Discrete Biorthogonal Wavelet Transform Based Convolutional Neural Network for Atrial Fibrillation Diagnosis from Electrocardiogram

Qingsong Xie[1,3] , **Shikui Tu**[2,3*] , **Guoxing Wang**[1,3] , **Yong Lian**[1,3] and **Lei Xu**[2,3†]
[1]Department of Micro-Nano Electronics, Shanghai Jiao Tong University, China
[2]Department of Computer Science and Engineering, Shanghai Jiao Tong University,China
[3]Centre for Cognitive Machines and Computational Health (CMaCH), School of SEIEE, Shanghai Jiao Tong University, China
{qingsongxie,tushikui,guoxing,eleliany,lxu}@sjtu.edu.cn

Abstract

For the problem of early detection of atrial fibrillation (AF) from electrocardiogram (ECG), it is difficult to capture subject-invariant discriminative features from ECG signals, due to the high variation in ECG morphology across subjects and the noise in ECG. In this paper, we propose an Discrete Biorthogonal Wavelet Transform (DBWT) Based Convolutional Neural Network (CNN) for AF detection, shortly called DBWT-AFNet. In DBWT-AFNet, rather than directly feeding ECG into CNN, DBWT is used to separate sub-signals in frequency band of heart beat from ECG, whose output is fed into CNN for AF diagnosis. Such sub-signals are better than the raw ECG for subject-invariant CNN representation learning because noisy information irrelevant to human beat has been largely filtered out. To strengthen the generalization ability of CNN to discover subject-invariant pattern in ECG, skip connection is exploited to propagate information well in neural network and channel attention is designed to adaptively highlight informative channel-wise features. Experiments show that DBWT-AFNet outperforms the state-of-the-art methods, especially for classifying ECG segments across different subjects, where no data from testing subjects have been used in training.

1 Introduction

Atrial fibrillation (AF) is one of the most common sustained arrhythmia observed in the clinical practice. The risk of the disease increases as age grows, and it influences approximately 0.4% of adult population [Simona *et al.*, 2006]. Disordered activation and irregular atrial contraction are the main causes of AF, usually accompanied by symptoms related to a rapid heart rate. The disease is associated with an increased risk of heart failure, dementia, stroke, coronary artery disease, cardiomyopathy, and congenital heart disease [Munger *et al.*, 2014]. Therefore, it is essential to develop a Computer-Aided Diagnosis system to detect AF at any time for early treatment even outside hospital. Long-term and continuous 12-lead electrocardiogram (ECG) signals are non-invasive and important tools to examine AF. However, it is cumbersome to wear 12-lead device and so inconvenient since long-term monitoring would lower the quality of daily life. An alternative is to exploit wearable devices which usually acquire one-lead dynamic ECG recording. It has advantages of low cost, ease of operation, and comfortable experience to users. However, high noises are particularly prevalent in wearable devices, and the morphology of such dynamic ECG shows high variations among different persons. As a result, it is of critical importance to find an effective way to detect AF, not only invariant to cross-subject differences but also robust against ECG noise.

Generally, ECG signal contains P, Q, R, S, and T waves for normal sinus mode (NSR) person while the ECG from AF patient is absent from P wave and shows irregular variability of R-R intervals. Therefore, most of existing methods for AF detection from ECG were divided into two categories: (1) absence of P wave; (2) the irregularity of R-R intervals. The algorithms [Guidera and Steinberg, 1993; Ladavich and Ghoraani, 2015] relied solely on the absence of P wave and thus were limited in real applications. It is because accurately detecting the fiducial position of a small P wave is difficult, particularly in the presence of noise and baseline drifting[Larburu *et al.*, 2011].

Compared to the first category, more algorithms utilized the irregularity of R-R intervals for automatic detection of AF. Turning points ratio and root mean square of consecutive RR differences, and Shannon entropy were used to detect AF [Dash *et al.*, 2009; Huang *et al.*, 2011; Lee *et al.*, 2012]. Combination of convolutional neural network (CNN) and recurrent neural network (RNN) was used to detect AF with a series of RR intervals [Andersen *et al.*, 2019; Dang *et al.*, 2019] as input. CNN by using RR intervals and F-wave frequency spectrum was also presented for AF detection [Lai *et al.*, 2019]. However, the performance would degrade when R peaks are detected by mistake in the case of noisy signals [Asgari *et al.*, 2015].

Several methods focused on manual feature extractors and feature selection to build a set of features relevant to AF [Martis *et al.*, 2013; Asgari *et al.*, 2015; Minggang *et al.*, 2018;

*Contact Author
†Contact Author

Teijeiro *et al.*, 2018; Rizwan *et al.*, 2018]. For example, higher order statistics [Martis *et al.*, 2013], peak-to-average power ratio and log-energy entropy [Asgari *et al.*, 2015], morphology information from P, Q, R, S, T waves and frequency-based features [Minggang *et al.*, 2018; Teijeiro *et al.*, 2018; Rizwan *et al.*, 2018]. AF was then diagnosed by passing these features into conventional machine learning classifiers, e.g., k-nearest neighbor and support vector machine [Asgari *et al.*, 2015], decision tree ensemble [Minggang *et al.*, 2018; Rizwan *et al.*, 2018] and XGBoost [Teijeiro *et al.*, 2018]. However, the fixed hand-crafted parameters may not be optimal for unknown signals.

Recently, deep learning was used for an end-to-end AF diagnosis by CNN [Xia *et al.*, 2018; Acharya *et al.*, 2017; Fujita and Cimr, 2019; Cao *et al.*, 2019], or by combining CNN and RNN [Zihlmann *et al.*, 2017; Warrick and Homsi, 2018; Xiong *et al.*, 2018], requiring neither wave detection nor hand-crafted feature extraction. However, they mainly deal with intra-subject scenarios, where the training and testing data contain ECG segments from the same subjects.

AF detection is based on the variations of heart beat. The heart beats are in a narrow frequency band (FB), but ECG has much wider FB. Thus, to facilitate CNN learning AF-relevant patterns from ECG, we are motivated to select most relevant FB in ECG to AF as input of CNN for accurate ECG classification. Inspired by [Xu, 2018; Xie *et al.*, 2019] which suggests turning time series into high-dimensional space by feature enrichment (FE), we decompose ECG signal into multiple sub-signals with different sub-FBs by Discrete Biorthogonal Wavelet Transform (DBWT), leading to discriminant input for better CNN representation learning.

To detect AF in cross-subject scenario in an end-to-end way, we propose an efficient DBWT Based Convolutional Neural Network (DBWT-AFNet) to discover subject-invariant features from AF ECG signals. Rather than directly feeding ECG into CNN, we segregate sub-signals with sub-FBs within FB of heart beat from ECG. Such sub-signals are propagated up to CNN for AF diagnosis, which reduces the difficulty of CNN to learn subject-invariant ECG representation as it removes the interaction with various features from irrelevant FB. To enhance the learning ability of CNN to capture the underlying features in ECG, CNN is designed to be guided by skip connection and channel attention (CA). Specifically, the identity skip connection is exploited which propagates information well in neural networks to avoid gradient vanishing. To select the relevant feature maps to AF, channel attention is explored to adaptively search the informative features across feature maps for AF detection. Experimental results demonstrate that our method outperforms state-of-the-art works in not only patient-specific modeling but also cross-patient setting. The proposed method provides a more accurate way to diagnose AF in real applications where no data have been collected from the new patients to train the model.

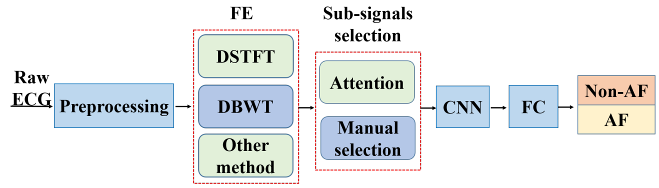

Figure 1: The overview of the proposed method.

2 Methodology

We aim at detecting AF from sing-lead ECG in cross-subject scenario. The challenge is to capture AF discriminant patterns from ECG signals, robust against noise and high variation in the same person or across different persons. Figure 1 shows the overview of the proposed method for AF detection. Raw ECG is first preprocessed to remove out-of-band frequencies. To generate discriminative input for CNN better capturing patient-invariant patterns in ECG, the preprocessed ECG can be transformed into high-dimensional space through FE [Xu, 2018], e.g., multiple sub-signals by DBWT, or 2-D image by discrete Short-Time Fourier Transform (DSTFT) as in [Xie *et al.*, 2019]. In this paper, we adopt DBWT considering the diverse contributions of different FBs to AF diagnosis. The most AF-relevant sub-signals selected according to prior knowledge are fed into our CNN to learn discriminative features. There are also other possible feature enrichment methods and sub-signals selection methods, e.g., adaptively selected through attention mechanism, and we leave them for future work. Finally, one fully-connected layer (FC) with softmax activation takes the learned features as input to distinguish between non-AF and AF.

2.1 Preprocessing

ECG signal is contaminated by various kinds of noises in real applications, e.g., respiration signal and muscle contraction. Since effective frequencies of ECG mainly lie in 1-20 Hz, the 4th order Butterworth bandpass filter is constructed for removing noises with cutoff frequencies of 1 and 20 Hz.

2.2 Discrete Biorthogonal Wavelet Transform

To investigate the contribution of different frequency bands to AF diagnosis, we exploit Discrete Biorthogonal Wavelet Transform (DBWT) suitable for non-stationary signal analysis [Averbuch *et al.*, 2014] to decompose ECG into multiple sub-signals with each sub-signal encompassing different frequency bands. DBWT is constructed by two scaling functions to generate different multi-resolution analyses as:

$$\mathrm{DBWT}(s(n)) = v_J(n) + \sum_{j=1}^{J} w_j(n), \tag{1}$$

where $s(n)$ is input signal, $v_J(n)$ and $w_j(n)$ denote different sub-signals with different frequency bands. $v_j(n)$ and $w_j(n)$ are calculated with:

$$v_j(n) = \sum_k 2^{-j/2} c_k^j \varphi(2^{-j}n - k) \tag{2}$$

$$w_j(n) = \sum_k 2^{-j/2} d_k^j \psi(2^{-j}n - k). \tag{3}$$

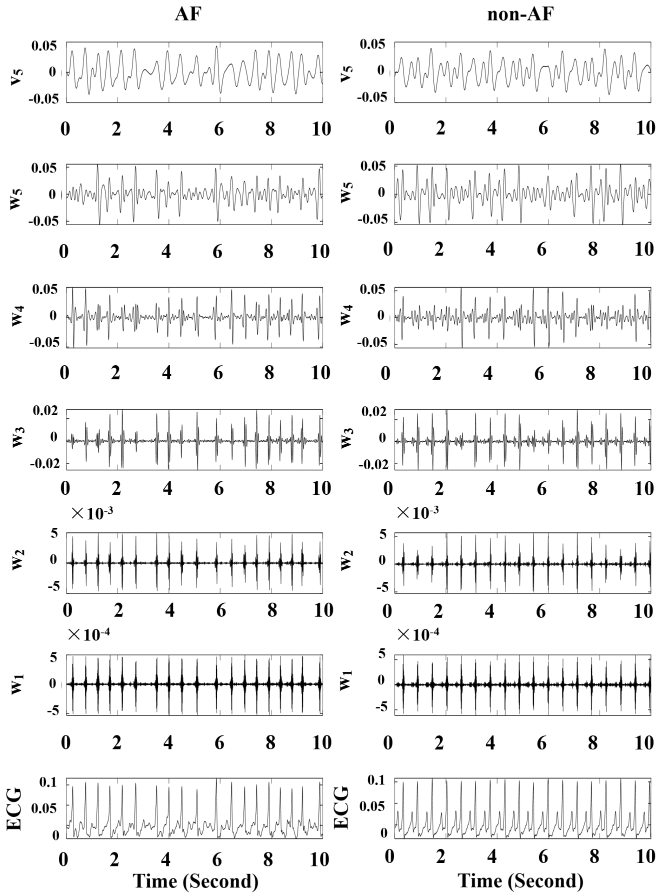

Figure 2: Plots of representative sub-signals of AF and non-AF decomposed by DBWT.

c_k^j, d_k^j, are recursively computed as:

$$c_0 = s \tag{4}$$

$$c_k^{j+1} = \sum_l h(l-2k)c_l^j, \tag{5}$$

$$d_k^{j+1} = \sum_l g(l-2k)c_l^j. \tag{6}$$

$\varphi(n)$, $\psi(n)$, $h(n)$, and $g(n)$ are scaling function, mother function, loss-pass filter coefficient, and high-pass filter coefficient, respectively, corresponding to 'bior 3.5'.

Through DBWT, the preprocessed ECG are decomposed into $J+1$ sub-signals. Concretely, we divide ECG into six sub-signals ($w_1, w_2, w_3, w_4, w_5, v_5$). Since the main frequencies of the preprocessed ECG are below 20 Hz, the sub-FBs of w_j are $[20/2^j, 20/2^{j-1}]$ Hz and v_5 are below 0.625 Hz. FB of heart beat contains w_4 and w_3, so we combine these two sub-signals by simply applying point-wise addition. Figure 2 illustrates the sub-signals of AF and non-AF by DBWT.

2.3 AFNet Network

To distinguish between AF and non-AF, CNN is constructed to learn distinct AF relevant features from sub-signals. The details of the CNN structure in AFNet network are given in the left of Figure 3. In this study, the convolutional layer with filter number of m, kernel size of k, and stride of 1 is denoted by Conv(m, k). AFNet is built based on Basic Blocks and each Basic Block contains 3-layer convolutional module with the same shape to reduce the free choice of hyper-parameters, where Basic Block(n) represents the filter number of all convolutional layers within it is n. In AFNet, all convolutional layers use Rectified linear unit (ReLU) to introduce nonlinearity and all the max pooling layers adopt pool size of 2 and stride of 2, denoted as Max pooling(2,2), to reduce computation complexity in the following processing. To improve CNN performance and help CNN capture patient-invariant features contained in sub-signals, a identity skip connection is employed to efficiently pass the information and to avoid gradient vanishing, forming Bottle Block, as suggested in Residual Network architecture [He *et al.*, 2016]. To further enhance the generalization ability of CNN for better AF diagnosis, one channel attention (CA) is added to Bottle Block to adaptively make emphasis on informative features, inspired by Squeeze-and-Excitation Networks [Hu *et al.*, 2018], yielding CA Block. Each CA Block includes one global averaging pooling, one ReLU, two convolutional layers with kernel size of 1, and one sigmoid function. The filter number of the first convolutional layer in CA is the filter number of Basic Block (n) divided by four while it is the same as n for the second convolutional layer. To reduce parameter number and capture the global information in every channel, global max pooling is used to generate channel-wise statistics (C_w). Note that the

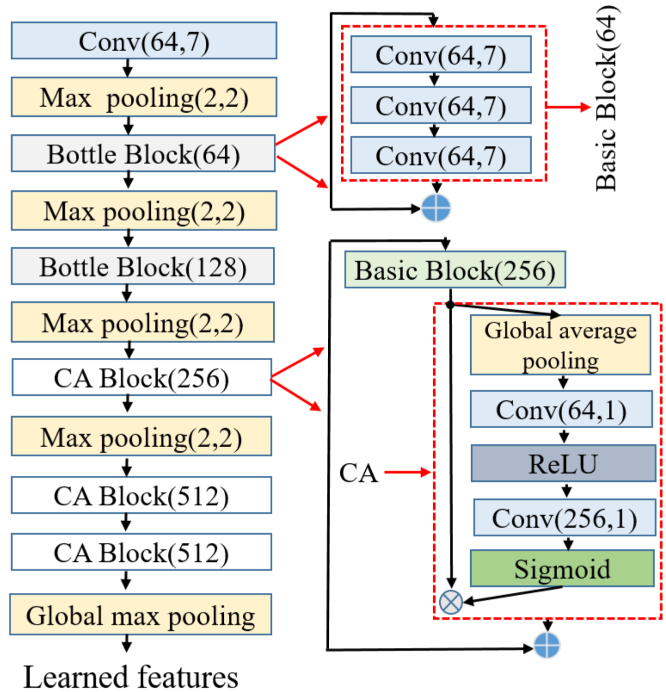

Figure 3: An overview of the proposed AFNet architecture. Shown in the left is the overall structure of the proposed CNN while shown in the upper right and bottom right are examples of Bottle Block(64) and CA Block(256). The Conv(m, k) denotes a convolutional layer with kernel number of m, kernel size of k, and stride of 1.

filter numbers of convolutional layer of Bottle Blocks and CA Blocks are given in the overall AFNet structure. For example, Bottle Block(64) and CA Block(256) represent the filter numbers as 64 and 256, respectively, for all three convolutional layers in Basic Block of Bottle Block and CA Block.

2.4 AF Diagnosis

To detect AF, channel-wise statistics $\boldsymbol{C}_w$ are fed into FC with softmax as activation to produce the probabilities $\boldsymbol{p}$ of the input ECG being non-AF (p_0) and AF (p_1) as:

$$\boldsymbol{p} = (p_0, p_1) = \text{softmax}(\boldsymbol{W}\boldsymbol{C}_w + \boldsymbol{b}) \tag{7}$$

where $\boldsymbol{W} \in \mathbb{R}^{2\times 512}$ and $b \in \mathbb{R}^2$ are parameters to learn.

2.5 Loss Function

We use cross-entropy between the ground truth and the predicted probability to calculate the loss for the total of M samples as:

$$Loss = -\frac{1}{M}\sum_{i=1}^{M} y_i \log p_1 + (1 - y_i)\log(p_0), \tag{8}$$

where $y_i \in (0, 1)$ is ground truth with $y_i = 0$ representing the corresponding ECG segment is non-AF and $y_i = 1$ denoting the ECG segment is AF.

3 Experimental Results

3.1 Dataset

All ECG data were collected from two databases in the publicly accessible PhysioNet [Goldberger *et al.*, 2000], MIT-BIH Atrial Fibrillation Database (AFDB) [Moody and Mark, 1983] and PhysioNet Challenge 2017[1].

AFDB provided 23 available two-lead ECG recordings from 23 subjects. Each recording has duration of approximately 10 hours with sampling frequency of 250 Hz. The annotation of start and ending time of AF for each recording was given by expert. In our study, only the lead-one ECGs were taken for AF diagnosis.

PhysioNet Challenge 2017 included 8528 single-lead short ECG recordings. The individual recordings were each 9-60 seconds in duration sampled by 300 Hz. All the recording were labelled into 5076 normal rhythm (N), 758 AF rhythm, 2415 other rhythm (O), and 279 noisy ($\sim$) recordings according to V3 version of labels.

All the ECG recordings were normalized into range [0,1]. Based on the two databases, three scenarios were constructed to test our method. We divided all ECGs in AFDB into fixed length of 10 seconds using sliding window without overlapping. A total of 84043 segments were obtained from AFDB. Our method was evaluated in both cross-subject and intra-subject scenarios based on AFDB as follows:

- *Cross-subject scenario*: For training the proposed network, the first 13 recordings were used, including 33824 non-AF segments and 13460 AF segments, while the remaining 10 recordings were used for the evaluation of the proposed classifier, including 16737 non-AF segments and 20022 AF segments. This process ensures that ECG data for training and testing sets do not contain recordings from the same subjects.
- *Intra-subject scenario*: To test our network in intra-subject scenario as other recent works [Asgari *et al.*, 2015; Xia *et al.*, 2018; Lai *et al.*, 2019], 5-fold cross-validation was used to evaluate the proposed method. In this procedure, all ECG segments were separated into 5 smaller datasets with almost equal ECG segments. The model was trained by 4 of 5 ECG segments while the remaining part (1 of 5) of the ECG segments was used to validate the performance of the proposed network. This approach was iterated 5 times by alternating the testing data. The performances were assessed based on testing data in each iteration. Finally, the overall performances of the proposed system were obtained by taking mean value of the performances acquired in all 5 iterations.

To compare with recent works [Cao *et al.*, 2019; Minggang *et al.*, 2018; Teijeiro *et al.*, 2018; Rizwan *et al.*, 2018; Zihlmann *et al.*, 2017; Warrick and Homsi, 2018; Xiong *et al.*, 2018; Teijeiro *et al.*, 2018] which tested their performance on PhysioNet Challenge 2017, we also assessed our method on this database. To make fair comparison with other works, we adopted the same procedure in [Cao *et al.*, 2019] to generate ECG segments. In this procedure, all ECGs were sliced into 9 seconds. Only one short segment is intercepted from the middle of each normal recording. For other rhythm those last less than 20 seconds, we intercepted a short segment from the middle while two samples were randomly intercepted without overlapping for those longer than 20 seconds. For AF rhythm and Noisy rhythm with fewer samples, we sliced the ECG with overlapping of 6 seconds and 8 seconds, respectively. Six-fold cross-validation was then used to evaluate the proposed method as [Cao *et al.*, 2019].

3.2 Training

We applied stochastic gradient descent to minimize cross-entropy loss function. Adam algorithm [Kingma and Ba, 2015] was used with a learning rate (lr) of 0.00002, and total epoch of 3 in cross-subject scenario. In intra-subject scenario and PhysioNet Challenge 2017 dataset, lr decayed as $0.0002/(1+2\times epoch)$ until epoch increased to 10. The proposed network was implemented on Nvidia Titan Xp GPU with Tensorflow framework and trained with batch size of 64.

3.3 Baseline Models

To make ablation analysis of the proposed DBWT-AFNet network in Figure 1, we implement several baseline models to evaluate the effectiveness of each step as follows:

- **AFNet**: AFNet straightly applies AFNet to the preprocessed ECG for AF detection. It is used to evaluate the effectiveness of DBWT which separates sub-signals with sub-FBs within FB of heart beat from ECG.
- **$\boldsymbol{w}_i$-AFNet** and **$\boldsymbol{v}_i$-AFNet**: w_i-AFNet and v_i-AFNet classify ECG through only feeding the sub-signal w_i or v_i to AFNet. These models are used to investigate the contribution of different sub-FBs to AF diagnosis.

[1]https://www.physionet.org/challenge/2017/

Method	Acc(%)	Sen(%)	Spe(%)	Ppr(%)	F1
[Acharya *et al.*, 2017]	50.22	20.14	86.19	63.58	0.306
[Fujita and Cimr, 2019]	51.47	16.90	92.83	73.82	0.273
[Fan *et al.*, 2018]	85.65	82.51	89.41	90.48	0.862
AFNet−CA	89.34	85.41	94.05	94.49	0.898
AFNet	90.51	90.64	90.35	91.89	0.912
DBWT-AFNet−CA	92.50	88.82	**96.90**	**97.21**	0.927
DBWT-AFNet	**95.81**	**95.69**	95.96	96.59	**0.962**

Table 1: Comparative results with state-of-the-art algorithms under cross-subject scenario.

Method	Acc(%)	Sen(%)	Spe(%)	Ppr(%)	F1
v_5-AFNet	59.73	29.75	95.60	89.04	0.445
w_5-AFNet	69.88	48.84	95.04	92.28	0.638
w_4-AFNet	**92.46**	**89.14**	96.43	96.76	**0.928**
w_3-AFNet	88.49	80.70	**97.83**	**97.81**	0.884
w_2-AFNet	82.39	72.86	93.80	93.39	0.818
w_1-AFNet	86.43	77.44	97.19	97.09	0.860

Table 2: Comparative results for different FBs under cross-subject scenario.

- **DBWT-AFNet−CA** and **AFNet−CA**: DBWT-AFNet−CA and AFNet−CA merely remove channel attention in DBWT-AFNet and AFNet, respectively. These two baselines are applied to prove the effectiveness of CA which automatically highlight useful channel-wise features.

3.4 Performance Metrics

Five statistical metrics are adopted to evaluate the effectiveness of the proposed classifier, e.g., accuracy (Acc), sensitivity (Sen), specificity (Spe), positive predictivity (Ppr), and F1 score. F1 score is the harmonic average of Ppr and Sen, which is widely used in the field of information retrieval and very useful to evaluate classifier performances in the case of class imbalance.

3.5 Results in Cross-Subject Scenario

We train all the methods starting from a random initialization on the parameters. To reduce the influence of random initialization on the model performance, we run the training and testing five times, and the final performance is measured by the average value of all metrics for five times under the cross-subject scenario. Table 1 lists the performance of the proposed method as well as other related methods.

It can be observed from Table 1 that the proposed DBWT-AFNet achieves the best performance. Compared to the highest scores for each index among the existing methods [Acharya *et al.*, 2017; Fan *et al.*, 2018; Fujita and Cimr, 2019] for AF detection, DBWT-AFNet improves from 85.65% to 95.81% for Acc, from 82.51% to 95.69% for Sen, from 92.83% to 95.96% for Spe, and from 90.48% to 96.59% for Ppr. The scores for Sen are low by most methods, indicating that the false negatives are high. In terms of F1 score, which trades off between Sen and Ppr, DBWT-AFNet improves from 0.862 to 0.961. To summarize, DBWT-AFNet learns the patient-invariant discriminant patterns for AF detection, much better than the existing methods do.

Method	Acc(%)	Sen(%)	Spe(%)	Ppr(%)	F1
[Asgari *et al.*, 2015]	96.4	96.6	96.3	–	–
[Xia *et al.*, 2018]	98.29	98.34	98.24	–	–
[Lai *et al.*, 2019]	97.5	97.8	97.2	–	–
DBWT-AFNet	**99.10**	**99.20**	**99.04**	98.57	**0.989**

Table 3: Comparative results with state-of-the-art algorithms in intra-subject scenario.

We test several ablation models to quantify the contribution of the key components to DBWT-AFNet on the performance improvement. We can see from Table 1 that the performances of AFNet and AFNet−CA are improved by DBWT-AFNet and DBWT-AFNet−CA, respectively. These results demonstrate it is important to separate low frequencies in the FB of heart beat from ECG through DBWT for more accurate AF diagnosis. The probable reason is that using such sub-signals as input of CNN removes the irrelevant information in ECG, which makes CNN become easier to learn discriminative features for AF detection. DBWT-AFNet achieves improvement compared to DBWT-AFNet−CA and AFNet shows superior performance to AFNet−CA. They indicate the effectiveness of CA which assists CNN to discover subject-invariant ECG representation for more accurate AF detection through adaptively emphasizing useful information across feature maps. The accuracy is very low for [Acharya *et al.*, 2017; Fujita and Cimr, 2019] which use four and three convolutional layers, respectively, to learn ECG representation. It is because it may be hard for shallow CNN to capture non-subject-specific features from complex variations of ECG across different subjects. The performance is improved by [Fan *et al.*, 2018] applying two-stream CNNs with each CNN including 13-layer convolution to discover AF relevant features, which shows that deep CNN is more capable of finding underlying features in ECG. The performance is further surpassed by AFNet−CA. This is probably due to that skip connection makes information well propagated through deep neural networks and better captures patient-invariant patterns in ECG.

To study the influence of each sub-FB for AF diagnosis, we apply the proposed AFNet to every sub-signal and all the performances are listed in Table 2. Heart beats in the sub-FB of [1.25, 2.5] is most frequent corresponding to w_4 and sub-FB of [2.5, 5] corresponding to w_3 is secondarily happened. It can be viewed that w_4-AFNet exploiting sub-signal (w_4) as input presents the best performance while w_3-AFNet with w_3 as input obtains the second ranking. These results further indicate that the discriminant features between AF and non-AF may concentrate on the FB of heart beat. Both w_4 and w_3 contains relevant features to ECG types, and hence combing them to recognise ECG classes can improve accuracy. It is reflected as the improvements of DBWT-AFNet, compared to w_4-AFNet and w_3-AFNet.

3.6 Results in Intra-Subject Scenario

The intra-subject scenario is easier than the cross-subject scenario for the problem of AF detection, because subject-specific patterns can be utilized in the former case to improve the classification performance. Specifically, in intra-

Method	N	AF	Other	Mean
[Zihlmann *et al.*, 2017]	0.888	0.764	0.726	0.792
[Cao *et al.*, 2019]	0.881	**0.966**	0.851	0.899
[Minggang *et al.*, 2018]	0.93	0.88	0.82	0.87
[Teijeiro *et al.*, 2018]	**0.953**	0.838	0.850	0.880
[Rizwan *et al.*, 2018]	0.889	0.791	0.702	0.794
[Warrick and Homsi, 2018]	0.910	0.810	0.780	0.853
[Xiong *et al.*, 2018]	0.919	0.858	0.816	0.864
DBWT-AFNet	0.927	0.950	**0.924**	**0.934**
AFNet	0.794	0.913	0.682	0.796
DBWT-AFNet−CA	0.923	0.950	0.920	0.931
AFNet−CA	0.796	0.913	0.660	0.789

Table 4: Comparative F1 score with state-of-the-art algorithms in PhysioNet Challenge 2017.

subject scenario, although the ECG segments from the training and testing datasets are not the same, they can come from the same person. The models trained in intra-subject scenario can be generalized better to testing data, by reducing the variation of ECG across different persons via including representative ECG beats of testing subjects in the training set. Table 3 shows the results in intra-subject scenario. In this scenario, our method also surpasses the existing methods by [Asgari *et al.*, 2015; Xia *et al.*, 2018; Lai *et al.*, 2019]. Since the performance of all methods is already beyond 96%, the improvement by the proposed method is marginally small.

3.7 Results in PhysioNet Challenge 2017

Table 4 summaries the state-of-the-art published research results of ECG classification based on PhysioNet Challenge 2017 public dataset. As these works only reported the F1 scores with respect to normal, AF, other classes, and mean F1, we use the same indicators to make direct comparison. It can be observed from Table 4 that the proposed DBWT-AFNet and DBWT-AFNet−CA outperform other methods in terms of mean F1, which is improved from 0.899 by [Cao *et al.*, 2019] to 0.934 by DBWT-AFNet. The performance of DBWT-AFNet and DBWT-AFNet−CA are much superior to AFNet and AFNet−CA. These results further prove that the features of different classes are mainly located in heart beat FB. Merely feeding these sub-signals into CNN reduces the feature interference among different sub-FBs, and hence deceases the difficulty of CNN to learn distinguishable information among different types. Compared to DBWT-AFNet−CA and AFNet−CA, DBWT-AFNet and AFNet also achieve slight improvement, respectively, which demonstrate that channel attention indeed facilitates CNN to recognise different ECG types.

4 Conclusion

This paper presents a new method for learning ECG representation from single-lead ECG to diagnose AF, without any peak detection or hand-crafted features. For this purpose, we explore DBWT to separate relevant sub-signals to AF with sub-FBs in FB of heart beat. Then, the CNN gated by skip connection and channel attention mechanism further learns the discriminant features for AF detection with such sub-signals as input. The experimental results show that the proposed method achieves significant improvement in cross-subject AF detection, indicating that the proposed method has indeed learnt subject-invariant discriminant patterns. The proposed method can be used in real applications where no data from new patients has ever been collected as training data.

Acknowledgments

This work was supported by National Science and Technology Innovation 2030 Major Project (2018AAA0100700) of the Ministry of Science and Technology of China, and National Natural Science Foundation of China (61874171).

References

[Acharya *et al.*, 2017] U. Rajendra Acharya, Hamido Fujita, Oh Shu Lih, Yuki Hagiwara, Jen Hong Tan, and Muhammad Adam. Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network. *Information Sciences*, 405:81–90, 2017.

[Andersen *et al.*, 2019] Rasmus S. Andersen, Abdolrahman Peimankar, and Sadasivan Puthusserypady. A deep learning approach for real-time detection of atrial fibrillation. *Expert Systems with Applications*, 115:465–473, 2019.

[Asgari *et al.*, 2015] Shadnaz Asgari, Alireza Mehrnia, and Maryam Moussavi. Automatic detection of atrial fibrillation using stationary wavelet transform and support vector machine. *Computers in Biology & Medicine*, 60:132–142, 2015.

[Averbuch *et al.*, 2014] Amir Z Averbuch, Pekka Neittaanmäki, and Valery A Zheludev. *Spline and spline wavelet methods with applications to signal and image processing*. Springer, 2014.

[Cao *et al.*, 2019] Xin-Cheng Cao, Bin Yao, and Bin-Qiang Chen. Atrial fibrillation detection using an improved multi-scale decomposition enhanced residual convolutional neural network. *IEEE Access*, 7:89152–89161, 2019.

[Dang *et al.*, 2019] Hao Dang, Muyi Sun, Guanhong Zhang, Xingqun Qi, Xiaoguang Zhou, and Qing Chang. A novel deep arrhythmia-diagnosis network for atrial fibrillation classification using electrocardiogram signals. *IEEE Access*, 7:75577–75590, 2019.

[Dash *et al.*, 2009] S. Dash, K. H. Chon, S. Lu, and E. A. Raeder. Automatic real time detection of atrial fibrillation. *Annals of Biomedical Engineering*, 37(9):1701–1709, 2009.

[Fan *et al.*, 2018] Xiaomao Fan, Qihang Yao, Yunpeng Cai, Fen Miao, Fangmin Sun, and Ye Li. Multiscaled fusion of deep convolutional neural networks for screening atrial fibrillation from single lead short ECG recordings. *IEEE Journal of Biomedical and Health Informatics*, 22(6):1744–1753, 2018.

[Fujita and Cimr, 2019] Hamido Fujita and Dalibor Cimr. Computer aided detection for fibrillations and flutters using deep convolutional neural network. *Information Sciences*, 486:231–239, 2019.

[Goldberger *et al.*, 2000] AL Goldberger, LAN Amaral, L Glass, JM Hausdorff, PCh Ivanov, RG Mark, JE Mietus, GB Moody, C-K Peng, and HE Stanley. Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. *Circulation*, 101(23):E215–E220, 2000.

[Guidera and Steinberg, 1993] Steven A Guidera and Jonathan S Steinberg. The signal-averaged P wave duration: A rapid and noninvasive marker of risk of atrial fibrillation. *Journal of the American College of Cardiology*, 21(7):1645–1651, 1993.

[He *et al.*, 2016] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Identity mappings in deep residual networks. In *European Conference on Computer Vision*, pages 630–645, 2016.

[Hu *et al.*, 2018] Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 7132–7141, 2018.

[Huang *et al.*, 2011] Chao Huang, Shuming Ye, Hang Chen, Dingli Li, Fangtian He, and Yuewen Tu. A novel method for detection of the transition between atrial fibrillation and sinus rhythm. *IEEE Transactions on Biomedical Engineering*, 58(4):1113–1119, 2011.

[Kingma and Ba, 2015] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv:1412.6980*, 2015.

[Ladavich and Ghoraani, 2015] Steven Ladavich and Behnaz Ghoraani. Rate-independent detection of atrial fibrillation by statistical modeling of atrial activity. *Biomedical Signal Processing & Control*, 18:274–281, 2015.

[Lai *et al.*, 2019] Dakun Lai, Xinshu Zhang, Yuxiang Bu, Ye Su, and Chang-Sheng Ma. An automatic system for real-time identifying atrial fibrillation by using a lightweight convolutional neural network. *IEEE Access*, 7:130074–130084, 2019.

[Larburu *et al.*, 2011] N Larburu, T Lopetegi, and I Romero. Comparative study of algorithms for atrial fibrillation detection. In *Computing in Cardiology*, pages 265–268, 2011.

[Lee *et al.*, 2012] Jinseok Lee, Bersain A. Reyes, D McManus, David D., Oscar Maitas, and Ki H. Chon. Atrial fibrillation detection using an iphone 4s. *IEEE Transactions on Biomedical Engineering*, 60:203–206, 2012.

[Martis *et al.*, 2013] Roshan Joy Martis, U. Rajendra Acharya, Hari Prasad, Chua Kuang Chua, Choo Min Lim, and Jasjit S Suri. Application of higher order statistics for atrial arrhythmia classification. *Biomedical Signal Processing & Control*, 8(6):888–900, 2013.

[Minggang *et al.*, 2018] Shao Minggang, Bin Guangyu, Wu Shuicai, Bin Guanghong, Huang Jiao, and Zhou Zhuhuang. Detection of atrial fibrillation from ECG recordings using decision tree ensemble with multi-level features. *Physiological Measurement*, 39(9):094008, 2018.

[Moody and Mark, 1983] Geoge B. Moody and Roger G. Mark. A new method for detecting atrial fibrillation using R-R intervals. In *Computers in Cardiology*, pages 227–230, 1983.

[Munger *et al.*, 2014] T M Munger, L-Q Wu, and W K Shen. Atrial fibrillation. *Journal of Biomedical Research*, 28(1):1–17, 2014.

[Rizwan *et al.*, 2018] Muhammed Rizwan, Bradley M Whitaker, and David V Anderson. AF detection from ECG recordings using feature selection, sparse coding, and ensemble learning. *Phyisological Measurement*, 39(12):124007, 2018.

[Simona *et al.*, 2006] Petrutiu Simona, Ng Jason, M. Nijm Grace, Al-Angari Haitham, Swiryn Steven, and V. Sahakian Alan. Atrial fibrillation and waveform characterization. *IEEE Engineering in Medicine & Biology Magazine*, 25(6):24–30, 2006.

[Teijeiro *et al.*, 2018] Tomás Teijeiro, Constantino A García, Daniel Castro, and Paulo Félix. Abductive reasoning as a basis to reproduce expert criteria in ecg atrial fibrillation identification. *Physiological Measurement*, 39(8):084006, 2018.

[Warrick and Homsi, 2018] Philip A Warrick and Masun Nabhan Homsi. Ensembling convolutional and long short-term memory networks for electrocardiogram arrhythmia detection. *Physiological Measurement*, 39(11):114002, 2018.

[Xia *et al.*, 2018] Yong Xia, Naren Wulan, Kuanquan Wang, and Henggui Zhang. Detecting atrial fibrillation by deep convolutional neural networks. *Computers in Biology & Medicine*, 93:84–92, 2018.

[Xie *et al.*, 2019] Qingsong Xie, Shikui Tu, Guoxing Wang, Yong Lian, and Lei Xu. Feature enrichment based convolutional neural network for heartbeat classification from electrocardiogram. *IEEE Access*, 7:153751–153760, 2019.

[Xiong *et al.*, 2018] Zhaohan Xiong, Martyn P Nash, Elizabeth Cheng, Vadim V Fedorov, Martin K Stiles, and Jichao Zhao. Ecg signal classification for the detection of cardiac arrhythmias using a convolutional recurrent neural network. *Physiological Measurement*, 39(9):094006, 2018.

[Xu, 2018] Lei Xu. Deep bidirectional intelligence: AlphaZero, deep IA-search, deep IA-infer, and tpc causal learning. *Applied Informatics*, 5(1), 2018.

[Zihlmann *et al.*, 2017] Martin Zihlmann, Dmytro Perekrestenko, and Michael Tschannen. Convolutional recurrent neural networks for electrocardiogram classification. In *2017 Computing in Cardiology (CinC)*, pages 1–4. IEEE, 2017.