PhysioZoo: The Open Digital Physiological Biomarkers Resource

Joachim A. Behar¹, Jeremy Levy¹,², Eran Zvuloni¹, Sheina Gendelman¹, Aviv Rosenberg¹,⁶, Shany Biton¹, Raphael Derman³, Jonathan A. Sobel¹, Alexandra Alexandrovich¹, Peter Charlton⁵ and Márton Á. Goda¹

¹Faculty of Biomedical Engineering, Technion, Technion–IIT, Haifa, Israel
²The Andrew and Erna Viterbi Faculty of Electrical & Computer Engineering, Technion–IIT, Haifa, Israel
³Department of Anesthesiology, Rambam Medical Center, Haifa, Israel
⁴University of Geneva, Geneva, Switzerland
⁵University of Cambridge, Cambridge, United Kingdom
⁶Faculty of Computer Science, Technion, Technion–IIT, Haifa, Israel

Abstract

PhysioZoo is a collaborative platform designed for the analysis of continuous physiological time series. The platform currently comprises four modules, each consisting of a library, a user interface, and a set of tutorials: (1) PhysioZoo HRV, dedicated to studying heart rate variability (HRV) in humans and other mammals; (2) PhysioZoo SPO2, which focuses on the analysis of digital oximetry biomarkers (OBM) using continuous oximetry (SpO2) measurements from humans; (3) PhysioZoo ECG, dedicated to the analysis of electrocardiogram (ECG) time series; (4) PhysioZoo PPG, designed to study photoplethysmography (PPG) time series. In this proceeding, we introduce the PhysioZoo platform as an open resource for digital physiological biomarker resource designed to facilitate research and unlock the potential wealth of information within physiological time series. Each PZ module offers essential functionalities, including prefiltering, handling of low-quality signals, precise fiducial point detection, and biomarker engineering. The primary objective of each module is to streamline the analysis of specific time series by providing a standardized pipeline, precise terminology, and a validated set of algorithms. Each module includes a code library and a user interface (Figure 2).

PZ is a pioneering platform that recognizes the importance of tailoring the extraction of features to the unique characteristics of signals in different physiological modalities such as ECG, SpO2, and PPG. Unlike many existing toolboxes that claim to offer “universal” features applicable to any time series, which often fall short in practice, PZ embraces the body of domain-specific knowledge associated with each signal or modality. By incorporating mathematically defined biomarkers with physiological relevance, specific to the nature of the analyzed signal, PZ ensures the extraction of accurate and meaningful information. This approach allows researchers and clinicians, who generally have less technical skills but more domain knowledge, to gain insights into various physiological phenomena in a reliable and informative manner. Furthermore, PZ empowers researchers from nonbiological fields, who may lack the domain knowledge required to work with such data, to quickly extract meaningful features, e.g. for training machine learning models.

Being an open-source initiative, PZ is committed to transparency and accessibility. Our user-centered approach incorporates intuitive toolboxes and a user interface.
Figure 1: The four modules available in PhysioZoo (PZ) enabling the processing of beat-to-beat, electrocardiogram (ECG), oximetry and photoplethysmography (PPG) time series. HRV: heart rate variability; OBM: digital oximetry biomarkers.

that simplifies data visualization and ensures accessibility for non-technical users to analyze data samples effectively. The PZ modules are compatible with various data formats, such as CSV, txt, and the PhysioNet WFDB format, allowing seamless integration with diverse datasets. Furthermore, PZ operates under the GNU General Public License, as published by the Free Software Foundation, fostering a collaborative and open community dedicated to advancing the field of physiological time-series analysis.

Each module is independently structured, documented, and managed by its original authors. This decentralized approach offers several advantages, empowering new contributors to take ownership of their projects while ensuring that all libraries remain part of an overarching initiative following a set of principles and scientific standards concerning the quality control of implemented functions. This includes validating source code against available benchmarks and/or annotated datasets, allowing the code to be open source, providing comprehensive documentation and the availability of a user interface. A set of video tutorials is available on the PZ YouTube channel (physiozoosoftware62).

2. PhysioZoo modules

A total of four PZ modules have been released (Figure 1 and Table I-I-3). The inaugural PZ module, PZ HRV, was launched in 2018, marking the beginning of this initiative. PZ HRV focuses on studying heart rate variability (HRV) in both humans and other mammals. The second module, PZ SPO2, is dedicated to the analysis of digital oximetry biomarkers (OBM) derived from continuous oximetry (SpO2) measurements in humans. The third module, PZ ECG, is designed for the analysis of electrocardiogram (ECG) time series data. Lastly, the fourth module, PZ PPG, is centered on studying photoplethysmography (PPG) time series. The release of these modules represents a significant step forward in the pursuit of advancing physiological time series analysis.

PZ HRV [1]. Heart rate variability, the variation of the time interval between heartbeats contains important information about cardiac health and the autonomic nervous system. HRV can be used to quantify the short- and long-term dynamics of heartbeats. A total of 26 HRV biomarkers were implemented as part of PZ HRV, in the mhrv toolbox. The biomarkers are divided into three standard categories from the HRV literature: time domain, frequency domain, and non-linear. We contributed a standardized, validated HRV toolbox, which can be used not only to analyze human interval data but also comes with configuration presets supporting the analysis of data from other mammals (dogs, rabbits, mice). The mhrv toolbox is implemented in MATLAB.

PZ SPO2 [2]. Pulse oximetry is routinely used to non-invasively monitor oxygen saturation levels. A low oxygen level in the blood means low oxygen in the tissues, which can ultimately lead to organ failure. The biomarkers were divided into five categories: (1) General statistics: time-based statistics describing the SpO2 data distribution, (2) Complexity: quantifies the presence of long-range correlations in non-stationary signal, (3) Periodicity: quantifies consecutive events to identify periodicity in the SpO2 signal, (4) Desaturations: time-based descriptive measures of the desaturation patterns occurring throughout the signal, and (5) Hypoxic burden: time-based measures quantifying the overall degree of hypoxemia imposed on the heart and other organs during the recording period. A total of 44 digital oximetry biomarkers (OBM) were implemented in the pbom toolbox. Studying the variability of the continuous oxygen saturation time series using pbom may provide information on the underlying physiological control systems and enhance our understanding of the manifestations and etiology of diseases, with an emphasis on respiratory diseases. The toolbox, pobm, is implemented in Python.

PZ ECG [3]. The electrocardiogram, or ECG, is a widely-used tool used in medical practice to identify cardiac disease. We implemented clinically important digital ECG biomarkers in the pecg toolbox to create a reference toolbox for ECG morphological analysis. The library and associated user interface can be used to process 12-lead ECG as well as long-term recordings, for example, from Holter monitors. A total 22 biomarkers were implemented. These were divided into two categories: (1) interval and
FIGURE 2: PhysioZoo user interface. The interface enables to load and process a single recording. Fiducial points are detected and digital biomarkers are computed over the specified window. The specific example displayed is for a three lead ECG (Holter) recording using the pecg toolbox.

Table 1: PhysioZoo modules, libraries, implementation language and year of the first public release. The libraries and user interface are available at physiozoo.com

<table>
<thead>
<tr>
<th>Module</th>
<th>Library</th>
<th>Language</th>
<th>Release</th>
</tr>
</thead>
<tbody>
<tr>
<td>PZ HRV</td>
<td>mhrv</td>
<td>MATLAB</td>
<td>2018</td>
</tr>
<tr>
<td>PZ SPO2</td>
<td>pobm</td>
<td>Python</td>
<td>2021</td>
</tr>
<tr>
<td>PZ ECG</td>
<td>pecg</td>
<td>MATLAB, Python</td>
<td>2023</td>
</tr>
<tr>
<td>PZ PPG</td>
<td>pyPPG</td>
<td>Python</td>
<td>2023</td>
</tr>
</tbody>
</table>

3. Discussion and future work

Within the field of biomedical signal processing, much work has been done, through initiatives such as PhysioNet [10], to make datasets openly available, standardized, and interoperable. Comparatively, less effort has been invested in making analytical tools openly available. Although there are some existing tools for each signal modality, they may often lack scientific validation or comprehensive documentation and generally do not have the benefit of a cohesive framework, which PZ provides. Thanks to the researchers who have contributed to PZ to date, four modules have been released. These enable the streamlined and reproducible analysis of HRV, ECG, SpO2, and PPG time series. PZ modules have been used in a number of scientific studies [11][20].
There are two main limitations of our efforts. The first is based on the mixture of MATLAB and Python across the constituents of the PZ platform. This is currently due to depending on the original wavedet algorithm, the mhrv toolbox, and the PZ graphical user interface, all of which are implemented in MATLAB. An effort to convert each module’s codebase to Python is needed. The second limitation is the need for engagement of new independent contributors to the platform. We encourage scientists and engineers to recognize the great potential of open source scientific tools to contribute to the growing set of PZ modules. In addition, there is ample opportunity to contribute new functions to the existing modules and toolboxes. Some ideas include detection of arrhythmias, cardiac abnormalities, abnormal beats in ECG or PPG, respiratory rate estimation and respiratory signal analysis or algorithms for electroencephalogram time series analysis. Finally, by contributing their existing software as a new PZ module or into an existing module, researchers can also benefit from wider exposure and user-base for their tools. We thus encourage researchers, scientists and engineers to utilize and contribute to this resource, thereby facilitating open-source collaboration and high standards of research within the domain of physiological time series analysis.

References


Address for correspondence:
PhysioZoo admin (physiozoolab@gmail.com)