

# Identification of Cardiac Arrhythmias From 12-lead ECG Using Beat-wise Analysis and a Combination of CNN and LSTM

Mohanad Alkhodari<sup>1</sup>, Leontios J Hadjileontiadis<sup>1</sup>, Ahsan H Khandoker<sup>1</sup>

<sup>1</sup> Healthcare Engineering Innovation Center, Khalifa University, Abu Dhabi, United Arab Emirates

## Abstract

*Throughout the years, there have been many attempts to develop an accurate cardiac arrhythmias identification algorithm. However, despite achieving acceptable results, they have been only applied on either small or homogeneous data-sets. A study was developed herein to identify cardiac arrhythmias from varied-length 12-lead ECG signals obtained from the PhysioNet/Computing in Cardiology Challenge 2020 and acquired from a wide set of sources. Our team, Care4MyHeart, developed an approach that starts by analyzing the labels of the database. Then, applying various signal processing techniques to denoise the 12-lead signals. After that a beat-by-beat segmentation procedure was followed to identify the most significant beats in exhibiting the arrhythmia within the signals. A CNN+BiLSTM model was then trained and evaluated on the training set using 10-fold cross-validation scheme as well as on hidden validation and testing sets. Our approach achieved a challenge validation score of 0.379 and full test score of 0.146 on the hidden validation and testing sets, respectively. Our team was ranked the 26th out of 41 entries in this year's Challenge.*

## 1. Introduction

Cardiovascular disease (CVD) is considered the most common cause of death with an overall 31% of total deaths worldwide. The majority of CVD deaths are due to cardiac arrhythmias [1]. Currently, electrocardiography (ECG) has been widely used by clinicians as an essential tool to represent the electrical activity of the heart. Despite being able of identifying and localizing arrhythmias successfully from an ECG signal, the process requires high efforts due to the interference and noise contaminated within the signals [2]. In addition, it is considered time-consuming for clinicians to diagnose a wide range of patients' on case by case basis.

In this work, we propose a deep learning approach using convolutional neural networks (CNN) and Bi-directional long short-term memory (BiLSTM) to automatically iden-

tify cardiac arrhythmias from 12-lead ECG obtained from the PhysioNet/Computing in Cardiology Challenge 2020 [3]. The novelty of the work presented herein lies in detecting the most significant beats within an ECG signal and extracting their corresponding spatial and temporal features using a CNN+BiLSTM network.

## 2. Material and methods

### 2.1. Database preparation

Initially, the database was fixed in terms of patients' arrhythmia types to include only the challenge-scored patients labels. The database comes originally with a total of 111 arrhythmia types, out of which only 27 were included for the scoring of the challenge. From these 27 types, 6 arrhythmia types were merged to have the same label; namely complete right bundle branch block and right bundle branch block, premature atrial contraction and supraventricular premature beats, and premature ventricular contractions and ventricular premature beats. The edited data-set included a total of 24 arrhythmia types/labels after the fixing procedure.

Furthermore, several patients had multiple labels due to having multiple arrhythmia types. To ensure the inclusion of all types, each patient data was duplicated for each arrhythmia type such that multiple signals might be the identical while representing different types of arrhythmia. It worth noting that due to the huge un-balance in the database in favor of the sinus rhythm (normal) labeled recordings (34.53%), it was randomly reduced to less than 25% of its total amount (20846) after the duplication.

### 2.2. Pre-processing

The pre-processing steps were performed to reduce the noise contaminated within ECG signals. Noise sources may be power-line interference, movement artifact, and base-line wander. It is important to have less noisy signals prior to any feature extraction methodology, while at the same time, maintain the original morphological information of each signal. To achieve this, the following sub-

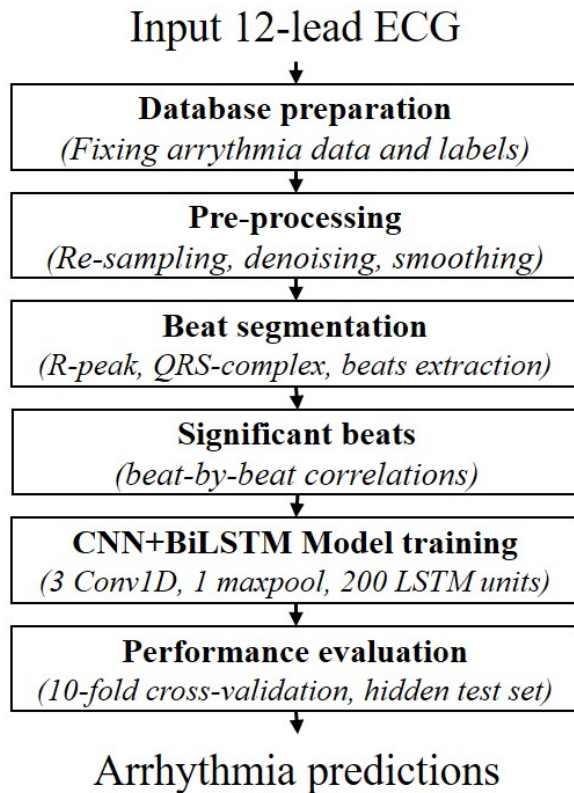


Figure 1. The complete procedure followed in 12-lead ECG arrhythmia prediction.

sections briefly describe the methods followed. By the end of the pre-processing stage, all signals were  $z$ -score normalized to ensure a mean of 0 and standard deviation of 1 for each signal.

### 2.2.1. Signals re-sampling

The database comes with signals sampled at 1000 Hz, 500 Hz, and 257 Hz due to the collection of different 12-lead ECG sources. Therefore, all signals were re-sampled to have a sampling rate of 500 Hz. Then, only the first 10 seconds segment from each lead signal was taken to reduce the complexity of the algorithm. If a signal was recorded for less than 10 seconds, it was padded with zeros.

### 2.2.2. Power-line filtering

Power-line interference is a common noise source in any ECG signal. It is characterized by having a sinusoidal pattern of a high frequency (50 Hz to 60 Hz). Such narrow-band noise ruins the extraction of useful information from any ECG signal. Therefore, a butterworth low-pass filter was used with passband of 50 Hz, stopband of 60 Hz,

less than 1 dB of passband ripple, and more than 2.5 dB of stopband attenuation. The filter was allowed to run in the forward and inverse directions to reduce the effects of phase-shifting.

### 2.2.3. Wavelet smoothing

Wavelet smoothing has been widely used in non-stationary signal analysis to decompose any signal into a group of mother wavelets. In this work, it was performed to make the signals less variable across each time step. Furthermore, it was followed to provide a smoother morphological information of each beat within the 10 seconds segment. Noise contaminated within signals maybe be due to frequent movements or muscular motion artifacts. To reduce this effect, the smoothing was based on the maximal overlap discrete wavelet transform (MODWT). In addition, a soft version of the MODWT smoother was used of level 4 with a Daubechies 5 (db5) mother wavelet.

### 2.2.4. Removal of baseline wandering

Baseline wander is defined as a low frequency artifact that corrupts any ECG signal and occurs due to strong movements, breathing, or a displacement of electrodes during the recording. It is essential to remove such effects without affecting the overall shape of beats. To achieve this, a local polynomial regression smoother (LOESS) was used for its fast and simple ability of reducing wandering effects. Each ECG signal was fitted with a weighted least-squares function of a 0.2 scaling factor.

## 2.3. Beat-wise analysis

Instead of using the whole ECG leads' signal separately to extract features, a step was required to detect the most significant beats representing the arrhythmia morphological details. To segment each ECG signal into multiple beats, the famous Pan-Tompkins algorithm [4] was followed. The algorithm detects the location of R-peaks using several adaptive filters and thresholds. To obtain the complete QRS-complex of a single beat, 30 samples were taken before and after the location of the R-peak. The complete beat is a combination of P, QRS-complex, and T waves, therefore, the detection of the whole beat segment was done by applying windows in the forward and backward direction from the QRS-complex. The forward window was of a length of  $2/3$  of the peak-to-peak (RR) interval between the current R-peak and the next one in the signal. On the other hand, the backward window was of a length of  $1/3$  of the RR interval between the current R-peak and the previous one in the signal. The complete beat segment included a total of 250 samples. It worth noting that the first and last beats in each signal were discarded, as

many signals were corrupted at the beginning and ending of the signals. In addition, they had less samples number compared to other beats due to not having an R-peak before or after.

To identify which beat is the most significant, a correlation was applied between each beat and the beat that follows. The beat with the lowest correlation value was selected as the beat that is highly representing the arrhythmia type within the signal. Then, the selected beat was duplicated three times (750 samples per signal) to increase the length of the segment. Due to the huge imbalance in the database, each arrhythmia type was balanced by selecting the most significant beats number that balances the number of samples. For example, if an arrhythmia type had only 1000 recordings compared to the highest number of recordings per class, i.e. 2000, two lowest correlation beats were selected per recording instead of one. Both beats were labeled with the same arrhythmia type to increase the number of samples in the database.

## 2.4. Neural network architecture

The architecture of the network followed in this work was based on a combination of 1D convolutional neural networks (CNNs) and long short-term memory (LSTM) units. The CNN allows for extracting features related to the morphology as well as the spatial structure of beats. In addition, the LSTM units capture temporal features from the beats in time-domain. The combination of both networks usually allows for a better performance.

Initially, an input layer was added to allow for multi-channel 1D inputs to the network. Then, the CNN was designed to include a total of 12 layers; namely 3 1D convolutional, 3 batch normalization (BN), 3 rectified linear unit (ReLU), 2 dropout, and 1 max pooling layers. The convolutional layer had 32, 64, and 96 filters with kernel sizes of 21, 11, 3, respectively. Each convolutional layer was followed by a BN, ReLU, and 30% dropout layers. Finally, the CNN ends with a maxpooling layer followed by a BN and ReLU to reduce the dimensionality of the network.

On the other hand, the LSTM network was selected to include a total of 200 hidden-units. Each unit is connected by four gates; namely the input, output, forget, and cell gates. The structure of the network was determined to process the data in the forward and backward direction as a Bi-directional LSTM functionality.

### 2.4.1. Training configuration

The training was allowed to run for 10 epochs with a mini-batch size of 1024 to include as much samples as possible. In addition, the optimization was based on the adaptive moment estimation (ADAM) solver with learning rate of 0.001 and  $L_2$ -regularization value of 0.0005.

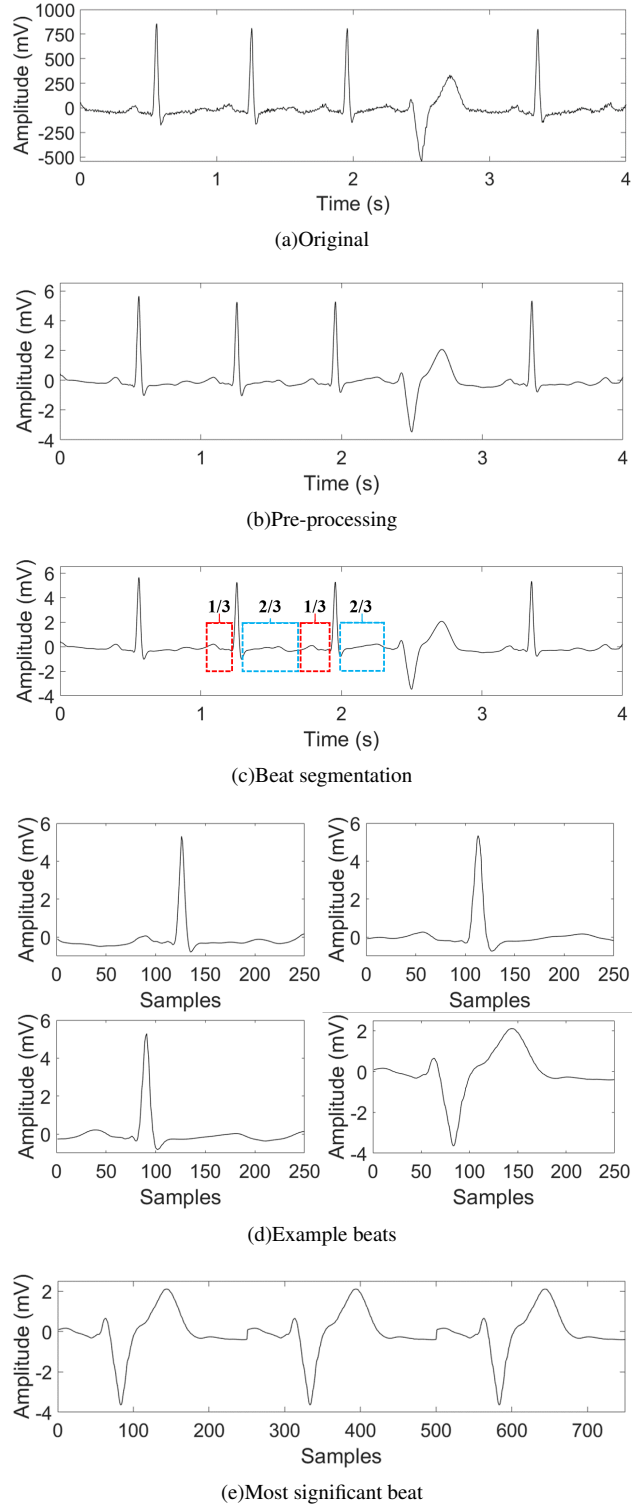


Figure 2. The preparation of the 12-lead ECG most significant beats showing: (a) original signal, (b) pre-processing, (c) beat segmentation, (d) example beats, and (e) most significant beat.

Table 1. The performance of the proposed approach applied on the training set using 10-fold cross-validation scheme.

| Folds                  | 1     | 2     | 3     | 4     | 5     | 6     | 7     | 8     | 9     | 10    | Average      |
|------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------------|
| <b>AUROC</b>           | 0.838 | 0.840 | 0.839 | 0.835 | 0.839 | 0.842 | 0.834 | 0.834 | 0.839 | 0.841 | <b>0.838</b> |
| <b>AUPRC</b>           | 0.347 | 0.360 | 0.350 | 0.331 | 0.332 | 0.349 | 0.336 | 0.342 | 0.348 | 0.361 | <b>0.346</b> |
| <b>Accuracy</b>        | 0.245 | 0.241 | 0.255 | 0.254 | 0.237 | 0.245 | 0.244 | 0.252 | 0.246 | 0.244 | <b>0.246</b> |
| <b>F-measure</b>       | 0.297 | 0.304 | 0.303 | 0.283 | 0.288 | 0.301 | 0.294 | 0.303 | 0.295 | 0.307 | <b>0.300</b> |
| <b>F-beta</b>          | 0.337 | 0.352 | 0.337 | 0.320 | 0.326 | 0.341 | 0.331 | 0.333 | 0.339 | 0.363 | <b>0.338</b> |
| <b>G-beta</b>          | 0.162 | 0.168 | 0.167 | 0.152 | 0.155 | 0.164 | 0.159 | 0.165 | 0.164 | 0.170 | <b>0.163</b> |
| <b>Challenge score</b> | 0.302 | 0.301 | 0.309 | 0.292 | 0.300 | 0.299 | 0.303 | 0.303 | 0.305 | 0.311 | <b>0.303</b> |

## 2.5. Performance evaluation

The performance of the algorithm was evaluated on both the training and testing data-sets. For the training set, a k-fold cross-validation scheme of 10 folds was followed. On the other hand, the testing set was completely hidden and evaluated by the challenge organizers. The evaluation metrics are described briefly in the challenge [3] and includes the area under the receiver-operating characteristic curve (AUROC), the area under the recall-precision curve (AUPRC), accuracy, macro F-measure, F-beta, G-beta, and the final challenge metric. Furthermore, a rank was given to each team by the end of the official phase based on the final challenge metric scoring.

## 3. Results

The results of the approach followed in extracting most significant beats is illustrated in Fig. 2. Furthermore, the performance of the k-fold cross-validation applied on the training set is provided in Table 1. The average challenge score across the 10 folds was 0.303 with an AUROC average of 0.838. On the other hand, the algorithm was evaluated on the hidden validation set and the final testing set and reached a challenge score of 0.379 and 0.146, respectively. Breaking up the final testing set, our algorithm achieved scores of 0.611, 0.342, and -0.027 on testing set 1, 2, and 3, respectively. Our team, Care4MyHeart, was ranked the 26th out of 41 entries.

## 4. Discussion and conclusions

The performance of the algorithm was not relatively high and this could be due to different reasons. Many signals did not exhibit the arrhythmia within its beats during the first 10 seconds of the recording. In addition, the imbalance in the data-set should be carefully dealt with to improve the performance. Furthermore, many patients did not show the arrhythmia within their recording due to ei-

ther being stable at the time of recording or due to a damaged recorded ECG. Future works will focus on the evaluation of the database in a signal-by-signal basis to ensure signals with an arrhythmia type truly exhibit this abnormality within the recording. In addition, developing an approach to analyze the recordings with only leads of high impact instead of the whole 12-lead signals.

## Acknowledgments

The authors would like to acknowledge Khwaja Y. Hasan from the Cardiology Department at Cleveland Clinic, Abu Dhabi. This work was supported by the Healthcare Engineering Innovation Center at Khalifa University, Abu Dhabi, UAE (Grant No: 8474000132).

## References

- [1] Mendis S, Puska P, Norrving B, World Health Organization, et al. Global atlas on cardiovascular disease prevention and control. World Health Organization, 2011. ISBN 9789241564373.
- [2] Yang W, Si Y, Wang D, Zhang G. A novel approach for multi-lead ECG classification using DL-CCANet and TL-CCANet. *Sensors* 2019;19(14):3214.
- [3] Perez Alday EA, Gu A, Shah A, Robichaux C, Wong AKI, Liu C, Liu F, Bahrami Rad A, Elola A, Seyedi S, Li Q, Sharma A, Clifford GD, Reyna MA. Classification of 12-lead ECGs: the Physionet/Computing in Cardiology Challenge 2020. *Physiol Meas In Press* 2020;.
- [4] Pan J, Tompkins WJ. A real-time qrs detection algorithm. *IEEE Transactions on Biomedical Engineering* 1985;BME-32(3):230–236.

Address for correspondence:

Ahsan H. Khandoker  
 Department of Biomedical Engineering, Khalifa University  
 PO Box 127788, Abu Dhabi, UAE  
 ahsan.khandoker@ku.ac.ae