Short-Term Research: Algorithm for arrhythmia detection

Nikolay Shvetsov

January 2019

1 Introduction

According World Health Organization(WHO), cardiovascular diseases is number one cause of death in the world, 17,5 mln. people died every year(31,6% of all death). Arrhythmia one of the representative of cardiovascular diseases. Arrhythmia refers to some irregular change from normal rhythms. It is very import to detect arrhythmia correct. One of the way to detect heart diseases is to use Electrocardiogram (ECG), ECG consist from 3 main part: P wave, QRS complex and T wave demonstrated electrical activity of heart. In order to detect arrhythmia and anomaly in ECG,machine learning algorithms are widely used. Such methods like Decision Trees, Random Forests, SVM, Naive Bayes, also CNN are widely used[1]. In this research we try use methods of sliding windows like application of Topological methods[2] of ECG analysis, and after try to apply Wasserstein distance for it, in order to compare normal ECG parts with ECG parts where arrhythmia were detected.

2 Metodology

2.1 Sliding window

First of all, we need to reorganized our Time Series to point cloud, for that reason we will use method described in "Sliding windows and Persistence:an application of topological methods to signal analysis" by Jose A. Perea and John Hare. This method called $SW1Per$, which stands for Sliding Windows and 1-dimensional Persistence Scoring. So,

$$SW_{M,\tau} f(t) = \begin{bmatrix} f(t) \\ f(t+\tau) \\ f(t+M\tau) \end{bmatrix}$$

So, $f(t)$ is function defined on some interval. The integer M and real number τ, both greater than 0. The sliding window embedding of f based at $t \in R$ into R^{M+1} is the point. Choosing different values of t gives a collection of points called a sliding window point cloud for f. A critical parameter for this embedding is the window-size $M\tau$.

So, as for result some graphs were made, that can show how, algorithms works.

The figure was created from 4 950 point(see part 4 Figure), or 137,5 sec. of measurements. This data is ECG with arrhythmia, which was taken from MIT-BIH database(num. 123). According description, on this data set arrhythmia observed, beginning from 1350 point. So, as we can see in the figure, the first 1349 points has some trend of location, which caused by normal heart beat, normal R-R distance. The points, with arrhythmia were highlighted blue color, and in graphs we can see some bias, that show us about some anomaly detection, in this case - arrhythmia.

The second goal of this research, was understand algorithm of Wasserstein Distance[3] working. Wasserstein Distance defined between probability distributions on a given metric space M. The main goal of apply Wasserstein distance, was to learn work with it and try to do it for some images, and after to the point's cloud. Wasserstein distance:

$$W_p(\mu, v) := (\inf_{\gamma \in \Gamma(\mu, v)} \int_{M*M} d(x,y)^p d\gamma(x,y))^{1/p}$$

Idea was to apply it for some images and learn how it works, after that try to apply it for the cloud of points.It was necessary to compare cloud of points from normal ECG and from ECG arrhythmia. The idea was, that Wasserstein distance must be smaller if clouds of points similar, and large if they are very different.

3 Conclusion

As a conclusion, this work was like first introduction with topic of my Master thesis, I made pool of articles, and tried some methods from it. The main result was implementation of $SW1Per$ method and Wasserstein distance. In the end, there was a vision for further work on the project appeared.

References

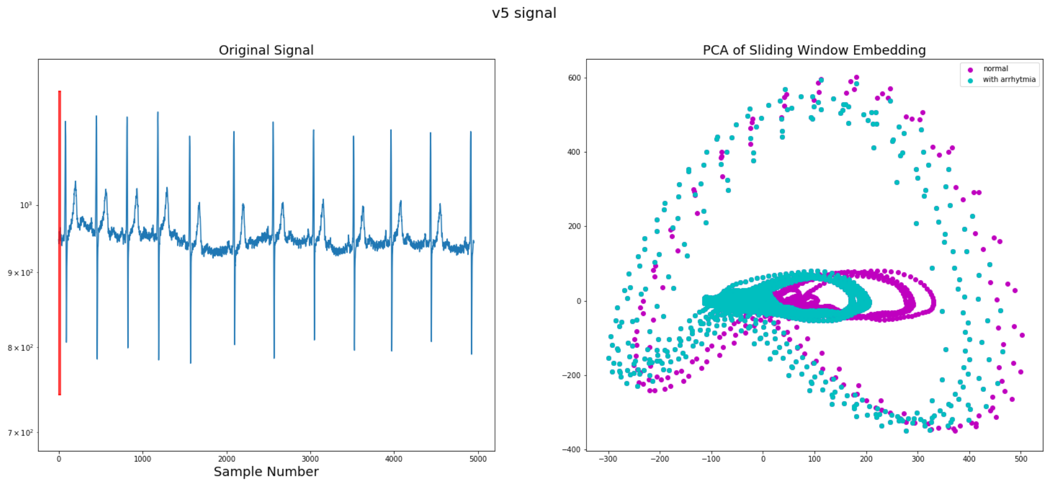
[1] Giulia Guidi and Manas Karandikar *Classification of Arrhythmia using ECG data*
`http://cs229.stanford.edu/proj2014/Manas`

[2] Jose A. Perea, John Hare *Sliding windows and Persistence:an application of topological methods to signal analysis*
`https://arxiv.org/pdf/1307.6188.pdf`

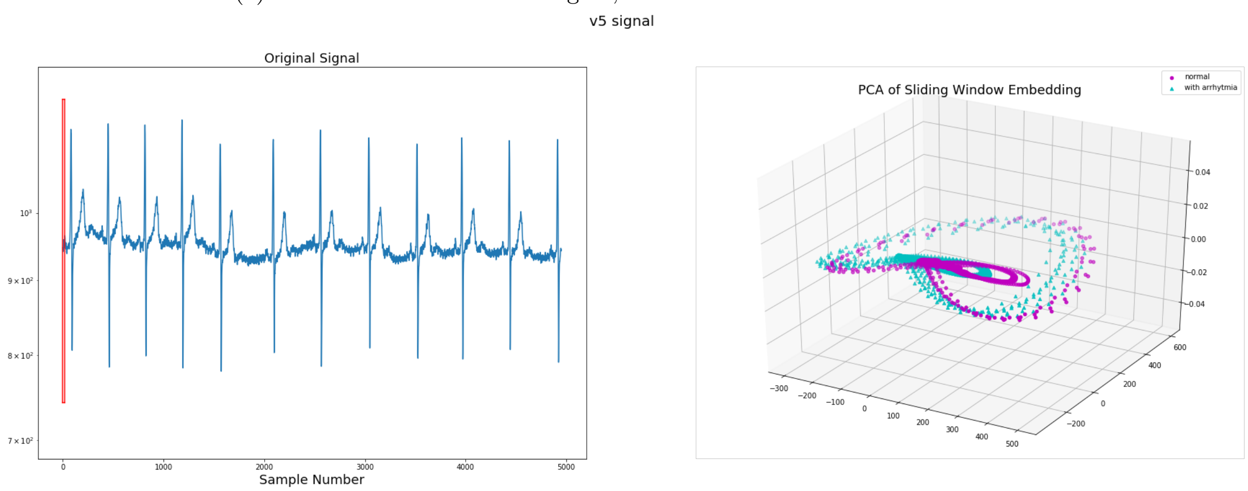
[3] Michael Muskulus,Sjoerd Verduyn-Lunel *Wasserstein distances in the analysis of time series and dynamical systems*
`http://www.math.leidenuniv.nl/reports/files/2009-12.pdf`

[4] Jonathan Hui *GAN—Wasserstein GAN WGAN*
`http://www.math.leidenuniv.nl/reports/files/2009-12.pdf`

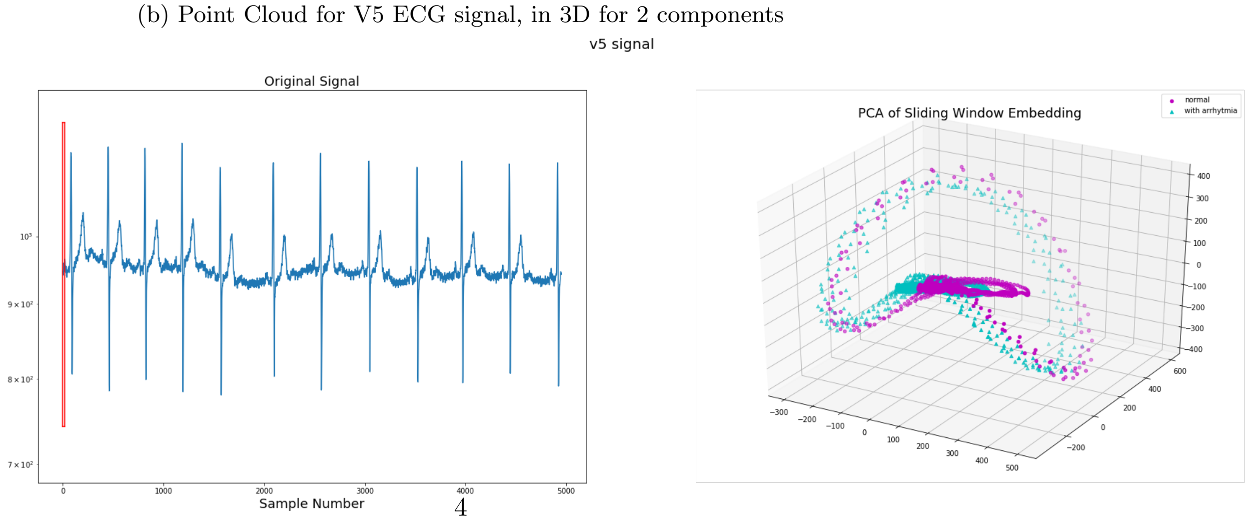
4 Figures



(a) Point Cloud for V5 ECG signal, in 2D

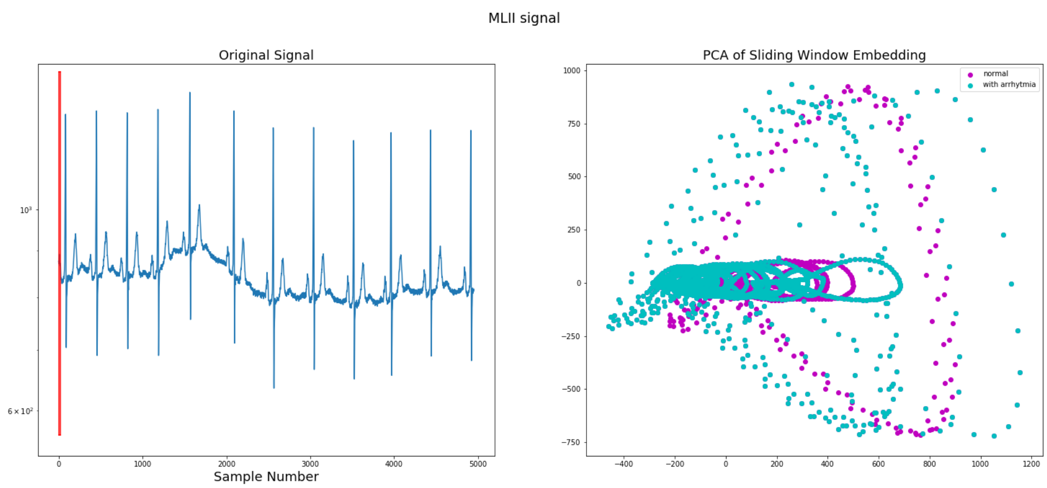


(b) Point Cloud for V5 ECG signal, in 3D for 2 components

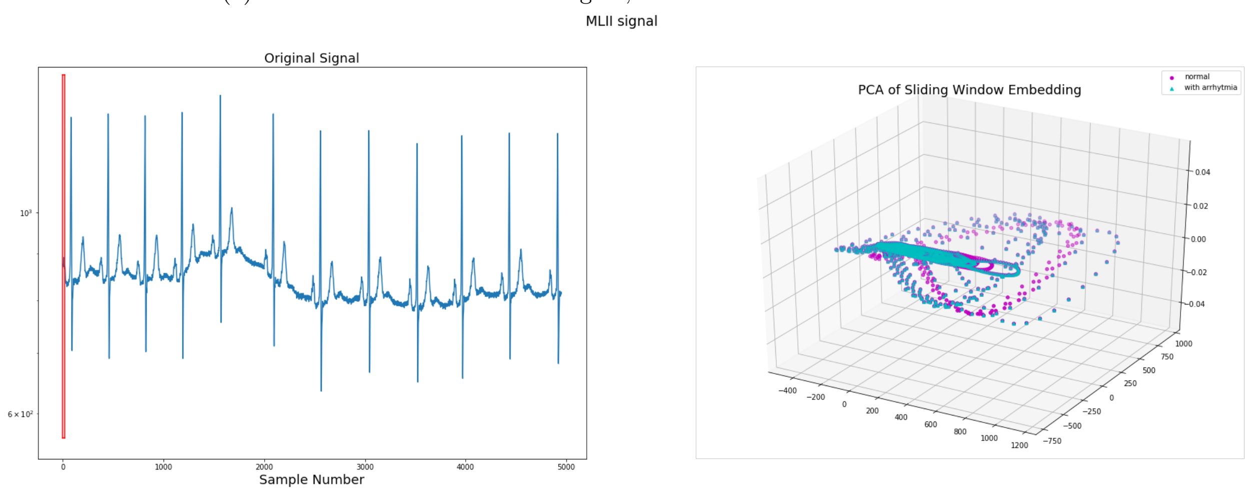


(c) Point Cloud for V5 ECG signal, in 3D for 3 components

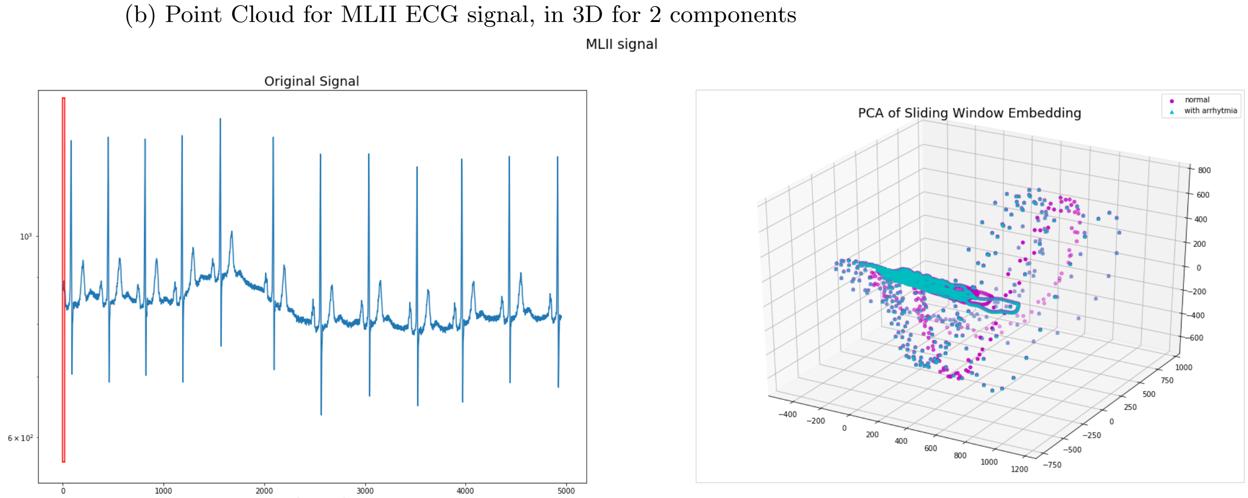
Figure 1: Example of working $SW1Per$



(a) Point Cloud for MLII ECG signal, in 2D



(b) Point Cloud for MLII ECG signal, in 3D for 2 components



(c) Point Cloud for MLII ECG signal, in 3D for 3 components

Figure 2: Example of working $SW1Per$