Volumetric-Guided Brain Image Synthesis with Diffusion Models

A. Cartaya Lathulerie, *A. Casamitjana, A. Oliver, X. Lladó*

Computer Vision and Robotics Research Group, University of Girona, Girona, Spain

MOTIVATION

Brain aging and neurodegeneration cause heterogeneous, region-specific structural changes, but modelling subject-specific trajectories is limited by scarce longitudinal data and reliance on paired scans or global conditioning.

PROPOSAL

BrainST is a latent diffusion model [1] trained on cross-sectional data that uses volumes from 18 brain regions of interest (ROIs) to enable anatomically controlled T1w MRI synthesis, counterfactual modification, and longitudinal prediction [2,3].

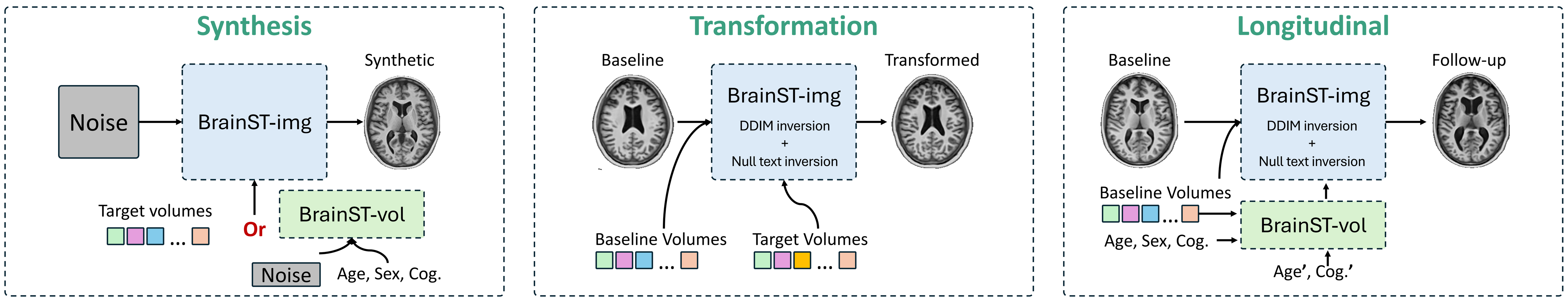

Figure 1. **Framework capabilities.** Left: Synthetic generation from random noise conditioned on target volumes, manually specified or generated from covariates. Centre: Anatomically controlled transformation of existing scans via volumetric manipulation. Right: Longitudinal MRI prediction driven by subject covariates.

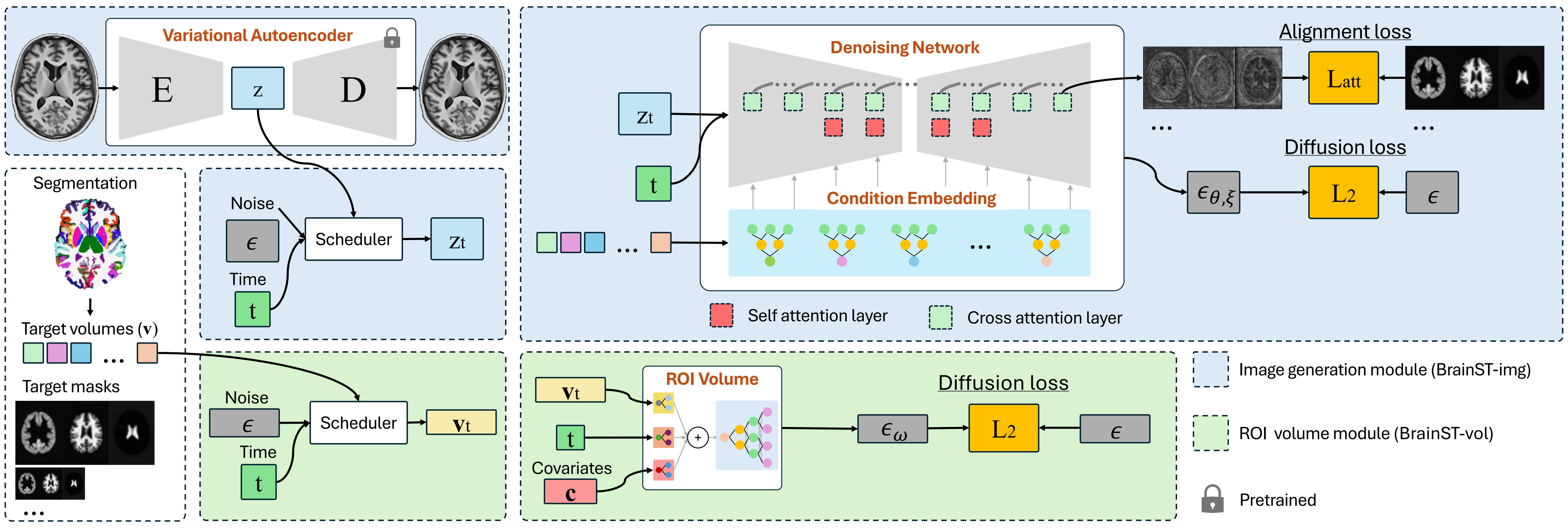

Figure 2. **Architecture and training.** The framework consists of an image generation module (blue) and an ROI volumetric generation module (green). The image module includes a pretrained VAE, a cross-attention denoiser, and an auxiliary alignment loss. The volumetric module models structural volumes conditioned on covariates. Both modules are jointly trained to predict noise and maintain anatomical consistency.

HIGHLIGHTS

- Flexible framework for anatomically controlled T1w MRI generation.
- Fine-grained volumetric conditioning for global and local guidance.
- Structure-aware loss for improved regional control and identity preservation.
- Training on cross-sectional data with application to longitudinal MRI generation in healthy and diseased populations, back-and-forth in time.
- Large-scale and targeted counterfactual transformations.

CONCLUSION

BrainST demonstrates improved structural fidelity and lower volumetric errors compared to prior methods. It also enables localized counterfactual changes while preserving subject identity, supporting its use in neuroimaging research with limited longitudinal data.

RESULTS

Synthesis

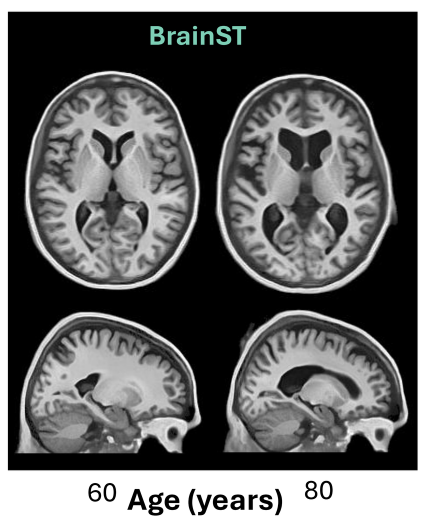

Figure 3. **Synthesis examples.** Left: 60-year-old healthy female. Right: 86-year-old AD female.

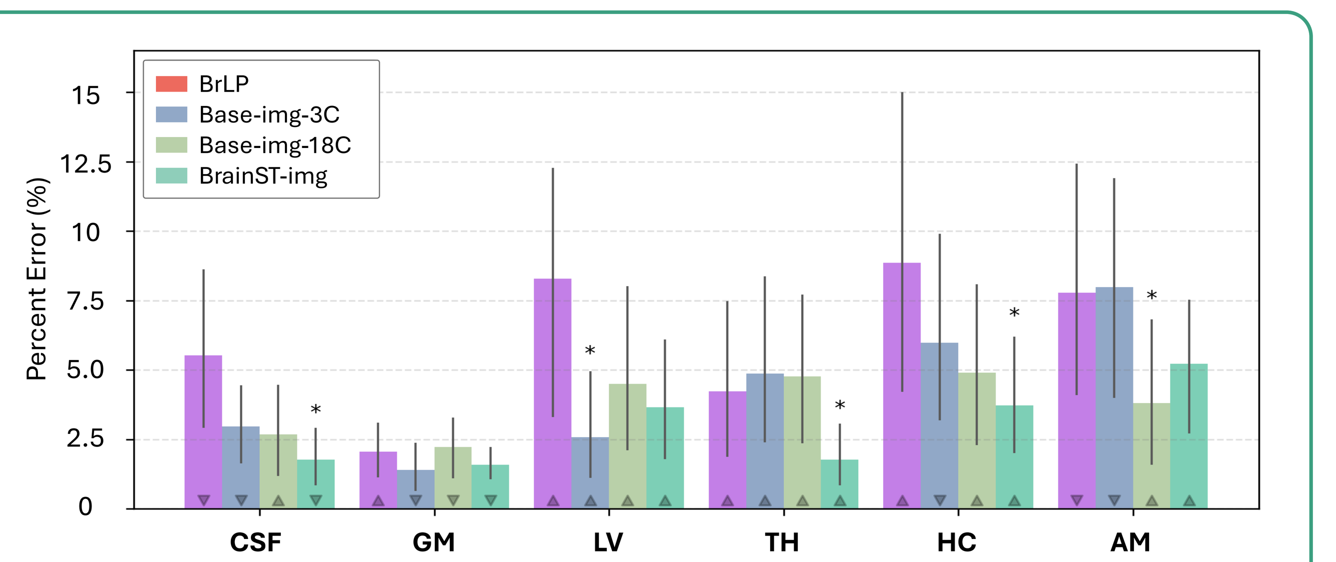

Figure 4. Volumetric synthesis error. Percentage error between synthetic and target volumes included for a representative image generation framework (BISLDM) [4], two ablated variants of our proposal (Base-img-3C, Base-img-18C), and the proposed image generator (BrainST-img). * denotes significant improvement.

Transformation

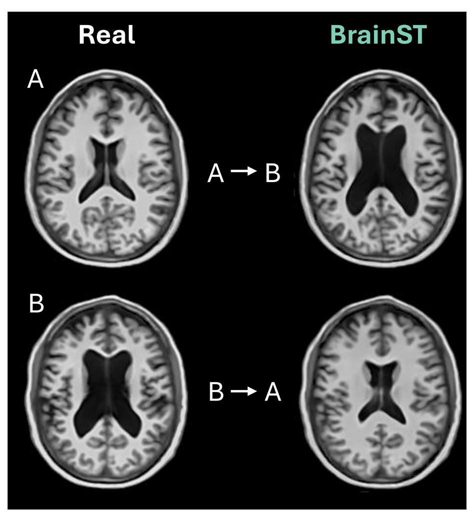

Figure 5. **Transformation examples.** Top: subject A with volumes from B. Bottom: subject B with volumes from A.

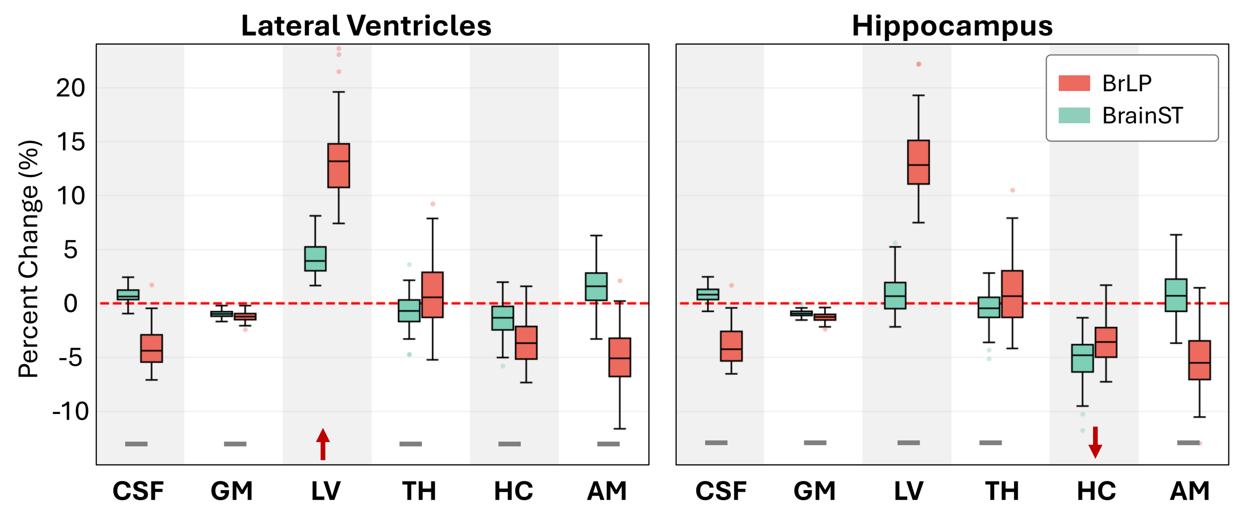

Figure 6. **Localized anatomical transformations.** Signed percentage volume change between transformed and baseline MRIs. Each panel shows a 5% controlled modification of one target structure (↑ = increase, ↓ = decrease and "–" denotes no modification). Results for baseline diffusion framework (BrLP) [5] and BrainST.

Longitudinal evolution

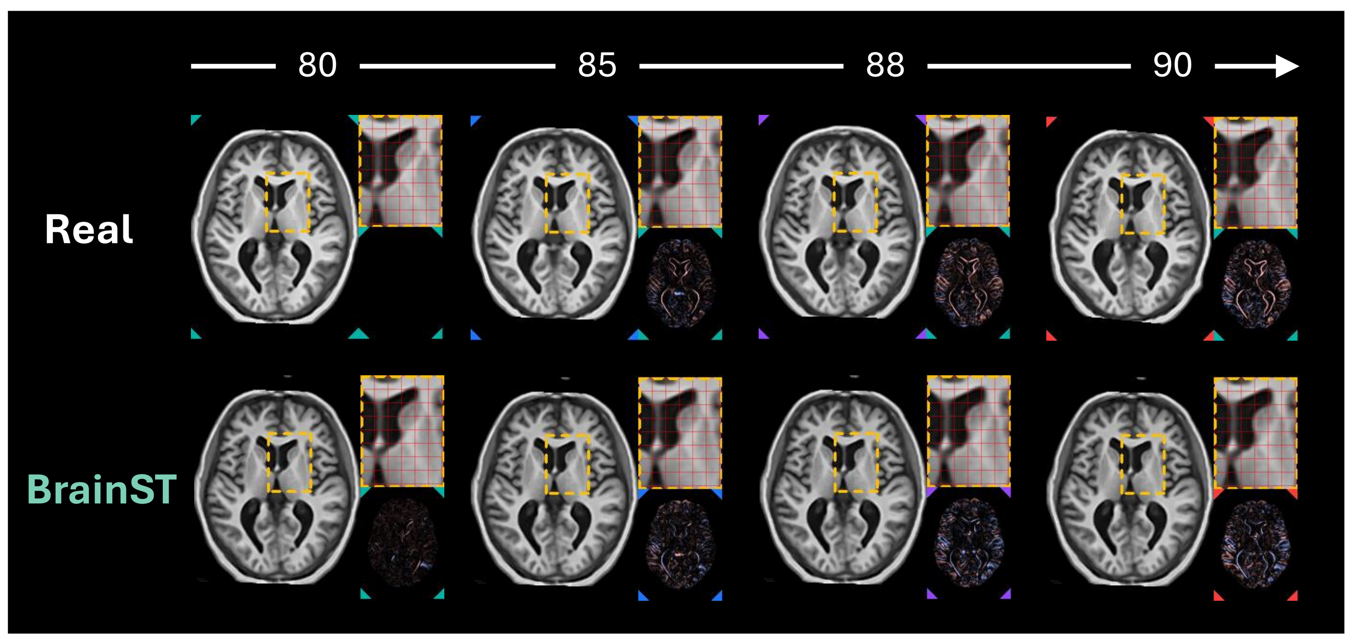

Figure 7. **Longitudinal evolution.** AD subject. Top: real scans; bottom: BrainST predictions. First column: baseline and reconstruction; subsequent columns: follow-ups. Each cell shows axial view, zoom, and difference from ground truth.

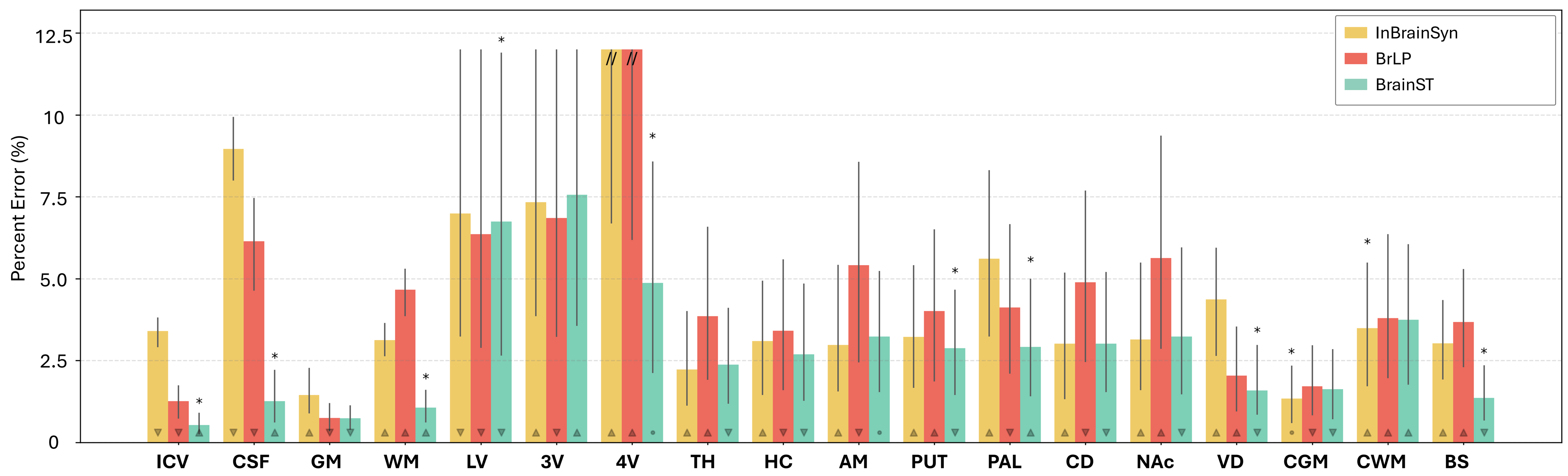

Figure 8. **Volumetric errors in longitudinal prediction.** Percentage error between predicted and expected follow-up volumes for two state of the art methos, BrLP [5] and InBrainSyn [6], and BrainST. Bars indicate median; whiskers, 25th–75th percentiles. * denotes significant improvement. // denotes cropped bars for display (4V InBrainSyn = 12.84%, BrLP = 14.04%).

ACKNOWLEDGEMENT

Agustin Cartaya Lathulerie holds an FPI grant from the Ministerio de Ciencia, Innovación y Universidades (PREP2023-001473) and Adrià Casamitjana a Ramón y Cajal grant (RYC2024-050753-I). This work was also supported by PID2023-146187OB-I00 and the ICREA Academia program.

References and PDF version

Volumetric-Guided Brain Image Synthesis with Diffusion Models

A. Cartaya Lathulerie, *A. Casamitjana, A. Oliver, X. Lladó*

Computer Vision and Robotics Research Group, University of Girona, Girona, Spain

REFERENCES

[1] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, B. Ommer, High-resolutionimage synthesis with latent diffusion models, in: 2022 IEEE/CVF Conferenceon Computer Vision and Pattern Recognition (CVPR), IEEE Computer Society,2022, pp. 10674–10685. doi: https://doi.org/10.1109/CVPR52688.2022.01042.

[2] J. Song, C. Meng, S. Ermon, Denoising diffusion implicit models, in: 9th International Conference on Learning Representations, ICLR 2021, Virtual Event,Austria, May 3-7, 2021, OpenReview.net, 2021. URL: https://openreview.net/forum?id=St1giarCHLP.

[3] R. Mokady, A. Hertz, K. Aberman, Y. Pritch, D. Cohen-Or, Null-text inver-sion for editing real images using guided diffusion models, in: 2023 IEEE/CVFConference on Computer Vision and Pattern Recognition (CVPR), 2023, pp.6038–6047. doi: https://doi.org/10.1109/CVPR52729.2023.00585.

[4] W. H. L. Pinaya, P.-D. Tudosiu, J. Dafflon, P. F. D. Costa, V. Fernandez,P. Nachev, S. Ourselin, M. J. Cardoso, Brain imaging generation with latent dif-fusion models, in: Deep Generative Models, Springer Nature Switzerland, 2022,pp. 117–126. doi: https://doi.org/10.1007/978-3-031-18576-2_12.

[5] L. Puglisi, D. C. Alexander, D. Ravì, Brain latent progression: Individual-based spatiotemporal disease progression on 3d brain mris via latent diffusion,Medical Image Analysis 106 (2025) 103734. doi: https://doi.org/10.1016/j.media.2025.103734.

[6] J. Fu, Y. Zheng, N. Dey, D. Ferreira, R. Moreno, Synthesizing individu-alized aging brains in health and disease with generative models and paralleltransport, Medical Image Analysis 105 (2025) 103669. doi: https://doi.org/10.1016/j.media.2025.103669.