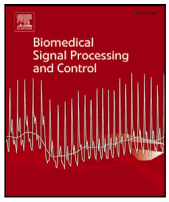
Biomedical Signal Processing and Control 79 (2023) 104067

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc



# Single-lead ECG recordings modeling for end-to-end recognition of atrial fibrillation with dual-path RNN

Mou Wang [a], Sylwan Rahardja [a], Pasi Fränti [a], Susanto Rahardja [b,c,*]

[a] School of Computing, University of Eastern Finland, Joensuu, Finland
[b] School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an, 710072, Shaanxi, China
[c] Infocomm Technology Cluster, Singapore Institute of Technology, 138683, Singapore

ARTICLE INFO

*Keywords:*
Atrial fibrillation
ECG
Deep learning
Recognition

ABSTRACT

Atrial fibrillation (AF) is the most common type of sustained cardiac arrhythmia, and is associated with stroke, coronary artery disease and mortality. Thus, early detection is crucial to avoid serious complications. Existing methods require specialized equipment and technical expertise, and accurate machine learning diagnosis of AF remains a dream. In this paper, we propose an end-to-end AF recognition method with dual-path recurrent neural network (DPRNN) from single-lead ECG. The model takes the whole ECG as input, and DPRNN splits the ECG into shorter segments and models the sequence between intra- and inter-segment iteratively. A mix-up operation is used for data augmentation, which overcomes the issue of limited data. We evaluated our method on the dataset from PhysioNet Challenge 2017. Experimental results shows that the proposed method can both effectively recognize AF with ECG signal without any human expertise, and outperforms state-of-the-art baseline methods. This demonstrates that dual-path model is effective for ECG analysis. We postulate that this framework can be generalized for other medical sequence signal.

## 1. Introduction

Cardiovascular disease is the leading cause of death globally according to the World Health Organization. Amongst cardiovascular pathology, Atrial Fibrillation (AF) is the most common type of sustained cardiac arrhythmia [1], characterized by rapid and irregular atrial activation with deterioration of atrial mechanical function [2]. Despite its prevalence, it is associated with thromboembolic complications such as stroke, hospitalization and mortality [3]. In most patients, AF progresses from relatively asymptomatic 'paroxysmal' form to a persistent arrhythmia.

With an aging of global population, projections estimate increasing prevalence of AF over the next few decades [4]. Although it is common, many patients have a completely asymptomatic course or occasional episodic symptoms of AF which can remain undetected [5], until complications occur. Due to its progressive nature, it is widely accepted that early AF detection and treatment may improve prognosis and prevent morbidity and mortality [6–8]. Hence, diagnosing AF early is important to tackling this increasing trend.

Detecting AF is commonly done through the observation and analysis of a 12 lead electrocardiogram (ECG) [9]. An ECG is a noninvasive test, captured at surface of the body, comprised of various electrical waveforms corresponding to the cardiac action potential. It is typically administered in a healthcare setting, and can provide valuable information such as rhythm and heart rate. However, ECG requires technical expertise to interpret accurately and reliably, and is thus manpower intensive. In addition, due to the silent nature of AF, it is unsurprising that patients do not seek medical attention until devastating complications occur.

With improving technology, hand-held and wearable ECG devices have become more common [10]. These devices provide a simple, convenient and non-invasive way of measurement of heart rate and rhythm using just a single lead as opposed to the 12 lead ECG. Although much less sophisticated than the ECG, it can potentially serve its purpose to highlight simple and common abnormalities. Despite this, such devices are still subjected to multiple limitations and are thus not widely adopted. ECG recordings are often limited by window of capture, and signals could be prone to artifacts due to inadvertent instrumental manipulation by wearers [11]. However, the technical issues pale against the benefit of longer monitoring period in an outpatient ambulatory setting. Single-lead ECG can be potentially and continuously monitored in real time at home or work. People can be reminded to go to hospital for further examination when AF is detected, instead of when there is

* Corresponding author.
*E-mail addresses:* wangmou21@mail.nwpu.edu.cn (M. Wang), sylwanrahardja@ieee.org (S. Rahardja), pasi.franti@uef.fi (P. Fränti), susantorahardja@ieee.org (S. Rahardja).

https://doi.org/10.1016/j.bspc.2022.104067
Received 26 March 2022; Received in revised form 29 July 2022; Accepted 8 August 2022

obvious and serious AF. Therefore, it is valuable to develop algorithms that can reliably detect AF and yet tackle its technical challenges by discounting artifacts.

Many automatic AF recognition algorithms have been developed to interpret ECG signals, including the traditional machine learning methods and deep learning methods. In traditional methods, the algorithms focus on extracting hand-crafted features manifested by AF from the ECG, such as heart rate based, morphology based, Wavelet based [12,13], graph based [14,15] and frequency domain features [16,17]. Morphology based indentification techniques capture abnormalities in P wave [18–20], R–R interval [20] and T wave [18,20], although *T* wave is less relevant in AF since AF has no *T* wave abnormalities. These features were fed into some well-known classifiers to recognize AF, such as Random Forests [16,18], linear discriminant analysis [12], support vector machine (SVM) [12–14,20,21] and AdaBoost [17].

Despite the above developments, hand-crafted devices still require expertise in the interpretation of ECG signals and cardiac arrhythmias. Moreover, these algorithms, which rely solely on those features, perform poorly because of limitations of the features. For example, noise and drifting of the signal baseline contaminating P-wave [22], and R–R intervals require long segments of data to identify long AF episodes [22]. Despite RR interval variations being a key feature of AF, this variation could be seen in normal patients with conditions such as sinus arrhythmia [23]. In addition, each feature reflects limited characteristics, and it was difficult, even impossible, to find the most appropriate features. These limitations significantly degrade performance of existing methods.

In recent years, deep learning had been used in a wide variety of fields as an effective tool for data analysis, such as image [24], audio [25], speech [26]. Many researchers have also applied deep learning on ECG signal classification [27–29], which had shown improvement in performance over the conventional machine learning methods. The application of machine learning in this field culminated in the Physionet Challenge 2017 [30]. Due to this challenge, researchers had an opportunity to create and benchmark their algorithms on relatively big datasets. The fruits of this challenge was a giant leap for the research on deep learning-based AF recognition with ECG.

Most of deep learning based methods employ a deep neural network to classify hand-crafted features, including convolutional neural network (CNN) and recurrent neural networks (RNN) [31]. For example, instant heart rate sequence [28], Short-term Fourier transform (STFT) [9,16,32,33] and stationary wavelet transform (SWT) [9] provided representations of ECGs in different domains before using CNN for classification. In addition, various other features had been considered. This include rhythm, morphological [34], and signal quality features [35,36], phase space and meta-level [16]. Akin to machine learning based methods, these algorithms still rely on aforementioned features, and are thus subjected to the similar performance limitations.

Currently, end-to-end architecture has been the trend for deep learning. These models perform predictions from raw data inputs owing to its capability of learning powerful features, especially for understanding highly complicated datasets without relevant expertise. In doing so, deep learning had brought data analyses into a whole new era. CNN [11,27,29,32,37,38], RNN [32,39,40] and Convolutional Recurrent Neural Network (CRNN) [41] had been utilized to conduct processing of raw ECG signals for ECG classification. These models predicted fixed-length segments of raw ECG data with final predictions being the classes of the highest votes on the respective segments.

Ensemble methods use either multiple features or learning algorithms to attain better performance in terms of accuracy of prediction. In Physionet challenge, participants had attempted to use ensemble method to achieve a good score. Some methods combined different type of features [42] or models [32,35] with an ensemble method, such as XGBoost [42], SVM. However, ensemble method significantly increased complexity. The increased complexity compromises its practicality in real world applications, especially for embedded systems. Therefore, it will be prudent to develop a one-step method to recognize the AF with a low complexity, rather than an ensemble system.

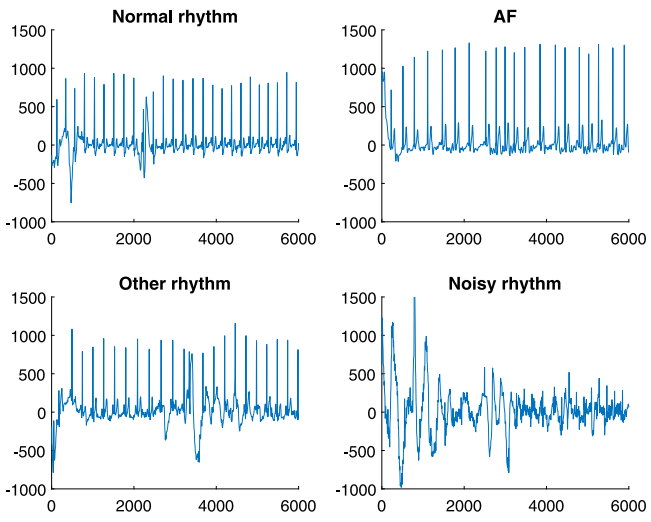



**Fig. 1.** 4 classes in the AF dataset: Normal rhythm (top left), AF (top right), Other rhythm(bottom right), Noisy rhythm (bottom right).

The ECG signal is intermittent and segmental, and there is a strong correlation between different period. However, model design of the aforementioned end-to-end methods did not consider this characteristic. In addition, data augmentation can overcome overfitting problem and improve the performance of model. However, the previous articles had not considered any data augmentation, even though the dataset of Physionet 2017 Challenge was limited.

In this paper, we propose a novel end-to-end AF automatic recognition method with dual-path recurrent neural network (DPRNN) from single-lead ECG signal. In contrast to end-to-end method in previous works, the proposed model takes the whole ECG with different length as input, and output the prediction without voting. The main contributions of this paper are summarized as follows:

- Considering periodic characteristics of ECG, we introduced DPRNN to split the ECG signal into multiple shorter segments and model the sequence between segments and within segments iteratively.
- Mix-up operation was used for data augmentation, which either mitigated or significantly reduced the problem of limited data. To the best of our knowledge, This article introduces the first attempt in combining dual-path model and mix-up operation in ECG processing.
- Different loss function were studied for AF recognition in this paper. We introduced large-margin softmax into AF recognition, which can further improve the performance.

## 2. The methods

The AF classification task aims to classify the ECG recordings into four classes: normal rhythm (*N*), AF rhythm (*A*), other rhythm (*O*) and noisy recordings (*R*) and Fig. 1 shows a typical wave depicting each class. A typical waveform for each class is shown in Fig. 1. We observed that the recordings of four classes have electrical artifacts, and the recordings of *N*, *A* and *O* were similar visually. This posed a challenge for accurate classification.

In our method, the raw ECG recordings were fed into the model which output the prediction. The model contains three components: encoder, DPRNN and classifier. First, an encoder used a linear one-dimensional (1D) convolution to transform the input ECG waveform into a learnable representation in an intermediate feature space. Then the data was fed to a dual-path RNN which models the sequence.

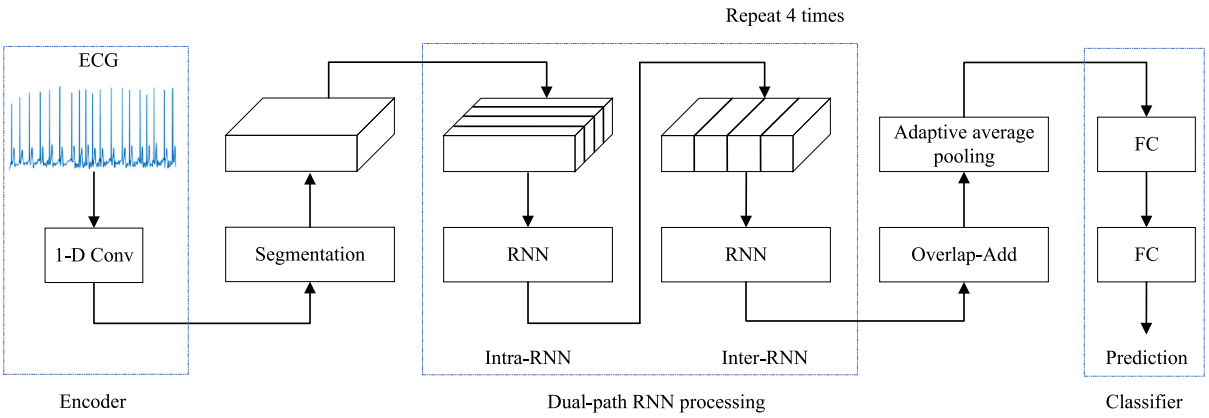



**Fig. 2.** Framework of end-to-end model for recognition of atrial fibrillation with dual-path RNN.

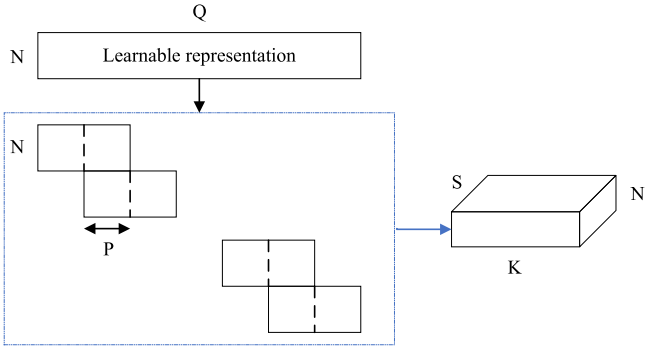



**Fig. 3.** The segmentation stage splits a sequential input into segments with overlaps of 50% and concatenates them to form a 3-D tensor.

Finally, the classifier, comprised of two fully-connected (FC) layers, computed the probability for each class to make a final decision. A block diagram of the proposed method is represented in Fig. 2 and is further elaborated in subsequent sections.

### 2.1. Encoder

In an end-to-end method, the first layer of deep neural network produces a learnable feature from the waveform. The layer is usually a convolutional layer, which is called an encoder. In contrast to hand-crafted features, the learnable feature is obtained by optimization of the network. It frees the developers from knowledge intensive capital requirements and hand-crafted features. Before using RNN to model a sequence, we need to use an encoder to transform the input sequence into a high-dimensional learnable feature.

We employed 1D convolutional layers as encoder because 1D convolution is essentially suitable to process time series data. The ECG recording is denoted by $x \in R^{1\times L}$. The recordings were then divided into overlapping vector $\mathbf{x} \in R^{L\times Q}$ of length $L$ samples, where $Q$ is the number of vectors. The encoder takes $\mathbf{x}$ as input and output the learnable representation $Z \in R^{M\times Q}$ as follows:

$$Z = W\mathbf{x}, \tag{1}$$

where the encoder can be characterized as filter-bank $W$ with $M$ filters of length $L$, and stride is half of the length. In this article, the number of filters $M$ is set to be 64, and $L$ is set to 8.

### 2.2. Dual-path RNN

RNN has its strengths and limitations. While it excels in sequence modeling, it performs poorly for longer sequences. Despite a single lead ECG being a long sequence, we can still divide it into segments, giving it features of shorter sequences. DPRNN can be applied to analysis of ECG sequence from both local (intra-segmental) as well as global (inter-segmental) aspects.

As shown in Fig. 2, DPRNN consists of 3 stages: segmentation, dual-path modeling, and overlap-add. First, a sequential representation from the encoder was split into a three-dimensional (3D) tensor segments in segmentation stage, and these segments was combined into a three-dimensional (3D) tensor. The tensor was then analyzed by stacking DPRNN blocks to systematically apply intra-segmental and inter-segmental modeling. Finally, using the overlap-add method, the output from the final DPRNN stage was constructed back into the original sequential form.

#### 2.2.1. Segmentation

As shown in Fig. 3, in segmentation stage, a sequential input such as the learnable representation $Z \in R^{M\times Q}$, was split into segments of length $K$ and hop size $P$, generating $S$ segments of equal size, where $P$ is usually set to a half of $K$, $T_s \in R^{M\times K}, s = 1,\dots,S$. Then, these segments were concatenated to form a 3D tensor $D = [T_1,\dots,T_S] \in R^{M\times K\times S}$.

#### 2.2.2. Dual-path modeling

The tensor $D$ from the segmentation was processed via the stacked $B$ DPRNN blocks. In every block, the input 3D tensor was transformed into a new tensor with similar shape. The input tensor of $b$th block is denoted as $D_b \in R^{M\times K\times S}$,where $b = 1,\dots,B$. Every block contains two sub-blocks corresponding to intra- and inter-segment processing. We first employ intra-segmental RNN to model the sequence with every $S$ segments, i.e, the second dimension of $D_b$. Considering the reverse sequence model is also important, bi-directional RNN is used in the intra-segment processing. The intra-segmental RNN can be represented by the following equation:

$$V_b = [f_b(D_b[:,:,i]), i = 1,\dots,S], \tag{2}$$

where $D_b[:,:,i] \in R^{M\times K}$ refers to the sequence of $i$th segment, $f_b(\cdot)$ is the intra-segmental RNN, and $V_b \in R^{H\times K\times S}$ is the output of the intra-segmental RNN.

To keep the length of sequence consistent, a linear FC layer is employed to convert $V_b$ into a new tensor $\hat{V}$, where the size of $\hat{V}$ and $D_b$ is the same. It can be expressed as

$$\hat{V}_b = [GV_b[:,:,i] + m, i = 1,\dots,S], \tag{3}$$

where $V_b[:,:,i] \in R^{H\times K}$ refers to sequence of $i$th segment in the output of the intra-segmental RNN, the weights and bias of the linear FC layer are denoted as $G \in R^{M\times H}$ and $m \in R^{M\times 1}$, respectively, and output of FC layer is $\hat{V}_b$.

Normalization has the potential of improving generalization of the deep learning model, and allow rapid and stable convergence during training. Instead of well-known batch normalization, Layer normalization (LN) has been a common normalization method [43]. Therefore, we applied LN to normalize the intermediate tensor $\hat{V}_b$, which can be expressed in the following equation:

$$LN(\hat{V}_b) = \frac{\hat{V}_b - \mu(\hat{V}_b)}{\sigma(\hat{V}_b) + \epsilon} \odot \omega + \kappa \tag{4}$$

where $\mu(\cdot)$ and $\sigma(\cdot)$ are the mean and standard deviation of the tensor in all dimensions, $\epsilon$ is a small positive constant for numerical stability, $\omega, \kappa \in R^{M\times 1}$ refer to the rescaling factors, and $\odot$ is the Hadamard product.

For a better back-propagation of the gradient, we added a residual connection between the input of intra-segmental RNN $D_b$ and the output of LN to produce a new tensor $\hat{D}_b$, i.e.,

$$\hat{D}_b = D_b + LN(\hat{V}_b). \tag{5}$$

Then, $\hat{D}_b$ produced by intra-segment modeling was fed into inter-segment RNN, which model the time sequence of every $k$ segments, i.e, the third dimension of $\hat{D}_b$. The inter-segment RNN can be expressed as:

$$U_b = [f'_b(\hat{D}_b[:, i, :]), i = 1, \ldots, K] \tag{6}$$

where $\hat{D}_b[:, i, :] \in R^{M\times S}$ refers to the sequence at the $i$th time point in all $S$ chunks, $f'_b(\cdot)$ refers to inter-segment RNN, and the output of inter-segment RNN is denoted as $U_b \in R^{H\times K\times S}$.

Similar to the intra-segmental RNN, we employed a linear FC layer to transform the $U_b$ into a new tensor with the same size as $\hat{D}_b$ and LN to normalize the tensor. We also added a residual connection between the input of inter-segmental RNN $\hat{D}_b$ and the output of LN to produce a new tensor, which is the output of whole dual-path modeling. With the ability of bi-directional intra-segmental RNN and inter-segmental RNN, DPRNN achieves superior sequence modeling.

To strike a balance between the intra- and inter-segmental RNN, the sequence length was expected to be similar during the training stage, i.e., $K \approx S$. The vector produced by encoder with length $Q$ can be split into $S$ segment of length $K$ and hop size of a half of $K$, then $S = \lceil 2Q/K \rceil + 1$, where $\lceil \cdot \rceil$ is the ceiling function. Therefore, $K \approx \sqrt{2Q}$.

RNN commonly includes standard RNN, Long Short-Term Memory (LSTM) and its variants, such Gate Recurrent Unit, Simple Recurrent Unit. In the proposed model, bidirectional LSTM was used as RNN with 64 hidden units in each direction. To improve the ability of sequence modeling, usually the intra- and inter-segment RNN needs to be stacked many times, and for our case we stacked four times and therefore $B$ is set to 4.

#### 2.2.3. Overlap-add

Assuming the output of stack DPRNN is $D_{B+1} \in R^{M\times K\times S}$, we applied overlap-add method to the $S$ segments to produce output $Z' \in R^{M\times Q}$. Therefore, overlap-add is an inverse operation of segmentation.

### 2.3. Classifier

After DPRNN, adaptive average pooling was used to adaptively reduce the size of $Z'$ and restrict the output to similar sizes even though the lengths of input were variable.

To produce a prediction, two FC layers were used, which transformed the outputs from DPRNN to a $4 \times 1$ vector having values that were essentially the outputs of each class ($N$, $A$, $O$ or $R$). Then, at the last stage of the classifier, a softmax function was applied to essentially adjust the values in accordance with the associated probability by normalization into the unity bound.

### 2.4. Loss function

In machine learning, the combination of Cross Entropy (CE) loss and Softmax has been widely used and is incorporated in the supervised portion of the system. Although this combination is simple and generally gives an excellence performance, it only supports limited discriminative learning of features. Large-margin softmax (LM) loss on the other hand explicitly creates compactness in intra-class and inter-class separability between learned features [44]. In addition, it has the benefit of adjusting desired margins while avoiding overfitting. Owing to these benefits, LM has brought the performance of classification and verification tasks to greater heights.

We denote the last FC layer as $W$, $i$th input as $x_i$ with the label $y_i$. Then, the output of FC layer can be written as $f_{y_i} = W_{y_i}^T x_i$, where $W_{y_i}$ is the $y_i$-th column of $W$. Because $f_j$ is the inner product between $W_j$ and $x_i$, it can be also formulated as $f_j = \|W_j\|\|x_i\|\cos(\theta_j)$ where $\theta_i (0 \le \theta_j \le \pi)$ is the angle between the vector $W_j$ and $x_i$.

The large-margin softmax loss can be defined as:

$$L = -\log\left(\frac{e^{\|W_{y_i}\|\|x_i\|\psi(\theta_{u_i})}}{e^{\|W_{y_i}\|\|x_i\|\psi(\theta_{u_i})} + \Sigma_{j\neq y_i} e^{\|W_j\|\|x_i\|\cos(\theta_j)}}\right), \tag{7}$$

in which $\psi(\theta)$ is constructed as:

$$\psi(\theta) = (-1)^k \cos(\alpha\theta) - 2k, \theta \in [\frac{k\pi}{\alpha}, \frac{(k+1)\pi}{\alpha}] \tag{8}$$

where $m$ is an integer that is closely related to the classification margin, and $k \in [0, \alpha - 1]$ is also an integer. With a larger value $\alpha$, the classification margin increases, resulting in increased resistance faced by the learning objectives. The detailed derivation had been explored in [44].

Based on naive large-margin softmax loss, a regularization was imposed on the logits to induce a large-margin classifier in a compatible form with the softmax loss [45]. This method provides a more effective and feasible loss than naive large-margin softmax loss. The large-margin loss in [45] can be expressed by:

$$L = -\log\frac{\exp(f_{y_i})}{\sum_{c=1}^{C}\exp(f_{y_c})} + \frac{\lambda}{2}\sum_{c\neq i}\left\{\frac{\exp(f_{y_c})}{\sum_{c'\neq i}\exp(f_{y_{c'}})} - \frac{1}{C-1}\right\}\log\left\{\frac{\exp(f_{y_c})}{\sum_{c'\neq i}\exp(f_{y_{c'}})}\right\} \tag{9}$$

where $C$ is the number of classes, and $\lambda$ is a regularization parameter, which is set to 0.3 in this paper. Similar to $m$, the learning objective will be harder with larger $\lambda$.

### 2.5. Data augmentation

Data augmentation is an effective way to improve generalization and prevent overfitting of the neural networks, especially when the data is insufficient. In this framework, mix-up was employed as the data augmentation method in the training stage [46]. Mix-up trains a neural network based on convex combinations from pairs of examples and associated labels. The mix-up training establishes a neural network that favors linear behavior between training examples. This limits presence of corrupt labels, increasing robustness of the examples and stabilizing the network training. Hence, it has been successfully applied to multiple fields such as image [46] and audio analytics [47].

In the analysis of two ECG samples $x_i$ and $x_j$ and the corresponding labels $y_i$ and $y_j$, the mix-up operations on the training set can produce the new sample $\hat{x}$ and its label $\hat{y}$. These equations can be expressed as:

$$\hat{x} = \lambda x_i + (1 - \lambda)x_j \tag{10}$$

and

$$\hat{y} = \lambda y_i + (1-\lambda) y_j \tag{11}$$

where $\lambda \in [0,1]$ is a random value drawn from the beta distribution. Then the loss with mix-up is defined as

$$L_{mix-up} = L(x_i, y_i) + (1-\lambda) L(x_j, y_j). \tag{12}$$

As the new samples in the mix-up operation were generated by linearly interpolating two real samples within a batch data, we can get more training samples without extra computational resources. The pseudocode of the proposed system is described in Algorithm 1.

**Algorithm 1** The proposed DPRNN-based method for AF recognition

**Input:** ECG Recording $x \in R^{1\times L}$
**Output:** Prediction of the ECG Recording
Initialize parameters of the model
**repeat**
  Get a mini-batch of ECG samples and label
  Mix-up operation on the batch according to Eqs. (10) and (11)
  Get the learnable representation $Z \in R^{M\times Q}$ through Encoder
  Get the 3D tensor $D \in R^{M\times K\times S}$ through segmentation
  DPRNN processing
  Get the new feature map $Z' \in R^{M\times Q}$ through Overlap-Add
  Get the prediction through Classifier
  Compute large-margin softmax loss
  Update the parameters of the model
**until** the validation loss had not descended after 10 consecutive epochs.

## 3. Experiments

### 3.1. Dataset

We evaluated the proposed method on PhysioNet 2017 dataset. This dataset contains 8,528 single-lead short-term ECG recordings lasting from 9 to 60 s with a sampling rate of 300 Hz [30]. In the dataset, all ECG recordings were collected from different patients using the AliveCor recording device. Next, the ECG recordings were manually labeled and classified by a team of expert into the four aforementioned classes, namely normal rhythm ($N$), AF rhythm ($A$), other rhythm ($O$) and noisy recordings ($R$). The sample sizes of $N$, $A$, $O$ and $R$ classes were 5,076, 758, 2,415 and 279, respectively.

The challenges of the dataset came from four main factors: relatively small-scale, imbalance of numbers in each class, unclear class definition, and variable lengths of ECG recordings. First, there were about 8,500 recordings, which was relatively small-scale for deep learning. Therefore, the common model was easily overfitted during the training. Moreover, the distribution of the different ECG classes was imbalanced. For example, about 60% of dataset is normal rhythm, while only 3.3% of dataset is noisy recordings. This introduced additional bias towards the class with more training data, thereby, increasing the complexity of the learning process for classification. Furthermore, the definition of class in this dataset is not entirely clear. The 'other rhythm' class contains many different type of recordings, which increased inconsistency within the class. Finally, the ECG signals in this dataset had lengths ranging from 2,714 to 18,260 data points. The variation of length further compounds the issue. For example, the data was fed into a deep learning model to train in mini-batch where same length of data is required in each batch. This created a challenge in determining correct classification based on a small signal segment.

To mitigate the challenges posed, data augmentation was considered to alleviate the problem of imbalance. As proposed in [11], we used different segment shifts during training stage, in which the recordings of different classes was split into several fixed-length segments with different shift length. The segment shifts $ss_i$ were manually determined based on the following relationship for $i$th class:

$$ss_i = \frac{l_i - l_s}{ns_i - 1} \tag{13}$$

where $ns_i$ is number of segments of $i$th class, $l_i$ is the total length of the recording of $i$th class and $l_s$ is segment length.

In the training dataset, the segment shift $ss_i$ of $i$th class was chosen to make $ns_i$ equal, which balanced the numbers of segments of the four classes. $l_s$ was fixed to 4,096, $ns_i$ was set to about 40,000. In the training data, we had four segment shifts being manually processed as $ss_{i=N,A,O,R} = [1142, 150, 560, 35]$.

During testing, the whole ECG signal without segmentation was fed into the model, which then output the prediction. The dataset was randomly split into 80% training set and 20% validation set for 5-fold cross validation. We selected the model with the best performance on the validation set for each fold. As for the test length of the original ECG, we recommend that the length of ECG data should be longer than 1s in order to recognize the AF effectively.

### 3.2. Experimental settings

The proposed model was implemented on NVIDIA GeForce 3090 GPU with PyTorch-lightning library. For all experiments, we trained the models for 25 epochs with AdamW optimizer. The learning rate was initialized to $1 \times 10^{-3}$, and exponentially decayed with a rate of 0.92 for each epoch. Early stopping was applied if the validation loss had not descended after 10 consecutive epochs.

### 3.3. Evaluation metrics

As the distribution was highly unbalanced, the F1-score was better suited to be the performance indicator, rather than as classification accuracy. Therefore, the F1-score was adopted as the main scoring function. We denote the recording of class $i$ predicted into class $j$ as $ij$, where $i \in [N, A, O, R]$ and $j \in [n, a, o, r]$. Note that capital and lower-case letters refer to the same class. The sum of recording of class $i$ is denoted as $\sum_i$, and the sum of recording predicted into class $j$ is denoted as $\sum_j$. Thus, the F1-scores of each class can be written as:

$$F_{1n} = \frac{2\times N}{\sum_N + \sum_n}, F_{1a} = \frac{2\times Aa}{\sum_A + \sum_a}, \tag{14}$$

and

$$F_{1o} = \frac{2\times Oo}{\sum_O + \sum_o}, F_{1r} = \frac{2\times Rr}{\sum_R + \sum_r}. \tag{15}$$

In PhysioNet 2017 challenge, the final evaluation rubrics focused on the class of $N$, $A$ and $O$. It is defined as

$$F1 = \frac{1}{3}\sum_{i\in n,a,o} F_{1i}. \tag{16}$$

In addition to F1-score, some common metrics are also used for evaluation, such as Accuracy, confusion matrix, Receiver Operating Characteristic (ROC) curve, and area under the ROC curve (AUC).

### 3.4. Comparison methods

As mentioned in Section 1, developing an end-to-end method was more practical than an ensemble system. However, most of the existing methods such as those in the submission of PhysioNet 2017 dataset utilized ensemble and therefore only the ensemble results were available for comparison. To evaluate the performance of this proposed framework, we compare the proposed method with the state-of-the-art methods from two aspects: a stand-alone end-to-end model without ensemble system, and an ensemble system having no standard definition

**Table 1**
The accuracy and detailed F1 score of cross validation on five folds. "Acc" refers to accuracy.

| Folds | Acc | $F_{1n}$ | $F_{1a}$ | $F_{1o}$ | $F_1$ |
|---|---|---|---|---|---|
| Fold 0 | 0.856 | 0.91 | 0.84 | 0.79 | 0.8452 |
| Fold 1 | 0.842 | 0.91 | 0.79 | 0.75 | 0.8174 |
| Fold 2 | 0.842 | 0.90 | 0.85 | 0.76 | 0.8346 |
| Fold 3 | 0.827 | 0.89 | 0.82 | 0.73 | 0.8126 |
| Fold 4 | 0.858 | 0.92 | 0.82 | 0.77 | 0.8356 |
| Average | 0.845 | 0.906 | 0.822 | 0.760 | 0.8291 |

in terms of the number of features and the number of methods being incorporated into the system.

Firstly, the state-of-the-art methods are CNN [11,37] and CRNN [41]. In [11], six 1D convolutional blocks were stacked for extraction of the deep feature from raw ECG data. Each convolutional block included a convolutional layer with kernel of 15, batch normalization layer, an activation function (ReLU), max-pooling with kernel of 4, and dropout with rate of 0.5. Then, three FC layers were used to output the prediction on segmentation level. Finally, the final prediction was decided by voting or SVM. These two methods are used as baseline systems, which are denoted as CNN_vote and Stacking SVM. In [37], 5 convolution blocks were used, where each block consisted of a convolutional layer, batch normalization and Relu. Then, a dense layer with softmax was used for prediction. In [41], 12 convolutional layers were used, followed by LSTM for sequence modeling. Lastly, a FC layer was used for classification. Another the state-of-the-art method for AF detection is multi-scale convolutional network [38]. However, because the dataset and category number in [38] are different with our paper, we are unable to compare directly with their published results.

For comparison with other ensemble systems, we also made an ensemble with the proposed model. Even though we know the possibility of introducing the hand-craft features of fusion with better machine learning methods, this study opted to fuse the output of the proposed end-to-end method with different settings through a very simple operation (averaging), and compare with several ensemble systems.

### 3.5. Results and discussions

The accuracy and detailed F1 scores of the proposed method for each class on five folds are shown in Table 1, and the corresponding precision, recall and AUC are shown in Table 2. The confusion matrix and ROC curve on fold 0 are shown in Figs. 4 and 5. We observed that F1 score for normal rhythm can achieve 0.9 because it has sufficient recordings with good consistency. Although AF rhythm has less data than other rhythm, the ECG recordings of AF have some characteristic with consistency, therefore F1 score for AF rhythm achieved 0.822. It proved that the proposed method can recognize the AF with good performance. It is reasonable that F1 score for other rhythm is less than that of other two classes because of the inconsistency within the class. The training and validation curves are shown in Fig. 6, which are the curve of large-margin softmax loss on the training set and the curve for F1 score on validation set. From the figure, we can observe that the proposed model is not overfitted.

Table 3 shows the performance comparison between different end-to-end methods. It is obvious that the proposed method outperformed most of state-of-the-art methods. Comparing to other methods, the main improvement of the proposed method came from better recognition for AF rhythm and other rhythm. However, most methods had a similar F1 score for normal rhythm because it was easily recognizable. Although [32] is also an end-to-end method based on RNN, our proposed method outperforms it by a large margin. It proved that the proposed dual-path is useful for modeling the sequence between intra- and inter-segment iteratively, and therefore is effective for analysis ECG signal.

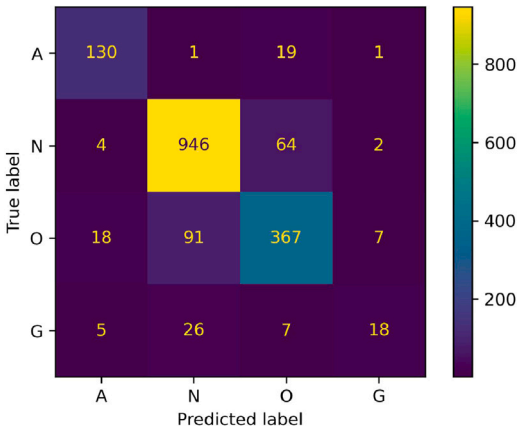



**Fig. 4.** Confusion matrix for the proposed method on fold 0 of cross validation.

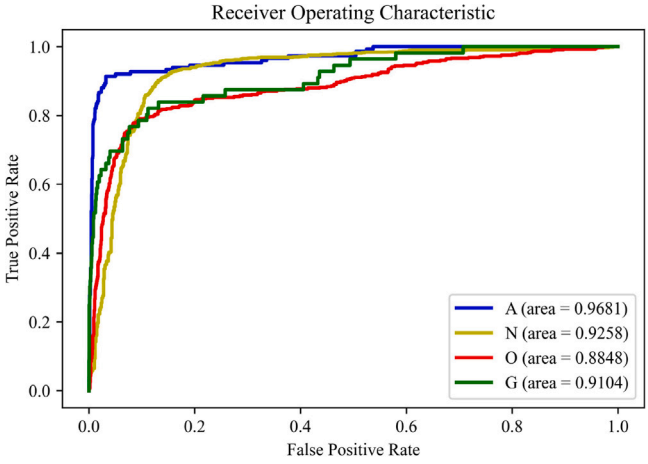



**Fig. 5.** ROC curves for the proposed method on fold 0 of cross validation. "area" refers to AUC.

As can be seen from Table 3, Stacking SVM outperformed the proposed method on recognition for normal rhythm and other rhythm, our method can better detect the AF rhythm. In addition, Stacking SVM needs to train both CNN and SVM, therefore it is considered a two-stage method. However, the proposed method is a one-stage method, which is easier for development and application. Therefore, considering both complexity and performance the proposed method is still attractive when compared with the Stack SVM.

#### 3.5.1. Comparison of loss functions

Cross entropy is the most commonly used loss function for classification. To compare the LM loss with CE loss, we introduced CE loss to our proposed model for training. Besides, we also used different regularization parameters $\lambda$ in LM loss. The results are shown in Table 4. We can observe that when $\lambda \leq 0.5$, LM loss outperforms CE loss, and the performance is best when $\lambda = 0.3$. This proved that introducing a modest large margin regularization improved the AF recognition. However, when $\lambda \geq 0.3$, the performance descends with $\lambda$ increasing, because large $\lambda$ will make the classification too hard. Especially, we found that the training cannot converge when $\lambda \geq 0.8$.

#### 3.5.2. Effects of mix-up

To evaluate the efficacy of mix-up, the performance of the proposed model without mix-up is shown in Table 5. We observed that the addition of mix-up improved the performance. During model training, we discovered that the existence of mix-up overcame the overfitting.

**Table 2**
Precision, recall and AUC for each class on five folds.

| Cross-validation | Normal | | | AF | | | Others | | |
|---|---|---|---|---|---|---|---|---|---|
| | Precision | Recall | AUC | Precision | Recall | AUC | Precision | Recall | AUC |
| Fold 0 | 0.89 | 0.94 | 0.926 | 0.84 | 0.84 | 0.968 | 0.80 | 0.77 | 0.885 |
| Fold 1 | 0.88 | 0.93 | 0.929 | 0.76 | 0.82 | 0.918 | 0.81 | 0.71 | 0.888 |
| Fold 2 | 0.90 | 0.89 | 0.918 | 0.82 | 0.88 | 0.965 | 0.74 | 0.78 | 0.886 |
| Fold 3 | 0.88 | 0.90 | 0.910 | 0.77 | 0.88 | 0.958 | 0.74 | 0.72 | 0.875 |
| Fold 4 | 0.90 | 0.93 | 0.938 | 0.75 | 0.90 | 0.955 | 0.81 | 0.75 | 0.885 |
| Average | 0.890 | 0.918 | 0.924 | 0.788 | 0.864 | 0.953 | 0.780 | 0.746 | 0.884 |

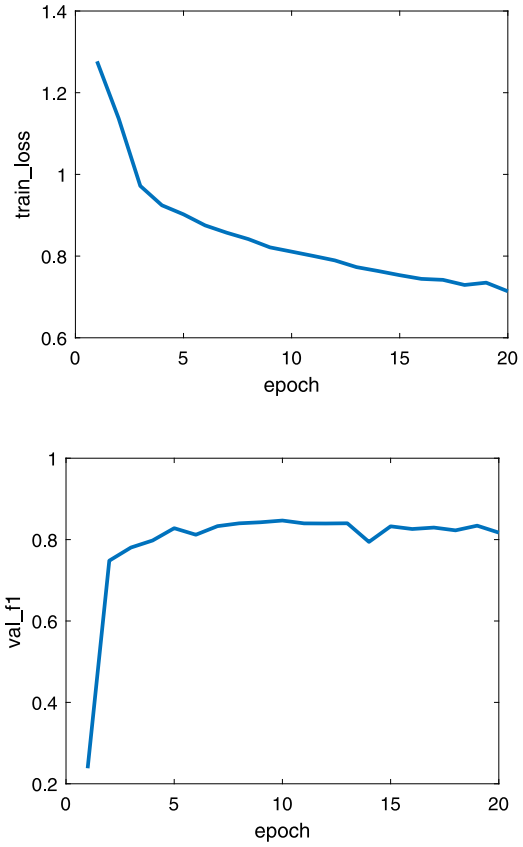



**Fig. 6.** The training and validation curves. "train_loss" refers to large-margin softmax loss on training set. "val_f1" refers to F1 score on validation set.

**Table 3**
The performance comparison between different end-to-end methods.

| Cross-validation | $F_{1n}$ | $F_{1a}$ | $F_{1o}$ | $F_1$ |
|---|---|---|---|---|
| RNN [32] | 0.800 | 0.720 | 0.640 | 0.720 |
| CNN [32] | 0.900 | 0.820 | 0.750 | 0.820 |
| CNN [41] | 0.878 | 0.790 | 0.701 | 0.790 |
| CRNN [41] | 0.888 | 0.764 | 0.726 | 0.792 |
| CNN [37] | 0.885 | 0.810 | 0.685 | 0.794 |
| CNN_vote [11] | 0.898 | 0.799 | 0.709 | 0.802 |
| Stacking SVM [11] | 0.930 | 0.787 | 0.793 | 0.837 |
| Proposed | 0.906 | 0.822 | 0.760 | 0.829 |

Hence, this proved that the inclusion of mix-up served as an effective data augmentation for AF recognition. In addition, performance of the system without mix-up was still better than baseline systems in Table 5, which again illustrates the superiority of DPRNN for AF.

### 3.5.3. Comparison of ensemble

As introduced in Section 1, ensemble methods could obtain better performance through integrating multiple features or models. It is known that having more complementarity between the features of the models, or having more numbers of integrated features in the models, would result in improvement of performance. Therefore, the methods which won good place in PhysioNet 2017 challenge used ensemble. However, the numbers and types of the features of models integrated in previous works were typically different. This made establishing a fair comparison for ensemble systems difficult.

For a qualitative comparison, we made an ensemble for the proposed end-to-end method with different settings, including with or without mix-up, CE loss or LM loss. We averaged the output of the four systems and attained the final labels. The result is shown in Table 6. Compared to ensembles with different features and integrated model, our model achieved a good performance. We believe that the performance of the proposed method can be further improved through ensemble with hand-craft features or other type of deep learning model, or using a better ensemble method, such as Adaboost and SVM.

**Table 4**
Performance comparison of loss functions with different parameters reflecting means of each class on five folds.

| Cross-validation | $F_{1n}$ | $F_{1a}$ | $F_{1o}$ | $F_1$ |
|---|---|---|---|---|
| CE loss | 0.896 | 0.823 | 0.749 | 0.8226 |
| LM loss ($\lambda = 0.1$) | 0.900 | 0.818 | 0.748 | 0.8226 |
| LM loss ($\lambda = 0.3$) | **0.906** | **0.822** | **0.760** | **0.8291** |
| LM loss ($\lambda = 0.5$) | 0.898 | 0.820 | 0.754 | 0.8245 |
| LM loss ($\lambda = 0.7$) | 0.892 | 0.816 | 0.746 | 0.8197 |

**Table 5**
The performance comparison between different systems with/without mix-up. "WO" refers to without.

| Cross-validation | $F_{1n}$ | $F_{1a}$ | $F_{1o}$ | $F_1$ |
|---|---|---|---|---|
| Proposed WO mix-up | 0.892 | 0.806 | 0.732 | 0.8100 |
| Proposed | **0.906** | **0.822** | **0.760** | **0.8291** |

**Table 6**
The performance comparison between different ensemble systems.

| Cross-validation | $F_{1n}$ | $F_{1a}$ | $F_{1o}$ | $F_1$ |
|---|---|---|---|---|
| Feature_ensemble [34] | 0.880 | 0.788 | 0.670 | 0.780 |
| RNN_ensemble [31] | 0.884 | 0.752 | 0.719 | 0.785 |
| Features_ensemble [21] | 0.914 | 0.805 | 0.891 | 0.803 |
| CNN_ensemble [33] | 0.910 | 0.800 | 0.760 | 0.820 |
| Feature_ensemble [17] | 0.910 | 0.798 | 0.772 | 0.826 |
| Feature_ensemble [16] | 0.909 | 0.835 | 0.734 | 0.830 |
| CNN+LSTM_ensemble [40] | 0.901 | 0.760 | 0.752 | 0.831 |
| Proposed | **0.930** | **0.842** | **0.782** | **0.851** |

### 3.6. Generalization of the model

To study the generalization of the proposed method, we tested the model on the dataset of China Physiological Signal Challenge (CPSC) 2018. The CPSC dataset has 6877 12-lead ECG recordings, including 1098 AF recordings and 5779 non-AF recordings. However, our model was trained with PhysioNet 2017, which includes single-lead ECG recordings of 4 classes. Therefore, we extracted one-lead from CPSC 2018 dataset and tested the ensemble model on that. When

**Table 7**
The performance comparison between different methods with five-fold ensemble on CPSC 2018 dataset.

| Methods | Accuracy | $F_1$ |
|---|---|---|
| CNN [37] | 0.947 | 0.891 |
| CNN_vote [11] | 0.958 | 0.927 |
| Proposed | **0.971** | **0.953** |

we computed the F1-score, the normal rhythm, other rhythm and noisy recordings were regarded as non-AF recordings. The results are shown in Table 7. Comparing to baseline systems, we can observe that the proposed method achieved better performance, and it has generalization ability on other dataset.

## 4. Conclusions

In this paper, a novel end-to-end AF recognition method that has superior performance compared to all existing methods is proposed. The method utilizes dual-path recurrent neural network (DPRNN) from single-lead ECG signal. The proposed model takes the whole ECG as input, used DPRNN to split the ECG into shorter segments and models the sequence between intra- and inter-segment iteratively. Mix-up operation was incorporated predominantly for data augmentation, which then eliminated the problem of limited data, thus giving the system an improved AF recognition performance. We evaluated our method on the dataset from PhysioNet Challenge 2017. Experimental results show that the proposed method can effectively recognize AF with ECG signal without any additional expertise, and outperforms state-of-the-art baseline methods. It demonstrates that dual-path model is effective for ECG analysis, which can potentially be applied to other types of medical data.

The proposed method have been validated on single-lead ECG signal. However, 12 lead ECG can provide more information for AF detection. In the future, we will develop multi-channel DPRNN method to explore the 12 lead ECG signal. In addition, we also plan to verify the proposed method on subjects clinically.

## Funding

Mou Wang gratefully acknowledges financial support from China Scholarship Council. The work of S. Rahardja was supported in part by the Overseas Expertise Introduction Project for Discipline Innovation, Finland (111 project: B18041). The publication costs were covered by the authors.

## CRediT authorship contribution statement

**Mou Wang:** Conceptualization, Investigation, Formal analysis, Validation, Visualization, Writing – original draft, Editing, Software. **Sylwan Rahardja:** Conceptualization, Investigation, Formal analysis, Validation, Visualization, Writing – original draft, Editing. **Pasi Fränti:** Formal analysis, Writing review & editing, Supervision. **Susanto Rahardja:** Conceptualization, Formal analysis, Writing review & editing, Supervision.

## Declaration of competing interest

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

## Data availability

Public data has been used for the performance comparison.

## References

[1] G.Y.H. Lip, L. Fauchier, S.B. Freedman, I.V. Gelder, A. Natale, C. Gianni, S. Nattel, T. Potpara, M. Rienstra, H.-F. Tse, D.A. Lane, Atrial. fibrillation, Atrial fibrillation, Nat. Rev. Dis. Primers 2 (16016) http://dx.doi.org/10.1038/nrdp.2016.16.

[2] V. Fuster, L.E. Rydén, D.S. Cannom, H.J. Crijns, A.B. Curtis, K.A. Ellenbogen, J.L. Halperin, J.-Y. Le Heuzey, G.N. Kay, J.E. Lowe, et al., Guidelines for the management of patients with atrial fibrillation executive summary, Rev. Espanola Cardiol. 59 (12) (2006) 1329.

[3] B.J. Hansen, J. Zhao, T.A. Csepe, B.T. Moore, N. Li, L.A. Jayne, A. Kalyanasundaram, P. Lim, A. Bratasz, K.A. Powell, et al., Atrial fibrillation driven by micro-anatomic intramural re-entry revealed by simultaneous sub-epicardial and sub-endocardial optical mapping in explanted human hearts, Eur. Heart J. 36 (35) (2015) 2390–2401.

[4] G.V. Naccarelli, H. Varker, J. Lin, K.L. Schulman, Increasing prevalence of atrial fibrillation and flutter in the united states, Am. J. Cardiol. 104 (11) (2009) 1534–1539.

[5] D.O. Arnar, G.H. Mairesse, G. Boriani, H. Calkins, A. Chin, A. Coats, J.-C. Deharo, J.H. Svendsen, H. Heidbüchel, R. Isa, et al., Management of asymptomatic arrhythmias: a european heart rhythm association (ehra) consensus document, endorsed by the heart failure association (hfa), heart rhythm society (hrs), asia pacific heart rhythm society (aphrs), cardiac arrhythmia society of southern africa (cassa), and latin america heart rhythm society (lahrs), EP Europace.

[6] S. I., K. N, K. A, C. AJ, Upstream therapies for management of atrial fibrillation: review of clinical evidence and implications for european society of cardiology guidelines part I: primary prevention, Europace 13 (3) (2011) 308–328, http://dx.doi.org/10.1093/europace/eur002.

[7] C. FG, A. E, B. GLJd, Delayed rhythm control of atrial fibrillation may be a cause of failure to prevent recurrences: reasons for change to active antiarrhythmic treatment at the time of the first detected episode, Europace 10 (1) (2008) 21–27, http://dx.doi.org/10.1093/europace/eum276.

[8] D.H. Lau, D. Linz, U. Schotten, R. Mahajan, P. Sanders, J.M. Kalman, Pathophysiology of paroxysmal and persistent atrial fibrillation: rotors foci and fibrosis, Heart Lung Circul. 26 (9) (2017) 887–893.

[9] Y. Xia, N. Wulan, K. Wang, H. Zhang, Detecting atrial fibrillation by deep convolutional neural networks, Comput. Biol. Med. 93 (2018) 84–92.

[10] A. Bansal, R. Joshi, Portable out-of-hospital electrocardiography: A review of current technologies, J. Arrhythmia 34 (2) (2018) 129–138.

[11] Q.H. Nguyen, B.P. Nguyen, T.B. Nguyen, T.T. Do, J.F. Mbinta, C.R. Simpson, Stacking segment-based cnn with svm for recognition of atrial fibrillation from single-lead ecg recordings, Biomed. Signal Process. Control 68 (2021) 102672, http://dx.doi.org/10.1016/j.bspc.2021.102672.

[12] T. Tuncer, S. Dogan, P. Plawiak, A. Subasi, A novel discrete wavelet-concatenated mesh tree and ternary chess pattern based ecg signal recognition method, Biomed. Signal Process. Control 72 (2022) 103331, http://dx.doi.org/10.1016/j.bspc.2021.103331, https://www.sciencedirect.com/science/article/pii/S1746809421009289.

[13] M.A. Kobat, O. Karaca, P.D. Barua, S. Dogan, Prismatoidpatnet54: An accurate ecg signal classification model using prismatoid pattern-based learning architecture, Symmetry 13 (10) http://dx.doi.org/10.3390/sym13101914 https://www.mdpi.com/2073-8994/13/10/1914.

[14] M. Baygin, T. Tuncer, S. Dogan, R.-S. Tan, U.R. Acharya, Automated arrhythmia detection with homeomorphically irreducible tree technique using more than 10000 individual subject ecg records, Inf. Sci. 575 (2021) 323–337, 10.1016/j.ins.2021.06.022.

[15] A. Subasi, S. Dogan, T. Tuncer, A novel automated tower graph based ecg signal classification method with hexadecimal local adaptive binary pattern and deep learning, J. Ambient Intell. Hum. Comput. http://dx.doi.org/10.1007/s12652-021-03324-4.

[16] M. Zabihi, A.B. Rad, A.K. Katsaggelos, S. Kiranyaz, S. Narkilahti, M. Gabbouj, Detection of atrial fibrillation in ecg hand-held devices using a random forest classifier, in: 2017 Computing in Cardiology (CinC), 2017, pp. 1–4.

[17] S. Datta, C. Puri, A. Mukherjee, R. Banerjee, A.D. Choudhury, R. Singh, A. Ukil, S. Bandyopadhyay, A. Pal, S. Khandelwal, Identifying normal, af and other abnormal ecg rhythms using a cascaded binary classifier, in: 2017 Computing in Cardiology (CinC), 2017, pp. 1–4, http://dx.doi.org/10.22489/CinC.2017.173-154.

[18] D. Sopic, E. De Giovanni, A. Aminifar, D. Atienza, Hierarchical cardiac-rhythm classification based on electrocardiogram morphology, in: 2017 Computing in Cardiology (CinC), 2017, pp. 1–4, http://dx.doi.org/10.22489/CinC.2017.343-119.

[19] S. Ladavich, B. Ghoraani, Rate-independent detection of atrial fibrillation by statistical modeling of atrial activity, Biomed. Signal Process. Control 18 (2015) 274–281, http://dx.doi.org/10.1016/j.bspc.2015.01.007.

[20] S. Yazdani, P. Laub, A. Luca, J.-M. Vesin, Heart rhythm classification using short-term ecg atrial and ventricular activity analysis, in: 2017 Computing in Cardiology (CinC), 2017, pp. 1–4, http://dx.doi.org/10.22489/CinC.2017.067-120.

[21] C. Liu, Q. Li, P.B. Suresh, A. Vest, G.D. Clifford, Multi-source features and support vector machine for heart rhythm classification, in: 2017 Computing in Cardiology (CinC), 2017, pp. 1–4.
[22] N. Larburu, T. Lopetegi, I. Romero, Comparative study of algorithms for atrial fibrillation detection, in: 2011 Computing in Cardiology, 2011, pp. 265–268.
[23] M. Soos, D. McComb, Sinus arrhythmia, 2021, https://www.ncbi.nlm.nih.gov/books/NBK537011/.
[24] M. Wang, M. Zhao, J. Chen, S. Rahardja, Nonlinear unmixing of hyperspectral data via deep autoencoder networks, IEEE Geosci. Remote Sens. Lett. 16 (9) (2019) 1467–1471, http://dx.doi.org/10.1109/LGRS.2019.2900733.
[25] M. Wang, R. Wang, X.-L. Zhang, S. Rahardja, Hybrid constant-q transform based cnn ensemble for acoustic scene classification, in: 2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), 2019, pp. 1511–1516.
[26] W. Zhu, M. Wang, X.-L. Zhang, S. Rahardja, A comparison of handcrafted, parameterized, and learnable features for speech separation, in: 2021 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), 2021, pp. 635–639.
[27] B. Pyakillya, N. Kazachenko, N. Mikhailovsky, Deep learning for ECG classification, J. Phys. Conf. Ser. 913 (2017) 012004, http://dx.doi.org/10.1088/1742-6596/913/1/012004.
[28] Z. Yao, Z. Zhu, Y. Chen, Atrial fibrillation detection by multi-scale convolutional neural networks, in: 2017 20th International Conference on Information Fusion (Fusion), 2017, pp. 1–6, http://dx.doi.org/10.23919/ICIF.2017.8009782.
[29] A.Y. Hannun, P. Rajpurkar, M. Haghpanahi, G.H. Tison, C. Bourn, M.P. Turakhia, A.Y. Ng, Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network, Nat. Med. 25 (1) (2019) 65–69.
[30] G.D. Clifford, C. Liu, B. Moody, L.-w. H. Lehman, I. Silva, Q. Li, A.E. Johnson, R.G. Mark, Af classification from a short single lead ecg recording: The physionet/computing in cardiology challenge 2017, in: 2017 Computing in Cardiology (CinC), 2017, pp. 1–4, http://dx.doi.org/10.22489/CinC.2017.065-469.
[31] D. Smoleń, Atrial fibrillation detection using boosting and stacking ensemble, in: 2017 Computing in Cardiology (CinC), 2017, pp. 1–4, http://dx.doi.org/10.22489/CinC.2017.068-247.
[32] Z. Xiong, M.K. Stiles, J. Zhao, Robust ecg signal classification for detection of atrial fibrillation using a novel neural network, in: 2017 Computing in Cardiology (CinC), 2017, pp. 1–4, http://dx.doi.org/10.22489/CinC.2017.066-138.
[33] J. Rubin, S. Parvaneh, A. Rahman, B. Conroy, S. Babaeizadeh, Densely connected convolutional networks and signal quality analysis to detect atrial fibrillation using short single-lead ecg recordings, in: 2017 Computing in Cardiology (CinC), 2017, pp. 1–4, http://dx.doi.org/10.22489/CinC.2017.160-246.
[34] S. Ghiasi, M. Abdollahpur, N. Madani, K. Kiani, A. Ghaffari, Atrial fibrillation detection using feature based algorithm and deep convolutional neural network, in: 2017 Computing in Cardiology (CinC), 2017, pp. 1–4, http://dx.doi.org/10.22489/CinC.2017.159-327.
[35] T. Teijeiro, C.A. García, D. Castro, P. Félix, Arrhythmia classification from the abductive interpretation of short single-lead ecg records, in: 2017 Computing in Cardiology (CinC), 2017, pp. 1–4, http://dx.doi.org/10.22489/CinC.2017.166-054.
[36] B.S. Chandra, C.S. Sastry, S. Jana, S. Patidar, Atrial fibrillation detection using convolutional neural networks, in: 2017 Computing in Cardiology (CinC), 2017, pp. 1–4, http://dx.doi.org/10.22489/CinC.2017.163-226.
[37] A. Ukil, L. Marin, S.C. Mukhopadhyay, A.J. Jara, Afsense-ecg: Atrial fibrillation condition sensing from single lead electrocardiogram (ecg) signals, IEEE Sens. J. (2022) 1, http://dx.doi.org/10.1109/JSEN.2022.3162691.
[38] P. Zhang, C. Ma, Y. Sun, G. Fan, F. Song, Y. Feng, G. Zhang, Global hybrid multi-scale convolutional network for accurate and robust detection of atrial fibrillation using single-lead ecg recordings, Comput. Biol. Med. 139 (2021) 104880, http://dx.doi.org/10.1016/j.compbiomed.2021.104880, https://www.sciencedirect.com/science/article/pii/S0010482521006740.
[39] V. Sujadevi, K. Soman, R. Vinayakumar, Real-time detection of atrial fibrillation from short time single lead ecg traces using recurrent neural networks, in: The International Symposium on Intelligent Systems Technologies and Applications, 2017, pp. 212–221.
[40] P. Warrick, M.N. Homsi, Cardiac arrhythmia detection from ecg combining convolutional and long short-term memory networks, in: 2017 Computing in Cardiology (CinC), 2017, pp. 1–4, http://dx.doi.org/10.22489/CinC.2017.161-460.
[41] M. Zihlmann, D. Perekrestenko, M. Tschannen, Convolutional recurrent neural networks for electrocardiogram classification, in: 2017 Computing in Cardiology (CinC), 2017, pp. 1–4, http://dx.doi.org/10.22489/CinC.2017.070-060.
[42] S. Hong, M. Wu, Y. Zhou, Q. Wang, J. Shang, H. Li, J. Xie, Encase: An ensemble classifier for ecg classification using expert features and deep neural networks, in: 2017 Computing in Cardiology (CinC), 2017, pp. 1–4, http://dx.doi.org/10.22489/CinC.2017.178-245.
[43] J.L. Ba, J.R. Kiros, G.E. Hinton, Layer normalization, arXiv preprint arXiv:1607.06450.
[44] W. Liu, Y. Wen, Z. Yu, M. Yang, Large-margin softmax loss for convolutional neural networks, in: International Conference on Machine Learning, 2016, pp. 507–516.
[45] T. Kobayashi, Large margin in softmax cross-entropy loss, in: British Machine Vision Conference 2019, 2019, pp. 1–12.
[46] H. Zhang, M. Cisse, Y.N. Dauphin, D. Lopez-Paz, mixup: Beyond empirical risk minimization, arXiv preprint arXiv:1710.09412.
[47] J. Bai, J. Chen, M. Wang, Multimodal urban sound tagging with spatiotemporal context, IEEE Trans. Cogn. Dev. Syst. (2022) 1, http://dx.doi.org/10.1109/TCDS.2022.3160168.