A Bio-toolkit for Multi-Cardiac Abnormality Diagnosis Using ECG Signal and Deep Learning

Akash Kirodiwal[1], Apoorva Srivastava [1], Ashutosh Dash[1], Ayantika Saha[1], Gopi Vamsi Penaganti[1], Sawon Pratiher[1], Sazedul Alam[2] , Amit Patra[1], Nirmalya Ghosh[1], Nilanjan Banerjee[2]

[1] Department of Electrical Engineering, IIT - Kharagpur, West Bengal, India
[2] CSEE, University of Maryland, Baltimore County, USA

Abstract

Automated cardiac abnormality detection from an ever-expanding number of electrocardiogram (ECG) records has been widely used to assist physicians in the clinical diagnosis of a variety of cardiovascular diseases. Over the last few years, deep learning (DL) architectures have achieved state-of-the-art performances in various biomedical applications. In this work, we propose a bio-toolkit based on the DL framework comprising stacked convolutional and long short term memory neural network blocks for multi-label ECG signal classification. Our team participated under the name "Cardio-Challengers" in the "PhysioNet/Computing in Cardiology Challenge 2020" and obtained a challenge metric score of 0.337 in the validation data set and 0.258 in the full test data, placing us 16^{th} *out of 41 teams in the official ranking.*

1. Introduction

Globally, cardiovascular disease (CVD) accounts for over 31% of deaths around the world [1], while sudden cardiac deaths statistics projects over 80% of them to be closely related to cardiac arrhythmias [2]. Electrocardiography (ECG) records the electrical activity of the heart using electrodes placed over the skin. ECG is a widely used technique for arrhythmia detection, and analyzing the abnormal ECG waveforms or rhythms can predict the morbid status of the cardiovascular system [3]. Early-stage clinical diagnosis of cardiac abnormalities such as cardiac arrhythmias can increase the chances of heart patient's survival by predicting cardiovascular morbidity and mortality. However, manual interpretation of ECG is subjective, tedious, and domain-expertise dependent. Deep learning allows unconventional ways to interpret ECG signal exploration computer-aided diagnosis by relieving hand-crafted feature engineering for clinical setting deployment.

1.1. Problem definition

The "PhysioNet/Computing in Cardiology Challenge 2020" seeks to classify 12-lead ECG recordings for diagnosis of CVDs [4]. The contestants are required to plan and device a working, open-source model which can automatically categorize the various multi-label cardiac abnormalities present in each of the 12-lead ECG signals based on the provided clinical data only. The ranking is based on the challenge metric score obtained on the hidden test set.

The paper is organized as follows: Section 2 presents related works. Section 3 describes methodology: preprocessing, ECG template extraction, and proposed deep learning model. Experimental results and discussion are presented in section 4 and section 5 concludes the paper.

2. Related works

Traditionally, ECG classification is done on the basis of expert features [5]. Using computer algorithms these features are automatically extracted. Several researchers are trying to come up with promising sets of expert features but they are limited by quality of data and human expert knowledge. Recently deep learning has achieved promising results in various domains like computer vision, image classification and speech recognition.

Zihlmann *et al.* studied the deep CNN (24 layer CNN) along with the combination of CNN and LSTM on the same ECG data set, but they achieved a better accuracy with the combination of CNN and LSTM (24 layer CNN + 3 layer LSTM) [6]. Warrick *et al.* reduced the complexity of the network and modelled it in two main components : representation learning (using CNN) and sequence learning (using LSTM)[7]. The network has three layers of LSTM stacked above the one layer of CNN. Riberio *et al.*, 12- lead ECG was used for the classification of six cardiac abnormalities using Deep Neural Network (DNN) called as residual networks which were adapted for unidimen-

Figure 1. System Architecture.

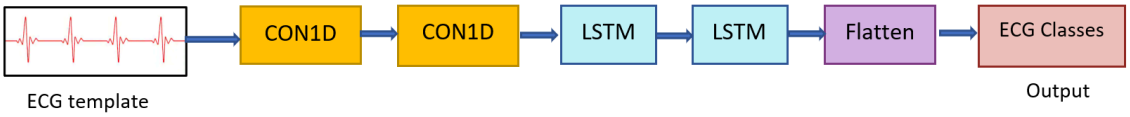

Figure 2. Deep Neural Network Architecture.

sional signals [8]. Our model is based on deep learning and is loosely inspired by [6], [7], [8] where the combinations of Convolution Neural Network (CNN) and Long Short term Memory (LSTM) were studied.

3. Methodology

3.1. ECG Pre-Processing

The training set provided by the Physionet/Computing in Cardiology Challenge 2020 [4] comprises 43101, 12-lead ECG recordings. Amongst the 12 leads, lead II being a bipolar lead with its electrodes placed on the right arm and left leg, is mostly used for identifying cardiac arrhythmias owing to its proximity to the cardiac axis and it encompasses the best view of the P and R waves [9], [10]. Hence, we selected lead II for further processing. Signals are pre-processed to remove the baseline wander, muscle artifacts and high frequency noise by using finite impulse response (FIR) bandpass filter with a passband in the 3-45 Hz frequency range.

3.2. Template Extraction

From Lead II, the R peaks were captured using Hamilton algorithm [11]. The signal between the R peaks is considered which incorporates the P and T waves of the ECG signal. Each cardiac cycle is re-sampled to 400 samples and averaged over each subject. This generates a subject wise averaged cardiac cycle. The average cardiac cycle is split in two equal parts and swapped to generate the templates.The templates were normalized in amplitude with respect to the maximum absolute value of the individual template. Representative ECG templates extracted from the Lead II ECG Signals are shown in Fig. 6.

3.3. Model Overview

In the paper we have proposed, an ECG classifier based 1-dimensional CNN (CNN-1D) and LSTM. We have implemented a DNN based model using the Keras library of the python programming language. The hidden layers of the model have two CNN-1D layers and two LSTM layers along with the input and the output layer. Figure 1 depicts the proposed pipeline and Fig. 2 shows the deep learning framework explored for automated ECG classification.

3.3.1. Input layer

In ECG pre-processing, the templates of shape (400 x 1) were generated and given as input to the CNN layer.

3.3.2. CNN-1D layer

CNN is one of the most popular neural network architectures and 1D CNN has been widely used for classification of time series data, for example, to classify cardiovascular abnormality using ECG signal [12], [13]. 1D CNN layers are capable to extract the features from a short segment of the raw signal, where the features do not depend upon the location within the segment. The CNN are used for image data but in unidimensional data it is efficient in capturing the spatial information. This has made it a prime choice for our model. The CNN model has n number of kernels (or filters) of size pxq, where p will be smaller than the input length. Each filter convolves with the input to create a feature map. In our model, we used two layers of CNN, where, each of the layers has 64 filters of size 4x1 and a stride of 1.

3.3.3. LSTM layer

LSTM is one kind of Recurrent Neural Network (RNN) introduced by Hochreiter and Schmidhuber [14]. It is also suitable for the classification of time series data, speech recognition, handwriting recognition, etc. LSTM layers have the ability to remember patterns in the data for long duration. They are efficient in capturing the temporal dependencies. In our model we have used two LSTM layers, the first layer has 32 units and the second layer has 12

units.

3.3.4. Output layer

The output layer is a densely connected neural network layer with an output size of 27x1. In order to get the probabilistic output from the model, the sigmoid activation function is used in the output layer. The output shows the probability of occurrence of each class. An optimum threshold is applied at the output to get output in binary fashion. 1 shows the presence of class and 0 shows the absence of the class. In this way we computed the multi label output.

4. Results and Discussion

4.1. Experiment Design

The training and testing data is split in 4:1 ratio. We have used the Adam optimizer with a learning rate of 0.005 and the binary cross-entropy loss function to compile the model. During the training of the model, the batch size is 256 and the number of epochs is 1000. We used an early stopping criterion based on the loss in each epoch. If the loss was not decreasing for 8 consecutive epochs the training ended. During the training of the model, we also calculated the best possible threshold value for each class that later we used during testing.

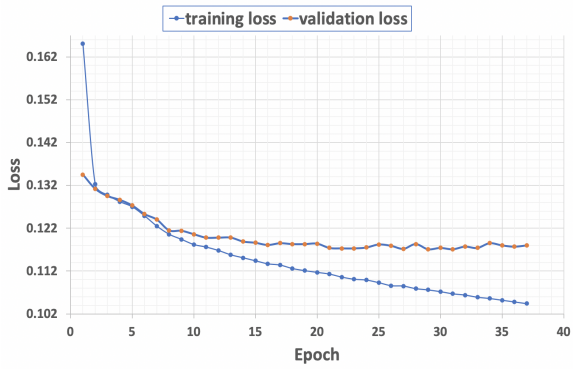

Figure 3. Variation of loss by epochs

Table 1. Different metric scores.

Scores	Validation Dataset	Test Dataset
Macro F-measure	0.105	0.116
Macro AUROC	0.498	0.498
Macro AUPRC	0.060	0.072
Challenge Metric	0.337	0.258

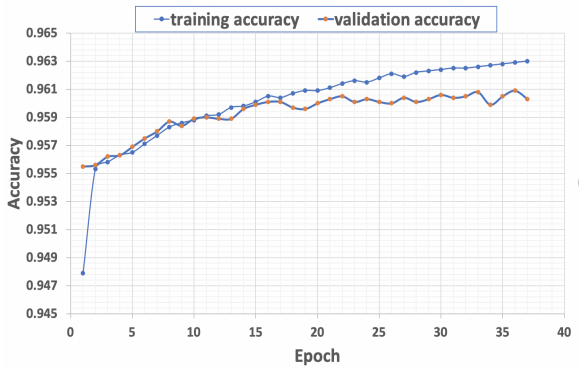

Figure 4. Variation of accuracy by epochs.

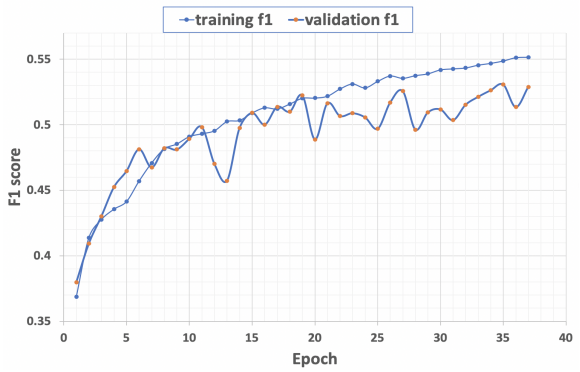

Figure 5. Variation of F1 score by epochs.

4.2. Model Compilation Results

The DNN based model received the challenge metric [4] score of 0.337 when we used a threshold optimizer. For a fixed threshold value of 0.005 the Challenge metric score was 0.314.

From the physionet challenge site we collected 43101 data files and extracted templates from them. After removing the samples with the NaN values, we had 42920 files. We used 64% of the available samples for training, 16% for validation of the model and the remaining 20% for testing. Figure 3 shows the loss for the training and validation dataset. Figure 5 shows the corresponding F1 score and Fig. 4 shows the variation in accuracy with the training epoch. Table 1 tabulates the score for different scoring metrics, including the challenge metric for the validation and test database.

4.3. Discussion

Figure 6 shows ECG templates for NSR (Normal synus rhythm), AF (Atrial fibrillation), I-AVB (1st degree AV block) and LAnFB (Left anterior fascicular block) classes. Here we have considered only four classes to explain the physiological features using the templates. AF, having