Epileptic Seizure Detection on TUSZ: Statistics and Channel-wise Approach

V. Stragier[1], P. Vanabelle[2] and R. El Tahry[3,4]

1. Information, Signal and Artificial Intelligence (ISIA) Lab, University of Mons (UMONS), Boulevard Dolez 31, 7000 Mons, Belgique
2. Data Science Department, Centre of Excellence in Information and Communication Technologies (CETIC), Avenue Jean Mermoz 28, 6041 Charleroi, Belgique
3. Refractory Epilepsy Centre, University Hospital of Saint-Luc, 1200 Bruxelles, Belgique
4. Institute of Neuroscience, Catholic University of Louvain, 1200 Bruxelles, Belgique

vincent.stragier@umons.ac.be, paul.vanabelle@cetic.be, riem.eltahry@uclouvain.be

The detection of epileptic seizures on electroencephalogram (EEG) signals presents a considerable challenge, whether for the diagnosis of epileptic seizures or for outpatient follow-up. Developing an automated system to assist neurologists in this task is essential [1]. Indeed, the task of analyzing an EEG is time-consuming, as well as costly, since it requires the expertise of neurologists specialized in interpreting epilepsy EEG signals.

Our epilepsy seizure detection pipeline [2] is built on feature extraction with an extreme gradient boosting (XGBoost) classifier. The particular interest of decision tree-based algorithms – random forest, XGBoost, etc. – is that the selection of features is not necessary and that models are interpretable by themselves. The dataset we used is the Temple University Hospital EEG Seizure Corpus (TUSZ) [3]. Recent changes to the pipeline were done to pre-processing, feature extraction and architecture. Version 1.5.2. of the TUSZ was considered as input. For pre-processing, the extraction of overlapping windows of EEG signals was added, i.e., a 4-second-wide window extracted from each channel every second. Concerning the features, some PyEEG [4] functions were corrected, and others, like the line length [5], were added to complete our set of features. Improvements to the pipeline can accelerate the exploration of new techniques and decrease the likeliness of implementation errors.

Moreover, to better understand past results, a detailed investigation of the onset and propagation of TUSZ seizures in relation to channels was carried out. This descriptive analysis is an important step in the methodology of data science. Besides some issues revealed in the corpus, the results demonstrate that our system, built on a global approach, is inadequate for detecting whether a seizure is occurring somewhere on the channels. Rather, the results suggest the utility of a channel-wise approach, due to the broad variety of seizure dynamotypes observed. Dynamotypes define the dynamic composition of the seizure, i.e., how they begin, evolve and end [6]. This method could answer whether seizure modeling is possible at the channel level. If successful, the advantages would be a better localization and segmentation of seizures. Hence, this is very interesting for research purposes and could help neurologists locate the origin of seizures within the brain, diagnose or decide on possible surgery more quickly. Furthermore, it is assumed that, if the seizure state of nearby channels is considered in a post-processing unit, seizure detection could be enforced with a reduced number of false alarms.

The rebuilt pipeline was tested for validation, giving us updated results for the v1.5.2. dataset with the global approach. Still, the highest F1 scores were low – at most 48 % – and the number of false detections was relatively elevated. After inspecting the results, artifacts were found to be the root of some false detections. This then advocates for proceeding with artifact removal during the pre-processing of the signals, which has not yet been implemented. On a few channels in the temporal area – F7-T3, T3-C3, and T3-T5 – (a topic of interest for temporal lobe epilepsy research for the neurologists consulted), the channel-wise approach was applied, and the combination of the results was tested in pairs. While the combination rules worked on the training set, performance on unseen data (dev set) did not improve. This observation is important, as it could imply that the extensive use of features as input of a classifier is inefficient to solve the detection task at the channel level. There are numerous explanations for this, from the artifact issues to the different distribution of seizure types between the training and dev set. However, this approach must be further investigated with a broader implementation and a new set of data.

ACKNOWLEDGMENTS

This work was carried out as part of the ARIAC project and supported by the Walloon Region.

REFERENCES

[1] S. Roy, I. Kiral, M. Mirmomeni, T. Mummert, A. Braz, J. Tsay, J. Tang, U. Asif, T. Schaffter, M. E. Ahsen, T. Iwamori, H. Yanagisawa, H. Poonawala, P. Madan, Y. Qin, J. Picone, I. Obeid, B. D. A. Marques, S. Maetschke, R. Khalaf, M. Rosen-Zvi, G. Stolovitzky, and S. Harrer, "Evaluation of artificial intelligence systems for assisting neurologists with fast and accurate annotations of scalp electroencephalography data," EBioMedicine, vol. 66, p. 103275, Apr. 2021. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2352396421000682.

[2] P. Vanabelle, P. De Handschutter, R. El Tahry, M. Benjelloun, and M. Boukhebouze, "Epileptic seizure detection using eeg signals and extreme gradient boosting," Journal of biomedical research, vol. 34, no. 3, p. 228, 2020.

[3] V. Shah, E. von Weltin, S. Lopez, J. McHugh, L. Veloso, M. Golmohammadi, I. Obeid, and J. Picone, "The Temple University Hospital Seizure Detection Corpus," *Frontiers in Neuroinformatics*, vol. 12, Jan. 2018.

[4] F. Bao, X. Liu, and C. Zhang, "PyEEG: An open source python module for EEG/MEG feature extraction," Computational intelligence and neuroscience, vol. 2011, p. 406391, Mar. 2011.

[5] J. G. Bogaarts, D. M. W. Hilkman, E. D. Gommer, V. H. J. M. van Kranen-Mastenbroek, and J. P. H. Reulen, "Improved epileptic seizure detection combining dynamic feature normalization with EEG novelty detection," Med Biol Eng Comput, vol. 54, no. 12, pp. 1883–1892, Dec. 2016.

[6] M. L. Saggio, D. Crisp, J. M. Scott, P. Karoly, L. Kuhlmann, M. Nakatani, T. Murai, M. Dumpelmann, ¨ A. Schulze-Bonhage, A. Ikeda, M. Cook, S. V. Gliske, J. Lin, C. Bernard, V. Jirsa, and W. C. Stacey, "A taxonomy of seizure dynamotypes," *eLife*, vol. 9, p. e55632, Jul. 2020. [Online]. Available: *https://doi.org/10.7554/eLife.55632*.