AF Detection by Exploiting the Spectral and Temporal Characteristics of ECG Signals with the LSTM Model

Yen-Chun Chang[1], Sau-Hsuan Wu[1], Li-Ming Tseng[1&2], Hsi-Lu Chao[1], Chun-Hsien Ko[1]

[1] National Chiao Tung University, Hsinchu, Taiwan
[2] Shin Kong Wu Ho-Su Memorial Hospital, Taipei, Taiwan

Abstract

This research reinvestigates the detection of atrial fibrillation (AF) from a recurrent neural network (RNN) viewpoint. In particular, a long short-term memory (LSTM) model of RNN is designed to exploit the high-order spectral and temporal features of the multi-lead electrocardiogram (ECG) signals of patients with AF. To verify the proposed method, the LSTM model is tested with ECG data available from the PhysioNet and some normal ECG data collected in our labs. The results show that not only the deviation of the so-called RR intervals of ECG signals but also its temporal variations are critical to AF detection. The accuracy of AF detection can reach up to 98.3 %, with an LSTM model of using 30 hidden units. Considering more realistic applications, we further tested the model with subjects different from that of the training data. The accuracy is about 87% with high sensitivity. The experimental results show that the proposed model allows us to extract both the long-term and short-term characteristics of the spectral content of the AF ECG signals, making it a good learning model for AF detection.

1. Introduction

With the growth of aging population, health care for the elderly has become an important welfare service. One of the threats to the elderly is heart disease, e.g. atrial fibrillation (AF), which could lead to stroke, heart failure, and blood clots. Studies show that the prevalence of AF is related to aging. For people over 60s, the prevalence of AF is 4 %, and it is 9 % for people over 80s [1]. It is estimated that over 2.3 million people suffer from AF in the U.S.

In clinical practices, AF could be classified as sustained or paroxysmal. The sustained AF is easier to diagnose for its consistent electrocardiogram (ECG) characteristics. The paroxysmal AF, whose intermittent episodes could last from minutes to hours, is, however, much harder to diagnose without long-term ECG monitoring.

Though more complete and accurate, the standard 12-lead ECG is impractical for long-term ECG monitoring. Alternatively, the Holter monitor or other ambulatory ECG devices are used to provide single or multi-leads ECG recording for 24 to 48 hours. With this type of continuous ECG monitoring, the diagnosis of AF can be improved in some cases [2]. To facilitate real-time diagnosis of AF, various algorithms have also been proposed for AF detection based on expert knowledge, e.g. the absence of P wave, the variability of RR interval, or the cross-entropy of ECG signals [3-4]. The sensitivity of such a knowledge-based detection approach can reach 96 %.

On the other hand, the abundant ECG records collected from telecardiology services have paved the way for AF detection with the new technology of artificial intelligence (AI), of which deep learning, such as Convolutional Neural Networks (CNN) or long short-term memory (LSTM) [5], achieves great success in learning the underlying features of big data. In view of the capability of deep learning in feature extraction, some apply it to replace the handcrafted features in AF detection [6-7], of which [6] develops a multi-scaled fusion of CNN (MS-CNN) that consists of 13 convolutional layers, 5 max-pooling layers, and 3 fully connected layers. In contrast, [7] proposes a rhythm classification model with Convolutional RNN (CRNN), which consists of 24 convolutional layers, 4 max-pooling layers, and one LSTM layer. The input of [7] is pre-processed and expressed as a spectrogram of the spectral features of ECG data. And the accuracies of [6] and [7] can attain 98.13% and 82.13%, respectively.

In contrast to existing AF detection methods that are primarily based on single-lead ECG measurements, we present herein an efficient LSTM model for AF detection with spectrograms of multi-lead ECG signals that can provide more comprehensive ECG information in different orientations of the heart. The accuracy of our proposed algorithm can achieve 98.3% with much lower complexity. Considering clinical practices, we further test the proposed model with a number of designed testing sets whose subjects are different from the training sets. The accuracy of the separated testing is about 87%, and the result suggests that the proposed approach can be applied in health care and provide AF alarms for the elderly.

2. Methods

2.1. Problem formulation

The proposed method is aimed to classify the AF and the Normal Sinus Rhythm (NSR) ECG signals from the collected ECG recordings, which are defined as $\boldsymbol{X} = [\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \dots, \boldsymbol{x}^{(N)}]$ of N ECG records. Each ECG record ($\boldsymbol{x}^{(i)}$) is a $K \times L$ measurement vector, including L samples of K-lead ECG signals. And $\boldsymbol{y}^{(i)}$ is the corresponding one-hot encoding label, indicating the incidence of AF syndrome. The classification problem can be defined as:

$$\boldsymbol{z}^{(i)} = F\left(\boldsymbol{x}^{(i)}; \boldsymbol{\theta}\right),$$

where $F(\cdot)$ is the classification model and $\boldsymbol{\theta}$ is the related parameters. The classification result $\boldsymbol{z}^{(i)}$ represents the probability of each class. The cost function would be the cross entropy of $\boldsymbol{z}^{(i)}$ and $\boldsymbol{y}^{(i)}$. For N observable data in the training set $\boldsymbol{X}$, the cost function is defined as:

$$E(\boldsymbol{X}) = -\frac{1}{N}\sum_{i=1}^{N}\left(\boldsymbol{y}^{(i)}\right)^{T}\ln\left(\boldsymbol{z}^{(i)}\right).$$

2.2. Model Architecture

The proposed RNN is based on a single layer LSTM model [5] whose input is the spectrogram of $\boldsymbol{X}$. Details of the model and the pre-processing procedure of data are described below.

1.) Data Pre-processing:

Considering that not only the heart rate but also the morphology of the ECG signals are critical to AF detection, we need to pre-process the ECG data in order to provide both the long-term and short-term temporal features of them to the NRR. One of our choices is the spectrogram, which can characterize the short-term spectral morphology and the long-term variation of the spectral content.

A spectrogram is a procedure that performs the Short-Term Fourier Transform (STFT) of data on a sliding window basis. Given the input ECG recording $\boldsymbol{X}$, the spectrogram is $\boldsymbol{S}_k^{(i)} \triangleq \left[\boldsymbol{s}_k^{(i)}(1), \boldsymbol{s}_k^{(i)}(2), \dots, \boldsymbol{s}_k^{(i)}(T)\right] \in \mathbb{R}^{M\times T}$, where M is the size of STFT output and T is the number of time bins, representing a 90% overlapping segment with a size of 64-point Hamming window of the input signal. And $\boldsymbol{s}_k^{(i)}(t)$ is the STFT result of the t^{th} time bin. The STFT sequence forms a spectrogram whose horizontal axis represents the temporal sequence and the vertical axis represents the frequency components.

For instance, Fig. 1 consists of the spectrograms converted from an AF and an NSR ECG signals. As the figure shows, the spectrogram of the NSR ECG signals are regular while that of the AF signals are messy. The results also show that the spectrogram can capture more features in each waveform of the ECG signal, not only limited to that of the R-wave.

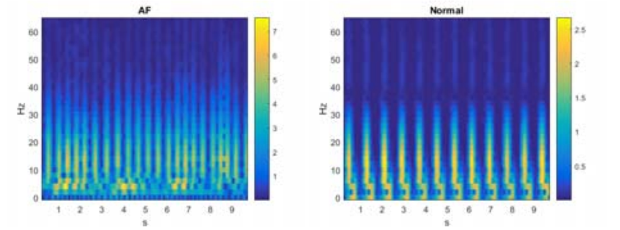

Fig.1 Spectrograms of AF (left) and NSR ECGs (right).

2.) Long Short-Term Memory

LSTM is an improved architecture of RNN, and is more effective at capturing the long-term temporal dependence of data. In addition, LSTM resolves the gradient vanishing or exploding problem by introducing the memory cell $\boldsymbol{c}(t)$ and the controls of the information flow by input gate ($\boldsymbol{v}^{i}$), output gate ($\boldsymbol{v}^{o}$), and forget gate ($\boldsymbol{v}^{f}$). The memory cell and gates are connected with weighting matrices, which are determined in the training phrase by visiting a huge amount of training data.

In the proposed model, we apply LSTM to learn the temporal correlation and convert the input into a 2-channel spectrogram for capturing the spatial and temporal correlations of the ECG signals. As shown in Fig.2, the proposed model consists of three layers: an input layer, an LSTM layer, and an output layer. In the input layer, $\boldsymbol{s}(t) \triangleq [\boldsymbol{s}_1(t),\ \boldsymbol{s}_2(t)]^T$ be the input vector, including the spectrogram of 2-lead ECG signals at the t^{th} time bin. In the LSTM layer, H is the hidden size of the memory cells. The output layer is the label of classifier. Several notations of the model are defined as follows:

Input Weights: $\boldsymbol{W}_c, \boldsymbol{W}_i, \boldsymbol{W}_f, \boldsymbol{W}_o \in R^{H\times 2M}$
Recurrent Weights: $\boldsymbol{R}_c, \boldsymbol{R}_i, \boldsymbol{R}_f, \boldsymbol{R}_o \in R^{H\times H}$
Bias Weights: $\boldsymbol{b}_c, \boldsymbol{b}_i, \boldsymbol{b}_f, \boldsymbol{b}_o \in R^{H}$
Output Weights: $\boldsymbol{W}_z \in R^{2\times H}$

Then the vector formulas for the LSTM layer forward pass can be written as:

$$\boldsymbol{v}^{j} = \sigma\left(\boldsymbol{W}_j\boldsymbol{s}(t) + \boldsymbol{R}_j\boldsymbol{h}(t-1) + \boldsymbol{b}_j\right), j \in \{i, f, o\}$$
$$\boldsymbol{net}_c(t) = \tanh(\boldsymbol{W}_c\boldsymbol{s}(t) + \boldsymbol{R}_c\boldsymbol{h}(t-1) + \boldsymbol{b}_c)$$
$$\boldsymbol{c}(t) = \boldsymbol{v}^{f}\cdot\boldsymbol{c}(t-1) + \boldsymbol{v}^{i}(t)\cdot g(\boldsymbol{net}_c(t))$$
$$\boldsymbol{h}(t) = \boldsymbol{v}^{o}\cdot g(\boldsymbol{c}(t))$$
$$\boldsymbol{z} = softmax(\boldsymbol{W}_z\boldsymbol{h}(T))$$

where $g(\cdot)$ is the hyperbolic tangent (tanh) activation function, $\sigma(\cdot)$ is the sigmoid activation function, and $\boldsymbol{h}(t)$ is the hidden vector of dimension H.

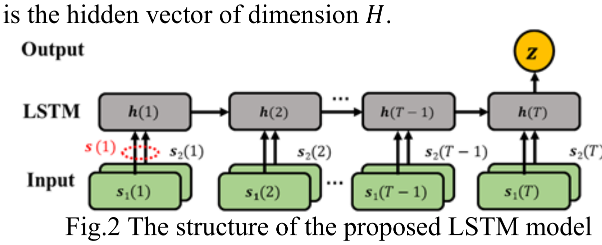

Fig.2 The structure of the proposed LSTM model

As shown in Fig. 2, the structure of our AF detection model consists of pre-processed 2-channel spectrogram and an LSTM classifier learning from huge training ECG data. The experimental results of the proposed model are presented in Section 3.

3. Experimental Results

The proposed LSTM model is developed under the supervised learning framework, where the model is learned from the training data. Therefore, the quality of the proposed model is directly related to the size and quality of training data. To prepare the AF and the NSR ECG data as complete as possible, we collect the AF ECG records from six standard databases [8-13] and obtain the NSR ECG data from three databases [13, 14] and some normal ECG signals collected from our lab and Shin Kong Wu Ho-Su Memorial Hospital, Taipei, Taiwan. All the ECG signals from databases are obtained from the PhysioNet [13] with long-term ECG records splitted into multiple 12-second segments. Each segment is downsampled to 128 Hz and pre-processed with an FIR high-pass filter for removing baseline wondering.

The entire ECG data are classified into 3 different categories: NSR, sustained AF, and paroxysmal AF. Paroxysmal AF episodes are self-terminated and followed by NSR. On the contrary, sustained AF episodes could not terminate untill treatment. In order to balance the classified labels, the number of NSR ECG recoders is the same to the total number of AF ECG records, which includes equal numbers of sustained and paroxysmal AF records. In summary, a total of 3,312 ECG segments are collected in our experiment, of which the total number of NSR segments is 1,656 extracted from 103 subjects, and the AF segment number is 1,656 extracted from 118 AF subjects.

3.1. Data Preparations

To verify the performance of the proposed model, two validation sets are designed. In the first set, named Common Set, 1,000 ECG segments are randomly selected to form a testing set, and the rest are the training data. In the Common Set, the ECG segments from the same subject may be categorized as training and testing data. As such, the subject-based pattern may lead to model overfitting and, hence, the reported accuracy may not imply a similar performance of AF detection on unknown subjects. Considering clinical practices, we propose an alternative validation approach named Separated Set. In this approach, the corresponding subjects of ECG segments in the training and the testing sets are separated. The Separated Set is designed by arbitrarily choosing the testing subjects with the number of ECG segments in the testing set is 1,000. Considering that the chosen subjects also influence the accuracy of the Separated Set, we form 10 different Separated Sets and average their outcomes.

3.2 Performance Evaluations

The proposed LSTM architecture is implemented with TensorFlow. The number of the hidden units is 30, which is determined by examining several trials on model settings. The results show that more numbers of hidden units do not improve the results, but rather increase the processing time. Furthermore, the gradient descent method with an Adam optimizer [15] is used to update the model parameters, whose learning rate is 0.001 and batch size is 34. For the spectrograms, they are genterated with a 64-point FFT out of an input vector of the same size. The ECG samples in the input vectors are overlapped by 90% with each other using a sliding Hamming window.

The performance of the proposed algorithm is evaluated with the following statistical measures: Accuracy (Acc), Positive predictive values (PPV), Sensitivity (Sen), and Specificity (Spec). Table 1 shows the performance of a single lead model (SL) and a double lead model (DL) with the common set. Table 2 shows the averaged performance of SL and DL of 10 different separated testing sets. The results show the performance of the Cmmon Set is much better than Separated Set for both SL and DL cases. It is because the model may separate subjects by the personal features, not the AF features. Moreover, the Separated Set is more practical in diagnosing some emergency patients. In this case, we cannot obtain the patient's previous ECG signals and need to infer the results based on the training ECG data from other subjects. And it is the reason why the separated set is important in our experimental results.

Table 1. Performance of the Common Set

	Acc	PPV	Sen	Spec
NSR (SL)	98.50	98.60	98.40	98.60
AF (SL)	98.50	98.40	98.60	98.40
NSR (DL)	99.10	99.00	99.20	99.00
AF (DL)	99.10	99.20	99.00	99.20

Table 2. Performance of the Separate Set

	Acc	PPV	Sen	Spec
NSR (SL)	83.21	83.08	86.88	79.55
AF (SL)	83.21	86.42	79.55	86.88
NSR (DL)	87.57	88.48	87.79	87.35
AF (DL)	87.57	87.96	87.35	87.79

In addition, DL performs better than SL in both the Common and the Separated Sets. According to the results, we can find that DL can provide more information to the AF detection problem and improve the performance. Comparing the misclassified results of SL and DL, we find that the additional ECG channel in DL provides stronger P waves in some NSR cases. Recall that the absence of P wave is an important feature for AF detection. Therefore, additional information is helpful to avoid the false alarms.

The experimental results also show that the signal quality is important to AF detection. In some cases, the noisy ECG signals and motion artifact in them can lead to misclassifications. Besides, considering that the proposed scheme is based on the ECG spectrograms, the frequency-based features may not fully characterize the morphology of the slightly changing ECG signals, e.g., the F waves in AF syndrome. This disadvantage also limits the accuracy of the proposed AF detection method.

4 Conclusions

In this paper, an AF detection method was proposed to exploit the spectral and temporal characteristics of AF ECG signals with a multi-lead LSTM model. No handcraft features were used in the proposed method. To verify the performance of the proposed AF detection method, two types of data sets, the common set and the separated set, were designed for testing the accuracy and robustness of the proposed LSTM model. Experimental results showed that the proposed method could achieve 98% accuracy in the common set and 85% accuracy in the separated set. This suggests that the proposed AF detection method has the potential to be applied in clinical practices.

Acknowledgments

This work was supported in part by the Ministry of Science and Technology (MOST), Taiwan, under the agreement of 107-2218-E-009-007, and in part by the Center for Open Intelligent Connectivity from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan.

References

[1] A. S. Go, et al. Prevalence of Diagnosed Atrial Fibrillation in Adults: National Implications for Rhythm Management and Stroke Prevention: the AnTicoagulation and Risk Factors in Atrial Fibrillation (ATRIA) Study 2001; 285:2370-2375.

[2] E. Fung, et al. Electrocardiographic Patch Devices and Contemporary Wireless Cardiac Monitoring. Frontiers in physiology 2015; pp. 6:149.

[3] C. Hung, et al. A Novel Method for Detection of the Transition Between Atrial Fibrillation and Sinus Rhythm. IEEE Transactions on Biomedical Engineering 2011; 58.4:1113-1119.

[4] B. M. Asl, et al. Support Vector Machine-Based Arrhythmia Classification Using Reduced Features of Heart Rate Variability Signal. Artificial Intelligence inMmedicine 2008; 44:51-64.

[5] F. A. Gers, et al. Learning to Forget: Continual Prediction with LSTM. 1999.

[6] X. Fan, et al. Multi-Scaled Fusion of Deep Convolutional Neural Networks for Screening Atrial Fibrillation from Single Lead Short ECG Recordings. IEEE Journal of Biomedical and Health Informatics 2018.

[7] M. Zihlmann, et al. Convolutional Recurrent Neural Networks for Electrocardiogram Classification. Computing 2017; 44:1.

[8] S. Petrutiu, et al. Abrupt Changes in Fibrillatory Wave Characteristics at the Termination of Paroxysmal Atrial Fibrillation in Humans. Europace 2007; 9:466-470.

[9] GB Moody, et al. A New Method for Detecting Atrial Fibrillation Using R-R Intervals. Computers in Cardiology 1983; 10:227-230.

[10] GB Moody, et al. The Impact of the MIT-BIH Arrhythmia Database. IEEE Engineering in Medicine and Biology Magazine 2011; 20:45-50

[11] A. Taddei, et al. The European ST-T Database: Standard for Evaluating Systems for the Analysis of ST-T Changes in Ambulatory Electrocardiography. European Heart Journal 1992; 13:1164-1172.

[12] F. Nolle, et al. A New Concept in Computerized Arrhythmia Monitoring Systems. Computers in Cardiology 1986; 13:515-518.

[13] A. Goldberger, et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. Circulation 2000; 101(23): e215-e220.

[14] R. Bousseljot, et al. A. Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das Internet. Biomedizinische Technik/Biomedical Engineering 1995; 40:317-318.

[15] D. P. Kingma, et al. A Method for Stochastic Optimization. International Conference on Learning Representations 2015.

National Chiao Tung University, Hsinchu, Taiwan
1001 University Road, Hsinchu, Taiwan 300
sauhsuan@nctu.edu.tw
hlchao@cs.nctu.edu.tw

Shin Kong Wu Ho-Su Memorial Hospital, Taipei, Taiwan
No.95, Wenchang Rd., Shilin Dist., Taipei City 111, Taiwan