Utilizing Multi-Modal Data to Enhance Epileptiform Classification by Deep Learning Model via Spatio-Temporal Reasoning

Poopa Kaewbuapan
School of Information Technology
King Mongkut's University of Technology Thonburi
Bangkok, Thailand
poopa.kaew@kmutt.ac.th

***Abstract*—Transformation of time-series EEG signal into spatial representation improves epileptiform classification performance of 2DCNN model. We employs an ablation study of an ensemble of deep learning models used in Kaggle competition "Harmful Brain Activity Classification". We compare between 3 settings of the model in which we isolate the components that utilize temporal information, spatial information, and finally the default setting where both information were utilized. We evaluate the models performance on KL divergence metric which show that both information model slightly improved (0.38) from temporal only (0.39) and spatial only (0.41). Overall, it is yet inconclusive about the real effect of spatio-temporal reasoning in spatially sensitive model due to that the interaction of both information were just on ensemble averaging during inference only. Possible future studies on training the both information model with concatenated flatten features has been addressed.**

Keywords—Spatio-temporal Reasoning, Epilepsy, Electroencephalography (EEG)

I. Introduction

Electroencephalography (EEG) is a critical tool for monitoring brain activity in critically ill patients, particularly for detecting seizures and other harmful brain activities. Traditional EEG analysis relies heavily on manual interpretation by specialized neurologists, which is both time-consuming and prone to errors due to fatigue and inter-rater variability. The increasing availability of large EEG datasets and advancements in deep learning have opened new possibilities for automating EEG analysis, potentially improving the speed and accuracy of diagnoses.

This study investigates the enhancement of epileptiform classification performance through the transformation of time-series EEG signals into spatial representations, leveraging the strengths of convolutional neural networks (CNNs). We conduct an ablation study using an ensemble of deep learning models originally developed for the Kaggle competition "Harmful Brain Activity Classification." Our analysis focuses on three model configurations: one utilizing temporal information, another using spatial information, and a combined model incorporating both. Performance is evaluated using the Kullback-Leibler (KL) divergence metric, with findings suggesting a marginal improvement when both temporal and spatial information are integrated. However, the results remain inconclusive regarding the true impact of spatio-temporal reasoning due to the simplistic ensemble averaging approach employed. Future research directions include training models with concatenated features to better harness cross-modality interactions.

II. Related works

A. *Harmful Brain Activity Classification*

" From stethoscopes to tongue depressors, doctors rely on many tools to treat their patients. Physicians use electroencephalography with critically ill patients to detect seizures and other types of brain activity that can cause brain damage.

Currently, EEG monitoring relies solely on manual analysis by specialized neurologists. While invaluable, this labor-intensive process is a major bottleneck. Not only can it be time-consuming, but manual review of EEG recordings is also expensive, prone to fatigue-related errors, and suffers from reliability issues between different reviewers, even when those reviewers are experts.

Your work in automating EEG analysis will help doctors and brain researchers detect seizures and other types of brain activity that can cause brain damage, so that they can give treatments more quickly and accurately. The algorithms developed in this contest may also help researchers who are working to develop drugs to treat and prevent seizures.

There are six patterns of interest for this competition: seizure (SZ), generalized periodic discharges (GPD), lateralized periodic discharges (LPD), lateralized rhythmic delta activity (LRDA), generalized rhythmic delta activity (GRDA), or "other".

The EEG segments used in this competition have been annotated, or classified, by a group of experts. In some cases experts completely agree about the correct label. On other cases the experts disagree. We call segments where there are high levels of agreement "idealized" patterns. Cases where ~1/2 of experts give a label as "other" and ~1/2 give one of the remaining five labels, we call "proto patterns". Cases where experts are approximately split between 2 of the 5 named patterns, we call "edge cases"." [3]

III. Methodology

Temporal Model: Training on 50 seconds time window EEG signal – Bandpass filter of 0.1-20 Hz, WaveNet 1DCNN backbone – Keras Tensorflow implemention (Trained on Kaggle's GPU T100 environment)

Spatial Model: Training on Mel spectrograms reconstructed from 50 seconds time window EEG signal – Bandpass filter of 0.1-20 Hz, EfficientNet-2B 2DCNN backbone (Trained on Kaggle's GPU T100 environment)

Both Information Model: Ensemble of trained Temporal model and trained Spatial model

We compare between 3 settings of the model in which we isolate the components that utilize temporal information, spatial information, and finally the default setting where both information were utilized. We then evaluate the models classification performance on KL divergence metric (lower is better) and do 5-fold cross validation.

IV. RESULTS

For the both information model slightly improved (0.38) from temporal only (0.39) and spatial only (0.4). For the cross-validation score difference to model score are 0.2285, 0.2388, and 0.2951 respectively.

V. DISCUSSION

It is possible for epileptiforms to regularly exhibit localized effect with consistent bursts of irregular activity onto specific band of frequencies, in which a spectrogram can extract further information gain from this spatially adjacent depencdency and so the spatial model focus on certain continuous "patches" [5].

VI. CONCLUSION

Overall, it is yet inconclusive about the real effect of spatio-temporal reasoning in spatially sensitive model due to that the interaction of both information were just on ensemble averaging during inference only. Possible future studies could be on training the both information model with concatenated flatten features to induce cross-modality interaction from the training phase also.

VII. FUTURE WORKS

The current approach of combining models trained on spectrograms and another model trained on EEG data presents promising opportunities for improving classification accuracy. However, several enhancements and new directions can be pursued to further optimize and validate the performance of the model.

1. **Integration of Multi-Modal Inputs**:
   - **Concatenation of Features**: Future work should focus on the integration of features from both the spectrograms and EEG data during the training phase. By concatenating the flattened features from both models and adding a shared classification head, we can enable the interplay of information during fine-tuning. This approach could reveal synergistic effects that improve overall model performance.
   - **Fine-Tuning**: Fine-tuning the combined model rather than only during inference could better leverage the shared information, leading to more robust feature representations. By allowing the weights of the combined model to adjust together, we may capture more complex relationships between the different types of data.
2. **Enhanced Model Architecture**:
   - **3D Convolutional Neural Networks (3D CNNs)**: Given that EEG data can be represented in three dimensions (time, electrodes, and frequency), employing a 3D CNN can capture the spatiotemporal dependencies more effectively. This could be especially beneficial in recognizing complex patterns across different electrodes. The use of 3D CNNs allows the model to consider the spatial relationships between electrodes over time, potentially leading to better performance in classification tasks.
   - **Electrode Data Augmentation**: Implementing data augmentation techniques that randomly swap electrodes could simulate a variety of spatial configurations. This would help the model generalize better by not relying on a fixed electrode arrangement, which aligns with the variability observed in practical scenarios. Such augmentation can make the model more robust to variations in electrode placement, improving its adaptability and performance in real-world applications.

By focusing on these enhancements, future work can significantly improve the performance and generalizability of models developed for EEG data analysis for epileptiform classification.

ACKNOWLEDGMENT

Thanks Kaggle and Harvard Medical School for sponsoring the competition and providing comprehensive EEG datsets for the competition.

REFERENCES

[1] *D. Ma, J. Zheng, and L. Peng, "Performance evaluation of epileptic seizure prediction using time, frequency, and time–frequency domain measures," Processes, vol. 9, no. 4, p. 682, 2021.*

[2] Kim, Kion (2018), WaveNet: Increasing reception field using dilated convolution, Medium https://medium.com/@kion.kim/wavenet-a-network-good-to-know-7caaae735435

Competition Page

[3] https://www.kaggle.com/competitions/hms-harmful-brain-activity-classification

Starting Notebooks (EDA, Training and Inference, Insights)

[4] https://www.kaggle.com/competitions/hms-harmful-brain-activity-classification/discussion/4734

About the characteristics of EEG - Octopus210 (Expert)

[5] https://www.kaggle.com/competitions/hms-harmful-brain-activity-classification/discussion/4729

Grad Cam : What is important in Spectrograms? - Chris Deotte (Grandmaster)

https://www.kaggle.com/datasets/nartaa/features-head-starter-models

Original source for all models that being used for benchmarking in this study. Based on Chris Deotte's works

https://www.kaggle.com/competitions/hms-harmful-brain-activity-classification/discussion/4792

Preprocessing in the Labeling Done by Experts - rafaelzimmermann1 (Expert) => Very important discussion regarding possibility of the competition organizer to disclose the experts methodology of preprocessing the dataset for reproducibility of the competition

https://www.kaggle.com/code/seanbearden/effnetb0-2-pop-model-train-twice-lb-0-39

EffNetB0 2 Pop Model Train Twice - [LB 0.39] - Sean Bearden (Expert) => Significant insight regarding distribution split within the dataset, suggesting a two pass training. Based on Chris Deotte's works.

Learning Resources

https://www.learningeeg.com/montages-and-technical-components

https://www.learningeeg.com/slowing-and-other-non-epileptiform-abnormalities

https://www.learningeeg.com/epileptiform-activity

https://www.learningeeg.com/seizures

chat.openai.com