Diffusion-based Perturbation Modeling for Unsupervised Detection of Sulcal Pattern Alterations

Hyeokjin Kwon[1,2 *] [0009-0005-6848-6670], Sarah U. Morton[2], Jane W. Newburger[2], Jong-Min Lee[3], P. Ellen Grant[1,2], and Kiho Im[1,2 *] [0000-0002-9239-8970]

[1] Fetal Neonatal Neuroimaging and Developmental Science Center, Boston, MA, USA
[2] Boston Children's Hospital and Harvard Medical School, Boston, MA, USA
[3] Hanyang University, Seoul, Republic of Korea
{hyeokjin.kwon, kiho.im}@childrens.harvard.edu

Abstract. Prenatal brain development involves rapid cortical growth and folding, and disruptions in sulcal patterns have been linked to neurodevelopmental disorders and long-term functional impairments. The spatial organization of early sulcal folds remains relatively stable over development, making sulcal patterns a promising substrate for detecting developmental deviations. In this work, we propose SynthSPA, an unsupervised framework for detecting sulcal abnormalities based on normal data. SynthSPA leverages latent diffusion-based generative modeling to synthesize diverse pseudo-abnormal (PA) sulcal pattern graphs through controlled perturbations of normal graphs and learns to discriminate them from normative samples. The synthesized PA graphs are perturbed while remaining close to the true normative graphs, preserving the overall organization while introducing subtle deviations. This enables SynthSPA to learn fine-grained decision boundaries that are sensitive to mild and heterogeneous sulcal abnormalities. Furthermore, to capture the unique geometric and topological properties of sulcal patterns, SynthSPA incorporates a newly proposed heterophily-aware variational graph autoencoder (HA-VGAE), tailored to low-homophily graph structures. We train SynthSPA on publicly available healthy brain magnetic resonance imaging datasets (total n=1,542), and evaluate it on congenital heart disease (CHD), the most common birth defect associated with altered brain development. Experimental results demonstrate that SynthSPA effectively captures subtle and heterogeneous deviations and outperforms existing baselines across hemispheres and CHD subtypes. To our knowledge, this work represents the first application of diffusion-based anomaly detection to sulcal pattern analysis. Code is available at: https://github.com/hookhy/SynthSPA

Keywords: Cortical Sulcal Patterns, Unsupervised Graph Anomaly Detection, Congenital Heart Disease

1 Introduction

Prenatal brain maturation is a dynamic process, with rapid cortical growth and folding beginning in the second trimester [1]. Disruptions in sulcal folding patterns are linked

to developmental disorders and long-term functional impairments [2]. The spatial organization of early sulcal folds, shaped by functional specialization, cortical connectivity, and genetic influences, remains stable with age [3-5]. Thus, sulcal patterns provide a robust substrate for modeling early neurodevelopment and detecting deviations [6].

Previous studies have used quantitative graph-matching method to characterize sulcal pattern alterations by representing sulcal pits as graph nodes and comparing individual graphs to predefined reference sets [7, 8]. Despite this framework has revealed sulcal abnormalities across diverse conditions, it requires computationally intensive pairwise comparisons, limiting the size and diversity of feasible reference sets. Consequently, prior studies have relied on small and cohort-specific reference sets, constraining scalability and generalizability. In addition, it has limited sensitivity for detecting localized patterns. Collectively, these challenges restrict sensitivity to individual deviations, particularly subtle and spatially localized abnormalities.

Motivated by the recent success of graph neural networks (GNNs), a supervised model has been proposed to overcome these limitations [9]. The method employs a prototype-based GNN to classify sulcal pattern graphs into disease and control groups by learning representative atypical sulcal prototypes. Notably, this approach represents the first application of deep learning to sulcal pattern graphs. However, it relies on a fully supervised framework, requiring a substantial number of labeled samples for both disease and control classes. In contrast, prior graph-matching methods quantify sulcal normality derived exclusively from healthy reference sets. Together, this motivates the development of unsupervised methods that rely solely on normal data.

A wide range of unsupervised anomaly detection approaches have been proposed, including kernel-based methods [10], one-class classification [11, 12], and reconstruction-based techniques [13]. These methods detect anomalies in unseen cases by learning a normative distribution from normal data. Despite their success across domains, their sensitivity can degrade when anomalies are subtle or intricate, as models may over-generalize and blur the boundary between normal and abnormal. Accordingly, medical applications have largely centered on focal, high-contrast anomalies (e.g., tumor), where deviations are visually salient and less susceptible to boundary blurring. This may limit suitability for sulcal pattern analysis, which is characterized by substantial inter-individual variability and requires modeling a broad spectrum of deviations ranging from subtle geometric distortions to large-scale topological alterations.

Recently, diffusion generative modeling has emerged as a powerful paradigm for unsupervised anomaly detection by providing a principled mechanism for capturing high-dimensional variability [14-16]. Among these approaches, Anomalous Graph Diffusion (AG-Diff) was proposed to generate pseudo-abnormal (PA) graphs that closely resemble normal data while introducing controlled perturbations [16]. By synthesizing diverse and subtle deviations and jointly training a classifier to distinguish them from normal samples, this enables effective anomaly detection without requiring explicit abnormal labels. Building upon this paradigm, we introduce SynthSPA, which uses the diffusion to synthesize sulcal PA graphs spanning a wide range of controlled alterations and a GNN classifier to distinguish them from the original ones. Crucially, SynthSPA incorporates a newly proposed heterophily-aware variational graph autoencoder (HA-VGAE) to capture the geometric and topological characteristics of sulcal pattern

graphs, enabling structure-preserving perturbations in the latent space. To evaluate the practical utility of the synthesized abnormal patterns, we train the model on public datasets and validate its anomaly detection capabilities using a clinical cohort of congenital heart disease (CHD), the most common birth defect and a condition associated with disrupted brain development and functional impairments [17, 18]. Ablation studies and comparisons with baseline methods demonstrate that the proposed model provides a sensitive solution for detecting sulcal abnormalities.

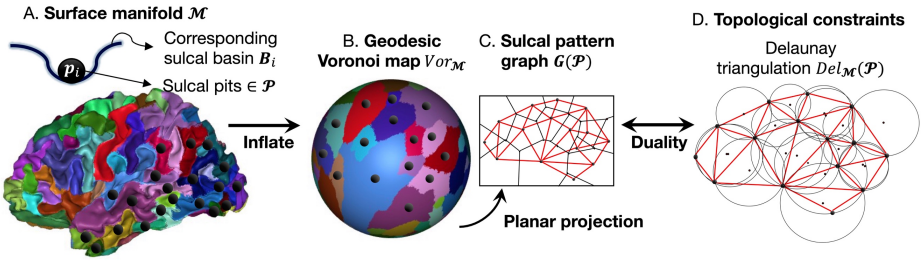

Fig. 1. Sulcal pattern graphs and topological constraints

2 Method

2.1 Preliminary: Sulcal Pattern Graphs

T1-weighted brain magnetic resonance images (T1w MRIs) were processed to extract left and right white matter surfaces using FreeSurfer v5.3.0 [19]. To ensure spatial consistency, individual surfaces were mapped to a standardized Talairach space using a linear transformation. Sulcal pits $\{p_i\}_{i=1}^{N} \in \mathcal{P}$ and their corresponding basins $\{B_i\}_{i=1}^{N} \in \mathcal{B}$ were identified from sulcal depth maps via watershed segmentation (Fig 1A) [20]. A sulcal pattern graph $G = (\mathcal{P}, E)$ was built by connecting pits with an edge $(p_i, p_j) \in E$ if their basins were adjacent. Each node was associated with a geometric feature $\boldsymbol{x}_i$, comprising sulcal area, depth, and 3D coordinates.

Topological Constraints. Each edge in $\boldsymbol{G}$ represents geodesic adjacency between sulcal basins on the cortical surface manifold $\boldsymbol{\mathcal{M}}$. Under this construction, the sulcal pattern graph is topologically equivalent to the geodesic Delaunay triangulation of the pits $\boldsymbol{\mathcal{P}}$, by the duality between Voronoi diagrams and Delaunay triangulations (Fig. 1B-D) [21]. On a genus-0 closed 2-manifold (topologically a sphere), such graphs satisfy $|\boldsymbol{E}| \leq \boldsymbol{3N - 6}$, implying an average node degree below six [22]. These properties impose necessary topological consistency constraints on sulcal pattern graphs.

Geometric Constraints. Beyond topology, sulcal graphs exhibit heterogeneous geometric feature behaviors. The 3D positions of pits vary smoothly along the graph struc-

ture, such that adjacent nodes are spatially proximate. In contrast, morphological attributes such as sulcal area and depth may differ substantially between neighboring pits. Consequently, sulcal graphs exhibit a geometrically low-homophily structure, characterized by locally homophilic spatial features but heterophilic morphological attributes.

2.2 SynthSPA

An overview of the proposed SynthSPA framework is illustrated in Fig. 2. SynthSPA models normality via variational inference pre-trained on normal graphs. Latent representations extracted by the encoder are subsequently perturbed through a conditioning module and a latent diffusion model to generate PA samples that resemble the original ones but exhibit controllable perturbations. These embeddings are then decoded to reconstruct corresponding PA graphs. Finally, an anomaly classifier is trained to distinguish between original and PA graphs, enabling the model to learn a decision boundary for normative sulcal patterns without requiring labeled abnormal data.

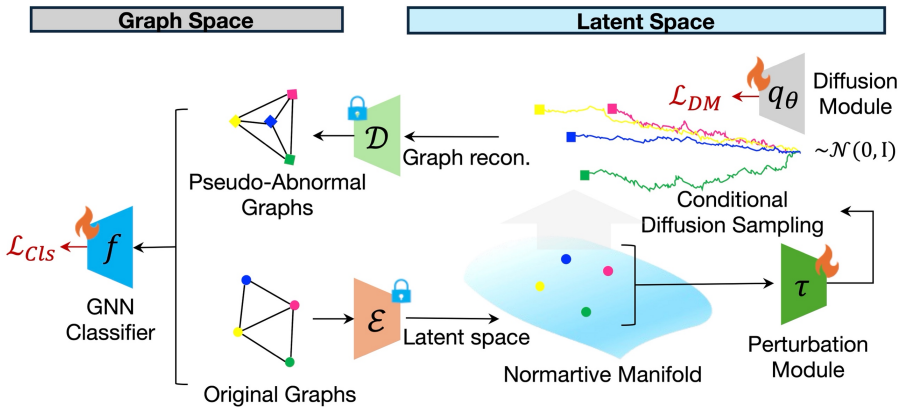

Fig. 2. Architecture of the proposed SynthSPA

Variational Inference. Given a graph, an encoder maps node features to a latent space, from which a decoder reconstructs graph structure and attributes. While standard (variational) graph autoencoders (GAEs and VGAEs) have been widely used, their underlying assumptions are not well aligned with the sulcal graphs [23]. To address this, we propose a HA-VGAE, which introduces modifications in both the encoder and decoder.

Encoder. Conventional VGAE encoders rely on message-passing layers, which can be interpreted as low-pass graph filters under a homophily assumption [24]. Such behavior may suppress high frequency, heterophilic sulcal features (e.g., sulcal area and depth), leading to suboptimal latent representations. To overcome this issue, HA-VGAE adopts a Beta-wavelet GNN encoder (BWenc), which captures multi-band spectral information through a combination of low-pass and band-pass filters [25]. Given a feature

vector $\boldsymbol{x}$, the Beta-wavelet transform constructed by wavelets $\{\mathcal{W}_{i,C-i}\}_{i=0}^{C}$ is defined as:

$$\mathcal{W}_{i,C-i}(\boldsymbol{x}) = \widehat{\boldsymbol{\beta}}_{i,C-i}(\boldsymbol{L})\boldsymbol{x} = 1/2\boldsymbol{\beta}_{i,C-i}(\boldsymbol{L}/2) \tag{1}$$

where $\widehat{\boldsymbol{\beta}}_{i,C-i}(\boldsymbol{L})$ is an adjusted Beta distribution over the spectral range $\lambda \in [0,2]$, and $\boldsymbol{L}$ denotes the normalized graph Laplacian. BWenc computes hidden representations, and the mean and log-variance of the latent variables are obtained by:

$$\boldsymbol{H}_i = \mathcal{W}_{i,C-i}(MLP(\boldsymbol{X})) \tag{2}$$

$$\boldsymbol{Z}_{\boldsymbol{\mu}}, \boldsymbol{log}(\boldsymbol{Z}_{\boldsymbol{\sigma}}) = MLP([\boldsymbol{H}_0; \boldsymbol{H}_1; \dots; \boldsymbol{H}_C]) \tag{3}$$

Latent embeddings $\boldsymbol{Z}$ are then sampled using the standard reparameterization trick.

Decoder. Standard VGAEs typically employ an inner-product edge decoder (IPdec) and a multi-layer perceptron (MLP)-based node decoder:

$$\widehat{\boldsymbol{A}} = \sigma(\boldsymbol{Z}\boldsymbol{Z}^T), \widehat{\boldsymbol{X}} = MLP(\boldsymbol{Z}) \tag{4}$$

where $\widehat{\boldsymbol{A}} \in \{0,1\}^{N\times N}$ is a reconstructed adjacency matrix and $\sigma(\cdot)$ denotes the sigmoid function. However, the IPdec implicitly favors homophilic structures and may be ill-suited for reconstructing low-homophily sulcal graphs. To mitigate this bias, HA-VGAE replaces it with an edge prediction decoder (EPdec) that estimates the probability of an edge between nodes i and j as:

$$\boldsymbol{p}_{i,j} = \sigma(MLP([\boldsymbol{Z}_i; \boldsymbol{Z}_j])) \tag{5}$$

This formulation allows the decoder to model heterogeneous edges explicitly. For node reconstruction, we retain the standard MLP-based decoder.

PA Graph Generation. Given the structured latent embeddings learned by HA-VGAE, a latent diffusion model synthesizes PA node embeddings through controlled perturbations. Gaussian noise ϵ is progressively added to $\boldsymbol{Z}_0 = \boldsymbol{Z}$ over T steps, and a neural network q_θ denoises the latent variables conditioned on the diffusion time step t and a perturbation condition $\boldsymbol{c}$:

$$d\boldsymbol{Z}/dt = (\boldsymbol{Z}_t - q_\theta(\boldsymbol{Z}_t, t, \boldsymbol{c}))/t \tag{6}$$

The perturbation condition $\boldsymbol{c}$ is generated by a conditioning module τ to introduce controlled degradations:

$$\boldsymbol{c} = \tau(\boldsymbol{Z}_0) = MLP(\boldsymbol{Z}_0 + \zeta) \tag{7}$$

where $\zeta \sim \mathcal{N}(\boldsymbol{0}, \rho\boldsymbol{I})$. The resulting embeddings are decoded by the trained HA-VGAE decoder to reconstruct PA graphs.

Anomaly Detection. A graph-level GNN classifier is trained to distinguish between original graphs and the generated PA graphs, thereby learning a decision boundary for normality. For each input graph $\boldsymbol{G}$, the classifier outputs a logit $\boldsymbol{h_G}$, which serves as the anomaly score. We employ an edge-weighted graph attention network (EGAT) [26] operating on the reconstructed node feature $\widehat{\boldsymbol{X}}$ and the edge probability $\boldsymbol{P} = \{\boldsymbol{p_{i,j}}\}_{\boldsymbol{i,j=0}}^{\boldsymbol{N-1}}$ predicted by the HA-VGAE decoder:

$$h_G = MLP(f(\widehat{\boldsymbol{P}}, \widehat{\boldsymbol{X}})) \tag{8}$$

To suppress spurious long-range connections that are inconsistent with sulcal geometry, we apply a spatial mask $\boldsymbol{M}$ to the edge probabilities:

$$\widehat{\boldsymbol{P}} = \boldsymbol{M} \odot \boldsymbol{P}, \boldsymbol{M}_{i,j} = \mathbb{1}(d(i,j) \leq \psi), \forall i,j \in \mathcal{P} \tag{9}$$

$d(i,j)$ denotes the Euclidean distance between pits i and j, $\mathbb{1}(\cdot)$ is the indicator function, and ψ is an empirically chosen threshold. The EGAT $f(\cdot)$ consist of L layers.

Learning Objectives. We first pre-train the HA-VGAE using variational inference on normal sulcal graphs. The objective consists of a node feature reconstruction loss, an edge reconstruction loss, and a Kullback-Leibler (KL) divergence regularization:

$$\mathbb{E}_i[\left\|\boldsymbol{X}_i - \widehat{\boldsymbol{X}}_i\right\|_F^2 + BCE\big(\boldsymbol{A}_i, \sigma(\boldsymbol{P}_i)\big) + KL(q(\boldsymbol{Z}_i; \boldsymbol{X}_i, \boldsymbol{A}_i)|p(\boldsymbol{Z})] \tag{10}$$

where $BCE(\cdot)$ denotes the binary cross-entropy loss and the KL regularizes the posterior distribution toward the normal prior $p(\boldsymbol{Z})$. After pre-training, the diffusion model and the anomaly classifier are jointly optimized. The objective combines a binary classification loss for anomaly detection with a score-matching loss for PA generation:

$$\mathcal{L} = \mathbb{E}_i[BCE\big(y_i, \sigma(h_{G_i})\big)] + \lambda_{DM}\mathbb{E}_{\boldsymbol{Z}_0}\mathbb{E}_{\boldsymbol{Z}_t}\|\epsilon - q_\theta(\boldsymbol{Z}_t, t, \boldsymbol{c})\|_2^2 \tag{11}$$

$y_i \in \{0,1\}$ denotes the graph label (0 for original and 1 for PA graphs), and λ_{DM} controls the balance between the two terms.

3 Experiments and Results

MRI Dataset. To train the SynthSPA, we used a retrospective dataset of T1w MRIs from three publicly available healthy datasets (total n=1,542 age: 5-35 years): Human Connectome Project (male/female: 332/452) [27], Child Mind Institute (341/339) [28], and Beijing Normal University (26/52) [29]. For evaluation, we used a retrospective dataset of T1-weighted MRIs (age: 10-19 years) from 248 patients (164/84) with CHD subtypes (single ventricle (SV; 67/48), transposition of the great arteries (TGA; 70/22), and tetralogy of Fallot (ToF; 27/14)) and 94 healthy controls (46/48) from previous studies and collected at Boston Children's Hospital (BCH). These studies were approved by the BCH institutional review board and written informed consent was obtained for all MRIs.

Baselines. We compared the proposed method against well-established unsupervised baselines. These included (1) graph-agnostic methods, namely Local Outlier Factor (LOF) [30], Isolation Forest (iF) [31], and one-class support vector machine (OCSVM) [32], and (2) graph reconstruction-based methods, including the GAE and VGAE. In addition, our HA-VGAE was evaluated as a reconstruction-based baseline. For graph-agnostic methods, each graph was represented either by a Weisfeiler-Lehman (WL) kernel or by mean embeddings extracted from the HA-VGAE encoder (Deep). For reconstruction-based methods, graph-level anomaly scores were computed from the average reconstruction errors on nodes (n), edges (e), and their combination (c).

Implementation Details. All models were implemented in PyTorch on Ubuntu 24.04 LTS and trained using NVIDIA RTX A5000 GPUs. Training samples were used, with 95% for training and 5% for validation. We employed the Adam optimizer with early stopping, an initial learning rate (LR) of 5e-3 decayed every 50 epochs. The batch size, minimum LR, decay factor γ, and maximum epochs were set to 24, 1e-4, 0.5, and 200, respectively. The BWenc used 3 wavelets (C=2) with a hidden dimension of 128. The diffusion process employed T=1000 time-steps, with λ_{DM} set to 0.1 and perturbation noise scale ρ=5. The EGAT classifier consisted of L=3 layers with a hidden dimension of 16, and an empirical threshold ψ=3 was applied to mask edge probabilities. All MLP-based modules used two-layer architectures. Performance was evaluated using the area under the receiver operating characteristic curve (AUC) on left (LH) and right (RH) hemispheric sulcal patterns across CHD subtypes.

Table 1. Comparison results. Results are reported as mean±standard deviation over four random seeds. The best/second best results are indicated in **boldface**/underlined.

Model	LH				RH			
	SV	TGA	ToF	All	SV	TGA	ToF	All
WL-LOF	0.53±0.01	0.54±0.03	0.47±0.02	0.52±0.02	0.49±0.02	0.46±0.04	0.52±0.04	0.49±0.03
WL-iF	0.50±0.01	0.49±0.02	0.43±0.01	0.48±0.01	0.51±0.01	0.45±0.03	0.46±0.04	0.48±0.01
WL-OCSVM	0.44±0.00	0.67±0.00	0.71±0.00	0.57±0.00	0.45±0.00	0.69±0.00	**0.76±0.00**	0.59±0.00
Deep-LOF	0.52±0.00	0.56±0.01	0.57±0.00	0.54±0.00	0.51±0.01	0.49±0.00	0.57±0.01	0.51±0.00
Deep-iF	0.52±0.02	0.55±0.02	0.57±0.02	0.54±0.02	0.50±0.01	0.54±0.03	0.61±0.01	0.54±0.02
Deep-OCSVM	0.52±0.00	0.56±0.00	0.56±0.00	0.54±0.00	0.51±0.00	0.49±0.00	0.59±0.00	0.51±0.00
GAE-n	0.64±0.02	0.53±0.03	0.52±0.02	0.58±0.02	**0.63±0.03**	0.49±0.01	0.52±0.02	0.56±0.01
VGAE-n	0.58±0.02	0.50±0.07	0.53±0.06	0.54±0.04	0.56±0.02	0.48±0.04	0.53±0.04	0.53±0.03
HA-VGAE-n	0.52±0.05	0.39±0.05	0.45±0.04	0.46±0.04	0.60±0.04	0.43±0.05	0.45±0.05	0.51±0.04
GAE-e	0.57±0.04	0.64±0.04	0.71±0.06	0.62±0.04	0.52±0.02	0.60±0.03	0.67±0.02	0.58±0.02
VGAE-e	0.53±0.06	0.44±0.03	0.46±0.01	0.48±0.03	0.57±0.02	0.46±0.02	0.48±0.05	0.51±0.02
HA-VGAE-e	0.62±0.01	0.65±0.04	0.59±0.08	0.63±0.03	0.53±0.02	0.50±0.06	0.53±0.08	0.52±0.03
GAE-c	0.61±0.03	0.62±0.04	0.68±0.06	0.63±0.03	0.58±0.03	0.56±0.03	0.62±0.03	0.58±0.02
VGAE-c	0.54±0.06	0.44±0.02	0.46±0.01	0.49±0.03	0.58±0.02	0.45±0.01	0.48±0.04	0.51±0.02
HA-VGAE-c	0.60±0.02	0.56±0.06	0.56±0.07	0.58±0.04	0.56±0.01	0.48±0.04	0.49±0.02	0.52±0.01
SynthSPA (*ours*)	**0.64±0.01**	**0.73±0.03**	**0.73±0.03**	**0.69±0.01**	0.54±0.04	**0.75±0.01**	0.73±0.00	**0.65±0.02**

Comparison Results. SynthSPA consistently outperformed baselines (Table 1). In LH, SPA-Diff achieved the highest AUC for all subtypes. In RH, SynthSPA achieved the best performance for TGA and the full CHD group and the second-best performance for ToF, while remaining competitive for SV. Among the baselines, graph-kernel methods, particularly WL-OCSVM, showed relatively strong performance. Their elevated AUC values for TGA and ToF suggest that these subtypes may involve prominent topology-driven alterations. Notably, SynthSPA outperformed WL-OCSVM in most settings, indicating that it captures such topological deviations while offering increased sensitivity to complex sulcal variations. Reconstruction-based baselines, especially GAE-n models, achieved comparatively higher AUCs for SV, suggesting that SV-related abnormalities may be dominated by localized geometric perturbations rather than global topological changes. HA-VGAE consistently outperformed standard GAE and VGAE variants. In particular, HA-VGAE-e achieved higher LH AUC across most subtypes, highlighting the importance of accurate modeling for detecting subtle structural deviations in sulcal graphs.

Table 2. An ablation experiments of each component of the HA-VGAE

Encoder	Decoder	LH			RH		
		AUC	mAD	mE	AUC	mAD	mE
GraphSage	IPdec	0.62±0.03	53.10	-1949.88	0.55±0.01	53.11	-1967.74
GraphSage	EPdec	0.65±0.02	7.38	-35.84	0.61±0.06	7.33	-40.77
BWenc	IPdec	0.60±0.03	53.87	-1982.00	0.56±0.03	53.93	-2002.25
BWenc	EPdec	0.69±0.01	6.01	2.77	0.65±0.02	5.86	7.01

Ablation Study. Table 2 summarizes ablation results for the proposed HA-VGAE components. To assess the quality of generated PA graphs, we measured the mean average degree (mAD) and the mean edge margin $\mathrm{mE} = mean(3N - 6 - |E|)$, reflecting topological consistency. Using the proposed EPdec produced PA graphs with Euler statistics (mAD≈6 and mE≈0) consistent with valid graphs. These topological characteristics of the synthesized graphs demonstrate that the model can generate structurally plausible patterns, thereby validating the utility of the proposed framework. In contrast, models trained with a conventional IPdec yielded suboptimal AUC, inflated mAD>50 and reduced mE<-1900, indicating a failure to preserve sulcal graph topology. Using a message passing GraphSAGE encoder [33] resulted in consistently lower AUCs, highlighting the importance of multi-band encoding for sulcal features.

4 Conclusion

We presented an unsupervised graph anomaly detection framework for sulcal patterns. By leveraging diffusion model to generate PA graphs, it effectively captures subtle and heterogeneous deviations that are difficult to detect using existing approaches. Future work will focus on scaling the framework to larger datasets to further improve robustness and generalization. Extending the proposed approach to other neurodevelopmental conditions will also be important for assessing its applicability. Overall, the proposed

method provides a generalizable and efficient tool for sulcal pattern analysis, with strong potential to support future studies of neurodevelopment and risk stratification.

Acknowledgments. This research was supported by the National Institute of Neurological Disorders and Stroke (R01NS114087), and the Thrasher Research Fund through the Thrasher Early Career Award Program.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Rakic, P.: Evolution of the neocortex: a perspective from developmental biology. Nature Reviews Neuroscience 10, 724-735 (2009)
2. Bae, B.-I., Tietjen, I., Atabay, K.D., Evrony, G.D., Johnson, M.B., Asare, E., Wang, P.P., Murayama, A.Y., Im, K., Lisgo, S.N.: Evolutionarily dynamic alternative splicing of GPR56 regulates regional cerebral cortical patterning. Science 343, 764-768 (2014)
3. O'Leary, D.D., Chou, S.-J., Sahara, S.: Area patterning of the mammalian cortex. Neuron 56, 252-269 (2007)
4. Im, K., Grant, P.E.: Sulcal pits and patterns in developing human brains. NeuroImage 185, 881-890 (2019)
5. Rakic, P.: Genetic control of cortical convolutions. Science 303, 1983-1984 (2004)
6. Meng, Y., Li, G., Lin, W., Gilmore, J.H., Shen, D.: Spatial distribution and longitudinal development of deep cortical sulcal landmarks in infants. Neuroimage 100, 206-218 (2014)
7. Im, K., Pienaar, R., Lee, J.-M., Seong, J.-K., Choi, Y.Y., Lee, K.H., Grant, P.E.: Quantitative comparison and analysis of sulcal patterns using sulcal graph matching: a twin study. Neuroimage 57, 1077-1086 (2011)
8. Tarui, T., Madan, N., Graham, G., Kitano, R., Akiyama, S., Takeoka, E., Reid, S., Yun, H.J., Craig, A., Samura, O.: Comprehensive quantitative analyses of fetal magnetic resonance imaging in isolated cerebral ventriculomegaly. NeuroImage: Clinical 37, 103357 (2023)
9. Kwon, H., Son, S., Morton, S.U., Wypij, D., Cleveland, J., Rollins, C.K., Huang, H., Goldmuntz, E., Panigrahy, A., Thomas, N.H.: Graph-based prototype inverse-projection for identifying cortical sulcal pattern abnormalities in congenital heart disease. Medical image analysis 102, 103538 (2025)
10. Shervashidze, N., Schweitzer, P., Van Leeuwen, E.J., Mehlhorn, K., Borgwardt, K.M.: Weisfeiler-lehman graph kernels. Journal of Machine Learning Research 12, (2011)
11. Qiu, C., Kloft, M., Mandt, S., Rudolph, M.: Raising the bar in graph-level anomaly detection. arXiv preprint arXiv:2205.13845 (2022)
12. Ruff, L., Vandermeulen, R., Goernitz, N., Deecke, L., Siddiqui, S.A., Binder, A., Müller, E., Kloft, M.: Deep one-class classification. In: International conference on machine learning, pp. 4393-4402. Proceedings of Machine Learning Research 80, (2018)
13. Kim, S., Lee, S.Y., Bu, F., Kang, S., Kim, K., Yoo, J., Shin, K.: Rethinking reconstruction-based graph-level anomaly detection: limitations and a simple remedy. Advances in Neural Information Processing Systems 37, 95931-95962 (2024)

14. Behrendt, F., Bhattacharya, D., Krüger, J., Opfer, R., Schlaefer, A.: Patched diffusion models for unsupervised anomaly detection in brain mri. In: Medical Imaging with Deep Learning, pp. 1019-1032. Proceedings of Machine Learning Research 227, (2023)
15. Karras, T., Aittala, M., Aila, T., Laine, S.: Elucidating the design space of diffusion-based generative models. Advances in neural information processing systems 35, 26565-26577 (2022)
16. Cai, J., Zhang, Y., Liu, F., Ng, S.-K.: Leveraging Diffusion Model as Pseudo-Anomalous Graph Generator for Graph-Level Anomaly Detection. In: Forty-second International Conference on Machine Learning. Proceedings of Machine Learning Research 267, (2025)
17. Marelli, A., Miller, S.P., Marino, B.S., Jefferson, A.L., Newburger, J.W.: Brain in congenital heart disease across the lifespan: the cumulative burden of injury. Circulation 133, 1951-1962 (2016)
18. Wernovsky, G.: Current insights regarding neurological and developmental abnormalities in children and young adults with complex congenital cardiac disease. Cardiology in the Young 16, 92-104 (2006)
19. Fischl, B.: FreeSurfer. Neuroimage 62, 774-781 (2012)
20. Im, K., Pienaar, R., Paldino, M.J., Gaab, N., Galaburda, A.M., Grant, P.E.: Quantification and discrimination of abnormal sulcal patterns in polymicrogyria. Cerebral cortex 23, 3007-3015 (2013)
21. Joe, B., Wang, C.A.: Duality of constrained Voronoi diagrams and Delaunay triangulations. Algorithmica 9, 142-155 (1993)
22. Diestel, R.: Graph theory. Springer Nature (2025)
23. Kipf, T.N., Welling, M.: Variational graph auto-encoders. arXiv preprint arXiv:1611.07308 (2016)
24. Wu, F., Souza, A., Zhang, T., Fifty, C., Yu, T., Weinberger, K.: Simplifying graph convolutional networks. In: International conference on machine learning, pp. 6861-6871. Proceedings of Machine Learning Research 97, (2019)
25. Tang, J., Li, J., Gao, Z., Li, J.: Rethinking graph neural networks for anomaly detection. In: International conference on machine learning, pp. 21076-21089. Proceedings of Machine Learning Research 162, (2022)
26. Monninger, T., Schmidt, J., Rupprecht, J., Raba, D., Jordan, J., Frank, D., Staab, S., Dietmayer, K.: Scene: Reasoning about traffic scenes using heterogeneous graph neural networks. IEEE Robotics and Automation Letters 8, 1531-1538 (2023)
27. Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E., Yacoub, E., Ugurbil, K., Consortium, W.-M.H.: The WU-Minn human connectome project: an overview. Neuroimage 80, 62-79 (2013)
28. Alexander, L.M., Escalera, J., Ai, L., Andreotti, C., Febre, K., Mangone, A., Vega-Potler, N., Langer, N., Alexander, A., Kovacs, M.: An open resource for transdiagnostic research in pediatric mental health and learning disorders. Scientific data 4, 1-26 (2017)
29. Yan, C., Gong, G., Wang, J., Wang, D., Liu, D., Zhu, C., Chen, Z.J., Evans, A., Zang, Y., He, Y.: Sex-and brain size–related small-world structural cortical networks in young adults: a DTI tractography study. Cerebral cortex 21, 449-458 (2011)
30. Breunig, M.M., Kriegel, H.-P., Ng, R.T., Sander, J.: LOF: identifying density-based local outliers. In: Proceedings of the 2000 ACM SIGMOD international conference on Management of data, pp. 93-104. (2000)
31. Liu, F.T., Ting, K.M., Zhou, Z.-H.: Isolation forest. In: 2008 eighth ieee international conference on data mining, pp. 413-422. IEEE, (2008)

32. Noumir, Z., Honeine, P., Richard, C.: On simple one-class classification methods. In: 2012 IEEE International Symposium on Information Theory Proceedings, pp. 2022-2026. IEEE, (2012)
33. Hamilton, W., Ying, Z., Leskovec, J.: Inductive representation learning on large graphs. Advances in neural information processing systems 30, (2017)