



# Post-processing Methods for Improving Accuracy in MRI Inpainting

Nishad Kulkarni<sup>1</sup>, Krithika Iyer<sup>1</sup>, Austin Tapp<sup>1</sup>, Abhijeet Parida<sup>1,2</sup>, Daniel Capellán-Martín<sup>1,2</sup>, Zhifan Jiang<sup>1</sup>, María J. Ledesma-Carbayo<sup>2</sup>, Syed Muhammad Anwar<sup>1,3</sup>, and Marius George Linguraru<sup>1,3</sup>(✉)

<sup>1</sup> Sheikh Zayed Institute for Pediatric Surgical Innovation, Children's National Hospital, Washington, DC, USA

[mlingura@childrensnational.org](mailto:mlingura@childrensnational.org)

<sup>2</sup> Universidad Politécnica de Madrid and CIBER-BBN, ISCIII, Madrid, Spain

<sup>3</sup> School of Medicine and Health Sciences, George Washington University, Washington, DC, USA

**Abstract.** Magnetic Resonance Imaging (MRI) is the primary imaging modality used in the diagnosis, assessment, and treatment planning for brain pathologies. However, most automated MRI analysis tools, such as segmentation and registration pipelines, are optimized for healthy anatomies and often fail when confronted with large lesions such as tumors. To overcome this, image inpainting techniques aim to locally synthesize healthy brain tissues in tumor regions, enabling the reliable application of general-purpose tools. In this work, we systematically evaluate state-of-the-art inpainting models and observe a saturation in their standalone performance. In response, we introduce a methodology combining model ensembling with efficient post-processing strategies such as median filtering, histogram matching, and pixel averaging. Further anatomical refinement is achieved via a lightweight U-Net enhancement stage. Comprehensive evaluation demonstrates that our proposed pipeline improves the anatomical plausibility and visual fidelity of inpainted regions, yielding higher accuracy and more robust outcomes than individual baseline models. By combining established models with targeted post-processing, we achieve improved and more accessible inpainting outcomes, supporting broader clinical deployment and sustainable, resource-conscious research. Our 2025 BraTS inpainting docker is available at [hub.docker.com/layers/aparida12/brats2025/inpt](https://hub.docker.com/layers/aparida12/brats2025/inpt).

**Keywords:** Brain tumor · Inpainting · MRI · Post-processing · Ensemble

## 1 Introduction

Magnetic Resonance Imaging (MRI) and automated analysis are critical for monitoring brain pathologies [1]. However, existing segmentation and registration

---

N. Kulkarni, K. Iyer and A. Tapp—Equal contribution.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2026  
S. Bakas et al. (Eds.): MICCAI 2025, LNCS 16377, pp. 146–157, 2026.  
[https://doi.org/10.1007/978-3-032-16370-7\\_13](https://doi.org/10.1007/978-3-032-16370-7_13)

tools, designed for healthy anatomies, perform poorly in the presence of large lesions such as glioma, limiting their clinical utility [1]. To address these limitations, recent efforts have explored the use of image inpainting to locally synthesize healthy-appearing brain tissue in regions affected by tumors. Inpainting, a well-established task in computer vision, involves reconstructing missing or corrupted parts of an image using contextual information [1, 13]. While traditionally applied to 2D natural images, in the context of neuroimaging, inpainting offers the potential to restore anatomical plausibility in 3D MR volumes, enabling the application of standard processing pipelines even in the presence of pathology [23, 29].

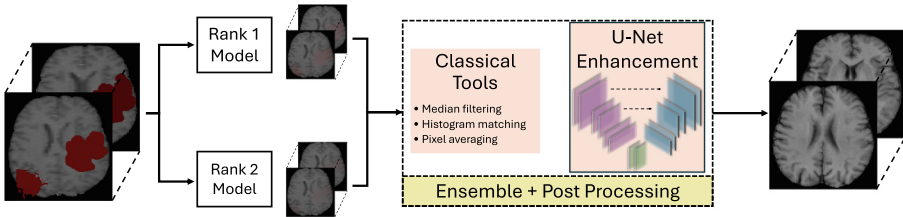
The Brain Tumor Segmentation (BraTS) challenge has historically focused on benchmarking algorithms for tumor segmentation in MRI scans of glioma patients [17, 18]. In 2023, the challenge introduced a new task focused on 3D inpainting of T1-weighted MRIs, where tumor regions were masked out and participants were asked to synthesize realistic, healthy tissue in their place. This task has both clinical and research relevance: in-painted images can improve image registration, facilitate whole-brain parcellation, and support studies of tumor – brain interactions by enabling analysis of the underlying brain anatomy in the absence of visible lesions [23]. Building on this foundation, the BraTS 2025 inpainting challenge continues to advance the field by encouraging the development of robust, generalizable algorithms for localized 3D brain tissue synthesis.

As the field advances, however, critical questions emerge regarding the direction and efficiency of model development for medical image inpainting. In recent years, deep learning models have reached a plateau, with leading algorithms showing only marginal differences in performance metrics [16, 26]. Moreover, even with architectural advances, issues like anatomical implausibility, blurred inpainted regions, and imperfect integration with surrounding tissue persist in many models [16, 21]. This diminishing return prompts reflection: does the field benefit from continually training larger, more complex models, or should efforts shift toward more intelligent utilization of existing resources and data? Medical imaging inherently suffers from data scarcity compared to natural image domains, making continued escalation in model complexity impractical and often counterproductive. Overly complex models not only demand significant computational resources but also risk overfitting, reduced interpretability, and high environmental cost factors that collectively limit real-world clinical translation, especially in resource-constrained clinical sites [5, 10].

A notable limitation of existing inpainting approaches lies in underutilizing abundant healthy brain scans [7, 15]. Current models often struggle with generating high-fidelity, realistic tissue, occasionally resulting in inpainted images that are blurry or anatomically implausible [24]. In contrast, adopting general purpose yet robust architectures such as denoising autoencoders trained on large, healthy image patches offers a promising, generalizable route for enhancing image quality [6, 24]. These models facilitate the restoration of anatomical consistency without needing large, complex networks by focusing on domain-specific feature extraction from healthy regions. Additionally, traditional image process-

ing techniques provide straightforward, computationally efficient post-processing strategies, including median filtering, pixel averaging (ensemble methods), and sharpening [25]. Combined with lightweight learning-based models, these tools have the potential to consistently outperform many computationally intensive approaches, eliminating the need for repeated, large-scale model training [3, 11]. Such solutions could dramatically lower GPU and memory requirements, reduce total energy consumption, and lend themselves to be readily deployable on-site—characteristics that address practical and sustainability considerations for clinical implementation [22]. Additionally, recent studies have demonstrated how intuitively designed methodologies demonstrate robust, generalizable pipelines across diverse clinical environments, as shown in recent efforts leveraging ensemble models for brain tumor segmentation and adaptation in under-resourced settings [2, 8, 9, 20]. Thus, streamlined, ensembled methodologies pave the way for building equity in model deployment [27]. Inequality in access to high-end computational infrastructure is a persistent barrier; therefore, optimizing and reusing established tools democratizes advanced neuroimaging workflows, enabling broader access and more consistent standards of care.

## 2 Methods



**Fig. 1.** Overview of the proposed post-processing pipeline for MRI inpainting. Inpainted outputs from the top two models from the BraTS 2024 inpainting challenge are first aggregated using classical tools like pixel-level averaging. The ensemble result is then refined by post-processing, a basic U-Net trained on synthetically blurred healthy brain regions, or a mix of both. This multi-stage pipeline improves anatomical fidelity and reconstruction sharpness while maintaining computational efficiency.

### 2.1 Dataset

The BraTS local inpainting dataset (2025) [12] is derived from the BraTS glioma segmentation dataset, which contains multi-modal scans (T1, T1ce, T2, and FLAIR) [19]. However, the local inpainting challenge exclusively utilizes the T1-weighted MRI scans for both training and validation. The training set comprises 1251 cases, with each case including the T1 scan, the mask segmenting healthy

and unhealthy regions, the full mask, and a voided T1 scan where the lesion region is removed. The validation set consists of 219 cases, each providing only the voided T1 scan and the mask segmenting healthy and unhealthy regions. This structure enables focused development and evaluation of inpainting algorithms on the T1 modality, distinct from segmentation challenges that use all four modalities.

## 2.2 Proposed Pipeline

Our pipeline integrates multiple stages to generate high-quality, healthy brain tissue inpaintings from masked MRI scans, as illustrated in Fig. 1. Pre-trained models, specifically the top-performing U-Net by Zhang et al. [28] and the 3D Wavelet Diffusion model by Ferreira et al. [4], produce inpainted predictions. These outputs then undergo post-processing steps, including classical pixel averaging filters and a dedicated U-Net-based enhancement module trained to denoise and refine anatomical details. The enhanced outputs from these combined procedures form the final synthesized healthy brain tissue volumes. In the following sections, we describe each component of this pipeline in detail.

### State-of-the-Art (SOTA) Inpainting Model Details.

*U-Net Based Healthy 3D Brain Tissue Inpainting, J.Zhang et al. [28]* As part of our model ensemble, we incorporate the U-Net-based brain tissue inpainting method developed by Zhang et al. Last year, this model was the top performer in the BraTS Local Synthesis of Tissue via Inpainting challenge, establishing it as the SOTA approach for healthy 3D brain tissue synthesis within masked regions of T1-weighted MRI volumes. This strong benchmark performance makes it a highly robust and justifiable choice for inclusion in our pipeline.

In our approach, we leverage the published implementation and utilize the pre-trained weights, ensuring faithful reproducibility with the original SOTA methodology. The architecture is a sophisticated 3D U-Net with three levels of downsampling and upsampling, skip connections, and a ReLU-activated bridge; regularization techniques include instance normalization and dropout (0.2). The loss function combines mean absolute error within the healthy mask regions and structural similarity index over the entire volume. For our ensemble, masked T1 images are input to the pre-trained U-Net model, with the output inpainted regions seamlessly integrated back into the original image context.

*Conditional Wavelet Diffusion Model, A. Ferreira et al. [4].* We incorporate the 3D Wavelet Diffusion Model (WDM) developed by Ferreira et al. as part of our brain tumor inpainting ensemble. This approach achieved second place in the BraTS 2024 Local Synthesis of Tissue via Inpainting challenge, demonstrating competitive performance and complementing the SOTA U-Net model.

The model leverages conditional diffusion processes applied in the wavelet domain, allowing full-resolution 3D brain MRI volumes to be processed without patching or downsampling. This preserves high spatial fidelity during inpainting synthesis while maintaining manageable GPU memory usage. Unlike conventional voxel-space diffusion models, the wavelet transform disentangles spatial frequency components, improving the learning efficiency and reconstruction quality of fine details. The conditional input includes masked MRI images paired with binary masks denoting regions for inpainting. The model was trained to iteratively denoise wavelet coefficients conditioned on these inputs, reconstructing healthy tissue in the tumor-masked regions.

Inference operates by feeding masked MRI data through the trained wavelet diffusion network to generate consistent, high-fidelity tissue inpaintings and synthesized modalities. The outputs are inverse wavelet transformed back into image space and seamlessly integrated into the original volumes for downstream analysis. By incorporating the pre-trained 3D Wavelet Diffusion Model, our ensemble benefits from an advanced generative approach optimized for high-resolution, volumetric brain MRI synthesis with efficient memory usage and strong inpainting accuracy. Together with the U-Net model, this provides complementary strengths in reproducing realistic healthy brain tissue across diverse pathological scenarios.

## Post-processing

**1. Pixel Averaging: Median/Geometric Approach:** After obtaining inpainted outputs from leading SOTA models, classical image processing techniques are employed to enhance visual consistency and reduce instability across the synthesized regions. Specifically, pixel-level aggregation strategies such as voxel-wise average, median, and geometric mean filters are applied to ensemble multiple model predictions. These techniques operate at the voxel level, aggregating pixel intensities across aligned predictions to produce a representative consensus output. Averaging provides a smooth blend of input predictions, minimizing high-frequency disagreement, while the geometric mean is more robust to multiplicative outliers. Notably, when only two predictions are available, average and median aggregations yield mathematically equivalent results. However, the pipeline is designed to accommodate richer ensemble configurations, enabling maximum intensity projection, variance-based adaptive fusion, or hybrid aggregation strategies in future expansions. Collectively, these pixel-level operations offer a robust and interpretable post-processing mechanism that harmonizes multi-model predictions and reduces synthesis variability across diverse MR volumes.

**2. Post-averaging Smoothing:** We employ two classical denoising strategies to refine the ensembled output. First, we apply a 3D median filter with a  $3 \times 3 \times 3$  kernel, which effectively removes localized outlier voxels while preserving edge structure and anatomical boundaries. This median filter is highly effective when ensembling three or more predictions, such as the top three ranked challenge submissions or hybrid combinations (e.g., two predictions from the first-place

model and one from the second-place model). In such cases, the median serves as an outlier-resistant estimator that preserves sharp transitions while rejecting spurious values introduced by any single model and makes it especially beneficial in areas where individual model outputs may hallucinate inconsistent textures or misaligned anatomical boundaries. Second, we apply Gaussian smoothing with a small standard deviation ( $\sigma = 0.5$ ), which gently suppresses high-frequency noise and improves voxel-wise consistency without overly blurring tissue interfaces.

**3. Histogram Matching:** Following smoothing, we perform histogram matching using the output of the Rank 1 model as the reference. This step aligns the intensity distribution of the ensembled prediction with that of the most reliable single-model output. By correcting for model-specific intensity shifts and ensuring consistency with the original intensity range, histogram matching enhances perceptual realism and improves cross-sample comparability, especially in downstream analyses where intensity drift may affect quantitative accuracy.

**4. Deep Learning Enhancement.** As a learned post-processing step, a U-Net model was trained using a synthetic inpainting dataset designed to enable supervised learning of image refinement. To construct this dataset, we started with the original inpainting training dataset and applied random Gaussian blurring to the healthy brain regions defined by the provided healthy masks. This localized degradation mimics the smooth, low-detail appearance of BraTS inpainting methods. The resulting blurred images serve as inputs to the U-Net, while the original, high-resolution MR scans provide the ground truth targets. The network minimizes mean squared error during this task in an attempt to improve ensemble scores. The U-Net effectively minimizes anatomical error, delineates structural boundaries, and optimizes residual smoothing artifacts by leveraging the ground truth healthy tissue appearance. This context-aware enhancement approach complements classical pixel-wise filtering by addressing limitations that ensemble averaging alone cannot resolve. As a result, the U-Net serves as a robust and flexible final-stage enhancement module, significantly improving the perceptual quality and clinical realism of inpainted brain MR volumes.

## 3 Experiments

### 3.1 Metrics

To assess image inpainting quality, we use three main metrics. Structural similarity (SSIM) measures how closely the synthetic image matches the real one in structure, contrast, and luminance; values closer to 1 reflect higher perceptual similarity. Peak signal-to-noise ratio (PSNR) quantifies the ratio between signal and noise based on pixel errors, with higher decibel values indicating better reconstruction quality. Mean squared error (MSE) calculates the average squared difference between predicted and ground truth images, where lower values mean closer pixel-wise agreement.

We optimized the thresholds and chose the best model with the ranking approach proposed by the BraTS team LaBella *et al.* [14]. The evaluation is done

in a hidden test set, computing lesion-wise metrics in all regions and comparing ranks of all the submissions rather than the metrics. To replicate this procedure, we built an internal ranking metric that produces a single score, where a smaller value means better performance on MSE, SSIM, and PSNR. The code is available on GitHub<sup>1</sup>.

The ranking metric approach is robust to outlier predictions, allowing us to optimize for a single value while aligning with the contest evaluation pipeline.

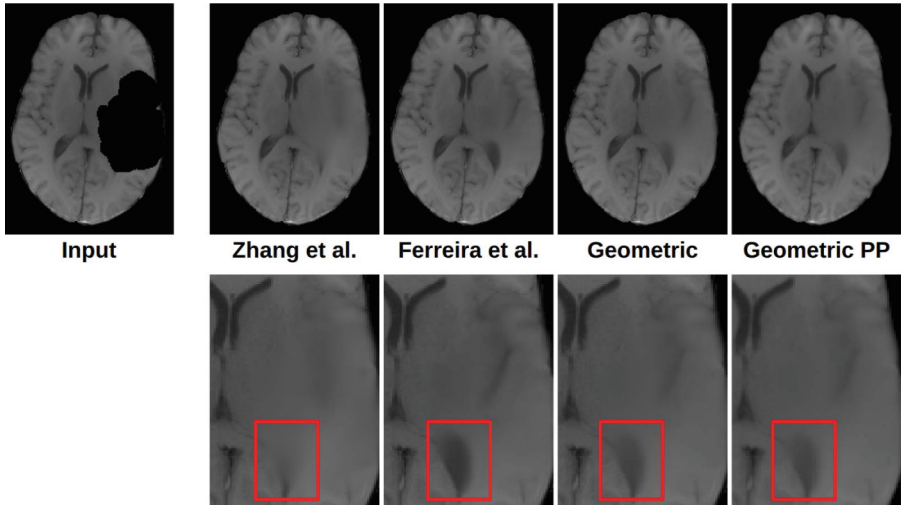
### 3.2 Models for Comparison

We compare five setups: (1) output from Zhang’s U-Net [28], (2) output from Ferreira’s 3D Wavelet Diffusion [4], (3) an ensemble of Zhang’s and Ferreira’s models via geometric or pixel-level averaging, (4) the same ensemble further refined with classical filters such as median or Gaussian, and (5) the filtered ensemble with an additional U-Net-based denoising stage for enhanced image quality. Filter implementation followed a straightforward scheme: model outputs from Zhang and Ferreira were combined voxelwise using equal 50/50 weighting, either via geometric averaging (primary setting) or alternative strategies (mean, median, max, min). The refined outputs were subsequently matched to their corresponding references using histogram matching to preserve intensity distributions. For the U-Net model, training was conducted for 1000 epochs with 250 iterations per epoch, including 50 validation iterations. We optimized with SGD (momentum = 0.99, Nesterov), weight decay =  $3 \times 10^{-5}$ , an initial learning rate of  $1 \times 10^{-2}$ , and polynomial decay. Batch size and patch size were 2 and  $96 \times 160 \times 160$ , respectively. Simple augmentations of rotations ( $\pm 30^\circ$  per axis or anisotropy-aware), uniform scaling [0.7, 1.4], Gaussian noise ( $p = 0.1$ ), Gaussian blur ( $\sigma \in [0.5, 1.0]$ ,  $p = 0.2$ ), brightness scaling ( $\times [0.75, 1.25]$ ,  $p = 0.15$ ), contrast adjustment ( $p = 0.15$ ), simulated low resolution (zoom  $\in [0.5, 1.0]$ ,  $p = 0.25$ ), and gamma corrections ( $p = 0.1/0.3$ ), with mirroring and oversampling of foreground regions were used (0.33). The refinement U-Net minimized mean squared error without deep supervision or additional augmentation. Training used a 5-fold cross-validation split, and results were reported as the equal-weight (50/50) ensemble of model probabilities across folds. This design enables us to assess the impact of model combination and each post-processing step against the state of the art (Fig. 2).

## 4 Results

Table 1 presents the quantitative results for our validation-phase experiments using the BraTS 2025 inpainting benchmark. The table summarizes how various post-processing strategies affect the perceptual and quantitative quality of inpainted brain MR volumes. All reported mean  $\pm$  standard deviation values are

<sup>1</sup> <https://github.com/Pediatric-Accelerated-Intelligence-Lab/BraTS-Unofficial-Ranker>.



**Fig. 2. Qualitative Results** Comparison of inpainted brain MR images from different methods visualized in ITK-SNAP. From left to right: Rank 1 (Zhang et al.), Rank 2 (Ferreira et al.), median ensemble, and geometric mean ensemble. All views are centered at the same voxel coordinate (164, 110, 78), but demonstrate differing image intensity values under the cursor: 931.1 (Rank 1), 664.2 (Rank 2), 797.6 (Geometric), and 796.4 (Geometric with median filter postprocessing). Notably, the bottom-left ventricle appears more faded and ill-defined in the Rank 1 output, while it is clearly delineated in the Rank 2 image. The ventricle is slightly more apparent in both the median and geometric ensemble outputs than in Rank 1, indicating partial recovery of anatomical detail through pixel-wise fusion strategies. These variations highlight differences in tissue reconstruction fidelity between individual models and ensemble-based approaches.

provided directly by the BraTS validation server, where the standard deviation reflects variation across individual validation cases rather than variation across cross-validation folds. Results on the testing set are included in Table XY.

The baseline Rank 1 and Rank 2 models (Ferreira et al. [4] and Zhang et al. [28], respectively) yield identical values for MSE ( $0.007 \pm 0.005$ ) and nearly identical SSIM ( $0.841 \pm 0.103$  vs.  $0.842 \pm 0.099$ ), but differ in PSNR:  $23.257 \pm 4.213$  for Rank 1 and  $22.463 \pm 3.776$  for Rank 2. Simple ensembling of these models yields strong improvements across all metrics, with MSE reduced to  $0.006 \pm 0.004$ , PSNR increased to  $23.652 \pm 4.092$ , and SSIM elevated to  $0.854 \pm 0.094$ . Importantly, this approach also lowers the standard deviations for all metrics, suggesting more stable and consistent performance. Adding classical post-processing filters preserves the MSE and maintains low variance, but leads to slightly reduced PSNR ( $23.385 \pm 3.908$ ) and SSIM ( $0.851 \pm 0.095$ ) relative to ensembling alone (Table 2).

In contrast, the U-Net-enhanced ensemble results in performance that, while still stable, shows a decline in mean values: MSE increases to  $0.007 \pm 0.004$ , PSNR



**Table 1. Validation Set Quantitative results** on the BraTS 2025 inpainting validation set. Performance is evaluated using Structural Similarity Index (SSIM,  $\uparrow$ ), Peak Signal-to-Noise Ratio (PSNR,  $\uparrow$ ), and Mean Squared Error (MSE,  $\downarrow$ ). Metrics reflect perceptual similarity, reconstruction quality, and pixel-level fidelity, respectively. The best performing configuration for each metric is highlighted in **bold**.

Model	Ranking Metric ( $\downarrow$ )	MSE( $\downarrow$ )	PSNR( $\uparrow$ )	SSIM( $\uparrow$ )
Rank 1 (Zhang et al.)	1.970	0.007 $\pm$ 0.005	23.257 $\pm$ 4.213	0.841 $\pm$ 0.103
Rank 2 (Ferreira et al.)	2.441	0.007 $\pm$ 0.005	22.463 $\pm$ 3.776	0.842 $\pm$ 0.099
Ensemble (R1 + R2)	<b>1.223</b>	<b>0.006<math>\pm</math>0.004</b>	<b>23.652<math>\pm</math>4.092</b>	<b>0.854<math>\pm</math>0.094</b>
Ensemble + Filters	<b>1.223</b>	<b>0.006<math>\pm</math>0.004</b>	23.385 $\pm$ 3.908	0.851 $\pm$ 0.095
Ensemble + U-Net	2.366	0.007 $\pm$ <b>0.004</b>	22.740 $\pm$ <b>3.670</b>	0.843 $\pm$ 0.096

**Table 2. Test Set Quantitative results** on the BraTS 2025 inpainting testing set. Performance is evaluated using Structural Similarity Index (SSIM,  $\uparrow$ ), Peak Signal-to-Noise Ratio (PSNR,  $\uparrow$ ), and Mean Squared Error (MSE,  $\downarrow$ ). Metrics reflect perceptual similarity, reconstruction quality, and pixel-level fidelity, respectively.

Model	MSE( $\downarrow$ )	PSNR( $\uparrow$ )	SSIM( $\uparrow$ )
Ensemble + Filters	0.007 $\pm$ 0.004	23.955 $\pm$ 4.989	0.867 $\pm$ 0.1312

drops to 22.740 $\pm$ 3.670, and SSIM falls to 0.843 $\pm$ 0.096. According to the overall ranking metric, the best configuration is the ensemble with classical filters (1.223), followed by Rank 1 (1.970), then the U-Net-enhanced ensemble (2.366). These results suggest that while the U-Net refiner introduces plausible structural detail and reduced variability, it may require further training to outperform simpler, well-tuned classical strategies.

## 5 Discussion

Motivated by the observed saturation in performance among state-of-the-art inpainting models, we proposed a modular pipeline that combines predictions from previous winners’ methods with classical image processing filters and a U-Net-based enhancement module trained on synthetically degraded healthy tissue. The synthetically degraded tissue uses a simple method that applies a Gaussian blur to healthy brain regions, and using these artificially degraded scans as inputs, we generated a rich supervised training dataset for the enhancement model without additional manual annotations. This allowed the U-Net to learn a targeted correction function aimed at restoring anatomical similarity in artifact-prone outputs from existing inpainting models.

Our results indicate that simple model ensembling leads to clear performance improvements across SSIM, PSNR, and MSE while reducing variability compared to individual model outputs. Adding classical post-processing filters preserves these gains with minimal overhead, and ranks best overall across quan-

titative metrics. Our metrics reflect a third place ranking in the BraTS 2025 inpainting challenge with a strategy that emphasizes using an ensemble of existing models and lightweight post-processing rather than creating and training increasingly complex architectures. For the 2025 validation set, the 2024 baseline Rank-1 (Ferreira et al. [4]) and Rank-2 (Zhang et al. [28]) systems yield identical MSE ( $0.007 \pm 0.005$ ) and near-identical SSIM ( $0.841 \pm 0.103$  vs.  $0.842 \pm 0.099$ ), but differ in PSNR ( $23.257 \pm 4.213$  vs.  $22.463 \pm 3.776$ ). A simple ensemble of these models produces consistent improvements across all metrics (MSE  $0.006 \pm 0.004$ , PSNR  $23.652 \pm 4.092$ , SSIM  $0.854 \pm 0.094$ ) and meaningfully reduces standard deviations, indicating more stable performance across slices/cases. Adding classical filters preserves the ensemble’s MSE and low variance but slightly lowers means (PSNR  $23.385 \pm 3.908$ , SSIM  $0.851 \pm 0.095$ ). The U-Net-based refinement showed slight reductions in mean SSIM and PSNR, along with an increase in MSE. These results suggest that while conceptually promising and effective at maintaining low standard deviation, the U-Net enhancement model in its current training configuration may require additional fine-tuning or regularization to outperform traditional averaging and filtering approaches.

Importantly, the components of our postprocessing pipeline are computationally lightweight, reproducible, and deployable in resource-constrained settings. This efficiency not only reduces the burden on GPU memory and energy consumption but also makes high-fidelity inpainting more accessible across diverse research and clinical environments. By strategically leveraging existing models, healthy data cohorts, and simple yet effective post-processing methods, our approach offers a sustainable and modular path forward for medical image inpainting and supports bridging the gap between benchmark performance and real-world applicability.

## 6 Conclusion

Our work shows that strong brain tissue inpainting performance can be achieved without resorting to ever larger or more complex models. By ensembling established, high-performing models and applying lightweight classical post-processing tools alongside a U-Net-based enhancement module trained on healthy scans, we obtain high-fidelity results in a resource-efficient manner. These approaches do not require extensive GPU resources and are simple to implement, demonstrating that carefully selected and combined methods, rather than sheer model scale, can provide reliable, clinically useful brain image synthesis. This strategy supports scalable and accessible solutions for medical image inpainting in both research and clinical practice.

## References

1. Bonato, B., Nanni, L., Bertoldo, A.: Advancing precision: a comprehensive review of MRI segmentation datasets from brats challenges (2012–2025). *Sens. (Basel, Switzerland)* **25**(6), 1838 (2025)
2. Capellán-Martín, D., et al.: Model ensemble for brain tumor segmentation in magnetic resonance imaging. In: *International Challenge on Cross-Modality Domain Adaptation for Medical Image Segmentation*, pp. 221–232. Springer (2023)
3. Dede, A., et al.: *Intelligent systems with applications* (2018)
4. Ferreira, A., Luijten, G., Puladi, B., Kleesiek, J., Alves, V., Egger, J.: Brain tumour removing and missing modality generation using 3D WDM. *arXiv preprint [arXiv:2411.04630](https://arxiv.org/abs/2411.04630)* (2024)
5. Guizard, N., Nakamura, K., Coupé, P., Fonov, V.S., Arnold, D.L., Collins, D.L.: Non-local means inpainting of MS lesions in longitudinal image processing. *Front. Neurosci.* **9**, 456 (2015)
6. He, K., Chen, X., Xie, S., Li, Y., Dollár, P., Girshick, R.: Masked autoencoders are scalable vision learners. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 16000–16009 (2022)
7. Iglesias, J.E., et al.: Synthsr: a public ai tool to turn heterogeneous clinical brain scans into high-resolution t1-weighted images for 3D morphometry. *Sci. Adv.* **9**(5), eadd3607 (2023)
8. Jiang, Z., et al.: Enhancing generalizability in brain tumor segmentation: model ensemble with adaptive post-processing. In: *2024 IEEE International Symposium on Biomedical Imaging (ISBI)*, pp. 1–4. IEEE (2024)
9. Jiang, Z., et al.: Magnetic resonance imaging feature-based subtyping and model ensemble for enhanced brain tumor segmentation. *arXiv preprint [arXiv:2412.04094](https://arxiv.org/abs/2412.04094)* (2024)
10. Kamraoui, R.A., Mansencal, B., Manjon, J.V., Coupé, P.: Longitudinal detection of new MS lesions using deep learning. *Front. Neuro.* **1**, 948235 (2022)
11. Kaur, A., Dong, G.: A complete review on image denoising techniques for medical images. *Neural Process. Lett.* **55**(6), 7807–7850 (2023)
12. Kofler, F., et al.: The brain tumor segmentation (brats) challenge: local synthesis of healthy brain tissue via inpainting. *arXiv preprint [arXiv:2305.08992](https://arxiv.org/abs/2305.08992)* (2023)
13. Kofler, F., et al.: Brats orchestrator: Democratizing and disseminating state-of-the-art brain tumor image analysis. *arXiv preprint [arXiv:2506.13807](https://arxiv.org/abs/2506.13807)* (2025)
14. LaBella, D., et al.: Analysis of the BraTS 2023 intracranial meningioma segmentation challenge. *J. Mach. Learn. Biomed. Imaging* **3**, 38–58 (2025)
15. Liu, X., Xing, F., Yang, C., Kuo, C.C.J., El Fakhri, G., Woo, J.: Symmetric-constrained irregular structure inpainting for brain MRI registration with tumor pathology. In: *International MICCAI Brainlesion Workshop*, pp. 80–91. Springer (2020)
16. Liu, X., Xiang, C., Lan, L., Li, C., Xiao, H., Liu, Z.: Lesion region inpainting: an approach for pseudo-healthy image synthesis in intracranial infection imaging. *Front. Microbiol.* **15**, 1453870 (2024)
17. Maleki, N., et al.: Analysis of the MICCAI brain tumor segmentation–metastases (brats-mets) 2025 lighthouse challenge: Brain metastasis segmentation on pre-and post-treatment MRI. *arXiv preprint [arXiv:2504.12527](https://arxiv.org/abs/2504.12527)* (2025)
18. Menze, B.H., et al.: The multimodal brain tumor image segmentation benchmark (brats). *IEEE Trans. Med. Imaging* **34**(10), 1993–2024 (2014)

19. Menze, B.H., et al.: The multimodal brain tumor image segmentation benchmark (brats). *IEEE Trans. Med. Imaging* **34**(10), 1993–2024 (2015)
20. Parida, A., et al.: Adult glioma segmentation in sub-Saharan Africa using transfer learning on stratified finetuning data. *arXiv preprint [arXiv:2412.04111](https://arxiv.org/abs/2412.04111)* (2024)
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**, imag\_a\_00446 (2025)
22. Prajwal, R., et al.: A study on energy consumption in ai-driven medical image segmentation. *J. Imaging* **11**(6), 174 (2025)
23. Santos, J.C., Tomás Pereira Alexandre, H., Seoane Santos, M., Henriques Abreu, P.: The role of deep learning in medical image inpainting: a systematic review. *ACM Trans. Comput. Healthcare* **6**(3), 1–24 (2025)
24. Sumathi, G., Devi, M.U.: High-resolution image inpainting using a probabilistic framework for diverse images with large arbitrary masks. *Front. Artif. Intell.* **8**, 1614608 (2025)
25. Susan, J., Subashini, P.: Deep learning inpainting model on digital and medical images-a review. *Int. Arab J. Inf. Technol.* **20**(6), 919–936 (2023)
26. Torrado-Carvajal, A., et al.: Inpainting as a technique for estimation of missing voxels in brain imaging. *Ann. Biomed. Eng.* **49**(1), 345–353 (2021)
27. Ueda, D., et al.: Fairness of artificial intelligence in healthcare: review and recommendations. *Jpn. J. Radiol.* **42**(1), 3–15 (2024)
28. Zhang, J., Weng, Y., Chen, K.: U-net based healthy 3D brain tissue inpainting. *arXiv preprint [arXiv:2507.18126](https://arxiv.org/abs/2507.18126)* (2025)
29. Zhu, R., Zhang, X., Pang, H., Xu, C., Ye, C.: Advancing brain tumor inpainting with generative models. *arXiv preprint [arXiv:2402.01509](https://arxiv.org/abs/2402.01509)* (2024)