Graph-based EEG Analysis

Eylül İpçi Filip Mikovíny Tuğba Tümer Yağız Gençer

***Abstract*—Epileptic seizure detection from EEG signals is essential for timely intervention and better patient outcomes. In this project, we explore graph-based models for classifying EEG time-series data using the Temple University Hospital Seizure Corpus (TUSZ), which contains multi-channel recordings from 19 electrodes (10–20 system). Each recording is segmented into 12-second windows labeled as seizure or non-seizure.**

As a baseline, we implement a spectrogram-based Convolutional Neural Network (CNN). We then develop two graph-based models, a Message Passing Neural Network (MPNN) and a Graph Transformer, that represent electrodes as nodes with edge weights based on inter-electrode distances. We detail the preprocessing, graph construction, and model architectures, and present a comparative analysis of their classification performance and design trade-offs.

***Index Terms*—EEG, seizure detection, graph neural networks, MPNN, graph transformer, deep learning, time-series classification, brain connectivity, spectrogram-based CNN**

I. Introduction

Epilepsy is a chronic neurological disorder affecting around 50 million people worldwide, making it one of the most common brain conditions [1]. It is marked by recurrent seizures, brief episodes of abnormal or excessive neuronal activity, that can severely impact quality of life [3]. Accurate and timely seizure detection is crucial, especially for patients who do not respond to medication.

Electroencephalography (EEG) is the primary non-invasive tool for monitoring brain activity. It records voltage fluctuations from scalp electrodes, producing multivariate time-series data. However, EEG signals are difficult to model due to their high dimensionality, low signal-to-noise ratio, and non-stationary nature.

We use a subset of the Temple University Hospital Seizure Corpus (TUSZ) [2], a large public EEG dataset. It includes recordings from 97 patients for training and 25 for testing, sampled at 250 Hz using 19 electrodes placed according to the 10–20 system. Recordings are segmented into 12-second non-overlapping windows labeled as seizure or non-seizure, producing a 19×3000 matrix per window.

Traditional deep learning models such as RNNs or transformers often treat channels independently, overlooking spatial relationships between electrodes and missing important connectivity patterns. Graph-based approaches address this by modeling EEG as a graph, where nodes are electrodes and edges reflect spatial or functional relationships. Graph Neural Networks (GNNs) can learn interactions across brain regions, improving generalization and interpretability for seizure detection.

In this project, we explore graph-based learning for EEG seizure classification and compare it with non-graph baselines. Our contributions are:

- A baseline spectrogram-based Convolutional Neural Network (CNN).
- Spatial EEG graph construction using inter-electrode distances.
- Implementation and evaluation of a Message Passing Neural Network (MPNN) and a Graph Transformer.
- Comparative analysis of classification performance and interpretability.

This study highlights the promise of geometric deep learning for EEG analysis and the advantages of incorporating spatial structure in seizure detection.

II. Spectrogram-Based CNN Baseline

A. Spectrogram Extraction

EEG data is naturally suited for time–frequency analysis due to the non-stationary nature of brain activity. Seizures often appear as transient changes in both time and frequency, making spectrograms a valuable representation for classification. We convert raw EEG time-series into spectrograms, treating each electrode as a separate channel, allowing us to retain both temporal and spectral characteristics in a compact 2D format.

A spectrogram shows how a signal's frequency content evolves over time using the Short-Time Fourier Transform (STFT) applied over sliding windows. This produces a 2D array where one axis represents time and the other frequency, with each value indicating signal power at a specific time–frequency point. This representation preserves both time and frequency localization, enabling the model to capture patterns that unfold over time and across frequency bands.

For each 12-second EEG window, we compute a spectrogram independently for all 19 electrodes. With data sampled at 250 Hz, we use 128-sample segments and 64-sample overlap, yielding a frequency resolution of 1.95 Hz and time resolution of 0.256 seconds. The overlap helps smooth transitions and reduce edge artifacts, leading to more stable spectrograms. We retain only frequencies from 0.5 to 30 Hz, which are most relevant for seizure detection, and apply a logarithmic transform to compress the dynamic range.

The result is a [19, F, T] tensor, where 19 denotes the EEG channels and F and T are the frequency and time bins. This serves as input to the CNN, which learns spatial patterns across electrodes and localized time–frequency features relevant to seizure classification.

B. Model Architecture

We designed a lightweight convolutional neural network (CNN) to classify EEG spectrograms. Each spectrogram is treated as a multi-channel image, with one channel per EEG electrode, enabling the model to learn spatial and spectral patterns across electrodes, frequencies, and time.

The network begins with a convolutional layer that expands the channels from 19 to 32, followed by ReLU activation and max pooling to reduce time–frequency resolution while retaining key features. The next two blocks increase the channels to 64 and then 128, gradually enriching the representation with higher-level features.

Instead of collapsing the feature map with global pooling, we apply an attention mechanism to highlight the most informative time–frequency regions. After flattening the spatial dimensions, a 1×1 convolution followed by softmax produces attention weights over all locations. These weights reweight the features, allowing each channel to focus on spectrogram regions most indicative of seizures. The resulting vector is passed to a final linear layer for classification.

The model is trained with binary cross-entropy loss and optimized using Adam. Its compact design helps prevent overfitting while leveraging attention to enhance performance.

III. EEG Graph Construction

To capture spatial relationships among EEG electrodes, we construct a graph where each of the 19 nodes corresponds to an electrode positioned according to the 10–20 international system [4]. The connections (edges) are determined using inter-electrode distances retrieved from the provided `distances.csv` file.

We transform these distances into edge weights by applying a softmax function to the squared distances scaled by a parameter β. This approach emphasizes nearby electrodes by assigning them higher weights, while more distant ones receive lower importance. After experimenting with several values, we found that $\beta = 5$ offers a good balance between local connectivity and graph sparsity. For each node, we retain only the strongest edges that cumulatively represent 90% of the total softmax weight. To ensure bidirectional information flow, we add reverse edges, resulting in an undirected graph.

This spatial graph forms the structural foundation for both our Message Passing Neural Network (MPNN) and Graph Transformer models, enabling them to leverage the brain's spatial layout. Figure 1 illustrates the resulting graph, with nodes representing electrodes and edges reflecting the most relevant spatial connections after thresholding.

IV. EEG Signal Preprocessing for Graph Models

To feed meaningful inputs into our graph-based models, we transform raw EEG time-series into fixed-size feature vectors for each electrode (i.e., graph node). Each 12-second window contains a 3000×19 matrix (time steps $\times$ channels), sampled at 250 Hz. Instead of using raw signals directly, we extract descriptive statistical and spectral features that summarize each channel's activity while preserving discriminative power for seizure classification.

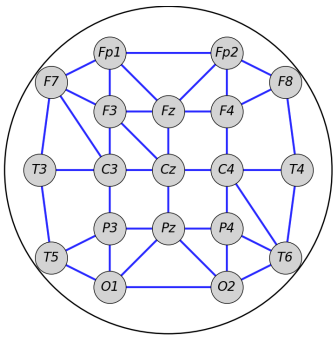

Fig. 1. EEG electrode graph constructed using softmax-weighted inter-electrode distances ($\beta = 5$) and a 90% cumulative threshold.

A. Handcrafted Feature Extraction [5]

We compute 18 handcrafted features per electrode, combining temporal, spectral, and complexity-based descriptors that capture the statistical structure of EEG signals without discarding key information.

- **Mean Absolute Differences:** First- and second-order mean absolute differences quantify signal variability and abrupt changes.
- **Statistical Moments:** Standard deviation, skewness, and kurtosis describe the distribution's shape.
- **Band Power and Relative Band Power:** Using FFT, we compute absolute and relative power in four frequency bands: delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), and beta (13–30 Hz), commonly associated with brain rhythms.
- **Hjorth Parameters:** *Mobility* estimates how rapidly the signal changes, while *complexity* measures how much rapidly mobility itself changes.
- **Zero Crossings:** Count of sign changes, indicative of high-frequency components.
- **Shannon Entropy:** Measures unpredictability from the signal amplitude histogram.
- **Spectral Entropy:** Computed from normalized power spectral density, reflecting the flatness or order in the frequency spectrum.

These features are concatenated into an 18-dimensional vector per electrode per window.

B. Feature Variants for Graph Models

We tested three node feature variants:

1) **Baseline FFT-based Features (354 features):** Precomputed from bandpass-filtered FFT coefficients (0.5–30 Hz), as in the provided example code.
2) **Handcrafted Features Only (18 features):** Our lightweight set emphasizes interpretability and efficiency.
3) **Combined Features (372 features):** Concatenation of the above two for maximum descriptive power.

V. Message Passing Neural Network (MPNN)

To capture spatial dependencies between electrodes, we implement an MPNN operating on the EEG graph. Each node represents an electrode and is initialized with a feature vector extracted from its 12-second window.

Each node updates its representation by receiving messages from its neighboring nodes. To generate these messages, each neighbor passes its embedding through an MLP. To determine how much influence each neighbor should have, we concatenate the neighbor's embedding with the target node's embedding and feed this pair into a separate MLP that outputs an attention score. These scores are normalized using a softmax across all neighbors. Each message is then scaled by its corresponding weight, and the node aggregates all weighted messages into a single vector. The node's original embedding is then concatenated with the aggregated message and passed through another MLP, followed by layer normalization and a ReLu activation, producing the updated node embedding. This process repeats for three message passing layers, enabling nodes to incorporate information from multi-hop neighbors.

After the final layer, we apply global mean pooling to combine all node embeddings into a single graph-level vector, summarizing brain activity in the window. This vector is passed through a linear layer to predict seizure presence. The model is trained using binary cross-entropy loss and optimized with Adam. This design enables adaptive spatial reasoning over EEG signals for effective seizure detection.

VI. Graph Transformer

To capture both spatial and higher-order interactions among electrodes, we implement a Graph Transformer that jointly updates node, edge, and global representations. Each node still begins with its feature vector, and each undirected edge carries a learned scalar weight reflecting inter-electrode proximity. We also maintain a small global feature vector summarizing the entire graph, which is initially set to 0. At each of the three Transformer layers, we first project every node's embedding into queries, keys, and values (via learnable linear maps), split across multiple attention heads. For each node pair, we compute a raw compatibility by taking element-wise products of their per-head query/key slices, normalized by the head dimension. We then modulate those scores with the edge features via FiLM: we linearly transform the edge embedding into a per-head "scale" and "shift", apply (1 + scale) × compatibility + shift (1+scale)×compatibility+shift, and mask out any padded or self-loops to respect graph sparsity and undirectedness. These modulated scores go through a masked softmax over neighbors to produce attention weights. We multiply those weights against each node's per-head value vectors, sum over neighbors, and recombine heads into a single updated node tensor. In parallel, the same FiLM mechanism (this time modulated by the global feature) refines each edge embedding, and two small aggregation MLPs (again via FiLM from the global) produce an updated global vector that pools both node- and edge-level statistics. Each of these streams—nodes, edges, and global—receives a residual connection + layer normalization, followed by a shared feed-forward sublayer (two linear layers with ReLU) and another residual+Norm. After stacking all layers, we optionally project back to any desired output dimensions: in our seizure-detection setup we skip updating nodes and edges (so they remain fixed), and only map the final global vector down to a single logit via a tiny MLP. This design—attention scores enriched by edges, global conditioning via FiLM, and tightly coupled residual updates, lets each electrode representation adaptively "attend" to relevant neighbors while maintaining a unified graph-level summary for classification.

VII. Experimental Results

We benchmarked our Graph Transformer and Message Passing Neural Network under a controlled set of hyperparameters across three distinct node-feature preprocessing pipelines (see Section IV). For all experiments involving the Graph Transformer, we held the architecture and optimization settings constant: 3 Transformer layers, 8 attention heads per layer, a learning rate of 10^{-4}, a batch size of 256, training for 500 epochs, and a decision threshold of 0.25 on the sigmoid outputs. This uniform setup isolates the impact of the input feature representation on validation F_1 performance. For MPNN, we used 3 message passing layers, batch size of 64 and a learning rate of 10^{-4}, training for 200 epochs, and a decision threshold of 0.50 on the sigmoid outputs. These hyperparameter choices are done by observing traning loss curves and finding a balance between best validation F1 performance and training stability.

To ensure generalizability and avoid data leakage, we adopted a patient-wise split where no patient appeared in both training and validation sets. The resulting splits preserved the dataset's class distribution while maintaining patient independence:

- Training set: 76 patients with a total of 10277 samples
- Validation set: 21 patients with a total of 2716 samples

For the two graph-based models, Graph transformer and MPNN, we listed the validation F1 scores for each of the preprocessing methods in Table I. LSTM (as taken from the example notebook provided alongside with the project description) and CNN models are not listed since the preprocessing considered in this table do not apply for them.

TABLE I
Validation F_1 scores across preprocessing pipelines

Preprocessing Pipeline	**Graph Transformer F_1**	**MPNN F_1**
Baseline FFT-based Features	0.647	0.669
Handcrafted Features Only	0.533	0.638
Combined Features	0.622	0.707

Table II also lists the best validation F1 scores we could get with each of the methods, including graph or non-graph models.

According to Table II, we see that CNN gives the best validation F1 with 0.756. For that reason, we also provide a

TABLE II
BEST VALIDATION F_1 SCORES ACROSS DIFFERENT MODELS

Model	Best F_1 Score
LSTM	0.642
CNN	0.756
MPNN	0.707
Graph Transformer	0.647

typical training curve of our spectrogram-based CNN model in Figure 2, which shows the evolution of training loss, validation loss and validation F1 across all epochs.

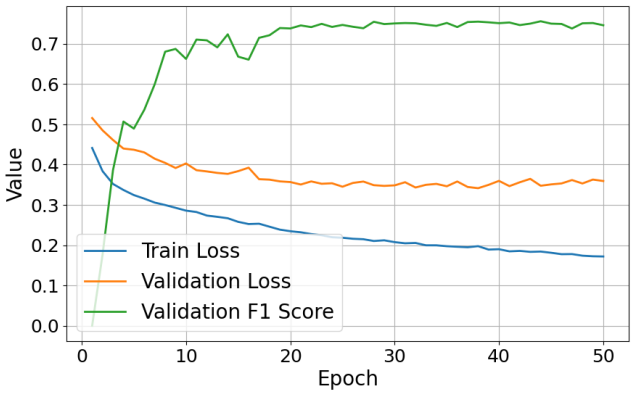

Fig. 2. CNN Loss Plot

VIII. Discussion of Results

Our experimental findings revealed several important insights and challenges during model development and evaluation.

First, despite our careful design of the spatial EEG graph using inter-electrode distances (Section III), graph-based models did not yield a clear performance advantage. Both the MPNN and Graph Transformer models underperformed compared to the CNN baseline. This suggests that encoding spatial proximity alone may not adequately capture the complex relationships between EEG signals across electrodes. Our literature review confirmed that no consistent or widely accepted graph structure exists for seizure detection using the 10–20 system, further highlighting the difficulty of finding a meaningful graph prior. Additionally, our assumption that smaller inter-electrode distances imply stronger signal similarity may not always hold, especially given the functional rather than purely anatomical nature of brain connectivity.

Second, we found that validation performance was highly sensitive to the specific patient split. Since our data was split such that no patient appeared in both training and validation sets, some splits led to better generalization than others. This likely reflects the fact that seizure patterns vary significantly between individuals, some patients' data are more informative or typical, while others may be outliers. As a result, a model trained on less diverse or less representative patients may struggle to generalize to unseen ones.

Third, label imbalance posed a consistent challenge. With only about 20% of the windows labeled as seizures, models had a tendency to favor the majority class. We tried weighted sampling and weighted loss functions to compensate, but neither improved generalization. Oversampling led to overfitting, while weighting the loss introduced optimization instability. Instead, we found that adjusting the sigmoid decision threshold (e.g., lowering it to 0.25 for some models) provided a simple yet effective way to address class imbalance during evaluation.

Table I shows that handcrafted features alone were insufficient for strong performance. While interpretable and computationally efficient, they lack the richness needed to capture complex seizure dynamics. In contrast, the FFT-based features and their combination with handcrafted features led to better performance, particularly with the MPNN model, which achieved its best F_1 score (0.707) using the combined feature set.

Among all models, the spectrogram-based CNN consistently achieved the highest validation F_1 score (0.756). This result aligns with our expectations, as spectrograms provide a dense and informative time–frequency representation well-suited to seizure detection. The CNN was also the fastest to train and required the fewest epochs to converge, showing a good trade-off between performance and efficiency. In contrast, more complex models like the Graph Transformer and LSTM required significantly longer training times and yielded lower validation performance, likely due to overfitting or their higher capacity being unnecessary for this relatively small dataset.

Overall, we observed that simpler models with strong inductive biases, such as the CNN, generalized better and trained more efficiently than more expressive but data-hungry architectures. This trend highlights the importance of model–data alignment, especially in medical applications with limited labeled data.

IX. Conclusion

In this project, we explored EEG-based seizure detection using both traditional and graph-based deep learning approaches. While graph models offer a promising way to incorporate spatial structure into learning, our experiments showed that they did not outperform a spectrogram-based CNN in terms of validation F_1 score, training stability, or efficiency.

The CNN, leveraging time–frequency information from spectrograms, achieved the best overall performance and required significantly less training time. Our results suggest that for relatively small datasets like TUSZ and binary classification tasks such as seizure detection, simpler models with well-chosen input representations remain highly effective.

Nonetheless, graph-based models may still hold value for more complex settings or when better graph priors, such as those based on functional connectivity, are available. Future work could explore dynamic graph construction, patient-specific graph learning, or multimodal data fusion to further unlock the potential of geometric deep learning in EEG analysis.

REFERENCES

[1] World Health Organization, *Epilepsy: A public health imperative*, Tech. Rep. WHO/MSD/MER/19.2, World Health Organization, 2019.

[2] V. Shah, E. von Weltin, S. Lopez, J. R. McHugh, L. Veloso, M. Golmohammadi, I. Obeid, and J. Picone, "The Temple University Hospital Seizure Detection Corpus," *Frontiers in Neuroinformatics*, vol. 12, Nov. 2018. [Online]. Available: https://doi.org/10.3389/fninf.2018.00083

[3] R. S. Fisher, W. van Emde Boas, W. Blume, C. Elger, P. Genton, P. Lee, and J. Engel Jr., "Epileptic seizures and epilepsy: Definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE)," *Epilepsia*, vol. 46, no. 4, pp. 470–472, 2005.

[4] Wikipedia, "10–20 system (EEG)," *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/w/index.php?title=10%E2%80%9320%20system%20(EEG)&oldid=1234311940, accessed June 9, 2025.

[5] K. Makkar, A. Bisen, and Editor Ijmtst, "EEG signal processing and feature extraction," *International Journal for Modern Trends in Science and Technology*, vol. 9, pp. 45–50, Oct. 2023. DOI: 10.46501/IJMTST0908008.