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[12], Hsi-Lu Chao[1], and 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 is able to effectively extract both the long-term and short-term characteristics of the spectral content of the AF ECG signals, making it a good candidate 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 may 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 multiple leads of 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. The sensitivity of such a knowledge-based detection approach can reach 96% [3-4].

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 LSTM [5], achieves great successes 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 such a separated testing is about 87%, which suggests that the proposed approach has the potential to applied for AF detection or screening in health care.

2. Methods

 ISSN: 2325-887X DOI: 10.22489/CinC.2018.266

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 × 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 is the cross entropy between $\boldsymbol{z}^{(i)}$ and $\boldsymbol{y}^{(i)}$. For N observation 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 for data are described below.

2.2.1 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 RNN. 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, obtained with a 64-point Hamming window that has 90% overlapped input data in each time bin. And $\boldsymbol{s}_k^{(i)}(t)$ is the result of the k^{th} lead of the t^{th} time bin of record i. 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 a NSR ECG signals. As the figure shows, the spectrogram of the NSR ECG signals is more regular while that of the AF signals is messy. The results also show that the spectrogram can capture more features in each waveform of the ECG signals, not only limited to that of the R-wave of ECG signals.

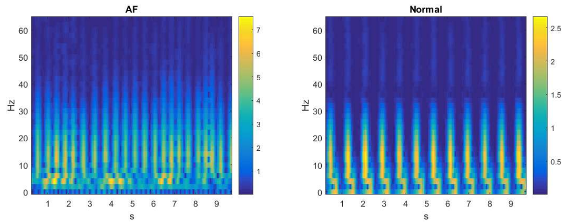

Fig.1 Spectrograms of AF (left) and NSR ECGs (right).

2.2.2 Long Short-Term Memory

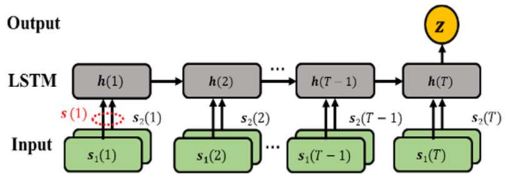

Fig.2 The structure of the proposed LSTM model

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 a memory cell $\boldsymbol{c}(t)$ and means for the control of the information flow by an input gate ($\boldsymbol{v}^{i}$), a output gate ($\boldsymbol{v}^{o}$), and a forget gate ($\boldsymbol{v}^{f}$). The memory cell and the gates are connected with weighting matrices, which are determined in the training phase by visiting a huge amount of training data.

In the proposed model, we apply the LSTM to learn the temporal correlation, and convert the inputs into 2 channel of spectrograms for characterizing 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, $\boldsymbol{s}(t) \triangleq [\boldsymbol{s}_1^T(t), \boldsymbol{s}_2^T(t)]^T$ represents the input vector, including the spectrograms of 2 leads of ECG signals at the t^{th} time bin. In the LSTM layer, H represents 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: $\boldsymbol{W}_z, \boldsymbol{W}_i, \boldsymbol{W}_f, \boldsymbol{W}_o \in R^{H\times 2M}$
Recurrent Weights: $\boldsymbol{R}_z, \boldsymbol{R}_i, \boldsymbol{R}_f, \boldsymbol{R}_o \in R^{H\times H}$
Bias Weights: $\boldsymbol{b}_z, \boldsymbol{b}_i, \boldsymbol{b}_f, \boldsymbol{b}_o \in R^{H}$
Output Weights : $\boldsymbol{W}_s \in R^{2\times H}$

Then, the forward pass of the LSTM layer is given by

$$\boldsymbol{v}^{j} = \sigma\left(\boldsymbol{W}_j\boldsymbol{s}(t) + \boldsymbol{R}_j\boldsymbol{h}(t-1) + \boldsymbol{b}_j\right), j \in \{i, f, o\} \quad (1)$$

$$\boldsymbol{net}_c(t) = \boldsymbol{W}_z\boldsymbol{s}(t) + \boldsymbol{R}_z\boldsymbol{h}(t-1) + \boldsymbol{b}_z \quad (2)$$

$$\boldsymbol{c}(t) = \boldsymbol{v}^{f}\cdot\boldsymbol{c}(t-1) + \boldsymbol{v}^{i}(t)\cdot g(\boldsymbol{net}_c(t)) \quad (3)$$

$$\boldsymbol{h}(t) = \boldsymbol{v}^{o}\cdot g(\boldsymbol{c}(t)) \quad (4)$$

$$\boldsymbol{z} = softmax(\boldsymbol{W}_s\boldsymbol{h}(T)) \quad (5)$$

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

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 collected 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 retrieved from the PhysioNet [13] with long-term ECG records splitted into multiple 12-second segments. Each segment is downsampled to 128 samples per Hz, and is pre-processed with a high-pass filter of finite impulse responses to remove 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 episodes. On the contrary, sustained AF episodes do not terminate untill intervention. In order to balance the classified labels, the number of NSR ECG recodings is the same to the total number of AF ECG recordings, which includes equal numbers of sustained and paroxysmal AF recordings. In summary, a total of 3312 ECG segments are collected in our experiments, of which a total of 1656 NSR segments are extracted from 103 subjects. The other 1656 AF segments are extracted from 118 subjects with AF.

3.1. Data Preparations

To verify the performance of the proposed model, two validation sets are designed. In the first set, named Common Set, 1000 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 or 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 also use an alternative set named Separate Set for validation. In this approach, the corresponding subjects of the ECG segments in the training set are separated from those in the testing set. The subjects for the Separate Set are arbitrarily chosen, with the number of ECG segments from a subject being 1000. Given that the chosen subjects will influence the representativeness of the Separate Set, we thus form 10 different Separated Sets and average their testing outcomes.

3.2 Performance Evaluations

The training and AF detection with the proposed LSTM architecture is implemented with TensorFlow [15]. The number of the hidden units is set to 30 in (1)~(5), which is determined by examining several trials on model settings. The results show more hidden units do not seem to improve the accuracy, but rather increase the processing time. Besides, a gradient descent method with the Adam optimizer [16] is used to update the model parameters, whose learning rate is 0.001 and batch size is 34.

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 comparisons of making use of single-lead (SL) and double-lead (DL) ECG measurements from the Common Set. Furthermore, to justify the use of the spectrogram as the LSTM input, we also present the results of directly using the time-series ECG data as the LSTM input. Similarly, a 64-point sliding window is used to process the time-series inputs, with data of adjacent windows overlapped by 90%.

Table 1. Performance with the Common Set

	Acc	PPV	Sen	Spec
NSR (SL)[1]	98.50	98.60	98.40	98.60
AF (SL)[1]	98.50	98.40	98.60	98.40
NSR (DL)[1]	99.10	99.00	99.20	99.00
AF (DL)[1]	99.10	99.20	99.00	99.20
NSR (SL)[2]	97.20	97.58	96.80	97.60
AF (SL)[2]	97.20	96.83	97.60	96.80
NSR (DL)[2]	98.50	97.83	99.20	97.80
AF (DL)[2]	98.50	99.19	97.80	99.20

1: The spectrogram input 2: The time-series ECG input

Table 2 shows the performance comparisons averaged over 10 different formations of the Separate Set. All the results show that the performance with the spectrogram inputs is better than that with the time-series inputs. The results also show that the performance of the Common Set is much better than that of the Separate Set in either cases. It is because the model may separate subjects with personal features beyond that of the AF.

We note that results with the Separate Set are important and more representative in some cases where there is no patient's previous ECG data and need to infer the results based on the training ECG data from other subjects.

Table 2. Performance with the Separate Set

	Acc	PPV	Sen	Spec
NSR (SL)[1]	83.21	83.08	86.88	79.55
AF (SL)[1]	83.21	86.42	79.55	86.88
NSR (DL)[1]	87.57	88.48	87.79	87.35
AF (DL)[1]	87.57	87.96	87.35	87.79
NSR (SL)[2]	79.46	81.07	79.98	78.95
AF (SL)[2]	79.46	79.98	78.95	79.98
NSR (DL)[2]	86.19	85.75	87.57	84.81
AF (DL)[2]	86.19	87.16	84.81	87.57

1: The spectrogram input 2: The time-series ECG input

Comparing the misclassified results with the SL and the DL measurements, we further find that the additional ECG channel in the DL measurements helps provide more obvious P waves in some NSR cases. Given that the absence of P waves is an important feature for AF detection, the additional ECG lead thus helps avoid the false alarms.

In addition to the performance tests with data from the PhysioNet, we also use data from the CinC challenge 2017 database to verify the robustness of the proposed method to measurement noises. When using data from the CinC challenge 2017 to test the LSTM model learned with data from the PhysioNet, the accuracy and the sensitivity become 75.6% and 70.17% in each case, degrading by 7% and 10%, respectively, compared to the results in Table II. This reveals that the quality of ECG measurements is important to AF detection. In some cases, the noisy ECG signals and motion artefacts in them can lead to misclassifications. Besides, considering that the proposed scheme is based on the spectrograms of ECG, 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 Separate set, were used 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] Go AS, Hylek EM, Phillips KA and Chang YC, 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. JAMA 2001;285:2370-2375.

[2] Fung E, Jarvelin MR, Doshi RN and Shinbane JS, et al. Electrocardiographic patch devices and contemporary wireless cardiac monitoring. Frontiers in Physiology 2015;6:1-10.

[3] Hung C, Ye S, Chen H and Li D, et al. A novel method for detection of the transition between atrial fibrillation and sinus rhythm. IEEE Transactions on Biomedical Engineering 2011;58:1113-1119.

[4] Asl BM, Setarehdan SK and Mohebbi M. Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal. Artificial Intelligence in Medicine 2008;44:51-64.

[5] Gers FA, Schmidhuber J and Cummins F. Learning to forget: Continual prediction with LSTM. Neural Computation 2000;12:2451-2471.

[6] Fan X, Yao Q, Cai Y and Miao F, et al. Multi-scaled fusion of deep convolutional neural networks for screening atrial fibrillation from single lead short ECG recordings. Available online at IEEE Journal of Biomedical and Health Informatics.

[7] Zihlmann M, Perekrestenko D and Tschannen M. Convolutional recurrent neural networks for electrocardiogram classification. Computing in Cardiology 2017;44:1-4.

[8] Petrutiu S, Sahakian AV and Swiryn S. Abrupt changes in fibrillatory wave characteristics at the termination of paroxysmal atrial fibrillation in humans. Europace 2007;9:466-470.

[9] Moody GB and Mark RG. A new method for detecting atrial fibrillation using R-R intervals. Computers in Cardiology 1983;10:227-230.

[10] Moody GB and Mark RG. The impact of the MIT-BIH arrhythmia database. IEEE Engineering in Medicine and Biology Magazine 2001;20:45-50.

[11] Taddei A, Distante G, Emdin M and Pisani P, 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] Nolle FM, Badura FK, Catlett JM and Bowser RW, et al. CREI-GARD, a new concept in computerized arrhythmia monitoring systems. Computers in Cardiology 1986;13:515-518.

[13] Goldberger AL, Amaral LAN, Glass L and Hausdorff JM, et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation 2000;101:215-220.

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

[15] Abadi M, Barham P, Chen J and Chen Z, et al. Tensorflow: A system for large-scale machine learning. Operating systems design and Implementation 2016;16:265-283.

[16] Kingma DP and Ba J. Adam: A method for stochastic optimization. International Conference on Learning Representations 2015.

1. National Chiao Tung University, 1001 University Road, Hsinchu, Taiwan 300.
2. Shin Kong Wu Ho-Su Memorial Hospital, No.95, Wenchang Rd., Shilin Dist., Taipei City 111, Taiwan
3. Contact email: sauhsuan@nctu.edu.tw