MTFNet: A Morphological and Temporal Features Network for multiple leads ECG Classification

Lebing Pan, Weibai Pan, Mengxue Li, Yuxia Guan, Ying An

Central South University , Changsha , China

Abstract

Cardiac abnormalities are the main cause of death and ECG is a key diagnostic tool to assess the cardiac condition of a patient. The goal of the 2021 PhysioNet/CinC Challenge was to develop algorithms to diagnose multiple cardiac abnormalities from reduced-lead ECG. In this work, We describe a new model that extracts morphological and temporal features to classify each ECG sequence into 26 cardiac abnormality classes. We use multiple double conv blocks with multiple layers of small convolution kernels to extract morphological features, each layer has similar hyperparameters. Then the Bidirectional LSTM is used to obtain the temporal characteristics. Our entry to the 2021 PhysioNet/CinC Challenge, using the official generalized weighted accuracy metric for evaluation, Our classifiers received scores of 0.61, 0.59, 0.60, 0.60, and 0.59 (ranked 14th, 12th, 12th, 13th, and 11th out of 60 teams) for the 12-lead, 6-lead, 4-lead, 3-lead, and 2-lead versions of the hidden validation set.

1. Introduction

ECG is a fundamental tool in the everyday practice of clinical medicine, that is pivotal for diagnosing a wide spectrum of cardiac abnormalities[1]. Automatic ECG interpretation algorithms as diagnosis support systems have become increasingly important in the clinical ECG workflow and promise large reliefs for the medical personnel. In the PhysioNet/Computing in Cardiology Challenge 2021, participants were to develop an algorithm that can classify cardiac abnormalities from either twelve-lead, six-lead, four-lead,three-lead, and two-lead ECGs[2–4]. Deep learning methods have recently been applied to this task and have achieved excellent performance. Compared to traditional methods, deep learning-based approaches can automatically learn to predict the information feature representation of cardiac abnormalities in an end-to-end manner. Research scholars widely use convolutional neural networks(CNN)[5, 6], recurrent neural networks (RNN)[7, 8], and hybrid networks[9]. At present, efficient hybrid models can provide more distinctive features from ECG signals [10]. In this article, we propose a new hybrid network classification method for the PhysioNet/Computing in Cardiology Challenge 2021.

2. Method

The objective of this study is to diagnose 26 types of cardiac abnormalities from 12-lead, 6-lead(I, II, III, aVR, aVL, aVF), 4-lead(I, II, II, V2), 3-lead(I, II, V2) and 2-lead (I and II) ECG recordings.The data set of this research includes 6 different data sets, namely CPSC Database and CPSC-Extra Database, INCART Database, PTB and PTB-XL Database, The Georgia 12-lead ECG Challenge (G12EC) Database, Augmented Undisclosed Database, Chapman-Shaoxing and Ningbo Database, a total of more than 88,000 ECG recordings, each recording has one or more cardiac lables.

2.1. Pre-processing

In order to ensure the consistency of the data, the data set with a sampling frequency higher than 500Hz is downsampled, so that the sampling frequency of all data sets are unified to 500Hz. We padded zeros to each record which is less than 5000 sample points and truncated the record that exceeds 5000 sample points so that the length of all records is 5000 sample points.

2.2. Overall Architecture

In this competition, we proposed an end-to-end ECG signal multi-label classification system. The system receives the pre-processed ECG segments at one end and outputs the multi-label classification decisions at the other end. The proposed model, illustrated in Fig.1, is constructed on a combined architecture of five Double Conv blocks, followed by four layers of bidirectional LSTM module, and with a fully connected layer at the end to produce the classification decision.

Since the ECG signal is composed of a sequence of waves on the one hand, on the other hand, there is a tempo-

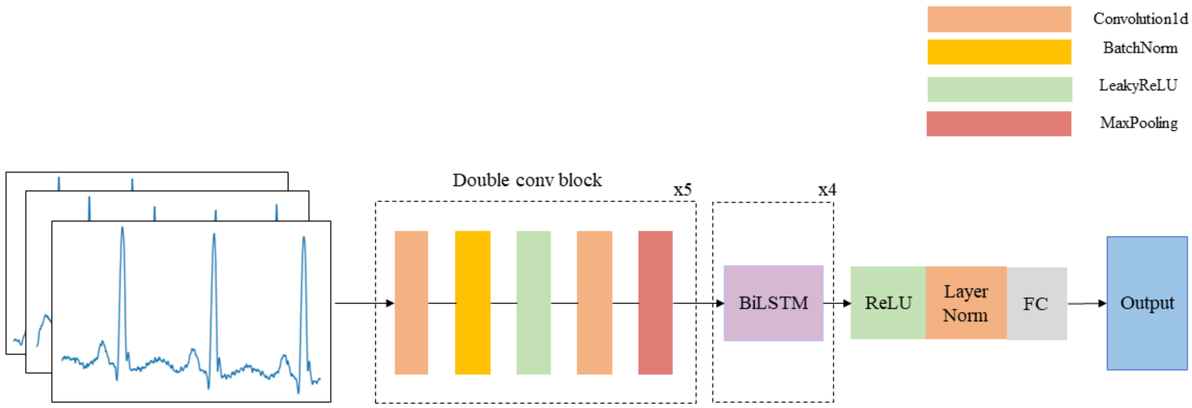

Figure 1. The architecture of MTFNet.

ral relationship in the heartbeat and between the heartbeats. Therefore, in the overall architecture, we use five Double Conv blocks to extract the high-level features from different waves at first and then use the LSTM module to learn the temporal information from the processed ECG signals.

Table 1, Table 2 and Table 3 show the detailed parameters of MTFNet.

Double conv	Input size	Output size	Kernel size	Padding
1	12/6/4/3/2	64	3	1
2-5	64	64	3	1

Table 1. Conv Blocks settings.

Bi-LSTM layer conv	Input size	Hidden size	Dropout size
1-4	64	64	0.2

Table 2. Bi-LSTM layer settings.

layer	Input	Output
LayerNorm	64	128
Fully connected	128	26

Table 3. LayerNorm and Fully connected layers setting.

2.2.1. Morphological Features

Specifically, the Double Conv block consists of a 2-layer 1D-CNN module with a batch normalization layer in between, and a leaky version of ReLU is applied after each 1D-CNN as the activation layer while the max-pooling layer is placed at the end of the Double Conv block to extract high-level features. We have made some attempts when setting the number of channels in 1D-CNN(Tabel 1), and finally set it in all 1D-CNN to 64.

Channels of 1D-CNN	12-lead Challenge score	3-lead Challenge score	2-lead Challenge score
12	0.6665	0.6616	0.6648
32	0.7005	0.6901	0.6872
64	**0.7095**	**0.6980**	**0.6882**

Table 4. Model performance for different channel numbers.

As can be seen in Table 4, increasing the number of channels of all 1D-CNN modules can significantly improve the effect of the model in different leads, the model score increased from 0.6665 to 0.7095 in 12-lead, 0.6616 to 0.6980 in 3- lead, and 0.6648 to 0.6882 in 2-lead.Since the number of channels in 1D-CNN increases, the Double Conv block can extract more high-level features, the Bi-LSTM module also receives more information, and the data compressed in the pooling layer is compensated in dimensionality.

2.2.2. Temporal Features

The hidden size in LSTM is set to 64, and dropped 20% of its paraments to reduce the effect of overfitting problem.When designing the model, we found that as the number of Bi-LSTM layers increased, the score of the model would be significantly improved (Table 5). When the number of Bi-LSTM layers was increased from 1 to 3, the scores were 0.6403, 0.6934, and 0.7069 respectively, which showed a significant improvement. And when the number of Bi-LSTM layers was increased to 4, although the score was not greatly raised, the convergence speed of the training process was much faster than that of the 3-layer Bi-LSTM. For Bi-LSTM with more than 5 layers, introducing more parameters will not bring more effect improvement, but will slow down the training process of the model.While Bi-LSTM can extract the temporal features from the ECG signal effectively, increasing the number of layers of Bi-LSTM can also increase the quantities of parameters of the model, making the model better fit the classification task.

Layers of Bi-LSTM	12-lead Challenge score
1	0.6463
2	0.6934
3	0.7069
4	**0.7095**

Table 5. Model performance for different Bi-LSTM layers.

2.3. Implementation Details

This model is implemented on the GPU version of the PyTorch framework. The training process adopted the Binary Cross Entropy loss and used AdamW as the optimizer. The initial learning rate was set to 0.0001, and the cosine annealing algorithm was used to dynamically adjust the learning rate during each epoch of training. Data were divided into training set and validation set at 8:2 and the models were trained with a batchsize of 128. There are 200 epochs for training in total, at the end of each epoch the score of the model on the validation set was calculated to save the optimal model.

3. Result

Challenge scores for our final selected model on the public training set repeated scoring on the hidden validation set, and one-time scoring on the hidden test set as well as the ranking on the hidden test set are shown in Tabel 6.

As can be seen in Table 6, there is a big gap between the score in our validation set and the official hidden validation set. The score of 12-leads dropped from 0.7095 to 0.602. This is because we actually trained for 200 epochs, but multiple submission failures made us changed our training epochs to 150 epochs at the final upload. On the other hand, during the training process we artificially divide the data set by 8:2 according to different data sources, but in the uploaded code, the entire data set is randomly divided, therefore the optimal model cannot be saved during training and validation process.

Leads	Training	Validation	Test	Ranking
12	0.7095	0.602		
6	0.6885	0.592		
4	0.7016	0.592		
3	0.6980	0.600		
2	0.6882	0.593		

Table 6. The resulte of the PhysioNet/CinC challenge.

4. Discussion and Conclusions

We developed a new neural network framework based on multiple double conv blocks and BiLSTM to accurately classify 26 cardiac abnormalities from reduced-lead ECG. Firstly, the competition evaluation metrics demonstrate that the double conv blocks can learn morphological features in different leads well from the ECG and integrate the temporal-morphological relationships by using the superimposed Bi-LSTM. Second, the overall structure of our model is very simple, and the number of parameters involved is small, so even with a large amount of competition data, our model can still run faster. And considering the heterogeneity of the competition data and the imbalance of the data, the algorithm should have good generalization ability.

However, our model does not explain what kind of ECG morphology it uses to make classification judgments. In future work, we might consider using attention mechanisms [11] or multimodal data such as gender, age, etc. to obtain better interpretability. In addition, we only intercepted all ECG data for 10 s during preprocessing, and the chance of data interception may lead to the extraction of ECG signals with insignificant morphological features and thus affect the performance of the model. We also noted that the AUROC of nonspecific intracardiac conduction disorders was lower in 26 categories compared with other categories, probably because the mild prolongation of QRS waves caused by it was so small that it was masked by noise.

In conclusion, we propose an open-source neural network framework to extract morphological and temporal features from reduced leads and to classify multiple cardiac abnormalities. And our model's challenge scores show that 2-lead can also provide helpful characteristic information as the 12-lead, which offers a theoretical basis for portable devices.

References

[1] Hannun AY, Rajpurkar P, Haghpanahi M, Tison GH, Bourn C, Turakhia MP, Ng AY. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nature medicine 2019; 25(1):65–69.

[2] Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. Circulation 2000;101(23):e215–e220.

[3] 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. Physiological Measurement 2020;41.

[4] Reyna MA, Sadr N, Perez Alday EA, Gu A, Shah A, Robichaux C, Rad BA, Elola A, Seyedi S, Ansari S, Li Q, Sharma A, Clifford GD. Will Two Do? Varying Dimensions in Electrocardiography: the PhysioNet/Computing in Cardiology Challenge 2021. Computing in Cardiology 2021;48:1–4.

[5] Acharya UR, Fujita H, Lih OS, Hagiwara Y, Tan JH, Adam M. Automated detection of arrhythmias using different intervals of tachycardia ecg segments with convolutional neural network. INFORMATION SCIENCES 2017;405:81–90.

[6] Acharya UR, Fujita H, Shu LO, Hagiwara Y, Adam M. Application of deep convolutional neural network for automated detection of myocardial infarction using ecg signals. Information Sciences 2017;415.

[7] Oliver F, Alex S, Murtadha K, Ru ST, Hamido F, Rajendra AU. Automated detection of atrial fibrillation using long short-term memory network with rr interval signals. Computers in Biology and Medicine 2018; 102:S0010482518301847–.

[8] Xu S, Li J, Liu K, Wu L. A parallel gru recurrent network model and its application to multi-channel time-varying signal classification. IEEE Access 2019;PP(99):1–1.

[9] Hong S, Xiao C, Ma T, Li H, Sun J. Mina: Multilevel knowledge-guided attention for modeling electrocardiography signals. 2019; .

[10] Fm A, Oy B, Mt C, Ubb D, Yd A, Uraef G. Application of deep learning techniques for heartbeats detection using ecg signals-analysis and review. Computers in Biology and Medicine 2020;120.

[11] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L, Polosukhin I. Attention is all you need 2017;.

Address for correspondence:

Ying An
Central South University , Changsha , China
anying@csu.edu.cn