Graph Attention Networks for Atrial Fibrillation Detection: Leveraging Inter-Lead Dependencies and Geographic Variations in ECG Signals

Rany Khirbawi
Technion - Israel Institute of Technology
ranykhirbawi@campus.technion.ac.il

Abstract

Atrial fibrillation (AF) is the most prevalent cardiac arrhythmia and a significant contributor to stroke, heart failure, and increased mortality rates globally. Early and accurate detection of AF is critical for timely medical intervention and improved patient outcomes. Electrocardiogram (ECG) data, which provides a non-invasive method for cardiac monitoring, holds vital spatial and temporal information. However, traditional machine learning models often overlook spatial dependencies between ECG leads, which encode crucial physiological relationships reflective of cardiac activity. As such, leveraging the micro-environmental spatial information embedded within ECG leads remains an underexplored avenue.

My approach extends foundational work, such as "Arrhythmia Detection by the Graph Convolution Network and a Proposed Structure for Communication Between Cardiac Leads" and "Graph Neural Networks for Topological Feature Extraction in ECG Classification." These studies highlighted the importance of graph-based representations and effective strategies for integrating ECG features into graph neural network (GNN) architectures. Building on these insights, I propose a novel framework utilizing a Graph Attention Network (GAT) with five layers and four attention heads, explicitly designed to capture inter-lead dependencies while dynamically weighting the importance of relationships using multi-head attention.

The research was conducted under constrained computational resources, necessitating the training of models locally. This motivated the iterative exploration of multiple architectures, including Random Forests, Support Vector Machines (SVMs), and simpler GAT variants (e.g., two and three-layer models). The models were trained and evaluated on datasets from the PhysioNet Challenge 2021, which included Chapman-Shaoxing (ECG data from China representing a Chinese population), PTB-XL (ECG data from Germany representing a European population), and Georgia (ECG data from the U.S., representing a Western population). Each dataset contained 12-lead ECG recordings, diagnostic labels for cardiac events like atrial fibrillation, and explicit geographic labels that i added to the preprocessed data.

The Graph Attention Network (GAT) model with five attention layers and two attention heads achieved an accuracy of 0.88. Similarly, the GAT model with five attention layers and eight attention heads also achieved an accuracy of 0.88. Despite these limitations, the proposed GAT model achieved remarkable performance metrics: an accuracy of 94.8%, a recall of 88.57%, and an ROC-AUC score of 97.28%, outperforming baseline models.

To optimize the model, early experiments investigated edge construction techniques, such as graph-based representations of signal similarities or transitions, Edge detection algorithms inspired by [6], such as Sobel and Canny, were applied to raw ECG signals to explore whether sharp transitions or local signal changes could better encode physiological inter-lead relationships. While promising, these methods were superseded by the use of Weighted Mutual Information (WMI), which provided a more robust basis for graph construction.

The proposed approach integrates both statistical and clinical features into graph nodes and utilizes the GAT's multi-head attention mechanism to dynamically prioritize critical inter-lead dependencies. This work highlights the potential of graph-based models in AF detection and paves the way for more robust representations of ECG data in arrhythmia classification tasks.

In addition, in accordance with the secondary hypothesis, the influence of geographic origin on ECG signals was rigorously investigated to validate the hypothesis that spatial characteristics of data vary across regions. Using datasets from China, Germany, and the United States, the proposed GAT model was trained on data from one region and tested on another, revealing significant accuracy degradation in cross-region evaluations. Statistical tests, including the Kolmogorov-Smirnov (K-S) Test and Mann-Whitney U Test, confirmed substantial differences in embedding distributions across regions, indicating region-specific dependencies in ECG signal relationships. Additionally, t-SNE visualizations demonstrated distinct clustering patterns based on geographic data, reinforcing these findings. To address these challenges, region-specific weight adjustments were implemented, which improved performance in challenging cross-region scenarios. This comprehensive analysis underscores the necessity of region-aware adaptations in graph-based models for ECG classification tasks and for machine learning models in general.

1 Introduction

Cardiovascular diseases remain a leading cause of global mortality, with atrial fibrillation (AF) being the most prevalent and clinically significant cardiac arrhythmia. AF is associated with a markedly increased risk of stroke, heart failure, and mortality, making its early detection critical for timely intervention and effective management. Electrocardiograms (ECGs), consisting of 12-lead recordings, are the primary diagnostic tool for identifying AF. These leads capture the heart's electrical activity from various spatial perspectives, providing valuable diagnostic insights.

Despite the widespread use of ECGs, existing computational methods face two major limitations. First, traditional models such as Random Forest (RF) and Support Vector Machines (SVM) fail to capture the spatial dependencies inherent in ECG data, as these methods treat leads independently, ignoring the relationships introduced by electrode placement. Second, predictive models typically rely on raw ECG signals or simple features like QRS duration, overlooking the rich statistical and clinical information embedded in the signals.

Addressing these challenges requires innovative approaches that incorporate spatial, statistical, and clinical insights. Recent advancements, such as [5], have demonstrated the potential of graph-based models in this domain. In this study, Weighted Mutual

Information (WMI) was employed to define inter-lead relationships, allowing Graph Convolutional Networks (GCNs) to leverage these spatial dependencies. However, GCNs are limited by their inability to dynamically prioritize inter-lead relationships, which led to the exploration of Graph Attention Networks (GATs).

This research builds on these advancements by utilizing a GAT-based framework to dynamically assign importance to inter-lead relationships. In this approach, leads are represented as graph nodes, and their dependencies are quantified as weighted edges using WMI. WMI accounts for both linear and non-linear dependencies between time-series signals, ensuring robust modeling of inter-lead relationships.

1.1 Past Studies

Over the years, several studies have employed advanced deep learning techniques for arrhythmia detection. [1] proposed a Deep Neural Network (DNN) to classify ECG rhythms, demonstrating the potential of neural networks in the domain. Shaker et al. introduced a Generative Adversarial Network (GAN) for rhythm classification [2], leveraging generative capabilities to improve detection accuracy. Similarly, Yao et al. developed an ATI-CNN model that utilized multi-channel ECG data to classify eight arrhythmia types [3], highlighting the advantages of leveraging spatial dependencies. Zhou & Tan extended this work by combining CNNs with Extreme Learning Machines (ELM) [4], emphasizing CNNs' ability to simultaneously perform feature extraction and classification.

Building on these advancements, the proposed framework leverages Weighted Mutual Information (WMI) to address spatial relationships, statistical dependencies, and geographic variations. By employing datasets from Chapman-Shaoxing (China), PTB-XL (Germany), and Georgia (United States), this study investigates the geographic influence on ECG data and develops region-specific models to enhance accuracy and robustness.

1.2 Weighted Mutual Information for ECG Graphs

In this study, the relationship between ECG leads is represented using a WMI-based adjacency matrix, which serves as the foundation for the graph representation of ECG data. ECG signals are time series, with each lead capturing distinct aspects of cardiac electrical activity. The goal of WMI is to measure the shared information between these time-series signals while accounting for both linear and non-linear dependencies.

WMI is particularly well-suited for ECG data as it avoids assumptions of linearity and independence that limit traditional correlation measures like Pearson or Spearman coefficients. By leveraging the concept of mutual information, WMI calculates the shared entropy between two time-series signals, capturing their statistical dependency in a robust manner.

Matrix Structure

A 12×12 WMI matrix is constructed for each ECG recording, corresponding to the 12-lead ECG setup. Each matrix element represents the mutual information between a pair of leads, where higher values indicate stronger dependency. To enhance interpretability and reflect physiological groupings, weights are applied to the matrix based on lead categories:

- **Precordial Leads:** Includes V1 through V6, which measure electrical activity across the chest. These leads emphasize local cardiac regions, and their relationships capture spatially localized activity.

- **Limb Leads:** Contains 6 leads divided into:
  - **Bipolar Leads (I, II, III):** Measure potential differences between two specific limb electrodes.
  - **Unipolar Leads (aVR, aVL, aVF):** Record potential differences relative to a virtual central point. Limb leads reflect global cardiac activity and provide a holistic perspective on cardiac function.

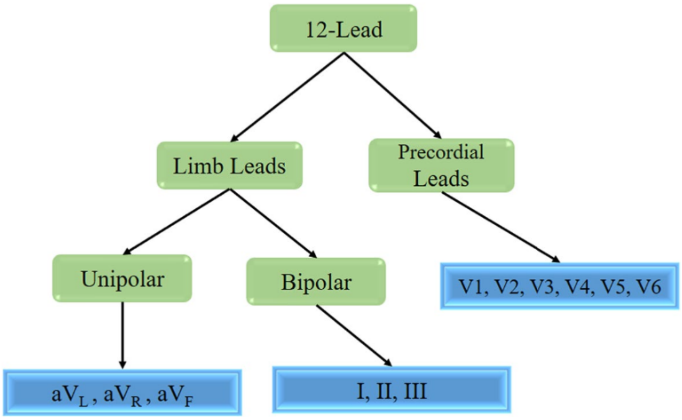


Figure 1: Clustering of heart leads based on the ECG pole.

Significance of Weighted Relationships

The weighting in WMI emphasizes interdependencies between specific lead categories. For instance:

- **Precordial vs Precordial:** Strong weights capture localized spatial relationships.
- **Precordial vs Limb:** Moderate weights reflect the transition between localized and global activity.
- **Limb vs Limb:** Relatively lower weights account for broad, systemic dependencies.

By incorporating these weights, the WMI matrix ensures that critical inter-lead relationships are preserved while minimizing noise from less significant connections. This weighted structure serves as the adjacency matrix for the Graph Attention Network (GAT), enabling the model to dynamically learn the importance of relationships between leads during training.

1.3 Graph Representation

In the graph representation:

- **Nodes:** Represent the 12 ECG leads, each enriched with clinical and statistical features extracted from the raw signal.

- **Edges:** Are defined by the WMI matrix, with weights determining the strength of inter-lead connections.

This approach bridges the gap between raw signal characteristics and graph-based representations, leveraging the spatial and statistical dependencies intrinsic to ECG data. The adjacency matrix enables the GAT to capture both local and global relationships, enhancing the model's ability to detect atrial fibrillation.

2 Methods and Background

Graphs are versatile data structures that can effectively model relationships between entities, consisting of nodes (vertices) and edges (connections). In various real-world problems, such as social networks, molecular interactions, and electrocardiogram (ECG) signal analysis, data can naturally be represented as graphs. Graph Neural Networks (GNNs) have emerged as powerful tools for learning from graph-structured data, leveraging the graph topology and node-specific features to model dependencies and relationships effectively.

2.1 Graph Neural Networks and Convolutions

GNNs propagate and aggregate information from neighboring nodes, enabling the representation of each node to be enriched by its local graph structure. Traditional GNNs update the feature vector of a node i by aggregating features from its neighbors $\mathcal{N}(i)$, often weighted by graph-based properties such as edge attributes.

Graph Convolutional Networks (GCNs) extend the convolutional operations from grid-like data (e.g., images) to irregular graph structures. In GCNs, the representation of node i at layer $l+1$ is updated as:

$$\mathbf{h}_i^{(l+1)} = \sigma\left(\sum_{j\in\mathcal{N}(i)} \frac{1}{\sqrt{\deg(i)\cdot\deg(j)}}\mathbf{W}^{(l)}\mathbf{h}_j^{(l)}\right),$$

where $\mathbf{h}_i^{(l)}$ represents the feature vector of node i at layer l, $\mathbf{W}^{(l)}$ is a learnable weight matrix, σ is an activation function (e.g., ReLU), and $\deg(i)$ is the degree of node i. While GCNs effectively propagate information across graph structures, they treat all neighboring nodes equally, limiting their ability to model diverse and dynamic relationships.

2.2 Graph Attention Networks (GATs)

Graph Attention Networks (GATs) overcome the limitations of GCNs by incorporating an attention mechanism to assign importance dynamically to neighboring nodes. In GATs, the relationship between nodes i and j is weighted by an attention coefficient α_{ij}, computed as:

$$\alpha_{ij} = \frac{\exp\left(\text{LeakyReLU}\left(\mathbf{a}^\top\left[\mathbf{W}\mathbf{h}_i \| \mathbf{W}\mathbf{h}_j\right]\right)\right)}{\sum_{k\in\mathcal{N}(i)}\exp\left(\text{LeakyReLU}\left(\mathbf{a}^\top\left[\mathbf{W}\mathbf{h}_i \| \mathbf{W}\mathbf{h}_k\right]\right)\right)},$$

where $\mathbf{W}$ is a learnable weight matrix, $\mathbf{a}$ is a learnable attention vector, $\|$ denotes concatenation, and LeakyReLU is the activation function. The updated representation for node i is:

$$\mathbf{h}_i^{(l+1)} = \sigma\left(\sum_{j\in\mathcal{N}(i)} \alpha_{ij}\mathbf{W}\mathbf{h}_j^{(l)}\right).$$

The use of multi-head attention in GATs enables the model to capture diverse patterns in node relationships, making it suitable for analyzing the complex dependencies in ECG data.

2.3 ECG Data and Preprocessing

Electrocardiograms (ECGs) are critical tools for diagnosing cardiac arrhythmias, including atrial fibrillation (AF). This research utilizes the Chapman-Shaoxing (China), PTB-XL (Germany), and Georgia (United States) datasets, each consisting of 12-lead ECG signals sampled at 500 Hz. These datasets provide a diverse representation of geographic populations.

To prepare the ECG data for graph-based modeling, the following preprocessing steps were applied:

- **Signal Filtering:** A bandpass filter (0.5–40 Hz) was used to eliminate baseline drift and high-frequency noise while preserving diagnostically relevant components.
- **Feature Extraction:**
  - *Statistical Features:* Mean (μ), variance (σ^2), skewness, and kurtosis.
  - *Clinical Features:* QRS duration (T_{QRS}) and P-wave amplitude (A_P).
- **Signal Normalization:** All signals were scaled to a consistent range to ensure uniform representation across leads (Min-Max Normalization).

2.4 Graph Representation and Weighted Mutual Information

The relationships among ECG leads are represented as a graph $G = (V, E)$, where V corresponds to the 12 leads (nodes) and E denotes edges weighted by the Weighted Mutual Information (WMI) between lead pairs.

WMI quantifies the dependency between time-series signals from two ECG leads. Unlike Pearson or Spearman correlations, WMI captures both linear and non-linear dependencies, making it suitable for complex physiological data. For two signals X_i and X_j, the mutual information is defined as:

$$MI(X_i; X_j) = \sum_{x\in X_i}\sum_{y\in X_j} p(x,y)\log\left(\frac{p(x,y)}{p(x)p(y)}\right),$$

where $p(x, y)$ is the joint probability distribution, and $p(x)$, $p(y)$ are the marginal distributions.

The adjacency matrix A for the graph is constructed as:

$$A_{ij} = w_{ij} \cdot MI(X_i; X_j),$$

where w_{ij} represents weights assigned based on lead categories:

- *Precordial vs. Precordial:* High weights to capture localized spatial dependencies.
- *Precordial vs. Limb:* Moderate weights to reflect complementary diagnostic roles.
- *Limb vs. Limb:* Lower weights for systemic dependencies.

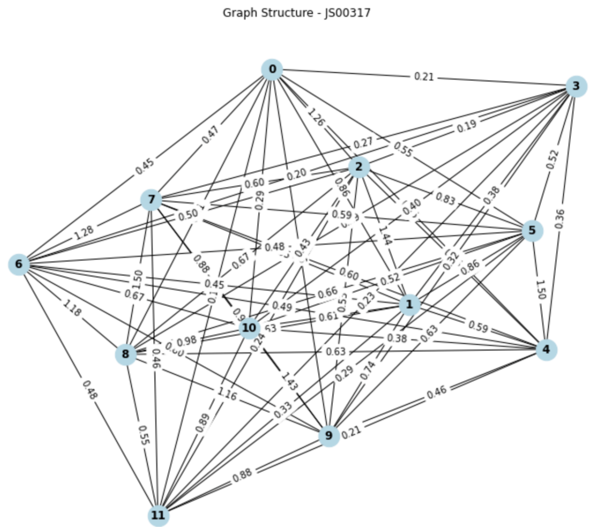


Figure 2: Graph representation of a single ECG record from the Chapman-Shaoxing dataset, visualized on jupyter notebook after applying the Weighted Mutual Information (WMI) matrix

2.5 Graph Attention Network Architecture

The GAT model was designed to dynamically learn inter-lead dependencies using the WMI-based graph representation. The architecture includes:

- **Input Features:** Each lead is represented by a vector combining statistical and clinical features.
- **Model Layers:** Five GAT layers, each with four attention heads, to capture diverse inter-lead relationships.
- **Global Mean Pooling:** Aggregates node embeddings into a single vector for graph-level classification.
- **Output Layer:** A fully connected layer for binary classification (AF vs. Normal).

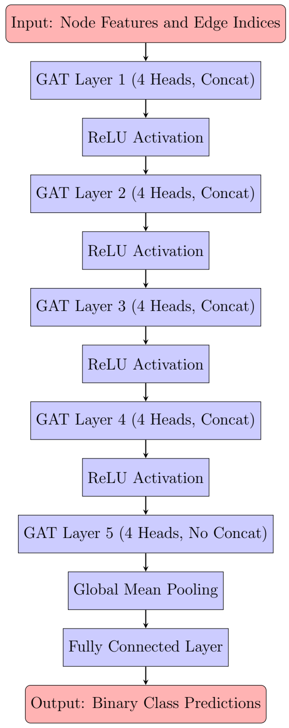


Figure 3: Architecture of the proposed Graph Attention Network (GAT) model. The model consists of five GAT layers, ReLU activations, global mean pooling, and a fully connected classification layer.

2.6 Training and Evaluation

The model was trained on data from the Chapman-Shaoxing, PTB-XL, and Georgia datasets. Training employed cross-entropy loss with the Adam optimizer at a learning rate of 0.001 and was limited to 300 epochs due to runtime constraints associated with computational resources. Evaluation metrics included Accuracy, Precision, Recall, F1-Score, and ROC-AUC. To validate regional influences, Kolmogorov-Smirnov (KS) and Mann-Whitney U tests were applied to compare embedding distributions across geographic datasets.

2.7 Background on t-SNE:

t-SNE (t-Distributed Stochastic Neighbor Embedding) is a widely used dimensionality reduction technique that projects high-dimensional data into a lower-dimensional space (typically 2D or 3D) while preserving the local structure of the data. It accomplishes this by minimizing the divergence between two probability distributions: one that measures pairwise similarities in the high-dimensional space and another that measures similarities in the low-dimensional space. This method is particularly effective for visualizing complex datasets, as it emphasizes clustering and separation of similar data points, making patterns and relationships more interpretable.

3 Results

3.1 Comparative Analysis of Model Performance

The proposed Graph Attention Network (GAT) outperformed traditional machine learning models, demonstrating its efficacy in capturing the complex spatial and statistical dependencies inherent in 12-lead ECG data. While the Random Forest model achieved a moderate accuracy of 87.65% with a ROC-AUC score of 71.16% and a poor F1-score of 0.0633, it lacked the ability to model the inter-lead relationships crucial for arrhythmia detection. In contrast, the GAT framework achieved a significantly higher accuracy. This demonstrates the strength of graph-based learning approaches over traditional methods.

```
Performance Metrics:
Accuracy       : 0.9482
Precision      : 0.7294
Recall         : 0.8857
F1-Score       : 0.8000
ROC-AUC Score  : 0.9728

Classification Report:
              precision    recall  f1-score   support

           0       0.98      0.96      0.97      1058
           1       0.73      0.89      0.80       140

    accuracy                           0.95      1198
   macro avg       0.86      0.92      0.89      1198
weighted avg       0.95      0.95      0.95      1198
```

Figure 4: Performance metrics and classification report for the GAT model with five attention layers and four heads

The confusion matrix further illustrates the robustness of the GAT model, with a high true positive rate for both normal (Class 0) and abnormal (Class 1) ECG signals. The Random Forest model, on the other hand, suffered from a substantial imbalance in its ability to correctly classify abnormal signals, highlighting its limitations in addressing the complexities of ECG data.

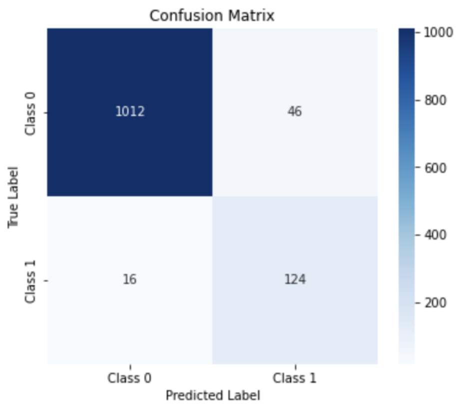


Figure 5: Confusion matrix for the GAT model

The superior performance of the GAT model can be attributed to its dynamic attention mechanism, which enables it to prioritize diagnostically significant inter-lead relationships, something traditional models like Random Forest are incapable of doing. Moreover, the integration of clinical and statistical features as node embeddings in the

GAT further enhanced its ability to discern subtle patterns indicative of atrial fibrillation, providing a more comprehensive diagnostic tool compared to traditional machine learning approaches.

3.2 Performance Evaluation And Generalization

The evaluation of the proposed Graph Attention Network (GAT) model highlighted both strengths and challenges, particularly concerning its generalization across datasets from different geographic regions. The analysis was conducted in two phases: baseline cross-region validation and validation using adjusted region-specific weights. Statistical tests and visualizations further substantiated the findings.

The GAT model was trained on region-specific graphs extracted from the Chapman-Shaoxing (China), PTB-XL (Germany), and Georgia (United States) datasets. Each dataset contained 12-lead ECG signals, which were converted into graph representations. Nodes represented ECG leads, enriched with clinical and statistical features, while edges were weighted using Weighted Mutual Information (WMI). A total of 1996 graphs for China, 1999 for Germany, and 1994 for the United States were used in the experiments.

The model architecture included five GAT layers, each with four attention heads, enabling it to dynamically assign importance to relationships between ECG leads. Node embeddings were aggregated using global mean pooling, and the final classification layer predicted the presence of atrial fibrillation (AF) or normal rhythm. The training process utilized the Adam optimizer with a learning rate of 0.01, and the cross-entropy loss function was employed to optimize model performance. To account for runtime constraints, training was limited to 50 epochs per experiment.

For cross-region validation, data from one geographic region were used for training, while data from another region were used for testing. This approach assessed the model's ability to generalize to datasets with varying demographic and geographic characteristics. For instance, when trained on Chinese data and tested on United States data, the model achieved an accuracy of 0.7693 in the baseline phase, which improved to 0.8365 after incorporating region-specific weights.

The results highlight the significance of leveraging graph-based models for ECG data and the importance of integrating region-specific adjustments to enhance generalizability. By adapting the model to account for regional differences in ECG characteristics, the proposed approach demonstrated superior performance compared to traditional machine learning models.

3.2.1 Baseline Cross-Region Validation

The initial evaluation utilized a standard GAT model without incorporating region-specific weights. The results, summarized in the first cross-region accuracy heatmap, revealed notable variations based on geographic pairings:

- **Same-Region Testing:**
  - Germany achieved the highest accuracy at 0.9750, likely due to consistent and spatially dense data.
  - The United States followed with 0.9424, reflecting balanced data characteristics.

  - China exhibited the lowest accuracy at 0.8575, likely influenced by higher intra-region variability.

- **Cross-Region Testing:**
  - Germany → China: Accuracy dropped to 0.8497, reflecting the model's difficulty in capturing China's unique data characteristics.
  - Germany → United States: Accuracy decreased to 0.8365 due to distributional mismatches.
  - China → United States: The lowest cross-region accuracy was 0.7693, highlighting challenges in transferring features across highly divergent datasets.
  - United States → Germany: Accuracy reached 0.9105, benefiting from shared structural features.
  - United States → China: A relatively high accuracy of 0.9048 was achieved, reflecting partial alignment in feature distributions.

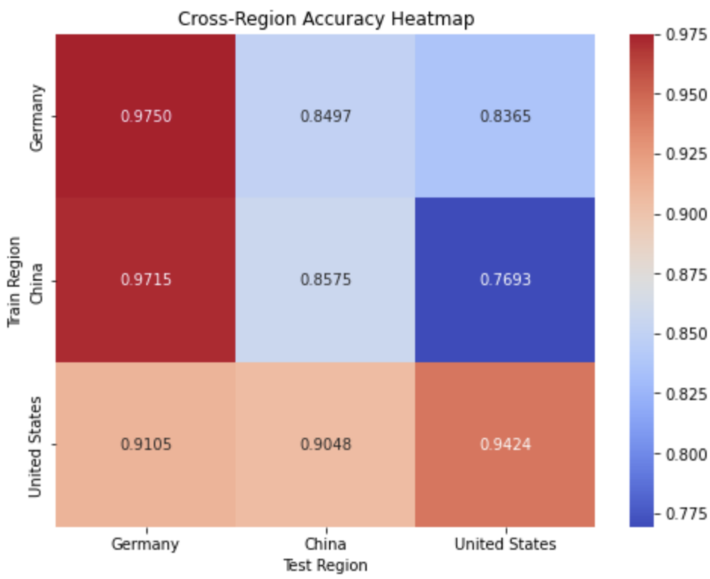


Figure 6: Cross-region accuracy heatmap illustrating the performance of the GAT model

The results reveal that the model's performance is highest in same-region scenarios, but cross-region generalization remains a significant challenge. This reinforces the hypothesis that geographic differences in ECG data influence model predictions.

3.2.2 Adjusted Cross-Region Validation

In the second phase, region-specific weights were applied to the GAT model to address geographic variability. This adjustment aimed to capture regional dependencies in ECG signal distributions and mitigate performance degradation observed in baseline cross-region testing. The heatmap below illustrates the improved cross-region accuracies after adjustments:

- **Key Improvements:**

– United States → Germany:Accuracy increased slightly to 0.9424, demonstrating that regional weighting also benefits datasets with shared structural features.

– China → United States: Accuracy improved from 0.7693 to 0.8365, highlighting the model's capacity to better adapt to the divergent distributions of the U.S. data when trained on China.

- **Same-Region Testing:**

  – Germany retained the highest accuracy at 0.9750, consistent with its dense and uniform data distribution.

  – China lags behind the other two regions in the same-region case (0.8575), indicating potential variability in the data structure or feature representation in this region.

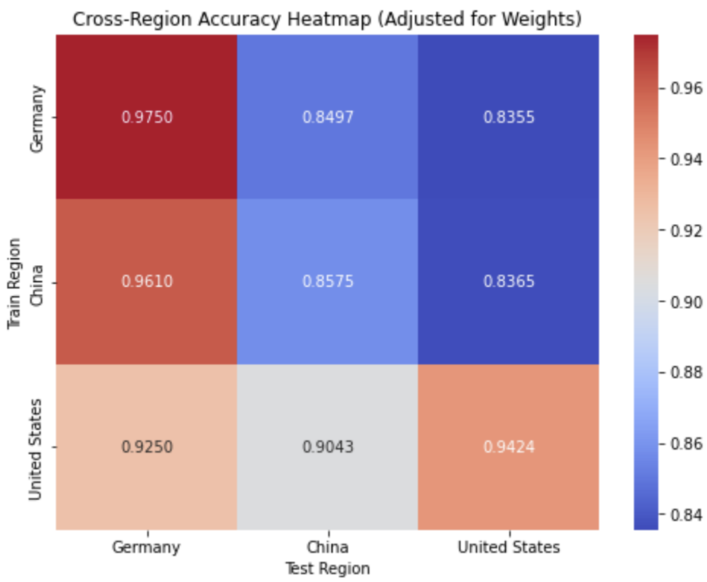


Figure 7: Cross-region accuracy heatmap after adjusted for weights illustrating the performance of the GAT model

The adjusted weights significantly enhanced cross-region performance, particularly for challenging train-test pairs, validating the importance of accounting for regional differences. This is consistent with findings in geographically weighted regression studies, where the incorporation of spatially adaptive weights improved model generalizability across regions [7]. Moreover, the results emphasize the importance of designing ECG models that are robust to demographic and geographic variations, paving the way for developing region-specific diagnostic tools to address population-level disparities effectively [8].

3.3 Statistical Validation

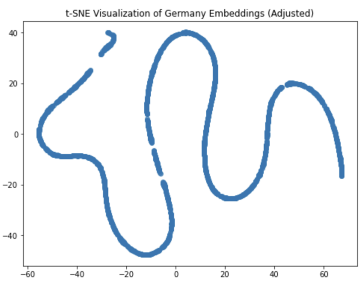
The statistical tests, including the Kolmogorov-Smirnov (KS) and Mann-Whitney U tests, confirmed significant differences in embedding distributions across geographic regions, validating the impact of regional variations on ECG signal relationships.

3.4 t-SNE Visualization

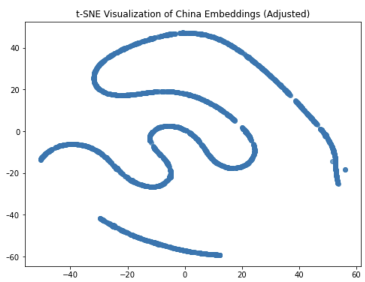
A t-SNE projection of embeddings provided additional insights into the impact of regional differences. Distinct clustering patterns were observed for data from Germany, the United States, and China. The embeddings for Germany exhibited a higher degree of separation from the other two regions, aligning with the statistical test results.

Extraction of t-SNE Embeddings: In the context of this study, the t-SNE projections were applied to embeddings generated by the trained Graph Attention Network (GAT). These embeddings are the learned representations of each ECG graph, captured in the final pooling layer of the GAT model. The embeddings are vector representations that encapsulate both the statistical and spatial dependencies of the ECG leads, as determined by the model.

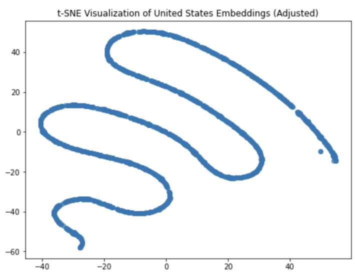
- **From High-Dimensional to Low-Dimensional Representation:**
  - The embeddings for each region (Germany, China, and the United States) were extracted after passing the ECG data through the GAT model.
  - Each graph in the dataset was mapped to a high-dimensional embedding (e.g., 128 or 256 dimensions, depending on the GAT architecture).
  - These embeddings were then projected into a 2D space using t-SNE for visualization and analysis.
- **Purpose of t-SNE in the Study:**
  - The goal of applying t-SNE was to explore how the learned embeddings from the GAT model clustered based on regional differences. This provides insight into whether the model can capture meaningful variations in ECG data attributed to geographic or population-specific factors.



(a) t-SNE visualization of Germany embeddings (after adding weights).



(b) t-SNE visualization of China embeddings (after adding weights).



(c) t-SNE visualization of United States embeddings (after adding weights).

Figure 8: t-SNE visualizations of adjusted embeddings for Germany, China, and the United States regions, demonstrating distinct patterns based on geographic variation.

Significance of t-SNE Results:

- The t-SNE projections visually confirm the findings from statistical tests, which showed significant differences between regions. For example, the greater separation observed in Germany's embeddings aligns with the statistical results that highlighted distinct characteristics of its ECG data.

- These projections also provide insights into cross-region challenges. For instance, the overlap or proximity of clusters between regions like China and the United States indicates potential difficulties in generalizing the model across populations.

- By emphasizing local structure, t-SNE highlights how well the GAT model learned to separate region-specific features, validating the use of Weighted Mutual Information (WMI) and region-specific adjustments.

Implications for Model Development: The t-SNE visualizations reinforce the importance of incorporating regional variations in model design. The ability to visually interpret the embeddings helps identify regions where the model struggles to generalize, guiding further refinements, such as adjusting WMI weights or introducing region-specific adaptations in the GAT architecture.

3.5 Limitations of the Study

While the proposed GAT-based model demonstrates notable improvements in cross-region arrhythmia detection and leverages region-specific weights to enhance generalizability, several limitations must be acknowledged. First, the reliance on preprocessed datasets from distinct geographic regions introduces potential biases due to differences in data collection protocols, population demographics, and clinical practices, which may not generalize to other unseen populations. Second, the model's performance is limited by the size and diversity of the available datasets, which could restrict its applicability in real-world scenarios with greater variability. Third, the computational requirements of GATs, particularly with multiple attention heads and deeper architectures, impose constraints on scalability and runtime efficiency, especially for large-scale datasets. Finally, while t-SNE visualizations and statistical tests provide valuable insights into embedding distributions, they do not fully capture the nuances of inter-lead dependencies or potential outliers, warranting further exploration with alternative dimensionality reduction and interpretability techniques.

4 Discussion and Future Work

4.1 Discussion

This study demonstrated the effectiveness of Graph Attention Networks (GATs) in leveraging Weighted Mutual Information (WMI) for arrhythmia detection across geographically diverse datasets. The use of region-specific weights improved cross-region accuracy, highlighting the importance of accounting for geographic variability in ECG data. However, challenges persisted in certain train-test pairs, underscoring the need for region-aware adaptations in machine learning models. t-SNE visualizations and statistical tests confirmed the distinctiveness of embeddings, though further exploration of their interpretability is necessary.

4.2 Future Work

Future research could focus on expanding datasets to include more regions for better generalizability, integrating advanced feature representations, and improving model interpretability by analyzing attention weights and embeddings. Optimizing the GAT architecture for efficiency and incorporating multi-modal patient data could enhance diagnostic capabilities. Additionally, alternative dimensionality reduction techniques such as UMAP may provide further insights into regional differences.

Acknowledgments

I would like to thank the Technion - Israel Institute of Technology for offering the Machine Learning for Health course, which provided invaluable insights and resources for this research. This course, designed exclusively for a selective group of students in Data Science and Engineering, was instructed by Dr. Uri Shalit. The engaging lectures, discussions, and rigorous project framework in the course were instrumental in shaping the methodology and execution of this study.

References

[1] Yıldırım, Ö., Talo, M., Tan, R. S., Acharya, U. R. (2020). *Deep Learning Techniques for Arrhythmia Detection in ECG Signals: A Review.* Computers in Biology and Medicine.

[2] Shaker, A. M., Al-Angari, H. M., Aldosari, S. A., et al. *Generative Adversarial Network for ECG Signal Enhancement and Classification.*

[3] Yao, Q., Wang, R., Fan, X., et al. *ATI-CNN for Multi-Channel ECG Arrhythmia Classification.*

[4] Zhou, X., Tan, S. *Combining Convolutional Neural Networks (CNNs) with Extreme Learning Machines (ELM) for ECG Feature Extraction and Classification.*

[5] Chen, T., Wang, X., Li, H., et al. *Arrhythmia Detection by Graph Convolutional Network and Proposed Structure for Communication Between Cardiac Leads.*

[6] Fusion of Edge Detection and Graph Neural Networks for Classifying Electrocardiogram Signals (2021). *Arrhythmia Detection Based on Sobel and Canny Edge Representations.*

[7] Laurikkala, J., Kentta, T., et al. *Geographically Weighted Regression with the Integration of Machine Learning for Spatial Prediction.*

[8] Ali, M., et al. *Regional Biases in Image Geolocation Estimation: A Case Study with the SenseCity Africa Dataset.*