IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

SPECIAL SECTION ON DATA-ENABLED INTELLIGENCE FOR DIGITAL HEALTH

Received June 13, 2019, accepted June 30, 2019, date of publication July 4, 2019, date of current version July 22, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2926749

Atrial Fibrillation Detection Using an Improved Multi-Scale Decomposition Enhanced Residual Convolutional Neural Network

XIN-CHENG CAO, BIN YAO, AND BIN-QIANG CHEN
School of Aerospace Engineering, Xiamen University, Xiamen 361005, China
Shenzhen Research Institute of Xiamen University, Shenzhen 518000, China

Corresponding author: Bin-Qiang Chen (cbq@xmu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 51605403, in part by the Natural Science Foundation of Guangdong Province, China, under Grant 2015A030310010, and in part by the Fundamental Research Funds for the Central Universities under Grant 20720190009.

ABSTRACT Atrial fibrillation, the most common sustained arrhythmia, is still a big challenge for researchers in the medical field. Many studies attempt to realize intelligent classification of AF based on deep learning methods. However, many of the studies focused on investigations of relatively simple datasets collected from a relatively small number of subjects. On the other hand, sophisticated preprocessing is usually adopted to analyse the ECG signals. These two factors significantly affect the generalization ability of the trained models for complicated data sets collected from a large number of subjects. In order to address this problem, an improved multi-scale decomposition enhanced residual convolutional neural network is proposed. The proposed method is applied to the large single-lead ECG dataset provided by the PhysioNet/CinC Challenge 2017, and good classification accuracy is suggested by the testing results. In the proposed method, the original ECG record with a large difference in length is re-segmented into short samples of 9 s. Then, using the derived wavelet frame decomposition, the segmented short samples are decomposed and reconstituted into sub-signal samples of different scales. We trained the fast down-sampling residual convolutional neural networks (FDResNets) with the original short-signal dataset and the reconstructed dataset of each scale. The transfer learning technique is then applied to couple the three FDResNets with good performance into a multi-scale decomposition enhanced residual convolutional neural network (MSResNet). The FDResNet trained by the [0, 9.375 Hz] reconstruction dataset achieved the best performance. After six-fold cross-validation, the average test accuracy reached 87.12%, and the average comprehensive F1 score reached 85.29%. The average test accuracy of the multi-scale residual neural network reached 92.1%, and the average overall F1 score reached 89.9%.

INDEX TERMS Electrocardiogram, atrial fibrillation, wavelet frame, deep learning, residual convolutional neural networks, transfer learning.

I. INTRODUCTION

Atrial fibrillation (AF), the most common sustained arrhythmia, represents a difficult scientific challenge and remains enigmatic even after more than one century of research [1]. The mechanisms responsible for AF are still not fully understood and the treatment is very complicated. Physiologically, the symptoms of AF are abnormal contractions of the upper atrium; on the electrocardiogram, there is a sinus P wave loss and a severe irregularity of the QRS complex sequence [2]. Therefore, ECG analysis has become an important means of AF diagnosis. An automated analysis and classification system for ECG records can provide physicians with diagnostic recommendations or help patients monitor their own health status, which is important for improving medical efficiency and reducing medical costs.

The traditional ECG classification methodology includes three steps, i.e., signal preprocessing, feature extraction and classification. The first step aims at eliminating various types of noises, including artifacts and baseline drift in the signal.

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

After noise reduction, the input ECG signal is split into separate heartbeat waveforms. Then, on the basis of medical and statistical knowledge, a series of feature indicators are extracted from the waveforms. Furthermore, feature selection algorithms such as principal component analysis and whale optimization algorithm [3] can be employed to generate a more efficient and compact feature vectors. Kumar utilized a flexible analytic wavelet transform framework to decompose the raw ECG into different frequency bands and extracted the sample entropy of the sub-signals as features to diagnosis myocardial infarction [4]. Finally, classification is implemented using machine learning such as support vector machines (SVM) [5], extreme learning machine (ELM) [6], and cellular neural networks [7], or threshold algorithm, such as Kolmogorov–Smirnov test [8] and AF/AT detector [9].

Machine learning techniques have been widely used in bioinformatics due to their ability to process large datasets and extract hidden patterns [10]. In recent years, the rapid growth of deep learning techniques was notable and convolutional neural networks (CNNs) have shown their success in various applications [11]–[13]. One dimension convolutions have proven their learning power for time series classification [14]. Andersen employed the RR intervals for training deep CNNs to identify AF, and obtained a specificity of 98.96% on a dataset consisting of 89 subjects [15]. Sellami proposed a deep convolutional neural network (DCNN) enhanced with batch-weighted loss function for accurate heartbeat classification, achieving 98.83% sensitivity and 96.97% specificity on the MIT-BIH arrhythmia database [16]. This CNN model takes single heartbeat waveforms as input and thus does not use any down-sampling. Tan developed a stacked convolutional and long short-term memory network to diagnose ECG signals for coronary artery disease with a recognition accuracy of 99.85% [17]. Faust utilized the heart rate sequence as the analysis object, and applied deep bidirectional long-short term memory networks to identify whether the sample had AF phenomenon. The average accuracy of 10-fold cross-validation reached 98.3% [18]. Erdenebayar designed a DCNN with an intermediate fully connected layer to identify atrial fibrillation, with a recognition accuracy reported at 98.7% [19].

In general, there have been many studies on ECG classification. However, most of the research is based on relatively simple datasets. There are few subjects or the data has undergone rigorous preprocessing, so the generalization ability of the algorithm is not ideal. The 2017 PhysioNet/CinC Challenge releases a large dataset containing 8528 short single-head ECG records [20]. Each ECG record in the dataset was collected from an individual subject. The durations of the ECG records are relatively short in duration, which range from 9 seconds to 60 seconds.

Many contestants have reported their classification results based on deep learning models. Warrick used this dataset to train a 13-layer convolutional neural network with an average F1 score of 0.83. The neural network directly processes the ECG record for 60 seconds without any pre-processing, and if the sample is small, it takes a repeated fill extension. The model uses batch normalization and random dropout techniques on a large scale to improve the training performance and generalization of the model [21]. Chandra trained two shallow CNNs with only one convolutional layer to identify ECG signals for atrial fibrillation. The former CNN is used to locate the R-peak in the ECG signal, and then divides the heartbeat waveform centered on the R-peak. The second CNN uses the 8-channel heartbeat vector as input to identify the ECG record category. The F1 scores for the Normal and AF signals are 86% and 73%, respectively [22]. Zihlmann trained the deep convolution residual neural network with the Logarithmic spectrogram of the ECG segment. The model employs a large number of convolutional layers to extract feature values, and uses the long-short term memory (LSTM) modules to further enhance the recognition after flatten layer. However, the cross-validation shows that the enhancement of LSTM is not obvious. Except improving the accuracy of AF samples on the augmented dataset, the overall accuracy and F1 scores decreased [23].

This paper presents a robust method capable of detecting AF from a single short ECG lead recording. First, the original ECG records are decomposed and reconstructed into sub-ECGs of different scales via derived wavelet framework. Then, a multi-scale residual convolutional neural network is constructed to classify the reconstructed sub-ECGs. The three sub-networks of this multi-scale convolutional neural network each learn features from a single-scale reconstructed ECG record, thereby bypassing the feature engineering and prior knowledge. Finally, through migration learning, a small fully connected neural network couples the three subnetworks into a parallel multi-scale residual convolutional neural network. Finally, applying transfer learning, a small fully connected neural network couples these sub-networks into a parallel multi-scale residual convolutional neural network. The effectiveness of the method was validated on public dataset from the PhysioNet/Computing in Cardiology Challenge 2017, achieving 92.1% and 89.9% for test accuracy and overall F1 score, respectively.

II. DATASET AND METHODS

In this study, the investigated dataset, from the PhysioNet Challenge 2017, is contributed by AliveCor, a manufacturer of single-channel ECG device. The dataset consists of 8528 single short ECG lead recordings of different lasting time, each of which is from individual customer of AliveCor. After expert identification, the recordings are divided into four categories: normal rhythm (N), AF rhythm (A), other rhythm (O) and noisy recordings (∼). The detailed data acquisition procedures can be found in [24], and the general profile are shown in Table 1.

A. RECORDING INTERCEPTION

The sampling frequency of all recordings is 300 Hz, but the lasting time of each record and the number of samples in each category are very different. To balance the dataset,

TABLE 1. Units for magnetic properties.

Label	Numbers	Lasting time (s)			
		Min	Max	Mean	Total
N	5154	9.0	61.0	31.9	$4.89*10^7$
A	771	10.0	60	31.6	$7.35*10^6$
O	2557	9.1	60.9	34.1	$2.48*10^7$
~	46	10.2	60	27.1	$2.04*10^6$

we re-segmented the samples by 9 seconds. It should be noted that some normal rhythm also have serious noise disruption at the beginning. For the entire record, experts can make accurate judgments based on the main high signal-to-noise ratio segments in the records. However, if a smaller sample is intercepted from the beginning, its label needs further identification. To balance the dataset and maintain sample diversity, only one secondary short sample is intercepted from the middle of each normal class record. For other rhythm those last less than 20 seconds, we intercept a short sample from the middle. For those last longer than 20 seconds, we intercept two samples randomly and without overlapping from the middle part. For AF rhythm and Noisy rhythm with fewer samples, we increase the numbers by overlapping interception. The overlap length of the former is 6 second, and the overlap length of the latter is 8 second. A more balanced secondary short-record dataset containing 19188 samples was eventually established.

B. SIGNAL DECOMPOSITION VIA DERIVED WAVELET FRAMES

The ECG signal is a non-stationary signal with strong impact characteristics. In order to ensure the integrity of the waveform, the sampling frequency is set to 300 Hz. In the frequency domain, the features are concentrated in the lower frequency band, and Afonso assert that the effective frequency band does not exceed 25 Hz [25]. From the perspective of signal processing, algorithms such as wavelet transform and local mean decomposition [26] show good performance for such mode mixing caused by intermittent impact superposition. Wavelet transform is an effective time-frequency analysis tool, which has achieved good results in baseline wandering elimination, QRS complex analysis and feature extraction [27]–[29]. Lacking shift-invariance, the traditional wavelet transform has a weak ability to identify repeated transient impacts [30]. Limited by the edge leakage of the filter, the ability of the dual-tree complex wavelet packet to decompose the information at the intersection of the band is not satisfactory.

In this paper, we apply derived wavelet frames (DWFs). It is based on the dyadic dual-tree complex wavelet packets (DDCWPs) and further enhances its ability to extract information in the transition band with additional derivative non-dyadic implicit wavelet packets. Dual-tree complex wavelet packet decomposition is constructed based on dual-tree complex wavelet basis, which consists of two scaling

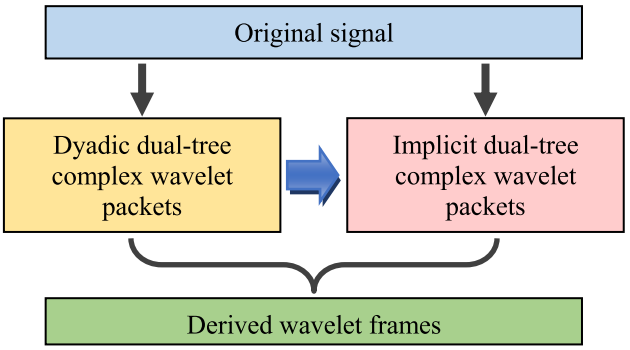


FIGURE 1. DWFs composed of DDCWPs and IDCWPs.

functions and two wavelet functions. There are many mature wavelet bases, but their frequency response functions are mostly attenuated at both ends, not ideal rectangles. Thus there is information loss or overlap at the band boundaries of the reconstructed sub-signals. In order to address this problem, as showed in figure 1, implicit dual-tree complex wavelet packets (IWPs) are constructed based on DDCWPs, and the two form the derived wavelet frames.

Assuming $\{x(n)\}$ is the input ECG signal; the following algorithm can derive the implicit wavelet packets (IWPs):

Step 1) Perform multi-scale decomposition based on dual-tree wavelet packet decomposition on the input signal. Assume that k is the number of decomposition layers and j is the sequence number of the sub-signal, such that $\{x(n)\}$ is transformed into a set of sub-signals:

$$D_k = \{D_k^j(n) \lfloor j = 1, 2, \ldots, 2^k\} \tag{1}$$

Step 2) Rearrange the elements in the set D_k according to the central frequency of wavelet packet. Let the generated set be $R_k = \{R_k^j(n) | \ j = 1, 2, \ldots, 2^k\}$, the mapping between D_k^j and R_k^j can be described as below:

For R_k^j, let the binary coding of the index j be

$$j = \sum_{m=0}^{k-1} 2^m n_m + 1 \tag{2}$$

Construct a new index as follows

$$j' = \sum_{m=0}^{k-1} 2^m n'_m + 1 \tag{3}$$

where the parameters n'_m is defined as

$$n'_m = \begin{cases} n_m, & m = k-1 \\ \mathrm{mod}(n_m + n_{m+1}, 2), & m = 0, 1, \ldots, k-2 \end{cases} \tag{4}$$

Step 3) Generate the implicit wavelet packet using the following equation

$$iwp_k^j(n) = R_k^{2j}(n) + R_k^{2j+1}(n), \quad 1 \le 2^{k-1} - 1 \tag{5}$$

The frequency-scale topology of the derived wavelet packet frame is shown in Figure 2. It can be seen that the center frequency of the derived wavelet packet is the band boundary of the traditional binary wavelet packet, thereby

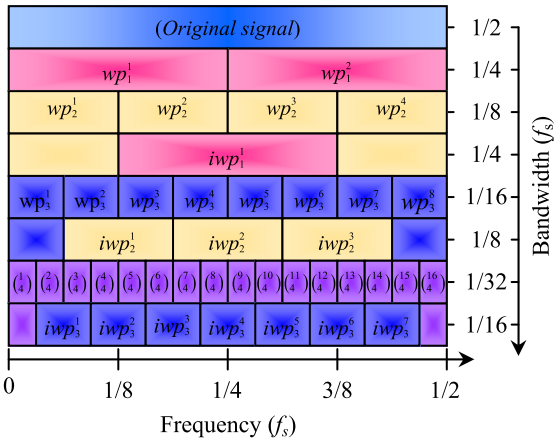


FIGURE 2. **Frequency-scale paving of DWFs.**

improving the ability of the algorithm to extract the information of the transition band. Referring to the sampling frequency and bandwidth of the original record, the number of decomposition layers is set to 4. The reconstructed sub signal is then normalized to the [0, 1] interval, eliminating interference at the recorded baseline position. Figure 3 shows samples representing each type.

C. PROPOSED CONVOLUTIONAL NETWORKS

This paper proposes two convolutional neural network model. One is the fast down-sampling residual convolutional neural network (FDResNet), as shown in Figure 4. The other is a multi-scale decomposition enhanced fast down-sampling residual convolutional neural network (MSResNet), as shown in Figure 5. FDResNet is mainly composed of a fast down-sampling module, a residual convolution module, and a classification module. Two convolutional layers with a stride of 3 are the principle part of the fast down-sampling module, which are followed by a random dropout layer and a batch-normalization layer to enhance the generalization of the model. The input sample length is 2700. The fast down-sampling module effectively reduces the calculation of subsequent deep networks, and on the other hand reduces data redundancy and facilitates model learning. Three residual convolution modules consisting of convolutional layers in series and residual short circuit follow this. The width of the three residual convolution modules is gradually increased, but all of them use the max-pooling layer to down-sample the feature vectors. The classification module consists of 1 flatten layer, 2 full connection layers and a softmax classifier. Before flatten layer, a convolution layer with a filter length of one is used to reduce the dimension of the feature vectors. After flatten layer, a random dropout layer is used to prevent overfitting.

The multis-cale residual neural network consists of three parallel FDResNets coupled by a small neural network. The three FDResNets have the same structure, but are independently trained by reconstructed samples of different scales. The trained weight matrix is directly transferred

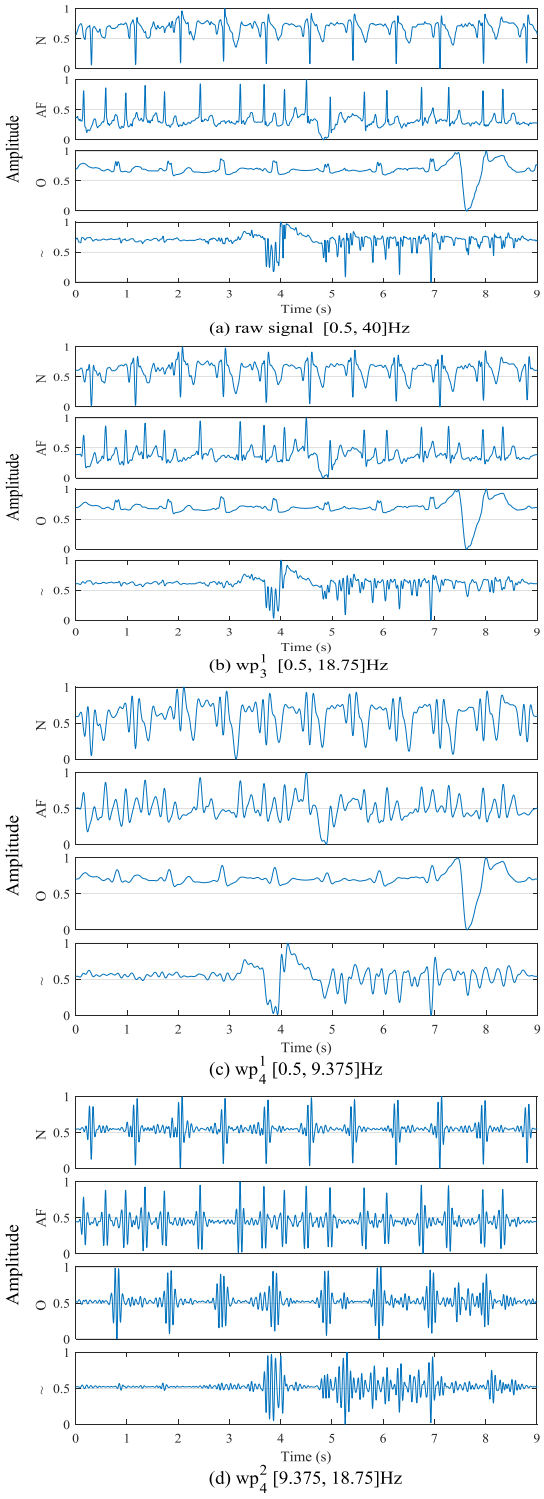


FIGURE 3. **Examples of short ECG recordings representing normal, AF, other, noisy rhythms and their reconstructed sub recordings.**

to MSResNet. Each of the three sub-networks has learned different features, and the independent classification capabilities are different. After the prediction vectors of the three sub-networks are connected into one feature vector, a small neural network learns the end-to-end characteristics of the three, and higher recognition accuracy can be obtained.

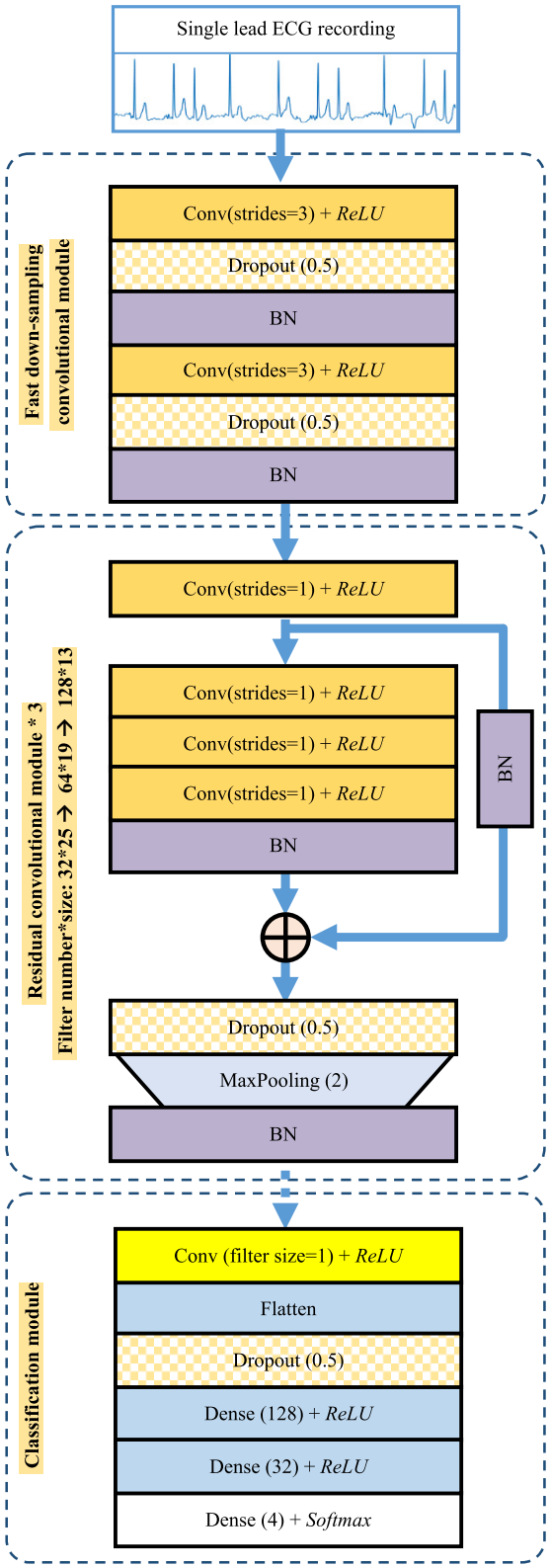


FIGURE 4. The structure of the proposed fast down-sampling ResNet.

III. EXPERIMENTAL RESULTS

In order to optimize the neural network, extensive experiments were conducted using the shot single lead ECG dataset.

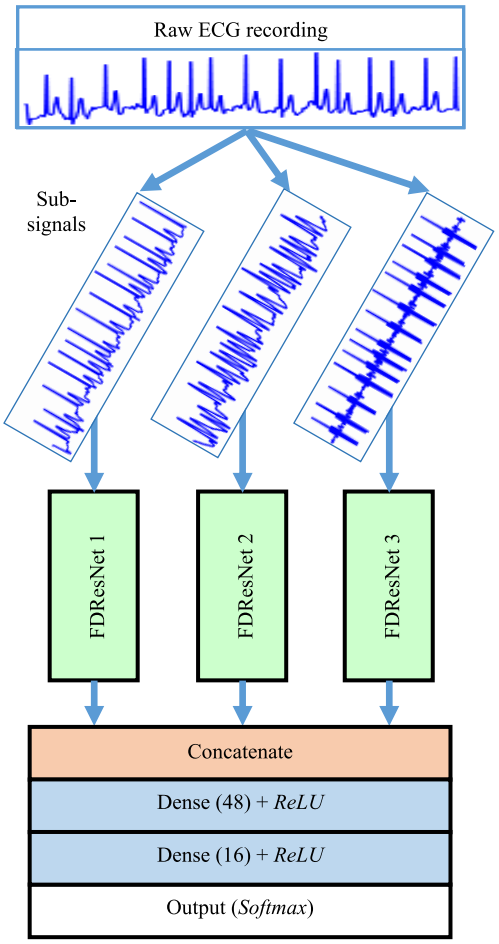


FIGURE 5. The structure of the proposed Multi-scale FDResNet.

Each different structure or combination of parameters is cross-validated by 6 fold. The experiment runs on a PC with 16GB of memory and 16GB of GPU memory.

A. PERFORMANCE EVALUATION OF DIFFERENT DOWN-SAMPLING MODULE IN FDRESNET

The down-sampling mode mainly undertakes two functions: first, quickly reducing the dimension of the feature vector and reducing the calculation of the entire model; second, concentrating the waveform features of the electrocardiogram to remove redundant details. The results of the six-fold cross-validation of the FDResNets using down-sampling modules containing different number of wide-stride convolution (WSConv) layer are shown in Figure 6, and the number of epoch is 75. The down-sampling module significantly speeds up training and improves test accuracy. It is particularly noteworthy that if the sample is directly processed by the residual convolution module, large-scale data redundancy is not conducive to the improvement of accuracy, but rather the over-fitting problem is exacerbated, and the loss value is quickly diverged. Regardless of the number of WSConvs, various down-sampling modules effectively improve accuracy and reduce loss values. Moreover, the more WSConv, the less

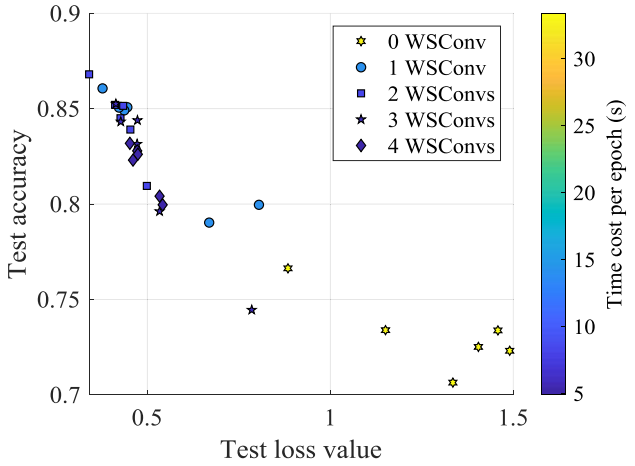


FIGURE 6. Performance of different fast down-sampling convolutional module.

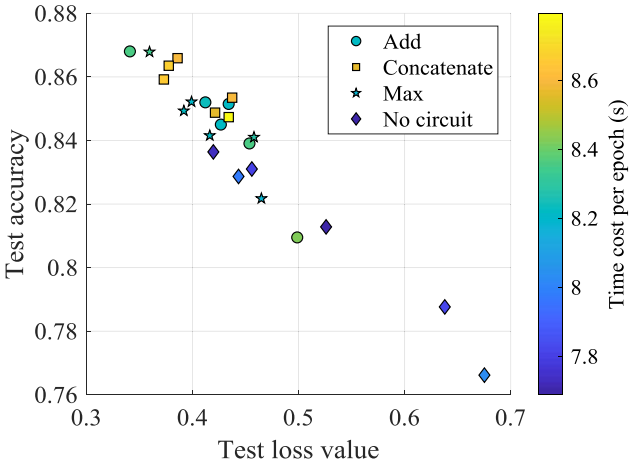


FIGURE 7. Performance of different short circuit.

the total calculation of the model, the faster the training. When only one WSConv is used or three WSConvs are used continuously, the cross-validation results of the model are scattered and the model is not stable enough. The cross-validation results of the four WSConv models are more concentrated, but the overall accuracy is lower. In summary, the fast down-sampling module containing two WSConvs is used herein.

B. PERFORMANCE EVALUATION OF DIFFERENT SHORT CIRCUITS IN FDRESNET

Adding layers can improve the learning ability of convolutional neural networks. However, feature map information and error gradients are gradually weakened as they continue to pass through deep networks. It almost disappears when it reaches the end or the initial point, and it is difficult for the model to train. In recent years, there have been many public papers focusing on this issue, in which the ResNets effectively solves this problem by using short-circuit paths [31]. Besides, several models use similar mechanisms. Liao Z introduced a competitive multi-scale CNN using a series-parallel hybrid structure in which parallel modules are jointed at the end with a maximum output unit [32]. Gao Huang introduced a densely connected CNN [33]. Each layer in the model is connected to each other, and the feature maps accumulate during the forward transmission of information.

In this paper, the experimental research on these three short-circuit methods is carried out. The experimental results of the six-fold cross-validation are shown in Figure 7 below. It can be seen that if the shorting path is not used, the training result of the model may be poor. As shown in the lower right corner of the picture, the test accuracy is only slightly higher than 0.76. On the other hand, all three short-circuit paths effectively improve the trainability of the model. For the study subjects, there is no significant difference in the applicability of the test in terms of test accuracy and test loss values. However, the concatenate unit increases the computational complexity of the model and reduces the computational efficiency. The test accuracy of the Add unit is slightly higher than that of the Max unit, and considering the optimization theory of the residual neural network, this paper finally adopts the typical residual unit.

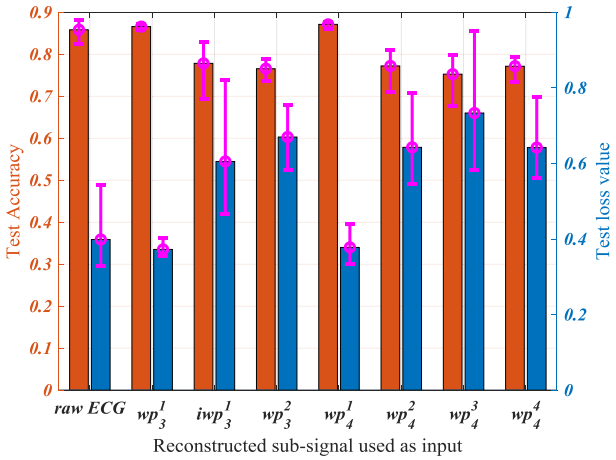


FIGURE 8. Performance comparison of FDResNets trained by reconstructed dataset of different sub-band.

C. PERFORMANCE EVALUATION OF RECONSTRUCTED DATASETS FROM DIFFERENT SUB-BAND IN TRAINING FDRESNETS

Different frequency bands of the ECG recordings carry different message. The derived wavelet frames can fidelity decompose the information of each frequency band, making the features of various ECGs more recognizable in each frequency band. In this paper, the performance of training CNN with reconstructed sub-signals in different frequency bands are studied experimentally. The results of the ten-fold crossover experiment are shown in Figure 8.

Deep convolutional neural networks can be trained using raw ECG records, resulting in an average accuracy of 85.83% in six-fold cross-validation. In contrast, the reconstructed ECG dataset of wp_3^1 ([0, 18.75 Hz]) achieves a better accuracy

TABLE 2. Test precision of the FDResNet trained with different reconstructed ECG dataset.

Sub-band	Precision				Overall F1 score
	Normal	AF	Other	Noisy	
raw signal	0.8059	**0.9816**	0.8557	0.9632	0.8702
wp_3^1	0.7982	0.9757	0.8954	**0.9767**	0.8766
wp_4^1	**0.8215**	0.9751	**0.9385**	0.9715	**0.8973**

and a lowest loss value. In addition, the difference of loss value is extremely small, indicating that the performance of the neural network is more stable. In terms of accuracy, the reconstructed ECG dataset of wp_4^1 ([0, 9.375 Hz]) achieves the best training results. Overall, the performance of the reconstructed sub-signal with lower frequency is better than the higher frequency.

The training results of the top three datasets for test accuracy are shown in Table 2. The reconstructed dataset obtained the best overall F1 score. In terms of the precision of Normal and Other samples, the training result of is better than the other two. In terms of the precision of the AF samples, the training result of the raw signal is best. At the same time, for the Noisy sample, FDResNet trained with the dataset has the highest recognition accuracy. This paper hypothesizes that the features of different types of ECG records are located in different frequency bands. In the raw record, various features interfere with each other, which increases the learning difficulty of the neural network model. The data features are sparser in datasets reconstructed from different frequency bands. For several types of ECG records, the neural network model has the potential to achieve higher recognition accuracy.

D. PERFORMANCE EVALUATION OF DIFFERENT COUPLING METHODS OF FDRESNETS IN MSRESNET

As mentioned earlier, FDResNet trained using datasets reconstructed from different frequency bands has its own advantages and disadvantages in identifying different types of ECG records. In this paper, multiple FDResNets independently trained by reconstructed datasets of different scale are coupled into one MSResNet for higher recognition capabilities. The cross-validation results of different coupling modes are shown in Figure 9.

A simple coupling method is that the three sub-networks vote on the label of the sample by the mechanism of prediction vector summation or maximum. As shown in Figure 9, the coupling method of unweighted summation has achieved a accuracy that is significantly better than that of the single-scale FDResNet in three validations, but the loss value increases significantly. The performance of unweighted extremes is relatively stable, and a certain performance improvement has also been achieved. Another coupling method is to process the predictions of three FDResNets using a small, fully connected neural network (NN). The three FDResNets use the weight matrix transferred from independent training. Only the weights of the last two fully connected

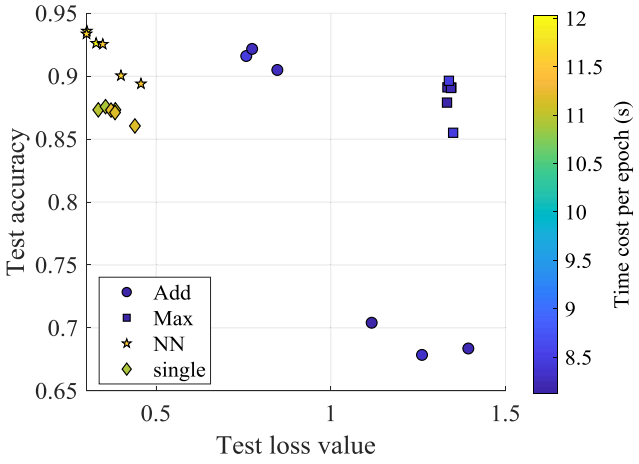


FIGURE 9. Performance of different coupling method of sub-FDResNet.

layers are updated during MSResNet training. This paper speculates that this coupling NN can autonomously learn the end-to-end output law of three single-scale FDResNet. As can be seen from Figure 9, this coupling method achieves the highest performance gains. Subsequent research can further optimize the coupling NN using algorithms such as PSO [34].

IV. DISCUSSION

A. THE EFFECTIVENESS OF THE PROPOSED METHOD

This paper aims at classifying the short single lead electrocardiogram arrhythmia, especially to identify whether the subject has atrial fibrillation based on their ECG recording. A multi-scale residual neural network is constructed by combining advanced deep learning techniques such as residual connection, random dropout, batch normalization and transfer learning. To verify the validity of the proposed method, a six-fold cross-validation is implemented after the model structure was completely determined. In each experiment, firstly, FDResNet is trained independently with training sets of different scales, then the trained weight matrix is migrated to MSResNet, and finally MSResNet is trained with these training sets. The sum of the six confusion matrices obtained by cross-validation is shown in Figure 10.

It can be seen from Figure 10 that the precision and recall rate of the model for AF and Noisy ECG records are significantly higher than the other two categories. One of the possible reasons is data imbalance. In fact, these two types of samples in the training set are augmented from a small number of original samples. In contrast, each of the Normal and Other samples comes from individual ECG collection objects. The Other class sample has the lowest recall, while some of them are recognized as the Normal record. Compared with Normal and AF electrocardiogram, Other ECG is undoubtedly the most diverse, and the recognition is the most difficult. Subdividing the Other class into multiple subclasses, or increasing the number of samples, will help improve the classification accuracy of the model.

TABLE 3. Comparison of previous study of ECG based on the PhysioNet/CinC Challenge 2017 public dataset.

method		F1 score			
Data processing	Classification	Normal	AF	Other	Overall
Original ECG records splitting by 5 seconds [35]	1D CNN containing residual blocks and recurrent layers.	0.919	0.858	0.816	0.864
Features extraction per beat and global aggregation [36]	XGBoost and LSTMs stacked by LDA	0.953	0.838	0.850	0.880
Feature extraction via R-peak detection & phasor transform PQRST and sparse coding [37]	Decision tree ensemble	0.889	0.791	0.702	0.794
Data splitting and augmentation; QRS complexes detection and STFT for spectrogram [38]	Signal quality assessment and dense convolutional neural networks	0.91	0.89	0.78	0.86
R-peak detection and RR & dRR interval extraction [39]	Decision tree and decision tree ensemble classifiers	0.93	0.88	0.82	0.87
Not applied [40]	16-layers 1D residual convolutional network	0.90	0.82	0.75	0.82
Not applied [21]	convolutional and LSTM networks ensemble	0.910	0.810	0.780	0.853
STFT for spectrogram [23]	2D convolutional network with LSTM layer	0.888	0.764	0.726	0.792
Proposed in this paper		0.881	0.966	0.851	0.899

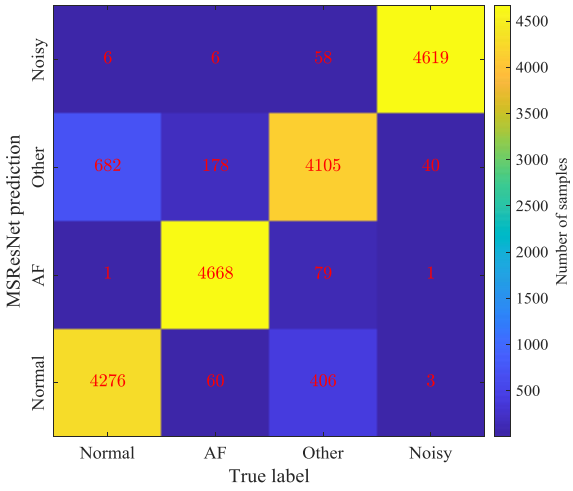


FIGURE 10. Confusion matrix.

B. COMPARISON WITH PREVIOUS STUDIES

To evaluate the proposed method, comparisons are also performed with previous studies. Table 3 lists some of the published research results of ECG classification based on the same dataset. There are both algorithms that take feature extraction plus machine learning strategies, as well as algorithms based on deep learning.

Overall, the classification model using the a priori feature extraction algorithm achieves higher recognition accuracy for the Normal ECG record. For example, Teijeiro used a large amount of expertise to extract 79 features from each individual heartbeat waveform in the ECG record, and then the recursive feature elimination is used to select 42 out of them. These single beat features are further fused into the global features of the sequence and then classified using the XGBoost algorithm. In parallel, ECG timing sequences are also classified using LSTMs. Finally, the ensemble algorithm achieves 95.3% recognition accuracy on Normal ECG record. On the other hand, the method of autonomously learning the characteristic law from ECG records using deep neural networks is also quite effective. The method proposed in this paper obtains the classification results comparable to the best research results without involving the professional knowledge of electrocardiogram. It is foreseeable that with the further accumulation of datasets, especially the increase of abnormal ECG samples and pattern subdivision, the deep learning model can achieve a more powerful classification ability.

V. CONCLUSION

This paper proposes an ECG record classification method based on the derived wavelet decomposition and the deep residual convolutional neural network. With the derived wavelet decomposition, the original ECG records are decomposed and reconstructed into sub-signals of different frequency scales. The fast down-sampled residual convolutional neural networks are trained using different scales of reconstructed datasets. Finally, multi-scale residual convolutional neural networks are trained using transfer learning. The proposed method was validated on the public short single lead ECG dataset from the 2017 PhysioNet/CinC Challenge with a test accuracy of 92.1% and an F1 score of 89.9%. The main findings of this study can be summarized as follows:

The one-dimensional deep residual convolutional neural network can learn effective classification features from the time domain ECG waveform. Wide-stride convolution can improve the stability of the model while greatly improving the

learning speed of the model. The residual short-circuit path is critical to the learning ability of the model.

In the frequency domain, the feature information of the ECG signal is mainly concentrated in the low frequency band. With the derived wavelet packet frame, the ECG signal can be decomposed into sub-signals of different scales with translation invariance. The model trained in the reconstructed dataset of the [0, 9.375 Hz] band obtains the highest recognition accuracy.

The waveform feature learned by neural networks trained using different scale reconstruction datasets are not the same. The multi-scale residual neural network trained by transfer learning outperforms any single-scale FDResNet. Compared with the previous research results, the proposed method has improved the comprehensive accuracy and F1 score, especially the F1 score of the AF ECG record is significantly improved.

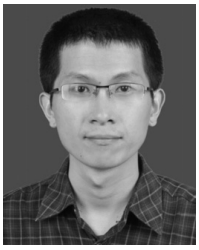
REFERENCES

[1] O. Berenfeld and J. Jalife, "Complex fractionated atrial electrograms: Is this the beast to tame in atrial fibrillation?" *Circulation, Arrhythmia Electrophysiol.*, vol. 4, no. 4, pp. 426–428, 2011.
[2] G. Y. H. Lip and H.-F. Tse, "Management of atrial fibrillation," *Lancet*, vol. 370, no. 9587, pp. 604–618, 2007.
[3] X. Zhang, Z. Liu, Q. Miao, and L. Wang, "Bearing fault diagnosis using a whale optimization algorithm-optimized orthogonal matching pursuit with a combined time–frequency atom dictionary," *Mech. Syst. Signal Process.*, vol. 107, pp. 29–42, Jul. 2018.
[4] M. Kumar, R. B. Pachori, and U. R. Acharya, "Automated diagnosis of myocardial infarction ECG signals using sample entropy in flexible analytic wavelet transform framework," *Entropy*, vol. 19, no. 9, p. 488, 2017.
[5] Y. Peng, Z. Wu, and J. Jiang, "A novel feature selection approach for biomedical data classification," *J. Biomed. Inform.*, vol. 43, no. 1, pp. 15–23, 2010.
[6] A. Diker, D. Avci, M. Gedikpinar, and E. Avci, "A new technique for ECG signal classification genetic algorithm Wavelet Kernel extreme learning machine," *Optik*, vol. 180, pp. 46–55, Feb. 2019.
[7] N. Zeng, Z. Wang, B. Zineddin, Y. Li, M. Du, L. Xiao, and X. Liu, "Image-based quantitative analysis of gold immunochromatographic strip via cellular neural network approach," *IEEE Trans. Med. Imag.*, vol. 33, no. 5, pp. 1129–1136, May 2014.
[8] C. Huang, S. Ye, H. Chen, D. Li, F. He, and Y. Tu, "A novel method for detection of the transition between atrial fibrillation and sinus rhythm," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 4, pp. 1113–1119, Apr. 2011.
[9] S. Sarkar, D. Ritscher, and R. Mehra, "A detector for a chronic implantable atrial tachyarrhythmia monitor," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 3, pp. 1219–1224, Mar. 2008.
[10] Y. Peng and X. Zhang, "Integrative data mining in systems biology: From text to network mining," *Artif. Intell. Med.*, vol. 41, no. 2, pp. 83–86, Oct. 2007.
[11] Y. Kim, "Convolutional neural networks for sentence classification," 2014, *arXiv:1408.5882*. [Online]. Available: https://arxiv.org/abs/1408.5882
[12] D. Palaz, R. Collobert, and M. Magimai-Doss, "Estimating phoneme class conditional probabilities from raw speech signal using convolutional neural networks," 2013, *arXiv:1304.1018*. [Online]. Available: https://arxiv.org/abs/1304.1018
[13] X.-C. Cao, B.-Q. Chen, B. Yao, and W.-P. He, "Combining translation-invariant wavelet frames and convolutional neural network for intelligent tool wear state identification," *Comput. Ind.*, vol. 106, pp. 71–84, Apr. 2019.
[14] Z. Wang, W. Yan, and T. Oates, "Time series classification from scratch with deep neural networks: A strong baseline," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, 2017, pp. 1578–1585.
[15] R. S. Andersen, A. Peimankar, and S. Puthusserypady, "A deep learning approach for real-time detection of atrial fibrillation," *Expert Syst. Appl.*, vol. 115, pp. 465–473, Jan. 2019.
[16] A. Sellami and H. Hwang, "A robust deep convolutional neural network with batch-weighted loss for heartbeat classification," *Expert Syst. Appl.*, vol. 122, pp. 75–84, May 2019.
[17] J. H. Tan, Y. Hagiwara, W. Pang, I. Lim, S. L. Oh, M. Adam, R. S. Tan, M. Chen, and U. R. Acharya, "Application of stacked convolutional and long short-term memory network for accurate identification of CAD ECG signals," *Comput. Biol. Med.*, vol. 94, pp. 19–26, Mar. 2018.
[18] O. Faust, A. Shenfield, M. Kareem, T. R. San, H. Fujita, and U. R. Acharya, "Automated detection of atrial fibrillation using long short-term memory network with RR interval signals," *Comput. Biol. Med.*, vol. 102, pp. 327–335, Nov. 2018.
[19] U Erdenebayar, H. Kim, J. U. Park, D. Kang, and K. J. Lee, "Automatic prediction of atrial fibrillation based on convolutional neural network using a short-term normal electrocardiogram signal," *J. Korean Med. Sci.*, vol. 34, no. 7, p. e64, 2019.
[20] 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, 2000.
[21] P. A. Warrick and M. N. Homsi, "Ensembling convolutional and long short-term memory networks for electrocardiogram arrhythmia detection," *Physiol. Meas.*, vol. 39, no. 11, 2018, Art. no. 114002.
[22] B. S. Chandra, C. S. Sastry, S. Jana, and S. Patidar, "Atrial fibrillation detection using convolutional neural networks," in *Proc. Comput. Cardiol. (CinC)*, 2017, pp. 1–4.
[23] M. Zihlmann, D. Perekrestenko, and M. Tschannen, "Convolutional recurrent neural networks for electrocardiogram classification," in *Proc. Comput. Cardiol. (CinC)*, 2017, pp. 1–4.
[24] G. D. Clifford, C. Liu, B. Moody, L.-W. H. Lehman, I. Silva, Q. Li, A. E. Johnson, and R. G. Mark, "AF classification from a short single lead ECG recording: The PhysioNet/computing in cardiology challenge 2017," in *Proc. Comput. Cardiol. (CinC)*, 2017, pp. 1–4.
[25] V. X. Afonso, W. J. Tompkins, T. Q. Nguyen, S. Trautmann, and S. Luo, "Filter bank-based processing of the stress ECG," in *Proc. 17th Int. Conf. Eng. Med. Biol. Soc.*, 1995, pp. 887–888.
[26] L. Wang, Z. Liu, Q. Miao, and X. Zhang, "Complete ensemble local mean decomposition with adaptive noise and its application to fault diagnosis for rolling bearings," *Mech. Syst. Signal Process.*, vol. 106, pp. 24–39, Jun. 2018.
[27] G McDarby, B. G. Celler, and N. H. Lovell, "Characterising the discrete wavelet transform of an ECG signal with simple parameters for use in automated diagnosis," in *Proc. 2nd Int. Conf. Bioelectromagnetism*, 1998, pp. 31–32.
[28] S. Banerjee and M. Mitra, "ECG feature extraction and classification of anteroseptal myocardial infarction and normal subjects using discrete wavelet transform," in *Proc. Int. Conf. Syst. Med. Biol.*, 2010, pp. 55–60.
[29] R. K. Tripathy and S. Dandapat, "Detection of cardiac abnormalities from multilead ECG using multiscale phase alternation features," *J. Med. Syst.*, vol. 40, no. 6, p. 143, 2016.
[30] Y. Wang, Z. He, and Y. Zi, "Enhancement of signal denoising and multiple fault signatures detecting in rotating machinery using dual-tree complex wavelet transform," *Mech. Syst. Signal Process.*, vol. 24, no. 1, pp. 119–137, 2010.
[31] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 770–778.
[32] Z. Liao and G. Carneiro, "Competitive multi-scale convolution," 2015, *arXiv:1511.05635*. [Online]. Available: https://arxiv.org/abs/1511.05635
[33] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 4700–4708.
[34] N. Zeng, Z. Wang, Y. Li, M. Du, and X. Liu, "A hybrid EKF and switching PSO algorithm for joint state and parameter estimation of lateral flow immunoassay models," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 9, no. 2, pp. 321–329, Mar./Apr. 2012.
[35] Z. Xiong, M. P. Nash, E. Cheng, V. V. Fedorov, M. K. Stiles, and J. Zhao, "ECG signal classification for the detection of cardiac arrhythmias using a convolutional recurrent neural network," *Physiol. Meas.*, vol. 39, no. 9, 2018, Art. no. 094006.
[36] T. Teijeiro, C. A. García, D. Castro, and P. Félix, "Abductive reasoning as a basis to reproduce expert criteria in ECG atrial fibrillation identification," *Physiol. Meas.*, vol. 39, no. 8, 2018, Art. no. 084006.

[37] M. Rizwan, B. M. Whitaker, and D. V. Anderson, "AF detection from ECG recordings using feature selection, sparse coding, and ensemble learning," *Physiol. Meas.*, vol. 39, no. 12, 2018, Art. no. 124007.
[38] S. Parvaneh, J. Rubin, A. Rahman, B. Conroy, and S. Babaeizadeh, "Analyzing single-lead short ECG recordings using dense convolutional neural networks and feature-based post-processing to detect atrial fibrillation," *Physiol. Meas.*, vol. 39, no. 8, 2018, Art. no. 084003.
[39] M. Shao, G. Bin, S. Wu, G. Bin, J. Huang, and Z. Zhou, "Detection of atrial fibrillation from ECG recordings using decision tree ensemble with multi-level features," *Physiol. Meas.*, vol. 39, no. 9, 2018, Art. no. 094008.
[40] Z. Xiong, M. K. Stiles, and J. Zhao, "Robust ECG signal classification for detection of atrial fibrillation using a novel neural network," in *Proc. Comput. Cardiol. (CinC)*, 2017, pp. 1–4.

XIN-CHENG CAO was born in Weifang, Shandong, China, in 1992. He received the bachelor's and master's degrees in mechanical engineering from the School of Aerospace Engineering, Xiamen University, in 2015 and 2018, respectively, where he is currently pursuing the Ph.D. degree. His current research interests include intelligent equipment, smart manufacturing, and structural health monitoring of equipment.

BIN YAO was born in Luoyang, Henan, China, in 1963. He received the Ph.D. degree from Xi'an Jiaotong University, in 2003. He is currently a Full Professor with the School of Aerospace Engineering, Xiamen University, China. His current research interests include shape of curved surface and NC technology, detection technology, and intelligent machining equipment.



BIN-QIANG CHEN was born in Fujian, China, in 1986. He received the bachelor's degree in mechanical engineering from the School of Manufacturing Science and Technology, Sichuan University, in 2008, and the Ph.D. degree in mechanical engineering from the School of Mechanical Engineering, Xi'an Jiaotong University, in 2013. His current research interests include intelligent equipment and smart manufacturing, structural health monitoring of equipment, and applied harmonic analysis.