VCDA-Net: Graph-Guided Voxel-Cluster Density Attention for Global-Regional Brain Age Gap Estimation

Quoc-Tien Nguyen[1 [0009−0004−0557−4123]], Myungeun Lee[2 [0000−0001−5032−0086]], Soo-Hyung Kim[1 [0000−0003−3575−5035]], and Hyung-Jeong Yang[1 [0000−0003−3024−5060]]

[1] Department of Artificial Intelligence Convergence, Chonnam National University, Buk-gu, Gwangju 61186, Republic of Korea
[2] Hyper-Wide Federated Medical AI Research Center, Chonnam National University, Buk-gu, Gwangju 61186, Republic of Korea
{tiennq27, melee, shkim, hjyang}@jnu.ac.kr

Abstract. Brain age gap (BAG) derived from structural MRI is a promising surrogate biomarker for quantifying neurodegenerative burden across the Alzheimer's disease (AD) continuum; however, most existing methods emphasize global accuracy while offering limited support for region-level interpretation. We propose VCDA-Net, a multi-scale brain age estimation framework that couples global morphological context with local geometric cues derived from voxel-cluster density. VCDA-Net introduces Voxel-Cluster Density Attention (VCDA) to encode stable regional geometry from high-saliency features and models inter-regional relations via a multi-attributed brain graph with Multi-Edge Attention. A Transformer aggregation module captures long-range dependencies across 32 anatomical structures, while an adaptive gated fusion integrates graph-derived signatures with UNet bottleneck priors for final age regression and structure-wise BAG profiling. Evaluated on multi-cohort normative modeling and the AD-spectrum, VCDA-Net achieves an MAE of 3.71 ± 0.31 years ($r = 0.97, R^2 = 0.94$). Our results reveal progressively increased BAG from MCI to AD, with regional BAG highlighting AD-vulnerable structures such as the bilateral hippocampus and the ventricular system. The code is available at https://github.com/quoctienchonnam25-collab/Voxel-Cluster_Density_Attention_Network.

Keywords: Voxel-Cluster Density · Brain Age Gap · Alzheimer's Disease · Graph Neural Network · Interpretability

1 Introduction

The brain age gap (BAG) measures the discrepancy between predicted and chronological age to quantify neurodegeneration. Although deep learning improves global accuracy, many models remain black boxes that lack anatomical transparency. By focusing on holistic features, these approaches often ignore

heterogeneous information within specific structures. For example, localized hippocampal atrophy contributes significantly to aging but is frequently obscured by single global representations.

While SFCN [18] offers resource efficiency, its shallow architecture often overlooks fine-grained anatomical details. Deeper models like ResNet50 [9] and DenseNet121 [12] extract complex features but struggle with boundary biases. The minimalist approach by Dartora et al. [3] enhances utility but remains susceptible to non-brain signals and fails to capture long-range structural dependencies. Similarly, despite the advances in cross-domain adaptation introduced by SynthBA [20], its reliance on rigid spatial transformations restricts the learning of natural anatomical variations. The ORDER loss function [22] addresses regression toward the average but faces systematic bias at age boundaries. Furthermore, voxel-level tasks [17] require immense computational resources for 3D Transformers and rely on idealized aging assumptions. These limitations underscore the necessity of a unified framework capable of capturing multi-scale variations while preserving both predictive accuracy and pathological interpretability.

We propose the Voxel-Cluster Density Attention (VCDA) mechanism to quantify neurodegeneration through spatial-density variations. Our adaptive dual-stream system integrates global UNet features [21,6] with local VCDA graphs via a Multi-Edge Attention mechanism, inspired by [1]. This mechanism incorporates dot-product, Euclidean distance, and cosine similarity to prioritize diagnostic connectivity. By combining a Transformer [24] with Gated Fusion [5], the model effectively captures early micro-scale changes. This approach accurately identifies global and regional BAGs across 32 structures, enabling quantitative stratification from MCI to AD.

The main contributions are: (i) the VCDA mechanism for enhanced sensitivity to early-stage neurodegeneration; (ii) a dual-stream architecture using Gated Fusion and Multi-Edge Attention to capture complex inter-regional dependencies; and (iii) a framework tracking global and regional BAG across 32 structures to characterize pathological shifts in the AD-spectrum.

2 Methods

As shown in Fig. 1, our proposed model integrates global morphological features with local voxel-cluster densities via a dual-stream gated fusion mechanism. After optimization, the model estimates multi-scale BAG values to differentiate between MCI and AD populations, thereby characterizing the neurodegenerative continuum at both global and regional levels.

2.1 Voxel-Cluster Density Attention Mechanism

As illustrated in Fig. 2, conventional regional volume measures exhibit limitations in tracking neurodegenerative progression due to inherent segmentation noise. In small structures such as the Right Hippocampus, inaccuracies in boundary delineation lead to suboptimal Dice scores (0.6188). Even in larger regions

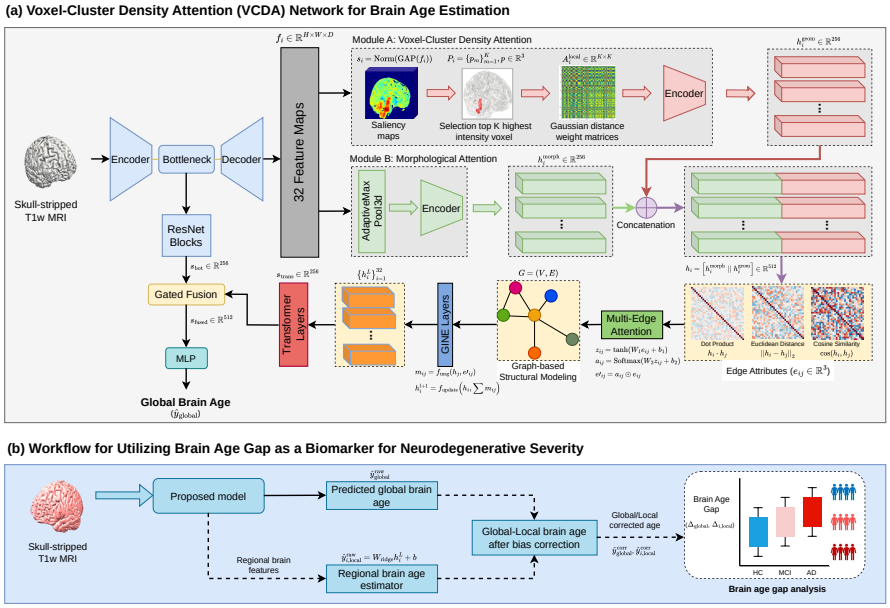

Fig. 1: Overview of the proposed VCDA-Net and its clinical application workflow.

like the Left Lateral Ventricle (0.8322), calculated volumetric metrics often misrepresent the true anatomical volume, as they remain highly susceptible to inconsistent segmentation masks.

To address this, regional saliency maps are generated from 3D feature maps $\mathbf{f}_i \in \mathbb{R}^{H\times W\times D}$ via Global Average Pooling (GAP), for each brain region index $i \in \{1,\ldots,32\}$. By deriving min-max normalized importance weights $s_i = \text{Norm}(\text{GAP}(\mathbf{f}_i)) \in [0,1]$, we identify stable, high-intensity voxel clusters robust to segmentation noise. These weights guide the selection of the top-K coordinates $\mathcal{P}_i = \{p_m\}_{m=1}^{K}$, where $p \in \mathbb{R}^3$. Analysis in Fig. 2(c) confirms that increasing K from 64 to 512 raises voxel density while maintaining spatial stability, as evidenced by converged centroids. While $K = 512$ introduces some peripheral, low-significance voxels, the increased granularity at this scale captures finer morphological details, ultimately yielding the optimal MAE as reported in Table 1. Consequently, the VCDA mechanism utilizes $\mathcal{P}_i$ to build a local geometric descriptor via a Gaussian-weighted Euclidean distance matrix $\mathbf{A}_i^{local} \in \mathbb{R}^{K\times K}$ with entries $a_{mn} = \exp(-\|p_m - p_n\|^2/2\sigma^2)$, where σ is experimentally fixed at 10 to balance spatial decay. This matrix is processed to extract fine-grained geometric features $\mathbf{h}_i^{geom} \in \mathbb{R}^{256}$. A ResNet-based stream extracts global features $\mathbf{h}_i^{morph} \in \mathbb{R}^{256}$, which are concatenated into $\mathbf{h}_i = [\mathbf{h}_i^{morph} \parallel \mathbf{h}_i^{geom}] \in \mathbb{R}^{512}$ to capture both morphological and geometric characteristics.

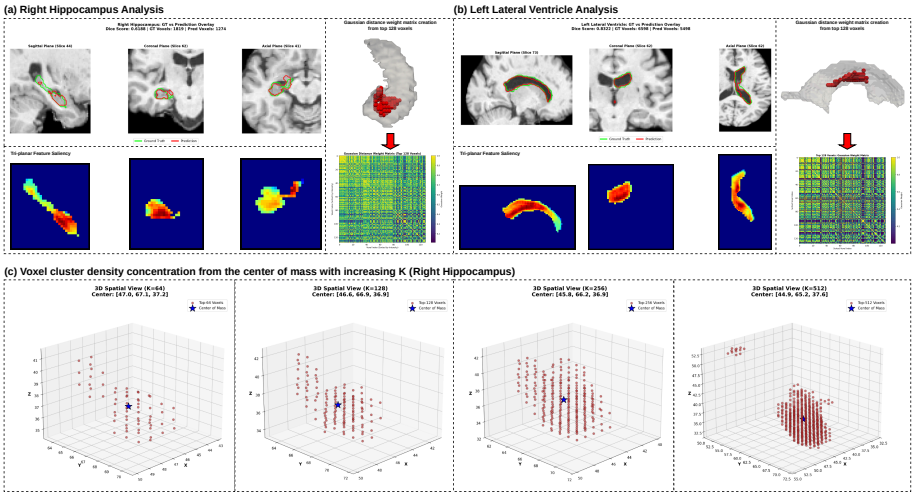

Fig. 2: Quantitative analysis of regional segmentation and voxel density. (a, b) UNet predictions vs. ground truth for the Right Hippocampus and Left Lateral Ventricle. (c) Cluster distribution across K scales; increasing K raises density while maintaining spatial stability via tightly converged centroids.

2.2 Graph-based Structural Modeling

We represent the brain as a multi-attributed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with $N = 32$ nodes, corresponding to anatomical regions segmented by SynthSeg [2]. The edge set $\mathcal{E}$ follows a k-NN topology ($k = 10$) in the latent space, where each edge (i, j) incorporates an attribute vector $\mathbf{e}_{ij} = [\mathbf{h}_i \cdot \mathbf{h}_j, \|\mathbf{h}_i - \mathbf{h}_j\|_2, \cos(\mathbf{h}_i, \mathbf{h}_j)]^\top$. These attributes comprise the dot product, Euclidean distance, and cosine similarity, providing diverse relational perspectives among node embeddings. To prioritize diagnostic connections, a Multi-Edge Attention mechanism re-weights these interactions as $\mathbf{e}'_{ij} = \text{Softmax}_{attr}(\mathbf{W}_2(\tanh(\mathbf{W}_1\mathbf{e}_{ij} + \mathbf{b}_1)) + \mathbf{b}_2) \odot \mathbf{e}_{ij}$, where the Softmax is computed across the attribute dimension to ensure weights for the three features sum to one per edge. Node features are then updated via GINEConv layers [11]:

$$\mathbf{h}_i^{(l+1)} = \text{MLP}^{(l)}\left((1+\epsilon^{(l)})\mathbf{h}_i^{(l)} + \sum_{j \in \mathcal{N}(i)} \text{ReLU}(\mathbf{h}_j^{(l)} + \mathbf{W}_e^{(l)}\mathbf{e}'_{ij})\right) \tag{1}$$

This integration embeds multi-dimensional edge information directly into the message-passing process to emphasize disease-specific connectivity.

2.3 Feature Aggregation and Gated Fusion

Following the graph-based message passing, we employ a Transformer Aggregation module to distill global structural markers from the final node embeddings

$\{\mathbf{h}_i^{(L)}\}_{i=1}^{32}$ after L message-passing iterations. Through multi-head self-attention, the model captures long-range dependencies across the 32 anatomical regions to produce a topological aging descriptor $\mathbf{s}_{trans} \in \mathbb{R}^{256}$. To integrate these regional associations with global contextual priors, an Adaptive Gated Fusion mechanism dynamically blends $\mathbf{s}_{trans}$ with the UNet Bottleneck embedding $\mathbf{s}_{bot} \in \mathbb{R}^{256}$ via a learnable gating network:

$$\mathbf{s}_{fused} = \mathbf{g} \odot \phi_1(\mathbf{s}_{trans}) + (1 - \mathbf{g}) \odot \phi_2(\mathbf{s}_{bot}), \tag{2}$$

where $\mathbf{g} = \text{Sigmoid}(\text{MLP}_{gate}([\mathbf{s}_{trans} \parallel \mathbf{s}_{bot}]))$ denotes sample-specific gating weights, and ϕ_1, ϕ_2 are linear projections. This strategy enables the model to selectively emphasize either graph-derived relational patterns or global morphological features. The final brain age prediction $\hat{y}$ is obtained via an MLP regression head applied to the fused representation $\mathbf{s}_{fused}$.

2.4 Multi-scale Age Prediction and Bias Correction

We employ a hierarchical framework to estimate brain age, followed by bias correction [23] to mitigate regression-to-the-mean effects. For each region $i \in \{1, \ldots, 32\}$, a Ridge decoder predicts the raw local age from node embeddings:

$$\hat{y}_{i,local}^{raw} = \mathbf{W}_{ridge}\mathbf{h}_i^{(L)} + b_{ridge}. \tag{3}$$

We prioritize Ridge regression over complex non-linear decoders to enhance model stability and interpretability while preventing overfitting through L_2 regularization. To ensure the BAG reflects true pathology, we apply a linear correction $\hat{y}_{corr} = (\hat{y}^{raw} - \beta)/\alpha$, where parameters are derived by fitting $(\hat{y}^{raw} - y) = \beta_1 y + \beta_0$ on the HC group (y denotes chronological age, $\alpha = 1 + \beta_1$, and $\beta = \beta_0$). This procedure isolates pathological accelerated aging by decoupling systematic prediction bias from true neurodegenerative-induced age advances. The final multi-scale BAGs are defined as:

$$\Delta_{global} = \hat{y}_{global}^{corr} - y, \quad \Delta_{i,local} = \hat{y}_{i,local}^{corr} - y. \tag{4}$$

This strategy minimizes age-related bias, enabling standardized identification of individuals exhibiting rapid neurodegeneration across clinical cohorts.

3 Experiments and results

3.1 Experimental Setup

Datasets. We utilized 1,710 structural MRI scans for normative modeling, defining this collection as the Multi-cohort Normative Dataset (MCN-Set). These were aggregated from ABIDE [4] (427), ADNI [19] (146), PPMI [16] (201), IXI [13] (449), and OASIS-1/-2 [15,14] (487; 90/397). To avoid longitudinal overlap in OASIS-2, we selected unique baseline scans to ensure a balanced age distribution (range: 6–97 years, mean±SD: 51.65 ± 25.70, sex: 849F/861M). Disease

Table 1: Quantitative comparison between VCDA-Net and existing methods on the MCN-Set. **Best** and second-best results are highlighted.

Methods	**MAE** ↓	**RMSE** ↓	R^2 ↑	r ↑
DenseNet121 [12]	6.27 ± 0.33	8.34 ± 0.44	0.90 ± 0.01	0.95 ± 0.01
ResNet50 [9]	8.16 ± 0.51	11.48 ± 0.81	0.80 ± 0.03	0.90 ± 0.01
SynthBA-G [20]	9.76 ± 0.60	13.77 ± 0.99	0.71 ± 0.05	0.90 ± 0.02
SynthBA-U [20]	10.89 ± 0.71	15.74 ± 1.10	0.62 ± 0.06	0.86 ± 0.02
Ordinal-DR [22]	5.14 ± 0.39	7.94 ± 0.72	0.90 ± 0.02	0.95 ± 0.01
Voxel-level BA [17]	6.96 ± 0.52	10.72 ± 1.23	0.80 ± 0.05	0.90 ± 0.02
MP-BA Net [3]	4.24 ± 0.33	6.78 ± 0.63	0.92 ± 0.02	0.96 ± 0.01
SFCN [18]	4.60 ± 0.30	6.49 ± 0.50	0.93 ± 0.01	**0.97 ± 0.01**
VCDA ($K = 64$)	5.34 ± 0.32	7.48 ± 0.57	0.92 ± 0.02	0.96 ± 0.01
VCDA ($K = 128$)	4.22 ± 0.32	6.55 ± 0.65	0.93 ± 0.01	**0.97 ± 0.01**
VCDA ($K = 256$)	4.05 ± 0.28	**6.03 ± 0.42**	**0.94 ± 0.01**	**0.97 ± 0.01**
VCDA ($K = 512$)	**3.71 ± 0.31**	6.21 ± 0.66	**0.94 ± 0.01**	**0.97 ± 0.01**

sensitivity was assessed using test sets from ADNI (427 MCI, 157 AD) and OASIS (514 Clinical Dementia Rating [CDR] 0.5: 70/444; 187 CDR 1.0: 28/159). These were evaluated against 257 independent Healthy Control (HC) samples, enabling multi-scale BAG quantification for clinical stratification across the AD-spectrum.

Implementation Details. The proposed framework was implemented in PyTorch on an RTX 3090; sMRI scans were skull-stripped via SynthStrip [10] and resampled to 128^3. The model was optimized end-to-end for 300 epochs (Adam/Huber loss) on a 70/15/15% data split. To preserve spatial features pre-trained on IXI, differential learning rates were used: 10^{-5} for the UNet backbone and 10^{-3} for remaining modules. Performance was validated via a 1,000-iteration bootstrap (95% CI) on the fixed test set. For regional analysis, the model served as a feature extractor to compute Δ_{local}, with significance and effect sizes assessed via one-sample t-tests and Cohen's d against a zero-deviation null hypothesis.

3.2 Performance Evaluation and Ablation Study

Performance Comparison. Our proposed model outperforms established baselines (Table 1), including DenseNet121 (6.27 ± 0.33), ResNet50 (8.16 ± 0.51), SFCN (4.60 ± 0.30), and MP-BA Net (4.24 ± 0.33) in MAE. The high accuracy ($r = 0.97, R^2 = 0.94$) confirms that integrating voxel-cluster features with GNNs effectively captures complex aging patterns. We observe a performance trade-off: $K = 512$ achieves the lowest MAE (3.71 ± 0.31) via fine-grained details, while $K = 256$ yields the optimal RMSE (6.03 ± 0.42), indicating better outlier robustness. Thus, while higher densities enhance precision, moderate cluster-

Table 2: Ablation studies of VCDA-Net on the MCN-Set.

Configuration	MAE ↓	RMSE ↓	R^2 ↑	r ↑
Edge Attributes				
Euclidean dist.	4.42 ± 0.30	6.55 ± 0.51	0.93 ± 0.01	0.97 ± 0.01
Cosine sim.	5.82 ± 0.34	7.99 ± 0.52	0.90 ± 0.01	0.95 ± 0.01
Dot prod.	4.40 ± 0.30	6.60 ± 0.60	0.93 ± 0.01	0.97 ± 0.01
Multi-edge attn.	**4.22 ± 0.32**	**6.55 ± 0.65**	**0.93 ± 0.01**	**0.97 ± 0.01**
Attention Modules				
Baseline	6.86 ± 0.38	9.01 ± 0.46	0.87 ± 0.01	0.94 ± 0.01
Baseline+(A)	6.58 ± 0.36	8.64 ± 0.44	0.88 ± 0.01	0.94 ± 0.01
Baseline+(B)	6.69 ± 0.38	8.88 ± 0.54	0.88 ± 0.02	0.94 ± 0.01
Baseline+(A)+(B)	**4.22 ± 0.32**	**6.55 ± 0.65**	**0.93 ± 0.01**	**0.97 ± 0.01**
Baselines				
UNETR	5.27 ± 0.46	8.96 ± 0.96	0.88 ± 0.03	0.94 ± 0.01
Swin UNETR	4.76 ± 0.39	7.84 ± 0.90	0.91 ± 0.02	0.95 ± 0.01
UNet	**4.22 ± 0.32**	**6.55 ± 0.65**	**0.93 ± 0.01**	**0.97 ± 0.01**

Table 3: Statistical analysis of global BAG (Δ_{global}) across different cohorts.

Dataset	Group	Mean ± SD	95% CI
MCN-Set	HC	0.00 ± 6.06	[−0.74, 0.74]
OASIS	CDR 0.5	2.29 ± 7.65	[1.63, 2.95]
	CDR 1.0	5.10 ± 8.06	[3.94, 6.27]
ADNI	MCI	2.68 ± 5.25	[2.18, 3.18]
	AD	4.09 ± 6.53	[3.06, 5.11]

ing stabilizes predictions, providing a more effective biomarker than traditional voxel-level or global approaches.

Ablation Studies. Ablation experiments, conducted at $K = 128$ (Table 2), validate the core components of our proposed model. The Multi-Edge Attention mechanism achieves optimal performance ($MAE = 4.22 \pm 0.32, r = 0.97 \pm 0.01$) by dynamically weighting structural covariance patterns. While VCDA (A) ensures feature stability, Morphological Attention (B) provides critical global context; their synergy reduces MAE from 6.86 ± 0.38 to 4.22 ± 0.32. Notably, our UNet-bottleneck backbone outperforms heavier models like UNETR [8] ($MAE = 5.27 \pm 0.46$) and Swin UNETR [7] ($MAE = 4.76 \pm 0.39$), demonstrating that a streamlined architecture can balance high accuracy with interpretability without the complexity of large Transformers.

3.3 Statistical Analysis of Brain Age Gaps

The statistical evaluation across clinical cohorts, summarized in Table 3, reveals that the global BAG serves as a sensitive biomarker for cognitive decline. In the MCN-Set, the HC group exhibits a mean Δ_{global} of 0.00 ± 6.06 with a narrow 95% CI [−0.74, 0.74], confirming that the bias correction effectively eliminates

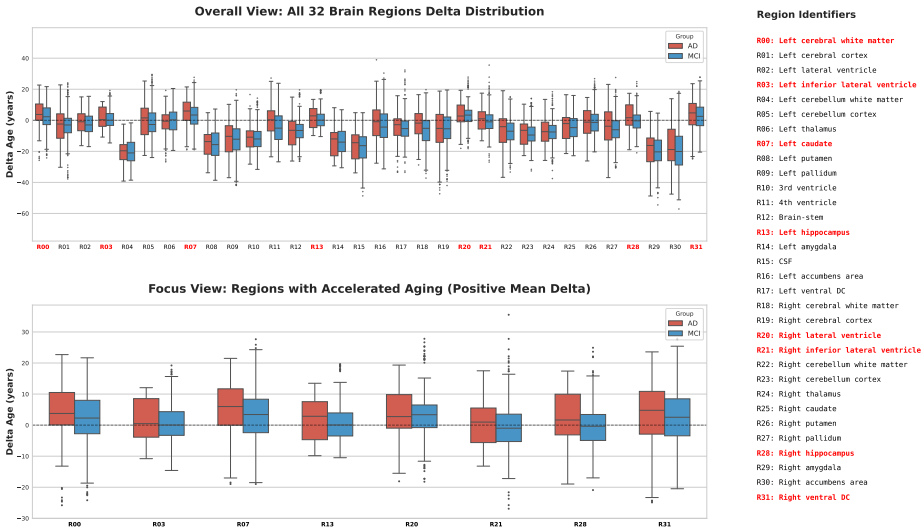

Fig. 3: Comparison of regional brain age gaps (Δ_{local}) between AD and MCI subjects across 32 anatomical regions defined by SynthSeg in the ADNI dataset.

system errors. We observe a progressive increase in Δ_{global} linked to impairment severity: in OASIS, the BAG rises from 2.29 ± 7.65 (CDR 0.5) to 5.10 ± 8.06 (CDR 1.0), while in ADNI, it increases from 2.68 ± 5.25 (MCI) to 4.09 ± 6.53 (AD). This global trend is further supported by regional BAG analysis, which identifies distinct neuroanatomical signatures across the AD-spectrum.

As illustrated in Fig. 3, the most significant regional differentiators are localized within the hippocampal and ventricular systems. Specifically, the Right and Left hippocampi exhibit positive gaps (Δ_{local}) of 2.11 ± 8.08 ($p < .005, d = 0.26$) and 1.41 ± 6.65 ($p < .009, d = 0.21$) in the AD group, respectively, while the MCI group remains relatively stable. Conversely, Right lateral ventricle expansion is significantly elevated in both AD ($3.94 \pm 7.39, p < .001, d = 0.53$) and MCI ($3.03 \pm 6.91, p < .001, d = 0.44$), suggesting that ventricular enlargement accelerates during the disease transition. Other notable AD-specific deviations include significant gaps in the Left caudate ($5.63 \pm 8.37, p < .001, d = 0.67$) and Left cerebral white matter ($4.24 \pm 9.24, p < .001, d = 0.46$). Collectively, these markers form a robust regional profile to distinguish pathological neurodegeneration from normal biological decline.

4 Conclusion

The proposed multi-scale framework achieves high predictive fidelity by integrating global morphological features with regional VCDA geometric extraction, Multi-Edge Attention, and Gated Fusion. This architecture effectively captures pathological signatures, exemplified by hippocampal atrophy and ventricular ex-

pansion, to facilitate precise clinical stratification from MCI to AD across the neurodegenerative continuum. By localizing neurodegenerative changes to specific anatomical structures, the framework enhances model interpretability and establishes a robust methodology for monitoring disease progression in clinical research.

Acknowledgments. This work was supported by the Institute of Information & communications Technology Planning & Evaluation (IITP) grants funded by the Korea government (MSIT) under the Artificial Intelligence Convergence Innovation Human Resources Development program (IITP-2023-RS-2023-00256629) and the ITRC (Information Technology Research Center) support program (IITP-2026-RS-2024-00437718). This research was also supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (RS-2025-25425780), and the Regional Innovation System & Education (RISE) Glocal University 30 Program through the Gwangju RISE Center, funded by the Ministry of Education (MOE) and the Gwangju Metropolitan City, Republic of Korea (2026-RISE(Glocal University 30)-05-011).

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

References

1. Basheer, R., Daraie, A.H., Mishra, D.: Hypercube-s4gnn: A multi-edge graph approach using state space models on multivariate eeg for seizure detection. In: 2024 IEEE 34th International Workshop on Machine Learning for Signal Processing (MLSP). pp. 1–6 (2024)
2. Billot, B., Greve, D., Puonti, O., Thielscher, A., Van Leemput, K., Fischl, B., Dalca, A., Iglesias, J.: Synthseg: Segmentation of brain mri scans of any contrast and resolution without retraining. Medical Image Analysis **86**, 102789 (2023)
3. Dartora, C., Marseglia, A., Mårtensson, G., et al.: A deep learning model for brain age prediction using minimally preprocessed T1w images as input. Frontiers in Aging Neuroscience **15**, 1303036 (2024)
4. Di Martino, A., Yan, C.G., Li, Q., et al.: The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. Molecular Psychiatry **19**(6), 659–667 (2014)
5. Dong, X., Zhang, H., Han, H., Xu, Z.: Dpgad: Structure-aware dual-path attention graph node anomaly detection. Symmetry **17**, 1452 (09 2025)
6. Gianchandani, N., Dibaji, M., Ospel, J., Vega, F., Bento, M., MacDonald, M.E., Souza, R.: A voxel-level approach to brain age prediction: A method to assess regional brain aging. Machine Learning for Biomedical Imaging **2**(April 2024), 761795 (Apr 2024)
7. Hatamizadeh, A., Nath, V., Tang, Y., Yang, D., Roth, H.R., Xu, D.: Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries (BrainLes 2021). Lecture Notes in Computer Science, vol. 12962, pp. 272–284. Springer International Publishing, Cham (2022)

8. Hatamizadeh, A., Tang, Y., Nath, V., Yang, D., Myronenko, A., Landman, B., Roth, H.R., Xu, D.: Unetr: Transformers for 3d medical image segmentation. In: 2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). pp. 1748–1758 (2022)
9. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 770–778 (June 2016)
10. Hoopes, A., Mora, J.S., Dalca, A.V., Fischl, B., Hoffmann, M.: Synthstrip: skull-stripping for any brain image. NeuroImage **260**, 119474 (2022)
11. Hu, W., Liu, B., Gomes, J., Zitnik, M., Liang, P., Pande, V., Leskovec, J.: Strategies for pre-training graph neural networks. In: International Conference on Learning Representations (2020)
12. Huang, G., Liu, Z., van der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2017)
13. Information eXtraction from Images: IXI dataset. https://brain-development.org/ixi-dataset/ (2026), accessed: 2026-02-09
14. Marcus, D.S., Fotenos, A.F., Csernansky, J.G., et al.: Open access series of imaging studies: longitudinal MRI data in nondemented and demented older adults. Journal of Cognitive Neuroscience **22**(12), 2677–2684 (2010)
15. Marcus, D.S., Wang, T.H., Parker, J., et al.: Open access series of imaging studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. Journal of Cognitive Neuroscience **19**(9), 1498–1507 (2007)
16. Marek, K., Chowdhury, S., Siderowf, A., et al.: The Parkinson's progression markers initiative (PPMI) - establishing a PD biomarker cohort. Annals of Clinical and Translational Neurology **5**(12), 1460–1477 (2018)
17. Nasser, T., Souza, R., El-Sheimy, N.: Investigating Voxel-level Brain Age Prediction as a Pretext Task for Brain MRI Segmentation . In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2025. vol. LNCS 15975, pp. 289 – 299. Springer Nature Switzerland (October 2025)
18. Peng, H., Gong, W., Beckmann, C.F., Vedaldi, A., Smith, S.M.: Accurate brain age prediction with lightweight deep neural networks. Medical Image Analysis **68**, 101871 (2021)
19. Petersen, R.C., Aisen, P.S., Beckett, L.A., et al.: Alzheimer's disease neuroimaging initiative (ADNI): clinical characterization. Neurology **74**(3), 201–209 (2010)
20. Puglisi, L., Rondinella, A., De Meo, L., Guarnera, F., Battiato, S., Ravì, D.: Synthba: Reliable brain age estimation across multiple mri sequences and resolutions. In: 2024 IEEE International Conference on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE). p. 559564. IEEE (Oct 2024)
21. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Lecture Notes in Computer Science, vol. 9351, pp. 234–241. Springer International Publishing, Cham (2015)
22. Shah, J., Rahman Siddiquee, M.M., Su, Y., Wu, T., Li, B.: Ordinal classification with distance regularization for robust brain age prediction. In: 2024 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). pp. 7867–7876 (2024)
23. Sihag, S., Mateos, G., McMillan, C., Ribeiro, A.: Explainable brain age prediction using covariance neural networks. In: Advances in Neural Information Processing Systems (NeurIPS). vol. 36, pp. 46958–46988 (2023)

24. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I.: Attention is all you need (2023)