

APPENDIX C - CREATION OF THE R PEAK DETECTION ALGORITHM

An important detail in the creation of this algorithm was the fact that it did not need to operate in real time, that is, the R peaks were identified during the ECG measurement. This happens because, since the classification of the machine learning model and generation of images of heartbeats requires that the signal is already ready, this can also be done with the algorithm. This algorithm is used both in the preparation of the signal for training the model, as well as in the microcontroller for the functioning of the prototype.

With that, the algorithm works as follows: Considering that the ECG record is sampled at 100Hz for 10s, which is the standard adopted by the German database, firstly, the highest value of the 1000 points that exist in the record is found. Then, the reference value of the algorithm is assigned as being half of the highest value found previously. The functioning of this assignment directly depends on the filtering performed in the previous step, as it will be possible to remove certain noisy components that distort the signal and make it oscillate, mainly baseband noise. With filtering, the record is approximately centered at 0 and with R peaks with similar amplitudes.

In this way, all the points of the signal in question are traversed, and if any of them is greater than the calculated reference value, the next step in the algorithm is evaluated, where the 5 points following the one identified as being greater are evaluated. than the reference value, and then determine which is the greater among them. This will correspond to an R peak. It then saves both the sample value of this R peak as well as the position of this sample in the signal as a whole.

The reason for evaluating the additional 5 points is to ensure that you get the highest value of the QRS complex according to the 100Hz sampling frequency used. As one can have values greater than the reference value within the QRS complex after the R peak, it is necessary to skip a certain number of samples after its identification. Knowing that heart rate varies within a range of 40 bpm to 200 bpm, it is possible to calculate the maximum number of samples without the occurrence of an R peak as a function of frequency, given by:
$$: \frac{60 f_s}{200} = \frac{60 \cdot 100}{200} = 30$$
 Therefore, once an R peak is found, 30 samples of signal analysis can be skipped without the possibility of identifying another R peak. This process is carried out during the entire length of the signal.

Consequently, it is concluded that the R peak detection algorithm needs to have the signal already sampled. Its operation consists of traversing the 10s ECG record, looking for points above a reference value within a certain range, where the largest sample in this region is an R peak. The algorithm, when tested, showed a huge efficiency when compared to the application of other similar methods, in addition to providing faster execution.

The tests that were carried out in order to verify the performance of the algorithm consisted of comparing the results obtained by the algorithm created from the signals captured from the SimMan simulator with the results considering the same input signals using a Python extension

tool known as BioSPPy, which can be accessed from the following link: <https://github.com/PIA-Group/BioSPPy>.

This extension brings together several signal processing and pattern recognition methods aimed at analyzing biosignals, having support for several biological signals such as ECG, EEG, EMG and respiration. Its operation includes several tried and tested signal analysis techniques such as filtering, frequency analysis and R peak detection, which in turn was created from the study of P.S. Hamilton in the article “Open Source ECG Analysis Software Documentation”.

As for the test itself, four base signals were used, composed of ECG recordings with SR with and without noise and in an analogous way for AFIB. By applying the BioSPPy algorithm, a total of 337 R peaks were detected.

After that, running the algorithm created in this article, the detection of each of the 337 R peaks identified in the previous algorithm was observed, as well as some extra indications in the AFIB signals, where when visually verifying the signal actually consisted of R peaks.

Therefore, the created algorithm was able to correctly identify the R peaks of SR and AFIB signals in comparison with the BioSPPy algorithm, as well as finding R peaks in AFIB signals that went unnoticed by the reference algorithm.

The tests and their results can be accessed in the project repository, more specifically, in the folder reserved for the development of this algorithm.