Channel self-Attention Deep Learning Framework for Multi-Cardiac Abnormality Diagnosis from Varied-lead ECG Signals

Apoorva Srivastava[1], Ajith Hari[1], Sawon Pratiher[1], Sazedul Alam[2],
Nirmalya Ghosh[1], Nilanjan Banerjee[2], Amit Patra[1]

[1] *Department of Electrical Engineering, Indian Institute of Technology Kharagpur, West Bengal, India.*
[2] *Computer Science and Electrical Engineering, University of Maryland-Baltimore County, Maryland, USA.*

Abstract

Electrocardiogram (ECG) is a widely used signal to diagnose heart health. Experts can detect multiple cardiac abnormalities using the ECG signal. In a clinical setting, 12-lead ECG is mainly used. But using a lower number of leads can make the ECG more pervasive as it can be integrated with wearable devices. At the same time, we need to build systems that can diagnose cardiac abnormality automatically. This work develops Channel self-Attention (CA) based deep neural network to diagnose cardiac abnormality using a different number of ECG leads. Our approach takes care of the temporal and spatial interdependence of multi-lead ECG signals. Our team participates under the name "cardiochallenger" in the "PhysioNet/Computing in Cardiology Challenge 2021". Our method achieves the challenge metric scores of 0.64, 0.64, 0.64, 0.63, and 0.63 for the 12-lead, 6-lead, 4-lead, 3-lead, and 2-lead cases, respectively, on the validation data set.

1. Introduction

With over 17.9 million deaths, cardiovascular diseases are the leading cause of mortality worldwide [1]. The heart's activity from different angles can be studied from 12-lead ECG. Detection of multiple cardiac abnormalities like coronary occlusion, myocardial infarction, etc can be done using 12-lead ECG.

Early-stage prognosis and timely interventions aid clinicians in the identification of different cardiac irregularities and provide improved clinical outcomes. The PhysioNet/CinC-2021 challenge is dedicated to cardiac abnormality classification (CAC) from 12-lead, 6-lead, 4-lead, 3-lead, and 2-lead ECG recordings [2]. Early and accurate detection of diseases with a lower number of leads makes ECG greater pervasive as it can be incorporated with wearable devices. Conventional CAC methods regularly employs machine learning models on the extracted domain-aware handcrafted features using raw ECG signal processing. Of late deep learning (DL) methods have democratized the CAC task with superior performance [3],[4],[5]. DL models can abstract explanatory ECG feature representations in an automated fashion and predict CACs in an end-to-end manner [6],[7],[8]. In this paper, an attention-based DL model is proposed, which will help medical practitioners to judiciously inspect and categorize the inter-beat and intra-beat patterns. The proposed model acknowledges the spatial interrelation among the channels and the temporal important segments of the signal.

The rest of the paper is organized as follows. Section 2 summarizes the data prepossessing and our channel self-attention based DL architecture. Experimental results are discussed in section 3 and 4. Section 5 concludes the paper.

2. Methodology

Cardiac abnormality detection using ECG signals can be formulated as a time-series classification problem. We aim to detect 29 multi-labeled cardiac abnormalities along with sinus rhythm using varying lead ECG signals [2]. The model is trained on 12-lead ECG and tested on:

- 12-lead: I,II,III,aVR,aVL,aVF,V1,V2,V3,V4,V5,V6
- 6-lead: I,II,III,aVR,aVL,aVF
- 4-lead: I,II,III,V2
- 3-lead: I,II,V2
- 2-lead: I,II

In this paper, the model is based on the channel self-attention (CA) framework for the diagnosis of multi-labeled cardiac abnormalities as shown in Figure 1. The model is inspired by squeeze and excitation network [9]. The global spatial information is squeezed and channel wise statistic is generated by CA framework. Higher weight is given to the more imperative channel which leads to enhanced performance. Here, it is applied with the inception and residual neural model. In the following section, we provide a detailed description of the system's components.

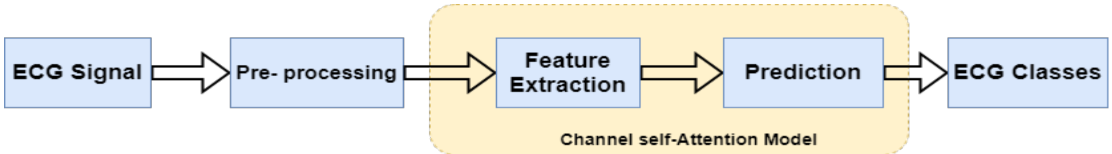

Figure 1. Pipeline for cardiac abnormalities detection

2.1. Data pre-processing

The publicly available challenge dataset consists of 88,253 twelve-lead ECGs recordings. The data is collected from 4 countries across 3 continents. The sampling frequency of the ECG signals varies from 257 Hz to 1000 Hz, and signal duration varies from 6 seconds to 30 minutes. All the signals are down-sampled to 125 Hz to handle the variable sampling frequency of the signals. If the signal is too long, it is truncated after 120 seconds.

2.2. Channel self-Attention Based Deep Learning Model

The input to the proposed CA based model is a variable-length ECG segment $(X = [x_1, ..., x_k])$ and prediction of 29 cardiac abnormalities along with sinus rhythm, is the output of the model. The proposed CA based architecture is shown in figure 2(a). The channel depth and the number of inception and residual blocks are experimentally decided and the tuning of network parameter is done by hit and trial. The channel self-attention based deep learning model is the ensemble of: 1) Inception and Residual architecture; 2) Channel self-Attention architecture; 3) Attention pooling.

2.2.1. Inception and Residual architecture

The idea of inception came from [10], where sparsely connected architecture was introduced to replace the fully connected connection of Convolution Neural Network (CNN) layers. In this paper, the Inception model has convolution layers with 1-D filter of size 3,4, and 5 with ReLu activation. The inputs to the CNN can be of a variable length, so the model is accustomed to handle variable length data. The channel number increases as we move forward in the architecture. Figure 2(c) demonstrated the interior of the Inception blocks used in the model.

The residual neural network helps in solving the problem of vanishing gradient [11]. The CNN layers in the Inception block maps the input x_j to low dimension embedding $h_k = f_\psi(x_j)$, where f_ψ is transformation function with parameter ψ. The output of the residual block is y = $\mathcal{F}$ (h_j) +$\mathcal{G}(x_j)$, where $\mathcal{F}(.)$ shows the residual mapping to be learned and $\mathcal{G}(.)$ is the convolution layer added to match the dimension. Figure 2(b) demostrates the residual block in the proposed deep learning model.

2.2.2. Channel self-Attention architecture

The spatial features are extracted by CNN. Each channel represents the information of the feature map extracted. Adaptive weights can be assigned to the channel to find the interrelation among the channels. Therefore, we built a channel self attention module to use the interdependence among the channels.

The idea of Channel self-Attention is derived from SE-Net [9] where inter-dependency among channels is captured as a function of channel description (global average). The main difference is that instead of finding explicit relation among channel descriptions, multiple channel descriptions are used and channel attention is computed only using the feature vector extracted from the corresponding channel. The feature vector is extracted by passing spatial features channel-wise to one dimensional CNN. 32 filters are used to extract 32 deep features as shown in figure 3. Here the interrelation among channels is captured through sharing the weights for calculating self-attention. Application of attention mechanism channel wise can be regarded as the method of choosing semantic attributes.

2.2.3. Attention Pooling layer

The attention pooling was introduced by [12] which is an adaptive multi instance pooling method. Depending on the number of classes it is modified to multi-head attention. Corresponding to every feature vector segment weights are generated by multi-head attention neural network. Softmax activation is used to ensure that the weight sums up to 1. The attention mechanism helps to make the model interpretable, and hence replaced the widely used Long Short-Term Memory (LSTM) or Bi-LSTMs. The higher weights give importance to that segment of the signal.

Let $V = \{v_1, v_2, \cdots, v_K\}$ is a bag of K feature vectors, then attention pooling is defined as: $p = \sum_{k=1}^{K} a_k v_k$. Here, $p \in \mathbb{R}^{\mathbb{N}\times\mathbb{L}}$ is the feature vector corresponding to N classes heart disease. $a_k = \frac{exp\{W^T\ tanh(Uv_k^T)\}}{\sum_{j=1}^{K} exp\{W^T\ tanh(Uv_j^T)\}}$, where, U and W are trainable parameters.

2.3. Threshold Optimisation

The output from the attention pooling layer is passed to the prediction layer for the detection of abnormalities. The sigmoid activation in prediction layer will give the proba-

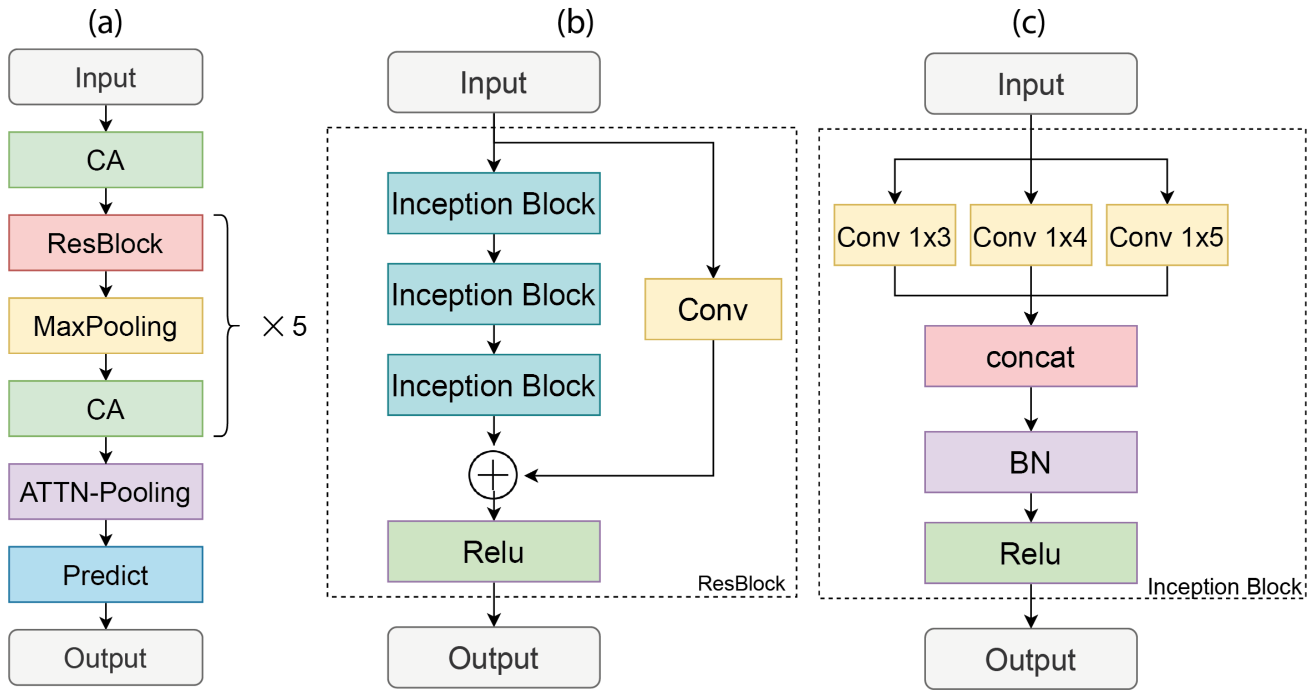

Figure 2. Channel self-Attention based deep learning architecture. One-hot encoding of cardiac abnormalities classes is the output (grey) of the model. The figure illustrates: (a) proposed CA based deep learning model; (b) Residual network block; (c) Inception network block

bilities of the occurrence of sinus rhythm and 29 cardiac abnormalities . For the evaluation of the challenge metric score, these probabilities are needed to be changed in binary format by applying the threshold value on these predictions. If the prediction crossed the threshold value, 1 is assigned to the corresponding class else 0. Genetic algorithm was used to optimise thresholds for each class that maximizes challenge metric on the validation dataset.

2.4. Implementation Details

The dataset consists of 111 abnormalities, but in challenge, it is required to detect 29 cardiac abnormalities along with sinus rhythm and their SNOMED CT codes are included in the challenge evaluation metric (W_{reward}) [13]. The loss function is described below:

$loss(x, y) = mf \cdot BCE(x, y) - S_{normalised}$. Here, BCE is Binary Cross-Entropy loss, mf is a multiplication factor which scales the BCE loss by a factor of 0.1 if the difference between true and predicted label is less than 0.3 and $S_{normalised}$ is the normalised challenge metric which is computed as, $S_{normalised} = \frac{S_{observed} - S_{inactive}}{S_{true} - S_{inactive}}$, where, $S_{(x,y)} = X^T \times (\frac{y}{norm}) \cdot W_{reward}$, $norm = max(x + y - xy, 1)$.

For training the model Adam optimizer [14] was used with a learning rate of 0.001. The parameters are initialized using Xaviar uniform initializer. The early stopping

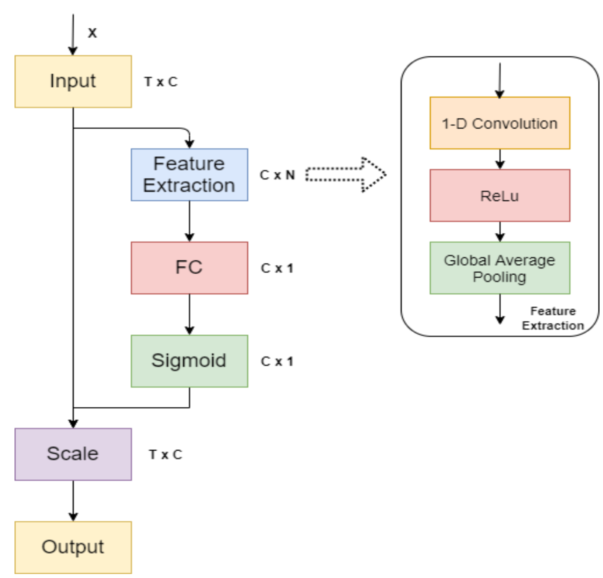

Figure 3. Schema of Channel self-attention architecture

method is also incorporated in the algorithm to avoid overfitting of the model. The model trained for 100 epochs with a batch size of 32. The complete model has 7,15,512 trainable parameters.

3. Results

The proposed model is trained on a publicly available dataset given by PhysioNet/CinC challenge and tested on the hidden 16,630 ECG recordings. The table 1 shows the ECG signal with variable lead number and the corresponding challenge metric score on test dataset. The proposed model requires 2498 minutes for training.

Table 1. Model performance on validation dataset

Number of Leads	Challenge Metric Score
12-Lead	0.64
6-Lead	0.64
4-Lead	0.64
3-Lead	0.63
2-Lead	0.63

4. Discussion

In this paper, we proposed a channel self-attention deep learning model to detect the cardiac abnormalities using 12-lead, 6-lead, 4-lead, 3-lead. and 2-lead. The attention mechanism was used to capture the informative segment of the signal along with spatial interdependence among the channels.

During the pre-processing of the raw signal, we truncated the ECG data that is more than two minutes long. Our decision to truncate affected 74 samples, which is relatively very low (< 0.01 % of the total samples).

The duration of the input signal varies from 6 seconds to 2 minutes. We made the length of the input to our CA model a variable length to make it robust for any length of data so that it can be used in real time. We tried to make a generalized model that can handle the heterogeneity in the training dataset and data of variable length.

5. Conclusion

In this paper, we have described a channel self-Attention neural network-based approach to classify cardiac abnormalities presented in the PhysioNet/Computing in Cardiology Challenge 2021. Our DL model can classify 29 cardiac abnormalities along with sinus rhythm with a challenge metric score of 0.64, 0.64, 0.64, 0.63 and 0.63 for the 12-lead, 6-lead, 4-lead, 3-lead, and 2-lead cases, respectively, placing us in the top ten team entries in the official phase. The efficacy of the proposed system on real-life datasets validates its competency for actual clinical practice.

References

[1] https://www.who.int/health-topics/cardiovascular-diseases/#tab=tab_1.

[2] Reyna M. SNPAEALCSS, Clifford GD. Will two do? varying dimensions in electrocardiography: the physionet - computing in cardiology challenge 2021. In 2021 Computing in Cardiology. IEEE, 2020; 1–4.

[3] 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.

[4] Ribeiro AH, Ribeiro MH, Paixão GM, Oliveira DM, Gomes PR, Canazart JA, Ferreira MP, Andersson CR, Macfarlane PW, Meira Jr W, et al. Automatic diagnosis of the 12-lead ECG using a deep neural network. Nature communications 2020;11(1):1–9.

[5] Hong S, Zhou Y, Shang J, Xiao C, Sun J. Opportunities and challenges of deep learning methods for electrocardiogram data: A systematic review. Computers in Biology and Medicine 2020;103801.

[6] Natarajan A, Chang Y, Mariani S, Rahman A, Boverman G, Vij S, Rubin J. A wide and deep transformer neural network for 12-lead ecg classification. In 2020 Computing in Cardiology. IEEE, 2020; 1–4.

[7] Zhu Z, Wang H, Zhao T, Guo Y, Xu Z, Liu Z, Liu S, Lan X, Sun X, Feng M. Classification of cardiac abnormalities from ECG signals using se-resnet. In 2020 Computing in Cardiology. IEEE, 2020; 1–4.

[8] Kirodiwal A, Srivastava A, Dash A, Saha A, Penaganti GV, Pratiher S, Alam S, Patra A, Ghosh N, Banerjee N. A bio-toolkit for multi-cardiac abnormality diagnosis using ecg signal and deep learning. In 2020 Computing in Cardiology. IEEE, 2020; 1–4.

[9] Hu J, Shen L, Sun G. Squeeze-and-excitation networks. In Proceedings of the IEEE conference on computer vision and pattern recognition. 2018; 7132–7141.

[10] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition. 2015; 1–9.

[11] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition. 2016; 770–778.

[12] Ilse M, Tomczak J, Welling M. Attention-based deep multiple instance learning. In International conference on machine learning. PMLR, 2018; 2127–2136.

[13] Alday EAP, Gu A, Shah AJ, Robichaux C, Wong AKI, Liu C, Liu F, Rad AB, Elola A, Seyedi S, et al. Classification of 12-lead ECGs: the physionet/computing in cardiology challenge 2020. Physiological measurement 2020; 41(12):124003.

[14] Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv preprint arXiv14126980 2014;.

Address for correspondence:

Apoorva Srivastava
EE, IIT Kharagpur, Kharagpur, West Bengal 721302, India
apoorva.srivastava.23@iitkgp.ac.in