This MICCAI paper is the Open Access version, provided by the MICCAI Society. It is identical to the accepted version, except for the format and this watermark; the final published version is available on SpringerLink.

Glioma Inpainting in Brain MRI via 3D Region Aware Diffusion

Zahra Karimaghaloo[1], Dumitru Fetco[1,3], Haz-Edine Assemlal[1,3], Hassan Rivaz[2], and Douglas L. Arnold[1,3]

[1] NeuroRx, A Clario Company, part of Thermo Fisher Scientific, Montreal, Canada
[2] Concordia University, Montreal, Canada
[3] Department of Neurology and Neurosurgery, McGill University, Montreal, Canada
zahrak@clario.com

Abstract. Automated neuroimaging pipelines often fail in the presence of gliomas, as their mass effect and abnormal signal intensities violate the healthy tissue assumptions required for reliable registration, segmentation, and volumetric analysis. Synthesizing healthy-appearing tissue over pathological regions restores the validity of these downstream tools. While deep generative models have shown promise for local tissue synthesis, most existing approaches rely on 2D slice-wise architectures that lack volumetric context and therefore introduce conspicuous inter-slice discontinuities in non-axial reformats. We present P3D-RAD, a pseudo-3D inpainting framework based on Denoising Diffusion Probabilistic Models (DDPMs). The model uses (2+1)D convolutions to capture volumetric continuity while retaining the computational efficiency of 2D generative kernels, and adapts a Region Aware Diffusion (RAD) to the medical domain by using spatially varying noise schedules and restricting the reverse process at inference to the masked region. This leaves the surrounding unmasked anatomy untouched and shortens inference. On the validation set of the ASNR-MICCAI BraTS Local Synthesis of Healthy Brain Tissue via Inpainting Challenge ($N = 219$), P3D-RAD (Model B) attains a Structural Similarity Index Measure (SSIM) of 0.832 ± 0.110, a Peak Signal-to-Noise Ratio (PSNR) of 24.45 ± 5.37 dB, and a Mean Squared Error (MSE) of 0.006 ± 0.005.

Keywords: Denoising Diffusion Probabilistic Models · Inpainting · Local Synthesis · Glioma · Pseudo-3D · Region Aware Diffusion · BraTS · MRI

1 Introduction

Structural brain MRI remains essential for the clinical assessment of gliomas, from surgical planning to treatment-response evaluation. The tumor itself, the surrounding edema, and any resection cavity, however, degrade the performance of automated neuroimaging pipelines: registration algorithms are confounded by intensity patterns that have no anatomical counterpart in reference atlases, and tools trained on healthy anatomy fail near pathology. Image inpainting, also

known as local tissue synthesis, mitigates these failures by replacing pathological regions with plausible healthy appearing tissue, restoring an image on which conventional tools operate as intended. Recognizing the importance of this task, the Brain Tumor Segmentation (BraTS) challenge [1,19,2,15] established a dedicated local synthesis sub-challenge. Recent results on this benchmark [23,3,6,17] show that generative approaches, and Denoising Diffusion Probabilistic Models (DDPMs) in particular, yield state-of-the-art synthetic fidelity.

Recent methods have addressed different aspects of the trade-off between anatomical fidelity, volumetric consistency, and computational efficiency. In the BraTS 2025 Inpainting Challenge, Zhang et al. achieved first place using a 3D U-Net with random masking augmentation [23], while Ferreira et al. (also ranked first) demonstrated that conditional diffusion sampling could be accelerated by more than an order of magnitude while maintaining competitive quantitative metrics [9]. Other recent approaches include FastSurfer-LIT, which employs a resolution-independent, multi-view DDPM and validates the resulting images through downstream whole brain segmentation and cortical surface reconstruction [21] and fastWDM3D, which performs wavelet-domain 3D inpainting using only a few diffusion steps [7]. Volumetric consistency has also been improved through hierarchical axial-coronal diffusion [18].

Building on our recent longitudinal RAD formulation for lesion inpainting in multiple sclerosis [12,13], which conditions on two timepoints, we adopt the single-timepoint setup of [8] and condition the U-Net on a voided T_1-weighted image together with its binary inpainting mask. To reduce the iterative sampling cost of conventional diffusion, we adapt the Region Aware Diffusion (RAD) mechanism [14] to the medical domain: a spatially varying noise schedule restricts the reverse trajectory strictly to the target region. This accelerates inference while leaving unmasked healthy tissue completely untouched, eliminating the need for post-hoc boundary blending. We evaluate the resulting framework on the BraTS 2026 benchmark, including direct comparison with leading BraTS 2025 approaches. Code is available at `https://github.com/zahra-aghaloo/3D-RAD-Inpainting`.

2 Methods

DDPMs [11] define a generative process by gradually corrupting data with Gaussian noise and learning to reverse it. The forward process is a fixed Markov chain that adds noise over T steps according to a variance schedule $\{\beta_t\}_{t=1}^{T}$,

$$q(\mathbf{x}_t \mid \mathbf{x}_{t-1}) = \mathcal{N}\Big(\mathbf{x}_t; \sqrt{1-\beta_t}\,\mathbf{x}_{t-1},\, \beta_t \mathbf{I}\Big), \tag{1}$$

which admits a closed-form marginal that allows sampling $\mathbf{x}_t$ directly from the clean sample $\mathbf{x}_0$,

$$\mathbf{x}_t = \sqrt{\bar{\alpha}_t}\,\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_t}\,\boldsymbol{\epsilon}, \qquad \boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}), \tag{2}$$

where $\alpha_t = 1 - \beta_t$ and $\bar{\alpha}_t = \prod_{s=1}^{t} \alpha_s$. A network is trained to approximate the reverse posterior $p_\theta(\mathbf{x}_{t-1} \mid \mathbf{x}_t)$, with the standard $\boldsymbol{\epsilon}$-prediction objective reducing

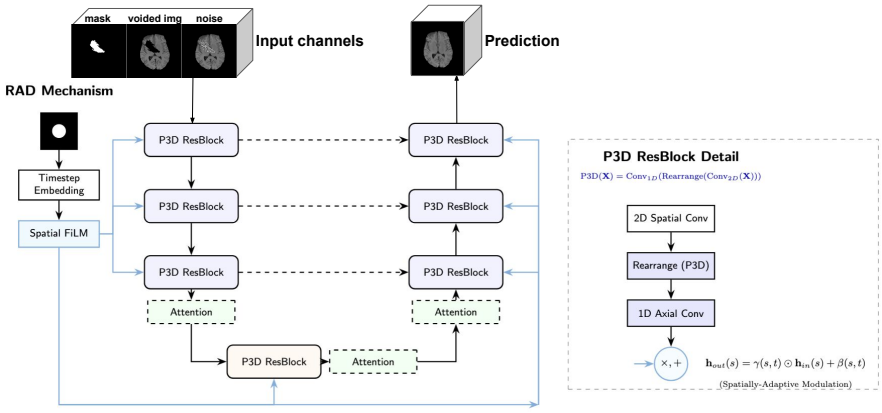

Fig. 1. P3D-RAD architecture. The model integrates pseudo-3D convolutions with a Region Aware Diffusion (RAD) mechanism. Spatial FiLM parameters are injected into both the encoder and decoder P3D ResBlocks (blue lines) to provide spatially aware feature modulation.

to $\mathbb{E}_{\mathbf{x}_0,\boldsymbol{\epsilon},t}\left[\|\boldsymbol{\epsilon}-\boldsymbol{\epsilon}_\theta(\mathbf{x}_t,t)\|^2\right]$. An alternative and often more stable parameterization [22] predicts the *velocity* $\mathbf{v}$,

$$\mathbf{v}_t \equiv \sqrt{\bar{\alpha}_t}\,\boldsymbol{\epsilon} - \sqrt{1-\bar{\alpha}_t}\,\mathbf{x}_0, \tag{3}$$

so that the network $\mathbf{v}_\theta$ is optimized with

$$\mathcal{L}_{\mathbf{v}} = \mathbb{E}_{\mathbf{x}_0,\,\boldsymbol{\epsilon},\,t}\left[\left\|\mathbf{v}_t - \mathbf{v}_\theta(\mathbf{x}_t,t)\right\|_2^2\right]. \tag{4}$$

The $\mathbf{v}$-parameterization interpolates smoothly between predicting noise and predicting the clean signal, which improves numerical conditioning at high noise levels and enables fast sampling.

Building on this formulation, we introduce a pseudo-3D conditional DDPM that exploits inter-slice context for glioma inpainting, and extend it into P3D-RAD by incorporating a Region Aware Diffusion mechanism. The input tensor $\mathbf{X} \in \mathbb{R}^{3\times D\times H\times W}$ concatenates three channels: the binary inpainting mask, the voided image (mask zero-filled), and the noisy image component (Fig. 1).

2.1 Pseudo-3D Architecture

Following the pseudo-3D design of [8,24], our model captures inter-slice context through 1D axial convolutions, a computationally efficient alternative to a fully 3D U-Net. A stack of adjacent axial slices is fed to the network, and each residual

block applies a (2+1)D convolution,

$$\mathrm{P3D}(\mathbf{h}) = \mathrm{Conv}_{1\mathrm{D}}(\mathcal{R}(\mathrm{Conv}_{2\mathrm{D}}(\mathbf{h}))), \tag{5}$$

where $\mathcal{R}$ denotes the dimension rearrangement required for axial processing. This factorization lets the model recognize 3D anatomical structures with the efficiency of 2D generative kernels, effectively resolving "staircase" artifacts in non-axial reformats.

2.2 Region Aware Diffusion (RAD)

We extend RAD [14] to the pseudo-3D setting. RAD uses a two-phase noise schedule over the $T = 1000$ training timesteps: in Phase 1 ($t \in [1, T/2]$) noise is added inside the inpainting mask, and in Phase 2 ($t \in [T/2 + 1, T]$) it is added to the background. The model therefore learns separate noise schedules for the target and surrounding regions, which is intended to improve fidelity at the mask boundary. At inference, we bypass Phase 2 and initialize the reverse process at $t = T/2$, so the healthy background is never perturbed and no boundary blending is required.

This regional conditioning is injected into the P3D ResBlock via Spatial FiLM [20] modulation (Fig. 1). After the 1D axial convolution, the spatial timestep map $t(s)$, which assigns a timestep to each spatial location s based on whether it lies inside or outside the inpainting mask, modulates the feature maps as

$$\mathbf{h}_{\mathrm{out}}(s) = \gamma(t(s)) \odot \mathbf{h}_{\mathrm{in}}(s) + \beta(t(s)), \tag{6}$$

where γ and β are scale and shift parameters learned from the local noise level. This makes the (2+1)D kernels aware of the inpainting boundary at every denoising step and focuses generative capacity on the intended region.

3 Experiments

3.1 Dataset

We use the BraTS 2026 local inpainting dataset (Synapse ID: `syn74274097`), a retrospective, multi-institutional collection of brain tumor MR scans acquired under standard clinical conditions with different scanners and protocols. For each training case, the organizers provide the T_1 scan of the tumor-affected brain, a tumor mask, a healthy tissue mask, a combined mask, and a "T_1-voided" scan in which the combined mask are zeroed out. A separate validation set of 219 subjects is also provided, for which only the voided image and the inpainting mask are released; this set is scored by the organizers and is the basis for the results in Section 3.5.

To train the model to synthesize plausible healthy tissue, we defined the candidate target region as all T_1 brain voxels outside the provided tumor mask.

Because the challenge's healthy masks are often small and cover a limited portion of the brain, we generated synthetic training masks by applying rigid transformations (rotation, flipping, and translation) to the tumor mask geometry, enforcing zero overlap with the actual tumor mask. This increases variability in mask shape, size, and location, so that the network learns healthy anatomical priors across a wide range of mask geometries.

Of the 1,251 training cases, we use 1,155 for training and 96 randomly selected cases as an internal validation set, used to monitor training and tune sampling hyperparameters; ground truth is available for these 96 cases, which is what makes the analyses in Sections. 3.2–3.4 possible.

3.2 Selecting the Number of Reverse Steps and Inference Runs

To select the number of reverse diffusion steps and stochastic inference runs, we performed a grid search on the internal validation set of 96 volumes, varying the number of reverse steps over $\{75, 100, 125, 150, 175\}$ and the number of averaged runs $N_{\text{avg}} \in \{1, 2, 3, 4\}$, where each run uses an independent random seed and the final prediction is the average of the N_{avg} reconstructions. Quality was assessed against the ground truth within the inpainted region using SSIM, PSNR, and MSE.

Two trends emerge (Fig. 3). First, for a fixed number of steps, increasing N_{avg} monotonically improves all three metrics, as averaging independent samples suppresses the high-frequency noise introduced by sampling; the gain is largest from $N_{\text{avg}} = 1$ to 2 and diminishes thereafter. Second, quality is non-monotonic in the number of reverse steps: it peaks at 100 and degrades beyond, with a pronounced drop at 175 steps (SSIM 0.712 ± 0.127 at $N_{\text{avg}} = 1$), indicating that excessive denoising steps accumulate sampling error rather than refining the reconstruction. The best configuration across all three metrics is 100 reverse steps with $N_{\text{avg}} = 4$, reaching SSIM 0.856 ± 0.090, PSNR 30.01 ± 5.63 dB, and MSE $(1.74 \pm 1.56) \times 10^{-3}$. The trade-off is visible qualitatively in Fig. 2: averaging improves the metrics but yields a smoother reconstruction.

3.3 Anatomical Precision

To assess how closely the generated tissue resembles the known healthy anatomy of the ground truth, we ran SynthSeg [4] on both the original T_1 image and the inpainted image, and computed the mean Dice score within the inpainting mask. Mean Dice increases monotonically with the number of averaged runs (Table 1). Two qualitative examples are shown in Fig. 4, where averaging over four runs yields better agreement with the ground-truth segmentation.

3.4 Robustness to Mask Size

Glioma inpainting masks vary considerably in size and shape across the BraTS cohort. To characterize robustness, we stratify the 96 internal validation cases

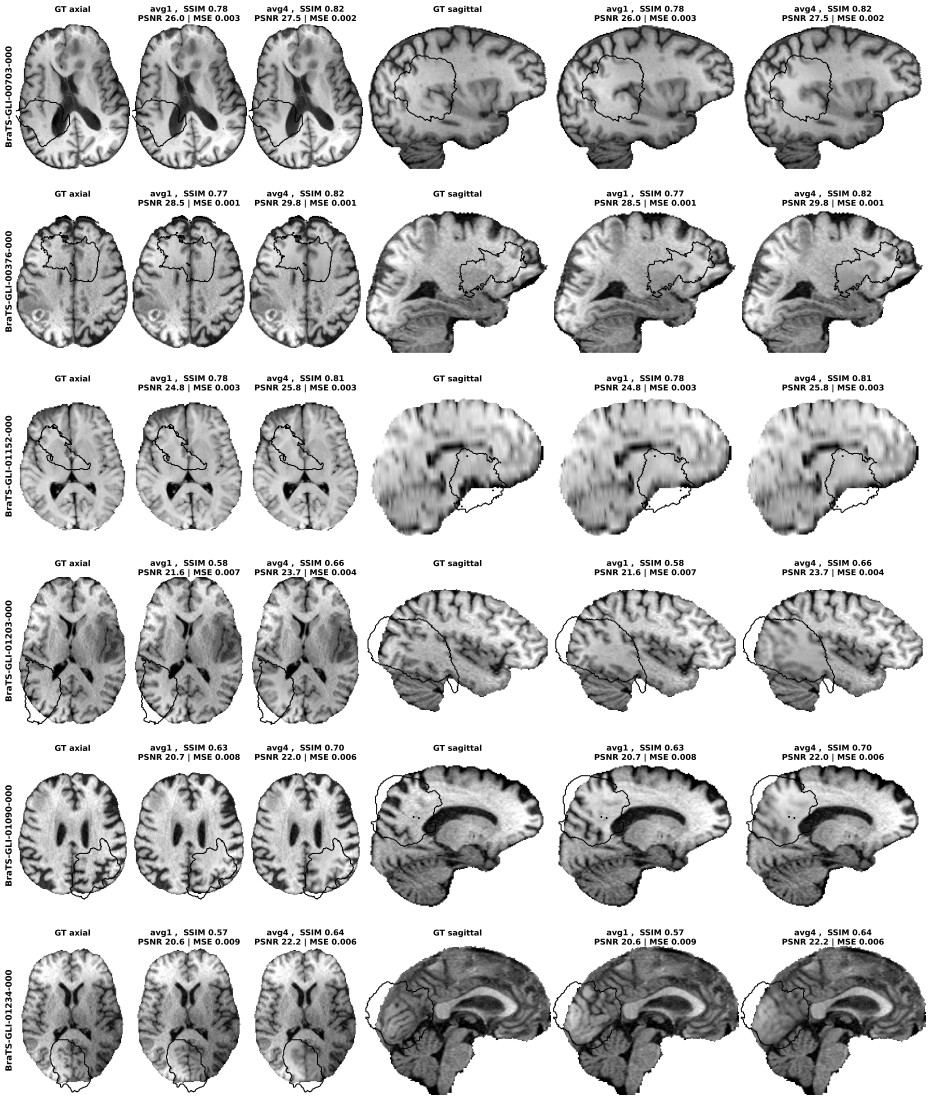

Fig. 2. Visual comparison of axial and sagittal inpainting outputs: ground truth, $N_{\text{avg}} = 1$, and $N_{\text{avg}} = 4$. Validation metrics improve from a single sample to the average of four, at the cost of a smoother appearance.

into mask-volume quartiles (Table 2). Performance degrades as the synthesis region grows and the surrounding context becomes less informative, with SSIM falling from 0.950 in the smallest quartile to 0.753 in the largest.

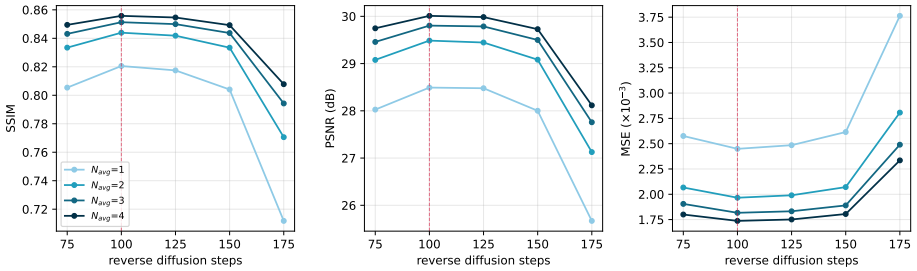

Fig. 3. Effect of the number of reverse diffusion steps and the number of averaged inference runs N_{avg} on inpainting quality, evaluated on 96 internal validation volumes. Each panel reports one metric (SSIM, PSNR, MSE) within the inpainted region as a function of the number of reverse steps, with one curve per N_{avg}. Quality peaks at 100 reverse steps (dashed line) and degrades for larger step counts, while increasing N_{avg} consistently improves all three metrics.

Table 1. Mean Dice between SynthSeg labels of the inpainted and original images, within the inpainting mask, as a function of the number of averaged runs.

Metric	$N_{\text{avg}} = 1$	$N_{\text{avg}} = 2$	$N_{\text{avg}} = 3$	$N_{\text{avg}} = 4$
Mean Dice ± std	0.845 ± 0.062	0.853 ± 0.062	0.857 ± 0.058	0.860 ± 0.058

3.5 Challenge Results and Perception–Distortion Trade-off

Table 3 summarizes the leaderboard performance of the methods considered in our comparison on the official BraTS validation set. In addition to our two P3D-RAD variants, we include three strong methods from the BraTS 2025 challenge: the two submissions that shared first place in the 2025 inpainting task, by Zhang et al. [23] and Ferreira et al. [9], and fastWDM3D [7], selected for its markedly fast inference time (≈ 2 seconds per volume).

For the qualitative comparison in Figure 5, predictions for these methods were obtained from their publicly available implementations or through the BraTS orchestrator [16]. The leaderboard scores in Table 3 are taken directly from the corresponding publications.

Model A is our primary P3D-RAD configuration. Model B uses the same architecture and training framework, but is initialized from the trained Model A checkpoint and further fine-tuned with more aggressive masking augmentation. As shown in Table 3, Model B achieves stronger leaderboard metrics than Model A.

An important observation from these experiments is that ranking according to SSIM, PSNR, and MSE does not always correspond to the perceived quality of the reconstruction over the complete inpainting region. This is particularly relevant to the BraTS evaluation protocol because the leaderboard metrics appear to be computed over the healthy-tissue portion of the mask, which is generally smaller than the full region that the model is asked to synthesize. A method may therefore achieve favorable ranking metrics on this evaluated region, while the

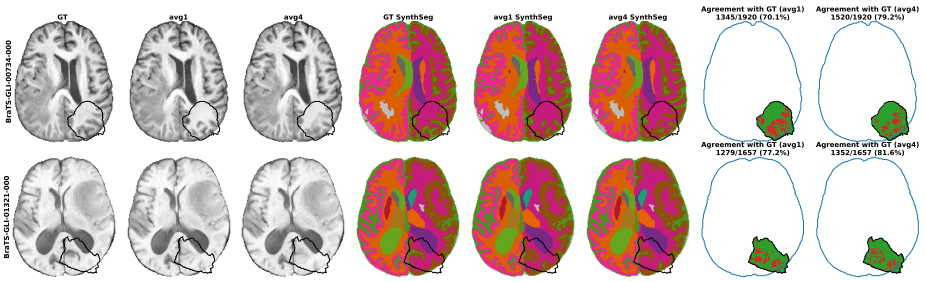

Fig. 4. Inpainting results and the corresponding tissue segmentations for two cases, shown for $N_{\text{avg}} = 1$ and $N_{\text{avg}} = 4$. Agreement with the ground truth (green in the last two columns) increases when four inpainting results are averaged.

Table 2. Inpainting performance on the internal validation set, stratified by mask size quartiles ($n = 24$ per bin).

Mask Size (voxels)	n	**SSIM** ↑	**PSNR (dB)** ↑	**MSE** ($\times 10^{-4}$) ↓
[824–5,650] (Small)	24	0.950 ± 0.039	35.91 ± 6.19	5.1 ± 5.0
[5,650–19,203]	24	0.863 ± 0.068	29.12 ± 4.14	17.5 ± 13.8
[19,203–53,049]	24	0.858 ± 0.063	29.61 ± 3.30	14.3 ± 10.4
[53,049–183,841] (Large)	24	0.753 ± 0.056	25.40 ± 2.23	32.6 ± 16.4

quality of its synthesis over the remaining, and typically much larger, portion of the mask is not directly constrained by the ranking.

This behavior is apparent in Figure 5. Model B produces qualitative results broadly comparable to those of the high-ranking BraTS 2025 methods, consistent with its strong leaderboard performance. Model A, however, more consistently produces plausible reconstructions throughout the full masked region, particularly within the larger region corresponding to the removed tumor. The figure additionally includes a single run of Model A, which is omitted from Table 3 because leaderboard metrics are not available for this configuration. The single-run result tends to preserve sharper anatomical detail, whereas averaging four runs yields somewhat smoother, and occasionally blurrier, reconstructions.

These results therefore highlight a perception–distortion trade-off [5] in the challenge evaluation: stronger distortion-based metrics do not necessarily identify the model that best reconstructs all portions of the missing anatomy. Because reconstruction of the complete mask, and particularly the tumor region, is our primary objective, we selected P3D-RAD Model A, averaged over four runs, as our final submission for evaluation on the challenge test set by the organizers.

3.6 Implementation Details

For preprocessing, voxel intensities are normalized to $[-1, 1]$. Initial 240×240 axial slices are center-cropped to 192×192. At evaluation time, predictions are

Table 3. Leaderboard performance on the BraTS validation set. Values reported for [23,9,7] are directly from their papers.

Method	SSIM ↑	PSNR (dB) ↑	MSE ↓
Zhang et al. [23]	0.87 ± 0.09	24.99 ± 4.69	0.005 ± 0.004
Ferreira et al. [9]	0.87 ± 0.09	24.53 ± 4.71	0.005 ± 0.004
fastWDM3D [7]	0.86 ± 0.12	22.26 ± 3.97	0.008 ± 0.006
P3D-RAD Model B			
4-run average	0.83 ± 0.11	24.45 ± 5.37	0.006 ± 0.005
P3D-RAD Model A			
4-run average	0.80 ± 0.09	21.18 ± 3.96	0.011 ± 0.010

mapped back to the original intensity scale and spatial dimensions of the raw scan.

The model was trained for 50 epochs on an NVIDIA A100 multi-GPU cluster using the Adam optimizer with a learning rate of 10^{-4}. To capture volumetric context during training, we sampled a window of up to 64 consecutive masked axial slices (using all available slices when a case contained fewer than 64). During inference, the model evaluates the entire 155-slice stack ($192 \times 192 \times 155$) in a single pass on a single GPU. On our hardware, inpainting a single volume with 100 reverse diffusion steps takes approximately 2.34 minutes, requiring a peak reserved memory of 27.30 GiB (FP32, batch size 1, on an 80 GB NVIDIA A100 after a warm-up pass).

We used the velocity ($\mathbf{v}$-prediction) objective with Min-SNR weighting ($\gamma = 5$) to balance training stability across timesteps [10], together with a cosine noise schedule, which provides a smoother variance ramp than the linear schedule and avoids signal over-saturation at high timesteps. The variance was kept fixed-small throughout, and Spatial FiLM conditioning injected the mask as a spatial prior into the U-Net. Although training used $T = 1000$ steps, 100 reverse steps with $N_{\text{avg}} = 4$ were sufficient at inference, as established in Sect. 3.2.

4 Conclusions

We presented P3D-RAD, an efficient brain MRI inpainting framework that couples (2+1)D convolutions with a Region Aware Diffusion mechanism. The 1D axial convolutions provide cross-slice anatomical continuity without the memory and compute cost of a full 3D U-Net, while RAD confines the reverse process to the masked region, leaving healthy tissue untouched. On an internal cohort of 96 held-out training volumes, 100 reverse steps combined with an ensemble of four stochastic samples gave the best synthesis quality, and this ensembling also improved downstream tissue segmentation agreement. The best quantitative metrics we achieved on the official challenge validation set (Model B) were SSIM of 0.832 ± 0.110, a PSNR of 24.45 ± 5.37 dB, and an MSE of 0.006 ± 0.005.

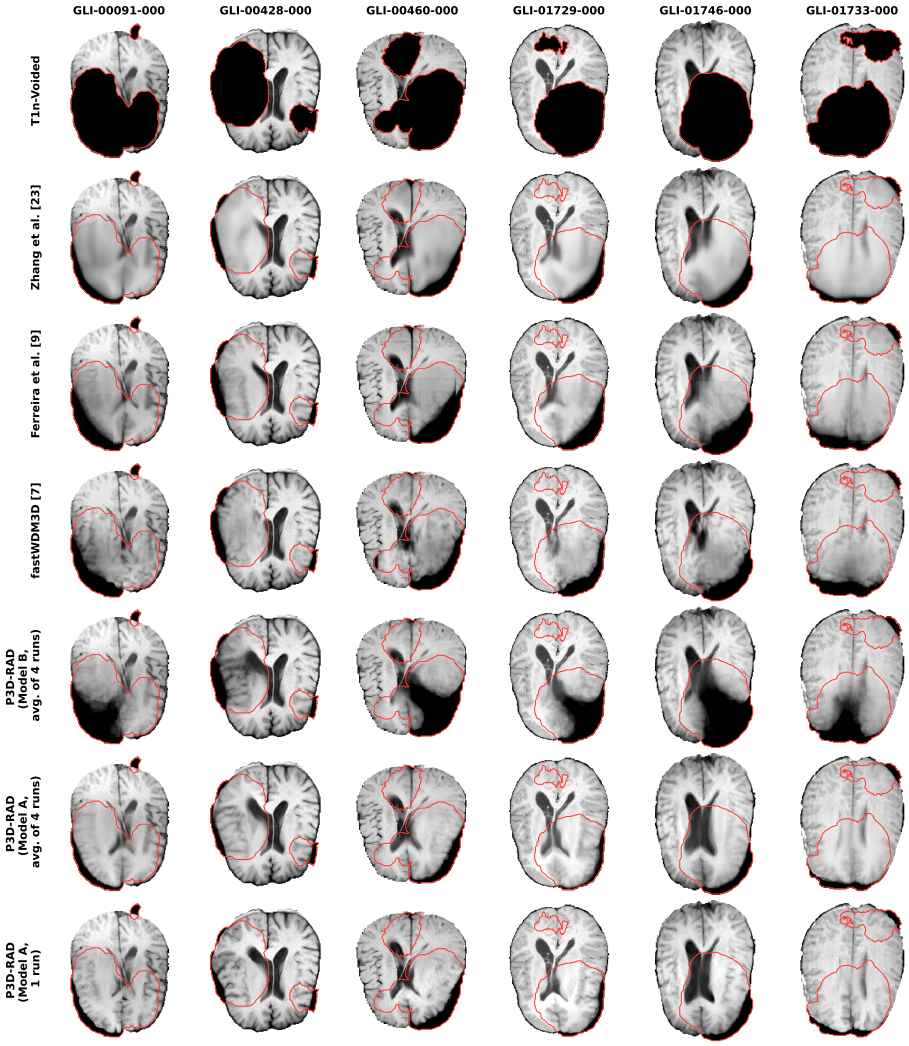

Fig. 5. Row 1: input T1n volumes with the region to be inpainted removed. Rows 2–4: Zhang et al. [23], Ferreira et al. [9], and fastWDM3D [7]. Rows 5–7: P3D-RAD Model B (4-run average), Model A (4-run average), and a single run of Model A. Corresponding leaderboard scores appear in Table 3. Model A reconstructs the full masked region more consistently than its leaderboard ranking suggests, particularly the larger region corresponding to the removed tumor.

References

1. Baid, U., Ghodasara, S., Mohan, S., Bilello, M., Calabrese, E., Colak, E., Farahani, K., Kalpathy-Cramer, J., Kitamura, F.C., Pati, S., et al.: The rsna-asnr-miccai

brats 2021 benchmark on brain tumor segmentation and radiogenomic classification. arXiv preprint arXiv:2107.02314 (2021)
2. Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J., Freymann, J., Farahani, K., Davatzikos, C.: Advancing the cancer genome atlas glioma mri collections with expert segmentation labels and radiomic features. sci. data 4, 170117 (2017) (2017)
3. Biller, V., Zimmer, L., Erdur, C., Nagar, S., Rückert, D., Bubeck, N., Weidner, J.: A biophysically-conditioned generative framework for 3d brain tumor mri synthesis. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 91–101. Springer (2025)
4. Billot, B., Greve, D.N., Puonti, O., Thielscher, A., Van Leemput, K., Fischl, B., Dalca, A.V., Iglesias, J.E., et al.: Synthseg: Segmentation of brain mri scans of any contrast and resolution without retraining. Medical image analysis **86**, 102789 (2023)
5. Blau, Y., Michaeli, T.: The perception-distortion tradeoff. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 6228–6237. IEEE (2018)
6. Dai, Y., Bi, Y., Navab, N., Jiang, Z.: Context-aware healthy brain inpainting: A multi-stage ddim approach for the brats 2025 challenge. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 158–167. Springer (2025)
7. Durrer, A., Bieder, F., Friedrich, P., Menze, B., Cattin, P.C., Kofler, F.: fastWDM3D: Fast and accurate 3D healthy tissue inpainting. arXiv preprint arXiv:2507.13146 (2025)
8. Durrer, A., Wolleb, J., Bieder, F., Friedrich, P., Melie-Garcia, L., Ocampo Pineda, M.A., Bercea, C.I., Hamamci, I.E., Wiestler, B., Piraud, M., et al.: Denoising diffusion models for 3d healthy brain tissue inpainting. In: MICCAI Workshop on Deep Generative Models. pp. 87–97. Springer (2024)
9. Ferreira, A., Luijten, G., Hinrichs-Puladi, B., Kleesiek, J., Alves, V., Egger, J.: Achieving over 10× faster sample generation with conditional denoising diffusion. In: Segmentation, Classification, and Synthesis for Brain Tumors and Traumatic Brain Injuries. pp. 123–132. Springer (2026). https://doi.org/10.1007/978-3-032-16370-7_11
10. Hang, T., Gu, S., Li, C., Bao, J., Chen, D., Hu, H., Geng, X., Guo, B.: Efficient diffusion training via min-snr weighting strategy. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 7441–7451 (2023)
11. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. Advances in neural information processing systems **33**, 6840–6851 (2020)
12. Karimaghaloo, Z., Assemlal, H.E., Rivaz, H., Arnold, D.L.: Topology-aware ddpm inpainting of evolving ms lesions in paired brain mri. In: Medical Imaging 2026: Image Processing. vol. 13925, pp. 403–407. SPIE (2026)
13. Karimaghaloo, Z., Fetco, D., Assemlal, H.E., Rivaz, H., Arnold, D.L.: Longitudinal lesion inpainting in brain MRI via 3D region aware diffusion. arXiv preprint arXiv:2603.05693 (2026)
14. Kim, S., Suh, S., Lee, M.: Rad: Region-aware diffusion models for image inpainting. In: Proceedings of the Computer Vision and Pattern Recognition Conference. pp. 2439–2448 (2025)
15. Kofler, F., Meissen, F., Steinbauer, F., Graf, R., Ehrlich, S.K., Reinke, A., Oswald, E., Waldmannstetter, D., Hoelzl, F., Horvath, I., et al.: The brain tumor segmentation (brats) challenge: Local synthesis of healthy brain tissue via inpainting. arXiv preprint arXiv:2305.08992 (2023)

16. Kofler, F., Rosier, M., Astaraki, M., Baid, U., Möller, H., Buchner, J.A., Steinbauer, F., Oswald, E., de la Rosa, E., Ezhov, I., et al.: Brats orchestrator: democratizing and disseminating state-of-the-art brain tumor image analysis. arXiv preprint arXiv:2506.13807 (2025)
17. Kulkarni, N., Iyer, K., Tapp, A., Parida, A., Capellán-Martín, D., Jiang, Z., Ledesma-Carbayo, M.J., Anwar, S.M., Linguraru, M.G.: Post-processing methods for improving accuracy in mri inpainting. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 146–157. Springer (2025)
18. Kwark, D.H., Luo, S., Zhu, X., Li, Y., Liang, Z.P., Kindratenko, V.: Hierarchical diffusion framework for pseudo-healthy brain MRI inpainting with enhanced 3d consistency. In: MICCAI Workshop on Deep Generative Models. pp. 78–88. Springer (2025)
19. Menze, B.H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., Burren, Y., Porz, N., Slotboom, J., Wiest, R., et al.: The multimodal brain tumor image segmentation benchmark (brats). IEEE transactions on medical imaging **34**(10), 1993–2024 (2014)
20. Perez, E., Strub, F., De Vries, H., Dumoulin, V., Courville, A.: Film: Visual reasoning with a general conditioning layer. In: Proceedings of the AAAI conference on artificial intelligence. vol. 32 (2018)
21. Pollak, C., Kügler, D., Bauer, T., Rüber, T., Reuter, M.: FastSurfer-LIT: Lesion inpainting tool for whole-brain MRI segmentation with tumors, cavities, and abnormalities. Imaging Neuroscience **3** (2025). https://doi.org/10.1162/imag_a_00446
22. Salimans, T., Ho, J.: Progressive distillation for fast sampling of diffusion models. arXiv preprint arXiv:2202.00512 (2022)
23. Zhang, J., Weng, Y., Chen, K.: Robust 3d brain mri inpainting with random masking augmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 102–109. Springer (2025)
24. Zhu, L., Xue, Z., Jin, Z., Liu, X., He, J., Liu, Z., Yu, L.: Make-a-volume: Leveraging latent diffusion models for cross-modality 3d brain mri synthesis. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 592–601. Springer (2023)