The Journal of Open Source Software

MNE-ICALabel: Automatically annotating ICA components with ICLabel in Python

Adam Li[1], **Jacob Feitelberg**[2], **Anand Prakash Saini**[3], **Richard Höchenberger**[6], **and Mathieu Scheltienne**[3,4,5]

1 Department of Computer Science, Columbia University, New York, United States **2** Department of Biomedical Engineering, Johns Hopkins University, Baltimore, United States **3** Human Neuroscience Platform, Fondation Campus Biotech Geneva, Geneva, Switzerland **4** Institute of Bioengineering, Center for Neuroprosthetics, EPFL, Geneva, Switzerland **5** Department of Radiology and Medical Informatics, University of Geneva (UNIGE), Geneva, Switzerland **6** Institute of Neuroscience and Medicine (INM-3), Research Center Jülich, Germany

DOI: 10.21105/joss.04484

Software

- Review
- Repository
- Archive

Editor: Elizabeth DuPre

Reviewers:

- @TomDonoghue
- @adswa

Submitted: 24 May 2022
Published: 26 August 2022

Summary

Scalp electroencephalography (EEG) and magnetoencephalography (MEG) analysis is typically very noisy and contains various non-neural signals, such as heartbeat artifacts. Independent component analysis (ICA) is a common procedure to remove these artifacts (Bell & Sejnowski, 1995). However, removing artifacts requires manual annotation of ICA components, which is subject to human error and very laborious when operating on large datasets. This work makes the popular ICLabel model (Pion-Tonachini et al., 2019) available in Python by creating a software package compatible with the MNE-Python [from v1.1; Gramfort et al. (2013)] software toolkit in a modern PyTorch format (Paszke et al., 2019). The ICLabel model was previously only available in an outdated version of TensorFlow that was no longer supported, and migrating the model now to an updated PyTorch version will ensure the model will not break due to unmaintained versions of software. This enables the automatic labeling of ICA components, improving the preprocessing and analysis pipeline of electrophysiological data.

The Python ICLabel model is fully tested against and matches exactly the output produced in its MATLAB counterpart (Pion-Tonachini et al., 2019). Moreover, this work builds the API on top of the robust MNE-Python ecosystem, enabling a seamless integration of automatic ICA analysis.

Statement of need

Typically EEG and MEG data have many artifacts due to a variety of signal sources, such as heartbeats, eye blinks, line noise, channel noise, or muscle artifacts. A common signal processing technique for assisting in separating true brain signals from noise sources is ICA (Bell & Sejnowski, 1995). ICA performs blind-source separation decomposing the observed noisy electrophysiological signals into maximally independent sources. For example, it allows researchers to decompose EEG signals into brain signals and a heartbeat signal. Then by manually annotating a component time-series as a heartbeat, one can remove the heartbeat source from the signal and obtain their signals without the heartbeat. This process is typically manually carried out and requires a human to label each component that comes out from an ICA decomposition. This process is subject to human error and is difficult to scale up when dealing with high-dimensional EEG or MEG recordings.

ICLabel is a proposed statistical model that uses neural networks and a crowdsourced training dataset to automatically label ICA components (Pion-Tonachini et al., 2019). However, this package was previously only available in MATLAB, limiting its usage among neuroscientists preferring Python-based tools. Moreover, the package relied on an outdated version of Tensorflow (Abadi et al., 2016), which makes the model difficult to build on. PyTorch is the most common Python deep-learning framework among researchers. Replicating the model in PyTorch enables future automatic ICA annotation research to easily build upon the model. MNE-Python is a general-purpose electrophysiology analysis package in Python that has a large core group of developers. Integration into MNE makes it likely that MNE-ICALabel will be maintained and continue to be improved. MNE-Python also has stability due to the funding it receives directly for development from institutions such as National Institutes of Health, the Chan Zuckerberg open-source initiative, the European Research Council and Agence Nationale de la Recherche in France.

Future Directions

In future versions, we plan on building on top of the ICLabel model to improve its robustness when operating on different types of recording hardware, sensor type, and sensor count. Moreover, we plan on building new types of models that improve their overall accuracy and performance with respect to auto-labeling neural and non-neural signals. The availability of a simple API for ICLabel will facilitate a strong benchmark to build future models and bring in new developers and contributors.

The developer team is excited to improve the state of the art in data handling for ICA preprocessing and looking forward to welcoming new contributors and users from the broader MNE, neuroscience and electrophysiology community.

Acknowledgements

`MNE-ICALabel` development is partly supported by the National Science Foundation under Grant # 2030859 to the Computing Research Association for the CIFellows Project and by the Human Neuroscience Platform, Fondation Campus Biotech Geneva, Geneva, Switzerland.

We acknowledge the work of (Pion-Tonachini et al., 2019), which was originally produced in MATLAB in the popular EEGLab package.

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., … Zheng, X. (2016). TensorFlow: A system for large-scale machine learning. *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, 265–283. https://www.usenix.org/system/files/conference/osdi16/osdi16-abadi.pdf

Bell, A. J., & Sejnowski, T. J. (1995). An Information-Maximization Approach to Blind Separation and Blind Deconvolution. *Neural Computation*, *7*(6), 1129–1159. https://doi.org/10.1162/neco.1995.7.6.1129

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., & Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-python. *Frontiers in Neuroscience*, *7*, 267. https://doi.org/10.3389/fnins.2013.00267

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E. Z., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., … Chintala, S. (2019). PyTorch:

An imperative style, high-performance deep learning library. *CoRR*, *abs/1912.01703*. https://doi.org/10.48550/ARXIV.1912.01703

Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage*, *198*, 181–197. https://doi.org/10.1016/j.neuroimage.2019.05.026