Incorporating Clinical and Heartbeat Level Features with Multichannel ECG for Cardiac Abnormality Detection Using Parallel CNN and GAP Network

Deepankar Nankani and Rashmi Dutta Baruah
Department of Computer Science and Engineering
Indian Institute of Technology Guwahati, Assam, India-781039

Abstract

Background: Early and correct diagnosis of cardiac arrhythmias from Multichannel Electrocardiogram (MECG) is a challenging problem. We address this problem as part of the PhysioNet/Computing in Cardiology Challenge 2021 by classifying cardiac abnormalities from MECG.

Method: The proposed method incorporates clinical features including patient age, gender and heartbeat features with MECG to detect cardiac abnormalities. Initially, MECG is cleaned from noise, followed by resampling and segmentation. Then R-peaks are extracted from Lead II signal using Pan Tompkins detector to obtain heartbeat level features such as heart rate, RR Intervals, Mean QRS Amplitude, Hermite polynomial coefficients, statistical features, and Wave Amplitude based features. The feature vector consisting of clinical and heartbeat features is combined with MECG for classification using a Parallel Convolution Neural Network with Global Average Pooling (PCNN-GAP) network. The model extracts local features using smaller convolution kernels and global patterns using larger kernels. Lastly, the last layer sigmoid activation function classifies the rhythm into one or more cardiac abnormalities.

Results: Our team, skylark, achieved a score of 0.466, 0.426, 0.5, 0.491, *and* 0.514 *(ranked* 131^{th}, 137^{th}, 116^{th}, 118^{th}, *and* 107^{th} *out of* 256 *teams) for the 12-lead, 6-lead, 4-lead, 3-lead, and 2-lead versions of the hidden validation set with the Challenge evaluation metric.*

1. Introduction

Cardiovascular diseases (CVD) such as arrhythmia are a leading cause of death worldwide [1]. Detecting arrhythmias before their occurrence using an Electrocardiogram (ECG) signal helps in risk stratification, better medical assistance, and patient treatment [2]. The standard 12-lead ECG is acquired non-invasively by placing electrodes to the patients' body. The recorded electrical activity of the heart is used to diagnose cardiac pathologies. The limited accessibility of 12-lead devices provides motivation to use reduced lead devices as they are cheap and easily accessible. However, reduced lead devices capture limited useful information compared to 12-lead devices [3]. This paper addresses the PhysioNet/Computing in Cardiology Challenge 2021 that focuses on automated, open-source approaches for classifying cardiac abnormalities from reduced-lead ECGs [4–6]. Our best entry in the challenge incorporates clinical features including age, gender and heartbeat features with MECG for cardiac abnormality detection using parallel convolution neural network with global average pooling (PCNN-GAP) network.

2. Methods

The workflow encompasses of a preprocessing and classification stage as described in Figure 1. The patient record consists of MECG with 12 lead ECGs, Analog to Digital conversion (ADC) gain, baseline for each lead, age, gender, patient history, symptoms, medical prescription, and diagnosis or cardiac rhythm information (disease labels).

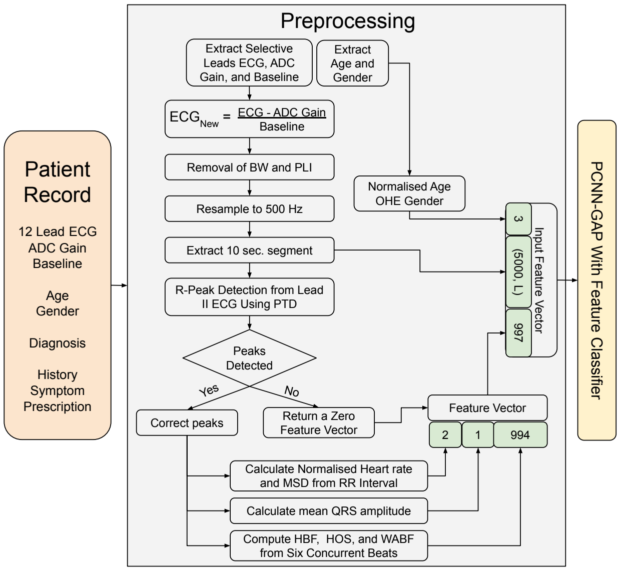

Figure 1. Proposed workflow consisting of patient record description, preprocessing, and classification stages.

2.1. Preprocessing

The MECG is normalized using the respective ADC gains and baselines. The resultant signal (ECG_{New}) is decluttered from low frequency noise ($\leq 0.6\ Hz$) such as baseline wander (BW) and high frequency noise ($50\ Hz$) such as powerline interference (PLI). The data is procured from different sources with different sampling frequencies, resampling to $500\ Hz$ is performed as $4/6$ sources are sampled at $500\ Hz$. Then, segments of 10-second are extracted from resampled signals as it is the most common duration among all the records. The longer length segments are divided into 10 second segments are short segments are padded with zeros. Then heartbeat level features and clinical features are extracted that includes age and gender encoding, followed by classification.

Heartbeat Level Feature Extraction: The Lead II signal is selected from segmented MECG for R-peak detection using the Pan Tompkins Detection (PTD) algorithm [7]. Lead II provides better depiction of characteristic waveforms present in ECG rhythm and is the common lead in all sets [8]. If PTD algorithm fails to extract R-peaks from the clean signal, then our method returns a 997 dimensional zero feature vector. If the peaks are detected using PTD, they are updated with the local maxima present in the close vicinity. The correct R-peaks are used to extract the following features: (i) Average Heart rate; (ii) Mean Square Difference between the RR intervals (MSDRR); (iii) Mean QRS Amplitude; and (iv) Beat Level Features (BLF). The average heart rate (HR_{avg}) and MSDRR are calculated using Eq. 1. The RR intervals (RRI) are calculated using the difference between adjacent R-peaks. Mean QRS amplitude is the mean of R-Peak amplitude.

$$\begin{aligned} HR &= \frac{60 \times Fs}{R - Peak_{Current} - R - Peak_{Previous}} \\ HR_{avg} &= \frac{\sum(HR)}{100 \times Len(HR)} \\ MSDRR &= \frac{Mean[Diff[RRI]^2]}{100} \end{aligned} \tag{1}$$

The BLF are calculated for up to six concurrent beats enclosed inside a window of 130 timestamps left to the R-peak and 230 timestamps right to the R-peak so that at least a heartbeat is enclosed in the window. The BLF include: (i) Hermite Polynomial Coefficients (HPC) [9]; (ii) Statistical Features [9, 10]; and (iii) Wave Amplitude-Based Features (WABF) [10]. The HPC exploits similarity between the beats by representing them with Hermite basis function coefficients. Assume, $x(t)$ denotes the beat, the Hermite series expansion yields Eq. 2. Here, c_n is the expansion coefficient, σ is width parameter, $H_0(x) = 1$ and $H_1(x) = 2x$. The higher the Hermite polynomial order, the higher is its frequency of changes within the time domain, and the better is its capability to reconstruct quick changes of the ECG paradigm.

$$\begin{aligned} x(t) &= \sum_{n=0}^{N-1} c_n \phi_n(t, \sigma) \\ \phi_n(t, \sigma) &= \frac{1}{\sqrt{\sigma 2^n n! \sqrt{\pi}}} e^{\frac{-t^2}{2\sigma^2}} H_n\left(\frac{t}{\sigma}\right) \\ H_n(x) &= 2xH_{n-1}(x) - 2(n-1)H_{n-2}(x) \end{aligned} \tag{2}$$

For computing the statistical features, the beat is divided into five intervals, and Kurtosis and Skewness value over each interval is computed. For computing WABF, euclidean distance is calculated between R-peak and the characteristic waves such as P, Q, S, T-wave. The BLF produce a 994 dimensional feature vector. If the number of beats are less and the features are less than 994, then zero padding is performed. The obtained feature vector is combined with HR_{avg}, MSDRR, and mean QRS amplitude to make a 997 dimensional feature vector.

Age and Gender Encoding: The age and gender values are extracted for each patient. The age is normalised between 0 and 1. The negative and empty age fields are replaced with zero. One hot encoding (OHE) is performed for the gender attribute, where the male and female are encoded as (1,0) and (0,1), respectively. The gender with Nan value is encoded as (0,0). The OHE age and normalised gender attributes account for three features.

The final feature vector consisting of cleaned 10 second MECG of 5000 timestamps with L leads, and a 1000 dimensional feature vector are provided to the proposed PCNN-GAP with feature classifier. L depends on number of leads available in the dataset. For 2-lead ECG, $L = 2$.

2.2. PCNN-GAP with Feature Classifier

The proposed architecture of Parallel Convolution Neural Network - Global Average Pooling (PCNN-GAP) incorporated with age, gender, and heartbeat level features is illustrated in Figure 2. The idea behind the application of parallel convolution layers is that MECG consists of local and global patterns. The global patterns are extracted using the large kernels embedded in the left branch and the local patterns are extracted using the small kernels embedded in the right branch of PCNN-GAP network. The number of filters, stride, and activation function of each convolution layer are mentioned in Figure 2. For instance, the first convolution layer in the left branch encompasses a 1-D convolution with 48 filters of size 19 and stride 4 followed by a Batch Normalization (BN) [11] layer and Rectified Linear Unit (ReLU) activation. Four more layers are added in a cascaded fashion in both branches with more number of filters of reduced size followed by 1-D GAP layer [12].

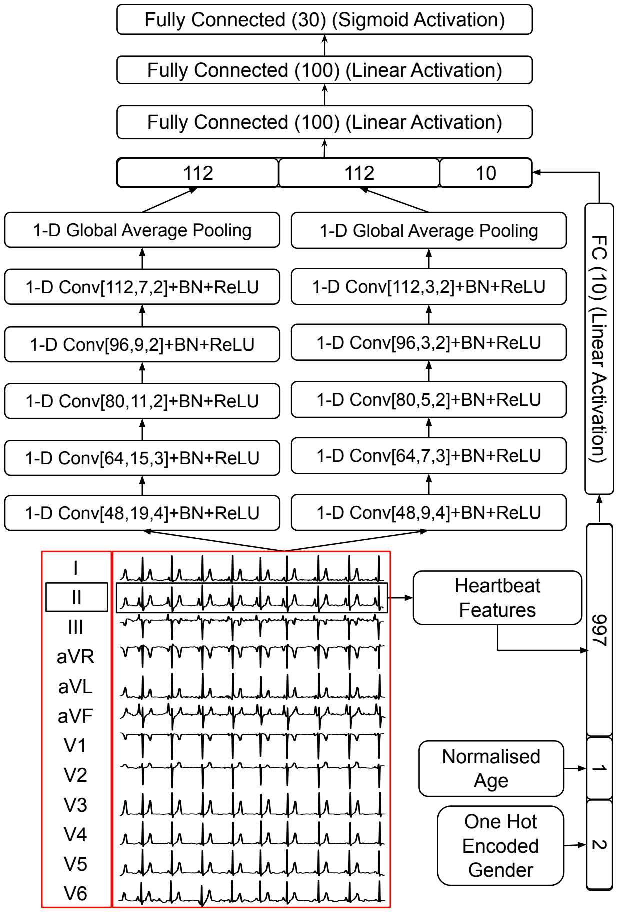

Figure 2. Architecture of Feature Fused PCNN-GAP.

The GAP layer calculates the spatial average of filters, making it robust to spatial translations of MECG. GAP layer has the following advantages over the combination of flatten and fully connected (FC) layer: (i) less prone to overfitting; (ii) no dependency on external regularization; (iii) no trainable parameters [12]. The reduced parameters lead to significantly faster training and reduced model size, making it suitable for low resource devices. The GAP layer reduces the last layer dimensions from $(53, 112)$ to (112). The branch outputs are concatenated with encoded feature vector obtained from FC layer and generates a 234-dimensional feature vector. This vector is provided to FC layers and scores ($s \in [0, 1]$) are produced by sigmoid activation function for 'scored' pathologies. The predicted scores are thresholded and labels are assigned only if the respective threshold is surpassed.

3. Results

Challenge scores for the final selected entry of the team **skylark** using 5-fold Cross Validation (CV) on the public training set, scoring and ranking on the hidden validation set are provided in Table 1. While performing CV, we did not include stratified holdout MECG's of the same patient in both the training and validation set. We tried multiple approaches, and were able to submit three approaches for testing on validation set. Table 2 describes the comparison between the approaches: PCNN-GAP; PCNN-GAP-Big; PCNN-GAP-Feature; and PCNN-GAP-Big-Feature. The PCNN-GAP method consists of parallel CNN and GAP layer, PCNN-GAP-Big consists of two additional convolution layers in each branch, both PCNN-GAP-Feature and PCNN-GAP-Big-Feature incorporates clinical and heartbeat features in PCNN-GAP and PCNN-GAP-Big model, respectively. The PCNN-GAP-Big-Feature was not submitted due to time constraints and therefore the validation results are not available (NAVL). The proposed methods improved over our 8-layer ResNet model that performed multi-class classification and predicted only single-label for MECG [13]. Introducing multiple labels improved the score to 0.5. Adding features increased the score to 0.51 on validation set for 12-lead ECG in the official phase.

The thresholding applied in sigmoid activation function affects the evaluation metrics as described in Figure 3. Thresholding the scores at 0.1 produces optimum results for all evaluation metrics and deteriorates at other intervals. During testing, the label predicted most number of times by MECG segments of a patient record is chosen as the final predicted label.

Leads	Training	Validation	Test	Ranking
12	0.5 ± 0.03	0.466	NA	131
6	0.19 ± 0.04	0.426	NA	137
4	0.56 ± 0.04	0.5	NA	116
3	0.53 ± 0.04	0.491	NA	118
2	0.59 ± 0.04	0.514	NA	107

Table 1. Challenge scores for our final selected entry (team **skylark**) using 5-fold CV on the public training set, scoring and ranking on the hidden validation set.

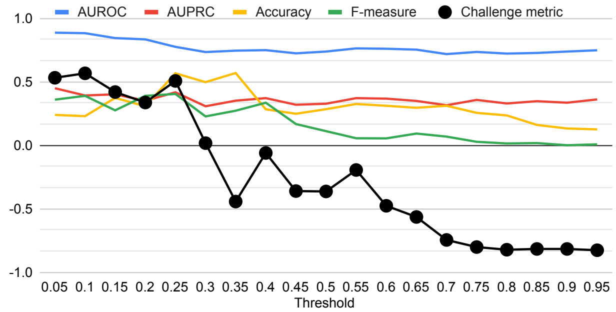

Figure 3. Deterioration in evaluation metrics with increasing threshold of last layer sigmoid activation function.

	PCNN-GAP		PCNN-GAP-Big		**PCNN-GAP-Feature**		PCNN-GAP-Big-Feature	
Leads	Training	Validation	Training	Validation	Training	Validation	Training	Validation
12	0.53±0.06	0.495	0.54±0.07	0.478	0.50±0.03	0.466	0.50±0.12	NAVL
6	0.18±0.07	0.516	0.19±0.05	0.467	0.19±0.04	0.426	0.05±0.05	NAVL
4	0.55±0.04	0.496	0.56±0.02	0.517	0.56±0.04	0.5	0.62±0.02	NAVL
3	0.57±0.02	0.493	0.53±0.10	0.532	0.53±0.04	0.491	0.57±0.01	NAVL
2	0.56±0.04	0.499	0.58±0.04	0.498	0.59±0.04	0.514	0.62±0.02	NAVL

Table 2. Comparison between various approaches tried alongside the final submitted (PCNN-GAP-Feature) approach.

4. Discussion and Conclusions

We implemented a multi-label strategy to detect 29 cardiac abnormalities and 'sinus rhythm' from MECG signals. During training, early stopping criteria was employed and validation AUC was monitored for 5 epochs to avoid overfitting. We believe that fusing age, gender, and other heartbeat features extracted from Lead II with MECG proved to be beneficial for classification for increased number of lead sets. Adding features improved the performance for more number of leads and reduced the performance for reduced lead sets. The model with less number of trainable parameters performed better for reduced lead sets and models with more number of trainable parameters performed better for standard 12-lead ECG. The six lead model performed poorly irrespective of the model. The training time of feature fused models was around 3000 seconds and 200 seconds for non-feature based models. Since noisy rhythms are not present in scored classes, the approach may produce high false alarm rates for noisy ECGs, limiting usability of the proposed algorithm in real-world applications.

Furthermore, the weight matrix provided by the challenge organisers was not used during training phase. We used single classifier for all lead configurations, namely PCNN-GAP-Feature, making the approach sub optimal. An ensemble approach could also be investigated for improved performance. In addition, feature selection is not performed for selecting most contributing hand crafted features. This might have deteriorated the performance for reduced lead set.

References

[1] Virani SS, Alonso A, Aparicio HJ, Benjamin EJ, Bittencourt MS, Callaway CW, et al. Heart Disease and Stroke Statistics – 2021 Update: a Report from the American Heart Association. Circulation 2021;143(8):e254–e743.

[2] Kligfield P. The Centennial of the Einthoven Electrocardiogram. Journal of Electrocardiology 2002;35(4):123–129.

[3] Drew BJ, Pelter MM, Brodnick DE, Yadav AV, Dempel D, Adams MG. Comparison of a new reduced lead set ecg with the standard ecg for diagnosing cardiac arrhythmias and myocardial ischemia. Journal of electrocardiology 2002; 35(4):13–21.

[4] Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. Circulation 2000;101(23):e215–e220.

[5] 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.

[6] 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.

[7] Pan J, Tompkins WJ. A real-time qrs detection algorithm. IEEE transactions on biomedical engineering 1985; 32(3):230–236.

[8] Meek S, Morris F. Abc of clinical electrocardiography: Introduction. i—leads, rate, rhythm, and cardiac axis. BMJ British Medical Journal 2002;324(7334):415.

[9] Lagerholm M, Peterson C, Braccini G, Edenbrandt L, Sornmo L. Clustering ecg complexes using hermite functions and self-organizing maps. IEEE Transactions on Biomedical Engineering 2000;47(7):838–848.

[10] Mondéjar-Guerra V, Novo J, Rouco J, Penedo MG, Ortega M. Heartbeat classification fusing temporal and morphological information of ecgs via ensemble of classifiers. Biomedical Signal Processing and Control 2019;47:41–48.

[11] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In International conference on machine learning. PMLR, 2015; 448–456.

[12] Lin M, Chen Q, Yan S. Network in network. arXiv preprint arXiv13124400 2013;.

[13] Nankani D, Saikia P, Baruah RD. Automatic concurrent arrhythmia classification using deep residual neural networks. In 2020 Computing in Cardiology. IEEE, 2020; 1–4.

Address for correspondence:

Deepankar Nankani
Department of Computer Science and Engineering
Indian Institute of Technology Guwahati, Assam, India
d.nankani@iitg.ac.in