventricular fibrillation (VFIB) patients collected from three public databases as experimental data. These databases are the MIT-BIH Normal Sinus Rhythm Database (NSRdb), MIT-BIH Atrial Fibrillation Database (AFdb), and MIT-BIH Malignant Ventricular Ectopy Database (VFdb). Details are provided in Table 2.

1) NSRdb includes ECG signals without abnormalities, which were obtained from 18 individuals.
2) AFdb contains records for 23 available AFIB patients.
3) In VFdb, ECGs were obtained from 4 patients with VFIB symptoms.

2.2. Dataset construction

In this work, ECGs collected from 45 individuals in NSRdb, AFdb, and VFdb were used as experimental data. For each patient, we randomly selected 400 heartbeats as the experimental data. Since only 99 and 100 R waves were detected from the ECG records of two

Table 3
Data distribution for sets A–F.

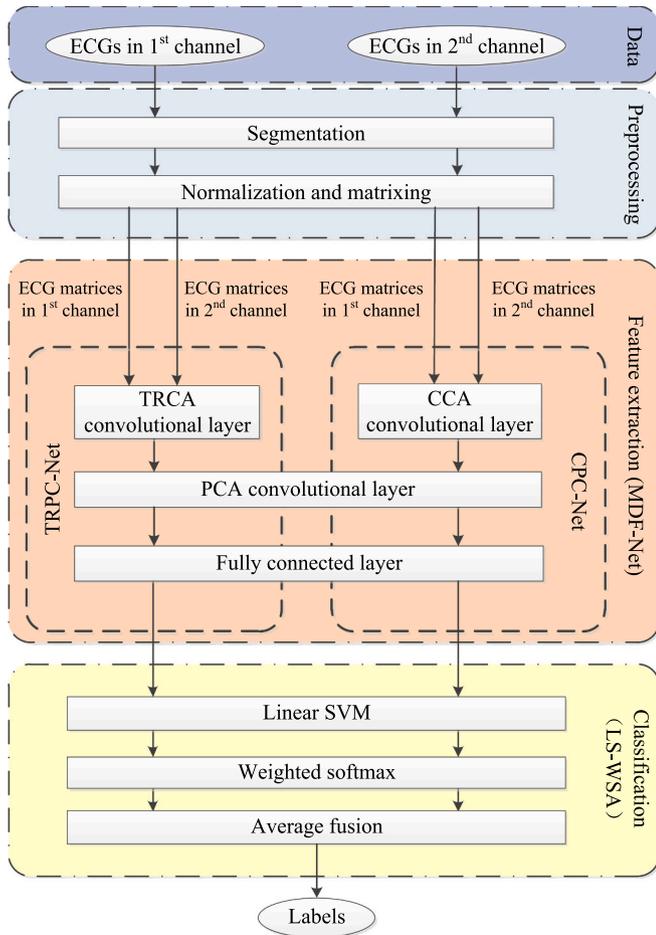| Datasets	Number of heartbeats		
	Normal	AFIB	VFIB
A/D	7200	8899	
B/E	7200		1300
C/F	7200	8899	1300

Fig. 1. The detailed process of the proposed method.

individuals (ID = 08405, 424), only 99 and 100 heartbeats were obtained. Then, we constructed six experimental datasets. The data distributions of these datasets are shown in Table 3.

1) Sets A–C were constructed to evaluate the performance of the proposed method on the intra-patient experiments. Among them, set A contains normal and AFIB heartbeats, while set B consists of normal and VFIB heartbeats. In addition, set C is composed of normal, AFIB and VFIB heartbeats. For each experiment, a random part of the heartbeats was used as training data, and the remainder was used as testing data.
2) Sets D–F were designed to verify the performance of the proposed method on the inter-patient experiments. Among them, set D contains normal and AFIB heartbeats, while set E consists of normal and VFIB heartbeats. In addition, set F includes normal, AFIB and VFIB patients. For each experiment, the training and testing data were obtained from different patients.

3. Methods

Fig. 1 shows the main flow of the proposed method. It consists of three main steps: preprocessing, feature extraction, and classification. In the preprocessing step, two-lead ECGs are first segmented into heartbeats. Then, all the heartbeats are normalized and converted to ECG matrices. At the step of feature extraction, a multi-angle dual-channel fusion network (MDF-Net) consisting of TRPC-Net and CPC-Net is designed to extract deep features of two-lead ECG matrices from different angles. TRPC-Net was developed to extract the inter-lead task-related features, while CPC-Net was employed to mine the inter-lead correlation features. At the classification step, the LSVM is adopted to process the above two types of features to yield two sets of decision values, which are then fused by the weighted softmax and average (WSA) method to obtain the final labels.

3.1. Preprocessing

3.1.1. ECG segmentation

In this work, we employ the Pan-Tompkins detection method to locate the positions of R peaks. Then, we use the ECG fragment alignment (EFA) method to divide the long-term ECG signals into short-term heartbeats. Specifically, for each R peak, we set the previous R peak as the R0 point and the next R peak as the R2 point. Then, the sample at 0.1 s after the R0 point is set as A1, while the sample at 0.1 s before the R1 point is set as A2. At the same time, the samples at 0.06 s before and after point R are set as points B1 and B2, respectively. Here, we set A1–B1, B1–B2, and B2–A2 as three segments. Next, we resample each segment to 100 points. Finally, these resampled segments are concatenated into a heartbeat with 300 sample points.

3.1.2. Normalization and matrixing process

Here, we use the min-max normalization method to convert the amplitude range of sample points in each heartbeat to [0,1]. Then, since the subsequent feature extraction method is designed to process two-dimensional (2-D) data, we convert each one-dimensional heartbeat to a 2-D ECG matrix with a size of by using the reshape function in MATLAB.

3.2. Feature extraction

3.2.1. MDF-Net

To effectively mine key discriminant information in two-lead ECGs, we adopt the multi-angle dual-channel fusion network (MDF-Net) as the feature extraction method. MDF-Net can be regarded as the combination of TRPC-Net and CPC-Net. TRPC-Net is used to maximize the reproducibility under the same task from two-lead ECGs, while CPC-Net is adopted to extract the correlation information between two-lead ECGs. Hence, we refer to TRPC-Net-based features and CPC-Net-based features as inter-lead task-related features and correlation features, respectively. The structure of MDF-Net is shown in Fig. 2. It consists of five processing layers: a scanning layer, a TRCA convolutional layer, a CCA convolutional layer, a PCA convolutional layer, and a fully connected layer.

1) Scanning layer

In this layer, a patch with a size of $k_1 \times k_2$ is used to scan the each-lead ECG matrix to obtain several matrices, which are then converted to vectors. Next, all vectors are combined into a new matrix, namely, the pending matrix.

2) TRCA convolutional layer

TRCA technology was developed to extract task-related components and task-unrelated components from data in different channels. In this work, the aim of the TRCA convolutional layer is to explore a series of
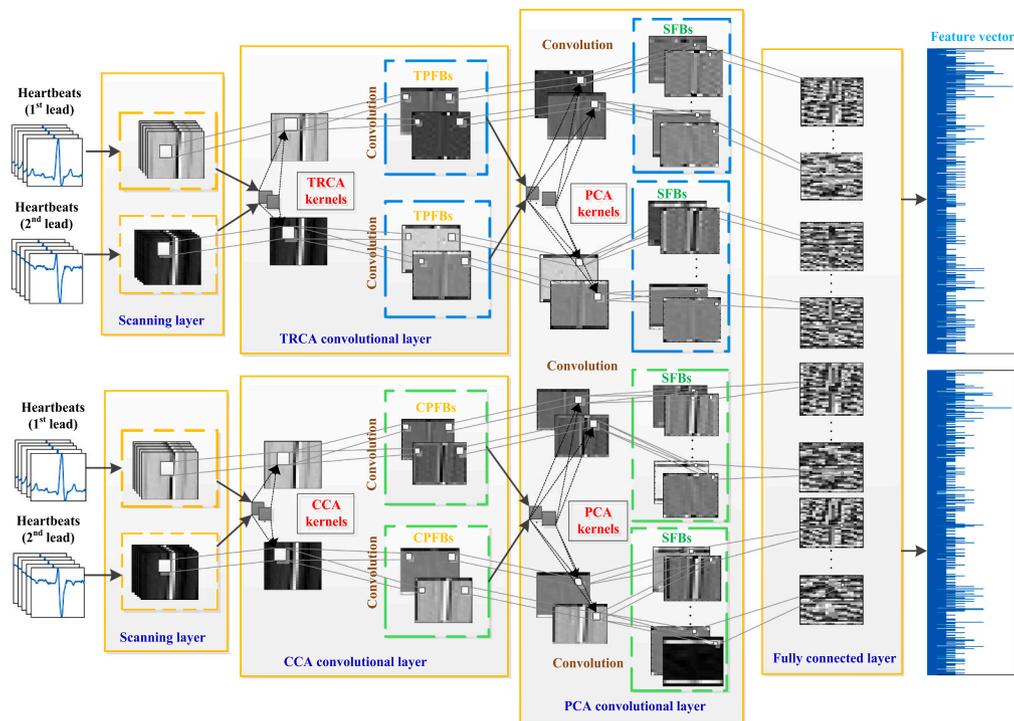
Fig. 2. The structure of MDF-Net.

task-related components from multi-lead ECG matrices. To address these issues, we first adopt TRCA technology to process the above pending matrix between multi-lead ECGs to yield L_1 TRCA filters. These filters are then convolved with the multi-lead ECG matrices to obtain TRCA-based primary feature blocks (TPFBs).

3) CCA convolutional layer

CCA technology was designed for mining correlations between data. In the CCA convolutional layer, we need to extract the correlation features from the multi-lead ECG matrices. Specifically, CCA technology is used to extract L_2 CCA filters by processing multi-lead ECGs. Then, these filters are convolved with the multi-lead ECG matrices to obtain CCA-based primary feature blocks (CPFBs).

4) PCA convolutional layer

Similar to the scanning layer, a $k_1 \times k_2$ patch scans TPFB or CPFB to obtain a series of mean-removed patches, which are then extended into vectors. To make the process clear, we only employ TPFBs as an example to describe the subsequent steps. After processing all the TPFBs by the above operation, all the obtained vectors are converted to second-order pending (SOP) matrices.

Then, PCA filters are used to extract L_3 PCA filters from the SOP matrices. Next, these filters are convolved with the multi-lead ECG matrices to obtain L_3 second-order feature blocks (SFBs). After processing all TPFBs and CPFBs, we obtain L_3 TPFB-based second-order feature blocks (TSFBs) and L_3 CPFB-based second-order feature blocks (CSFBs).

5) Output layer

In this layer, by using hash coding, each TSFB or CSFB is converted to a binary matrix, which is then calculated to a decimal matrix. Next, a $u_1 \times u_2$ window with overlap R is used to scan the decimal matrix to obtain a series of blocks. Finally, by using the histogram statistical method, all blocks are processed to yield the feature vector. After pro-

cessing all TSFBs and CSFBs, we obtain two sets of feature vectors $f_{i,TRC}$ and $f_{i,CCA}$.

A detailed description of the MDF-Net algorithm is shown in Appendix A with Table 13.

3.3. Classification

Here, we proposed the LSVM-weighted softmax function and average method (LS-WSA) to process the features $f_{i,TRC}$ and $f_{i,CCA}$ to obtain the final labels. The specific process of LS-WSA is as follows.

3.3.1. LSVM

SVM is a common supervised learning model and has been widely used in pattern recognition as a classifier model. It can establish the boundary of data between categories by building a decision hyperplane. In general, to cope with different situations, SVMs can employ different kernel functions, such as polynomial kernels, radial basis function (RBF) kernels, Laplace kernels, and linear kernels. Since the features extracted by MDF-Net have high dimensions, the SVM with a linear kernel function, namely, LSVM, is adopted as the classifier model. In our experiments, the LSVM is implemented using the Liblinear toolkit in MATLAB. In addition, the penalty parameter C is set to 1. After processing all the $f_{i,TRC}$ or $f_{i,CCA}$ with the original label h_i, we will obtain two sets of decision values $v_{i,h,TRC}$ or $v_{i,h,CCA}, h_i = 1, 2, \cdots, H$, where h_i expresses the corresponding original label, and decision values $v_{i,h,TRC}$ and $v_{i,h,CCA}$ express the distance between the sample and the decision hyperplane.

3.3.2. Decision fusion

(1) Weight assignment

Due to differences in the expression abilities of features $f_{i,TRC}$ and $f_{i,CCA}$, the decision values $v_{i,h,TRC}$ and $v_{i,h,CCA}$ should be assigned different importance. Hence, the LSVM-based sensitivities (SEs) of $f_{i,TRC}$ or $f_{i,CCA}$ for each class are used as the reference weights for $v_{i,h,TRC}$ or $v_{i,h,CCA}$, by which we will obtain $v_{i,h,TRC}^{weighted} = SE_{z,TRC} \times v_{i,h,TRC}$ or $v_{i,h,CCA}^{weighted} = SE_{z,CCA} \times$

Table 4
The optimal parameters of the proposed method.

Model	Layer	Parameters	Search range	Optimized values
Preprocessing	Matrixing process	m	{2:30}	15
		n	{2:30}	20
MDF-Net	Scanning layer	$k_1(k_2)$	{3:2:9}	3(3)
	TRCA convolutional layer	L_1	{3:1:10}	6
	CCA convolutional layer	L_2	{3:1:10}	6
	PCA convolutional layer	L_3	{3:1:10}	7
	Output layer	$u_1 \times u_2$	{3:2:9}	7×7
		R	{0.34:0.2:0.9}	0.5

$v_{i,h,CCA}$, where z represents the serial number of the label.

(2) Probability assignment

Here, according to Eq. (15), the softmax function (SF) is adopted to project $v_{i,h,TRC}^{weighted}$ and $v_{i,h,CCA}^{weighted}$ to two sets of probabilities $s_{i,h,TRC}$ and $s_{i,h,CCA}$, respectively.

$$s_{i,h}^k = \frac{e^{v_{i,h}}}{\sum_h e^{v_{i,h}}} > 0 \qquad (15)$$

(3) Decision fusion

The average method is employed to fuse the SF-based probabilities to obtain the overall decision probability $s_{i,h}$ as per Eq. (16).

$$s_{i,h} = \frac{s_{i,h,TRC} + s_{i,h,CCA}}{2} > 0 \qquad (16)$$

For each category, the label with the highest probability $s_{i,h}$ is used as the final label.

3.4. K-fold cross validation

1) Intra-patient experiments: We adopted ten-fold cross validation to ensure the reliability of the experimental results. Specifically, we randomly divided the data into ten parts. For each-fold experiment, a rotating part of the data was used as the testing data, while the others were employed as the training data.
2) Inter-patient experiments: Leave-one-out cross validation was adopted for the inter-patient experiments. For each experiment, the heartbeats of a rotating patient were used as the testing data, while the others were employed as the training data.

3.5. Evaluation indicators

In this work, we adopted the confusion matrix including true positives (TPs), true negatives (TNs), false positives (FPs), and false negatives (FNs) as the primary results. Based on them, we calculated the overall accuracy (OA), positive predictive value (PPV), sensitivity (SE), specificity (SPE), F1-score, and accuracy (AC) as the evaluation indicators, which were calculated as per Eqs. (17)–(22).

$$OA = \frac{\text{Correctly classified heartbeats}}{\text{Total number of heartbeats}} \times 100\% \qquad (17)$$

$$PPV = \frac{TP}{TP + FP} \times 100\% \qquad (18)$$

$$SE = \frac{TP}{TP + FN} \times 100\% \qquad (19)$$

$$SPE = \frac{TN}{TN + FP} \times 100\% \qquad (20)$$

$$F1 - score = \frac{2 \times PPV \times SE}{PPV + SE} \times 100\% \qquad (21)$$

$$AC = \frac{TP + TN}{TP + FP + TN + FN} \times 100\% \qquad (22)$$

4. Results and discussion

4.1. Experimental setup

In this work, a personal computer with an Intel® Core(TM) i7-8750H and a 16G RAM CPU was used as the experimental equipment. All experiments were conducted on MATLAB 2018a. For the proposed method, all optimized parameters, which were obtained from validation data through the grid search method, are shown in Table 4.

4.2. The results based on the intra-patient experiments

4.2.1. The results on sets A–C
Tables 5–6 record the classification results on intra-patient experiments. Among them, set A contains normal and AFIB heartbeats, while set B consists of normal and VFIB heartbeats. According to Table 5, the proposed method realized OA, AC, SE, PPV, SPE, and F1-score of over 99.9 % in distinguishing normal and AFIB heartbeats. In terms of classifying normal and VFIB heartbeats, Table 6 shows that most of the obtained indicators exceeded 99.9 % except for the PPV of 99.69 % and F1-score of 99.8 % when recognizing AFIB. Figs. 3–4 show the results for different folds on sets A–B. According to Figs. 3–4, all of the OA, average SE, average PPV, average SPE, and average F1-score obtained by our method are more than 99 %, indicating that the performance for all the folds is at a high level.

Table 7 records the classification results on set C for distinguishing normal, AFIB and VFIB heartbeats. According to Table 7, an OA of 99.39 % was achieved with misclassification rates of 0.07 % for normal heartbeats, 0.74 % for AFIB heartbeats, and 2.69 % for VFIB heartbeats. For each class, the obtained AC and SPE were greater than 99 %. In terms of recognizing normal and AFIB heartbeats, all of the obtained indicators exceeded 99.2 %. For identifying VFIB heartbeats, the proposed method achieved an SE of 97.31 %, a PPV of 95.26 %, and an F1-score of 96.27 %. Fig. 5 shows the results for different folds on set C. It can be seen that all the values are over 97 %, meaning that the performance of the proposed method is reliable and stable. The above results indicate that the proposed method exhibited excellent performance in classifying normal, AFIB and VFIB in intra-patient experiments.

4.2.2. The results on sets A–C with multi-level extremely imbalanced data
Here, we verified the performance of the proposed method in processing multi-level extremely imbalanced heartbeats. Fig. 6 shows the classification results on sets A–C with multi-level extremely imbalanced data. The number of normal heartbeats was N times that of AFIB and VFIB heartbeats, where N represents the level of imbalance. As shown in Fig. 6, most of the indicators decreased slightly as N increased. However, all the values of SE, PPV, SPE, and F1-score achieved by the proposed method were more than 94 % on sets A–C when the N value changed from 6 to 40. In particular, when N was 20, all the indicators for AFIB and VFIB still exceeded 99 % on sets A–B. Although the SE, PPV, and F1-score for AFIB and VFIB on set C were slightly less than 98 % (over 96 %), the obtained SPE was still greater than 99.8 %. Hence, the proposed method has excellent performance in detecting normal, AFIB and VFIB in intra-patient experiments containing multi-level extremely

Table 5
Confusion matrix for set A (normal and AFIB classes).

N = 1		Predicted		AC (%)	SE (%)	PPV (%)	SPE (%)	F1-score (%)
		Normal	AFIB					
Original	Normal	7193	7	99.94	99.9	99.97	99.98	99.93
	AFIB	2	8897	99.94	99.98	99.92	99.9	99.95
OA (%)				99.4				

Table 6
Confusion matrix for set B (normal and VFIB classes).

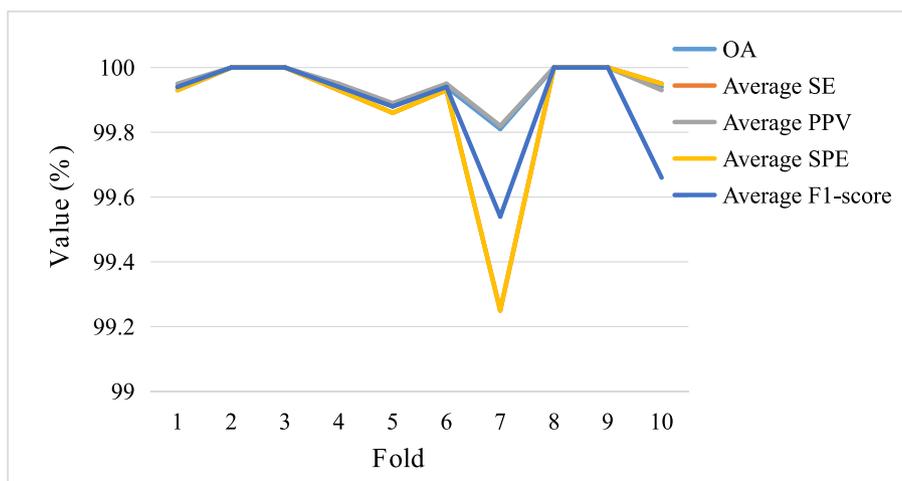| N = 1		Predicted		AC (%)	SE (%)	PPV (%)	SPE (%)	F1-score (%)
		Normal	AFIB					
Original	Normal	7196	4	99.94	99.94	99.99	99.92	99.96
	VFIB	1	1299	99.94	99.92	99.69	99.94	99.8
OA (%)				99.94				

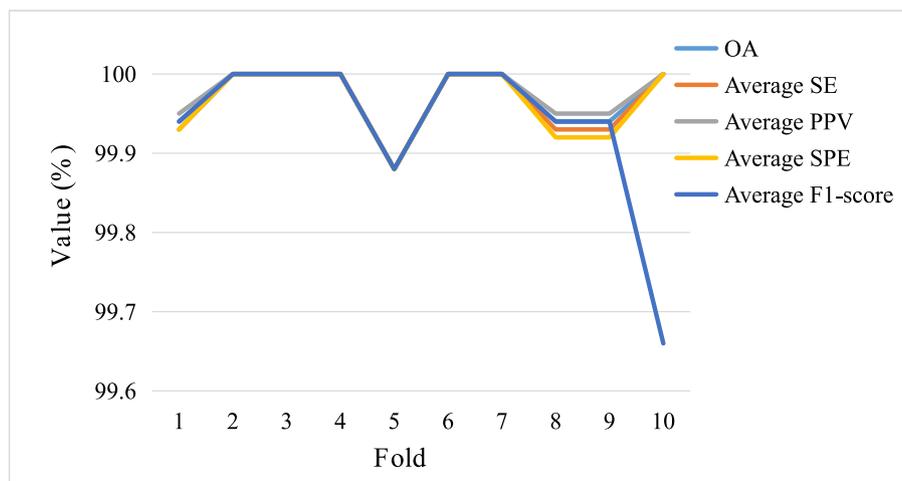**Fig. 3.** Results for different folds on set A.

Fig. 4. Results for different folds on set B.

imbalanced data, in which abnormal heatbeats are much less than normal heartbeats.

4.3. The results based on the inter-patient experiment

Tables 8–10 record the classification results on sets D–F. According to Table 8, the proposed method achieved an OA of 99.84 % in distinguishing normal and AFIB heartbeats. In addition, all the values of PPV, SE, SPE and F1-score for normal and AFIB heartbeats exceeded 99.7 %. Table 9 shows the confusion matrix obtained on set E, including normal and VFIB heartbeats. In total, the proposed method only misclassified 0.61 % of heartbeats containing 0.08 % of normal heartbeats

Table 7

Confusion matrix for set C (normal, AFIB and VFIB classes).

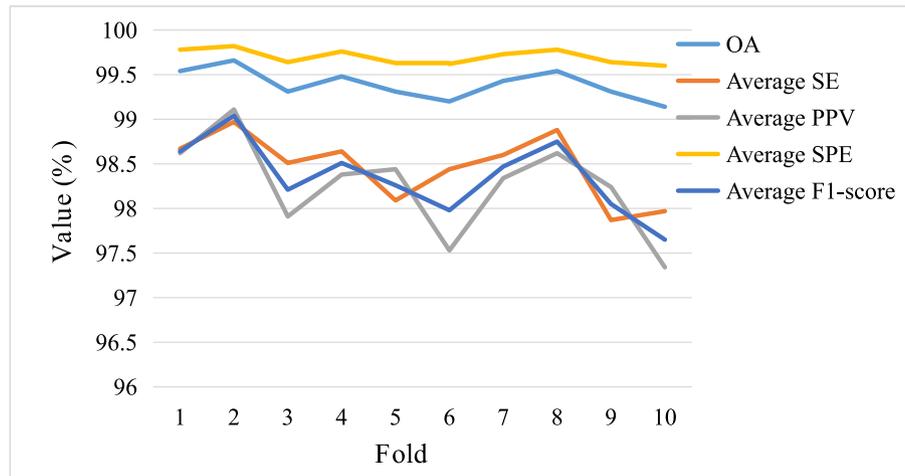|		Predicted			AC (%)	SE (%)	PPV (%)	SPE (%)	F1-score (%)
		Normal	AFIB	VFIB					
Original	Normal	7195	2	3	99.93	99.93	99.92	99.94	99.92
	AFIB	6	8833	60	99.41	99.26	99.58	99.56	99.42
	VFIB	0	35	1265	99.44	97.31	95.26	99.61	96.27
OA (%)					99.39				

Fig. 5. Results for different folds on set C.

and 3.54 % of VFIB heartbeats. In terms of PPV and F1-score, the values for all classes exceeded 97 %.

Table 10 shows the classification results on set F containing normal, AFIB, and VFIB heartbeats. Overall, an OA of 97.17 % was obtained by the proposed method. In recognizing normal heartbeats, the proposed method yielded values of over 99.6 % for all the indicators. In terms of identifying AFIB heartbeats, the obtained AC, SE, and F1-score exceeded 97 %, while PPV and SPE reached 95.97 % and 95.66 %, respectively. Although the SE for recognizing VFIB heartbeats was only 72.23 %, which may stem from the small number of VFIB patients (only three) used to train the model, the AC, PPV, and SPE for recognizing VFIB heartbeats still exceeded 90 %. Based on the above results, the proposed method can effectively classify normal, AFIB, and VFIB heartbeats in inter-patient experiments.

4.4. The results of the proposed method in processing multi-level noisy heartbeats

In this work, we verified the ability of the proposed method to process heartbeats with multi-level noise. To achieve this goal, we added different levels of noise to the heartbeats in sets A–F using the awgn function in MATLAB, as shown in Fig. 7. The ∞ db means that no additional noise was added to the ECG. According to Fig. 7, with the reduction in the SNR, the distortion of the heartbeat gradually increased. In particular, the morphological characteristics of ECG waveforms of normal, AFIB and VFIB patients are hardly recognized by the naked eye when the SNR is 6 dB. Fig. 8 shows the classification results on multi-level noisy data. The OAs obtained for all the sets exceeded 90 % when the SNR was 18 dB. Although the P wave, QRS wave and T wave of all the ECGs were severely distorted when the signal-to-noise ratio was 12 dB, the proposed method still achieved OAs of over 85 % on set F and over 90 % on sets A–E. In summary, ECGs with different levels of noise can be effectively classified, indicating that the proposed method has acceptable noise robustness.

4.5. Comparison and analysis

4.5.1. Comparison of results on intra-patient data

Table 11 shows the comparison of the performance in previous studies in identifying AFIB and VFIB heartbeats using K-fold cross validation. All of these studies adopt the AFIB and VFIB heartbeats in AFdb and VFdb as the experimental data. In Table 11, the SE, PPV, SPE, and F1-score, which is the harmonic mean of SE and PPV, were adopted as the indicators for comparison. According to Table 11, Radhakrishnan et al. [15] and Mohanty et al. [33] achieved the highest SE, PPV, SPE, and F1-scores for recognizing VFIB or AFIB heartbeats, respectively. Compared with these studies, the proposed method achieved a significantly higher indicators in detecting AFIB or VFIB heartbeats, respectively. In terms of classifying AFIB and VFIB heartbeats, the proposed method realized the highest F1-scores of 99.42 % for AFIB and 96.27 % for VFIB among all the studies [26–28]. For the AFIB, although references [28] and [26,27] achieved slightly higher SE than our method, our method realized significantly higher PPV, F1-score, and SPE than their methods. For the VFIB, compared to the references [28] and [26,27] our method achieved similar SPE and higher SE and F1-score. In summary, compared with state-of-the-art methods, this approach has significantly better performance in recognizing AFIB and VFIB.

4.5.2. Comparison of results on inter-patient data

In terms of the performance on the inter-patient experiments, we found that the obtained OAs on sets D, E, and F are only 0.1 %, 0.55 %, and 2.22 % lower than those on sets A, B, and C, respectively. This indicates that the proposed method can achieve excellent performance on both intra- and inter-patient experiments. Since most studies did not consider the effects of their proposed methods on inter-patient experiments, we compare the performance of our method with ResNet-18 [34], ResNet-34 [34], DenseNet [35], EFAP-Net [36], and the ECG-Convolution-Vision Transformer Network (ECVT-Net) [37]. To ensure the fairness of comparison, we further test the performance of the above methods in set F with inter-patient normal, AFIB and VFIB heartbeats. In
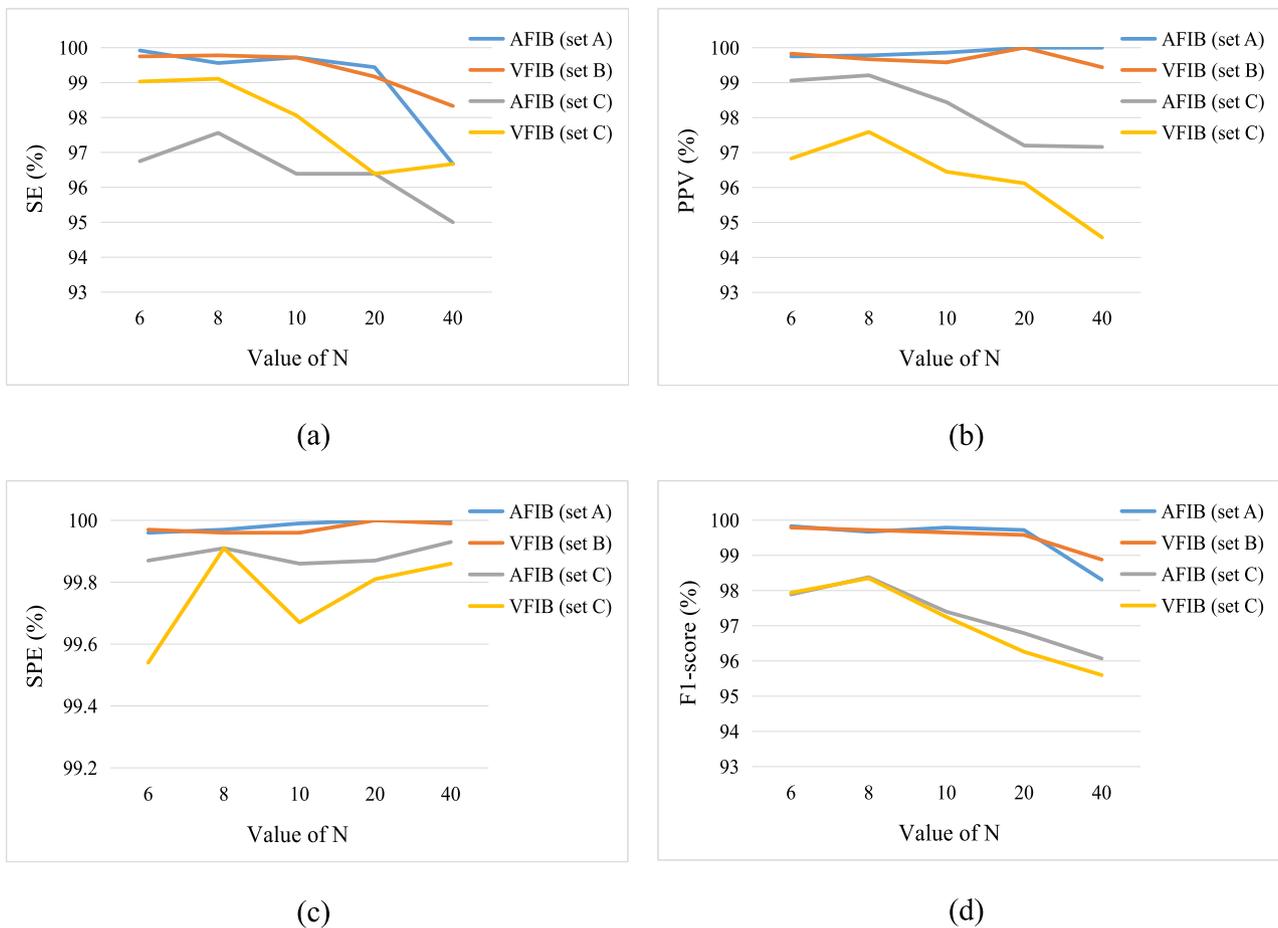
Fig. 6. Curve drawn obtained on multi-level extremely imbalanced data: (a) SEN; (b) PPV; (c) SPE; (d) F1-score.

Table 8
Confusion matrix for set D (normal and AFIB classes).

		Predicted		AC (%)	SE (%)	PPV (%)	SPE (%)	F1-score (%)
		Normal	AFIB					
Original	Normal	7181	19	99.84	99.74	99.92	99.93	99.83
	AFIB	6	8893	99.84	99.93	99.79	99.74	99.76
OA (%)				99.84				

Table 9
Confusion matrix for set E (normal and VFIB classes).

		Predicted		AC (%)	SE (%)	PPV (%)	SPE (%)	F1-score (%)
		Normal	VFIB					
Original	Normal	7194	6	99.39	99.92	99.36	96.46	99.63
	AFIB	46	1254	99.39	96.46	99.52	99.92	97.97
OA (%)				99.39				

Table 10
Confusion matrix for set F (normal, AFIB and VFIB classes).

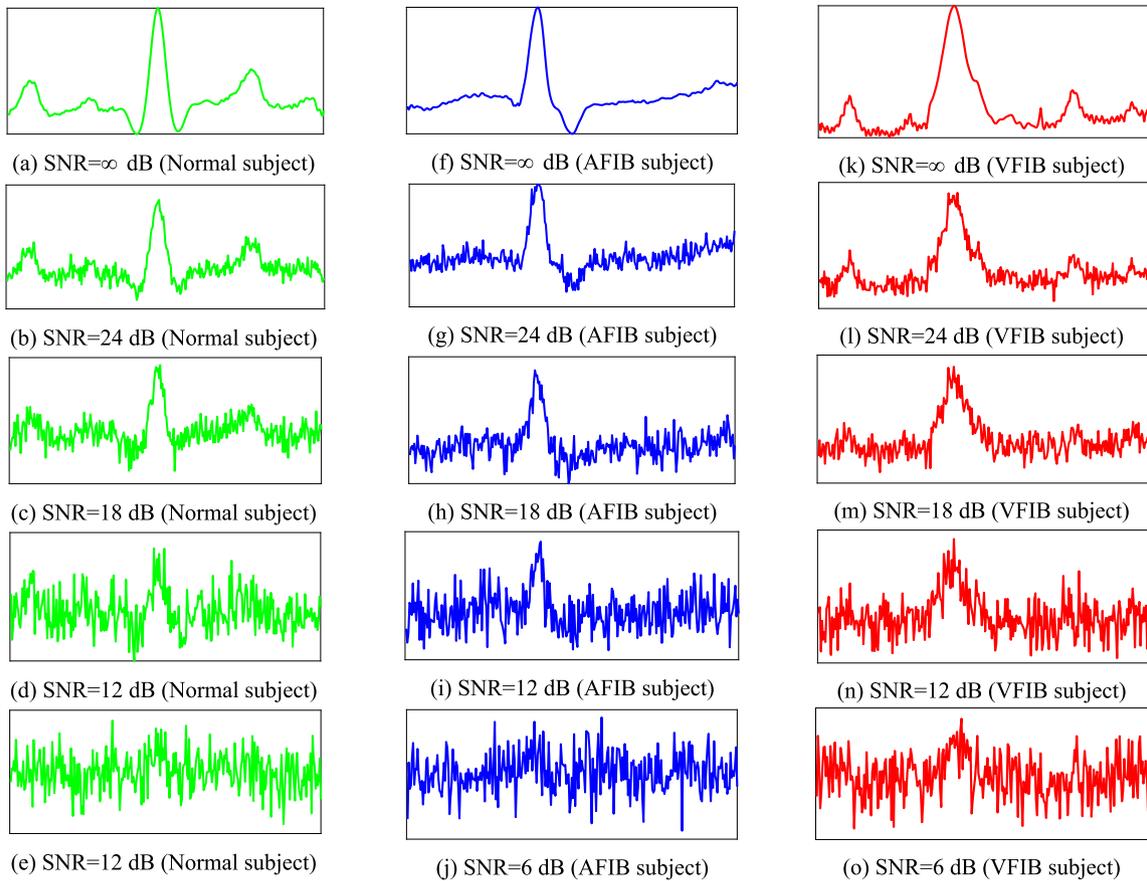|		Predicted			AC (%)	SE (%)	PPV (%)	SPE (%)	F1-score (%)
		Normal	AFIB	VFIB					
Original	Normal	7177	18	5	99.76	99.68	99.74	99.81	99.72
	AFIB	9	8791	99	97.26	98.79	95.97	95.66	97.36
	VFIB	10	351	939	97.33	72.23	90.02	99.35	80.15
OA (%)					97.17				

Fig. 7. Noisy heartbeats of the 1st lead with multi-level noise.
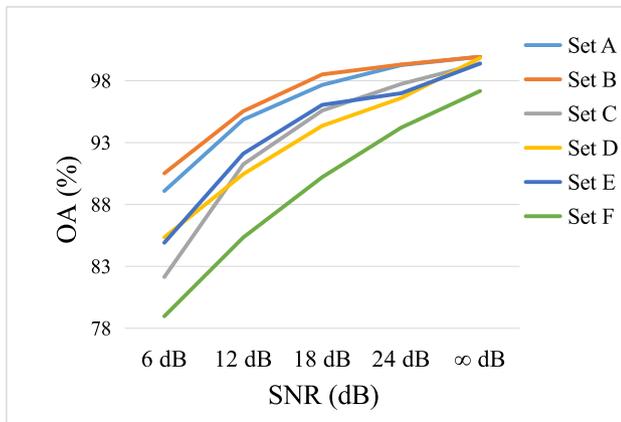
Fig. 8. The curves of OAs obtained on noisy data (SNR = ∞, 24, 18, 12, 6).

Table 11
Relevant studies in detecting AFIB and VFIB heartbeats in intra-patient experiments.

	SE/PPV/F1-score/SPE (%)	
	AFIB	VFIB
Andersen et al. [21]	98.98/95.76/97.31/ 96.95	–
Rahul et al. [20]	98.9/98.8/98.85/98.80	–
Radhakrishnan et al. [17]	99.17/–/99.42/99.18	–
Mohanty et al. [33]	–	97.97/99/98.48/99.15
Fujita et al. [26]	99.57/97.87/98.72/ 93.24	96.07/92.63/94.32/ 99.86
Fujita et al. [27]	99.43/98.74/99.08/ 96.07	96.17/95.82/95.99/ 99.83
Tripathi et al. [28]	100/92.1/95.8/95.7	92.8/–/95.6/99.2
Our method on Set A	99.98/99.92/99.95/99.9	–
Our method on Set B	–	99.92/99.69/99.8/99.94
Our method on Set C	99.26/99.58/99.42/ 99.56	97.31/95.26/96.27/ 99.61

terms of the evaluation indicators, the OA and the F1-scores for AFIB and VFIB, which can comprehensively reflect the SE and PPV, were adopted for the performance comparison. Fig. 9(a–c) shows the results for the above methods. According to Fig. 9(a–c), the F1-scores of AFIB heartbeats are higher than those of VFIB heartbeats for most methods. Among these methods, the highest F1-score of approximately 93 % for recognizing AFIB was realized by DenseNet, while EFAP-Net achieved the highest F1-score of approximately 75 % for identifying VFIB heartbeats. However, the OA (97.17 %) and the F1-scores for AFIB (97.36 %) and VFIB (80.15 %) heartbeats of MDF-Net are significantly higher than those of DenseNet and EFAP-Net. Hence, compared with other methods, MDF-Net has significantly better performance in classifying AFIB and

VFIB heartbeats in inter-patient experiments.

4.5.3. Comparison of results on multi-level extremely imbalanced data

Here, we compared the results of the proposed method the above methods in processing extremely imbalanced data. It is known that these methods have the ability to process imbalanced ECGs or other types of data to a certain extent. In terms of the evaluation indicators, the F1-scores for AFIB and VFIB, which can comprehensively reflect the SE and PPV, were adopted for the performance comparison. Fig. 10 shows the curve of the recognition results of VFIB and AFIB as a function of N values. According to Fig. 10(a–b), the highest and lowest F1-scores were achieved when N = 8 and N = 40, respectively. Except for MDF-Net, the
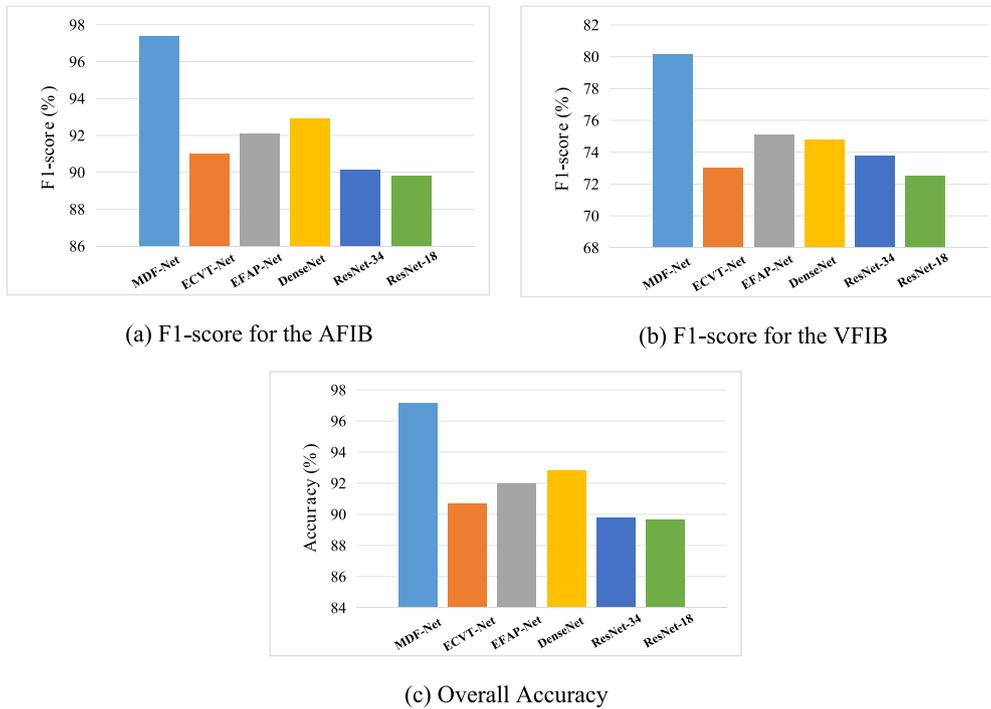
(a) F1-score for the AFIB

(b) F1-score for the VFIB

(c) Overall Accuracy

Fig. 9. Classification results using MDF-Net and other methods on set F.

(a) F1-score for AFIB

(b) F1-score for VFIB
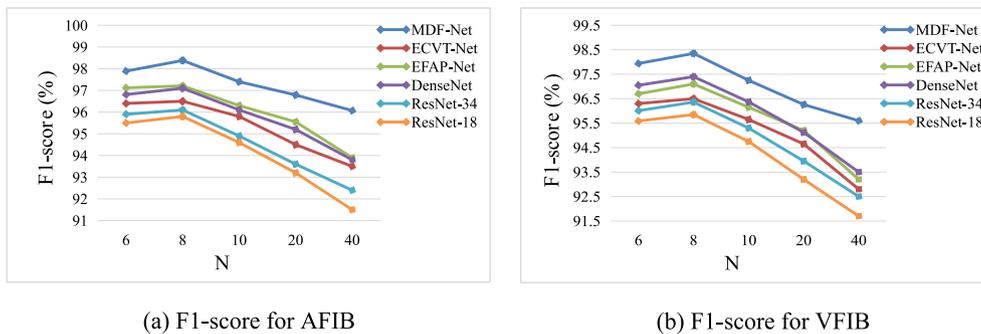
Fig. 10. Classification results using MDF-Net and other methods on set C.

(a) Results on the intra-patient experiment on Set C
(Normal+AFIB+VFIB)

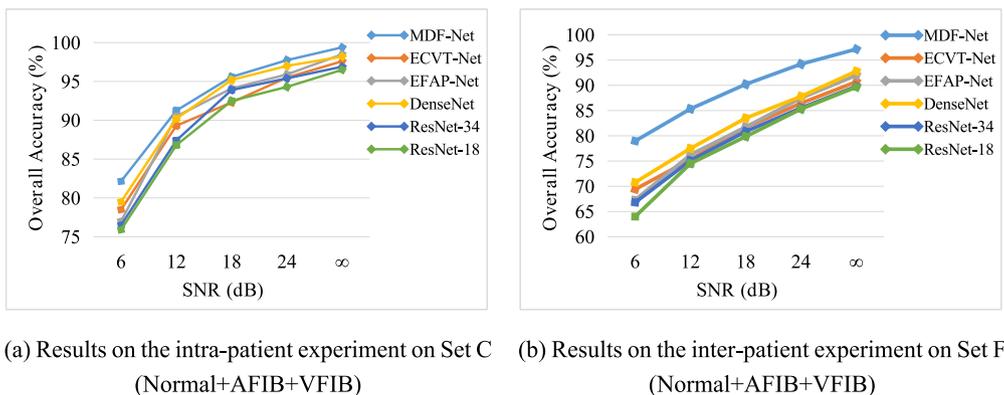(b) Results on the inter-patient experiment on Set F
(Normal+AFIB+VFIB)

Fig. 11. Classification accuracies using MDF-Net and other methods.

highest F1-score of approximately 97 % for recognizing AFIB was realized by DenseNet, while EFAP-Net achieved the highest F1-score of approximately 97.8 % for identifying VFIB heartbeats. However, when N is 40, the F1-scores of these methods for AFIB and VFIB are less than

93.5 % except for MDF-Net, which is significantly lower than that achieved by MDF-Net (95.5 %–96 %). Specifically, ResNet-18 only achieved F1-scores of approximately 92 % for AFIB and VFIB, which is far lower than that of MDF-Net. In summary, compared with other methods, the
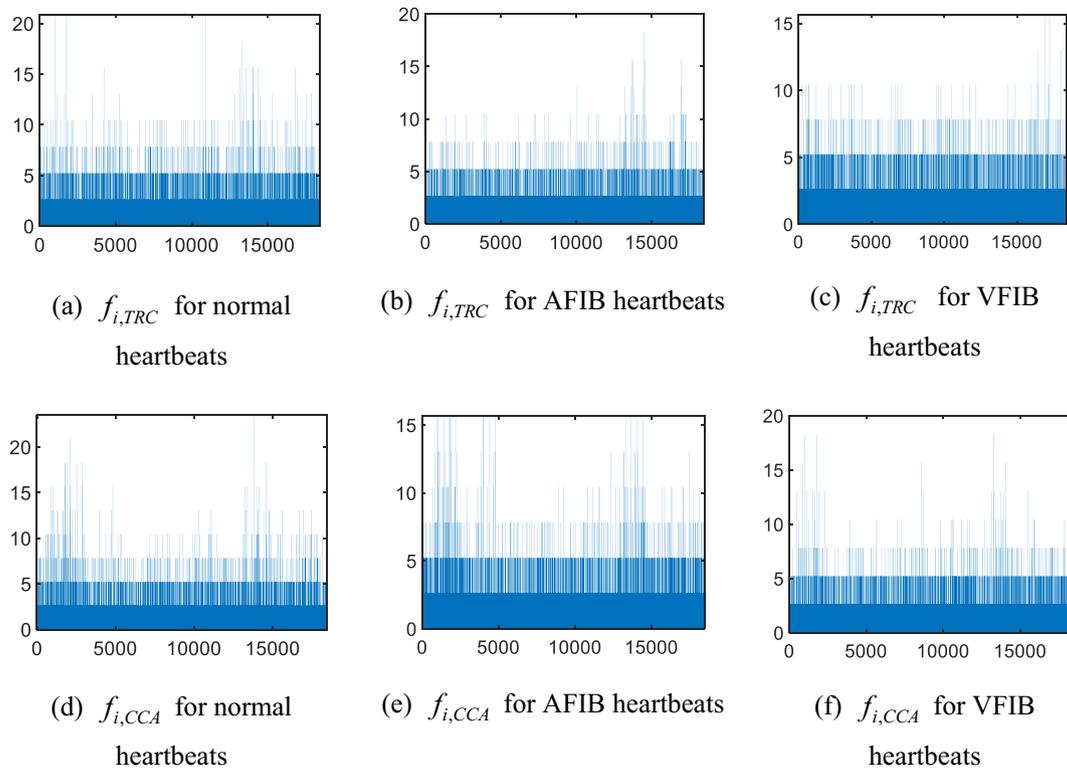
11

(a) $f_{i,TRC}$ for normal heartbeats

(b) $f_{i,TRC}$ for AFIB heartbeats

(c) $f_{i,TRC}$ for VFIB heartbeats

(d) $f_{i,CCA}$ for normal heartbeats

(e) $f_{i,CCA}$ for AFIB heartbeats

(f) $f_{i,CCA}$ for VFIB heartbeats

Fig. 12. Statistical histogram feature vectors of MDF-Net for normal, AFIB, and VFIB heartbeats.

Table 12
The training time of different CNNs and MDF-Net.

Method	ResNet-18 [34]	ResNet-34 [34]	VGGNet-16 [40]	MDF-Net
Time per epoch	90 s	170 s	320 s	540 s
Number of epochs	Over 20	Over 20	Over 20	1
Total time	1800 s	3400 s	6400 s	540 s

proposed method has a better ability to process extremely imbalanced data, in which abnormal heatbeats are much less than normal heartbeats.

4.5.4. Comparison of results on inter-patient data with noise

Fig. 11 shows the comparison of the performance between our method and other methods. To ensure the fairness of comparison, the overall accuracies shown in Fig. 11 were yielded by the experiments on Set C and Set F with different degrees of noise. According to Fig. 11(a–b), it can be seen that the accuracy of the proposed method is only slightly higher than that of other methods in intra-patient experiments, while there are significant differences in performance in inter-patient experiments. This may be because the discriminative information in intra-patient data is significantly more similar than that in inter-patient data, causing these methods to easily distinguish intra-patient data even if noise is added. For the inter-patient experiments, according to Fig. 11(b), when the SNR is 6 dB, ResNet-18 achieved an OA of only approximately 64 %, which is lower than that of the other methods. In addition, although ECVT-Net, DenseNet and EFAP-Net achieved OAs of approximately 90 % in processing noise-free heartbeats, the obtained OAs are only 75 %–78 % when the SNR is 12 dB, indicating that the noise in heartbeats seriously affected their classification performance. However, when SNR is 12 dB, the proposed method (MDF-Net) still achieved an OA of 85.34 %, which is significantly higher than that achieved by other methods. Hence, compared with other methods, our

method has significantly better noise robustness.

4.5.5. Analysis of MDF-Net

The above excellent performance stems from the fact that MDF-Net can achieve two stages of feature-level information fusion from different angles to yield two sets of features, which are then fused at the decision level by using the LS-WSA method. Specifically, MDF-Net can be seen as the fusion of two methods: TRPC-Net and CPC-Net, which can extract deep features of multi-lead ECGs from different angles. TRPC-Net consists of a TRCA convolutional layer, a PCA convolutional layer, and a fully connected layer. For TRPC-Net, the TRCA convolutional layer is designed to extract the task-related component by maximizing the reproducibility of the data under the same task, while the PCA convolutional layer removes the correlation of data between categories. Then, the fully connected layer is adopted to fuse the outputs of the PCA convolutional layer to obtain the feature vectors. Hence, TRPC-Net can extract inter-lead task-related features from multi-lead ECGs, by which the first stage of feature-level information fusion is achieved. Furthermore, CPC-Net, which is composed of a CCA convolutional layer, a PCA convolutional layer, and a fully connected layer, is employed to realize the second stage of feature-level information fusion. Among them, the CCA convolutional layer is used to extract the correlation information from two-lead ECGs, while the functions of the PCA convolutional layer and the fully connected layer are similar to those in TRPC-Net. Hence, we refer to the CPC-Net-based features as the inter-lead correlation features. To realize decision making, LSVM is used to process the TRPC-Net-based features and CPC-Net-based features to obtain two sets of decision values. These decision values are then fused at the decision level using the WSA method, by which the inter-lead task-related features and inter-lead correlation features are fully utilized. Hence, the proposed method can comprehensively utilize the above multi-angle features, thus performing well on intra-patient data, extremely imbalanced data, and inter-patient data in detecting AFIB and VFIB. In addition, TRCA, PCA and CCA methods have the ability to remove noise from data. After integrating TRCA, CCA and PCA methods into the

Table 13
The MDF-Net algorithm.

Input:	Raw ECG signals
Output:	$f_{i,TRC}$ and $f_{i,CCA}$
1:	Form ECG matrix $I_i^c, c = 1, \cdots, C$
2:	**for** Scanning layer **do**
3:	Extract patches using a window of size $k_1 \times k_2$
4:	Construct the first-order pending matrices $X^c, c = 1, 2$
5:	**end for**
6:	**for** the TRCA convolutional layer **do**
7:	Construct the matrix $S =$ $$(S_{c_1 c_2})_{1 \le c_1, c_2 \le C} = \left(\sum_{\substack{h_1, h_2 = 1 \\ h_1 \ne h_2}}^{N_t} Cov\big(x^{c_1,(h_1)}(t), x^{c_2,(h_2)}(t)\big) \right)_{1 \le c_1, c_2 \le C}$$
8:	Compute the matrix $Q =$ $$(S_{c_1 c_2})_{1 \le c_1, c_2 \le C} = \left(\sum_{h=1}^{N_t} Cov\big(x^{c_1,(h)}(t), x^{c_2,(h)}(t)\big) \right)_{1 \le c_1, c_2 \le C}$$
9:	Compute L_1 eigenvectors from $Q^{-1}S$
10:	Calculate TRCA filters $W_l^1, l = 1, 2, \cdots, L_1$
11:	Construct the first-order feature (FOF) matrix as per $\Gamma_{i,l,TRC}^c = I_i^c * W_l^1, c = 1, 2, l = 1, 2, \cdots, L_1$
12:	**end for**
13:	**for** the CCA convolutional layer **do**
14	Calculate the cross-covariance matrix s_{ij} of X^i and X^j
15	Calculate the two-lead project directions α_l and $\beta_l, l = 1, 2, \cdots, L_2$ by the Lagrange multiplier technique
16	Compute CCA filters $U_l^c, l = 1, 2, \cdots, L_2$
17:	Calculate the primary feature blocks (PFBs) as per $\Gamma_{i,l,CCA}^c = I_i^c * U_l^c, c = 1, 2, l = 1, 2, \cdots, L_2$
18:	**end for**
19:	**for** the PCA convolutional layer **do**
20:	Calculate the second-order pending matrices $Y^c, c = 1, 2$
21:	Construct the covariance matrix $(Y^c)(Y^c)^T$
22:	Compute L_3 eigenvectors
23:	Construct PCA filters $W_{\ell,c}^2, \ell = 1, 2, \cdots, L_3$
24:	Calculate the second-order feature (SOF) matrices $O_{i,l}^c = \left\{ \Gamma_{i,l}^c * W_{\ell,c}^2 \right\}_{\ell=1}^{L_3}$
25:	**end for**
26:	Compute the decimal image $\Psi_{i,l,TRC}^2 = \sum_{\ell=1}^{L_3} 2^{\ell-1} H\big(\Gamma_{i,l,TRC}^1 * W_{\ell,1}^2, \Gamma_{i,l,TRC}^2 * W_{\ell,2}^2\big)$ and $\Psi_{i,l,CCA}^2 = \sum_{\ell=1}^{L_3} 2^{\ell-1} H\big(\Gamma_{i,l,CCA}^1 * W_{\ell,1}^2, \Gamma_{i,l,CCA}^2 * W_{\ell,2}^2\big)$
27:	Calculate two histogram vector $f_{i,TRC}$ and $f_{i,CCA}$ as features of MDF-Net

cascaded and convoluted network structure of MDF-Net, their ability to remove noise is enhanced, by which the proposed method has good noise robustness. Based on the above analysis, the proposed method has the potential to be applied in the clinic to detect AFIB and VFIB patients.

CNN models have been used in some studies to process multi-lead ECG data. However, these methods are mostly adopted for detecting myocardial infarction (MI) rather than AFIB and VFIB. For example, Liu et al. [38] proposed a multiple-feature-branch (MFB) CNN based ECG classification method. In this study, they first extracted features from each-lead ECGs using multiple sets of branches composed of convolutional and pooling layers, and these features were then connected using the fully connected layer. In addition, He et al. [39] designed a CNN and active learning-based ECG classification method to process multi-lead ECGs. In their study, each-lead ECGs were first processed by the MFB model and lead attention mechanism (LAM) to yield primary features. Then, multi-lead ECG-based primary features were cascaded by the concatenate and flattening methods using a fully connected layer. The above methods have successfully integrated multi-lead ECG features. However, there are several issues that need to be addressed in the above methods. First, these methods can only extract features through MFB or MBF-LAM models, making it impossible to mine ECG features from different perspectives. Second, these methods cannot enhance the correlation information that records cardiac states between multi-lead ECG features during the feature extraction process of branches. Third, these methods only adopted feature-level information fusion methods and did not use multi-channel decision-level information fusion methods, by

which the uncertainty information of multiple channels cannot be comprehensively utilized. However, as we mentioned earlier, our method can achieve two angles of feature-level and decision-level information fusions through the use of MDF-Net and D-S fusion methods. At the same time, it can also enhance the information correlation and task correlation between multi-lead ECGs through the CCA and TRCA algorithms, thereby achieving the extraction of multi-angle correlation features. To sum up, compared with the methods in the above references, our method has a more comprehensive performance in processing multi-lead ECGs.

4.5.6. Features of MDF-Net

Fig. 12(a–f) shows the histogram statistical feature vectors $f_{i,TRC}$ and $f_{i,CCA}$ obtained using MDF-Net by processing normal, AFIB, and VFIB ECGs. Among them, Fig. 12(a–c) shows the $f_{i,TRC}$, while Fig. 12(d–f) shows the $f_{i,CCA}$ obtained from the above three types of ECGs. As we know, the LSVM with the D-S method can effectively analyze $f_{i,TRC}$ and $f_{i,CCA}$ for classifying ECGs. However, according to Fig. 12, we find that the features extracted by MDF-Net are very abstract, by which the MDF-Net based features have difficulty corresponding to the physical meaning of the original ECG waveforms. Although the MDF-Net features of different types of ECG can be found to have certain differences through human eyes, it is difficult to determine the types of these differences. Therefore, the features acquired by the MDF-Net model exhibit a relatively weak level of interpretability.

4.5.7. Comparison of training times between MDF-net and CNNs

In terms of training time, since the proposed MDF-Net has a multi-layer cascaded structure similar to the CNN, we compare the classification time between MDF-Net and CNNs. Table 12 shows the training time using 16,999 heartbeats on set F using three types of CNNs: ResNet-18 [34], Resnet-34 [34], and VGGNet-16 [40]. According to Table 12, the one-cycle training time of all CNN models exceeded 90 s. Because CNNs with excellent capabilities depend on multiple iterations, their epochs usually exceeded 20, and their total training time exceeded 1800 s. However, for MDF-Net, the training time was only 540 s. This stem from the fact that the convolutional kernels of MDF-Net can be directly extracted by the TRCA, CCA, and PCA algorithms, by which MDF-Net does not need the iterative optimization process for the entire network. Hence, compared to CNNs, MDF-Net had a significantly faster training speed.

4.6. Limitations

3) Only Gaussian white noise yielded by AWGN was added to original ECGs to test the noise robustness of the proposed method. However, other types of noise, such as correlated noise and colored noise, were not adopted.
4) Only two-lead ECGs were used to test the performance of our algorithm. This is because the dimensions of features extracted by MDF-Net are very high. When processing multi-lead ECG data, MDF-Net will occupy a large amount of memory, causing our device to not support MDF-Net for processing 12-lead ECG data.
5) The number of patients is limited, causing us to not verify the performance of the proposed method in processing large-scale ECG data and ECGs with a higher degree of imbalance than the one used in this work.

5. Conclusion

Cardiovascular disease (CVD) has become a main cause of death. Among various CVDs, atrial fibrillation (AFIB) and ventricular fibrillation (VFIB) are the two most common cardiac disorders. To diagnose AFIB and VFIB in a timely manner, it is important to develop an accurate ECG automatic classification method with fast speed. In this work, we

proposed a novel ECG recognition method using MDF-Net for distinguishing AFIB and VFIB. MDF-Net, which can be seen as the fusion of TRPC-Net and CPC-Net, is used to extract multi-angle deep features from two-lead ECGs. TRPC-Net was developed to mine inter-lead task-related features, while CPC-Net was adopted to calculate inter-lead correlation features. Then, we used the LSVM to classify these TRPC-Net-based and CPC-Net-based features to obtain two sets of decision values. Finally, these decision values were then fused by the weighted softmax with average (WSA) to obtain the final labels, by which inter-lead task-related features and correlation features were fully utilized. In classifying normal, AFIB, and VFIB heartbeats, the proposed method achieved accuracies of 99.39 % and 97.17 % with a classification time of 0.0138 s for each heartbeat in intra- and inter-patient experiments, respectively. In addition, the proposed method performed well on noisy data and extremely imbalanced data, in which abnormal heatbeats are much less than normal heartbeats. In conclusion, the proposed method is expected to be applied in the clinic to detect AFIB and VFIB patients.

In future work, we will improve the MDF-Net algorithm to reduce the feature dimensions by using the pooling layer or other methods, thereby achieving the processing of 12-lead ECG data. At the same time, other types of noise will be adopted to verify the noise robustness of the proposed method. In addition, we will recruit new healthy subjects and AFIB and VFIB patients through Jiangsu Provincial People's Hospital to further verify the performance of the proposed method.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Appendix A

Here, we present a detailed explanation MDF-Net, consisting of five processing layers: a scanning layer, a TRCA convolutional layer, a CCA convolutional layer, a PCA convolutional layer, and a fully connected layer

1) Scanning layer

We adopt $I_i^c \in \Re^{m \times n}, i = 1, 2, \cdots, N, c = 1, 2$ to represent two-lead ECG matrices, where N represents the number of ECG matrices and C represents the number of channels. Each I_i^c is scanned by a $k_1 \times k_2$ patch to extract the sample blocks with a step of one. We then reshape these sample blocks into mean-removed vectors $X_i^c = \left[\overline{x}_{i,1}^c, \overline{x}_{i,2}^c, \cdots, \overline{x}_{i,mn}^c \right] \in \Re^{k_1 k_2 \times mn}$. After processing all the I_i^c, the first-order pending (FOP) matrices are calculated as $X^c = \left[X_1^c, X_2^c, \cdots, X_N^c \right] \in \Re^{k_1 k_2 \times Nmn}$ for each-lead ECG.

2) TRCA convolutional layer

TRCA was developed to mine task-related components $r(t)$ and task-unrelated components $u(t)$ from data in different channels. In this work, we need to extract $r(t)$ and $u(t)$ from X^c using mixing coefficients $a_{1,c}$ and $a_{2,c}$ according to eq. (1).

$$X^c(t) = a_1^c r(t) + a_2^c u(t), c = 1, 2 \tag{1}$$

where t represents the tth sample. Hence, $r(t)$ should be extracted from $X^c, c = 1, 2$ according to Eq. (2).

$$y(t) = \sum_{c=1}^{C} w^c x^c(t) = \sum_{c=1}^{C} \left(w^c a_1^c r(t) + w^c a_2^c u(t) \right) \tag{2}$$

where $\sum_{c=1}^{C} w^c a_1^c$ and $\sum_{c=1}^{C} w^c a_2^c$ are 1 and 0, respectively. Here, inter-trial covariance maximization is adopted to address the above issues. Specifically, we adopt $X^{c,h}, c = 1, 2$ and $y^h(t), h = 1, 2, \cdots, N_t$ as the hth trial in $X^c, c = 1, 2$ and the corresponding TRC, respectively. Next, all the possible combinations of the inter-trial covariance of $y(t)$ are shown in Eqs. (3) and (4).

$$S = (S_{c_1 c_2})_{1 \leq c_1, c_2 \leq C} = \left(\sum_{\substack{h_1, h_2 = 1 \\ h_1 \neq h_2}}^{N_t} Cov \left(x^{c_1, (h_1)}(t), x^{c_2, (h_2)}(t) \right) \right)_{1 \leq c_1, c_2 \leq C} \tag{3}$$

$$\sum_{\substack{h_1,h_2=1 \\ h_1 \neq h_2}}^{N_t} C_{h_1 h_2} = \sum_{\substack{h_1,h_2=1 \\ h_1 \neq h_2}}^{N_t} \mathrm{Cov}\left(y^{(h_1)}(t), y^{(h_2)}(t)\right)$$

$$= \sum_{\substack{h_1,h_2=1 \\ h_1 \neq h_2}}^{N_t} \sum_{c_1,c_2=1}^{C} w^{c_1} w^{c_2} \mathrm{Cov}\left(x^{c_1,(h_1)}(t), x^{c_2,(h_2)}(t)\right) \tag{4}$$

$$= w^T S w$$

Next, Eq. (5) is used as the constraint for $Var(y(t))$ representing the variance of $y(t)$.

$$Var(y(t)) = \sum_{c_1,c_2=1}^{C} w_{c_1} w_{c_2} \mathrm{Cov}(x^{c_1}(t), x^{c_2}(t))$$
$$= w^T Q w \tag{5}$$
$$= 1$$

Furthermore, Eq. (6) is employed to optimize Eq. (4).

$$\widehat{w} = \underset{w}{\mathrm{argmax}} \frac{w^T S w}{w^T Q w} \tag{6}$$

Then, we calculate the TRCA filters W_l^1 according to Eq. (7).

$$W_l^1 = mat_{k_1,k_2}\left(q_l(Q^{-1}S)\right), l = 1,2,\cdots,L_1 \tag{7}$$

where $q_l()$ is employed to calculate eigenvectors of L_1 maximum eigenvalues from $Q^{-1}S$, and each eigenvector is reconstructed to W_l^1 by $mat_{k_1,k_2}()$. Finally, as per $\Gamma_{i,l,TRC}^c = I_i^c * W_l^1, c = 1,2, l = 1,2,\cdots,L_1$, we acquire L_1 TRCA-based primary feature blocks (TPFBs) $\Gamma_{i,l,TRC}^c$ for each lead ECG.

3) CCA convolutional layer

Here, we need to calculate a series of α_l and $\beta_l, l = 1,2,\cdots,L_2$ for constructing the CCA filters. To address this issue, we calculate the first set of vectors α_1 and β_1 according to Eq. (8).

$$\begin{aligned} &\textit{Maximize } \alpha_1^T S_{12} \beta_1 \\ &s.t. \alpha_1^T S_{11} \alpha_1 = 1, \beta_1^T S_{22} \beta_1 = 1 \end{aligned} \tag{8}$$

where the autocovariance of $X^c, c = 1,2$ is expressed by S_{11} and S_{22}, while the cross-covariance between X^1 and X^2 is expressed by S_{12}. Hence, we adopt the Lagrange multipliers λ and ν to optimize Eq. (9).

$$J(\alpha_1, \beta_1) = \alpha_1^T S_{12} \beta_1 - \frac{\lambda}{2}\left(\alpha_1^T S_{11} \alpha_1 - 1\right) - \frac{\nu}{2}\left(\beta_1^T S_{22} \beta_1 - 1\right) \tag{9}$$

For Eq. (9), we need to maximize $J(\alpha_1, \beta_1)$ to obtain α_1 and β_1. To address this problem, Eq. (10) will be acquired by computing the partial derivative of $J(\alpha_1, \beta_1)$.

$$\begin{cases} S_{11}^{-1} S_{12} S_{22}^{-1} S_{21} \alpha_1 = \lambda^2 \alpha_1 \\ S_{22}^{-1} S_{21} S_{11}^{-1} S_{12} \beta_1 = \lambda^2 \beta_1 \end{cases} \tag{10}$$

where α_1 and β_1 can be obtained by calculating the eigenvectors from $S_{11}^{-1} S_{12} S_{22}^{-1} S_{21}$ and $S_{22}^{-1} S_{21} S_{11}^{-1} S_{12}$, respectively. Moreover, we can calculate the vectors α_l and $\beta_l, 1 < l < L_2$ as per Eq. (11).

$$\begin{aligned} &\textit{Maximize } \alpha_l^T S_{12} \beta_l \\ &s.t. \alpha_l^T S_{11} \alpha_l = 1, \beta_l^T S_{22} \beta_l = 1, l = 2,3,\cdots,L_2 \\ &\alpha_{l-1}^T S_{11} \alpha_l = 0, \beta_{l-1}^T S_{22} \beta_l = 0 \end{aligned} \tag{11}$$

Here, we calculate the CCA filters $U_l^c, c = 1,2$ according to Eq. (12).

$$\begin{cases} U_l^1 = mat_{k_1,k_2}(\alpha_l) \in \Re^{k_1 \times k_2} \\ U_l^2 = mat_{k_1,k_2}(\beta_l) \in \Re^{k_1 \times k_2} \end{cases}, l = 1,2,\cdots,L_2 \tag{12}$$

where $mat_{k_1,k_2}(\bullet)$ reconstructs α_l and β_l to matrices $U_l^c, c = 1,2$. Finally, according to $\Gamma_{i,l,CCA}^c = I_i^c * U_l^c, c = 1,2, l = 1,2,\cdots,L_2$, we calculate L_2 CCA primary feature blocks (CPFBs) $\Gamma_{i,l,CCA}^c$ for each-lead ECG, where * expresses the convolution operation.

6) PCA convolutional layer

Similar to the scanning layer, a $k_1 \times k_2$ patch scans TPFB $\Gamma^c_{i,l,TRC}$ or CPFB $\Gamma^c_{i,l,CCA}$ matrices to obtain a series of mean-removed patches, which are then reconstructed to vectors $\bar{y}^c_{i,l,1}, \bar{y}^c_{i,l,2}, \cdots, \bar{y}^c_{i,l,mn} \in \Re^{k_1 \times k_2}$. To make the process clear, we only employ TPFBs $\Gamma^c_{i,l,TRC}$ as pending data to describe the subsequent steps. After processing all the $\Gamma^c_{i,l,TRC}$ by the above operations, $\bar{Y}^c_{i,l} = \left[\bar{y}^c_{i,l,1}, \bar{y}^c_{i,l,2}, \cdots, \bar{y}^c_{i,l,mn} \right] \in \Re^{k_1 k_2 \times mn}$, $\bar{Y}^c_i = \left[\bar{Y}^c_{i,1}, \bar{Y}^c_{i,2}, \cdots, \bar{Y}^c_{i,L_1} \right] \in \Re^{k_1 k_2 \times L_1 mn}$, and $Y^c = \left[\bar{Y}^c_1, \bar{Y}^c_2, \cdots, \bar{Y}^c_N \right] \in \Re^{k_1 k_2 \times L_1 Nmn}$ are obtained. Among them, Y^c represents the second-order pending (SOP) matrix.

Then, we calculate the L_3 PCA filters $W^2_{\ell,c}$ as Eq. (13).

$$W^2_{\ell,c} = mat_{k_1,k_2} \left(q_\ell \left((Y^c)(Y^c)^T \right) \right) \in \Re^{k_1 \times k_2}, c = 1, \cdots, C, \ell = 1, \cdots, L_3 \tag{13}$$

where $q_l()$ is employed to calculate eigenvectors of L_3 maximum eigenvalues from the covariance matrix of Y^c, and each eigenvector is reconstructed to $W^2_{\ell,c}$ by $mat_{k_1,k_2}()$. Finally, according to $O^c_{i,l} = \left\{ \Gamma^c_{i,l} * W^2_{\ell,c} \right\}^{L_3}_{\ell=1}$, we acquire L_3 second-order feature blocks (SFBs) $O^c_{i,l}$ for each-lead ECGs. After processing all $\Gamma^c_{i,l,TRC}$ and $\Gamma^c_{i,l,CCA}$, $O^c_{i,l,TRC}$ and $O^c_{i,l,CCA}$ will be obtained.

7) Output layer

In this layer, we convert each SFB $O^c_{i,l,TRC}$ to a binary matrix according to Eq. (14).

$$H(d) = \begin{cases} 1 & if\ d > 0 \\ 0 & if\ d \leq 0 \end{cases} \tag{14}$$

where d represents an element in the SBF. Then, decimal matrix $\Psi_{i,l}$ is calculated as per $\Psi_{i,l} = \sum^{L_3}_{\ell=1} 2^{\ell-1} H \left(\Gamma^1_{i,l} * W^2_{\ell,1}, \Gamma^2_{i,l} * W^2_{\ell,2} \right)$, of which the values are in the range of $\left[0, \ 2^{L_3} - 1 \right]$. Next, a $u_1 \times u_2$ window with overlap R is used to scan decimal matrix $\Psi_{i,l}$ to obtain the blocks $b_{l,p}, l = 1, \cdots, L_1, p = 1, \cdots P$. Next, by using the histogram statistical method, all blocks $b_{l,p}$ are processed to yield the feature vector $f_i = \left[hist \left(b_{1,1} \right), \cdots, hist \left(b_{L_2,P} \right) \right]$. After processing all $O^c_{i,l,TRC}$ and $O^c_{i,l,CCA}$, we obtain two sets of feature vectors $f_{i,TRC}$ and $f_{i,CCA}$. The MDF-Net algorithm is shown in Table 13.

References

[1] Akhmerov A, Parimon T. Extracellular vesicles, inflammation, and cardiovascular disease. Cells 2022;11:2229.

[2] World Health Organization. Global health estimates. Available from: https://www.who.int/zh/news/item/09-12-2020-who-reveals-leading-causes-of-death-and-disability-worldwide-2000-2019; 2019.

[3] Li XY, Shi XL, Handa BS, et al. Classification of fibrillation organisation using electrocardiograms to guide mechanism-directed treatments. Front Physiol 2021;12:712454.

[4] Tripathi PM, Kumar A, Kumar M, Komaragiri R. Multilevel classification and detection of cardiac arrhythmias with high-resolution superlet transform and deep convolution neural network. IEEE Trans Instrum Measure 2022;71:4006113.

[5] Aksa B, Ma A, Gbbb C. An intelligent learning approach for improving ECG signal classification and arrhythmia analysis. Artif Intell Med, 103, 101788.

[6] Xuan Z A, Hui W B, Tc A, et al. Automatic diagnosis of arrhythmia with electrocardiogram using multiple instance learning: From rhythm annotation to heartbeat prediction. Artif Intell Med, 132, 102379.

[7] Feng PP, Fu J, Ge ZY, et al. Unsupervised semantic-aware adaptive feature fusion network for arrhythmia detection. Inform Sci 2022;582:509–28.

[8] Li YW, Zhang Y, Zhao LN, et al. Combining convolutional neural network and distance distribution matrix for identification of congestive heart failure. IEEE Access 2018;6:39734–44.

[9] Jahmunah V, Oh SL, Wei JKE, et al. Computer-aided diagnosis of congestive heart failure using ECG signals — a review. Phys Med Eur J Med Phys 2019;62:95–104.

[10] Wang HY, Zhou YJ, Zhou B, et al. Interactive ECG annotation: an artificial intelligence method for smart ECG manipulation. Inform Sci 2021;581:42–59.

[11] Li WQ, Tang YM, Yu KM, et al. SLC-GAN: an automated myocardial infarction detection model based on generative adversarial networks and convolutional neural networks with single-lead electrocardiogram synthesis. Inform Sci 2022; 589:738–50.

[12] He ZY, Yuan SY, Zhao JH, et al. A novel myocardial infarction localization method using multi-branch DenseNet and spatial matching-based active semi-supervised learning. Inform Sci 2022;606:649–68.

[13] Fan W, Si YJ, Yang WY, et al. Class-specific weighted broad learning system for imbalanced heartbeat classification. Inform Sci 2022;610:525–48.

[14] Shu LO, Vicnesh J, Ru ST, et al. Comprehensive electrocardiographic diagnosis based on deep learning. Artif Intell Med 2020;103:101789.

[15] Martis RJ, Acharya UR, Prasad H, et al. Automated detection of atrial fibrillation using Bayesian paradigm. Knowl-Based Syst 2013;52:269–75.

[16] Jiang FF, Xu JG, Lu ZY, et al. A transfer learning approach to detect paroxysmal atrial fibrillation automatically based on ballistocardiogram signal. J Med Imaging Health Inform 2019;9(9):1943–9.

[17] Radhakrishnan T, Karhade J, Ghosh SK, et al. AFCNNet: automated detection of AF using chirplet transform and deep convolutional bidirectional long short term memory network with ECG signals. Comput Biol Med 2021;137:104783.

[18] Maghawry E, Ismail R, Gharib TE, et al. An efficient approach for paroxysmal atrial fibrillation events prediction using extreme learning machine. J Intell Fuzzy Syst 2021;40(3):5087–99.

[19] Ting-Ruen Wei, Lu Senbao, Yuling Yan. Automated atrial fibrillation detection with ECG. Bioengineering 2022;9(10).

[20] Rahul J, Sharma LD, et al. Artificial intelligence-based approach for atrial fibrillation detection using normalised and short-duration time-frequency ECG. Biomed Sign Process Cont 2022;71:103270.

[21] Andersen RS, Peimankar A, Puthusserypady S. A deep learning approach for real-time detection of atrial fibrillation. Exp Syst With Appl 2019;115:465–73.

[22] Parsi A, Glavin M, Jones E, et al. Prediction of paroxysmal atrial fibrillation using new heart rate variability features. Comput Biol Med 2021;133:104367.

[23] Lee K, Kim S, Choi HO, et al. Analyzing electrocardiogram signals obtained from a nymi band to detect atrial fibrillation. Multimed Tools Appl 2020;79(23–24): 15985–99.

[24] Lee K, Kim JY, Choi HO, et al. Detection of atrial fibrillation in short-lead electrocardiogram recordings obtained using a smart scale. J Electr Eng Technol 2021;16(2):1109–18.

[25] Panigrahy D, Sahu PK, Albu F. Detection of ventricular fibrillation rhythm by using boosted support vector machine with an optimal variable combination. Comput Electr Eng 2021;91:107035.

[26] Fujita H, Cimr D. Decision support system for arrhythmia prediction using convolutional neural network structure without preprocessing. Appl Intell 2019;49 (9):3383–91.

[27] Fujita H, Cimr D. Computer aided detection for fibrillations and flutters using deep convolutional neural network. Inform Sci 2019;486:231–9.

[28] Tripathi PM, Kumar A, Kumar M, Komaragiri R. Multilevel classification and detection of cardiac arrhythmias with high-resolution superlet transform and deep convolution neural network. IEEE Trans Instrum Measure 2022;71:4006113.

[29] Nakanishi M, Wang YJ, Chen XG, et al. Enhancing detection of SSVEPs for a high-speed brain speller using task-related component analysis. IEEE Trans Biomed Eng 2018;65(1):104–12.

[30] Komorowski D, Mika B. A new approach for denoising multichannel electrogastrographic signals. Biomed Sign Cont 2017;45:213–24.

[31] Melgani F, Bazi Y. Classification of electrocardiogram signals with support vector machines and particle swarm optimization. IEEE Trans Inf Technol Biomed 2008; 12(5):667–77.

[32] Osowski S, Hoai LT, Markiewicz T. Support vector machine-based expert system for reliable heartbeat recognition. IEEE Trans Biomed Eng 2004;51(4):582–9.

[33] Mohanty M, Biswal P, Sabut S. Machine learning approach to recognize ventricular arrhythmias using VMD based features. Multidimen Syst Sign Process 2020;31(1): 49–71.

[34] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June; 2016. p. 770–8.

[35] Zhang G, Si Y J, Yang W Y, Wang D. A robust multilevel DWT densely network for cardiovascular disease classification. Sensors, 20(17), 4777.

[36] Yang WY, Si YJ, Wang D, et al. Automated intra-patient and inter-patient coronary artery disease and congestive heart failure detection using EFAP-Net. Knowl-Based Syst 2020;201:106083.

[37] Liu T, Si Y, Yang W, Huang J, Yu Y, et al. Inter-patient congestive heart failure detection using ECG-convolution-vision transformer network. Sensors. 2022;22(9): 3283.

[38] Liu WH, Huang QJ, Chang S, et al. Multiple-feature-branch convolutional neural network for myocardial infarction diagnosis using electrocardiogram. Biomed Sign Process Cont 2018;45:22–32.

[39] He ZY, Yuan ZY, An PF, et al. MFB-LANN: a lightweight and updatable myocardial infarction diagnosis system based on convolutional neural networks and active learning. Comput Meth Progr Biomed 2021;210:106379.

[40] Zhang Y, Wang Z, Zhang J, Wang C, Wang Y, Chen H, et al. Deep learning model for classifying endometrial lesions. J Transl Med 2021;19(1):10.