Anatomical Consistency and Adaptive Prior-informed Transformation for Multi-contrast MR Image Synthesis via Diffusion Model

Supplementary Material

In this supplementary document, we provide additional analyses and ablations to support our main findings, specifically focusing on anatomical consistency and the efficacy of our two-stage inference strategy.

1. Anatomical consistency and meaningful representation

Ensuring anatomical consistency across synthesized multi-contrast MR images is crucial for clinical reliability. To clearly illustrate how effectively the anatomical structural consistency is maintained, Figure 1 provides synthesized FLAIR images conditioned solely on the anatomical feature vector, extracted using our Mutual Information Fusion (MIF) module. Starting purely from Gaussian noise and relying exclusively on the learned anatomical representation, our diffusion model successfully synthesizes images that preserve essential anatomical details and maintain structural alignment with input multi-contrast MR images.

When compared to the image synthesized from a random latent vector (Random z), which shows significant anatomical distortion and lacks clinically relevant details, the synthesis guided by the MIF module maintains key anatomical structures, ensuring realistic anatomical fidelity. This result highlights the robustness and meaningfulness of the anatomical feature vector. By capturing rich anatomical information through the mutual information fusion approach, the feature vector ensures the synthesized images inherently reflect the structural integrity needed for clinical applications without requiring additional guidance or iterative refinements during inference.

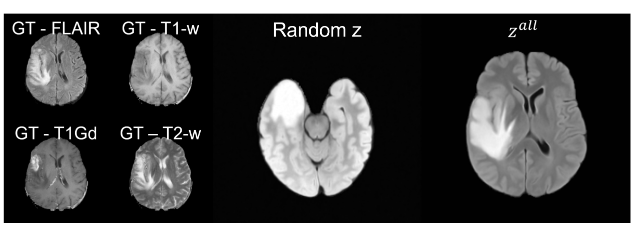

Figure 1. Ablation results for z^{all}.

2. Two-stage inference strategy

The proposed two-stage inference strategy is meticulously designed to enhance the quality of synthesized MR images by explicitly separating and optimizing for global structural alignment and fine anatomical details. The first stage employs adaptive prior-informed transformation, providing a structurally coherent initialization derived from our anatomically-guided codebook. By initiating the synthesis from a structured representation rather than purely random noise, the model effectively captures global anatomical consistency, significantly reducing initial variance and anatomical misalignments.

Following this structured initialization, the second stage employs iterative unrolling at a relatively early timestep. This approach progressively refines the synthesized image through weighted averaging across iterative predictions, addressing any residual inconsistencies or minor anatomical inaccuracies. As shown in Figure 2, the two-stage approach produces more valuable MR images compared to the single-pass diffusion approach.

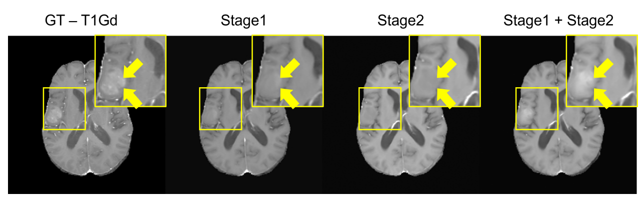

Figure 2. Ablation results for the inference stage.