Arrhythmia Detection Based on Semantic Segmentation for Multi-lead ECG

Hanshuang Xie[1], Mengna Zheng[1], Huaiyu Zhu[2], Fan Wu[1], Yun Pan[2]

[1]Research and Development Department, Hangzhou Proton Technology Co. Ltd., Hangzhou, China
[2]College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, China

Abstract

In order to detect multi-class arrhythmias with high accuracy using multi-lead electrocardiogram (ECG) signals, we propose an arrhythmia classification method based on semantic segmentation. In our framework, ECG signals are firstly filtered and normalized, and divided into 30-second segments. Then, a convolutional neural network (CNN) with different dilation rates is designed to extract and integrate the multi-scale features of ECG signals. Particularly, we apply squeeze-and-excitation blocks to assign weights to features, and heartbeats are finally classified by Softmax function. Aiming at the problem of class-imbalance, the method of overlap between segments is futher adopted to increase the samples, and probability threshold values are set. We evaluate the performance of the proposed method on five public databases. The precision, sensitivity and F1 score for fusion of ventricular contraction and normal beat (F), supraventricular escape beat (AE) and ventricular escape beat (VE) are all over than 90%. The proposed method combines CNN and semantic segmentation could be helpful for automated ECG diagnosis in clinical practice.

1. Introduction

With the rapid pace of human life, heart disease has become an important issue threatening human health, and most heart patients are accompanied by arrhythmia. Currently, the electrocardiogram (ECG) screening for arrhythmia is usually done by medical staff, which is time consuming. Therefore, accurate automated detection and diagnosis of arrhythmia is of great significance for the prevention, monitoring and treatment of heart disease, as well as improving the efficiency of ECG interpretation by doctors.

At present, there are a variety of methods to classify heartbeats, including expert system, traditional machine learning and deep learning methods. The classification method of expert system [1] extracts relevant features according to the position of R peak, and classify heartbeats by setting threshold. Expert systems usually require professional knowledge and have certain subjective limitations. Support vector machine[2], decision tree [3] and other machine learning need to extract the time domain, frequency domain, morphology and other features of ECG signals [4]. Due to the dependence on these features, it is difficult to ensure the classification accuracy on unknown data.

Deep learning based methods, on the other hand, could autonomously acquire features according to data and find out the internal relationship of ECG signals. In these methods, ECG signals are firstly divided into single [5-6] or three-heartbeat [7-8] segments and then classified through deep learning method. However, most of these methods only retain the morphological characteristics of heartbeats, where rhythm information among heartbeats is lost. Meanwhile, these methods usually rely on the accuracy of heartbeat detection, resulting in performance loss in real clinical scenarios.

In recent years, semantic segmentation technology has been applied in the field of ECG, such as the detection of ECG key waves [9-11] and arrhythmia classification [12] to improve the robustness and generalization of the model. In this paper, we propose a classification method of heartbeats based on semantic segmentation. In our framework, we divide the ECG signals into continuous 30-second segments to introduce the information among the segments and retain the rhythm information of heartbeats. Three parallel convolution blocks are used to construct a convolutional neural network (CNN) model and process the above ECG segments for heartbeat positioning and classification.

We evaluated the performance of the proposed method on five open datasets. The results show that the precision, sensitivity and F1 score are generally good, which improves the arrhythmia classification performance and realizes an automatic classification of multiple arrhythmias.

2. Data preparation

2.1. Database construction

Data used in this paper is from MIT-BIH Arrhythmia

 ISSN: 2325-887X DOI: 10.22489/CinC.2022.173

Database [13], European ST-T Database [14], Sudden Cardiac Death Holter Database, MIT-BIH Long-Term ECG Database, and St Petersburg INCART 12-lead Arrhythmia Database [15]. All heartbeats are divided into eight classes and represented by eight labels. Table 1 shows the number of heartbeats of each type in the reconstructed database.

Table 1. Number of heartbeats

Type	Number	Our label	Data label
Baseline	-	BG	-
Normal	2355998	N	N
Bundle branch block beat	73234	B	L, R, B
Atrial premature contraction	9229	A	A, a, J, S
Premature ventricular contraction	1119305	V	V, r
Fusion of ventricular and normal beat	4593	F	F
Supraventricular escape beat	374	AE	e, j, n
Ventricular escape beat	122	VE	E
Question	31677	Q	/, f, q, ?, 0

2.2. Preprocessing

Since the sampling rate of each database is different, we resample ECG signals to 200 Hz. The signals are then processed by a bandpass filter of 0.67-40 Hz. The filtered ECG signals are normalized using the min-max scaling technique as shown in Equation (1):

$$x' = \frac{x - min(x)}{max(x) - min(x)} \tag{1}$$

In addition, the position labels of the databases should be converted into the position values corresponding to the 200 Hz signals. The conversion formula is shown in Equation (2), where $point_{new}$ is the label position after resampling, $point_{old}$ is the label position before resampling, and f_{old} is the original sampling rate.

$$point_{new} = point_{old} / f_{old} \times 200 \tag{2}$$

2.3. ECG segmentation

Usually, the ECG signal is divided by single or three heartbeats, which leads to the loss of information among heartbeats. However, if an ECG signal is too long, there will be data redundancy issue, which results in excessive computation. Therefore, on the basis of objectively reflecting the basic heart rate of the signal and retaining the rhythm information of heartbeats, we choose to segment ECG signals with a time length of 30 seconds.

We divided ECG signals into 9 regions, including 8 kind of cardiac beat regions and a baseline region, where BG represents baseline region, and the width of the cardiac beat region is 0.2 second, as shown in Figure 1.

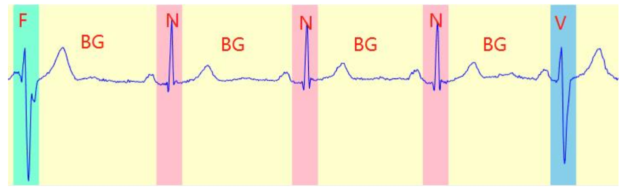

Figure 1. Example of ECG segmentation

2.4. Data balancing

Due to the very small number of some types such as AE, VE, and F, we increase the number of these types by an overlap window. And if there are only normal cardiac beats in a 30-second segment, then this segment will be discarded. Table 2 shows the number of each heart type after data balancing.

Table 2. Number of heartbeats after data balancing

Type	Number
BG	-
N	905679
B	62861
A	11273
V	140508
F	18433
AE	3518
VE	1248
Q	27322

3. Model and train

We use the CNN model structure in [16] for reference and make some modifications. The model structure is shown in Figure 3. The input layer of the model are followed by three parallel convolutional blocks, denoted as Block1, Block2 and Block3 respectively. Each convolutional block is composed of three layers of neural networks, each layer of neural network consists of a 1D convolutional layer, a BN layer and a Max pooling layer. BN layer is used to speed up the training process and prevent overfitting. The three blocks adopt different expansion rates respectively.

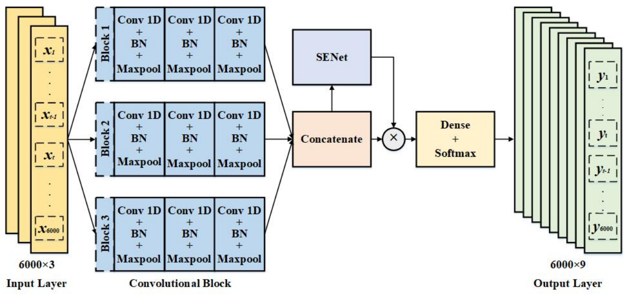

Figure 2. Framework of the proposed method

There are different receptive fields to obtain multi-scale information and improve the performance of the model. The output results of Block1, Block2, and Block3 are concatenated for feature fusion, and fed into SENet to obtain the weight result to get the weighted fusion features. Finally, the result is output through a fully connected layer with Softmax activation function. Since the temporal dimension of each ECG signal is 6000, the output has the same dimension of size 6000. The output sequence represents the probability distribution of ECG data over different types as $K = \{0, 1, 2, 3, 4, 5, 6, 7, 8\}$ for BG, N, B, A, V, F, AE, VE, and Q, respectively.

The labels are adopted one-hot encoding for model training and test, as shown in Figure 3. The model with the minimum loss value is saved in the training process. If the loss value does not decrease after m epochs, the learning rate will be dynamically adjusted. When the learning rate reaches the set minimum value and the loss value does not decrease any more, the training will be finished and the best saved model will be taken as the final trained model.

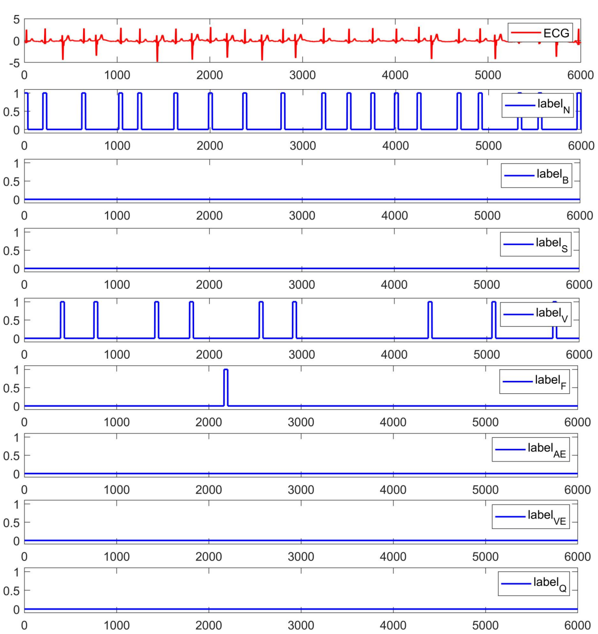

Figure 3. One-hot encoding for data labeling

4. Results and discussion

The data set is a matrix of $M \times 6000 \times 3$, where M represents the number of ECG segments, 6000 represents the length of each segment, and 3 is the number of ECG leads. It was further divided into a training set and a test set with 70% and 30%, respectively. Test data was input into the trained model to get the prediction probability value. The position of the heartbeat type is obtained by setting the threshold.

Due to the data imbalance of different types, the categories with small sample size will contain too few features, and it is difficult to extract rules from them. It is also easy to overfit the samples with a large proportion, resulting in poor generalization ability and low accuracy of the model. When the trained model is applied to the test data, categories with a small proportion are likely to be misclassified if the type of heartbeat is classified with the maximum probability value.

To solve this problem, we set a threshold for classification, and set a lower probability threshold for categories with a small number. As long as the probability value of a category reaches its threshold, it will be assigned to the corresponding category, so as to improve the accuracy of classification.

In this experiment, the number of escape beats is small, so the probability threshold of escape beats is set as 0.4. Similarly, the number of fusion beats is larger than that of escape beats, but less than that of other heart beats. The probability threshold of fusion beats is set as 0.55, and the probability value of other data is set as 0.65. The metrics of the model on the test set are shown in Table 3.

Table 3. Results of classification

Type	Precision (%)	Se (%)	F1 (%)
N	96.14%	95.62%	95.88%
B	94.46%	93.83%	94.14%
A	88.58%	84.19%	86.33%
V	90.86%	89.16%	90.00%
F	90.86%	91.79%	91.32%
AE	97.61%	91.82%	94.63%
VE	98.57%	96.50%	97.53%
Q	94.71%	94.07%	94.39%

It can be seen from Table 2 that the minimum value of precision is 88.58% and the maximum value of precision is 98.57%, the lowest sensitivity is 84.19% and the highest sensitivity is 96.50%, the lowest F1 score is 86.33% and the highest F1 score is 97.53%, indicating that the classification performance of the model is generally good. Besides, the precision, sensitivity and F1 score of fusion of ventricular and normal beat (F), supraventricular escape beat (AE) and ventricular escape beat (VE) reached more than 90%, especially the precision, sensitivity and F1 score of VE reached more than 95%, 98.57%, 96.50 and 97.53%, respectively. It shows that our method also provides promising results on the imbalanced dataset.

5. Conclusion

In this paper, an arrhythmia classification method based on CNN network with semantic segmentation was proposed. The application of semantic segmentation avoids the dependence of arrhythmia classification on QRS detection algorithm accuracy, and could accurately identify the location of heartbeat, as well as automatically divide heartbeats into 8 types (N, F, B, A, V, AE, VE and Q). The influence of data imbalance was decreased by using the probabilistic threshold and the model achieved high performance parameters, i.e., precision, sensitivity and F1 score. Different from the methods based on single or three-heartbeat segments, we segmented the ECG signals into segments with the length of 30 seconds, which retained the basic heart rate and rhythm information among heartbeats. Five open source databases were used to evaluate our model, and the experimental results showed that our ECG classification model is an effective tool for arrhythmia detection.

In the future, we plan to expand our database. In order to examine the robustness of the classification system, we would use one database for training and the other for testing, so that the data are completely independent and the generalization ability of the model could be further checked. Another worthy exploration is to apply the semantic segmentation method on the detection of waves in ECG signals.

Acknowledgments

This work was supported by Zhejiang Provincial Key Research and Development Program of China (2021C03027).

References

[1] Y. H. Hu, S. Palreddy, and W. J. Tompkins, “A patient-adaptable ECG beat classifier using a mixture of experts approach,” *IEEE Transactions on Biomedical Engineering*, vol. 44, no. 9, pp. 891–900, Sep. 1997.

[2] A. Mustaqeem, S. M. Anwar, and M. Majid, “Multiclass classification of cardiac arrhythmia using improved feature selection and SVM invariants,” *Computational and Mathematical Methods in Medicine*, pp. 1–10, Mar. 2018.

[3] E. Alickovic, and A. Subasi, “Medical decision support system for diagnosis of cardiovascular diseases using DWT and random forests classifier,” *J. MED. SYST.*, vol. 40, no. 4, pp. 108, Feb. 2016.

[4] H. M. Shirin, A. Mohsen, and S. Reza, “Automated method for discrimination of arrhythmias using time, frequency, and nonlinear features of electrocardiogram signals,” *Sensors*, vol. 18, no. 7, pp. 2090, Jun. 2018.

[5] S. S. Xu, M. M. Mak, and C. C. Cheung, “Towards end-to-end ECG classification with raw signal extraction and deep neural networks,” *Biomedical Signal Processing and Control*, vol. 23, no. 4, pp. 1574-1584, Jul. 2019.

[6] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, M. Adam et al., “A deep convolutional neural network model to classify heartbeats,” *Computers in Biology and Medicine*, vol. 89, pp. 389-396, Oct. 2017.

[7] H. W. Liang. T. Y. Yan, L. Wei, X. Lu, X. X. Chao et al., “Three-heartbeat multilead ECG recognition method for arrhythmia classification,” *IEEE Access*, vol. 10, pp. 44046-44061, Apr. 2022.

[8] M. Jangra, S. K. Dhull, and K. K. Singh, “ECG arrhythmia classification using modified visual geometry group network (mVGGNet),” *Journal of Intelligent & Fuzzy Systems*, vol. 38, no. 3, pp. 3151-3165, Mar. 2020.

[9] K. Duraj, N. Piaseczna, P. Kostka, and E. Tkacz, “Semantic segmentation of 12-lead ECG using 1D residual U-net with squeeze-excitation blocks,” *Applied Sciences*, vol. 12, no. 7, pp. 3332, Mar. 2022.

[10] A. N. Londhe, and M. Atulkar, “Semantic segmentation of ECG waves using hybrid channel-mix convolutional and bidirectional LSTM,” *Biomedical Signal Processing and Controls*, vol. 63, pp. 102162, Sep. 2021.

[11] M. Aman, H. Srinidhi, G. M. Siddesh, K. G. Srinivasa and N. T. Sanjay, “Supervised ECG wave segmentation using convolutional LSTM,” *ICT Express*, vol. 6, no. 3, pp. 166-169, Sep. 2020.

[12] S. L. Oh, E. Y. Ng, R. S. Tan, and U. R. Acharya, “Automated beat-wise arrhythmia diagnosis using modified U-net on extended electrocardiographic recordings with heterogeneous arrhythmia types,” *Computers in Biology and Medicine*, vol. 105, no. 3, pp. 92-101, Feb. 2019.

[13] G. B. Moody, and R. G. Mark, “The impact of the MIT-BIH arrhythmia database,” *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45-50, Feb. 2001.

[14] A. Taddei, G. Distante, M. Emdin, P. Pisani, G. B. Moody et al., “The European ST-T Database: standard for evaluating systems for the analysis of ST-T changes in ambulatory electrocardiography,” *European Heart Journal*, vol. 13, pp. 1164-1172, Sep. 1992.

[15] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov et al., “PhysioBank, physioToolkit, and physioNet: components of a new research resource for complex physiologic signals,” *Circulation*, vol. 101, no. 23, pp. 215-220, Jun. 2000.

[16] W. J. Cai, and D. Q. Hu, “QRS complex detection using novel deep learning neural networks,” *IEEE Access*, vol. 8, pp. 97082-97089, May. 2020.

Address for correspondence:

Hanshuang Xie
Room 405, Chuangxin Building, East Software Park, No.90 Wensan Road, Xihu District, Hangzhou, Zhejiang, China
hanshuang.xie@protontek.com