Paper: ECG Signal Data Processing Analysis

Introduction

ECGs, or Electrocardiograms, measure electrical activity from the heart. As a result of the widespread use of ECGs in research settings, an abundance of data exists, providing insight into cardiac behavior under varying physical conditions. However, raw ECG data isn't fit for analysis. In this project, public ECG datasets were processed using a bandpass filter and R-peak detection, exactly how real biomedical engineers transform ECG data into usable and interpretable measurements.

Data Sources

The specific ECG data samples used in the program were sourced from PhysioNet. The resting sample was specifically found in the MIT-BIH Normal Sinus Rhythm Database and the stressed sample was found in the MIT-BIH ST Change Database. The samples were processed and analyzed in Python after importing NumPy, SciPy, Matplotlib, and WFDB.

Methods

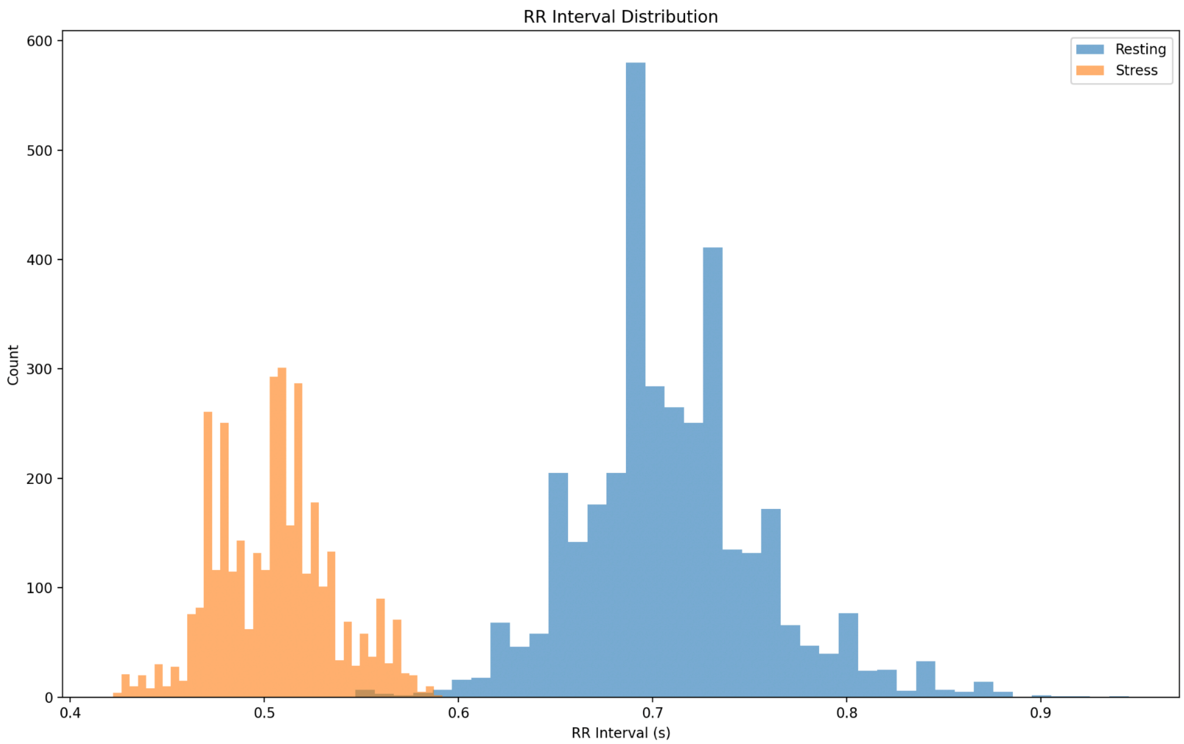
Using a bandpass filter, a filter that suppresses interference, noise, and error in the data due to motion, the program removes excess noise from the data sample. For higher frequencies, a fourth-order butterworth filter was chosen, cleaning high-frequency noise while not heavily distorting the dataset. It also gives enough of a steep roll-off while not distorting the ECG data excessively. QRS complexes have energy that spans around 4-30 Hz, and the bandpass filter filters from 5-15 Hz in order to suppress noise while maintaining the data itself. This range also mitigates the effects of possible error in the data, like movement and breathing.

The largest magnitudes of each QRS complex were detected with the 'gqrs_detect' tool sourced from the WFDB library. This tool was used for its strong QRS detection and capability for use in both resting and stress datasets. It labels R-peaks when the data is R-dominant and labels S-peaks when the data is S-dominant. Using the QRS peaks detected, the program calculates the heart rate, SDNN (standard deviation of normal-to-normal intervals), and heart rate variability using RR intervals. RR intervals are the intervals between two consecutive R-peaks.

Initial data segments in the stress dataset were not fit for processing and therefore were removed prior to graphing. To match the two samples, the program also removes the same amount of data in the resting dataset and removes data from the longer dataset in order to match them in length for processing.

Results

The final data showed that at rest, while heart rate stayed lower, the data also had a higher HRV (heart rate variability). During exercise, heart rate was elevated but had lower variability. The RR histogram generated visualizes this difference:



Histogram of RR interval distribution for the stress/rest datasets

These findings can be understood when considering the parasympathetic and sympathetic nervous systems and the dominance of one or the other. The autonomic nervous system itself controls the heart. The parasympathetic and sympathetic nervous systems refer to the two branches of this system. The parasympathetic branch is in charge of slowing heart rate and increasing HRV, giving flexible variation to the heart rate if needed. On the sympathetic branch, the heart speeds up and gets “locked into” a faster rhythm, with evenly spaced RR intervals. When one of these branches is in control, it is called dominance, and determines how the heart will behave. During parasympathetic dominance, extremely small physical or mental changes can add more variation to heart rate. This explains why there is higher RR interval variation in the resting sample and why the stress sample shows more consistent RR intervals

Real-World Relevance

In real-world biomedical engineering, the same techniques are applied to real datasets for patient care, as data must be processed, filtered, and prepared for analysis. These techniques, like QRS detection and bandpass filtering are also utilized in heart rate monitoring devices like workout heart monitors, ECG machines, and important machines used for stress testing. This processed data is especially important when finding arrhythmia and monitoring stress in high-risk patients.