COUPLED SPATIAL–SPECTRAL LEARNING FOR GENERATIVE MRI RECONSTRUCTION WITH ENHANCED DIAGNOSTIC SEPARABILITY

Anonymous authors
Paper under double-blind review

ABSTRACT

Accelerated Magnetic Resonance Imaging (MRI) reconstruction is fundamentally an ill-posed inverse problem where aggressive undersampling of the frequency domain (k-space) leads to severe aliasing and the loss of high-frequency structural details. While deep generative adversarial networks (GANs) have established new benchmarks for this task, they frequently struggle with frequency-domain consistency, yielding reconstructions with oversmoothed textures and blurred semantic boundaries. To address this and improve diagnostic separability, we introduce AUSGAN, a novel generative framework that couples spatial attention with spectral regularization to enforce structural fidelity across dual domains. Specifically, the generator integrates soft attention mechanisms to dynamically gate high-frequency features along semantic edges, while the discriminator employs an azimuthal integration module to explicitly penalize distributional shifts in the polar spectral profile. This coupled spatial-spectral synergy prevents the hallucination of isotropic noise while strictly constraining high-frequency recovery to valid anatomical regions. We benchmark AUSGAN against state-of-the-art reconstruction baselines under highly aggressive undersampling rates (up to 10x acceleration) across three diverse clinical cohorts: the ISBR dataset, and multi-modal datasets for Brain Metastasis and Liver Cirrhosis. Going beyond standard pixel-wise metrics, we introduce a boundary-aware evaluation framework leveraging the Bhattacharyya distance to quantitatively measure the feature separability of the reconstructed outputs. Extensive experiments demonstrate that AUSGAN establishes a new state-of-the-art, delivering superior perceptual quality, robust artifact suppression, and significantly enhanced semantic contrast at critical tissue interfaces.

1 INTRODUCTION

Magnetic resonance imaging (MRI) is a preferred clinical modality due to its non-invasive nature, absence of ionizing radiation, and strong soft-tissue contrast. However, long acquisition times remain a major limitation, causing patient discomfort and motion artifacts, particularly in dynamic imaging and pediatric scans. Accelerating MRI while preserving diagnostic fidelity is therefore a critical challenge.

A common acceleration strategy is k-space undersampling, where only a subset of Fourier-domain measurements is acquired. Non-uniform sampling, which prioritizes low-frequency central k-space regions, enables reconstruction from as little as 10% of measurements Yang et al. (2020). Despite reducing scan time, recovering high-frequency components remains difficult, resulting in blurred edges and loss of fine anatomical detail under aggressive undersampling.

Deep learning methods, especially Generative Adversarial Networks (GANs), have been widely adopted to address this limitation. However, purely data-driven approaches lack adherence to MRI acquisition physics, often producing hallucinated high-frequency textures that appear realistic but are anatomically incorrect. This mismatch between spatial realism and frequency-domain consistency limits clinical reliability.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

Recent advances have explored alternative generative paradigms, but significant clinical barriers remain. While diffusion models claim to achieve high perceptual quality, their iterative denoising process results in prohibitively long inference times. More critically, they often produce visually pleasing but clinically irrelevant images that lack the diagnostic separability required by radiologists to differentiate critical tissues. Conversely, multi-modal architectures (e.g., AUNet (Du et al., 2026)) depend heavily on the availability of multiple aligned imaging modalities to produce reliable results, rendering them ineffective in real-world clinical scenarios characterized by data scarcity.

To address these limitations, we propose AUSGAN, an Attention U-Net–based spectral GAN that operates jointly in k-space and image domains. The framework integrates: (1) attention mechanisms to emphasize anatomically salient regions without requiring multi-modal inputs, (2) an azimuthal spectral discriminator to penalize shifts in high-frequency distributions and prevent hallucination, and (3) a multi-term loss enforcing joint spectral consistency and spatial realism, all while maintaining the rapid inference speeds characteristic of feed-forward GANs.

We evaluate AUSGAN on brain and Liver MRI dataset with undersampling rates of 10%–50% (10×–2× acceleration). AUSGAN consistently outperforms state-of-the-art methods in PSNR and SSIM, with stronger gains under aggressive undersampling. Beyond global metrics, we assess tissue contrast using Bhattacharyya distance along anatomical boundaries, showing improved gray matter–white matter and gray matter–cerebrospinal fluid, tumor and peritumoral regions as well as liver boundary separation. These results demonstrate improved preservation of clinically relevant structures without introducing artifacts.

2 Related Work

2.1 Model-Driven and Physics-Informed Networks

Model-driven MRI reconstruction embeds acquisition physics into learning to reduce hallucination and improve stability. Classical methods such as SENSE (Pruessmann et al., 1999), GRAPPA (Griswold et al., 2002), and ESPIRiT (Uecker et al., 2014) exploit coil sensitivity and calibration structure, while later work incorporated symmetry priors (Heidemann et al., 2001) and equivariant learning (Chen et al., 2021a). A major advance was algorithm unrolling, where iterative solvers are converted into trainable networks. MoDL (Aggarwal et al., 2018), ADMM-CSNet (Yang et al., 2020), variational networks (Hammernik et al., 2018), and end-to-end variational approaches (Sriram et al., 2020) combine learned regularization with explicit data consistency, offering improved interpretability and reliability.

However, these methods often depend on fixed priors, predefined optimization structure, and assumptions that may not generalize across anatomies, contrasts, or scanners. Recent multimodal unrolling methods such as the alignment-aware framework of Du *et al.* (Du et al., 2026) further improve reconstruction by leveraging complementary information across modalities, but such pipelines require well-aligned multimodal acquisitions and sufficiently paired training data. This makes them less practical for scarce-data settings or single-modality cohorts, where paired multimodal supervision is limited or unavailable. These limitations motivate hybrid frameworks such as AUSGAN, which preserve physical grounding while retaining the flexibility of data-driven representation learning.

2.2 Data-Driven Architectures

Data-driven approaches were introduced to overcome the computational burden and manual tuning of classical reconstruction. Early CNN-based methods learned direct mappings from undersampled measurements to high-quality images (Wang et al., 2016), while cascaded CNNs (Schlemper et al., 2017) mimicked iterative refinement. To improve fidelity at high acceleration factors, later work adopted residual and dense U-Net variants (Jin et al., 2017; Ding et al., 2019) and multi-scale architectures (Dhengre & Sinha, 2022). GAN-based methods further improved perceptual sharpness and texture recovery (Shaul et al., 2020), often incorporating k-space priors, local-global context, or cyclic consistency losses (Deora et al., 2020; Murugesan et al., 2019; Quan et al., 2018). Transformer-based models, including Swin-based designs (Huang et al., 2022; Zhao et al., 2023) and self-supervised dual-domain methods (Zhou et al., 2023), addressed the limited receptive field of CNNs by modeling long-range dependencies.

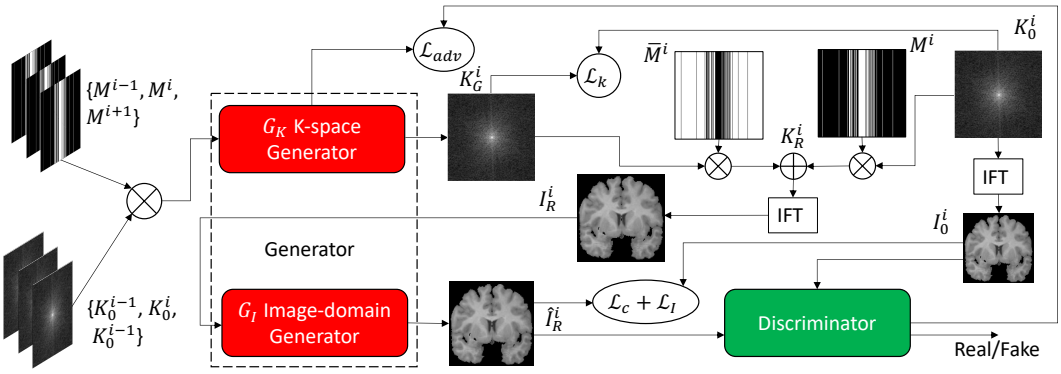

Figure 1: Overview of the proposed dual-domain AUSGAN architecture.

More recently, diffusion models such as AdaDiff (Güngör et al., 2023) used score-based denoising (Song et al., 2020) to achieve strong fidelity, while implicit neural representations and neural fields (Huang et al., 2023; Spieker et al., 2023; Catalán et al., 2025) modeled MRI as a continuous function rather than a discrete grid. NeRF-inspired methods (Mildenhall et al., 2021; Jang & Hyun, 2024; Iddrisu et al., 2023) similarly reformulated reconstruction as continuous 3D rendering. Despite these advances, many GAN, Transformer, Diffusion, and implicit approaches prioritize perceptual realism or spatial context without explicitly constraining high-frequency spectral fidelity, which can lead to smoothing or hallucination of subtle diagnostic boundaries. AUSGAN addresses this gap by coupling rapid spatial reconstruction with explicit frequency-domain regularization through an attention-guided spectral adversarial design.

3 Method

Accelerated MRI reconstruction is fundamentally an ill-posed inverse problem. By violating the Nyquist-Shannon sampling theorem, aggressive k-space undersampling manifests as severe coherent aliasing and the suppression of high-frequency structural details in the spatial domain. To constrain this vast solution space without hallucinating artificial textures, we propose AUSGAN. In this section, we formulate our dual-domain reconstruction pipeline, detail the coupled spatial-spectral mechanisms that govern it, and introduce our boundary-aware clinical evaluation framework.

3.1 Problem Formulation and Dual-Domain Architecture

Let $K_0 \in \mathbb{C}^N$ represent the fully sampled complex k-space signal, and $I_0 = |F^H K_0|$ be the corresponding image, where F is the orthonormal 2D discrete Fourier transform and H is the Hermitian transpose. The undersampled k-space measurements are acquired as $K_p = M \odot K_0$, where M is a binary undersampling mask. Given K_p, we define a generative model $\mathbf{G}$, parameterized by Φ, to solve the inverse problem of estimating the missing frequency samples. As illustrated in Figure 1, $\mathbf{G}$ is a cascaded dual-domain architecture comprising a k-space generator ($\mathbf{G}_K$) and an image-space generator ($\mathbf{G}_I$). To exploit inter-slice spatial dependencies, $\mathbf{G}_K$ utilizes three adjacent slices $M \odot K^{i-1}, M \odot K^i, M \odot K^{i+1}$ to reconstruct the central slice K_G^i. The generated frequencies are merged with the acquired measurements via the inverse mask $\bar{M} = \mathbf{1}_M - M$. The resulting refined k-space is transformed into the spatial domain as an intermediate image $I_R^i = |F^H(K_p + \bar{M} \odot K_G^i)|$, which serves as the input for $\mathbf{G}_I$ to perform final spatial refinement, outputting $\hat{I}_R^i$.

3.2 Spatially-Gated High-Frequency Recovery

Standard Convolutional Neural Networks (CNNs) suffer from limited receptive fields, often smoothing over the high-frequency boundaries necessary for diagnostic accuracy. To strictly constrain feature recovery to anatomically valid regions, we incorporate soft attention mechanisms into the skip connections of our generators (Figure 2). Rather than treating all aliasing artifacts equally, these attention gates act as dynamic spatial filters. For a skip-connection activation map x_l and a gating

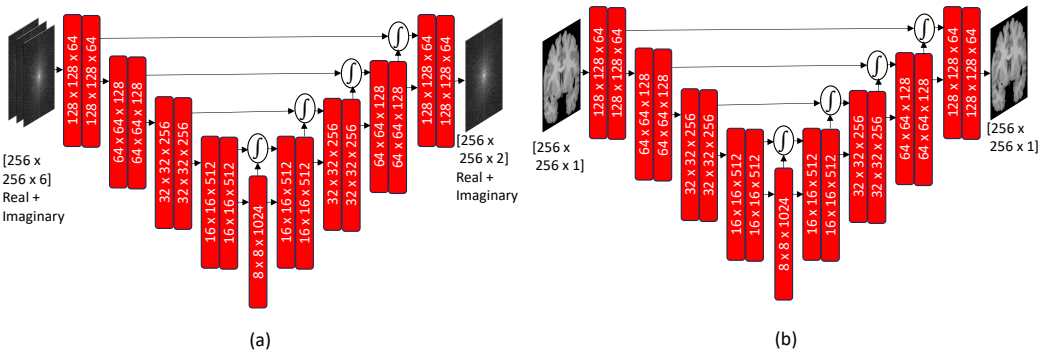

Figure 2: Detailed architecture of the spatially-gated Attention U-Net generators.

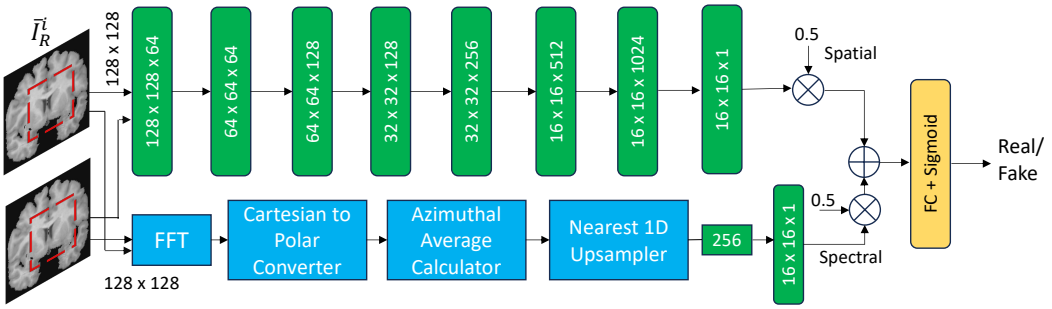

Figure 3: Architecture of the azimuthal spectral discriminator.

signal g from a coarser scale, the attention coefficients $\alpha \in [0, 1]$ are computed as:

$$\alpha = \sigma(\psi^T(\delta(W_g^T g + W_x^T x_l))),$$

where δ is the ReLU function, σ is the sigmoid function, and W_g, W_x, ψ represent 1×1 convolutions. The gated output is defined by the element-wise product:

$$\hat{x}_l = x_l \cdot \alpha.$$

By dynamically weighting features across spatial and channel dimensions, this mechanism forces the generator to prioritize anatomically salient edges while suppressing the isotropic noise inherent to missing k-space data.

3.3 Azimuthal Spectral Regularization

While our attention mechanism resolves spatial blurring, standard GAN discriminators primarily evaluate pixel-wise spatial distributions, ignoring frequency-domain physics. This allows the generator to hallucinate high-frequency textures that appear visually realistic but are clinically false. To mathematically penalize these frequency-domain discrepancies, we employ a spectral discriminator (Figure 3). The discriminator evaluates the generator output $\hat{I}_R^i$ and the ground truth I_0^i simultaneously in both domains. The spatial features are extracted via a cascade of 8 convolutional layers with instance normalization. Concurrently, the spectral features are extracted by computing the discrete Fourier transform of the image, converting it to polar coordinates $r = \sqrt{k^2 + l^2}, \theta = \text{atan2}(l, k)$, and deriving a 1D spectral profile by azimuthally averaging over all angles θ at each radius r. This azimuthal integration explicitly maps the distributional shifts in the high-frequency spectrum. The resulting vector is upsampled to a 16×16 map via nearest-index upsampling and merged with the spatial feature map via point-wise averaging. By forcing the generator to satisfy this dual-domain discriminator, AUSGAN prevents texture hallucination and enforces strict spectral consistency.

3.4 SYNERGISTIC OBJECTIVE FUNCTION

We optimize the AUSGAN framework through a multi-loss objective that synergizes data consistency, perceptual fidelity, and adversarial spectral regularization.

The k-space loss L_k enforces strict fidelity to the unmeasured frequency components:

$$L_k = \|\mathbf{G}_K(K_p) - K_0\|_2.$$

The image-space loss L_I combines ℓ_1 and ℓ_2 penalties to ensure broad structural accuracy:

$$L_I = \alpha\|\hat{I}_R - I_0\|_1 + \beta\|\hat{I}_R - I_0\|_2.$$

While L_I provides a stable global reconstruction, standard pixel-wise metrics notoriously penalize slight spatial shifts in high-frequency textures, leading to oversmoothing. To explicitly preserve these fine diagnostic details, we enforce perceptual fidelity in a high-dimensional feature space. Utilizing the first 35 layers of a pre-trained VGG19 network as a feature extractor, we compute a perceptual content loss L_C. Let $\gamma_{i,j}$ denote the feature representation at the j-th convolution (after activation) before the i-th maxpooling layer, with spatial dimensions $W_{i,j}$ and $H_{i,j}$. The content loss is defined as the squared Euclidean distance between the feature maps of the reconstructed and ground truth images:

$$L_C = \frac{1}{W_{i,j}H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} \left(\gamma_{i,j}(I_0)_{x,y} - \gamma_{i,j}(\hat{I}_R)_{x,y}\right)^2.$$

Finally, the adversarial min-max optimization is defined as:

$$\min_{\mathbf{G}} \max_{\mathbf{D}} V(\mathbf{G}, \mathbf{D}) = \mathbb{E}_{K_0} \log[\mathbf{D}(I_0)] + \mathbb{E}_{K_p} \log[1 - \mathbf{D}(\hat{I}_R)].$$

The total synergistic loss driving the AUSGAN architecture is a weighted sum of these components, mathematically balancing spatial reconstruction with spectral and perceptual constraints: $L_{total} = \lambda_1 L_k + \lambda_2 L_I + \lambda_3 L_C + \lambda_4 V(\mathbf{G}, \mathbf{D})$.

3.5 BOUNDARY-AWARE CLINICAL EVALUATION

Standard quantitative metrics, such as Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity (SSIM), measure global pixel correlations but fail to capture the statistical separability of distinct anatomical tissues required for clinical diagnosis.

To directly quantify clinical diagnostic utility, we introduce a boundary-aware evaluation framework utilizing the Bhattacharyya distance. By overlaying expert manual segmentation masks onto the reconstructed outputs $\hat{I}_R$, we isolate the pixel intensity distributions of critical adjacent tissues. This ensures evaluations are strictly mapped to anatomical ground truths—such as isolating gray matter from white matter in neurological scans, or strictly delineating liver parenchyma without interference from adjacent organs in abdominal cohorts.

Let $p(x)$ and $q(x)$ represent the normalized discrete probability distributions (histograms) of pixel intensities within two adjacent masked regions defined over the same domain X. The Bhattacharyya distance D_B measuring the separability between these two tissue classes is defined as:

$$D_B(p, q) = -\ln\left(\sum_{x \in X} \sqrt{p(x)q(x)}\right),$$

where the argument of the logarithm represents the Bhattacharyya coefficient, measuring the exact amount of overlap between the two distributions. A higher Bhattacharyya distance explicitly correlates to reduced distribution overlap, demonstrating enhanced tissue contrast and superior anatomical separability, thereby validating the clinical viability of the AUSGAN reconstruction.

To rigorously validate this boundary-aware evaluation framework beyond healthy neuroanatomy, we extend our methodology to two complex pathological cohorts: a Liver Cirrhosis dataset and a

Brain Metastasis dataset. These datasets introduce severe real-world reconstruction challenges, such as inherently lower signal-to-noise ratios in abdominal imaging and the critical risk of smoothing over oncological anomalies. For these cohorts, our evaluation strictly isolates the liver boundary interface (crucial for staging fibrosis) and the contrast between the dense tumor core and surrounding peritumoral edema, ensuring AUSGAN is penalized if it fails to preserve diagnostic separability in diseased tissues.

4 Experimental Evaluation

4.1 Datasets and Preprocessing

We evaluated AUSGAN on three MRI datasets: the ISBR brain dataset, a Liver Cirrhosis dataset, and a Brain Metastasis dataset. The ISBR dataset contains T1-weighted brain MRI scans acquired in coronal view with a 1.5 Tesla scanner. For the clinical evaluations, we randomly selected 30 patients from the Liver Cirrhosis dataset (Jha et al., 2025) and 20 patients from the Brain Metastasis dataset (Flouri et al., 2025). All datasets were partitioned at the subject level to avoid data leakage. Images were pre-processed with skull stripping and resizing when applicable and normalized to $[0, 1]$.

For k-space simulation, we applied a 2D Fourier transform and retrospectively undersampled along the phase-encoding direction using variable-density random masks at $R = \{10, 5, 3.3, 2.5, 2\}$, corresponding to 10%, 20%, 30%, 40%, and 50% sampling rates. Different random seeds were used to generate diverse sampling masks. Separate masks were generated for each training sample using different random seeds to ensure diversity.

4.2 Implementation Details

We implemented the proposed method using the PyTorch framework and tested on a variety of systems with NVIDIA GPUs, including a TITAN X 12G, a V100 32G, an A100 80G and a GH200 96G GPU. The network was trained for $\sim$350 epochs with the batch size set to 4. We used Adam as the optimizer for the generator as well as the discriminator network and a onecyle learning rate scheduler for smoother training, with the discriminator learning rate being twice that of the generator.

Code for our method will be made available on GitHub as open-source upon the paper acceptance.

4.3 Evaluation Metrics

We use the standard Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) as evaluation metrics. Larger PSNR and SSIM values indicate better reconstruction. PSNR is used to measure the reconstruction quality of the image, focusing on the loss of details in the reconstructed image.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(I_0^i - \hat{I}_0^i)^2, PSNR = 10\log(\frac{L^2}{MSE}) \tag{1}$$

where L is the maximum value of any image pixel and N is the number of all pixels in I_0^i and $\hat{I}_R^i$. On the other hand, SSIM focuses on differences in brightness, contrast and structure between the reconstructed and ground truth images.

$$SSIM(I_0, \hat{I_R}) = \frac{(2\mu_{I_0}\mu_{\hat{I_R}} + C_1)(\sigma_{I_0\hat{I}_r} + C_1)(\sigma_{I_0\hat{I_R}} + C_2)}{(\mu_{I_0}^2 + \mu_{\hat{I_R}}^2 + C_1)(\sigma_{I_0}^2 + \sigma_{\hat{I_R}}^2 + C_2)} \tag{2}$$

where μ and σ^2 are the mean and variance, $\sigma_{I_0\hat{I_R}}$ is the error between I_0 and $\hat{I_R}$, and C_1 and C_2 are small constants for numerical stability.

While PSNR and SSIM assess global fidelity and perceptual similarity, they do not explicitly clinical validation. To address this limitation, we additionally evaluate reconstruction quality using the Bhattacharyya distance, which measures the dissimilarity between intensity distributions of different tissue types. Using expert-provided manual segmentation masks, we extract edge regions corresponding to tissue interfaces and compute the Bhattacharyya distance between intensity distributions along these boundaries. In particular, we use the manual segmentation masks for Gray Matter

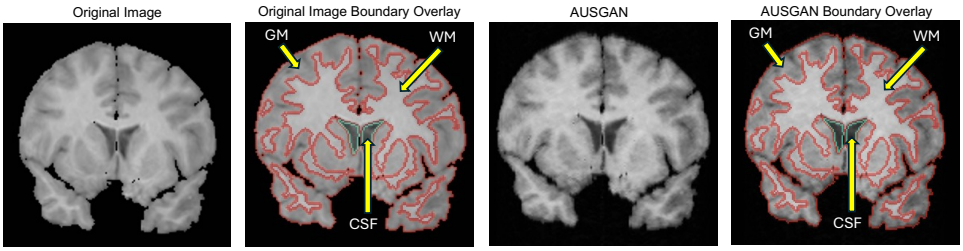

Figure 4: Comparison of tissue boundary segmentation for contrast analysis between the original MRI slice and AUSGAN. Different slices illustrate GM–WM and GM–CSF boundary regions. Segmentation masks highlight Gray Matter (GM), White Matter (WM), and Cerebrospinal Fluid (CSF).

(GM), White Matter (WM), and Cerebrospinal Fluid (CSF) with voxel values other=0, CSF=128, gray=192, and white=254 to extract the boundary regions where tissues interface (GM-WM and GM-CSF boundaries) by applying morphological dilation and subtraction operations (see Figure 4). Additonal overlay figures are provided in the supplementary material.

For each boundary region B_{XY} between tissues X and Y, we compute the intensity distributions P_X and P_Y from pixels within a 3-pixel radius of the boundary. The Bhattacharyya distance D_B between these distributions is calculated as

$$D_B(P_X, P_Y) = -\ln\left(\sum_{i=1}^{N}\sqrt{P_X(i)P_Y(i)}\right), \tag{3}$$

where $P_X(i)$ and $P_Y(i)$ represent the probability densities of intensity bin i for tissues X and Y, respectively. A higher D_B value indicates better separation between tissue intensity distributions, corresponding to improved tissue contrast in the reconstructed image. We compare the D_B values obtained from our AUSGAN reconstructions against those from baseline methods and the ground truth images to demonstrate enhanced tissue differentiation capability. This metric enables a targeted assessment of reconstruction quality at gray matter, white matter, and cerebrospinal fluid boundaries, providing a clinically meaningful measure of anatomical differentiation.

5 Results

To evaluate the performance of AUSGAN, we compare it against zero-filling, the conventional compressed sensing approach CS-MRI (Lustig et al., 2008), and several state-of-the-art learning-based reconstruction methods, including ADMM-CSNet (Yang et al., 2020), DualGAN (Shaul et al., 2020), AUNet (Du et al., 2026), and the score-based diffusion prior AdaDiff (Güngör et al., 2023).

We conduct evaluations across our three distinct clinical cohorts (ISBR, Liver Cirrhosis, and Brain Metastasis) using five aggressive undersampling rates: 10%, 20%, 30%, 40%, and 50%. To best illustrate the comparative performance and diagnostic stability of the models across these acceleration factors, visual trend analyses of tissue-edge contrast are presented in this section (Figure **??**). The comprehensive quantitative tabular results—including all exact PSNR, SSIM, and Bhattacharyya distance measurements for all datasets and ablation studies are provided in the supplementary material.

5.1 Reconstruction Accuracy

Figure **??** shows the PSNR and SSIM values achieved by AUSGAN and the baseline methods across all undersampling rates. Representative reconstruction examples, along with magnified central, are visualized in Figure 5 to facilitate visual comparison.

Across all undersampling rates, AUSGAN consistently achieves the highest PSNR and SSIM values. Improvements are particularly evident at lower sampling rates, where reconstruction is most challenging. Compared to the second-best method(s) (Dual GAN & AUNet), AUSGAN yields up to 19.75% improvement in SSIM and 8.38% improvement in PSNR. Dual GAN & AUNet consistently

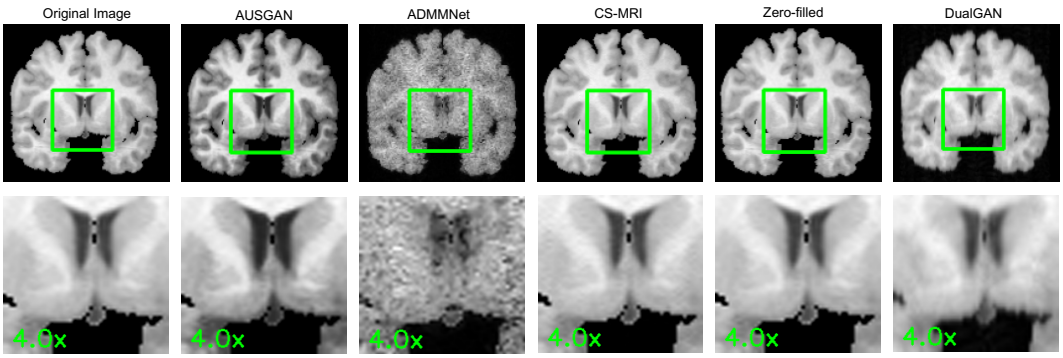

Figure 5: Reconstruction of a representative brain ISBR healthy subject MRI image at 50% undersampling. The first row compares full images from the proposed model with all baseline models, noted above the images. The second row shows 4× magnified views of the brain matter region (green box).

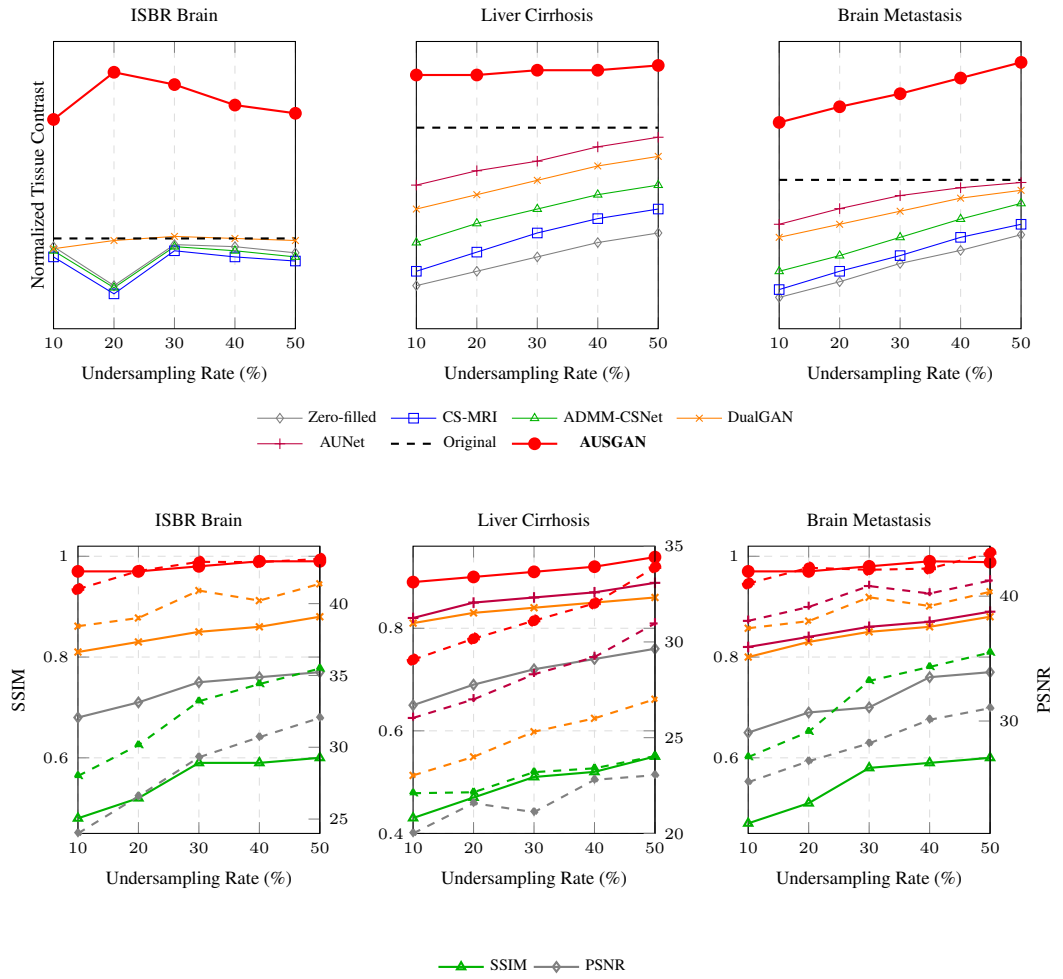

Figure 6: Quantitative reconstruction evaluation across undersampling rates. **Top Row:** Normalized tissue-contrast recovery for ISBR Brain (GM-CSF), Liver Cirrhosis, and Brain Metastasis. **Bottom Row:** Combined SSIM (solid lines, left axis) and PSNR (dashed lines, right axis) trends.

ranks second, while diffusion-based reconstruction using AdaDiff shows competitive performance at higher sampling rates but degrades noticeably under more aggressive undersampling.

Qualitatively, AUSGAN reconstructions exhibit sharper tissue boundaries and reduced blurring compared to the baselines. This behavior can be attributed to the larger effective receptive field provided by the attention-guided U-Net generator, as well as the spectral discriminator, which encourages the preservation of high-frequency image content. These effects are particularly visible in fine anatomical structures that are either smoothed or partially lost in competing methods.

5.2 Tissue-Level Contrast Evaluation

While PSNR and SSIM quantify global reconstruction fidelity, they do not fully capture how well anatomical boundaries are preserved. To address this, we evaluate tissue-level contrast at gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) interfaces using an edge-based Bhattacharyya distance. Higher Bhattacharyya distance values indicate improved tissue separability and clearer anatomical delineation.

As shown in Figure **??** and Figure 4, AUSGAN consistently achieves higher tissue-edge contrast than all competing methods across all undersampling rates. The most pronounced gains are observed for ISBR at the GM–CSF interfaces, where AUSGAN improves contrast by up to 25.0%. Improvements of up to 22.5% are also observed for GM–WM boundaries.

These gains are visually supported in Figure 4 by the overlaid edge masks, which show that AUSGAN better preserves sharp and continuous tissue boundaries, particularly in regions where other methods exhibit blurring or boundary leakage. This result is clinically significant, as GM–CSF contrast is critical for identifying cortical structures and subtle anatomical changes. Moreover, the segmentation masks used were validated by clinicians. Together, these findings indicate that AUSGAN not only improves global reconstruction metrics but also preserves fine-grained anatomical information that is essential for reliable clinical interpretation.

6 Conclusion

In this work, we proposed AUSGAN, an attention-based U-Net GAN with spatial-spectral supervision for accelerated MRI reconstruction from undersampled k-space. Unlike many existing CNN- and GAN-based methods that struggle under aggressive undersampling, AUSGAN is designed to better preserve fine anatomical details and tissue contrast by combining attention-guided feature learning with frequency-aware adversarial training.

The attention modules help the generator focus on anatomically important regions, leading to sharper boundaries and more consistent structures, especially at low sampling rates where reconstruction is most challenging. Across all undersampling factors, AUSGAN consistently outperforms competing methods in PSNR and SSIM, with the largest improvements observed at higher acceleration levels. The spectral discriminator further improves reconstruction quality by encouraging consistency in the frequency domain, which helps retain fine textures and edge details, as confirmed by qualitative results and magnified visual comparisons.

Beyond standard global metrics, we evaluated tissue-level contrast using an edge-based Bhattacharyya distance along expert-annotated tissue boundaries. This analysis shows that AUSGAN significantly improves contrast at GM–CSF and GM–WM interfaces, with gains of up to 25.0% and 22.5%, respectively, and showed modest improvement for Liver tissue boundary and brain metastasis tumor-peritumoral region, indicating better preservation of clinically relevant anatomical distinctions. It is found to work well with both single modality and multi-modality datasets.

Future work will extend AUSGAN to multi-coil reconstruction and task-driven evaluation. Overall, our results show that combining attention mechanisms with spectral adversarial learning is a clinically relevant and practical approach for high-quality accelerated MRI reconstruction.

AI use statement

In this work, we used generative AI tools (specifically, large language models) exclusively for reformatting the manuscript to comply with the ICLR submission guidelines and for minor text polishing

to improve readability. We have not used generative AI tools for generating the core research ideas, experimental design, data analysis, coding, or drafting the primary scientific claims. We have carefully reviewed all AI-assisted text and formatting to ensure absolute accuracy and fidelity to our original findings. We take full responsibility for the final content of this work, including all text, claims, and artifacts produced with the aid of generative AI.

ETHICS STATEMENT

This research utilizes retrospective, fully anonymized MRI datasets (ISBR, Liver Cirrhosis, and Brain Metastasis cohorts) obtained from properly cited and established sources. No new human subject data was collected directly for this study, and all utilized data comply with standard privacy regulations and the original data providers' ethical guidelines. Furthermore, while our generative model (AUSGAN) significantly improves diagnostic separability, we acknowledge that deep learning-based medical reconstruction carries the inherent risk of hallucinating artificial textures. We actively addressed this ethical concern by introducing an azimuthal spectral discriminator and a boundary-aware clinical evaluation framework that specifically penalizes clinically false hallucinations. We stress that this framework is designed to assist, not replace, trained radiologists in clinical decision-making.

REPRODUCIBILITY STATEMENT

To ensure the rigorous reproducibility of our results, we have provided comprehensive details regarding our experimental setup, optimization strategy, and network architectures in Section 4. We clearly specify the data preprocessing steps, k-space simulation parameters (including the variable-density random masks), and the exact formulations for our evaluation metrics (including the boundary-aware Bhattacharyya distance). To facilitate future research and allow the community to verify our claims, our complete PyTorch source code, including the training scripts, network weights, and testing pipelines, will be released publicly on GitHub upon the acceptance of this paper. The clinical datasets utilized in our evaluations are thoroughly referenced to guide proper access.

REFERENCES

Hemant K Aggarwal, Merry P Mani, and Mathews Jacob. MoDL: Model-based deep learning architecture for inverse problems. *IEEE Transactions on Medical Imaging*, 38(2):394–405, 2018.

Amir Aghabiglou and Ender M Eksioglu. Projection-based cascaded U-Net model for MR image reconstruction. *Computer Methods and Programs in Biomedicine*, 207:106151, 2021.

Tabita Catalán, Matías Courdurier, Axel Osses, Anastasia Fotaki, René Botnar, Francisco Sahli-Costabal, and Claudia Prieto. Unsupervised reconstruction of accelerated cardiac cine MRI using neural fields. *Computers in Biology and Medicine*, 185:109467, 2025.

Elena Celledoni, Matthias J Ehrhardt, Christian Etmann, Brynjulf Owren, Carola-Bibiane Schönlieb, and Ferdia Sherry. Equivariant neural networks for inverse problems. *Inverse Problems*, 37(8):085006, 2021.

Dongdong Chen, Julián Tachella, and Mike E Davies. Equivariant imaging: Learning beyond the range space. In *Int. Conf. Comput. Vis.*, pp. 4379–4388, Nashville, 2021a. IEEE.

Yuanqi Chen, Ge Li, Cece Jin, Shan Liu, and Thomas Li. SSD-GAN: Measuring the realness in the spatial and spectral domains. In *AAAI*, volume 35, pp. 1105–1112, Virtually, 2021b. AAAI Press.

Inc CiteDrive. Imperial college london website, 2014. `https://brain-development.org/ixi-dataset/` [Accessed: 02/08/2024].

Puneesh Deora, Bhavya Vasudeva, Saumik Bhattacharya, and Pyari Mohan Pradhan. Structure preserving compressive sensing MRI reconstruction using generative adversarial networks. In *IEEE Conf. Comput. Vis. Pattern Recog. Worksh.*, pp. 522–523, Seattle, 2020. IEEE.

Arjun D Desai, Beliz Gunel, Batu M Ozturkler, Harris Beg, Shreyas Vasanawala, Brian A Hargreaves, Christopher Ré, John M Pauly, and Akshay S Chaudhari. VORTEX: Physics-driven data augmentations for consistency training for robust accelerated MRI reconstruction. *arXiv preprint arXiv:2111.02549*, arXiv–2111(arXiv–2111):arXiv–2111, 2021.

Nikhil Dhengre and Saugata Sinha. Multiscale U-Net-based accelerated magnetic resonance imaging reconstruction. *Signal, Image and Video Processing*, 16:1–8, 2022.

Pak Lun Kevin Ding, Zhiqiang Li, Yuxiang Zhou, and Baoxin Li. Deep residual dense U-Net for resolution enhancement in accelerated MRI acquisition. In *Medical Imaging 2019: Image Processing*, volume 10949, pp. 110–117, Seattle, 2019. SPIE, SPIE.

Yi Du, Tingting Wang, Ziyi Wang, Sunzhe Yang, Faming Fang, and Guixu Zhang. Alignment-aware unrolling network for multimodal mri reconstruction. In *ICASSP 2026-2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 7106–7110. IEEE, 2026.

Zalan Fabian, Reinhard Heckel, and Mahdi Soltanolkotabi. Data augmentation for deep learning based accelerated MRI reconstruction with limited data. In *International Conference on Machine Learning*, pp. 3057–3067, Virtual, 2021. PMLR, PMLR.

Jeffrey A Fessler. Model-based image reconstruction for MRI. *IEEE Signal Processing Magazine*, 27(4):81–89, 2010.

Dimitra Flouri, Christos-Panagiotis Papanikas, Georgios C. Manikis, Eleftherios Ioannou, Georgia Karaoli, Elisavet Papageorgiou, Anastasia Constantinidou, Loizos Siakallis, Marilena Theodorou, and Vasileios Vavourakis. Diagnostic and follow-up mris, cts, radiotherapy and radiomics data of brain metastases, 2025. URL `https://doi.org/10.5281/zenodo.15499111`.

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks. *Communications of the ACM*, 63(11):139–144, 2020.

Mark A Griswold, Peter M Jakob, Robin M Heidemann, Mathias Nittka, Vladimir Jellus, Jianmin Wang, Berthold Kiefer, and Axel Haase. Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magnetic Resonance in Medicine*, 47(6):1202–1210, 2002.

Beliz Gunel, Arda Sahiner, Arjun D Desai, Akshay S Chaudhari, Shreyas Vasanawala, Mert Pilanci, and John Pauly. Scale-equivariant unrolled neural networks for data-efficient accelerated MRI reconstruction. *arXiv preprint arXiv:2204.10436*, 13436:737–747, 2022.

Alper Güngör, Salman UH Dar, Şaban Öztürk, Yilmaz Korkmaz, Hasan A Bedel, Gokberk Elmas, Muzaffer Ozbey, and Tolga Çukur. Adaptive diffusion priors for accelerated MRI reconstruction. *Medical Image Analysis*, 88:102872, 2023.

Kerstin Hammernik, Teresa Klatzer, Erich Kobler, Michael P Recht, Daniel K Sodickson, Thomas Pock, and Florian Knoll. Learning a variational network for reconstruction of accelerated MRI data. *Magnetic Resonance in Medicine*, 79(6):3055–3071, 2018.

Robin M Heidemann, Mark A Griswold, Axel Haase, and Peter M Jakob. VD-AUTO-SMASH imaging. *Magnetic Resonance in Medicine*, 45(6):1066–1074, 2001.

Jiahao Huang, Yingying Fang, Yinzhe Wu, Huanjun Wu, Zhifan Gao, Yang Li, Javier Del Ser, Jun Xia, and Guang Yang. Swin transformer for fast MRI. *Neurocomputing*, 493:281–304, 2022.

Wenqi Huang, Hongwei Bran Li, Jiazhen Pan, Gastao Cruz, Daniel Rueckert, and Kerstin Hammernik. Neural implicit k-space for binning-free non-cartesian cardiac MR imaging. In *International Conference on Information Processing in Medical Imaging*, pp. 548–560. Springer, 2023.

Khadija Iddrisu, Sylwia Malec, and Alessandro Crimi. 3D reconstructions of brain from MRI scans using neural radiance fields. In *International Conference on Artificial Intelligence and Soft Computing*, pp. 207–218. Springer, 2023.

Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pp. 1125–1134, Honolulu, 2017. IEEE.

Tae Jun Jang and Chang Min Hyun. NeRF solves undersampled MRI reconstruction. *arXiv preprint arXiv:2402.13226*, 2024.

Debesh Jha, Onkar Kishor Susladkar, Vandan Gorade, Elif Keles, Matthew Antalek, Deniz Seyithanoglu, Timurhan Cebeci, Halil Ertugrul Aktas, Gulbiz Dagoglu Kartal, Sabahattin Kaymakoglu, Sukru Mehmet Erturk, Yuri Velichko, Daniela P. Ladner, Amir A. Borhani, Alpay Medetalibeyoglu, Gorkem Durak, and Ulas Bagci. Large scale mri collection and segmentation of cirrhotic liver. *Scientific Data*, 12(1):896, 2025. doi: 10.1038/s41597-025-05201-7. URL https://doi.org/10.1038/s41597-025-05201-7.

Kyong Hwan Jin, Michael T McCann, Emmanuel Froustey, and Michael Unser. Deep convolutional neural network for inverse problems in imaging. *IEEE Transactions on Image Processing*, 26(9): 4509–4522, 2017.

Florian Knoll, Kristian Bredies, Thomas Pock, and Rudolf Stollberger. Second order total generalized variation (TGV) for MRI. *Magnetic Resonance in Medicine*, 65(2):480–491, 2011.

Yilmaz Korkmaz, Salman UH Dar, Mahmut Yurt, Muzaffer Özbey, and Tolga Cukur. Unsupervised MRI reconstruction via zero-shot learned adversarial transformers. *IEEE Transactions on Medical Imaging*, 41(7):1747–1763, 2022.

Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pp. 4681–4690, Honolulu, 2017. IEEE.

Zhongnian Li, Tao Zhang, Peng Wan, and Daoqiang Zhang. SEGAN: Structure-enhanced generative adversarial network for compressed sensing MRI reconstruction. In *AAAI*, volume 33, pp. 1012–1019, Honolulu, 2019. AAAI Press.

Michael Lustig, David L Donoho, Juan M Santos, and John M Pauly. Compressed sensing MRI. *IEEE Signal Processing Magazine*, 25(2):72–82, 2008.

Ben Mildenhall, Pratul P Srinivasan, Matthew Tancik, Jonathan T Barron, Ravi Ramamoorthi, and Ren Ng. NeRF: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM*, 65(1):99–106, 2021.

Balamurali Murugesan, S Vijaya Raghavan, Kaushik Sarveswaran, Keerthi Ram, and Mohanasankar Sivaprakasam. Recon-GLGAN: A global-local context based generative adversarial network for MRI reconstruction. In *Machine Learning for Medical Image Reconstruction*, pp. 3–15, Shenzen, 2019. Springer, Springer.

Ozan Oktay, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Mattias Heinrich, Kazunari Misawa, Kensaku Mori, Steven McDonagh, Nils Y Hammerla, Bernhard Kainz, et al. Attention U-Net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999*, abs/1804.03999:1–8, 2018.

Klaas P Pruessmann, Markus Weiger, Markus B Scheidegger, and Peter Boesiger. SENSE: Sensitivity encoding for fast MRI. *Magnetic Resonance in Medicine*, 42(5):952–962, 1999.

Tran Minh Quan, Thanh Nguyen-Duc, and Won-Ki Jeong. Compressed sensing MRI reconstruction using a generative adversarial network with a cyclic loss. *IEEE Transactions on Medical Imaging*, 37(6):1488–1497, 2018.

JH Rick Chang, Chun-Liang Li, Barnabas Poczos, BVK Vijaya Kumar, and Aswin C Sankaranarayanan. One network to solve them all–solving linear inverse problems using deep projection models. In *Int. Conf. Comput. Vis.*, pp. 5888–5897, Venice, 2017. IEEE.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention*, pp. 234–241, Munich, 2015. Springer, Springer.

Jo Schlemper, Jose Caballero, Joseph V Hajnal, Anthony N Price, and Daniel Rueckert. A deep cascade of convolutional neural networks for dynamic MR image reconstruction. *IEEE Transactions on Medical Imaging*, 37(2):491–503, 2017.

Roy Shaul, Itamar David, Ohad Shitrit, and Tammy Riklin Raviv. Subsampled brain MRI reconstruction by generative adversarial neural networks. *Medical Image Analysis*, 65:101747, 2020.

Ohad Shitrit and Tammy Riklin Raviv. Accelerated magnetic resonance imaging by adversarial neural network. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pp. 30–38, Canada, 2017. Springer, Springer.

Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. In Yoshua Bengio and Yann LeCun (eds.), *Int. Conf. Learn. Represent.*, San Diego, CA, USA, 2015. ICLR. URL http://arxiv.org/abs/1409.1556.

Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. *arXiv preprint arXiv:2011.13456*, 2020.

Veronika Spieker, Wenqi Huang, Hannah Eichhorn, Jonathan Stelter, Kilian Weiss, Veronika A Zimmer, Rickmer F Braren, Dimitrios C Karampinos, Kerstin Hammernik, and Julia A Schnabel. Iconik: Generating respiratory-resolved abdominal MR reconstructions using neural implicit representations in k-space. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 183–192. Springer, 2023.

Anuroop Sriram, Jure Zbontar, Tullie Murrell, Aaron Defazio, C Lawrence Zitnick, Nafissa Yakubova, Florian Knoll, and Patricia Johnson. End-to-end variational networks for accelerated MRI reconstruction. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 64–73, Lima, 2020. Springer, Springer.

Martin Uecker, Peng Lai, Mark J Murphy, Patrick Virtue, Michael Elad, John M Pauly, Shreyas S Vasanawala, and Michael Lustig. ESPIRiT—an eigenvalue approach to autocalibrating parallel MRI: where SENSE meets GRAPPA. *Magnetic Resonance in Medicine*, 71(3):990–1001, 2014.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Adv. Neural Inform. Process. Syst.*, NIPS'17, pp. 6000–6010, Red Hook, NY, USA, 2017. Curran Associates Inc. ISBN 9781510860964.

Shanshan Wang, Zhenghang Su, Leslie Ying, Xi Peng, Shun Zhu, Feng Liang, Dagan Feng, and Dong Liang. Accelerating magnetic resonance imaging via deep learning. In *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pp. 514–517, Prague, 2016. IEEE, IEEE.

Guang Yang, Simiao Yu, Hao Dong, Greg Slabaugh, Pier Luigi Dragotti, Xujiong Ye, Fangde Liu, Simon Arridge, Jennifer Keegan, Yike Guo, et al. DAGAN: Deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction. *IEEE Transactions on Medical Imaging*, 37(6):1310–1321, 2017.

Yan Yang, Jian Sun, Huibin Li, and Zongben Xu. ADMM-CSNet: A deep learning approach for image compressive sensing. *IEEE Trans. Pattern Anal. Mach. Intell.*, 42(3):521–538, 2020. doi: 10.1109/TPAMI.2018.2883941.

Kai Zhang, Wangmeng Zuo, Shuhang Gu, and Lei Zhang. Learning deep CNN denoiser prior for image restoration. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pp. 3929–3938, Honolulu, 2017. IEEE.

Lei Zhang and Wangmeng Zuo. Image restoration: From sparse and low-rank priors to deep priors [lecture notes]. *IEEE Signal Processing Magazine*, 34(5):172–179, 2017.

Can Zhao, Muhan Shao, Aaron Carass, Hao Li, Blake E Dewey, Lotta M Ellingsen, Jonghye Woo, Michael A Guttman, Ari M Blitz, Maureen Stone, et al. Applications of a deep learning method for anti-aliasing and super-resolution in MRI. *Magnetic Resonance Imaging*, 64:132–141, 2019.

Xiang Zhao, Tiejun Yang, Bingjie Li, and Xin Zhang. SwinGAN: A dual-domain Swin Transformer-based generative adversarial network for MRI reconstruction. *Computers in Biology and Medicine*, 153:106513, 2023.

Bo Zhou, Neel Dey, Jo Schlemper, Seyed Sadegh Mohseni Salehi, Chi Liu, James S Duncan, and Michal Sofka. DSFormer: A dual-domain self-supervised transformer for accelerated multi-contrast MRI reconstruction. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 4966–4975, Waikoloa, 2023. IEEE.

A Ablation Study

To understand their individual contributions, we evaluated AUSGAN without some of its modules and with different input number of slices and show the results in Tables 10, 11, and 12. The full AUSGAN framework consistently achieves the highest tissue-edge contrast, SSIM, and PSNR, highlighting the strong synergy among its constituent modules. Removing the attention mechanism leads to a clear reduction in tissue-edge contrast, most notably at GM–CSF boundaries, confirming the role of attention in emphasizing anatomically relevant regions. Likewise, ablating the spectral profile–based classification or the perceptual loss results in moderate but consistent performance degradation, underscoring their importance in preserving fine structural details.

The results highlight the role of each component in the AUSGAN framework. The most notable observation is the sharp performance degradation when either the image-domain or k-space generator is removed. Eliminating image-domain refinement leads to a substantial drop in structural fidelity, as reflected by lower SSIM values, while removing k-space consistency results in unstable reconstructions with significantly reduced PSNR. These results emphasize that the two domains are complementary and cannot be treated independently—accurate spatial structure requires consistent frequency-domain modeling, and vice versa.

The slice-wise analysis further demonstrates the importance of contextual information. Moving from a single-slice input to multi-slice configurations provides the model with richer anatomical context, enabling more reliable reconstruction of missing structures. While increasing the context to five slices improves performance over the single-slice setting, it does not surpass the three-slice configuration used in the full AUSGAN model. This suggests that an optimal balance exists between contextual information and model complexity.

B Evaluation on Additional Clinical Datasets

To evaluate the robustness, generalizability, and diagnostic separability of AUSGAN beyond the primary healthy brain cohort (ISBR), we conducted extensive evaluations on two additional pathological datasets: Liver Cirrhosis and Brain Metastasis. Due to specific submission constraints regarding the inclusion of additional datasets in the main manuscript, the comprehensive quantitative tabular results for these cohorts are provided in this supplementary material.

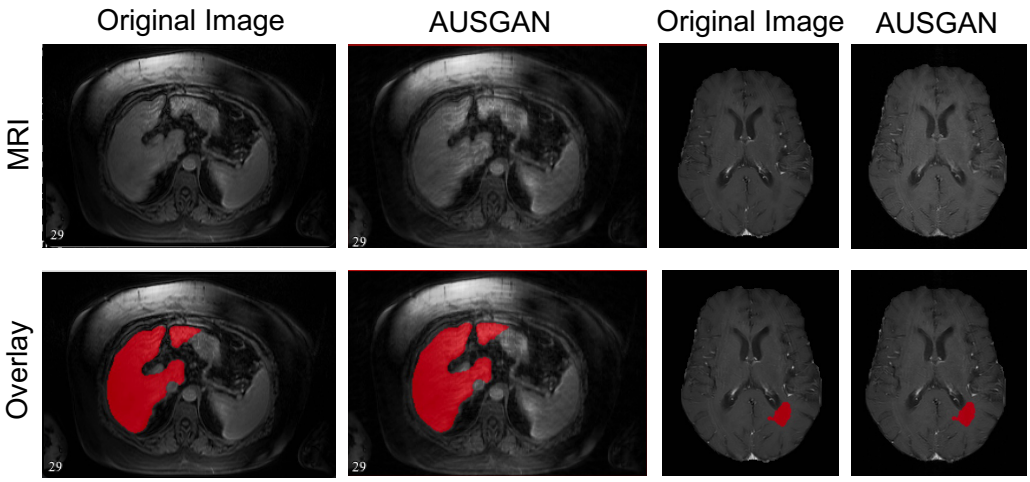

Figure 7: Reconstruction of a representative Liver Cirrhosis and Brain Metastasis MRI image at 50% undersampling. The first row compares MRI images from AUSGAN and original images, noted above the images. The second row shows overlay image of the manually annotated segmentation ground truth by clinicians.

Liver Cirrhosis Dataset: Reconstructing abdominal MRI is notoriously challenging due to inherently lower signal-to-noise ratios and the prevalence of motion artifacts. We evaluated our model's

ability to maintain structural fidelity in this domain, with a specific focus on the liver boundary interface. Precise tissue contrast at this boundary is clinically imperative for identifying fibrotic morphological changes and staging cirrhosis.

Brain Metastasis Dataset: To ensure the model does not smooth over critical pathological anomalies, we evaluated AUSGAN on a neuro-oncology dataset. Specifically, we measured the feature separability between the dense tumor core and the surrounding peritumoral edema. Maintaining high contrast in this region is a vital biomarker for tumor delineation and subsequent surgical or radiological treatment planning.

As demonstrated in Tables 4 through 9, AUSGAN consistently establishes the state-of-the-art across both pathological datasets, maintaining superior PSNR and SSIM while demonstrating unprecedented diagnostic separability at critical tissue interfaces, even under aggressive 10% sampling rates.

Table 1: Comparison of Reconstruction Models on the ISBR Brain Dataset: SSIM Across 5 Undersampling Rates

Method	10%	20%	30%	40%	50%
CS-MRI (Lustig et al., 2008)	0.15 ± 0.005	0.18 ± 0.009	0.21 ± 0.021	0.21 ± 0.012	0.23 ± 0.012
AdaDiff (Güngör et al., 2023)	0.35 ± 0.025	0.37 ± 0.019	0.37 ± 0.030	0.38 ± 0.052	0.39 ± 0.132
ADMM-CSNet (Quan et al., 2018)	0.48 ± 0.030	0.52 ± 0.025	0.59 ± 0.020	0.59 ± 0.028	0.60 ± 0.031
Zero-filled	0.68 ± 0.020	0.71 ± 0.018	0.75 ± 0.019	0.76 ± 0.022	0.77 ± 0.019
DualGAN (Shaul et al., 2020)	$\underline{0.81 \pm 0.001}$	$\underline{0.83 \pm 0.007}$	$\underline{0.85 \pm 0.008}$	$\underline{0.86 \pm 0.009}$	$\underline{0.88 \pm 0.009}$
AUNet (Du et al., 2026)	-	-	-	-	-
AUSGAN	$\mathbf{0.97 \pm 0.004}$	$\mathbf{0.97 \pm 0.002}$	$\mathbf{0.98 \pm 0.013}$	$\mathbf{0.99 \pm 0.001}$	$\mathbf{0.99 \pm 0.012}$

Note: Bold indicates best performing method and underline represents the second best. SSIM = Structural Similarity Index Metric. Since AUNet is a multi-modal model, hence could not be used.

Table 2: Comparison of Reconstruction Models on the ISBR Brain Dataset: PSNR Across 5 Undersampling Rates

Method	10%	20%	30%	40%	50%
CS-MRI (Lustig et al., 2008)	24.50 ± 0.820	32.51 ± 0.971	34.67 ± 0.221	37.85 ± 0.100	38.67 ± 0.690
AdaDiff (Güngör et al., 2023)	22.10 ± 0.910	23.19 ± 0.592	23.27 ± 0.415	21.01 ± 0.950	21.47 ± 0.821
ADMM-CSNet (Quan et al., 2018)	28.00 ± 1.501	30.15 ± 1.100	33.20 ± 1.850	34.40 ± 1.931	35.48 ± 1.730
Zero-filled	24.00 ± 1.010	26.58 ± 1.730	29.30 ± 1.810	30.70 ± 1.450	32.04 ± 1.690
DualGAN (Shaul et al., 2020)	$\underline{38.43 \pm 0.521}$	$\underline{39.00 \pm 0.013}$	$\underline{40.91 \pm 0.432}$	$\underline{40.21 \pm 0.176}$	$\underline{41.38 \pm 0.435}$
AUNet (Du et al., 2026)	-	-	-	-	-
AUSGAN	$\mathbf{41.01 \pm 0.330}$	$\mathbf{42.27 \pm 0.110}$	$\mathbf{42.91 \pm 0.120}$	$\mathbf{42.89 \pm 0.002}$	$\mathbf{43.11 \pm 0.128}$

Note: Bold indicates best performing method and underline represents the second best. PSNR = Peak Signal-to-Noise Ratio.

Table 3: Comparison of Reconstruction Models on the ISBR Brain Dataset Tissue-Edge Contrast Using 5 Undersampling Rates

Method	10%		20%		30%		40%		50%	
	GM-WM	GM-CSF	GM-WM	GM-CSF	GM-WM	GM-CSF	GM-WM	GM-CSF	GM-WM	GM-CSF
CS-MRI (Lustig et al., 2008)	0.105 ± 0.036	0.315 ± 0.099	0.110 ± 0.056	0.297 ± 0.147	0.112 ± 0.041	0.318 ± 0.107	0.109 ± 0.041	0.315 ± 0.098	0.107 ± 0.056	0.313 ± 0.081
ADMM-CSNet (Quan et al., 2018)	0.108 ± 0.036	0.318 ± 0.099	0.112 ± 0.056	0.300 ± 0.147	0.114 ± 0.041	0.320 ± 0.107	0.111 ± 0.041	0.318 ± 0.098	0.109 ± 0.056	0.315 ± 0.081
Zero-filled	0.110 ± 0.042	0.320 ± 0.116	$\underline{0.113 \pm 0.066}$	0.301 ± 0.173	$\underline{0.115 \pm 0.048}$	0.321 ± 0.126	0.112 ± 0.048	0.320 ± 0.115	0.110 ± 0.066	0.317 ± 0.095
DualGAN (Shaul et al., 2020)	0.109 ± 0.084	0.319 ± 0.151	0.111 ± 0.120	0.323 ± 0.112	0.113 ± 0.078	$\underline{0.325 \pm 0.188}$	$\underline{0.118 \pm 0.105}$	0.324 ± 0.119	$\underline{0.113 \pm 0.101}$	0.323 ± 0.251
AUNet (Du et al., 2026)	-	-	-	-	-	-	-	-	-	-
Original Image	$\underline{0.111 \pm 0.056}$	$\underline{0.324 \pm 0.148}$	0.111 ± 0.056	$\underline{0.324 \pm 0.148}$	0.111 ± 0.056	0.324 ± 0.148	0.111 ± 0.056	$\underline{0.324 \pm 0.148}$	0.111 ± 0.056	$\underline{0.324 \pm 0.148}$
AUSGAN	$\mathbf{0.128 \pm 0.049}$	$\mathbf{0.382 \pm 0.137}$	$\mathbf{0.136 \pm 0.078}$	$\mathbf{0.405 \pm 0.204}$	$\mathbf{0.135 \pm 0.056}$	$\mathbf{0.399 \pm 0.148}$	$\mathbf{0.133 \pm 0.056}$	$\mathbf{0.389 \pm 0.135}$	$\mathbf{0.131 \pm 0.078}$	$\mathbf{0.385 \pm 0.112}$

Note: Bold indicates best performing method and underline represents the second best. GM = Gray Matter, WM = White Matter, CSF = Cerebrospinal Fluid.

Table 4: Comparison of Reconstruction Models on the Liver Cirrhosis Dataset: PSNR Across 5 Undersampling Rates

Method	10%	20%	30%	40%	50%
CS-MRI (Lustig et al., 2008)	24.50 ± 0.820	32.51 ± 0.971	34.67 ± 0.221	37.85 ± 0.100	38.67 ± 0.690
AdaDiff (Güngör et al., 2023)	22.10 ± 0.910	23.19 ± 0.592	23.27 ± 0.415	21.01 ± 0.950	21.47 ± 0.821
ADMM-CSNet (Quan et al., 2018)	22.10 ± 0.341	22.15 ± 0.100	23.20 ± 0.650	23.40 ± 0.931	24.04 ± 0.030
Zero-filled	20.00 ± 0.910	21.58 ± 0.030	21.11 ± 1.010	22.80 ± 1.110	23.04 ± 0.22
DualGAN (Shaul et al., 2020)	23.03 ± 0.011	24.00 ± 0.022	25.31 ± 0.041	26.01 ± 0.054	27.008 ± 0.034
AUNet (Du et al., 2026)	$\underline{26.03 \pm 0.021}$	$\underline{27.00 \pm 0.013}$	$\underline{28.31 \pm 0.432}$	$\underline{29.21 \pm 0.076}$	$\underline{30.9648 \pm 0.6336}$
AUSGAN	$\mathbf{29.07 \pm 0.230}$	$\mathbf{30.17 \pm 0.210}$	$\mathbf{31.11 \pm 0.190}$	$\mathbf{32.01 \pm 0.002}$	$\mathbf{33.8991 \pm 0.9975}$

Note: Bold indicates best performing method and underline represents the second best. PSNR = Peak Signal-to-Noise Ratio.

Table 5: Comparison of Reconstruction Models on the Liver Cirrhosis Dataset: SSIM Across 5 Undersampling Rates

Method	10%	20%	30%	40%	50%
CS-MRI (Lustig et al., 2008)	0.14 ± 0.021	0.15 ± 0.012	0.17 ± 0.021	0.16 ± 0.012	0.17 ± 0.027
AdaDiff (Güngör et al., 2023)	0.30 ± 0.015	0.31 ± 0.019	0.33 ± 0.040	0.34 ± 0.051	0.35 ± 0.032
ADMM-CSNet (Quan et al., 2018)	0.43 ± 0.021	0.47 ± 0.015	0.51 ± 0.030	0.52 ± 0.029	0.55 ± 0.031
Zero-filled	0.65 ± 0.032	0.69 ± 0.019	0.72 ± 0.029	0.74 ± 0.012	0.76 ± 0.032
DualGAN (Shaul et al., 2020)	0.81 ± 0.001	0.83 ± 0.007	0.84 ± 0.008	0.85 ± 0.009	$\underline{0.86 \pm 0.0129}$
AUNet (Du et al., 2026)	$\underline{0.82 \pm 0.001}$	$\underline{0.85 \pm 0.007}$	$\underline{0.86 \pm 0.021}$	$\underline{0.87 \pm 0.010}$	$\underline{0.8885 \pm 0.013}$
AUSGAN	$\mathbf{0.89 \pm 0.014}$	$\mathbf{0.90 \pm 0.023}$	$\mathbf{0.91 \pm 0.025}$	$\mathbf{0.92 \pm 0.005}$	$\mathbf{0.9387 \pm 0.0149}$

Note: Bold indicates best performing method and underline represents the second best. SSIM = Structural Similarity Index Metric.

Table 6: Comparison of Reconstruction Models: Tissue Contrast of Liver Boundary Improvement Across 5 Undersampling Rates for Liver Cirrhosis dataset

Method	10%	20%	30%	40%	50%
CS-MRI (Lustig et al., 2008)	0.622 ± 0.051	0.626 ± 0.050	0.630 ± 0.049	0.633 ± 0.048	0.635 ± 0.048
ADMM-CSNet (Quan et al., 2018)	0.628 ± 0.052	0.632 ± 0.051	0.635 ± 0.049	0.638 ± 0.047	0.640 ± 0.046
Zero-filled	0.619 ± 0.060	0.622 ± 0.058	0.625 ± 0.055	0.628 ± 0.054	0.630 ± 0.052
DualGAN (Shaul et al., 2020)	0.635 ± 0.049	0.638 ± 0.048	0.641 ± 0.048	0.644 ± 0.047	0.646 ± 0.047
AUNet (Du et al., 2026)	$\underline{0.640 \pm 0.048}$	$\underline{0.643 \pm 0.047}$	$\underline{0.645 \pm 0.046}$	$\underline{0.648 \pm 0.046}$	$\underline{0.650 \pm 0.045}$
Original Image	0.652 ± 0.045	0.652 ± 0.045	0.652 ± 0.045	0.652 ± 0.045	0.652 ± 0.045
AUSGAN	$\mathbf{0.663 \pm 0.049}$	$\mathbf{0.663 \pm 0.048}$	$\mathbf{0.664 \pm 0.048}$	$\mathbf{0.664 \pm 0.047}$	$\mathbf{0.665 \pm 0.047}$

Note: Bold indicates best performing method and underline represents the second best. Contrast measured at the liver boundary interface.

Table 7: Comparison of Reconstruction Models on the Brain Metastasis Dataset: PSNR Across 5 Undersampling Rates

Method	10%	20%	30%	40%	50%
CS-MRI (Lustig et al., 2008)	21.15 ± 0.130	23.51 ± 0.171	24.67 ± 0.231	25.81 ± 0.060	25.97 ± 0.590
AdaDiff (Güngör et al., 2023)	23.01 ± 0.110	24.12 ± 0.56	24.97 ± 0.015	24.01 ± 0.910	24.26 ± 0.212
ADMM-CSNet (Quan et al., 2018)	27.11 ± 1.010	29.14 ± 1.201	33.21 ± 1.058	34.330 ± 1.51	35.48 ± 0.990
Zero-filled	25.10 ± 1.01	26.78 ± 1.00	28.21 ± 0.110	30.10 ± 1.102	31.03 ± 1.050
DualGAN (Shaul et al., 2020)	$\underline{37.43 \pm 0.521}$	$\underline{38.00 \pm 0.013}$	$\underline{39.91 \pm 0.432}$	$\underline{39.21 \pm 0.176}$	$\underline{40.35 \pm 0.412}$
AUNet (Du et al., 2026)	$\underline{38.02 \pm 0.131}$	$\underline{39.13 \pm 0.22}$	$\underline{40.81 \pm 0.031}$	$\underline{40.21 \pm 0.016}$	$\underline{41.25 \pm 1.873}$
AUSGAN	$\mathbf{41.01 \pm 0.330}$	$\mathbf{42.27 \pm 0.110}$	$\mathbf{42.11 \pm 0.10}$	$\mathbf{42.21 \pm 0.042}$	$\mathbf{43.45 \pm 1.52}$

Note: Bold indicates best performing method and underline represents the second best. PSNR = Peak Signal-to-Noise Ratio.

Table 8: Comparison of Reconstruction Models on the Brain Metastasis Dataset: SSIM Across 5 Undersampling Rates

Method	10%	20%	30%	40%	50%
CS-MRI (Lustig et al., 2008)	0.14 ± 0.105	0.17 ± 0.209	0.22 ± 0.033	0.22 ± 0.702	0.24 ± 0.904
AdaDiff (Güngör et al., 2023)	0.34 ± 0.109	0.35 ± 0.022	0.36 ± 0.180	0.37 ± 0.041	0.38 ± 0.390
ADMM-CSNet (Quan et al., 2018)	0.47 ± 0.020	0.51 ± 0.130	0.58 ± 0.120	0.59 ± 0.009	0.60 ± 0.031
Zero-filled	0.65 ± 0.120	0.69 ± 0.015	0.70 ± 0.0289	0.76 ± 0.022	0.77 ± 0.019
DualGAN (Shaul et al., 2020)	$\underline{0.80 \pm 0.001}$	$\underline{0.83 \pm 0.007}$	$\underline{0.85 \pm 0.018}$	$\underline{0.86 \pm 0.029}$	$\underline{0.88 \pm 0.109}$
AUNet (Du et al., 2026)	$\underline{0.82 \pm 0.021}$	$\underline{0.84 \pm 0.027}$	$\underline{0.86 \pm 0.218}$	$\underline{0.87 \pm 0.039}$	$\underline{0.89 \pm 0.209}$
AUSGAN	$\mathbf{0.97 \pm 0.004}$	$\mathbf{0.97 \pm 0.002}$	$\mathbf{0.98 \pm 0.013}$	$\mathbf{0.99 \pm 0.001}$	$\mathbf{0.9883 \pm 0.004437}$

Note: Bold indicates best performing method and underline represents the second best. SSIM = Structural Similarity Index Metric.

Table 9: Comparison of Reconstruction Models: Tumor-Peritumoral Edema Tissue Contrast Across 5 Undersampling Rates for Brain Metastasis dataset

Method	10%	20%	30%	40%	50%
CS-MRI (Lustig et al., 2008)	0.605 ± 0.051	0.612 ± 0.050	0.618 ± 0.049	0.625 ± 0.048	0.630 ± 0.048
ADMM-CSNet (Quan et al., 2018)	0.612 ± 0.052	0.618 ± 0.051	0.625 ± 0.049	0.632 ± 0.047	0.638 ± 0.046
Zero-filled	0.602 ± 0.060	0.608 ± 0.058	0.615 ± 0.055	0.620 ± 0.054	0.626 ± 0.052
DualGAN (Shaul et al., 2020)	0.625 ± 0.049	0.630 ± 0.048	0.635 ± 0.048	0.640 ± 0.047	0.643 ± 0.047
AUNet (Du et al., 2026)	$\underline{0.630 \pm 0.048}$	$\underline{0.636 \pm 0.047}$	$\underline{0.641 \pm 0.046}$	$\underline{0.644 \pm 0.046}$	$\underline{0.646 \pm 0.045}$
Original Image	0.647 ± 0.045	0.647 ± 0.045	0.647 ± 0.045	0.647 ± 0.045	0.647 ± 0.045
AUSGAN	$\mathbf{0.669 \pm 0.049}$	$\mathbf{0.675 \pm 0.048}$	$\mathbf{0.680 \pm 0.048}$	$\mathbf{0.686 \pm 0.047}$	$\mathbf{0.692 \pm 0.047}$

Note: Bold indicates best performing method and underline represents the second best. Contrast measured between tumor core and surrounding peritumoral edema.

Table 10: Effects of Proposed Modules on ISBR Dataset: PSNR Comparison

Method	10%	20%	30%	40%	50%
AUSGAN − Image Gen.	21.56 ± 0.201	21.98 ± 0.541	22.02 ± 0.102	22.23 ± 0.113	22.87 ± 0.695
AUSGAN − K-space Gen.	0.29 ± 0.123	0.33 ± 0.522	0.34 ± 0.815	0.36 ± 0.151	0.35 ± 0.112
AUSGAN (1-slice)	20.28 ± 0.271	21.18 ± 0.761	22.01 ± 0.652	22.98 ± 0.13	23.31 ± 0.227
AUSGAN (3-slice)	$\mathbf{41.01 \pm 0.330}$	$\mathbf{42.27 \pm 0.110}$	$\mathbf{42.91 \pm 0.120}$	$\mathbf{42.89 \pm 0.002}$	$\mathbf{43.11 \pm 0.128}$
AUSGAN (5-slice)	$\underline{34.22 \pm 0.341}$	$\underline{36.38 \pm 0.761}$	$\underline{37.21 \pm 0.109}$	$\underline{38.55 \pm 0.113}$	$\underline{40.84 \pm 0.124}$

Note: Bold indicates best performing method and underline represents the second best. PSNR = Peak Signal-to-Noise Ratio.

Table 11: Effects of Proposed Modules on ISBR Dataset: SSIM Comparison

Method	10%	20%	30%	40%	50%
AUSGAN – Image Gen.	0.81 ± 0.331	0.84 ± 0.161	0.85 ± 0.651	0.86 ± 0.211	0.87 ± 0.005
AUSGAN – K-space Gen.	0.40 ± 0.841	0.42 ± 0.651	0.45 ± 0.281	0.47 ± 0.321	0.49 ± 0.651
AUSGAN (1-slice)	0.70 ± 0.334	0.72 ± 0.512	0.74 ± 0.213	0.75 ± 0.430	0.75 ± 0.332
AUSGAN (3-slice)	$\mathbf{0.97 \pm 0.004}$	$\mathbf{0.97 \pm 0.002}$	$\mathbf{0.98 \pm 0.013}$	$\mathbf{0.99 \pm 0.001}$	$\mathbf{0.99 \pm 0.012}$
AUSGAN (5-slice)	$\underline{0.83 \pm 0.001}$	$\underline{0.86 \pm 0.841}$	$\underline{0.87 \pm 0.341}$	$\underline{0.88 \pm 0.761}$	$\underline{0.90 \pm 0.654}$

Note: Bold indicates best performing method while underline represents the second best. SSIM = Structural Similarity Index Metric.

Table 12: Comparison of Proposed Modules on the ISBR Brain Dataset Tissue-Edge Contrast Across Five Undersampling Rates

Method	10%		20%		30%		40%		50%	
	GM-WM	GM-CSF	GM-WM	GM-CSF	GM-WM	GM-CSF	GM-WM	GM-CSF	GM-WM	GM-CSF
AUSGAN	$\mathbf{0.128 \pm 0.049}$	$\mathbf{0.382 \pm 0.137}$	$\mathbf{0.136 \pm 0.078}$	$\mathbf{0.405 \pm 0.204}$	$\mathbf{0.135 \pm 0.056}$	$\mathbf{0.399 \pm 0.148}$	$\mathbf{0.133 \pm 0.056}$	$\mathbf{0.389 \pm 0.135}$	$\mathbf{0.131 \pm 0.078}$	$\mathbf{0.385 \pm 0.112}$
AUSGAN – Attn.	0.120 ± 0.027	0.340 ± 0.022	0.113 ± 0.034	0.345 ± 0.116	0.117 ± 0.224	0.351 ± 0.187	0.120 ± 0.041	0.348 ± 0.016	0.115 ± 0.072	0.360 ± 0.171
AUSGAN – Spec.	$\underline{0.123 \pm 0.072}$	$\underline{0.360 \pm 0.159}$	$\underline{0.127 \pm 0.080}$	$\underline{0.363 \pm 0.187}$	$\underline{0.125 \pm 0.011}$	$\underline{0.361 \pm 0.445}$	$\underline{0.124 \pm 0.078}$	$\underline{0.366 \pm 0.154}$	$\underline{0.130 \pm 0.081}$	$\underline{0.373 \pm 0.186}$
AUSGAN – Perc.	0.112 ± 0.340	0.349 ± 0.231	0.121 ± 0.210	0.353 ± 0.101	0.117 ± 0.038	0.351 ± 0.381	0.118 ± 0.144	0.354 ± 0.222	0.120 ± 0.080	0.358 ± 0.172

Note: Bold = best, underline = second best. GM = Gray Matter, WM = White Matter, CSF = Cerebrospinal Fluid.