Harmful Brain Activity Classification

[AUTHORS] Daria Chylak (duc9244), Keen Huei Liew (khl456) - Team 9

1. Motivation and Related Work

The Harmful Brain Activity Classification project presents an opportunity to leverage the power of machine learning to improve the field of neurology, particularly in the accurate detection and classification of seizures and other harmful brain activities. The project is motivated by the need to address the limitations of current electroencephalography (EEG) analysis practices, which rely on manual interpretation by neurologists. This traditional approach introduces significant challenges, i.e. human error, high cost, variability in EEG interpretation by neurologists, and a labor intensive process (Chung et al., 2023).

Our hypothesis is: "Utilizing EfficientNet for deep learning analysis of EEG recordings and spectrograms will significantly improve the accuracy, reduce human error, and decrease the time and cost associated with manual interpretation of harmful brain activities, such as seizures, in neurological diagnostics." The purpose of the Harmful Brain Activity Classification project is to harness advanced machine learning techniques, specifically EfficientNet, to improve the accuracy and efficiency of detecting and classifying seizures and other harmful brain activities through EEG and spectrogram analysis.

2. Data

The dataset consists of 17,089 EEGs recordings and 11,138 spectrograms taken from 1,950 patients. Experts reviewed 50 second long EEG samples along with matched spectrograms covering a 10-minute window centered at the same time, and labeled the central 10 seconds. To be more specific, with a center T, the EEG time window is 50 seconds long $[T\text{-}25 : T\text{+}25]$ and the spectrogram time window is 600 seconds long $[T\text{-}300 : T\text{+}300]$, and both the EEG and spectrogram have the same center timestamp. The task is to predict the event occurring in the middle 10 seconds of both these time windows $[T\text{-}5 : T\text{+}5]$. This dataset is ready to use with no access restrictions.

3. Proposed Method

We plan to use EfficientNet, a type of CNN, due to its scalability, efficiency and accuracy in handling image and time-series data. EfficientNet is based on compound scaling, which is known for its ability to scale the depth, width, and resolution of the network in a balanced way. We plan on starting with EfficientNetB0, establishing a performance benchmark before scaling up to a larger EfficientNet model for increased accuracy if feasible. We believe that this method can manage the complexities of EEG data, offering a promising solution to the challenges of manual interpretation in a clinical setting.

4. Evaluation Metrics

Model performance will be evaluated using metrics such as accuracy, sensitivity (recall), specificity and the area under the ROC curve (AUC), which are standard processes for classification tasks. These metrics will help in assessing the model's ability to correctly identify harmful brain activities and differentiate them from normal activity.

5. Timeline

The due date for this project is April 30th, 2024. Our timeline includes data preprocessing (2 weeks), model development and training (2 weeks), validation (2 weeks), and testing and refinement (2 weeks), totaling an 8-week project duration, allowing for iterative improvements and rigorous evaluation of the deep learning model's effectiveness.

References

Y.G. Chung, W.J. Lee, S.M. Na, H. Kim, H. Hwang, C. H. Yun, and K.J. Kim. Deep learning-based automated detection and multiclass classification of focal interictal epileptiform discharges in scalp electroencephalograms. *Scientific Reports*, 13(1):6755, 2023. doi: 10.1038/s41598-023-33906-5.

J. Jing, Z. Lin, C. Yang, A. Chow, S. Dane, and M. B. Westover J. Sun. Hms - harmful brain activity classification, 2024. URL `https://kaggle.com/competitions/hms-harmful-brain-activity-classification`.

Jing et al. (2024) Chung et al. (2023)