2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Automated Signal Quality Assessment of Electroencephalography Data in Preclinical Mouse Models of Brain Injury

Eric Qi, Leslie M. Collins, *Senior Member, IEEE*, Bradley J. Kolls, Brian E. Mace, Eduardo Chaparro, Eric Lassiter, Boyla O. Mainsah, *Member, IEEE*

***Abstract*— In this work, we investigate the use of machine learning as an alternative to manual review and annotation of electroencephalography (EEG) data in preclinical mouse models of brain injury. Preliminary results from classifiers trained and tested on independent datasets from two preclinical mouse models of brain injury show high performance in predicting labels of signal patterns of interest. These results demonstrate the potential utility of an automated tool to assess the quality of EEG data in preclinical models of brain injury.**

***Clinical Relevance*— High quality EEG data ensures the development and assessment of reliable EEG-based biomarkers based on EEG features that reflect brain function and not noise.**

I. INTRODUCTION

Preclinical studies are crucial to pre-assess the safety and effectiveness of candidate drugs and to develop reliable biomarkers that can facilitate the evaluation of therapies on outcomes prior to clinical trials [1]. Our research focuses on using preclinical mouse models to develop a biomarker of brain injury severity to predict outcomes based on electroencephalography (EEG) data analysis. To ensure high data quality, the typical approach is to manually review EEG recordings to annotate patterns of interest; however, this process is very time-consuming. In this work, we leverage machine learning to develop an automated tool for EEG data quality assessment in preclinical models of brain injury.

II. METHODS

We induced traumatic brain injury (TBI) in male and female C57Bl/6 mice and induced epilepsy with systemic pilocarpine injection in male C57Bl/6 mice [2]. Two EEG electrode leads were placed within 3 hours after injury and data were recorded (8 hours/day) at a sampling rate of 200 Hz. In the TBI model, EEG data were recorded over the first 6 days after injury. In the epilepsy model, EEG data were recorded for 6 days at 4 weeks after injury. EEG signals from a subset of the mice from both preclinical models were manually labeled as non-seizure, seizure, and artifact (Fig. 1). EEG signals were bandpass filtered (0.5-100 Hz), and notch filtered (20, 40, 60 and 80 Hz) to attenuate electrical noise. Features extracted from 30 second segments include delta, theta, alpha, and beta power bands, regularity, and Shannon entropy. K-nearest neighbor (KNN), random forest (RF), Gaussian naïve Bayes (GNB) and voting ensemble classifiers to predict the signal label were trained on 85% of the epilepsy dataset and tested on the remaining epilepsy dataset and the TBI dataset.

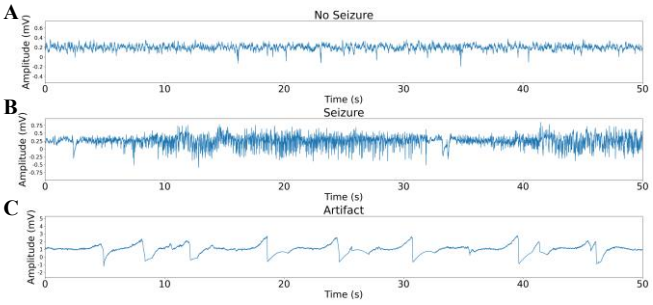

Figure 1. Example electroencelaphalography (EEG) signals with **(A)** non-seizure activity, **(B)** seizure activty and **(C)** artifact.

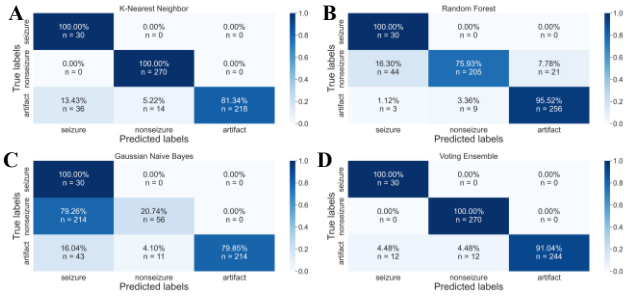

Figure 2. Confusion matrices of EEG signal label predictions of trained classifiers applied to the test dataset (15% epilepsy dataset and 100% TBI dataset). **(A)** K-nearest neighbor (KNN, k = 5), **(B)** Random forest (RF) **(C)** Gaussian naïve Bayes classifiers, and **(D)** A voting ensemble of the KNN and RF classifiers for better aggregate performance.

III. RESULTS

Fig. 2A-D show confusion matrices of the KNN, RF, GNB, and the best voting ensemble classifier, respectively. The average accuracies for the classifiers were KNN 91%, RF 86.44%, GNB 52.82% and voting ensemble 95.77%.

IV. DISCUSSION & CONCLUSION

These early findings show the potential of a machine learning approach to assess data quality in preclinical mouse models that generate EEG data. We are currently annotating the full dataset to develop the EEG signal label classifier on a larger dataset. Our goal is to replace the time-consuming manual review process with automated annotation of EEG data.

REFERENCES

[1] Cavagnaro, J. A. Preclinical safety evaluation of biotechnology-derived pharmaceuticals. *Nat. Rev. Drug Discov*. 2002.1(6): 469-75.

[2] Laskowitz, D.T., *et al.* Traumatic brain injury exacerbates neurodegenerative pathology: improvement with an apolipoprotein E-based therapeutic. *J Neurotrauma*, 2010. **27**(11): 1983-95.

*Research supported by Corticare, Inc.

E. Q., L. M. C. and B. O. M. are with the Department of Electrical and Computer Engineering, Duke University, USA. B. J. K., B. E. M. and E. L. are with the Department of Neurology; E. C. is with the Department of Neurosurgery, Duke University Medical Center, USA.