PhysioZoo ECG: Digital electrocardiography biomarkers to assess cardiac conduction

Sheina Gendelman[1], Shany Biton[1], Raphaël Derman[2], Eran Zvuloni[1], Jeremy Levy[1,3], Snir Lugassy[1]
Alexandra Alexandrovich[1] and Joachim A. Behar[1]

[1]Faculty of Biomedical Engineering, Technion-IIT, Haifa, Israel
[2] Rambam Health Care Campus, Haifa, Israel
[3] Faculty of Electrical Engineering, Technion-IIT, Haifa, Israel

Abstract

Introduction: The electrocardiogram (ECG) is a standard tool used in medical practice for identifying cardiac pathologies. Because the necessary expertise to interpret this tracing is not readily available in all medical institutions or at all in some large areas of developing countries, there is a need to create a data-driven approach that can automatically capture the information contained in this physiological time series. Yet, contrary to heart rate variability measures, a field which has seen the development of standards, advanced toolboxes and software, very little open tools exist for ECG morphological analysis. The primary objective of this work was to identify and implement clinically important digital ECG biomarkers for the purpose of creating a reference toolbox and software for ECG morphological analysis. Methods: The epltd algorithm was used for R-peak detection. We used a zero-phase filter with passband 0.67Hz - 100Hz to remove baseline wander and high frequency noise. We used a Notch filter at 50/60Hz to remove the power-line interference. ECG fiducial points were detected using the well-known open source wavedet algorithm. A total of 22 biomarkers were engineered including 14 extracted from intervals and segments duration and 8 from waves characteristics. Results and discussion: the result of this work consists of a Python toolbox termed "pebm" and its user interface termed "PhysioZoo ECG" for data visualization and analysis. The software is available at physiozoo.com under a GNU GPL licence. The pebm toolbox may be used to provide new physiological information on cardiac conduction as well as used as a source of readily handcrafted features for training machine learning models.

1.   Introduction

The electrocardiogram (ECG) is a standard tool used in medical practice for identifying cardiac pathologies. Be-cause the necessary expertise to interpret this tracing is not readily available in all medical institutions or at all in some large areas of developing countries, there is a need to create a data-driven approach that can automatically capture the information contained in this physiological time series. Yet, contrary to heart rate variability measures, a field which has seen the development of standards and advanced toolboxes and software [1, 2], very little open tools exist for ECG morphological analysis. The primary objective of this work was to identify and implement clinically important digital ECG biomarkers ("pebm") for the purpose of creating a reference toolbox for ECG morphological analysis.

2.   Methods

Overview: For an individual ECG lead the process consists of (Figure 1): (1) Preprocessing; (2) ECG fiducial points detection; (3) Engineering of ECG biomarkers and (4) Summary statistics

Preprocessing: Before computing the ECG morphological biomarkers, prefiltering of the raw ECG time series is performed to remove the baseline wander as well as remove high frequency noise. Specifically, we include a zero phase second-order infinite impulse response bandpass filter with the passband of 0.67Hz - 100Hz to remove baseline wander and high frequency noise. We include an optional Notch filter that can be set to 50 or 60Hz to remove the power-line interference. Because wavedet [3] was originally designed to implicitly perform frequency filtering while performing the discrete wavelet transform then this step is to be considered as optional. In other words prefiltering may be enabled or disabled in the toolbox and software.

ECG fiducial points detection: For a given channel, the epltd algorithm was used for R-peak detection. The ECG fiducial points are detected using the well-known wavedet algorithm [3]. Duration biomarkers include segments and
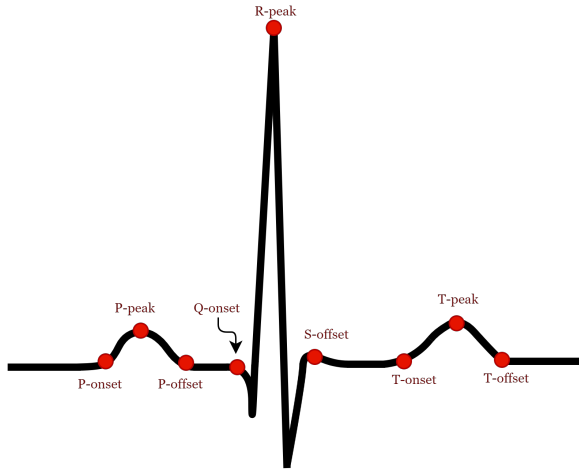
Figure 1. Fiducials points detected by wavedet [3].

intervals. An ECG segment is the region between two waves. An interval is a duration of time that includes one segment and one or more waves. Briefly, wavedet uses the dyadic wavelet transform of the ECG signal to identify the location of the peaks of the P, Q, R, S R' and T waves as well as the P, QRS and T waves boundaries. The fiducials and their notations used in PhysioZoo ECG are shown on Figure 1.

Engineering of ECG biomarkers Using the fiducial points ECG biomarkers are engineered for individual ECG cycles. When a biomarker cannot be engineered because some fiducial points could not be detected by wavedet then the feature was marked as a NaN. For an ECG channel a total of 14 features are extracted from intervals duration (Table 1) and 8 from waves characteristics (Table 1) to describe the ECG morphology. Some of these features were described in the work of Biton et. al. [4]. In addition we included a measure of signal quality, bsqi [5].

Summary statistics: For a specified time window the five summary statistics (median, min, max, Q1 and Q3) are computed for all ECG biomarkers. This is performed by considering all cycles available in this window and for which a cycle specific biomarker was not marked as NaN.

Notations: Since we computed MOR features for each channel we use the index (i) to specify on what channel it was computed on in the instance of an input with multiple channels. For example, PR^{V1} will refer to the PR interval computed on the lead V1.

3.   Results

The results of this work consists of the pebm toolbox and PhysioZoo ECG user interface. The software is avail-

able at physiozoo.com under a GNU GPL licence. Figure 2 shows an example of a 2-lead ECG loaded in PhysioZoo ECG and for which the fiducial points characterizing the waves peaks are shown. Summary statistics are shown in the lower panel under "Duration" and "Amplitude". The software supports three input formats (mat, txt, wfdb).

Figure 3 shows violin plots for 100, 30sec windows computed from each recording of the MIT NSR database. The density is mainly contained within the normality ranges (reported in Table 1) which provides a first layer of quality control on the toolbox implementation.

4.   Discussion and conclusion

In the case of multi-channel ECG and with the current implementation we decided to run epltd on each individual channel to obtain channel dependant R-peak locations. It is also possible to consider detecting the R-peak on a single channel to save computational time and input that to wavedet on the individual channel for extracting the other channel dependant fiducials. The tradeoff between delniation accuracy and computing time is to be evaluated.

This work presents a first step towards creating a standard library and software for ECG morphological analysis. The pebm toolbox may be used to provide new physiological information on cardiac conduction as well as used as a source of readily handcrafted features for training machine learning models. The engineered biomarkers need to be validated against human annotations. The list of biomarkers presented may also be augmented as many other ECG biomarkers, albeit less standard in clinical practice have been used in research for the purpose of analysing the ECG signal. Additional contributions from the research community are welcomed.

Acknowledgements

References

[1] Behar JA, Rosenberg AA, Weiser-Bitoun I, et al. Physiozoo: a novel open access platform for heart rate variability analysis of mammalian electrocardiographic data. Frontiers in physiology 2018;9:1390.

[2] Vest AN, Da Poian G, Li Q, Liu, et al. An open source benchmarked toolbox for cardiovascular waveform and interval analysis. Physiological measurement 2018; 39(10):105004.

[3] Martínez JP, Almeida R, Olmos S, Rocha AP, Laguna P. A wavelet-based ECG delineator: evaluation on standard databases. IEEE Transactions on Biomedical Engineering 2004;51(4):570–581.
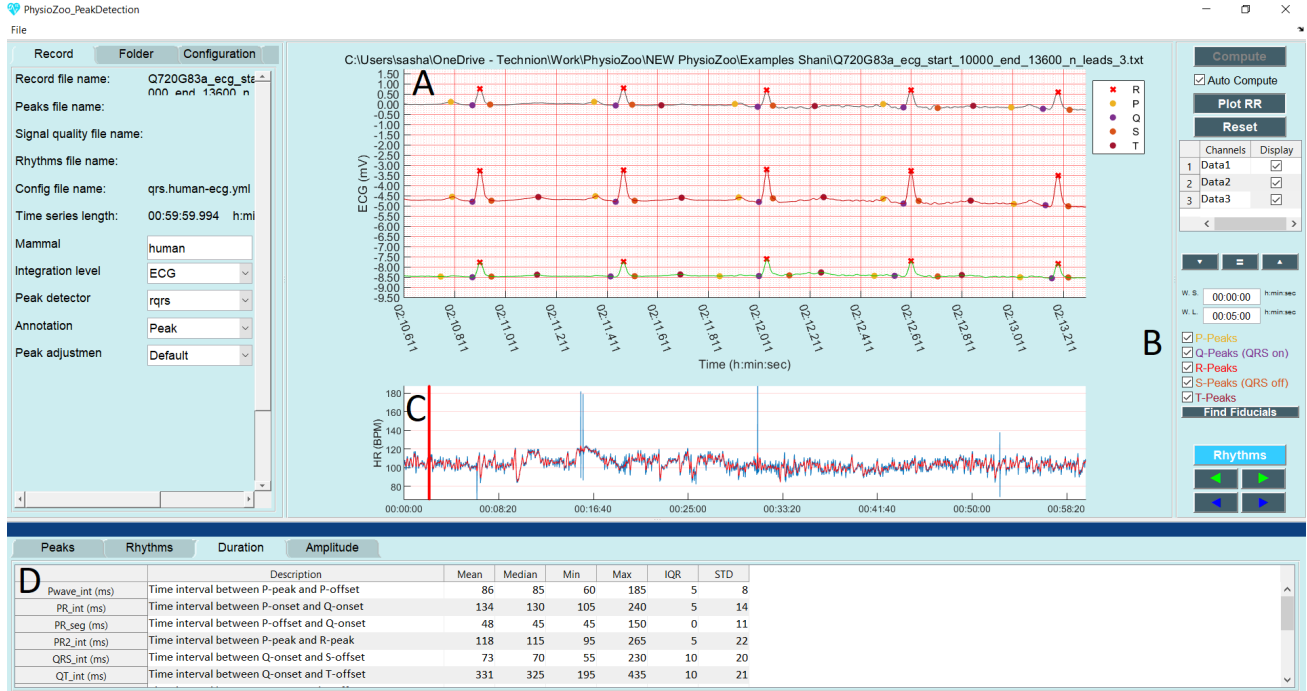
Figure 2. Screenshot of the PhysioZoo ECG interface. (A) ECG signal shown for the selected window; (B) setting panel for the fiducial points detection; (C) beat-to-beat time series and (D) summary statistics for the duration biomarkers.

	Biomarkers	Definition	Normality range
Interval and segments			
1	$Pwave_{int}$	Time interval between P-peak and P-offset	<120ms
2	PR_{int}	Time interval between P-onset and Q-onset	120-200ms
3	$PR2_{int}$	Time interval as defined by Mao et al. [6]	
4	PR_{seg}	Time interval between P-offset and Q-onset	
5	QRS_{int}	Time interval between Q-onset and S-offset	<110ms
6	QT_{int}	Time interval between Q-onset and T-offset	
7	QT_{cB}	Corrected QT interval (QTc) by Bazett [7]	<450ms (w), <430ms (m)
8	QT_{cFrid}	QTc by Fridericia [8]	
9	QT_{cF}	QTc by Framingham [9]	
10	QT_{cH}	QTc by Hodges [10]	
11	$Twave_{int}$	Time interval between T-onset and T-offset	
12	TP_{seg}	Time interval between T-offset and P-onset	
13	RR_{int}	Time interval between consecutive R peaks	
14	R_{dep}	Time interval between Q-onset and R-peak	
Waves characteristics			
15	$Pwave$	Amplitude difference between P-peak and P-offset	
16	$Twave$	Amplitude difference between T-peak on and T-offset	
17	$Rwave$	R peak amplitude	
18	$Pwave_{Area}$	P wave interval area defined as integral between P-onset and P-offset	
19	$Twave_{Area}$	P wave interval area defined as integral between T-onset and T-offset	
20	QRS_{Area}	QRS interval area defined as integral between Q-onset and S-offset	
21	ST_{seg}	Amplitude difference between S-offset and T-onset	
22	$Jpoint$	Amplitude 40ms after S-offset as defined by Hollander et al. [11]	

Table 1. Interval and segments duration and waves characteristics biomarkers. Normality range are from [12].
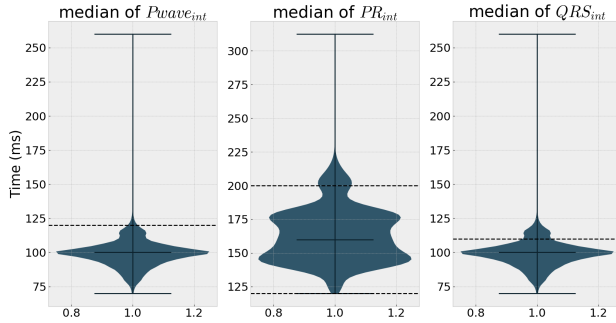
Figure 3. Violin plot for the distribution of the intervals $Pwave_{int}$, PR_{int} and QRS_{int} computed on the PhysioNet NSR database.

[4] Biton S, Gendelman1 Sheina amd Ribeiro AH, et al. Atrial fibrillation risk prediction from the 12-lead ecg using digital biomarkers and deep representation learning. European Heart Journal Digital Health 2021;.

[5] Li Q, Mark RG, Clifford GD. Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a Kalman filter. Physiological Measurement 2007;29(1):15.

[6] Mao L, Chen H, Bai J, et al. Automated Detection of First-Degree Atrioventricular Block Using ECGs. International Conference on Health Information Science 2018;156–167.

[7] Bazett H. An analysis of the time-relations of electrocardiograms. Annals of Noninvasive Electrocardiology 1997; 2(2):177–194.

[8] Fridericia L. Die systolendauer im elektrokardiogramm bei normalen menschen und bei herzkranken. Acta Medica Scandinavica 1921;54(1):17–50.

[9] Sagie A, Larson MG, Goldberg RJ, Bengtson JR, Levy D. An improved method for adjusting the qt interval for heart rate (the framingham heart study). The American journal of cardiology 1992;70(7):797–801.

[10] Hodges M. Bazett's qt correction reviewed: evidence that a linear qt correction for heart rate is better. J Am Coll Cardiol 1983;1:694.

[11] Hollander J, Blomkalns A, Brogan G, et al. Standardized reporting guidelines for studies evaluating risk stratification of emergency department patients with potential acute coronary syndromes. Annuals of emergency medicine 2004; 44(6):589–598.

[12] Surawicz B, Knilans T. Chou's Electrocardiography in Clinical Practice E-Book: Adult and Pediatric. Elsevier Health Sciences, 2008.

Address for correspondence:

Joachim A. Behar (jbehar@technion.ac.il)
Technion-IIT Technion City, Haifa 32000, Israel