Adaptive Diffusion Priors for Accelerated MRI Reconstruction

Alper Gungör\textsuperscript{a,b,c,f}, Salman UH Dar\textsuperscript{a,b,f}, Şaban Öztürk\textsuperscript{a,b,d,f}, Yılmaz Korkmaz\textsuperscript{a,b}, Hasan A Bedel\textsuperscript{a,b}, Gokberk Elmas\textsuperscript{a,b}, Muzaffer Oz bey\textsuperscript{a,b}, Tolga Çukur\textsuperscript{a,b,e,}∗

\textsuperscript{a}Department of Electrical and Electronics Engineering, Bilkent University, Ankara 06800, Turkey
\textsuperscript{b}National Magnetic Resonance Research Center (UMRAM), Bilkent University, Ankara 06800, Turkey
\textsuperscript{c}ASELSAN Research Center, Ankara 06200, Turkey
\textsuperscript{d}Department of Electrical and Electronics Engineering, Amasya University, Amasya 05100, Turkey
\textsuperscript{e}Neuroscience Program, Bilkent University, Ankara 06800, Turkey
\textsuperscript{f}denotes equal contribution

A R T I C L E I N F O
diffusion, adaptive, MRI, reconstruction, generative, image prior

A B S T R A C T

Deep MRI reconstruction is commonly performed with conditional models that de-alias undersampled acquisitions to recover images consistent with fully-sampled data. Since conditional models are trained with knowledge of the imaging operator, they can show poor generalization across variable operators. Unconditional models instead learn generative image priors decoupled from the operator to improve reliability against domain shifts related to the imaging operator. Recent diffusion models are particularly promising given their high sample fidelity. Nevertheless, inference with a static image prior can perform suboptimally. Here we propose the first adaptive diffusion prior for MRI reconstruction, AdaDiff, to improve performance and reliability against domain shifts. AdaDiff leverages an efficient diffusion prior trained via adversarial mapping over large reverse diffusion steps. A two-phase reconstruction is executed following training: a rapid-diffusion phase that produces an initial reconstruction with the trained prior, and an adaptation phase that further refines the result by updating the prior to minimize data-consistency loss. Demonstrations on multi-contrast brain MRI clearly indicate that AdaDiff outperforms competing conditional and unconditional methods under domain shifts, and achieves superior or on par within-domain performance.

1. Introduction

Magnetic resonance imaging (MRI) is a preferred modality in diagnostic applications due to its exceptional soft-tissue contrast, yet canonically long exams hinder its clinical use. A fundamental solution is to shorten scan times by undersampling k-space acquisitions and solve an ill-posed inverse problem to reconstruct images (Lustig et al., 2007; Gu et al., 2021). In recent years, deep learning methods have become a gold standard in MRI reconstruction, given their ability to solve complex inverse problems based on data-driven priors (Liang et al., 2020; Yang et al., 2016; Wang et al., 2016; Hammernik et al., 2017; Schlemper et al., 2017; Dar et al., 2020b; Kwon et al., 2017; Yang et al., 2020). Many proposed methods are based on conditional models that process undersampled acquisitions provided as input to recover output images that are consistent with fully-sampled acquisitions (Zhu et al., 2018; Aggarwal et al., 2019; Hyun et al., 2018; Lee et al., 2018; Knoll et al., 2019; Yoon et al., 2018; Ye et al., 2018; Quan et al., 2018; Yu et al., 2018; Adler and Oktem, 2018; Guo et al., 2021). This conditional mapping can be learned explicitly from a training dataset of paired undersampled and fully-sampled acquisitions (Dar et al., 2020a; Mardani et al., 2019; Biswas et al., 2019; Wang et al., 2019; Yang et al., 2020; Eo et al., 2018; Cheng et al., 2019; Qin et al., 2019; Hosseini et al., 2020). To alleviate requirements on training data, the mapping can also be learned implicitly on undersampled acquisitions via self-supervision (Tamir et al., 2019; Wang et al., 2020a; Cole et al., 2020; Yaman et al., 2020; Huang et al., 2019; Liu et al., 2020; Aggarwal et al., 2021) or cycle-consistency approaches (Quan et al., 2018; Oh et al., 2020; Lei et al., 2021; Chung et al., 2021; Lu et al., 2021). Regardless of the learning strategy, conditional models capture a de-aliasing prior to suppress undersampling artifacts, so they have explicit knowledge of the imaging operator that reflects the choice of sampling patterns and coil sensitivities for acceleration (Polak et al., 2020; Feng et al., 2021; Küstner et al., 2020; Sriram et al., 2020). Deep reconstruction models are commonly...
trained based on a relatively standardized imaging operator to help maximize performance (Knoll et al., 2020). However, in the testing stage, an end user might have to prescribe spontaneous changes to the imaging operator (e.g. changes in acceleration rate, sampling density or number of coils) in order to meet practical considerations on image quality or scan time. Since the imaging operator inherently determines the characteristics of aliasing artifacts in undersampled acquisitions, such domain shifts in the operator can compromise reconstruction performance and necessitate re-training of conditional models (Narnhofer et al., 2019; Liu et al., 2021). To avoid potential losses in generalization performance, deep reconstruction models that are resilient against variations in the imaging operator between the training and test sets are direly needed.

An alternative framework employs unconditional models that are not trained to perform the reconstruction task (i.e., mapping undersampled to fully-sampled data), but instead to capture generative image priors through auxiliary tasks such as additive noise removal (Ahmad et al., 2020), image autoencoding (Tezcan et al., 2019; Liu et al., 2020; Tezcan et al., 2022) or image generation (Narnhofer et al., 2019; Darestani and Heckel, 2021; Luo et al., 2020; Korkmaz et al., 2022; Elmas et al., 2022). Image priors are only combined with the imaging operator during inference, so they improve generalization against variable operators as they are agnostic to undersampling (Tezcan et al., 2019; Narnhofer et al., 2019). Adversarial priors are particularly prominent as they offer elevated sensitivity to detailed tissue structure (Narnhofer et al., 2019; Korkmaz et al., 2022), but they might manifest poor diversity in generated image samples (Dhariwal and Nichol, 2021). As a promising surrogate, diffusion models enhance sample diversity while maintaining comparable sample quality (Ho et al., 2020). Recent studies have reported remarkable reconstructions with alternated projections through diffusion priors and through the imaging operator to enforce consistency to acquired data (Jalal et al., 2021; Chung and Ye, 2022; Song et al., 2022; Chung et al., 2022; Luo et al., 2022; Xie and Li, 2022; Peng et al., 2022). Still, static diffusion priors can limit model performance under domain shifts in the MR image distribution, which can result from changes in the pulse sequence or scanner settings.

Here we introduce a novel diffusion-based method, AdaDiff, to improve performance and reliability against domain shifts in accelerated MRI reconstruction. AdaDiff learns an unconditional diffusion prior for high-fidelity image generation (Fig. 1), and adapts the diffusion prior during inference for enhanced performance (Fig. 2). Vanilla diffusion models generate images through a long sequence of inference steps, resulting in prolonged image sampling (Ho et al., 2020). We instead propose a diffusion model based on an adversarial mapper to generate images in few, large reverse diffusion steps for a notable speed up in image sampling. A two-phase reconstruction is then employed during inference: a rapid-diffusion phase that produces an initial reconstruction by fast image sampling based on the trained prior, and an adaptation phase that produces a refined reconstruction by updating the prior to minimize data-consistency loss. The adversarial diffusion prior enables AdaDiff to reconstruct high-quality images in fewer inference iterations compared to static, untrained or non-adversarial diffusion priors.

The proposed method is demonstrated for reconstruction of multiple contrasts in brain MRI based on a unified model trained on mixed contrasts. Experiments are reported for within-domain cases with matched imaging operator and image distribution between training-test sets, and cross-domain cases with a domain shift in the operator or the MR image distribution. Comparisons are provided against state-of-the-art traditional, conditional and unconditional deep models. In general, AdaDiff achieves superior or on par performance in within-domain cases, and outperforms competing methods in cross-domain cases. Code to implement AdaDiff is available at https://github.com/icon-lab/AdaDiff.

Contributions

- To our knowledge, AdaDiff is the first prior adaptation method based on diffusion models in literature for accelerated MRI reconstruction.
- The proposed method leverages a rapid diffusion process with an adversarial mapper for efficient sampling from the diffusion prior.
- Inference adaptation is performed on the trained diffusion prior to improve performance and reliability against domain shifts.

2. Related Work

Unconditional models that decouple the image prior from the imaging operator promise enhanced generalization in MRI reconstruction (Narnhofer et al., 2019; Tezcan et al., 2019; Liu et al., 2020; Korkmaz et al., 2022; Ahmad et al., 2020; Darestani and Heckel, 2021; Luo et al., 2020). In this generative modeling (GM) framework, image priors are typically learned to capture information regarding the distribution of high-quality MRI data (Narnhofer et al., 2019; Tezcan et al., 2019; Liu et al., 2020; Korkmaz et al., 2022; Elmas et al., 2022; Luo et al., 2020). A common approach rests on generative adversarial networks (GANs) that indirectly characterize the data distribution (Narnhofer et al., 2019; Liu et al., 2020; Korkmaz et al., 2022; Elmas et al., 2022). Despite the realism of generated image samples, GAN-based methods can be susceptible to low representational diversity that can hamper reconstruction performance.

A recent class of GMs based on diffusion promise enhanced representational diversity over GANs. Diffusion models use a multi-step process to gradually transform Gaussian noise into image samples. Unlike GANs, they directly characterize correlates of the data distribution (e.g., derivative or lower bound of log-likelihood). The learned priors are coupled with the imaging operator at time of inference (Jalal et al., 2021; Chung and Ye, 2022; Song et al., 2022; Chung et al., 2022; Luo et al., 2022; Peng et al., 2022). Reconstruction can then be performed via repeated projections through the diffusion prior and the operator. Projections through the diffusion prior involve generation of image samples. For instance, Jalal et al. (2021); Luo et al. (2022) proposed sampling with score-based functions and Langevin dynamics; Chung and Ye (2022); Song et al. (2022)
used a predictor to solve a stochastic differential equation followed by Langevin sampling. High image quality has typically been reported with diffusion-based MRI reconstruction (Peng et al., 2022).

Despite their prowess, diffusion-based methods are not without limitation. Vanilla diffusion models use hundreds of reverse steps for image generation (Jalal et al., 2021), elevating computational burden. Peng et al. (2022) considered rescaling the diffusion step size during inference to accelerate image sampling, but this can potentially reduce the accuracy of reverse diffusion steps (Ho et al., 2020). Chung et al. (2022) proposed to obtain an initial reconstruction via a separate method, and then initiate the reverse diffusion process with this initial reconstruction for faster inference. While promising, this approach implements a second reconstruction method. Furthermore, existing diffusion methods learn a static prior that is kept fixed during inference. In turn, a trained prior might be rendered suboptimal by domain shifts in the operator.

3. Theory

3.1. MRI Reconstruction

Accelerated MRI entails recovery of a subject’s MR image $x$ from undersampled k-space acquisitions $y$:

$$Ax = y$$ (1)

where $A = \Omega F B$ is the imaging operator that captures the influence of the k-space undersampling pattern ($\Omega$) and coil sensitivities ($B$), and $F$ denotes Fourier transform. Since the inverse problem in Eq. 1 is ill-posed, prior information is typically incorporated to obtain a reconstruction $\hat{x}$:

$$\hat{x} = \min_x ||Ax - y||_2 + R(x, y)$$ (2)

where $R(x, y)$ denotes the regularization term that enforces the prior. Given a training set of MRI data, conditional models capture a de-aliasing prior often conditioned on the inverse Fourier transform of undersampled data as input $R(x/F^{-1}(y))$. Instead, unconditional models capture a generative image prior agnostic to undersampling, $R(x)$. Since image priors are not tied to specific imaging operators, they promise improved reliability against domain shifts in the operator.

3.2. Diffusion Models

Diffusion models are likelihood-based GMs that express image generation as a temporal Markov process (Fig. 1a). Modeling involves forward and reverse processes that conventionally comprise hundreds of steps (Ho et al., 2020). The forward process adds a small amount of isotropic Gaussian noise $\epsilon \sim N(0, I)$ in each step to modify an actual image $x_0 \sim q(x_0)$ at time step 0, and produce a sequence of noisy samples $x_{1:T}$ where $T$ is the final time step. At step $t$, the relationship between $x_t$ and $x_{t-1}$, and the corresponding conditional distribution $q(x_t|x_{t-1})$ can be described as follows:

$$x_t = \sqrt{1-\beta_t} x_{t-1} + \sqrt{\beta_t} \epsilon,$$ (3)

$$q(x_t|x_{t-1}) = N(x_t; \sqrt{1-\beta_t} x_{t-1}, \beta_t I)$$ (4)

where $\beta_t$ is the noise scaling. After a large number of forward steps, $x_T$ approaches an isotropic Gaussian sample.

The reverse process gradually removes the added noise in $x_T$ to recollect $x_0$. Diffusion models operationalize each reverse step as $x_t = f\theta(x_{t+1})$, where $f\theta$ denotes projection through a network with parameters $\theta$. The network can be trained to minimize a lower bound on the negative log-likelihood:

$$L_{ab} = \sum_{t=0}^{T-1} D_{KL}(q(x_t|x_{t+1}) \parallel p_\theta(x_t|x_{t+1}))$$ (5)

where $p_\theta(x_t|x_{t+1})$ denotes the parametrization of $q(x_t|x_{t+1})$, and $D_{KL}$ is the Kullback-Leibler (KL) divergence. Note that $q(x_t|x_{t+1})$ is generally unknown, so Eq. 5 cannot be evaluated and an alternative formulation is adopted:

$$L_{ab} = -\log p_\theta(x_0|x_1) + \sum_{t=1}^{T-1} D_{KL}(q(x_t|x_{t+1}, x_0) \parallel p_\theta(x_t|x_{t+1}))$$ (6)

where the auxiliary distribution $q(x_t|x_{t+1}, x_0)$ has a closed-form expression (Ho et al., 2020). For small step sizes, $q(x_t|x_{t+1}, x_0) \approx q(x_t|x_{t-1})$, so a generator network can learn the desired mapping by minimizing Eq. 6. A common approach to minimize $L_{ab}$ is to predict the additive noise $\epsilon$ given $x_{t+1}$ and $x_0$ as input (Ho et al., 2020):

$$\min_\theta \mathbb{E} \left\| \sqrt{\beta_{t+1}} \epsilon - (x_{t+1} - \sqrt{1-\beta_{t+1}} f_\theta(x_{t+1})) \right\|_2$$ (7)

where $\mathbb{E}$ is the expectation of the difference norm between the scaled noise added onto $x_{t+1}$ in the forward step and the noise predicted by the network in the reverse step.

A trained diffusion model can be used to generate random image samples during inference. Starting from a noise image $x_T \sim N(0, I)$, hundreds of reverse steps are performed using $p_\theta(x_t|x_{t+1})$ to obtain an image sample $x_0$. For MRI reconstruction, reverse diffusion projections are data-consistency projections to align the generated image sample with the acquired k-space data for each subject (Luo et al., 2022; Peng et al., 2022). However, recent reconstruction methods use static diffusion priors that can potentially elicit suboptimal performance.
3.3. AdaDiff
Here we propose to perform prior adaptation on a diffusion model for enhanced performance in MRI reconstruction. AdaDiff is trained to efficiently generate high-quality image samples via a rapid diffusion process with substantially fewer steps than typically used (Fig. 1b). Vanilla diffusion models assume approximate normality for reverse transition probabilities \( q(x_t|x_{t+1}) \), but this assumption breaks down with increasing step size (Ho et al., 2020; Xiao et al., 2022). To improve accuracy, we propose an adversarial mapper to parametrize the reverse diffusion steps as inspired by a recent study on natural image synthesis (Xiao et al., 2022). Given a trained diffusion prior, a two-phase reconstruction is employed to recover a subject’s images during inference (Fig. 2). In the rapid diffusion phase, an initial reconstruction is obtained via interleaved reverse diffusion and data-consistency projections. In the prior adaptation phase, the prior is combined with the imaging operator to evaluate a data-consistency loss based on the difference between generated and measured k-space data. The parameters of the prior are iteratively updated to minimize the data-consistency loss. The image generated by the adapted prior is taken as the final reconstruction.

3.3.1. Training of the Prior
Vanilla diffusion models prescribe small step sizes to approximate \( q(x_t|x_{t+1}) \) with an auxiliary Gaussian distribution, necessitating computationally-intensive inference with a large number of diffusion steps. We instead adopt a rapid adversarial diffusion model with a large step size of \( k \) as described in (Xiao et al., 2022).
et al., 2022):

\[ x_t = \sqrt{1 - \gamma_t} x_{t-k} + \sqrt{\gamma_t} z, \]  \quad \text{(8)}

\[ q(x_t|\{x_{t-k}\}) = \mathcal{N}(x_t; \sqrt{1 - \gamma_t} x_{t-k}, \gamma_t I) \]  \quad \text{(9)}

where the noise variance \( \gamma_t \) has to be greater than \( \beta_t \) to compensate for large \( k \). Because the normality assumption does not hold for \( q(x_t|\{x_{t-k}\}) \), the original likelihood formulation must be considered for the lowerbound:

\[ L_B = \sum_{t=0}^{T-k} D_{KL}(q(x_t|x_{t+k}) \parallel p_0(x_t|x_{t+k})) \]  \quad \text{(10)}

There is no closed-form expression for \( q(x_t|\{x_{t+k}\}) \), so we introduce an adversarial mapper to implicitly capture the conditional distribution for the reverse diffusion steps. A generator \( G_{\theta_i} \) is used to parametrize a sampling distribution \( p_{\theta_i}(x_t|x_{t+k}) \) that synthesizes \( \hat{x}_t \). Meanwhile, a discriminator \( D_{\theta_j} \) differentiates between synthetic samples (\( \hat{x}_t \)) drawn from \( p_{\theta_j}(x_t|x_{t+k}) \) and actual samples (\( x_t \)) drawn from the true denoising distribution \( q(x_t|x_{t+k}) \). Both \( G_{\theta_i} \) and \( D_{\theta_j} \) receive time index \( t+k \) as input. In this framework, the discriminator is trained to minimize an adversarial loss (Dar et al., 2019) coupled with a gradient penalty to improve learning (Mescheder et al., 2018):

\[ L_D = \sum_{t=0}^{T-k} [\mathbb{E}_{q(x_t|x_{t+k})} \left[- \log (D_{\theta_j}(x_t, x_{t+k}, t + k))\right] \\
+ \mathbb{E}_{p_{\theta_i}(x_t|x_{t+k})} \left[- \log (1 - D_{\theta_j}(\hat{x}_t, x_{t+k}, t + k))\right] \\
+ \mathbb{E}_{q(x_t|x_{t+k})} \left[\frac{1}{2} \| \nabla_x D_{\theta_j}(x_t, x_{t+k}, t) \|_2^2 \right]] \]  \quad \text{(11)}

To avoid saturation, the generator is trained accordingly to maximize the following loss function:

\[ L_G = \sum_{t=0}^{T-k} [\mathbb{E}_{q(x_t|x_{t+k})} p_{\theta_i}(x_t|x_{t+k}) \left[- \log (D_{\theta_j}(\hat{x}_t, x_{t+k}, t + k))\right]] \]  \quad \text{(12)}

While the first and third terms in Eq. 11 require sampling from the unknown \( q(x_t|x_{t+k}) \), an equivalent formulation can be de-
rived in terms of the known forward distribution \( q(x_{t+1} | x_t) \):

\[
\mathbb{E}_{q(x_{t+1} | x_t)}[q(x_{t+1} | x_t)] \approx \mathbb{E}_{q(x_{t+1} | x_t)}[q(x_{t+1} | x_t)] = \mathbb{E}_{q(x_{t+1} | x_t)}[q(x_{t+1} | x_t)]
\] (13)

Meanwhile, Eq. 12 and the second term in Eq. 11 require sampling from the network parameterized distribution as \( \hat{x}_t \sim p_{\theta_t}(x_t | x_{t+1}) \). Although it is possible to use \( G_{\theta_t} \) to predict \( x_t \) estimates from an insufficiently trained generator at intermediate stages can yield suboptimal results. Thus, here we operationalize the sampling distribution based on the generator as:

\[
p_{\theta_t}(x_t | x_{t+1}) := q(x_t | x_{t+1}, \hat{x}_0)
\] (14)

where the generator is used to estimate the denoised image sample at \( t = 0 \) as \( \hat{x}_0 = G_{\theta_t}(x_{t+1}, t + k) \). Assuming that \( \hat{x}_0 \) is a reasonable estimate of \( x_t \), \( q(x_t | x_{t+1}, \hat{x}_0) \) can be shown to have a closed form expression \( \mathcal{N}(\mu, \bar{\gamma}) \) with:

\[
\mu = \frac{\sqrt{\alpha_t} y_t}{1 - \alpha_t} \hat{x}_0 + \frac{\sqrt{\alpha_t} (1 - \alpha_t)}{1 - \alpha_t} x_{t+1}
\] (15)

\[
\bar{\gamma} = \frac{1 - \alpha_t}{1 - \alpha_t} y_{t+1}
\] (16)

where \( \alpha_t := 1 - \gamma_t \) and \( \alpha_t := \prod_{r=0}^{t-1} \alpha_r \) (Ho et al., 2020).

Since our proposed formulation does not involve a normality assumption on \( q(x_t | x_{t+1}) \), it can improve accuracy of reverse diffusion mappings at large step sizes. Finally, the discriminator and generator losses can be expressed as:

\[
L_D = \sum_{t \geq 0} \left[ \mathbb{E}_{q(x_{t+1} | x_t)}[q(x_{t+1} | x_t)] - \log(D_{\theta_t}(x_t, x_{t+1}, t + k)) \right]
\]

\[
+ \mathbb{E}_{q(x_{t+1} | x_t)}[q(x_{t+1} | x_t)] - \log(1 - D_{\theta_t}(x_t, x_{t+1}, t + k))
\]

\[
+ \mathbb{E}_{q(x_{t+1} | x_t)}[q(x_{t+1} | x_t)] \left[ \frac{1}{2} \left\| \nabla_x D_{\theta_t}(x_t, x_{t+1}, t + k) \right\|_2^2 \right]
\] (17)

\[
L_G = \sum_{t \geq 0} \left[ \mathbb{E}_{q(x_{t+1} | x_t)}[q(x_{t+1} | x_t)] - \log(D_{\theta_t}(x_t, x_{t+1}, t + k)) \right]
\] (18)

### 3.3.2. Reconstruction with the Prior

The diffusion prior is trained to capture the distribution of high-quality MR images so as to generate random image samples. However, these synthetic images do not correspond to a test subject as they are not informed about the imaging operator and the resultant acquired data (Namhoffer et al., 2019). Thus, reconstruction requires a solution at the intersection of the image sets spanned by the diffusion prior versus the operator. Here we propose a two-phase reconstruction with rapid diffusion and prior adaptation stages (Fig. 2). In rapid diffusion, an initial reconstruction is computed as a fast, compromise solution between the trained diffusion prior and the imaging operator. Note that the image sets for the diffusion prior and the operator can weakly intersect due to distributional shifts between training and test subjects. To improve performance, prior adaptation refines the reconstruction by updating the diffusion prior to better conform it to the distribution of individual test subjects.

**Rapid diffusion:** The rapid diffusion phase calculates an initial reconstruction \( \tilde{x}_{init} \) that is a compromise solution between the image sets spanned by the imaging operator and the trained diffusion prior. This compromise solution can be obtained by alternating between data-consistency projections that sample images consistent with the operator, and reverse diffusion projections that sample images consistent with the trained diffusion prior (Jalal et al., 2021). Starting with \( \tilde{x}_T \) at time step \( T \) randomly drawn from a Gaussian noise distribution, the two projections can be performed progressively across time steps.

**Prior adaptation:** Taking the initial reconstruction \( \tilde{x}_{init} \) as input, the adaptation phase further refines the diffusion prior to improve the reconstruction. To do this, the generator parameters \( \theta_G \) are fine-tuned to minimize a data-consistency loss between synthesized and acquired k-space data:

\[
\theta_G = \min_{\theta_G} \| AG_{\theta_G}(\tilde{x}_{init}, 0) - y \|_2
\] (21)

Starting with the trained generator parameters at iteration \( j = 1 \), Eq. 21 is solved by iteratively updating \( \theta_G \). At the \( j^{th} \) iteration, synthetic multi-coil k-space data are obtained by projecting the synthetic coil-combined image produced by the generator (i.e., \( \hat{x}_j^{\theta} = G_{\theta_G}(\tilde{x}_{init}, 0) \)) through the imaging operator, where \( \theta_j^G \) represents the parameters of the generator at \( j^{th} \) iteration. The imaging operator is derived from estimated coil sensitivities \( B \) (Uecker et al., 2014) and the subject’s prescribed sampling mask \( \Omega \). After \( J \) iterations, \( \theta_G^f \) is taken as \( \theta_G^f \) and the final reconstruction \( \tilde{x}_{fin} \) is computed as the synthetic image produced by the generator:

\[
\tilde{x}_{fin} = G_{\theta_G^f}(\tilde{x}_{init}, 0)
\] (22)

### 4. Methods

#### 4.1. Datasets

Demonstrations were performed on multi-contrast brain MRI from IXI (http://brain-development.org/ixi-dataset/) and fastMRI (Knoll et al., 2020) datasets. In IXI, coil-combined magnitude images for T1w, T2w and PD-weighted acquisitions were analyzed as single-coil data. Sequence parameters were repetition time (TR)=9.813ms, echo time (TE)=4.063ms, flip angle=8° for T1w scans, TR=1817ms, TE=100ms, flip angle=90° for T2w scans, TR=1817ms, TE=8ms, flip angle=90° for PD scans, and a voxel size of 0.94x0.94x1.2mm³ for all. The training, validation and test sets contained (21, 15, 30) subjects resulting in (2268, 1620, 3240) cross-sections across the
three contrasts. In fastMRI, multi-coil complex k-space data for $T_1$, $T_2$- and FLAIR-weighted acquisitions were analyzed. Since MRI scans were conducted at separate sites with heterogeneous protocols, only data with at least 10 cross-sections and 20 coil elements were selected. To improve the computational efficiency of reconstruction models, geometric coil compression was used to map multi-coil data onto 5 or 10 virtual coils that preserved over 95% and 98% of the energy in the original data, respectively (Zhang et al., 2013). The training, validation, test sets included (240, 60, 120) subjects resulting in (2400, 600, 1200) cross-sections across contrasts. Data were retrospectively undersampled in the transverse plane (i.e., anterior-posterior and left-right dimensions) using variable-density random undersampling (Lustig et al., 2007). A normal sampling density was assumed with covariance adjusted to achieve acceleration rates of $R=4x$, 8x or 12x. Coil sensitivities were estimated from a fully-sampled central calibration region via ESPiRiT (Uecker et al., 2014). Volumetric k-space data undersampled in the transverse plane were inverse Fourier transformed across the fully-sampled superior/inferior dimension. Then, 2D
cross-sections across the fully-sampled dimension were individually reconstructed (Haldar and Zhuo, 2016; Aggarwal et al., 2019).

4.2. Network Architecture

AdaDiff implemented reverse diffusion steps via an adversarial mapper comprising a generator and a discriminator (Fig. 3). The generator followed a residual encoder-decoder structure to help project the image sample between consecutive time steps, given time index $t$ (i.e., the current time step value in the diffusion process) and a set of random normal variables $z$ (Xiao et al., 2022). A total of 6 encoder stages were used, each stage containing 2 flat residual blocks followed by a downsampling residual block (by a factor of 2). Attention layers were used in the last two stages, and downsampling was omitted in the final stage. A total of 6 decoder stages were used, each stage containing 3 flat residual blocks followed by an upsampling residual block (by a factor of 2). An attention layer was used in the second stage, and upsampling was omitted in the final stage. Feature maps in each generator block were modulated via adaptive group normalization given a latent embedding vector (Zheng et al., 2020). This vector was computed via a multi-layer perceptron (MLP) containing 8 fully-connected (FC) layers to embed random normal variables $z$ (Karras et al., 2019). The discriminator followed a residual encoder structure (He et al., 2016) to distinguish actual versus synthetic image samples. A total of 6 encoder stages were used, each containing a downsampling residual block (by a factor of 2). A final MLP with 1 FC layer was used for discrimination. All generator and discriminator blocks received a time embedding vector computed by projecting a sinusoidal encoding of the time index through a 2-layer MLP (Song et al., 2020), and this vector was added as a channel-specific bias term onto feature maps. Upsampling was performed by inserting intermediate zero-valued pixels and convolving with a learnable finite-impulse-response (FIR) filter, and downsampling was performed by convolving with a learnable FIR and discarding intermediate pixels as described in (Karras et al., 2020). SiLU activation functions were used in generator blocks and MLP layers, and leaky ReLU functions with negative slope 0.2 were used in discriminator blocks. Both the generator and discriminator employed 2 channels to represent the real and imaginary parts of images.

4.3. Competing Methods

AdaDiff was demonstrated for MRI reconstruction against a traditional method (LORAKS), conditional models (rGAN, MoDL), and unconditional models (GAnprior, DDPM, DiffReccon). Conditional models were trained to map inverse Fourier transform of undersampled data onto ground-truth images derived from fully-sampled data. Unconditional models were trained to generate coil-combined MR image samples derived from fully-sampled data. An exponentially decreasing noise scheduler with parameters $\beta_{\text{min}}, \beta_{\text{max}} = [0.1, 20]$ was adopted for all diffusion models (Song et al., 2020). Hyperparameter selection was performed via one-fold cross-validation. Deep models were trained via the Adam optimizer using the decay rates $\beta_1=0.5$ and $\beta_2=0.9$. Prior adaptation during inference was performed via the Adam optimizer at $\beta_1=0.5$ and $\beta_2=0.9$. All deep models were executed on Nvidia RTX 3090 GPUs via PyTorch.

**AdaDiff**: Hyperparameters for AdaDiff were $6 \times 10^{-3}$ learning rate, 500 epochs, $k=125$ step size, $T/k=8$ diffusion steps for training; 8 iterations combining a reverse diffusion step and a data-consistency projection for rapid diffusion; and $10^{-3}$ learning rate, 1000 iterations for prior adaptation.

**LORAKS**: An autocalibrated low-rank reconstruction was implemented (Haldar and Zhuo, 2016). The k-space neighborhood radius and the rank of the system matrix were selected as: (2.6) for IXI, and (2.30) for fastMRI (Elmas et al., 2022).

**rGAN**: A conditional GAN model was implemented with architecture and loss functions in (Dar et al., 2020a). Hyperparameters were selected as $2 \times 10^{-4}$ learning rate, 100 epochs, adversarial and pixel-wise loss weights of (1,100) for training (Dar et al., 2020a).

**MoDL**: A conditional unrolled model that interleaves data-consistency blocks with convolutional layers was implemented.
Our proposed method performed prior adaptation by minimizing a data-consistency loss as in AdaDiff (Aggarwal et al., 2019). The architecture and loss function were selected as Unconditional Di ffusion steps, 65 epochs for training; 1000 iterations combining a reverse diffusion step and a data-consistency projection for inference (Ho et al., 2022). Hyperparameters were selected as $10^{-4}$ learning rate, $k=1$ step size, $T/k=4000$ diffusion steps, 300 epochs for training; 400 coarse and 20 fine iterations of reverse diffusion and data-consistency projection for inference (Peng et al., 2022).

4.4. Analyses

Retrospectively undersampled acquisitions in IXI and fastMRI were reconstructed. Here, a single unified model was trained on aggregate data from multiple distinct contrasts to improve practicality given the pervasiveness of multi-contrast protocols. In each dataset, model training was accordingly performed on data pooled across multiple contrasts: (T1, T2, PD) in IXI and (T1, T2, FLAIR) in fastMRI. Training samples were randomly drawn from the pooled data, and the model was not informed regarding the contrast of the samples. Conditional models receive undersampled data as input so they are informed regarding the acceleration rate during training. We observed that training a unified model with undersampled data from mixed R

**Table 3:** Cross-domain performance for T1-, T2-, and FLAIR- (FL.) weighted contrasts in fastMRI. Results listed for training at R=4x, testing at R=8x and R=12x.

| Dataset | LORAKS | rGAN | MoDL | GANprior | DiffRecon | AdaDiff | R=8x | R=12x |
|---------|--------|------|------|----------|-----------|--------|------|-------|
| T1 | PSNR 28.0±1.7 | 34.7±1.3 | 28.2±1.2 | 35.3±1.2 | 36.0±0.8 | 36.3±1.5 |
| SSIM 77.4±4.5 | 93.2±2.4 | 97.5±0.4 | 97.6±0.6 |
| T2 | PSNR 28.8±1.7 | 34.2±1.4 | 34.6±1.0 | 34.9±0.4 | 36.0±1.5 |
| SSIM 72.9±3.6 | 91.1±4.0 | 95.8±0.6 | 97.3±0.6 |
| PD | PSNR 28.0±2.2 | 34.9±1.7 | 35.2±1.2 | 35.4±0.5 | 36.6±1.8 |
| SSIM 66.6±4.7 | 92.7±2.9 | 96.5±0.6 | 97.6±0.6 |

**Table 4:** Cross-domain performance for T1-, T2-, and FLAIR- (FL.) weighted contrasts in fastMRI. Results listed for training at R=4x, testing at R=8x and R=12x.

| Dataset | LORAKS | rGAN | MoDL | GANprior | DiffRecon | AdaDiff | R=8x | R=12x |
|---------|--------|------|------|----------|-----------|--------|------|-------|
| T1 | PSNR 34.1±2.0 | 36.0±0.9 | 36.4±1.0 | 26.7±1.2 | 34.7±1.0 | 37.2±1.5 |
| SSIM 85.3±4.5 | 93.3±1.6 | 57.1±4.5 | 91.1±2.1 | 89.4±1.8 | 93.5±2.2 |
| T2 | PSNR 33.9±0.7 | 32.9±0.8 | 25.9±0.8 | 34.8±0.6 | 35.2±0.5 | 35.3±0.8 |
| SSIM 92.6±1.0 | 92.2±0.7 | 93.4±0.8 | 94.2±0.6 | 93.9±0.7 | 94.4±0.6 |
| FL | PSNR 28.9±2.2 | 32.8±0.1 | 34.9±0.5 | 35.4±0.9 | 36.6±1.3 |
| SSIM 77.5±10.2 | 87.7±4.9 | 88.8±4.8 | 87.3±4.8 | 88.8±6.2 |

MoDL and DDPM show high noise amplification, rGAN shows residual aliasing, and MoDL shows visible spatial blurring despite high performance in quantitative metrics. Among diffusion models, DDPM has relatively higher noise and DiffRecon shows local ringing artifacts near tissue boundaries. AdaDiff produces high-quality reconstructions with lower artifacts/noise and clearer tissue depiction than competing methods.
values did not have a notable influence on performance. Thus, separate models were trained at each individual R value to prevent biases in performance assessments under domain shifts in R between training and test sets. Reconstruction quality was measured via peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) metrics between the recovered and ground-truth images derived from fully-sampled acquisitions. In ablation studies, image fidelity was additionally characterized via Frechet inception distance (FID; Heusel et al. (2017)) and learned perceptual image patch similarity (LPIPS; Zhang et al. (2018)) metrics. Images were normalized to unity mean prior to measurements. Significance of differences in PSNR, SSIM and LPIIPS between models were evaluated with non-parametric signed-rank tests. Note that FID quantifies the overall similarity of distributions across the examined set of samples as a single metric value, so significance testing was not conducted for FID.

5. Results

5.1. Within-Domain Reconstruction

AdaDiff was first demonstrated for within-domain reconstructions where the imaging operator and the MR image distribution were matched between the training and test sets (e.g., trained and tested for R=4x in fastMRI). Comparisons were performed against a traditional method (LORAKS), conditional models (rGAN, MoDL), an unconditional GAN that performs prior adaptation (GANprior), and unconditional diffusion models that use static priors (DDPM, DiffRecon). PSNR and SSIM for competing methods are listed in Table 1 for IXI, and in Table 2 for fastMRI. In IXI, AdaDiff achieves the highest performance among competing methods across contrasts and acceleration rates (p<0.05), except for MoDL that performs similarly on T1 in general and on T2 at R=8x in PSNR. In fastMRI, AdaDiff again outperforms competing methods across contrasts and acceleration rates (p<0.05), except for MoDL that performs similarly on T1, FLAIR in general, and DiffRecon that yields higher performance on T2 at R=4x and similar PSNR on T2 at R=8x. On average, AdaDiff outperforms the traditional method by 6.8dB PSNR and 15.8% SSIM, conditional models by 2.0dB PSNR and 2.5% SSIM, the adaptive GAN by 6.6dB PSNR and 15.0% SSIM, and static diffusion models by 1.3dB PSNR and 1.4% SSIM. These results indicate that the adaptive diffusion prior in AdaDiff helps improve reconstruction quality over both an adaptive GAN prior and static diffusion priors. Representative reconstructions are displayed in Fig. 4. LORAKS and GANprior show high noise amplification. While rGAN and MoDL have relatively low noise levels, rGAN shows residual reconstruction artifacts and MoDL suffers from spatial blurring, which can be attributed to its pixel-wise loss function. Among diffusion models, DDPM has relatively high noise whereas DiffRecon is effective in noise suppression via repeated averaging of multiple diffusion samples. DiffRecon tends to produce sharper images than AdaDiff likely due to its fine iteration steps, but this refinement can also introduce ringing artifacts near tissue boundaries by emphasizing high spatial frequencies. In contrast, AdaDiff adapts its diffusion prior to better conform to the distribution of test data, enabling it to produce high-quality reconstructions that clearly depict tissues with lower artifacts and noise than competing methods.

5.2. Domain Shifts in the Imaging Operator

We then demonstrated performance in cross-domain reconstructions where the MR image distribution was matched, albeit the imaging operator was mismatched between the training and test sets. To this end, several studies were conducted to examine the influence of varying operator attributes on reconstruction performance. First, we assessed the influence of acceleration rate by training conditional models at R=4x while testing all models at R=8x and R=12x. Note that unconditional models and LORAKS were not informed about undersampling during training. PSNR and SSIM for competing methods are listed in Table 3 for IXI, and in Table 4 for fastMRI. In IXI,
AdaDiff achieves the highest performance across contrasts and acceleration rates ($p<0.05$), except for DiffRecon that yields similar SSIM on $T_2$ at $R=8x$. In fastMRI, AdaDiff again outperforms competing methods across contrasts and acceleration rates ($p<0.05$), except for MoDL that yields similar SSIM on FLAIR at $R=8x$ and higher SSIM on $T_1$ at $R=12x$, and DiffRecon that yields similar PSNR on $T_2$ at $R=8x$. On average, AdaDiff outperforms the traditional method by 4.9dB PSNR and 16.0% SSIM, conditional models by 2.3dB PSNR and 4.4% SSIM, the adaptive GAN by 6.8dB PSNR and 20.6% SSIM, and static diffusion models by 0.9dB PSNR and 1.5% SSIM. Note that, at $R=8x$, the performance benefit of AdaDiff over MoDL is 1.5dB PSNR, 2.3% SSIM under domain shift in acceleration rate, versus 0.3dB PSNR, 0.4% SSIM in within-domain reconstruction. This difference suggests that AdaDiff’s unconditional prior is more reliable against variations in acceleration rate compared to MoDL’s conditional prior. Representative reconstructions are displayed in Fig. 5. LORAKS and GAN$_{prior}$ suffer from noise amplification, rGAN and MoDL show residual aliasing artifacts and blurring. Similar to the within-domain case, DDPM has relatively high noise levels and DiffRecon shows local ringing artifacts in comparison to AdaDiff that maintains the closest appearance to the reference images.

Next, we assessed the influence of domain shifts in sampling trajectory and number of coils on reconstruction performance. Under fixed acceleration rate and number of coils, sampling trajectory was varied by training conditional models based on 2D undersampling patterns while testing all models on 1D undersampling patterns. Under fixed acceleration rate and sampling trajectory, number of coils was varied by training conditional models based on 5 virtual coils and testing all models on 10 virtual coils. PSNR and SSIM for competing methods are listed in Table 5 for both assessments. When the sampling trajectory is varied, AdaDiff achieves the highest performance among competing methods across tissue contrasts ($p<0.05$), except for MoDL that yields higher PSNR on FLAIR, and DDPM that yields similar PSNR on FLAIR. When the number of coils is varied, AdaDiff yields higher performance across tissue contrasts ($p<0.05$), except for MoDL that yields similar SSIM on FLAIR, and DiffRecon that yields higher performance on $T_2$. On average, AdaDiff outperforms the traditional method by 5.0dB PSNR and 10.0% SSIM, conditional models by 1.6dB PSNR and 1.2% SSIM, the adaptive GAN by 7.2dB PSNR and 20.2% SSIM, and static diffusion models by 1.0dB PSNR and 2.0% SSIM. Note that, at $R=4x$ in fastMRI, the performance benefit of AdaDiff over MoDL is 1.6dB PSNR, 1.2% SSIM under domain shift in sampling trajectory, 0.6dB PSNR, 0.2% SSIM under domain shift in number of coils, versus 0.5dB PSNR, 0.1% SSIM in within-domain reconstruction. These findings suggest that AdaDiff is notably more reliable than MoDL against shifts in the sampling trajectory, whereas it is similarly affected by shifts in the number of coils. Representative reconstructions are displayed in Fig. 6. LORAKS and GAN$_{prior}$ suffer from noise amplification, rGAN and MoDL show residual aliasing artifacts and blurring. Similar to the within-domain case, DDPM has relatively high noise levels and DiffRecon shows local ringing artifacts in comparison to AdaDiff that maintains the closest appearance to the reference images.

5.3. Domain Shifts in the Image Distribution

We also examined cross-domain reconstruction where the imaging operator was matched, albeit the MR image distribution was mismatched between the training and test sets. For this purpose, training was performed on fastMRI and testing was performed on IXI. PSNR and SSIM for competing methods are listed in Table 6. In general, AdaDiff achieves the highest performance among competing methods across tissue contrasts and acceleration rates ($p<0.05$), except for DiffRecon that yields modestly higher SSIM. On average, AdaDiff outperforms the traditional method by 7.7dB PSNR and 20.3% SSIM, condi-

Fig. 6: Cross-domain reconstructions at $R=4x$ under domain shifts in the sampling trajectory and number of coils. Results are shown for (a) $T_2$-weighted acquisitions with 1D undersampling, and (b) PD-weighted acquisitions with 10 virtual coils in fastMRI. Reconstructed images are given along with the reference image derived from fully-sampled acquisitions, and zoom-in windows and arrows are included to highlight differences among methods. Conditional models were trained for 2D undersampling and 5 virtual coils. LORAKS and GAN$_{prior}$ show high noise amplification, rGAN and MoDL show residual reconstruction artifacts and blurring. DDPM has relatively high noise and DiffRecon has local ringing artifacts. AdaDiff reconstructs images with low artifacts/noise.
tional models by 8.0dB PSNR and 21.8% SSIM, the adaptive GAN by 12.7dB PSNR and 40.7% SSIM, and static diffusion models by 3.1dB PSNR and 5.7% SSIM. Note that, compared to domain shifts in the imaging operator, a domain shift in the image distribution induces more notable performance losses for competing methods including static diffusion models. Although the adaptive GAN model uses prior adaptation, its relatively poor performance can be attributed to the low representational diversity of adversarial priors. In contrast, the adaptive diffusion prior in AdaDiff maintains high reconstruction performance. Representative reconstructions are displayed in Fig. 7. High noise amplification in LORAKS and GANprior, substantial residual artifacts in rGAN, and spatial blurring in MoDL are observed. While DDPM shows elevated noise and DiffRecon yields ringing artifacts, AdaDiff achieves high-fidelity reconstructions with clear tissue depiction.

5.4. Ablation Studies

We conducted a series of ablation studies to demonstrate the main elements in AdaDiff. First, we examined the effects of number of diffusion steps, and number of training epochs on the fidelity of images synthesized by the diffusion prior during reconstruction. Table 7 lists results for varying number of diffusion steps $T/k$, and Table 8 lists results for varying number of epochs $N_e$ across the validation set. In general, modest improvements in FID, LPIPS, PSNR and SSIM are observed with increasing $T/k$, yet the benefits in all metrics are marginal beyond $T/k = 8$. The results are more intermixed for $N_e$, with FID and LPIPS showing a degree of degradation towards high $N_e$. That said, the observed differences across $N_e$ between 500 to 2000 are relatively minute for all metrics. Therefore, these results suggest that the cross-validated values of $T/k$ and $N_e$ yield near-optimal performance.

Next, we assessed the importance of adaptive normalization of feature maps within the generator. To do this, we built a variant model that removed the latent variables $z$ to perform non-adaptive normalization (w/o $z$). Table 9 lists performance metrics for AdaDiff and the variant. AdaDiff outperforms the variant across reconstruction tasks, except for $T_1$ where ‘w/o $z$’ yields similar SSIM. This result indicates the importance of using random latents for adaptive normalization in AdaDiff.

Lastly, we examined the importance of the prior adaptation phase, the rapid diffusion phase, and the use of an adversarial mapper in reverse diffusion steps. For this purpose, we built a variant with a static diffusion prior that omitted the prior adaptation phase (w/o adapt.), a variant with an untrained diffusion prior (w/o train.), and a variant with a non-adversarial mapper by ablating the discriminator in AdaDiff and replacing the

![Fig. 7: Cross-domain reconstructions under domain shifts in the MR image distribution. Results are shown for $T_1$-weighted acquisitions at $R=8x$. All models were trained on fastMRI and tested on IXI. The imaging operator matched between training and testing for conditional models. Reconstructed images are given along with the reference image derived from fully-sampled acquisitions, and zoom-in windows and arrows are included to highlight differences among methods. High noise amplification in LORAKS and GANprior, substantial residual artifacts in rGAN, and spatial blurring in MoDL are observed. While DDPM shows elevated noise and DiffRecon yields ringing artifacts, AdaDiff achieves high-fidelity reconstructions with clear tissue depiction.](image)
Table 7: Performance of AdaDiff for varying number of diffusion steps, T/k. The prescribed T/k for the main experiments is marked in bold font; results are listed at J=1000 iterations. FID, LPIPS, PSNR, SSIM are reported across the validation set for R=4x4 in IWI.

| T/k = 4 | T/k = 8 | T/k = 16 | T/k = 32 |
|--------|--------|---------|---------|
| FID | 38.33 | 31.99 | 32.99 | 32.78 |
| LPIPS | 4.61 | 4.27 | 4.20 | 4.31 |
| PSNR | 40.56 | 40.89 | 41.01 | 40.93 |
| SSIM | 98.79 | 98.88 | 98.92 | 98.89 |

Table 8: Performance of AdaDiff for varying number of training epochs, N_e. The prescribed N_e for the main experiments is marked in bold font; results are listed at J=1000 iterations.

| N_e=500 | N_e=1000 | N_e=1500 | N_e=2000 |
|---------|---------|---------|---------|
| FID | 31.99 | 31.63 | 29.88 | 32.26 |
| LPIPS | 4.27 | 4.20 | 4.22 | 4.34 |
| PSNR | 40.89 | 40.91 | 40.88 | 40.89 |
| SSIM | 98.88 | 98.88 | 98.87 | 98.88 |

adversarial loss with a pixel-wise \( \ell_1 \) loss (w/o adv.). The untrained variant omitted the rapid diffusion phase to start the reconstruction with a randomly initialized generator analogously to the deep image prior method (Ulyanov et al., 2018). The non-adversarial variant performed prior adaptation on a rapid diffusion prior, unlike vanilla diffusion models that interleave reverse diffusion mappings and data-consistency projections based on a slow diffusion process. Table 10 lists performance metrics for AdaDiff and variant models at J=500 and 1000 iterations. AdaDiff outperforms all variants across reconstruction tasks (p<0.05 for LPIPS, PSNR, SSIM), except for T_1 at J=1000 where the non-adversarial variant yields similar PSNR, SSIM. Among the examined components, the prior adaptation phase has the largest contribution to reconstruction performance as AdaDiff attains the most substantial improvement levels over the static variant. On average, AdaDiff improves FID by 363.54 (164.97% change), LPIPS by 55.43 (146.00%), PSNR by 10.72dB, SSIM by 20.09% over the static variant. In theory, the untrained and non-adversarial variants that undergo prior adaptation can converge onto an equivalent solution to AdaDiff given a very large number of iterations, as they share the same generator architecture. In practice, however, the adversarial diffusion prior in AdaDiff enables prior adaptation to start at a more favorable point and reach high performance levels in fewer iterations. AdaDiff improves FID by 18.95 (39.42% change), LPIPS by 1.59 (14.40%), PSNR by 0.94dB, SSIM by 0.48% over the untrained variant; and it improves FID by 11.02 (24.98%), LPIPS by 1.74 (15.65%), PSNR by 0.42dB, SSIM by 0.29% over the non-adversarial variant. These results indicate that the rapid diffusion phase and the adversarial mapper also have important contributions to reconstruction performance, albeit at relatively modest levels.

Notable improvements in AdaDiff’s performance are apparent in Table 10 when J is increased from 500 to 1000. While prescribing J>1000 further elevates PSNR and SSIM slightly, differences in perceptual quality metrics (FID, LPIPS) and visual appearance between reconstructions at J=1000 and J=1000 become indiscernible (unreported). Thus, J=1000 offers a decent trade-off between reconstruction time and image quality. We also find that the static variant has suboptimal performance metrics compared to ZF and DDPM reconstructions. Both DDPM and the static variant inject data-consistency projections in between reverse diffusion steps for image reconstruction. This results in a compromise between a solution that carries realistic features of high-quality MR images based on the diffusion prior, and a solution that is anatomically consistent with the subject’s acquired k-space data based on the imaging.

Table 9: Performance of AdaDiff and a variant that omits random latent variables (w/o z). Results listed at J=1000 iterations.

| T_1 | w/o z | AdaDiff |
|-----|-------|---------|
| FID | 34.05 | 31.99 |
| LPIPS | 4.47 | 4.27 |
| PSNR | 40.84 | 40.89 |
| SSIM | 98.74 | 98.88 |

Table 10: Performance of AdaDiff and ablated variants. A variant omitting the prior adaptation phase (w/o adv.), a variant with an untrained prior (w/o train.), and a variant with a non-adversarial mapper (w/o adv.) were considered. Zero-filled (ZF) reconstruction results are included as reference.
operator. Since performance assessments involve comparisons between reconstructed and corresponding ground-truth images, they primarily reflect the anatomical consistency of reconstructions. Note that DDPM uses 1000 small steps gradually integrated with an equivalent number of data-consistency projections resulting in enhanced consistency to acquired data, and ZF natively satisfies full consistency to acquired data. In contrast, the static variant only uses 8 large steps inherently limiting consistency to acquired data and the anatomical consistency between reconstructed and ground-truth images.

5.5. Computation Times

A practical concern regarding MRI reconstruction methods involves training and inference times. Table 11 lists the computation times for competing methods, along with the static, untrained and non-adversarial variants of AdaDiff (at J=1000). In general, conditional models have shorter training times than unconditional models, and GAN models have shorter training times than diffusion models. Among diffusion-based methods, AdaDiff and the static variant have the longest training times due to the introduction of adversarial components. Note that LORAKS and the untrained variant have no training overhead. Meanwhile, LORAKS, conditional models and the non-adapted variant have notably shorter inference times than unconditional models in general. Among unconditional methods, GAN$_{prior}$, AdaDiff and its adapted variants (w/o train., w/o adv.) have comparable inference times as they all involve inference optimization procedures.

| Method | lGAN | MoDL | GAN$_{prior}$ | DDPM | DiffRecon | w/o adapt | w/o train. | w/o adv. | AdaDiff |
|--------|------|------|--------------|------|-----------|-----------|------------|----------|---------|
| Training | –    | 4.0  | 26.5         | 53.8 | 15.5      | 77.3      | 131.3      | –        | 87.5    |
| Inference | 3.0  | 0.03 | 0.05         | 129.6 | 57.5      | 12.0      | 0.4        | 131.0    | 131.4   |

AdaDiff learns an image prior that generates an initial reconstruction via diffusion sampling, and then adapts the prior to the test subject with an inference optimization. Despite AdaDiff’s rapid diffusion process, prior adaptation naturally elevates run times over conditional models that recover images in a single forward pass. While AdaDiff is closer to regular diffusion models with iterative image sampling, it still yields relatively longer inference due to backward passes involved in prior adaptation. As expected, AdaDiff has similar inference time to the GAN-based prior adaptation method that employs a similar inference optimization. Another practical concern regarding computational complexity is memory load during inference. Among competing methods, conditional models and regular diffusion models that only leverage forward passes have relatively limited memory load. In contrast, prior adaptation methods including AdaDiff involve both forward and backward passes through the network, so they introduce additional memory load to store model gradients. To improve practicality of methods that use inference optimization, efficiency can be increased by sharing optimized model parameters across spatially proximate cross-sections within an MRI volume (Korkmaz et al., 2022), or by parallel computations on multiple GPUs. Here, we adopted the Adam algorithm observed to perform well in prior adaptation for consistency between training-inference procedures and among competing methods. Note that Namhofer et al. (2019) originally implemented GAN$_{prior}$ based on the iPALM algorithm (Pock and Sabach, 2016). It remains important future work to systematically investigate the relative benefits of different algorithms in prior adaptation, including computational efficiency and reconstruction performance.

6. Discussion

The proposed AdaDiff method was demonstrated against conditional and unconditional baselines for MRI reconstruction. Within-domain tasks were considered with matching acceleration rate and image distribution across the training-test sets. Several cross-domain tasks were also examined with mismatched acceleration rates, mismatched sampling trajectories, mismatched number of coils, or mismatched image distribution. We find that AdaDiff offers improved reliability against domain shifts in the imaging operator against conditional models, and against domain shifts in the MR image distribution against all competing models. Importantly, the adaptation procedure in AdaDiff notably improves reconstruction performance over competing diffusion models based on static priors for both within- and cross-domain scenarios. Yet, it remains important future work to examine reliability against broader changes in anatomy such as generalization across different body parts, and other changes in the imaging operator such as generalization across Cartesian versus non-Cartesian trajectories and across different coil arrays.

Acceleration techniques have recently been considered to speed up the characteristically slow sampling process in regular diffusion models. An elegant approach is to initiate sampling with the image obtained from a separate reconstruction method, including zero-filled reconstructions of undersampled k-space data (Chung et al., 2022). In unreported analyses, we observed that a variant model that initiated prior adaptation with zero-filled reconstructions does not offer notable benefits in performance or inference time against AdaDiff. While employing reconstructions from learning-based methods might help accelerate prior adaptation, it also necessitates independent training of secondary reconstruction models. Another powerful approach is to train diffusion models with small step size and to rescale to large step sizes during inference (Peng et al., 2022). This method shortens the sampling process, but reverse diffusion steps can potentially have suboptimal accuracy. Instead, AdaDiff implements reverse diffusion over large step sizes via an adversarial mapper for improved accuracy. That said, combining the adversarial mapper in AdaDiff with the abovementioned acceleration approaches might offer further benefits.

Several recent studies have proposed adaptation of image
priors for MRI reconstruction. A group of methods reconstruct with untrained priors that map low-dimensional latent variables onto images (Jin et al., 2019). Convolutional architectures with randomly initialized weights are adopted for this purpose, wherein convolution operators serve to regularize synthesized images (Arora et al., 2020; Ke et al., 2020; Zou et al., 2021; Darestani and Heckel, 2021). For inference, the untrained priors are combined with the imaging operator and adapted to enforce consistency between synthesized and acquired data. While performant MRI reconstruction has been reported with this approach, markedly longer inference optimization is typically required to intersect the image set reflecting the untrained prior with the image set reflecting the imaging operator (Darestani and Heckel, 2021). An alternative group of methods instead employ priors pre-trained on MR images to provide an improved initialization point for adaptation. Previous studies in this group have predominantly proposed priors based on GAN models that implicitly characterize the MR image distribution (Narnhofer et al., 2019; Korkmaz et al., 2022), while some rely on patch-based auto-encoder models (Tezcan et al., 2019) or convolutional models (Aggarwal and Jacob, 2021; Han et al., 2018). Our work differs from recent methods based on untrained priors in that AdaDiff leverages a prior pre-trained on high-quality coil-combined MR images to improve efficiency during inference optimization. It also differs from methods with pre-trained priors since AdaDiff leverages a novel diffusion prior to improve fidelity of image samples.

In theory, prior adaptation can be performed based on regular diffusion models instead of the adversarial diffusion model considered here. However, regular diffusion models might elicit several limitations in the context of prior adaptation. Note that generation of the initial reconstruction with regular diffusion models involves image sampling across hundreds of reverse diffusion steps interleaved with data-consistency projections (Chung et al., 2022). Thus, computing the initial reconstruction with regular diffusion models requires comparable inference time to the prior adaptation stage (e.g., see DDPM and AdaDiff in Table 11), significantly elevating the overall computational burden. Furthermore, the initial reconstruction based on hundreds of data-consistency projections naturally yields an enhanced match to acquired k-space data. In turn, regular diffusion models already yield low data-consistency loss, significantly limiting the added benefit that can be achieved via prior adaptation. To improve prior adaptation with regular diffusion models, data-consistency projections might be partly omitted during the initial reconstruction stage to ensure a reasonably high level of data-consistency loss. It remains future work to examine the optimal training and inference procedures for prior adaptation with regular diffusion models.

Reconstruction performance for AdaDiff might be improved through several lines of technical development. First, all reported models were trained by pooling acquisitions across multiple distinct contrasts in each dataset. The trained models were then used for independently reconstructing individual MRI contrasts. Performance might be improved by training separate models on each contrast, at the expense of computational burden (Dar et al., 2020b). When a multi-contrast accelerated MRI protocol is available in each subject, joint reconstruction models can also be used to exploit structural correlations among contrasts to improve performance (Dar et al., 2020a; Polak et al., 2020; Gailliochet et al., 2020; Xuan et al., 2022). For AdaDiff, this would involve training of a multi-contrast diffusion prior. Alternatively, contrast type can be provided as side information to maintain specificity to individual contrasts in a unified model (Liu et al., 2022; Dalmaz et al., 2022a). Second, cycle-consistent learning strategies can be adopted to alleviate the dependence of AdaDiff on datasets comprising fully-sampled acquisitions (Quan et al., 2018; Ok et al., 2020; Ozbeyl et al., 2022). Here, AdaDiff was implemented to perform prior adaptation by optimizing the generator parameters at the final time step for computational efficiency. In theory, performing prior adaptation on the entire set of diffusion steps could improve reconstruction performance. In practice, however, inference optimization over multiple diffusion steps requires computation and storage of gradients across all steps. In turn, this would substantially elevate the memory load and inference time for AdaDiff, limiting practical utility.

The primary focus of the current study was on a learning strategy to improve generalization in diffusion-based MRI reconstruction. Thus, we adopted convolutional generator and discriminator architectures reported to offer high performance in previous studies on generative modeling (Song et al., 2020; Xiao et al., 2022; Karras et al., 2020). Further work is warranted to assess the contributions of various design elements in the employed architectures to AdaDiff’s performance. Future studies should also be conducted to explore the utility of alternative architectures such as transformer backbones (Dalmaz et al., 2022a; Gungor et al., 2022; Guo et al., 2022), and the influence of different normalization layers on model performance. To represent complex MRI data, we used separate network channels for the real and imaginary components following common practice in learning-based MRI reconstruction (Schlemper et al., 2017; Eo et al., 2018; Sriram et al., 2020; Aggarwal et al., 2019; Yang et al., 2020). Recent studies suggest that complex-valued network operations might offer benefits particularly in phase-oriented reconstruction tasks (Dedmari et al., 2018; Kustner et al., 2020; Cole et al., 2021; Wang et al., 2020). It remains important future work to explore the potential benefits of adopting complex-valued operations in AdaDiff.

7. Conclusion

In this study, we introduced the first prior adaptation method based on diffusion modeling for MRI reconstruction. AdaDiff leverages an adversarial mapper for reverse diffusion that enables efficient image generation in few steps. During inference, an initial reconstruction is obtained via rapid projection through the trained diffusion prior. The final reconstruction is then computed by further adapting the prior to the test subject. Compared against state-of-the-art baselines, AdaDiff performs competitively in within-domain tasks, and achieves superior reconstructions in cross-domain tasks. Therefore, AdaDiff holds great promise for high-performance MRI reconstruction.
References

Adler, J., Oktem, O., 2018. Learned primal-dual reconstruction. IEEE Transactions on Medical Imaging 37, 1322–1332.

Aggarwal, H.K., Jacob, M., 2021. Model adaptation for image reconstruction using generalized Stein’s unbiased risk estimator. arXiv:2102.00047.

Aggarwal, H.K., Mani, M.P., Jacob, M., 2019. MoDL: Model-Based deep learning architecture for inverse problems. IEEE Transactions on Medical Imaging 38, 394–405.

Aggarwal, H.K., Pramanik, A., Jacob, M., 2021. Ensure: Ensemble stein’s unbiased risk estimator for unsupervised learning, in: IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 1160–1164.

Ahmad, R., Bouman, C.A., Buzzard, G.T., Chan, S., Liu, S., Reehor, E.T., Schniter, P., 2020. Plug-and-play methods for magnetic resonance imaging: Using denoisers for image recovery. IEEE Signal Processing Magazine 37, 105–116.

Arora, S., Roeloffs, V., Lustig, M., 2020. Untrained modified deep decoder for joint denoising and parallel imaging reconstruction, in: Proceedings of ISMRM, p. 3585.

Biswas, S., Aggarwal, H.K., Jacob, M., 2019. Dynamic MRI using model-based deep learning and STORM priors: MoDL-STORM. Magnetic resonance in medicine 82, 485–494.

Cheng, J., Wang, H., Ying, L., Liang, D., 2019. Model learning: Primal dual networks for fast MR imaging, in: Proceedings of MICCAI, pp. 21–29.

Chung, H., Cha, E., Sunwoo, L., Ye, J.C., 2021. Two-stage deep learning for accelerated 3D time-of-flight MRA without matched training data. Medical Image Analysis 71, 102047.

Chung, H., Sim, B., Ye, J.C., 2022. Come-closer-di. Chung, H., Cha, E., Sunwoo, L., Ye, J.C., 2021. Two-stage deep learning for image reconstruction in MRI, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 12143–12142.

Chung, H., Ye, J.C., 2022. Score-based diffusion models for accelerated MRI. Medical Image Analysis 80, 102479.

Cole, E., Cheng, J., Pauly, J., Vasanawala, S., 2021. Analysis of deep complex-valued convolutional neural networks for mri reconstruction and phase-focused applications. Magnetic Resonance in Medicine 86, 1093–1109.

Cole, E.K., Pauly, J.M., Vasanawala, S.S., Ong, F., 2020. Unsupervised MRI reconstruction with generative adversarial networks. arXiv:2008.13065.

Dalmaz, O., Mirza, U., Elmas, G., Ozbey, M., Dar, S.U., Ceyani, E., Aves-timehr, S., Çukur, T., 2022a. One model to unite them all: Personalized federated learning of multi-contrast mri synthesis. arXiv preprint arXiv:2207.06509.

Dalmaz, O., Yurt, M., Çukur, T., 2022b. ResViT: Residual vision transformers for multi-modal medical image synthesis. IEEE Transactions on Medical Imaging 41, 2598–2614.

Dar, S.U., Yurt, M., Çukur, T., 2021. A few-shot learning approach for accelerated MRI via fusion of data-driven and subject-driven priors, in: Proceedings of ISMRM, p. 1949.

Dar, S.U., Yurt, M., Shahdloo, M., Ildiz, M.E., Tinaz, B., Çukur, T., 2020a. Prior-guided image reconstruction for accelerated multi-contrast MRI via generative adversarial networks. IEEE Journal of Selected Topics in Signal Processing 14, 1072–1087.

Dar, S.U.H., Özbeş, M., Çatlı, A.B., Çukur, T., 2020b. A transfer-learning approach for accelerated MRI using deep neural networks. Magnetic Resonance in Medicine 84, 663–685.

Dar, S.U.H., Yurt, M., Karacan, L., Erdem, A., Erdem, E., Çukur, T., 2019. Image Synthesis in Multi-Contrast MRI with Conditional Generative Adversarial Networks. IEEE Transactions on Medical Imaging 38, 2375–2388.

Darestanti, M.Z., Heck, R., 2021. Accelerated mri with un-trained neural networks. IEEE Transactions on Computational Imaging 7, 724–733.

Dedmari, M.A., Conjeti, S., Estrada, S., Ehses, P., Stocker, T., Reuter, M., 2018. Complex fully convolutional neural networks for mri image reconstruction, in: Machine Learning for Medical Image Reconstruction, pp. 30–38.

Dhariwal, P., Nichol, A., 2021. Diffusion models beat gans on image synthesis, in: Advances in Neural Information Processing Systems, Curran Associates, Inc., pp. 8780–8794.

Elmas, G., Dar, S.U., Korkmaz, Y., Ceyani, E., Susam, B., Özbeş, M., Avestimehr, S., Çukur, T., 2022. Federated learning of generative image priors for mri reconstruction. IEEE Transactions on Medical Imaging doi:10.1109/TMI.2022.3220757.

Eo, T., Jun, Y., Kim, T., Jang, J., Lee, H.J., Hwang, D., 2018. KKI-net: cross-domain convolutional neural networks for reconstructing undersampled magnetic resonance images. Magnetic Resonance in Medicine 80, 2188–2201.

Feng, C.M., Yang, Z., Fu, H., Xu, Y., Yang, J., Shao, L., 2021. Donet: Dual-octave network for fast mr image reconstruction. IEEE Transactions on Neural Networks and Learning Systems, 1–11.

Gaillochet, M., Tezcan, K.C., Konukoglu, E., 2020. Joint reconstruction and bias field correction for undersampled mri imaging, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, pp. 44–52.

Gu, H., Yaman, B., Uguribil, K., Moeller, S., Akçaşakaya, M., 2021. Compressed sensing mri with /1-wavelet reconstruction revisited using modern data science tools, in: International Conference of the IEEE Engineering in Medicine & Biology Society, pp. 3596–3600.

Güngör, A., Askın, B., Soydan, D.A., Saritas, E.U., Top, C.B., Çukur, T., 2022. TransSMS: Transformers for Super-Resolution Calibration in Magnetic Particle Imaging. IEEE Transactions on Medical Imaging 41, 3562–3574.

Guo, P., Mei, Y., Zhou, J., Jiang, S., Patel, V.M., 2022. Reconformer: Accelerated mri reconstruction using recurrent transformer. arXiv preprint arXiv:2201.09376.

Guo, P., Valanarasu, J.M.J., Wang, P., Zhou, J., Jiang, S., Patel, V.M., 2021. Over-and-under complete convolutional mri for mri reconstruction, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, pp. 13–23.

Haldar, J.P., Zhou, J., 2016. P-LORAKS: Low-Rank Modeling of Local k-Space Neighborhoods with Parallel Imaging Data. Magnetic resonance in medicine 75, 1499.

Hammermik, K., Klatzer, T., Kohler, E., Recht, M.P., Sodickson, D.K., Pock, T., Knoll, F., 2017. Learning a variational network for reconstruction of accelerated MRI data. Magnetic Resonance in Medicine 79, 3055–3071.

Han, Y., Yoo, J., Kim, H.H., Shin, H.J., Sung, K., Ye, J.C., 2018. Deep learning with domain adaptation for accelerated projection-reconstruction MR. Magnetic Resonance in Medicine 80, 1189–1205.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., Hochreiter, S., 2017. Gans trained by a two time-scale update rule converge to a local nash equilibrium, in: Advances in Neural Information Processing Systems.

Ho, J., Jain, A., Abbeel, P., 2020. Denoising diffusion probabilistic models, in: Advances in Neural Information Processing Systems, pp. 6840–6851.

Hosseini, S.A.H., Yaman, B., Moeller, S., Hong, M., Akccakaya, M., 2020. Dense recurrent neural networks for accelerated MRI: History-cognizant unrolling of optimization algorithms. IEEE Journal of Selected Topics in Signal Processing 14, 1280–1291.

Huang, P., Li, C.H., Gaire, S.K., Liu, R., Zhang, X., Li, X., Ying, L., 2019. Deep MRI reconstruction without ground truth for training, in: Proceedings of ISMRM, p. 4668.

Hyun, C.M., Kim, H.P., Lee, S.M., Lee, S., Seo, J.K., 2018. Deep learning for undersampled MRI reconstruction. Physics in Medicine and Biology 63, 135007.

Jalal, A., Arvinte, M., Daras, G., Price, E., Dimakis, A.G., Tamir, J., 2021. Robust compressed sensing mri with deep generative priors, in: Advances in Neural Information Processing Systems, pp. 14938–14954.

Jin, K.H., Gupta, H., Herly, J., Stuber, M., Unser, M., 2019. Time-dependent deep image prior for dynamic MRI. arXiv:1910.01684.

Karras, T., Laine, S., Aila, T., 2019. A style-based generator architecture for generative adversarial networks, in: Advances in Neural Information Processing Systems, pp. 4401–4410.

Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., Aila, T., 2020. Analyzing and improving the image quality of StyleGAN, in: IEEE CVPR, pp. 8107–8116.

Ke, Z., Zhu, Y., Cheng, J., Ying, L., Liu, X., Zheng, H., Liang, D., 2020. Assessment of the generalization of learned unsupervised deep learning method, in: Proceedings of ISMRM, p. 3630.
Ye, J.C., Han, Y., Cha, E., 2018. Deep convolutional framelets: A general deep learning framework for inverse problems. SIAM Journal on Imaging Sciences 11, 991–1048.

Yoon, J., Gong, E., Chatnuntawech, I., Bilgic, B., Lee, J., Jung, W., Ko, J., Jung, H., Setompop, K., Zaharchuk, G., Kim, E.Y., Pauly, J., Lee, J., 2018. Quantitative susceptibility mapping using deep neural network: QSMnet. NeuroImage 179, 199–206.

Yu, S., Dong, H., Yang, G., Slabaugh, G., Dragotti, P.L., Ye, X., Liu, F., Arridge, S., Keegan, J., Firmin, D., Guo, Y., 2018. DAGAN: Deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction. IEEE Transactions on Medical Imaging 37, 1310–1321.

Zhang, R., Isola, P., Efros, A.A., Shechtman, E., Wang, O., 2018. The unreasonable effectiveness of deep features as a perceptual metric, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Zhang, T., Pauly, J.M., Vasanawala, S.S., Lustig, M., 2013. Coil compression for accelerated imaging with Cartesian sampling. Magnetic Resonance in Medicine 69, 571–82.

Zheng, H., Fu, J., Zeng, Y., Luo, J., Zha, Z.J., 2020. Learning semantic-aware normalization for generative adversarial networks, in: Advances in Neural Information Processing Systems, pp. 21853–21864.

Zhu, B., Liu, J.Z., Rosen, B.R., Rosen, M.S., 2018. Image reconstruction by domain transform manifold learning. Nature 555, 487–492.

Zou, Q., Ahmed, A.H., Nagpal, P., Kruger, S., Jacob, M., 2021. Deep generative SToRM model for dynamic imaging. arXiv:2101.12566.