An Explainable Geometric-Weighted Graph Attention Network for Identifying Functional Networks Associated with Gait Impairment

Favour Nerrise[1], Qingyu Zhao[2], Kathleen L. Poston[3], Kilian M. Pohl[2], and Ehsan Adeli[2(✉)]

[1] Department of Electrical Engineering, Stanford University, Stanford, CA, USA
fnerrise@stanford.edu
[2] Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA, USA
eadeli@stanford.edu
[3] Department of Neurology and Neurological Sciences, Stanford University, Stanford, CA, USA

Abstract. One of the hallmark symptoms of Parkinson's Disease (PD) is the progressive loss of postural reflexes, which eventually leads to gait difficulties and balance problems. Identifying disruptions in brain function associated with gait impairment could be crucial in better understanding PD motor progression, thus advancing the development of more effective and personalized therapeutics. In this work, we present an explainable, geometric, weighted-graph attention neural network **(xGW-GAT)** to identify functional networks predictive of the progression of gait difficulties in individuals with PD. **xGW-GAT** predicts the multi-class gait impairment on the MDS-Unified PD Rating Scale (MDS-UPDRS). Our computational- and data-efficient model represents functional connectomes as symmetric positive definite (SPD) matrices on a Riemannian manifold to explicitly encode pairwise interactions of entire connectomes, based on which we learn an attention mask yielding individual- and group-level explainability. Applied to our resting-state functional MRI (rs-fMRI) dataset of individuals with PD, **xGW-GAT** identifies functional connectivity patterns associated with gait impairment in PD and offers interpretable explanations of functional subnetworks associated with motor impairment. Our model successfully outperforms several existing methods while simultaneously revealing clinically-relevant connectivity patterns. The source code is available at https://github.com/favour-nerrise/xGW-GAT.

Keywords: Resting-state fMRI · Geometric learning · Attention mechanism · Gait impairment · Explainability · Neuroimaging biomarkers

1   Introduction

Parkinson's Disease (PD) is an age-related neurodegenerative disease with complex symptomology that significantly impacts the quality of life, with nearly 90,000 people diagnosed each year in North America [29]. Recent research has shown that gait difficulty and postural impairment symptoms of PD are highly correlated with alterations in various brain networks, including the motor, cerebellar, and cognitive control networks [25]. Understanding brain functional networks associated with an individual's gait impairment severity is essential for developing targeted interventions, such as physical therapy or brain stimulation techniques. However, most prior works have *either* focused only on a binary diagnosis (PD vs. Control) [16] (ignoring the progression and heterogeneity of the disease symptoms) *or* only used sensor- and vision-based technologies [8,18] to quantify PD symptoms (abstaining from identifying brain networks associated with gait impairment severity).

Graph Neural Networks (GNNs) have been highly successful in inferring neural activity patterns in resting-state fMRI (rs-fMRI) [23]. These models represent functional connectivity matrices as weighted graphs, where each node is a brain region of interest (ROI), and the edges between them capture the magnitude of connectivity, i.e., interactions, as weights. Changes in the connectivity strengths can reflect intrinsic representations in a high-dimensional space that correlate with symptom or disease severity. Assuming that edges with higher weights exert greater functional connectivity (and vice versa), GNNs can encode how ROIs and their neighbors across various individuals can possess similar attributes. GAT [26] is a well-known GNN model that encodes pairwise interactions (edges) into an attention mechanism and uses eigenvectors and eigenvalues of each node as positional embeddings for local structures. However, since each node or ROI in a brain network has the same degree and connects to every other node, standard graph representations are limited in modeling functional connectivity differences in a high-dimensional space that can be used for inter-subject functional covariance comparison. Riemannian geometry [14] is another robust, mathematical framework for rs-fMRI analysis that projects a functional, connectivity matrix in a manifold of symmetric positive-definite (SPD) matrices, making it possible to model high-dimensional, edge interactions and dependencies. It has been applied to analyzing gait patterns [20] in Parkinson's disease and to functional brain network analysis in other neurological disorders (e.g., Mild Cognitive Impairment [7] and autism [30]).

Addressing the problem of identifying brain functional network alterations related to the severity of gait impairments presents several challenges: (**i**) clinical datasets are often sparse or highly imbalanced, especially for severely impaired disease states; (**ii**) although substantial progress has been made in modeling functional connectomes using graph theory, few studies exist that capture the individual variability in disease progression and they often fall short of generating clinically relevant explanations that are symptom-specific.

In this work, we propose a novel, explainable, geometric weighted-graph attention network (**xGW-GAT**) that embeds functional connectomes in a

learnable, graph structure that encodes discriminative edge attributes used for attention-based, transductive classification tasks. We train the model to predict a gait impairment rating score (MDS-UPDRS Part 3.10) for each PD participant. To mitigate limited clinical data across all different classes of gait impairment and data imbalance (challenge **i**), we propose a stratified, learning-based sample selection method that leverages non-Euclidean, centrality features of connectomes to sub-select training samples with the highest predictive power. To provide clinical interpretability (challenge **ii**), **xGW-GAT** innovatively produces individual and global attention-based, explanation masks per gait category and soft assigns nodes to functional, resting-state brain networks. We apply the proposed framework on our dataset of 35 clinical participants and compare it with existing methods. We observe significant improvements in classification accuracy while enabling adequate clinical interpretability.

In summary, our contributions are: (1) we propose a novel, geometric attention-based model, **xGW-GAT**, that uses edge-weights to depict neighborhood influence from local node embeddings during dynamic, attention-based learning; (2) we develop a multi-classification pipeline that mitigates sparse and imbalanced sampling with stratified, learning-based sample selection during training on real-world clinical datasets; (3) we provide an explanation generator to interpret attention-based, edge explanations that highlight salient brain network interactions for gait impairment severity states; (4) we establish a new benchmark for PD gait impairment assessment using brain functional connectivities.

2 XGW-GAT: Explainable, Geometric-Weighted GAT

Problem definition. Assume a set of functional connectomes, $\mathcal{G}_n \in \mathbb{R}^{d \times d}, \mathcal{G}_2, \ldots, \mathcal{G}_N$ are given, where N is the number of samples and d is the number of ROIs. Each connectome is represented by a weighted, undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{W})$, where $\mathcal{V} = \{v_i\}_{i=1}^d$ is the set of nodes, $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the edge set, and $\mathbf{W} \in \mathbb{R}^{|\mathcal{V}| \times |\mathcal{V}|}$ denotes the matrix of edge weights. The weight w_{ij} of an edge $e_{ij} \in \mathcal{E}$ represents the strength of the functional connection between nodes v_i and v_j, i.e., the Pearson correlation coefficient of the time series of the pair of the nodes. Each $\mathcal{G}_n$ contains node attributes X_n and edge attributes H_n. We develop a model that predicts a gait impairment score, $\mathcal{Y}_n$ and outputs an individual explanation mask $\mathbf{M}_c \in \mathbb{R}^{d \times d}$ per class c to assign ROIs to functional brain networks.

2.1 Connectomes in a Riemannian Manifold

Functional connectivity matrices belong to the manifold of symmetric positive-definite (SPD) matrices [31]. We leverage Riemannian geometry to perform principled comparisons between different connectomes, such as prior work [24]. To highlight connections between adjacent nodes, each weight matrix $\mathbf{W}_n \in \mathbb{R}^{d \times d}$ can be represented as a symmetric, adjacency matrix with zero, non-negative
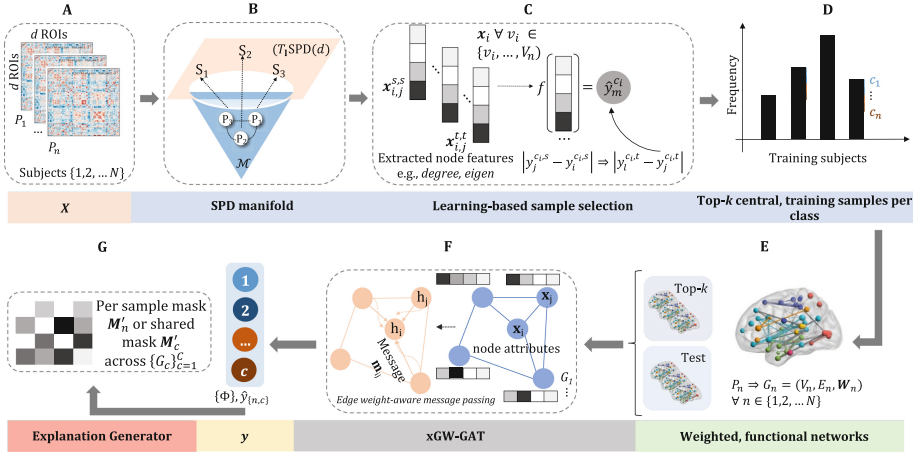
Fig. 1. xGW-GAT. (A) Input: functional connectomes. (B) Extract pairwise tangent matrices in SPD(d). (C) Compress tangent matrices into weighted graphs (connectomes) (D) Use linear regression to train a mapping, f, on training samples to learn pairwise differences between target and record scores. (E) Group top-k samples per class across N-fold cross-validation runs with the lowest predicted difference and oversample for imbalance. (F) Represent samples as weighted, graphs and use edge weight-aware attention to encode and propagate learning; predict gait score. (G) Produce explanation masks for each class or individual participants within functional brain networks.

eigenvalues, where each element of the adjacency matrix is the edge weight, w_{ij} between nodes i and j. We then consider $\mathbf{W}_n$ to be a point, $\mathbf{S}_n$, in the manifold of SPD matrices Sym_d^+ that locally looks like a topological Euclidean space. However, Sym_d^+ does not form a vector space; thus, we project each SPD matrix $\mathbf{S}_n$ onto a common tangent space using parallel transport. Given a reference point $\mathbf{S}_i \in Sym_d^+$, we transport a tangent vector $v \in T_{\mathbf{S}_j}$ from $\mathbf{S}_j$ to $\mathbf{S}_i$ along the geodesic connecting $\mathbf{S}_j$ and $\mathbf{S}_i$ (see Fig. 1-B). This process is performed for each subject $n = 1, 2, \ldots, N$, yielding a set of tangent vectors in a common tangent space that can be analyzed using traditional Euclidean methods.

To calculate the geodesic distance between two SPD matrices $\mathbf{S}_i$ and $\mathbf{S}_j \in Sym_d^+$, we adopt the Log-Euclidean Riemannian Metric (LERM) [1] distance, d_{LERM} as follows:

$$d_{\text{LERM}}(\mathbf{S}_i, \mathbf{S}_j) = \|log(\mathbf{S}_i) - log(\mathbf{S}_j)\|_F^2 \tag{1}$$

where $\|\cdot\|_F^2$ is the Frobenius norm. LERM is invariant to similarity transformations (scaling and orthogonality) and is computationally efficient for high-dimensional data. See the Supplementary Material for results with other distance metrics.

2.2 Stratified Learning-Based Sample Selection

Data availability and dataset imbalance are re-occurring challenges with real-world clinical datasets, often leading to bias and overfitting during model train-

ing. We address this by expanding a learning-based sample selection method [11] to weight per-class distributions. We assume that similar brain connectivity networks are correlated with disease severity whereas connectomes that vary in topological patterns might elicit different gait impairment scores. Our subsampling technique selects training samples containing the highest representative power, i.e., contributing the least amount of pairwise differences for predicting a gait score. We divide our training samples into subgroups: train-in, n_s, and holdout, n_t using N-fold cross-validation. For each pair of symmetric d-by-d tangent matrices, $\mathbf{S}_{i,j}^{s,s} \in T_{\mathbf{I}}\mathrm{SPD}(d)$, we encode the pairwise differences between the connectomes from the train-in, n_s, to obtain a set of $n_s(n_s-1)/2$ tangent matrices in $T_{\mathbf{I}}\mathrm{SPD}(d)$. Each tangent matrix represents the "difference" between two connectomes. We affix a threshold of k samples to be selected from each class c to identify l central training samples with the highest expected predictive power, i.e., the *lowest average difference* in target gait impairment scores per class, $\hat{y}_c$, between samples j from the train-in and holdout group. We select degree, closeness, and eigenvector centrality as our topological features that encode information on *changes* in node connectivity. We train a linear regression mapping f on the Riemannian geometric distances $\mathcal{D}_{le}(\mathbf{S}_i^s, \mathbf{S}_j^s)$ between the connectomes from n_s using the vectorized upper triangular portion (including the diagonal) of the tangent matrices. The *absolute difference in target score, per class*, between samples i and j from the train-in group n_s is denoted by $|\hat{y}_{c,j}^s - \hat{y}_{c,i}^s|$ (see Fig. 1-C). The top-k samples per class with the highest predictive power are sub-selected from the total training set, oversampled for class imbalance with RandomOverSampler [15], and used for training xGW-GAT layers (see Fig. 1-D).

2.3   Dynamic Graph Attention Layers

Attention. We employ Graph Attention Network version 2 (GATv2) [2], a GAT [26] variant to perform dynamic, multi-head, edge-weight attention message passing for classifying each $\mathbf{S}_n$. We assume that every node $i \in \mathcal{V}$ has an initial representation $\mathbf{h}_i^{(0)} \in \mathbb{R}^{d_0}$. GATv2 updates each node representation, $\mathbf{h}$ based on the features of neighboring nodes and the edge weights between nodes by computing attention scores α_{ij} for every edge (i,j) by normalizing attention coefficients $e(\mathbf{h}_i, \mathbf{h}_j)$. α_{ij} measures the importance of node j's features to node i's at layer l by performing a weighted sum over the neighboring nodes $j \in \mathcal{N}_i$:

$$e(\mathbf{h}_i, \mathbf{h}_j) \coloneqq \mathrm{LeakyReLU}(\mathbf{a}^{(l)\top}[\mathbf{\Theta}^{(l)}\mathbf{h}_i^{(l-1)} \| \mathbf{\Theta}^{(l)}\mathbf{h}_j^{(l-1)}]) \tag{2}$$

$$\alpha_{ij} \coloneqq softmax_j(e(\mathbf{h}_i, \mathbf{h}_j)) \tag{3}$$

$$\mathbf{h}_i^{(l+1)} \coloneqq \sigma\left(\sum_{j \in \mathcal{N}_i} \alpha_{ij}^{(l)} \mathbf{h}_j^{(l-1)}\right). \tag{4}$$

where $\mathbf{a}^{(l)} \in \mathbb{R}^{2F}$ and $\mathbf{\Theta}^{(l)}$ are trainable parameters and learned, $\mathbf{h}_i^{(l)} \in \mathbb{R}^F$ is the embedding for node i, σ represents a non-linearity activation function, and $\|$ denotes vector concatenation. As conventional graph attention mechanisms for

transductive tasks typically do *not* incorporate edge attributes, we introduce an attention-based, message-passing mechanism incorporating edge weights, similar to [6]. The algorithm uses a message vector $\mathbf{m}_{ij} \in \mathbb{R}^F$ by concatenating node features of neighboring nodes i, j, and edge weight $\mathbf{W}_{i,j}$:

$$\mathbf{m}_{ij}^{(l)} = \mathrm{MLP}_1([\mathbf{h}_i^{(l)}; \mathbf{h}_j^{(l)}; \mathbf{W}_{ij}]), \tag{5}$$

where MLP_1 is a Multi-Layer Perceptron. Accordingly, an update of each ROI representation is influenced by its neighboring regions weighted by their connectivity strength. After stacking L layers, a readout function summarizing all node embeddings is employed to obtain a graph-level embedding $\mathbf{g}$:

$$\mathbf{z} = \sum_{i \in V} \mathbf{h}_i^{(L)}, \mathbf{g} = \mathrm{MLP}_2(\mathbf{z}) + \mathbf{z}. \tag{6}$$

Loss Function. xGW-GAT layers (Fig. 1-F) are trained with a supervised, weighted negative log-likelihood loss function to mitigate class imbalance across classes, C, defined as:

$$\mathcal{L}_{\mathrm{NLL}} := \frac{-1}{N} \sum_{p=1}^{N} \sum_{q=1}^{C} r_q y_{pq} log(\hat{y}_{pq}), \tag{7}$$

where r_q is the rescaling weight for the q-th class, y_{pq} is the q-th element of the true label vector y_p for the p-th sample, and $\hat{y}_{pq}$ is the predicted label vector.

2.4 Individual- And Global-Level Explanations

We define an attention explanation mask for each sample, $n \in 1, 2, \ldots, N$ and for each class $c \in 1, 2, \ldots, C$ that identifies the most important node/ROI connections contributing to the classification of subjects. We return a set of attention coefficients $\alpha^n = [\alpha_1^n, \alpha_2^n, \ldots, \alpha_S^n]$ for each sample n, where S is the number of attention heads. We aggregate trained, attention coefficients per sample used for predicting each $\hat{y}$ using a *max* operation that returns $\alpha_{\max}^n \in \mathbb{R}^{d \times d}$. An explanation mask per class, M_c, or per sample, M_n, can be derived using the max attention coefficients, $\alpha_{\max}$ (Fig. 1-G):

$$\mathbf{M}_c = \frac{1}{N} \sum_{n=1}^{N} \alpha_{\max}; \quad \mathbf{M}_n = \frac{1}{C} \sum_{c=1}^{C} \alpha_{\max}. \tag{8}$$

$\mathbf{M}$ can be soft-thresholded to retain the top-L most positively attributed attention weights to the mask as follows:

$$\mathbf{M}'[i] = \begin{cases} \mathbf{M}[i] & \text{if } \mathbf{M}[i] \in \text{Top-L}(\mathbf{M}) \\ 0, & \text{otherwise}, \end{cases} \tag{9}$$

where Top-L($\mathbf{M}$) represents the set of top-L elements in $\mathbf{M}$.

3   Experiments

Dataset. We obtained data from a private dataset ($n = 35$, mean age 69 ± 7.9) defined in [19], which contains MDS-UPDRS exams from all participants. Following previously published protocols [21], all participants are recorded during the off-medication state. Participants were evaluated by a board-certified movement disorders specialist on a scale from 0 to 4 based on MDS-UPDRS Sect. 3.10 [10]. The dataset includes 22 participants with a score 1, 4 participants with a score 2, 4 participants with a score of 3, 4 participants with a score 4, and 1 participant with a score 0 on MDS-UPDRS item 3.10. The single score-0 participant (normal) was combined with the score-1 participants (minor gait impairment) to adjust for severe class imbalance. We pre-processed functional connectivity matrices and corrected them for possible motion artifacts using the CONN toolbox [28]. The FC matrices were obtained using a combined Harvard-Oxford and AAL parcellation atlas [28] with 165 ROIs, where each entry in row i and column j in the matrix is the Pearson correlation between the average rs-fMRI signal measured in ROI i and ROI j. We imputed any missing ROI network scores with the mean score per column and Z-transformed FC matrices [$\mu = 0, \sigma = 1$]. This dataset (like other clinical datasets in practice) poses highly imbalanced distributions for classes with severe impairment, which makes it useful to demonstrate our method's capability in an imbalanced and limited-data scenario. In addition, most existing studies focus on differentiating participants from controls, while the severity of specific impairments is understudied (our focus).

Software. All experiments were implemented in Python 3.10 and ran on Nvidia A100 GPU runtimes. We used PyTorch Geometric [9], PyTorch, and Scikit-learn for machine learning methods. We used the SPD class from the Morphometrics package to compute Riemannian geometrics and NetworkX to extract the topological features of the graphs from the tangent matrices. Hyper-parameters are tuned automatically with the open-source AutoML toolkit NNI (https://github.com/microsoft/nni).

Setup. We used the mean, connectivity profile, i.e., **W** [5], as the node feature for xGW-GAT layers, a weighted ADAM optimizer, a learning rate of $1e - 4$, a batch size of 2 for training and 1 for test, and 100 training epochs. We used 2 GATv2 layers, dropout rate $= 0.1$, hidden_dim $= 8$, heads$= 2$, and a global mean pooling layer. We used 4-fold cross-validation to partition training and holdout sets and selected $k = 4$ as the optimal number of selected training samples between 2 and 15. We report weighted, macro average scores for F_1, area under the ROC curve (AUC), precision (Pre), and recall (Rec) over 100 trials.

3.1   Results

We perform a multi-class classification task of *Slight(1), Mild(2), Moderate(3), Severe(4)* gait impairment severity. To benchmark our method, we compare our

Table 1. Comparison with baseline and ablated methods. * indicates statistical difference by Wilcoxon signed rank test at ($p < 0.05$) compared with our method.

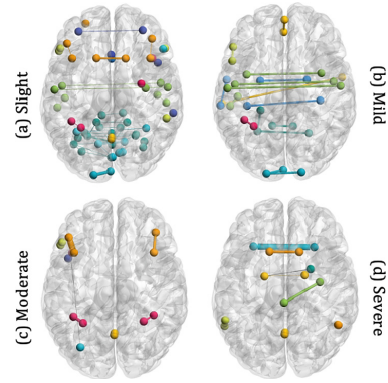| Method	Pre	Rec	F_1	AUC
GCN* [13]	0.46	0.48	0.47	0.54
PNA* [4]	0.52	0.54	0.53	0.56
BrainNetCNN*[12]	0.62	0.71	0.66	0.57
BrainGNN*[17]	0.66	0.53	0.59	0.62
GAT* [26]	0.70	0.58	0.64	0.71
xGW-GAT (dc)*	0.61	0.65	0.63	0.51
xGW-GAT (ec)*	0.64	0.62	0.63	0.72
xGW-GAT (cc)*	0.61	0.53	0.57	0.57
xGW-GAT (!ss)*	0.55	0.47	0.51	0.54
xGW-GAT (ss)	**0.75**	**0.77**	**0.76**	**0.83**

Fig. 2. Salient ROI connections on explanation brain networks across the four classes of gait impairment (DMN, SMN, VN, SN, DAN, FPN, LN, CN, and BLN).

results with several state-of-the-art classifiers: GAT [26], GCN [13], PNA [4], and two state-of-the-art deep models design for brain networks: BrainNetCNN [12] and BrainGNN [17]. We also perform an ablation study on sample selection and the type of topological features used in training our method: (!ss) no [stratified, learning-based] sample selection, (dc) node degree centrality, (cc) node closeness centrality, and (ec) eigenvector centrality. Results for the highest-performing settings of **xGW-GAT** are displayed in Table 1 and node feature descriptions are included in the Supplementary Material.

The results (Table 1) show that **xGW-GAT** yields significant improvement in performance over SOTA graph-based models, including models designed for brain network analysis. xGW-GAT with our stratified, learning-based selection method combined with the RandomOverSampler technique to temper the effects of class imbalance outperforms a standard xGW-GAT by 42%. Compared with SOTA deep models like GCN and PNA, our model also outperforms them by large margins, with up to 29% improvement for AUC. These predictions are promising for an explainable analysis of PD gait impairment while also minimizing random uncertainties introduced in individual participant graphs.

4   Discussion

Brain Networks Mapping. As shown in Fig. 2, we aid interpretability for clinical relevance by partitioning the ROIs into nine "networks" based on their functional roles: Default Mode Network (DMN), SensoriMotor Network (SMN), Visual Network (VN), Salience Network (SN), Dorsal Attention Network (DAN), FrontoParietal Network (FPN), Language Network (LN), Cerebellar Network (CN), and Bilateral Limbic Network (BLN) are colored accordingly, while edges

across different systems are colored gray. Edge widths here are the attention weights.

Salient ROIs. We provide per-class and individual-level interpretations for understanding how ROIs contribute to predicting gait impairment scores. We build the node and edge files with the thresholded attention explanation masks, $\mathbb{M}'$ per PD participant or per class and plot glass brains using BrainNet Viewer ((https://www.nitrc.org/projects/bnv/)).

We observe that rich interactions decrease significantly for the *Mild* class, Fig. 2(b), within the CN, primarily associated with coordinating voluntary movement, the SN, responsible for thought, cognition, and planning behavior, and the VN, the center for visual processing during resting and task states. These observations are consistent with existing neuroimaging findings, which support that PD is positively associated with the severity of cognitive deficits and neuromotor control for inter-network and intra-network interactions within the salience network, cerebellar lobules, and visual network [22,32]. Similarly, there are significantly lower connections within CN and VN and sparser connections within the SMN for the *Moderate* and *Severe* classes, Fig. 2(c)-(d). Existing studies show functional connectivity losses within the sensorimotor network (SMN) [3] are correlated with disruptions in regional and global topological organization for SMN areas for people with PD, resulting in loss of motor control. For *Mild*, *Moderate*, and *Severe* PD participants, abrupt connectivity is also observed for the frontoparietal network, FPN, known for coordinating behavior and associated with connectivity alterations correlated with motor deterioration [27].

5   Conclusion

This study showcases a novel benchmark for using an explainable, geometric-weighted graph attention network to discover patterns associated with gait impairment. The framework innovatively integrates edge-weighted attention encoding and explanations to represent neighborhood interactions in functional brain connectomes, providing interpretable functional network clustering for neurological analysis. Despite a small sample size and imbalanced settings, the lightweight model offers stable results for quick inference on categorical PD neuromotor states. Future work includes new experiments, an expanded, multi-modal dataset, and sensitivity and specificity analysis to discover subtypes associated with the severity of PD gait impairment.

References

1. Arsigny, V., Fillard, P., Pennec, X., Ayache, N.: Geometric means in a novel vector space structure on symmetric positive-definite matrices. SIAM J. Matrix Anal. Appl. **29**(1), 328–347 (2007)

2. Brody, S., Alon, U., Yahav, E.: How attentive are graph attention networks? arXiv preprint arXiv:2105.14491 (2021)
3. Caspers, J., et al.: Within-and across-network alterations of the sensorimotor network in Parkinson's disease. Neuroradiology **63**(12), 2073–2085 (2021)
4. Corso, G., Cavalleri, L., Beaini, D., Liò, P., Veličković, P.: Principal neighbourhood aggregation for graph nets. NeurIPS **33**, 13260–13271 (2020)
5. Cui, H., et al.: Braingb: a benchmark for brain network analysis with graph neural networks. IEEE TMI 2022 (2022)
6. Cui, H., Dai, W., Zhu, Y., Li, X., He, L., Yang, C.: Interpretable graph neural networks for connectome-based brain disorder analysis. In: MICCAI 2022, pp. 375–385. Springer (2022). https://doi.org/10.1007/978-3-031-16452-1_36
7. Dodero, L., Minh, H.Q., Biagio, M.S., Murino, V., Sona, D.: Kernel-based classification for brain connectivity graphs on the riemannian manifold of positive definite matrices. In: 2015 IEEE ISBI, pp. 42–45 (2015)
8. Endo, M., Poston, K.L., Sullivan, E.V., Fei-Fei, L., Pohl, K.M., Adeli, E.: GaitForeMer: Self-Supervised Pre-Training of Transformers via Human Motion Forecasting for Few-Shot Gait Impairment Severity Estimation. MICCAI, pp. 130–139 (2022). https://doi.org/10.1007/978-3-031-16452-1_13
9. Fey, M., Lenssen, J.E.: Fast graph representation learning with pytorch geometric. arXiv preprint arXiv:1903.02428 (2019)
10. Goetz, C.G., et al.: The MDS-sponsored revision of the unified Parkinson's disease rating scale. Official MDS Dutch Translation (2019)
11. Hanik, M., Demirtaş, M.A., Gharsallaoui, M.A., Rekik, I.: Predicting cognitive scores with graph neural networks through sample selection learning. Brain Imaging Behav. **16**(3), 1123–1138 (2022)
12. Kawahara, J., et al.: Convolutional neural networks for brain net-works; towards predicting neurodevelopment. Neu-roImage (2017)
13. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907 (2016)
14. Klingenberg, W.: Contributions to Riemannian geometry in the large. Ann. Math. **69**(3), 654–666 (1959)
15. Lemaître, G., Nogueira, F., Aridas, C.K.: Imbalanced-learn: a python toolbox to tackle the curse of imbalanced datasets in machine learning. J. Mach. Learn. Res. **18**(17), 1–5 (2017)
16. Li, K., Su, W., Li, S.H., Jin, Y., Chen, H.B.: Resting state fMRI: a valuable tool for studying cognitive dysfunction in pd. Parkinson's Disease 2018 (2018)
17. Li, x, et al.: Braingnn: interpretable brain graph neural network for fMRI analysis. Med. Image Anal. **74**, 102233 (2021)
18. Lu, M., et al.: Vision-based estimation of MDS-UPDRS gait scores for assessing Parkinson's disease motor severity. MICCAI **2020**(12263), 637–647 (2020)
19. Lu, M.: Quantifying Parkinson's disease motor severity under uncertainty using MDS-UPDRS videos. Med. Image Anal. **73**, 102179 (2021)
20. Olmos, J., Galvis, J., Martínez, F.: Gait patterns coded as Riemannian mean covariances to support Parkinson's disease diagnosis. In: IBERAMIA, pp. 3–14 (2023)
21. Poston, K.L., et al.: Compensatory neural mechanisms in cognitively unimpaired Parkinson disease. Ann. Neurol. **79**(3), 448–463 (2016)
22. Ruan, X., et al.: Impaired topographical organization of functional brain networks in parkinson's disease patients with freezing of gait. Front. Aging Neurosci. **12**, 580564 (2020)

23. Rubinov, M., Sporns, O.: Complex network measures of brain connectivity: uses and interpretations. Neuroimage **52**(3), 1059–1069 (2010)
24. Shahbazi, M., Shirali, A., Aghajan, H., Nili, H.: Using distance on the Riemannian manifold to compare representations in brain and in models. Neuroimage **239**, 118271 (2021)
25. Togo, H., Nakamura, T., Wakasugi, N., Takahashi, Y., Hanakawa, T.: Interactions across emotional, cognitive and subcortical motor networks underlying freezing of gait. NeuroImage: Clin. **37**, 103342 (2023)
26. Velickovic, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., Bengio, Y., et al.: Graph attention networks. stat **1050**(20), 10–48550 (2017)
27. Vervoot, G., et al.: Functional connectivity alterations in the motor and fronto-parietal network relate to behavioral heterogeneity in parkinson's disease. Parkinsonism Related Disorders **24**, 48–55 (2016)
28. Whitfield-Gabrieli, S., Nieto-Castanon, A.: Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. Brain connectivity **2**(3), 125–141 (2012)
29. Willis, A., et al.: Incidence of Pakinson disease in north America. NPJ Parkinson's Disease **8**(1), 170 (2022)
30. Wong, E., Anderson, J.S., Zielinski, B.A., Fletcher, P.T.: Riemannian regression and classification models of brain networks applied to autism. In: CNI 2018, Held in Conjunction with MICCAI 2018,pp. 78–87 (2018)
31. You, K., Park, H.J.: Re-visiting Riemannian geometry of symmetric positive definite matrices for the analysis of functional connectivity. Neuroimage **225**, 117464 (2021)
32. Zhu, H., et al.: Abnormal dynamic functional connectivity associated with subcortical networks in Parkinson's disease: a temporal variability perspective. Front. Neurosci. **13**, 80 (2019)