Article Published on: Neural Approaches to Dynamics of Signal Exchanges, 2019;395-402.

https://doi.org/10.1007/978-981-13-8950-4_35

Chapter 35
PySiology: A Python Package for Physiological Feature Extraction

Giulio Gabrieli, Atiqah Azhari, and Gianluca Esposito

G. Gabrieli (B) · A. Azhari · G. Esposito
Psychology Program, School of Social Sciences, Nanyang Technological University, Singapore, Singapore
e-mail: GIULIO001@e.ntu.edu.sg

G. Esposito
Department of Psychology and Cognitive Science, University of Trento, Trento, Italy

Article Published on: Neural Approaches to Dynamics of Signal Exchanges, 2019;395-402.

https://doi.org/10.1007/978-981-13-8950-4_35

Abstract

Physiological signals have been widely used to measure continuous data from the autonomic nervous system in the fields of computer science, psychology, and human–computer interaction. Signal processing and feature estimation of physiological measurements can be performed with several commercial tools. Unfortunately, those tools possess a steep learning curve and do not usually allow for complete customization of estimation parameters. For these reasons, we designed *PySiology*, an open-source package for the estimation of features from physiological signals, suit- able for both novice and expert users. This package provides clear documentation of utilized methodology, guided functionalities for semi-automatic feature estimation, and options for extensive customization. In this article, a brief introduction to the features of the package, and to its design workflow, are presented. To demonstrate the usage of the package in a real-world context, an advanced example of image valence estimation from physiological measurements (ECG, EMG, and EDA) is described. Preliminary tests have shown high reliability of feature estimated using *PySiology*.

Article Published on: Neural Approaches to Dynamics of Signal Exchanges, 2019;395-402.

https://doi.org/10.1007/978-981-13-8950-4_35

Introduction

Physiology, the science of the functions of living organisms and their parts, is a broad scientific discipline that encompasses elements of psychology, computer sciences, and human–computer interaction. As a subject, physiology spans across molecular and cellular components through to the levels of tissues, organs, and whole systems. Physiology provides the bridge between scientific discoveries and their application to medical science.

Analysis of physiological signals can bring advantages to several different fields. First, physiological signals are directly controlled by the autonomic nervous system (ANS), and therefore, analysis based on physiological features allows for the exclusion of potentially biased social inferences from collected data. Second, using wearable sensors and devices, it is possible to collect continuous measurements that can be used in pervasive computing and continuous medical investigation [1]. On the downside, physiological measurements are highly influenced by the presence of artifacts and noise. In order to be effectively used, hardware setup for data collection, signal processing, and feature extraction have to be carefully executed in order to maximize the signal-to-noise ratio.

From the variety of available physiological signals, electrocardiography (ECG), electromyography (EMG), and electrodermal activity (EDA) play an exceptionally important role in clinical application and have been widely adopted in emotion recognition, affective computing, and neuroscientific research [2–9].

Various software have been developed in recent years for automatic or semi- automatic analysis of physiological signals. These tools provide support for signal pre-processing and feature extraction, as well as data analysis and data reporting. Unfortunately, these programs are usually proprietary (closed-source), limiting ac- cess to the methodology used in the analysis, which subsequently makes comparison across different methods impossible. Furthermore, some of them do not allow for configuration of estimation parameters, precluding the possibility of testing different setups and tunings during feature estimation.

For these reasons, we have decided to develop a new tool for physiological feature estimation, in the form of a free and open-source Python package, released under the name *PySiology*.

The aim of *PySiology* is to provide researchers with a tool that can be easily configured for feature estimation of physiological signals, suitable for both novice and expert users. The open nature of the project allows users to assess and modify the source code, share new features, and make comparisons across new and old techniques. Also, because of the high degree of customization and ease of use, *PySiology* can be easily integrated into complex analysis structures from other soft- ware. *PySiology* (version 0.0.9) provides modules for the estimation of features from three different types of signals: ECG, EMG, and EDA.

In this paper, a brief introduction to the development workflow and status of our library is provided.

Development Workflow

The workflow of *PySiology* has been specifically developed to enhance collaboration between users. The source code of the project is provided through GitHub, a public repository that allows for external contributions from users with regard to bug tracking, the proposal of novel physiological features, and general discussion. This platform also allows users to enter sample data into the package, upload tutorials, improve the precision of estimated features, and enhance the quality of modules documentation.

Documentation is automatically generated using Sphinx, and provided in both an online and printable format, on the GitHub Repository and on Read the Docs.

Tutorials and examples are preferably uploaded in the form of *Jupyter notebooks*, while no specific format is suggested for the sample data. This allows for the presentation of different experimental situations (e.g., data collected from different devices) while maintaining high readability and reproducibility of example codes. Feature estimation is conducted using functions, classes, and objects. Prior to being added to the project, parameters should be customizable and have predefined values, methods of feature estimation should be referenced, and all functions must be documented. Data are preferred in the form of Numpy array.

Modules, Features, and Installation

Package Structure

PySiology's code is organized into sub packages, each one referring to a specific signal (e.g., ECG) or functionality (e.g., loading of sample data). Version 0.0.9 is structured as follows:

- pysiology.**electrocardiography**: ECG signal processing and feature estimation
- pysiology.**electrodermalactivity**: EDA signal processing and feature extraction
- pysiology.**electromyography**: EMG signal processing and feature estimation
- pysiology.**sampledata**: sample data for testing and educational purposes.

Feature Estimation

PySiology allows for the estimation of different physiological features, both in time and frequency domains. For each feature, expert users can customize all the used parameters (e.g., thresholds and cut-off frequencies), while standard values are provided for quick analysis and to assist novice users. A brief list of estimable features for each signal is reported in Table 35.1. More detailed information on estimable features, references, and methodology used for estimation is available in the document.

Functions and methods for signal pre-processing are provided. Basic pre-processing pipelines, including bandpass filters, phasic filters, and down sampling, are suggested.

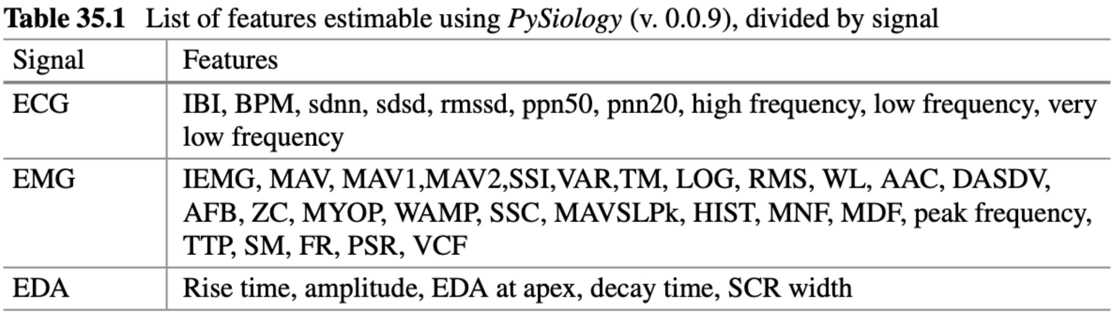


Table 35.1 List of features estimable using *PySiology* (v. 0.0.9), divided by signal

Signal	Features
ECG	IBI, BPM, sdnn, sdsd, rmssd, ppn50, pnn20, high frequency, low frequency, very low frequency
EMG	IEMG, MAV, MAV1,MAV2,SSI,VAR,TM, LOG, RMS, WL, AAC, DASDV, AFB, ZC, MYOP, WAMP, SSC, MAVSLPk, HIST, MNF, MDF, peak frequency, TTP, SM, FR, PSR, VCF
EDA	Rise time, amplitude, EDA at apex, decay time, SCR width

Installation

PySiology is available via the Python package index PyPi or from the GitHub repository. Detailed instruction for installation, including required packages and their versions, is available in the document. Installation using a package manager, such as *pip* or *conda*, is recommended.

Advanced Example: Predicting Valence of an Image from Physiological Features.

There is a longstanding tradition for emotion recognition to be investigated from physiological signals. For example, in a study on musically induced aesthetic pleasure by Ray et al. [10], morphological variation in ECG and EEG signatures was found, suggesting a correlation between a mild increase in sympathetic activity of the autonomic nervous system with the valence of the stimuli. Similar results were obtained by de Jong et al. [11, 12] in which ECG and EDA signals were collected during experiments on music and painting perception.

In this example, physiological recordings of viewers in order to estimate the valence of emotional images selected from the International Affective Picture System (IAPS) dataset are used. IAPS is a dataset consisting of 1180 emotionally evocative pictures, developed by NIMH Centre for the Study of Emotion and Attention, University of Florida [13]. Affective norms (ratings of valence, arousal, and dominance) for IAPS pictures were obtained from 18 separate studies and are provided together with the dataset.

Because of the high arousal level of selected images, we expect a high correlation between physiological activity and emotional experience [14]. To estimate the valence of an image, *PySiology* has been used, paired with classifiers provided in *scikit- learn*, a python package for machine learning and neural networks implementation.

Data Collection

From the IAPS dataset, 50 stimuli were selected for presentation and divided into two groups: low valence (N = 25, mean valence = 1.79 ± 0.14) and high valence (N = 25, mean valence = 8.06 ± 0.13). Stimuli were presented to university students (N = 58, 26 males, 33 females, mean age 21.5 ± 2, 3) for 8 s each, with an interval of 6 s between two stimuli. ECG, EDA, and EMG (*corrugator supercilii*) signals were collected on a Bitalino Revolution BT board, a low-cost device designed for physiological data collection (sampling rate: 1000 Hz, Wireless Biosignals S.A.) [15], at a resolution of 1279 × 800. Data used in this manuscript are available online [16].

Preprocessing

ECG signal was first preprocessed for noise removal, using a fifth-order bandpass filter with a high-pass cutoff at 0.5Hz and a low-pass cutoff at 2.5Hz. Baseline correction was conducted with reference to the 20 s of recording preceding the presentation of the first stimulus.

GSR signal was preprocessed for noise remotion. First, a second-order bandpass filter, with a high-pass cutoff at 0.05 Hz and a low-pass cutoff at 1 Hz, was applied. Then, the signal was down sampled using a scaling factor of 100–10 Hz. Using the median value, a median filter was then applied to remove the tonic component from the signal.

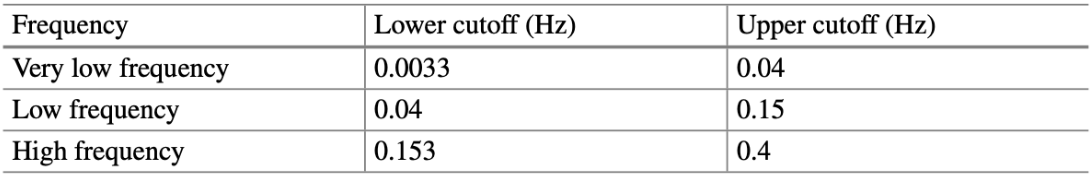
Raw EMG signal was first preprocessed for noise correction, using a second-order bandpass filter with a high-pass cutoff at 20Hz and a low-pass cutoff at 50Hz. A basic median filter was applied to separate the phasic from the tonic components, subtracting the median amplitude calculated four seconds before and after each time point.

Feature Extraction

For feature estimations, preprocessed raw signals were first segmented into epochs. Functionalities for feature estimation were then applied to each epoch.

ECG. ECG features were estimated using the submodule *electrocardiography*. From the ECG signals, both time-domain and frequency-domain features were estimated. For time-domain frequencies, peak detection was done by utilizing a minimum distance between peaks of 500 ms. Frequency-domain features were estimated using cutoff frequency values adapted from blood [17]. Used values are reported in Table 35.2.

Table 35.2 Cutoff values for ECG frequency analysis



Frequency	Lower cutoff (Hz)	Upper cutoff (Hz)
Very low frequency	0.0033	0.04
Low frequency	0.04	0.15
High frequency	0.153	0.4

GSR. GSR features were estimated using the submodule *electrodermalactivity*. The first peak with an onset of at least 1 s from stimuli presentation was used for feature estimation.

EMG. EMG features were estimated using the submodule *electromyography*. For the evaluation of zero crossing (ZC), average myopulse output (MYOP), Willison amplitude (WAMP), and slope sign changes (SSC), the threshold value was set to 0.01 (***threshold*). While estimating the mean absolute value slope (MAVLSPk), the number of segments used (***nseg*) was 3.

Classification

Recursive partitioning (decision tree) and multi-layer perceptron classifier (MLP) have been used for classification. Estimated physiological features were first reduced using principal component analysis (PCA) to six components, while IAPS valence values have been used as targets. Before being fed to the neural network classifier, input values were standardized by removing the mean and values were scaled to unit variance.

For both the classifiers, standard scikit-learn (version 0.19.1) parameter was used.

For each classifier, 100 tests have been done, from which 45 pictures were chosen for training and 5 used as test cases.

Results & Discussion

Average accuracy results, reported in Table 35.3, show that both classifiers were able to achieve an accuracy above chance with a small number of features and samples. Classifier parameters were not tuned according to the input data, suggesting that enhancement of accuracy level can be achieved. Also, having used data recorded from low-cost devices, feature extraction from clinical-grade instrumentation seems promising for high-quality estimation from raw signals.

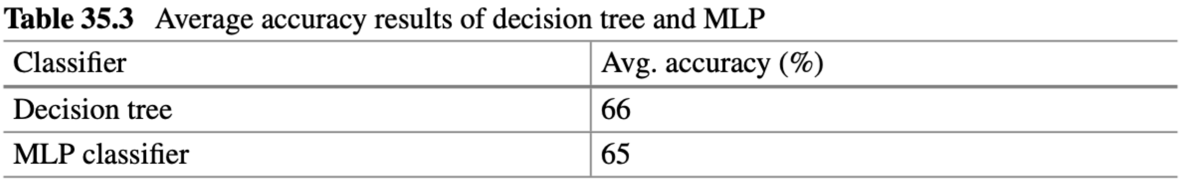


Table 35.3 Average accuracy results of decision tree and MLP

Classifier	Avg. accuracy (%)
Decision tree	66
MLP classifier	65

Future Development

PySiology is still under continuous development, despite alpha and beta versions having already been released and tested. Future development for the package will include improvement in computational duration and accuracy, implementation of novel methodologies for physiological feature estimation, modules for guided statistical analysis, and data modelling with report generation. Furthermore, we aim to add more open-source sample datasets and tutorials, with detailed instructions for every step of the conducted analysis.

Conclusion

PySiology is an actively developing open-source package for the estimation of physiological features. Its primary goal is to provide researchers with a tool that can easily be used for feature estimation and that allows for the customization of every parameter for ad hoc tuning of used algorithms. Furthermore, the package is designed to be employed by non-expert users.

Due to the open nature of this software, *PySiology* allows for comparison across different methodologies. Aside from its user-friendly features, such as tutorials, examples, suggested parameters, and analysis procedures, this package has also been developed to promote customization, high-quality documentation, and sharing of open data. In this paper, a brief introduction to the package, demonstrated within the context of an example case, has been provided, thus showcasing the accuracy of estimated features.

Article Published on: Neural Approaches to Dynamics of Signal Exchanges, 2019;395-402.

https://doi.org/10.1007/978-981-13-8950-4_35

References

1. Wagner, J., Kim, J., André, E.: From physiological signals to emotions: implementing and comparing selected methods for feature extraction and classification. In: IEEE International Conference on Multimedia and Expo. ICME 2005, pp. 940–943. IEEE (2005)
2. Pitt, M.: The use of stimulated EMG in the diagnosis of neuromuscular junction abnormality. In: Pediatric Electromyography, pp. 123–136. Springer, Berlin (2017)
3. Ahuja, N.D., Agarwal, A.K., Mahajan, N.M., Mehta, N.H., Kapadia, H.N.: GSR and HRV: its application in clinical diagnosis. In: 16th IEEE Symposium on Computer-Based Medical Systems, 2003. Proceedings, pp. 279–283. IEEE (2003)
4. Xun, L., Zheng, G.:ECG signal features election for emotion recognition. Indonesian J. Electr. Eng. Comput. Sci. **11**(3), 1363–1370 (2013)
5. Esposito, G., Yoshida, S., Ohnishi, R., Tsuneoka, Y., del Carmen Rostagno, M., Yokota, S., Okabe, S., Kamiya, K., Hoshino, M., Shimizu, M., et al.: Infant calming responses during maternal carrying in humans and mice. Curr. Biol. **23**(9), 739–745 (2013)
6. Castroflorio, T., Bargellini, A., Rossini, G., Cugliari, G., Deregibus, A., Manfredini, D.: Agreement between clinical and portable EMG/ECG diagnosis of sleep bruxism. J. Oral Rehabil. **42**(10), 759–764 (2015)
7. Truzzi, A., Setoh, P., Shinohara, K., Esposito, G.: Physiological responses to dyadic interactions are influenced by neurotypical adults' levels of autistic and empathy traits. Physiol. Behav. **165**, 7–14 (2016)
8. Truzzi, A., Bornstein, M.H., Senese, V.P., Shinohara, K., Setoh, P., Esposito, G.: Serotonin transporter gene polymorphisms and early parent-infant interactions are related to adult male heart rate response to female crying. Front. Physiol. **8**, 111 (2017)
9. Truzzi, A., Poquérusse, J., Setoh, P., Shinohara, K., Bornstein, M.H., Esposito, G.: Oxytocin receptor gene polymorphisms (rs53576) and early paternal care sensitize males to distressing female vocalizations. Dev. Psychobiol. **60**(3), 333–339 (2018)

10. Ray, G., Kaplan, A.Y., Jovanov, E.: Morphological variations in ECG during music-induced change in consciousness. In: Proceedings of the 19th Annual International Conference of the Engineering in Medicine and Biology Society, 1997, vol. 1, pp. 227–230. IEEE (1997)
11. de Jong, M.A.: A physiological approach to aesthetic preference—I. Paintings. Psychother. Psychosom. **20**(6), 360–365 (1972)
12. de Jong, M.A., Van Mourik, K., Schellekens, H.: A physiological approach to aesthetic preference. Psychother. Psychosom. **22**(1), 46–51 (1973)
13. Lang, P., Bradley, M.M.: The international affective picture system (IAPS) in the study of emotion and attention. In: Handbook of Emotion Elicitation and Assessment, 29 (2007)
14. Mauss, I.B., Levenson, R.W., McCarter, L., Wilhelm, F.H., Gross, J.J.: The tie that binds? coherence among emotion experience, behavior, and physiology. Emotion **5**(2), 175 (2005)
15. daSilva,H.P.,Guerreiro,J.,Lourenço,A.,Fred,A.L.,Martins,R.:Bitalino: a novel hardware framework for physiological computing. In: PhyCS, pp. 246–253. Citeseer (2014)
16. Esposito, G., Gabrieli, G.: Replication data for: pysiology: a python package for physiological feature extraction. Dataset Published on DR-NTU (Data) (2019). https://doi.org/10.21979/N9/GZSQT7
17. Blood, J.D., Wu, J., Chaplin, T.M., Hommer, R., Vazquez, L., Rutherford, H.J., Mayes, L.C., Crowley, M.J.: The variable heart: high frequency and very low frequency correlates of depressive symptoms in children and adolescents. J. Affect. Disord. **186**, 119–126 (2015)