FilterDiff: Noise-free Frequency-domain Diffusion Models for Accelerated MRI Reconstruction

Tao Song[1,2], Fang Nie[3], Yi Guo[1], Feng Xu[1], and Shaoting Zhang[2]

[1]Fudan University, Shanghai, China
[2]Sensetime Research, Shanghai, China
[3]Department of Radiology, Zhongda Hospital, Nanjing, China
tsong22@m.fudan.edu.cn

Abstract. Accelerated MRI reconstruction has garnered increasing attention due to its significant clinical value. Recently, the exceptional capabilities of diffusion models in image generation have led to their widespread application in accelerated MRI reconstruction. However, the inherent noisy diffusion process in these models introduces uncertainty during the reverse diffusion restoration, which can compromise the consistency of the results. Moreover, adding Gaussian noise contradicts the actual MRI imaging process. To address these issues, we propose FilterDiff, a noise-free frequency-domain diffusion framework. In FilterDiff, the diffusion process is modeled as a filtering operation, similar to the MRI acquisition process, thereby eliminating the dependence on noise and simplifying the diffusion procedure. To better capture frequency-domain long-range information, we proposed a Swin-DiTs network, which modifies the DiT transformer network by replacing the self-attention mechanism with Swin-attention to reduce computational cost, and removing the position embedding to mitigate feature artifacts. Extensive experiments on two public datasets demonstrate that our model achieves state-of-the-art performance in accelerated MRI reconstruction, both in in-distribution and out-of-distribution scenarios.

Keywords: MRI Reconstruction · Noise-free Diffusion · Filter

1 Introduction

Accelerated MRI reconstruction has become a important issue due to its significant clinical application value. Recent years have witnessed considerable advancements in data-driven methods [20,11,13,12,6,10,15,4,23,5,21]. Deep neural network models have proven effective in learning the distribution transformation between under-sampled and fully-sampled images (or k-space). Compared to traditional techniques, these models not only enhance the quality of the reconstructed images but also provide the added benefit of enabling real-time imaging.

Diffusion models, known for their high-fidelity generation capabilities, have garnered widespread application in accelerated MRI reconstruction in recent

years. Peng et al. [18] introduced an unconditional diffusion model trained to generate coil-combined MRI images derived from fully-sampled data. Gungor et al. [7] proposed a conditional diffusion model for accelerated MRI reconstruction based on adaptive diffusion priors. To exploit the invariance of the sampled frequency region in k-space during accelerated sampling, [3,22] present a conditional diffusion model that focuses on the diffusion of non-sampled frequencies. In [16], these authors developed a diffusion process in the frequency domain by employing a circular mask-based filter and introducing noise outside the filtered region, thereby combining two degradation mechanisms: filtering and noise addition. Additionally, [19] proposed a bridge diffusion model, where the diffusion process initiates from a fully-sampled image and terminates at an under-sampled image, with Gaussian noise added exclusively to the intermediate state images.

Although diffusion models in accelerated MRI reconstruction have started incorporating acquisition priors and imaging knowledge, they still often rely on degradation methods that involve adding noise to construct the diffusion process. These methods conflict with the actual MRI imaging procedure. Moreover, noise-based diffusion introduces uncertainty during the inverse sampling process, which leads to variability in the generated results. In contrast, medical image generation requires a highly precise and deterministic process, distinct from the diversity sought in natural image generation. Recently, ColdDiff [2], a general degradation diffusion model, was proposed in the natural image domain. This model allows the construction of the diffusion process to move beyond the sole reliance on noise addition, incorporating other forms of degradation such as blurring, snow and pixelation, etc. In the context of accelerated MRI acquisition in Cartesian coordinates, the acquisition prioritizes dense sampling of the center in the phase-encode direction while sparsely sampling the peripheral non-central regions to preserve significant information. Therefore, the variation in the size of the sampling width in the phase-encode direction (also referred to as the filter in the phase-encode direction) can be regarded as a form of frequency-domain degradation. Although the method proposed by Huang et al. [9] incorporates the sampling mask during the forward diffusion process, it ultimately degrades to a highly corrupted k-space rather than an under-sampled one. Moreover, both the reverse diffusion sampling and the network training in their approach are conducted in the image space, rather than directly in the k-space domain.

Inspired by the MRI acquisition process, we propose a noise-free diffusion model in the frequency domain, named **FilterDiff**. Unlike traditional diffusion models, which rely on noise addition, FilterDiff employs a deterministic diffusion process, ensuring that no uncertainty factors arise during the reverse process. As a result, the images generated through multiple iterations of reverse diffusion remain consistent. Specifically, we define the initial state of the forward process as the fully-sampled k-space data, while the terminal state is the under-sampled k-space data obtained through central sampling using a mask filter. To mitigate the loss of information from non-central regions, the original under-sampled k-space data serves as conditional input. The intermediate states consist of under-sampled k-space data filtered with varying widths of the central sampling mask.

During the reverse diffusion process, the initial under-sampled k-space data, filtered with the central mask, is progressively predicted outward in a step-by-step manner, ultimately restoring it to fully sampled k-space data.

To further enhance the reconstruction capability of FilterDiff, we modify a pure transformer Swin-DiTs network based on DiT [17], replacing the self-attention mechanism with Swin-attention [14] and removing the position embedding to reduce feature artifacts. This modified architecture serves as our restoration network. Extensive experiments on publicly available datasets, such as fastMRI [25] and IXI [1], demonstrate that our model achieves state-of-the-art performance in both in-distribution and out-of-distribution scenarios.

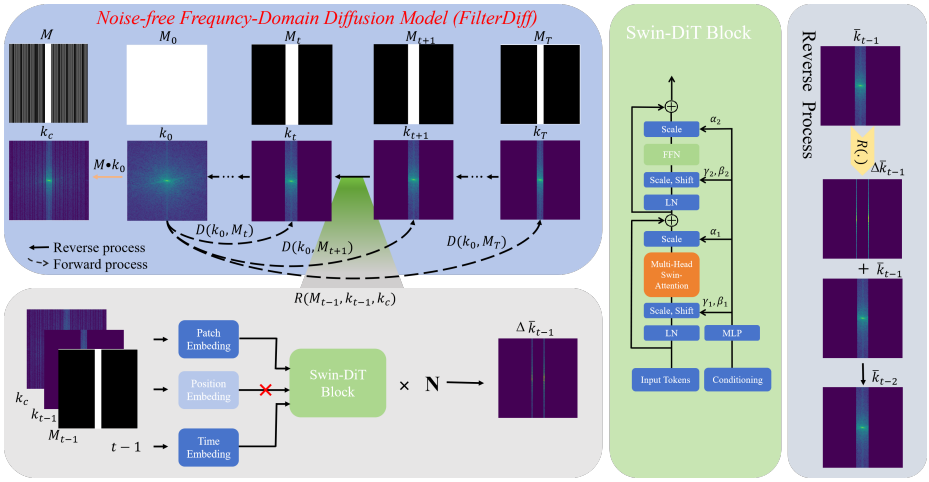

Fig. 1. Overview of the proposed FilterDiff for accelerated MRI reconstruction.

2 Background

2.1 Cold Diffusion

The Cold Diffusion Model is a generalized diffusion model [2], which extends the traditional noise-based diffusion to any type of degradation, such as blurring, snow and pixelation, etc. Specifically, given an image $\boldsymbol{x}_0$ from the training data distribution Q, the image $\boldsymbol{x}_0$ is gradually degraded using a custom degradation operator $\mathcal{D}(\cdot)$ into an image $\boldsymbol{x}_T$ sampled from a random initial distribution P, where T is the total number of diffusion time steps. The intermediate state image $\boldsymbol{x}_t$ in the diffusion process is then defined as $\boldsymbol{x}_t = \mathcal{D}(\boldsymbol{x}_0, \boldsymbol{x}_T, t)$. In accelerated MRI reconstruction, adding noise is the most commonly used degradation operator. If a noise-adding diffusion process is employed, the forward process in the Cold Diffusion Model can be defined as:

$$\boldsymbol{x}_t = \mathcal{D}(\boldsymbol{x}_0, \boldsymbol{x}_T, t) = \sqrt{\alpha_t}\boldsymbol{x}_0 + \sqrt{(1-\alpha_t)}\boldsymbol{x}_T, \tag{1}$$

where $\boldsymbol{x}_T$ is random noise with specific distribution. In the reverse process, we need a restoration operator $\mathcal{R}(\cdot)$ to approximate the inverse of the degradation operator $\mathcal{D}(\cdot)$. This restoration operator can be expressed as:

$$\overline{x}_0 = \mathcal{R}(\boldsymbol{x}_t, t) \approx \boldsymbol{x}_0, \tag{2}$$

where $\mathcal{R}(\cdot)$ is a neural network parameterized by θ, which can be optimized by the following objective function:

$$\min_{\boldsymbol{\theta}} \mathbb{E}_{\boldsymbol{x}_0 \sim Q, \boldsymbol{x}_T \sim P} \|\mathcal{R}_{\boldsymbol{\theta}}(\mathcal{D}(\boldsymbol{x}_0, \boldsymbol{x}_T, t), t) - \boldsymbol{x}_0\|. \tag{3}$$

After obtaining the trained restoration operator $\mathcal{R}(\cdot)$, the restoration process in its inverse diffusion process is given by the following formula:

$$\boldsymbol{x}_{t-1} = \boldsymbol{x}_t - \mathcal{D}(\overline{\boldsymbol{x}}_0, \boldsymbol{x}_T, t) + \mathcal{D}(\overline{\boldsymbol{x}}_0, \boldsymbol{x}_T, t-1), \tag{4}$$

where $\overline{\boldsymbol{x}}_T = \left(\boldsymbol{x}_t - \sqrt{\alpha_t}\overline{\boldsymbol{x}}_0\right) / \sqrt{(1-\alpha_t)}$. Although ColdDiff[2] is a generalized diffusion method in the image domain, the exploration of frequency-domain degradation operators in accelerated MRI reconstruction tasks has not been conducted.

2.2 Under-sampled MRI Reconstruction

Typically, under-sampled MRI reconstruction can be formulated as the following equation:

$$\boldsymbol{y}_{\boldsymbol{M}} = \boldsymbol{M}\mathcal{A}\boldsymbol{x} + \boldsymbol{\epsilon}, \tag{5}$$

where $\boldsymbol{x} \in \mathbb{R}^n$ represents an MRI image, $\boldsymbol{y}_{\boldsymbol{M}} \in \mathbb{R}^m$ $(m < n)$ is the under-sampled k-space measurements, $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ is the measuring matrix, $\boldsymbol{M} \in \mathbb{R}^{m \times n}$ is the under-sampling matrix with the specific sampling pattern, and $\boldsymbol{\epsilon}$ is the noise. Typically, the sampling pattern of $\boldsymbol{M}$ is performed in the Cartesian coordinate system, where the sampling ensures dense full-frequency acquisition along the central part of the phase-encode direction (the x-axis) to minimize information loss, denoted as $\boldsymbol{M}_{\text{core}} \in \mathbb{R}^{m \times n}$. The non-central part of the phase-encode direction is sampled with sparse full-frequency acquisition (usually using equidistant or random sampling) and is denoted as $\boldsymbol{M}_{\text{other}} \in \mathbb{R}^{m \times n}$. The elements of matrix $\boldsymbol{M}$ are either 0 or 1, where 0 represents positions that are not sampled, and 1 represents positions that are sampled. Therefore, $\boldsymbol{M} = \boldsymbol{M}_{\text{core}} + \boldsymbol{M}_{\text{other}}$.

3 Method: FilterDiff

Previous diffusion-based accelerated MRI reconstruction methods [3,22,19] typically describe the diffusion process as the addition of Gaussian noise and use the under-sampled image (or k-space) as a condition for predicting the corresponding fully-sampled image (or k-space). However, it is important to note that the statistical characteristics of noise in MRI images (or k-space) are complex and

Algorithm 1 Training for FilterDiff

Input: Paired fully-sampled/under-sampled k-space sets $I = \{(\boldsymbol{k}_0, \boldsymbol{k}_T)_i\}_{i=1}^{N}$, total time steps T
Output: Trained Restoration Net $\mathcal{R}_\theta$
Initialization: Randomly initializes Restoration Net $\mathcal{R}_\theta$
repeat
  Sample Paired k-space data $(\boldsymbol{k}_0, \boldsymbol{k}_T) \sim I$
  Sample Timestep $t \sim Uniform(1, ..., T)$
  Calculate $\boldsymbol{k}_t$ by Eq. (6)
  $\Delta k_t \leftarrow \mathcal{R}_\theta(\boldsymbol{M_t}, \boldsymbol{k}_t, \boldsymbol{k}_c, t)$
  Update θ by Eq. (8)
until converged

cannot be simply modeled using a Gaussian distribution. Moreover, performing diffusion based on noise deviates from the actual physical process of MRI acquisition and reconstruction.

Instead, the variation in the sampling matrix $\boldsymbol{M}$ of k-space inherently introduces a degradation process, which can be interpreted as a form of filtering diffusion in k-space. Based on this observation, we propose **FilterDiff** in k-space. Specifically, we define the fully sampled k-space data as $\boldsymbol{k}_0$ and the under-sampled k-space data as $\boldsymbol{k}_T$. To ensure that the diffusion process accurately mimics the physical degradation process of MRI k-space, we introduce a novel degradation operator $\mathcal{D}(\cdot)$, defined as follows:

$$\boldsymbol{k}_t = \mathcal{D}(\boldsymbol{k}_0, \boldsymbol{M_t}) = \boldsymbol{M_t} * \boldsymbol{k}_0, \tag{6}$$

where $\boldsymbol{M}_t$ is an intermediate state between $\boldsymbol{M}_0$ and $\boldsymbol{M}_T$. Here, $\boldsymbol{M}_0$ represents an identity matrix, and $\boldsymbol{M}_T$ corresponds to $\boldsymbol{M}_{\text{core}}$. This degradation operator simulates degradation in k-space in the form of filtering, and thus, we refer to it as the k-space filter degradation operator. Typically, both Cold Diffusion [2] and classical Gaussian diffusion models [22,3], begin their sampling process from random Gaussian noise, gradually reducing noise levels until generating $\overline{\boldsymbol{x}}_0$(or $\overline{\boldsymbol{k}}_0$). As a result, they require a large number of sampling steps to transform an image with a similar noise level into an under-sampled image (or k-space) that retains meaningful semantic information.

It is important to note that, in our proposed diffusion model, the endpoint of the diffusion process is the under-sampled k-space, which does not consist entirely of pure Gaussian noise nor contain partial Gaussian noise. Therefore, we can directly sample from the under-sampled k-space using a smaller T, while our proposed diffusion model eliminates the randomness introduced by Gaussian noise. Consequently, multiple repeated samplings yield identical and fully reproducible $\overline{\boldsymbol{k}}_0$ values. This characteristic is critical for real-world applications in accelerated MRI reconstruction-based medical imaging diagnostics, ensuring stability and consistency in clinical settings. In the reverse process, we constructed a new restoration operator $\mathcal{R}_\theta(\cdot)$ to approximate the inverse of the degradation

Algorithm 2 Reverse Process for FilterDiff

1: **Input:** under-sampled k-space data sets $(\boldsymbol{k}_c, \boldsymbol{k}_T, \boldsymbol{M}_T)$, where $\boldsymbol{k}_T$ can be computed from $\boldsymbol{k}_c$ and $\boldsymbol{M}_T$
2: **Output:** Generated k-space data $\boldsymbol{k}_0$
3: **Load** the trained Restoration Net $\mathcal{R}_\theta$
4: **for** $t = T, T-1, \ldots, 1$ **do**
5: $\quad \Delta\overline{\boldsymbol{k}}_t \leftarrow \mathcal{R}_\theta(\boldsymbol{M_t}, \overline{\boldsymbol{k}}_t, \boldsymbol{k}_c, t)$
6: $\quad$ **Calculate** $\overline{\boldsymbol{k}}_{t-1}$ by Eq. (9)
7: **end for**

operator $\mathcal{D}(\cdot)$. This restoration operator can be expressed as follows:

$$\Delta\overline{\boldsymbol{k}}_{t-1} = \mathcal{R}_{\boldsymbol{\theta}}(Cond) = \Delta\boldsymbol{M}_{t-1} * FFT(\boldsymbol{\phi_\theta}(Cond)) \approx \Delta\boldsymbol{M}_{t-1} * \boldsymbol{k}_0, \tag{7}$$

where $Cond$ represents $(\boldsymbol{M_{t-1}}, \boldsymbol{k}_{t-1}, \boldsymbol{k}_c, t)$, $\boldsymbol{k}_c$ represents the under-sampled k-space data, FFT represents the Fourier transform, $\boldsymbol{\phi}$ represents a neural network with parameters $\boldsymbol{\theta}$. And $\Delta\overline{\boldsymbol{k}}_{t-1}$ represents the predicted difference between $\boldsymbol{k}_t$ and $\boldsymbol{k}_{t-1}$, and $\Delta\boldsymbol{M}_{t-1} = \boldsymbol{M_{t-1}} - \boldsymbol{M_t}$ represents the difference between $\boldsymbol{M_t}$ and $\boldsymbol{M_{t-1}}$. Thus, the $\mathcal{R}_{\boldsymbol{\theta}}$ can be optimized using the following objective function:

$$\min_{\boldsymbol{\theta}} \mathbb{E}\{\|\mathcal{R}_{\boldsymbol{\theta}}(\boldsymbol{M_t}, \boldsymbol{k}_t, \boldsymbol{k}_c, t) - \Delta\boldsymbol{k}_{t-1}\| + \lambda\|\boldsymbol{\phi_\theta}(\boldsymbol{M_t}, \boldsymbol{k}_t, \boldsymbol{k}_c, t) - \boldsymbol{x}_0\|\}. \tag{8}$$

Once the restoration operator $\mathcal{R}_{\boldsymbol{\theta}(\cdot)}$ is trained, the restoration process initiates the inverse diffusion process, which is given by the following formula:

$$\overline{\boldsymbol{k}}_{t-1} = \overline{\boldsymbol{k}}_t + \mathcal{R}_\theta(\boldsymbol{M_t}, \overline{\boldsymbol{k}}_t, \boldsymbol{k}_c, t) = \overline{\boldsymbol{k}}_t + \Delta\overline{\boldsymbol{k}}_t. \tag{9}$$

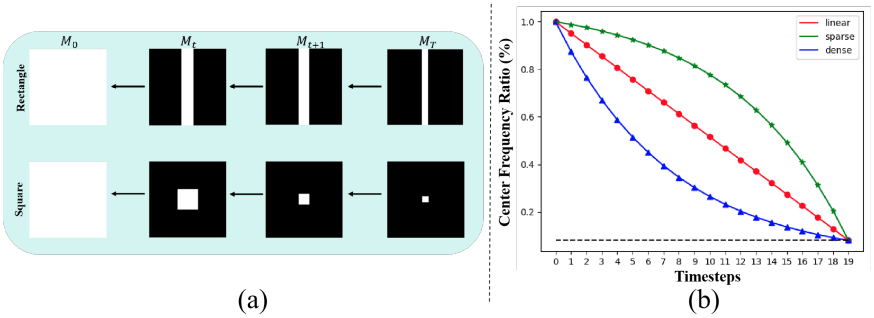

Fig. 2. (a) Different filter shapes and (b) The trend of the filter variation across different time steps.

The Fig. 1 provides a detailed illustration of the restoration process. Finally, the training and restoration procedures are shown in Algs. 1 and 2, respectively. Regarding the network architecture, we modify the original DiT [17] framework

to a Swin-DiTs architecture. Since the original DiT [17] is designed for the latent space, directly applying it to images results in high computational costs due to the self-attention mechanism. To address this issue, we replace self-attention with Swin-attention [14]. Additionally, inspired by [24], where position embedding introduces feature artifacts in transformer-based architectures, we remove position embedding in our design, as illustrated in Fig. 1.

4 Experimental Results

Data and Implementation Details. The experiments were conducted on the public fastMRI [25] and IXI datasets [1]. Fellowing [19] setting, for the fastMRI dataset, 360 individuals from knee dataset were selected for training, 8 for validation, and 20 for testing. Additionally, 20 individuals in brain dataset were used as an out-of-distribution set. And for the IXI dataset, 577 individuals' T1 images were split into training (500 individuals), validation (37 individuals), and test (40 individuals) sets. K-space under-sampled was simulated with equally spaced Cartesian under-sampled in the phase-encode direction, with acceleration factors of 4 and 8. For training, 200k iterations were performed using the AdamW optimizer, a learning rate of 1×10^{-4}, and a batch size of 8. For our FilterDiff, both training and restoration steps were set to 20, with a loss weight $\lambda = 1$. All experiments were conducted using 4 NVIDIA GeForce RTX 4090 GPUs.

Table 1. Comparison of methods in fastMRI and IXI experiments.

		knee(in-distribution)		brain(out-of-distribution)		IXI(in-distribution)	
factor	Method	PSNR(dB)	SSIM(%)	PSNR(dB)	SSIM(%)	PSNR(dB)	SSIM(%)
4×	DDPM[8]	32.78±2.87	86.84±8.06	29.71±1.85	90.38±2.11	32.70±2.77	95.32±2.10
	AdaDiff[7]	28.16±1.94	78.77±5.03	24.93±1.30	81.21±3.07	25.84±1.28	80.09±3.71
	HFS-SDE[3]	30.70±2.10	82.19±5.36	26.03±1.35	79.69±3.58	30.76±2.80	89.22±2.15
	FDB[16]	31.68±2.33	82.99±4.17	27.10±1.41	80.36±3.40	31.89±3.10	90.02±2.08
	MC-DDPM[22]	33.84±3.38	88.24±3.57	29.93±1.76	90.47±2.14	36.20±3.11	97.49±1.31
	CBDM[19]	34.84±2.54	89.73±6.20	30.67±1.71	91.97±1.42	37.56±3.07	98.10±1.02
	Ours	**37.65±3.01**	**93.69±1.22**	**33.68±1.56**	**94.78±1.22**	**40.76±3.77**	**98.38±0.87**
8×	DDPM[8]	30.36±2.79	81.45±9.06	27.47±1.30	86.93±2.23	32.56±2.72	95.02±2.21
	AdaDiff[7]	28.85±2.15	77.53±6.15	25.67±1.37	80.08±3.13	23.74±1.05	75.45±3.87
	HFS-SDE[3]	28.14±2.03	75.52±5.20	24.81±1.39	73.87±4.14	30.47±2.46	92.96±2.27
	FDB[16]	28.93±2.19	76.34±5.30	25.84±1.33	74.06±4.01	31.01±2.55	93.64±2.36
	MC-DDPM[22]	31.30±2.45	84.58±5.84	26.10±2.21	82.52±3.56	35.53±2.96	97.04±1.53
	CBDM[19]	32.80±2.20	85.83±6.97	28.65±1.24	89.02±1.94	37.09±2.96	97.70±1.15
	Ours	**34.20±2.49**	**88.04±5.19**	**31.62±1.65**	**91.80±1.74**	**39.21±3.53**	**97.82±1.03**

Results Analysis. Our FilterDiff is compared with multiple diffusion models in terms of generation quality. As shown in Tab. 1, under equally spaced Cartesian under-sampling with acceleration factors of 4 and 8, our FilterDiff consistently outperforms all other methods both in-distribution and out-of-distribution of the fastMRI dataset and in-distribution of the IXI dataset in terms of PSNR

and SSIM metrics. Compared to the advanced diffusion model CBDM [19], our method achieves an increase of approximately 3.0 dB in PSNR and an increase of approximately 4% in SSIM. These results demonstrate that our method significantly enhances the performance and generalization of advanced noise-based diffusion models.

Table 2. Ablation studies of our FilterDiff.

Filter Shape		Filter Variation			Network	Metrics	
Rectangle	Square	Linear	Sparse	Dense	Swin-DiTs	PSNR(dB)	SSIM(%)
✓		✓			✓	37.53±3.05	93.66±1.25
✓			✓		✓	33.20±2.55	86.78±5.43
✓				✓	✓	**37.65±3.01**	**93.69±1.22**
	✓			✓	✓	37.24±2.96	93.26±3.34
✓				✓	✓	**37.65±3.01**	**93.69±1.22**
✓				✓		37.01±2.43	93.03±2.96
✓				✓	✓	**37.65±3.01**	**93.69±1.22**

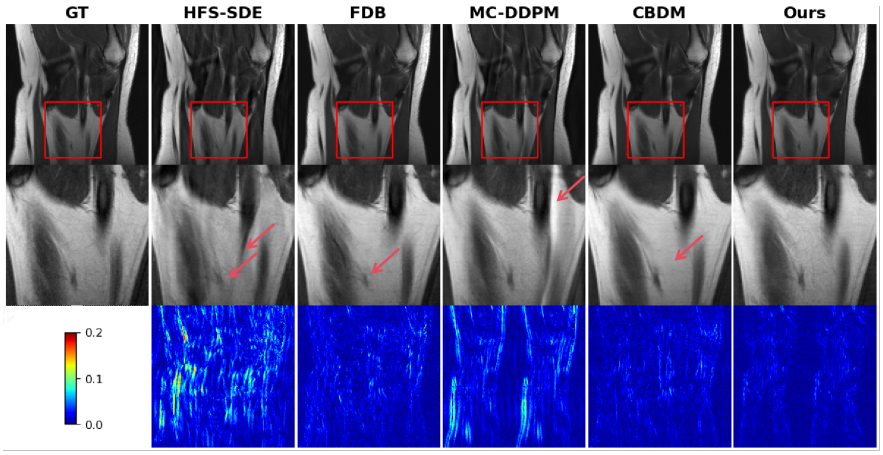

Fig. 3. Accelerated MRI reconstructions with different methods.

For the ablation study, we conducted experiments exclusively on images with 4x acceleration using the equally spaced Cartesian under-sampling method. As illustrated in Tab. 2 and Fig. 2, we first investigated the impact of filter variations across different time steps and filter shapes. The results indicate that the dense mode with a rectangular filter achieves the best performance, suggesting that information in the central region is more critical. Additionally, the rectangular mask outperforms the square mask as a filter, which aligns with the principles of MRI data acquisition. Furthermore, we replaced the Swin-DiTs network with ResUnet [19] from other diffusion models and observed that Swin-DiTs provides

an improvement of approximately 0.6 dB in reconstruction performance. This demonstrates that Swin-DiTs effectively captures long-range relationships in the frequency domain. Finally, we visualized the reconstruction effects of different methods, as shown in Fig. 3. In terms of specific reconstruction details, our FilterDiff surpasses other diffusion models.

5 Conclusion

In this paper, we propose FilterDiff, a novel noise-free frequency-domain diffusion model designed for accelerated MRI reconstruction. Unlike traditional diffusion models that rely on stochastic noise addition, FilterDiff employs a deterministic diffusion process, ensuring consistent and reliable image generation through multiple reverse diffusion iterations. By defining the forward process as a transition from fully-sampled to under-sampled k-space data using a central mask filter, the diffusion process becomes independent of noise and better aligns with the MRI acquisition process. Furthermore, we introduce a modified Swin-DiTs network, which replaces self-attention with Swin-attention and removes position embedding to reduce feature artifacts, significantly enhancing reconstruction performance. Extensive experiments on the fastMRI and IXI datasets demonstrate that FilterDiff achieves state-of-the-art results in both in-distribution and out-of-distribution scenarios, highlighting its superior performance and generalization capability.

6 Disclosure of Interest

The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Biomedical image analysis group imperial college london centre for the developing brain king's college london. (2018), https://brain-development.org/ixi-dataset, accessed December 15, 2021
2. Bansal, A., Borgnia, E., Chu, H.M., Li, J., Kazemi, H., Huang, F., Goldblum, M., Geiping, J., Goldstein, T.: Cold diffusion: Inverting arbitrary image transforms without noise. In: NeurIPS 2024. vol. 36 (2024)
3. Cao, C., Cui, Z.X., Wang, Y., Liu, S., Chen, T., Zheng, H., Liang, D., Zhu, Y.: High-frequency space diffusion model for accelerated mri. IEEE Transactions on Medical Imaging (2024)
4. Chen, Y., Firmin, D., Yang, G.: Wavelet improved gan for mri reconstruction. In: Medical Imaging 2021: Physics of Medical Imaging. vol. 11595, pp. 285–295 (2021)
5. Cole, E.K., Pauly, J.M., Vasanawala, S.S., Ong, F.: Unsupervised mri reconstruction with generative adversarial networks. arXiv preprint arXiv:2008.13065 (2020)
6. Fabian, Z., Heckel, R., Soltanolkotabi, M.: Data augmentation for deep learning based accelerated mri reconstruction with limited data. In: ICML 2021. pp. 3057–3067 (2021)

7. Güngör, A., Dar, S.U., Öztürk, Ş., Korkmaz, Y., Bedel, H.A., Elmas, G., Ozbey, M., Çukur, T.: Adaptive diffusion priors for accelerated mri reconstruction. Medical Image Analysis p. 102872 (2023)
8. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. In: NeurIPS 2020. vol. 33, pp. 6840–6851 (2020)
9. Huang, J., Aviles-Rivero, A.I., Schönlieb, C.B., Yang, G.: Cdiffmr: Can we replace the gaussian noise with k-space undersampling for fast mri? In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 3–12. Springer (2023)
10. Huo, Y., Xu, Z., Bao, S., Assad, A., Abramson, R.G., Landman, B.A.: Adversarial synthesis learning enables segmentation without target modality ground truth. In: ISBI 2018. pp. 1217–1220 (2018)
11. Hyun, C.M., Kim, H.P., Lee, S.M., Lee, S., Seo, J.K.: Deep learning for undersampled mri reconstruction. Physics in Medicine & Biology **63**(13), 135007 (2018)
12. Jeelani, H., Martin, J., Vasquez, F., Salerno, M., Weller, D.S.: Image quality affects deep learning reconstruction of mri. In: ISBI 2018. pp. 357–360 (2018)
13. Knoll, F., Hammernik, K., Zhang, C., Moeller, S., Pock, T., Sodickson, D.K., Akcakaya, M.: Deep-learning methods for parallel magnetic resonance imaging reconstruction: A survey of the current approaches, trends, and issues. IEEE Signal Processing Magazine **37**(1), 128–140 (2020)
14. Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., Guo, B.: Swin transformer: Hierarchical vision transformer using shifted windows. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 10012–10022 (2021)
15. Lv, J., Wang, C., Yang, G.: Pic-gan: a parallel imaging coupled generative adversarial network for accelerated multi-channel mri reconstruction. Diagnostics **11**(1), 61 (2021)
16. Mirza, M.U., Dalmaz, O., Bedel, H.A., Elmas, G., Korkmaz, Y., Gungor, A., Dar, S.U., Çukur, T.: Learning fourier-constrained diffusion bridges for mri reconstruction. arXiv preprint arXiv:2308.01096 (2023)
17. Peebles, W., Xie, S.: Scalable diffusion models with transformers. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 4195–4205 (2023)
18. Peng, C., Guo, P., Zhou, S.K., Patel, V.M., Chellappa, R.: Towards performant and reliable undersampled mr reconstruction via diffusion model sampling. In: Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 623–633. Springer (2022)
19. Song, T., Wu, Y., Hu, M., Luo, X., Luo, G., Wang, G., Guo, Y., Xu, F., Zhang, S.: Cycle-consistent bridge diffusion model for accelerated mri reconstruction. arXiv preprint arXiv:2412.09998 (2024)
20. Sriram, A., Zbontar, J., Murrell, T., Zitnick, C.L., Defazio, A., Sodickson, D.K.: Grappanet: Combining parallel imaging with deep learning for multi-coil mri reconstruction. In: CVPR 2020 (2020)
21. Wu, Y., Song, T., Wu, Z., Ge, Z., Chen, Z., Cai, J.: Codebrain: Impute any brain mri via instance-specific scalar-quantized codes. arXiv preprint arXiv:2501.18328 (2025)
22. Xie, Y., Li, Q.: Measurement-conditioned denoising diffusion probabilistic model for under-sampled medical image reconstruction. In: Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 655–664. Springer (2022)

23. Yang, G., Lv, J., Chen, Y., Huang, J., Zhu, J.: Generative adversarial networks (gan) powered fast magnetic resonance imaging–mini review, comparison and perspectives. arXiv preprint arXiv:2105.01800 (2021)
24. Yang, J., Luo, K.Z., Li, J., Deng, C., Guibas, L., Krishnan, D., Weinberger, K.Q., Tian, Y., Wang, Y.: Denoising vision transformers. In: European Conference on Computer Vision. pp. 453–469. Springer (2024)
25. Zbontar, J., Knoll, F., Sriram, A., Murrell, T., Huang, Z., Muckley, M.J., Defazio, A., Stern, R., Johnson, P., Bruno, M., et al.: fastmri: An open dataset and benchmarks for accelerated mri. arXiv preprint arXiv:1811.08839 (2018)