Received August 20, 2019, accepted September 3, 2019, date of publication September 6, 2019, date of current version September 24, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2939822

An Automatic System for Real-Time Identifying Atrial Fibrillation by Using a Lightweight Convolutional Neural Network

DAKUN LAI[1], (Member, IEEE), XINSHU ZHANG[1], YUXIANG BU[1], YE SU[2], AND CHANG-SHENG MA[3]

[1]School of Electronic Science and Engineering, University of Electronic Science and Technology of China, Chengdu 610054, China
[2]Department of Cardiovascular Ultrasound and Cardiology, Sichuan Academy of Medical Sciences and Sichuan Provincial People's Hospital, Chengdu 610072, China
[3]Department of Cardiology, Beijing Anzhen Hospital, Capital Medical University, Beijing 100029, China

Corresponding author: Dakun Lai (dklai@uestc.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 61771100.

ABSTRACT A lightweight convolutional neural network (CNN) is presented in this study to automatically indentify atrial fibrillation (AF) from single-lead ECG recording. In contrast to existing methods employing a deeper architecture or complex feature-engineered inputs, this work presents an attempt to employ a lightweight CNN to confront current drawbacks such as higher computational requirement and inadequate training dataset, by using representative rhythms features of AF rather than raw ECG signal or hand-crafted features without any electrophysiological considerations. The experimental results suggested that this method presents the following significant advantages: (1) higher performances for indentifying AF in terms of accuracy, sensitivity, and specificity that are 97.5%, 97.8%, and 97.2%, respectively; (2) It is capable of automatically extracting the shared features of AF episodes of different patients and would be much robust and reliable; (3) with the cardiac rhythm features as input dataset, rather than complex transforming and classifying the raw data, thus requiring a lower computational resource. In conclusion, this automated method could analyze large amounts of data in a short time while assuring a relative high accuracy, and thus would potentially serve to provide a comfortable single-lead monitoring for patients and a clinical useful tool for doctors.

INDEX TERMS Atrial fibrillation, cardiac rhythms, convolutional neural network, deep learning, electrocardiogram, single-lead recording.

I. INTRODUCTION

Atrial fibrillation (AF) is the most common sustained cardiac arrhythmia in humans [1], occurring in 1 – 2% of the general population [2], increasing with each decade of life to a prevalence of 6% in the population older than 65 years [3], [4]. Importantly, thromboembolic stroke is one of the most serious complications of AF [5], which often leads to long-term disability or death [6], [7]. The electrocardiographic (ECG) characterization is an increasingly important measure of AF diagnosis [8]. AF is manifested on the ECG signal as an absence of the P-waves and presence of the fibrillatory waves (F-waves), which are continuous irregular and low-amplitude atrial waveforms on baseline [2], [9]. Another characteristic of AF is an irregular ventricular rate representing the variability of R-R intervals [2], [10], [11]. An early and automated identification of AF from patients' ECG signals is highly significant for effective treatments to reduce the incidence of AF-induced complications and enhancing the quality of life as well as increasing the survival rate.

In the past decade, signal processing and machine learning advocated for new and advanced method for detection of cardiac arrhythmias, and a wide variety of algorithms have been developed for automatic AF detection [12]. These works are either based on the atrial activity analysis in terms of the presence of F-waves [13]–[15] or the ventricular response in terms of the R-R intervals [16]–[19], most of which have been reported to automatically detect AF based on manual extraction of various cardiac features in time domain [13]–[16], spectrum domain [11], [20], time-frequency domain [21],

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

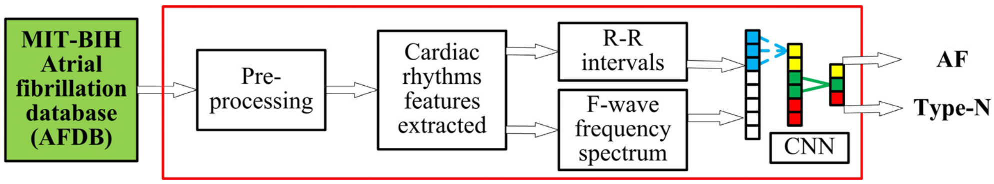

FIGURE 1. Schematic diagram of the proposed method for automated detection of atrial fibrillation (AF).

and/or nonlinear domain [22]–[24]. Nevertheless, the design of such traditionally automatic detectors is still based on manually extracted features or characteristics from recorded ECG signals during AF episodes, which is not only a highly subjective way for various groups due to currently limited knowledge about the mechanism of AF, but also tends to be a good performance if appropriate features used, and vice versa.

Deep learning is a machine learning technique based image learning where the system automatically learns and discovers the complex features from the raw data and performs the classification in an end-to-end manner [25]. Most recently, a few preliminary studies on convolutional neural network (CNN) with the raised hope that we get a deeper understanding by posing our questions to a well-trained network for detecting AF in single-lead ECG [26]–[32]. Rubin *et al.* [29] converted ECG signals into spectrogram images and input into a densely CNN to classify AF, normal, other arrhythmia and noise. Fan *et al.* [30] reported a multi-scaled fusion of deep 13-layer CNN to detect AF from single-lead ECG by employing two streams of convolutional networks with different filter size. Xiong *et al.* [31] introduced a 16-layer CNN with skip connections lead to fast training. Furthermore, Awni *et al.* [32] developed a 34-layer CNN-based method to a variety of arrhythmias from a single-lead ECG and their performance goes beyond cardiologist-level detection. Although the above mentioned studies are effective on solving the problem of AF detection and have demonstrated that CNN has ability to extract sophisticated nonlinear features from raw data or converted data without manual intervention at different levels, training such deeper and deeper CNN is a complicated task which still requires a large amount of data and a higher and higher computational requirement would also limit its application in real-time detection of AF.

In this study, a lightweight CNN by using cardiac rhythm features as input dataset is proposed to automatically indentify AF from single-lead ECG recording. In contrast to existing methods employing a deeper architecture of CNN, this work presents an attempt to employ a less 8-layer CNN to confront current drawbacks such as higher computational requirement and inadequate training dataset, by using representative rhythms features of AF rather than raw ECG signal or other hand-crafted features without any electrophysiological considerations. We examined the performance of the proposed method in an open access database, and compared with that of other recent methods based on CNN in terms of the sensitivity, specificity, and accuracy and computation complexity. This automated method could analyze large amounts of data in a short time while assure a relative high accuracy, and thus would potentially serve to provide a comfortable single-lead, real-time monitoring for patients and a clinical useful tool for doctors.

II. MATERIAL AND METHODS

A. DATASET USED

The experimental study was conducted based on data from the AFDB database, which is publicly accessible from the Physionet [33]. This database includes 25 long-term ECG Holter recordings from different subjects (mostly paroxysmal). Note that two of the 25 recordings (records 00735 and 03665) have been excluded from our study as the signals are unavailable. Each recording has a duration of ten hours containing two ECG signals labeled 'ECG1' and 'ECG2' and acquired with a sample rate of 250 Hz and 12-bit resolution over a range of ± 10 mV. All of 23 recordings were marked by various annotations manually by expert clinicians including atrial fibrillation (AF), atrial flutter (AFL), atrial-ventricular junctional rhythm (AVJ) and all other rhythms (type-N), which was given as the golden standard to represent changes of the heart rhythm in this work. Furthermore, the R peaks are labeled and the R-R interval sequence was extracted based on these labels. In this work, only ECG1 signal with marks of AF (291 episodes of AF) and type-N are used to evaluate the AF detection methods.

B. THE AUTOMATIC SYSTEM FOR AF DETECTION

The schematic diagram of the proposed automatic AF detection system on the basis of a lightweight CNN is illustrated in Fig. 1. Briefly, a single-lead ECG signal is firstly preprocessed for noise reduction and segmented by a certain length of time (10 s in this study). Secondly, an initial traditional features extraction of AF by considering its representative cardiac rhythms is performed to detect the corresponding R-R intervals and F-waves frequency spectrum. Thirdly, the results of initial features extraction are taken as the input of a lightweight CNN based classifier for a final identification of AF. Additionally, the performance of the proposed method on the open access MIT-BIH Atrial Fibrillation database (AFDB) is assessed so as to evaluate its clinical applicability.

1) PRE-PROCESSING

Noise reduction can greatly increase the precision and efficiency of detection algorithm, especially for the performance

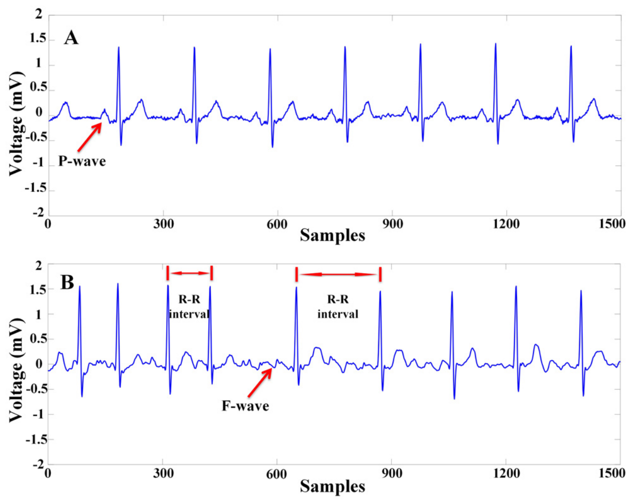

FIGURE 2. Comparison of Electrocardiographic characteristics between of (A) a normal subject and (B) an patient with atrial fibrillation (AF), where two representative cardiac rhythm features of AF are shown in (B) including a obvious variety in R-R intervals among heart beats and the presence of numerous low-amplitude F-waves instead of the P-wave.

of neural network that is highly dependent on the quality of input data. In this work, all ECG data collected from the open access AFDB database was pre-processed firstly. For each recording, continuous signal segmentation with a ten-second length was carried out. Furthermore, the baseline drift and the high frequency noise of each ECG fragment were removed by median filter and a band-pass filter (0.5-100Hz), respectively. Additionally, a sliding filter was developed to remove power frequency interferences.

2) LOW LEVEL FEATURES EXTRACTION

An ECG is one the best features to describe heart diseases, which carries valuable information about electrophysiological properties of the heart [1], [2]. Mechanisms of AF have been comprehensively studied during the last decades, and data suggest that common underlying electrophysiological substrates could be found in all AF [34]. For a normal person, the recorded ECG is usually a regular electrical signal with periodic P-QRS-T waves which is significantly different when AF occurs. As described in Introduction of this paper, two representative cardiac rhythm features including the R-R intervals and the F-waves can be easily found in ECG signal recorded in an AF patient, as shown in Fig. 2. In detail, there is obviously variety in R-R intervals among heart beats where an R-R interval is of much longer than the other one. Meanwhile, the absence of P-wave can be found in the AF signal while numerous low-amplitude F-wave appeared near where the P-wave should be. For the purpose of the low level features extraction of cardiac rhythms with physiological significance, a series of ECG signal analysis were used in this study such as R-wave detection, R-R interval calculation, and F-wave transformation, which were reported by our previous study [35]. As shown in the Fig. 3, examples of two low level features extraction of cardiac rhythm in normal subject and AF patient are listed in filtered ECG with R-wave location, Q-T signal, F-wave signal, F-wave frequency spectrum and R-R interval series.

3) CNN MODEL FOR TUNING DEEP FEATURES

Although the traditional features of AF can be extracted and directly used to distinguish an AF signal from a normal signal as mentioned above, a better and reliable performance could be expected with deep and comprehensive features hidden in these traditional features, especially for single-lead, patch-based ambulatory ECG recordings highly disturbed by varieties of noise and moving artifact. Therefore, we further explore deep features of AF that have diversity with different levels for a much accurate detection of AF by employing a CNN model as followings.

Deep learning methods try to develop the model by using all available information from the input. Extracting this information yield the implicit knowledge which underpins the robust decision making process. CNN is a deep feed forward artificial neural network which can extract deep features from input data. As shown in Fig. 4, the features extracting part comprises of convolutional layer and pooling layer. A convolutional layer applies a set of weights to process small local parts of the features from the raw input, then the result of convolution were fed into an activation function such as ReLU to obtains feature maps as the input of next layer as the below equation:

$$c = \sigma(b + \sum_{n} wx_{i+k}) \tag{1}$$

where σ is the activation function producing nonlinearity, b is the bias of activation map, n is the size of convolutional kernel, w is the weight of convolutional kernel and k is the convolutional stride. Then the result of convolutional layer was fed into a pooling layer to reduce the dimension of feature maps and the number of parameters. The equation is as follows:

$$P = \max_{t \in T} c_{i+s} \tag{2}$$

where t is the size of pooling window and s is the pooling stride. At the end of CNN, fully-connected layer was usually used to classify the previous features and got the result of current task. Then the back propagation algorithm was used in CNN to compare the result with golden standard to reduce loss in CNN. Therefore, the weight could be adjusted through numbers of iteration like the back propagation algorithm and improve the performance of CNN. Particularly, batch normalization was used to accelerate training in this work. Since the method of CNN has to adapt to the inhomogeneous data at every iteration, it was trained slowly and hard to get good performance. Early normalization method was computed by the below equation:

$$\hat{x} = (x - \mu)/\sigma \tag{3}$$

where x is the input, μ and σ are the mean and the variance of input. However, this method changes the data distribution and

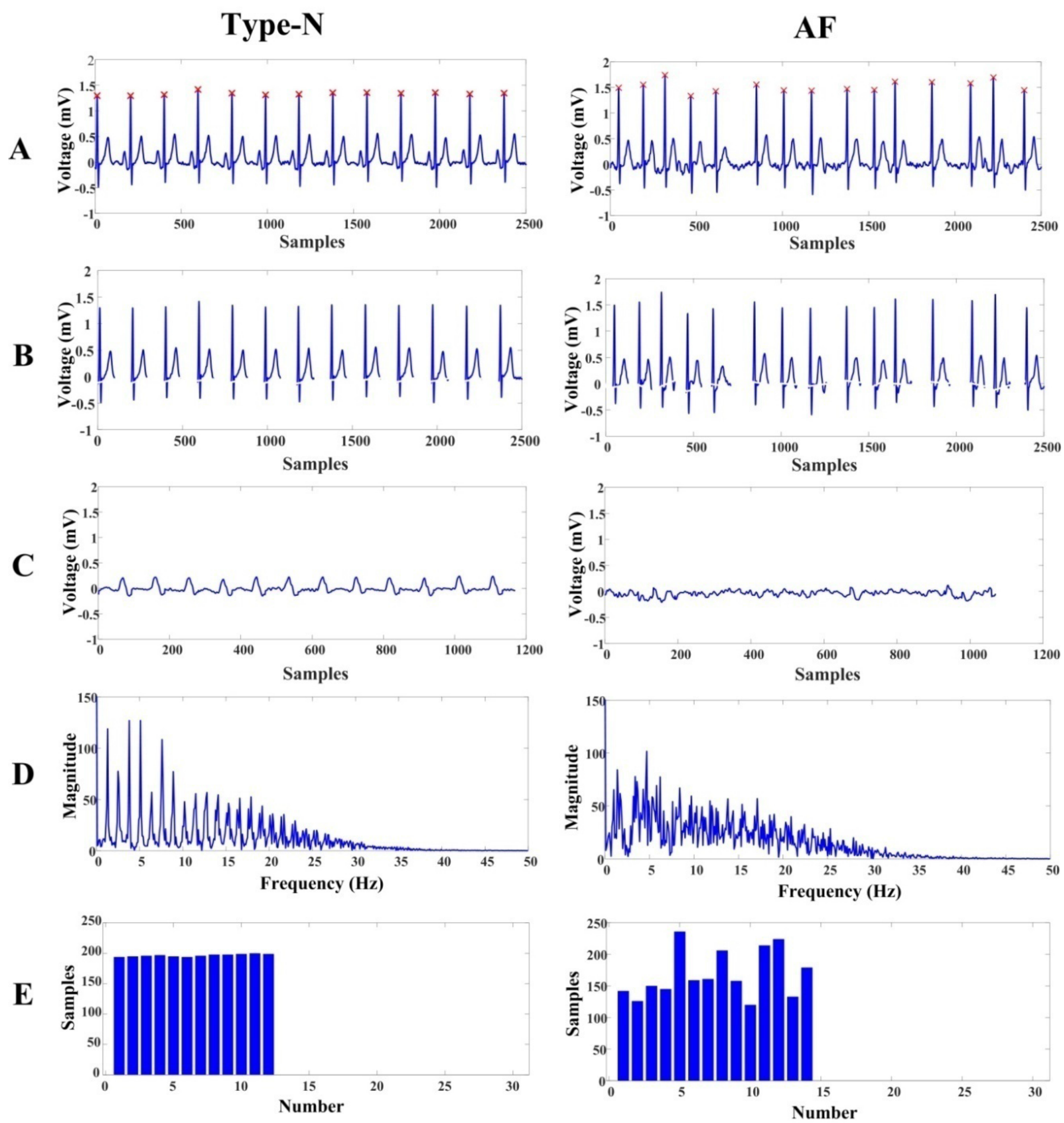

FIGURE 3. Low level features extraction of cardiac rhythms both of normal subject (type-N) and patient with atiral fibrillation (AF): (A) filtered ECGs with R-wave location, (B) Q-T signal, (C) F-wave signal, (D) F-wave frequency spectrum, and (E) R-R interval series.

relative features. To solve this problem, the method of batch normalization is updated through the equation:

$$y^{(k)} = \gamma^{(k)}\hat{x}^{(k)} + \beta^{(k)} \tag{4}$$

where the γ and β are trained at each iteration, k means the number of iteration. The CNN judged the effect of normalization and then reversed normalization to some degree through the two parameters. Therefore, the CNN could achieve a good effect of normalization.

4) CLASSIFICATION

As the final step of the proposed method, a lightweight CNN is designed as a binary classifier which distinguishes between AF and other signals in this work. The details of the proposed lightweight CNN for automatic AF detection are shown in Fig. 5. As shown in Fig. 5, there are actually two 8-layer CNNs including the R-R intervals based CNN (named RRI_CNN) and the F-wave spectrum based CNN (named FWS_CNN), each of which comprises two convolution layers, two pooling layers, one batch normalization layer, and one fully-connected layer besides the input and output layers. For each of the designed CNN, either the extracted R-R interval or F-wave frequency spectrum from an original ECG segment is used as its input dataset separately for a deep learning of the hidden features of AF. After the fully-connected layer, both of the RRI_CNN and the FWS_CNN can individually distinguish an AF ECG fragment from normal ECGs and correspondingly output a probability of AF detection. Moreover, these two probabilities of AF detection obtained from both of the RRI_CNN and the FWS_CNN

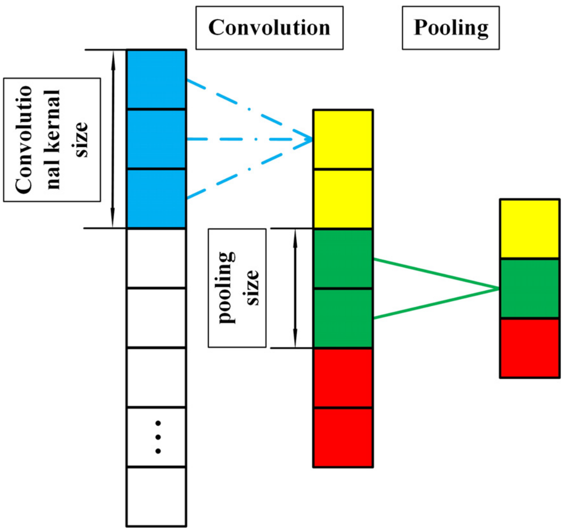

FIGURE 4. Illustration of a basic construction of CNN with an input layer, a convolutional layer and a pooling layer.

are summarized to achieve a combinational probability so as to finally determine whether this ECG segment is AF or not by using a percentage threshold parameter P: a data segment is considered to be AF only when the summarized percentage of two probabilities is greater than or equal to P. In contrast, a non-AF data segment means that the summarized percentage is less than P. According to our preliminary results as well as the experimental results reported by Asgari *et al.* [36], we set the value of P to be 50% in this study, where the combination of both the RRI_CNN and the FWS_CNN is named as the COM_CNN that has a summarized percentage of two probabilities. The parameters of both CNN models will be determined experimentally in the next section.

C. TRAINING AND OPTIMIZATION OF CNN MDOELS

In this section, we trained and analyzed the best CNN model by using the R-R intervals series and the F-wave frequency spectrum as input datasets for AF detection. There are numerous parameters in a CNN that have a significant impact on its classification accuracy. The settings used tend to be based on experience and practical considerations [37]. For the RRI_CNN proposed above, a total of eight CNN models were trained with different parameters to select the best one in this study. We carried out experiments using five-fold cross validation with all eight models on the same dataset, with a total of 33,509 AF segments and 49,952 type-N segments. All Models were designed on the basis of the traditional concept that the number of kernels increases in each layer with increasing network depth. Whereas, the size of kernels in models M1-M4 was increased from 1×3 to 1×13, respectively. For models M4-M6, the size of its fully-connected layer was set to 1,024, 512, and 2,048, respectively. Moreover, two more attempts on various numbers of kernels were conducted on M7 and M8. For the FWS_CNN, similar experiments on eight CNN models were performed by using the F-wave spectrum of each ECG segment as input dataset.

D. EVALUATION PROTOCOL

As the target dataset is relatively small, a stratified five-fold cross-validation strategy was used to tune both the model architecture and the hyperparameters and then to evaluate the model performance in this study. Specifically, the original annotated ECG signal is divided continuously every 10 s segments as described in the signal pre-processing stage. For all of 23 ECG recordings collected from the AFDB in this work, the dataset of AF therefore consisted of 33,509 AF segments and the dataset of type-N included a total of 49,952 normal segments. All extracted data segments were random divided into five subsets, each time four subsets were used for training the CNN and the remaining one subset was used for test. Therefore, five CNN models were obtained and the performance of the proposed method could be estimated by the average performance of these five CNNs. Results from this stratified five-fold cross-validation are given in Section 3.

As a consequence, three indexes including the sensitivity (Se), the specifically (Sp), and the accuracy (Acc) are calculated to evaluate the performance of the proposed CNN classifier with and without cardiac rhythms features on the open access AFDB dataset. Results from the experimental study are given in Section 3, respectively. Based on true positive (TP), true negative (TN), false positive (FP) and false negative (FN), three performances can be calculated with the formulas:

$$Sensitivity = TP/(TP + FN) \tag{5}$$

$$Specificity = TN/(TN + FP) \tag{6}$$

$$Accuracy = (TN + TP)/(TP + TN + FP + FN) \tag{7}$$

III. RESULTS

A. EXPERIMENTAL ENVIROMENT

In this work, all of the designed lightweight CNNs ran on the deep learning framework Tensorflow 1.6, using the Microsoft Windows 7 operating system. The platform of Tensorflow was deployed on a desktop computer with an Intel E3-1230 processor with 32 GB memory. This computer was also equipped with an NVIDIA RTX-2070 GPU with 8 GB memory to speed up the process and allow for a faster classification of long-term recordings.

B. OPTIMIZED LIGHTWEIGHT CNN MODELS

Tables 1 and 2 give the obtained average performance results for five-fold cross-validation of different CNN models. Details of optimized lightweight models of the RRI_CNN and FWS_CNN in this work are described as following:

(1) Input: two separate inputs with a size of 1×30 and a size of 1×500 are designed to input the calculated R-R intervals series and the converted F-wave frequency spectrum, respectively.

(2) Conv1 layer: In this study, effectively the convolution layer acts to learn and extract the deep features from both

TABLE 1. Optimization of FWS_CNN models with F-wave frequency spectrum features and the mean performance using five-fold cross-validation[1].

Model		M1	M2	M3	M4	M5	M6	M7	M8
CONV1	No. of kernels	32	32	32	**32**	32	32	32	64
	size	1×3	1×5	1×11	**1×13**	1×13	1×13	1×13	1×13
	stride	1×2	1×2	1×2	**1×2**	1×2	1×2	1×2	1×2
POOL1	size	1×2	1×2	1×2	**1×2**	1×2	1×2	1×2	1×2
	stride	1×2	1×2	1×2	**1×2**	1×2	1×2	1×2	1×2
CONV2	No. of kernels	96	96	96	**96**	96	96	64	128
	size	1×3	1×5	1×11	**1×13**	1×13	1×13	1×13	1×13
	stride	1×2	1×2	1×2	**1×2**	1×2	1×2	1×2	1×2
POOL2	size	1×2	1×2	1×2	**1×2**	1×2	1×2	1×2	1×2
	stride	1×2	1×2	1×2	**1×2**	1×2	1×2	1×2	1×2
Dropout		0.8	0.8	0.8	**0.8**	0.8	0.8	0.8	0.8
FC_1		1024	1024	1024	**1024**	512	2048	1024	1024
FC_2		2	2	2	**2**	2	2	2	2
Acc		93.4%	93.9%	94.0%	**94.5%**	94.2%	94.0%	94.1%	94.4%
Se		95.4%	96.2%	95.0%	**93.2%**	93.5%	92.8%	93.9%	94.9%
Sp		92.1%	92.4%	93.3%	**95.3%**	94.7%	94.8%	94.1%	94.1%

[1]FC=Fully-connected layer, Acc= accuracy, Se= sensitivity, Sp=specificity

TABLE 2. Optimization of RRI_CNN models with R-R interval features and the mean performance using five-fold cross-validation[1].

Model		M1	M2	M3	M4	M5	M6	M7	M8
CONV1	numbers	32	**32**	32	32	32	32	32	64
	size	1×3	**1×3**	1×5	1×7	1×3	1×3	1×3	1×3
	stride	1×2	**1×1**	1×1	1×1	1×1	1×1	1×1	1×1
POOL1	size	1×2	**1×2**	1×2	1×2	1×2	1×2	1×2	1×2
	stride	1×2	**1×2**	1×2	1×2	1×2	1×2	1×2	1×2
CONV2	numbers	96	**96**	96	96	96	96	64	128
	size	1×3	**1×3**	1×5	1×7	1×3	1×3	1×3	1×3
	stride	1×2	**1×1**	1×1	1×1	1×1	1×1	1×1	1×1
POOL2	size	1×2	**1×2**	1×2	1×2	1×2	1×2	1×2	1×2
	stride	1×2	**1×2**	1×2	1×2	1×2	1×2	1×2	1×2
Dropout		0.8	**0.8**	0.8	0.8	0.8	0.8	0.8	0.8
FC_1		1024	**1024**	1024	1024	512	2048	1024	1024
FC_2		2	**2**	2	2	2	2	2	2
Acc		93.3%	**95.8%**	93.8%	92.6%	95.5%	95.5%	95.3%	95.7%
Se		93.8%	**95.7%**	93.5%	93.8%	95.4%	95.2%	95.8%	95.9%
Sp		93.1%	**95.9%**	94.1%	91.8%	95.6%	95.7%	95.0%	95.6%

[1]FC=Fully-connected layer, Acc= accuracy, Se= sensitivity, Sp=specificity

the input R-R intervals data sequence and F-wave frequency spectrum. Specifically, this layer contains 32 features maps for RRI_CNN and FWS_CNN, respectively. The kernel size and stride of the RRI_CNN is 1×3 and 1×1, while those of the FWS_CNN are 1×13 and 1×2, respectively. Therefore, numbers of features map of the RRI_CNN and the FWS_CNN are $1 \times 28((30\text{-}3)/1+1 = 28)$ and 1×244 $((500\text{-}13)/2+1 = 244)$, respectively.

(3) Pooling1 layer: One pooling layer is designed for both of the RRI_CNN and the FWS_CNN in this work, which is aimed at progressively reducing the dimension of feature maps in order to avoid overfitting and reduce the amount of calculation. With the kernel size of 1×2 and the stride of 1×2 in this pooling layer, sizes of feature maps of the RRI_CNN and the FWS_CNN are transformed into 1×14 $(28/2 = 14)$ and 1×122 $(244/2 = 122)$, respectively.

(4) Conv2 layer: One more convolutional layer after the Pooling1 layer is designed similarly to the Conv1 layer which has the same size and stride of Conv1. Therefore, sizes of current feature maps of the RRI_CNN and the FWS_CNN are reduced as 1×12 $((14\text{-}3)/1+1 = 12)$ and 1×55 $((122\text{-}13)/2+1 = 55)$, respectively.

(5) Pooling2 layer: One more pooling layer after Conv2 is also designed similarly to the Pooling1 layer with the same size and stride. Sizes of feature maps of the RRI_CNN

TABLE 3. Experimental results of the proposed automatic system with various CNN models for AF detection in the open access AFDB database by using five-fold cross-validation, including the FWS_CNN with raw data, the RRI_CNN with R-R intervals features, the FWS_CNN with F-wave spectrum features, and FWS_CNN and RRI_CNN with a combination of two features[1].

Fold	Raw data			R-R intervals features			F-wave spectrum features			Combined features		
	Se (%)	Sp (%)	Acc (%)	Se (%)	Sp (%)	Acc (%)	Se (%)	Sp (%)	Acc (%)	Se (%)	Sp (%)	Acc (%)
1	89.4	82.8	86.2	95.8	96.0	95.9	92.4	95.6	94.3	97.7	97.4	97.5
2	90.2	83.0	87.0	95.1	96.0	95.6	93.6	95.7	94.8	97.7	97.3	97.5
3	88.7	83.1	86.1	96.2	96.2	96.2	92.0	95.1	93.9	97.7	97.4	97.5
4	89.0	84.1	86.7	96.4	95.1	95.6	94.0	95.1	94.7	98.0	96.8	97.3
5	90.2	80.6	85.6	94.9	96.4	95.8	93.8	95.2	94.7	97.7	97.3	97.5
Mean	**89.5**	**82.7**	**86.3**	**95.7**	**95.9**	**95.8**	**93.2**	**95.3**	**94.5**	**97.8**	**97.2**	**97.5**

[1]Acc= accuracy, Se= sensitivity, Sp=specificity.

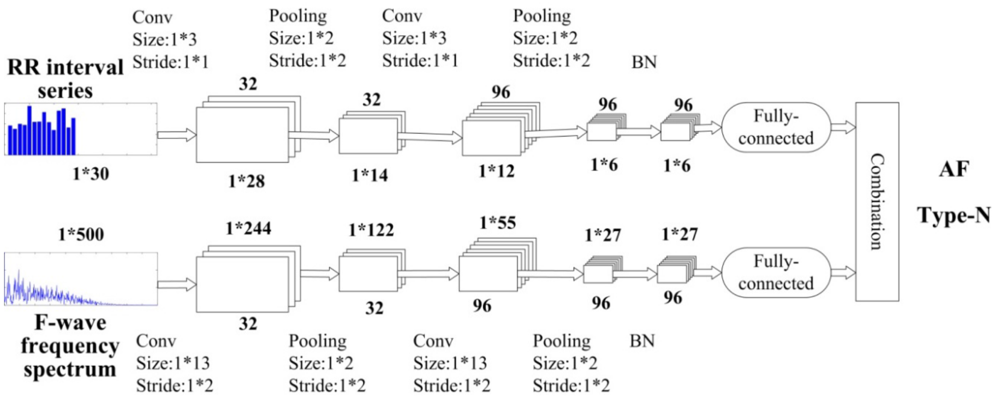

FIGURE 5. Illustration of the proposed lightweight CNN for automatic AF detection.

and the FWS_CNN are of 1 × 6 (12/2 = 6) and 1 × 27 (55/2 = 27), respectively.

(6) BN layer (batch normalization): Moreover, a normalization layer is set for each of CNN to accelerate the training process of CNN.

(7) Fully-connected layer: This layer is designed individually for the RRI_CNN and the FWS_CNN to combines the previous feature maps so as to calculate the probability of AF, correspondingly.

(8) Output: In this work, two probabilities of AF obtained from the fully-connected layer are combined by summation so as to get a final result of AF detection as described above.

Fig. 5 shows the proposed CNN models with optimized parameters for AF detection. Based on the overall results, the model M4 for the RRI_CNN showed a better performance than others, which worked better with a dropout of 0.8 and 1,024 neurons (M4) in the fully connected layer rather than 512 (M5) or 2048 neurons (M6). For the FWS_CNN, the model M2 shows a best performance for AF detection in comparison with the other seven models.

C. EVALUATED PERFORMANCE OF THE PROPOSED SYSTEM FOR AF DETECTION

Table 3 gives the experimental results of the performance of the proposed method for AF detection by using the stratified five-fold cross-validation strategy. As shown in Table 3, there is an obvious difference of classification performance between of the proposed AF detectors with and without consideration of cardiac rhythms features. In this paper, four validation experiments of the trained CNN models on a total of 83,461 segments ECG signal collected from the AFDB database were implemented to address AF screening problem for comparison, which include the raw data, the R-R interval series only, the F-wave frequency spectrum only and the combination of these two features. Although the best model M4 of FWS_CNN as given in Table 1 was used to test the raw data, a significant increase of the computation cost for the training process (14 minutes) and an obvious decrease of the performance (approximately 10%) both were found in comparisons with the same FWS_CNN model with the F-wave features input and the simpler RRI_CNN model with the R-R intervals. It could be addressed the reason why the previous works [32], [28], [39] designed a deeper and deeper network with the raw data so as to achieve an accepted performance. Moreover, the evaluated average accuracies of various CNN models including the RRI_CNN with the R-R interval series, the FWS_CNN with F-wave frequency spectrum and their combination are 95.8%, 94.5%, and 97.5%, respectively. Importantly, the averaged sensitivity, specificity and accuracy of the combinational CNNs by using a five-fold

TABLE 4. The comparison of the performance and computational cost of AF detection between our proposed detector with and without cardiac rhythms feature and the other recently reported detectors.

Author (year)	Database used	Methodology		Best performance and computational costs				
		Low level Features	Classifier	Acc (%)	Se (%)	Sp (%)	Training time	Test time
Wei et al. [38] (2019)	AFDB	RCN	CNN (6 layers)	94.59	94.28	94.91	9.65 h	0.1186 s for 1 sample (one heartbeat)
Xu et al. [39] (2018)	AFDB	2D MFSWT	CNN (12 layers)	84.85	79.05	89.99	Not available	Not available
Xia et al. [40] (2018)	AFBD	2D STFT, 2D SWT	CNN (14 layers)	98.29 98.63	98.34 98.79	98.24 97.87	40 min 16min	Not available
Yanun et al. [32] (2019)	Clinical data	Raw data	CNN (34 layers)	-	86.1	94.1	Not available	Not available
Andersen et al. [41] (2019)	AFDB MITDB NSRDB	RRI	LSTM+CNN (6 layers)	87.40	98.96	86.04	40 min	0.92 s for 24 h data[2] (exclusive of the pre-processing time)
Present study	**AFDB**	**RRI+FWFS**	**CNN (8 layers)**	**97.5**	**97.8**	**97.2**	**10 min**	**0.004 s for 1 sample (10s) and 0.25s for 24 h data[2]**
		Raw data	**CNN (8 layers)**	**86.3**	**89.5**	**82.7**	**22 min**	**0.003 s for 1 sample (10s) and 0.23 s for 24 h data[2]**

[1]AFDB means the MIT-BIH AF Database, MITDB means the MIT-BIH Arrhythmia Database, NSRDB means the MIT-BIH NSR Database, All are open access databases (Physionet) [33], RCN=recurrence complex network, 2D=two dimensional, MFSWT=modified slice wavelet transform, CNN=convolutional neural network, LSTM=long short-term memory, STFT=short-term Fourier transform, SWT=stationary wavelet transform, RRI=R-R intervals, FWFS=F-wave frequency spectrum.
[2]Computational costs for testing 24 h data are not include the signal pre-processing and initial feature extraction process, which is around 35 s for the whole test procedure in this work, whereas it is not available in [41].

validation are 97.8%, 97.2%, and 97.5, respectively. This finding could indicate that either the R-R interval series or the F-wave frequency spectrum is a relative independent feature, which could complement one another and improve the combinational performance for AF detection. As such, the combination of the FWS_CNN and RRI_CNN with both of two rhythm features proposed in this work would provide a higher performance for AF detection than either one of them individually.

IV. DISCUSSIONS

A. HIGHER PERFORMANCE COMPARING TO EXISTING AF DETECTORS

To illustrate the effectiveness of the proposed lightweight CNNs by using cardiac rhythms features as input dataset, its performance for AF detection on the same database is compared with those of other recent AF detectors, as shown in Table 4. It can be seen that the proposed lightweight CNN tends to outperform those algorithms without considerations of cardiac rhythms. Wei *et al.* [38] achieved an accuracy of 94.59% for AF detection by using a CNN based classifier with initial features of recurrence complex network, while a longer training time of 9.65 h is needed for this 6-layer CNN. Meanwhile, Xu *et al.* [39] and Xia *et al.* [28] reported two CNN based detectors with more than 10 layers with an accuracy of 84.5% and 98.63%, respectively. Both of them ran on converted two dimensional images from ECG signal by using various wavelet or Fourier transform. Although the training time around 40 minutes can be significantly saved, their low level features extraction in the subsequent test process would be the most time-consuming process. Additionally, compared with the proposed CNN models with cardiac rhythms features, an obvious decrease in the performance of the same model (FWS_CNN) by using the raw data as input dataset rather than rhythms features was observed in term of the accuracy (86.3%), sensitivity (89.5%), and specificity (82.7%), as shown in Table 3.

Apart from these previous methods with complex feature-engineered inputs, both of a 34-layers CNN with the raw data [32] and a 6-layer CNN with the R-R intervals [40] were reported in 2019 to successfully distinguish AF from others rhythms. The former achieved a sensitivity of 86.1% and a specificity of 80%, and the accuracy of the latter one was given as 87.4%. Looking beyond the narrow area of AF detection with deep leaning applied on the same database of AFDB, we presented an attempt to decrease computational cost and simultaneously keep an accepted accuracy by introducing cardiac rhythm features instead of feature-engineered characteristics into two lightweight 6-layer CNNs. The superiority of the proposed method can be addressed as that the initial extraction of rhythm features including R-R interval series and F-wave frequency spectrum could reduce the complexity for AF detection of CNN and that the combined two features could also improve the performance of each other, whereas some of the other method used one of these features such as R-R interval only. Therefore, the proposed combination of two lightweight 6-layer CNNs for AF detection could effectively solve the problem between performance and computational cost, especially for a practical capability in real-time, ambulatory applications.

B. REAL-TIME IDENTIFICATION OF AF WITH LOWER COMPUTATIONAL COST

For a CNN based classifier, training the CNN model is a time-consuming process. However, the training process can be carried out off-line. In this work, the proposed lightweight CNNs by using two cardiac rhythm features as input can be trained in less than 10 minutes to converge with 2,000 back propagation iteration, and the training time is also less than others previous works [28], [38], [40], as shown in Table 4. Two of the important reasons that we used less time for training possibly include there two initially extracted representative rhythm features and a lightweight CNN model, which both could speed up the deep learning of hidden features and obtain a faster training of our proposed model. As for the testing process of our lightweight CNNs, several experiments were performed on the recorded 24 h clinic data to obtain the time cost, and it was revealed that the testing process is about 0.004 s for one sample of 10 s data and 0.25 s for 24 h data, which shows a practical capability in real-time AF detection. Most recently, Wei *et al.* [38] used a CNN based method for AF detection to process one sample (one heartbeat) in 0.1186 s, and Andersen *et al.* [40] reported that a 24 h single-lead data can be classified in 0.92 s. Both of these two studies also reported that their method could be used to detect AF in real-time processing. It should be noted that the low level features extraction in previous works and the rhythms features extraction in this work would be the most time-consuming process (35 s for 24 h data in our work), especially for the earlier works [28], [39] by using two dimensional wavelet transform or recurrence complex network.

C. LIMITATIONS AND FUTURE WORKS

However, this study has several important limitations. First, the relatively small data was collected in the present work. Further work would be interested to collect a great number of clinical data to train and validate the proposed CNN for a higher accuracy. Second, it should be noted that the problem of overfitting may appear during the training process that may reduce the accuracy of the identification of real AF. Which means the production is too close to the training set and it may therefore fail to fit additional data or predict future observations reliably. A cross-validation with other clinic data like physical symptoms would be useful to handle the problem of overfitting. Additionally, the training process of CNN classifier could take a relatively long time even under an offline situation. However, the currently computational ability of power computer would help on saving time of training. By the way, the time consuming of testing process for clinical application is much short and efficient, the accepted computation cost of 0.004 s for 1 sample (10s) and 0.25 s for 24 h data in this work would be potentially accepted for a real-time clinic use.

V. CONCLUSION

This study suggested that the lightweight convolutional neural network by using cardiac rhythms features as input dataset could serve as an accurate and fast tool for automatic detection of AF from single-lead ECG recording. The effectiveness of the proposed method is evaluated by the open access AFDB database. The proposed methodology would present the following significant advantages: (1) compared with the recently reported AF detectors based on either a lightweight CNN using the feature of R-R intervals only or a deeper CNN using raw data, the proposed method achieves a higher accuracy using both features of the R-R intervals and F-wave frequency spectrum as input dataset for lightweight CNNs, thus considering more specific information not only in the time domain, but also in the spectrum domain and related electrophysiological significances; (2) it is capable of automatically extracting the shared features of AF episodes of different patients and would be much robust, unlike other traditional automated methodologies proposed in the literature where characteristics of AFs were extracted manually on basis of researchers' knowledge, that probably is prone to observer bias; (3) with the proposed electrophysiological features extraction, all suspicious AF could be initially found out and input into the CNN for deep feature extraction and subsequent classification, rather than all raw data or its two dimensional image from complex Fourier or wavelet transformation, thus requiring a lower computational resource. As such, this AF detector combing the initial features extraction and the CNN classification could analyze large amounts of data in a short time while assuring a relative high accuracy, and thus would potentially serve to provide a comfortable single-lead, real-time monitoring for patients and a clinical useful tool for doctors.

ACKNOWLEDGMENT

The authors would like to thank Biomedical Imaging and Electrophysiology Laboratory (BMI-EP), University of Electronic Science and Technology of China for providing computational and biomedical equipments.

REFERENCES

[1] S. T. Mathew, J. Patel, and S. Joseph, "Atrial fibrillation: Mechanistic insights and treatment options," *Eur. J. Internal Med.*, vol. 20, no. 7, pp. 672–681, 2009.

[2] H. Calkins *et al.*, "2017 HRS/EHRA/ECAS/APHRS/SOLAECE expert consensus statement on catheter and surgical ablation of atrial fibrillation," *Heart Rhythm*, vol. 14, no. 10, pp. e275–e444, 2017.

[3] S. Colilla, A. Grow, W. Petkun, D. E. Singer, T. Simon, and X. Liu, "Estimates of current and future incidence and prevalence of atrial fibrillation in the U.S. adult population," *Amer. J. Cardiol.*, vol. 112, no. 8, pp. 1142–1147, 2013.

[4] B. P. Krijthe, A. Kunst, E. J. Benjamin, G. Y. H. Lip, O. H. Franco, A. Hofman, J. C. M. Witteman, B. H. Stricker, and J. Heeringa, "Projections on the number of individuals with atrial fibrillation in the European Union, from 2000 to 2060," *Eur. Heart J.*, vol. 34, no. 35, pp. 2746–2751, 2013.

[5] R. G. Hart and J. L. Halperin, "Atrial fibrillation and stroke: Concepts and controversies," *Stroke*, vol. 32, no. 3, pp. 803–808, 2001.

[6] E. J. Benjamin, P. A. Wolf, R. B. D'Agostino, H. Silbershatz, W. B. Kannel, and D. Levy, "Impact of atrial fibrillation on the risk of death: The framingham heart study," *Circulation*, vol. 98, no. 10, pp. 946–952, Sep. 1998.

[7] C. W. Israel, G. Grönefeld, J. R. Ehrlich, Y.-G. Li, and S. H. Hohnloser, "Long-term risk of recurrent atrial fibrillation as documented by an implantable monitoring device: Implications for optimal patient care," *J. Amer. College Cardiol.*, vol. 43, no. 1, pp. 47–52, 2004.

[8] K. Harris, D. Edwards, and J. Mant, "How can we best detect atrial fibrillation?" *J. Roy. College Phys. Edinburgh*, vol. 42, pp. 5–22, Jan. 2012.

[9] S. Petrutiu, J. Ng, G. M. Nijm, H. Al-Angari, S. Swiryn, and A. V. Sahakian, "Atrial fibrillation and waveform characterization," *IEEE Eng. Med. Biol. Mag.*, vol. 25, no. 25, pp. 24–30, Nov./Dec. 2006.

[10] P.-S. Yang, S. Ryu, D. Kim, E. Jang, H. T. Yu, T.-H. Kim, J. Hwang, B. Joung, and G. Y. H. Lip, "Variations of prevalence and incidence of atrial fibrillation and oral anticoagulation rate according to different analysis approaches," *Sci. Rep.*, vol. 8, May 2018, Art. no. 6856.

[11] T. Hennig, P. Maass, J. Hayano, and S. Heinrichs, "Exponential distribution of long heart beat intervals during atrial fibrillation and their relevance for white noise behaviour in power spectrum," *J. Biol. Phys.*, vol. 32, no. 5, pp. 383–392, 2006.

[12] N. Larburu, T. Lopetegi, and I. Romero, "Comparative study of algorithms for atrial fibrillation detection," in *Proc. Comput. Cardiol.*, Sep. 2011, pp. 265–268.

[13] S. A. Guidera and J. S. Steinberg, "The signal-averaged P wave duration: A rapid and noninvasive marker of risk of atrial fibrillation," *J. Amer. College Cardiol.*, vol. 21, no. 7, pp. 1645–1651, 1993.

[14] S. Ladavich and B. Ghoraani, "Developing an atrial activity-based algorithm for detection of atrial fibrillation," in *Proc. 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2014, pp. 54–57.

[15] S. Mehta, N. Lingayat, and S. Sanghvi, "Detection and delineation of P and T waves in 12-lead electrocardiograms," *Expert Syst.*, vol. 26, no. 1, pp. 125–143, 2009.

[16] K. Tateno and L. Glass, "A method for detection of atrial fibrillation using RR intervals," in *Proc. Comput. Cardiol.*, Sep. 2000, pp. 391–394.

[17] N. Nuryani, M. Solikhah, A. S. Nugoho, A. Afdala, and E. Anzihory, "RR-interval variance of electrocardiogram for atrial fibrillation detection," *J. Phys., Conf. Ser.*, vol. 776, no. 1, 2016, Art. no. 012105.

[18] J. Lian, L. Wang, and D. Muessig, "A simple method to detect atrial fibrillation using RR intervals," *Amer. J. Cardiol.*, vol. 107, pp. 1494–1497, Feb. 2011.

[19] A. Kennedy, D. D. Finlay, D. Guldenring, R. R. Bond, K. Moran, and J. McLaughlin, "Automated detection of atrial fibrillation using R-R intervals and multivariate-based classification," *J. Electrocardiol.*, vol. 49, no. 6, pp. 871–876, 2016.

[20] J. Lee, D. D. McManus, P. Bourrel, L. Sörnmo, and K. H. Chon, "Atrial flutter and atrial tachycardia detection using Bayesian approach with high resolution time–frequency spectrum from ECG recordings," *Biomed. Signal Process. Control*, vol. 8, no. 6, pp. 992–999, 2013.

[21] M. Stridh, L. Sornmo, C. J. Meurling, and S. B. Olsson, "Sequential characterization of atrial tachyarrhythmias based on ECG time-frequency analysis," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 1, pp. 100–114, Jan. 2004.

[22] E. K. Roonizi and R. Sassi, "An extended Bayesian framework for atrial and ventricular activity separation in atrial fibrillation," *IEEE J. Biomed. Health Inform.*, vol. 21, no. 6, pp. 1573–1580, Nov. 2017.

[23] S. Dash, K. H. Chon, S. Lu, and E. A. Raeder, "Automatic real time detection of atrial fibrillation," *Ann. Biomed. Eng.*, vol. 37, no. 9, pp. 1701–1709, Sep. 2009.

[24] D. T. Linker, "Accurate, automated detection of atrial fibrillation in ambulatory recordings," *Cardiovascular Eng. Technol.*, vol. 7, no. 2, pp. 182–189, 2016.

[25] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.

[26] R. Kamaleswaran, R. Mahajan, and O. Akbilgic, "A robust deep convolutional neural network for the classification of abnormal cardiac rhythm using single lead electrocardiograms of variable length," *Physiol. Meas.*, vol. 39, no. 3, 2018, Art. no. 035006.

[27] U. R. Acharya, H. Fujita, S. L. Oh, Y. Hagiwara, J. H. Tan, and M. Adam, "Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network," *Inf. Sci.*, vol. 405, no. 1, pp. 81–90, Sep. 2017.

[28] Y. Xia, N. Wulan, K. Wang, and H. Zhang, "Detecting atrial fibrillation by deep convolutional neural networks," *Comput. Biol. Med.*, vol. 93, pp. 84–92, Feb. 2018.

[29] J. Rubin, S. Parvaneh, A. Rahman, R. Conroy, and S. Babaeizadeh, "Densely connected convolutional networks for detection of atrial fibrillation from short single-lead ECG recordings," *J. Electrocardiol.*, vol. 51, no. 6, pp. S18–S21, 2018.

[30] X. Fan, Q. Yao, Y. Cai, F. Miao, F. Sun, and Y. Li, "Multiscaled fusion of deep convolutional neural networks for screening atrial fibrillation from single lead short ECG recordings," *IEEE J. Biomed. Health Inform.*, vol. 22, no. 6, pp. 1744–1753, Nov. 2018.

[31] 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.*, vol. 44, Sep. 2017, pp. 1–4.

[32] Y. H. Awni, P. Rajpurkar, M. Haghpanahi, G. H. Tison, C. Bourn, M. P. Turakhia, and A. Y. Ng, "Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network," *Nature Med.*, vol. 25, pp. 65–69, Jan. 2019.

[33] A. 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.

[34] S. Oral and R. Latchamsetty, "Atrial fibrillation: Proxysmal, persistent, and permanent," in *Cardiac Electrophysiology: From Cell to Bedside*, D. P. Zipes and J. Jalife, Eds., 6th ed. Philadelphia, PA, USA: Saunders, 2013, pp. 739–754.

[35] D. Lai, X. Zhang, Y. Zhang, and M. Heyat, "Convolutional neural network based detection of atrial fibrillation combing R-R intervals and F-wave frequency spectrum," presented at the 41st Int. Conf. Eng. Med. Biol., Berlin, Germany, to be published.

[36] S. Asgari, A. Mehrnia, and M. Moussavi, "Automatic detection of atrial fibrillation using stationary wavelet transform and support vector machine," *Comput. Biol. Med.*, vol. 50, pp. 132–142, May 2005.

[37] D. Lai, X. Zhang, K. Ma, Z. Chen, W. Chen, H. Zhang, H. Yuan, and L. Ding, "Automated detection of high frequency oscillations in intracranial EEG using the combination of short-time energy and convolutional neural networks," *IEEE Access*, vol. 7, pp. 82501–82511, 2019.

[38] X. Wei, J. Li, C. Zhang, M. Liu, P. Xiong, X. Yuan, Y. Li, F. Lin, and X. Liu, "Atrial fibrillation detection by the combination of recurrence complex network and convolution neural network," *J. Probab. Statist.*, vol. 2019, Jan. 2019, Art. no. 8057820.

[39] X. Xu, S. Wei, C. Ma, K. Luo, L. Zhang, and C. Liu, "Atrial fibrillation beat identification using the combination of modified frequency slice wavelet transform and convolutional neural networks," *J. Healthcare Eng.*, vol. 2018, Jul. 2018, Art. no. 2102918.

[40] 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.

DAKUN LAI (M'11) received the Ph.D. degree in biomedical engineering from Fudan University, Shanghai, China, in 2008, and the Postdoctoral Fellowship in biomedical engineering from the University of Minnesota, Minneapolis, MN, USA.

Since 2012, he has been on the faculty of the University of Electronic Science and Technology of China (UESTC), China, where he was appointed as an Associate Professor in electrical engineering and biomedical engineering. He is currently the Director of the Biomedical Imaging and Electrophysiology Laboratory, UESTC. He is currently involved in the field of heart disease, neurological disorder, and sleep disorders. His current research interests include bioelectromagnetism, neuroengineering, and cardiac electrophysiology. He has pioneered the development of noninvasive cardiac electric source imaging, and made significant contributions to deep learning-based bioelectrical signal analysis, and detection and prediction of severe cardiac arrhythmias and neuro disorder, and numerical application of bio electromagnetism. He is an Editor and a Reviewer of several international journals.

XINSHU ZHANG received the bachelor's degree in electronic engineering from the University of Electronic Science and Technology of China (UESTC), in 2016, where he is currently pursuing the master's degree with the School of Electronic Science and Engineering. His current research interests include biomedical signal processing, deep learning technology, and atrial fibrillation detection.

YUXIANG BU received the bachelor's degree in electronic information engineering from the Hubei University of Technology (HBUT), Wuhan, China, in 2018. He is currently pursuing the master's degree with the School of Electronic Science and Engineering, University of Electronic Science and Technology of China (UESTC). His current research interests include ECG monitoring systems and atrial fibrillation detection and prediction using deep learning technology.

YE SU received the Ph.D. degree in biomedical engineering from the University of Electronic Science and Technology. She is currently an attending Physician of cardiovascular ultrasound and cardiology with the Sichuan Academy of Medical Sciences and Sichuan Provincial People's Hospital. Her current research interests include the analysis and diagnosis of stress electrocardiogram, electrocardiogram, dynamic electrocardiogram, and dynamic blood pressure. In recent years, she mainly participated in scientific research with the Science and Technology Department of Sichuan Province and the Health Department.

CHANG-SHENG MA received the M.D. degree from the School of Medicine, Wuhan University, Wuhan, China, in 2004, and then completed his Resident and Chief-resident at the Cardiology Department, Beijing Anzhen Hospital, and Peking University Hospital, Beijing, China, in 1993.

Since 2001, he has been on the faculty with Capital Medical University, and the Director of the Cardiology Department, Beijing Anzhen Hospital, where he was appointed as a Professor of medicine. He serves or served as an Associate Editor or Editorial Board Member of more than 30 international journals at home and abroad, such as *Circulation, EUROPACE, Journal of Cardiovascular Electrophysiology, Journal of Interventional Cardiac Electrophysiology*, and *Chinese Journal of Interventional Cardiology*. His research has been recognized by a number of awards. He was a recipient of the Ministry of Health Outstanding Contributions to the Young and Middle-aged Experts, the Science and Technology Beijing 100 Leading Talents, the Beijing High-Level Health Technology Leaders, and the Wu Jieping-Paul Yang Sen Medical and Pharmacy Award. He also served as the Chairman of Great Wall International Congress of Cardiology, the President of the Chinese College of Cardiovascular Physicians, and the Vice-President of the Chinese Society of Cardiac Pacing and Electrophysiology. He is engaged in the prevention and treatment of cardiovascular diseases and clinical effectiveness research, good at atrial fibrillation, and complex arrhythmia catheter ablation. He is the first person who carried out Radiofrequency catheter ablation of atrial fibrillation in China. He has assisted more than 200 hospitals in China to carry out catheter ablation of tachyarrhythmia and trained a large number of professionals in interventional therapy. He published 730 articles and received the Second Prize of National Scientific and Technological Progress for three times. His books, such as *Interventional Cardiology*, *Radiofrequency Ablation Map of Arrhythmia and Practice of Cardiology* have a wide influence in China.