ELE670 - Assignement 1

Lunelli Riccardo, Parola Riccardo, Santini Gabriele

September 25, 2023

Abstract

Objective. Classify patient pathology from EEG.

Approach. The classification task is performed using convolutional neural networks (CNNs), which have been widely used in computer vision and speech recognition to perform automatic feature extraction and classification but they also have successfully been applied to EEG signals. For this project we use EEGNet, a pretrained CNN model for EEG signals.

Main results. Our efforts focused on understanding the dataset and implementing the EEGnet model, resulting in significant performance improvements by transitioning to ReLU activation. However, the small dataset size underscores the need for a larger and more diverse dataset to enhance model accuracy. Future work involves partitioning data for disease-focused learning, enhancing model generalizability

1 Introduction

Working memory is a central component of our cognitive system, enabling us to temporarily hold and manipulate information. This function is instrumental in tasks such as language comprehension, reasoning, and learning. To gain a deeper understanding of the brain's activity during working memory tasks, it's essential to have detailed data on its electrical patterns. In this paper, we delve into a dataset derived from fifteen subjects participating in a verbal working memory task. These participants, all epilepsy patients, were under observation for their medical condition. They were engaged in a version of the Sternberg task, tailored to assess various facets of memory, including encoding, maintenance, and recall. In our research, we use the power of deep learning, specifically employing EEGNet, a convolutional neural network for EEG signals. Our efforts are concentrated on understanding the dataset and fine-tuning the EEGNet model. We explore two primary classification tasks: multiclass and binary. The multiclass classification aims to identify specific pathologies, while the binary classification focuses on determining the presence of hippocampal sclerosis. Notably, our implementation resulted in significant performance enhancements, especially when transitioning to the ReLU activation function. However, it's essential to note that the limited size of our dataset underscores the need for more extensive data to further refine our model's accuracy. Through our analysis, we aim to provide a deeper understanding of the neural dynamics associated with working memory, leveraging the capabilities of deep learning to achieve this goal.

2 Data descriptions

The dataset Dimakopoulos et al. (2023) used for the assignment has been recorded from fifteen subjects during a verbal working memory task. Subjects were epilepsy patients undergoing intracranial monitoring for localization of epileptic seizures. Subjects performed a modified Sternberg task Corbin and Marquer (2013) in which the encoding of memory items, maintenance, and recall were temporally separated. Subject characteristics and information on sessions (set size, match/mismatch, correct/incorrect, response, response time for each trial) are also provided. This dataset enables the investigation of working memory by providing simultaneous scalp EEG and iEEG recordings, which can be used for connectivity analysis, alongside reconstructed beamforming EEG sources that can enable further cognitive analysis such as replay of memory items.

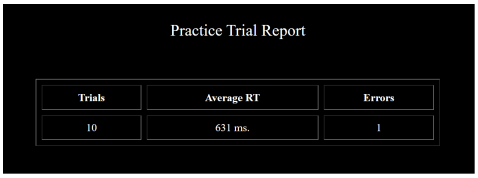


Figure 1: Our score at Sternberg Task

2.1 Subjects

The dataset was recorded from fifteen subjects during a verbal working memory task. Each patient is identified using 3 different id system: usz, sciadv and elife. Additional recorded metadata includes the age of the participant at the time of the task, the self-reported gender of the participant, and the clinical diagnosis of the patient as shown in Figure 2. The shared characteristic among all patients is their confirmed diagnosis of focal epilepsy.

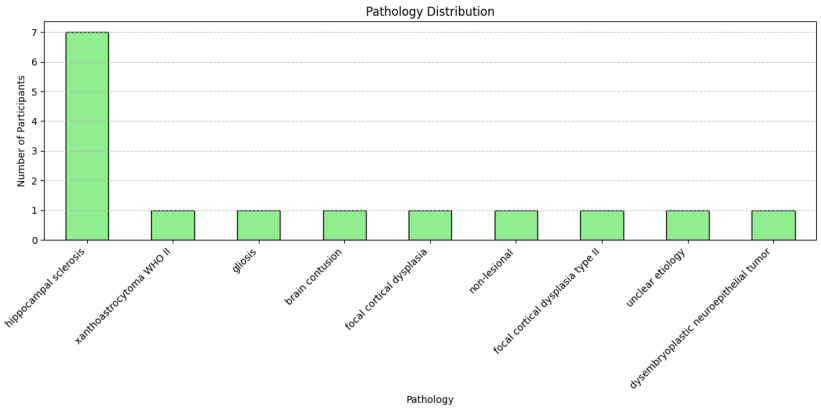
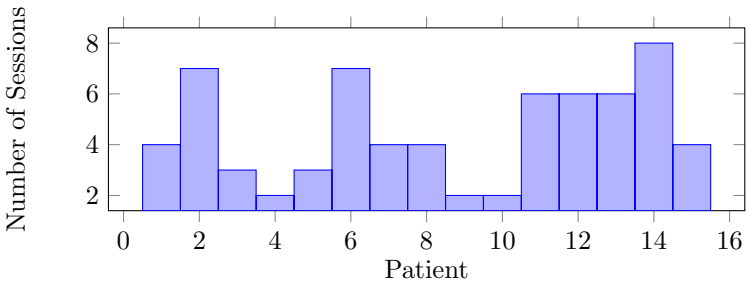


Figure 2: Pathology List Distribution

Patients participated in multiple test sessions conducted on different days, Figure 3 visually represents the distribution of the number of sessions in which each patient participated.

Figure 3: Distribution of Number of Sessions



2.2 Signal parameters

The dataset includes simultaneously recorded scalp EEG with the 10-20 system, intracranial EEG (iEEG) recorded with depth electrodes, waveforms, and the MNI coordinates and anatomical labels of all intracranial electrodes. The dataset includes also reconstructed virtual sensor data that were created by performing LCMV beamforming on the EEG at specific brain regions including, temporal superior lobe, lateral prefrontal cortex, occipital cortex, posterior parietal cortex, and Broca. The data sampling frequency varied for different patients, with values of 200, 4000, or 4096 Hz, while the power line frequency was set at 50 Hz. Each subject participated in multiple working memory sessions, each

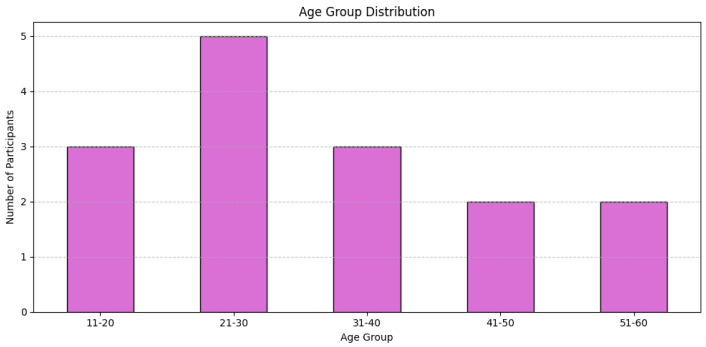


Figure 4: Age Group Distribution

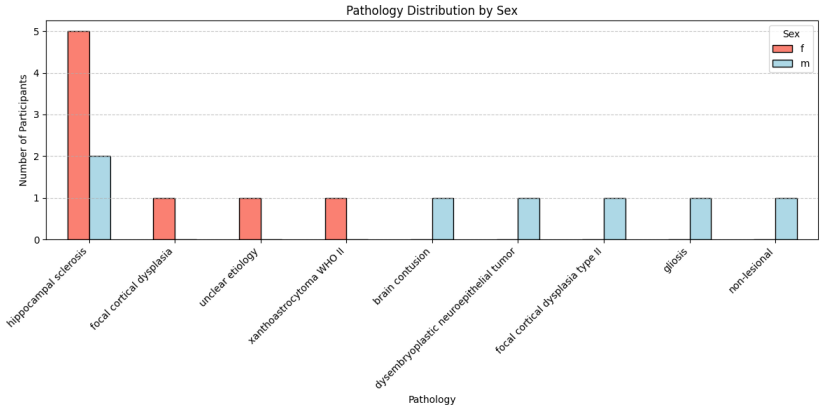


Figure 5: Pathology Distribution by sex

comprising approximately 50 trials. Each trial included an encoding period of 2 seconds, a maintenance period of 3 seconds, and a subsequent probe letter presentation, during which participants indicated whether the probe was part of the memorized letter string. The EEG data was epoched with an epoch length of 8 seconds and was recorded using a high-pass filter with a cutoff frequency of 0.5 Hz and a low-pass filter with a cutoff frequency of 1000 Hz. No information about electrode placement protocols followed, useful to maintain consistency EEG studies, is given.

3 Pre-processing

In this section, we will discuss the data pre-processing and its reasoning, motivations, and challenges. The dataset provided is very heterogeneous and part of the aim of this phase is to make it homogeneous and to create better conditions for learning.

3.1 Data Loading

The dataset follows the Brain Imaging Data Structure standard also called BIDS: Neuroimaging experiments generate complex data, often organized differently among researchers, causing confusion and wasted time. The Brain Imaging Data Structure (BIDS) Gorgolewski et al. (2016) provides a simple and widely accepted framework for organizing and sharing neuroimaging and behavioral data, promoting consistency and efficiency in research practices.

This type of standard allows to use of tools such as the MNE library to easily import and apply some standard methods to the data.

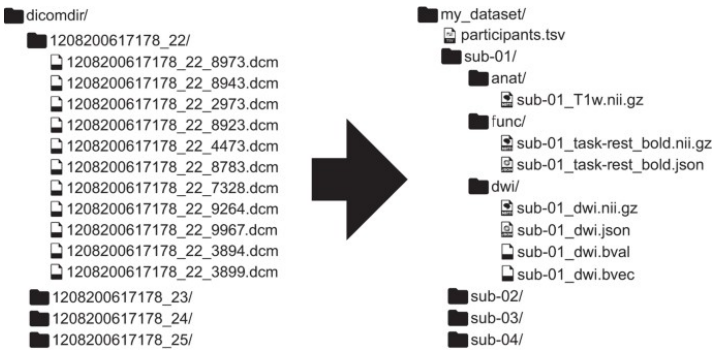


Figure 6: BIDS standard structure

3.2 Frequency and Cannel Selection

Not all patient data were collected the same way: for different patients in this dataset there might be different numbers of sensors or different frequencies of acquisition. Most of the data were collected at 200hz and some were collected at a much higher frequency of 4000hz. We decided to convert all to 200hz since is the minimum possible to use. The Nyquist-Shannon sampling theorem, a fundamental principle in signal processing, asserts that to accurately represent a signal, the sampling rate must be at least twice the frequency of the highest component in the signal. In the context of EEG data, this theorem holds significant implications. Gamma waves, which are associated with higher cognitive functions, including memory and problem-solving, are typically found in the 30-100 Hz range of frequency. By adhering to the Nyquist criterion, we recognize that the sampling rate should be a minimum of twice the highest frequency component of interest. Therefore, in the case of gamma waves, using a 200 Hz sampling rate is justified. This choice allows us to accurately capture and represent the essential information contained within these high-frequency brain oscillations.

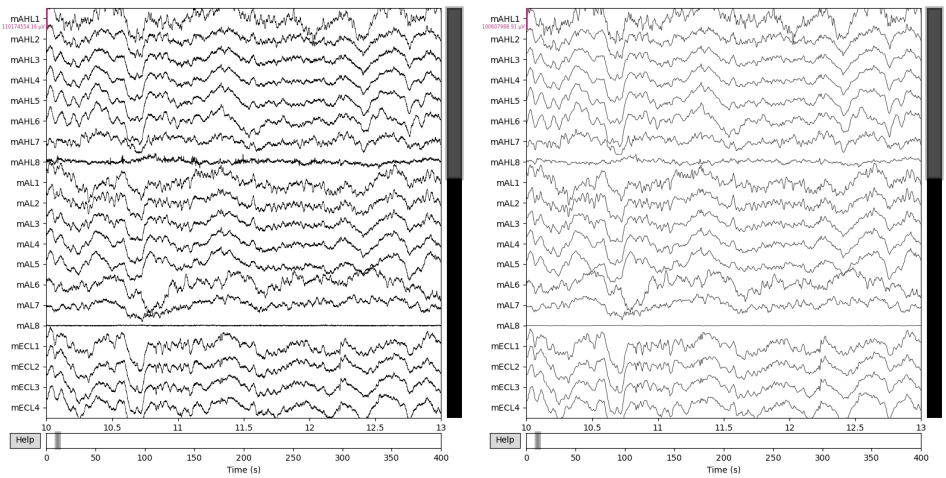
Not only was the frequency inconsistent across the dataset, but the number of channels or sensors used for EEG recordings also varied among different patients. This variation in the number of sensors adds another layer of complexity when standardizing the data for analysis.

We decided to use a consistent set of 8 channels for EEG sensors across all patient data. This decision was made for several reasons:

- Data Consistency: By using the same set of 8 channels for all patients, we ensure data consistency, making it easier to compare and analyze EEG signals across different individuals.
- Reduced Dimensionality: While EEG systems can have a large number of channels (often 32 or more), using a reduced number of channels simplifies data processing and analysis.
- Consistency in Analysis: Researchers often focus their analyses on specific regions of the brain associated with the cognitive functions under investigation. By choosing a consistent set of 8 channels, we can ensure that the data we analyze are relevant to our research questions and hypotheses.

3.2.1 Filtering and Artifact Removal

Filtering is a very useful tool in data and signal analysis in our case we applied a low-pass filter with a lower cutoff frequency of 1 Hz to the EEG data. It removes low-frequency noise or drifts while preserving higher-frequency components. At the same time, we applied a high-pass filter with a frequency of 50hz. This allowed us to create much more clean and clear data as you can see in the following images and resulting in better performances for the model.



(a) Unfiltered EEG (b) Filtered EEG

Figure 7: Comparison of EEG data

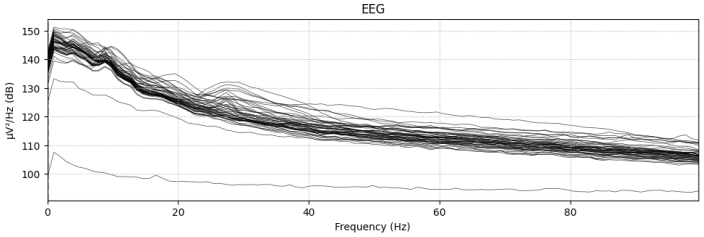


Figure 8: Unfiltered Power Spectral Density (PSD) EEG

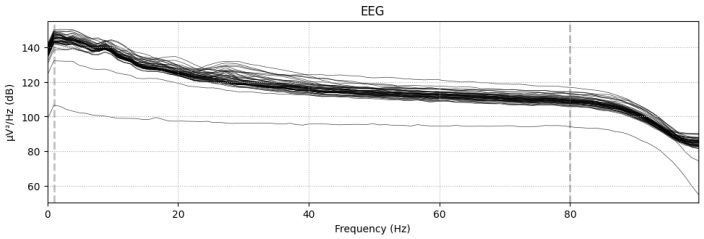


Figure 9: Filtered Power Spectral Density (PSD) EEG

We than divided the data in epochs of 8 seconds each, segmenting the continuous EEG data into shorter epochs or time windows. Each epoch typically corresponds to a specific event or stimulus in the experiment. This step allows you to analyze EEG responses to specific events.

Some of the measurings might present some artifacts which are signaled in the dataset so we eliminated those types of epochs in order to get better and more clean results.

4 Flowchart of the method

In this section, we describe the implementation of EEGNet for classification and provide comprehensive details about the training process. We utilized PyTorch as the foundational library for our training pipeline.

We explored two distinct types of classification:

- **Multiclass Classification:** The aim here is to determine the specific pathology affecting the patient. This classification poses a more complex challenge due to the presence of an unbalanced

dataset. Given that for all diseases, except for hippocampal sclerosis, we have data from only one patient, there's a potential risk that the network might learn to identify the patient rather than generalizing to the disease.

- **Binary Classification:** This ascertains whether a patient has hippocampal sclerosis. While this model might offer better generalization, it is limited in its ability to classify patients based on other pathologies.

For evaluating our models, we employed the F1 score and accuracy as our primary metrics.

4.1 Dataset and Dataloader

The initial phase of our training pipeline involved the creation of a Dataset class. This class is designed to transform EEG data into a format compatible with PyTorch.

The class iterates over all patients and their respective sessions. Each session is subsequently divided into epochs, adhering to the BIDS standard. We treated each epoch as an individual sample for our model, with the label determined by the patient's pathology.

Prior to segmenting the data into epochs, we applied the preprocessing techniques previously detailed in an earlier section. For the epoch division, we utilized the `mne_bids` library, which automated the process and concurrently eliminated epochs containing artifacts.

The concluding step involved partitioning the dataset into training, validation, and testing subsets. Subsequently, we used the dataloaders to allow batch training of our models.

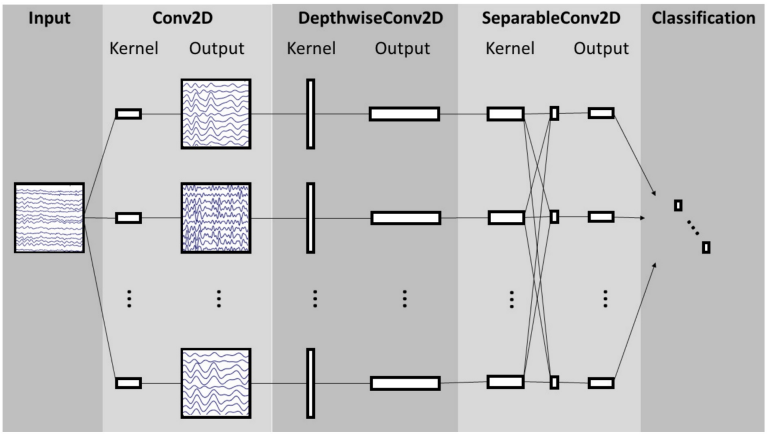


Figure 10: EEGNet structure Lawhern et al. (2018)

4.2 EEGNet

We now delve into the implementation of EEGNet as described in the original paper: Lawhern et al. (2018). This network comprises an initial layer of standard 2D convolution, succeeded by a depthwise convolution and another layer utilizing separable convolution. The primary function of these initial three layers is to serve as a feature extractor. Subsequently, a dense layer is appended for classification purposes. The architecture of the network is illustrated in Fig.10. To mitigate the risk of overfitting, a dropout rate of 0.5 is employed. Looking into the implementation of the proposed EEGNet and observed its utilization of the ELU activation function Clevert et al. (2016), which doesn't consistently exhibit optimal convergence properties. Consequently, we opted to modify the original model against a custom variant that employs the ReLU 1Nair and Hinton (2010) activation function.

4.3 Training Phase

With the dataset prepared and the classification model implemented, we moved on the training process. For all training sessions, we employed the Adam optimizer in conjunction with the Cross Entropy loss. To counteract overfitting, we adopted an early stopping strategy, specifically saving the model iteration that achieved the highest F1 score on the validation dataset.

4.4 Multiclass classification

We first trained the original multiclass model with a learning rate of 0.001. After that, we trained our custom model with the ReLU activation function using the same settings. The learning curves for both models are presented in Fig.11 and Fig.12.

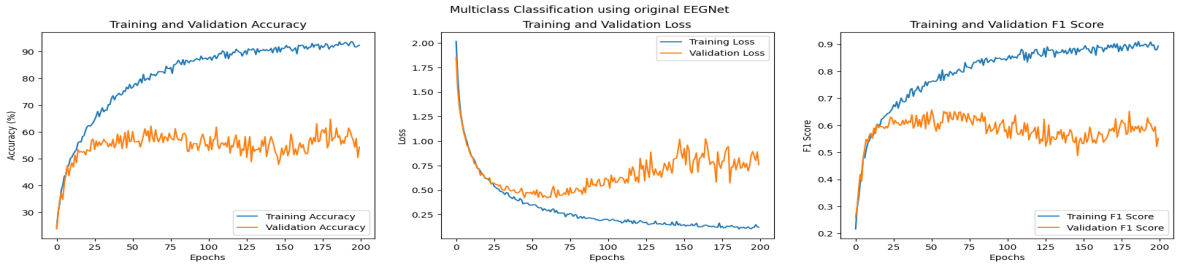


Figure 11: Training curves for Multiclass Classification using original EEGNet

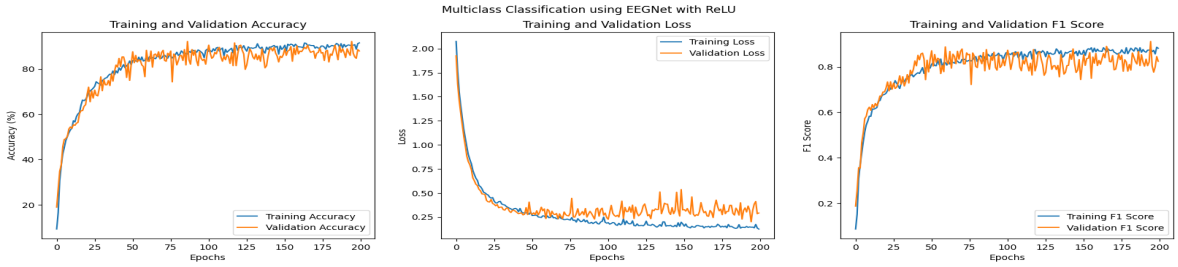


Figure 12: Training curves for Multiclass Classification using modified EEGNet

4.5 Binary classification

We trained the binary models using a reduced learning rate of 0.0001, as it yielded better performance in this scenario.

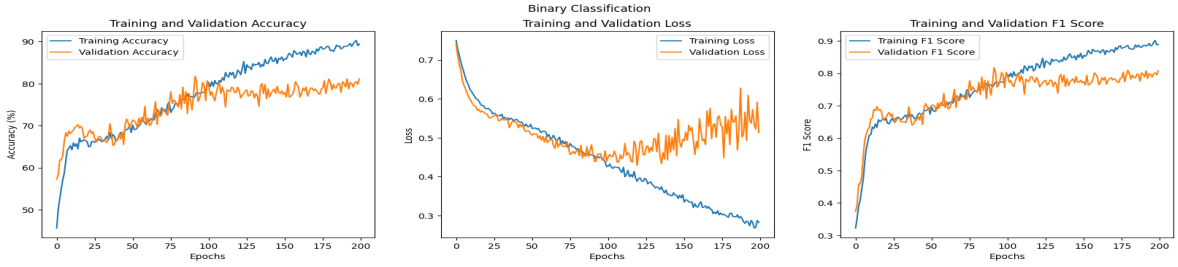


Figure 13: Training curves for Binary Classification using original EEGNet

4.6 Experimental Results

Given the settings described above, we achieved satisfactory results, especially considering the challenges posed by the small and unbalanced dataset. It's noteworthy to observe the training curves: the original model tends to overfit earlier compared to our proposed model. The original model, when applied to the binary task, struggles to learn a robust representation of the data. The validation dataset's loss begins to overfit after just a few epochs, and this issue isn't attributed to an high learning rate since the training loss decreases at a very slow pace.

As we hypothesized, using ReLU (as shown in Fig.12 and Fig.14) enabled the model to converge to more optimal solutions compared to the version using ELU (illustrated in Fig.11 and Fig.13).

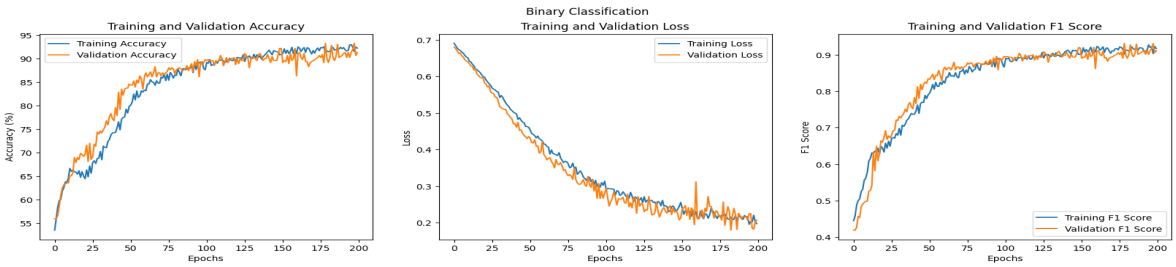


Figure 14: Training curves for Binary Classification using modified EEGNet

In Table 1, the results from the test set further corroborate our findings, with our models consistently outperforming the original.

Table 1: Accuracy and F1 score on test dataset

	Multiclass Acc.	Multiclass F1	Binary Acc.	Binary F1
EEGNet	61.04%	0.62	79.39%	0.79
EEGNet with ReLU (our)	**89.61%**	**0.83**	**87.42%**	**0.87**

5 Results and Conclusion

In conclusion, our efforts primarily revolved around gaining a comprehensive understanding of the dataset, including its construction, recording process, sampling techniques, and various parameters. With this understanding of the data, we set out to implement the EEGnet model in order to identify pathologies within individual epochs of EEG signals. Our initial findings indicated that the model was already performing well; however, we were able to significantly enhance its performance by transitioning from the ELU activation function to ReLU.

It is worth noting that the statistical population of the dataset was relatively small and not ideally suited for our specific task. To further refine the accuracy and robustness of our model, a more extensive and diverse dataset is imperative. While a larger volume of data per patient may not be necessary, a greater number of patients are indeed needed for comprehensive and reliable results.

As we look toward future endeavors for this project, it is crucial to ensure that our model is learning to identify diseases rather than individual patients. To achieve this, an idea is partitioning the dataset into distinct training and testing subsets, where the testing subset comprises patients who have never been encountered by the network during training. This approach will help mitigate the risk of overfitting to individual patients and advance the generalizability of our disease identification model.

References

Clevert, D.-A., Unterthiner, T., and Hochreiter, S. (2016). Fast and accurate deep network learning by exponential linear units (elus).

Corbin, L. and Marquer, J. (2013). Is sternberg's memory scanning task really a short-term memory task? *Swiss Journal of Psychology*, 72:181.

Dimakopoulos, V., Mégevand, P., Stieglitz, L., Imbach, L., and Sarnthein, J. (2023). Dataset of intracranial eeg, scalp eeg and beamforming sources from epilepsy patients performing a verbal working memory task.

Gorgolewski, K. J., Auer, T., Calhoun, V. D., Craddock, R. C., Das, S., Duff, E. P., Flandin, G., Ghosh, S. S., Glatard, T., Halchenko, Y. O., Handwerker, D. A., Hanke, M., Keator, D., Li, X., Michael, Z., Maumet, C., Nichols, B. N., Nichols, T. E., Pellman, J., Poline, J.-B., Rokem, A.,

Schaefer, G., Sochat, V., Triplett, W., Turner, J. A., Varoquaux, G., and Poldrack, R. A. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data*, 3(1):160044.

Lawhern, V. J., Solon, A. J., Waytowich, N. R., Gordon, S. M., Hung, C. P., and Lance, B. J. (2018). EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces. *Journal of Neural Engineering*, 15(5):056013.

Nair, V. and Hinton, G. E. (2010). Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th international conference on machine learning (ICML-10)*, pages 807–814.