Harmful Brain Activity Classification
CS6190 - Spring 2024

Gnanesh Rasineni, UID : u1470943

Abstract

Electroencephalograms (EEG) are instrumental in diagnosing neurological disorders due to their ability to record the brain's electrical activity through electrodes placed on the scalp. These recordings capture various frequency bands (delta, theta, alpha, beta, and gamma), each associated with different brain states and functions. However, the manual interpretation of EEG data is not only time-consuming and costly but also subject to inter-observer variability. This project aims to automate the classification of EEG spectrograms using the HMS dataset, which includes categories such as seizure (SZ), generalized periodic discharges (GPD), lateralized periodic discharges (LPD), lateralized rhythmic delta activity (LRDA), generalized rhythmic delta activity (GRDA), and "other." We employ preprocessing and augmentation techniques, like the Short-Time Fourier Transform and horizontal flips, to prepare the data. An EfficientNet model, optimized with KL-Divergence loss, is trained on these processed images. The performance of our approach is validated through k-fold cross-validation, highlighting the potential of machine learning in enhancing the precision and efficiency of EEG analysis.

A. Introduction

Project Description

The primary objective of this project is to leverage machine learning models to detect and classify seizures and other harmful brain activities including generalized periodic discharges (GPD), lateralized periodic discharges (LPD), lateralized rhythmic delta activity (LRDA), generalized rhythmic delta activity (GRDA) and others [6] [9] detailed in the appendix section. The data utilized for training these models are derived from EEG signals [7] recorded from critically ill patients in a hospital setting. EEG, or electroencephalography, measures the electrical activity in the brain via electrodes placed on the scalp. A portion of this data has been preprocessed into spectrogram images, which are included in the dataset used for model training.

Significance of the Problem

This work holds great significance with potential for various medical domains [12], particularly in areas like neurocritical care, epilepsy treatment, and pharmaceutical development. Advances in the accuracy of EEG pattern classification could revolutionize neurocritical care by enabling quicker and more precise interventions for patients experiencing seizures or other neurological disturbances. Enhancements in EEG analysis [1] can empower physicians and neuroscientists to detect and respond to abnormal brain activities swiftly. This capability is critical not only for immediate patient care but also for advancing our understanding of neurological disorders. Improvements in this field could lead to more effective treatment protocols, optimizing patient outcomes in clinical settings.

Rationale for using Machine Learning

The intricate and variable nature of EEG data [3] presents substantial challenges for traditional analysis methods, which are typically manual and subject to significant observer bias [8]. Traditional EEG interpretation relies on handcrafted feature extraction that can vary widely among different technicians and is prone to subjectivity. Additionally, the conversion of raw EEG data into spectrogram images can result in variations that traditional methods are ill-equipped to handle uniformly.

Machine learning techniques offers a robust solution to these issues. By employing a variety of techniques including Probabilistic modelling approaches [11], systems can be trained to recognize and classify complex patterns within EEG spectrograms automatically. This approach minimizes human error and increases the reliability and efficiency of EEG analysis. These techniques are capable to handle large datasets and learn from data iteratively makes them ideal for adapting to new information and improving diagnostic accuracy over time. brain activity patterns.

B. Methodology

Dataset Overview

The project utilizes the Harmful Brain Activity dataset provided by CCEMRC [6]. This dataset contains 50-second long EEG samples covering a broader 10-minute window, as well as spectrograms matching each sample. Expert annotators labeled the EEG samples, determining which of the six categories each spectrogram belonged to, highlighting the subjective nature of manual classification. The dataset also stratifies data based on the consensus among labelers: samples with unanimous labeling are categorized as "Idealized", those split between one of the five conditions and "other" as "Proto", and those split between two or more conditions as "Edge". This stratification leverages the uncertainty levels in the data to fine-tune the training of our model.

Dataset Structure and Preprocessing

The training data, contained in `train.csv`, consists of 106,800 rows, but there are only 17,089 unique EEG IDs, 11,138 unique spectrogram IDs, and 1,950 unique patients. Each row represents a window of time from a specific patient, with corresponding EEG and spectrogram data available in the respective parquet files. The EEG time window is 50 seconds long, and the spectrogram time window is 600 seconds long, both centered on the same timestamp. The EEG files, comprising 19 EEG time signals, are transformed using the commonly-used Bipolar Double Banana Montage [10] as shown in figure 1, combined with the Short Time Fourier Transform (STFT) [4] as shown in the figure 2 to generate LL, LP, RP, and RR spectrograms. These spectrograms, along with the original spectrograms as shown in the figure 3 from the dataset, form an 8-channel spectrogram feature image used for model predictions. Additionally, data augmentation including horizontal flips is performed on the spectrogram, since the double banana montage representations LL, LP, RP, RR are symmetrical, reflecting the left and right sides of the brain.

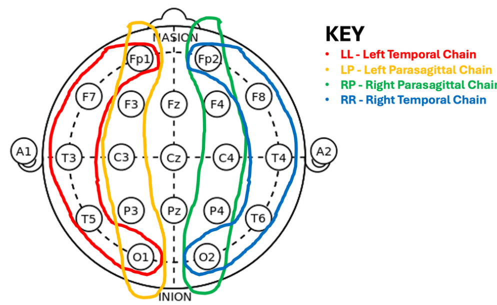

Figure 1. Bipolar Double Banana Montage used for EEG signal processing. This montage technique involves placing electrodes in predefined positions on the scalp to capture the electrical activity from specific brain regions. The montage is designed to enhance the detection of temporal lobe abnormalities, commonly associated with epilepsy and other neurological disorders.

Model Architecture and Training

Dataset Splits: The full dataset was split into a training set, a testing set, and a validation set. The training and validation sets made up 80% of the total data, while the testing set accounted for 20%. K-fold cross-validation [2] as shown in the figure 4 was employed to obtain an accurate estimate of the model's predictive ability. The training data is divided into five folds, and each fold is stratified based on labels, ensuring that each fold contains a uniform distribution of class labels. This strategy prevents data leakage and ensures that the model's evaluations are robust and reliable.

Loss Function: Because the output of the model is an array of predicted probabilities for each label class, the Kullback-Leibler divergence [5] is used to evaluate the model's performance. This metric measures the difference between two probability distributions—ideal for comparing the predicted probabilities with the actual label distribution. The KL divergence

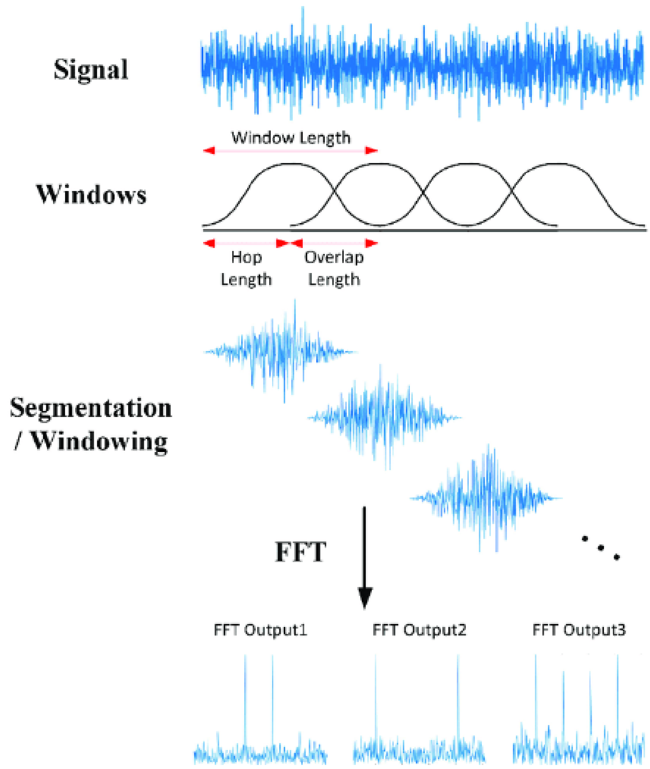

Figure 2. Illustration of the Short Time Fourier Transform (STFT) process used for converting EEG time signals into frequency domain. STFT is a mathematical technique that divides a longer time signal into smaller segments of equal length and computes the Fourier transform separately on each segment. This approach helps analyze the frequency components of the signal over time, providing a dynamic spectrum that is essential for understanding the nature of EEG signals.

is defined as:

$$KL(P \parallel Q) = \sum_{x} P(x) \log \frac{P(x)}{Q(x)} \tag{1}$$

Model Architecture: Transfer learning was utilized with the pre-trained model EfficientNet_B0 from torchvision models over the processed spectrograms with 8 channels. The rationale for choosing the efficientnet model is that these are designed with optimal model scaling, balanced network depth, width, and resolution, which has been proven to be highly effective for improving accuracy and efficiency. The structure of the model includes an input layer, an EfficientNet backbone layer, a linear classification layer, and a softmax layer to convert the output into probabilities summing to 1. The Adam optimizer was used with an adaptive learning rate. The neural network was trained for 10 epochs with a batch size of 32.

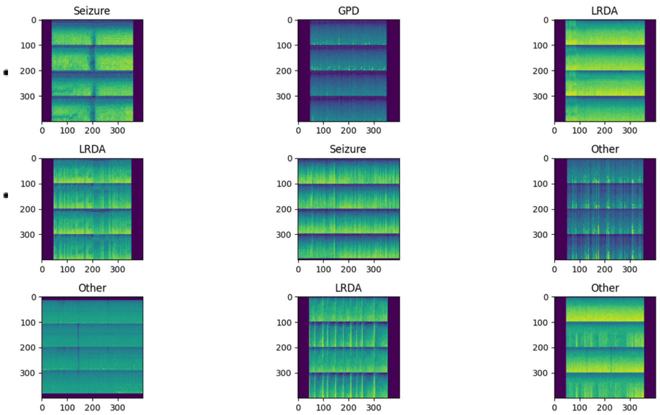

Figure 3. Visual representation of various harmful brain activity conditions using stacked LL, LP, RP, RR spectrograms for each condition. From left to right: Seizure, GPD, LRDA, Seizure, Other, LRDA, Other. Each subplot displays a visualization crucial for distinguishing between these conditions.

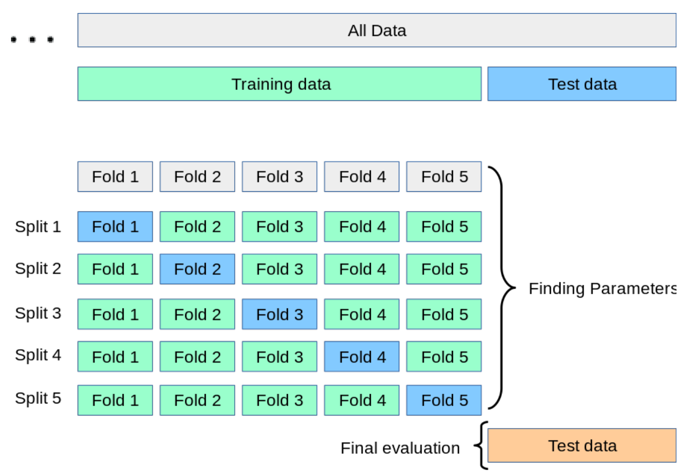

Figure 4. Schematic of the k-fold cross-validation process used in model training. This diagram shows how the dataset is divided into five folds, with each fold used once as a test set while the other four serve as training data. This method is critical for ensuring that the model's performance is not biased towards a particular subset of the data, providing a robust estimate of the model's predictive ability.

C. Results and Discussion

The model achieved a peak Kullback-Leigler Divergence score of 0.513, demonstrating effective classification of harmful brain activities with minimal domain-specific adjustments. This score indicates the model's capability to efficiently learn complex patterns in EEG spectrogram data.

The utilization of the EfficientNet model as a backbone, augmented by the probabilistic modeling techniques of k-fold cross-validation and KL divergence loss, enhances the effectiveness, accuracy and reliability of predictive models. These techniques ensure robust handling of the inherent variability and complexity of medical data, suggesting that machine learning can significantly advance diagnostic processes in neurology and related fields.

D. Conclusion and Future work

By leveraging a robust dataset provided by CCEMRC, which includes labeled EEG data across several brain conditions, we successfully employed an EfficientNet model, optimized with Kullback-Leibler divergence as a loss function, to learn and predict EEG patterns with notable accuracy. The utilization of k-fold cross-validation ensured that our model's performance was reliably assessed, mitigating the risk of overfitting and providing a robust estimate of its predictive capabilities. To further enhance the accuracy and robustness of this EEG classification model, following techniques could be explored: By leveraging the strengths of different model types, such as variations of CNN architectures or other EfficientNet scales, an ensemble could potentially reduce variance and bias, leading to better generalization and accuracy in classifying EEG data. Integrating Bayesian neural networks could provide significant advantages by treating network weights as distributions rather than fixed values, thereby offering a measure of uncertainty in predictions—a feature invaluable in medical diagnostics to aid clinical decision-making. Further refinement of the learning process could be achieved through Bayesian backpropagation, a technique that dynamically adjusts learning by accounting for the uncertainty in weights, potentially leading to more reliable and interpretable models.

References

[1] Hafeez Ullah Amin, Wajid Mumtaz, Ahmad Rauf Subhani, Mohamad Naufal Mohamad Saad, and Aamir Saeed Malik. Classification of eeg signals based on pattern recognition approach. *Frontiers in computational neuroscience*, 11:103, 2017. 1

[2] D. Anguita, L. Ghelardoni, A. Ghio, L. Oneto, and S. Ridella. The 'k' in k-fold cross validation. In *European Symposium on Artificial Neural Networks, Computational Intelligence, and Machine Learning, ESANN*, Apr. 2012. 2

[3] Martin Dinov and Robert Leech. Modeling uncertainties in eeg microstates: Analysis of real and imagined motor movements using probabilistic clustering-driven training of probabilistic neural networks. *Frontiers in human neuroscience*, 11:534, 2017. 1

[4] H. Jeon, Y. Jung, S. Lee, and Y. Jung. Area-efficient short-time fourier transform processor for time-frequency analysis of non-stationary signals. *MDPI*, Oct. 2020. 2

[5] S. Ji, Z. Zhang, S. Ying, L. Wang, X. Zhao, and Y. Gao. Kullback-leibler divergence metric learning. *IEEE Transactions on Cybernetics*, 52(4):2047–2058, Apr. 2022. 2

[6] Jin Jing, Zhen Lin, Chaoqi Yang, Ashley Chow, Sohier Dane, Jimeng Sun, and M. Brandon Westover. Hms - harmful brain activity classification. https://kaggle.com/competitions/hms-harmful-brain-activity-classification, 2024. 1, 2

[7] Johns Hopkins Medicine. Electroencephalogram (eeg). https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/electroencephalogram-eeg, 2019. 1

[8] Saeid Sanei and Jonathon A Chambers. *EEG signal processing and machine learning*. John Wiley & Sons, 2021. 1

[9] A. A. Ein Shoka, M. M. Dessouky, A. El-Sayed, and E. E.-D. Hemdan. Eeg seizure detection: concepts, techniques, challenges, and future trends. *Multimedia Tools and Applications*, Apr. 2023. 1

[10] C. Sun and C. Mou. Survey on the research direction of eeg-based signal processing. *Frontiers in Neuroscience, Neural Technology*, July 2023. 2

[11] Wei Wu, Srikantan Nagarajan, and Zhe Chen. Bayesian machine learning: Eeg\/meg signal processing measurements. *IEEE Signal Processing Magazine*, 33(1):14–36, 2015. 2

[12] Yuhao Zhang, Yue Yao, Zakir Hossain, Shafin Rahman, and Tom Gedeon. Eeg feature significance analysis. In *Neural Information Processing: 28th International Conference, ICONIP 2021, Sanur, Bali, Indonesia, December 8–12, 2021, Proceedings, Part VI 28*, pages 212–220. Springer, 2021. 1

A. Appendix

Electroencephalography (EEG) data typically appears as a series of wavy lines, each representing the electrical activity recorded from different electrodes placed on the scalp. These lines, called traces, show the voltage changes over time. The patterns observed in EEG data are influenced by the brain activity of the individual, and they can vary significantly depending on the state of consciousness, activity, or any neurological conditions.

Patterns Observed in EEG Data

EEG data exhibits a variety of wave patterns, amplitude and frequency variations, artifacts, and abnormal patterns:

- **Wave Patterns:** EEG data includes alpha, beta, delta, and theta waves, each associated with different brain states, such as alpha waves linked to relaxation and beta waves to active thinking.
- **Amplitude and Frequency:** The amplitude indicates the strength of the signal, while the frequency shows the speed of brain wave oscillation.
- **Artifacts:** These are non-brain waveforms that can appear due to muscle movements, eye blinks, or electrical interference.
- **Abnormal Patterns:** In cases of neurological disorders like epilepsy, the EEG may show spikes, sharp waves, or other unusual patterns that indicate abnormal brain activity.

Relation of Spectrograms to EEG

Spectrograms provide a visual representation of the spectrum of frequencies of a signal as they vary with time and are closely related to EEG in the context of analyzing brain waves:

- **Frequency Analysis:** Spectrograms can visually display various brain wave frequencies, like alpha, beta, theta, and delta waves, and how they change over time during an EEG recording.
- **Identifying Patterns:** Spectrograms help identify patterns that might not be easily discernible in the raw EEG waveforms.
- **Temporal and Frequency Resolution:** They provide crucial information about both the timing (temporal resolution) and the frequency (frequency resolution) of brain waves.
- **Data Visualization:** Spectrograms transform EEG's time-domain data into a more accessible frequency-domain representation, making it easier to analyze and understand.

Explanation of LPD/GPD/LRDA/GRDA and Seizure Conditions

Different EEG patterns are associated with specific conditions:

- **Lateralized Periodic Discharges (LPDs):** Associated with acute or subacute brain dysfunction, often related to structural brain lesions or acute brain injuries, featuring lateralization and periodicity, with sharp waveforms or complexes distinguishable from the background EEG activity.
- **Generalized Periodic Discharges (GPDs):** Linked to diffuse or generalized brain dysfunction, characterized by a repeating pattern across both hemispheres, showing repetitive, sharply contoured waveforms that are more synchronized.
- **Lateralized Rhythmic Delta Activity (LRDA):** Involves rhythmic, slow-wave activity in the delta frequency range, typically localized to one hemisphere, often observed in patients with focal brain lesions.
- **Generalized Rhythmic Delta Activity (GRDA):** Characterized by rhythmic delta activity that is distributed more uniformly across both hemispheres, associated with various clinical conditions, including encephalopathies and diffuse brain disorders.
- **Seizures:** EEG patterns during seizures may display sudden, dramatic changes including spikes and sharp waves, indicative of sudden bursts of electrical activity in the brain. This can be seen in various forms such as generalized tonic-clonic seizures, where the pattern shows widespread rhythmic activity, or focal seizures, where disturbances are localized.