Big Data Medical Diagnostics System

Tabitha Beavers, Andrew Binns, and Michael Wells

Faculty Advisors: Dr. Stansifer and Dr. Grossman Dept of CSS and ECE, Florida Institute of Technology

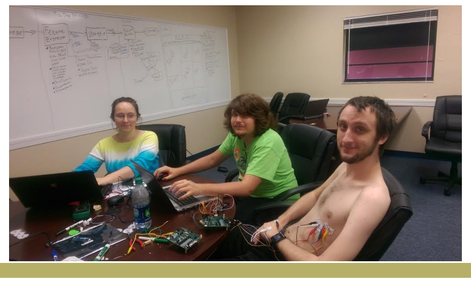
Introduction

Misplaced ECG leads are estimated to be the cause of as much as 4% of misleading results[1] given to patients, occasionally causing fatalities[2].

This project seeks to replace current leads with a pad of ECG electrodes, then use machine learning algorithms to select the optimal signal; additionally running diagnoses on acquired data.

Hardware Interface

The system has 16 individual 24-bit ADC channels feeding 2 Audino Due units, which handle buffering. These units in turn feed an embedded computer which uploads to our central server.

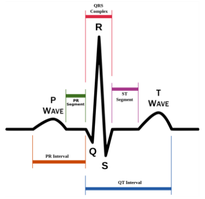


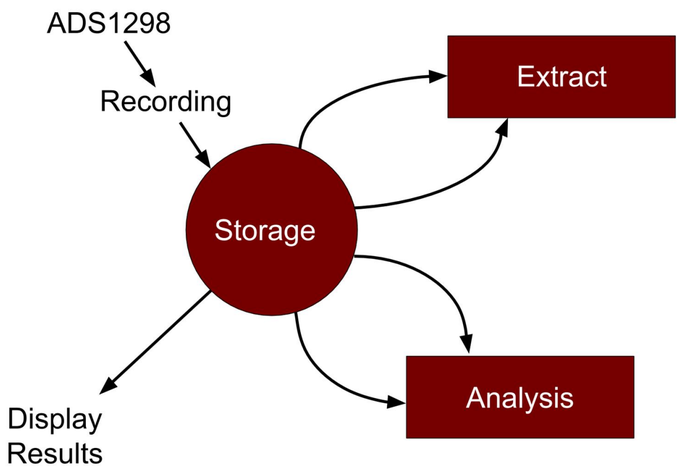
Early test run with an 8 lead patch

Feature Extraction

In order to use machine learning algorithms, certain characteristics (features) are extracted from a windowed set of data. These include:

- Selective wavelet transformation
- Length of time of standard cardiac segments
- Weight, age, and gender
- Discrete Fourier Transform (DFT) results



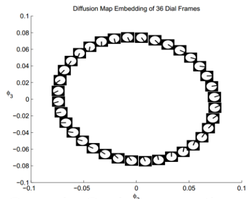


Implementation

Storage is implemented using a MySQL database on a central server. We then have two seperate server processes which handle Feature Extraction and Analysis methods respectively. This is designed to allow for highly scalable systems.

Analysis Methods

Diffusion maps describe the similarity between data points by embedding data from a higher dimensional space onto a lower one.



Sample Code separating images using diffusion maps

Diffusion maps use the eigen-decomposition of a Markov probability matrix which is created a distance metric based on the feature set[3].

References

1. Tang, Stephen, et al. "Electrical Features Associated With the Occurrence of Ventricular Fibrillation in STEMI Patients." *Circulation* 130.Suppl 2 (2014): A16174-A16174.
2. Rudiger, Alain, et al. "Electrocardiographic artifacts due to electrode misplacement and their frequency in different clinical settings." *The American journal of emergency medicine* 25.2 (2007): 174-178.
3. Fox, J. and Beavers, T., Smith, K. and Peter, A.M. and Tenali, G.B. (2014) Non-Linear Dimensionality Reduction: An Evaluation of Diffusion Based Analysis, Technical Report TR-2014-03, The AMALTHEA REU Program, Summer 2014