Harmful Brain Activity Classification using Ensemble Deep Learning

Marija Habijan, Robert Šojo, Ivana Hartmann Tolić, Irena Galić
Faculty of Electrical Engineering, Computer Science and Information Technology Osijek
Osijek, Croatia
marija.habijan@ferit.hr

***Abstract*—Electroencephalography (EEG) is a crucial tool for monitoring electrical brain activity and diagnosing neurological conditions. Manual analysis of EEG signals is time-consuming and prone to variability, necessitating automated methods for accurate and efficient interpretation. This paper presents an ensemble learning approach for the automated classification of harmful brain activity from EEG data. The model combines the strengths of ResNet1d GRU, EfficientNetB0, and EfficientNetB1 to predict brain activity from spectrograms and raw EEG signals. By leveraging the unique capabilities of each model and employing a weighted averaging scheme, the ensemble achieves a competitive Kullback-Liebler divergence score of 0.3 on the testing dataset. This approach demonstrates the potential of ensemble learning to improve the accuracy and robustness of automated EEG signal classification, offering a valuable tool for early detection and diagnosis of neurological conditions.**

***Keywords*—Brain Activity Classification; EEG Signal Classification; Ensemble Learning; Deep Learning**

I. Introduction

Electroencephalography (EEG) signals are critical indicators used by physicians to detect abnormal brain activity in patients and diagnose potential signs of brain damage, such as epilepsy, sleep disorders, and brain tumors. These signals provide a non-invasive way to monitor and understand the brain's electrical activity, offering invaluable insights into various neurological conditions. Traditionally, neurologists manually analyze EEG results to identify abnormalities [1]. This manual process involves careful examination of the EEG recordings and interpreting the patterns observed, which are often extensive. Such analysis is labor-intensive, often requiring healthcare providers to review approximately 100 pages of activity for each patient [2]. This is time-consuming and prone to inter-observer variability, as neurologists may interpret the same signals differently [3]. The subjective nature of EEG signal interpretation can lead to inconsistent assessments, necessitating multiple neurologists for review to reach a consensus, thereby increasing patient costs [4].

Given the essential role of EEG in measuring electrical disturbances, there is a need for an automatic and efficient methods to classify these signals [5]. Such methods can facilitate the early recognition of epileptic seizures and other brain disorders, enabling timely intervention and treatment. Automated classification systems can significantly reduce the workload on healthcare providers, minimize human error, and provide more consistent and reliable results. The use of deep learning to classify spectrograms of brain activity into various categories, including seizures, generalized periodic discharges (GPD), lateralized periodic discharges (LPD), lateralized rhythmic (LRDA) delta activity, generalized rhythmic delta activity (GRDA), and other patterns are well-suited for this task [6–8]. Moreover, accurate classification of these brain activity patterns can standardize interpretations, aiding clinicians in diagnosing such conditions more reliably. This standardization is crucial, as it helps in reducing inter-observer variability and ensures that patients receive a consistent and accurate diagnosis regardless of the neurologist or healthcare facility. Furthermore, automated systems can process large volumes of data quickly, providing real-time or near-real-time analysis. This is particularly beneficial in critical care settings where timely decisions are essential.

In this work, we employed an ensemble learning approach that combined three distinct pre-trained models: ResNet1d GRU, EfficientNetB0 and EfficientNetB1 to accurately predict brain activity from spectrograms and EEG signals. This approach has the potential to significantly reduce the time and costs associated with manual EEG signal diagnosis. We have structured our research into three main components: data preprocessing, signal processing and transforms, and model training. By normalizing and standardizing the spectrograms, we aim to remove irrelevant noise and decrease computational complexity. After fine-tuning the dataset and model, we achieved a peak Kullback-Liebler divergence score of 0.3 on the testing dataset, demonstrating the effectiveness of our approach.

II. Methods

A. *ResNet1d GRU*

ResNet1d GRU is a hybrid deep learning model that combines the strengths of Residual Networks (ResNet) [9] and Gated Recurrent Units (GRU) [10] tailored for processing one-dimensional (1D) sequential data [11], [12]. This model structure effectively captures both spatial and temporal features within sequences, making it particularly useful for time-series analysis. The ResNet part of the model utilizes residual blocks specifically designed for 1D data. These blocks help learn complex patterns in the data without the hindrance of the vanishing gradient problem, thanks to their skip connections, which allow gradients to flow through the network more effectively. The use of 1D convolutions within these blocks makes

979-8-3503-7542-8/24/31.00 ©2024 IEEE

2024 International Symposium ELMAR | 979-8-3503-7542-8/24/$31.00 ©2024 IEEE | DOI: 10.1109/ELMAR62909.2024.10694603

the model adept at handling sequences like audio, sensor data, or any other form of temporal data where the input is typically a single-dimensional array of values. Following the ResNet layers, the GRU component comes into play. GRUs are a type of recurrent neural network (RNN) that manages sequence dependencies, where the output is not only determined by the current input but also by the previous outputs. GRUs simplify the standard Long Short-Term Memory (LSTM) [13] design with fewer gates and parameters, leading to efficient learning of temporal dependencies with less computational overhead. Together, the integration of ResNet1d for spatial feature extraction and GRU for capturing temporal dynamics offers a robust architecture for tasks such as speech recognition, anomaly detection in time-series data, and physiological signal analysis, providing accurate predictions influenced by both the immediate and historical context of the data.

B. *EfficientNetB0*

EfficientNetB0 [14], [15], a baseline model in the EfficientNet family [16], is a convolutional neural network (CNN) designed for image classification tasks. It strikes a balance between accuracy and efficiency by employing a compound scaling method that uniformly scales the network's width, depth, and resolution. The architecture of EfficientNetB0 is based on the mobile inverted bottleneck convolution (MBConv) block, which uses depthwise separable convolutions to reduce the number of parameters and computations compared to traditional convolutions. These MBConv blocks are arranged in a specific sequence with varying kernel sizes and expansion ratios, allowing the network to capture features at different scales. EfficientNetB0 has been pre-trained on the ImageNet dataset [17], a large collection of labeled images. This pre-training enables transfer learning, where the learned features can be utilized for other image-related tasks with limited data. The model's advantages include its high accuracy, parameter efficiency, and versatility in handling various image sizes. Its compact size and computational efficiency make it suitable for deployment on devices with limited resources. However, like other CNNs, EfficientNetB0 can be computationally demanding to train from scratch, and its performance might be less optimal for tasks that deviate significantly from ImageNet's image distribution.

C. *EfficientNetB1*

Building upon the baseline EfficientNetB0, EfficientNetB1 [14], [18] achieves higher accuracy by scaling up the model's depth, width, and resolution using a compound scaling method. This method ensures balanced scaling, avoiding overemphasizing any single dimension. The architecture of EfficientNetB1, like other EfficientNets, relies on the MBConv block, which employs depthwise separable convolutions for parameter and computational efficiency. The MBConv blocks are arranged in a specific sequence with varying kernel sizes and expansion ratios, allowing the network to effectively capture features at multiple scales. EfficientNetB1 has been pre-trained on the extensive ImageNet dataset, enabling transfer learning for various image-related tasks. By leveraging the knowledge gained from this pre-training, the model can be fine-tuned for specific tasks with relatively small amounts of labeled data. This model's advantages include its improved accuracy compared to EfficientNetB0, while still maintaining a relatively small model size and computational efficiency. It offers good performance across a range of image sizes and can be adapted to different tasks. While EfficientNetB1 demonstrates strong performance, it does require more computational resources than smaller models like EfficientNetB0. Nevertheless, EfficientNetB1 provides a compelling balance between accuracy and efficiency, making it a valuable tool for image classification in various domains.

D. *Ensemble Model*

The ensemble model employed a strategic combination of three distinct convolutional neural network architectures: ResNet1d GRU, EfficientNetB0, and EfficientNetB1, as shown in Figure 1. This approach aimed to leverage the unique strengths of each architecture while mitigating their individual weaknesses. EfficientNetB0, a baseline model known for its efficiency, provided a robust foundation for feature extraction. EfficientNetB1, a scaled-up version, contributed to capturing more intricate patterns within the EEG spectrograms, potentially enhancing overall accuracy. ResNet1d GRU, with its

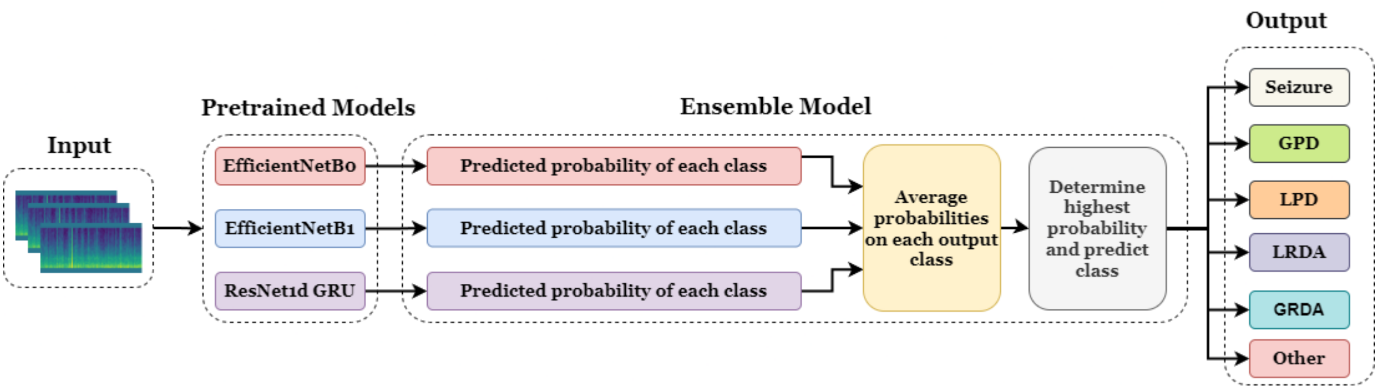

Figure 1. An illustration of ensemble learning approach used for brain activity prediction.

deep residual connections, further augmented the ensemble's capacity to learn complex representations. By averaging the predictions of these diverse models, we aimed to create a more robust and generalizable classifier for harmful brain activity.

III. Implementation Details

A. Dataset Description

The dataset used for this research is the Harmful Brain Activity Classification (HMS) dataset provided by CCEMRC [19]. This dataset includes EEG recordings and their corresponding spectrograms, segmented into 50-second-long samples over a 10-minute window. Each EEG sample and spectrogram pair is labeled by expert annotators, with labels indicating six types of brain activity: seizures, generalized periodic discharges (GPD), lateralized periodic discharges (LPD), lateralized rhythmic delta activity (LRDA), generalized rhythmic delta activity (GRDA), and other. An examples of EEG spectrograms of each category are shown in Figure 2.

B. Data Aggregation and Spectrogram Generation

The initial preprocessing stage involved aggregating relevant information at the EEG recording level. For each unique EEG recording, we extracted key features such as the identifiers of the first and last spectrograms, the temporal boundaries of the recording, the corresponding patient identifier, and the cumulative sum of target labels. This aggregation step was essential for organizing the data and ensuring that the model could learn patterns associated with entire EEG recordings. To further optimize data loading and processing, we implemented a flexible mechanism for accessing both training and test spectrograms. If preprocessed spectrograms were available, they were loaded directly from NumPy arrays to expedite the process. Otherwise, the spectrograms were read from their original Parquet files. For the training data, a custom DataGenerator class was developed to efficiently load, preprocess, and augment the spectrogram data in batches, enhancing the model's training efficiency.

C. Training Details

The deep learning model was trained using the Keras framework with TensorFlow as the backend, leveraging the computational power of multiple GPUs for acceleration. To optimize performance and memory usage, mixed precision training was employed, enabling the utilization of both 16-bit and 32-bit floating-point representations during computations. A 5-fold cross-validation strategy was implemented to ensure robust model evaluation and mitigate potential data leakage. To optimize the training process, we explored two learning rate schedules: cosine annealing and step decay. These schedules dynamically adjusted the learning rate during training, allowing for a balance between exploration and exploitation of the parameter space. Since the model outputs an array of predicted probabilities for each label class, we employ the Kullback-Leibler (KL) divergence metric, defined with 1, to evaluate its performance.

$$D_{KL}(P\|Q) = \sum_i P(i) \log \left(\frac{P(i)}{Q(i)} \right) \tag{1}$$

This metric facilitates the comparison of two distributions—in this context, the predicted probabilities and the actual labels. Moreover, the differentiability of the KL divergence allows it to be used directly as a loss function during model training, eliminating the need for an alternative function.

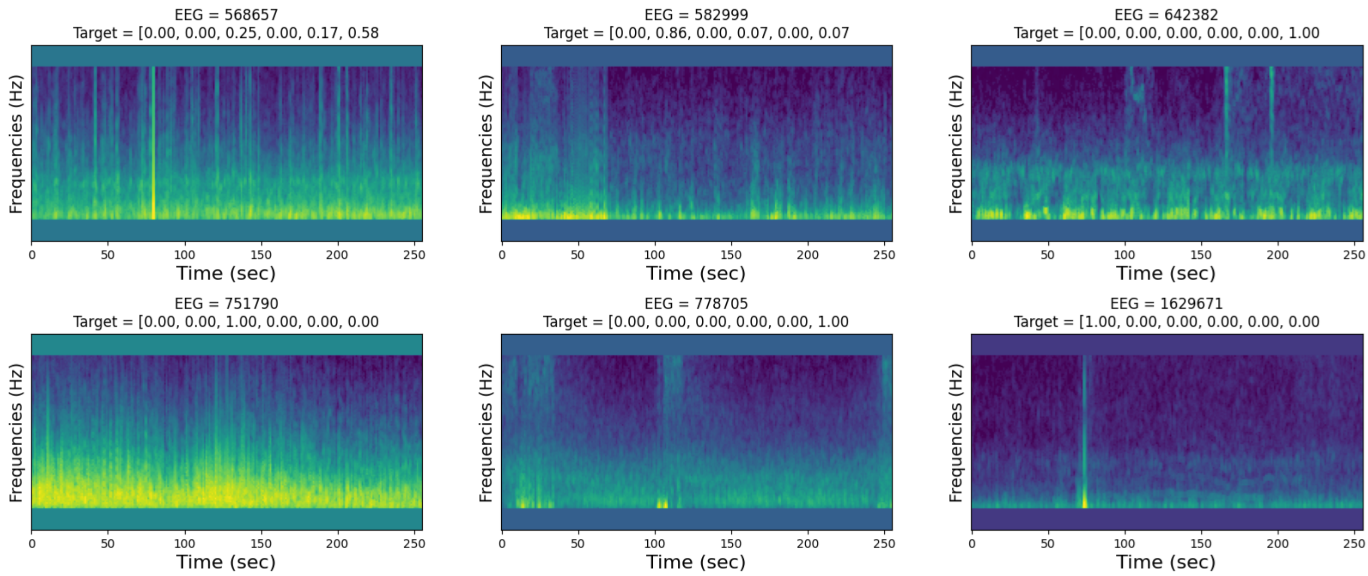

Figure 2. Visualization of EEG spectrogram. The associated target values beneath each spectrogram represent the probability distribution across six classes of harmful brain activity: Seizure, LPD, GPD, LRDA, GRDA, and Other. Notably, some spectrograms exhibit a single dominant class (e.g., EEG 642382, 778705), while others display a more mixed distribution (e.g., EEG 568657).

IV. Results and Discussion

The ensemble model, combining predictions from three distinct models, ResNet1d GRU, EfficientNetB1, and EfficientNetB0, achieved a competitive Kullback-Liebler divergence score of 0.3 on the testing dataset. This demonstrates the effectiveness of integrating diverse model architectures to enhance predictive performance. The ResNet1d GRU model, designed to capture temporal dependencies in EEG signals, likely contributed to the ensemble's ability to discern patterns over time. The EfficientNet models, known for their efficiency and accuracy in image classification, likely leveraged their pre-trained knowledge to extract relevant features from the spectrograms.

The ensemble's success can be attributed to several factors. Firstly, the use of diverse architectures allowed the ensemble to capture a wider range of features and patterns in the data, compared to relying on a single model. Secondly, the weighted averaging of predictions from different models helped to mitigate the impact of individual model biases and errors. The weights assigned to each model (0.45 for ResNet1d GRU, 0.15 for EfficientNetB1, and 0.40 for EfficientNetB0) were determined through experimentation and optimization to maximize the ensemble's performance.

However, there is room for further improvement. The current ensemble relies solely on model averaging, which might not fully exploit the potential diversity of the models. Exploring more advanced ensemble techniques, such as stacking or boosting, could lead to even better performance. Additionally, fine-tuning the weights assigned to each model or incorporating additional models into the ensemble could further enhance the results. Therefore, the ensemble model demonstrates the power of combining diverse models to achieve competitive performance in the challenging task of classifying harmful brain activity from EEG data. Future work will focus on refining the ensemble approach and exploring additional techniques to improve the accuracy and robustness of the predictions.

V. Conclusion

In this study, we demonstrated the effectiveness of ensemble modeling in classifying harmful brain activity from EEG data. By combining the strengths of diverse model architectures, namely ResNet1d GRU, EfficientNetB1, and EfficientNetB0, we achieved a competitive Kullback-Liebler divergence score of 0.3 on the testing dataset. This highlights the potential of ensemble methods to improve predictive performance in complex classification tasks. Future research will focus on refining the ensemble approach, exploring additional model architectures, and incorporating more sophisticated ensemble techniques to enhance further the accuracy and robustness of harmful brain activity classification.

References

[1] N.-H. Liu, C.-Y. Chiang, and H.-C. Chu, "Recognizing the degree of human attention using eeg signals from mobile sensors," *Sensors (Basel, Switzerland)*, vol. 13, pp. 10273 – 10286, 2013.

[2] S. Ganesan, Y. Kiran, and S. Ram, "Harmful brain activity classification of spectrograms with transfer deep learning," 04 2024.

[3] K. Juneja and C. Rana, "Individual and mutual feature processed elm model for eeg signal based brain activity classification," *Wireless Personal Communications*, vol. 108, pp. 659 – 681, 2019.

[4] G. Altan, Y. Kutlu, and N. Allahverdi, "Deep belief networks based brain activity classification using eeg from slow cortical potentials in stroke," *International Journal of Applied Mathematics, Electronics and Computers*, pp. 205–210, 2016.

[5] A. A. E. Shoka, M. M. Dessouky, A. El-Sayed, and E. E. Hemdan, "Eeg seizure detection: concepts, techniques, challenges, and future trends," *Multimedia Tools and Applications*, pp. 1 – 31, 2023.

[6] A. Haider and B. Guragain, "Challenges and future trends of eeg as a frontier of clinical applications," *2023 IEEE International Conference on Electro Information Technology (eIT)*, pp. 484–498, 2023.

[7] R. Jana, S. Bhattacharyya, and S. Das, "Epileptic seizure prediction from eeg signals using densenet," *2019 IEEE Symposium Series on Computational Intelligence (SSCI)*, pp. 604–609, 2019.

[8] Z. Wan, M. Li, S. Liu, J. Huang, H. Tan, and W. Duan, "Eegformer: A transformer–based brain activity classification method using eeg signal," *Frontiers in Neuroscience*, vol. 17, 2023.

[9] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, 2015.

[10] K. Cho, B. van Merrienboer, Çaglar Gülçehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder–decoder for statistical machine translation," in *Conference on Empirical Methods in Natural Language Processing*, 2014.

[11] G. Toderici, D. Vincent, N. Johnston, S. J. Hwang, D. C. Minnen, J. Shor, and M. Covell, "Full resolution image compression with recurrent neural networks," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5435–5443, 2016.

[12] A. Foltyn, J. Deuschel, N. R. Lang-Richter, N. Holzer, and M. P. Oppelt, "Evaluating the robustness of multimodal task load estimation models," *Frontiers in Computer Science*, 2024.

[13] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, pp. 1735–1780, 1997.

[14] Keras, "Efficientnet b0 to b7," 2024.

[15] C. Patel, D. Undaviya, H. S. Dave, S. Degadwala, and D. Vyas, "Efficientnetb0 for brain stroke classification on computed tomography scan," *2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*, pp. 713–718, 2023.

[16] M. Tan and Q. V. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," *ArXiv*, vol. abs/1905.11946, 2019.

[17] D. Toumpanakis, "Imagenet dataset," *Radiopaedia.org*, 2021.

[18] S. Benkrama and N. E. H. Hemdani, "Deep learning with efficientnetb1 for detecting brain tumors in mri images," *2023 International Conference on Advances in Electronics, Control and Communication Systems (ICAECCS)*, pp. 1–6, 2023.

[19] J. Jing, Z. Lin, C. Yang, A. Chow, S. Dane, J. Sun, and M. B. Westover, "Hms - harmful brain activity classification," 2024.