applied sciences

MDPI

Article

A 3D Global-Patch Transformer for Brain Age Prediction Using T1-Weighted MRI with Gray and White Matter Maps

Seung-Jun Lee [1], Myungeun Lee [2], Yoo Ri Kim [3,4] and Hyung-Jeong Yang [1,2,*]

[1] Department of Artificial Intelligence Convergence, Chonnam National University, Gwangju 61186, Republic of Korea
[2] Hyper-Wide Federated Medical AI Research Center, Chonnam National University, Gwangju 61186, Republic of Korea
[3] Department of Cardiovascular Medicine, Chonnam National University Hospital, Gwangju 61469, Republic of Korea
[4] Division of Cardiology, Department of Internal Medicine, Chonnam National University Medical School, Gwangju 61469, Republic of Korea
* Correspondence: hjyang@jnu.ac.kr

Abstract

With the increasing prevalence of neurodegenerative diseases driven by population aging, imaging-based biomarkers are needed to quantify brain aging at an early stage. Brain age, which estimates structural brain aging relative to chronological age, has emerged as a useful indicator. Prior work has mainly used T1-weighted MRI with deep learning models such as convolutional neural networks (CNNs) or transformers; however, many approaches insufficiently capture three-dimensional structural continuity and localized anatomical patterns, and tissue-specific aging in gray matter (GM) and white matter (WM) is often treated as auxiliary. To address these limitations, we propose a 3D Global–Patch Transformer framework for brain age prediction that directly processes volumetric data while jointly learning global brain structure and local anatomical features. Our model runs global and patch pathways in parallel and explicitly incorporates GM and WM structural maps alongside T1-weighted MRI to encode tissue-specific aging signals. Experiments on multiple public datasets, including IXI and OASIS, show that the proposed method reduces mean absolute error (MAE) by approximately 10–15% compared with CNN-based and single-input transformer baselines, with notably improved performance in older populations, highlighting the value of tissue-level structural information for brain age estimation.

Keywords: magnetic resonance imaging (MRI); brain age prediction; deep learning; global–patch transformer; neuroimaging

Check for updates

Academic Editors: Valeria Sacca and Fabiana Novellino

Received: 27 February 2026
Revised: 16 March 2026
Accepted: 17 March 2026
Published: 20 March 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Population aging has led to a steady increase in the prevalence of neurodegenerative diseases such as Alzheimer's and Parkinson's disease, motivating the development of imaging-based biomarkers that can characterize brain aging at an early stage. Brain age estimation, which quantifies the degree of structural brain aging relative to chronological age, has emerged as a promising biomarker for assessing abnormal aging patterns and potential neurodegenerative risk [1–3]. Structural magnetic resonance imaging (MRI), particularly T1-weighted MRI (T1w MRI), has been widely adopted for this task due to its ability to capture macroscopic anatomical changes associated with aging, including cortical thinning and volumetric alterations in gray matter (GM) and white matter (WM) [4,5].

Recent advances in deep learning have substantially improved brain age prediction performance. Convolutional neural network (CNN)–based approaches using 3D T1w MRI have demonstrated strong predictive accuracy across large cohorts [3,6–15]. More recently, attention mechanisms and transformer-based models have been introduced to better capture long-range dependencies and region-specific relevance [16,17]. Despite these advances, a large portion of existing studies process 3D MRI volumes as collections of 2D slices, which may weaken inter-slice spatial continuity and limit the modeling of volumetric anatomical context. Furthermore, such slice-based formulations can introduce plane-dependent biases, potentially affecting the stability and interpretability of learned representations.

Recent studies have increasingly adopted 3D Vision Transformers and hybrid architectures to capture complex aging trajectories [18–20]. However, directly extending massive multiscale 3D transformers to simultaneously process three distinct volumetric inputs (T1, GM, WM) imposes prohibitive GPU memory constraints. Therefore, our model preserves an efficient CNN extractor and utilizes a multi-head cross-attention bottleneck to circumvent the $O(N^2)$ memory explosion.

Another important limitation of existing brain age prediction frameworks lies in the utilization of tissue-derived structural information. Gray matter and white matter exhibit distinct and complementary aging trajectories, providing valuable cues for brain age estimation [4,8,11,21]. However, in many prior studies, GM and WM information is either indirectly inferred from T1w MRI or incorporated as auxiliary covariates alongside the primary imaging input. As a result, tissue-specific aging patterns may not be fully exploited during model learning, particularly when global brain context and localized anatomical changes need to be considered jointly.

We propose a 3D Global–Patch Transformer framework for brain age prediction that directly operates on volumetric MRI data while explicitly integrating GM and WM structural maps as primary inputs to address these limitations. The proposed framework is designed to preserve full three-dimensional spatial continuity and to simultaneously model global brain structure and localized anatomical patterns. The model enables effective interaction between whole-brain context and region-level features by employing parallel global and patch pathways within a unified 3D architecture. GM and WM maps are combined with T1w MRI using a simple and stable fusion strategy, allowing tissue-specific aging signals to be incorporated without introducing additional fusion parameters.

The main contributions of this work are summarized as follows:

- **Fully 3D global–local modeling:** We present a brain age prediction framework that processes MRI data entirely in three dimensions, preserving volumetric spatial continuity and reducing limitations associated with slice-based approaches.
- **Explicit use of tissue-specific structural information**: GM and WM structural maps are incorporated as primary inputs rather than auxiliary features, enabling direct learning of tissue-specific aging patterns and their complementary roles in brain age estimation.
- **Global–patch transformer architecture:** A parallel global and patch pathway design is adopted to jointly capture whole-brain structural characteristics and localized anatomical variations within a unified transformer-based framework.
- **Comprehensive empirical evaluation:** The proposed method is evaluated on multiple public datasets using a unified preprocessing pipeline, and extensive ablation studies are conducted to analyze architectural choices and input configurations.

The remainder of this paper is organized as follows. Section 2 reviews related work on brain age prediction and deep learning–based neuroimaging analysis. Section 3 describes the proposed 3D Global–Patch Transformer framework in detail. Section 4 presents the

experimental setup, datasets, evaluation metrics, and comprehensive experimental analyses. Finally, Section 5 concludes the paper by summarizing the findings and directions for future work.

2. Related Work

2.1. Brain Age Prediction Using Structural MRI

Brain age prediction has been widely studied as an imaging-based biomarker for characterizing brain aging and identifying abnormal aging patterns. Early studies primarily relied on handcrafted features extracted from structural MRI, while more recent approaches have leveraged deep learning models to directly learn representations from T1-weighted MRI (T1w MRI). In particular, three-dimensional convolutional neural networks (3D CNNs) have demonstrated strong performance by modeling volumetric anatomical information across the brain [3,6–15]. Large-scale studies have shown that CNN-based frameworks trained on T1w MRI can achieve reliable brain age estimation across wide age ranges and diverse cohorts [9–11,13].

Despite their effectiveness, CNN-based methods are inherently limited in capturing long-range dependencies and complex global relationships across distant brain regions. To address this limitation, attention mechanisms and transformer-based architectures have recently been introduced into brain age prediction tasks. These models aim to improve feature representation by selectively emphasizing age-relevant regions and modeling interactions between spatially separated anatomical structures [16,17].

2.2. Global–Local Modeling and Transformer-Based Approaches

Global–local modeling strategies have gained attention as a means to jointly capture whole-brain context and localized anatomical patterns [22]. He et al. [16] proposed a global–local transformer network (GLTN) for brain age estimation, demonstrating that combining global contextual features with local patch-level representations can improve prediction accuracy. Related transformer-based approaches [23] have further explored attention mechanisms to enhance robustness and interpretability in brain age prediction [17].

However, many of these transformer-based methods process 3D MRI volumes as collections of 2D slices. While slice-based formulations reduce computational complexity, they may compromise volumetric spatial continuity and weaken the modeling of three-dimensional anatomical context. Such limitations can lead to plane-dependent representations and restrict the effective integration of global and local structural information when operating on inherently volumetric brain data.

2.3. Incorporation of Gray and White Matter Information

Several studies have investigated the use of tissue-derived structural information to improve brain age prediction. Gray matter (GM) and white matter (WM) exhibit distinct aging trajectories, including cortical thinning and white-matter degeneration, which provide important cues for estimating brain age [4,8,11,21]. Prior work has shown that incorporating GM and WM information alongside T1w MRI can enhance prediction performance [8,11,21].

In most existing frameworks, GM and WM features are treated as auxiliary inputs or are implicitly inferred from T1w MRI rather than being explicitly modeled as primary structural signals. As a result, tissue-specific aging patterns may not be fully exploited, particularly in architectures that do not preserve three-dimensional spatial structure or that separate global and local representations without unified modeling.

Previous studies demonstrate the effectiveness of deep learning–based approaches for brain age prediction and highlight the potential benefits of attention mechanisms and tissue-

derived structural information. However, there remains a need for a unified framework that (i) preserves full three-dimensional spatial continuity, (ii) jointly models global brain structure and localized anatomical patterns, and (iii) explicitly leverages GM and WM information as primary structural inputs. These considerations motivate the proposed 3D Global–Patch Transformer framework described in the following section.

3. Proposed Method

This section introduces the proposed 3D Global–Patch Transformer framework for brain age prediction from structural MRI. The proposed method is designed to address two key technical challenges in brain age estimation: (i) preserving three-dimensional spatial continuity when modeling volumetric brain anatomy and (ii) effectively integrating tissue-specific structural information that reflects heterogeneous aging patterns across gray matter (GM) and white matter (WM). To this end, the framework directly operates on 3D MRI volumes and jointly models global brain structure and localized anatomical patterns within a unified architecture.

As illustrated in Figure 1, the model takes T1-weighted MRI (T1w MRI) together with GM and WM structural maps as inputs. The proposed framework explicitly incorporates these tissue-derived maps as primary structural inputs, unlike conventional approaches that implicitly infer tissue information from raw intensity or treat GM/WM as auxiliary covariates. This design choice enables the model to directly capture tissue-specific aging characteristics while maintaining consistent representation and learning across different structural modalities.

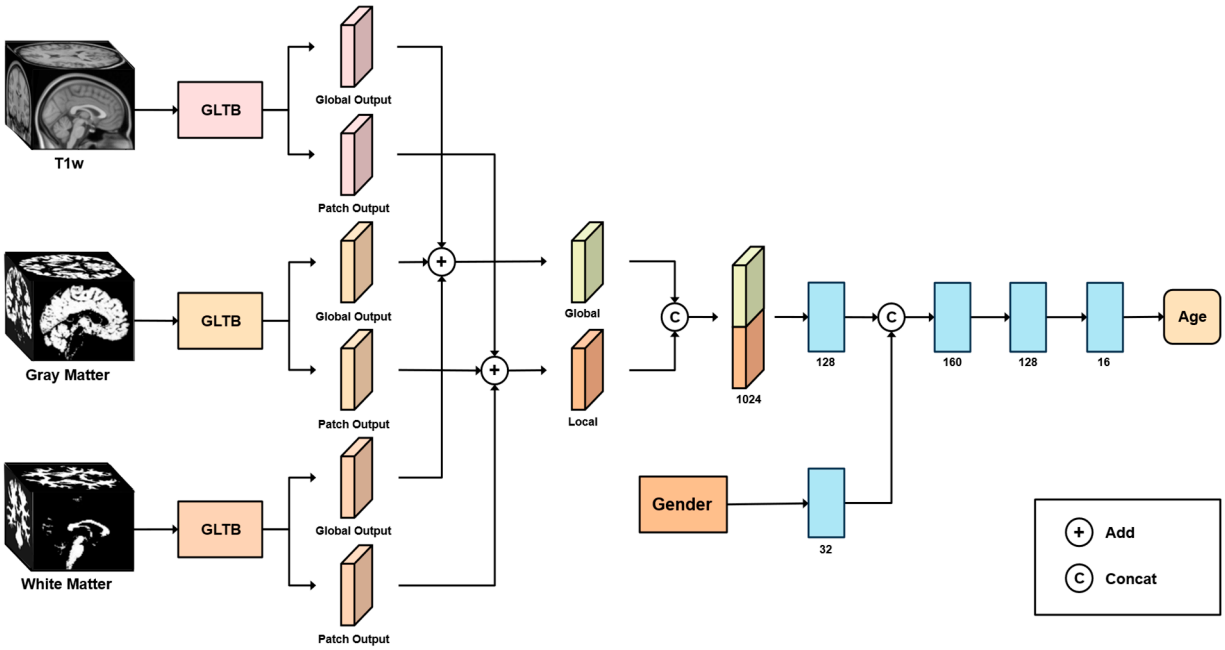


Figure 1. Overview of the Proposed Model Architecture.

The proposed framework adopts a global–patch strategy to capture complementary information at different spatial scales. A global representation encodes whole-brain anatomical context and long-range spatial dependencies, while patch-level representations focus on fine-grained local structural variations extracted from overlapping volumetric regions. These representations are coupled through transformer-based attention mechanisms, allowing localized features to be interpreted in the context of global brain structure. The overall architecture is optimized for stable training and computational efficiency in high-dimensional 3D settings, providing a practical and scalable solution for volumetric brain

age prediction. The following subsections describe the individual components of the proposed framework in detail, including the 3D global–patch modeling strategy, the feature extraction and transformer design, and the integration of structural and demographic information together with the training objective.

3.1. 3D Global–Patch Modeling

The proposed framework adopts a 3D global–patch modeling strategy to jointly capture whole-brain structural context and localized anatomical variations from volumetric MRI data. Given that age-related brain changes manifest at multiple spatial scales, ranging from global volumetric atrophy to region-specific tissue degeneration, the model is designed to explicitly represent both global and local information within a unified three-dimensional architecture.

A global representation is first extracted to encode overall brain morphology and long-range spatial dependencies across the entire 3D field of view for each structural input volume. This global feature serves as a compact summary of whole-brain structure and provides contextual guidance for interpreting localized patterns. In parallel, the same input volume is partitioned into overlapping three-dimensional patches, enabling the model to focus on fine-grained anatomical details while preserving spatial continuity between neighboring regions.

Patch-level representations are obtained independently for each volumetric patch and subsequently refined through interaction with the global representation. Specifically, a transformer-based attention mechanism is employed to model global–local interactions, where patch features attend to the global feature to selectively incorporate whole-brain contextual information. Localized structural variations can be interpreted in relation to global anatomical trends, which is particularly important for disambiguating region-specific aging patterns that may appear similar in isolation.

This global–patch interaction is implemented in a fully 3D manner, avoiding slice-based processing and maintaining volumetric consistency throughout the network. By coupling global and patch representations within the transformer module, the proposed modeling strategy enables effective information exchange across spatial scales while retaining computational efficiency for high-dimensional MRI data. The resulting global–local representation forms the basis for subsequent feature refinement and structural integration stages described in the following subsections. This global–patch modeling is implemented as a unified Global–Local Transformer Block (GLTB).

3.2. Feature Extraction and Transformer Design

The proposed framework employs a 3D feature extraction backbone followed by a transformer-based global–patch attention module to effectively represent complex anatomical structures in volumetric MRI data. The design of this component is guided by two considerations: (i) preserving volumetric spatial context while capturing cross-channel interactions and (ii) enabling efficient global–local information exchange within a unified representation.

We adopt a 3D convolutional architecture inspired by ConvNeXt [24] for feature extraction. While ConvNeXt was originally developed for 2D image recognition, its core design principles—including large receptive fields, hierarchical feature representation, and simplified block structure—are adapted to three-dimensional MRI data. In particular, depthwise convolutions are replaced with standard 3D convolutions to jointly model spatial and inter-channel relationships, which are critical for capturing volumetric anatomical patterns in brain MRI. The backbone consists of multiple stages with progressively increasing channel dimensions and receptive fields, allowing the network to encode both low-level structural details and higher-level anatomical abstractions.

The extracted volumetric features are subsequently processed by a transformer module designed to model interactions between global and patch-level representations. Therefore, the global feature serves as a contextual reference, while patch-level features act as localized descriptors. Patch features are treated as queries, and the global representation is used as keys and values within a multi-head attention mechanism.

The interaction within the GLTB is strictly based on a multi-head cross-attention mechanism, where localized patch features serve exclusively as queries (Q), and the global representation acts as keys (K) and values (V).

Let $Z_{patch} \in R^{N \times d}$ denote the sequence of localized patch features, where N is the number of patches and d is the feature dimension. Similarly, let $Z_{global} \in R^{1 \times d}$ represent the global feature vector encapsulating the whole-brain context. These features are projected into query (Q), key (K), and value (V) spaces through learnable linear projections:

$$Q = Z_{patch} W_Q,\ K = Z_{global} W_K,\ V = Z_{global} W_V \tag{1}$$

where $W_Q, W_K, W_V \in R^{d \times d_k}$ are the projection matrices, and d_k is the dimension of the key vectors. The cross-attention operation is then formalized as:

$$Attention(Q,\ K,\ V) = softmax\left(\frac{QK^T}{d_k}\right)V \tag{2}$$

The dot product QK^T computes the relevance scores between each local patch and the global context. The scaling factor d_k is applied to prevent gradient vanishing in the softmax function. By using the global feature as the key and value, each patch representation is refined by explicitly attending to the global anatomical structure, allowing the model to interpret localized aging signals (e.g., cortical thinning in a specific region) within the context of overall brain atrophy.

Multi-head attention extends this mechanism by performing the attention operation h times in parallel, allowing the model to jointly attend to information from different representation subspaces. The outputs of the attention heads are concatenated and linearly projected to form the final refined patch representations:

$$MHCA(Z_{patch,}\ \ Z_{global}) = Concat(head_1, \ \ldots\ , head_h)W_O \tag{3}$$

where $head_i = Attention\left(Z_{patch} W_Q^{(i)},\ Z_{global} W_K^{(i)},\ Z_{global} W_V^{(i)}\right)$ and $W_O \in R^{hd_k \times d}$ is the output projection matrix.

The attention outputs are passed through position-wise feed-forward layers with residual connections and normalization to stabilize training and enhance representational capacity. Multiple transformer blocks are stacked to progressively refine the global–local representations, allowing iterative integration of contextual information across spatial scales. Importantly, all operations are performed in a fully three-dimensional manner, preserving volumetric consistency and avoiding the loss of inter-slice relationships inherent to slice-based approaches.

This module provides a flexible and expressive representation of brain structure that captures both distributed anatomical context and localized aging-related variations by combining a 3D convolutional backbone with a global–patch transformer design. The resulting features serve as the foundation for subsequent structural integration and regression stages described in the following subsection.

3.3. Structural Integration and Training Objective

Gray matter (GM) and white matter (WM) maps derived from T1-weighted MRI are integrated into the proposed framework as primary inputs to explicitly incorporate tissue-specific structural information. Each structural input, such as T1w MRI, GM, and WM, is processed independently through the same 3D feature extraction and global–patch transformer pipeline described in the previous subsections. This shared-parameter design ensures consistent representation learning across different structural modalities while avoiding modality-specific bias.

The resulting global–local representations from the T1w, GM, and WM streams are combined using an element-wise addition operation. This fusion strategy is intentionally chosen to provide a simple and stable mechanism for aggregating complementary tissue information without introducing additional trainable parameters. This fusion strategy does not introduce additional trainable parameters, thereby preserving computational efficiency and stable optimization when handling high-dimensional 3D volumetric features. Given the high dimensionality of volumetric MRI features, this design helps maintain computational efficiency and stable optimization while allowing tissue-specific aging signals to be directly integrated into a unified representation.

Demographic information is incorporated as auxiliary input to provide complementary contextual cues for brain age estimation in addition to imaging-derived features. Specifically, gender is encoded as a low-dimensional vector and concatenated with the fused global–local representation prior to the regression layers. This design allows demographic factors to contribute to the prediction while ensuring that structural MRI features remain the dominant source of information. The contribution of demographic information is further examined in the experimental analysis through feature-combination studies.

The fused representation is projected to a compact latent space via fully connected layers, resulting in a 128-dimensional feature embedding that summarizes both global and localized structural characteristics. The 128-dimensional feature embedding is concatenated with the encoded demographic information and further processed by a fully connected layer to produce a 160-dimensional representation for final brain age regression. A regression head is then applied to estimate brain age as a continuous scalar output. During training, the framework is optimized using a composite loss function designed to balance absolute prediction accuracy and relative ranking consistency across subjects. Mean squared error (MSE) is employed as the primary regression objective, minimizing the discrepancy between predicted brain age and chronological age. In addition, a rank-based loss inspired by Spearman's rank correlation coefficient (SRCC) is incorporated to encourage consistent ordering of predictions. To enable gradient-based optimization, a differentiable approximation of the ranking operation is adopted following the SoDeep [12] framework.

The Combined Loss function is mathematically defined as follows:

Let N denote the batch size during training. For the i-th subject in a batch, let y_i be the chronological age and $\hat{y}_i$ be the predicted brain age. The training objective consists of two components: a regression loss to minimize prediction error and a ranking loss to enforce ordinal consistency. The Mean Squared Error (MSE) loss is employed as the primary regression objective and is defined as:

$$\mathcal{L}_{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{4}$$

This term ensures that the predicted brain age values are quantitatively close to the ground-truth chronological ages.

We incorporate a differentiable ranking loss to enforce ordinal consistency among predictions. We also utilize a pretrained proxy network (a continuously differentiable sorter) to approximate the ranks, following the approach proposed in [12] since the standard rank operator is non-differentiable (due to the sorting operation). The proxy-based ranking objective was adopted to provide a differentiable approximation to the sorting operation, thereby enabling end-to-end optimization of relative age ordering within the proposed 3D framework. In contrast to conventional margin-based pairwise ranking formulations, this design is intended to preserve broader ordinal relationships among predicted ages without depending solely on a manually specified margin threshold. By incorporating the proxy network into the ranking formulation, the model can optimize age-order consistency jointly with the regression objective in a differentiable manner.

Let $y = [y_1, \ldots, y_N]^T$ and $\hat{y} = [\hat{y}_1, \ldots, \hat{y}_N]^T$ denote the vectors of chronological ages and predicted brain ages in a batch, respectively. The differentiable ranking loss is defined as the mean absolute difference between the ground-truth ranks and the approximated predicted ranks:

$$\mathcal{L}_{rank} = \frac{1}{N} \sum_{i=1}^{N} \left| r(y_i) - \hat{r}(\hat{y}_1) \right| \quad (5)$$

where $r(y_i)$ represents the fixed ground-truth rank of the i-th subject within the batch, and $\hat{r}(\cdot)$ denotes the output of the pre-trained proxy network that approximates the ranking operation. This formulation allows gradients to be back-propagated through the ranking step, encouraging the model to preserve the relative order of brain ages even when absolute values fluctuate.

The final objective function combines these two terms:

$$\mathcal{L}_{Total} = (1 - \alpha) L_{MSE} + a L_{Rank} \quad (6)$$

where α is a hyperparameter balancing absolute error minimization and ranking consistency.

Although this formulation provides a practical approximation to the ranking objective, the present study did not systematically evaluate the sensitivity of final performance to the pre-training condition or initialization of the proxy network. Therefore, the robustness of the ranking component with respect to proxy-network configuration remains an important topic for future investigation.

4. Experiments and Analysis

This section presents the experimental evaluation of the proposed 3D Global–Patch Transformer for brain age prediction. Experiments are conducted using four publicly available brain MRI datasets under a unified preprocessing and training protocol. The proposed method is evaluated through quantitative comparison with representative convolutional neural network (CNN)–based and transformer-based approaches reported in prior brain age prediction studies.

4.1. Datasets, Preprocessing, and Experimental Setup

Experiments are conducted using four publicly available T1-weighted brain MRI datasets: ABIDE [25], CoRR [26], IXI [27], and OASIS-3 [28]. Only healthy subjects are included in all experiments. The number of samples, age ranges, and gender distributions for each dataset are summarized in Table 1. For the multi-site datasets ABIDE and CoRR, site-wise dataset compositions are further reported in Tables 2 and 3.

Table 1. Datasets.

Dataset	N_{Sample}	$N_{Subject}$	Age Range	Gender (Female/Male)
ABIDE [25]	1070	1070	7–39	163/939
CoRR [26]	1378	1378	6–88	684/694
IXI [27]	527	527	21–74	292/235
OASIS-3 [28]	2214	1057	42–96	588/469

Table 2. ABIDE Datasets.

Dataset	N_{Sample}	Age Range	Gender (Female/Male)
NYU	184	6–39	37/147
UM	145	8–29	28/117
USM	101	9–50	0/101
UCLA	99	8–18	12/87
Leuven	64	12–32	8/56
Pitt	57	9–35	8/49
MaxMun	57	7–58	7/50
Yale	56	7–18	16/40
KKI	55	8–13	13/42
Trinity	49	12–26	0/49
Stanford	40	8–13	8/32
Caltech	38	17–56	8/30
SDSU	36	9–17	7/29
Olin	36	10–24	5/31
SBL	30	20–64	0/30
OHSU	28	8–15	0/28
CMU	27	19–40	6/21
Overall	1102	6–64	163/939

All MRI scans are processed using a unified preprocessing pipeline implemented with SPM [29]. Each T1-weighted image is nonlinearly registered to the MNI152 template and resampled to a voxel resolution of 2 mm, resulting in standardized volumes of size $79 \times 95 \times 79$. Gray matter (GM) and white matter (WM) structural maps are derived from the preprocessed T1-weighted images using tissue segmentation provided by SPM, following the same procedure across all datasets. This unified SPM-based preprocessing was adopted to ensure consistency across datasets; however, robustness to alternative preprocessing settings and more heterogeneous clinical acquisition protocols warrants further investigation.

Five-fold cross-validation is employed for all experiments unless otherwise specified. For datasets containing multiple scans per subject, such as OASIS-3, subject-level splitting is applied to ensure that images from the same subject do not appear in both training and evaluation sets.

Table 3. CoRR Dataset.

Dataset	N Sample	Age Range	Gender (Female/Male)
SWU	304	17–27	142/162
IPCAS	226	6–62	102/124
NYU	212	6–55	126/86
BNU	136	18–30	70/66
UPSM	99	10–20	51/48
LMU	92	20–88	47/45
UM	80	55–84	22/58
BMB	50	20–60	24/26
IBATRT	36	19–48	18/18
HNU	30	20–30	15/15
IACAS	28	19–43	13/15
Utah	27	8–39	27/0
UWM	25	21–32	14/11
XHCUMS	23	36–62	13/10
NKI	10	22–57	0/10
Overall	1378	6–88	684/694

All models are trained and evaluated using identical data splits and preprocessing settings to ensure a fair comparison. Model training follows the configuration described in Section 3. The patch size was empirically set to $48 \times 48 \times 48$ with a stride of 12. Optimization is performed using the Adam optimizer. To mitigate overfitting, we applied a dropout rate of 0.5 within the transformer modules, utilized strictly held-out validation sets for early stopping, and ensured no data leakage occurred during hyperparameter tuning. Mean absolute error (MAE) is used as the primary evaluation metric, and Pearson correlation coefficient (PCC) is additionally reported to assess the correlation between predicted and chronological age.

4.2. Quantitative Results and Comparison with Existing Methods

The quantitative performance of the proposed method is summarized in Table 4. Experiments are conducted on the ABIDE, CoRR, IXI, and OASIS datasets using five-fold cross-validation under the same preprocessing and training protocol. Performance is evaluated using mean absolute error (MAE) and Pearson correlation coefficient (PCC).

The mathematical formulation of MAE is given as:

$$MAE = \frac{1}{N_{test}} \sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{7}$$

where y_i is the actual value, $\hat{y}_i$ is the predicted value, and N_{test} denotes the total number of samples in the test set.

PCC evaluates the strength and direction of the linear relationship between two variables, providing a measure of the correlation between predicted and actual values. The coefficient ranges from -1 to 1, where values closer to 1 indicate a stronger positive linear relationship.

The mathematical definition of PCC is as follows:

$$PCC = \frac{\sum_{i=1}^{n}(y_i - \overline{y})(\hat{y}_i - \hat{\overline{y}})}{\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}\sqrt{\sum_{i=1}^{n}(\hat{y}_i - \hat{\overline{y}})^2}} \quad (8)$$

Table 4. Performance comparison on ABIDE, CoRR, IXI, and OASIS.

Model \ Metrics	ABIDE		CoRR		IXI		OASIS	
	MAE ↓	PCC ↑	MAE ↓	PCC ↑	MAE ↓	PCC ↑	MAE ↓	PCC ↑
3D CNN	3.61 ± 0.41	0.61 ± 0.10	4.36 ± 0.15	0.83 ± 0.11	14.03 ± 0.57	0.39 ± 0.08	6.09 ± 0.14	0.69 ± 0.02
3D ResNet	2.56 ± 0.32	0.85 ± 0.04	4.15 ± 0.68	0.89 ± 0.04	6.48 ± 0.64	0.81 ± 0.06	2.97 ± 0.19	0.89 ± 0.01
3D DenseNet	3.34 ± 0.30	0.63 ± 0.04	9.06 ± 1.12	0.48 ± 0.11	6.12 ± 0.21	0.75 ± 0.12	3.22 ± 0.14	0.87 ± 0.02
Joo et al. [21]	3.15 ± 0.70	0.67 ± 0.14	7.52 ± 2.08	0.63 ± 0.19	5.94 ± 0.87	0.79 ± 0.16	3.17 ± 0.11	0.88 ± 0.01
SFCN [11]	3.24 ± 0.69	0.67 ± 0.17	5.64 ± 0.65	0.83 ± 0.04	6.65 ± 0.71	0.81 ± 0.06	2.87 ± 0.15	0.90 ± 0.02
GLTN [16]	2.47 ± 0.22	0.83 ± 0.03	3.89 ± 0.34	0.91 ± 0.02	6.47 ± 0.61	0.85 ± 0.02	2.75 ± 0.06	0.91 ± 0.01
Global	2.32 ± 0.18	0.83 ± 0.08	3.77 ± 0.22	0.94 ± 0.01	5.06 ± 0.61	0.88 ± 0.02	2.67 ± 0.17	0.91 ± 0.02
Local	2.39 ± 0.13	0.86 ± 0.01	3.72 ± 0.14	0.93 ± 0.01	4.73 ± 0.45	0.90 ± 0.03	2.63 ± 0.20	0.91 ± 0.02
Proposed Model	2.28 ± 0.22	0.86 ± 0.02	3.58 ± 0.23	0.93 ± 0.03	4.46 ± 0.25	0.91 ± 0.01	2.59 ± 0.16	0.92 ± 0.01

MAE: Mean Absolute Error, PCC: Pearson Correlation Coefficient; Organized as MAE: (Mean ± std) through 5-fold cross-validation.

Here, $\overline{y}$ and $\hat{\overline{y}}$ represent the mean of the actual and redicted values, respectively.

MAE was used to evaluate the model's accuracy by quantifying the magnitude of absolute prediction errors, whereas PCC was employed to assess the strength of the linear relationship between the predicted and actual values. Together, these metrics offer complementary perspectives on model performance, encompassing both error magnitude and predictive correlation.

The proposed 3D Global–Patch Transformer consistently achieves lower MAE and comparable or higher PCC values across all four datasets compared with representative CNN-based and transformer-based baseline methods, including 3D CNN, 3D ResNet, 3D DenseNet, SFCN [11], and GLTN [16]. On the ABIDE dataset, the proposed method achieves an MAE of 2.28 ± 0.22 with a PCC of 0.86 ± 0.02. On the CoRR dataset, an MAE of 3.58 ± 0.23 and a PCC of 0.93 ± 0.03 are obtained. For the IXI dataset, the proposed model reports an MAE of 4.46 ± 0.25 with a PCC of 0.91 ± 0.01, while on the OASIS dataset, an MAE of 2.59 ± 0.16 and a PCC of 0.92 ± 0.01 are achieved.

The proposed method yields lower MAE across all datasets while maintaining similar levels of correlation compared to slice-based transformer approaches such as GLTN [16]. Our proposed framework also shows consistent improvements in absolute prediction error, particularly on datasets with broader age distributions relative to conventional 3D CNN-based architectures, including ResNet and DenseNet variants.

In addition, Table 4 reports the performance of global-only and local-only variants of the proposed architecture. While both variants achieve competitive results, neither matches the performance of the full global–patch configuration, indicating that jointly modeling global and localized anatomical information contributes to improved brain age estimation accuracy.

4.3. Computational Complexity

Regarding computational complexity, the proposed model contains approximately 1.58 billion parameters. Despite this massive scale, the cross-attention bottleneck effectively restricts the estimated peak GPU memory consumption per sample to approximately 9.95 GB. This O(N) attention complexity allows the model to be trained on stan-

dard high-end GPUs without the severe memory explosion typical of standard 3D self-attention mechanisms.

Although the overall model remains computationally demanding due to simultaneous processing of T1-weighted MRI, GM, and WM volumes, the asymmetric global-local cross-attention design constrains contextual interaction to global-guided local refinement and thereby avoids the severe growth in memory usage typically associated with dense volumetric self-attention.

4.4. Ablation and Structural Input Analysis

A series of ablation experiments are conducted to analyze the effects of architectural design choices and input configurations. The results are summarized in Tables 5 and 6. Unless otherwise specified, all ablation experiments are performed on the IXI dataset using the same preprocessing and training protocol described in Section 4.1.

Table 5. Experiment on Parameter α.

	IXI	
Parameter \ **Metrics**	**MAE ↓**	**PCC ↑**
$\alpha = 0.1$	4.46 ± 0.25	0.91 ± 0.01
$\alpha = 0.5$	4.69 ± 0.37	0.90 ± 0.02
$\alpha = 1$	4.83 ± 0.56	0.90 ± 0.02

MAE: Mean Absolute Error, PCC: Pearson Correlation Coefficient; Organized as MAE: (Mean ± std) through 5-fold cross-validation.

Table 6. Experimental Results Based on Different Feature Combinations.

	ABIDE		CoRR		IXI		OASIS	
Model \ **Metrics**	**MAE ↓**	**PCC ↑**	**MAE ↓**	**PCC ↑**	**MAE ↓**	**PCC ↑**	**MAE ↓**	**PCC ↑**
w/T1	2.94 ± 0.19	0.75 ± 0.06	4.67 ± 0.47	0.84 ± 0.08	6.45 ± 0.46	0.81 ± 0.01	3.02 ± 0.07	0.89 ± 0.01
w/T1, GM	2.38 ± 0.01	0.83 ± 0.03	4.04 ± 0.30	0.89 ± 0.01	4.98 ± 0.57	0.89 ± 0.03	2.76 ± 0.25	0.90 ± 0.02
w/T1, WM	2.59 ± 0.12	0.84 ± 0.01	4.08 ± 0.45	0.93 ± 0.01	5.24 ± 0.28	0.87 ± 0.01	2.84 ± 0.09	0.90 ± 0.01
w/T1, Gender	2.80 ± 0.09	0.76 ± 0.04	4.45 ± 0.37	0.90 ± 0.05	6.55 ± 0.46	0.80 ± 0.01	3.05 ± 0.11	0.89 ± 0.02
w/T1, GM, WM	2.38 ± 0.16	0.85 ± 0.01	3.82 ± 0.30	0.93 ± 0.03	4.79 ± 0.47	0.89 ± 0.01	2.76 ± 0.08	0.91 ± 0.01
w/T1, GM, Gender	2.38 ± 0.15	0.86 ± 0.01	4.04 ± 0.21	0.89 ± 0.01	4.98 ± 0.35	0.89 ± 0.01	2.76 ± 0.10	0.90 ± 0.01
w/T1, WM, Gender	2.56 ± 0.11	0.83 ± 0.01	4.02 ± 0.31	0.93 ± 0.02	5.24 ± 0.52	0.84 ± 0.02	2.81 ± 0.14	0.89 ± 0.01
w/GM	2.37 ± 0.09	0.86 ± 0.02	3.79 ± 0.53	0.89 ± 0.09	4.79 ± 0.41	0.89 ± 0.03	2.72 ± 0.12	0.90 ± 0.01
w/WM	2.49 ± 0.18	0.84 ± 0.02	3.90 ± 0.27	0.91 ± 0.01	5.33 ± 0.58	0.86 ± 0.02	2.79 ± 0.15	0.91 ± 0.02
w/GM, WM	2.33 ± 0.18	0.86 ± 0.01	3.68 ± 0.18	0.94 ± 0.01	4.46 ± 0.25	0.91 ± 0.01	2.59 ± 0.16	0.91 ± 0.01
Proposed Model	2.28 ± 0.22	0.86 ± 0.02	3.58 ± 0.23	0.93 ± 0.03	4.46 ± 0.25	0.91 ± 0.01	2.59 ± 0.16	0.92 ± 0.01

MAE: Mean Absolute Error, PCC: Pearson Correlation Coefficient; Organized as MAE: (Mean ± std) through 5-fold cross-validation.

Table 5 reports the effect of the loss-weight parameter α in the composite training objective. When α is set to 0.1, the proposed model achieves the lowest MAE of 4.46 ± 0.25 with a PCC of 0.91 ± 0.01. Increasing α to 0.5 or 1.0 results in higher MAE values and slightly reduced PCC, indicating that the rank-based loss term is most effective when used with a relatively small weight.

Table 6 presents the performance of different structural input combinations. When only T1-weighted MRI is used as input, the model shows the highest MAE across all datasets. Incorporating either GM or WM information improves prediction accuracy compared with the T1-only configuration. Among the single-tissue inputs, GM-only consistently yields

lower MAE than WM-only. When both GM and WM are jointly used, the model achieves further improvements in MAE and PCC across datasets, reaching performance comparable to the full configuration that additionally includes demographic information.

The ablation results indicate that GM and WM maps carry a substantial proportion of the predictive signal for brain-age estimation. In particular, the strong performance of GM + WM-based configurations suggests that tissue-derived structural representations contain highly informative age-related features. However, this should not be interpreted as implying that T1-weighted MRI is redundant. Rather, T1-weighted MRI may provide complementary global anatomical context that is less directly encoded in tissue-specific maps. Accordingly, within the proposed framework, the role of T1-weighted MRI is interpreted as complementary rather than dominant.

4.5. Qualitative Analysis and Visualization

Qualitative analyses are conducted to examine the spatial characteristics of the learned representations and to support the quantitative results. Visualization results are obtained using Gradient-weighted Class Activation Mapping (Grad-CAM) applied to the proposed model. Representative examples are shown in Figures 2–4.

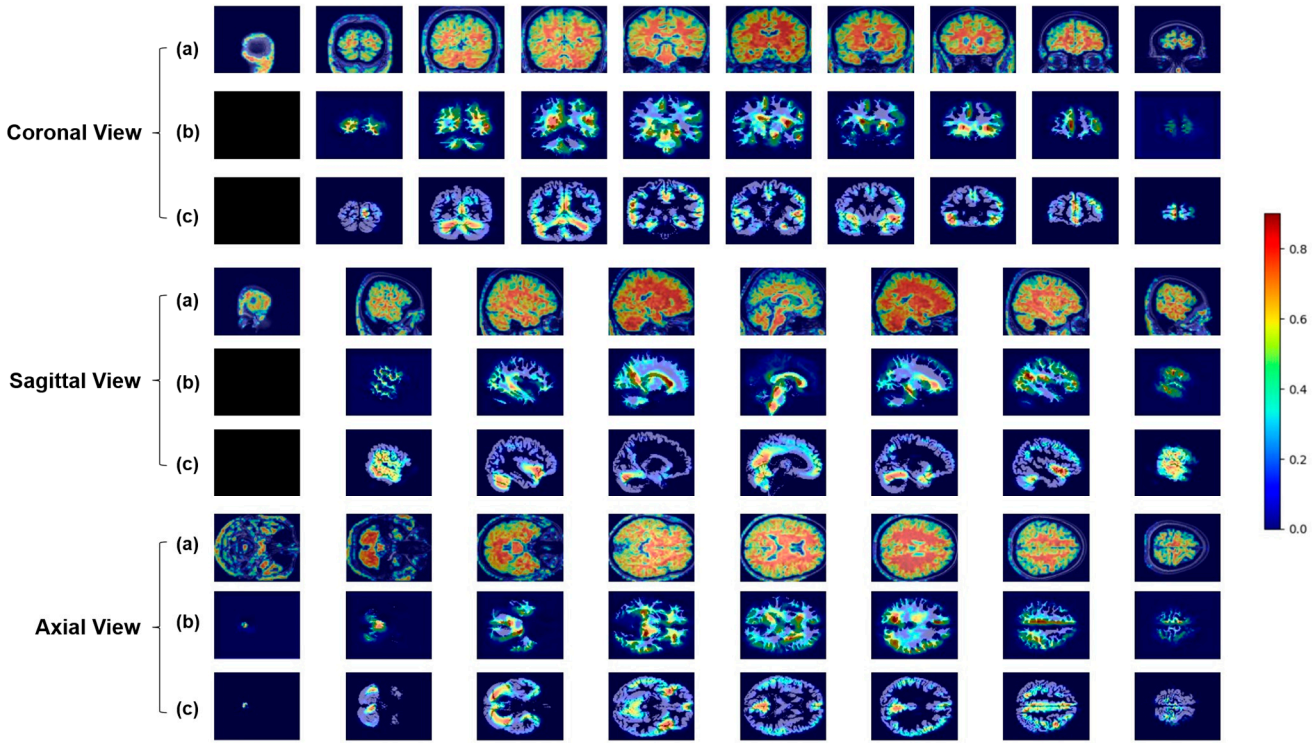


Figure 2. MRI Visualization with Grad-CAM: (**a**) T1-weighted MRI; (**b**) WM; (**c**) GM.

Figure 2 illustrates Grad-CAM activation maps overlaid on T1-weighted MRI, gray matter (GM), and white matter (WM) inputs for a representative subject. The activation patterns highlight central brain regions and areas associated with GM and WM structures, indicating that the model attends to spatial regions relevant to brain age estimation. Activation maps are observed consistently across different anatomical planes, including axial, sagittal, and coronal views.

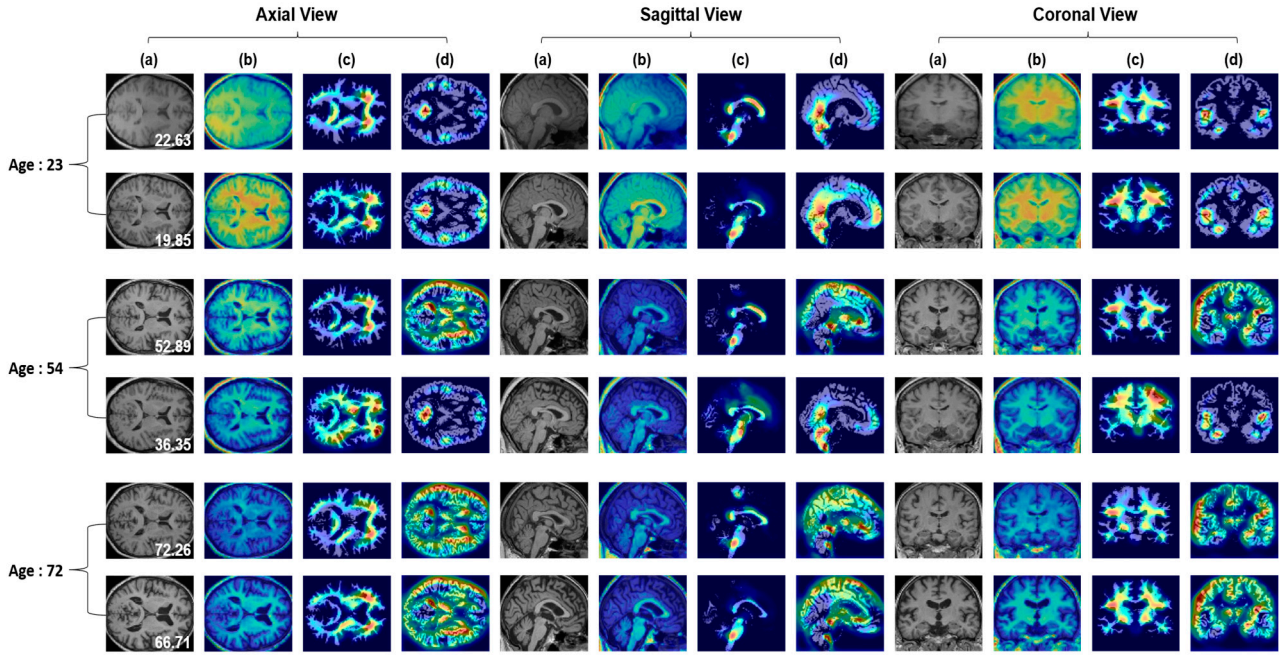


Figure 3. Grad-CAM visualizations of brain age prediction for representative individuals aged 23, 54, and 72 years. (**a**) Original T1-weighted MRI; (**b**) preprocessed T1-weighted MRI; (**c**) white matter (WM); and (**d**) gray matter (GM). Colors denote Grad-CAM activation intensity, with warmer colors (e.g., red) indicating greater contribution to the model's prediction and cooler colors (e.g., blue) indicating lower contribution.

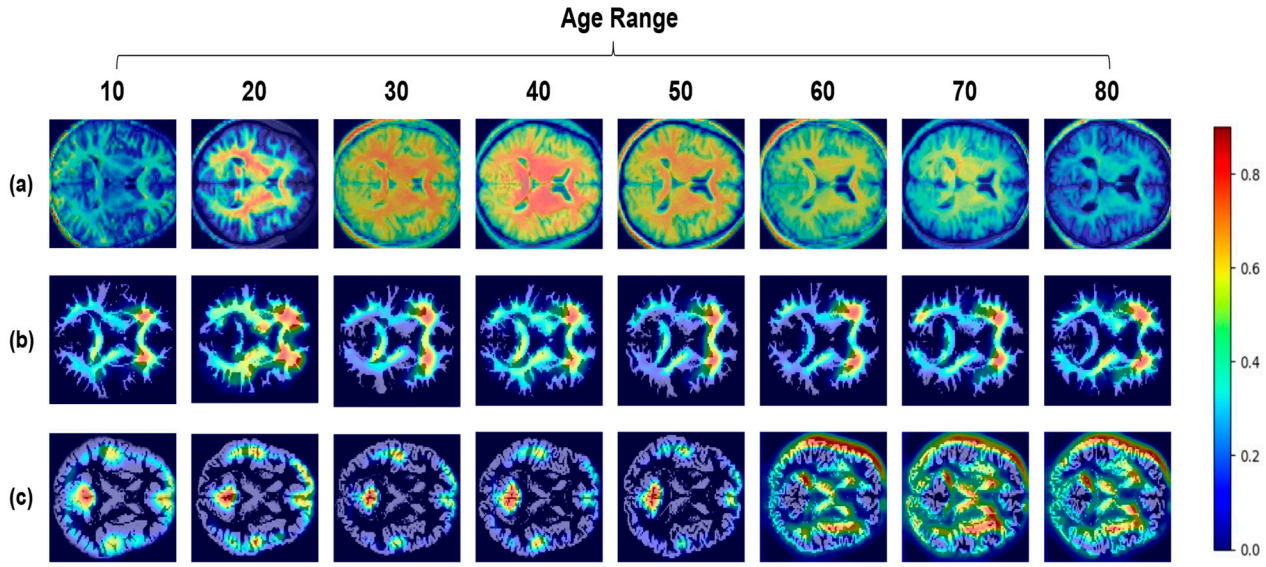


Figure 4. MRI Visualization by Age Range with Grad-CAM: (**a**) T1-weighted MRI; (**b**) WM; (**c**) GM.

Figure 3 presents examples of brain age prediction results for subjects from different age groups (23, 54, and 72 years). For each age group, cases with accurate predictions and cases with larger prediction errors are visualized. In cases with lower prediction error, Grad-CAM activations tend to be more spatially focused, whereas in cases with higher prediction error, broader and less localized activation patterns are observed.

Figure 4 shows Grad-CAM visualizations across age ranges at 10-year intervals. As age increases, changes in activation patterns are observed across T1-weighted MRI, GM, and WM inputs. In particular, GM-related activations become more pronounced in older age groups, while WM-related activations show relatively less variation in the presented

examples. These observations are consistent with the quantitative results and reflect age-dependent structural differences captured by the model.

5. Discussion

The present study showed that the proposed global–local transformer framework provided consistent performance improvements for multimodal brain-age estimation across multiple datasets. By combining whole-brain contextual information with local patch refinement, the model was designed to capture both global anatomical structure and tissue-specific variation from T1-weighted MRI, GM, and WM inputs. The ablation results suggested that GM and WM contributed substantially to prediction performance, whereas T1-weighted MRI played a complementary role by providing broader anatomical context. These results support the usefulness of jointly modeling global and local information in multimodal 3D neuroimaging.

The Grad-CAM results provided qualitative evidence that the model captured age-related structural patterns across tissue types and age groups. In particular, the activation maps suggested tissue-specific differences in salient regions, with stronger GM-related responses in older groups. However, these findings should be interpreted cautiously because the present study did not include atlas-level quantitative validation, cortical thinning correspondence analysis, or region-wise comparison with established neuroanatomical biomarkers. Therefore, the biological interpretation of the activation maps remains preliminary.

Several limitations should be noted. First, because the GM and WM maps were generated by automated segmentation, segmentation errors may have affected downstream prediction performance. Second, although the same SPM-based preprocessing pipeline was applied across datasets, residual inter-site and inter-scanner heterogeneity may still have influenced the learned representations, and harmonization methods were not examined. Third, although the proposed model showed consistent improvements under the same evaluation setting, formal post hoc statistical analyses were not performed. Fourth, the computational evaluation was incomplete because detailed training time and full runtime benchmarking against all baseline models were not systematically recorded. Finally, the sensitivity of the proxy-based ranking objective to the pre-training condition of the proxy network was not evaluated.

For future clinical application, improving parametric reliability and operational robustness will be important. In this regard, hybrid AI optimization strategies, such as the AI-Taguchi-ANOVA approach [30,31], may help improve hyperparameter tuning and system stability. Future work should address statistical uncertainty, interpretability validation, scanner harmonization, broader benchmarking, and optimization for more reliable clinical translation.

6. Conclusions

This study presented a 3D Global–Patch Transformer framework for brain-age prediction from structural MRI. The proposed model directly processes volumetric data and incorporates GM and WM structural maps within a unified architecture, enabling joint learning of whole-brain context and localized anatomical patterns.

Under a consistent preprocessing and training protocol, the proposed method achieved lower mean absolute error and comparable or higher correlation than representative baseline models. The ablation results further supported the contribution of multimodal tissue inputs and the global–patch design to prediction performance.

Overall, these findings suggest that combining global and local structural information in multimodal 3D MRI is an effective strategy for brain-age estimation and provides a useful foundation for future neuroimaging and clinical applications.

Author Contributions: Conceptualization, S.-J.L., M.L. and H.-J.Y.; Methodology, S.-J.L. and M.L.; Software, S.-J.L.; Validation, S.-J.L. and M.L.; Data curation, S.-J.L.; Writing—original draft preparation, S.-J.L. and M.L.; Writing—review and editing, M.L., Y.R.K. and H.-J.Y.; Project administration, H.-J.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (RS-2025-25425780). Additionally, it was supported by the Institute of Information Communications Technology Planning Evaluation (IITP) under the Artificial Intelligence Convergence Innovation Human Resources Development (IITP2023-RS-2023-00256629) and the Information Technology Research Center (ITRC) support program (IITP-2026-RS-2024-00437718) supervised by the IITP, both funded by the Ministry of Science and ICT (MSIT), Korea.

Institutional Review Board Statement: Ethics review and informed consent were not required for this study because only anonymized/de-identified data released for research use were analyzed (no participant contact; no direct identifiers).

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are openly available in the Autism Brain Imaging Data Exchange (ABIDE) repository, reference number [25]; the Consortium for Reliability and Reproducibility (CoRR) repository, reference number [26]; the IXI dataset at https://brain-development.org/ixi-dataset/ (10 January 2026), reference number [27]; and the Open Access Series of Imaging Studies (OASIS-3) database, reference number [28].

Acknowledgments: During the preparation of this manuscript, the authors used ChatGPT (OpenAI, 2025 version, GPT-5.2 model) to review the text for potential improvements in clarity and grammar. The authors have carefully reviewed and edited all output and take full responsibility for the final content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Patterson, C. *World Alzheimer Report 2018: The State of the Art of Dementia Research: New Frontiers*; Alzheimer's Disease International: London, UK, 2018.
2. Cole, J.H.; Franke, K. Predicting age using neuroimaging: Innovative brain ageing biomarkers. *Trends Neurosci.* **2017**, *40*, 681–690. [CrossRef]
3. Cole, J.H.; Poudel, R.P.; Tsagkrasoulis, D.; Caan, M.W.; Steves, C.; Spector, T.D.; Montana, G. Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. *NeuroImage* **2017**, *163*, 115–124. [CrossRef]
4. Fjell, A.M.; Walhovd, K.B. Structural brain changes in aging: Courses, causes and cognitive consequences. *Nat. Rev. Neurosci.* **2010**, *11*, 673–685. [CrossRef]
5. Sun, J.; Tu, Z.; Meng, D.; Gong, Y.; Zhang, M.; Xu, J. Interpretation for individual brain age prediction based on gray matter volume. *Brain Sci.* **2022**, *12*, 1517. [CrossRef] [PubMed]
6. Huang, T.W.; Chen, H.-T.; Fujimoto, R.; Ito, K.; Wu, K.; Sato, K.; Taki, Y.; Fukuda, H.; Aoki, T. Age estimation from brain MRI images using deep learning. In *Proceedings of the IEEE 14th International Symposium on Biomedical Imaging (ISBI), Melbourne, VIC, Australia, 18–21 April 2017*; IEEE: New York, NY, USA, 2017; pp. 849–852.
7. Ueda, M.; Ito, K.; Sato, K.; Taki, Y.; Fukuda, H.; Aoki, T. An age estimation method using 3D-CNN from brain MRI images. In *Proceedings of the IEEE 16th International Symposium on Biomedical Imaging (ISBI), Venice, Italy, 8–11 April 2019*; IEEE: New York, NY, USA, 2019; pp. 1017–1020.
8. Jonsson, B.A.; Bjornsdottir, G.; Thorgeirsson, T.E.; Ellingsen, L.M.; Helgason, A.; Gudbjartsson, D.F.; Stefansson, K.; Magnusson, M.O. Brain age prediction using deep learning uncovers associated sequence variants. *Sci. Rep.* **2019**, *10*, 5409. [CrossRef]
9. Feng, X.; Lipton, Z.C.; Yang, J.; Small, S.A.; Provenzano, F.A.; Alzheimer's Disease Neuroimaging Initiative; Frontotemporal Lobar Degeneration Neuroimaging Initiative. Estimating brain age based on a uniform healthy population with deep learning and structural magnetic resonance imaging. *Neurobiol. Aging* **2020**, *91*, 15–25. [CrossRef] [PubMed]
10. Bashyam, V.M.; Erus, G.; Doshi, J.; Habes, M.; Nasrallah, I.M.; Truelove-Hill, M.; Srinivasan, D.; Mamourian, L.; Pomponio, R.; Fan, Y.; et al. MRI signatures of brain age and disease over the lifespan based on a deep brain network and 14,468 individuals worldwide. *Brain* **2020**, *143*, 2312–2324. [CrossRef] [PubMed]

11. Peng, H.; Gong, W.; Beckmann, C.F.; Vedaldi, A.; Smith, S.M. Accurate brain age prediction with lightweight deep neural networks. *Med. Image Anal.* **2021**, *68*, 101871. [CrossRef]
12. Cheng, J.; Liu, Z.; Guan, H.; Wu, Z.; Zhu, H.; Jiang, J.; Wen, W.; Tao, D.; Liu, T. Brain age estimation from MRI using cascade networks with ranking loss. *IEEE Trans. Med. Imaging* **2021**, *40*, 3400–3412. [CrossRef]
13. Dinsdale, N.K.; Bluemke, E.; Smith, S.M.; Arya, Z.; Vidaurre, D.; Jenkinson, M.; Namburete, A.I. Learning patterns of the ageing brain in MRI using deep convolutional networks. *NeuroImage* **2021**, *224*, 117401. [CrossRef]
14. Bellantuono, L.; Marzampo, L.; La Rocca, M.; Duca, D.; Lombardi, A.; Maiquiddi, G.; Monaco, A.; Tangaro, S.; Amoroso, N.; Bellotti, R.; et al. Predicting brain age with complex networks: From adolescence to adulthood. *NeuroImage* **2021**, *225*, 117458. [CrossRef]
15. Gupta, U.; Lam, P.K.; Ver Steeg, G.; Thompson, P.M. Improved brain age estimation with slice-based set networks. In *Proceedings of the IEEE 18th International Symposium on Biomedical Imaging (ISBI), Nice, France, 13–16 April 2021*; IEEE: New York, NY, USA, 2021; pp. 1237–1241.
16. He, S.; Grant, P.E.; Ou, Y. Global-local transformer for brain age estimation. *IEEE Trans. Med. Imaging* **2022**, *41*, 213–224. [CrossRef]
17. He, S.; Feng, Y.; Grant, P.E.; Ou, Y. Deep relation learning for regression and its application to brain age estimation. *IEEE Trans. Med. Imaging* **2022**, *41*, 2304–2317. [CrossRef]
18. Siegel, N.T.; Kainmueller, D.; Deniz, F.; Ritter, K.; Schulz, M.A. Do transformers and CNNs learn different concepts of brain age? *Hum. Brain Mapp.* **2025**, *46*, e70243. [CrossRef] [PubMed]
19. Bintsi, K.M.; Rahman, M.N.; Rahman, A.H.; Rahman, M.S. BrainRotViT: Transformer-ResNet Hybrid for Explainable Modeling of Brain Aging from 3D sMRI. *arXiv* **2025**, arXiv:2511.15188.
20. Zhang, Y.; Zhang, C.; Ma, X.; Zhu, X.; Lin, L.; Tian, M. ds-FCRN: Three-dimensional dual-stream fully convolutional residual networks and transformer-based global-local feature learning for brain age prediction. *Comput. Biol. Med.* **2026**, *169*, 107921.
21. Joo, Y.; Namgung, E.; Jeong, H.; Kang, I.; Kim, J.; Oh, S.; Lyoo, I.; Jeong, S.; Yoon, S.; Hwang, J. Brain age prediction using combined deep convolutional neural network and multi-layer perceptron algorithms. *Sci. Rep.* **2023**, *13*, 16825. [CrossRef]
22. He, L.; Wang, S.; Chen, C.; Wang, Y.; Fan, Q.; Chu, C.; Fang, L.; Xu, J. Network occlusion sensitivity analysis identifies regional contributions to brain age prediction. *Hum. Brain Mapp.* **2025**, *46*, e70239. [CrossRef]
23. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 5998–6008.
24. Liu, Z.; Mao, H.; Wu, C.Y.; Feichtenhofer, C.; Darrell, T.; Xie, S. A convnet for the 2020s. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 18–24 June 2022*; IEEE: New York, NY, USA, 2022; pp. 11976–11986.
25. Di Martino, A.; Yan, C.-G.; Li, Q.; Denio, E.; Castellanos, F.X.; Alaerts, K.; Anderson, J.S.; Assaf, M.; Bookheimer, S.Y.; Dapretto, M.; et al. The autism brain imaging data exchange: Towards a large-scale evaluation of the intrinsic brain architecture in autism. *Mol. Psychiatry* **2014**, *19*, 659–667. [CrossRef]
26. Zuo, X.N.; Anderson, J.S.; Bellec, P.; Birn, R.M.; Biswal, B.B.; Blautzik, J.; Breitner, J.; Buckner, R.L.; Calhoun, V.D.; Castellanos, F.X. An open science resource for establishing reliability and reproducibility in functional connectomics. *Sci. Data* **2014**, *1*, 140049. [CrossRef]
27. IXI Dataset. Available online: https://brain-development.org/ixi-dataset/ (accessed on 10 January 2026).
28. LaMontagne, P.J.; Benzinger, T.L.S.; Morris, J.C.; Keefe, S.; Hornbeck, R.; Xiong, C.; Grant, E.; Hassenstab, J.; Moulder, K.; Vlassenko, A.G.; et al. OASIS-3: Longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and Alzheimer's disease. *Alzheimers Dement.* **2018**, *14*, P1097.
29. Penny, W.D.; Friston, K.J.; Ashburner, J.T.; Kiebel, S.J.; Nichols, T.E. *Statistical Parametric Mapping: The Analysis of Functional Brain Images*; Elsevier: Amsterdam, The Netherlands, 2011.
30. Laganà, F.; Faccì, A.R. Parametric optimisation of a pulmonary ventilator using the Taguchi method. *J. Electr. Eng.* **2025**, *76*, 265–274. [CrossRef]
31. Laganà, F.; Pratticò, D.; Quattrone, M.F.; Pullano, S.A.; Calcagno, S. Hybrid AI–Taguchi–ANOVA Approach for Thermographic Monitoring of Electronic Devices. *Eng* **2026**, *7*, 28. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.