Supplementary Material
Enhancing Spatiotemporal Disease Progression Models via Latent Diffusion and Prior Knowledge

Lemuel Puglisi[1], Daniel C. Alexander[2], and Daniele Ravì[2,3,4]

[1] Dept. of Math and Computer Science, University of Catania, Italy
[2] Centre for Medical Image Computing, University College London, UK
[3] School of Physics, Engineering and CS, University of Hertfordshire, UK
[4] MIFT Department, University of Messina, Italy

Table 1. Pipeline Settings. The training of our pipeline is conducted on a GeForce RTX 4090 (24GB VRAM) and the code is implemented using the MONAI library [2].

	Autoencoder	**UNet**	**ControlNet**
Architecture	[1]	[6]	[3]
Input size	$122 \times 146 \times 122$	$3 \times 16 \times 20 \times 16$	$3 \times 16 \times 20 \times 16$
Initialization	Pretrained model[4]	Random	Random
Optimizer	Adam	AdamW	AdamW
Learning rate	10^{-4}	2.5×10^{-5}	2.5×10^{-5}
Batch size	8	16	16
Diffusion steps	-	$T = 1000$	$T = 1000$
Noise schedule	-	Linear [.0015, .0205]	Linear [.0015, .0205]
Sampling strategy	-	DDIM [5] (25 steps)	DDIM [5] (25 steps)
Training time	$\sim$51 hours	$\sim$12 hours	$\sim$12 hours

References

1. https://github.com/Project-MONAI/GenerativeModels/blob/main/generative/networks/nets/autoencoderkl.py
2. Cardoso, M.J., Li, W., Brown, R., Ma, N., Kerfoot, E., Wang, Y., Murrey, B., Myronenko, A., Zhao, C., Yang, D., et al.: Monai: An open-source framework for deep learning in healthcare. arXiv preprint arXiv:2211.02701 (2022)
3. https://github.com/Project-MONAI/GenerativeModels/blob/main/generative/networks/nets/controlnet.py
4. https://github.com/Project-MONAI/GenerativeModels/tree/main/model-zoo/models/brain_image_synthesis_latent_diffusion_model
5. Song, J., Meng, C., Ermon, S.: Denoising diffusion implicit models. In: International Conference on Learning Representations (2021)
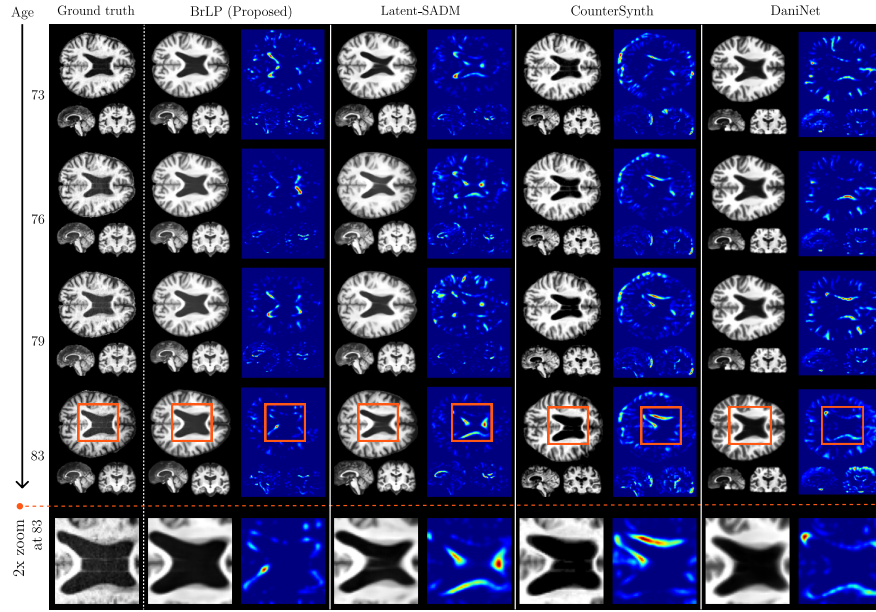6. https://github.com/Project-MONAI/GenerativeModels/blob/main/generative/networks/nets/diffusion_model_unet.py

Fig. 1. Another comparison between the real progression of a 69 y.o. subject with MCI over 14 years and the predictions obtained by BrLP and the baselines. Each method shows a predicted MRI (left) and its deviation from the subject's real brain MRI (right).
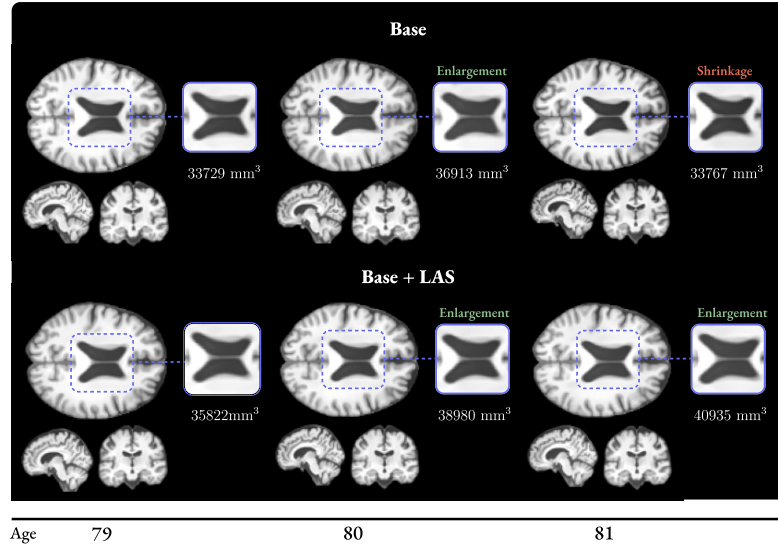

Fig. 2. An example of spatiotemporal consistency improvement achieved with LAS, where we observe an adequate monotonic expansion of the lateral ventricles.