Toward Standard User-Centered EEG BMI Performance Evaluation

Sumit Soman, *Senior Member, IEEE,* Martijn De Neeling, and Zach McKinney, *Senior Member, IEEE*

***Abstract*—Despite extensive development and wide adoption, the successful application of BMIs remains limited by the persistent lack of standard performance metrics. As part of IEEE Standards Working Group P2794, this paper summarizes state-of-the-art evaluation metrics and suggests next steps and future directions towards the standardization of BMIs.**

I. Background and Motivation

Electroencephalogram (EEG)-based Brain-Computer and Brain-Machine Interface (BCI/BMI) systems have evolved significantly and been widely adopted in recent years for both medical and consumer applications, providing rich non-invasive measures of brain activity, mental state, and user intention. Such systems typically acquire multi-channel time-series data through electrode caps placed over the scalp while the user performs specific cognitive tasks. The EEG signal is then passed through signal processing and machine learning-based feature extraction modules to detect user intention, for communication or control of assistive devices.

A wide variety of metrics have been used for performance evaluation of BMI systems [1], that have conventionally focused aspects such as signal quality or statistical measures of machine learning classifier performance, including accuracy, precision, recall, Information Transfer Rate (ITR), Cohen's Kappa coefficient, and others. However, subject-specific training and user's skill play a significant role in determining the practical use of such BMI systems, including hybrid BMIs. For example, ITR estimates have been found to vary between offline and online use cases (due to time lag in shifting between targets) [2]. It also assumes lack of memory and equal probability of outcomes that do not reflect the user- and context-dependent nature of BMI performance. These effects create ITR heterogeneity across studies, confounding systematic review and meta-analyses [3].

As a result, the reliable interpretation, comparison, and meta-analysis of BMI performance across separate systems and studies remains limited by the lack of standard evaluation protocols and metrics that address human factors. Here, we highlight the need to expand on conventional BMI metrics to include user- and clinically-related metrics.

II. Current and Future Work

To address the need, IEEE Standards Working Group (SWG) P2794[1] is developing Reporting Standards for in vivo Neural Interface Research (RSNIR), for which it recently published a set of preliminary minimum reporting requirements for implantable interfaces [4] and remains open to collaboration. Here, we focus on metrics for EEG-based BMIs motivated by [5,6] to move from signal quality and machine learning model generalization metrics towards more subject-specific and clinically relevant metrics.

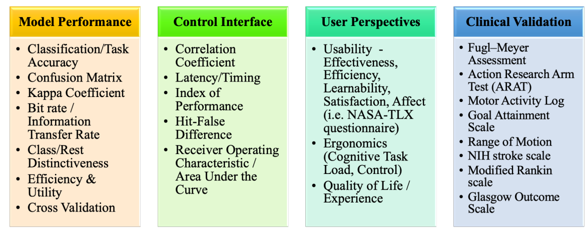

Figure 1 : Proposed integration of technical & human-centered BMI Metrics

By accounting for human factors, metrics oriented towards user function, usability, and/or clinical outcomes can complement the limitations of classical metrics, by providing more comprehensive, integrated measures of BMI performance. Several such metrics exist in common BMI applications such as stroke rehabilitation. However, the successful application of BMIs to increasingly diverse, heterogenous user populations and use scenarios demands the rigorous integration of technology- and user-oriented performance metrics, as proposed in Fig. 1.

Despite the custom development of BMI-specific clinical outcome measures [1] there remains no consensus on which metric(s) are most accurate, reliable, or functionally meaningful. The RSNIR SWG aims to address these issues by developing a BMI-specific scientific reporting standard that ensures interpretability, transparency, and replicability of BMI-related research. Together with the BCI functional model [7], glossary, and ongoing work of IEEE SWG P2731 (Unified Terminology for BCIs), RSNIR aims to build the foundations for a more robust, interoperable neurotechnology ecosystem capable of serving a wider range of human needs, all with more precise attention to emerging ethical issues [8].

[1] SWG P2794 is supported by IEEE Technical Activities Board Committee on Standards (CoS), IEEE EMBS Technical CoS, and the IEEE Brain Initiative, via Industry Connections Activity on Neurotech for BMI.
Sumit Soman (sumitsoman@ieee.org) is with the Center for Development of Advanced Computing, India. Martijn De Neeling with the Neurosciences Dept, KU Leuven, Belgium. (martijn.dene@gmail.com). Zach McKinney (z.mckinney@ieee.org) is with Scuola Superiore Sant'Anna (Pisa, Italy).

References

[1] Choi I, *et al*. "A systematic review of hybrid brain-computer interfaces: Taxonomy and usability perspectives." *PloS One* 12.4 (2017).
[2] Yuan P, *et al*. "A study of the existing problems of estimating the information transfer rate in online brain–computer interfaces." *J. Neural Engg.* 10.2 (2013).
[3] Marchetti M, *et al.* "Effectiveness of the P3-speller in brain–computer interfaces for amyotrophic lateral sclerosis patients: a systematic review and meta-analysis" *Frontiers in Neuroengineering* 7.12 (2014).
[4] Eiber C.D., *et al.* "Preliminary Minimum Reporting Requirements for Reporting In-Vivo Neural Interface Research: Implantable Neural Interfaces." *IEEE Open J. Engg. Med Bio* (2021).
[5] De Neeling M, *et al*. "Single-paradigm and hybrid brain computing interfaces and their use by disabled patients." *J. Neural Eng* 16.6 (2019).
[6] Lotte F, *et al.* "Defining and quantifying users' mental imagery-based BCI skills: a first step." *J. Neural Engg.* 15.4 (2018).
[7] Easttom C, Bianchi L, *et al.* "A Functional Model for Unifying Brain Computer Interface Terminology". *IEEE Open J. Engg. Med Bio* (2021).
[8] Klein E. "Ethics and the emergence of brain-computer interface medicine." *H. book of clinical neurology*, 168(2020). Elsevier.