Automatic Classification of Arrhythmias by Residual Network and BiGRU With Attention Mechanism

Runnan He [1, 2], Kuanquan Wang [2], Na Zhao [1, 2], Qiang Sun [3], Yacong Li [2], Qince Li [1, 2*], Henggui Zhang[1, 4, 5, 6*]

[1] Peng Cheng Laboratory, Shenzhen, China
[2] School of Computer Science and Technology, Harbin Institute of Technology (HIT), Harbin, Heilongjiang, 150001, China
[3] The Department of Pharmacology, Beijing Electric Power Hospital, Beijing, China
[4] School of Physics and Astronomy, The University of Manchester, Manchester M13 9PL, UK
[5] Pilot National Laboratory of Marine Science and Technology, Qingdao, China
[6] International Laboratory for Smart Systems and Key Laboratory of Intelligent of Computing in Medical Image, Ministry of Education, Northeastern University, Shenyang 110004, China

Abstract

Aims: Over the last decade, many attempts have been implemented for automatic detection of cardiac arrhythmias, however, their performances are still not ideal due to unreliable extracted features of designed models or limited small public datasets. In this study, we investigate automatic detection of arrhythmias from 12-lead electrocardiogram (ECG) recordings using an attention-based Res-BiGRU model. Methods: We train a deep neural networks (DNNs) to identify eight types of arrhythmias. The constructed model contains residual convolutional modules and bidirectional Gated Recurrent Unit (BiGRU) layers to extract features from preprocessed ECG signals. The mechanism of attention is applied to learn an attention distribution on the list of extracted features, and sum the weighted features by the attention into a single feature vector, that is responsible for the final classification. We train our model on the training set of the PhysioNet/Computing in Cardiology Challenge 2020 with time length ranging from 6 s to 60 s. The proposed model can handle signals with variable lengths. Results: The overall score with five-cross validation of training set is 0.543 for our team (DeepHeart), which demonstrates a good efficacy of our model, demonstrating that it may have potential practical applications.

1. Introduction

Cardiac arrhythmias are a group of conditions in which the electrical activity of the heart is not regular, behaving faster or slower rhythm than that under normal condition [1]. Many types of arrhythmias are possibly caused by different cardiac diseases that threaten human health [2-4]. Electrocardiogram (ECG) can be employed to analyze and recognize arrhythmias, which serves as the most popular diagnostic tool in clinical practices for its convenience, non-invasiveness and low cost. In general, there are two steps for analyzing ECG features. The first part is to extract features of ECGs. Then, the ECGs are classified into various conditions according to these extracted features [5]. However, manually analyzing ECG records is tedious and error-prone. Therefore, many researchers have been devoted to developing some methods for automatic arrhythmia classification which can improve the accuracy of diagnosis and reduce costs.

The aim of the challenge is using information from the available ECGs to correctly classify arrhythmias [6]. Previously, many researchers were dedicated themselves to extracting hand-crafted features for automatic arrhythmia detection, which were extracted based on QRS detection [7] or captured from time domain [8], frequency domain [9], and time-frequency domain [10]. Nowadays, deep learning methods have also been applied in the detection of arrhythmias. Different from traditional methods, DNNs can automatically extract useful features from the raw input by learning the probability distribution of the dataset. Therefore, based on sufficient training samples, features extraction by deep learning algorithms can be more effective than hand-crafted algorithms. For example, a convolutional neural network (CNN) has been widely used for automatic diagnosis of cardiac arrhythmias. In the previous study, we have proposed a CNN for atrial fibrillation (AF) detection using the frequency information of ECGs [11]. Furthermore, a multiscale fusion of deep convolutional neural network was also proposed to detect AF, which employs two types of convolutional networks with different filter sizes to extract features of different scales [12]. In another study, a CNN with the residual network model was built to classify 12 rhythm classes,

which exceeds the level of experts [13].

In this paper, we proposed a model including residual network and BiGRU with the attention mechanism for automatic classification of cardiac arrhythmias. The proposed model divides a long ECG signal into several short signal segments. Then multi-scale features are extracted from these signal segments. Based on the extracted features, cardiac diseases the ECG signals are classified into different types.

2.   Method

Figure 1 outlines the architecture of our proposed algorithm which included preprocessing of dataset and a sequence-to-sequence deep neural networks (DNNs) trained to classify arrhythmias. Each major step was explained in details in the two following subsections. Then signal segments divided from a long ECG signal are projected into attention-based Resnet to obtain multi-scale features. These multi-scale features are received by our attention-based BiGRU in chronological order. Since an ECG signal may contain more than one abnormity, our attention-based BiGRU will output which categories this long ECG signal belongs to.
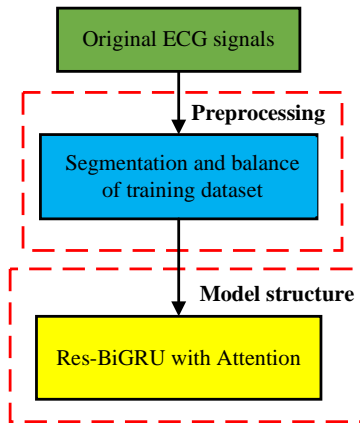

Figure 1. Flowchart diagram of the proposed method for the classification of arrhythmias.

2.1.   Preprocessing

In this challenge, the 12 leads ECG recordings from the PhysioNet dataset consist of about 40,000 recordings with four categories of data sets, which are used for the training of arrhythmia classification. Each recording has an uncertain length ranging from a few seconds to tens of minutes with different sampling rates. The ECG recordings contain about more than one hundred types such as atrial fibrillation (AF) and premature atrial contraction (PAC), which has a labelled annotation. Although most of the recordings only have one label, some recordings have two or three labels. Therefore, it is a multi-label classification

problem.

As it is difficult for training the model with non-identical length of ECG recordings, in the preprocessing phase, each ECG recording was divided into segments with a length of 30 second. If a recording length is less than 30 seconds, we pad the recording into 30 seconds by adding zero values at the beginning period. If the length is more than 30 seconds and less than 60 s, we cut off the extra data after 30 seconds. It is possible that the padding of zero for records shorter of 30 second may have some influence on the performance of the model, however, it is necessary for solving the problem with different input lengths of data and more convenient for training the model. Furthermore, it can reduce the training time of model.

After the preprocessing, the balance of dataset is also crucial for the classification of arrhythmias. In this study, some attention needs to be paid to deal with imbalance training dataset. After analyzing the training set, it is obvious that training dataset is not balanced. To deal with this problem, we have randomly divided the training set into five subsets for each class. After that, each class was copied to make its number equal to the largest number of this category. This operation addressed the issue of dataset imbalance.

2.2.   The design of network structure

Deep learning algorithm can effectively extract features from the raw data and generally achieve a better performance than the traditional hand-crafted features [14, 15]. In this paper, we proposed a new sequence-to-sequence model to learn features from the processed ECG signals. The overall structure of our model is shown in Figure 2. Our proposed network mainly contains residual models, BiGRU and attention layer, allowing extraction of the spatial and temporal information to identify arrhythmias.

In the first part, we utilize stacked residual convolutional modules to compress the long ECG signal into a much shorter sequence which is used to learn spatial features. The input of this part is the raw ECG signal, which is represented as a three-dimensional matrix. The batch size is set to 128, and the length of the other two dimensions is the signal length and lead number, respectively. The applied residual convolutional module in our model was adopted s from Andrew Ng [13]. As shown in Figure 2, the network consists of several type of modules, such as 1-dimensional convolutional (1D Conv) layers, batch normalization (BN) [16] layers, rectified linear units (ReLU) [17] of activation layer, dropout [18] layers and max-pooling [19] layers, accordingly. There are a total of 14 convolutional layers and 7 max-pooling layers in this part. As the margin of the input will be lost during a convolutional operation, the input feature maps are padded before each convolutional layer so that the output has the same length as the original input. The feature maps are

compressed in length only when they go through a pooling layer.
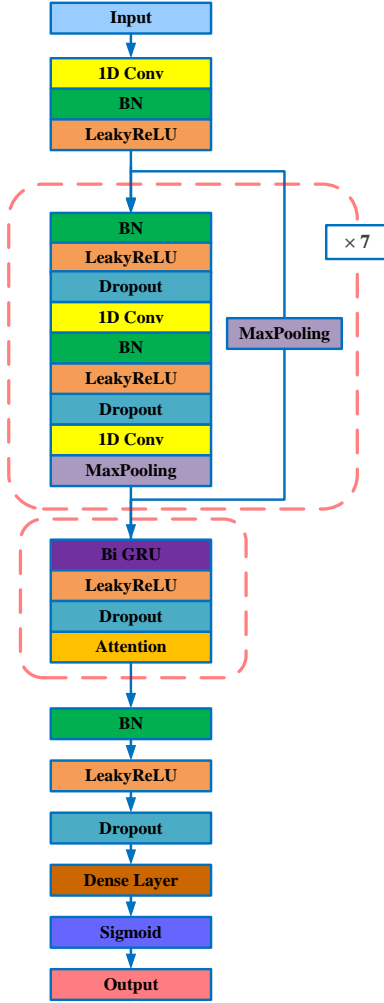

Figure 2. The structure of the proposed model.

After the residual model part, the extracted feature vectors are input into a BiGRU [20] layer. The ECG signal is essentially a periodic signal with certain regularity. Thus, the cardiac condition corresponding to the current sampled value is not only related to the information at previous time point, but also to the information at next time point. To efficiently learn the temporal correlation of ECG signals in each lead, BiGRU with attention mechanism is accordingly employed to further improve the performance of the proposed model. GRU is designed to improve the three-gate structure of LSTM by removing cell state and conflating the forget gate and input gate to an update gate, which has fewer parameters and performs more efficiently than LSTM. GRU can keep the properties of a longer temporal sequence spanning tens to hundreds of time steps in its internal state. In order to use the past and future information, BiGRU is applied that contains a forward GRU layer and a backward GRU layer. At each time step, the output of the forward GRU summarizes temporal features from all previous steps, while the output of the backward GRU summarizes temporal features from all subsequent steps. Finally, the outputs of the BiGRU layer is input into an attention layer that packages features from the context of the current step in both forward and backward directions as the final feature.

The feature classification section processes each of the feature vectors learned as described above. This part contains one fully connected layer. Since it is a multi-label classification problem, the activation function of the second layer is Sigmoid, and its formula is as follows:

$$S(t) = \frac{1}{1 - e^{-x}} \qquad (1)$$

For each training sample, it is obvious that the amount of data in each category is very unbalanced. Therefore, the objective function to be optimized during the model training process is the Focal Loss function [21]. Its formula is as follows:

$$FL(p_t) = -\alpha_t (1 - p_t)^{\gamma} \log(p_t) \qquad (2)$$

where p_t is the probability of ground truth class, α_t is the balance factor that is used to balance the ratio of positive and negative samples, r is the focusing parameter that is greater or equal to zero, $(1-p_t)$ is the modulating factor. The DNNs apply the Adam optimization method and were trained using Keras based on the TensorFlow engine. According to the method described above, the classification in test dataset of PhysioNet database was estimated.

3. Results and discussions

Results showed that the overall score with five-cross validation of training set is 0.543 for our team (DeepHeart), which is acceptable in the existence different noises of different ECG signals in the test dataset.

It is obvious that the advantages of the proposed algorithm include the proposed method of segmentation and balance of data, and the novel model structure. The good performance shows it may have the potential for clinical automatic arrhythmias classification in the future.

4. Conclusion

In this paper, we proposed an attention-based Res-BiGRU model for cardiac arrhythmias classification, which has an end-to-end classification structure. In summary, our model have two main advantages: (1) In the process of feature extraction, residual convolutional modules are introduced to compress and extract spatial features from ECG signals among different leads. Meanwhile, BiGRU is applied to extract temporal features for each lead. Moreover, the attention layer is used for integrating the features in the end. Finally, the spatial and

temporal features extracted from two modules are combined as global features for the final classification process. (2) The model uses raw ECG signals without filtering, proving the capability of the model for handling with noisy signals. In conclusion, the proposed method demonstrated good performance in anti-noise interference, automatic feature extraction and low classification error, making it a potential choice for clinical application in the future.

Acknowledgements

References

[1] Kass, R. E., & Clancy, C. E. (2005). Basis and treatment of cardiac arrhythmias. Springer Science & Business Media, 171.

[2] Afonso, V. X., & Tompkins, W. J. (1995) Detecting ventricular fibrillation. IEEE Engineering in Medicine and Biology Magazine, 14(2), 152-159.

[3] Barro, S., Ruiz, R., Cabello, D., & Mira, J. (1989). Algorithmic sequential decision-making in the frequency domain for life threatening ventricular arrhythmias and imitative artefacts: a diagnostic system. Journal of Biomedical Engineering, 11(4), 320-328.

[4] Minami, K. I., Nakajima, H. & Toyoshima, T. (1999). Real-time discrimination of ventricular tachyarrhythmia with fourier-transform neural network. Journal of Biomedical Engineering, vol. 2(2), 179-185.

[5] Kastor, J. A. (2000). Arrhythmias. WB Saunders Company.

[6] Perez Alday EA, Gu A, Shah A, Robichaux C, Wong AKI, Liu C, Liu F, Rad BA, Elola A, Seyedi S, Li Q, Sharma A, Clifford GD, & Reyna MA. Classification of 12-lead ECGs: the PhysioNet/Computing in Cardiology Challenge 2020. Physiol. Meas. 2020 (Under Review)

[7] Pan, J., & Tompkins, W. J. (1985). A real-time QRS detection algorithm. IEEE Trans. Biomed. Eng, 32(3), 230-236.

[8] Mazomenos, E., Chen, T., Acharyya, A., Bhattacharya, A., & Maharatna, K. (2012). A time-domain morphology and gradient based algorithm for ECG feature extraction. IEEE International Conference on Industrial Technology, 117-122.

[9] Lin, C. H. Frequency-domain features for ECG beat discrimination using grey relational analysis-based classifier. (2008). Computers & Mathematics with Applications, 55(4), 680-690.

[10] Kulkas, A., Huupponen, E., Virkkala, J., Tenhunen, M., Saastamoinen, A., Rauhala, E., & Himanen, S. L. (2009). New tracheal sound feature for apnoea analysis. Medical & Biological Engineering & Computing, 47(4), 405-412.

[11] He, R., Wang, K., Zhao, Na., Liu, Yang., Yuan, Y., Li, Q., & Zhang, H. (2018). Automatic detection of atrial fibrillation based on continuous wavelet transform and 2d convolutional neural networks. Frontiers in Physiology, 9.

[12] Fan, X., Yao, Q., Cai, Y., Miao, F., Sun, F., & Li, Y. (2018). Multiscaled fusion of deep convolutional neural networks for screening atrial fibrillation from single lead short ECG recordings. IEEE Journal of Biomedical and Health Informatics, 22(6), 1744-1753.

[13] Hannun, A., Rajpurkar, P., Haghpanahi, M., Tison, G. H., Bourn, C., Turakhia, M. P., & Ng, A. Y. (2019). Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nature Medicine, 25(1), 65-69.

[14] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 770-778.

[15] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Identity mappings in deep residual networks. In European Conference on Computer Vision. Springer, Cham, 630-645.

[16] Ioffe, S., & Szegedy, C. (2015). Batch Normalization: accelerating deep network training by reducing internal covariate shift. arXiv: Learning.

[17] Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted boltzmann machines. International Conference on Machine Learning, Haifa, Israel, 807-814.

[18] Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2012). Improving neural networks by preventing co-adaptation of feature detectors. arXiv: Neural and Evolutionary Computing.

[19] Ciresan, D., Meier, U., Masci, J., Gambardella, L. M., & Schmidhuber, J. (2011). Flexible, high performance convolutional neural networks for image classification. international joint conference on artificial intelligence, Barcelona, Catalonia, Spain, 1237-1242.

[20] Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. arXiv: Computation and Language.

[21] Li, B., Liu, Y., & Wang, X. (2018). Gradient harmonized single-stage detector. arXiv: Computer Vision and Pattern Recognition.

Address for correspondence.

Henggui Zhang
Room 3.07, Shuster building
The School of Physics and Astronomy
The University of Manchester, Oxford Road
Manchester, M13 9PL, UK
E-mail:henggui.zhang@manchester.ac.uk