PRGNN: Pyramidal Region Graph Neural Network for Region-Based Brain PET Classification

Daesung Kim[1][0009−0006−7966−947X], Seungbeom Seo[2][0009−0006−9779−3757], Boosung Kim[3], Kyobin Choo[1], Youngjun Jun[1], and Mijin Yun[2]

[1] Yonsei University, Seoul, Republic of Korea
[2] Yonsei University College of Medicine, Seoul, Republic of Korea
[3] Dickinson College, Pennsylvania, USA
daesungkim@yonsei.ac.kr, yunmijin@yuhs.ac

Abstract. Brain positron emission tomography (PET) has been widely used for the diagnosis of various neurodegenerative diseases. To assist physicians, convolutional neural networks (CNNs) and transformers have been explored for prediction of diseases based on brain PET images. While these models show promising performance, they are designed to process the entire image, which facilitates shortcut learning by extracting irrelevant features. To alleviate shortcut learning, we observe that brain images share the same structure, and regions of interest (ROIs) can be defined for relevant regions. In this regard, we propose Pyramidal Region Graph Neural Network (PRGNN), which employs a 3D convolutional backbone to learn multi-level feature representations and constructs nodes that correspond to anatomical ROIs. Using ROI-based node embeddings, PRGNN extracts metabolic patterns in functionally relevant regions and performs explicit inter-regional reasoning. We evaluate PRGNN on classifying 18F-fluorodeoxyglucose (FDG) and amyloid PET, outperforming models based on CNN, transformer, and GNN. Moreover, interpretability analyses highlight disease-relevant regions that align with clinical observations, demonstrating PRGNN's potential for improving diagnostic performance and reliability. Code is available at https://github.com/Treeboy2762/PRGNN.

Keywords: Classification · Graph Neural Network · Positron Emission Tomography · Explainable AI.

1 Introduction

The brain undergoes distinct pathological and metabolic changes during the progression of neurodegenerative disorders, driven by disease-specific molecular pathology and organizational patterns [1]. In this regard, positron emission tomography (PET), such as 18F-fluorodeoxyglucose (FDG) and amyloid PET, has gained significant attention for its ability to detect disease-specific abnormalities even before structural changes become evident on MRI [2, 3]. Given the critical

role of early and accurate diagnosis in improving patient outcomes, various classification models have been explored for PET-based disease classification, aiming to extract meaningful patterns and enhance diagnostic precision.

Among them, CNNs and transformers have demonstrated strong performance in PET-based disease classification [4–7]. However, these models unintentionally learn *shortcut features*, spurious correlations that boost accuracy in training but do not reflect true pathology [8]. In particular, traditional image classification models process all pixels of an image, inadvertently referring to irrelevant features or background context as shortcuts. This redundancy degrades the model's ability to generalize to images in the unseen domain, especially when limited data are available to learn features that robustly distinguish the classes.

Alternatively, graph neural network (GNN)-based architectures are uniquely suited to incorporate anatomical structure into the model, addressing the limitations of CNNs and transformers. Traditional CNNs utilize a grid image input and use fixed local neighborhoods, which limit their capacity to capture long-range relationships [9]. On the other hand, transformers do capture long-range interactions using attention, but they still treat the image as an unstructured sequence of patches, and they require larger datasets for effective training [10]. GNNs, by contrast, can be designed to treat anatomical regions as nodes in a graph with edges encoding relationships. To elaborate, by integrating CNN-based feature extraction that captures detailed local patterns within each region with graph operations that connect these regions, GNNs can combine the advantages of CNN's localized feature extraction and the ability of GNN to model long-range dependencies.

A few works have incorporated anatomical knowledge into brain image classification [7, 11–13]. For instance, BrainGNN [11] groups nodes and uses an MLP for classification but depends on explicit edges from modality-specific sources (e.g., fMRI functional connectivity), which are harder to define in PET images. Another study [12] uses separate CNNs for each region of interest (ROI), standardizing their dimensions before feature extraction and subsequent GNN processing. Although this captures local features, it overlooks the multiscale hierarchical maps typically produced by backbone networks. In contrast, another work [13] constructs graphs based on ROI data and employs a GNN with Transformer-guided Adaptive Diffusion (GTAD) for Alzheimer's classification; however, it relies on averaged SUVR values within ROIs, which do not fully capture the complex metabolic patterns in PET images.

To address these issues, we propose the Pyramidal Region Graph Neural Network (PRGNN), a novel framework that integrates the strengths of 3D CNN and GNN to capture both local feature hierarchies and global structural relationships in medical imaging data. PRGNN uses an ROI-based node embedding method to define nodes in the feature map extracted by a CNN backbone. By constraining the model to extract features from anatomically relevant regions, PRGNN alleviates the shortcut problem. Moreover, by using graph convolution between nodes, PRGNN allows the model to utilize long-range dependencies, modeling intricate relationships of the brain. Our contributions are as follows:

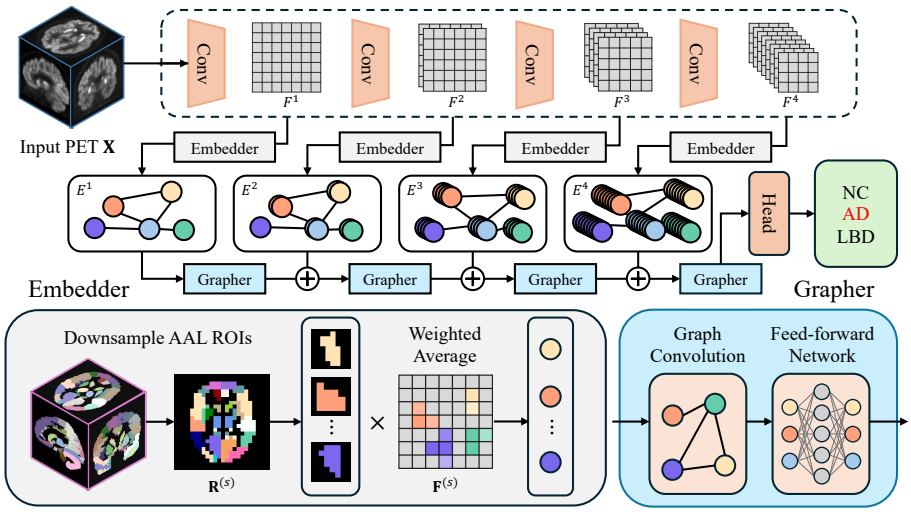

Fig. 1. Model Architecture of PRGNN. A 3D CNN backbone extracts hierarchical features from PET images, which are then embedded to node representations using ROIs of automatic anatomical labeling. The node embeddings are then processed by grapher modules and fed into a classifier head for prediction.

- We present PRGNN, which integrates a 3D CNN backbone with a GNN to capture both local hierarchical features and global anatomical relationships in brain PET images.
- We provide comprehensive experimental validation on FDG and amyloid PET datasets, demonstrating that PRGNN outperforms existing classification approaches in brain PET classification. We provide an ablation study to validate that graph convolutions between regions improve performance.
- We show that PRGNN makes an interpretable diagnosis by summing up logit contributions from each region. We visualize the average node contributions of PRGNN and confirm the clinical relevance of the model predictions.

2 Methodology

Our PRGNN (Fig. 1) is composed of the following components: 1) 3D CNN backbone to extract hierarchical features, 2) node embedder to project CNN features to node representations using anatomical ROIs, 3) grapher block to perform graph convolution between nodes, and 4) classifier head to perform final classification. We trained three versions of PRGNN with different model sizes: tiny (Ti), small (S), and medium (M). The details are provided in Table 1.

Table 1. Detailed descriptions of parameters for PRGNN. C: feature dimension, K: number of neighbors in GCN, N: number of nodes, $H \times W \times D$: input image size. 'Ti' stands for tiny, 'S' for small, and 'M' for medium.

Stage	Output size	PRGNN-Ti	PRGNN-S	PRGNN-M
Stem	$\frac{H}{2} \times \frac{W}{2} \times \frac{D}{2}$	Conv7 × 7 × 7	Conv7 × 7 × 7	Conv7 × 7 × 7
Stage 1	N	$\begin{bmatrix} C=48 \\ K=9 \end{bmatrix} \times 2$	$\begin{bmatrix} C=72 \\ K=9 \end{bmatrix} \times 2$	$\begin{bmatrix} C=96 \\ K=9 \end{bmatrix} \times 2$
Downsample	$\frac{H}{4} \times \frac{W}{4} \times \frac{D}{4}$	Conv3 × 3 × 3	Conv3 × 3 × 3	Conv3 × 3 × 3
Stage 2	N	$\begin{bmatrix} C=48 \\ K=9 \end{bmatrix} \times 2$	$\begin{bmatrix} C=72 \\ K=9 \end{bmatrix} \times 2$	$\begin{bmatrix} C=96 \\ K=9 \end{bmatrix} \times 2$
Downsample	$\frac{H}{8} \times \frac{W}{8} \times \frac{D}{8}$	Conv3 × 3 × 3	Conv3 × 3 × 3	Conv3 × 3 × 3
Stage 3	N	$\begin{bmatrix} C=96 \\ K=9 \end{bmatrix} \times 6$	$\begin{bmatrix} C=144 \\ K=9 \end{bmatrix} \times 6$	$\begin{bmatrix} C=192 \\ K=9 \end{bmatrix} \times 16$
Downsample	$\frac{H}{16} \times \frac{W}{16} \times \frac{D}{16}$	Conv3 × 3 × 3	Conv3 × 3 × 3	Conv3 × 3 × 3
Stage 4	N	$\begin{bmatrix} C=240 \\ K=9 \end{bmatrix} \times 2$	$\begin{bmatrix} C=288 \\ K=9 \end{bmatrix} \times 2$	$\begin{bmatrix} C=384 \\ K=9 \end{bmatrix} \times 2$
Head	1	Pooling & FC	Pooling & FC	Pooling & FC

2.1 Hierarchical feature extraction

Let $\mathbf{X} \in \mathbb{R}^{C \times H \times W \times D}$ represent the 3D PET input image, where C is the number of channels (e.g., 1 for a single PET image), and H, W, D represent the spatial dimensions. A CNN backbone is used to extract four stages of hierarchical feature map, denoted $\mathbf{F}^{(s)} \in \mathbb{R}^{C_s \times H_s \times W_s \times D_s}$, where $s \in \{1, 2, 3, 4\}$ represents the four stages of the CNN. Each stage is downsampled by a factor of 2.

2.2 Node Embedder

To generate node embeddings from the hierarchical feature map, we use Automatic Anatomical Labeling (AAL) atlas that provides 116 unique brain regions [14]. To prevent overfitting, we merge AAL labels along gyral boundaries, removing arbitrary anterior-posterior splits within the same gyrus. Moreover, we merge fragmented areas such as the cerebellum. The resulting atlas had total of 56 ROIs. We then one-hot encode the merged ROI map and downsample it via bilinear interpolation to match the resolution of the hierarchical feature map.

Let $\mathbf{E}^{(s)} \in \mathbb{R}^{C_s \times N}$ be the node embeddings and $\mathbf{R}^{(s)} \in \mathbb{R}^{N \times H_s \times W_s \times D_s}$ be the downsampled AAL atlas, where N is the number of ROIs. The node embedding process can be written as:

$$E_i^{(s)} = \sum_{h,w,d} R_{i,h,w,d}^{(s)} \cdot F_{:,h,w,d}^{(s)} \quad \text{for each node } i \in \{1, ..., N\}. \tag{1}$$

This represents a weighted sum of the feature map values, where the ROI mask $\mathbf{R}$ acts as a set of weights to aggregate features from relevant spatial locations.

2.3 Grapher

To construct the graph $\mathcal{G} = G(\mathbf{E})$, we follow the scheme proposed in Vision GNN [15]. For each node embedding vector E_i, we identify its K-nearest neighbors $\mathcal{N}(E_i)$ and establish directed edges ξ_{ij} from E_i to every $E_j \in N(E_i)$. The resulting graph is then processed through a Grapher block, which serves as the core computational unit for extracting graphical features. The Grapher block consists of two main components: a graph convolution (GC) operator and a feed-forward network (FFN). First, GC operator updates node representations by aggregating feature information from their neighbors. This process is defined as

$$\mathbf{E}' = \sigma(\mathcal{G}'(\mathbf{E}W_{\text{in}}))W_{\text{out}} + \mathbf{E}, \tag{2}$$

where W_{in} and W_{out} are the weights of fully-connected (FC) layers, $\mathcal{G}'$ is the GC operator, and σ is the activation function. Specifically, we employ max-relative graph convolution [16] for the aggregation step, followed by update operation to merge the aggregated features within the GC operator:

$$E_i' = \sigma\left(\left(E_i \oplus \max_{E_j \in \mathcal{N}(E_i)} (E_i - E_j)\right) W_{\text{update}}\right), \tag{3}$$

where W_{update} is the learnable weight of the update operation. Next, to mitigate over-smoothing and enhance the feature transformation capacity, we apply a feed-forward network comprising two FC layers:

$$\mathbf{E}'' = \sigma(\mathbf{E}'W_1)W_2 + \mathbf{E}', \tag{4}$$

where W_1 and W_2 are the weights of FC layers, and the bias term is omitted. By combining the GC operator with FFN, the grapher block effectively captures and transforms graph-structured information.

2.4 Classifier head

Before passing the node embeddings to an FC layer, we apply average pooling across the channels to generate final node embeddings of shape N, where N represents the number of ROIs—different from vision GNN in which pooling is performed across the nodes. This way, we can uniquely isolate the contributions of each node for the final prediction. The pooled embeddings are then fed into an FC layer to predict the final label.

3 Experiments

Dataset. We evaluated the performance of the proposed PRGNN on two tasks: **(1)** classification of FDG PET images into normal condition (NC), Alzheimer's disease (AD), Lewy body dementia (LBD), and progressive supranuclear palsy (PSP), which are representative conditions that can be differentiated using FDG PET; **(2)** classification of amyloid positivity in amyloid PET images. For FDG classification, 234 FDG PET images were retrospectively collected internally,

Table 2. Comparison of model performance for FDG and FBB classification tasks.

Classifier	Params (M)	FLOPs (G)	FDG			FBB		
			Acc.	AUC	F1	Acc.	Sens.	Spec.
ResNet-18 [20]	33.2	106.5	88.2	97.9	88.3	96.2	**97.7**	94.6
ResNet-50 [20]	46.2	138.3	85.9	96.9	85.9	97.2	97.2	97.3
M3T [6]	27.0	615.4	84.5	95.6	84.5	96.2	**97.7**	95.7
AAGN [7]	1.5	505.4	85.9	96.7	85.8	97.2	**97.7**	96.8
ViG-Ti [15]	12.4	5.3	84.5	96.1	84.5	95.5	95.8	95.2
ViG-S [15]	17.6	9.9	86.3	96.6	86.4	95.5	95.3	95.7
ViG-M [15]	39.0	16.8	87.3	96.5	87.3	96.5	97.2	95.7
ViG-B [15]	68.5	29.6	88.4	97.3	88.4	97.0	96.2	97.9
PRGNN-Ti (ours)	5.1	5.1	91.7	98.5	91.7	96.5	95.8	97.3
PRGNN-S (ours)	8.6	9.6	91.9	98.4	91.9	96.7	96.2	97.3
PRGNN-M (ours)	19.7	15.9	**92.6**	**98.6**	**92.6**	**97.5**	96.7	**98.4**

consisting of 37 NC, 71 AD (42 MCI, 29 Dementia), 71 LBD, and 55 PSP cases (mean age ± SD (in years): NC, 62 ± 15; AD, 71 ± 8; LBD, 66 ± 9; PSP, 68 ± 6; sex distribution (in % female): NC, 62%; AD, 69%; LBD, 35%; PSP, 38%). In addition, 198 FDG PET images were collected from Alzheimer's Disease Neuroimaging Initiative (ADNI), which comprises 137 NC and 61 AD patients. AD patients in the ADNI cohort were confirmed amyloid positive. For amyloid positivity classification, 399 18F-florbetaben (FBB) PET images were collected internally, comprising 186 amyloid negative and 213 positive cases.

Labeling of the in-house dataset was performed according to predefined criteria [17–19], and all images were visually verified by two physicians. All PET images were spatially normalized to the MNI space and resampled to 2 mm isotropic voxels using SPM. The intensity of the images was normalized using the 99.9th percentile.

Baselines. As baselines, we categorized the models into three groups: **1) Convolutional models**: ResNet-18 and ResNet-50 [20] were selected for comparison with standard convolutional neural networks. **2) Attention models**: For comparison with attention models, we evaluated M3T [6], which efficiently integrates CNN and Transformer for 3D medical image classification. Additionally, we included AAGN [7], which incorporates an anatomy-aware squeeze-and-excite mechanism to extract region-specific features from brain MRI scans. **3) GNN models**: Vision GNN [15] was used as a representative model for comparison in GNN-based image classification.

Implementation Details The model was trained using the AdamW optimizer with an initial learning rate of 10^{-4} and cosine annealing learning rate scheduler. The training was performed for 100 epochs using Cross-Entropy loss with a batch size of 8 on two NVIDIA RTX A6000 48GB GPU. The hyperparameter K was chosen as 9 after empirical tuning from 3 to 12.

Table 3. Top ROIs for disease prediction. For each disease category, we report the top five ROIs (nodes) with the highest average logit contributions from the final FC layer. We only include cases where the predicted label matches the ground truth.

AD		LBD		PSP	
Region	**Logit**	**Region**	**Logit**	**Region**	**Logit**
Lt. Thalamus	3.17	Rt. Lingual	2.66	Lt. Putamen	4.00
Lt. Sup. Parietal	2.66	Lt. Cuneus	2.14	Lt. Sup. Parietal	2.05
Rt. Caudate Nucleus	1.98	Lt. Thalamus	2.08	Lt. Inf. Parietal	1.82
Lt. Sup. Temporal	1.21	Lt. Mid. Cingulate	2.03	Lt. Thalamus	1.37
Lt. Precuneus	1.19	Lt. Lingual	1.28	Lt. Mid. Cingulate	0.88

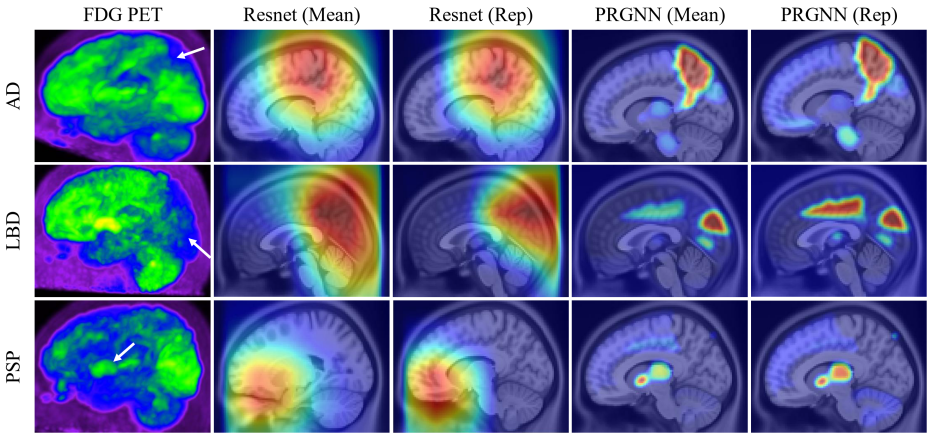

Fig. 2. Visualization of model inference. The "Mean" map displays the average across the test dataset, while the "Rep" map shows the activation of a representative image, whose FDG PET image is displayed in the left. These maps, derived from the final pooled node embeddings weighted by the fully-connected layer, are projected onto the AAL atlas to highlight region-specific contributions. For comparison, Grad-CAM results from ResNet-18 are visualized.

3.1 Evaluations

We followed standard 5-fold cross-validation setup, but performed an additional stratified 8:2 for train/val split within each training set. Accuracy (Acc.), Area under the receiver-operating characteristic curve (AUC), and F1 Score (F1) are reported for FDG classification, while sensitivity (Sens.) and specificity (Spec.) are reported in lieu of AUC and F1-score for FBB classification.

Qualitative Results. Table 2 presents the performance comparison of our PRGNN models against various baselines for FDG and FBB classification tasks. Our PRGNN-M model achieved the highest accuracy (92.6%) and AUC (98.6%) for FDG classification, outperforming ViG-B, the best-performing baseline. For FBB classification, PRGNN-M also attained the highest accuracy (97.5%) and

Table 4. Ablation study. Classification accuracies for truncated CNN backbones across four stages with and without node embedding-based graph convolution (GC). Stage 4 with GC corresponds to our PRGNN-M model.

	Stage 1	**Stage 2**	**Stage 3**	**Stage 4**
Without GC	44.2	61.4	75.2	89.3
With GC	82.4	83.3	84.5	**92.6**

specificity (98.4%), demonstrating superior robustness. Despite having significantly fewer parameters and lower computational cost compared to CNN-based models like ResNet-50 and transformer-based variants, our PRGNN models consistently achieved better classification performance, highlighting the effectiveness of our graph-based approach.

Model Interpretation. For interpretation of our model, we visualize node contribution, which is a measure of how much each node (ROI) contributes to the final prediction. In our approach, the model's final pooled node embeddings are multiplied by the following FC weights for each class to compute node contribution scores. These scores are then aggregated over all samples and folds. For a specific class (e.g., AD), the average contribution of each node (ROI) is calculated and mapped to the corresponding regions of the AAL atlas, as visualized in Figure 2. In addition, we provide the Grad-CAM visualization of the ResNet model for comparison of the visualization.

The visualization of node contributions closely aligns with characteristic FDG uptake patterns for each disease, highlighting the parietal and temporal regions in AD, the occipital region in LBD, and both the putamen and superior parietal regions in PSP. In contrast, while Grad-CAM analysis does provide relevant activations, the activation maps are coarse and do not explicitly indicate how much the activated areas contributed to the logits of the final prediction.

The top nodes selected by PRGNN for each class have been summarized in Table 3. Regions such as superior parietal and superior temporal exhibited high logit contribution for AD prediction [21], lingual and cuneus for LBD [22], and putamen and parietal for PSP [23]. Interestingly, thalamus appeared among the top five regions, which was also selected for AD classification in a previous work [7, 24].

3.2 Ablation Study

We evaluated our node embedding-based graph convolution (GC) across CNN stages with a seven-step ablation: four backbone-only models (Stages 1–4) followed by each stage augmented with GC. Table 4 reports the resulting accuracies, with Stage 4+GC (PRGNN-M) achieving the best performance. In every stage, GC consistently improved classification, highlighting the value of region-specific feature aggregation.

4 Conclusion

We propose PRGNN, a multi-scale and anatomy-aware graph neural network that integrates hierarchical feature extraction and graph-based modeling for brain PET image classification. Our approach effectively captures both local metabolic patterns and global anatomical relationships, outperforming existing CNN, transformer, and graph-based models in FDG and FBB PET classification. The interpretability analysis demonstrates that PRGNN identifies disease-relevant brain regions, validating the node-based approach for model interpretation. Future work will explore incorporating additional imaging modalities and refining the graph structure to further enhance diagnostic performance.

Acknowledgments. This work was supported by the Technology development Program (RS-2024-00508913) funded by the Ministry of SMEs and Startups (MSS, Korea), Testbed Seoul for Innovative Technology 2024 through Seoul Business Agency funded by Seoul Metropolitan Government (TE240215), and Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (RS-2024-00465306).

Disclosure of Interests. The authors have no competing interests.

References

1. Matej Perovnik, Tomaž Rus, Katharina A Schindlbeck, and David Eidelberg. Functional brain networks in the evaluation of patients with neurodegenerative disorders. *Nature Reviews Neurology*, 19(2):73–90, 2023.
2. Satoshi Minoshima, Donna Cross, Tanyaluck Thientunyakit, Norman L Foster, and Alexander Drzezga. 18f-fdg pet imaging in neurodegenerative dementing disorders: insights into subtype classification, emerging disease categories, and mixed dementia with copathologies. *Journal of nuclear medicine*, 63(Supplement 1):2S–12S, 2022.
3. Lyduine E Collij, Gemma Salvadó, Viktor Wottschel, Sophie E Mastenbroek, Pierre Schoenmakers, Fiona Heeman, Leon Aksman, Alle Meije Wink, Bart NM Berckel, Wiesje M van de Flier, et al. Spatial-temporal patterns of β-amyloid accumulation: a subtype and stage inference model analysis. *Neurology*, 98(17):e1692–e1703, 2022.
4. Shuyang Fan, Maria Rosana Ponisio, Pan Xiao, Sung Min Ha, Satrajit Chakrabarty, John J Lee, Shaney Flores, Pamela LaMontagne, Brian Gordon, Cyrus A Raji, et al. Amyloidpetnet: classification of amyloid positivity in brain pet imaging using end-to-end deep learning. *Radiology*, 311(3):e231442, 2024.
5. Kobra Etminani, Amira Soliman, Anette Davidsson, Jose R Chang, Begoña Martínez-Sanchis, Stefan Byttner, Valle Camacho, Matteo Bauckneht, Roxana Stegeran, Marcus Ressner, et al. A 3d deep learning model to predict the diagnosis of dementia with lewy bodies, alzheimer's disease, and mild cognitive impairment using brain 18f-fdg pet. *European journal of nuclear medicine and molecular imaging*, pages 1–22, 2022.
6. Jinseong Jang and Dosik Hwang. M3t: three-dimensional medical image classifier using multi-plane and multi-slice transformer. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 20718–20729, 2022.

7. Hongchao Jiang and Chunyan Miao. Anatomy-aware gating network for explainable alzheimer's disease diagnosis. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 90–100. Springer, 2024.
8. Brandon G Hill, Frances L Koback, and Peter L Schilling. The risk of shortcutting in deep learning algorithms for medical imaging research. *Scientific Reports*, 14(1):29224, 2024.
9. Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.
10. Satoshi Takahashi, Yusuke Sakaguchi, Nobuji Kouno, Ken Takasawa, Kenichi Ishizu, Yu Akagi, Rina Aoyama, Naoki Teraya, Amina Bolatkan, Norio Shinkai, et al. Comparison of vision transformers and convolutional neural networks in medical image analysis: a systematic review. *Journal of Medical Systems*, 48(1):84, 2024.
11. Xiaoxiao Li, Yuan Zhou, Nicha Dvornek, Muhan Zhang, Siyuan Gao, Juntang Zhuang, Dustin Scheinost, Lawrence H Staib, Pamela Ventola, and James S Duncan. Braingnn: Interpretable brain graph neural network for fmri analysis. *Medical Image Analysis*, 74:102233, 2021.
12. Kevin Mueller, Anke Meyer-Baese, and Gordon Erlebacher. Combining graph neural networks and roi-based convolutional neural networks to infer individualized graphs for alzheimer's prediction. In *Medical Imaging 2023: Biomedical Applications in Molecular, Structural, and Functional Imaging*, volume 12468, pages 25–30. SPIE, 2023.
13. Jaeyoon Sim, Minjae Lee, Guorong Wu, and Won Hwa Kim. Multi-modal graph neural network with transformer-guided adaptive diffusion for preclinical alzheimer classification. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 511–521. Springer, 2024.
14. Nathalie Tzourio-Mazoyer, Brigitte Landeau, Dimitri Papathanassiou, Fabrice Crivello, Octave Etard, Nicolas Delcroix, Bernard Mazoyer, and Marc Joliot. Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain. *Neuroimage*, 15(1):273–289, 2002.
15. Kai Han, Yunhe Wang, Jianyuan Guo, Yehui Tang, and Enhua Wu. Vision gnn: An image is worth graph of nodes. *Advances in neural information processing systems*, 35:8291–8303, 2022.
16. Guohao Li, Matthias Muller, Ali Thabet, and Bernard Ghanem. Deepgcns: Can gcns go as deep as cnns? In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 9267–9276, 2019.
17. Guy M McKhann, David S Knopman, Howard Chertkow, Bradley T Hyman, Clifford R Jack Jr, Claudia H Kawas, William E Klunk, Walter J Koroshetz, Jennifer J Manly, Richard Mayeux, et al. The diagnosis of dementia due to alzheimer's disease: recommendations from the national institute on aging-alzheimer's association workgroups on diagnostic guidelines for alzheimer's disease. *Alzheimer's & dementia*, 7(3):263–269, 2011.
18. Ian G McKeith, Bradley F Boeve, Dennis W Dickson, Glenda Halliday, John-Paul Taylor, Daniel Weintraub, Dag Aarsland, James Galvin, Johannes Attems, Clive G Ballard, et al. Diagnosis and management of dementia with lewy bodies: Fourth consensus report of the dlb consortium. *Neurology*, 89(1):88–100, 2017.
19. Günter U Höglinger, Gesine Respondek, Maria Stamelou, Carolin Kurz, Keith A Josephs, Anthony E Lang, Brit Mollenhauer, Ulrich Müller, Christer Nilsson, Jen-

nifer L Whitwell, et al. Clinical diagnosis of progressive supranuclear palsy: the movement disorder society criteria. *Movement disorders*, 32(6):853–864, 2017.

20. Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
21. Heidi IL Jacobs, Martin PJ Van Boxtel, Jelle Jolles, Frans RJ Verhey, and Harry BM Uylings. Parietal cortex matters in alzheimer's disease: an overview of structural, functional and metabolic findings. *Neuroscience & Biobehavioral Reviews*, 36(1):297–309, 2012.
22. Jennifer L Whitwell, Jonathan Graff-Radford, Tarun D Singh, Daniel A Drubach, Matthew L Senjem, Anthony J Spychalla, Nirubol Tosakulwong, Val J Lowe, and Keith A Josephs. 18f-fdg pet in posterior cortical atrophy and dementia with lewy bodies. *Journal of Nuclear Medicine*, 58(4):632–638, 2017.
23. Philipp T Meyer, Lars Frings, Gerta Rücker, and Sabine Hellwig. 18f-fdg pet in parkinsonism: differential diagnosis and evaluation of cognitive impairment. *Journal of Nuclear Medicine*, 58(12):1888–1898, 2017.
24. Laura W de Jong, Karin van der Hiele, Ilya M Veer, JJ Houwing, RGJ Westendorp, ELEM Bollen, Paul W de Bruin, HAM Middelkoop, Mark A van Buchem, and Jeroen van der Grond. Strongly reduced volumes of putamen and thalamus in alzheimer's disease: an mri study. *Brain*, 131(12):3277–3285, 2008.