ECG-Based Abnormal Heartbeat Classification: A Deep Learning Approach for Arrhythmia Detection

Aditya Kendre

Cumberland Valley High School

Rationale

Electrocardiograms (ECG) have created a profound impact in the field of cardiology, specify in recognizing of heart arrhythmias. Non-invasive arrhythmia analysis is based on 10 electrodes that reflect the electrical activity on ECGs. An estimated three million cases of arrhythmia occur in the United States yearly (Mayo Clinic). Diagnosing this disease early is the key to one’s wellness, yet 18% of cardiologists misinterpreted ECGs containing atrial fibrillation (Anh et al, 2006). With the recent advancements in technology, Machine Learning algorithms such as Deep Neural Networks (DNNs), allow a computer to learn features and identify patterns within a given dataset. On the basic level, DNNs receive input data, and through a series of weights and biases, outputs a confidence value in all possible labels of the dataset, similar to a human’s neural network. Furtherance in the accuracy of abnormal heartbeat classification will allow cardiologists to accurately, and efficiently recognizing arrhythmia before becoming prevalent in one’s wellbeing.

Research

Research Question: This research project will examine whether a classifier will be able to accurately identify abnormal heartbeat in ECGs.

Hypothesis: If an image classifier received a supervised dataset of heart arrhythmia of ECGs, then the image classifier will allow an accurate identification of arrhythmia.

Expectation: The image classifier should reach an accuracy of above 82%.

After gathering data from medial databases that provide the public access to annotated ECGs (e.g., PhysioNet 2017), the data must be processed before it can be fed into a Convolutional Neural Network (CNN) (computational model), commonly termed Data Pre-processing. Data Pre-processing evolves normalizing, augmenting, and randomizing, the dataset to ensure the model learns and not memorizes the given training data. Normalizing data evolves scaling all values in a single ECG with a certain range (-1 to 1). The process involves calculating the absolute value of all values (in millivolts) in the single ECG, taking the absolute max of the single, and diving all values by the max. Data augmenting is the process of increasing the amount and diversity of data (e.g., Zero Bursts). Zero Bursts were added by randomly setting values in a sample to 0, representing the machines failure to detect any electrical active from the heart. The data is then fed into the CNN using the Adam optimizer and the Cross-Entropy Loss function, which adjusts the weights of the CNN to increase the accuracy of CNN’s predictions.

Randomize data

* + Normalizing data
  + Data augmentation
    - Zero Bursts - implementing random zeros in data to replace data with errors
    - Random Resampling - changing the sample rate of data collection to increase dataset size and avoid overfitting
* Convolutional Neural Network Model
  + Time dependent features
  + Frequency dependent features

=

1. Gather a dataset of annotated ECGs
2. Determine type of classifier used to learn dataset features
3. Analyze results using Gradient Decent and Mean Loss function

Risks and Safety:

This research project involves no risks or safety concerns.

References

Alfaras, Miquel, Soriano, & Silvia. (2019, July 3). A Fast Machine Learning Model for ECG-Based Heartbeat Classification and Arrhythmia Detection. Retrieved October 30, 2019, from https://www.frontiersin.org/articles/10.3389/fphy.2019.00103/full.

Mayo Clinic. (2019, April 2). Heart arrhythmia. Retrieved October 30, 2019, from https://www.mayoclinic.org/diseases-conditions/heart-arrhythmia/symptoms-causes/syc-20350668?utm\_source=Google&utm\_medium=abstract&utm\_content=Cardiac-arrhythmia&utm\_campaign=Knowledge-panel.

Srinivasan, N. T., & Schilling, R. J. (2018, June). Sudden Cardiac Death and Arrhythmias. Retrieved October 30, 2019, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6020177/.