Graph-based EEG Analysis

Madeleine Hueber
Computer Science
EPFL
Lausanne, Switzerland
madeleine.hueber@epfl.ch

Ilias Marwane Merigh
Computer Science
EPFL
Lausanne, Switzerland
ilias.merigh@epfl.ch

Alfred Arnaud
MTE
EPFL
Lausanne, Switzerland
alfred.arnaud@epfl.ch

Mehdi Abdallahi
Energy Science
EPFL
Lausanne, Switzerland
mehdi.abdallahi@epfl.ch

***Abstract*—Accurate seizure detection from EEG data is critical for improving the quality of life for individuals with epilepsy. While traditional approaches often overlook the spatial relationships between electrodes, we investigate a graph-based framework to better model the structural connectivity of brain regions. Using a subset of the Temple University Hospital EEG Seizure Corpus [1], we construct graphs based on the 10–20 electrode layout. We evaluate multiple deep learning models and preprocessing strategies, including graph construction and class imbalance handling. Our results show that a Gated Graph Neural Network (GatedGNN) outperforms alternative models. An ensemble of the top three configurations further boosts performance, achieving a Macro-F1 score of 84.8% on the Kaggle test set provided for the competition [2]. These findings highlight the effectiveness of geometric deep learning for robust, interpretable EEG-based seizure detection.**

***Index Terms*—EEG, Seizure Detection, Graph Neural Networks, Spatial Modeling, Ensemble Learning**

I. Introduction

Epilepsy is a neurological disorder that affects nearly 50 million people worldwide and is characterized by recurrent seizures caused by abnormal, excessive neuronal activity in the brain. Electroencephalography (EEG) is the most widely used non-invasive tool for monitoring this activity. However, EEG signals are notoriously challenging to model due to their high dimensionality, noise, and non-stationary nature.

Traditional approaches often neglect the spatial relationships between EEG electrodes, treating them as independent channels. In contrast, spatial topology plays a critical role in seizure propagation. Graph-based modeling enables encoding of electrode spatial layout, allowing Graph Neural Networks (GNNs) to learn structured temporal-spatial patterns. The contributions of our work are listed below:

1) Designed multiple inter-electrode graph configurations, edge weighting schemes, and signal transformations.
2) Benchmarked various model architectures.
3) Performed hyperparameter tuning on the best model configuration.
4) Ensembled the top-performing configurations to achieve higher score.

The remainder of this paper is organized as follows: Section II discusses related work. Section III presents our approach, including graph construction, signal transformations, model design, and ensemble strategy. Section IV covers our experiments, results, and analysis. Section V concludes the paper. Additional materials are provided in the Appendix.

II. Related Work

EEG-based seizure detection has drawn growing interest due to its important medical application on epilepsy monitoring and intervention. Early deep learning efforts focused on:

- Temporal models: LSTM networks captured seizure dynamics over time [3], but operated on each channel independently, ignoring spatial context.
- Spatial model: CNNs were then applied to extract localized patterns across electrodes arranged in 2D grids [4], but their fixed convolutional kernels cannot flexibly mirror the electrodes' topology.

Graph-based methods address these shortcomings by encoding inter-electrode connectivity in a graph and performing message passing along its edges [5]. Among these, Gated Graph Neural Networks (GGNNs) [6] advance standard graph convolution by embedding gated recurrent units at each node. This design:

1) models long-range dependencies through multi-step message updates,
2) controls information flow via gating, preventing noisy signals from overwhelming the hidden state.

III. Method

A. EEG Montage and Graph Construction

To capture the spatial structure of EEG signals, we explored several graph-based representations, varying node layouts, edge weighting schemes, and signal preprocessing techniques.

Nodes: *Standard 10–20* system provided with the original dataset. It is a clinically established electrode placement scheme that ensures uniform scalp coverage (as shown in Figure 2).

Edges:

1) Inverse distance: $w = \frac{1}{d}$
2) Inverse-squared distance: $w = \frac{1}{d^2}$
3) Gaussian kernel: $w = \exp\left(-\frac{d^2}{2\sigma^2}\right)$

Graph Construction: EEG electrodes relationships were modeled with different approaches:

1) Fully Connected (Baseline): We construct a fully connected graph using the distances from `distances_3d.csv`, with edge weights defined using Gaussian kernel. This setup includes all pairwise electrode connections, allowing the model to learn global relationships across the scalp.

2) k-Nearest Neighbor (k-NN) Graph: We construct a graph where each electrode is connected to its k nearest neighbors based on 3D distance. This approach emphasizes local spatial relationships while keeping the graph sparse enough to reduce potential overfitting.
3) Temporal Emphasis Graph: In this variant, we start from the same fully connected Gaussian 10–20 graph but boost edge weights that involve temporal electrodes (T3, T4, T5, T6) by a factor α. This strategy is biologically motivated: "Approximately 60 percent of all forms of epilepsy are focal in origin, with the majority originating in the temporal lobe" [7]. By amplifying connections to temporal nodes, the model is pushed to focus more on these critical regions during seizure detection.

Signal Transformations:

1) *Fast Fourier Transform (FFT)*, which converts time-domain signals into the frequency domain:

$$X(f) = \sum_{n=0}^{N-1} x(n) e^{-j2\pi f n/N}$$

2) *Discrete Wavelet Transform (DWT)*, which captures both time and frequency information using localized basis functions:

$$W[a,b] = \sum_{n=0}^{N-1} x[n]\, \psi^*_{a,b}[n]$$

B. Models

Table I lists every architecture we implemented. For clarity, Figure 1 provides block diagrams of the most complex network (Generative GNN) and the highest scoring one (Gated GNN). Finally, we summarize the different model components in the Table I. Below, models are grouped into *sequence baselines*, *static-graph baselines*, and *advanced / dynamic graph models*.

Sequence baselines

1) **LSTM Classifier:** [3] A single-layer LSTM reads each channel's time-series and feeds the final hidden state to a fully-connected (FC) layer, there are no spatial coupling between electrodes.
2) **Transformer Classifier:** [8] Treats each electrode as a token. Per-node features are linearly projected to d_{model}, enriched with positional encodings, processed by a two-layer Transformer encoder, they are then mean-pooled, and passed to a FC head.

Static-graph baselines

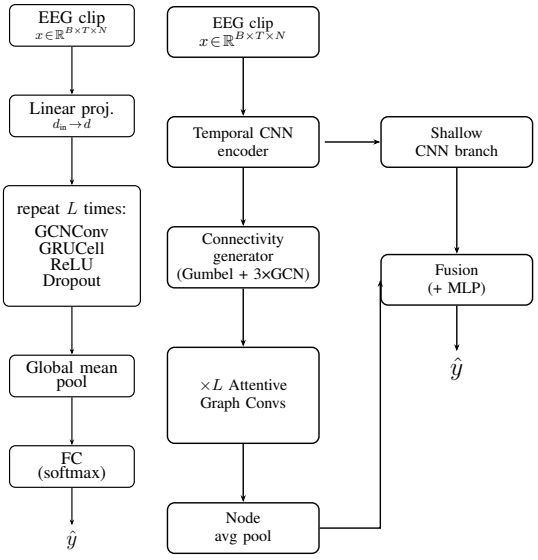
1) **GCN Classifier:** [9] The FFT node features flow through L `GCNConv` layers with batch-norm and ReLU, they then go through global mean pooling and a FC layer.

Advanced / dynamic graph models

1) **Generative GNN Classifier:** [10] Contains the following:
   (i) Temporal CNN encoder for node features.
   (ii) Connectivity generator (three parallel GCNs and Gumbel-softmax) for a dynamic adjacency matrix.
   (iii) Spatial decoder with L attentive graph-conv layers.
   (iv) Shallow CNN branch treats the EEG as a $N \times T$ image.
   (v) Fusion and two-layer MLP classifier.
2) **GatedGNN Classifier:** [6] Adds a GRU gate after every graph convolution, i.e. $h^{\ell+1} = \mathrm{GRU}(\texttt{GCNConv}(h^\ell), h^\ell)$, which improves long-range credit assignment.
3) **GatedGNN$^{\text{Pos}}$ Classifier:** Extends GatedGNN with sine–cosine positional encodings of node indices [11], concatenated before the initial projection to endow the network with an explicit notion of electrode order.
4) **GatedGNN$^{\text{Att}}$ Classifier:** Includes attention-based aggregation after each convolution. Specifically, node updates follow $h^{\ell+1} = \mathrm{GRU}(\texttt{AttnAgg}(\texttt{GCNConv}(h^\ell)), h^\ell)$, where `AttnAgg` denotes multi-head self-attention. This improves the model's capacity to focus on relevant neighbors in irregular graphs.

TABLE I: Summary of model components. "Temp CNN" = temporal convolution; "GRU-gate" = gated update; "Dyn. adj." = dynamically generated adjacency; "Attn. agg." = attention-based neighbor aggregation.

Model	Temp CNN	GCN	GRU-gate	Dyn. adj.	Attn. agg.
LSTM	–	–	–	–	–
Transformer	–	–	–	–	–
GCN	–	✓	–	–	–
Generative GNN	✓	✓	–	✓	✓
GatedGNN	–	✓	✓	–	–
GatedGNN$^{\text{Pos}}$	–	✓	✓	–	–
GatedGNN$^{\text{Attention}}$	–	✓	✓	–	✓



(a) GatedGNN (b) Generative GNN

Fig. 1: **Architectural schematics.** (a) GatedGNN uses L gated message-passing layers (GCNConv + GRU) followed by global pooling. (b) Generative GNN first encodes temporal features, then *learns* a dynamic adjacency via a connectivity generator; attentive graph convolutions and a shallow CNN branch are fused before classification.

C. Ensemble Learning

We employed ensemble learning [12] by aggregating the predictions of the top three models from the best-performing

configuration using majority voting, an effective technique shown to improve robustness and generalization. An odd number of models (three) was selected to avoid tie cases and ensure decisive outcomes. In our case, ensemble learning is implemented as a straightforward majority vote over the output predictions of the individual models. Further details on the ensemble setup and its effectiveness are provided in Section IV-C.

IV. Experiments

A. Dataset

We used a pre-defined subset of the Temple University Hospital EEG Seizure Corpus (TUSZ), one of the largest publicly available EEG seizure datasets. This subset includes recordings from 50 patients for training and 25 for testing, as provided by the dataset authors. EEG signals were recorded at 250 Hz using 19 electrodes placed according to the international 10–20 system, with inter-electrode distances also included, as shown in Figure 2.

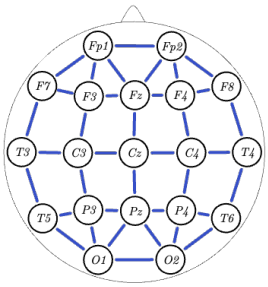


Fig. 2: Standard 10-20 montage for EEG acquisition. Blue edges represent the graph based on distances between electrodes on the scalp.

The recordings were segmented into non-overlapping 12-second windows, each labeled as either seizure or non-seizure. We followed the provided training set and randomly split it into 80% for training and 20% for validation. The resulting training set contains 10,394 samples and exhibits a pronounced class imbalance: 80.6% non-seizure and 19.4% seizure. In addition to the validation set, we used the 57%-available Kaggle test set provided for the competition as a complementary benchmark to assess the generalization performance of our models.

B. Data Processing

To address strong class imbalance, we explored oversampling the seizure class as an alternative to weighted cross-entropy loss. We tested both simple duplication and duplication with added Gaussian noise. While noise injection reduced the gap between training and validation loss, it did not improve the validation macro-F1 score. We therefore retained oversampling without noise in the final setup.

Performance differences across the various distance-based graph constructions and signal transformations were negligible (within 0.1 macro-F1), leading us to retain the 10–20 electrode layout, Gaussian edge weighting, and Fast Fourier Transform for spectral representation.

C. Results

We evaluated all models on the validation set using an 80/20 random split, with a fixed seed (42) to ensure reproducibility. The results, summarized in Table II, compare the performance of various models after manual hyperparameter tuning. Macro-F1 score was used as the primary evaluation metric.

Model Architecture	Val Macro-F1
GCN	69.7%
GatedGNN	**78.9%**
GatedGNN$^{\text{Att}}$	78.7%
GatedGNN$^{\text{Pos}}$	78.5%
GenerativeGNN	72.3%
LSTM	67.9%
Transformer	74.3%

TABLE II: Validation Macro-F1 scores of different model architectures.

Since GatedGNN yielded the best initial results, we performed a grid search over the following hyperparameter space:

- `hidden_channels`: 32, 64, 128, 256
- `gnn_layers`: 3, 4, 5
- `dropout`: 0.05, 0.10, 0.20

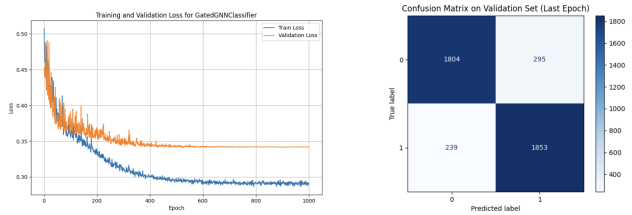
The top three configurations were then retrained on the **full data**, and their scores on the test set are reported in Table III. Finally, we combined the top three configurations using an ensemble model to enhance robustness of the prediction.

Hidden Channels	GNN Layers	Dropout	Val Macro-F1	Test Macro-F1
256	4	0.10	78.9%	82.2%
128	4	0.10	78.8%	82.1%
128	5	0.20	78.6%	82.1%
Ensemble (Top 3)			**80.3%**	**84.8%**

TABLE III: Macro-F1 scores of the top 3 GatedGNN configurations and their ensemble

To assess robustness, we ran additional experiments with different random seeds. Results varied within ±1.5%, confirming the stability of our approach.

During grid search, we trained each configuration for 500 epochs, as losses typically converged around epoch 400, allowing us to reduce resource usage. We used a batch size of 512 across all experiments. For the final training on the full dataset, we extended training to 1000 epochs to maximize performance. Figure 3 shows the loss curves and confusion matrix on the validation set for the best model configuration (row 1 in Table III).



(a) Training and validation loss curves

(b) Confusion matrix on the validation set

Fig. 3: Diagnostics for the best model configuration (Table III), row 1

Notes. The validation set is here augmented, hence the equal number of samples in each class.

Since GatedGNN showed the best results, we investigated the impact of different graph construction strategies on its performance. Table IV reports the test Macro-F1 scores for each graph configuration

Graph Configuration	Val Macro-F1
Fully Connected 10-20	**78.9**%
k-NN (k=2)	76.8%
k-NN (k=3)	**78.9**%
k-NN (k=4)	74.6%
Temporal Emphasis ($\alpha = 2$)	77.7%

TABLE IV: Validation Macro-F1 scores of different graphs configurations.

These results show that graph construction strategies (explained in section III) can influence model performance. Both the fully connected 10-20 graph and the k-NN (k=3) graph achieved the highest Macro-F1 score, suggesting that either fully connected or moderately sparse local graphs can be effective in capturing EEG dynamics. The k-NN (k=2) and k-NN (k=4) configurations showed slightly lower performance, indicating that too few or too many local connections can reduce model effectiveness. The temporal emphasis configuration ($\alpha = 2$) is close to the top-performing graphs but did not surpass them, suggesting that emphasizing temporal regions might help but doesn't consistently outperform the default fully connected graph in this dataset. Overall, these results suggest that both fully connected and carefully balanced k-NN graphs can be effective for seizure detection.

D. Discussion

Our results demonstrate that graph-based modeling of EEG signals using a Gated Graph Neural Network (GatedGNN), significantly improves seizure detection compared to traditional approaches, outperforming both static-graph baselines like GCN (II) and sequential models such as LSTM and Transformer. Notably, it also outperformed the more complex GenerativeGNN , highlighting the benefit of controlled message propagation over architectural complexity alone. This approach was further validated by an ensemble of top-performing GatedGNN configurations (Table III).

Ablation studies on graph constructions showed that alternative strategies, including k-nearest neighbor settings and time-weighted edges, did not enhance the model's performance. Similarly classical oversampling methods to address class imbalance performed comparably to more sophisticated augmentation techniques. In particular, duplicating seizure signals with added Gaussian noise increased the training-validation loss gap without yielding macro-F_1 improvements.

These findings indicate that the main performance gains can be attributed to the GRU-style gating mechanism.

Despite these promising results, our study has several limitations :

- Evaluation was limited to a single dataset with a fixed 80/20 train-validation split, leaving generalizability to other EEG datasets and clinical settings uncertain. We observed this limitation on the final Kaggle competition test set, where some inconsistencies in scores appeared, potentially due to differences in patients data and limited diversity in our data.
- While our model captures spatial relationships between electrodes, it does not reflect dynamic changes in functional connectivity, especially around seizure onset. Future work could explore other signal transforms and architectures to better capture these dynamic changes.
- Although ensembling enhances robustness, it increases inference latency and system complexity. This tradeoff poses challenges for real-time or resource-constrained deployment scenarios, such as wearable or embedded medical devices.

Overall, our work demonstrates that integrating gated message-passing into GNNs is a robust and effective approach for EEG classification tasks, offering a compelling direction for future research and clinical applications.

V. Conclusion and Future Work

In this study, we introduced a graph-based deep learning framework for EEG seizure detection, showing that a Gated Graph Neural Network leveraging FFT spectral features and a 10–20 electrode layout outperforms existing approaches. Our findings underscore the importance of spatial modeling and gated message-passing mechanisms in effectively capturing seizure-related patterns within noisy EEG data. Furthermore, ensemble learning using the top 3 configurations increased performance, achieving a Macro-F1 score of **84.8%** on the available Kaggle test set.

Future work should focus on validating this approach across diverse and clinically sourced datasets, incorporating dynamic graph structures to model temporal changes in brain connectivity, and aligning training objectives more closely with evaluation metrics such as macro-F_1. Additionally, reducing model complexity for real-time deployment and improving interpretability will be essential for clinical integration.

Building on these results, our approach offers a strong foundation for developing reliable, efficient, and clinically relevant seizure detection systems.

REFERENCES

[1] "Shah, v., von weltin, e., lopez, s., mchugh, j.r., veloso, l., golmohammadi, m., obeid, i., picone, j.: The temple university hospital seizure detection corpus. frontiers in neuroinformatics 12 (nov 2018).,"

[2] "Epfl network machine learning 2025." https://www.kaggle.com/competitions/epfl-network-machine-learning-2025, 2025. Accessed: 2025-06-10.

[3] A. M. Abdelhameed, H. G. Daoud, and M. Bayoumi, "Deep convolutional bidirectional lstm recurrent neural network for epileptic seizure detection," in *2018 16th IEEE International New Circuits and Systems Conference (NEWCAS)*, pp. 139–143, 2018.

[4] M. T. Avcu, Z. Zhang, and D. W. Shih Chan, "Seizure detection using least eeg channels by deep convolutional neural network," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, p. 1120–1124, IEEE, May 2019.

[5] H. Mohseni, A. Maghsoudi, and M. Shamsollahi, "Seizure detection in eeg signals: A comparison of different approaches," pp. 6724–6727, 01 2006.

[6] Y. Li, D. Tarlow, M. Brockschmidt, and R. Zemel, "Gated graph sequence neural networks," 2017.

[7] McIntosh WC, Das JM., "Temporal seizure. [updated 2023 jul 4]. in: Statpearls [internet]. treasure island (fl): Statpearls publishing; 2025 jan-.,"

[8] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in *Advances in Neural Information Processing Systems (NeurIPS)*, pp. 5998–6008, 2017.

[9] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *International Conference on Learning Representations (ICLR)*, 2017.

[10] Z. Li, K. Hwang, K. Li, J. Wu, and T. Ji, "Graph-generative neural network for eeg-based epileptic seizure detection via discovery of dynamic brain functional connectivity," *Scientific Reports*, vol. 12, no. 1, p. 18998, 2022.

[11] V. P. Dwivedi and X. Bresson, "Graph neural networks with learnable positional encodings," in *ICML*, 2021.

[12] W. H. Beluch, T. Genewein, J. M. Nürnberger, and J. Köhler, "The power of ensembles for active learning in image classification," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021.

APPENDIX

A. Hyperparameter Tuning

Hidden Channels	GNN Layers	Dropout	Macro-F1
256	**4**	**0.10**	**78.9%**
128	**4**	**0.10**	**78.8%**
128	**5**	**0.20**	**78.6%**
256	4	0.05	78.6%
128	5	0.10	78.5%
256	3	0.05	78.3%
256	5	0.10	78.3%
128	3	0.10	78.1%
128	3	0.20	78.0%
128	3	0.05	77.8%
128	5	0.05	77.8%
256	5	0.05	77.7%
256	4	0.20	77.7%
256	3	0.10	77.5%
256	5	0.05	77.3%
128	4	0.20	77.2%
64	5	0.10	77.1%
32	4	0.10	76.9%
256	3	0.20	76.8%
64	3	0.10	76.6%
64	3	0.20	76.4%
32	5	0.10	76.5%
256	5	0.20	76.2%
64	4	0.10	76.0%
32	3	0.05	76.0%
64	5	0.05	75.8%
64	5	0.20	75.5%
64	4	0.20	75.4%
32	3	0.10	75.1%
32	5	0.05	74.9%
64	3	0.05	74.9%
128	4	0.10	74.8%
32	4	0.05	74.8%
64	5	0.10	74.7%
32	3	0.20	74.4%
32	4	0.20	74.4%
128	4	0.20	74.4%
256	4	0.20	73.8%
32	5	0.20	73.7%

TABLE V: Validation Macro-F1 scores for all GatedGNN hyperparameter combinations