Autoregressive Image Diffusion: Generation of Image Sequence and Application in MRI [1]

Guanxiong Luo, Shoujin Huang, Martin Uecker

University Medical Center Göttingen & Graz University of Technology

NeurIPS 2024

[1] https://openreview.net/forum?id=jIh4W7r0rn

MRI Reconstruction Challenges

- MRI requires high-quality imaging with rapid acquisition.
- Faster imaging through undersampling introduces artifacts, compromising diagnostic value.

Why We Need Sequence Modeling in MRI

- MRI in clinical practice often involves acquiring volumetric image sequences to monitor disease progression and response to treatment.
- Modeling these sequences is challenging due to the need for coherence across frames.
- Autoregressive models capture dependencies between frames by modeling the joint distribution of image sequences.
- Standard diffusion processes treat images independently, which limits coherence across a sequence.

Combining Autoregressive and Diffusion Models

- To address this, we propose the Autoregressive Image Diffusion (AID) model.
- AID combines the strengths of both autoregressive and diffusion models to generate temporally coherent image sequences.
- This approach captures inter-image dependencies while effectively handling noise, producing coherent MRI sequences ideal for clinical assessment.

Theory - Conditional Probability in AID

- AID models the joint distribution of images as conditional probabilities, ensuring temporal coherence:

$$p(x) = p(x_1) \prod_{t=2}^{N} p(x_t | x_{<t})$$

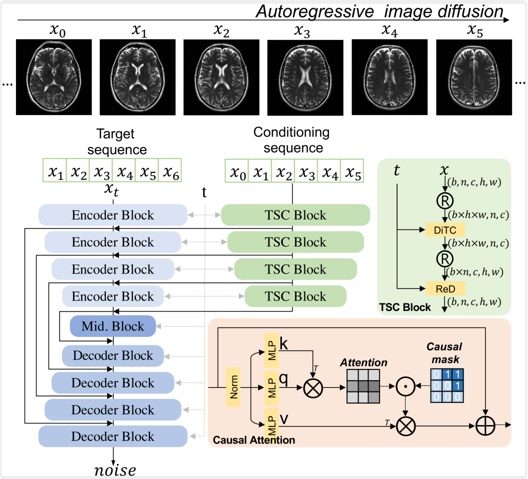
- Each frame depends on the previous frames $x_{<t}$, creating a coherent sequence.
- Forward process adds Gaussian noise iteratively, and the reverse process refines each frame.

$$p_\theta(x_{t-1} | x_t, x_{<t}) = N(x_{t-1}; \mu_\theta(x_t, x_{<t}), \Sigma_\theta(x_t, x_{<t}))$$

- μ_θ and Σ_θ depend on prior frames $x_{<t}$, progressively denoising each frame.

Model Architecture

- Temporal-Spatial Conditioning (TSC) blocks captures inter-frame dependencies.
- Causal attention module ensures each frame depends on previous frames.



Results: Sequence Generation

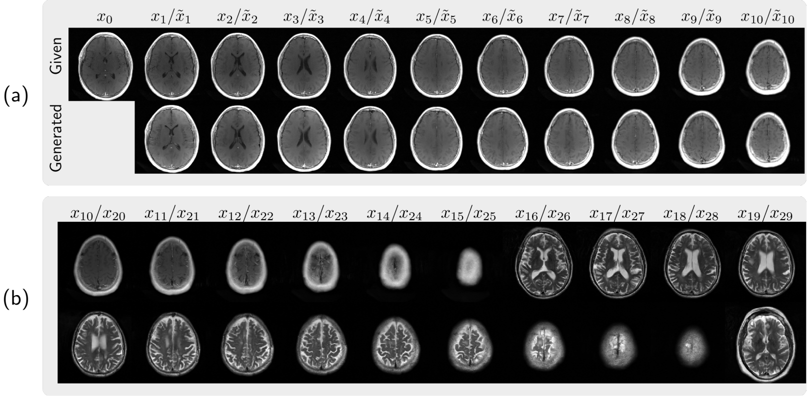


Figure: (a): A sequence of images from dataset is shown in the first row and is used as conditioning to generate retrospective samples that are shown in the second row. (b): With the given sequence in (a) as a warm start, prospective samples extending it are shown.

Results: Cold Start Generation

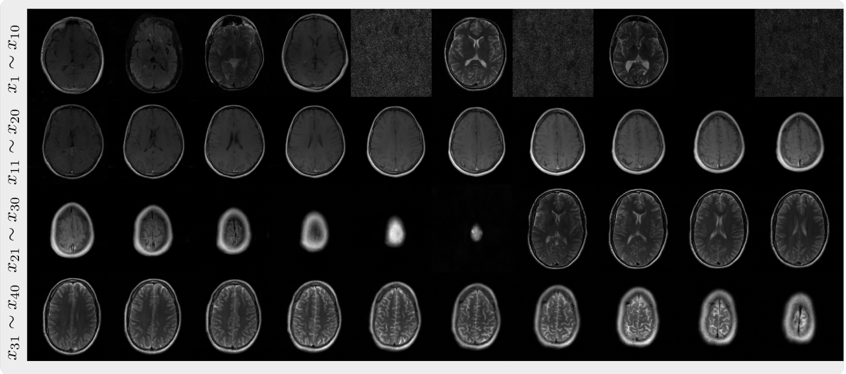


Figure: Prospective samples with cold start. The initial images generated in the cold start are not sequentially coherent, but as the sampling process continues, the model progressively generates more sequentially coherent and realistic images.

Results: MRI Reconstruction Statistics

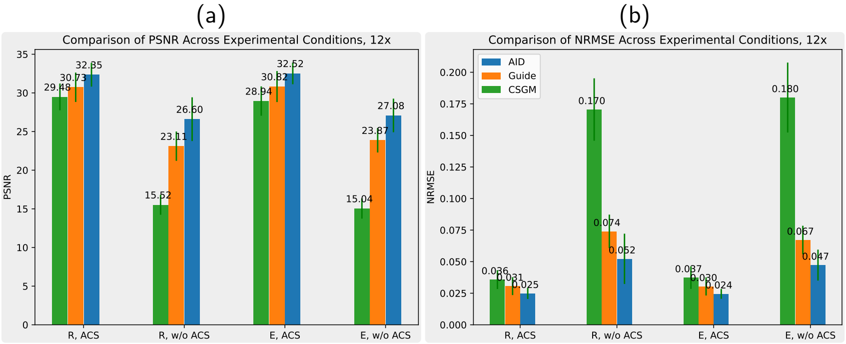


Figure: E: equispaced, R: random. (a): PSNR and (b): NRMSE of the images reconstructed from the twelve-times undersampled k-space data using the autoregressive diffusion model (AID), the standard diffusion model (Guide), and the baseline method CSGM. PSNR higher is better, and NRMSE lower is better.

Results: MRI Reconstruction Quality

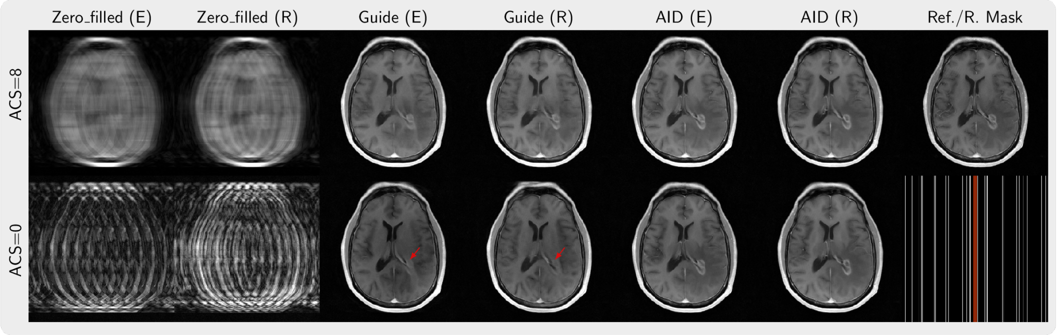


Figure: E: equispaced, R: random. The last column shows the reference and the random sampling mask in k-space. The red lines are autocalibration signal (ACS) and equispaced mask is not shown. Zero-filled images are computed by inverse Fourier transform of the zero-filled k-space data. The hallucinations are pointed with red arrows.

Conclusion

- The proposed autoregressive image diffusion model offers an approach to generating image sequences, with significant potential as a trustworthy prior in accelerated MRI reconstruction.
- In various experiments, it outperforms the standard diffusion model in terms of both image quality and robustness by taking the advantage of the prior information on inter-image dependencies.