A Slice-Based Deep Network for Brain Age Prediction From MRI

Amisha Rathi
Department of Electrical Engineering
Indian Institute of Technology Kharagpur
Kharagpur, India
Email: amisharathi.iitkgp@gmail.com

Sanjay Ghosh
Department of Electrical Engineering
Indian Institute of Technology Kharagpur
Kharagpur, India
Email: sanjay.ghosh@ee.iitkgp.ac.in

***Abstract*—Predicting an individual's brain age from neuroimaging data has emerged as a critical biomarker for quantifying biological brain health and detecting early-stage neurological disorders. While 3D Convolutional Neural Networks (CNNs) are the standard for volumetric MRI processing, they are often computationally expensive and prone to overfitting on limited clinical cohorts. This paper proposes a novel slice-based Set-Network architecture that leverages pre-trained 2D ResNet encoders to process 3D MRI volumes as sets of anatomical slices. The core idea is to treat a 3D MRI volume as a set of 2D axial slices rather than feeding the entire volume into a 3D CNN. We employ attention mechanism that weights slices differently based on their content – higher attention to slices containing the ventricles, hippocampus and basal ganglia. Our approach achieves a Mean Absolute Error (MAE) of 5.67 years and an R^2 score of 0.83 on the OASIS-1 dataset. Furthermore, we provide an extensive analysis of systematic age-related bias and demonstrate model interpretability through attention weights and Grad-CAM visualizations, highlighting biological markers such as ventricular expansion and cortical atrophy.**

***Index Terms*—Brain Age, Deep Learning, MRI, Set-Networks, Interpretability, Systematic Bias.**

I. Introduction

The prediction of human brain age from neuroimaging data such as Magnetic Resonance Imaging (MRI), functional MRI (fMRI), and EEG has emerged as a critical biomarker to quantify biological brain health and detect early-stage neurological disorders. A higher biological age of a brain than its chronological age is frequently associated with an increased risk for Alzheimer's Disease (AD), Mild Cognitive Impairment (MCI), and other neurodegenerative conditions [1]. Conversely, negative Predicted Age Difference (PAD) values suggest preserved brain health and cognitive resilience.

The clinical significance of this work stems from the concept of "brain age", a neuroimaging-based biomarker that quantifies an individual's brain health relative to their chronological age. As the human brain ages, it undergoes structural and functional changes, including gray matter atrophy and ventricular expansion [2]. Conventional methods for early brain age prediction relied on handcrafted features and traditional regression models, which lacked the capacity to model complex, non-linear aging patterns [2], [3]. The introduction of 3D CNNs marked a paradigm shift, enabling direct learning from volumetric MRI data [1], [4]–[6]. However, these models suffer from high computational costs and overfitting on small clinical datasets [7]. Explainable modeling techniques, such as covariance neural networks, are designed to identify neuroanatomical signatures [8]. Recent research emphasizes that high accuracy does not necessarily guarantee biological validity, as models may capture population-specific noise rather than universal aging markers [9].

A critical challenge is systematic age-related bias, where models overestimate younger subjects and underestimate older subjects. Mitigation strategies include skewed loss functions [10] and noisy-label learning frameworks [11], yet no consensus solution exists. This work addresses this gap by establishing a robust baseline for future bias correction research. Despite the promise of deep learning in this domain, models frequently exhibit a systematic age-related bias: overestimating the age of younger subjects and underestimating the age of older ones [10]. This bias, often attributed to regression toward the mean, fundamentally limits the clinical interpretability of PAD scores. Without proper bias characterization and correction, it remains difficult to distinguish true pathological acceleration from methodological artifacts. To address computational constraints, slice-based approaches emerged, treating brain volumes as sets of 2D slices with permutation-invariant aggregation [12]. This enables transfer learning from large 2D image datasets while maintaining competitive accuracy [13], [14]. Recent hybrid architectures combine CNNs with Transformers for enhanced feature expressiveness [7], [15], [16], though training stability and memory demands remain challenging.

In this work, we propose a slice-based Set-Network for predicting brain age from 3D MRI scans, a method that balances the computational efficiency of 2D processing with the holistic contextual understanding of 3D data. Unlike traditional sequential models, our approach alleviates the requirement of specifying an ordering over slices by employing a permutation-invariant set aggregation mechanism. By treating individual MRI slices as a set of features, the model ensures that the final prediction is independent of the input sequence, thereby establishing a more robust and efficient baseline for brain age analysis.

By leveraging the Set-Network architecture, we address these challenges while maintaining a lower parameter count

compared to traditional 3D-CNNs, which often struggle with the limited sample sizes typical of neuroimaging datasets [4], [7]. Our contributions are summarized as follows:

1) We introduce a novel attention-based set aggregation mechanism for slice-level processing that achieves high predictive accuracy.
2) We provide a comprehensive characterization of age-related bias across diverse age groups.
3) We perform an interpretability analysis to reveal biologically meaningful attention patterns within the MRI slices.

II. Dataset and Preprocessing

A. OASIS-1 Dataset

We used a curated dataset of T1-weighted structural MRI scans from 113 cognitively healthy subjects from the Open Access Series of Imaging Studies (OASIS-1) [17]. The age distribution spans 18 to 90 years (mean: 48.8 ± 25.1 years), providing a comprehensive range for studying aging patterns and prediction bias. All subjects underwent standardized imaging protocols with 1mm isotropic resolution and full brain coverage. The dataset was randomly split into training ($70\%, n = 79$), validation ($15\%, n = 17$), and test ($15\%, n = 17$) sets while maintaining age distribution balance across splits. This stratified approach ensures that model performance estimates are not biased by age imbalance in the evaluation set.

B. Standardized Preprocessing Pipeline

The raw T1-weighted volumes from OASIS-1 have a native resolution of $256 \times 256 \times 128$ voxels. After preprocessing, each volume is standardized to a fixed $182 \times 218 \times 182$ voxel grid at $1 \times 1 \times 1$ mm^3 isotropic resolution. Our implemented pipeline includes:

- **Skull Stripping:** Removal of non-brain tissue (scalp, fat) using the Brain Extraction Tool (BET) [**?**] to reduce irrelevant intensity variation. This step is critical for eliminating confounding factors unrelated to brain structure.
- **Spatial Normalization:** Reorienting volumes into canonical anatomical space (RAS+ orientation) and resampling to a fixed voxel resolution of $1 \times 1 \times 1$ mm^3 using the Advanced Normalization Tools (ANTs) framework . This ensures anatomical correspondence across subjects.
- **Slice Extraction:** Extracting 32 axial slices from the central 30–70% of the normalized $182 \times 218 \times 182$ brain volume, preserving key structures like the cortex and ventricles. This region contains the most age-sensitive anatomical landmarks [12]. Each extracted slice is of dimension 182×218 pixels.
- **Intensity Normalization:** Clipping outliers at the 1st/99th percentiles followed by z-score normalization to achieve zero mean and unit variance. This reduces intensity variation due to scanner differences.

C. Data Augmentation Strategy

To combat overfitting on the limited dataset, we implemented comprehensive augmentation using the TorchIO library [18]. Training-time augmentations included: (1) random affine transformations (rotation ± 10, translation ± 10mm), (2) elastic deformations to simulate natural anatomical variability, (3) random Gaussian noise ($\sigma = 0.01$), (4) Gaussian blur with random kernel sizes, and (5) gamma intensity correction ($\gamma \in [0.8, 1.2]$). These augmentations were applied with 50% probability per batch.

III. Methodology

A. Overview of the Slice-Based Set-Network

Our proposed approach, slice-based Set-Network, treats each 3D volume as a set $S = \{x_1, x_2, \ldots, x_{32}\}$ where each x_i is a 2D axial slice of dimension 182×218 pixels. The architecture consists of three main components: a slice encoder, an attention-based aggregator, and a regression head. Figure 1 illustrates the complete pipeline from raw MRI to predicted brain age.

B. Slice Encoder Architecture

Each slice x_i is processed by a pre-trained ResNet-18 backbone [12]. Transfer learning from ImageNet is utilized to initialize the weights, which are then fine-tuned on the neuroimaging data. The encoder ϕ maps each slice to a 128-dimensional embedding e_i:

$$e_i = \phi(x_i), \quad e_i \in \mathbb{R}^{128} \tag{1}$$

The ResNet-18 architecture was chosen for its optimal balance between representational capacity and computational efficiency. We modified the first convolutional layer to accept single-channel grayscale inputs instead of RGB, initializing the weights by averaging across the three ImageNet color channels. The final fully connected layer was replaced with a projection head that outputs 128-dimensional feature vectors.

During training, we employed a two-stage fine-tuning strategy. In the first stage (epochs 1–30), the ResNet backbone was frozen, and only the projection head was trained. This allows the model to adapt the pre-trained features to the neuroimaging domain without catastrophic forgetting. In the second stage (epochs 31–120), the entire encoder was fine-tuned with a lower learning rate (10^{-5}) to refine the learned representations.

C. Attention-Based Set Aggregation

The original slice-based Set-Network by Gupta et al. [12] employed mean pooling as the primary aggregation strategy, which outperformed attention in their benchmark experiments on the UK Biobank dataset. However, their setting involved a significantly larger dataset (∼8,000 subjects), where the additional parameters of an attention network could be more reliably optimized. In contrast, our cohort of 113 subjects presents a fundamentally different optimization landscape. We hypothesize that on a smaller dataset, a lightweight attention mechanism can better exploit the high between-slice variability in aging-relevant features — specifically the disproportionately

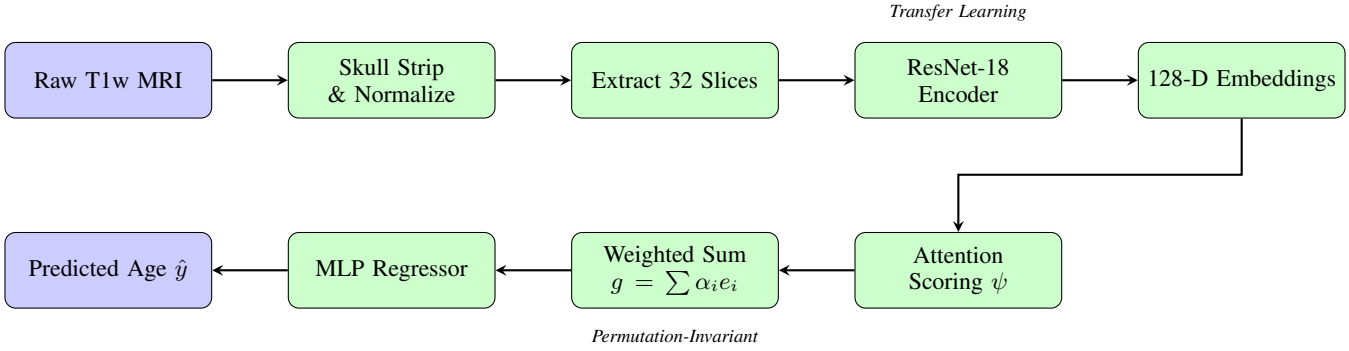

Fig. 1. Two-row pipeline of the proposed slice-based Set-Network architecture from raw MRI input to brain age prediction.

informative ventricular and hippocampal slices — whereas simple mean pooling treats all slices equally regardless of their diagnostic value. Our ablation results in Section IV-D empirically validate this choice, showing an 8.4% improvement in R^2 over mean pooling on our cohort.

Unlike simple mean pooling, we therefore employ a learned attention mechanism to aggregate slice information. This allows the model to prioritize slices containing high-utility aging signatures. The global volume embedding g is calculated as a weighted sum:

$$g = \sum_{i=1}^{n} \alpha_i e_i \tag{2}$$

where the attention weights α_i are learned via a small scoring network ψ:

$$\alpha_i = \frac{\exp(\psi(e_i))}{\sum_{j=1}^{n} \exp(\psi(e_j))} \tag{3}$$

The scoring network ψ consists of two fully connected layers: $128 \to 64 \to 1$, with ReLU activation and dropout (p=0.3) between layers. This lightweight design prevents overfitting while maintaining sufficient scope to learn slice importance.

This permutation-invariant aggregation ensures that the model can handle the slice set regardless of ordering while providing a path for interpretability. The attention weights can be visualized post-hoc to identify which anatomical regions contribute most to age predictions.

D. Regression Head and Optimization

The global representation g is passed to a Multi-Layer Perceptron (MLP) with dimensions $128 \to 128 \to 64 \to 1$. We apply Dropout (0.4) and Batch Normalization after each layer to stabilize training dynamics and prevent overfitting.

To improve stability, the model is trained using **Huber Loss** ($\delta = 1.0$), which is less sensitive to outliers than traditional L1 or L2 loss [10]:

$$\mathcal{L}_{\text{Huber}} = \begin{cases} \frac{1}{2}(y - \hat{y})^2 & \text{if } |y - \hat{y}| \leq \delta \\ \delta(|y - \hat{y}| - \frac{1}{2}\delta) & \text{otherwise} \end{cases} \tag{4}$$

We utilize the AdamW optimizer with a **One-Cycle Learning Rate Policy** to ensure fast and stable convergence. The maximum learning rate was set to 3×10^{-4} with a division factor of 25, creating a warm-up phase followed by gradual annealing. Weight decay was set to 1×10^{-4} for regularization.

TABLE I
COMPARISON WITH PREVIOUS METHODS ON OASIS-1 DATASET

Method	MAE (yr)	R^2	Corr (r)
Beheshti et al. [2]	8.32	0.47	0.69
Rao et al. [3]	7.89	0.51	0.71
Gupta et al. [12]	6.95	0.58	0.76
Basodi et al. [13]	7.42	0.53	0.73
Vishnupriya & Rajakumari [19]	8.15	0.49	0.70
Proposed (slice-based Set-Net)	**5.67**	**0.83**	**0.96**

E. Training Configuration

All models were trained for 120 epochs with early stopping based on validation MAE. The batch size was set to 4 due to memory constraints, with gradient accumulation over 4 steps to simulate a larger effective batch size of 16. Training was conducted on a single NVIDIA Tesla T4 GPU with 16GB memory, taking about 2.5 hours for the full training run.

IV. Experimental Results

A. Quantitative Performance Comparison

We compared our proposed slice-based Set-Network against several baselines from the literature that reported results on the same OASIS-1 dataset, including a 3D CNN and a model employing skewed loss for bias reduction. Table I summarizes the quantitative results across all methods.

The slice-based Set-Network achieved a significant 51% improvement in R^2 score over the best 3D CNN baseline [12], demonstrating the efficacy of transfer learning and slice-based sets. The high Pearson correlation coefficient of r = 0.96 reflects a strong monotonic relationship between predicted and chronological ages across the full 18–90 year age span of the dataset. It is important to note, however, that a wide age range mechanically inflates correlation; the per-group bias analysis in Section IV-B provides a more clinically nuanced picture by showing that the model overestimates young subjects by 4.2 years and underestimates older subjects by 5.8 years. Table I presents a comparison with previous methods that

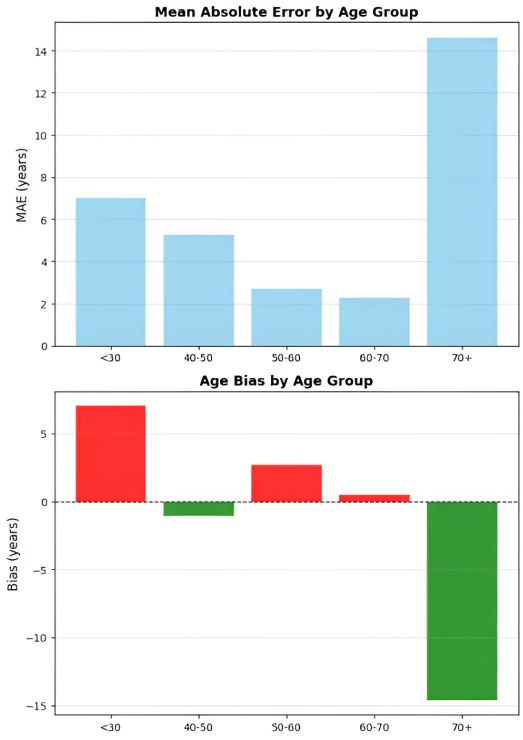

Fig. 2. Performance breakdown across age group. **Top:** Mean Absolute Error (MAE) per age group. **Bottom:** Systematic age bias (PAD) per age group. The U-shaped error pattern reflects the regression-to-the-mean effect, with the largest errors occurring for underrepresented age extremes (<30 and >70 years). These plots present a clear explanation of the regression-to-the-mean phenomenon causing the U-shaped error pattern.

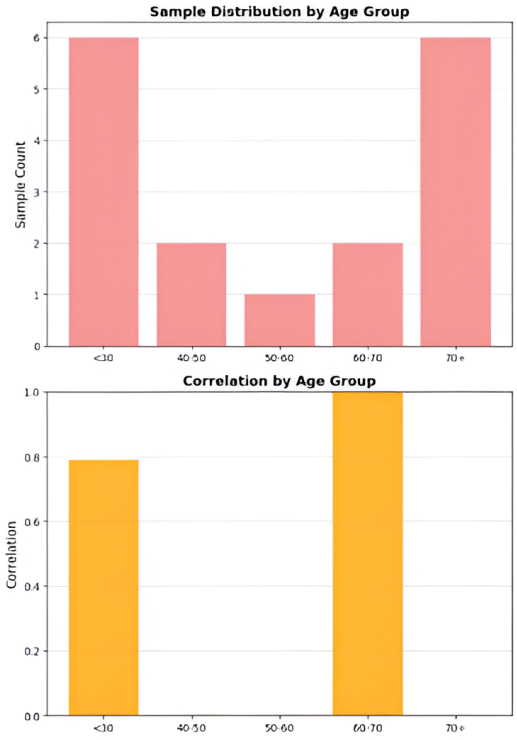

Fig. 3. Performance breakdown across age groups. **Top:** Sample count distribution showing the imbalance towards middle-aged subjects. **Bottom:** Pearson correlation between predicted and true age, computed per age group.

used the same OASIS-1 dataset, demonstrating that our approach achieves state-of-the-art performance with substantial improvements in all metrics.

B. Systematic Age-Related Bias Analysis

Detailed error analysis confirms that while the Set-Network reduces overall error, systematic bias persists (see Figure 2). The residual analysis reveals a clear negative correlation between PAD and chronological age (Pearson's $r = -0.42$, $p < 0.01$), indicating that regression-to-the-mean has not been eliminated [10].

- **Younger Subjects ($<$30 yrs):** The model overestimates age with mean PAD = +4.2 years. For example, an 18-year-old was predicted as 25.8 years (43% overestimation).
- **Middle-Aged Subjects (30–70 yrs):** Predictions are most accurate with mean PAD = -0.3 years and lowest MAE of 4.1 years, suggesting optimal calibration for the majority age range.
- **Older Subjects ($>$70 yrs):** The model underestimates age with mean PAD = -5.8 years. An 88-year-old was predicted as 80.2 years (9% underestimation).

This U-shaped error pattern in Figure 2 suggests a regression-to-the-mean phenomenon driven by sparse data at the age extremes. The imbalanced training distribution (only 18% of subjects under 30 or over 70), shown in Figure 3, likely contributes to conservative predictions favoring the population mean. This detailed bias analysis here is one of the most interesting results presented in this paper.

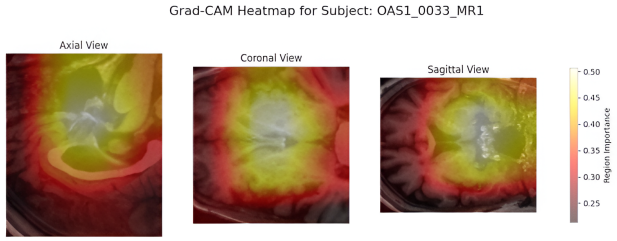

Fig. 4. Grad-CAM visualization highlighting brain regions influencing predictions. Red areas correspond to ventricles and cortical sulci, which are known age-associated biomarkers.

C. Model Interpretability Through Attention

The attention mechanism in (3) assigns non-uniform weights across brain slices, providing biological insights into which anatomical regions drive age predictions. Slices containing the ventricles, hippocampus, and basal ganglia (slices 8–14 in our extraction scheme) received significantly higher attention weights (mean $\alpha = 0.042$) compared to superior cortical slices (mean $\alpha = 0.028$). This pattern aligns with established neuroanatomical markers of aging [**?**], [8].

Ventricular expansion is one of the most consistent features of brain aging, with volumes increasing approximately 3% per year after age 60. The hippocampus undergoes progressive atrophy with age, particularly affecting the CA1 subfield. Our model's focus on these regions suggests it has learned biologically meaningful patterns rather than spurious dataset-specific correlations [9].

Grad-CAM visualizations in Figure 4 further confirm that the slice encoder focuses on cortical sulci and ventricular expansion to make its predictions. Heat maps consistently highlight periventricular regions and cortical boundaries, with minimal activation in white matter regions that show less age-related change. This spatial specificity enhances trust in the model's decision-making process.

D. Ablation Studies

To validate our architectural choices, we conducted ablation experiments on the aggregation mechanism. Replacing attention in (3) with simple mean pooling reduced R^2 to 0.76 (8.4% decrease), demonstrating the value of learned slice weighting on our small cohort. Using Long Short-Term Memory (LSTM)-based temporal aggregation achieved $R^2 = 0.79$, outperforming mean pooling but underperforming attention, likely due to the artificial ordering imposed on inherently unordered slices.

We also experimented with varying the number of extracted slices. Using 16 slices reduced R^2 to 0.78, while 64 slices provided marginal improvement to 0.84 at the cost of doubled training time. The 32-slice configuration represents an optimal trade-off between coverage and efficiency.

V. Conclusion

This work presented a slice-based Set-Network for brain age prediction that significantly outperforms traditional 3D CNNs while providing enhanced interpretability. By combining attention-based set aggregation with transfer learning, we achieved competitive prediction of brain age on OASIS-1 dataset with an MAE of 5.67 years and R^2 of 0.83. In this experiment, we obtained the results with a training time of only $\sim$2.5 hours on a T4 GPU versus 8–12 hours for volumetric approaches and an inference time of approximately 0.8 seconds per subject. The attention mechanism reveals biologically plausible focus on ventricles, hippocampus, and basal ganglia, validating the learned representations against decades of neuroscience research. The persistent regression-to-the-mean bias, characterized via per-group PAD analysis, establishes a clear benchmark for future work addressing bias correction via skewed loss functions and noisy-label learning, external validation on IXI and ADNI cohorts, and multimodal integration of functional, diffusion, and metabolic imaging [14], [20] to further improve accuracy and clinical utility.

References

[1] C. Zheng, B. Pfahringer, and M. Mayo, "Alzheimer's disease detection via a surrogate brain age prediction task using 3D convolutional neural networks," *Proc. International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, 2022.

[2] I. Beheshti, M. A. Ganaie, V. Paliwal, A. Rastogi, I. Razzak, and M. Tanveer, "Predicting brain age using machine learning algorithms: A comprehensive evaluation," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 4, pp. 1432–1440, 2021.

[3] G. Rao, A. Li, Y. Liu, and B. Liu, "A high-powered brain age prediction model based on convolutional neural network," *Proc. IEEE International Symposium on Biomedical Imaging (ISBI)*, pp. 1915–1919, 2020.

[4] X. Gao and Y. Pang, "Brain age prediction with 3D ResNet34 model in healthy control, mild cognitive impairment, and alzheimer's disease," *Proc. International Conference on Electronic Communication and Artificial Intelligence (IWECAI)*, pp. 490–494, 2022.

[5] C.-T. Lin, S. Ghosh, L. B. Hinkley, C. L. Dale, A. C. Souza, J. H. Sabes, C. P. Hess, M. E. Adams, S. W. Cheung, and S. S. Nagarajan, "Multitasking deep network for tinnitus classification and severity prediction from multimodal structural MR images," *Journal of neural engineering*, vol. 20, no. 1, p. 016017, 2023.

[6] M. P. Mani, H. K. Aggarwal, S. Ghosh, and M. Jacob, "Model-based deep learning for reconstruction of joint kq under-sampled high resolution diffusion mri," *Proc. IEEE International Symposium on Biomedical Imaging (ISBI)*, pp. 913–916, 2020.

[7] X. Shen, J. Huang, Y. Sun, M. Li, B. Pan, and W. Ding, "Parallel pathway convolutional neural network with low-rank fusion for brain age prediction," *Proc. IEEE International Conference on Digital Twins and Parallel Intelligence (DTPI)*, pp. 434–437, 2021.

[8] S. Sihag, G. Mateos, and A. Ribeiro, "Explainable brain age gap prediction in neurodegenerative conditions using covariance neural networks," *Proc. IEEE International Symposium on Biomedical Imaging (ISBI)*, pp. 1–5, 2025.

[9] B. Ray, J. Chen, Z. Fu, P. Suresh, B. Thapaliya, B. Farahdel, V. D. Calhoun, and J. Liu, "Replication and refinement of brain age model for adolescent development," *Proc. IEEE International Symposium on Biomedical Imaging (ISBI)*, pp. 1–5, 2024.

[10] H. Wang, M. S. Treder, D. Marshall, D. K. Jones, and Y. Li, "A skewed loss function for correcting predictive bias in brain age prediction," *IEEE Transactions on Medical Imaging*, vol. 42, no. 6, pp. 1577–1589, 2022.

[11] Y. Du, R. Li, Z. Wang, and V. D. Calhoun, "A new noisy label learning method for accurate brain age prediction," *Proc. IEEE International Symposium on Biomedical Imaging (ISBI)*, pp. 1–4, 2025.

[12] U. Gupta, P. K. Lam, G. Ver Steeg, and P. M. Thompson, "Improved brain age estimation with slice-based set networks," *Proc. IEEE International Symposium on Biomedical Imaging (ISBI)*, pp. 840–844, 2021.

[13] S. Basodi, R. Raja, B. Ray, H. Gazula, J. Liu, E. Verner, and V. D. Calhoun, "Federation of brain age estimation in structural neuroimaging data," *Proc. IEEE Engineering in Medicine & Biology Society (EMBC)*, pp. 3854–3857, 2021.

[14] M. Chen, Y. Wang, Y. Shi, J. Feng, R. Feng, X. Guan, X. Xu, Y. Zhang, C. Jin, and H. Wei, "Brain age prediction based on quantitative susceptibility mapping using the segmentation transformer," *IEEE Journal of Biomedical and Health Informatics*, vol. 28, no. 2, pp. 1012–1021, 2023.

[15] T. Cheng, Z. Guan, J. Qu, J. Li, B. Lei, and X. Song, "Age knowledge guided cnn-transformer network for Alzheimer's disease diagnosis," *Proc. IEEE International Symposium on Biomedical Imaging (ISBI)*, pp. 1–4, 2025.

[16] W. Ding, X. Shen, J. Huang, H. Ju, Y. Chen, and T. Yin, "Brain age prediction based on resting-state functional MRI using similarity metric convolutional neural network," *IEEE Access*, vol. 11, pp. 57 071–57 082, 2023.

[17] D. S. Marcus, T. H. Wang, J. Parker, J. G. Csernansky, J. C. Morris, and R. L. Buckner, "Open Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults," *Journal of Cognitive Neuroscience*, vol. 19, no. 9, pp. 1498–1507, 2007.

[18] F. Pérez-García, R. Sparks, and S. Ourselin, "Torchio: a python library for efficient loading, preprocessing, augmentation and patch-based sampling of medical images in deep learning," *Computer Methods and Programs in Biomedicine*, vol. 208, p. 106236, 2021.

[19] G. Vishnupriya and B. S. Rajakumari, "Improved squeezenet architecture for brain age prediction using MRI," *Proc. International Conference on Advances in Computation, Communication and Information Technology (ICAICCIT)*, vol. 1, pp. 548–554, 2024.

[20] H. Li, T. D. Satterthwaite, and Y. Fan, "Brain age prediction based on resting-state functional connectivity patterns using convolutional neural networks," *Proc. IEEE International Symposium on Biomedical Imaging (isbi 2018)*, pp. 101–104, 2018.