GEOMETRIC DEEP LEARNING FOR SULCAL GRAPHS

R. Yadav, François-Xavier Dupé, S Takerkart, G Auzias

GEOMETRIC DEEP LEARNING FOR SULCAL GRAPHS

R.Yadav[⋆‡†], *F.X. Dupé*[†], *S. Takerkart*[⋆], *G. Auzias*[⋆]

[⋆]Institut de Neurosciences de la Timone UMR 7289, Aix-Marseille Université, CNRS
[‡]Aix Marseille Université, Institut Marseille Imaging, Marseille, France
[†]Laboratoire d'Informatique et Systèmes UMR 7020, Aix-Marseille Université, CNRS

The human cerebral cortex features intricate, convoluted sulci forming a complex and variable geometry. Characterizing these variations in the sulcal patterns is critical for assessing cerebral structure abnormalities.However, the uniqueness of sulcal patterns in individuals poses challenges in dissociating abnormal from normal variations. In this study, we use deep graph representation learning models to analyze sulcal graphs from a population of 1090 healthy adults. We leverage sulcal graphs for encoding the surface geometry, combined with graph neural networks(GNNs) for predicting the gender of the subjects, and analyze the underlying variability. These methods show potential in offering valuable insights into biomarker discovery for neurological and psychiatric disorders, relying on structured representations of cortical geometry.

1. INTRODUCTION

The highly convoluted geometry of the cerebral cortex is difficult to analyze and compare across individuals. Various methods have been proposed to unravel the complicated patterns of cortical folding. Recently, deep learning techniques have reached the forefront in numerous neuroimaging applications [1]. They have showed a remarkable capability to autonomously harness hierarchical feature representations from large datasets, surpassing traditional approaches[2].

A limited number of studies have explored the potential of deep learning methods for the analysis of folding patterns. One of the first study by [3], evaluated convolutional neural networks for classifying folding patterns. More recently, [4] used the well established β-VAE [5] for the detection of abnormal patterns. Similarly, another work by [6] compared the former β-VAE approach with a recent self-supervised deep contrastive model(SimCLR)[7]. These studies introduced a framework for conducting exploratory analysis of cortical folds by leveraging existing representation learning techniques. These studies aim at capturing the intrinsic characteristics of cortical folds using unsupervised generative methods, that learn a latent space under the constraint of reconstructing the input data. In parallel, studies such as [2] demonstrated that supervised learning settings yield robust and discriminative latent representations that are relevant for diagnostic classification and disease characterisation. In this work, we explore the potential of this strategy for learning representations of a population of sulcal graphs, taking the gender classification as a pretext task. The relevance of gender classification as a pretext task in the examination of brain morphology is supported by many studies reporting that gender shapes the spatial pattern of cortical alterations in many conditions such as autism [8] and dyslexia[9].

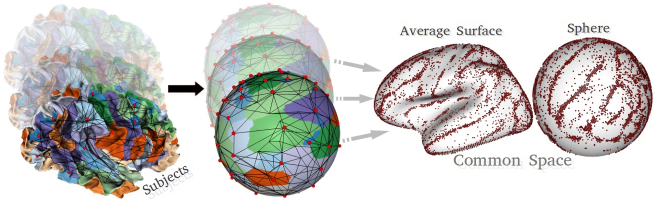

Fig. 1. The sulcal graph from each subject is transferred onto a common sphere using the inflation and spherical registration tools from *freesurfer*[2]. The sulcal graphs from every subjects can be mapped in an average surface for visualization. Note that the spatial distribution of nodes in common spaces is uneven, forming dense clusters in cortical regions where individual variations are minimal.

Additionally, in contrast with [3, 4, 6], we explicitly take into account the spatial relationships between cortical folds by using sulcal graphs to represent patterns of cortical folds. As proposed in [10], a sulcal graph is constructed by decomposing the cortical surface into *sulcal basins*. Illustrated on Fig.1, sulcal graphs effectively capture the spatial relationship of the sulci, that was ignored in former methods. Note that since the sulcal graph extraction approach is applied to each individual separately, the number of nodes and graph topology vary across individuals. In contrast, in most approaches for designing brain graphs from the literature, the graphs are constructed based on a predefined atlas that ensure an equal number of nodes for all subjects. While this technique simplifies the manipulation of the graphs with GNNs, it does disregard the essential consideration of individual variations in

[2]https://surfer.nmr.mgh.harvard.edu/

cortical folding patterns. On the contrary, our sulcal graphs do exhibit topological variations, necessitating the identification of suitable GNN architectures.

In the present work, we examine for the first time the relevance GNNs for the analysis of sulcal graphs. We compare two alternative GNN designs : a shallow Graph attention network (GAT) and a shallow GCN model with a hierarchical graph pooling strategy. Both of these architectures were selected for their capacity to offer interpretable maps and with the anticipation of producing complementary insights. All source code and data are publicly available on our github repository[3].

2. METHODS

Formally, a sulcal graph $\mathcal{G}$ is given as a triplet of vertices, edges and attributes: $\mathcal{G} = (V, E, X)$, with $v_1, v_2 \dots v_M \in V$ being the nodes in $\mathcal{G}$ and $|\mathcal{G}| = M$ is the number of nodes in the graph.

Graph Neural Networks

Graph Neural Networks (GNNs) have gained prominence for their ability to integrate node features and graph topology for end-to-end learning, with applications to analyzing complex brain data and predicting brain diseases[11]. Notable advancements have resulted in more interpretable models and improved outcomes. For instance, BrainGNN, introduced by [12] uses fMRI data analysis to define biomarkers in Autism Spectrum Disorder using specialized graph convolutional and pooling layers. Additionally, BrainNetCNN, developed by [13], employs a variety of convolutional filters to exploit topological relationships in brain networks. Furthermore, [14] conducted a comprehensive benchmarking of diverse GNN architectures and feature engineering methods for brain network analysis. These studies provide a robust framework for applying GNNs to the analysis of brain graphs with various architectures and learning strategies [11]. However, all the above methods are intended for graphs with uniform number of nodes across subjects, which makes them unsuitable for direct application in the present study.

The basic process underlying all of the GNN designs is its message passing operation. During the learning process it allows nodes to gather and integrate information from surrounding nodes. As a result, the model is able to capture the complex connections and relationships inherent in data arranged as a graph. The message passing operation typically consists of two main steps: *message aggregation* and *message update*.

- *Message aggregation*: In this step, for each node in the graph a message is computed based on its own features and the features of its neighboring nodes:

$$m_{v_i}^{(q)} = \phi^{(q)}(h_{v_i}^{(q)}, \{h_{v_j}^{(q)}\}_{j \in \mathcal{N}_{v_i}}) \tag{1}$$

where $m_{v_i}^{(q)}$ is the computed message for node v_i, $h_{v_i}^{(q)}$ are the features of node v_i at q^{th} layer and $\mathcal{N}_{v_i}$ denotes the neighborhood of v_i. $\phi^{(q)}$ is a differentiable function that computes the message based on the given input features of the nodes in the graph. The features of nodes at layer $q = 0$ correspond to initial node attributes, $h_{v_i}^{(0)} = x_{v_i}, \forall v_i \in V$, where x_{v_i} is the node attribute associated to v_i.

- *Message update*: The messages are then updated to obtain a summary of the neighborhood information:

$$h_{v_i}^{(q+1)} = U^{(q)}(h_{v_i}^{(q)}, m_{v_i}^{(q)}) \tag{2}$$

where $h_{v_i}^{(q+1)}$ is the updated message for node v_i at layer $q+1$ and U is the learnable update function. After running for Q layers of this message passing, the output of the final layer can be used as node embeddings, $z_{v_i} = h_{v_i}^{Q}$.

In practice, most of the GNNs models differ depending on how they perform the aggregation of messages from its neighbouring nodes. For a standard GCN, an output message for node v_i is computed as :

$$m_{v_i}^{(q)} = \sigma(\sum_{j \in \mathcal{N}_{v_i}} \alpha_{ij} h_{v_j}^{(q)}) \tag{3}$$

where σ is some activation function, α_{ij} is a factor that determines the importance of node features of node v_j to node v_i. In a standard message passing of GCN, the α_{ij} is often set to edge weights. However in a Graph attention network(GAT)[15], α_{ij} is defined implicitly using *self-attention mechanism* that computes the coefficients e_{ij} based on the features for each pair of nodes i and j in the graph:

$$e_{ij} = LeakyReLU(a^{\top}[\Theta h_{v_i} \parallel \Theta h_{v_j}]) \tag{4}$$

Where a is a learnable weight vector, and Θ is a learnable linear transformation matrix, usually estimated using a MLP and finally $LeakyReLU$ for imposing non-linearity. In order to be comparable across different neighbourhoods, these coefficients are normalised using the softmax function:

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{p \in \mathcal{N}_{v_i}} \exp(e_{ip})} \tag{5}$$

In similar manner, multiple α's can be computed in parallel where each attention head can automatically attend to various aspects of the input data. GAT models have significant advantages in terms of interpretation as the learnt attention scores are indicative of nodes (in our case brain regions) that are most influential in the propagation of information.

Graph Pooling

Graph pooling operations are used for generating coarsened representations of the given graphs in order to achieve graph-level tasks while preserving topological information. As detailed in the review [16], the approaches for graph pooling can fall into two main categories: *Flat* and *Hierarchical Pooling*. In this work, we use *mean pooling* and incorporate a

[3] https://github.com/Rohit3594/Deep_sulcal_graphs.git

learnable *Top-k pooling* strategy [17] for better interpretation. For a graph $\mathcal{G}$, the graph-level representation g is obtained through *mean pooling*, i.e. by averaging the node features across nodes of a graph: $g = \frac{1}{M}\sum_{v_i=1}^{M} z_{v_i}$.

The *Top-k* pooling operates by selecting the K nodes with the highest importance scores from original graph in order to create a coarsened representation of the graph. This selection of nodes is particularly interesting because it provides interpretable insights on most informative nodes in the sulcal graphs for the current learning task.

Formally, at a given pooling layer l, the *Top-k* pooling method is defined by computing:

$$\begin{aligned} y &= X^{(l)}\mathbf{p}^{(l)}/\|\mathbf{p}^{(l)}\|, \\ id &= rank(y,k), \\ \tilde{y} &= sigmoid(y(id)), \end{aligned}$$

Here, a scalar projection on attribute matrix $X^{(l)} \in \mathbb{R}^{M\times d}$ is computed using a learnable projection vector $\mathbf{p} \in \mathbb{R}^d$ which produces scalar projection scores $\tilde{y}$ corresponding to each node. k is a hyper-parameter defining the number of nodes to select. Next, a reduced attribute matrix corresponding to the selected nodes is computed for subsequent pooling or GNN layers until g is obtained.

3. EXPERIMENTS

Imaging data and sulcal graph extraction

This study relies on the data from 1090 healthy adults subjects from the S1200 release of the open dataset shared by the Human Connectome Project (HCP) [18]. Structural images were processed using the HCP structural processing pipeline, which has been described in [19]. The sulcal graphs were computed as detailed in [10]. We considered the following attributes as features for the nodes of the graphs: 1) *Sphere 3dcoords*: 3D coordinates of the deepest point in each sulcal basin. 2) *Basin area*: area of corresponding sulcal basin. 3) *Max depth*: sulcal depth of the deepest point in the sulcal basin. 4) *Basin mean depth*: average of the sulcal depth across all the vertices located in the sulcal basin. 5) *Basin var depth*: variance of the sulcal depth across all the vertices located in the sulcal basin. 6) *Basin int curvature*: integral across all the vertices located in the sulcal basin of the absolute value of the mean curvature. 7) *Pit thickness*: cortical thickness at the location of the deepest point. 8) *Basin mean thickness*: average of the cortical thickness across all the vertices located in the sulcal basin.

Classification task and Model description

As proposed in [2], we learn a representation space for sulcal graphs in a supervised setting consisting in the binary classification of the subjects according to their gender. The GAT architecture used in this work consists of a three-layer GAT network with double head attention in the first layer followed by single head in subsequent layers. Drawing inspiration from [12], the "Top-k" model in this work uses a two-layers GCN network. After each GCN layer, a top-k pooling operation is applied where we set the ratio $k = 0.5$, resulting in a 50% reduction of irrelevant nodes after each pooling layer. Finally, we also designed a baseline MLP model, that is also capable to learn the significance of attributes but independently from the topology. The MLP has two linear layers with ReLU activation that translates node attributes to node embeddings. For all these models, the node embeddings are subsequently processed by a mean pooling layer to obtain the graph representations, which is then processed through an additional learnable layer that performs a linear transformation to translate these graph representations g_n into class scores. All models are trained on the 1090 sulcal graphs from the left hemisphere, using a 10-fold cross validation framework in a fully-supervised manner. We define the loss using a softmax classification and negative log-likelihood :

$$\mathcal{L} = \sum_{n\in N_{train}} -\log(\text{softmax}(g_n, y_n)) \tag{6}$$

where N_{train} denotes the graphs in training set, and $\text{softmax}(g_n, y_n)$ denotes the predicted probability corresponding to class y_n :

$$\text{softmax}(g_n, y_n) = \sum_{i=1}^{C} y_n[i]\frac{e^{g_n^\top w_i}}{\sum_{j=1}^{C} e^{g_n^\top w_j}} \tag{7}$$

with w being the learnable parameters with C classes.

All models were implemented using the *pytorch-geometric*[4] package. We empirically fixed the embedding size in the models to 16-dimensional embedding. All models were trained for 100 epochs with Adam optimizer[20] and 0.01 as learning rate. Along with the comparison of GNN models, we also assess the relevance of sulcal graph attributes by constructing two different attribute sets. One constitutes of only 3D coordinates providing spatial information, while the second set corresponds to all the node attributes with both the spatial and the geometrical information. All models achieved the task of gender classification above chance throughout multiple trials.

4. RESULTS

We report the average and standard deviation across all folds of test scores in Table 1. We observe that the model GCN + Top-K outperforms both GAT and the baseline MLP in both experiments. Our interpretation is that the addition of the Top-k pooling operation improves the reduction of noise and integration of structural information. When comparing between the two experiments (and thus the two columns in the table), we observe that the enrichment of the attributes on nodes results in an increase in performance for all models. The gain

[4] https://pytorch-geometric.readthedocs.io/en/latest/index.html

Fig. 2. 2-dim t-SNE visualization for the different models at best folds. MLP(left), GAT(middle) and GCN+TopK(right)

is higher for MLP than for GAT. The lower performance of GAT which suggests that this model did not learn relevant information from the topology for this specific classification task.

Table 1. Average test accuracy across 10 folds

	Cross-val accuracy	
Model	**3D coords**	**3D Coords + Geometric features**
MLP	62 ± 4	76 ± 3
GAT	64 ± 3	69 ± 3
GCN + Top-K	65 ± 3	78 ± 2

To complement the quantitative results from Table 1, we visualize the 2-dim *t-SNE* for the whole population corresponding to the best models on Fig. 2. For the three methods, the two sub-populations are visually separable, which is consistent with Table 1. The spatial pattern obtained with the model GCN + Top-K suggests that the underlying structure of the information could be 1-dimensional, hence confirming the potential of this method.

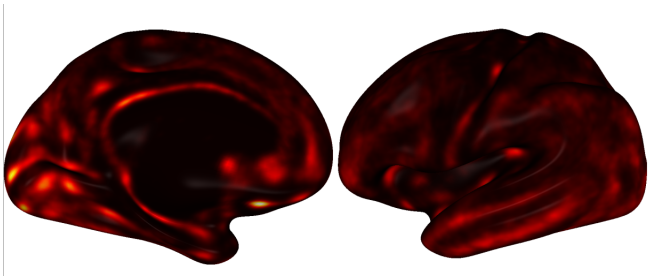

Fig. 3. For each node of each graph, we compute the mean attention from incoming edges to visualize the smooth map across all subjects.

To better interpret the underlying learning mechanism of the GAT model we visualize the attention map corresponding to last layer of the model on Fig. 3. The attention scores are computed for each incoming edge to a node in the graph, where a softmax score range the value between 0 and 1. We compute the mean attention by averaging these scores for each node. These scores are then summed across subjects

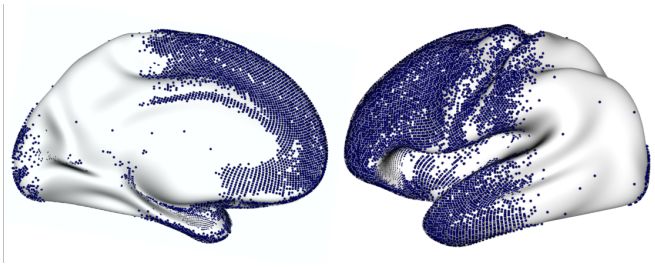

Fig. 4. Top-k nodes after the first layer of pooling operation.

to create a map showing the spatial distribution of the relevant information across cortical regions. The map obtained suggest that the GAT models automatically emphasize the regions located in deep cortical folds which are known to be less variable across individuals as pivotal to learn the global characteristics of the graphs. For GCN + TopK, we visualize on Fig.4 the top-k nodes extracted from the first pooling layer selected by the model that were the most relevant for the gender discrimination task. The most informative nodes for classifying the subjects according to their gender are concentrated in three large cortical regions: the frontal lobe, the occipital lobe and the anterior part of the temporal lobe. The spatial extent of informative regions is much larger than for GAT (Fig.3). In addition, these regions are highly consistent with the observations from [10]. This consistent localization across studies support the relevance of the representation learnt by the GCN + TopK method.

5. CONCLUSION

We explored the potential of recent geometric deep learning approaches to learn representations in the context of sulcal graphs, specifically focusing on the variations in cortical folding patterns analyzed through supervised gender classification task. We identified two methods having particularly attractive features in this context. Our experiments on the sulcal graphs from 1090 subjects confirmed the relevance of these approaches allowing better interpretation than a classical MLP.

6. ACKNOWLEDGMENT

Data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. The project leading to this publication has received funding from Excellence Initiative of Aix-Marseille University - A*MIDEX, a french "Investissements d'Avenir" programme (AMX-19-IET-002). The research leading to these results has also been supported by the ANR SulcalGRIDS Project, Grant ANR-19-CE45-0014 and the ERA-NET NEURON MULTI-FACT Project, Grant ANR-21-NEU2-0005 funded by the French National Research Agency.

7. COMPLIANCE WITH ETHICAL STANDARDS

This research study was conducted retrospectively using human subject data made available in open access by `https://humanconnectome.org`. Additional ethical approval specific to the current study was not required as confirmed by the license attached with the open access data.

8. REFERENCES

[1] Geert Litjens and Thijs Kooi et al., "A survey on deep learning in medical image analysis," *Medical image analysis*, vol. 42, pp. 60–88, 2017.

[2] Anees Abrol and Zening Fu et al., "Deep learning encodes robust discriminative neuroimaging representations to outperform standard machine learning," *Nature communications*, vol. 12, no. 1, pp. 353, 2021.

[3] Léonie Borne and Denis Rivière et al., "Automatic labeling of cortical sulci using patch-or cnn-based segmentation techniques combined with bottom-up geometric constraints," *Medical Image Analysis*, vol. 62, pp. 101651, 2020.

[4] Louise Guillon and Cagna et al., "Detection of abnormal folding patterns with unsupervised deep generative models," in *MLCN*. Springer, 2021, pp. 63–72.

[5] Irina Higgins and Loic Matthey et al., "beta-vae: Learning basic visual concepts with a constrained variational framework," in *ICLR*, 2016.

[6] Joël Chavas and Guillon et al., "Unsupervised representation learning of cingulate cortical folding patterns," in *MICCAI*. Springer, 2022, pp. 77–87.

[7] Ting Chen and Simon Kornblith et al, "Simclr: A simple framework for contrastive learning of visual representations," in *ICLR*, 2020, vol. 2.

[8] Christopher Hammill, Jason P Lerch, Margot J Taylor, Stephanie H Ameis, M Mallar Chakravarty, Peter Szatmari, Evdokia Anagnostou, and Meng-Chuan Lai, "Quantitative and qualitative sex modulations in the brain anatomy of autism," *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, vol. 6, no. 9, pp. 898–909, 2021.

[9] Hannah Kiesow, Robin IM Dunbar, Joseph W Kable, Tobias Kalenscher, Kai Vogeley, Leonhard Schilbach, Andre F Marquand, Thomas V Wiecki, and Danilo Bzdok, "10,000 social brains: sex differentiation in human brain anatomy," *Science advances*, vol. 6, no. 12, pp. eaaz1170, 2020.

[10] Sylvain Takerkart and Auzias et al., "Structural graph-based morphometry: A multiscale searchlight framework based on sulcal pits," *Medical Image Analysis*, vol. 35, pp. 32–45, Jan. 2017.

[11] Alaa et al. Bessadok, "Graph neural networks in network neuroscience," *IEEE PSPB*, vol. 45, no. 5, pp. 5833–5848, 2022.

[12] Xiaoxiao Li and Zhou et al., "Braingnn: Interpretable brain graph neural network for fmri analysis," *Medical Image Analysis*, vol. 74, pp. 102233, 2021.

[13] Jeremy Kawahara and et al., "Brainnetcnn: Convolutional neural networks for brain networks; towards predicting neurodevelopment," *NeuroImage*, vol. 146, pp. 1038–1049, 2017.

[14] Hejie Cui and Wei Dai et al., "Braingb: a benchmark for brain network analysis with graph neural networks," *IEEE Transactions on Medical Imaging*, 2022.

[15] Petar Velickovic and Guillem Cucurull et al., "Graph attention networks," *stat*, vol. 1050, no. 20, pp. 10–48550, 2017.

[16] Chuang Liu and et al., "Graph pooling for graph neural networks: Progress, challenges, and opportunities," in *IJCAI 2023*, 2023, pp. 6712–6722.

[17] Hongyang Gao and Shuiwang Ji, "Graph u-nets," in *international conference on machine learning*. PMLR, 2019, pp. 2083–2092.

[18] David C. et al. Van Essen, "The Human Connectome Project: A data acquisition perspective," vol. 62, no. 4, 2012, NeuroImage.

[19] Matthew F. et al Glasser, "The minimal preprocessing pipelines for the Human Connectome Project," *NeuroImage*, vol. 80, pp. 105–124, Oct. 2013.

[20] Diederik P et al. Kingma, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.