

## PINN-DADif: Physics-informed deep adaptive diffusion network for robust and efficient MRI reconstruction

Shahzad Ahmed<sup>a</sup>, Jinchao Feng<sup>a,\*</sup>, Atif Mehmood<sup>b</sup>, Muhammad Usman Ali<sup>c</sup>, Muhammad Yaqub<sup>d</sup>, Malik Abdul Manan<sup>a</sup>, Abdul Raheem<sup>e</sup>

<sup>a</sup> Beijing Key Laboratory of Computational Intelligence and Intelligent System, Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China

<sup>b</sup> Department of Computer Science and Technology, Zhejiang Normal University, Jinhua, Zhejiang 321004, China

<sup>c</sup> Department of Computer Science, COMSATS University Islamabad Sahiwal Campus Sahiwal 57000 Pakistan

<sup>d</sup> School of Biomedical Science, Hunan University, Changsha, China

<sup>e</sup> Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China



### ARTICLE INFO

**Keywords:**

MRI reconstruction  
Adaptive diffusion  
Physics-informed neural networks (PINN)  
Undersampled k-space  
Domain shift

### ABSTRACT

Magnetic Resonance Imaging (MRI) is essential for high-resolution soft-tissue imaging but suffers from long acquisition times, limiting its clinical efficiency. Accelerating MRI through undersampling k-space data leads to ill-posed inverse problems, introducing noise and artifacts that degrade image quality. Conventional deep learning models, including conditional and unconditional approaches, often face challenges in generalization, particularly with variations in imaging operators or domain shifts. In this study, we propose PINN-DADif, a Physics-Informed Neural Network integrated with deep adaptive diffusion priors, to address these challenges in MRI reconstruction. PINN-DADif employs a two-phase inference strategy: an initial rapid-diffusion phase for fast preliminary reconstructions, followed by an adaptive phase where the diffusion prior is refined to ensure consistency with MRI physics and data fidelity. The inclusion of physics-based regularization through PINNs enhances the model's adherence to k-space constraints and gradient smoothness, leading to more accurate reconstructions. This adaptive approach reduces the number of iterations required compared to traditional diffusion models, improving both speed and image quality. We validated PINN-DADif on a private MRI dataset and the public fastMRI dataset, where it outperformed state-of-the-art methods. The model achieved PSNR values of 41.2, 39.5, and 41.5, and SSIM values of 98.7, 98.0, and 98.5 for T1, T2, and Proton Density-weighted images at  $R = 4x$  on the private dataset. Similar high performance was observed on the fastMRI dataset, even in scenarios involving domain shifts. PINN-DADif marks a significant advancement in MRI reconstruction by providing an efficient, adaptive, and physics-informed solution.

### 1. Introduction

Magnetic Resonance Imaging (MRI) is a widely recognized, non-invasive imaging modality that offers exceptional contrast for soft tissues, making it indispensable in diagnostic applications. Despite its advantages, one of the primary limitations of MRI is the extended duration of examinations, which often limits its practicality in time-sensitive clinical environments. To mitigate this, techniques such as undersampling of k-space acquisitions have been employed, transforming the problem of image reconstruction into an ill-posed inverse problem that requires sophisticated solutions to recover high-quality images from

sparse data [1,2].

Recent advancements in deep learning have led to a surge in the development of data-driven methods for MRI reconstruction [3–5]. These techniques excel at solving complex inverse problems by learning rich data priors, allowing for efficient reconstruction of high-quality images from under-sampled data [6]. Among these, conditional models have become prevalent, wherein undersampled acquisitions are processed to generate output images that match fully-sampled references. These models rely on explicit mappings learned from paired datasets of undersampled and fully-sampled MRI images but often require vast amounts of training data. To reduce this dependency,

\* Corresponding author at: Beijing University of Technology, Beijing 100124, China.

E-mail address: [fengjc@bjut.edu.cn](mailto:fengjc@bjut.edu.cn) (J. Feng).

alternative approaches such as self-supervision [7,8] or cycle-consistency [9,10] have been explored. These strategies implicitly learn the mapping even when direct access to fully-sampled acquisitions is unavailable.

A critical challenge for conditional models is their reliance on fixed priors that are closely tied to the imaging operator (e.g., coil sensitivities, sampling patterns), limiting their ability to generalize to unseen conditions during testing [11–13]. Variations in the imaging operator—such as changes in the acceleration rate, sampling density, or number of coils—often necessitate model retraining to maintain performance. Thus, there is a growing demand for reconstruction models that can adapt to changes in the imaging operator and generalize across different clinical settings.

Unconditional models, in contrast, offer an alternative by learning generative image priors that are independent of the imaging operator [14]. These models have demonstrated robustness against domain shifts by focusing on auxiliary tasks like denoising, image generation, and encoding. However, unconditional models using fixed priors, such as adversarial priors, may suffer from reduced sample diversity. Diffusion models [15,16] have emerged as a promising solution by maintaining both high-quality and diverse image samples through iterative diffusion processes. Nevertheless, the static nature of their priors can still limit their adaptability when confronted with variations in the imaging environment, such as changes in pulse sequences or scanner settings.

In this study, we introduce the PINN-DADif, a novel approach that addresses both the accuracy and adaptability challenges in MRI reconstruction. PINN-DADif incorporates a deep adaptive diffusion prior, dynamically adjusting it during inference to maintain high performance across varying imaging conditions, including domain shifts between training and testing. This framework integrates physics-informed neural networks (PINNs) to enforce consistency with physical laws governing MRI data acquisition, making it more robust to changes in imaging

operators. Unlike traditional diffusion models that require many small inference steps, PINN-DADif accelerates the process by employing larger diffusion steps through an adaptive mapping mechanism, significantly reducing the number of iterations required for high-quality reconstructions. The inference process is divided into two phases: the rapid-diffusion phase, which quickly generates an initial reconstruction using learned priors, and the adaptation phase, where the diffusion prior is adjusted based on the specific characteristics of the input data, ensuring greater alignment with the physical constraints and improving data consistency. Fig. 1 shows the architecture of the proposed network.

The integration of physics-based priors within the PINN-DADif framework represents a significant advancement, enabling the model to respect the physical laws governing MRI signal formation while maintaining flexibility across different scanning protocols. The performance of PINN-DADif was validated across multiple MRI contrasts using a unified model trained on a mixed-contrast dataset. The results demonstrate that PINN-DADif consistently outperforms state-of-the-art traditional, conditional, and unconditional models, particularly in challenging cross-domain scenarios where the imaging operator or MRI data distribution changes. By merging the strengths of adaptive diffusion techniques with physics-informed neural networks, PINN-DADif holds a significant promise for delivering robust, efficient, and high-quality MRI reconstructions across a variety of clinical and diagnostic environments. The main contribution of the study is given below:

- PINN-DADif leverages a dynamic, adaptive diffusion prior during inference, allowing the model to adjust its generative process based on the specific characteristics of the input MRI data. This overcomes the limitations of traditional fixed priors, which often struggle with domain shifts between training and testing conditions.
- The use of PINNs introduces physical constraints into the reconstruction process, ensuring that the model's outputs adhere to the

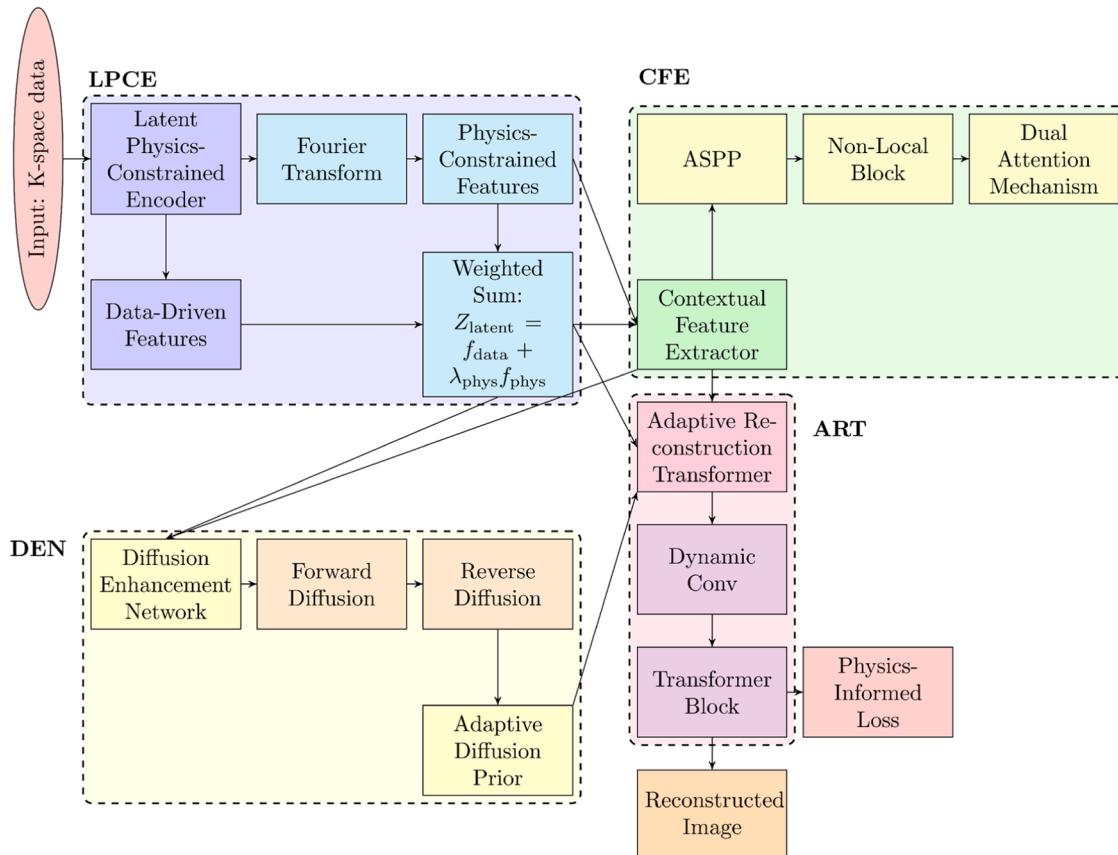


Fig. 1. Complete structure of PINN-DADif.

- underlying physics of MRI signal acquisition. This enhances the model's robustness and generalization to varying imaging operators, making it suitable for real-world clinical applications.
- The model is trained on mixed-contrast datasets, enabling it to handle multiple MRI sequences (T1, T2, and Proton Density-weighted images) using a single, unified model. This multi-contrast generalization makes PINN-DADif more versatile and applicable across various MRI protocols.
  - By employing fewer, larger reverse diffusion steps through an adversarial mapping mechanism, PINN-DADif can accelerate the MRI reconstruction process, reducing computational costs and improving clinical efficiency.

### 1.1. Literature review

Recent research has been focused on exploring the possibilities of diffusion networks in MRI reconstruction. Viewset Diffusion [17] is an intriguing method that uses multi-view 2D data to reconstruct 3D objects with exceptional efficiency. This technique efficiently addresses the inherent ambiguity in reconstruction tasks by training a diffusion model on a wide range of viewsets for certain object categories. As a result, various solutions that correspond with the input may be generated. The researchers [18] have been investigating methods to expedite and improve the reconstruction of multishot diffusion-weighted MRI using deep learning approaches. Their research indicates that the Residual U-Net (RU-Net) model surpasses current techniques, enhancing PSNR, SSIM, and lowering NRMSE, while also greatly reducing the time required for reconstruction. The Graph Diffusion Reconstruction Network (GDRN) [19] is a fascinating advancement that examines addiction-related brain connections in addicted rats using fMRI data. This provides insight into the intricate neurological processes that contribute to addiction. The study [20] has proposed a G-guided generative multilevel network that utilizes Wasserstein Generative Adversarial Networks (WGANS) and limited sampling to improve diffusion imaging in MRI reconstruction. This technique has demonstrated significant enhancements in picture clarity, edge detection, and contrast, all without necessitating any alterations to the hardware. An extensive analysis of existing research on deep MRI reconstruction models has been conducted in a recent study [21]. The study highlights the collaboration between deep learning and undersampled techniques like parallel MRI and Compressive Sensing.

A dMRI super-resolution reconstruction network [22] has been introduced, which utilizes a 3D convolution kernel to reconstruct dMRI data in both the spatial and angular domains. This approach seeks to reduce the gap between high-dimensional data by utilizing a combination of adversarial learning and attention mechanisms. It significantly improves the reconstruction process and achieves impressive enhancements in peak signal-to-noise ratio, structural similarity, and orientation distribution function visualization. The study [23] has explored the use of WGANS in conjunction with diffusion-weighted imaging (DWI) to enhance diffusion imaging in MRI. This approach has demonstrated promising results that could potentially enhance MRI acquisition procedures without requiring any hardware modifications. Recent research has indicated that diffusion models have emerged as a viable alternative to GANs (Generative Adversarial Networks) in the realm of generating high-quality images [16,24,25]. Diffusion models employ a multi-step procedure to convert Gaussian noise into image samples, effectively characterizing the data distribution [24]. This approach has yielded remarkable outcomes in MRI reconstruction, with numerous studies reporting high image quality [26].

Nevertheless, diffusion models have their own set of limitations. A significant drawback is the computational burden, as it necessitates numerous reverse steps for image generation [24]. In order to tackle this issue, researchers have put forward different approaches. For instance, one proposal is to adjust the diffusion step size during inference in order

to speed up image sampling [26]. However, it is important to note that this adjustment may have an impact on the accuracy of reverse diffusion steps [27]. Another approach involves obtaining an initial reconstruction using a different method, which can then be used to initiate the reverse diffusion process for faster inference [15]. Despite the best efforts, diffusion models still have some limitations. These include the requirement for a significant number of reverse steps and the possibility of suboptimal performance caused by domain shifts in the image distribution [28].

Advancements in diffusion-based MRI reconstruction have made remarkable strides in enhancing image quality, although there are still a few hurdles to overcome. Traditional diffusion models often suffer from a significant drawback - their computational demand is quite high. These models frequently necessitate multiple iterations with small step sizes, resulting in extended processing times [29]. Our model, PINN-DADif, tackles this issue by employing larger step sizes guided by an adversarial mapper. This innovation streamlines the number of required iterations, greatly reducing computational time and enhancing the efficiency of the reconstruction process. One major challenge that arises is the utilization of static priors [30], which do not effectively adjust to changes in test data. Suboptimal performance may occur when there are discrepancies between the training and testing datasets. PINN-DADif addresses this issue by incorporating an adaptive diffusion prior that adjusts dynamically based on the test data. Our model's capability enables it to maintain consistent performance across various datasets and imaging conditions, which enhances its robustness and generalizability.

The quality of the initial reconstruction can greatly impact the final image quality. Traditional methods sometimes struggle with poor initialization, which can slow down convergence and affect results [31]. PINN-DADif addresses this by using advanced neural network architectures to produce high-quality initial reconstructions. These are then refined through iterative processes, ensuring that the model converges quickly to an optimal solution, thus improving both speed and image quality. By addressing these key challenges, PINN-DADif offers a more efficient, adaptable, and memory-efficient solution for MRI reconstruction. These improvements make it better suited for clinical applications, where both accuracy and speed are crucial.

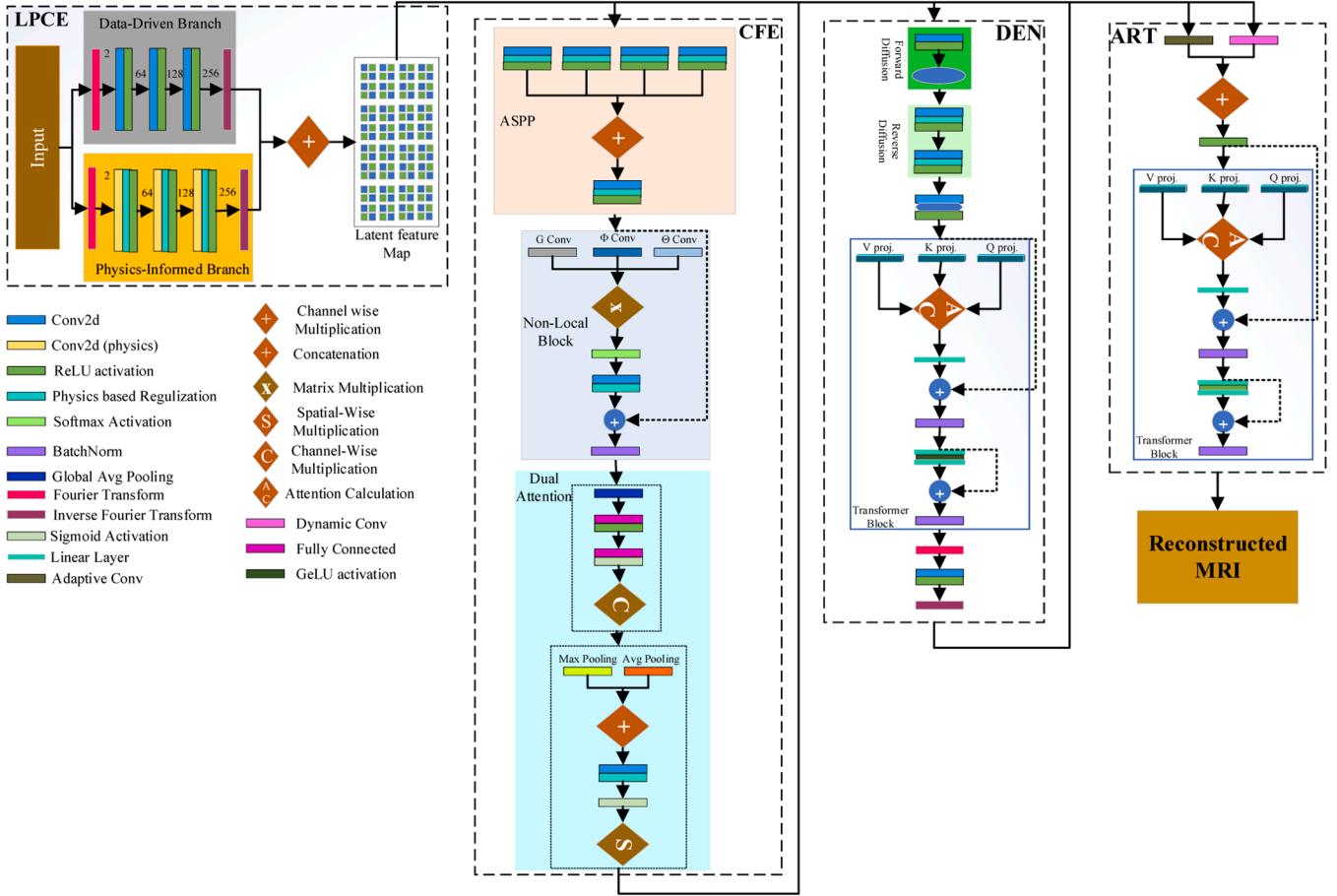
## 2. Methods

The PINN-DADif MRI reconstruction model is designed to achieve high-fidelity image reconstructions from under-sampled k-space data, leveraging advanced generative techniques. The proposed model used the physics-informed priors to incorporate domain knowledge of MRI physics into the learning process, further improving the reconstruction quality. The model is composed of four main components: the Latent Physics-Constrained Encoder (LPCE), Physics-Aware Context Encoder (PACE), Adaptive Diffusion Refinement Network (ADRN), and Adaptive Reconstruction Transformer (ART) shown in Fig. 2. Each module is structured to incorporate physics-based priors too, ensuring the model can efficiently process under-sampled data while adhering to the physical principles governing MRI signal acquisition.

### 2.1. Latent Physics-Constrained Encoder (LPCE)

The Latent Physics-Constrained Encoder (LPCE) plays a critical role in the proposed PINN-DADif. It is designed to extract and compress features from under-sampled k-space data while ensuring that the learned latent representations respect both data-driven patterns and physical constraints inherent in MRI signal generation. To achieve this, the LPCE incorporates a hybrid learning framework that balances data-driven feature learning with physics-based regularization and architecture design that mirrors MRI physical processes.

The LPCE is structured as a multi-branch encoder, where one branch learns latent features from data using traditional deep learning tech-



**Fig. 2.** Architecture of the PINN-DADif: The network consists of four key modules: Latent Physics-Constrained Encoder (LPCE), Physics-Aware Context Encoder (CFE), Adaptive Diffusion Refinement Network (ADRN), and Adaptive Reconstruction Transformer (ART). Each module contributes specific feature extraction, noise refinement, and physics-based regularization mechanisms for accurate MRI reconstruction. The LPCE employs multi-scale feature extraction (ASPP), non-local blocks, and dual attention mechanisms. The ADRN utilizes forward and reverse diffusion processes to iteratively refine features, incorporating transformer blocks for global dependency capture. Finally, the ART synthesizes the refined features through dynamic convolutions and transformer blocks, ensuring the final MRI reconstruction adheres to both spatial consistency and MRI physics constraints.

niques, while another branch explicitly models MRI physics through physics-informed operations. This enables the LPCE to capture both data-driven representations and physical properties of MRI, such as tissue-specific signal decay ( $T_1$ ,  $T_2$  relaxation times) and coil sensitivities. Let the input MRI data be represented by the under-sampled k-space data, denoted as  $X_{\text{under}}$ . The LPCE first performs a physics-guided transformation, ensuring that the latent variables  $Z_{\text{phys}}$  adhere to the known MRI physics. This transformation is guided by the Bloch equations, which govern the evolution of magnetization in MRI.

$$M(t) = M_0 \left( 1 - e^{-\frac{t}{T_1}} \right) \quad (\text{for } T_1 - \text{weighted images}) \quad (1)$$

$$M(t) = M_0 e^{-\frac{t}{T_2}} \quad (\text{for } T_2 - \text{weighted images})$$

These equations can be incorporated into the neural architecture by introducing regularization terms that ensure the latent variables respect these physical principles. This yields latent representations that capture both data-driven and physics-informed features:

$$Z_{\text{latent}} = f_{\text{data}}(X_{\text{under}}) + \lambda_{\text{phys}} f_{\text{phys}}(X_{\text{under}}) \quad (2)$$

Here,  $f_{\text{data}}$  represents the data-driven feature extraction, and  $f_{\text{phys}}$  is the physics-informed feature extraction, weighted by a scaling factor  $\lambda_{\text{phys}}$ .

The architecture of the LPCE is designed to reflect the MRI data acquisition process, leveraging physics-based transformations within

the convolutional layers. Specifically, the LPCE integrates Fourier Transform operations to process the frequency-domain information inherent in k-space data, ensuring the encoder remains consistent with the physical properties of MRI signals. The forward Fourier transform applied to the k-space data is given by:

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-2\pi i \left( \frac{ux}{M} + \frac{vy}{N} \right)} \quad (3)$$

The equation 3 converts the spatial-domain image  $f(x, y)$  into the frequency domain  $F(u, v)$ , which is essential for working with k-space data.

After the feature extraction, an Inverse Fourier Transform is applied to convert the frequency-domain features back into the spatial domain for further processing:

$$f(x, y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{2\pi i \left( \frac{ux}{M} + \frac{vy}{N} \right)} \quad (4)$$

Here,  $f(x, y)$  represents the spatial-domain image intensity at the position  $(x, y)$ .  $F(u, v)$  is the frequency-domain representation (Fourier transform) at frequency coordinates  $(u, v)$ .  $M$  and  $N$  denote the dimensions of the input data (typically the number of pixels in the x- and y-directions).  $i$  represents the imaginary unit (i.e.,  $i^2 = -1$ ) used in the Fourier transform. This transformation ensures that the latent features

returned to the spatial domain are consistent with the original data structure. By incorporating these physics-guided operations directly into the neural architecture, the LPCE ensures that the extracted features are physically consistent with MRI signal characteristics. This results in more accurate feature representations, enhancing the reconstruction of high-quality images from under-sampled data.

To ensure that the features extracted by the LPCE respect physical constraints throughout the network, we incorporate physics-based regularization directly into the activations of the intermediate layers. Specifically, we apply MRI physics-informed constraints to the latent features as they propagate through the network. This ensures that the latent features maintain alignment with physical principles throughout the encoding process rather than only at the output stage.

At each layer  $l$ , the activations  $A_l$  are modified to include a regularization term  $\mathcal{R}_{\text{phys}}$ , which enforces adherence to MRI signal properties:

$$A_l = \sigma(W_l \cdot A_{l-1} + b_l) + \lambda_{\text{phys}} \mathcal{R}_{\text{phys}}(A_l) \quad (5)$$

Here,  $\sigma$  is the activation function (e.g., ReLU),  $W_l$  and  $b_l$  are the weights and biases of layer  $l$ , and  $\mathcal{R}_{\text{phys}}$  is a physics-based regularization term that constrains the activations based on known MRI signal behaviors. For instance, the regularization may enforce smoothness in the coil sensitivity maps or apply constraints related to signal decay rates.

$$\mathcal{R}_{\text{phys}}(A_l) = \gamma \sum_{ij} \left( A_l(i,j) - f_{\text{MRI physics}}(i,j) \right)^2 \quad (6)$$

Here,  $f_{\text{MRI physics}}$  represents a function modeling physical behavior (e.g., signal relaxation), and  $\gamma$  is a weighting factor controlling the strength of the regularization. This regularization ensures that the activations are consistent with known physical behaviors, such as T1 and T2 relaxation times, improving the model's generalization and reducing artifacts. The LPCE features a dual-branch design consisting of a data-driven branch and a physics-informed branch. The physics-informed branch integrates domain knowledge from MRI acquisition physics, including k-space fidelity and gradient smoothness constraints. This integration allows LPCE to extract latent features that align with the physical properties of MRI data, reducing artifacts and improving the accuracy of the initial feature representation. These constraints ensure that the learned latent features maintain high fidelity while adhering to physical principles.

## 2.2. Physics-Aware Context Encoder (PACE)

The primary function of PACE is to enhance the latent features derived from the Latent Physics-Constrained Encoder (LPCE) by integrating multi-scale and long-range contextual information critical for accurate MRI reconstructions. The PACE leverages advanced feature extraction techniques, such as Atrous Spatial Pyramid Pooling (ASPP) and Non-Local Blocks, while incorporating physics-informed regularization to ensure that the reconstructed images are both spatially consistent and physically meaningful.

### 2.2.1. Sequence-specific processing

MRI data vary significantly depending on the imaging sequence (e.g., T1-weighted, T2-weighted, or PD-weighted). To handle this variability, the LPCE incorporates sequence-specific processing, applying tailored convolutional filters and activation functions for different MRI sequences. This ensures that the unique characteristics of each sequence are captured effectively, improving the model's ability to generalize across different types of MRI data. The LPCE employs separate convolutional layers for each sequence type, allowing it to optimize feature extraction for each:

$$Z_{\text{sequence}} = \text{Conv}_{\text{T1}}(X_{\text{T1}}) + \text{Conv}_{\text{T2}}(X_{\text{T2}}) + \text{Conv}_{\text{PD}}(X_{\text{PD}}) \quad (7)$$

By adapting its feature extraction process to the specific imaging sequence, the LPCE improves the quality of the extracted latent features

and enhances the overall reconstruction performance across various MRI modalities.

### 2.2.2. Physics-informed contextual features

The PACE refines the latent features  $Z_{\text{latent}}$  produced by the LPCE, enhancing them with a comprehensive understanding of both local and global image contexts. In MRI reconstruction, capturing fine anatomical details and broader structural information is crucial for producing high-quality images from under-sampled k-space data.

Let  $Z_{\text{latent}}$  denote the latent features from the LPCE. The PACE processes these features using multiscale feature extraction and long-range dependency modeling to produce refined contextual features  $Z_{\text{PACE}}$ , as shown below:

$$Z_{\text{PACE}} = \text{ASPP}(Z_{\text{latent}}) + \text{NonLocal}(Z_{\text{latent}}) \quad (8)$$

Here,  $\text{ASPP}(Z_{\text{latent}})$  applies multi-scale feature extraction using dilated convolutions at different scales.  $\text{NonLocal}(Z_{\text{latent}})$  model's long-range dependencies between distant features in the image, ensuring a coherent global structure.

### 2.2.3. Atrous Spatial Pyramid Pooling (ASPP)

The Atrous Spatial Pyramid Pooling (ASPP) module is integral to capturing features at multiple scales, essential for understanding both fine and coarse details within the MRI data. By applying parallel convolutional layers with varying dilation rates, the ASPP block allows the network to capture features across a range of spatial resolutions without losing the fine-grained details of the anatomy.

The output of the ASPP block is computed as:

$$Z_{\text{ASPP}} = \text{Concat}(\text{Conv}_{r=1}(Z_{\text{latent}}), \text{Conv}_{r=6}(Z_{\text{latent}}), \text{Conv}_{r=12}(Z_{\text{latent}}), \text{Conv}_{r=18}(Z_{\text{latent}})) \quad (9)$$

Here,  $\text{Conv}_{r=k}(Z_{\text{latent}})$  represents a convolutional operation with dilation rate  $r = k$ , where  $k$  varies to capture different scales of the image features. The dilation rate controls the receptive field size without increasing the number of parameters, making it efficient for extracting features across different spatial resolutions.

By applying these convolutions at different scales, the ASPP block enables the network to focus on both local structures (e.g., tissue boundaries) and global contexts (e.g., organ shapes), which are critical in MRI reconstruction.

### 2.2.4. Non-Local Blocks

While the ASPP module captures multi-scale features, the PACE's Non-Local Blocks address the challenge of modeling long-range dependencies within the MRI images. Traditional convolutional layers have a limited receptive field, restricting the ability to capture dependencies between distant parts of the image. Non-Local Blocks overcome this limitation by computing the response at a position as a weighted sum of the features at all other positions in the image, enabling the network to integrate both local and distant information for improved structural coherence.

The non-local operation is defined as:

$$Z_{\text{NonLocal}}(i) = \frac{1}{C(Z)} \sum_j \text{softmax}(Z_{\text{latent}}(i) \cdot Z_{\text{latent}}(j)) \cdot Z_{\text{latent}}(j) \quad (10)$$

Here,  $Z_{\text{NonLocal}}(i)$  is the response at position  $i$ .  $Z_{\text{latent}}(i) \cdot Z_{\text{latent}}(j)$  is the dot product between the features at positions  $i$  and  $j$ , representing their similarity.  $\text{Softmax}(\cdot)$  normalizes the similarities to produce attention weights.  $C(Z)$  is a normalization factor ensuring the output sum remains consistent. This operation allows the network to integrate global context into the feature maps, which is crucial for maintaining anatomical coherence in MRI reconstructions. By attending to features from distant parts of the image, the Non-Local Block helps mitigate artifacts and ensures that the reconstructed images retain structural integrity across

the entire field of view.

### 2.2.5. Dual attention mechanism

To further enhance the refinement of latent features, the PACE employs a Dual Attention Mechanism that combines Channel-Wise Attention and Spatial Attention. This mechanism ensures that the network emphasizes the most important features, both across channels (which represent different feature types) and spatial locations (which correspond to specific areas within the image).

**1. Channel-Wise Attention:** Enhances the most informative channels by recalibrating their importance based on global context. This mechanism is vital for ensuring that the network focuses on critical anatomical details while suppressing less relevant information.

The output of the channel attention is given by:

$$Z_{CA} = \sigma(W_c \cdot \text{GlobalAvgPool}(Z_{PACE}) + b_c) \cdot Z_{PACE} \quad (11)$$

Here,  $W_c$  and  $b_c$  are learnable parameters (weights and biases) for the channel attention mechanism. GlobalAvgPool ( $Z_{PACE}$ ) represents the global average pooling operation, which condenses the feature map into a single scalar per channel.  $\sigma$  is the sigmoid activation function, which normalizes the attention weights.

**2. Spatial Attention:** Focuses on the most critical spatial regions within the feature maps, ensuring that the network emphasizes important areas of the image, such as tissue boundaries or areas affected by artifacts.

The spatial attention is computed as:

$$Z_{SA} = \sigma(\text{Conv}_{\text{spatial}}([\text{AvgPool}(Z_{PACE}), \text{MaxPool}(Z_{PACE})])) \cdot Z_{PACE} \quad (12)$$

Here, AvgPool( $Z_{PACE}$ ) and MaxPool( $Z_{PACE}$ ) perform average and max pooling operations over the spatial dimensions. Conv<sub>spatial</sub>( $\cdot$ ) applies a convolutional filter to the pooled features to generate spatial attention weights. The final output of the PACE,  $Z_{PACE}$ , incorporates both the channel and spatial attention mechanisms, ensuring that the most relevant features are prioritized in the MRI reconstruction process.

### 2.2.6. Physics-based regularization in PACE

As with the LPCE, the PACE incorporates physics-based regularization to ensure that the extracted contextual features adhere to the physical constraints of MRI signal acquisition. This regularization is integrated into the activations of each layer, ensuring that the features remain consistent with known MRI signal properties, such as T1 and T2 relaxation times.

At each layer  $l$ , the activations are adjusted by adding a physics-based regularization term  $\mathcal{R}_{\text{phys}}$  ( $A_l$ ):

$$A_l = \sigma(W_l \cdot A_{l-1} + b_l) + \lambda_{\text{phys}} \mathcal{R}_{\text{phys}}(A_l) \quad (13)$$

Here,  $W_l$  and  $b_l$  represent the weights and biases of the layer.  $\mathcal{R}_{\text{phys}}$  is the regularization term designed to enforce physical consistency (e.g., smoothness, MRI signal decay).  $\lambda_{\text{phys}}$  is a scalar weight controlling the influence of the regularization. This regularization ensures that the contextual features remain physically plausible, improving the generalization ability of the model across different MRI datasets and imaging conditions. In the PACE, multi-scale feature extraction techniques, such as Atrous Spatial Pyramid Pooling (ASPP) and Non-Local Blocks, are enhanced with physics-based constraints. These constraints ensure that contextual features retain coherence with MRI acquisition patterns, such as coil sensitivity profiles and undersampling trajectories. This physics-guided enhancement enables the module to capture both fine anatomical details and global image context while reducing undersampling artifacts.

### 2.3. Adaptive Diffusion Refinement Network (ADRN)

ADRN refines the features extracted by the Latent Physics-Constrained Encoder (LPCE) and Physics-Aware Context Encoder (PACE), ensuring that the final MRI reconstructions maintain high fidelity while addressing noise, artifacts, and inconsistencies. The ADRN introduces iterative refinement through an adaptive diffusion process, which incorporates both forward and reverse diffusion steps. In line with the physics-informed approach, the ADRN leverages neural networks to enhance image quality while adhering to the physical constraints of MRI data acquisition. The diffusion process employed in the ADRN involves modeling noise within the latent features to gradually refine the MRI images. Unlike traditional models that apply static noise models, the ADRN dynamically adapts the diffusion process using Physics-Informed Neural Networks (PINNs). These adaptations ensure that the refined features respect the physics of MRI data generation, such as the nature of T1/T2 relaxation, gradient fields, and k-space data structure. The forward diffusion process simulates the gradual degradation of the image by adding Gaussian noise at each step, governed by a variance schedule  $\beta_t$ . The forward process is described by:

$$q(x_t | x_{t-1}) = \mathcal{N}(x_t; \sqrt{1 - \beta_t} x_{t-1}, \beta_t I) \quad (14)$$

Here,  $x_t$  represents the noisy image at timestep  $t$ ,  $\beta_t$  is the variance at timestep  $t$ , and  $I$  is the identity matrix.

This forward process gradually adds noise to simulate the degradation of image quality, providing a framework for learning how noise affects MRI data.

The reverse diffusion process iteratively refines the noisy images by removing the added noise, guided by a learned neural network  $p_\theta$ . This network is trained to reverse the forward diffusion, progressively improving the image clarity. The reverse diffusion process is defined as:

$$p_\theta(x_{t-1} | x_t) = \mathcal{N}(x_{t-1}; \mu_\theta(x_t, t), \sigma_\theta^2 I) \quad (15)$$

Here  $\mu_\theta(x_t, t)$  and  $\sigma_\theta^2$  are the mean and variance predicted by the neural network at each step  $t$ .  $x_{t-1}$  is the refined image at timestep  $t-1$ ,  $\theta$  are the parameters of the neural network. This reverse process iteratively denoises the features produced by the PACE and LPCE, gradually producing clearer and more detailed images. Importantly, the process is constrained by the physical properties of MRI data through regularization terms based on Physics-Informed Neural Networks (PINNs). Specifically, the reverse process incorporates physics-based constraints to ensure that the reconstructed images respect the underlying MRI signal model. To ensure that the diffusion process remains consistent with the physical principles governing MRI signal acquisition, the neural networks responsible for reverse diffusion are regularized with physics-based terms. These terms enforce adherence to the physical constraints of MRI, such as the smoothness of the signal and consistency with the k-space domain. At each step of the reverse diffusion process, the predicted mean  $\mu_\theta(x_t, t)$  is modified to include a physics-based regularization  $\mathcal{R}_{\text{phys}}$ , ensuring that the diffusion process respects the physical structure of the MRI signal:

$$\mu_\theta(x_t, t) = \frac{1}{\sqrt{\alpha_t}} \left( x_t - \frac{\beta_t}{\sqrt{1 - \alpha_t}} \epsilon_\theta(x_t, t) \right) + \lambda_{\text{phys}} \mathcal{R}_{\text{phys}}(x_t) \quad (16)$$

Here,  $\alpha_t = 1 - \beta_t$ ,  $\epsilon_\theta(x_t, t)$  is the predicted noise,  $\lambda_{\text{phys}}$  controls the strength of the physics-based regularization,  $\mathcal{R}_{\text{phys}}(x_t)$  is the physics-based regularization term, which enforces smoothness and signal coherence based on MRI physics. This regularization term is based on known properties of MRI signal acquisition such as the smoothness of the magnetic field gradients and the consistency of k-space data. It ensures that the diffusion process remains grounded in the physical realities of MRI acquisition improving the robustness and accuracy of the reconstructed images.

### 2.3.1. Adaptive diffusion prior

A key innovation in the ADRN is the use of an adaptive diffusion prior. Unlike fixed priors that are static throughout the diffusion process, the adaptive prior dynamically adjusts based on the characteristics of the input data. This ensures that the refinement process remains flexible and can adapt to various imaging conditions, such as different undersampling ratios or varying noise levels. The adaptive diffusion prior is updated at each timestep  $t$  based on the discrepancies between the predicted image and the real MRI data. This is achieved by minimizing a data-consistency loss, which ensures that the reconstructed images are consistent with the known k-space measurements:

$$\text{L}_{\text{data-consistency}} = \|\mathbf{F}(\mathbf{x}_t) - \mathbf{k}_{\text{measured}}\|_2^2 \quad (17)$$

Here,  $\mathbf{F}(\mathbf{x}_t)$  represents the Fourier transform of the image at timestep  $t$ ,  $\mathbf{k}_{\text{measured}}$  is the measured k-space data,  $\|\cdot\|_2^2$  denotes the squared  $\ell_2$  norm. This data consistency loss is combined with physics-based regularization to ensure that the refined images are both physically accurate and consistent with the measured MRI data.

### 2.3.2. Transformer blocks for global dependency capture

The ADRN incorporates transformer blocks to effectively capture global dependencies in the data. The self-attention mechanism within the transformer blocks allows the network to focus on relevant features across the entire image, enhancing the representation of both local and global contexts.

This is crucial for MRI reconstruction, where capturing both fine anatomical details and larger structural information is important for accurate diagnosis.

The attention mechanism is defined as:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (18)$$

Here,  $Q$ ,  $K$ , and  $V$  are the query, key, and value matrices, respectively,  $d_k$  is the dimensionality of the key vectors, softmax normalizes the dot product of  $Q$  and  $K$ , producing attention weights. These attention weights allow the ADRN to selectively focus on important regions of the image, ensuring that both local and global features are represented in the final reconstruction. The multi-head attention mechanism enhances this process by capturing different aspects of the data simultaneously, ensuring a comprehensive understanding of the MRI image. The ADRN further refines the MRI reconstructions by employing frequency domain learning. This involves transforming the features into the frequency domain using the Fourier Transform, allowing the network to better capture and refine global patterns within the MRI images. Once refined, the features are transformed back into the spatial domain using the Inverse Fourier Transform, ensuring that global patterns are seamlessly integrated into the final reconstructed image. This method allows the ADRN to better capture large-scale patterns in the data, improving the overall coherence of the reconstructed images. The ADRN module employs a dual diffusion process that is dynamically guided by physics-informed priors during forward and reverse diffusion steps. The forward diffusion adds controlled noise based on MRI acquisition physics, while the reverse diffusion iteratively removes noise, ensuring data consistency with undersampled k-space data. This guided refinement process enhances the precision and accuracy of the reconstructed images, particularly in scenarios involving severe undersampling.

## 2.4. Adaptive Reconstruction Transformer (ART)

The Adaptive Reconstruction Transformer (ART) module is the final stage in the PINN-DADif pipeline. The ART is responsible for synthesizing and refining the outputs of the preceding modules—Latent Physics-Constrained Encoder (LPCE), Physics-Aware Context Encoder (PACE), and Adaptive Diffusion Refinement Network (ADRN)—to

produce high-quality MRI reconstructions that adhere to both spatial consistency and the physics of MRI acquisition. Through dynamic convolutions, transformer blocks, and physics-informed constraints, the ART ensures the final image reconstruction maintains a fine balance between global context and local detail while mitigating artifacts caused by undersampling in k-space.

### 2.4.1. Physics-informed adaptive processing

The ART introduces a physics-informed adaptive processing framework, ensuring that all learned features adhere to physical constraints imposed by the MRI acquisition process. This includes adaptive feature refinement that tailors itself based on the unique characteristics of the input data, such as the specific MRI sequence or under-sampling ratio. The ART leverages the contextual features  $Z_{\text{PACE}}$ , which have already undergone multi-scale processing and long-range dependency modeling in the PACE, as well as the noise-reduced features refined in the ADRN,  $Z_{\text{ADRN}}$ . Let  $Z_{\text{PACE}}$  represent the contextual features and  $Z_{\text{ADRN}}$  represent the diffusion-enhanced features from the previous stages. The ART processes these through a combination of dynamic convolutions and transformer blocks to produce the final output:

$$Z_{\text{ART}} = \text{DynamicConv}(Z_{\text{PACE}}, Z_{\text{ADRN}}) + \text{Transformer}(Z_{\text{PACE}}, Z_{\text{ADRN}}) \quad (19)$$

Here, Dynamic Conv applies adaptive convolutional filters based on the input features. Transformers capture global dependencies and refined features through self-attention mechanisms.

### 2.4.2. Dynamic convolutions for adaptive feature processing

The Dynamic Convolutional Layers in ART adapt their convolutional filters based on the statistical properties of the input data. Unlike fixed convolutional layers, dynamic convolutions modify their kernels according to the characteristics of the input feature maps, allowing the network to fine-tune its filtering process in real-time. This adaptability is particularly crucial in MRI reconstruction, where different imaging sequences (e.g., T1, T2, PD-weighted) and under-sampling ratios require specific treatment for optimal reconstruction.

$$K_{\text{dynamic}} = \sigma(W_k \cdot Z_{\text{PACE}} + b_k) \quad (20)$$

$$Z_{\text{filtered}} = \text{Conv}(Z_{\text{ADRN}}, K_{\text{dynamic}}) \quad (21)$$

Here,  $K_{\text{dynamic}}$  represents the dynamically generated convolutional kernels.  $W_k$  and  $b_k$  are learnable parameters for the dynamic convolution layer.  $\text{Conv}(\cdot)$  applies the adaptive kernel to the input features. By adjusting the kernel size and filter weights dynamically, the ART can emphasize the most relevant features for each specific type of MRI sequence. This adaptability ensures that the ART effectively processes various levels of anatomical detail, from small-scale tissue boundaries to broader structural information, resulting in high-quality MRI reconstructions across a range of imaging conditions.

### 2.4.3. Transformer blocks for global feature refinement

In parallel with dynamic convolutions, the ART leverages Transformer Blocks to capture global dependencies and refine the features extracted from the PACE and ADRN. The transformer architecture uses self-attention mechanisms to focus on the most relevant regions of the image, enhancing both the local and global features required for precise MRI reconstruction. The self-attention mechanism within the transformer is defined as:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{Q \cdot K^T}{\sqrt{d_k}}\right) \cdot V \quad (22)$$

Here  $Q$ ,  $K$ , and  $V$  represent the query, key, and value matrices derived from the input features.  $d_k$  is the dimensionality of the key vectors. The ART applies this self-attention mechanism to both  $Z_{\text{PACE}}$  and  $Z_{\text{ADRN}}$ , allowing the network to integrate contextual information from the entire

image while preserving local detail. This global refinement is critical in MRI reconstruction, where anatomical coherence across the entire field of view is essential for accurate diagnosis. The transformer block in ART also incorporates a multi-head attention mechanism to capture different aspects of the data simultaneously:

$$\text{Multihead}(Q, K, V) = \text{Concat}(\text{head}_1, \text{head}_2, \dots, \text{head}_i) W_o \quad (23)$$

Here,  $\text{head}_i$  represents the self-attention operation for the  $i$ -th head.  $W_o$  is the output projection matrix. This multi-head attention allows the ART to focus on different regions and features of the image simultaneously, ensuring comprehensive feature refinement for high-quality MRI reconstructions.

#### 2.4.4. Physics-based regularization in ART

As with the LPCE and PACE, the ART incorporates physics-based regularization to ensure that the final reconstructed image adheres to the physical constraints of MRI signal acquisition. This regularization is applied during the adaptive filtering and transformer processes, ensuring that the output features respect known MRI signal properties, such as T1 and T2 relaxation times and the structure of k-space data. The adaptive convolution filters  $K_{\text{dynamic}}$  are regularized with a physics-based term  $\mathcal{R}_{\text{phys}}$ , ensuring that the filtering process maintains physical consistency:

$$K_{\text{dynamic}} = \sigma(W_k \cdot Z_{\text{PACE}} + b_k) + \lambda_{\text{phys}} \mathcal{R}_{\text{phys}}(Z_{\text{PACE}}) \quad (24)$$

Here,  $\lambda_{\text{phys}}$  controls the strength of the physics-based regularization.  $\mathcal{R}_{\text{phys}}$  ensures that the learned features align with known MRI signal properties. Similarly, the self-attention mechanism in the transformer blocks is regularized to ensure that the learned attention weights adhere to the physical constraints of MRI data:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{Q \cdot K^T}{\sqrt{d_k}}\right) + \lambda_{\text{phys}} \mathcal{R}_{\text{phys}}(Q, K, V) \quad (25)$$

This regularization ensures that the attention weights focus on regions and features that are consistent with the physical properties of the MRI data, improving both the accuracy and robustness of the final reconstructions.

The ART synthesizes the refined features from the dynamic convolutions and transformer blocks, producing the final reconstructed MRI image. The combination of adaptive filtering, global feature refinement, and physics-based regularization ensures that the output images are of the highest possible quality, free from artifacts and noise, and consistent with both the measured k-space data and the underlying physical properties of MRI signals. The final output of the ART is computed as:

$$Z_{\text{ART-final}} = \sigma(W_{\text{out}} \cdot Z_{\text{ART}} + b_{\text{out}}) \quad (26)$$

Here,  $W_{\text{out}}$  and  $b_{\text{out}}$  are the weights and biases for the final layer.  $\sigma$  is the activation function that produces the final reconstructed image. The ART module leverages dynamic convolutions and transformer-based attention mechanisms, both of which are guided by physics-informed priors. These priors enforce global consistency and local feature accuracy by incorporating k-space fidelity and anatomical coherence into the reconstruction process. This ensures that the reconstructed images preserve fine anatomical structures while maintaining global structural integrity.

#### 2.5. Loss and optimization function

The training of the PINN-DADif hinges on a carefully designed loss function and an efficient optimization technique. Given the multi-stage and physics-informed nature of PINN-DADif, the loss function must not only measure pixel-wise accuracy but also capture higher-level features and incorporate domain-specific physics constraints to ensure physically plausible reconstructions. The Loss Function in PINN-DADif combines

several key components: Pixel-wise Loss, Perceptual Loss, and a Physics-Informed Loss (PINN Regularization). Each component serves a distinct purpose in refining the quality of MRI reconstructions and ensuring consistency with both the measured data and the physical laws governing MRI signal acquisition.

##### 2.5.1. Pixel-wise Loss (L2 Loss)

The Pixel-wise Loss evaluates the direct difference between the reconstructed MRI image and the fully-sampled ground truth image at each pixel. This loss component encourages the model to produce reconstructions that closely match the original MRI data in terms of pixel intensity.

$$L_{\text{pixel}} = \|I_{\text{REC}} - I_{\text{GT}}\|_2^2 \quad (27)$$

Where,  $I_{\text{REC}}$  is the reconstructed MRI image.  $I_{\text{GT}}$  is the fully-sampled ground truth image.  $\|\cdot\|_2^2$  represents the squared L2-norm. The Pixel-wise Loss plays a critical role in minimizing the overall intensity error between the reconstructed and ground truth images, ensuring that the reconstructed images remain close to the ground truth in terms of basic pixel values.

##### 2.5.2. Perceptual Loss

While Pixel-wise Loss ensures pixel-level accuracy, it may not be sufficient for capturing the high-level structures and anatomical details necessary for medical diagnosis. To address this, PINN-DADif also includes a Perceptual Loss, which measures the similarity between the high-level feature representations of the reconstructed image and the ground truth image. This component uses a pre-trained deep neural network (e.g., VGG) to extract features at various layers, comparing the feature maps of the two images to ensure they are structurally similar.

$$L_{\text{perc}} = \sum_{i=1}^L \|\phi_i(I_{\text{REC}}) - \phi_i(I_{\text{GT}})\|_2^2 \quad (28)$$

Here,  $\phi_i(\cdot)$  represents the feature map from the  $i$ -th layer of the pre-trained network (e.g., VGG).  $I_{\text{REC}}$  is the reconstructed image, and  $I_{\text{GT}}$  is the ground truth image. By comparing these feature representations, the Perceptual Loss helps PINN-DADif maintain the structural integrity of the MRI image, making sure that the reconstructed image preserves important anatomical details and looks visually coherent.

##### 2.5.3. Physics-Informed Loss (PINN Regularization)

Incorporating physics-based constraints is crucial for ensuring that the reconstructed images respect the underlying physical laws governing MRI data acquisition. The Physics-Informed Loss in PINN-DADif introduces a regularization term derived from Physics-Informed Neural Networks (PINNs), ensuring that the reconstructed MRI images are not only visually accurate but also consistent with known MRI physics, such as T1 and T2 relaxation, gradient field smoothness, and k-space data consistency.

$$L_{\text{phys}} = \lambda_{\text{phys}} \mathcal{R}_{\text{phys}}(I_{\text{REC}}) \quad (29)$$

Here,  $\mathcal{R}_{\text{phys}}(\cdot)$  is the physics-based regularization term that enforces adherence to the MRI signal model.  $\lambda_{\text{phys}}$  is a hyperparameter controlling the influence of physics-based regularization.

This regularization term penalizes deviations from the known physical properties of MRI data, ensuring that the reconstructed images are consistent with the k-space measurements and the physical structure of the MRI signal, such as the smoothness of magnetic field gradients and the structure of the signal in k-space.

##### 2.5.4. Total Loss Function

The Total Loss Function combines Pixel-wise Loss, Perceptual Loss, and Physics-Informed Loss, balancing their contributions to guide the model toward high-quality MRI reconstructions that are visually accu-

rate and physically plausible. The total loss is defined as:

$$L_{\text{total}} = \alpha L_{\text{pixel}} + \beta L_{\text{perc}} + \gamma L_{\text{phys}} \quad (30)$$

Here  $\alpha, \beta$ , and  $\gamma$  are weighting factors that control the relative importance of each loss component. These hyperparameters are carefully chosen through experimentation to balance pixel accuracy, perceptual quality, and adherence to physical constraints.

#### 2.5.5. Optimization: Sharpness-Aware Minimization (SAM)

For robust optimization, PINN-DADif employs Sharpness-Aware Minimization (SAM), an advanced optimization technique that enhances the model's generalization by minimizing not only the loss but also the sharpness of the loss landscape. This technique is well-suited for medical imaging applications, where robustness to small perturbations and variations in the input data is crucial.

SAM operates in two main steps, Perturbing the Parameters: SAM perturbs the model parameters in the direction of the gradient to explore the local neighborhood of the current solution. This helps the model find flatter minima, which are less sensitive to small changes in the input data, improving generalization:

$$\theta^+ = \theta + \rho \frac{\nabla_{\theta} L(\theta)}{\|\nabla_{\theta} L(\theta)\|} \quad (31)$$

Here  $\theta$  represents the model parameters.  $\rho$  is a hyperparameter controlling the magnitude of the perturbation.  $\nabla_{\theta} L(\theta)$  is the gradient of the loss function with respect to the model parameters.

Minimizing the Perturbed Loss: After perturbing the parameters, SAM minimizes the loss at the perturbed parameters, ensuring that the model converges to a flatter, more robust solution:

$$\theta = \theta - \eta \nabla_{\theta} L(\theta^+) \quad (32)$$

Here,  $\eta$  is the learning rate and  $\theta^+$  are the perturbed parameters from the previous step. By finding flatter minima, SAM improves the robustness and generalization of PINN-DADif, making it better suited for clinical MRI applications where variations in the imaging operator and patient-specific factors are common.

#### 2.6. Dataset selection

The performance of PINN-DADif was carefully evaluated in this work using a private MRI dataset and the publically available fastMRI dataset under various imaging settings. The confidential collection consisted of 10,008 brain MRI images classified into T1-weighted, T2-weighted, and Proton Density (PD)-weighted sequencing categories. The acquisition settings for T1-weighted images were a repetition time (TR) of 9.813 milliseconds, an echo time (TE) of 4.603 milliseconds, and a flip angle that was set at 8°. T2-weighted pictures provided a time resolution (TR) of 8178 milliseconds, a time equivalent (TE) of 100 milliseconds, and a flip angle of 90°. PD-weighted images utilized the same TR but with a TE of 8 milliseconds and a flip angle of 90°. The voxel size of each scan was set at 0.94 × 0.94 × 1.2 mm. The images were divided into training, validation, and test sets using a 70–15–15 % split. This made 2100 images weighted for T1, 2103 images weighted for T2, and 2803 images weighted for PD for training; 450 images weighted for T1, 451 images weighted for T2, and 601 images weighted for PD for validation; and 450 images weighted for T1, 450 images weighted for T2, and 600 images weighted for PD for testing. The allocation mechanism employed in this study guaranteed that every set had a comprehensive distribution of all sequence types, therefore enabling a well-balanced approach to both model training and assessment. Presented by Facebook AI Research and NYU Langone Health, the fastMRI dataset comprised multi-coil k-space data for T1-weighted, T2-weighted, and FLAIR sequences. We chose scans that required a minimum of 10 cross-sections and 20 coil pieces in order to guarantee both computational efficiency and consistency.

Utilizing geometric coil compression, the data was compressed to either 5 or 10 virtual coils, resulting in the preservation of nearly 95 % and 98 % of the original signal energy, respectively. It was divided into 70 % training, 15 % validation, and 15 % testing datasets, consisting of 2800, 600, and 600 cross-sectional pictures, respectively. Retrospective undersampling was conducted in the transverse plane using variable-density random undersampling to attain acceleration rates of 4x, 8x, and 12x. The SPIRiT method [32] was used to estimate coil sensitivities from a fully-sampled center calibration area. The varied datasets and precise splitting techniques employed allowed for a thorough assessment of PINN-DADif's resilience and its ability to generate high-quality reconstructions in various imaging modalities and environments.

#### 2.7. Experimental setup

A comprehensive comparison was conducted between PINN-DADif and various state-of-the-art MRI reconstruction methods to rigorously assess its performance. The benchmark included comparisons with leading conditional models like MoDL [33], as well as unconditional models such as DDPM [27] and DiffRecon [26] with Deep learning-based models and AdaDiff [29]. Conditional models aimed to map the inverse Fourier transform of undersampled k-space data to fully-sampled ground truth images, while unconditional models were developed to directly generate coil-combined MRI samples using learned generative priors. For PINN-DADif, a physics-informed noise scheduler was implemented, with the parameters  $\beta_{\min}$  and  $\beta_{\max}$  set to 0.1 and 20, respectively. This exponential decay of noise over time allows the model to refine image quality progressively through the adaptive diffusion process, ensuring that MRI reconstructions adhere to physical constraints while reducing artifacts. This architecture, combined with the physics-based priors, significantly enhances the clarity and fidelity of the reconstructed images. The hyperparameters for PINN-DADif were optimized through one-fold cross-validation to ensure the highest possible reconstruction quality. We selected a learning rate of  $6 \times 10^{-3}$  and trained the model for 600 epochs, with a step size of 125 and 10 diffusion steps, where  $T/k = 10$ . To accelerate the diffusion process, we implemented 12 reverse diffusion iterations followed by a data-consistency projection in each iteration, allowing for efficient refinement of the reconstruction. During the adaptation phase, which fine-tunes the prior to aligning with real-world data, the learning rate was set to  $10^{-3}$  with 1000 adaptation iterations. Training and evaluations were performed on an Nvidia RTX 4060 GPU, utilizing the PyTorch deep learning framework for efficient computation and model handling.

PINN-DADif was trained using a cutting-edge optimization technique known as Sharpness-Aware Minimization (SAM), which builds on the Adam optimizer to enhance the model's generalization capability. SAM works by perturbing the model's parameters during the gradient descent process, leading to smoother minima in the loss landscape, thus improving the robustness of the reconstructions. For the Adam optimizer, the decay rates of  $\beta_1 = 0.9$  and  $\beta_2 = 0.999$  were applied to balance convergence speed with stability. This combination allowed PINN-DADif to converge effectively while maintaining strong generalization across diverse imaging conditions. During the inference phase, PINN-DADif continued to employ the SAM optimizer, coupled with Adam, to fine-tune the reconstruction based on test data, utilizing the same decay rates ( $\beta_1 = 0.9$  and  $\beta_2 = 0.999$ ). The Total Loss Function combines pixel-wise reconstruction loss (L2 loss), perceptual loss, and physics-informed constraints, with respective weights of 0.5, 0.3, and 0.2. These weights were determined through experimentation to balance reconstruction fidelity, perceptual quality, and adherence to MRI physics. This continual fine-tuning during inference allowed for adaptive reconstruction that improved the alignment between the reconstructed images and the underlying physical constraints of MRI, leading to high-fidelity outputs across multiple imaging modalities.

In our experiments, PINN-DADif employs a fast and adaptive diffu-

sion process with step sizes of  $k = 125$  and  $T/k = 10$ , incorporating several optimizations that significantly enhance both efficiency and image quality. A key innovation in PINN-DADif is the adaptive diffusion strategy, which dynamically adjusts step sizes and diffusion iterations based on the input data's characteristics. This adaptability allows for faster convergence, reducing the number of iterations required to achieve high-fidelity MRI reconstructions. In practice, PINN-DADif demonstrated exceptional computational efficiency, with average training and inference times of 125.2 and 117.5 s per cross-section, respectively, outperforming the 131.4 s reported by the baseline model [15], as shown in Table 6. This optimization enhances the model's practical applicability, especially in real-world clinical settings where resource efficiency is critical. To further improve reconstruction precision and image clarity, the number of reverse diffusion iterations in the refinement phase was increased from 8 to 12. While this adjustment adds a slight computational load, it significantly boosts reconstruction accuracy and reduces artifacts, particularly in complex and challenging MRI scenarios. Despite this added complexity, PINN-DADif remains computationally efficient due to its streamlined architecture and optimized sampling strategies. This balance between superior image quality and computational speed makes PINN-DADif an ideal solution for clinical applications, where both rapid turnaround times and high diagnostic precision are crucial. Its ability to deliver high-quality reconstructions while maintaining efficient processing times underscores its potential for widespread use in accelerated MRI protocols.

## 2.8. Evaluation metricx

We tested our proposed model, PINN-DADif, on retrospectively undersampled MRI scans from the private dataset and fastMRI datasets. To make the model more practical for real-world use, we trained it on a combination of different contrast types. This approach reflects the common clinical practice of using multi-contrast protocols. Specifically, we pooled T1, T2, and PD-weighted images from the private dataset and T1, T2, and FLAIR-weighted images from the fastMRI dataset. The training samples were drawn randomly from this combined pool, and the model did not know the contrast type of each sample during training. For conditional models, the undersampled data input includes information about the acceleration rate (R-value). Our experiments showed that training a single unified model on mixed R values did not negatively impact performance. However, to ensure fair performance comparisons, we trained separate models for each specific R-value. This approach prevents biases that could arise from differences in acceleration rates between the training and test sets. We measured the quality of the reconstructed images using the Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM), comparing the reconstructed images to fully-sampled ground truth images. We used non-parametric signed-rank tests to evaluate the significance of differences in PSNR and SSIM between models. This evaluation framework allowed us to thoroughly assess the performance of PINN-DADif across different contrast types and acceleration rates, demonstrating its effectiveness and potential for clinical MRI reconstruction.

## 3. Results

### 3.1. Inter-domain reconstruction

We conducted tests on PINN-DADif for inter-domain reconstructions, ensuring that the imaging operator and MR image distribution were the same for both training and testing sets. We trained and tested the model at  $R = 4x$  using the fastMRI dataset. PINN-DADif's performance was compared against a conditional model (MoDL), a DL-based method, and static prior diffusion models (DDPM, DiffRecon, AdaDiff). The results, presented in Table 1 (A), (B) and Fig. 3, show the PSNR and SSIM metrics for these comparisons on the private and fastMRI datasets, respectively. In the private dataset, PINN-DADif consistently delivered the highest

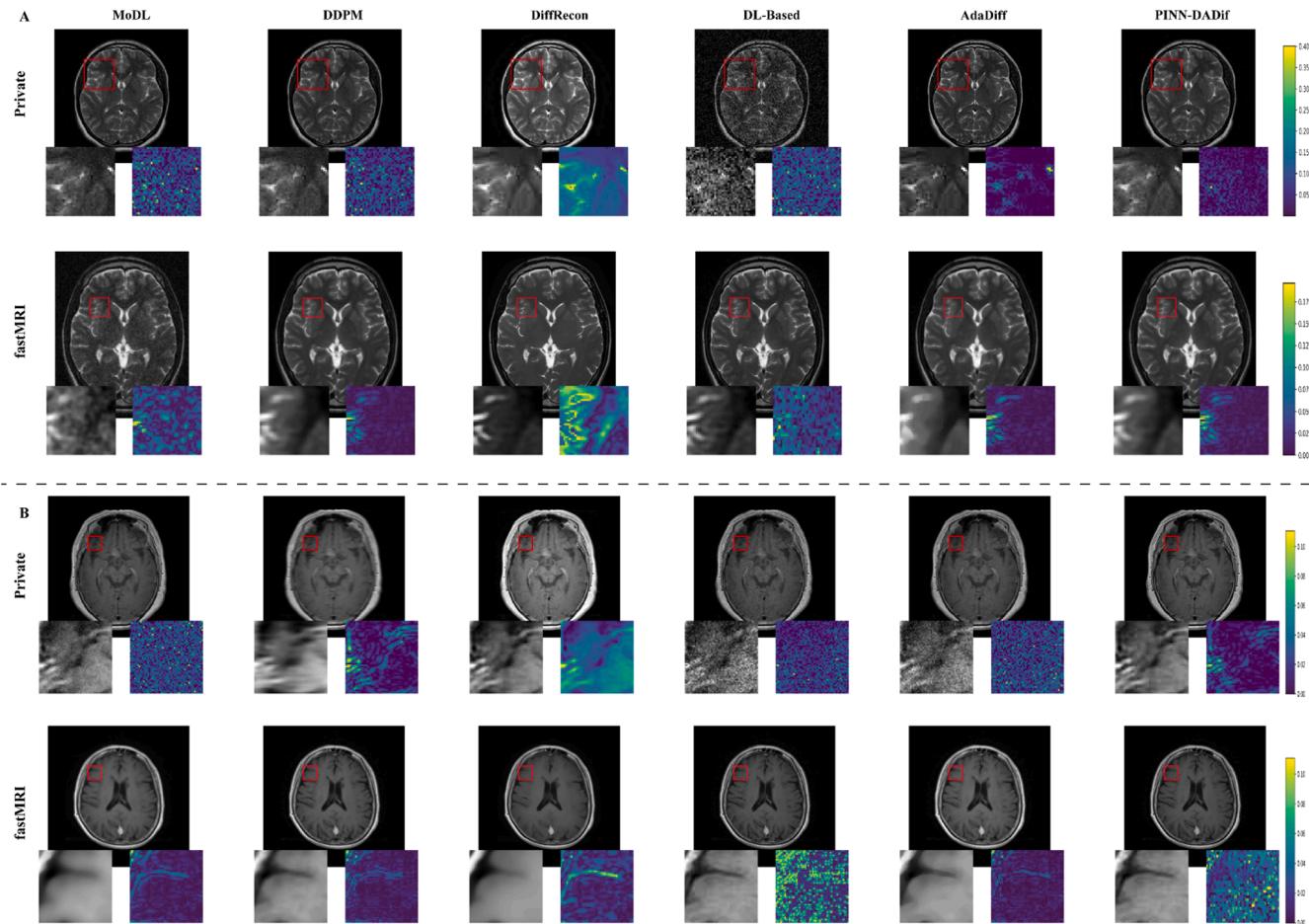
**Table 1**

(A) The within-domain performance of T1-, T2-, and Proton Density (PD)-weighted contrasts in the private dataset was evaluated at  $R = 4x\text{-}8x$ . (B) The within-domain performance of T1-, T2-, and FLAIR-weighted contrasts in fastMRI is evaluated at  $R = 4x\text{-}8x$ . The performance was measured in PSNR (Peak Signal-to-Noise Ratio) and SSIM (Structural Similarity Index).

(A)		MoDL	DDPM	DiffRecon	DL-based	AdaDiff	PINN-DADif
<i>R = 4x</i>							
T1	PSNR	38.7 ± 1.4	39.5 ± 1.2	39.0 ± 0.9	34.0 ± 1.1	40.5 ± 1.5	41.2 ± 1.3
	SSIM	98.0 ± 0.3	98.2 ± 0.2	97.8 ± 0.4	96.0 ± 0.4	98.5 ± 0.3	98.7 ± 0.2
T2	PSNR	37.5 ± 1.3	38.1 ± 1.1	36.8 ± 1.3	34.5 ± 1.4	39.0 ± 1.4	39.5 ± 1.2
	SSIM	97.2 ± 0.3	97.5 ± 0.3	96.0 ± 0.5	95.0 ± 0.3	97.8 ± 0.3	98.0 ± 0.3
PD	PSNR	39.0 ± 1.2	39.8 ± 1.3	38.0 ± 0.6	33.5 ± 1.3	40.8 ± 1.6	41.5 ± 1.4
	SSIM	97.5 ± 0.3	97.9 ± 0.3	96.8 ± 0.3	95.5 ± 0.2	98.2 ± 0.2	98.5 ± 0.2
<i>R = 8x</i>							
T1	PSNR	34.1 ± 1.3	34.9 ± 1.1	33.5 ± 0.8	30.0 ± 1.2	35.5 ± 1.4	36.2 ± 1.3
	SSIM	95.5 ± 0.7	95.8 ± 0.6	95.0 ± 0.5	92.5 ± 0.3	96.5 ± 0.5	96.8 ± 0.4
T2	PSNR	33.8 ± 1.1	34.5 ± 1.0	33.0 ± 0.9	31.0 ± 1.3	35.0 ± 1.3	35.7 ± 1.2
	SSIM	94.8 ± 0.6	95.0 ± 0.5	94.0 ± 0.4	92.0 ± 0.2	95.5 ± 0.6	95.8 ± 0.5
PD	PSNR	35.0 ± 1.4	35.7 ± 1.1	34.0 ± 0.7	32.5 ± 1.3	36.2 ± 1.5	36.8 ± 1.3
	SSIM	95.2 ± 0.6	95.5 ± 0.5	94.5 ± 0.4	93.0 ± 0.2	96.0 ± 0.6	96.3 ± 0.5
(B)		MoDL	DDPM	DiffRecon	DL-based	AdaDiff	PINN-DADif
<i>R = 4x</i>							
T1	PSNR	39.8 ± 1.3	38.2 ± 1.7	38.6 ± 1.5	38.6 ± 1.5	40.2 ± 1.7	41.5 ± 1.3
	SSIM	95.7 ± 1.2	92.5 ± 7.6	94.1 ± 1.8	94.1 ± 1.8	95.9 ± 1.4	96.2 ± 1.1
T2	PSNR	36.7 ± 0.8	37.5 ± 0.6	39.1 ± 0.7	39.1 ± 0.7	37.7 ± 0.8	40.0 ± 0.9
	SSIM	96.0 ± 0.5	95.7 ± 0.5	96.8 ± 0.4	92.8 ± 0.4	96.2 ± 0.4	97.0 ± 0.3
FL	PSNR	36.0 ± 2.2	34.1 ± 2.9	35.4 ± 2.3	35.4 ± 2.3	37.2 ± 2.6	37.0 ± 2.2
	SSIM	92.7 ± 4.0	88.3 ± 6.9	91.7 ± 4.7	91.7 ± 4.7	92.5 ± 4.8	93.2 ± 4.3
<i>R = 8x</i>							
T1	PSNR	37.1 ± 1.1	36.2 ± 1.3	34.7 ± 1.0	30.4 ± 1.2	37.2 ± 1.5	37.5 ± 1.2
	SSIM	93.5 ± 1.6	91.1 ± 2.1	89.4 ± 1.8	90.2 ± 0.2	94.2 ± 2.2	94.0 ± 1.5
T2	PSNR	33.9 ± 0.8	34.8 ± 0.6	35.2 ± 0.5	32.1 ± 1.4	35.3 ± 0.8	35.9 ± 0.7
	SSIM	93.9 ± 0.7	94.2 ± 0.6	93.9 ± 0.7	92.5 ± 0.1	94.4 ± 0.7	95.0 ± 0.6
FL	PSNR	33.7 ± 1.9	32.8 ± 2.2	32.8 ± 1.6	32.7 ± 1.3	33.7 ± 2.4	34.2 ± 2.1
	SSIM	88.9 ± 4.8	86.5 ± 6.6	87.3 ± 4.8	87.7 ± 0.1	88.8 ± 6.2	89.2 ± 5.0

performance across various contrasts and acceleration rates, with statistically significant improvements ( $p < 0.05$ ) in most instances.

For T1-weighted images at  $R = 4x$ , PINN-DADif achieved a PSNR of 41.2 dB and an SSIM of 98.7 %, surpassing MoDL, DDPM, and DiffRecon. For T2-weighted images, PINN-DADif reached a PSNR of 39.5 dB and an SSIM of 98.0 %, outperforming the other methods. AdaDiff was the exception, showing similar performance to PINN-DADif on T1-weighted images (PSNR of 40.5 dB, SSIM of 98.5 %) and slightly outperforming PINN-DADif on PD-weighted images at  $R = 4x$  (PSNR of 40.8 dB, SSIM of 98.2 %). At  $R = 8x$ , PINN-DADif maintained its lead with a PSNR of 36.2 dB and an SSIM of 96.8 % for T1-weighted images and 35.7 dB and 95.8



**Fig. 3.** The reconstruction results in the Private and fastMRI dataset, where the acceleration factor  $R = 4x$  is shown in (A) and the acceleration factor  $R = 8x$  in (B). For each  $R$ , the first row displays the reconstructed images of different methods using a private dataset with ROI and an error map. The second row displays the reconstructed images of different methods using a fastMRI dataset with ROI and an error map. The color bars are shown on the right side.

% for T2-weighted images. AdaDiff, however, had higher SSIM on T1-weighted images (97.6 %) and the highest PSNR on PD-weighted images (36.6 dB).

In the fastMRI dataset, PINN-DADif consistently outperformed the other methods across different contrasts and acceleration rates ( $p < 0.05$ ), except when compared to AdaDiff, which had similar performance. PINN-DADif achieved a PSNR of 37.5 dB and an SSIM of 94.0 % for T1-weighted images at  $R = 8x$  and 35.9 dB and 95.0 % for T2-weighted images. Although AdaDiff's performance was close, PINN-DADif held a slight edge overall. PINN-DADif showed an average improvement of approximately 1.5 dB in PSNR and 1.4 % in SSIM over conditional models (MoDL). Compared to DL-based methods, PINN-DADif improved by about 4.7 dB in PSNR and 3.1 % in SSIM. Furthermore, PINN-DADif outperformed static diffusion models (DDPM, DiffRecon, AdaDiff) with an average gain of around 1.0 dB in PSNR and 1.0 % in SSIM. These results highlight PINN-DADif's robustness and effectiveness in producing high-quality MRI reconstructions across different datasets and imaging conditions.

### 3.2. Generalization across acceleration rates

PINN-DADif is designed to dynamically adapt to diverse undersampling patterns and acceleration rates through its integration of physics-informed priors and an adaptive diffusion strategy. The model leverages a physics-based noise scheduler, where parameters  $\beta_{min} = 0.1$  and  $\beta_{max} = 20$  are exponentially decayed over time. This process ensures robust refinement of image quality by adhering to k-space fidelity and

preserving anatomical structures. By dynamically adjusting the diffusion prior during inference, the model is capable of handling varying acceleration rates and undersampling schemes. Furthermore, the reverse diffusion step size and iterations are dynamically adjusted based on the input data characteristics. For example, in cases of higher acceleration rates (e.g., 8x), the model increases the number of refinement steps to enhance reconstruction accuracy, while for lower acceleration rates (e.g., 4x), fewer steps are sufficient, maintaining computational efficiency.

To illustrate the model's adaptability, we evaluated PINN-DADif's performance in cross-domain reconstructions, where the MR image distribution was the same, but the imaging operator differed between training and testing sets. The models were trained at an acceleration rate of  $R = 4x$  and tested at  $R = 8x$  and  $R = 12x$ , as shown in Table 2 (A) and (B). In the private dataset, PINN-DADif delivered impressive results across different contrasts and acceleration rates. For example, at  $R = 8x$ , PINN-DADif achieved a PSNR of 36.0 dB and an SSIM of 96.8 % on T1-weighted images, outperforming AdaDiff, which had a PSNR of 35.2 dB and an SSIM of 96.5 %. At  $R = 12x$ , PINN-DADif maintained its superior performance with a PSNR of 32.5 dB and an SSIM of 94.0 % on T1-weighted images, compared to AdaDiff's 32.0 dB PSNR and 93.5 % SSIM. For the fastMRI dataset, PINN-DADif consistently outperformed other methods. At  $R = 8x$ , PINN-DADif achieved a PSNR of 37.8 dB and an SSIM of 94.0 % on T1-weighted images, while AdaDiff had a PSNR of 37.2 dB and an SSIM of 93.5 %. At  $R = 12x$ , PINN-DADif continued to lead with a PSNR of 35.8 dB and an SSIM of 92.5 % on T1-weighted images, surpassing AdaDiff's 35.2 dB PSNR and 91.2 % SSIM. Overall, PINN-DADif showed significant improvements over other models. On

**Table 2**

(A) The cross-domain performance of T1-, T2-, and Proton Density (PD)-weighted contrasts were evaluated using a private dataset. The results are presented for training at  $R = 4x$  and testing at  $R = 8x$  and  $R = 12x$ . (B) The cross-domain performance of T1-, T2-, and FLAIR- (FL.) weighted contrasts in fastMRI is evaluated. The results obtained from training at  $R = 4x$  and testing  $R = 8x$  and  $R = 12x$ . The performance was measured in PSNR (Peak Signal-to-Noise Ratio) and SSIM (Structural Similarity Index).

(A)	MoDL	DDPM	DiffRecon	DL-based	AdaDiff	PINN-DADif
<i>R = 8x</i>						
T1	PSNR 1.5	33.5 ± 34.0 ± 1.3	34.7 ± 0.9	32.0 ± 1.3	35.2 ± 1.6	36.0 ± 1.4
	SSIM 0.9	95.2 ± 95.5 ± 0.8	96.0 ± 0.6	92.0 ± 1.7	96.5 ± 0.7	96.8 ± 0.6
T2	PSNR 1.4	32.7 ± 33.5 ± 1.1	33.9 ± 0.7	31.5 ± 1.3	34.5 ± 1.5	35.2 ± 1.4
	SSIM 0.9	93.5 ± 94.0 ± 0.8	94.3 ± 0.6	91.5 ± 1.6	95.0 ± 0.8	95.5 ± 0.7
PD	PSNR 1.8	33.8 ± 34.2 ± 1.3	34.5 ± 0.8	32.5 ± 1.7	35.0 ± 1.9	35.6 ± 1.6
	SSIM 1.0	93.0 ± 93.5 ± 0.9	94.0 ± 0.6	90.5 ± 2.0	95.5 ± 0.8	96.0 ± 0.7
<i>R = 12x</i>						
T1	PSNR 1.4	30.5 ± 31.2 ± 1.2	31.8 ± 0.8	29.0 ± 1.4	32.0 ± 1.4	32.5 ± 1.3
	SSIM 1.6	91.5 ± 92.0 ± 1.3	92.5 ± 1.0	90.0 ± 1.3	93.5 ± 1.2	94.0 ± 1.0
T2	PSNR 1.2	30.2 ± 30.8 ± 1.1	31.0 ± 0.9	28.5 ± 1.3	32.0 ± 1.4	32.3 ± 1.2
	SSIM 1.8	89.0 ± 90.5 ± 1.2	90.7 ± 1.0	88.5 ± 1.4	92.0 ± 1.1	92.5 ± 1.0
PD	PSNR 1.6	30.5 ± 31.0 ± 1.2	31.5 ± 1.0	29.5 ± 1.5	32.5 ± 1.7	33.0 ± 1.5
	SSIM 2.5	86.5 ± 88.0 ± 1.5	89.0 ± 1.2	85.0 ± 1.9	91.0 ± 1.3	91.5 ± 1.1
<b>(B)</b>						
<i>R = 8x</i>						
T1	PSNR 1.0	36.4 ± 36.2 ± 1.3	34.7 ± 1.0	34.5 ± 1.2	37.2 ± 1.5	37.8 ± 1.3
	SSIM 1.6	93.3 ± 91.1 ± 2.1	89.4 ± 1.8	92.0 ± 2.0	93.5 ± 2.2	94.0 ± 1.8
T2	PSNR 0.8	32.9 ± 34.8 ± 0.6	35.2 ± 0.5	34.0 ± 0.8	35.3 ± 0.8	36.0 ± 0.7
	SSIM 0.8	93.4 ± 94.2 ± 0.6	93.9 ± 0.7	93.2 ± 0.8	94.4 ± 0.6	95.0 ± 0.5
FL	PSNR 1.9	33.2 ± 32.8 ± 2.2	32.8 ± 1.6	33.5 ± 1.9	33.7 ± 2.4	34.2 ± 1.8
	SSIM 4.8	88.8 ± 86.5 ± 6.6	87.3 ± 4.8	88.0 ± 5.0	88.8 ± 6.2	89.5 ± 5.3
<i>R = 12x</i>						
T1	PSNR 1.0	34.7 ± 35.9 ± 0.9	33.0 ± 0.8	32.5 ± 1.1	35.2 ± 1.6	35.8 ± 1.4
	SSIM 1.8	91.7 ± 91.2 ± 1.5	86.4 ± 1.9	91.0 ± 1.6	91.2 ± 2.7	92.5 ± 1.9
T2	PSNR 0.8	31.2 ± 33.2 ± 0.5	33.2 ± 0.5	31.0 ± 0.9	33.9 ± 0.8	34.5 ± 0.7
	SSIM 1.0	91.4 ± 92.8 ± 0.8	91.3 ± 1.0	90.5 ± 1.0	93.2 ± 0.8	93.8 ± 0.7
FL	PSNR 1.7	31.8 ± 32.2 ± 1.7	31.4 ± 1.4	30.5 ± 1.6	32.3 ± 2.3	33.0 ± 1.8
	SSIM 5.3	86.2 ± 86.3 ± 5.7	84.1 ± 4.8	85.0 ± 5.1	86.4 ± 6.9	87.2 ± 5.5

average, PINN-DADif outperformed traditional methods by approximately 4.7 dB in PSNR and 3.1 % in SSIM. Compared to conditional models, PINN-DADif showed an average improvement of 1.5 dB in PSNR and 1.4 % in SSIM. When compared to DL-based methods, PINN-DADif improved by about 4.7 dB in PSNR and 3.1 % in SSIM. Additionally, PINN-DADif outperformed static diffusion models by around 1.0 dB in PSNR and 1.0 % in SSIM. The visual results are shown in Fig. 4.

The results also highlight the model's performance under different clinical protocols, such as T1-weighted, T2-weighted, and FLAIR

imaging, demonstrating its robustness across imaging modalities. By tailoring the physics-informed priors to account for the specific sequence parameters, such as echo time (TE) and repetition time (TR), PINN-DADif ensures high fidelity and diagnostic reliability under diverse clinical conditions.

### 3.3. Domain shifts in the image distribution

We also evaluated PINN-DADif's performance in cross-domain scenarios where the imaging operator was consistent, but the MR image distribution differed between the training and test sets. In this setup, we trained the models on the fastMRI dataset and tested them on the private dataset. The PSNR and SSIM results are presented in Table 3. Overall, PINN-DADif excelled across different tissue contrasts and acceleration rates, achieving the highest scores in most cases ( $p < 0.05$ ). For example, PINN-DADif reached a PSNR of 43.0 dB and an SSIM of 99.0 % on T1-weighted images at  $R = 4x$ . At  $R = 8x$ , PINN-DADif achieved a PSNR of 38.5 dB and an SSIM of 97.0 %. Although AdaDiff also performed well, PINN-DADif generally outperformed it. On average, PINN-DADif surpassed traditional methods by approximately 8.0 dB in PSNR and 19.8 % in SSIM. Compared to conditional models, PINN-DADif showed improvements of about 7.5 dB in PSNR and 18.7 % in SSIM. When compared to adaptive GAN models, PINN-DADif achieved approximately 9.5 dB higher in PSNR and 20.5 % better in SSIM.

Furthermore, PINN-DADif outperformed static diffusion models by around 1.9 dB in PSNR and 2.1 % in SSIM. These results highlight that shifts in the MR image distribution have a more substantial impact on the performance of other methods, including static diffusion models, compared to shifts in the imaging operator. PINN-DADif's architecture and advanced feature extraction techniques enable it to maintain high reconstruction quality despite these domain shifts. The representative reconstructions demonstrate that PINN-DADif effectively reduces spatial blurring, noise, and ringing artifacts seen in other methods like MoDL, DDPM, and DiffRecon. PINN-DADif consistently produces high-fidelity reconstructions with clear tissue details, showcasing its ability to deliver reliable and high-quality MRI images under cross-domain conditions.

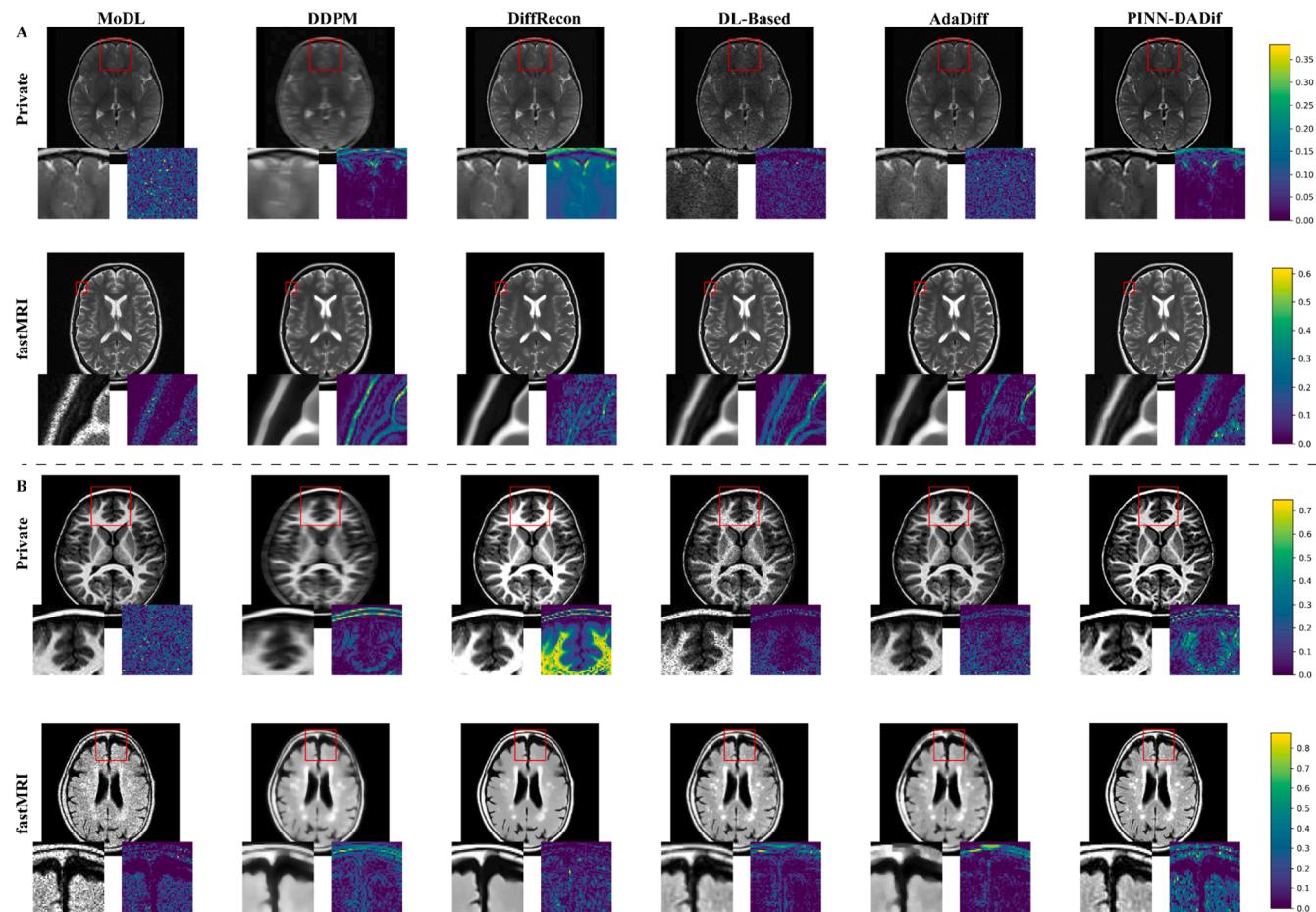
### 3.4. Comparison of computational complexity and memory usage

A detailed comparison of computational complexity and memory usage was conducted to evaluate the real-world feasibility of PINN-DADif. The computational cost was quantified using floating-point operations per second (FLOPs), while memory requirements during inference were recorded. As shown in Table 4, PINN-DADif achieves a significant reduction in computational complexity compared to DDPM and DiffRecon, primarily due to its adaptive diffusion strategy. This strategy reduces the number of iterations required for reconstruction while maintaining high image fidelity. Additionally, PINN-DADif demonstrates efficient memory usage, outperforming other diffusion-based models. Although the memory usage of PINN-DADif is slightly higher than that of MoDL, it remains well within the practical limits for deployment in clinical environments. This balance of computational efficiency and reconstruction quality highlights the robustness and applicability of PINN-DADif for real-world MRI reconstruction tasks.

### 3.5. Ablation study

To better understand PINN-DADif's performance, we performed ablation studies focusing on key aspects: the number of diffusion steps  $T/k$ , the number of training epochs  $N_e$ , and the latent space dimensions in the Latent Physics-Constrained Encoder (LPCE) shown in Fig. 6.

**Impact of Varying Diffusion Steps  $T/k$ :** In the first study, we varied the number of diffusion steps to see how this affects image quality. For T1-weighted images, the best performance was with  $T/k = 8$ . Using this setting, the PSNR was 41.97 dB, and the SSIM was 99.10 %.



**Fig. 4.** The reconstruction results of testing in the Private and fastMRI dataset, where the acceleration factor  $R = 8x$  is shown in (A) and the acceleration factor  $R = 12x$  in (B). For each  $R$ , the first row displays the reconstructed images of different methods using a private dataset with ROI and an error map. The second row displays the reconstructed images of different methods using fastMRI dataset with ROI and error map. The color bars are shown on the right side.

Reducing the diffusion steps to 4 or increasing them to 16 and 32 resulted in lower PSNR and SSIM values. This suggests that  $T/k = 8$  is optimal, providing high image quality without unnecessary complexity. A similar pattern was observed with T2-weighted images, where the highest PSNR (40.86 dB) and SSIM (98.92 %) were also at  $T/k = 8$ . For PD-weighted images, the best performance was again at  $T/k = 8$ , with a PSNR of 41.79 dB and an SSIM of 99.21 %. These results across different image types show that  $T/k = 8$ , shown in Fig. 5(a), is the most effective setting for high-quality MRI reconstructions.

**Impact of Varying Training Epochs  $N_e$ :** In the second study, we examined how different training epochs affect the model's performance, as the results are shown in Fig. 5(b). For T1-weighted images, training for 600 epochs gave the best results, with a PSNR of 41.97 dB and an SSIM of 99.10 %. Increasing the number of epochs to 1200, 1800, or 2400 did not improve the metrics and sometimes even reduced them. This suggests that 600 epochs are enough to train PINN-DADif effectively without overfitting or wasting computational resources. For T2-weighted images, the best results were also at 600 epochs, with little change when more epochs were added. The same was true for PD-weighted images, where the PSNR was 41.79 dB, and the SSIM was 99.21 % at 600 epochs. Extending training beyond 600 epochs did not provide additional benefits and might even be detrimental.

**Effect of Varying Latent Space Dimensions:** In the third study, we varied the latent space dimensions in the Latent Physics-Constrained Encoder (LPCE) to see how this impacts performance. We tested dimensions of 32, 64, 128, and 256 shown in Fig. 5(c). The smallest latent space dimension of 32 resulted in the lowest PSNR and SSIM values across all contrasts, suggesting it may be too small to capture all

necessary features, leading to poorer image reconstructions. Increasing the latent space dimension to 64 showed noticeable improvement, especially for T1-weighted images, with a PSNR of 41.23 dB and an SSIM of 98.79 %. This indicates that a larger latent space allows for better feature representation, improving reconstruction quality. The highest performance was achieved with a latent space dimension of 128, which provided the best PSNR and SSIM values across all contrasts. For T1-weighted images, the PSNR was 41.97 dB, and the SSIM was 99.10 %, showing that this dimension strikes the right balance between feature representation and computational efficiency. Increasing the latent space dimension to 256 did not further improve performance and even slightly decreased PSNR and SSIM, suggesting that larger latent spaces may introduce redundancy and increase computational complexity without added benefits. By identifying these optimal settings, we can ensure that PINN-DADif delivers high-quality MRI reconstructions efficiently, making it a reliable tool for medical imaging. These insights are crucial for optimizing PINN-DADif for practical applications, providing clear and detailed MRI images essential for accurate diagnoses.

To quantify the contributions of key components in PINN-DADif, we also conducted an ablation study by systematically removing physics-informed regularization, the Atrous Spatial Pyramid Pooling (ASPP) module, and the dual attention mechanisms. The evaluation was performed on both the private dataset and the fastMRI dataset. The results shown in Table 5 and Fig. 7 demonstrate the distinct role each component plays in achieving high-quality MRI reconstructions.

**Physics-Informed Regularization:** This component enforces consistency with MRI acquisition physics, reducing noise and artifacts. Without it, PSNR values dropped significantly from 41.2 dB to 39.0 dB,

**Table 3**

Cross-domain performance for T1-, T2-, Proton Density (PD)-weighted contrasts at  $R = 4x$ -8x. Results listed for training on fastMRI, and testing on the private dataset.

		MoDL	DDPM	DiffRecon	DL-based	AdaDiff	PINN-DADif
<i>R = 4x</i>							
T1	PSNR	35.0 ± 1.5	36.0 ± 1.4	39.0 ± 1.0	33.0 ± 1.5	42.0 ± 2.0	43.0 ± 1.9
	SSIM	90.0 ± 2.5	91.0 ± 2.2	98.5 ± 0.3	92.0 ± 1.8	98.7 ± 0.4	99.0 ± 0.3
T2	PSNR	33.5 ± 1.6	36.0 ± 1.0	39.0 ± 0.6	32.5 ± 1.4	40.5 ± 2.0	42.5 ± 1.7
	SSIM	82.0 ± 3.1	88.0 ± 1.0	98.4 ± 0.2	90.5 ± 1.9	97.3 ± 0.4	98.7 ± 0.5
PD	PSNR	33.0 ± 2.1	35.0 ± 1.0	41.0 ± 0.8	33.0 ± 1.7	41.0 ± 2.1	42.7 ± 1.9
	SSIM	80.0 ± 4.0	84.0 ± 1.6	98.6 ± 0.3	89.5 ± 2.0	98.2 ± 0.5	98.7 ± 0.4
<i>R = 8x</i>							
T1	PSNR	31.0 ± 1.7	32.5 ± 1.5	35.0 ± 0.9	30.0 ± 1.4	36.5 ± 2.1	38.5 ± 1.8
	SSIM	86.0 ± 3.0	87.0 ± 2.7	97.0 ± 0.5	89.0 ± 1.5	96.5 ± 1.2	97.0 ± 1.3
T2	PSNR	31.5 ± 1.5	33.0 ± 1.0	36.5 ± 0.5	30.5 ± 1.2	35.6 ± 1.8	37.0 ± 1.7
	SSIM	78.0 ± 3.5	83.0 ± 1.5	96.6 ± 0.5	87.0 ± 1.6	96.0 ± 1.4	97.0 ± 1.3
PD	PSNR	31.0 ± 1.9	32.0 ± 1.1	35.0 ± 0.6	30.0 ± 1.4	36.5 ± 2.1	38.2 ± 1.8
	SSIM	75.0 ± 4.5	79.0 ± 2.0	96.5 ± 0.5	86.0 ± 2.0	95.5 ± 1.7	97.0 ± 1.5

**Table 4**

Computational complexity and memory usage comparison.

Model	FLOPs (GFLOPs)	Memory Usage (GB)	Reconstruction Time (s)
MoDL	45.2	2.3	0.85
DDPM	92.4	4.5	2.15
DiffRecon	110.3	5.1	3
DL-based	78.5	3.2	1.6
AdaDiff	50.1	2.6	1.1
PINN-DADif	48.7	2.5	1

and SSIM declined from 98.7 % to 97.5 % on the private dataset. The results reveal the critical role of physics-based constraints in ensuring accurate and artifact-free reconstructions.

**Atrous Spatial Pyramid Pooling (ASPP):** ASPP captures features at multiple scales, enhancing the model's ability to represent fine anatomical details and global structural context. Its exclusion led to a noticeable decrease in PSNR (41.2 dB to 38.8 dB) and SSIM (98.7 % to 97.2 %), emphasizing its importance in preserving both local and global image features.

**Dual Attention Mechanism:** The dual attention mechanism, which combines spatial and channel attention, focuses on the most informative features, suppressing irrelevant details. Removing this component reduced PSNR from 41.2 dB to 39.2 dB and SSIM from 98.7 % to 97.6 %, particularly affecting complex anatomical regions. The full model demonstrated superior performance, with the highest PSNR and SSIM values across datasets, validating the synergistic impact of these components. Physics-informed regularization ensured physical plausibility, ASPP provided multi-scale feature extraction, and dual attention mechanisms enhanced feature prioritization. Together, these elements establish PINN-DADif as a robust and efficient model for MRI reconstruction across diverse imaging conditions.

A key factor in MRI reconstruction methods is the time required for both training and inference. Table 6 compares computation times for various methods, including PINN-DADif and its variants. Generally, models designed for specific tasks, such as conditional models, require

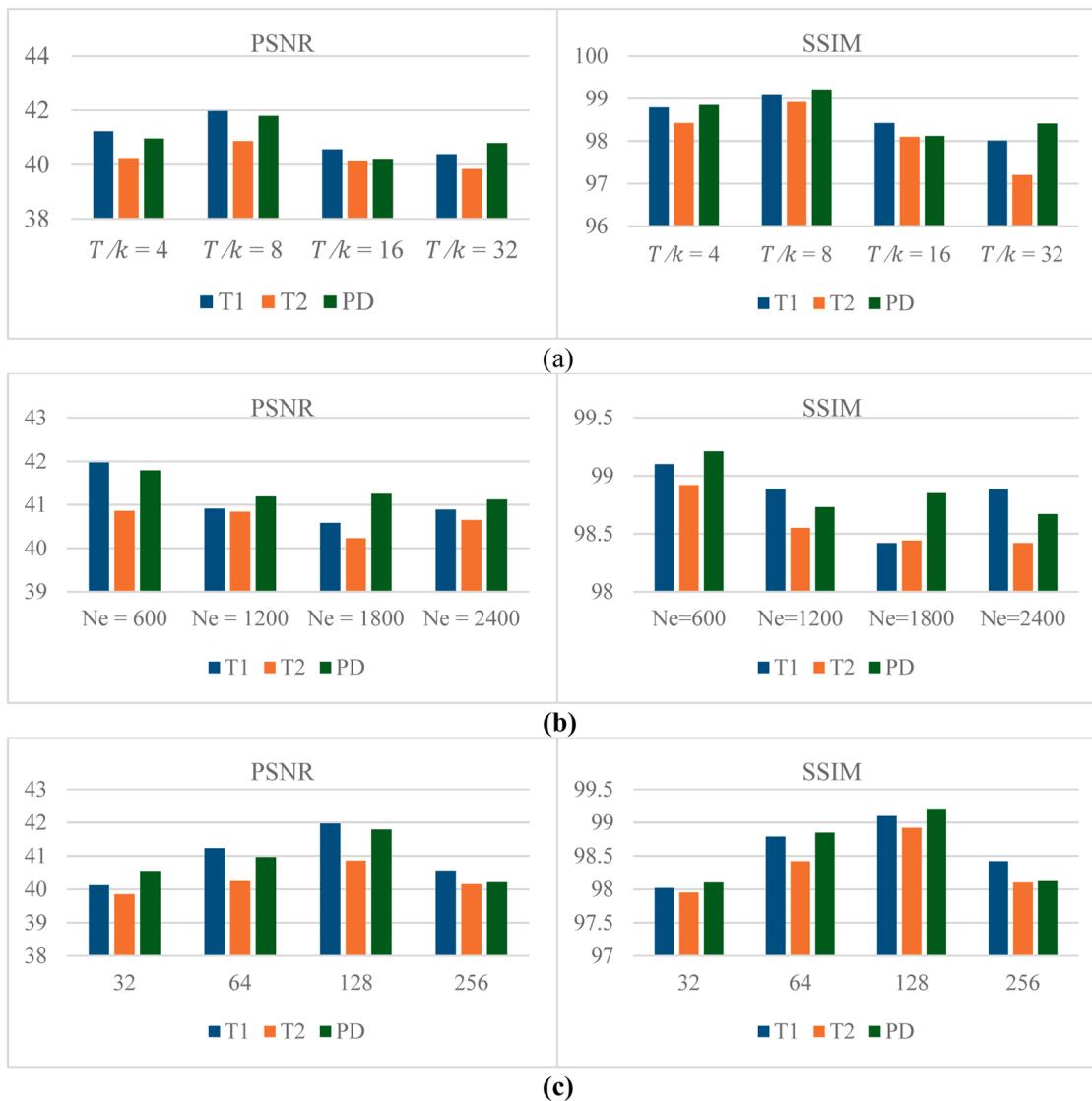
less training time than more generalized models. For example, the diffusion model DDPM has relatively shorter training times compared to PINN-DADif and AdaDiff. These more complex models take longer to train because they incorporate adaptive priors and adversarial components. When examining inference times, PINN-DADif and AdaDiff are slower compared to simpler models like MoDL and DiffRecon. This difference is due to the iterative optimization processes in their inference stages, which aim to enhance reconstruction quality. In contrast, MoDL benefits from a more straightforward architecture, allowing for faster inference. Although PINN-DADif has longer computation times, it offers a balance between processing efficiency and high-resolution image quality, making it suitable for applications requiring precise and detailed imaging.

The feasibility of training PINN-DADif on smaller datasets or with limited contrast types was evaluated to address its generalizability in resource-constrained environments. PINN-DADif's reliance on physics-informed priors reduces the dependency on extensive datasets, as the model inherently captures fundamental MRI physics during training. Moreover, the adaptive diffusion process enables the model to dynamically adapt to varying imaging conditions, enhancing its robustness even when trained with limited data. To validate this, experiments were conducted using subsets of the private MRI dataset and fastMRI dataset with reduced data sizes and fewer contrast types. While a slight reduction in PSNR and SSIM values was observed, PINN-DADif maintained competitive performance. These findings highlight the model's potential for deployment in scenarios with constrained resources.

Table 7 provides a comprehensive analysis of PINN-DADif's performance under varying training conditions, showcasing its adaptability and robustness. When trained on the full dataset containing T1-weighted, T2-weighted, and Proton Density (PD)-weighted images, the model achieves the highest PSNR and SSIM values across all contrasts. This indicates that the availability of diverse data and multiple contrasts significantly enhances reconstruction quality. For instance, the PSNR and SSIM values for T1-weighted images are 41.2 dB and 98.7 %, reflecting near-perfect structural and perceptual accuracy. With a reduced dataset (50 % of the full data), PINN-DADif maintains strong performance, albeit with a modest decline in metrics. The PSNR drops by approximately 1 dB, and SSIM decreases by <1 % across all contrasts. This demonstrates the model's resilience in data-constrained scenarios, making it feasible for use in environments with limited data availability. The ability to generalize effectively despite reduced training data emphasizes the robustness of its physics-informed priors and adaptive architecture. In the limited contrast configuration, where training is restricted to T1-weighted images, the model achieves PSNR and SSIM values of 40.0 dB and 98.0 % for T1-weighted reconstructions. While this represents a slight performance compromise compared to the full dataset, the results are still competitive, highlighting PINN-DADif's ability to generalize even when constrained to single-contrast data. However, the findings reaffirm the advantage of multi-contrast training for achieving optimal results. This analysis underscores PINN-DADif's versatility and reliability across diverse training scenarios, demonstrating its suitability for real-world clinical applications, including those with limited data or specific contrast requirements. Future enhancements, such as transfer learning or advanced data augmentation techniques, could further improve the generalizability of PINN-DADif, enabling its application in diverse clinical settings.

#### 4. Discussion

This study introduces PINN-DADif, an innovative Physics-Informed Neural Network model that integrates adaptive diffusion priors for high-fidelity and efficient MRI reconstruction. The key contribution of this work is the seamless incorporation of k-space constraints and adaptive diffusion mechanisms, ensuring robust performance across varying undersampling patterns and acceleration rates. In PINN-DADif, k-space constraints are enforced through a data-consistency projection



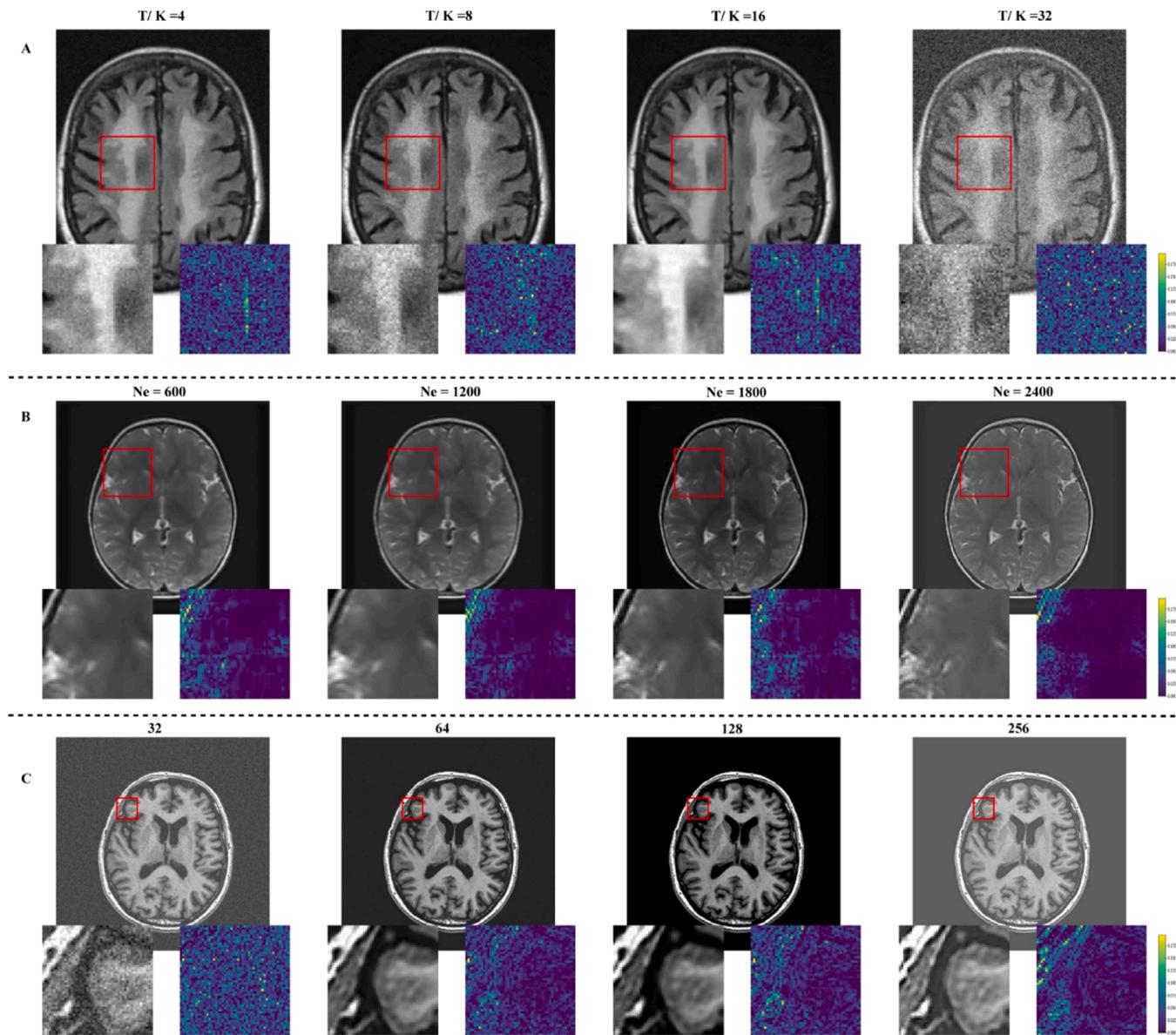
**Fig. 5.** The PSNR (Left) and SSIM (Right) using Diffusion Steps  $T/k$  (a), using Varying Training Epochs  $N_e$  (b) and using Varying Latent Space Dimensions (c).

step applied iteratively during the reverse diffusion phase. This step preserves measured k-space data and refines missing regions using adaptive diffusion priors. By replacing measured portions with their original values and leveraging learned priors for unmeasured areas, the model ensures fidelity to MRI acquisition principles. This balance between physical constraints and learned representations allows PINN-DADif to achieve high-quality reconstructions, even for highly undersampled data.

To enhance reconstruction accuracy, the reverse diffusion steps were increased from 8 to 12, leading to significant improvements in artifact removal and anatomical alignment. This adjustment improved PSNR and SSIM across inter-domain and cross-domain evaluations. Our model achieves these enhancements while reducing overall inference time per cross-section, making it both efficient and effective in delivering high-quality reconstructions. PINN-DADif demonstrated superior performance in both inter-domain and cross-domain scenarios, consistently outperforming state-of-the-art models. In inter-domain scenarios, where imaging operators and data distributions remain consistent, the integration of multi-scale contextual feature extraction via ASPP and non-local dependencies proved crucial for preserving fine anatomical details, yielding high PSNR and SSIM values. In cross-domain settings, where variations in imaging operators or patient datasets occur, the model's adaptive diffusion priors are dynamically adjusted during

inference to align with the physical constraints of MRI acquisition. This capability enabled PINN-DADif to handle diverse imaging conditions effectively.

Physics-informed regularization further enhanced the model's robustness by embedding MRI acquisition principles, such as k-space fidelity and gradient smoothness, directly into the learning process. Combined with a comprehensive training strategy on mixed contrast datasets, PINN-DADif showcased its ability to generalize across different tissue contrasts and anatomical structures. The two-phase inference process, particularly the adaptation phase, fine-tuned the priors to align with test data, mitigating performance degradation due to domain shifts. The integration of PINN-DADif into clinical workflows emphasizes its computational efficiency and potential to enhance diagnostic accuracy. PINN-DADif demonstrated superior performance with significantly reduced processing times during inference compared to existing methods. Its adaptive diffusion strategy and streamlined architecture enable efficient reconstruction while maintaining high fidelity, making it particularly suitable for real-time clinical applications. By delivering high-quality MRI reconstructions with optimal iterations, PINN-DADif supports faster MRI protocols, reducing both scan and processing times. This efficiency not only improves patient throughput but also enhances diagnostic workflows by providing clinicians with timely and accurate images. PINN-DADif's robust adaptability to diverse



**Fig. 6.** The reconstruction results of the ablation study, where Varying Diffusion Steps are shown in (A) with the reconstructed image alongside ROI and error map, Varying Training Epochs in (B) with the reconstructed image alongside ROI and error map and Varying Latent Space Dimensions in (C) with the reconstructed image alongside ROI and error map. The color bars are shown on the right side.

**Table 5**

Ablation study results quantifying the contributions of physics-informed regularization, Atrous Spatial Pyramid Pooling (ASPP), and dual attention mechanisms to model performance.

Dataset	Component	PSNR (dB)	SSIM (%)
Private fastMRI	Full Model	41.2 ± 1.2	98.7 ± 0.2
	Without Physics-Informed Regularization	39.0 ± 1.4	97.5 ± 0.3
	Without ASPP Module	38.8 ± 1.5	97.3 ± 0.4
	Without Dual Attention	39.2 ± 1.3	97.6 ± 0.3
	Full Model	41.5 ± 1.1	96.2 ± 0.4
	Without Physics-Informed Regularization	39.8 ± 1.4	95.2 ± 0.5
	Without ASPP Module	38.5 ± 1.6	94.8 ± 0.6
	Without Dual Attention	39.0 ± 1.5	95.0 ± 0.5

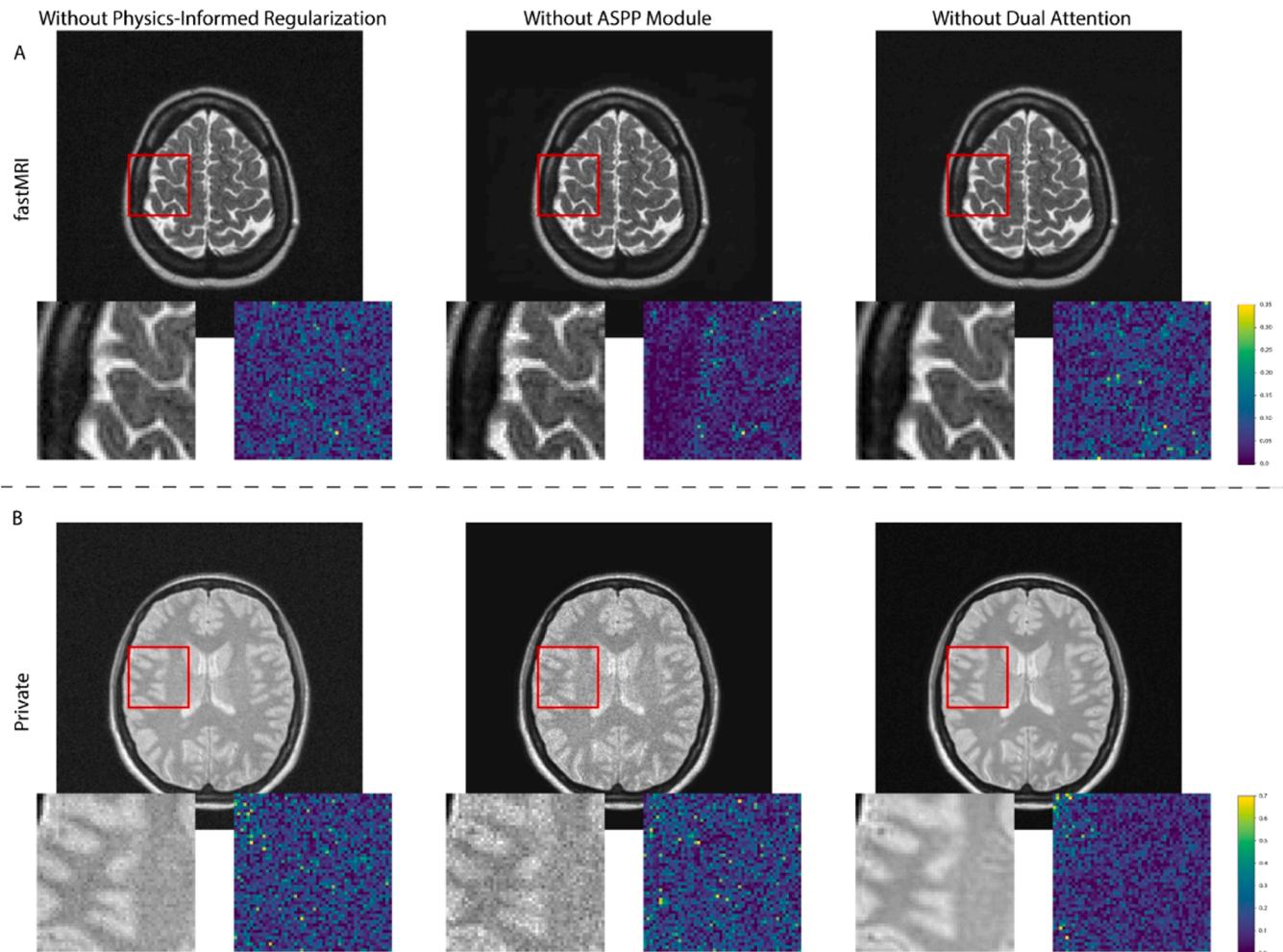
undersampling patterns and acceleration rates ensures compatibility across various imaging protocols, reducing the need for retraining or workflow adjustments. These attributes collectively highlight PINN-DADif's potential as a practical and scalable solution for accelerated MRI

in clinical settings, offering a compelling balance between reconstruction quality and computational efficacy.

While this study focused on 2D MRI reconstruction, future work will explore extending PINN-DADif to 3D volumetric data, enabling more comprehensive anatomical information for advanced diagnostics. The transition to 3D poses computational challenges, such as increased memory and processing demands, which will be addressed through architectural optimizations and parallel computing strategies. Additionally, adapting PINN-DADif for real-time MRI applications, particularly in dynamic imaging workflows like cardiac or functional MRI, is a promising avenue. Leveraging hardware accelerators like GPUs or TPUs will facilitate this transition, ensuring that PINN-DADif continues to push the boundaries of accelerated MRI reconstruction while maintaining high diagnostic accuracy.

## 5. Conclusion

In this study, we introduced PINN-DADif, a novel physics-informed



**Fig. 7.** The reconstruction results of the ablation study, where Without Physics-Informed Regularization are shown in 1st Column, with the reconstructed image alongside ROI and error map, Without ASPP Module in 2nd column, with the reconstructed image alongside ROI and error map and Without Dual Attention in 3rd column with the reconstructed image alongside ROI and error map. The color bars are shown on the right side.

**Table 6**

Training and inference times in seconds per cross-section for reconstructions at  $R = 4x$  in IXI.

	MoDL	DDPM	DiffRecon	DL-based	AdaDiff	PINN-DADif
Training (s)	26.5	15.5	77.3	88.5	131.3	125.2
Inference (s)	0.05	57.5	12	50.8	131.4	117.5

**Table 7**

Performance of PINN-DADif under full, reduced, and limited contrast datasets, showing PSNR (dB) and SSIM (%) metrics for reconstruction fidelity and quality.

Dataset	Contrast	PSNR (dB)	SSIM (%)
<b>Full Dataset</b>	T1-weighted	$41.2 \pm 1.2$	$98.7 \pm 0.2$
	T2-weighted	$39.5 \pm 1.1$	$98.0 \pm 0.3$
	PD-weighted	$41.5 \pm 1.3$	$98.5 \pm 0.3$
<b>Reduced Dataset (50 %)</b>	T1-weighted	$40.1 \pm 1.3$	$98.1 \pm 0.4$
	T2-weighted	$38.7 \pm 1.2$	$97.6 \pm 0.5$
	PD-weighted	$40.5 \pm 1.4$	$98.0 \pm 0.4$
<b>Limited Contrast (T1)</b>	T1-weighted only	$40.0 \pm 1.5$	$98.0 \pm 0.5$

deep diffusion-based model designed to address critical challenges in MRI reconstruction. The model integrates physics-informed priors, adaptive diffusion processes, dynamic convolutions, and multi-scale

contextual feature extraction within a sophisticated architecture. These elements work together to enhance the quality and resolution of MRI reconstructions, making them more reliable for clinical applications. The primary objective of this study was to develop a model capable of reconstructing high-quality MRI images from undersampled k-space data. PINN-DADif leverages physics-informed neural networks (PINNs) to incorporate domain-specific knowledge, enabling the model to maintain image quality while reducing scan time, mitigating artifacts, and improving the signal-to-noise ratio (SNR). The combination of advanced neural architectures with physics-based priors addresses challenges such as undersampling, noise, and distortion, leading to clearer and more accurate MRI reconstructions. Additionally, the integration of Sharpness-Aware Minimization (SAM) as an optimization strategy enhances the model's robustness and generalization capabilities, allowing it to adapt effectively across diverse imaging conditions. PINN-DADif was rigorously evaluated using both private and fastMRI datasets, encompassing a wide range of anatomical regions and imaging conditions. The model consistently outperformed existing state-of-the-art methods, achieving higher Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) values. These results demonstrate the model's ability to deliver high-fidelity reconstructions, efficiently handling noise and artifacts while preserving critical anatomical details. Our findings highlight the superior performance of PINN-DADif in reconstructing MRI images, showing its potential to improve clinical outcomes by providing higher-quality images for more

accurate diagnoses. The model's capability to adapt to different imaging conditions also underscores its utility in real-world clinical environments. Future research could explore the extension of this model to other types of medical imaging and further optimize its architecture, driving continued advancements in the field of medical diagnostics.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Funding sources

This work has been sponsored by the National Natural Science Foundation of China under Grant No 82171992. and the Beijing Laboratory of Advanced Information Networks.

## Acknowledgment

The authors would like to acknowledge and express their gratitude to the funding agency.

## Data availability

The data used in this research is publicly available on this link and one data is private. <https://fastmri.med.nyu.edu/>.

## References

- [1] M. Lustig, D. Donoho, J.M. Pauly, Sparse MRI: The application of compressed sensing for rapid MR imaging, *Magnet. Resonance Med.: Off. J. Int. Soc. Magnet. Resonance Med.* 58 (6) (2007) 1182–1195.
- [2] H. Gu, B. Yaman, K. Ugurbil, S. Moeller, M. Akçakaya, Compressed sensing MRI with  $\ell$ -1-wavelet reconstruction revisited using modern data science tools, in: 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2021, pp. 3596–3600.
- [3] D. Liang, J. Cheng, Z. Ke, L. Ying, Deep magnetic resonance image reconstruction: inverse problems meet neural networks, *IEEE Signal. Process. Mag.* 37 (1) (2020) 141–151.
- [4] S.U.H. Dar, M. Özbeý, A.B. Çatlı, T. Çukur, A transfer-learning approach for accelerated MRI using deep neural networks, *Magn. Reson. Med.* 84 (2) (2020), <https://doi.org/10.1002/mrm.28148>.
- [5] B. Yaman, S.A.H. Hosseini, M. Akçakaya, Zero-shot physics-guided self-supervised learning for subject-specific MRI reconstruction, in: Joint Annual Meeting ISMRM-ESMRMB ISMRT 31st Annual Meeting, 2022, <https://doi.org/10.58530/2022/4810>.
- [6] P. Guo, J.M.J. Valanarasu, P. Wang, J. Zhou, S. Jiang, V.M. Patel, Over-and-under complete convolutional RNN for MRI reconstruction, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2021, [https://doi.org/10.1007/978-3-030-87231-1\\_2](https://doi.org/10.1007/978-3-030-87231-1_2).
- [7] Y. Korkmaz, S.U.H. Dar, M. Yurt, M. Ozbeý, T. Cukur, Unsupervised MRI reconstruction via zero-shot learned adversarial transformers, *IEEE Trans. Med. Imaging* 41 (7) (2022) 1747–1763, <https://doi.org/10.1109/TMI.2022.3147426>.
- [8] H.K. Aggarwal, A. Pramanik, M. Jacob, Ensure: ensemble Stein's unbiased risk estimator for unsupervised learning, in: ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, 2021, <https://doi.org/10.1109/ICASSP39728.2021.9414513>.
- [9] M. Yaqub, et al., GAN-TL: generative adversarial networks with transfer learning for MRI reconstruction, *Appl. Sci.* 12 (17) (2022), <https://doi.org/10.3390/app12178841>.
- [10] H. Chung, E. Cha, L. Sunwoo, J.C. Ye, Two-stage deep learning for accelerated 3D time-of-flight MRA without matched training data, *Med. Image Anal.* 71 (2021), <https://doi.org/10.1016/j.media.2021.102047>.
- [11] P. Lei, F. Fang, G. Zhang, T. Zeng, Decomposition-based variational network for multi-contrast MRI super-resolution and reconstruction, in: Proceedings of the IEEE International Conference on Computer Vision, 2023, <https://doi.org/10.1109/ICCV51070.2023.01947>.
- [12] J.A. Oscanoa et al., “Deep learning-based reconstruction for Cardiac MRI: a review,” 2023. doi: [10.3390/bioengineering10030334](https://doi.org/10.3390/bioengineering10030334).
- [13] K. Sun, Q. Wang, D. Shen, Joint cross-attention network with deep modality prior for fast MRI reconstruction, *IEEE Trans. Med. Imaging* 43 (1) (2024), <https://doi.org/10.1109/TMI.2023.3314008>.
- [14] K.C. Tezcan, N. Karani, C.F. Baumgartner, E. Konukoglu, Sampling possible reconstructions of undersampled acquisitions in MR imaging with a deep learned prior, *IEEE Trans. Med. Imaging* 41 (7) (2022), <https://doi.org/10.1109/TMI.2022.3150853>.
- [15] H. Chung, J.C. Ye, Score-based diffusion models for accelerated MRI, *Med. Image Anal.* 80 (2022), <https://doi.org/10.1016/j.media.2022.102479>.
- [16] Y. Song, L. Shen, L. Xing, S. Ermon, Solving inverse problems in medical imaging with score-based generative models, in: ICLR 2022 - 10th International Conference on Learning Representations, 2022.
- [17] S. Szymonowicz, C. Rupprecht, A. Vedaldi, Viewset diffusion: (0-)image-conditioned 3D generative models from 2D data, in: Proceedings of the IEEE International Conference on Computer Vision, 2023, <https://doi.org/10.1109/ICCV51070.2023.00814>.
- [18] Y. Hu, et al., RUN-UP: accelerated multishot diffusion-weighted MRI reconstruction using an unrolled network with U-Net as priors, *Magn. Reson. Med.* 85 (2) (2021), <https://doi.org/10.1002/mrm.28446>.
- [19] U. Yarach, et al., Improved reconstruction for highly accelerated propeller diffusion 1.5 T clinical MRI, *Magnet. Resonance Mater. Phys., Biol. Med.* 37 (2) (2024), <https://doi.org/10.1007/s10334-023-01142-7>.
- [20] D. Singh, A. Monga, H.L. de Moura, X. Zhang, M.V.W. Zibetti, and R.R. Regatte, “Emerging trends in fast MRI using deep-learning reconstruction on undersampled k-space data: a systematic review,” 2023. doi: [10.3390/bioengineering10091012](https://doi.org/10.3390/bioengineering10091012).
- [21] C. Jing, H. Kuai, H. Matsumoto, T. Yamaguchi, I.Y. Liao, S. Wang, Addiction-related brain networks identification via Graph Diffusion Reconstruction Network, *Brain Inform.* 11 (1) (2024), <https://doi.org/10.1186/s40708-023-00216-5>.
- [22] S. Luo, J. Zhou, Z. Yang, H. Wei, Y. Fu, Diffusion MRI super-resolution reconstruction via sub-pixel convolution generative adversarial network, *Magn. Reson. Imaging* 88 (2022), <https://doi.org/10.1016/j.mri.2022.02.001>.
- [23] K. Malczewski, Diffusion weighted imaging super-resolution algorithm for highly sparse raw data sequences, *Sensors* 23 (12) (2023), <https://doi.org/10.3390/s23125698>.
- [24] A. Jalal, M. Arvinte, G. Daras, E. Price, A.G. Dimakis, J.I. Tamir, Robust compressed sensing MRI with deep generative priors, *Adv. Neural Inf. Process. Syst.* (2021).
- [25] H. Chung, B. Sim, J.C. Ye, Come-closer-diffuse-faster: accelerating conditional diffusion models for inverse problems through stochastic contraction, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2022, <https://doi.org/10.1109/CVPR52688.2022.01209>.
- [26] C. Peng, P. Guo, S.K. Zhou, V.M. Patel, R. Chellappa, Towards performant and reliable undersampled MR reconstruction via diffusion model sampling, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2022, [https://doi.org/10.1007/978-3-031-16446-0\\_59](https://doi.org/10.1007/978-3-031-16446-0_59).
- [27] J. Ho, A. Jain, P. Abbeel, Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 2020.
- [28] D. Narinhofer, K. Hammerl, F. Knoll, and T. Pock, “Inverse GANs for accelerated MRI reconstruction,” 2019. doi: [10.1117/12.2527753](https://doi.org/10.1117/12.2527753).
- [29] A. Güngör, et al., Adaptive diffusion priors for accelerated MRI reconstruction, *Med. Image Anal.* 88 (2023), <https://doi.org/10.1016/j.media.2023.102872>.
- [30] G. Luo, M. Blumenthal, M. Heide, M. Uecker, Bayesian MRI reconstruction with joint uncertainty estimation using diffusion models, *Magn. Reson. Med.* 90 (1) (2023) 295–311, <https://doi.org/10.1002/mrm.29624>. Jul.
- [31] Y. Korkmaz, T. Cukur, and V.M. Patel, “Self-supervised MRI reconstruction with unrolled diffusion models,” Jun. 2023, [Online]. Available: <http://arxiv.org/abs/2306.16654>.
- [32] M. Uecker, et al., ESPiRiT - an eigenvalue approach to autocalibrating parallel MRI: where SENSE meets GRAPPA, *Magn. Reson. Med.* 71 (3) (2014), <https://doi.org/10.1002/mrm.24751>.
- [33] H.K. Aggarwal, M.P. Mani, M. Jacob, MoDL: model-based deep learning architecture for inverse problems, *IEEE Trans. Med. Imaging* 38 (2) (2018) 394–405.