 DOI: 10.3217/978-3-85125-962-9-157

Advancing artifact handling in BCI research: from filtering to non-neuronal artifacts

T. Memmott[1*], D. Klee[1], B. Celik[2], B. Peters[1], N. Smedemark-Margulies[2], M. Fried-Oken[1] & B. Oken[1] on behalf of the Consortium for Accessible Multimodal Brain-Body Interfaces (CAMBI)

[1]Oregon Health & Science University, Portland, OR, USA; [2]Northeastern University, Boston, MA, USA
*707 SW Gaines St #1290, Portland, OR, 97239. E-mail: memmott@ohsu.edu

*Introduction***:** Research studies relying on electroencephalography (EEG) require some handling of non-neuronal data to function with sufficient data resolution for interpretation. For non-implantable BCIs, this requirement becomes even more important as the underlying potentials used for classification are now real-time instead of averaged data, and every inquiry into a user's state costs time. Furthermore, each BCI system operates under different working assumptions. For instance, a blink may constitute an artifact or an enhancer of classification performance in P300 spellers, depending on the time and consistency of the response [1]. The data may not be recoverable under certain circumstances due to presence of artifacts or a user's inability to perceive a target stimulus, and it may be necessary to recollect data in real-time (ask again; retry) as opposed to undergoing costly signal reconstructions and attempting to make inferences with it. Therefore, bridging the gap between customary neurophysiological methods of artifact handling requires nuance and a systematic understanding of the operating limits of the BCI system utilized [2, 3]

*Methods, Materials, and Results***:** In this study, we leveraged rapid serial visual presentation (RSVP) calibration data collected for a pilot study of thirty-one non-disabled participants (age 49.4 ± 19.99 years). EEG data were collected using a DSI-24 cap (Wearable Sensing) at a sampling rate of 300 Hz. RSVP calibration occurred with letters presented at a rate of 5 Hz, for 110 inquiries of ten letters (resulting in 1100 total trials for calibration) using BciPy [4]. Two neurophysiology researchers labelled the data for the following artifact types: peak voltage, blink, EOG, EMG, ECG, flat voltage, and event (unknown origin, but non-neuronal, such as movement). Two filtering pipelines were tested on the calibration data: 1) conventional training, where the whole dataset was filtered, then epoched into trials for training, and 2) inquiry-based training, where the data were epoched as it would be in real-time, filtered, then epoched again into trials for training. A bandpass filter of 1-20Hz, order 2, 60Hz notch was used. In addition to this question, we trained our models on data conventionally filtered with all prelabeled artifacts described above removed. Trained models from calibrations were generated in BciPy using PCA pre-processing and a regularized discriminant analysis [4]. Using a paired t-test, the inquiry-based filtering approach ($M = .782$, $SD = .015$) worked significantly better for classification (AUC) than the conventional filtering ($M = .768$, $SD = .014$) ($t(30) = 3.34$, $p < 0.002$). Removing all artifacts ($M = 210$ epochs with artifact dropped) resulted in no meaningful change in classification performance (AUC) in this population ($p < 0.1$) using a conventionally filtered dataset.

*Discussion***:** Artifact rejection in BCI systems remains a challenge. Our data suggest that upfront categorization of signals and noise per paradigm may benefit the overall system, as a one--fits-all-approach to artifact seems unlikely across BCIs [1, 2, 3]. Future studies should attempt training on subsets of specific noise types as described above and add simulated and/or real typing data to see how these models perform when encountering real-time data with artifacts. This will be necessary to deduce how each type, with varying amplitudes and morphologies and time relationships to targets/non-targets, will impact classifier performance. A limitation of this study was that additional samples could not be collected after artifact removal, and the use of fewer training samples may have its own cost, particularly when concerning the target trials, which are much less frequent (<=100) vs. nontargets (<=1000). Furthermore, while preliminary data suggest that removing all non-neuronal signals may not result in a significant classifier improvement, it may still be better to continue without the artifacts depending on the research question and the temporal proximity of artifacts to the signals of interest.

Significance: Artifact detection and handling are fundamental components of a functional BCI system. These procedures can cause a benefit or detriment. We describe an approach and considerations to improve this discourse.

Acknowledgements: This work was supported by NIH R01DC009834.

References: [1] Thompson, D. E., Mowla, M. R., Dhuyvetter, K. J., Tillman, J. W., & Huggins, J. E. (2019). Automated Artifact Rejection Algorithms Harm P3 Speller Brain-Computer Interface Performance. *Brain computer interfaces (Abingdon, England)*, *6*(4), 141–148. https://doi.org/10.1080/2326263X.2020.1734401
[2] McDermott, E. J., Raggam, P., Kirsch, S., Belardinelli, P., Ziemann, U., & Zrenner, C. (2021). Artifacts in EEG-Based BCI Therapies: Friend or Foe? *Sensors (Basel, Switzerland)*, *22*(1), 96. https://doi.org/10.3390/s22010096
[3] Minguillon, J., Lopez-Gordo, M. A., & Pelayo, F. (2017). Trends in EEG-BCI for daily-life: Requirements for artifact removal. *Biomedical Signal Processing and Control*, *31*, 407-418.
[4] Memmott T, Kocanaogullari A, Lawhead M, Klee D, Dudy S, Fried-Oken M, Oken B. BciPy: Brain-computer interface software in Python. *Brain-Computer Interfaces*, 8:4: 137-153, 2021.