Spatio-Temporal ECG Network for Detecting Cardiac Disorders from Multi-Lead ECGs

Long Chen[1], Zheheng Jiang[2], Tiago P Almeida[1], Fernando S Schlindwein[1], Jakevir S Shoker[1], G André Ng[1], Huiyu Zhou[1], Xin Li[1]

[1]University of Leicester, Leicester, UK
[2]University of Lancaster, Lancaster, UK

Abstract

Automatic detection and classification of cardiac disorders play a critical role in the analysis of clinical electrocardiogram (ECG). Deep learning methods are effective for automated feature extraction and have shown promising results in ECG classification. In this work, we proposed a deep spatio-temporal ECG network (ST-ECGNet) to extract robust spatio-temporal features for detecting multiple cardiac disorders from the multi-lead ECG data. The proposed ST-ECGNet combines a Convolutional Neural Network (CNN) module for extracting local spatial features, an attention module for capturing global spatial features, and a Bi-directional Gated Recurrent Unit (Bi-GRU) module for extracting temporal features from ECG data. Specifically, the attention mechanism enables our deep learning architecture to focus on the most important and useful parts of the input to make more accurate predictions. In PhysioNet/Computing in Cardiology Challenge 2021, our entry was not officially ranked and scored on the test data of the Challenge, because our code was not successfully processed during the official phase and failed to run with errors.

1. Introduction

Cardiovascular diseases (CVDs) are the main cause of death in the world nowadays, taking an estimated 17.9 million human lives each year. Therefore, detecting and treating heart diseases are of great importance and attract more attention worldwide. The electrocardiogram (ECG) that measures the electrical activity of the heart diseases is a clinical tool widely utilised for the clinical diagnosis of multiple cardiac diseases [1]. However, manual interpretation of ECG is a time-consuming task, and requires experienced cardiologists. Thus, computer-aided interpretation has become increasingly adopted in the process of clinical diagnosis, assisting the cardiologist with health care decision making.

Many traditional machine learning methods have been employed for ECG signal classification. In these methods, a variety of features are firstly extracted from ECG recordings using different techniques, such as Discrete Wavelet Transform (DWT) [2] and Pan Tompkins algorithm [3]. Then, a classification method, such as Support Vector Machine (SVM) [4] and Hidden Markov model (HMM) [5], is employed for classification. However, these approaches have two main drawbacks: 1) they rely heavily on the carefully selected features, which has been reported insufficient to handle multi-class classification tasks using these approches. 2) The accuracy of machine learning algorithm is much lower than that by a cardiologist due to the poor feature representation capability and high complexity of ECG classification tasks. Convolutional neural networks (CNNs) have recently achieved great success in detecting cardiovascular abnormalities from ECG data [6]. The major advantage of CNNs is that they are able to automatically learn discriminative features from raw input data without requiring data preprocessing and feature engineering [6]. However, CNNs cannot capture sufficient global spatial information in many cases due to the small-size convolutional filters they use when extracting feature representations. For example, CNNs commonly used 3x3 convolutional filters that have 9 pixels, the value of an output pixel is calculated with only referring to the 9 surrounding pixels which means only local information have been applied to compute an output pixel. This will bring some bias as global information is not seen. Using larger convolution filters may mitigate the problem, however, the computational overhead gets heavier and the performance has not been improved remarkably in practice. Therefore, introducing efficient modules to capture global information is of vital importance for CNNs.

In this work, we aim to 1) develop a novel end-to-end multi-label cardiac disease detection framework, where a deep CNN module, an attention module, and a Bi-

* Long Chen and Zheheng Jiang contributed equally to this work.

 ISSN: 2325-887X DOI: 10.22489/CinC.2021.194

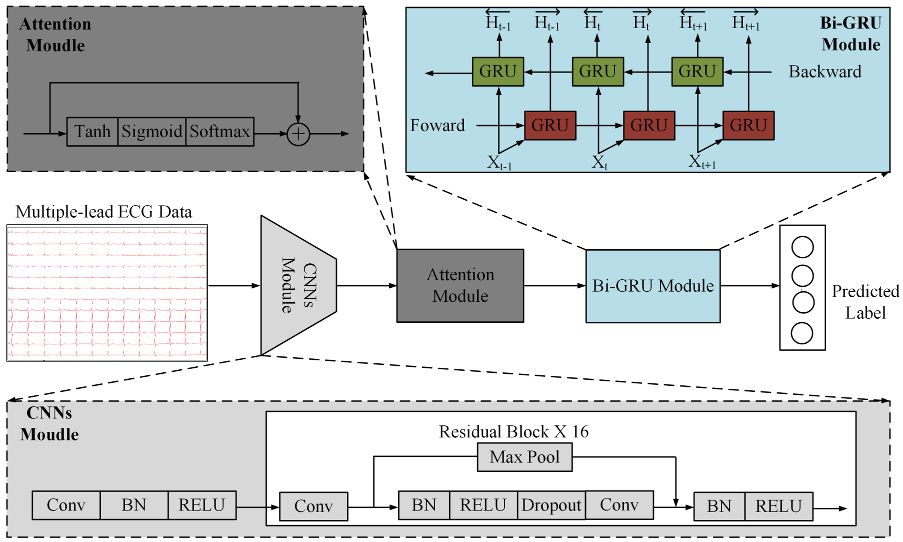

Figure 1. Overall framework of the proposed ST-ECGNet.

directional gated recurrent unit (GRU) are combined to learn roubust spatial-temporal features of ECG. 2) design an efficent attention model to caputure global information that is a supplement to the CNN model. 3) demonstrate the effectiveness and efficacy of our architecture on the ECG dataset of PhysioNet/CinC Challenge 2021 [7].

2. Model Architecture

Fig. 1 presents the overall framework of our proposed ST-ECGNet, which consisits of a CNN module for extracting local spatial features, an attention module for capturing global spatial features, and a Bi-directional GRU module for extracting temporal features from ECG data.

2.1. CNN Module

In the proposed architecture, a CNN module is first applied to learn high-level feature representations of ECG recordings. To facilite the optimisation of the CNN module, a Residual Neural Networks similar to [6] was adopted to add a short-cut connection that skips two convolutional layers. The short-cuts construct direct connections between shallow layers and deep layers, enabling the Residual Neural Networks to solve the gradient vanishing problem that is commonly found during the training stage. Specifically, the network consists of a convolutional layer (Conv) followed by 16 residual blocks with two convolutional layers per block. The width of filters is fixed with 16 in every convolutional layer. The number of filters per convolutional layer starts with 32, and after the first four residual blocks, it doubles at the first convolutional layers in every fourth residual block. Every second residual block subsamples its inputs by a factor of 2. In addition, the Batch Normalisation (BN) algorithm is adopted for rescaling the output of each convolutional layer and a rectified linear activation unit (RELU) as a nonlinear activation function. The dropout layers with a rate of 0.2 after RELU were used to prevent overfitting. The detailed structure of the CNN module can be seen in Fig. 1. The CNN module can be defined as a mapping function f_{cnn} and the extracted CNN feature is denoted as F_{CNN}:

$$F_{CNN} = f_{CNN}(E; \theta_{CNN}) \in \mathbb{R}^D \quad (1)$$

where θ_{CNN} and D denote the parameters of the CNN module and the dimension of the output CNN feature, and E is an input of multi-lead ECG recording.

2.2. Attention Module

In CNN, small-size convolution filters are commonly used to extract local features while the global features have been ignored. However, global features describe the input as a whole that are important for the high-level classification task. The global features represent the high-level context information of the whole input while the local features describe a set of local regions. Combining the global and local features is able to improve the accuracy of the classification task. To capture more global information, larger convolution filters have been employed, however, the computational overhead gets heavier and the performance may not improve remarkably in practice. In our proposed architecture, we introduced an attention module to focus on learning the most important parts of the whole input. As shown in Fig. 1, the first layer of the attention module is a tanh activation layer followed by a sigmoid layer and a softmax layer. The attention module uses additive attention to capture global context information. The attention module can be defined as a mapping function f_{AT} and the extracted feature is denoted as F_{AT}:

$$F_{AT} = f_{AT}(f_{CNN}(E; \theta_{CNN}); \theta_{AT}) \in \mathbb{R}^{D_a} \quad (2)$$

where θ_{AT} denotes the parameters of the attention module, and D_a denotes the output dimension of the attention module.

2.3. Bi-GRU Module

Spatio-temporal feature learning is of central importance for sequential data classification. In the proposed architecture, CNN and attention modules are able to extract spatial features. To further extract the temporal features from the time series of spatial features, we introduced Bi-GRU module into our framework. In our experiments, we selected the Bi-GRU since it makes full use of the context information from two directions, including forward direction and backward direction as shown in Fig. 1. We defined the Bi-GRU module as a mapping function f_{GRU}, the output feature as F_{GRU} can be calculated as:

$$F_{GRU} = f_{GRU}(f_{AT}(f_{CNN}(E; \theta_{CNN}); \theta_{AT}); \theta_{GRU}) \quad (3)$$

where θ_{GRU} denotes the parameters of the GRU module.

2.4. Training

We train separate models for 12 / 6 / 4 / 3 / 2 lead ECG data. 6 / 4 / 3 / 2 lead ECG data are extracted from the 12-lead ECG data. We padded zeros into the ECG recordings which are shorter than 18 seconds. Our network took this signal as input and output one prediction every 512 samples. We apply the class-aware binary cross-entropy loss to optimise our network. The proposed framework was trained using Adam stochastic gradient descent (SGD) optimiser with random initialisation of the weights. The training ran 30 epochs in total, the batch size was set to 32, and the learning rate is set to 0.001. The learning rate was reduced by a factor of 10 when the validation loss stopped improving for three consecutive epochs.

3. Experiment

3.1. Dataset

The dataset was provided by the PhysioNet/CinC Challenge 2021. It includes twelve-lead ECG recordings from six sources, including the CPSC database [8], the INCART database [9], the PTB database [10, 11], the Chapman-Shaoxing Database [12], the Ningbo Database [13] and other databases [7, 14]. These databases include over 100,000 twelve-lead ECG recordings with over 88,000 ECGs shared publicly as training data, 6,630 ECGs retained privately as validation data, and 16,630 ECGs retained privately as test data. Since the test data is not public available, we locally split 20% of training data as the validation dataset and the other 80% as the training dataset, and conduct the experiments on our locally split dataset.

3.2. Evaluation Metric

PhysioNet/CinC Challenges 2021 [7] has extended the 2020 Challenge scoring metric [14] to incorporate additional data and diagnoses. In total, there are five evaluation metrics including the area under the receiver-operating characteristic curve (AUROC), the area under the recall-precision curve (AUPRC), accuracy (fraction of correct recordings), macro F-measure, and the Challenge metric, which assigns different weights to different misclassification errors.

4. Results

In PhysioNet/Computing in Cardiology Challenge 2021, our entry was not officially ranked because our source code encountered the error caused by the incompatible Docker Image. The Docker Image installed on the official machines did not allow our code to access the GPU. We had also found the same error on our machine and solved this error, the solution to this error is that setting the -runtime=nvidia flag explicitly when install the Docker Image, i.e., using 'docker run -it -runtime=nvidia -v ...' to install the Docker Image.

We first conduct four ablation experiments to verify the effectiveness of each component in the proposed ST-ECGNet, including CNN module, attention module, and Bi-GRU module. In our experiments, we compare different methods using the 12-lead ECG data. We report the experimental results on the validation dataset in Table 1. Among all the compared architectures, the CNN module solely performs much worst than the other architectures, achieving 0.219 challenge metric score and 0.322 accuracy. This is because CNN module cannot extract sufficient global spatial and temporal information that support the accurate classification of ECG recordings. When we augmented the CNN architecture with the attention module or the Bi-GRU module, the performance greatly increased. The gains came from the enhancement of the final features that the attention module is able to introduce more global information, and the Bi-GRU module is able to introduce more temporal information into the final features. Our ST-ECGNet consisting of three modules achieves the best performance (0.414 challenge metric score and 0.481 accuracy). Second, we also report the performance of ST-ECGNet on different multi-lead ECG data in Table 2, from which, we observe that the ST-ECGNets

Methods	AUROC	AUPRC	Accuracy	F-measure	Challenge Metric
CNN	0.749	0.213	0.322	0.129	0.219
CNN + Attention	0.797	0.306	0.415	0.198	0.376
CNN + Bi-GRU	0.788	0.291	0.403	0.185	0.361
CNN+Attention+Bi-GRU	**0.838**	**0.338**	**0.481**	**0.232**	**0.414**

Table 1. The ablation experimental results on the local validation set drawn from the training set.

Leads	AUROC	AUPRC	Acc	F-m	Challenge
12	0.838	0.338	0.481	0.232	0.414
6	0.838	0.340	0.482	0.233	0.417
4	0.836	0.354	0.495	0.240	0.427
3	0.836	**0.361**	**0.505**	**0.251**	**0.434**
2	**0.838**	0.348	0.484	0.234	0.419

Table 2. The experimental results of the proposed framework for ECG data with different leads on the local validation set drawn from the training set.

trained on on different multi-lead ECGs achieve similar performance, and among them the ST-ECGNet trained on the 3-lead ECGs achieves the best performance (0.836 AUROC, 0.361 AUPRC, 0.505 accuracy, 0.251 F-measure score, and 0.434 challenge metric score). From the observation of our experimental results, it seems more leads have not brought better performance.

5. Conclusion

In this paper, we proposed a novel deep ST-ECGNet for detecting multiple cardiac disorders from the multi-lead ECG data. ST-ECGNet integrates a CNN module for extracting local spatial features, a attention module for capturing global spatial features, and a Bi-directional GRU module for extracting temporal features, therefore it is able to extract robust spatio-temporal features and enhances the performance in automated ECG clinical diagnosis of multiple cardiac disorders.

References

[1] Wilson FN, Kossmann CE, Burch GE, Golderger E, Graybiel A, Hecht HH, et al. Recommendations for Standardization of Electrocardiographic and Vectorcardiographic Leads. Circulation 1954;10(4):564–573.

[2] Vijayavanan M, Rathikarani V, Dhanalakshmi P. Automatic Classification of Ecg Signal for Heart Disease Diagnosis using Morphological Features. International Journal of Computer Science Engineering Technology 2014;5(4):449–455.

[3] Korürek M, Doğan B. Ecg Beat Classification using Particle Swarm Optimization and Radial Basis Function Neural Network. Expert Systems with Applications 2010; 37(12):7563–7569.

[4] Park K, Cho B, Lee D, Song S, Lee J, Chee Y, et al. Hierarchical Support Vector Machine based Heartbeat Classification using Higher Order Statistics and Hermite Basis Function. In 2008 Computers in Cardiology. IEEE, 2008; 229–232.

[5] Andreao RV, Dorizzi B, Boudy J. Ecg Signal Analysis through Hidden Markov Models. IEEE Transactions on Biomedical Engineering 2006;53(8):1541–1549.

[6] Ribeiro AH, Ribeiro MH, Paixão GM, Oliveira DM, Gomes PR, Canazart JA, et al. Automatic Diagnosis of the 12-lead Ecg using a Deep Neural Network. Nature Communications 2020;11(1):1–9.

[7] Reyna MA, Sadr N, Perez Alday EA, Gu A, Shah A, Robichaux C, et al. Will Two Do? Varying Dimensions in Electrocardiography: the PhysioNet/Computing in Cardiology Challenge 2021. Computing in Cardiology 2021;48:1–4.

[8] Liu F, Liu C, Zhao L, Zhang X, Wu X, Xu X, et al. An Open Access Database for Evaluating the Algorithms of Electrocardiogram Rhythm and Morphology Abnormality Detection. Journal of Medical Imaging and Health Informatics 2018;8(7):1368—1373.

[9] Tihonenko V, Khaustov A, Ivanov S, Rivin A, Yakushenko E. St Petersburg INCART 12-lead Arrhythmia Database. PhysioBank PhysioToolkit and PhysioNet 2008;Doi: `10.13026/C2V88N`.

[10] Bousseljot R, Kreiseler D, Schnabel A. Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das Internet. Biomedizinische Technik 1995;40(S1):317–318.

[11] Wagner P, Strodthoff N, Bousseljot RD, Kreiseler D, Lunze FI, Samek W, et al. PTB-XL, a Large Publicly Available Electrocardiography Dataset. Scientific Data 2020;7(1):1–15.

[12] Zheng J, Zhang J, Danioko S, Yao H, Guo H, Rakovski C. A 12-lead Electrocardiogram Database for Arrhythmia Research Covering More Than 10,000 Patients. Scientific Data 2020;7(48):1–8.

[13] Zheng J, Cui H, Struppa D, Zhang J, Yacoub SM, El-Askary H, et al. Optimal Multi-Stage Arrhythmia Classification Approach. Scientific Data 2020;10(2898):1–17.

[14] Perez Alday EA, Gu A, Shah A, Robichaux C, Wong AKI, Liu C, et al. Classification of 12-lead ECGs: the Physionet/Computing in Cardiology Challenge 2020. Physiological Measurement 2020;41.

Address for correspondence:

Long Chen
Department of Cardiovascular Sciences, University of Leicester, Glenfield Hospital Leicester, LE3 9QP
lc408@leicester.ac.uk