Department of BioMedical Engineering
Technion University

# A feature extraction approach to the classification of 12-lead ECGs

*Assaraf David, Intermediate report*

AIMLab.

04-05-2020

# Contents

# 1. Summary

In this report, I am going to expose the several steps of my work in order to be didactic in my approach. The macro-approach I will be leading is a feature extraction approach. I will be extracting ECG morphological features in order to compute relevant features specific to every pathology considered in this challenge (9 pathologies). This extraction of ECG morphological characteristics is for now done thanks to the wavedet algorithm. In this first report, I will focus on this wavedet algorithm and the problems I face with the extraction of morphological characteristic points.

# 2. The wavedet algorithm

The wavedet algorithm is a state-of-the are algorithm in the recognition of ECG morphological features. This algorithm is only available on Matlab. We have therefore implemented a Matlab wrapper in order to use this code in Python

**Input:** digital ecg, frequence of recording, R-peaks of the ecg (*optional*)

**Output:** matlab struct containing morphological features of the ECG

The morphological features output by the wavedet algorithm are: **POn**, **P**, **POff**, Pprima, Pscale, Ptipo, **QRSOn**, **Q**, **R**, Rprima, **S**, **QRSoff**, qrs, RinQRSOff, RinQRSOn, **TOn**, **T**, **TOff**, Tprima, Ttipo, Tscale, Ttipoon, Ttipooff, contadorToff, QRSonsetcriteria, QRSoffcriteria, QRSpa, QRSpp, QRSmainpos, QRSmaininv. I have hihglighted the morphological features that we use until now. There are two choices to be made when we use the above algorithm: whether to use our own R-peaks or not (and then the algorithm will implement its own peaks detection algorithm), and whether we preprocess or not the signal before using the wavedet algorithm. We will talk about both choices in what follows, and see how our choices affect or not the detection of morphological features of our ECG.

## 2.1 Preprocessing of the signal

The signal we get as input from the competition is a raw ecg. Therefore, there are several artifacts that lower the quality of our signal: powerline interferences, tremor noise...

### 2.1.1 Powerline interferences

#### 2.1.1.1 Bandpass filter

The powerline interferences problem is being solved thanks to a bandpass filter implementing IIR/FIR filtering at 50/60Hz. Since we aim at removing high frequence residual noises, we applied a pass-band filter that leveraged IIR/FIR lfilter. The lowcut frequency has been set to 0.1Hz whereas the highcut has been set to 15Hz. The results are displayed here, drawn from leads randomly selected from the first 20 samples:

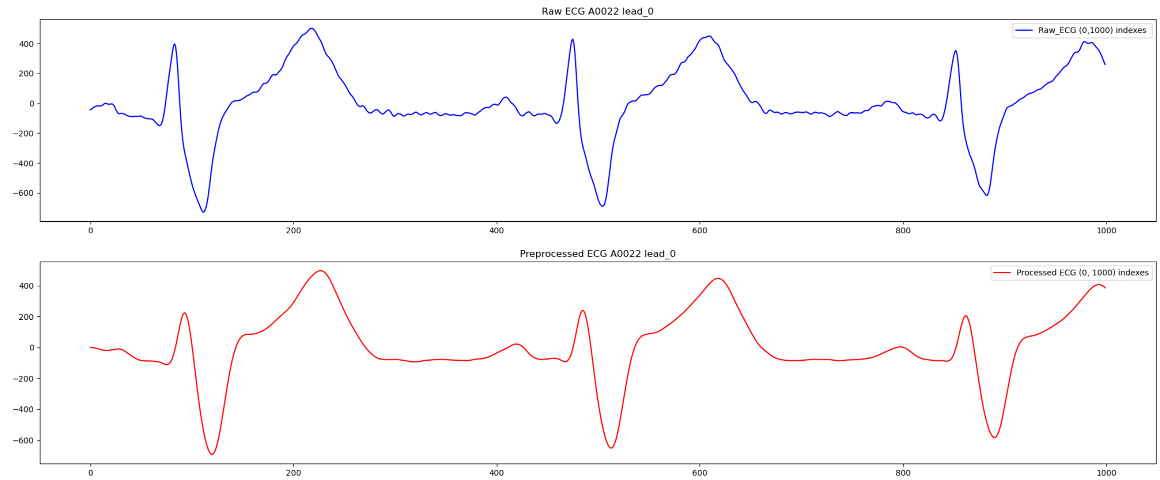


Figure 2.1: (0:1000) indexes of an ecg before and after bandpass filtering

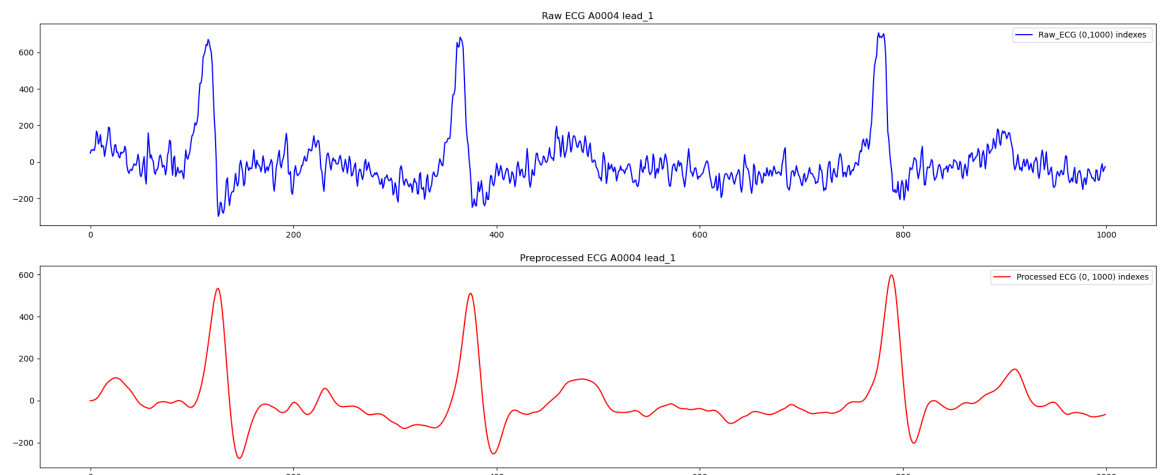


Figure 2.2: (0:1000) indexes of an ecg before and after bandpass filtering

#### 2.1.1.2 Wavelet denoising technique and integration

We used the denoisewavelet function from the skimage.restoration package in python.
Our initial signal is:

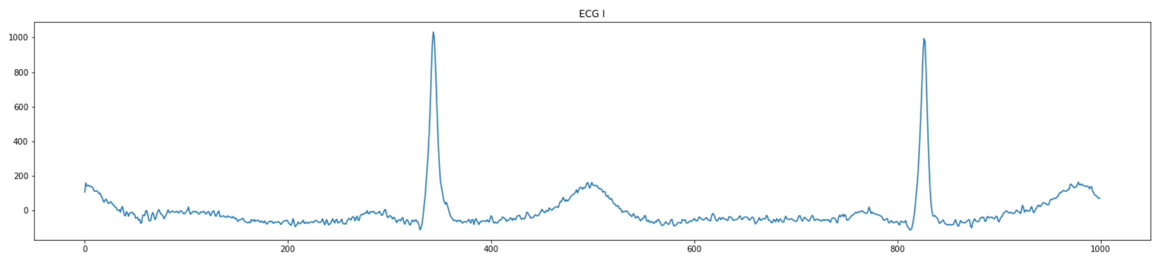


Figure 2.3: (0:1000) indexes of a raw ecg

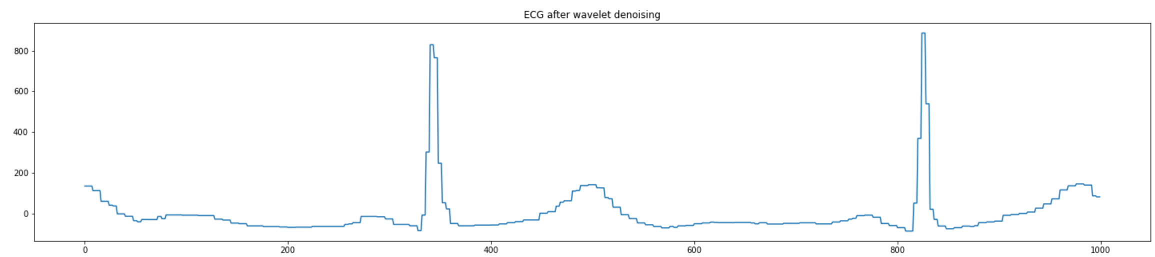


Figure 2.4: (0:1000) indexes of an ecg after applying wavelet denoising

However, as we can see, our signal is not smooth. This is why we have further integrated it.

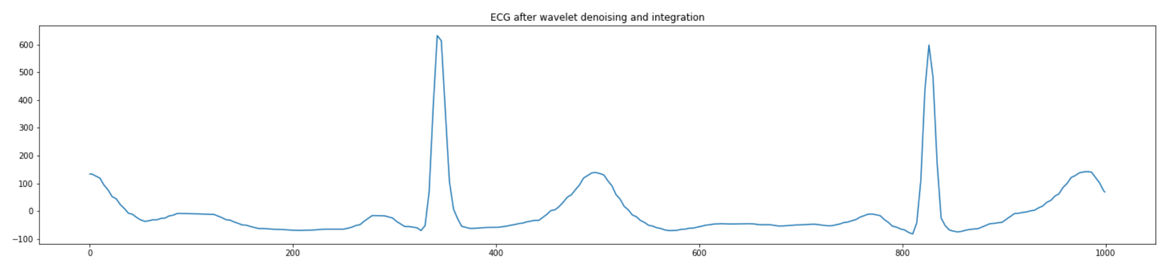


Figure 2.5: (0:1000) indexes of an ecg after applying wavelet denoising and integrating it

#### 2.1.1.3 Wavelet denoising, integration and bandpass filter

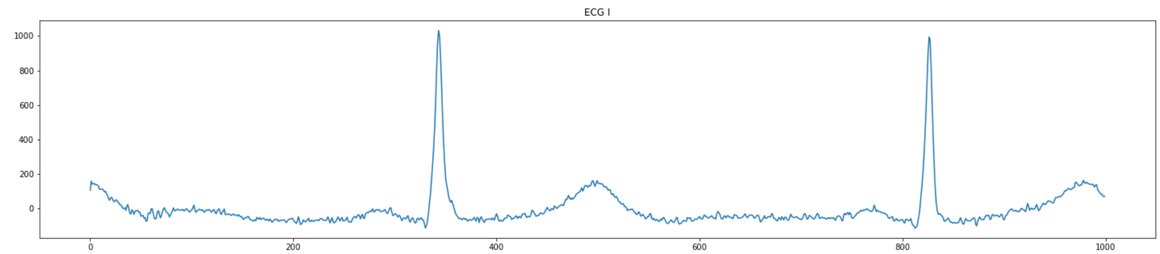


Figure 2.6: (0:1000) indexes of a raw ecg

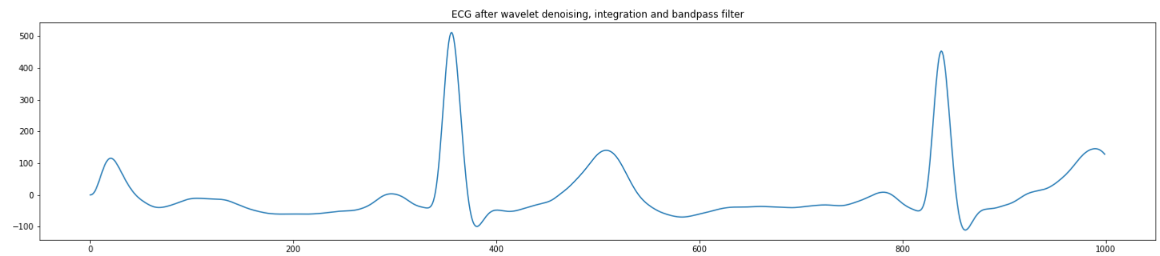


Figure 2.7: (0:1000) indexes of a preprocessed ecg

Therefore, this will for now be our final preprocessing technique, since it is the one where remain all the interesting variations of our signal while suppressing the artifacts.

### 2.1.2 Tremor noise

For now, I have found no artifacts in the signals I have inspected after preprocessing.

## 2.2 The wavedet algorithm

In this section, we are going to present preliminary considerations on the output of the wavedet algorithm. Our results will be based on the first 500 samples of the competition, from which we extracted morphological features from 6 leads, for the characteristic points aforementioned.

### 2.2.1 Number of characteristic points extracted

#### 2.2.1.1 When we perform ecg preprocessing or not

Let us say that we take in every lead the R-peaks as reference, meaning that if we detect 35 R-peaks in a sample, we expect to get 35 points of each category (POn, ...). The table presented here displays the proportion of morphological features detected accross the 3000 first leads we have inspected. We will conduct this experiment when performing preprocessing of the signal or not.

| | POn | P | POff | QRSOn | Q | R | S | QRSOff | TOn | T | TOff |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $filteredecg$ | 0.86727 | 0.86727 | 0.86727 | 1 | 0.16 | 1 | 0.85 | 1 | 0.9984 | 1 | 1 |
| $rawecg$ | 0.86727 | 0.86727 | 0.86727 | 1 | 0.16 | 1 | 0.85 | 1 | 0.9984 | 1 | 1 |

Figure 2.8: Results of the proportion of morphological features extracted

Therefore, we can see that choosing whether performing prior ecg preprocessing or not has no impact in the quantity of morphological features detected. Another pain point that can be seen thanks to this experiment is the **low number of Q points** detected. Indeed, from this table, we can assume that, if we do not consider the Q points, at least 7/10 qrs complexes are completely discovered.

#### 2.2.1.2 When we change the R-peaks detector

Another experience we can be conducting is try to find out is the qrs detector we are using has an influence on the detection of morphological features. For this experience, we will be using the default R-peaks detector of the wavedet algorithm and the gqrs detector from the wfdb package, and we will be using both on filtered ecg, since the preprocessing has no influence on the detection of complexes.

| | POn | P | POff | QRSOn | Q | R | S | QRSOff | TOn | T | TOff |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $Wavedetdefault$ | 0.86727 | 0.86727 | 0.86727 | 1 | 0.16 | 1 | 0.85 | 1 | 0.9984 | 1 | 1 |
| $gqrs$ | 0.86727 | 0.86727 | 0.86727 | 1 | 0.16 | 1 | 0.85 | 1 | 0.9984 | 1 | 1 |

Figure 2.9: Results of the proportion of morphological features extracted

### 2.2.2 Visualisation of the morphological features extracted

We are going to visualize some morphological features extracted, by differenciating whether the signal has been preprocessed or not.

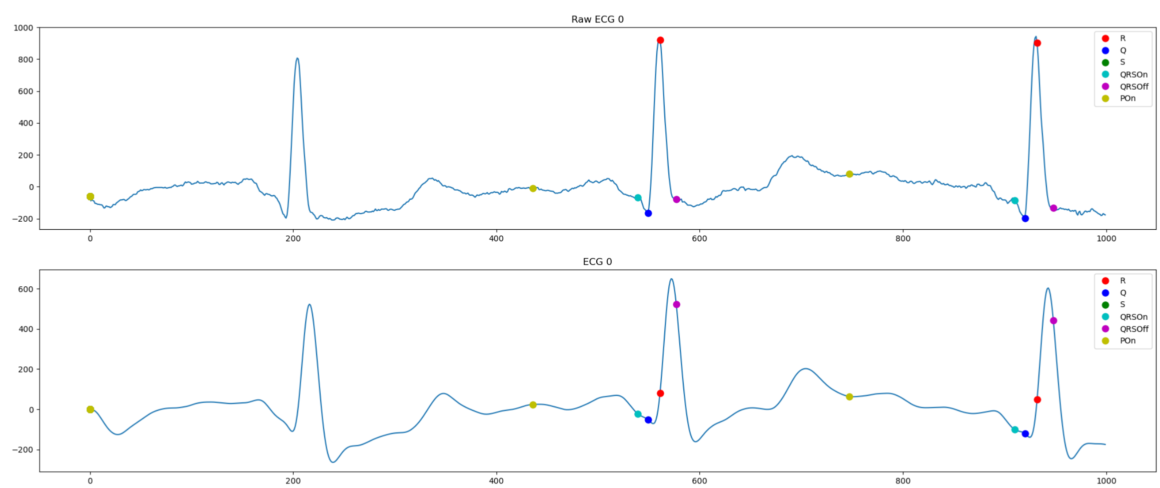


Figure 2.10: (0:1000) indexes of morphological features extracted for a lead preprocessed or not

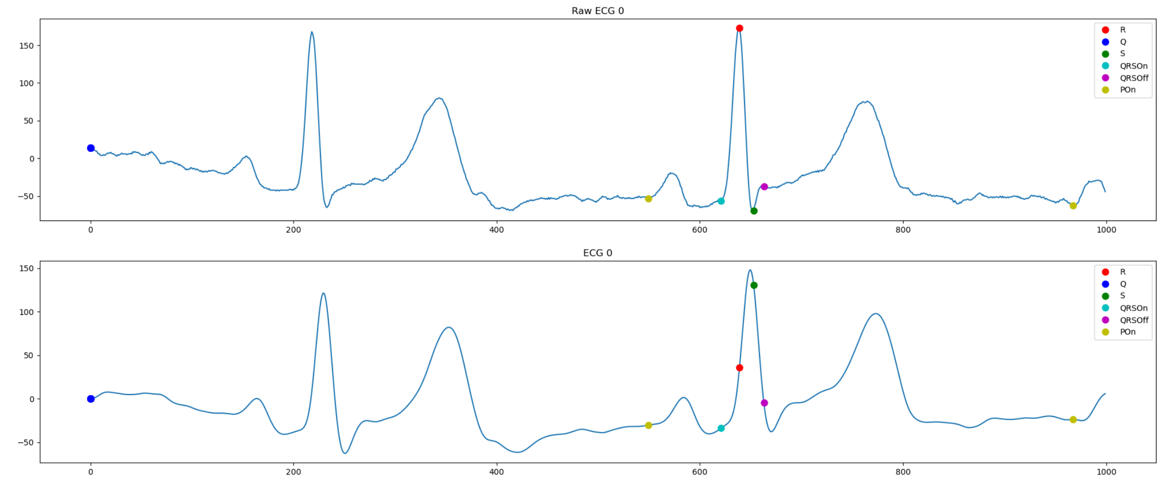


Figure 2.11: (0:1000) indexes of morphological features extracted for a lead preprocessed or not

### 2.2.3 Questions

Why is it so that we detect such a low quantity of Q points? Why is it that it appears that the morphological feature detection is more accurate on the raw ecg than on the preprocessed ecg?