Fuzzy Systems and Data Mining IX
A.J. Tallón-Ballesteros and R. Beltrán-Barba (Eds.)

doi:10.3233/FAIA231104

Multi-Class Arrhythmia Classification and R-Peak Detection Method of ECG Signal Based on One-Dimensional U-Net with Skip-Connection and Data Augmentation

Yamin LIU[a], Hanshuang XIE[a,1], Jiayi YAN[a], Zhen WANG[a], Mengna ZHENG[a], Fan WU[a] and Yun PAN[b]

[a] *Research and Development Department, Hangzhou Proton Technology Co. Ltd., Hangzhou, China*

[b] *College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, China*

Abstract. Automatic arrhythmia analysis techniques and QRS detection provide convenience for the prevention and diagnosis of cardiac disease. The existing studies generally study arrhythmia classification and QRS recognition separately, which requires two different models and may result in time and resource wasting. To realize the goal of arrhythmia classification and R-peak detection of electrocardiograms at the same time, we proposed a method for multi-class arrhythmia classification and R-peak detection method based on one-dimensional U-net with skip-connection and data augmentation. First, the ECG signals were preprocessed by filtering and segmentation, and the ECG annotations were also processed into pixel labels of equal length. Then, we applied data augmentation techniques such as changing gain, adding noise, and flipping signals up and down to increase the diversity of the data. Finally, a modified one-dimensional U-net with skip-connection layers was built to adaptively extract deep features and to detect the arrhythmia type and R peaks of ECG at the same time. We set up an 8-category experiment using five publicly available datasets, and the experimental results show that the macro average F1 value is 94.57%, which is over 4.3% and 3% higher than that of original U-net and the skip-connection U-net without data augmentation, respectively. Meanwhile, the F1 of R-peak detection is 99.64%.

Keywords. ECG, arrhythmia classification, R-peak detection, 1-D U-net, data augmentation

1. Introduction

Cardiovascular diseases are one of the major diseases that threaten human life [1]. Therefore, a rapid and accurate diagnosis of arrhythmia is of great importance for the prevention and treatment of heart disease. The electrocardiogram (ECG) has been widely used for cardiovascular disease diagnosis as physiological signals generated by the heart's excitement.

In recent years, with the rapid development of artificial intelligence techniques, researches on arrhythmia classification methods based on machine learning continue to emerge. Compared to the limitations of traditional machine learning [2-4], deep learning always has better performance. Xu et al. [5] first divided ECG signals into single heartbeat and then use deep neural network (DNN) for end-to-end arrhythmia classification. Similarly, Acharya, et al [6] utilized a deep convolutional neural network (CNN) to automatically identify different categories of heartbeats in ECG signals. The division of a single heartbeat is simple, but it may cause the loss of the rhythm

[1] Corresponding Author, Hanshuang XIE, Research and Development Department, Hangzhou Proton Technology Co. Ltd., Hangzhou, China; E-mail: hanshuang.xie@protontek.com.

information of the preceding and following heartbeats. To avoid this problem, L. -H et al. [7] proposed a novel classification method for arrhythmia that uses three-heartbeat ECG signals based on one-dimensional CNN. The rationale for using recurrent neural network (RNN) in ECG analysis lies on its ability to interpret the idea of time. They can hence learn complex temporal dynamics within time-varying data. In this way, Lu et al. [8] proposed a gated recurrent unit (GRU) and decision tree fusion model to explore the problem of arrhythmia recognition and to improve the credibility of deep learning methods. Luo et al. [9] proposed an arrhythmia classification method based on a hybrid convolutional recurrent neural network (CRNN) for the time-series signal of ECG. The above methods are used to identify ECG arrhythmia from a classification point of view. Importantly, the prediction models are based on ECG signal readouts with single arrhythmic conditions. However, there are always contain mixed arrhythmic conditions in real-life. Hence, Oh et al. [10] proposed an improved U-net model, which is usually used in the semantic segmentation domain to perform beat-wise analysis of ECG segments of different lengths. Although the above methods provide some solutions for ECG arrhythmia diagnosis, there are still problems. For example, the input ECG signal of these models is usually one or three heartbeats. Even in [10], the ECG signals segment is only 2.7 seconds, which contains about three to five heartbeats. In the absence of fragmented rhythm information, certain arrhythmia categories may be difficult to classify, such as normal beats and atrial premature contraction.

To solve such problem mentioned above, we propose an ECG arrhythmia multiple classification and R-peak detection method based on data augmentation and U-net with skip-connection, for convenience, we call it SDA U-net. In detail, we split the ECG signal into consecutive 30-second segments, introducing effective information among the segments while retaining the rhythmic information of the heartbeat. We verify the proposed method on five public datasets and the results show that the method has good performance in terms of precision, recall, and F1 score.

2. Proposed Method

2.1. Data Description

The ECG arrhythmia recordings used in this paper are obtained from 5 public databases: St Petersburg INCART 12-lead Arrhythmia Database (INCART12) [11], MIT-BIH Arrhythmia Database (MIT-BIH) [12], Sudden Cardiac Death Holter Database (SCDHD) [13], MIT-BIH Long-Term ECG Database (MIT-BIH LTDB) [12] and European ST-T Database (ESTT) [14].

We use the heartbeats processing methods in [15] as reference, the 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	905679	N	N
Bundle branch block beat	62861	B	L, R, B
Atrial premature contraction	11273	A	A, a, J, S
Premature ventricular contraction	140508	V	V, r
Fusion of ventricular and normal beat	18433	F	F
Supraventricular escape beat	3518	AE	E, j, n
Ventricular escape beat	1248	VE	E
Question	27322	Q	/, f, q, ?, 0

2.2. Data Preprocessing

First, due to the sampling rate of the 5 datasets is different, we resampled all signals to 200 Hz. The optimal lead of the multi-lead data is selected according to the criterion of minimum abnormal interference and the signal is as smooth as possible. For example, the lead of a raw ECG signal in MIT-BIH is {V5, II} shown in the left two blue signals in Figure 1, after the optimal lead selection algorithm, the sorted lead is {II, V5} shown in the right two red signals in Figure 1, in this case, the lead II is our selected signal. Second, we segment ECG signals into 30 seconds as one data sample. At the same time, ECG labels are also processed into pixel labels of equal length as shown in Figure 2. Specifically, we take the labeled peak R as the base point, then take the time before and after 0.1 seconds of R peak as the label of this beat. BG, N, B, A, V, F, AE, VE, Q are mapped to {0, 1, 2, 3, 4, 5, 6, 7, 8} as shown in Figure 2. During model training, the pixel labels are converted into one-hot labels. Third, we apply a 7-level wavelet transform with db6 to remove the electromyographic noise information and baseline drift of ECG data. Figure 3 shows the ECG signal comparison before and after filtering. Finally, the data set is divided according to the ratio of approximately 8:1:1 for the training set, validation set, and test set.

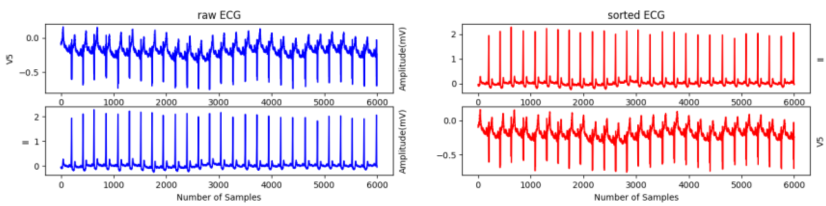

Figure 1. The leads of ECG signal in MIT-BIH are sorted before and after

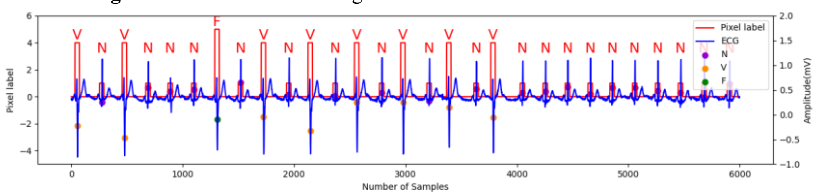

Figure 2. The ECG signal and pixel label

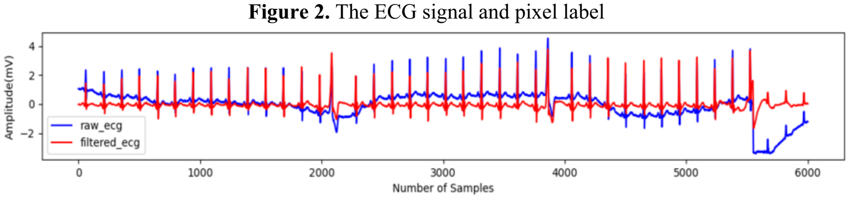

Figure 3. The ECG comparison before and after filtering, blue: before filtering, red: after filtering

2.3. Data Augmentation

To enrich the information of ECG data, we apply data augmentation before model training. Specifically, the data augmentation includes methods such as changing gain, adding noise, and flipping signals up and down as shown in Figure 4. The purpose of changing the gain of ECG signals is to avoid missed R peaks of small amplitude. The reason of adding noise is to enhance the robustness of the model. The signal is flipped up and down to enrich the data distribution of the data set.

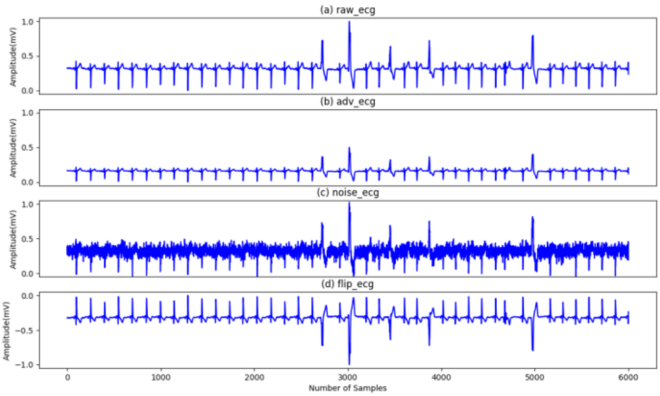

Figure 4. The ECG after data augmentation, (a) raw ECG, (b) change gain, (c) add noise and (d) flip signals

2.4. Model Build and Train

U-net has been proved its superior performance in images segmentation processing [16]. We use U-net as the base network and make some improvements to get better results. Unlike 2-dimensional images, ECG signals are 1-D signals, so we build U-net with a 1-D layer.

The model structure diagram is shown in Figure 5. As we can see, we add eight skip-connections (black dotted lines in Figure 5) from the shallow layer to the deep layer of the model, and the skip-connections are directly added after the corresponding convolution layer. As we all know, spatial domain information is very important for segmentation tasks. However, the pooling layers in the encoder part of the U-net have reduced the feature map resolution to a very small size, which is not conductive to accuracy segmentation mask. Skips-connection can be used to introduce shallow convolution layer features with high resolution to reduce the information loss for a better segmentation mask. The input and output of the build model are both 1*6000. The model with the minimum loss during training is saved as the best-trained model for the next test process.

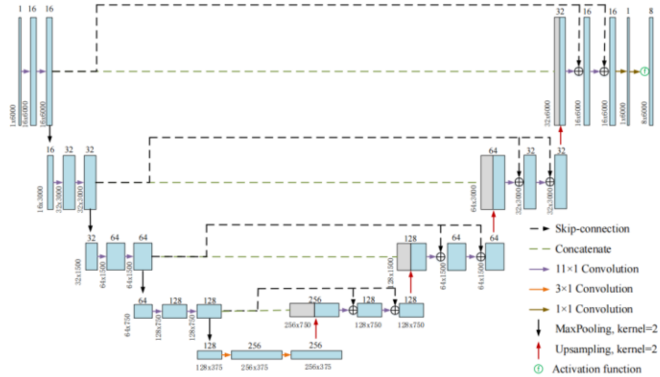

Figure 5. The structure of 1-D SDA U-net

3. Result and Discussion

The test set is put into the trained model to judge model performance. The output of the trained model is the probability corresponding to the input, as is shown in Figure 6, and we determine the label according to the set threshold, which is set to 0.5. Meanwhile, in order to detect the R-peak, we add the prediction probability of each category of heartbeat as the probability of R-peak and also use the threshold to determine whether it is R-peak. In this way, we can obtain both arrhythmia classification results and R-peak detection results.

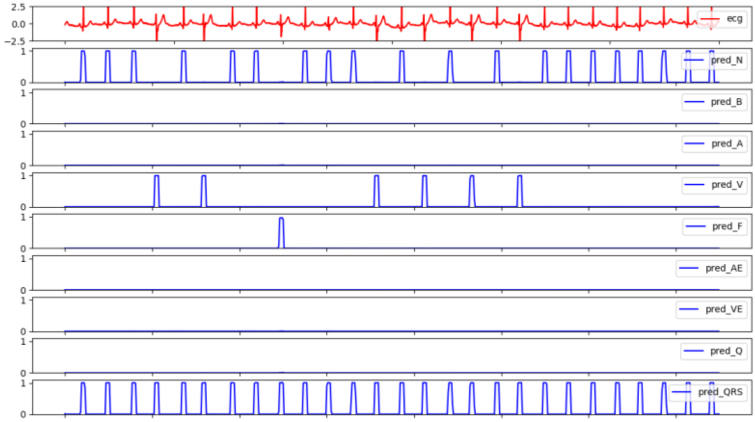

Figure 6. The red line represents the ECG signal, and the blue line represents the one-hot labels of different categories

The metrics of the proposed method on the test set are shown in Table 2. We also provide the comparison metrics of the original U-net and skip-connection U-net without data augmentation. From Table 2, we can see that all categories except Q showed significant improvement, especially the category of AE and VE using SDA U-net (our method). Combined with data volume, we think data augmentation can improve the recognition effect of categories with little data volume by enriching data information. In conclusion, the macro average F1 of SDA U-net is over 4.3% and 3% higher than that of U-net and U-net with skip-connection, respectively. The R-peak detection (QRS) of the SDA U-net also outperforms U-net and skip-connection U-net without data augmentation. Thus, we can draw the conclusion that the skip-connection in U-net could increase the performance of U-net to some extent. In our opinion, this may be because the skip-connection could reduce the loss of feature information extracted from shallow convolution layers. What’ s more, the data augmentation could enrich the data distribution for better model training.

Table 2. The comparison metrics of different methods (unit: %)

Type	U-net			U-net with skip-connection			SDA U-net (Our method)		
	precision	recall	F1	precision	recall	F1	precision	recall	F1
N	96.38	97.51	96.94	97.71	96.61	97.16	97.01	97.31	97.16
B	92.48	91.02	91.75	98.01	97.24	97.63	98.21	97.97	98.09
A	75.79	70.11	72.84	85.64	77.27	81.24	87.78	80.97	84.24
V	95.66	91.45	93.51	96.54	90.55	93.44	94.09	92.99	93.54
F	88.54	77.21	82.48	92.82	83.79	88.07	92.97	91.36	92.16
AE	89.64	90.51	90.07	95.95	95.29	95.62	96.42	96.46	96.44
VE	90.04	90.11	90.07	93.81	92.69	93.24	96.38	95.92	96.15
Q	94.99	96.29	95.64	94.75	93.02	93.87	94.84	93.12	93.73
QRS	99.13	99.35	99.24	99.78	99.08	99.43	99.91	99.36	99.64
Macro average	91.41	89.28	90.28	95.01	91.72	93.30	95.29	93.94	94.57

4. Conclusion

In this paper, an arrhythmia multiple classification and R-peak detection method based on SDA U-net with semantic segmentation was proposed. The application of semantic segmentation not only introduces effective information among the ECG segments but also retains the rhythmic information of the heartbeat. What's more, data augmentation techniques also make the proposed method more robust and generalization performance, which has been verified on five public datasets.

The validation of the methods presented in this paper is based on inner-patient analysis, and we will try more methods to explore more robust models for different databases in the future.

References

[1] The American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart Disease and Stroke Statics 2009 Update: A Report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee.Circulation, 2009, 119.

[2] Kohli N, Verma N K. SVM based Methods Optimized by Feature Selection for Arrhythmia Classification in ECG. International Conference on Computer and Communication Technology (ICCCT). 2010: 486-490.

[3] Lee G B, Cho B J. EMD based Cardiac Arrhythmia Classification using Multi-class SVM. Journal of the Korea Institute of Information and Communication Engineering. 2010. 14(1).

[4] Marasović T, Papić V. A Comparative Study of FFT, DCT, and DWT for Efficient Arrhytmia Classification in RP-RF Framework. International Journal of E-Health and Medical Communications, 2018; 9(1):15.

[5] Xu SS et al. Towards End-to-End ECG Classification with Raw Signal Extraction and Deep Neural Networks. IEEE Journal of Biomedical & Health Informatics, 2018; 23(4):1574-1584.

[6] Rajendra Acharya U et al. A deep convolutional neural network model to classify heartbeats. Computers in Biology and Medicine, 2017 Oct; 89:389-396.

[7] Wang LH et al. Three-Heartbeat Multilead ECG Recognition Method for Arrhythmia Classification. IEEE Access, 2022, 10:44046-44061.

[8] Lu P, Zhang Y, Zhou B et al. Identification of Arrhythmia by Using a Decision Tree and Gated Network Fusion Model. Computational and Mathematical Methods in Medicine, 2021.

[9] Luo X, Yang L, Cai H, et al. Multi-classification of Arrhythmias Using a HCRNet on Imbalanced ECG Datasets. Computer Methods and Programs in Biomedicine, 2021, 208(9):106258.

[10] Oh, Shu Lih., et al. Automated beat-wise arrhythmia diagnosis using modified U-net on extended electrocardiographic recordings with heterogeneous arrhythmia types. Computers in biology and medicine. 2019: 105:92-101.

[11] Goldberger AL, Amaral LA, Glass L et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation. 2000 Jun 13; 101(23):E215-20.

[12] Moody GB, Mark RG. The impact of the MIT-BIH Arrhythmia Database. IEEE Engineering in Medicine and Biology Magazine. 2001; 20(3):45-50.

[13] Greenwald SD. The development and analysis of a ventricular fibrillation detector. Massachusetts Institute of Technology. 2015.

[14] Taddei A, Distante G, Emdin M 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(9):1164-1172.

[15] Hanshuang Xie, Mengna Zheng, Huaiyu Zhu, Fan Wu, Yun Pan. Arrhythmia Detection Based on Semantic Segmentation for Multi-lead ECG. 2022 Computing in Cardiology (CinC). 2022: 1-4.

[16] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham. 2015, 9351.