Draft version March 28, 2026
Typeset using LaTeX **twocolumn** style in AASTeX7.0.1

ScattNet-MR: A Wavelet Scattering Transform-Based Deep Residual Network for Accelerated MRI Reconstruction

Diego Villalba-González,[1] Boris Escalante-Ramírez,[2] and Jimena Olveres[2]

[1] *Facultad de Ciencias, Universidad Nacional Autónoma de México, Ciudad de México, Mexico*
[2] *Facultad de Ingeniería, Universidad Nacional Autónoma de México, Ciudad de México, Mexico*

Submitted to Computers in Biology and Medicine

ABSTRACT

Magnetic resonance imaging (MRI) is a cornerstone of diagnostic medicine, yet its inherently long acquisition times limit clinical throughput and increase sensitivity to motion and noise. While deep learning has enabled accelerated MRI reconstruction, many existing approaches rely on black-box architectures that lack interpretability and may exhibit unstable behavior under distribution shifts, particularly in the presence of acquisition noise. In this work, we introduce **ScattNet-MR**, an interpretable deep residual network that integrates the **Wavelet Scattering Transform (WST)** to incorporate analytically defined multiscale priors into the reconstruction process. The model jointly processes zero-filled aliased images and their corresponding fixed scattering coefficients, leveraging WST features as stable, orientation-sensitive descriptors that act as implicit spectral regularizers.

We evaluate ScattNet-MR on the public *fastMRI* brain dataset under $\times 2$ and $\times 4$ variable-density Cartesian undersampling. Robustness is assessed by progressively injecting Gaussian noise in k-space up to $\sigma_k/\mathrm{STD}_k = 1.0$, and performance is quantified using PSNR, SSIM, and L1 error. The proposed method demonstrates smooth and consistent performance degradation as noise increases, maintaining competitive reconstruction quality relative to strong convolutional baselines. In the severe noise regime, ScattNet-MR achieves PSNR values of 18.90 ± 1.87 dB and 19.19 ± 1.88 dB, SSIM values of 0.698 ± 0.008 and 0.690 ± 0.012, and mean L1 errors of 0.184 ± 0.040 and 0.178 ± 0.037 for $\times 2$ and $\times 4$ acceleration, respectively.

These results indicate that integrating fixed, mathematically grounded scattering representations within trainable neural networks yields a transparent and robust reconstruction framework. By embedding stable multiscale structure into the model, ScattNet-MR mitigates sensitivity to noise and distribution shifts, offering a promising direction toward clinically reliable and interpretable deep learning methods for accelerated MRI.

Keywords: Magnetic Resonance Imaging (MRI), Deep Learning, Wavelet Scattering Transform (WST), Image Reconstruction, Acceleration, Interpretability, Robustness, Medical Imaging

1. INTRODUCTION

Magnetic Resonance Imaging (MRI) is a cornerstone of modern medical diagnostics, providing high-resolution, non-invasive visualization of anatomy and physiology across diverse applications such as neurology, cardiology, and oncology Bushong & Clarke (2014). Unlike modalities that use ionizing radiation, MRI relies on nuclear magnetic resonance, making it exceptionally safe for repeated examinations and functional studies. However, its intrinsically long acquisition times remain a fundamental limitation: prolonged scans increase motion artifacts, reduce patient comfort, and raise operational costs—factors that hinder accessibility, particularly in resource-limited healthcare systems elax Organisation for Economic Co-operation and Development (2023).

Over the past decades, substantial research has focused on reducing MRI scan time while maintaining diagnostic image quality. Classical acceleration techniques such as parallel imaging (e.g., SENSE, GRAPPA) and compressed sensing (CS) exploit coil sensitivities and signal sparsity, respectively. Although these approaches

have achieved significant reductions in acquisition time, their performance degrades at high acceleration factors or under noise and motion conditions Rani et al. (2024); Bushong & Clarke (2014).

1.1. *Acceleration Principles in MRI*

MRI acquisition differs fundamentally from direct imaging modalities such as X-ray, as spatial information is encoded in the frequency domain known as k-space. Each point in k-space represents a specific spatial frequency, and the complete image is reconstructed through an inverse Fourier transform. Importantly, most of the useful information in k-space is concentrated in its low-frequency center, while high-frequency components encode fine structural details Sprawls (2000). This property underlies the development of acceleration strategies that aim to reduce the number of acquired samples while preserving diagnostic quality.

In practice, acceleration can be achieved by:

- **Parallel-sampling** the k-space using multiple receiver coils simultaneously, increasing data acquisition efficiency.
- **Subsampling** the k-space, deliberately omitting certain lines to shorten acquisition time.

These strategies exploit the redundancy and sparsity inherent to MRI signals. Reconstruction algorithms such as parallel imaging and compressed sensing subsequently recover the missing information by enforcing coil sensitivity or sparsity priors, thereby suppressing aliasing artifacts. Nevertheless, such methods often struggle to generalize across anatomical regions and acquisition protocols.

1.2. *Deep Learning-Based MRI Reconstruction*

The emergence of deep learning (DL) has transformed MRI reconstruction by enabling models to learn complex, non-linear mappings from undersampled k-space data to high-fidelity images, frequently surpassing classical CS-based techniques in both reconstruction speed and accuracy Singh et al. (2023); Rani et al. (2024).

DL-based approaches can be broadly grouped into three paradigms:

1. **Loss-driven convolutional networks:** These optimize pixel-wise or perceptual losses (e.g., L1/L2, SSIM, or perceptual VGG-based terms) to enhance image fidelity and perceptual realism Kusakunniran et al. (2021); Zhang et al. (2022).
2. **Adversarial networks:** Generative adversarial networks (GANs) promote visual realism by training discriminators to distinguish true from synthetic reconstructions. However, they can produce *hallucinated* structures—anatomically plausible but false details—and exhibit training instability Quan et al. (2021); Oh et al. (2022).
3. **Physics-informed and probabilistic methods:** These integrate acquisition models or uncertainty estimation, including plug-and-play priors, Bayesian formulations, and diffusion-based generative models Shen et al. (2024); Zhu et al. (2018); Chen et al. (2022).

Despite remarkable quantitative performance, several limitations persist that hinder clinical adoption: (1) a lack of interpretability in feature extraction and decision mechanisms, (2) difficulty in quantifying reconstruction uncertainty, and (3) sensitivity to noise, sampling pattern, and anatomical variability. These gaps motivate the development of architectures that are not only accurate but also robust and physically interpretable.

1.3. *Evaluation Challenges and Clinical Reliability*

A major issue in validating DL-based reconstructions lies in the discrepancy between conventional image quality metrics and clinical perceptual relevance. The most widely used measures—**Peak Signal-to-Noise Ratio (PSNR)** and **Structural Similarity Index (SSIM)**—quantify pixel fidelity but may not correlate with diagnostic usefulness Wang et al. (2004); Hore & Ziou (2010). Reconstructions with high PSNR/SSIM values can appear overly smooth, lacking subtle textural and contrast cues vital for pathology detection.

Furthermore, DL models can inadvertently generate false yet plausible features. Bhadra et al. (2021) formalized the problem of **hallucinations** in tomographic reconstruction, showing that data-driven methods can produce anatomically inconsistent details that are invisible to the acquisition system. Such errors raise legitimate concerns about clinical trust and safety.

Consequently, the community has emphasized complementary evaluation strategies that go beyond classical fidelity metrics:

- **Perceptual quality metrics** that align more closely with radiologist visual assessment.
- **No-reference metrics** for settings lacking fully sampled ground truth data.
- **Task-based evaluations** that measure diagnostic consistency rather than pixel accuracy.
- **Uncertainty estimation** to identify regions where model predictions are unreliable.

These considerations underscore the need for reconstruction frameworks that combine high image fidelity with explicit mechanisms for interpretability and robustness.

1.4. *Wavelet Scattering Transform as a Statistical Prior*

The **Wavelet Scattering Transform (WST)** provides a mathematically principled means of extracting hierarchical, translation-invariant, and deformation-stable signal representations Mallat (2012). Constructed as a cascade of wavelet convolutions, modulus nonlinearities, and averaging operations, the WST yields multiscale features that resemble those of CNNs but with fixed, analytically defined filters. This structure ensures stability to small geometric deformations and noise while preserving essential spectral statistics Andén & Mallat (2014); Angles & Mallat (2018); Cheng et al. (2024).

In recent years, WST has been increasingly adopted as a statistical prior in physics-based and generative modeling tasks, owing to its ability to capture long-range correlations with modest computational cost. For MRI reconstruction, this framework offers an appealing route toward integrating interpretable, mathematically grounded priors into otherwise opaque deep models.

1.5. *Motivation and Contributions of this Work*

This study introduces **ScattNet-MR**, a dual-input deep residual network that integrates the interpretability of the Wavelet Scattering Transform with the expressivity of data-driven learning. The model processes two parallel representations of the same image: (1) a zero-filled reconstruction from undersampled k-space, and (2) its corresponding WST coefficients. These features are fused via Squeeze-and-Excitation residual blocks to predict a residual correction that restores fine structural detail while suppressing artifacts.

Our central hypothesis is that embedding WST-based statistical priors within a trainable network improves robustness and interpretability without the instability associated with adversarial training. ScattNet-MR explicitly aligns the multiscale scattering statistics of reconstructed and ground-truth images, yielding reconstructions that are both perceptually coherent and statistically stable under noise.

In summary, the key contributions of this work are as follows:

1. We propose a hybrid **wavelet scattering–guided residual network** that leverages analytically defined multiscale priors for accelerated MRI reconstruction.
2. We demonstrate that WST regularization improves robustness and statistical consistency across undersampling factors and acquisition noise levels.
3. We validate the model on the *FastMRI* benchmark, showing competitive fidelity relative to strong CNN baselines while offering enhanced interpretability and deterministic behavior compared to adversarial approaches.

For quantitative evaluation, image fidelity is reported using **PSNR** and **SSIM**, whose mathematical definitions are provided in Appendix.

2. MATERIALS AND METHODS

This section details the data, computational setup, and methodological design used to develop and evaluate the proposed **ScattNet-MR** framework. All experiments were performed under reproducible conditions to ensure transparency and facilitate future research.

2.1. *Dataset and Preprocessing*

All experiments were conducted using the publicly available **FastMRI** dataset developed by NYU Langone Health in collaboration with Meta AI (Zbontar et al. 2018). The dataset provides raw k-space measurements and corresponding fully sampled reference reconstructions for multiple anatomies, including brain, knee, prostate, and breast MRI acquired under clinical conditions. In this work, we focus exclusively on the **multi-coil brain MRI subset**. We follow the *official FastMRI data partitioning*, which consists of **5193 training slices**, **1592 validation slices**, and **2956 test slices**. Model training and hyperparameter selection were performed using the training split, with early stopping and checkpoint selection guided by performance on the validation split. All quantitative metrics reported in this study (PSNR, SSIM, MSE, stability, and uncertainty) are computed on the **held-out validation set**, which is completely unseen during parameter optimization. The official FastMRI test set remains hidden and is reserved exclusively for leaderboard evaluation. Brain volumes include T1-weighted, T2-weighted, and Proton Density (PD) contrasts, with and without fat suppression. Acquisitions were performed using multi-coil Cartesian sampling, reconstructed into matrices of 320×320 or 640×320 pixels. To simulate accelerated acquisitions, we applied **1D variable-density Cartesian undersampling** along the phase-encoding direction using the official FastMRI masks implemented in the `fastmri` Python library (Facebook AI Research and NYU Langone Health 2020). With the same mask applied across slices of each volume to maintain spatial coherence. The central low-frequency region of k-space

was fully retained to preserve contrast, corresponding to approximately **4% of total lines for** $\times 2$ and **2% for** $\times 4$. Missing data were **zero-filled** prior to inverse Fourier reconstruction.

Multi-coil data were combined using a root-sum-of-squares (RSS) method, and magnitudes were normalized to the [0,1] range. Each undersampled image was paired with its fully sampled reference. Afterward, all images were **center-cropped** to remove acquisition padding and then **upscaled to** 256×256 **pixels via bicubic interpolation** for consistency across modalities and acceleration factors.

The FastMRI dataset is publicly available for non-commercial research at https://fastmri.org/ under a CC BY-NC 4.0 license.

2.2. *Experimental Setup*

The model was implemented in **Python** using the **PyTorch** framework (Paszke et al. 2019). Wavelet scattering features were computed with **Kymatio** (Andreux et al. 2020), and visualizations were produced using **Matplotlib** (Hunter 2007). Training was performed on a workstation equipped with an **NVIDIA RTX 4090 GPU** (24 GB VRAM). The network was optimized with the **Adam** optimizer (learning rate 5×10^{-4}), trained for 200 epochs per acceleration factor, with a batch size of 8. Data augmentation consisted of random flips and rotations. Each model converged after approximately 14 hours of training. All code, preprocessing scripts, and trained weights will be released publicly upon publication. The final ScattNet-MR configuration comprises approximately 3.4 million trainable parameters, placing it in a moderately lightweight regime suitable for deployment on standard GPU hardware.

2.3. *Training Objective*

To connect learning with a physically interpretable formulation, we follow an energy-based perspective where an MR image $x \in \mathcal{X}$ is treated as a sample from a Gibbs distribution parameterized by network weights θ:

$$p_\theta(x) = \frac{1}{Z_\theta} \exp\big(-U_\theta(x)\big), \tag{1}$$

with $U_\theta(x)$ representing an energy function and Z_θ the partition function. Instead of matching pixel-level distributions directly, we minimize divergence in a feature space defined by the **Wavelet Scattering Transform (WST)** (Cheng et al. 2024). This formulation constrains the model to learn reconstructions that preserve the intrinsic statistics of real MRI data, enhancing stability and interpretability.

The final training objective combines pixel-wise fidelity, perceptual similarity, and statistical consistency into a composite loss:

$$\begin{aligned}\mathcal{L}_{\text{total}} = \lambda_1 \|\hat{x} - x\|_1 \\ + \lambda_2 \|\phi_{\text{VGG}}(\hat{x}) - \phi_{\text{VGG}}(x)\|_1 \\ + \lambda_3 D_{\text{KL}}\big(P_S(\hat{x}) \,\|\, P_S(x)\big).\end{aligned} \tag{2}$$

where $\hat{x}$ and x are reconstructed and reference images, $\phi_{\text{VGG}}(\cdot)$ denotes high-level features from a pretrained VGG network (Johnson et al. 2016), and $P_S(\cdot)$ represents the empirical distribution of WST coefficients. The KL term encourages statistical alignment between the scattering representations of generated and reference images, acting as a **stability prior** that improves robustness to noise and undersampling artifacts. The experiments used $\lambda_1 = 0.5$, $\lambda_2 = 0.5$, $\lambda_3 = 0.5$.

2.4. *Model Architecture: ScattNet-MR*

The proposed **ScattNet-MR** operates as a residual dual-branch network (Figures 1 and 2). One branch processes the zero-filled image, extracting local spatial features, while the second branch computes and projects **WST coefficients** (S_0, S_1, S_2) to a compatible latent space. These representations are fused through $N = 10$ **Squeeze-and-Excitation (SE)** residual blocks (Hu et al. 2018) that reweight feature channels adaptively and facilitate residual learning.

The network predicts a residual correction Δx, which is added to the interpolated input to form the high-resolution output:

$$x_{HR} = \text{Interp}(x_{LR}) + \Delta x.$$

This formulation directs learning toward the reconstruction of missing high-frequency structures rather than reproducing the entire image, improving both efficiency and perceptual realism.

The architecture integrates analytical (WST) and learned (CNN) feature pathways, providing a form of **structural interpretability**: the model's learned mapping is constrained by mathematically defined multiscale priors that promote stability and explainable behavior.

3. RESULTS

This section presents the experimental evaluation of the proposed **ScattNet-MR** model on the FastMRI dataset. Results are organized as follows: (1) training convergence, (2) quantitative fidelity (SSIM/PSNR) on the held-out test set, (3) representative qualitative reconstructions, (4) stability under additive k-space noise, (5) uncertainty estimation via Monte Carlo dropout, and (6) comparison with established baselines and selected state-of-the-art methods.

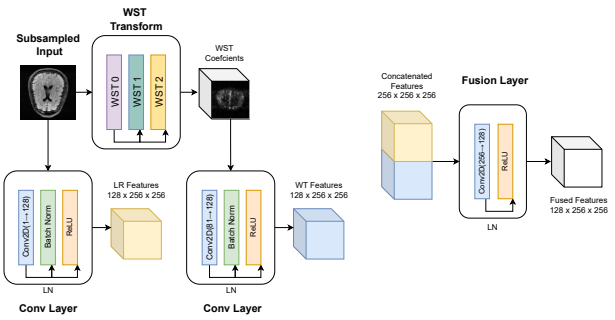

Figure 1. Dual-branch feature extraction in **ScattNet-MR**. The left branch processes the zero-filled input, while the right extracts multiscale WST coefficients (S_0, S_1, S_2) prior to projection and fusion.

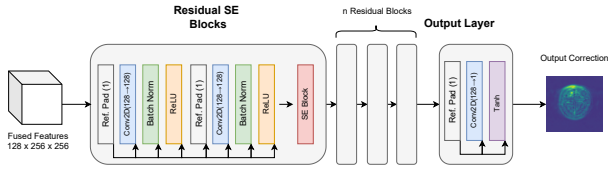

Figure 2. Residual feature fusion and reconstruction stage of **ScattNet-MR**. Concatenated features are passed through SE-Residual blocks to produce the residual correction Δx.

Unless otherwise stated, quantitative results are reported on the **held-out FastMRI test set**, consisting of volumes not used for model selection or training. The **training set** was used to optimize network parameters, while the **validation set** was used for checkpoint selection and hyperparameter tuning. The test set therefore contains **completely unseen images** for all reported evaluations.

3.1. *Training and Validation Loss Analysis*

To evaluate the convergence behavior of our proposed **ScattNet-MR** model, we monitored the evolution of training and validation loss for both ×2 and ×4 acceleration factors. Figure 3 shows the complete loss history, highlighting the stable decrease in both training and validation loss over 200 epochs. The validation curves remain close to the training curves throughout training, showing no pronounced divergence consistent with severe overfitting under this training regime.

3.2. *Quantitative Results*

To evaluate reconstruction fidelity, we compute **SSIM** and **PSNR** between reconstructions and fully sampled references, and report results for both the zero-filled inputs and ScattNet-MR outputs on the held-out test set.

Table 1 summarizes the descriptive statistics for the ×4 and ×2 acceleration scenarios. We report mean, standard deviation, minimum, and maximum values for each metric.

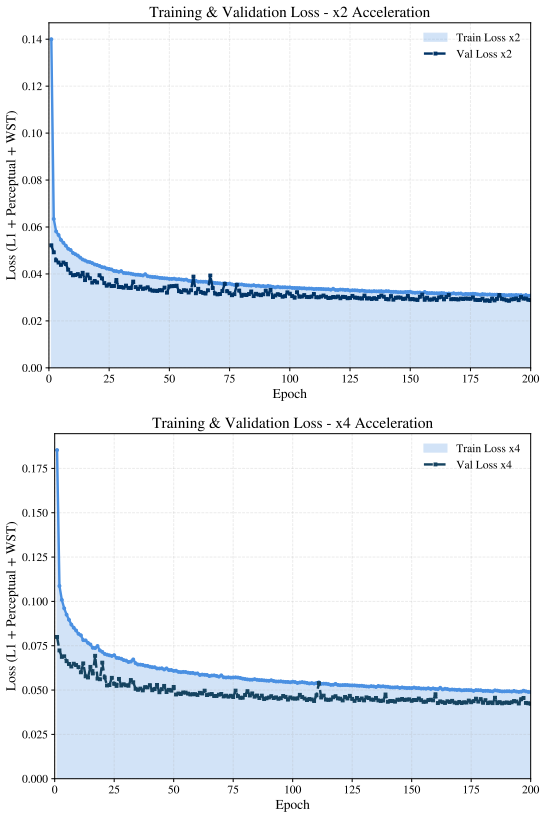

Figure 3. Training and validation loss for *ScattNet-MR* under ×2 (left) and ×4 (right) acceleration. Shaded regions indicate the training loss, while dashed lines correspond to the validation loss.

Table 1. SSIM and PSNR (mean ± std) for ×2 and ×4 accelerations .

Accel.	Input	ScattNet-MR
×2	0.859±0.043 / 29.7±2.6	**0.945±0.019** / **35.8±2.2**
×4	0.694±0.063 / 23.4±2.2	**0.894±0.030** / **31.5±2.0**

3.3. *Qualitative Evaluation*

Representative reconstructions from the *FastMRI* brain dataset under ×2 and ×4 acceleration are shown in Fig. 4. Each panel displays the zero-filled input, the ScattNet-MR reconstruction, and the fully sampled reference, together with the SSIM difference map (1 − SSIM). ScattNet-MR effectively suppresses aliasing artifacts and restores fine anatomical details, particularly in high-frequency white-matter regions. These examples illustrate the model's stable perceptual performance across noise levels and sampling factors.

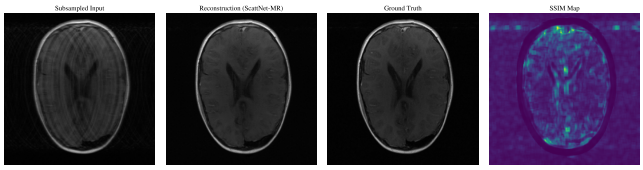

Figure 4. Qualitative visual comparison for the `AXT1` sequence. SSIM = **0.9408**. The reconstruction closely matches the ground truth with minimal structural differences.

3.4. *Stability Under Acquisition Noise*

To assess robustness against acquisition noise, we simulate perturbations directly in the k-space domain. For each slice, complex-valued Gaussian noise is added to the measured k-space coefficients, with magnitude controlled by the normalized noise ratio $\sigma_k/\mathrm{STD}_k \in \{0,\ 0.1,\ 0.5,\ 1.0\}$, where STD_k denotes the standard deviation of the original k-space for the corresponding scan. The corrupted k-space is then inverse Fourier-transformed to obtain a noisy image, which is subsequently processed by the trained reconstruction models.

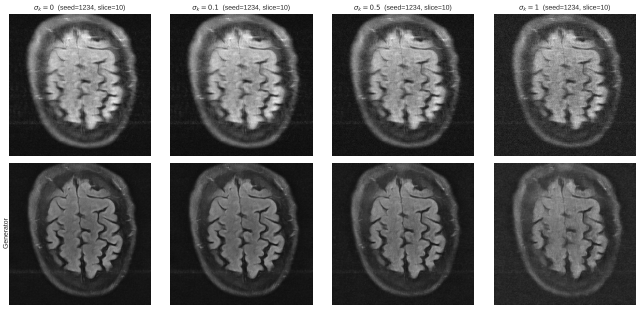

Figure 5. Qualitative effect of increasing k-space noise on reconstructions at $\times 2$ acceleration. Columns correspond to $\sigma_k/\mathrm{STD}_k \in \{0, 0.1, 0.5, 1.0\}$.

The stability analysis is conducted over **all slices from selected near-average technical cases** in the test set for each sampling configuration. Each slice is evaluated under multiple independent noise realizations to account for stochastic variability. Reported metrics correspond to mean±standard deviation aggregated across slices and noise draws. In total, eight reconstructions are computed per metric–configuration pair, and representative qualitative examples are shown in Figures 5 and 6.

Quantitative stability results.—Figure 7 reports the evolution of SSIM as a function of increasing acquisition noise for acceleration factors $\times 2$ and $\times 4$, comparing all evaluated reconstruction architectures. Across all methods and acceleration factors, SSIM decreases monotonically with increasing noise level, with larger dispersion observed at $\times 4$ acceleration.

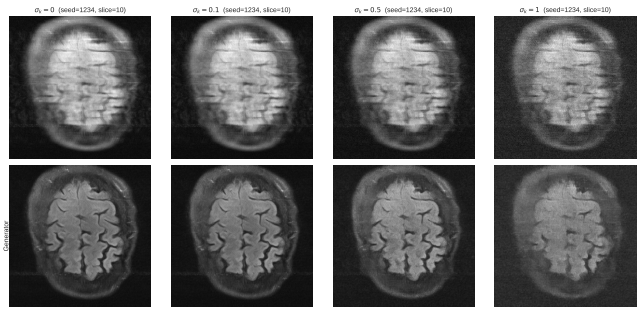

Figure 6. Qualitative effect of increasing k-space noise on reconstructions at $\times 4$ acceleration. Columns correspond to $\sigma_k/\mathrm{STD}_k \in \{0, 0.1, 0.5, 1.0\}$.

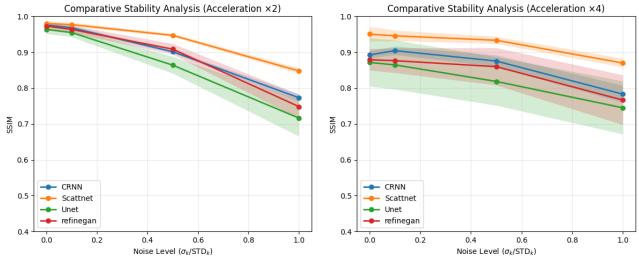

Figure 7. Structural Similarity (SSIM) degradation trajectories under progressively increasing Gaussian k-space noise levels $\sigma_k/\mathrm{STD}_k \in \{0, 0.1, 0.5, 1.0\}$ for acceleration factors $\times 2$ (left) and $\times 4$ (right). Shaded bands correspond to ± 1 standard deviation computed across slices and independent noise realizations. The curves characterize the relative robustness and stability of each reconstruction model by illustrating their performance sensitivity and degradation behavior under acquisition noise perturbations.

Tables 2 and 3 summarize endpoint stability by reporting SSIM values under clean conditions and under severe noise ($\sigma_k/\mathrm{STD}_k = 1.0$).

Table 2. Stability benchmarks at $\times 2$ acceleration (SSIM mean±std) under clean inputs and severe noise ($\sigma_k/\mathrm{STD}_k = 1.0$).

Technique	**SSIM (Clean)**	**SSIM** ($\sigma = 1.0$)
ScattNet-MR	$\mathbf{0.9801 \pm 0.0062}$	$\mathbf{0.8478 \pm 0.0076}$
Unet	0.9636 ± 0.0126	0.7164 ± 0.0511
refinegan	0.9728 ± 0.0077	0.7483 ± 0.0334
CRNN	0.9760 ± 0.0033	0.7730 ± 0.0103

Effect of removing WST features.—For the ScattNet-MR architecture, suppressing the WST input leads to reduced reconstruction quality as noise increases. Representative qualitative examples are shown in Figs. 8 and 9.

Table 3. Stability benchmarks at ×4 acceleration (SSIM mean±std) under clean inputs and severe noise ($\sigma_k/\mathrm{STD}_k = 1.0$).

Technique	SSIM (Clean)	SSIM ($\sigma = 1.0$)
ScattNet-MR	$\mathbf{0.9505 \pm 0.0208}$	$\mathbf{0.8698 \pm 0.0122}$
Unet	0.8714 ± 0.0671	0.7450 ± 0.0738
refinegan	0.8790 ± 0.0298	0.7666 ± 0.0692
CRNN	0.8926 ± 0.0098	0.8230 ± 0.0237

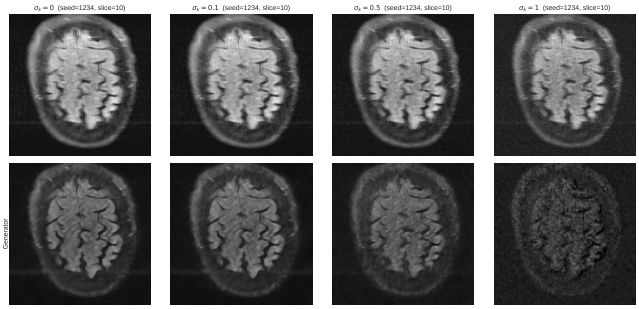

Figure 8. Reconstruction behavior without WST input at ×2 acceleration across increasing noise levels.

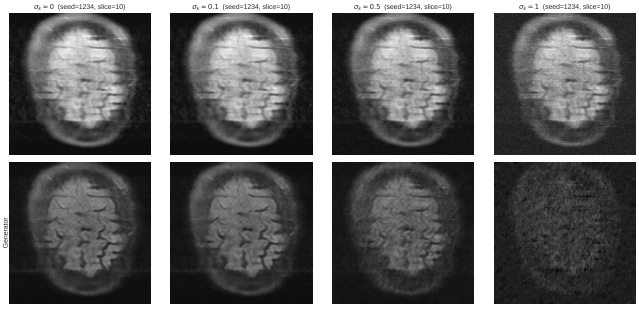

Figure 9. Reconstruction behavior without WST input at ×4 acceleration across increasing noise levels.

3.5. *Model Uncertainty Estimation via Monte Carlo Dropout*

To evaluate not only reconstruction accuracy but also predictive confidence, we employ a **Monte Carlo Dropout (MCD)** strategy following Nair et al. (2018). Dropout layers are activated at inference time and each input is processed through 100 stochastic forward passes, yielding a distribution of reconstructions from which both a mean prediction and a pixel-wise uncertainty map (standard deviation) are computed. This procedure enables a direct assessment of reconstruction variability under identical acquisition conditions.

For each reconstruction model, we report three complementary metrics: (i) the L1 error between the MCD mean reconstruction and the ground truth, (ii) the average uncertainty magnitude, and (iii) the correlation between uncertainty and reconstruction error. Tables 4 and 5 summarize these quantities averaged across representative test slices.

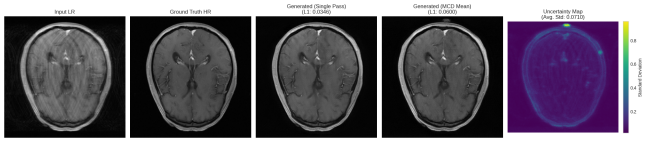

Figure 10. Example of uncertainty-aware reconstruction at ×2 acceleration using Monte Carlo Dropout. From left to right: low-resolution input, ground-truth high-resolution image, MCD mean reconstruction, and pixel-wise uncertainty map (standard deviation across 100 stochastic passes). The displayed metrics correspond to the illustrated slice.

Table 4. Monte Carlo Dropout mean L1 reconstruction error at ×2 acceleration (100 stochastic forward passes).

Model	MCD Mean L1	Std.
ScattNet-MR	**0.033**	**0.004**
U-Net	0.037	0.006
CRNN	0.059	0.010
RefineGAN	0.096	0.018

Table 5. Monte Carlo Dropout uncertainty and error metrics at ×2 acceleration.

Model	Avg. Uncertainty	Corr. (Unc., Error)
ScattNet-MR	$\mathbf{0.021 \pm 0.006}$	0.61 ± 0.07
CRNN	0.071 ± 0.014	0.65 ± 0.06
U-Net	0.094 ± 0.022	0.60 ± 0.08
RefineGAN	0.118 ± 0.031	0.81 ± 0.05

All models exhibit positive correlation between uncertainty magnitude and reconstruction error, as reflected by consistently positive correlation values. Among the compared methods, **ScattNet-MR achieves the lowest MCD mean L1 error and the lowest average uncertainty**, indicating that its predictions are both accurate and consistent across stochastic realizations. Other architectures exhibit larger reconstruction error and/or higher uncertainty magnitude, despite in some cases achieving comparable single-pass reconstruction quality.

3.6. *Computational performance*

In addition to predictive performance, we also evaluated the computational aspects of the proposed framework. The generator network comprises a total of **3,367,376 trainable parameters**, situating it in a moderately lightweight regime compared to larger med-

ical imaging models. This parameter count provides a good trade-off between representational capacity and computational efficiency. Inference experiments further

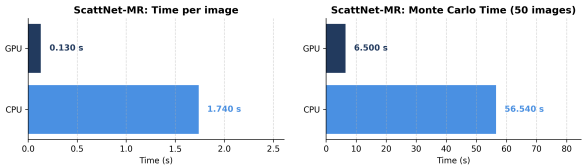

Figure 11. Inference time comparison of ScattNet-MR on GPU vs CPU for single-pass (Left) and Monte Carlo Dropout (100 passes) (Right).

confirm the model's practicality. As shown on figure 11, on a single GPU, the average inference time per image is **0.130 seconds**, while CPU execution requires **1.740 seconds**. For Monte Carlo Dropout inference with 100 stochastic forward passes, the GPU achieves a total time of **6.500 seconds**, compared to **56.540 seconds** on CPU. These results demonstrate that ScattNet-MR enables rapid reconstruction and uncertainty quantification, making it well suited for real-time or near-real-time clinical applications, where timely feedback on both reconstructions and confidence maps is crucial.

3.7. *Comparison with State-of-the-Art Methods*

Following the comparison protocol commonly reported for FastMRI reconstruction studies (e.g., Xia et al. **?**), we contextualize ScattNet-MR against representative deep learning baselines and selected state-of-the-art methods on the *FastMRI* brain dataset (Fig. 12). The comparison set includes SSIM-optimized U-Net Hyun et al. (2018), CRNN Qin et al. (2019), and GAN-based approaches such as RefineGAN Quan et al. (2021), DR-CAM-GAN **?**, and ESSGAN Zhou et al. (2021). These models span convolutional, recurrent, and adversarial training paradigms, providing a diverse benchmark for assessing fidelity under accelerated acquisition.

Fair comparison protocol.—For consistency with our preprocessing and evaluation pipeline, we retrained U-Net, CRNN, and RefineGAN using the same data split, undersampling masks, and normalization procedures as ScattNet-MR. Results for additional methods (e.g., DR-CAM-GAN and ESSGAN) are reported from the corresponding publications when retraining was not performed in this study.

As shown in Fig. 12, ScattNet-MR yields substantial gains over the zero-filled inputs for both acceleration factors. At ×2 acceleration, the average SSIM improves from 0.859 to 0.945 and PSNR from 29.7 dB to 35.8 dB.

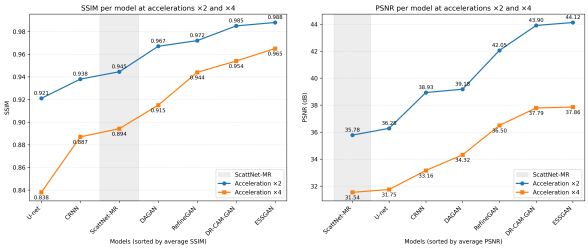

Figure 12. SSIM and PSNR comparison of ScattNet-MR against representative baselines and selected state-of-the-art methods on the FastMRI brain dataset for ×2 and ×4 acceleration.

At ×4 acceleration, ScattNet-MR achieves 0.894 SSIM and 31.5 dB PSNR on average, remaining competitive under more aggressive undersampling.

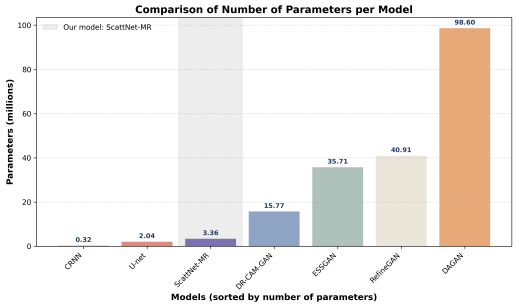

Figure 13. Model size comparison (trainable parameters) between ScattNet-MR and the compared reconstruction architectures.

Figure 13 further shows that ScattNet-MR operates in a moderately lightweight regime, with a parameter count comparable to established CNN-based baselines. Together with the inference benchmarks in Section 3.6, this supports practical deployment in resource-constrained clinical settings where low latency and on-premise execution are preferred.

3.8. *Experimental Conclusions*

- ScattNet-MR improves SSIM/PSNR relative to zero-filled inputs for both ×2 and ×4 acceleration on the held-out test set (Table 1).
- Under additive k-space noise, SSIM decreases monotonically across all methods, with performance summarized in Fig. 7 and Tables 2–3.
- Monte Carlo dropout analysis indicates that ScattNet-MR yields the lowest mean L1 error and lowest average uncertainty among the evaluated architectures (Tables 4–5).

- The proposed model supports fast single-pass inference on GPU and feasible uncertainty estimation via MCD (Section 3.6).

Overall, these results indicate that ScattNet-MR supports accurate reconstruction and uncertainty estimation under accelerated acquisition settings.

4. DISCUSSION

The results presented in Section 3 demonstrate that the proposed **ScattNet-MR** framework achieves competitive performance for accelerated magnetic resonance imaging (MRI) reconstruction under Gaussian undersampling. Across both $\times 2$ and $\times 4$ acceleration factors, ScattNet-MR consistently improves over zero-filled inputs and classical CNN-based baselines, while remaining slightly below the peak quantitative performance reported by recent GAN-based approaches in terms of raw fidelity metrics such as SSIM and PSNR (Fig. 12).

Importantly, these differences in peak metric values must be interpreted in the broader context of robustness, stability, and clinical reliability. GAN-based reconstruction methods are known to exhibit several limitations that are particularly relevant for accelerated MRI, including training instability, higher computational cost at inference, and increased sensitivity to distribution shift due to the narrow data manifold enforced by the discriminator of "Medical Imaging Complexity & its Effects on GAN" (2024); Feng et al. (2024). At high acceleration factors, these issues may manifest as noise amplification, the introduction of **hallucinated structures**, and reduced interpretability of the reconstruction process—phenomena that raise significant concerns for clinical deployment (Bhadra et al. 2021; Singh et al. 2023).

In contrast, ScattNet-MR offers a more balanced trade-off between reconstruction fidelity, robustness, and interpretability. The observed performance gains relative to conventional CNNs can be attributed to the explicit integration of **Wavelet Scattering Transform (WST)** coefficients as fixed, multiscale statistical priors. By embedding WST features within a residual SE–ResNet generator, the model enforces **statistical consistency** between reconstructed and ground-truth images, guiding the learning process toward physically meaningful representations rather than unconstrained texture synthesis. The ablation results presented in Section 3.4 corroborate this interpretation: suppressing the WST branch leads to pronounced degradation in PSNR and SSIM under acquisition noise, confirming that scattering coefficients encode structurally relevant information essential for stable MRI reconstruction.

Stability, robustness, and hallucination risk.—A key empirical observation emerging from the stability experiments is that, although ScattNet-MR attains lower absolute SSIM values than some competing methods in the most severe noise regimes, its degradation trajectory remains smooth, monotonic, and highly consistent across noise levels and acceleration factors. By contrast, architectures that do not incorporate wavelet scattering features exhibit sharper transitions between moderate and high noise conditions, particularly at $\times 4$ acceleration.

This behavior suggests that ScattNet-MR responds to increasing acquisition noise primarily through a gradual loss of fine detail rather than abrupt structural changes. Such a response is consistent with the theoretical stability properties of wavelet scattering representations and is desirable in safety-critical reconstruction settings. From a practical perspective, this form of controlled degradation may reduce the likelihood of generating visually plausible but unsupported anatomical structures when the model is applied to previously unseen or heavily corrupted inputs.

The Monte Carlo Dropout analysis further supports this conclusion. ScattNet-MR exhibits the lowest mean reconstruction error and the lowest average predictive uncertainty among the evaluated architectures, while maintaining meaningful alignment between uncertainty and reconstruction error. Together, these findings indicate that the model not only produces accurate reconstructions, but also expresses calibrated confidence under uncertainty; an essential property for clinical trust.

Overall, these results highlight a fundamental trade-off between peak reconstruction fidelity and stability under acquisition noise. While purely data-driven architectures may achieve higher apparent quality under specific conditions, hybrid models incorporating analytically grounded priors such as WST favor more predictable, conservative, and physically grounded behavior under distribution shift.

4.1. *Broader Implications*

The implications of this work extend beyond the FastMRI benchmark:

1. **Clinical relevance:** Accelerated MRI acquisition with controlled and predictable degradation can reduce patient discomfort, mitigate motion artifacts, and enable higher-throughput clinical workflows. These benefits are particularly important in pediatric imaging and longitudinal studies of neurodegenerative diseases, where repeated scans are often required.

2. **Interpretability and trustworthiness:** ScattNet-MR exemplifies a hybrid modeling paradigm aligned with the growing demand for **clinically interpretable AI**. By incorporating WST features grounded in multiscale signal statisticsthe model's behavior can be partially interpreted in terms of energy distributions across spatial scales, offering greater transparency than purely latent CNN representations.

3. **Generality across domains:** Although this study focuses on brain MRI, the proposed framework is naturally extensible to other imaging modalities, including cardiac and dynamic MRI. More broadly, the underlying principles are applicable to inverse problems involving structured spatial fields, such as cosmological field reconstruction, which conceptually motivated aspects of the model design.

4.2. *Limitations and Future Directions*

Despite the promising results, several limitations should be acknowledged.

First, training and evaluation were restricted to the FastMRI dataset. While this benchmark provides a standardized comparison, validation on multi-center and multi-vendor datasets is necessary to assess robustness across different scanners, protocols, and patient populations.

Second, although the hybrid loss formulation promotes statistical fidelity, achieving explicit **probabilistic interpretability** remains an open challenge. Future work could integrate WST-based priors with diffusion or flow-based generative models to provide more principled uncertainty estimates.

Finally, the current implementation assumes static Cartesian undersampling. Extending the framework to **non-Cartesian trajectories** or **adaptive acquisition strategies** may further improve achievable acceleration while preserving reconstruction quality.

Future research should therefore focus on three complementary directions: (1) combining wavelet scattering representations with uncertainty-aware generative models; (2) extending the architecture to multi-contrast and dynamic MRI settings; and (3) conducting prospective clinical validation studies to assess diagnostic equivalence in real-world scenarios.

4.3. *Concluding Perspective*

Within the broader landscape of AI-driven MRI acceleration, ScattNet-MR represents a **hybrid, physics-informed paradigm** that integrates deep residual learning with interpretable multiscale statistics. Rather than maximizing peak fidelity at all costs, the framework emphasizes stability, robustness, and transparency—properties that are essential for safe clinical deployment. This work therefore contributes a principled pathway toward explainable and trustworthy accelerated MRI reconstruction.

ACKNOWLEDGMENTS

This work was supported by the Universidad Nacional Autónoma de México (UNAM) through the PAPIIT grants IN108624 and IT101624.

REFERENCES

elax Organisation for Economic Co-operation and Development. 2023, Health at a Glance 2023: OECD Indicators, OECD Publishing, doi: 10.1787/4dd50c09-en

Andén, J., & Mallat, S. 2014, IEEE Transactions on Signal Processing, 62, 4114, doi: 10.1109/TSP.2014.2326991

Andreux, M., Lostanlen, V., Andén, J., & Mallat, S. 2020, Journal of Machine Learning Research, 21, 1

Angles, T., & Mallat, S. 2018, in Proceedings of the International Conference on Learning Representations (ICLR)

Bhadra, S., Kelkar, V. A., Brooks, F. J., & Anastasio, M. A. 2021, IEEE Transactions on Medical Imaging, 40, 3249, doi: 10.1109/TMI.2021.3077857

Bushong, S. C., & Clarke, G. D. 2014, Magnetic Resonance Imaging: Physical and Biological Principles, 4th edn. (St. Louis, MO: Elsevier Mosby)

Chen, Y., Yang, G., & Wang, S. 2022, Magnetic Resonance in Medicine, 87, 3161, doi: 10.1002/mrm.29188

Cheng, S., Morel, R., Allys, E., Ménard, B., & Mallat, S. 2024, PNAS Nexus, 1, doi: 10.1093/pnasnexus/pgae103

Facebook AI Research and NYU Langone Health. 2020, fastmri: PyTorch library for the fastMRI dataset and models, https://pypi.org/project/fastmri/

Feng, Y., et al. 2024, arXiv preprint

Hore, A., & Ziou, D. 2010, in 20th International Conference on Pattern Recognition, 2366–2369, doi: 10.1109/ICPR.2010.579

Hu, J., Shen, L., & Sun, G. 2018, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 7132–7141, doi: 10.1109/CVPR.2018.00745

Hunter, J. D. 2007, Computing in Science & Engineering, 9, 90, doi: 10.1109/MCSE.2007.55

Hyun, C. M., Kim, H. P., Lee, S. M., Lee, S., & Seo, J. K. 2018, Physics in Medicine & Biology, 63, 135007, doi: 10.1088/1361-6560/aac71a

Johnson, J., Alahi, A., & Fei-Fei, L. 2016, in European Conference on Computer Vision (ECCV), 694–711, doi: 10.1007/978-3-319-46475-6_43

Kusakunniran, W., Liu, M., & Wu, S. 2021, Magnetic Resonance Imaging, 80, 10, doi: 10.1016/j.mri.2021.03.002

Mallat, S. 2012, Communications on Pure and Applied Mathematics, 65, 1331, doi: 10.1002/cpa.21413

Nair, T., Precup, D., Arnold, D. L., & Arbel, T. 2018, arXiv preprint

of "Medical Imaging Complexity, A., & its Effects on GAN". 2024, arXiv preprint

Oh, G.-B., Lim, H., & Kim, H.-I. 2022, IEEE Transactions on Medical Imaging, 41, 375, doi: 10.1109/TMI.2021.3098564

Paszke, A., Gross, S., Massa, F., et al. 2019, in Advances in Neural Information Processing Systems 32 (NeurIPS 2019)

Qin, C., Schlemper, J., Caballero, J., et al. 2019, IEEE Transactions on Medical Imaging, 38, 280, doi: 10.1109/TMI.2018.2863670

Quan, T., Chang, Y., & Huang, C. 2021, Medical Image Analysis, 68, 101907, doi: 10.1016/j.media.2020.101907

Rani, B., Singh, P. K., et al. 2024, Magnetic Resonance Imaging

Shen, G., Li, M., Farris, C. W., Anderson, S., & Zhang, X. 2024, Scientific Reports, 14, 21877, doi: 10.1038/s41598-024-72820-2

Singh, A., Yadav, S., Sharma, N., Pachori, R. B., & Acharya, U. R. 2023, Bioengineering, 10, 1012, doi: 10.3390/bioengineering10101012

Sprawls, P. 2000, Magnetic Resonance Imaging: Principles, Methods, and Techniques (Madison, WI: Medical Physics Publishing)

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. 2004, IEEE Transactions on Image Processing, 13, 600, doi: 10.1109/TIP.2003.819861

Zbontar, J., Knoll, F., Sriram, A., et al. 2018, arXiv preprint arXiv:1811.08839

Zhang, C., Wang, Y., & Chen, Y. 2022, Medical Image Analysis, 76, 102301, doi: 10.1016/j.media.2021.102301

Zhou, W., Du, H., Mei, W., & Fang, L. 2021, Neurocomputing, 422, 51, doi: 10.1016/j.neucom.2020.09.008

Zhu, B., Liu, J. Z., Cauley, S. F., Rosen, B. R., & Rosen, M. S. 2018, Nature, 555, 487, doi: 10.1038/nature25988