MT-REBA: ANATOMICALLY GROUNDED REGIONAL BRAIN AGE REPRESENTATION LEARNING VIA WEAK SUPERVISION

Anonymous authors
Paper under double-blind review

ABSTRACT

Brain age estimation provides a measure of brain aging, but most existing methods produce only global estimates, obscuring regional aging heterogeneity. Learning regional aging representations remains challenging because chronological age is defined at the subject level, while regional biological ages are neither directly observable nor supported by regional annotations. We introduce **M**ulti-**T**ask **R**egional **B**rain **A**ge Estimation (MT-ReBA), a weakly supervised framework for learning anatomically grounded regional aging representations from T1-weighted MRI without regional labels. The key idea is to replace one atlas-defined parcel with its counterpart from another healthy subject, thereby coupling a localized image modification with a donor-age weak target, converting subject-level age labels into localized weak supervision. Specifically, an atlas-defined donor parcel is integrated into a recipient brain, and the donor's chronological age supervises the corresponding regional head as a weak proxy for regional aging. We evaluate MT-ReBA across six public cohorts. MT-ReBA achieves a global brain MAE of 3.10 years and a regional age-ranking accuracy of $69.63 \pm 0.11\%$, shows agreement with constructed weak targets on mixed images, and exhibits anatomically specific responses and exploratory disease associations on unspliced scans.

1 INTRODUCTION

Brain aging occurs throughout the lifespan and is associated with cognitive decline, structural remodeling, and morphological changes (Azzam et al., 2025). Quantifying this process is important for understanding healthy aging and deviations from typical aging, including accelerated aging that can occur in Alzheimer's disease (AD) and Parkinson's disease (PD) (Baecker et al., 2021b). Brain age estimation (BAE) addresses this need by predicting age from neuroimaging data and is a widely used approach for assessing brain health and investigating disease risk (Baecker et al., 2021b).

However, most BAE methods provide a single global estimate, obscuring regional aging heterogeneity. First, even in healthy individuals, brain regions can follow distinct aging trajectories associated with different genetic, physiological, and functional processes (Smith et al., 2020). Second, brain disorders can preferentially affect specific regions, exhibiting disease-associated patterns of accelerated aging (Kaufmann et al., 2019). Together, these observations motivate regional brain age estimation (ReBA) for characterizing aging at the level of individual anatomical regions. Our goal is to learn anatomically grounded regional aging representations that reveal heterogeneity obscured by global estimates and support fine-grained assessment of brain health.

Prior work has explored complementary approaches to regional brain aging. Local BrainAge (Popescu et al., 2021) uses a patch-based U-Net to predict voxel-wise age maps, evaluated through test–retest reliability and disease-group comparisons. Gianchandani et al. (2024) jointly predict voxel-wise and global age alongside tissue segmentation, examining regional age deviations across diagnostic groups. Kalc et al. (2024) train independent Gaussian-process regressors on lobar gray-matter inputs and assess regional prediction accuracy and disease-associated age gaps. Building on these advances, we focus on two challenges in joint regional modeling. 1. Regional supervision and evaluation. The lack of ground-truth regional brain-age labels limits direct supervision and validation, as subject-level ages do not specify how aging differs across brain regions.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

Moreover, well-established region-specific evaluation metrics remain lacking, making it difficult to assess whether regional predictions capture anatomically specific aging patterns rather than shared global age information. 2. Practical utility of regional representations. The practical value of regional brain-age representations remains insufficiently explored, particularly their contribution to brain-age estimation and downstream disease classification. Beyond revealing disease-associated regional aging patterns, their ability to provide useful predictive features and offer added value over global brain-age measures requires explicit evaluation.

To address these challenges, we propose MT-ReBA, a weakly supervised framework that first learns a global aging reference from healthy controls (HCs) and then learns regional aging representations through atlas-guided regional replacement between HCs. To address challenge 1, donor chronological age serves as a weak proxy target for the output corresponding to the replaced region, enabling region-specific supervision without ground-truth regional age labels. We further develop a dedicated, multidimensional evaluation framework to assess complementary aspects of regional predictions. To address challenge 2, we assess the practical utility of the learned representations for global brain-age estimation and investigate disease-associated regional aging patterns in AD and PD.

Our contributions are summarized as follows.

1. Given the lack of regional brain-age labels and the limited region-specific training strategies in prior work, we design atlas-guided region mixing to provide regional brain-age supervision.
2. We develop a joint global–regional framework combining multi-scale anatomical readouts, real-image global supervision, and localized mixed-image losses to promote anatomical specificity while preserving global age-prediction accuracy.
3. We develop a dedicated evaluation suite for regional brain aging across six public cohorts, assessing weak-target agreement, anatomical specificity, spatial confinement, and regional brain-age ranking accuracy, alongside global age-prediction performance and disease-related associations.

2 Related Work

2.1 Global Brain Age Estimation

Early BAE methods used SVR, GPR, XGBoost, and PLSR (Franke et al., 2010; Cole et al., 2017; Baecker et al., 2021a; Niu et al., 2020; Blake et al., 2023; de Lange et al., 2022; Guan et al., 2024), performing well on small or homogeneous datasets but struggling with age and disease heterogeneity. Recent 3D CNNs learn global morphological features from MRI (Pardakhti & Sajedi, 2020) and can incorporate AD-sensitive parahippocampal and temporal regions (Poloni et al., 2022). Cascade networks and ranking losses improve accuracy and age-ordering consistency (Cheng et al., 2021). However, these methods do not explicitly characterize regional aging heterogeneity.

2.2 Regional Brain Age Estimation

Global brain age can obscure regional aging heterogeneity, motivating ReBA to characterize spatially distinct aging and disease-associated patterns (Nguyen et al., 2024; Riccardi et al., 2025).

Existing approaches include U-Net-based local age estimation (Popescu et al., 2021), voxel-wise multitask prediction (Gianchandani et al., 2024), anatomical aggregation of receptive-field-restricted predictions (Nguyen et al., 2024), and ROI- or lobe-specific regression (Lee et al., 2022; Kalc et al., 2024). Recent work further links local age deviations to cognition and neurodegeneration (Chaudhari et al., 2026). ReBA-Pred-Net introduces teacher-derived regional targets, functional-consistency regularization, and indirect regional evaluation (Shao et al., 2026). Complementary saliency methods and coVariance neural networks explain global predictions through input attribution and anatomically indexed contributions, respectively (Yin et al., 2023; Sihag et al., 2023).

Despite these advances, subject-level ages do not specify within-subject regional differences, while noise-perturbed or teacher-derived targets remain proxies rather than measurements of regional biological age. Moreover, the field still lacks dedicated computational training pipelines, established

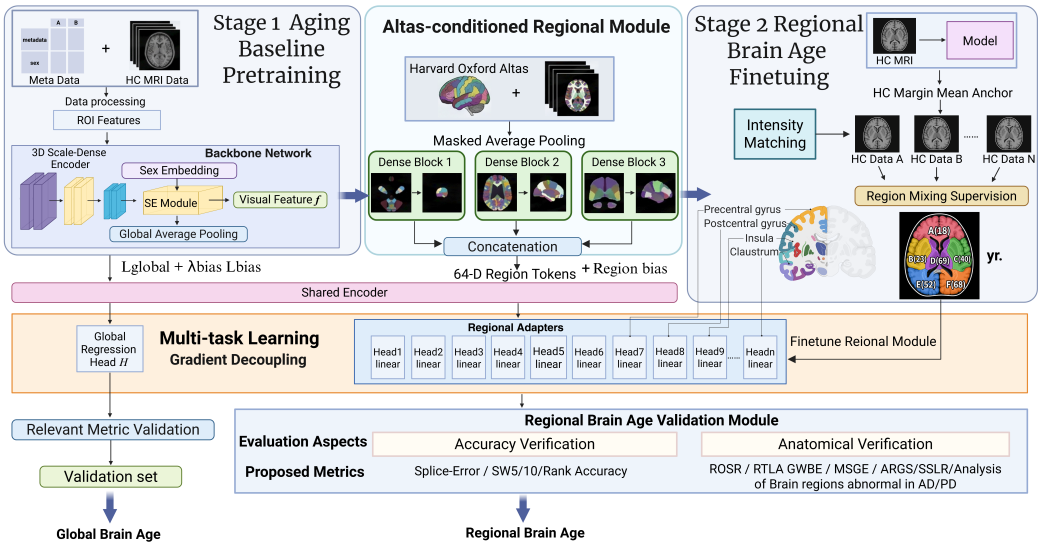


Figure 1: MT-ReBA overview. Stage 1 learns a global aging representation from healthy controls (HCs), while the atlas-conditioned regional module extracts multi-scale ROI features via Harvard–Oxford atlas-masked pooling to produce anatomically indexed regional representations. Stage 2 introduces region mixing supervision by integrating an atlas-defined parcel from an age-matched donor into a recipient image, using donor age as a weak supervisory signal for the corresponding regional head. The learned representations are evaluated through splice fidelity, anatomical specificity, and disease-related analyses.

evaluation metrics, and standardized benchmarks for regional brain-age learning. For models sharing whole-brain representations, anatomical output assignment and clinical associations alone do not establish region-specific sensitivity. These limitations motivate region-differentiated weak supervision together with explicit training and evaluation protocols for anatomical specificity.

3 METHOD

3.1 OVERVIEW OF MT-REBA

As illustrated in Figure 1, MT-ReBA learns anatomically indexed regional aging representations without regional labels by converting subject-level age information into localized weak supervision through atlas-guided region mixing. In Stage 1, a 3D-ScaleDenseNet encoder learns normative global aging from HCs. Given an MRI volume $X \in \mathbb{R}^{H\times W\times D}$ and biological sex $g \in \{0,1\}$, it combines visual features with a sex embedding to predict global age through a regression head.

The regional module applies multi-scale masked average pooling using downsampled Harvard–Oxford cortical atlas masks from FSL (Jenkinson et al., 2012). Pooled ROI features are combined with learnable region embeddings and processed by lightweight region-specific adapters and prediction heads. In Stage 2, region mixing replaces an HC recipient's atlas parcel with its counterpart from an intensity-matched HC donor, providing weak supervision from the donor age (Eq. 4).

3.2 STAGE 1: AGING BASELINE PRETRAINING

3.2.1 BACKBONE NETWORK

We use a 3D multi-scale dense encoder with anisotropic blocks (Huang et al., 2017) and squeeze-and-excitation attention (Hu et al., 2018) to learn aging features from T1w MRI. After global average pooling, f_i is concatenated with a learnable sex embedding $e_{g_i} \in \mathbb{R}^8$ to form the shared representation h_i. The global head predicts $\widehat{y}_i$ from h_i using the smooth ℓ_1 (Huber) loss with $\beta = 2$ (Huber, 1964), establishing the normative global aging representation used in both stages.

3.2.2 STAGE 1 OPTIMIZATION OBJECTIVE

Stage 1 minimizes the objective in Eq. 1:

$$\mathcal{L}^{(1)} = \mathcal{L}_{\text{global}} + \lambda_{\text{bias}} \mathcal{L}_{\text{bias}}, \tag{1}$$

where $\mathcal{L}_{\text{bias}}$ is the absolute empirical covariance between prediction residuals and mean-centred chronological age, and $\lambda_{\text{bias}} \geq 0$ is its weight. This term penalizes the age dependence of residuals associated with regression-to-the-mean bias in BAE (de Lange & Cole, 2020); its expanded definition and the separate post-hoc calibration rule are provided in Appendix K.

3.3 ATLAS-CONDITIONED REGIONAL REPRESENTATION LEARNING

After global baseline pretraining, MT-ReBA activates an atlas-conditioned regional module to learn anatomically indexed aging representations. Let $F_i^{(\ell)} \in \mathbb{R}^{C_\ell \times D_\ell \times H_\ell \times W_\ell}$ denote the feature map of subject i at encoder level ℓ, and let $M_r^{(\ell)}$ be the soft occupancy mask obtained by applying the encoder's pooling geometry to atlas parcel r. At the first three encoder levels, the atlas-masked regional descriptor is computed as:

$$z_{i,r}^{(\ell)} = \frac{\sum_v M_r^{(\ell)}(v) F_i^{(\ell)}(v)}{\sum_v M_r^{(\ell)}(v) + \varepsilon}, \qquad z_{i,r} = \underset{\ell=1}{\overset{3}{\text{concat}}} z_{i,r}^{(\ell)}. \tag{2}$$

Here, v indexes voxels, r indexes atlas parcels, and $\varepsilon > 0$ prevents division by zero. The resulting multi-scale regional descriptor $z_{i,r}$ is processed by a shared regional encoder and combined with a learnable region bias, a 64-dimensional region token, and the sex embedding. A region-specific prediction head then produces an anatomically indexed, age-related score $\hat{y}_{i,r}$. The mean-anchor objective constrains the aggregate age scale.

The atlas masks $M_r^{(\ell)}$ are implemented as soft occupancy maps, yielding fractional weights near parcel borders and reducing sensitivity to small mismatches from registration, resampling, and anatomical variability. Accordingly, atlas-masked pooling provides each regional representation with a consistent anatomical index . The regional module is inactive in Stage 1 and optimized in Stage 2.

3.4 STAGE 2: REGIONAL REPRESENTATION FINE-TUNING

Stage 2 introduces region mixing to provide localized weak supervision for anatomically specific regional learning. Let $\widehat{y}_{i,r}^{\,p}$ denote the prediction of regional head r in pass $p \in \{\text{real}, \text{mix}\}$. We define the corresponding regional target vector as

$$t_{i,r}^{\text{real}} = y_i, \qquad t_{i,r}^{\text{mix}} = \begin{cases} y_{j(i,r)}, & (i,r) \in \mathcal{C}, \\ y_i, & (i,r) \notin \mathcal{C}, \end{cases} \tag{3}$$

where $\mathcal{C}$ denotes the changed subject–region pairs and $j(i,r)$ identifies the donor providing parcel r to recipient i. Donor age provides a weak supervisory signal only for the transplanted parcel, while unchanged heads receive no element-wise supervision from the mixed image. To reduce splice-specific cues, donor intensities are matched to the recipient and parcel boundaries are blended.

3.4.1 LOCALIZED WEAK SUPERVISION VIA REGION MIXING

To provide localized weak supervision, we construct a mixed image by replacing an atlas-defined parcel $r_i^\star$ of recipient X_i with the corresponding parcel from donor X_j:

$$X_i^{\text{mix}} = \left(1 - \boldsymbol{\alpha}_{r_i^\star}\right) \odot X_i + \boldsymbol{\alpha}_{r_i^\star} \odot \text{match}(X_j; X_i), \tag{4}$$

where $\boldsymbol{\alpha}_{r_i^\star}$ is a soft atlas mask with smooth boundaries, $\text{match}(X_j; X_i)$ aligns donor intensities with the recipient, and $\odot$ denotes element-wise multiplication. Donor age provides a weak supervisory signal only for the regional head of the transplanted parcel:

$$\mathcal{L}_{\text{changed}} = \frac{1}{|\mathcal{C}|} \sum_{(i,r) \in \mathcal{C}} \left| \widehat{y}_{i,r}^{\text{mix}} - y_{j(i,r)} \right|, \tag{5}$$

where $\mathcal{C}$ denotes the changed subject–region pairs. Donor chronology associates the replaced parcel with age-related morphology observed in healthy subjects; by altering a specified parcel while retaining the recipient context, region mixing provides a localized age-proxy learning signal for the corresponding regional head. Unchanged regions receive no element-wise supervision from the mixed image, preventing donor or recipient age from being propagated across unrelated regional heads. Donor intensities are affine-matched within the selected parcel, and the atlas mask is Gaussian-feathered to reduce intensity and boundary cues unrelated to regional aging. Additional details on regional identifiability and synthetic-gradient isolation are provided in Appendices A and B, respectively, with full mixing and optimization settings in Appendix C.

3.4.2 Stage 2 Optimization Objective

To stabilize regional learning while preserving head-specific variation, we introduce a changed-region margin and a regional mean anchor:

$$\mathcal{L}_{\mathrm{HC}} = \frac{1}{|\mathcal{C}|} \sum_{(i,r)\in\mathcal{C}} \left[\left| \widehat{y}_{i,r}^{\mathrm{mix}} - t_{i,r}^{\mathrm{mix}} \right| - \delta \right]_+, \; \mathcal{L}_{\mathrm{anchor}}^{p} = \frac{1}{B_p} \sum_{i=1}^{B_p} \left| \frac{1}{R} \sum_{r=1}^{R} \widehat{y}_{i,r}^{p} - \frac{1}{R} \sum_{r=1}^{R} t_{i,r}^{p} \right|, \tag{6}$$

where $[u]_+ = \max(0, u)$, B_p is the number of samples in pass p. The margin is applied only to changed HC parcels, whereas the anchor calibrates the overall regional age scale through the subject-level mean without imposing element-wise targets on unchanged heads.

The Stage 2 objectives are then

$$\mathcal{L}_{\mathrm{real}} = \lambda_g \mathcal{L}_{\mathrm{global}} + \lambda_b \mathcal{L}_{\mathrm{bias}} + \lambda_a^{\mathrm{real}} \mathcal{L}_{\mathrm{anchor}}^{\mathrm{real}}, \; \mathcal{L}_{\mathrm{mix}} = \lambda_r \mathcal{L}_{\mathrm{changed}} + \lambda_{\mathrm{HC}} \mathcal{L}_{\mathrm{HC}} + \lambda_a^{\mathrm{mix}} \mathcal{L}_{\mathrm{anchor}}^{\mathrm{mix}}. \tag{7}$$

Here, $\mathcal{L}_{\mathrm{global}}$ denotes the real-image global-age loss and $\mathcal{L}_{\mathrm{bias}}$ follows Stage 1. The two objectives are optimized in separate steps: the real-image pass updates the backbone and regional pathway, whereas mixed-image gradients are restricted to the regional pathway.

4 Experiments

4.1 Experimental Setup

We evaluate MT-ReBA on six public cohorts spanning healthy aging and neurodegenerative disease. Of the 4,311 participants, 3,238 HCs are used for model training and testing, while the remaining 1,073 patients with AD or PD are reserved exclusively for downstream evaluation. All T1-weighted scans undergo FSL-based preprocessing: affine and nonlinear registration to MNI152 2-mm space, brain extraction, bias-field correction, and within-brain z-score normalization. Dataset composition and participant-level splits are detailed in Appendix E. Both stages run for 150 epochs on eight NVIDIA RTX A6000 GPUs, using Adam in Stage 1 and AdamW in Stage 2. Checkpoints are selected by the lowest validation global MAE (Willmott & Matsuura, 2005). Hyperparameters, compute details, and convergence analyses appear in Appendices C, H, and F. The global–regional objective ablation is reported in Appendix G.1 and Table 7, while the backbone architecture comparison is reported in Appendix G.2. Our metric suite assesses region-mixing fidelity using Splice-Error and SW_t on held-out mixed images, anatomical specificity through ROI occlusion and input-gradient locality on unspliced HCs, and spatial confinement using SSLR. Finally, the frozen HC-trained model is evaluated on original AD and PD scans to examine disease-associated regional aging patterns.

4.2 Experimental Results

4.2.1 Comparison Experiments on Global BAE

To comprehensively evaluate global brain age estimation, we employ four standard regression metrics: MAE, RMSE, R^2, and Pearson's correlation coefficient (ρ) (Willmott & Matsuura, 2005). Let N denote the total number of subjects, y_i the chronological age, $\hat{y}_i$ the predicted brain age, and $\bar{y}$ and $\bar{\hat{y}}$ their respective sample means. We construct a large-scale multi-site T1w MRI dataset by integrating six publicly available cohorts (GSP, ADNI, PPMI, OULU, OpenNeuro, and ICBM), covering diverse age ranges and acquisition sites. Using 3,238 HCs from this dataset, we establish a rigorously controlled benchmark to evaluate global brain age estimation.

Method	MAE ± SD	95% CI	RMSE	R^2	ρ
Dufumier et al. (Dufumier et al., 2021)	4.04 ± 0.37	$[3.35, 4.81]$	6.75	0.88	0.94
Barbano et al. (Barbano et al., 2023)	3.93 ± 0.37	$[3.24, 4.66]$	6.67	0.89	0.94
Zha et al. (Zha et al., 2023)	3.87 ± 0.37	$[3.18, 4.63]$	6.65	0.89	0.94
Träuble et al. (Träuble et al., 2024)	3.95 ± 0.36	$[3.28, 4.69]$	6.55	0.89	0.94
Barbano et al. (Barbano et al., 2026)	3.64 ± 0.32	$[3.04, 4.29]$	5.92	0.91	0.96
MT-ReBA (Ours)	$\mathbf{3.10 \pm 0.44}$	$[2.31, 4.05]$	**5.71**	**0.94**	**0.97**

Table 1: Comparison of global BAE on HCs. MAE uncertainty is estimated from 10,000 subject-level bootstrap resamples and reported as ± SD with 95% CIs. R^2 and ρ are point estimates.

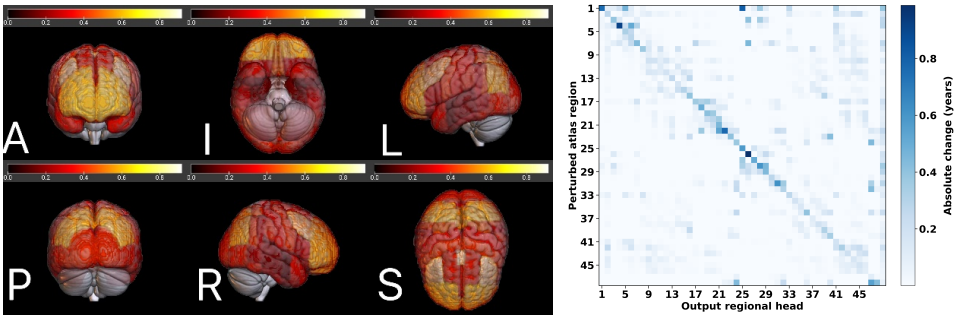


Figure 2: Anatomical specificity on unspliced HC MRI. (a) Self-response after occluding each atlas parcel. (b) The 48×48 response matrix, where rows are occluded parcels and columns are regional heads. A strong diagonal indicates parcel-to-head specificity.

MT-ReBA achieves an MAE of 3.10 ± 0.44 years in Table 1, improving over the strongest re-implemented baseline by 14.84% (Barbano et al., 2026), with RMSE $= 5.71$, $R^2 = 0.94$, and $\rho = 0.97$. Additional analysis is provided in Appendix F.

4.2.2 Regional Anatomical Specificity

To determine whether the learned regional specificity extends beyond the region-mixing task rather than being driven solely by splice-specific cues, we evaluate anatomical specificity on held-out, unspliced HC scans. ROI occlusion tests whether perturbing parcel r preferentially affects its corresponding head, while input-gradient locality examines whether the same head draws sensitivity from that parcel on the original image.

For ROI occlusion, voxels within parcel r are replaced by the held-out HC cohort-mean intensities, with all other voxels unchanged. We measure the response of regional head k as

$$\Delta_{r,k} = \frac{1}{N_{\mathrm{HC}}} \sum_{i=1}^{N_{\mathrm{HC}}} \left| \widehat{y}_{i,k}\left(X_i^{(r)}\right) - \widehat{y}_{i,k}(X_i) \right|, \tag{8}$$

where $X_i^{(r)}$ denotes the image with parcel r occluded. Repeating Eq. 8 over all $R = 48$ parcels and heads yields the response matrix in Figure 2. Regional Occlusion Specificity Ratio (ROSR) compares matched and mismatched responses, while Regional Top-1 Localization Accuracy (RTLA) measures whether the strongest response occurs in the corresponding head:

$$\mathrm{ROSR} = \frac{R^{-1} \sum_{r=1}^{R} \Delta_{r,r}}{[R(R-1)]^{-1} \sum_{r=1}^{R} \sum_{k \neq r} \Delta_{r,k}}, \qquad \mathrm{RTLA} = \frac{1}{R} \sum_{r=1}^{R} \mathbf{1}\left[r = \arg\max_{k} \Delta_{r,k} \right]. \tag{9}$$

The mean matched response (0.3536 years) exceeds the mismatched response (0.0245 years), yielding ROSR $= 14.45\times$. RTLA reaches 75.0%, compared with a 2.1% chance level, with both metrics significant under parcel-label permutation tests ($p = 0.0005$). These results support atlas-aligned regional responses on unspliced scans, independent of donor replacement.

Method	Seed 1	Seed 2	Seed 3	Mean ± SD
Voxel-level (Gianchandani et al., 2024)	55.94	60.79	60.94	59.22 ± 2.84
Global CNN	56.47	55.76	54.76	55.67 ± 0.86
SFCN (Peng et al., 2021)	56.07	55.61	56.19	55.95 ± 0.31
MT-ReBA (Ours)	**69.69**	**69.69**	**69.50**	$\mathbf{69.63 \pm 0.11}$

Table 2: Regional brain age ranking accuracy. Mean and sample SD are computed across three training seeds.

Because occlusion itself modifies the input, we further assess locality using input gradients computed directly on intact scans. Let $G_{i,r}(v) = |\partial \widehat{y}_{i,r}/\partial X_i(v)|$. For voxel set Ω, we define

$$m_{i,r}(\Omega) = \sum_{v \in \Omega} G_{i,r}(v), \qquad \gamma_{i,r}(\Omega) = \frac{m_{i,r}(\Omega)}{m_{i,r}(V)}, \qquad d_{i,r}(\Omega) = \frac{m_{i,r}(\Omega)}{|\Omega|}. \tag{10}$$

We compare self-ROI gradient density against whole-brain, volume-matched random-ROI, and adjacent-ROI controls using Gradient-to-Whole-Brain Enrichment (GWBE), Matched-Size Gradient Enrichment (MSGE), and Adjacent-ROI Gradient Specificity (ARGS):

$$\mathrm{GWBE} = \mathbb{E}_{i,r}\left[\frac{d_{i,r}(M_r)}{d_{i,r}(M_{\mathrm{brain}})}\right], \ \mathrm{MSGE} = \mathbb{E}_{i,r}\left[\frac{d_{i,r}(M_r)}{\bar{d}^{\mathcal{Q}}_{i,r}}\right], \ \mathrm{ARGS} = \mathbb{E}_{i,r}\left[\frac{d_{i,r}(M_r)}{d_{i,r}(M_r^{\mathrm{adj}})}\right]. \tag{11}$$

Here, $\bar{d}^{\mathcal{Q}}_{i,r}$ is the mean density of $K = 3$ volume-matched random ROIs and M_r^{adj} denotes adjacent atlas parcels. We obtain $\mathrm{GWBE} = 14.29\times$, $\mathrm{MSGE} = 768.22\times$, and $\mathrm{ARGS} = 3.05\times$; all regional heads also exceed their ROI-volume null, with a mean self-ROI gradient fraction of 9.40%. Together with the occlusion results, these complementary tests support anatomically localized regional representations on unspliced MRI. Additional analyses are provided in Appendix M.

4.2.3 Assessing Localized Weak Supervision

We evaluate agreement with localized weak targets on held-out mixed images. For replaced instances $(i, r) \in \mathcal{S}$, let $e_{i,r} = |\widehat{y}_{i,r} - y_{j(i,r)}|$ denote the error relative to the donor age. We compute

$$\text{Splice-Error} = \frac{1}{|\mathcal{S}|} \sum_{(i,r) \in \mathcal{S}} e_{i,r}, \qquad \mathrm{SW}_t = \frac{1}{|\mathcal{S}|} \sum_{(i,r) \in \mathcal{S}} \mathbf{1}[e_{i,r} \le t]. \tag{12}$$

Lower Splice-Error and higher SW5/SW10 indicate better weak-target agreement. Across mixed regions, MT-ReBA obtains a Splice-Error of 8.13 years, SW5 of 61.96%, and SW10 of 76.93%.

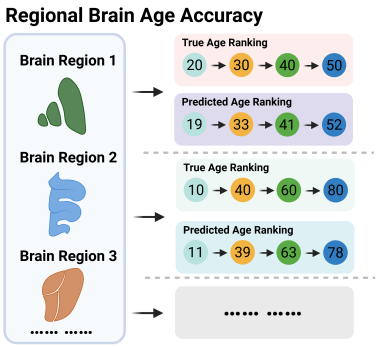
4.2.4 Regional Age Ranking Accuracy

Because there are no established evaluation metrics or standardized benchmarks for assessing regional brain-age estimates, we introduce an intervention-based regional age ranking accuracy to evaluate whether regional predictions preserve age ordering across valid donor pairs (Figure 3). Let $\hat{a}_{ir}^{(d)}$ denote the regional prediction for recipient i, region r, and donor d. We define $\mathcal{P}_i = \{(d, e) : a_e - a_d \ge 2\}$ as the set of valid donor pairs ordered from younger to older, and compute

$$\mathrm{RankAcc} = \frac{100}{NR} \sum_{i=1}^{N} \sum_{r=1}^{R} \frac{1}{|\mathcal{P}_i|} \sum_{(d,e) \in \mathcal{P}_i} \mathbf{1}\left[\hat{a}_{ir}^{(e)} - \hat{a}_{ir}^{(d)} > \epsilon\right], \tag{13}$$

where N is the number of evaluated recipients, $R = 48$, and $\epsilon = 10^{-6}$ years.

We reimplement several existing regional brain-age methods and evaluate them under our proposed evaluation framework. Table 2 reports the matched comparison on a held-out test set of 356 HCs using three training seeds. MT-ReBA achieves $69.63 \pm 0.11\%$, exceeding the voxel-level baseline by 10.41 percentage points and attaining the highest accuracy for every seed.



Pairwise Comparisons

Pair	True Order	Predicted Order	Correct
(20,30)	20<30	21<31	✓
(20,40)	20<40	21>42	✗
(20,50)	20<50	21<50	✓
.....	.....	.....	.....
(40,50)	40<50	44<53	✓
(50,60)	50<60	48<62	✓
.....	.....	.....	.....
(60,70)	60<70	63<69	✓

Figure 3: Evaluation of regional age ranking consistency. We evaluate all valid donor pairs and count a pair as correct when the predicted regional ages preserve the donors' chronological age ordering.

4.2.5 Validating Spatial Confinement of Region Mixing

We use the Splice Spatial Leakage Rate (SSLR) to verify whether detectable image changes caused by region mixing remain inside the selected atlas parcel. For each mixed instance $(i, r) \in \mathcal{A}_{\text{splice}}$, we define the voxel-wise intensity difference and SSLR jointly as

$$d_{i,r}(v) = \left| X_{i,r}^{\text{mix}}(v) - X_i(v) \right|, \ \text{SSLR} = \frac{\sum_{(i,r)\in\mathcal{A}_{\text{splice}}} \sum_{v\in M_{\text{brain}}\setminus M_r} \mathbf{1}[d_{i,r}(v) > \tau]}{\sum_{(i,r)\in\mathcal{A}_{\text{splice}}} \sum_{v\in M_{\text{brain}}} \mathbf{1}[d_{i,r}(v) > \tau]} \tag{14}$$

where M_r and M_{brain} are the target-parcel and brain masks, and τ is the detection tolerance.

As shown in Figure 4, the mixing weights and intensity changes are confined to the selected parcel, yielding $\text{SSLR} = 0\%$. This verifies spatial confinement of the mixing operation. To reduce splice-related intensity and boundary cues within the target parcel, we use intensity matching and mask feathering (Section 3.4.1). The preceding analyses on unspliced MRI test whether regional heads respond preferentially to their corresponding parcels even without donor splicing.

4.2.6 Regional Aging Patterns in Neurodegenerative Cohorts

To examine whether the learned regional representations capture disease-associated aging patterns, we apply the HC-trained model to unspliced AD and PD scans. We compare age- and sex-matched groups within each study (ADNI–AD vs. ADNI–HC and PPMI–PD vs. PPMI–HC). Figure 5 shows the six regions with the largest group differences in regional BAG.

In AD, the largest shift occurs in the posterior parahippocampal cortex (Region 35; $+10.93$ years), followed by anterior parahippocampal and temporal regions, consistent with established medial-temporal involvement in AD (Dickerson et al., 2009; Desikan et al., 2009). In PD, significant effects concentrate in motor-related cortex, particularly the juxtapositional lobule/supplementary motor area (Region 26; 6.48 years, $q = 0.00232$) and precentral gyrus (Region 7; 6.04 years, $q = 0.00232$) (Kalia & Lang, 2015). These findings provide exploratory disease-related validation of the learned regional representations on MRI.

5 Conclusion and Future Work

Conclusion. MT-ReBA learns anatomically grounded regional aging representations without regional labels through two-stage weak supervision. Atlas-guided region mixing provides localized supervision, while gradient isolation preserves global representations. Occlusion and gradient analyses on unspliced HCs support anatomical specificity, and AD/PD analyses highlight the value of regional representations for characterizing disease-associated aging patterns.

Limitations and Future Work. The current regional analysis is restricted to 48 cortical parcels. Future work will extend MT-ReBA to finer cortical and subcortical parcellations for more detailed and comprehensive regional aging analysis.

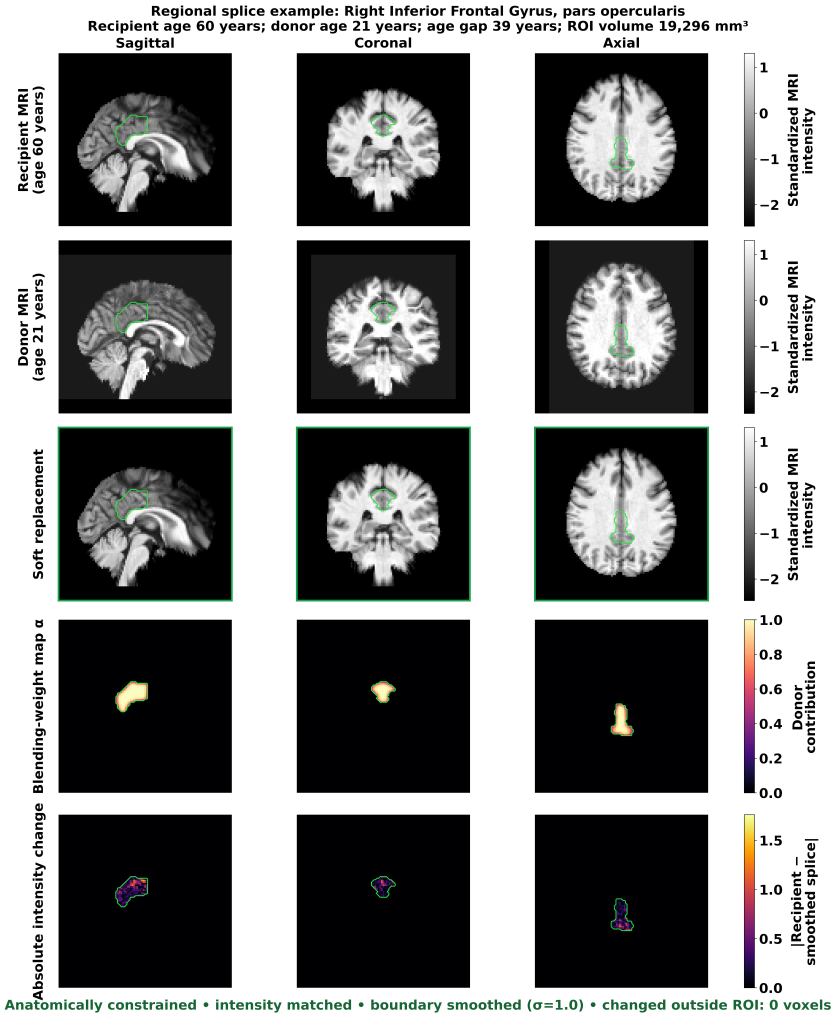


Figure 4: Region mixing across sagittal, coronal, and axial views. Rows show host, donor, mixed image, mixing map α, and absolute difference.

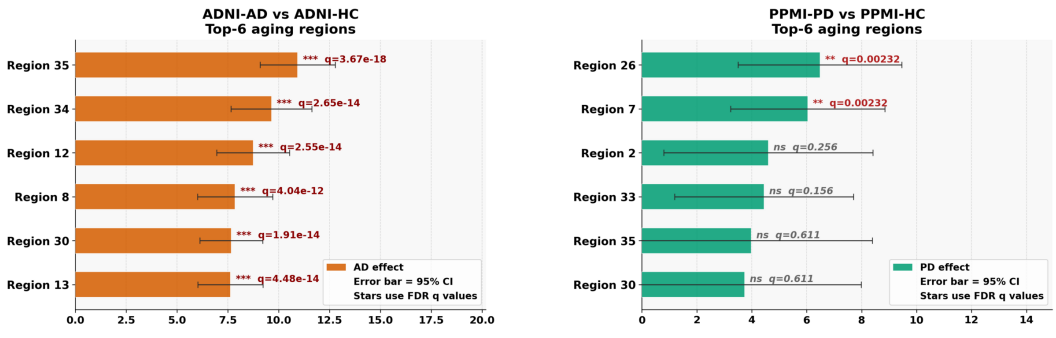


Figure 5: Disease-associated regional BAG patterns in age- and sex-matched comparisons. Left: ADNI–AD vs. ADNI–HC. Right: PPMI–PD vs. PPMI–HC. Bars show mean differences with 95% CIs; q-values are FDR-adjusted across 48 parcels.

AI USE STATEMENT

The authors utilized large language models (LLMs) during the preparation of this manuscript for proofreading, correcting grammatical errors, and improving sentence structure and readability. All LLM-assisted suggestions were manually reviewed and edited by the authors to ensure that the final text accurately reflects the research and its claims. The scientific contributions, methodology, experimental design, results, and conclusions remain the responsibility of the authors.

ETHICS STATEMENT

This study uses pre-existing T1-weighted MRI data from publicly available neuroimaging resources, including GSP, ADNI, PPMI, OULU, OpenNeuro, and ICBM. No new data were collected for this study. Data access and reuse are subject to the applicable data-use terms and privacy requirements of the respective providers.

REPRODUCIBILITY STATEMENT

We provide implementation and experimental details throughout the main paper and appendix to facilitate reproducibility. Section 3 describes the two-stage training procedure, atlas-conditioned regional representation, region-mixing construction, and optimization objectives. Appendix C reports the architecture and optimization hyperparameters, including the random seed and settings for both training stages. Appendix E details the preprocessing pipeline, dataset composition, and participant-level data splits. Appendix H specifies the computational environment and hardware used for training. The sources of all neuroimaging datasets are cited in the paper.

REFERENCES

M. Azzam et al. A review of artificial intelligence-based brain age estimation and its applications for related diseases. *Briefings in Functional Genomics*, 2025. doi: 10.1093/bfgp/elae042.

L. Baecker, J. Dafflon, P. F. Da Costa, et al. Brain age prediction: A comparison between machine learning models using region- and voxel-based morphometric data. *Human Brain Mapping*, 42: 2332–2346, 2021a.

L. Baecker, R. Garcia-Dias, S. Vieira, et al. Machine learning for brain age prediction: Introduction to methods and clinical applications. *EBioMedicine*, 72:103600, 2021b. doi: 10.1016/j.ebiom.2021.103600.

C. A. Barbano, B. Dufumier, E. Duchesnay, M. Grangetto, and P. Gori. Contrastive learning for regression in multi-site brain age prediction. In *2023 IEEE 20th International Symposium on Biomedical Imaging*, pp. 1–4. IEEE, 2023. doi: 10.1109/ISBI53787.2023.10230733.

C. A. Barbano, B. Dufumier, E. Duchesnay, M. Grangetto, and P. Gori. Robust brain age estimation from structural MRI with contrastive learning. *Pattern Recognition Letters*, 203:78–84, 2026. doi: 10.1016/j.patrec.2026.02.032.

Bharat B. Biswal, Maarten Mennes, Xi-Nian Zuo, Suril Gohel, Clare Kelly, et al. Toward discovery science of human brain function. *Proceedings of the National Academy of Sciences*, 107(10): 4734–4739, 2010. doi: 10.1073/pnas.0911855107.

K. V. Blake, Z. Ntwatwa, T. Kaufmann, et al. Advanced brain ageing in adult psychopathology: A systematic review and meta-analysis of structural MRI studies. *Journal of Psychiatric Research*, 157:180–191, 2023. doi: 10.1016/j.jpsychires.2022.11.011.

N. N. Chaudhari, O. M. Vega Huerta, S. Bhattacharya, N. F. Chowdhury, A. Irimia, and Alzheimer's Disease Neuroimaging Initiative. Deep learning maps local brain aging in relation to cognition across human adulthood. *Proceedings of the National Academy of Sciences*, 123(32): e2532233123, 2026. doi: 10.1073/pnas.2532233123.

J. Cheng, Z. Liu, H. Guan, Z. Wu, H. Zhu, J. Jiang, W. Wen, D. Tao, and T. Liu. Brain age estimation from MRI using cascade networks with ranking loss. *IEEE Transactions on Medical Imaging*, 40 (12):3400–3412, 2021. doi: 10.1109/TMI.2021.3085948.

J. H. Cole, R. P. K. Poudel, D. Tsagkrasoulis, et al. Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. *NeuroImage*, 163:115–124, 2017. doi: 10.1016/j.neuroimage.2017.07.059.

A.-M. G. de Lange and J. H. Cole. Commentary: Correction procedures in brain-age prediction. *NeuroImage: Clinical*, 26:102229, 2020. doi: 10.1016/j.nicl.2020.102229.

A.-M. G. de Lange, M. Anatürk, J. Rokicki, et al. Mind the gap: Performance metric evaluation in brain-age prediction. *Human Brain Mapping*, 43:3113–3129, 2022. doi: 10.1002/hbm.25837.

R. S. Desikan, H. J. Cabral, C. P. Hess, et al. Automated MRI measures identify individuals with mild cognitive impairment and alzheimer's disease. *Brain*, 132(8):2048–2057, 2009. doi: 10.1093/brain/awp123.

B. C. Dickerson, A. Bakkour, D. H. Salat, et al. The cortical signature of alzheimer's disease: Regionally specific cortical thinning relates to symptom severity in very mild to mild AD dementia and is detectable in asymptomatic amyloid-positive individuals. *Cerebral Cortex*, 19(3):497–510, 2009. doi: 10.1093/cercor/bhn113.

B. Dufumier, P. Gori, J. Victor, A. Grigis, M. Wessa, P. Brambilla, et al. Contrastive learning with continuous proxy meta-data for 3D MRI classification. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*, pp. 58–68. Springer, 2021. doi: 10.1007/978-3-030-87196-3_6.

K. Franke, G. Ziegler, S. Klöppel, C. Gaser, and Alzheimer's Disease Neuroimaging Initiative. Estimating the age of healthy subjects from T1-weighted MRI scans using kernel methods: Exploring the influence of various parameters. *NeuroImage*, 50(3):883–892, 2010. doi: 10.1016/j.neuroimage.2010.01.005.

Neha Gianchandani, Mahsa Dibaji, Johanna Ospel, Fernando Vega, Mariana Bento, M. Ethan MacDonald, and Roberto Souza. A voxel-level approach to brain age prediction: A method to assess regional brain aging. *Machine Learning for Biomedical Imaging*, 2:761–795, 2024. doi: 10.59275/j.melba.2024-4dg2.

S. Guan, R. Jiang, C. Meng, et al. Brain age prediction across the human lifespan using multimodal MRI data. *GeroScience*, 46:1–20, 2024. doi: 10.1007/s11357-023-00924-0.

A. J. Holmes, M. O. Hollinshead, T. M. O'Keefe, et al. Brain genomics superstruct project initial data release with structural, functional, and behavioral measures. *Scientific Data*, 2:150031, 2015. doi: 10.1038/sdata.2015.31.

J. Hu, L. Shen, and G. Sun. Squeeze-and-excitation networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7132–7141, 2018.

G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4700–4708, 2017.

P. J. Huber. Robust estimation of a location parameter. *The Annals of Mathematical Statistics*, 35 (1):73–101, 1964. doi: 10.1214/aoms/1177703732.

M. Jenkinson, C. F. Beckmann, T. E. J. Behrens, M. W. Woolrich, and S. M. Smith. FSL. *NeuroImage*, 62(2):782–790, 2012. doi: 10.1016/j.neuroimage.2011.09.015.

Polona Kalc, Robert Dahnke, Felix Hoffstaedter, and Christian Gaser. BrainAGE: Revisited and reframed machine learning workflow. *Human Brain Mapping*, 45(3):e26632, 2024. doi: 10.1002/hbm.26632.

L. V. Kalia and A. E. Lang. Parkinson's disease. *The Lancet*, 386(9996):896–912, 2015.

T. Kaufmann, D. van der Meer, N. T. Doan, E. Schwarz, M. J. Lund, I. Agartz, et al. Common brain disorders are associated with heritable patterns of apparent aging of the brain. *Nature Neuroscience*, 22(10):1617–1623, 2019.

P.-L. Lee et al. Regional rather than global brain age mediates cognitive function in cerebral small vessel disease. *Brain Communications*, 4(5):fcac233, 2022. doi: 10.1093/braincomms/fcac233.

K. Marek, D. Jennings, S. Lasch, A. Siderowf, C. Tanner, T. Simuni, et al. The parkinson progression marker initiative (PPMI). *Progress in Neurobiology*, 95(4):629–635, 2011. doi: 10.1016/j.pneurobio.2011.09.005.

C. J. Markiewicz, K. J. Gorgolewski, F. Feingold, R. Blair, Y. O. Halchenko, E. Miller, et al. The OpenNeuro resource for sharing of neuroscience data. *eLife*, 10:e71774, 2021. doi: 10.7554/eLife.71774.

John C. Mazziotta, Arthur W. Toga, Alan Evans, Peter Fox, and Jack Lancaster. A probabilistic atlas of the human brain: Theory and rationale for its development: The international consortium for brain mapping (ICBM). *NeuroImage*, 2(2):89–101, 1995. doi: 10.1006/nimg.1995.1012.

S. G. Mueller, M. W. Weiner, L. J. Thal, R. C. Petersen, C. Jack, W. Jagust, J. Q. Trojanowski, A. W. Toga, and L. Beckett. The alzheimer's disease neuroimaging initiative. *Neuroimaging Clinics of North America*, 15(4):869–877, 2005. doi: 10.1016/j.nic.2005.09.008.

H.-D. Nguyen, M. Clément, B. Mansencal, and P. Coupé. Brain structure ages—a new biomarker for multi-disease classification. *Human Brain Mapping*, 45(1):e26558, 2024. doi: 10.1002/hbm.26558.

X. Niu, F. Zhang, J. Kounios, et al. Improved prediction of brain age using multimodal neuroimaging data. *Human Brain Mapping*, 41:1626–1643, 2020. doi: 10.1002/hbm.24899.

N. Pardakhti and H. Sajedi. Brain age estimation based on 3D MRI images using 3D convolutional neural network. *Multimedia Tools and Applications*, 79:25051–25065, 2020.

Han Peng, Weikang Gong, Christian F. Beckmann, Andrea Vedaldi, and Stephen M. Smith. Accurate brain age prediction with lightweight deep neural networks. *Medical Image Analysis*, 68:101871, 2021. doi: 10.1016/j.media.2020.101871.

K. M. Poloni, R. J. Ferrari, and Alzheimer's Disease Neuroimaging Initiative. A deep ensemble hippocampal CNN model for brain age estimation applied to alzheimer's diagnosis. *Expert Systems with Applications*, 195:116622, 2022. doi: 10.1016/j.eswa.2022.116622.

S. G. Popescu, B. Glocker, D. J. Sharp, and J. H. Cole. Local brain-age: A U-Net model. *Frontiers in Aging Neuroscience*, 13:761954, 2021. doi: 10.3389/fnagi.2021.761954.

N. Riccardi, A. Teghipco, S. Newman-Norlund, R. Newman-Norlund, I. Rangus, C. Rorden, J. Fridriksson, et al. Distinct brain age gradients across the adult lifespan reflect diverse neurobiological hierarchies. *Communications Biology*, 8:802, 2025. doi: 10.1038/s42003-025-08228-z.

S. Shao et al. ReBA-Pred-Net: Weakly-supervised regional brain age prediction on MRI. *arXiv preprint arXiv:2602.12751*, 2026. doi: 10.48550/arXiv.2602.12751.

S. Sihag, G. Mateos, C. T. McMillan, and A. Ribeiro. Explainable brain age prediction using covariance neural networks. In *Advances in Neural Information Processing Systems*, volume 36, pp. 46958–46988, 2023. doi: 10.52202/075280-2034.

S. M. Smith, L. T. Elliott, F. Alfaro-Almagro, P. McCarthy, T. E. Nichols, G. Douaud, and K. L. Miller. Brain aging comprises many modes of structural and functional change with distinct genetic and biophysical associations. *eLife*, 9:e52677, 2020. doi: 10.7554/eLife.52677.

J. Träuble, L. V. Hiscox, C. L. Johnson, C.-B. Schönlieb, G. S. Kaminski Schierle, and A. I. Aviles-Rivero. Contrastive learning with adaptive neighborhoods for brain age prediction on 3D stiffness maps. *Transactions on Machine Learning Research*, 2024.

C. J. Willmott and K. Matsuura. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, 30(1):79–82, 2005. doi: 10.3354/cr030079.

C. Yin, P. Imms, M. Cheng, A. Amgalan, N. F. Chowdhury, R. J. Massett, N. N. Chaudhari, X. Chen, P. M. Thompson, P. Bogdan, A. Irimia, and Alzheimer's Disease Neuroimaging Initiative. Anatomically interpretable deep learning of brain age captures domain-specific cognitive impairment. *Proceedings of the National Academy of Sciences*, 120(2):e2214634120, 2023. doi: 10.1073/pnas.2214634120.

S. Yun, D. Han, S. J. Oh, S. Chun, J. Choe, and Y. Yoo. CutMix: Regularization strategy to train strong classifiers with localizable features. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 6023–6032, 2019.

K. Zha, P. Cao, J. Son, Y. Yang, and D. Katabi. Rank-n-contrast: Learning continuous representations for regression. In *Advances in Neural Information Processing Systems*, volume 36, 2023.

H. Zhang, M. Cissé, Y. N. Dauphin, and D. Lopez-Paz. mixup: Beyond empirical risk minimization. In *International Conference on Learning Representations*, 2018.

A Theoretical Supplement on Multi-task Learning in MT-ReBA

MT-ReBA combines global age regression on real HCs with region-addressable donor-age proxy supervision on mixed HCs. Equations 3–7 in the main paper define every active Stage 2 loss, the optional inactive regularizers, and their weights.

For a mixed sample with changed parcel $r^\star$, only head $r^\star$ receives the donor-age proxy target. Sampling interventions across all parcels breaks the permutation symmetry that would arise if every regional head were supervised by the same subject age and links each output index to its atlas-defined parcel.

B Synthetic-Gradient Isolation Safeguard

Region mixing can introduce intensity and boundary cues that differ from those in real MRI. To limit their direct influence on the pretrained global age-predictive representation, Stage 2 separates real-image and mixed-image optimization through explicit gradient routing and normalization-statistic control. This mechanism serves as an optimization safeguard; the region-specific supervision comes from the localized donor-age targets, rather than from gradient isolation itself.

Let θ denote the full parameter vector and Θ_{ROI} the set of parameter blocks in the atlas-conditioned regional pathway:

$$\Theta_{\mathrm{ROI}} = \{\theta_{\text{reg-enc}}, \theta_{\text{bias}}, \theta_{\text{token}}, \theta_{\text{adapter}}, \theta_{\text{cal}}, \theta_{\text{head}}\}. \tag{15}$$

These blocks comprise the regional encoder, region biases and tokens, adapters, calibration layers, and regional prediction heads.

Real and mixed images are processed in separate backward passes. For each parameter block p, the real-image pass retains

$$g_{\mathrm{real}}(p) = \nabla_p \mathcal{L}_{\mathrm{real}}, \tag{16}$$

whereas the mixed-image pass retains only regional gradients:

$$g_{\mathrm{mix}}(p) = \begin{cases} \nabla_p \mathcal{L}_{\mathrm{mix}}, & p \in \Theta_{\mathrm{ROI}}, \\ \mathbf{0}, & p \notin \Theta_{\mathrm{ROI}}. \end{cases} \tag{17}$$

These expressions specify loss-gradient routing, rather than the optimizer's parameter-update rule. The backbone and global head receive no direct gradient from $\mathcal{L}_{\mathrm{mix}}$, while the regional pathway receives supervision from both passes. Stage 2 uses AdamW with the parameter-group settings in Appendix C, including learning rates of 2.5×10^{-6} for the backbone and 10^{-5} for the regional pathway.

Parameter	Value	Parameter	Value
Input grid	$91 \times 109 \times 91$	Voxel resolution	$2 \times 2 \times 2\,\mathrm{mm}^3$
Regional heads	48	Regional ROI levels	3
Backbone blocks	4	Base filters	8
Dropout	0.3	Sex embedding dimension	8
Region token dimension	64	Random seed	42

Table 3: Shared data and architecture settings used in both training stages.

Parameter	Value	Parameter	Value
Epochs	150	Global batch size	8
LR reduction factor	0.5	Initial learning rate	1×10^{-4}
Backbone weight decay	5×10^{-4}	Maximum gradient norm	5.0
Warm-up epochs	10	Warm-up start LR	1×10^{-5}
Scheduler patience	10 epochs	Minimum learning rate	1×10^{-6}
Huber parameter β	2.0	Bias-loss weight λ_{bias}	0.05

Table 4: Stage 1 global aging baseline settings.

During each mixed-image forward pass, all BatchNorm layers temporarily use their stored running statistics without updating them. This prevents synthetic images from modifying the running statistics and complements the explicit gradient mask.

The safeguard restricts direct synthetic-loss gradients without freezing the backbone throughout Stage 2, allowing continued adaptation to real HC data. However, the regional pathway still processes mixed-image features, so gradient isolation alone does not exclude reliance on splice-specific cues. We therefore assess anatomical specificity separately through ROI occlusion and input-gradient locality on unspliced MRI.

C Key Hyperparameters

The settings used in the submitted experiments are summarized in Table 3 for shared data and architecture choices, Table 4 for Stage 1, and Table 5 for Stage 2. Implementation-only quantities that do not affect the reported protocol are omitted. All scans and atlas masks use the same MNI152 2-mm grid, and each of the 48 regional heads is assigned to one Harvard–Oxford cortical parcel.

Stage 1 optimization settings are summarized in Table 4.

Table 5 reports the Stage 2 settings. The backbone learning rate applies only to the real-HC pass; mixed-image gradients remain restricted to the regional pathway as defined in Appendix B.

D Pathological Interpretation of AD- and PD-related Regional Aging

The main paper reports age- and sex-matched ADNI–AD/HC and PPMI–PD/HC comparisons. The HC-trained model is frozen and applied to unspliced disease scans to evaluate downstream transfer beyond synthetic training interventions.

E Data Processing and Training

Table 6 summarizes the six cohorts, including cohort size, age range, and sex distribution. The data sources are GSP (Holmes et al., 2015), PPMI (Marek et al., 2011), ADNI (Mueller et al., 2005), the OULU cohort distributed through the 1000 Functional Connectomes Project (Biswal et al., 2010), OpenNeuro (Markiewicz et al., 2021), and ICBM (Mazziotta et al., 1995).

Parameter	Value	Parameter	Value
Epochs	150	Global batch size	8
Gradient accumulation	1	Optimizer	AdamW
Regional/head learning rate	1×10^{-5}	Backbone LR scale	0.25
Backbone learning rate	2.5×10^{-6}	Weight decay	1×10^{-4}
Minimum learning rate	1×10^{-6}	Maximum gradient norm	1.0
Splice probability	0.5	Replaced parcels/sample	1
Minimum donor–recipient BAG	8 years	Blend sigma	1.0
HC margin δ	3 years	Changed-region weight λ_r	2.0
HC-margin weight λ_{HC}	5.0	Anchor weights $\lambda_a^{\mathrm{real}}, \lambda_a^{\mathrm{mix}}$	1.0, 1.0

Table 5: Stage 2 regional fine-tuning and region-mixing settings.

Dataset	Total	Age range	Male
GSP (Brain Genomics Superstruct Project)	1,476	19–35	41.9%
PPMI (Parkinson's Progression Markers Initiative)	441	31–85	59.6%
ADNI (Alzheimer's Disease Neuroimaging Initiative)	845	56–93	52.1%
OULU (Oulu Neuroimaging Dataset)	102	21	35.3%
OpenNeuro (OpenNeuro)	1,427	8–92	41.7%
ICBM (International Consortium for Brain Mapping)	20	13–20	35.0%
Total	**4,311**	**8–93**	**45.4%**

Table 6: Dataset composition used in the submitted paper. Total denotes cohort size, and Male denotes the cohort-level percentage.

All T1w scans are processed with the same FSL-based pipeline (Jenkinson et al., 2012). Let I_{raw} denote a raw MRI volume. Affine and nonlinear registration map each scan to MNI152 2-mm space:

$$I_{\mathrm{lin}} = T_{\mathrm{affine}}(I_{\mathrm{raw}}), \qquad I_{\mathrm{nonlin}} = T_{\mathrm{nonlinear}}(I_{\mathrm{lin}}). \tag{18}$$

Brain extraction and bias-field correction are then applied, and the resulting brain image is resampled to $91 \times 109 \times 91$. Within-brain z-score normalization is performed as

$$x(v) = \frac{I_{\mathrm{brain}}(v) - \mu_I}{\sigma_I}, \qquad v \in M_{\mathrm{brain}}, \tag{19}$$

where μ_I and σ_I are computed over brain voxels only.

We partition the data at the participant level into training, validation, and test sets in a 70:15:15 ratio. The same split is used in both stages and held fixed across runs, with region-mixing donors sampled exclusively from the training partition. Our evaluation protocol comprises three independently seeded training runs per experimental setting, using a common seed set across methods to assess variability arising from initialization and stochastic optimization. Training schedules and hyperparameters for both stages are detailed in Appendix C.

F Convergence Analysis

Convergence is assessed over the submitted 150-epoch protocol. The numerical optimization settings for both stages are listed in Appendix C.

Controlled global brain age results

Table 1 in the main paper reports the controlled comparison on the 3,238-HC benchmark under identical experimental and data settings.

In the controlled comparison, MT-ReBA achieves the lowest point-estimate MAE among all evaluated methods, indicating strong global BAE performance under identical experimental settings.

Experiment	Global MAE	Splice-Error	SW5	HC BAG Std	PD-prior BAG	AD-prior BAG
Global only	4.39	18.95	6.0	13.48	2.55	1.40
Regional only	24.78	12.82	54.0	18.05	**10.17**	1.43
Global+Regional	**3.10**	**8.13**	**62.0**	**13.03**	9.13	**2.88**

Table 7: Comparison of global-only, regional-only, and joint task configurations. Metrics report global accuracy, region-mixing fidelity, HC BAG dispersion, and disease-prior BAG.

Model	Valid MAE	Valid RMSE	Pearson ρ	Valid Loss
UNet3D	4.01	6.59	0.94	7.46
ResNet3D	3.61	**5.46**	0.96	**6.51**
3D-ScaleDenseNet	**3.10**	5.71	**0.97**	6.56

Table 8: Comparison of backbone architectures under an identical training setting.

While the 95% confidence intervals of leading methods overlap, MT-ReBA provides the best point-estimate MAE in this controlled benchmark.

G Ablation Results Under the Submitted Protocol

We report only ablations evaluated under the same submitted protocol.

G.1 Ablation of Global and Regional Objectives

We compare three configurations in Table 7—global-only, regional-only, and joint—to evaluate complementary supervision. Global-only yields a Splice-Error of 18.95 years (SW5 = 6.0%). Regional-only improves these to 12.82 years and 54.0%, but degrades global MAE from 3.10 to 24.78 years. Joint training preserves global MAE at 3.10 years while improving Splice-Error to 8.13 years and SW5 to 62.0%. These objectives provide complementary supervision, with joint optimization achieving the best global–regional balance.

G.2 Comparison of Backbone Architectures

We compare our backbone against 3D U-Net and 3D ResNet-18 under identical settings (Table 8).

Our model achieves the lowest validation MAE (3.10 years), outperforming U-Net (4.01 years) and ResNet-18 (3.61 years) by 22.7% and 14.1%, respectively. ResNet-18 achieves better RMSE (5.46 vs. 5.71 years) and lower validation loss, whereas our backbone achieves higher Pearson correlation ($\rho = 0.97$ vs. 0.96). Thus, although ResNet-18 better controls large prediction errors, our backbone obtains the best MAE and correlation. Since MAE is the primary metric and regional decomposition is our main objective, we adopt our backbone.

H Compute Reporting

Training is performed on eight NVIDIA RTX A6000 GPUs with 48 GB memory per GPU, 64 CPU cores, and 200 GB system memory. Each stage runs for 150 epochs. The same atlas grid is supplied to all model replicas, and mixed-image gradients are restricted to the regional pathway.

I Task-Configuration and Backbone Notes

The task-configuration comparison in Table 7 and the backbone comparison in Table 8 confirm that global and regional supervision are strongly complementary and that 3D-ScaleDenseNet achieves the best validation MAE under the controlled setup. These selections support the joint global–regional objective of MT-ReBA.

J Anatomical Addressability of the Regional Decomposition

A weakly supervised regional decomposition can collapse if all heads are trained against the same subject age. MT-ReBA addresses this failure mode by keeping the regional pathway inactive in Stage 1 and assigning the donor-age proxy only to the head corresponding to the changed parcel in Stage 2. Sampling interventions across all $R = 48$ parcels links each output index to its atlas mask, while unchanged heads receive no synthetic age target.

The held-out ROI-occlusion analysis evaluates this correspondence on unspliced HC images. Its pronounced response-matrix diagonal, ROSR, and RTLA results support anatomical addressability: each head responds preferentially to its corresponding parcel even when no donor tissue or splice boundary is present. Input-gradient locality provides a second real-image test. These results argue against an artifact-only explanation. Appendices L and M provide the corresponding controls.

K Expanded Form of the Bias Regularization Loss

With residuals $\varepsilon_i = \widehat{y}_i - y_i$ and mean chronological age $\bar{y}$, the Stage 1 bias term is

$$\mathcal{L}_{\text{bias}} = \left| \frac{1}{N} \sum_{i=1}^{N} \varepsilon_i \, (y_i - \bar{y}) \right|. \tag{20}$$

This is the absolute empirical covariance between residual and age (the residual need not be explicitly centred because $\sum_i (y_i - \bar{y}) = 0$). Minimizing it penalizes the regression-to-the-mean bias common in BAE (de Lange & Cole, 2020; de Lange et al., 2022). Separately, when linear calibration is used for downstream BAG analysis, the BAG–age relation is fitted on validation HCs only and applied unchanged to the test set; test labels are not used for calibration.

L Relation to Input Mixing and Mixing Artifacts

Relation to mixup and CutMix

Region mixing is related to input-mixing methods such as mixup (Zhang et al., 2018) and CutMix (Yun et al., 2019), but differs in its mixing unit and target semantics. The mixup method blends whole images, whereas CutMix replaces an axis-aligned rectangle; both produce a single global mixed label. MT-ReBA instead replaces an anatomically defined atlas parcel and routes donor chronology only to the corresponding regional head. This creates anatomically addressable weak supervision: the donor age is an age-linked proxy.

Artifact controls and interpretation

All images undergo the same registration, brain extraction, bias-field correction, and within-brain normalization. Donor intensities are matched to the recipient, and the parcel mask is smoothly blended. The SSLR analysis in the main paper shows that detectable changes remain inside the selected parcel, ruling out spatial leakage beyond the target ROI.

Evidence from untouched HC images is therefore essential. ROI occlusion tests whether head r responds preferentially to parcel r, and input-gradient locality tests whether the sensitivity of head r is enriched inside the same parcel. Because these inputs contain no transplanted tissue, the results provide evidence against a model driven solely by splice artifacts. Frozen-model AD/PD analyses on original scans provide a further downstream check. These analyses support anatomically aligned sensitivity while remaining distinct from biological regional-age validation.

M Input-gradient Locality on Held-out HCs

The main paper defines the whole-grid gradient fraction, gradient density, and the GWBE, MSGE, and ARGS metrics. The three enrichment metrics compare self-ROI gradient density with whole-brain density, the mean density of three compact volume-matched random ROIs, and the density in

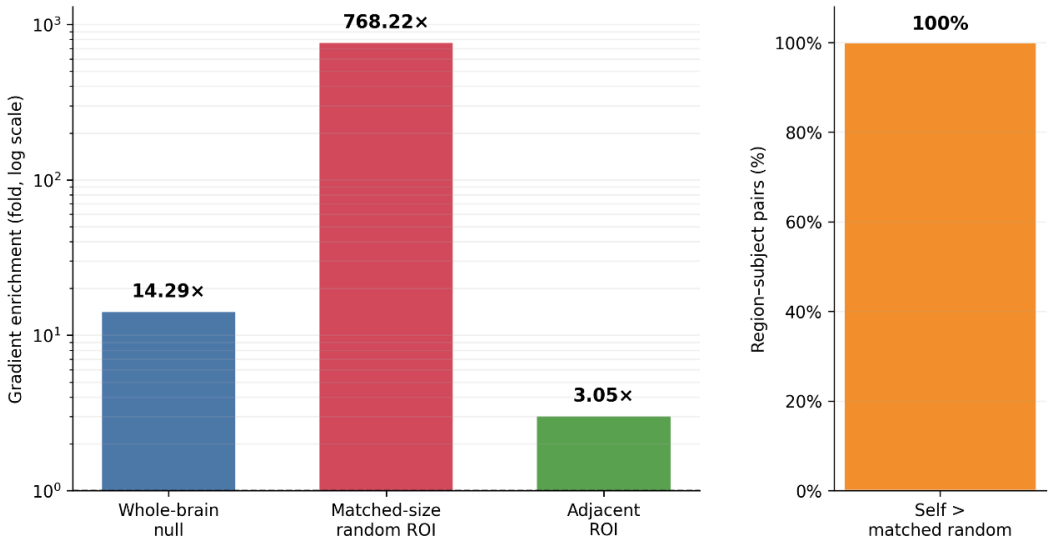


Figure 6: Input-gradient locality controls on held-out HC images. The panels compare self-ROI gradient density with whole-brain, volume-matched random-ROI, and adjacent-ROI density controls.

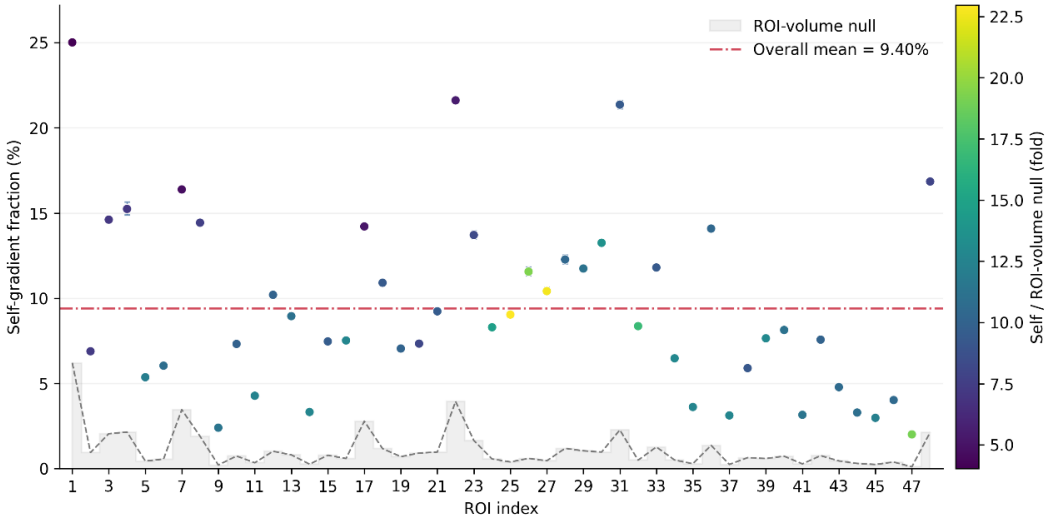


Figure 7: Self-ROI gradient fraction by Harvard–Oxford parcel. The grey curve denotes the ROI-volume null. All 48 parcels exceed this null, with a mean self-ROI gradient fraction of 9.40%.

the union of six-neighbour adjacent atlas parcels, respectively. Figure 6 summarizes these controlled comparisons, while Figure 7 reports parcel-wise self-ROI fractions.

The controlled comparisons demonstrate that regional heads concentrate their input sensitivity on anatomically corresponding parcels rather than following generic volume-driven patterns. In Figure 6, self-ROI gradient density consistently exceeds whole-brain, volume-matched random, and adjacent-ROI controls, with GWBE, MSGE, and ARGS reaching $14.29\times$, $768.22\times$, and $3.05\times$, respectively. Figure 7 further shows that this localization is consistent across all 48 atlas parcels, with every regional head exceeding its ROI-volume null and achieving a mean self-ROI gradient fraction of 9.40%. Together with the perturbation-based analyses, these results indicate that MT-ReBA learns anatomically aligned regional representations rather than merely relying on shared global image features.