Comparing time-series and graph-based models in EEG seizure detection

Vincent Roduit
MSc student, EPFL
Electrical Engineering
Sciper: 325140

Caspar Henking
MSc student, EPFL
Electrical Engineering
Sciper: 323825

Jan Zgraggen
MSc student, EPFL
Computational Science
Sciper: 344351

Aurel Mäder
PhD student, EPFL
Social Computing Lab
Sciper: 327599

***Abstract*—Epilepsy is a neurological disorder characterized by abnormal neuronal activity that can cause sudden disruptions in brain function, known as seizures [9]. Electroencephalography (EEG) is the primary tool for detecting these events. Machine learning methods can be effectively applied to detect seizures in EEG data. In this project, we compare the performance of time-series-based and graph-based machine learning architectures for seizure prediction. To this end, we experiment with different signal representations, graph construction methods, and model architectures. Our results show that hybrid models combining both temporal and graph-based processing of EEG data (such as GAT-LSTM) perform particularly well, achieving F1 scores of up to 80%.**

***Keywords*—** Seizure detection, Network Machine Learning, GCN, GAT, LSTM, CNN, ResNet

I. Introduction

Diagnosing and treating epilepsy at an early stage can significantly improve the quality of life for those affected. The International League Against Epilepsy (ILAE) defines an epileptic seizure as “a transient occurrence of signs and/or symptoms due to abnormal excessive or synchronous neuronal activity in the brain” [9]. According to the World Health Organization (WHO), approximately 50 million people were living with epilepsy in 2019 [26]. Encouragingly, studies have shown that with simple and low-cost medications, up to 70% of affected individuals can achieve seizure-free lives. However, reducing costs and improving the accuracy of epilepsy diagnoses remain critical public health challenges.

Diagnosing epileptic seizures from EEG scans using machine learning systems can significantly reduce costs compared to manual inspection. Electroencephalography (EEG) remains the primary tool for diagnosing and monitoring epileptic seizures. However, EEG sessions often last 30 minutes or longer, making manual analysis both time-consuming and prone to human error. Machine learning offers a promising solution by automating the interpretation of EEG data. Recent studies have demonstrated encouraging results in this area [30].

In this project, we compare geometric (Graph based e.g. GCN,GAT) to time series (Signal Based e.g. CNN, LSTM) learning architectures for detecting epileptic seizures in EEG data. The EEG data consists of signals from 19 scalp electrodes (see Section III), which can be treated either as independent channels of time series or as functionally connected nodes in a graph. By incorporating the relationships between channels, we hope to improve model performance.

II. Motivation and Related Work

A variety of geometric and non-geometric machine learning methods have been proposed for detecting seizures from EEG data. Although studies vary in data formats—such as segment lengths, channel configurations, and prediction targets (e.g., seizure type classification vs. binary detection) they offer valuable insights into effective modeling strategies. EEG data can be viewed as multi-channel time series. In this case, temporal features can be captured using deep learning architectures like 1D Convolutional Neural Networks (1D-CNNs) [6], Residual Networks (ResNet) [5], Long Short-Term Memory networks (LSTMs) [13], self attention-based models [3] or combinations of the latter [1], [23], [28]. These approaches model the temporal evolution of brain activity. If EEG signals are treated as graph-structured data, with scalp electrodes represented as interconnected nodes, Graph Neural Networks (GNNs), such as Graph Convolutional Networks (GCNs) [27], [29] and Graph Attention Networks (GATs) [31], can be used. These models capture spatial relationships across channels through learning inter-channel dependencies with graph convolution or attention mechanisms [10]. Moreover, hybrid models combining GNNs with temporal architectures like LSTMs have demonstrated strong performance in distinguishing seizure and non-seizure events [11], [15].

III. Dataset

For the following analysis, we use a subset of EEG recordings from the Temple University Hospital EEG Seizure Corpus [21]. The dataset is divided into training and test sets, consisting of 97 training patients and 25 test patients. Each patient had one or more EEG recording sessions. Each session was further divided into 12-second segments of non-overlapping brain activity. To capture brain activity, 19 electrodes were placed on the patients’ scalps according to the international 10–20 system [12] as seen in Fig. 1. Each electrode recorded electrical signals at a sampling rate of 250 Hz, resulting in 19 channels with 3,000 data points per segment. Each segment is labeled to indicate whether it contains a seizure episode or normal brain activity.

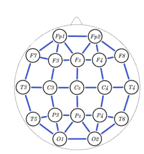


Fig. 1: Standard 10-20 montage for EEG acquisition. Blue edges represent the graph based on distances between electrodes on the scalp.

In total, the dataset comprises 12,993 EEG segments, of which 2,517 are labeled as seizure episodes. Approximately 80% of all segments represent normal brain activity. Summary statistics, such as the mean and median number of sessions per patient and segments per session, are shown in Table I.

IV. Prediction Pipeline

To predict seizure events from the raw EEG data, we developed a pipeline consisting of three main steps. First, we extracted various signal representations. Second, if Graph based models were used, we

TABLE I: Summary statistics of the data per set. Session is abbreviated with *sess.* and Patient with *pat.*

Set	#Patients	Stat	Mean	Median	Min	Max
Train	97	sess./pat.	1.82	1.00	1	7
		seg./sess.	73.41	62.00	1	285
Test	25	sess./pat.	2.00	1.00	1	7
		seg./sess.	72.28	54.00	3	470

constructed graph-structured data using multiple methods. Finally, we applied several model architectures to perform the seizure detection task.

Prior to running this pipeline, we preprocessed the raw data by excluding segments containing too many leading zeros.

A. Signal Representation

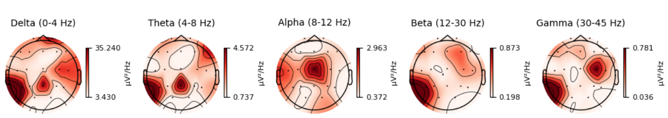
To obtain a cleaner and often lower-dimensional representation of the electrical signal at each electrode, we employed several signal representations, including the raw time-domain signal, fast Fourier transform (FFT), power spectral density (PSD), and wavelet decomposition.

1) No Processing `-`*:* As a baseline, we used the raw time-domain signal. Previous studies have shown that the time-domain signal contains valuable temporal and amplitude information that is predictive of seizure events [27].

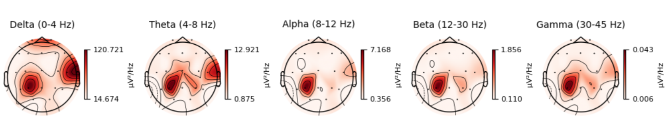
2) Bandpass Filter `BP`*:* While keeping the signal in the time domain, we cleaned the signal applying a bandpass filter to retain only signal frequencies between 0.5 and 30 Hz. Moreover, the signal is per-channel standardized.

3) Fourier Transform `FFT`*:* We used the first 354 components arising from the fast fourier transform (FFT) to represent the signal in the frequency domain. This frequency domain representation reduces the input dimension and captures potential frequency shifts associated with epilepsy and is commonly applied in the literature [27].

4) Power Spectral Density `PSD`*:* We explored the PSD representation, which estimates the power density at each frequency component based on the FFT. PSD provides a low-dimensional and noise-robust signal representation that has been widely utilized for seizure detection [27]. A visualization of band widths of PSD during normal brain activity and a seizure epsidode can be seen in Fig. 2.



(a) Normal brain activity



(b) Seizure episode

Fig. 2: Scalp topology with band widths of power spectral density (PSD) of normal brain activity and a seizure episode

5) Additional Processing Experiments: The following Signal Processing steps were additionally tested, however due to inferior performance not reported in section V.

a) Wavelet Transform: We experimented with the Wavelet Transform, which decomposes the signal into time-localized frequency components at multiple scales. This approach enables effective analysis of both short-term transients and long-term patterns by capturing detail coefficients at different resolutions. This technique has been previously shown to be effective in biomedical signal processing, including seizure detection [2], [27].

b) Fusion: Additionally we experimented with concatenation of differently processed signals.

B. Graph Construction

To construct graph-structured data from the EEG dataset, we must define nodes and a notion of connectivity between them, which determines the edge weights in the graph. The experimental design using 19-channel EEG defines the graph nodes as the individual electrode channels. We try three methods to define the edges and their weights: distance-based, correlation-based, and coherence-based.[1]

1) Distance Graph `dst`*:* First, we use the inverse of the physical distance between the 19 EEG sensors as a measure of connectivity. Applying a threshold results in sparsely connected graphs. Prior research has shown that spatial proximity between electrodes correlates with signal similarity and can improve seizure outcome prediction [24].

2) Correlation Graph `cor`*:* Second, we compute the Pearson correlation between EEG channels—either in the time or frequency domain—to define the edge weights. Thresholding is again used to sparsify the graph. Correlation captures functional relationships between brain regions and has been effectively applied in seizure detection algorithms [7], [22]. Unlike distance-based graphs, correlation-based graphs vary between data samples.

3) Coherence Graph `coh`*:* Third, we use coherence to define connectivity between EEG signals. Coherence is a frequency-domain measure capturing the linear phase relationship between signals. It reflects whether two brain regions exhibit similar neuronal oscillatory activity [4], and has also proven effective for seizure detection [7]. Like correlation-based graphs, coherence-based graphs vary across samples.

Figure 3 shows the resulting mean normalized adjacency matrices from the three methods from the train data. Interestingly, all three matrices exhibit similar connectivity patterns; however, the coherence- and correlation-based methods appear to capture more complex relationships between distant brain regions.

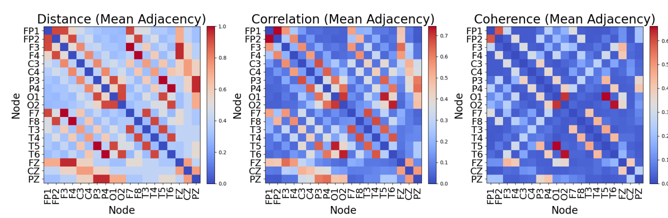


Fig. 3: Resulting mean adjacency matrices from the training set using inverse distance, correlation and coherence as measure of edge weight. Diagonals are set to zero.

C. Model Architectures

Lastly, we define a set of different time-series and graph-based machine learning architectures to perform the final seizure detection task.

Time-series architectures:

1) Residual Network `ResNet`*:* The `ResNet` architecture here employs a sequence of three residual blocks, each featuring a shortcut connection to mitigate vanishing gradients. Within each block, the input passes through three convolutional layers with batch normalization and rectified linear unit (`ReLU`) activations. The residual output is formed by adding a convolutional layer to the main path before

[1] Additionaly Graph Generative Networks have been proposed for Graph creation EGG-epilepsy seizure detection. [16] In order to keep our models simple, we only experimented with induced graph creation.

a final `ReLU`. After processing through all blocks, the network uses global average pooling to collapse temporal dimensions, followed by a single linear layer for classification. This design emphasizes identity mapping through residual connections while maintaining hierarchical feature learning.

2) Fully Convolutional Network `FCN`: The Fully Convolutional Network is a well-known architecture used for time series classification [25]. The model used in this project consists of three layers of 1D convolutions, with filter sizes varying from 128 to 256 and then back to 128. Each convolutional layer is followed by a batch normalization layer, and `ReLU` is used as the activation function. After these three convolutional layers, adaptive average pooling is applied. Binary cross-entropy loss is used as the loss function.

3) Convolutional Neural Network `CNN`: Another widely adopted architecture is the Convolutional Neural Network. `CNNs` gained prominence following the success of AlexNet in the ImageNet competition [14]. While originally developed for image classification, recent work has demonstrated that `CNN`'s can also be effectively applied to time series classification tasks [8].

The architecture employed in this project consists of four 1D convolutional layers with increasing numbers of filters (32, 32, 64, and 128, respectively). Each convolutional layer is followed by a `ReLU` activation function and a 1D max pooling operation. A dropout layer is applied after the final convolutional block to mitigate overfitting. The resulting feature maps are then passed through two fully connected layers: the first with 64 units, followed by a second with a single output unit. An additional dropout layer is applied before the final output layer.

4) Long Short Time Memory Network `LSTM`: An `LSTM` network processes sequences of data by using memory cells and gates to selectively remember or forget information, allowing it to capture patterns over long intervals [13]. The baseline `LSTM` consists of a single `LSTM` layer, followeid by a classification layer.[2].

Graph-based architectures:

5) Graph Convolutional Network `GCN`: Our Graph Convolutional Network (GCN) architecture employs a two-layer graph convolutional structure with residual connections for graph-level prediction. It begins with a `GCN` convolution layer followed by batch normalization, `LeakyReLU` activation, and dropout. The second `GCN` layer similarly applies convolution and batch normalization, then incorporates a residual connection by adding the output from the first layer before applying `LeakyReLU` and dropout again. This residual design helps maintain gradient flow. Finally, graph-level features are obtained through global mean pooling, followed by a final dropout and linear layer for prediction.

6) Long Short Term memory & GCN `LSTM-GCN`: This architecture integrates graph convolutional networks with recurrent processing for temporal graph data. It begins with two sequential `GCN` layers, each followed by `ReLU` activation, to extract spatial features from the graph structure. These layers' outputs are then stacked as a temporal sequence representing two time steps. This sequence is processed by an `LSTM` layer to capture neighboring patterns, with the final `LSTM` output step retained for each node. Node-level features are aggregated into graph-level representations via global mean pooling. Finally, a linear layer transforms these pooled features into prediction logits.

7) Long Short Term memory & GAT `LSTM-GAT`: This architecture combines `GATs` with recurrent sequence modeling for temporal graph data. It employs three consecutive `GAT` layers, each using multi-head attention with averaged outputs and followed by ReLU activation. These layers process spatial dependencies within the graph structure. The outputs from all three `GAT` layers are then stacked as a temporal sequence representing distinct time steps. This sequence is fed into an `LSTM` layer to capture time-evolving patterns, with the final `LSTM` output step retained for each node. Node-level features are aggregated into graph-level representations via global mean pooling. A final linear layer transforms these pooled features into prediction logits. The model integrates spatial attention mechanisms with temporal sequence modeling through stacked `GAT` outputs and `LSTM` processing.

[2]No Experiments were conducted using the Baseline `LSTM`, the reference score is provided via Kaggle [18]

V. Results

To ensure robust evaluation, the dataset described in Section III is further partitioned into a training and a validation set. We evaluated our models with a 5-fold stratified cross-validation to enhance the reliability of the results [19]. To leverage the available data, the model is retrained on the entire training set, and the resulting model is used to generate the final Kaggle submission [18]. Given the strong class imbalance in the data, the macro F1-score [20] is used as the primary evaluation metric, complemented by accuracy. Furthermore, we report the indicative score of relevant models on the test set achieved via the Kaggle competition [18]. In Table II we report evaluation metrics of a subset of the conducted experiments.

TABLE II: Comparison of Cross-validated Model Performance and Public Kaggle F1 Score [18] in Percentages: Non-Graph vs. Graph-Based Architectures.

Non-Graph-Based Models

Model	**Signal**	**F1**(%)	**Acc.**(%)	**F1**$_K$(%)
ResNet	FFT	87 ± 6	93 ± 4	70.35
ResNet	PSD	19 ± 2	22 ± 1	-
FCN	BP	79 ± 6	85 ± 7	-
FCN	FFT	77 ± 6	83 ± 7	-
FCN	PSD	18 ± 0	20 ± 0	-
FCN	-	75 ± 1	84 ± 1	69.62
CNN	-	71 ± 4	81 ± 3	-
CNN	BP	69 ± 2	79 ± 3	-
CNN	FFT	68 ± 9	74 ± 13	-
LSTM (Base)	-	$67, 9$	-	67.9

Graph-Based Models

Model	**Graph[th]**	**Signal**	**F1**(%)	**Acc.**(%)	**F1**$_K$ (%)
GCN	dst [0.8]	FFT	64 ± 1	72 ± 1	-
GCN	dst [0.8]	BP	62 ± 1	73 ± 0	-
GCN	dst [0.5]	BP	61 ± 0	73 ± 0	-
GCN	cor [0.5]	FFT	65 ± 1	74 ± 2	76.09
GCN	coh [0.8]	PSD	67 ± 2	75 ± 1	78.55
LSTM-GCN	coh [0.5]	FFT	65 ± 1	72 ± 1	-
LSTM-GCN	coh [0.5]	PSD	77 ± 1	84 ± 1	**83.55**
LSTM-GCN	cor [0.8]	PSD	78 ± 0	85 ± 1	80.32
LSTM-GAT	coh [0.5]	PSD	$\mathbf{80 \pm 1}$	87 ± 1	72.34

In terms of signal representation, Table II shows no clear advantage between training models in the time domain (with or without band-pass filtering) and in the frequency domain (using FFT). This suggests that the utility of frequency transformation is highly dependent on the specific model architecture. However, when focusing solely on graph-based models, power spectral density (PSD) features significantly outperform both time- and frequency-domain representations. Wavelet transform results are not reported due to consistent overfitting, despite mitigation efforts such as dropout, reduced model complexity, and early stopping. Similarly, signal fusion methods are excluded, as they did not yield performance improvements over individual signal representations.

Regarding graph construction, coherence- and correlation-based graphs perform comparably and both outperform the distance-based graph.

For time series models, we observe that `ResNet` outperforms both `FCN` and `CNN`. Among graph-based architectures, `LSTM-GAT` and LSTM−GCN achieve the best results, while the simpler GCN lags behind in performance.

VI. Discussion

A. Signal Representation

Both time and frequency domain signal representation perform well with convolutional architectures In the time domain, convolution acts to "smooth" temporal variations, while in the frequency domain (via FFT), convolution highlights relevant frequency bands by aggregating neighboring coefficients. This flexibility makes CNN-based models well-suited for a variety of input representations.

However, when using models that lack inherent temporal modeling capabilities—such as pure GCN's or feed-forward architectures—the performance with time-domain signals degrades noticeably. This is likely due to the model's inability to capture sequential dependencies, which are crucial for time-series classification tasks. In contrast, transformations like FFT may implicitly offer a form of dimensionality reduction and structure that compensates for this lack of temporal modeling, especially in shallow models. Lower-dimensional signal representation such as PSD seemed to improve model performance, likely because the tested architectures were prone to overfitting with the higher-dimensional signal representation. However, PSD signal processing did under-perform in conjunction with convolution methods.

B. Graph Construction

We observe that coherence- and correlation-based graphs consistently outperform distance-based graph methods. This suggests that sharing information between channels with similar signal patterns is more effective than relying on physical proximity alone. The advantage of coherence and correlation lies in their ability to capture dynamic functional relationships between brain regions, reflecting the synchronization of neuronal activity that occurs during seizures.

C. Model Architectures

The comparison of results between model architectures underscores the critical role of temporal and spatial modeling in EEG-based seizure detection. Among time series models, deeper convolutional architectures such as ResNet consistently outperform shallower CNNs and fully convolutional networks (FCN's), highlighting the advantage of hierarchical feature extraction in both time and frequency domains.

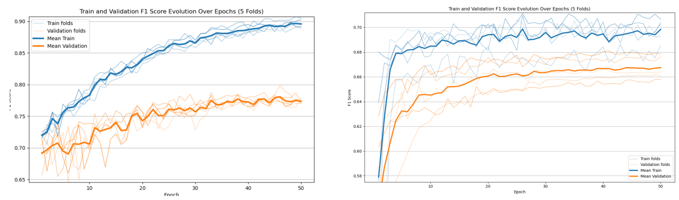
A significant performance improvement is expected when incorporating graph-based architectures that explicitly model inter-channel relationships. Pure graph convolutional models (GCN's), while incorporating spatial information through graph structures, fall short compared to both CNNs and hybrid models in cross-validation however seem to generalize well. We suspect that the lack of modeling the temporal dependencies of EEG data leads to the comparatively bad performance of purely graph-based architectures.

In contrast, hybrid architectures that combine graph neural networks with temporal modeling — such as LSTM−GCN and LSTM−GAT - achieve the best results in cross-validation. These models successfully integrate sequential signal dynamics with graph-based spatial relationships, allowing for more expressive representations of the EEG data. However due to LSTM being parameter heavy and combination increasing size of the model, it becomes more likely to overfit as seen in Figure 4, Notably, LSTM−GCN models using coherence or correlation-based graphs with PSD inputs deliver the highest macro F1-scores, both in cross-validation and on the Kaggle test set [18]. This indicates a strong synergy between frequency-domain signal representations, functional connectivity graphs, and recurrent temporal processing.

The LSTM−GAT model achieves the best cross-validation score, which may reflect the benefits of attention mechanisms in selectively weighting relevant neighbors in the graph. However, its generalization to the Kaggle test set [18] is more limited compared to LSTM−GCN, potentially due to sensitivity to noise in the sparse dataset (overfitting).

D. Inference and Generalization

Throughout the experiments, we consistently observe a noticeable gap between training and validation performance as can be seen in Figure 4. This gap is indicative of overfitting, likely exacerbated by the limited size and class imbalance of the dataset. While techniques such as dropout, reducing model complexity, and early stopping, were employed to mitigate overfitting, they only partially resolve the issue.



(a) LSTM-GCN, ∼126k params, 0.78 F1 (b) GCN, ∼500 params, 0.67 F1

Fig. 4: Evolution of Training and Validation Accuracy during Cross-validation

These regularization methods helped reduce the variance between training and validation F1-scores, but at the cost of overall classification accuracy and final Kaggle [18]performance[3]. This trade-off suggests a delicate balance between preventing overfitting and preserving the model's expressive power—particularly for complex architectures like LSTM-GNN hybrids that benefit from greater capacity but are more prone to memorization.

The evolution of training and validation curves (see Figure 4) reinforces this observation: while training accuracy steadily increases, validation accuracy plateaus or slightly declines, especially for deeper or hybrid models. This behavior highlights the necessity for larger, more diverse datasets or data augmentation strategies in future work to fully leverage the capacity of advanced architectures without sacrificing generalization. [17]

VII. Conclusion

This study shows that hybrid models combining temporal dynamics (LSTM) with graph-based spatial processing (GCN/GAT) outperform both standalone graph and non-graph approaches for EEG seizure detection. Functional connectivity-based graphs (correlation/coherence) significantly outperform distance-based ones, underscoring the value of dynamic inter-channel relationships. Signal representation effectiveness varies by model: PSD works best in graph-based settings, while time-domain signals suit temporal models. Persistent overfitting, despite regularization, points to dataset limitations and the need for more data or augmentation. Overall, these results highlight the importance of integrated temporal-spatial modeling. Future work could investigate dynamic graph construction or improved generalization for clinical applications.

VIII. Data and Code Availability

A. Code

The code implementation for this project is available online at *github.com/roduit/EE452-Network-Machine-Learning/project*.

B. Data

The data [18] for the project was retreived from *kaggle.com/competitions/epfl-network-machine-learning-2025*.

[3]We note that the indicative score did not prove representative for the evaluation of test set. While the indicative Score favored models that slightly overfit, the final performance favored lower indicative performances with better generalization

References

[1] Tala Abdallah, Nisrine Jrad, Fahed Abdallah, Anne Humeau-Heurtier, and Patrick Van Bogaert. A self-attention model for cross-subject seizure detection. *Computers in Biology and Medicine*, 165:107427, 2023. doi:10.1016/j.compbiomed.2023.107427.

[2] U Rajendra Acharya, Seong Lee Oh, Yuki Hagiwara, Jen Hong Tan, and Hojjat Adeli. Automated eeg-based screening of depression using deep convolutional neural network. *Computer Methods and Programs in Biomedicine*, 161:103–113, 2018.

[3] Ghita Amrani, Amina Adadi, and Mohammed Berrada. An attention mechanism-based interpretable model for epileptic seizure detection and localization with self-supervised pre-training. *IEEE Access*, 13:60213–60232, 2025. doi:10.1109/ACCESS.2025.3556811.

[4] Susan M Bowyer. Coherence a measure of the brain networks: past and present. *Neuropsychiatric Electrophysiology*, 2:1–12, 2016.

[5] Bibina Vadakke Chalil, Sakthivel Vellaiswamy, and Sanjay Mathilakath. A deep learning paradigm with residual networks for enhanced epileptic seizure prediction. In *AIP Conference Proceedings*, volume 3253, page 030001. AIP Publishing, 2025. doi:10.1063/5.0248386.

[6] W. Chen, Y. Wang, Y. Ren, et al. An automated detection of epileptic seizures eeg using cnn classifier based on feature fusion with high accuracy. *BMC Medical Informatics and Decision Making*, 23(1):96, 2023. doi:10.1186/s12911-023-02180-w.

[7] Alan A Díaz-Montiel and Milad Lankarany. Graph representations of ieeg data for seizure detection with graph neural networks. *bioRxiv*, pages 2023–06, 2023.

[8] Hassan Ismail Fawaz, Germain Forestier, Jonathan Weber, Lhassane Idoumghar, and Pierre-Alain Muller. Deep learning for time series classification: a review. *Data Mining and Knowledge Discovery*, 33(4):917–963, 2019. doi:10.1007/s10618-019-00619-1.

[9] Robert S. Fisher, Walter van Emde Boas, Warren Blume, Christian Elger, Pierre Genton, Phillip Lee, and Jerome Engel Jr. Epileptic seizures and epilepsy: Definitions proposed by the international league against epilepsy (ilae) and the international bureau for epilepsy (ibe). *Epilepsia*, 46(4):470–472, 2005. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.0013-9580.2005.66104.x, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.0013-9580.2005.66104.x, doi:10.1111/j.0013-9580.2005.66104.x.

[10] William L. Hamilton. *Graph Representation Learning*, volume 14 of *Synthesis Lectures on Artificial Intelligence and Machine Learning*. Morgan & Claypool Publishers, 2020. doi:10.2200/S00980ED1V01Y202009AIM046.

[11] Jiatong He, Jia Cui, Gaobo Zhang, Mingrui Xue, Dengyu Chu, and Yanna Zhao. Spatial–temporal seizure detection with graph attention network and bi-directional lstm architecture. *Biomedical Signal Processing and Control*, 78:103908, 2022. URL: https://www.sciencedirect.com/science/article/pii/S1746809422004098, doi:10.1016/j.bspc.2022.103908.

[12] Herbert H. Jasper. The ten-twenty electrode system of the international federation. *Electroencephalography and Clinical Neurophysiology*, 10:371–375, 1958.

[13] S.U. Khan, S.U. Jan, and I. Koo. Robust epileptic seizure detection using long short-term memory and feature fusion of compressed time-frequency eeg images. *Sensors*, 23(23):9572, 2023. doi:10.3390/s23239572.

[14] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. *Commun. ACM*, 60(6):84–90, May 2017. doi:10.1145/3065386.

[15] Zhejun Kuang, Simin Liu, Jian Zhao, Liu Wang, and Yunkai Li. Epilepsy eeg seizure prediction based on the combination of graph convolutional neural network combined with long- and short-term memory cell network. *Applied Sciences*, 14(24), 2024. URL: https://www.mdpi.com/2076-3417/14/24/11569, doi:10.3390/app142411569.

[16] Z. Li, K. Hwang, K. Li, and et al. Graph-generative neural network for eeg-based epileptic seizure detection via discovery of dynamic brain functional connectivity. *Scientific Reports*, 12:18998, 2022. doi:10.1038/s41598-022-23656-1.

[17] Eneko Lopez, Jaione Etxebarria-Elezgarai, Jose Manuel Amigo, and Andreas Seifert. The importance of choosing a proper validation strategy in predictive models. a tutorial with real examples. *Analytica Chimica Acta*, 1275:341532, September 2023. Under a Creative Commons license. doi:10.1016/j.aca.2023.341532.

[18] LTS4. Epfl network machine learning 2025. https://kaggle.com/competitions/epfl-network-machine-learning-2025, 2025. Kaggle.

[19] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[20] PyTorch Team. torcheval.metrics.functional.multiclass_f1_score. https://docs.pytorch.org/torcheval/stable/generated/torcheval.metrics.functional.multiclass_f1_score.html, 2023. Accessed: 2025-06-09.

[21] Vinit Shah, Eva Von Weltin, Silvia Lopez, James Riley McHugh, Lillian Veloso, Meysam Golmohammadi, Iyad Obeid, and Joseph Picone. The temple university hospital seizure detection corpus. *Frontiers in neuroinformatics*, 12:83, 2018.

[22] Chongrui Tian and Fengbin Zhang. Eeg-based epilepsy detection with graph correlation analysis. *Frontiers in Medicine*, 12:1549491, 2025.

[23] X. Wang, Y. Wang, D. Liu, et al. Automated recognition of epilepsy from eeg signals using a combining space–time algorithm of cnn-lstm. *Scientific Reports*, 13:14876, 2023. doi:10.1038/s41598-023-41537-z.

[24] Yujiang Wang, Nishant Sinha, Gabrielle M Schroeder, Sriharsha Ramaraju, Andrew W McEvoy, Anna Miserocchi, Jane de Tisi, Fahmida A Chowdhury, Beate Diehl, John S Duncan, et al. Interictal intracranial electroencephalography for predicting surgical success: The importance of space and time. *Epilepsia*, 61(7):1417–1426, 2020.

[25] Zhiguang Wang, Weizhong Yan, and Tim Oates. Time series classification from scratch with deep neural networks: A strong baseline, 2016. URL: https://arxiv.org/abs/1611.06455, arXiv:1611.06455.

[26] World Health Organization. Epilepsy: a public health imperative. summary, 2019. WHO/MSD/MER/19.2. Licence: CC BY-NC-SA 3.0 IGO. URL: https://apps.who.int/iris/handle/10665/325300.

[27] Hui Xiong, Yan Yan, Yimei Chen, and Jinzhen Liu. Graph convolution network-based eeg signal analysis: a review. *Medical & Biological Engineering & Computing*, pages 1–17, 2025.

[28] Gaowei Xu, Tianhe Ren, Yu Chen, and Wenliang Che. A one-dimensional cnn-lstm model for epileptic seizure recognition using eeg signal analysis. *Frontiers in Neuroscience*, Volume 14 - 2020, 2020. URL: https://www.frontiersin.org/journals/neuroscience/articles/10.3389/fnins.2020.578126, doi:10.3389/fnins.2020.578126.

[29] Jie Xu, Shasha Yuan, Junliang Shang, Juan Wang, Kuiting Yan, and Yankai Yang. Spatiotemporal network based on gcn and bigru for seizure detection. *IEEE Journal of Biomedical and Health Informatics*, 28(4):2037–2046, 2024.

[30] Ye Yuan, Guangxu Xun, Kebin Jia, and Aidong Zhang. A multi-view deep learning framework for eeg seizure detection. *IEEE journal of biomedical and health informatics*, 23(1):83–94, 2018.

[31] Yanna Zhao, Gaobo Zhang, Changxu Dong, Qi Yuan, Fangzhou Xu, and Yuanjie Zheng. Graph attention network with focal loss for seizure detection on electroencephalography signals. *International journal of neural systems*, 31(07):2150027, 2021.