Adaptive diffusion priors for accelerated MRI reconstruction

Alper Güngör [a,b,c,1], Salman UH Dar [a,b,d,1], Şaban Öztürk [a,b,e,1], Yilmaz Korkmaz [a,b], Hasan A. Bedel [a,b], Gokberk Elmas [a,b], Muzaffer Ozbey [a,b], Tolga Çukur [a,b,f,*]

[a] Department of Electrical and Electronics Engineering, Bilkent University, Ankara 06800, Turkey
[b] National Magnetic Resonance Research Center (UMRAM), Bilkent University, Ankara 06800, Turkey
[c] ASELSAN Research Center, Ankara 06200, Turkey
[d] Department of Internal Medicine III, Heidelberg University Hospital, Heidelberg 69120, Germany
[e] Department of Electrical and Electronics Engineering, Amasya University, Amasya 05100, Turkey
[f] Neuroscience Program, Bilkent University, Ankara 06800, Turkey

ARTICLE INFO

ABSTRACT

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., 2020a; Kwon et al., 2017; Yaman et al., 2021). 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., 2020b; 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., 2020b; 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) and Luo et al. (2022) proposed sampling with score-based functions and Langevin dynamics; Chung and Ye (2022) and 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 involves implementation of 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 image distribution between the training-test sets (Narnhofer et al., 2019).

Here we propose an adaptive diffusion prior, AdaDiff, for MRI reconstruction. AdaDiff differs from recent GM-based reconstruction methods in several key aspects. Unlike GAN-based methods that use adversarial learning for single-shot mapping from noise variables onto images, AdaDiff is based on a multi-step diffusion process to improve fidelity in generated image samples. Unlike methods based on static diffusion priors, AdaDiff performs subject-specific adaptation of its prior during inference to increase conformity of its prior to the distribution of the test data. Finally, unlike diffusion methods based on a long-chain of sampling steps, AdaDiff performs diffusion modeling in few large steps implemented via adversarial mapping for enhanced reliability.

a) Regular Diffusion Model

b) AdaDiff

Forward Diffusion
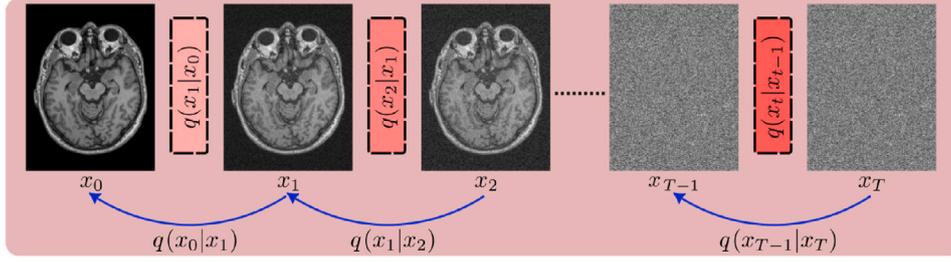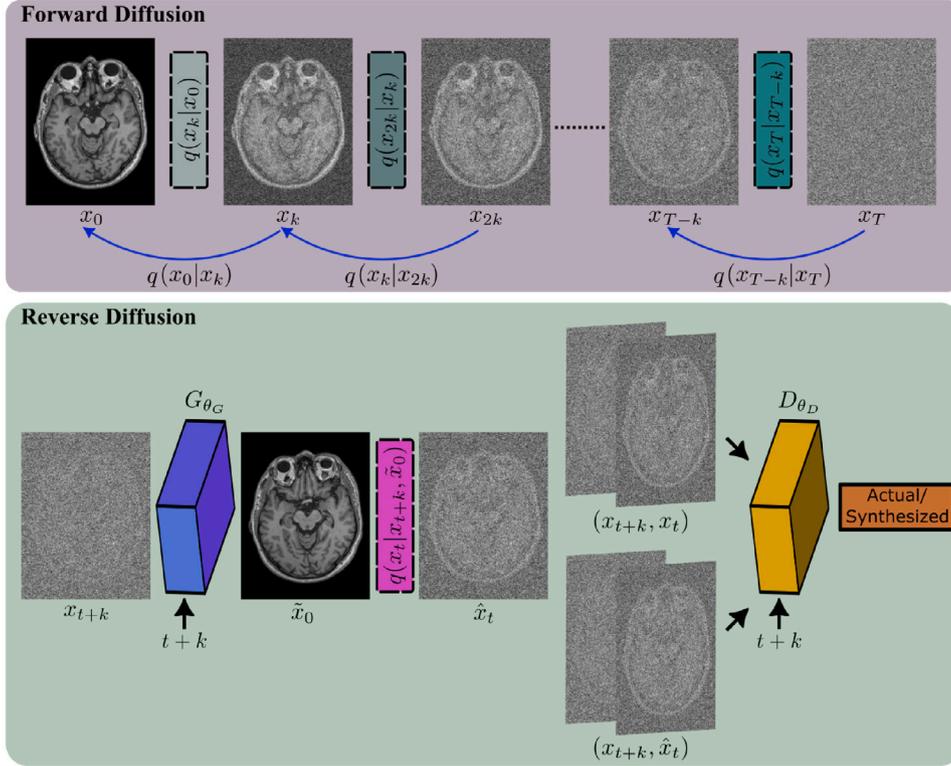
Reverse Diffusion

Fig. 1. (a) Diffusion models generate an actual image (x_0) starting from isotropic Gaussian noise (x_T) through a gradual process with forward and reverse steps. In a forward step, scaled Gaussian noise is added onto the previous sample x_{t-1} to obtain a noisier sample x_t (Eq. (3)). In a reverse step, additive noise on x_{t+1} is suppressed to obtain x_t. This reverse mapping is parameterized as projection through a neural network, $p_\theta(x_t|x_{t+1})$. Vanilla diffusion models use small step sizes to ensure approximate normality of the reverse transition probability $q(x_t|x_{t+1})$, resulting in prolonged sampling. (b) AdaDiff leverages rapid diffusion with a large step size k to transform between x_0 and x_T in few steps (Eq. (8)). Because the added noise in each step is scaled up to account for large step size, the normality assumption for the reverse transition probability $q(x_t|x_{t+k})$ breaks down. To address this issue, AdaDiff employs an adversarial mapper that implicitly characterizes the distribution of the reverse diffusion steps. The generator estimates denoised image samples (Eq. (12)), whereas the discriminator distinguishes actual samples based on the forward diffusion process from synthesized samples produced by the generator (Eq. (11)).

3. Theory

3.1. MRI reconstruction

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

$$Ax = y \tag{1}$$

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

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

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 $R(x|\mathcal{F}^{-1}(y))$. 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

a) Rapid Diffusion

b) Prior Adaptation
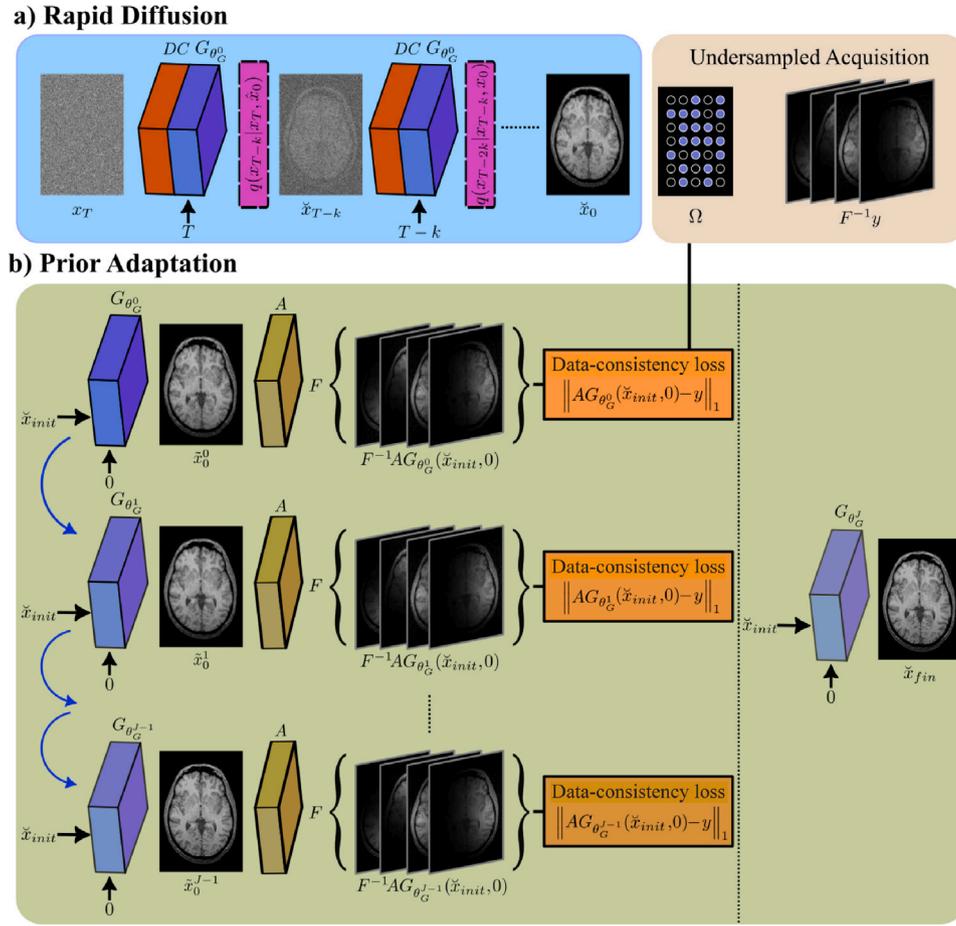
Fig. 2. AdaDiff employs a two-phase reconstruction given a learned diffusion prior with a trained generator $G_{\theta_G^0}$. **(a)** The rapid diffusion phase calculates a fast, initial solution as a compromise between consistency with the learned prior and consistency with the imaging operator. Starting with a Gaussian noise sample $\breve{x}_T$, interleaved projections are performed through data-consistency (DC) blocks (Eq. (19)) and reverse diffusion steps (Eq. (20)). The sample at time step 0 is taken as the initial reconstruction, $\breve{x}_0 = \breve{x}_{init}$. **(b)** The adaptation phase refines the diffusion prior per test subject to further improve the initial reconstruction. To do this, the generator parameters (θ_G) are iteratively optimized to minimize a data-consistency loss (Eq. (21)). At the jth iteration, the generator receives the initial reconstruction $\breve{x}_{init}$ to synthesize a coil-combined image $\tilde{x}_0^j$. Synthetic multi-coil images are obtained by projecting $\tilde{x}_0^j$ through the imaging operator A that encapsulates estimated coil sensitivities and undersampling in k-space with the subject's prescribed sampling mask Ω. The data-consistency loss is taken as the difference between synthesized and acquired data in k-space. The generator output at the end of J iterations is taken as the final reconstruction $\breve{x}_{fin}$.

isotropic Gaussian noise $z \sim \mathcal{N}(0, \mathrm{I})$ 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:

$$x_t = \sqrt{1 - \beta_t} x_{t-1} + \sqrt{\beta_t} z, \tag{3}$$

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

where β_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_θ denotes projection through a network with parameters θ. The network can be trained to minimize a lower bound on the negative log-likelihood:

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

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:

$$
\begin{aligned}
L_{\mathrm{lb}} = & -\log p_\theta(x_0|x_1) + \\
& \sum_{t=1}^{T-1} D_{KL}\left(q\left(x_t|x_{t+1}, x_0\right) \parallel p_\theta\left(x_t|x_{t+1}\right)\right)
\end{aligned} \tag{6}
$$

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_{lb} is to predict the additive noise z given x_{t+1} and x_0 as input (Ho et al., 2020):

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

where $\mathbb{E}$ is expectation of the difference norm between the scaled noise added onto x_t 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 $x_T \sim \mathcal{N}(0, \mathrm{I})$, 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 interleaved with 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 $q\left(x_t|x_{t+1}\right)$, 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 $q\left(x_t|x_{t+1}\right)$ 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):

$$x_t = \sqrt{1-\gamma_t}x_{t-k} + \sqrt{\gamma_t}z, \tag{8}$$

$$q\left(x_t|x_{t-k}\right) = \mathcal{N}\left(x_t; \sqrt{1-\gamma_t}x_{t-k}, \gamma_t I\right) \tag{9}$$

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

$$L_{\text{lb}} = \sum_{\substack{t=rk \\ r=[0,\dots,T/k-1]}} D_{KL}\left(q\left(x_t|x_{t+k}\right) \| p_\theta\left(x_t|x_{t+k}\right)\right) \tag{10}$$

There is no closed-form expression for $q\left(x_t|x_{t+k}\right)$, so we introduce an adversarial mapper to implicitly capture the conditional distribution for the reverse diffusion steps. A generator G_{θ_G} is used to parametrize a sampling distribution $p_{\theta_G}\left(x_t|x_{t+k}\right)$ that synthesizes $\hat{x}_t$. Meanwhile, a discriminator D_{θ_D} differentiates between synthetic samples ($\hat{x}_t$) drawn from $p_{\theta_G}\left(x_t|x_{t+k}\right)$ and actual samples (x_t) drawn from the true denoising distribution $q\left(x_t|x_{t+k}\right)$. Both G_{θ_G} and D_{θ_D} 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):

$$\begin{aligned} L_D = \sum_{t\geq 0} &\mathbb{E}_{q(x_{t+k})}[\mathbb{E}_{q(x_t|x_{t+k})}\left[-\log\left(D_{\theta_D}\left(x_t, x_{t+k}, t+k\right)\right)\right] \\ &+ \mathbb{E}_{p_{\theta_G}(x_t|x_{t+k})}\left[-\log\left(1 - D_{\theta_D}\left(\hat{x}_t, x_{t+k}, t+k\right)\right)\right] \\ &+ \mathbb{E}_{q(x_t|x_{t+k})}[\frac{1}{2}\left\|\nabla_{x_t} D_{\theta_D}(x_t, x_{t+k}, t)\right\|_2]] \end{aligned} \tag{11}$$

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

$$L_G = \sum_{t\geq 0} \mathbb{E}_{q(x_{t+k})p_{\theta_G}(x_t|x_{t+k})}\left[-\log\left(D_{\theta_D}\left(\hat{x}_t, x_{t+k}, t+k\right)\right)\right] \tag{12}$$

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 derived in terms of the known forward distribution $q(x_{t+k}|x_t)$:

$$\mathbb{E}_{q(x_{t+k})q(x_t|x_{t+k})} \approx \mathbb{E}_{q(x_0)q(x_t|x_0)q(x_{t+k}|x_t)} = \mathbb{E}_{q(x_0,x_t)q(x_{t+k}|x_t)} \tag{13}$$

Meanwhile, Eq. (12) and the second term in Eq. (11) require sampling from the network parameterized distribution as $\hat{x}_t \sim p_{\theta_G}(x_t|x_{t+1})$. Although it is possible to use G_{θ_G} 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:

$$p_{\theta_G}\left(x_t|x_{t+k}\right) := q\left(x_t|x_{t+k}, \tilde{x}_0\right) \tag{14}$$

where the generator is used to estimate the denoised image sample at $t=0$ as $\tilde{x}_0 = G_{\theta_G}(x_{t+k}, t+k)$. Assuming that $\tilde{x}_0$ is a reasonable estimate of x_0, $q\left(x_t|x_{t+k}, \tilde{x}_0\right)$ can be shown to have a closed form expression $\mathcal{N}_{\theta_G}(\bar{\mu}, \bar{\gamma})$ with:

$$\bar{\mu} = \frac{\sqrt{\bar{\alpha}_t}\gamma_t}{1-\bar{\alpha}_{t+k}}\tilde{x}_0 + \frac{\sqrt{\alpha_{t+k}}\left(1-\bar{\alpha}_t\right)}{1-\bar{\alpha}_{t+k}}x_{t+k} \tag{15}$$

$$\bar{\gamma} = \frac{1-\bar{\alpha}_t}{1-\bar{\alpha}_{t+k}}\gamma_{t+k} \tag{16}$$

where $\alpha_t := 1 - \gamma_t$ and $\bar{\alpha}_t := \prod_{\substack{\tau=rk \\ r=[0,\dots,t/k]}} \alpha_\tau$ (Ho et al., 2020). Since our proposed formulation does not involve a normality assumption on $q\left(x_t|x_{t+k}\right)$, it can improve accuracy of reverse diffusion mappings at large step sizes. Finally, the discriminator and generator losses can be expressed as:

$$\begin{aligned} L_D = \sum_{t\geq 0} &\left(\mathbb{E}_{q(x_0,x_t)}\mathbb{E}_{q(x_{t+k}|x_t)}\left[-\log\left(D_{\theta_D}\left(x_t, x_{t+k}, t+k\right)\right)\right]\right. \\ &+ \mathbb{E}_{q(x_{t+k})}\mathbb{E}_{\mathcal{N}_{\theta_G}(\bar{\mu},\bar{\gamma})}\left[-\log\left(1 - D_{\theta_D}\left(\hat{x}_t, x_{t+k}, t+k\right)\right)\right] \\ &+ \left.\mathbb{E}_{q(x_0,x_t)}\mathbb{E}_{q(x_{t+k}|x_t)}[\frac{1}{2}\left\|\nabla_{x_t} D_{\theta_D}(x_t, x_{t+k}, t+k)\right\|_2]\right) \end{aligned} \tag{17}$$

$$L_G = \sum_{t\geq 0} \mathbb{E}_{q(x_{t+k})}\mathbb{E}_{\mathcal{N}_{\theta_G}(\bar{\mu},\bar{\gamma})}\left[-\log\left(D_{\theta_D}\left(\hat{x}_t, x_{t+k}, t+k\right)\right)\right] \tag{18}$$

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 (Narnhofer 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 ($\breve{x}_{init}$) 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 $\breve{x}_T$ at time step T randomly drawn from a Gaussian noise distribution, the two projections can be performed progressively across time steps. Given $\breve{x}_{t+k}$, the data-consistency projection at time step $t+k$ can be implemented as in Peng et al. (2022):

$$\dot{x}_{t+k} = \breve{x}_{t+k} + A^H(y - A\breve{x}_{t+k}) \tag{19}$$

where A^H denotes the Hermitian adjoint of A. The reverse diffusion projection can then be performed by sampling $\breve{x}_T$ from $q(x_t|x_{t+k}, \tilde{x}_0)$ as described in Eq. (14), where x_{t+k} is taken as $\dot{x}_{t+k}$, and $\tilde{x}_0$ is computed via the generator as:

$$\tilde{x}_0 = G_{\theta_G}(\dot{x}_{t+k}, t+k) \tag{20}$$

Unlike conventional diffusion methods, AdaDiff leverages a rapid diffusion model with large step size. Thus, the initial reconstruction can be computed as $\breve{x}_{init} = \breve{x}_0$ in a few steps.

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

$$\theta_G^* = \min_{\theta_G} \left\| AG_{\theta_G}(\breve{x}_{init}, 0) - y \right\|_1 \tag{21}$$

Starting with the trained generator parameters at iteration $j = 1$, Eq. Eq. (21) is solved by iteratively updating θ_G. At the jth iteration, synthetic multi-coil k-space data are obtained by projecting the synthetic coil-combined image produced by the generator (i.e., $\breve{x}_0^j = G_{\theta_G^j}(\breve{x}_{init}, 0)$) through the imaging operator, where θ_G^j represents the parameters of the generator at jth iteration. The imaging operator is derived from estimated coil sensitivities B (Uecker et al., 2014) and the subject's prescribed sampling mask Ω. After J iterations, θ_G^* is taken as θ_G^J, and the final reconstruction ($\breve{x}_{fin}$) is computed as the synthetic image produced by the generator:

$$\breve{x}_{fin} = G_{\theta_G^J}(\breve{x}_{init}, 0) \tag{22}$$

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 T_1-, T_2- and PD-weighted acquisitions were analyzed as single-coil data. Sequence parameters were repetition time (TR) = 9.813 ms, echo time (TE) = 4.603 ms, flip angle=8° for T_1 scans, TR = 8178 ms, TE = 100 ms, flip angle=90° for T_2 scans, TR = 8178 ms, TE = 8 ms, flip angle=90° for PD scans, and a voxel size of $0.94 \times 0.94 \times 1.2$ mm³ 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 (GAN$_{prior}$, DDPM, DiffRecon). 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 $\beta_{min}, \beta_{max} = \{0.1, 20\}$ 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 6×10^{-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. (2020b). Hyperparameters were selected as 2×10^{-4} learning rate, 100 epochs, adversarial and pixel-wise loss weights of (1,100) for training (Dar et al., 2020b).
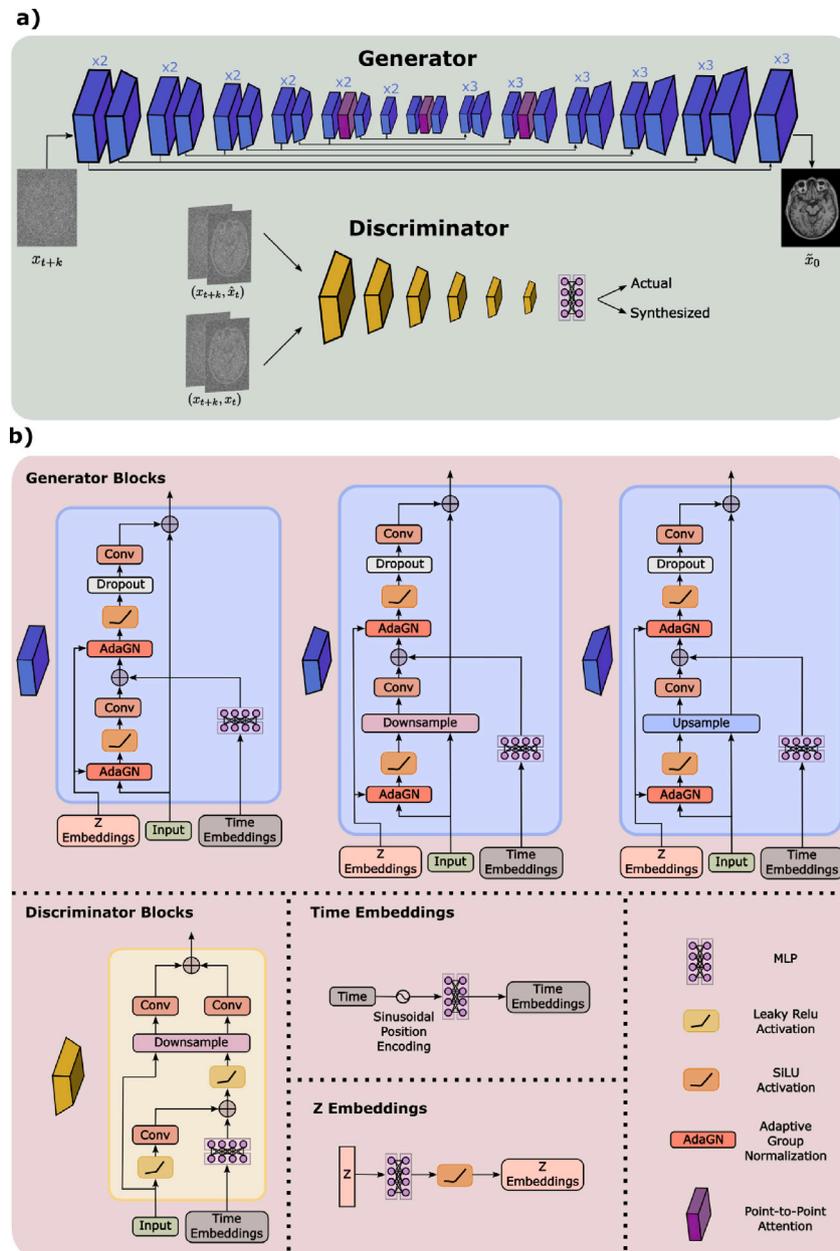
Fig. 3. Architectural overview of AdaDiff. **(a)** The generator processes the noisy image sample x_{t+k} via an encoder–decoder architecture comprising residual (blue) and attentional (purple) blocks to predict the denoised image $\tilde{x}_0$. Long-range skip connections are used to enhance information flow across blocks. The discriminator processes a pair of noisy image samples, either $(x_{t+k}, \hat{x}_t)$ or (x_{t+k}, x_t), with residual downsampling blocks to distinguish actual versus synthesized samples. An MLP block is used at the output layer. **(b)** Detailed structure of the blocks used in adversarial network. Each discriminator block convolutionally encodes the input image samples and MLP encodes the time embeddings, followed by downsampling and convolutional filtering. Three types of residual blocks are used in the generator based on feature map resolution: flat blocks (rectangular), downsampling blocks (trapezoidal), upsampling blocks (inverted trapezoidal). Each block convolutionally encodes the input image and MLP encodes the time embeddings, followed by convolutional filtering and/or downsampling/upsampling. Feature maps are subjected to adaptive normalization based on Z (i.e., latent variable) embeddings. Time embeddings are obtained by processing sinusoidal encoding of time step with an MLP. Z embeddings are obtained by processing a random latent vector z with an MLP.

MoDL: A conditional unrolled model that interleaves data-consistency blocks with convolutional layers was implemented (Aggarwal et al., 2019). The architecture and loss function were adopted from Dar et al. (2021). Hyperparameters were selected as 10^{-3} learning rate, 200 epochs for training (Dar et al., 2021).

GAN$_{prior}$: An unconditional GAN model that performs prior adaptation was implemented (Narnhofer et al., 2019). The architecture and loss function were adopted from Karras et al. (2019). GAN$_{prior}$ performed prior adaptation by minimizing a data-consistency loss as in AdaDiff. Hyperparameters were selected as 10^{-3} learning rate, 3000 epochs for training; 10^{-2} learning rate, 1000 iterations for inference (Elmas et al., 2022).

DDPM: An unconditional diffusion model was implemented with architecture and loss functions in Ho et al. (2020). Hyperparameters were selected as 10^{-4} learning rate, $k = 1$ step size, $T/k = 1000$ diffusion steps, 65 epochs for training; 1000 iterations combining a reverse diffusion step and a data-consistency projection for inference (Ho et al., 2020).

DiffRecon: An unconditional diffusion model was implemented with architecture and loss functions in Peng et al. (2022). Hyperparameters were selected as 10^{-4} learning rate, $k = 1$ step size, $T/k = 4000$ diffusion steps, 300 epochs for training; 100 coarse and 20 fine iterations of reverse diffusion and data-consistency projection for inference (Peng et al., 2022).

Table 1
Within-domain performance for T_1-, T_2-, PD-weighted contrasts in IXI at R = 4x-8x. PSNR (dB) and SSIM (%) are reported as mean±std across subjects. Boldface marks the method with the highest performance metric.

		LORAKS	rGAN	MoDL	GAN_{prior}	DDPM	DiffRecon	AdaDiff
		R = 4x						
T_1	PSNR	31.2 ± 2.2	36.2 ± 1.0	41.7 ± 1.3	38.0 ± 1.4	40.8 ± 1.3	40.5 ± 1.0	**42.1 ± 1.6**
	SSIM	83.0 ± 3.7	95.0 ± 1.0	99.0 ± 0.2	96.8 ± 1.3	98.7 ± 0.3	98.5 ± 0.3	**99.1 ± 0.3**
T_2	PSNR	32.3 ± 2.1	35.0 ± 0.7	41.4 ± 1.2	36.6 ± 1.8	40.2 ± 1.0	37.7 ± 1.4	**41.9 ± 1.6**
	SSIM	79.5 ± 2.9	90.3 ± 0.7	98.6 ± 0.2	94.4 ± 3.2	98.1 ± 0.2	96.5 ± 0.7	**98.9 ± 0.2**
PD	PSNR	31.3 ± 2.6	35.8 ± 1.0	42.0 ± 1.5	37.6 ± 1.9	40.9 ± 1.4	39.0 ± 0.4	**42.6 ± 1.9**
	SSIM	73.9 ± 3.8	91.0 ± 0.9	98.8 ± 0.2	95.8 ± 2.1	98.5 ± 0.2	97.4 ± 0.3	**99.1 ± 0.2**
		R = 8x						
T_1	PSNR	28.0 ± 1.7	32.4 ± 1.1	36.2 ± 1.2	32.8 ± 1.3	35.3 ± 1.2	36.0 ± 0.8	**36.3 ± 1.5**
	SSIM	77.4 ± 4.5	91.8 ± 1.5	97.5 ± 0.6	93.2 ± 2.4	96.8 ± 0.7	97.5 ± 0.4	**97.6 ± 0.6**
T_2	PSNR	28.8 ± 1.7	31.2 ± 0.7	35.9 ± 1.2	32.1 ± 1.6	34.6 ± 1.0	34.9 ± 0.4	**36.0 ± 1.5**
	SSIM	72.9 ± 3.6	85.5 ± 1.0	96.6 ± 0.5	91.1 ± 4.0	95.8 ± 0.6	95.5 ± 0.4	**97.3 ± 0.6**
PD	PSNR	28.0 ± 2.2	32.0 ± 1.1	36.3 ± 1.5	32.9 ± 1.8	35.2 ± 1.2	35.4 ± 0.5	**36.6 ± 1.8**
	SSIM	66.6 ± 4.7	86.4 ± 1.7	96.7 ± 0.7	92.7 ± 2.9	96.5 ± 0.6	96.4 ± 0.3	**97.6 ± 0.6**

Table 2
Within-domain performance for T_1-, T_2-, FLAIR- (FL.) weighted contrasts in fastMRI at R = 4x-8x.

|		LORAKS	rGAN	MoDL	GAN_{prior}	DDPM	DiffRecon	AdaDiff
		R = 4x						
T_1	PSNR	34.0 ± 2.3	38.0 ± 1.0	39.8 ± 1.3	31.7 ± 1.6	38.2 ± 1.7	38.6 ± 1.5	**40.2 ± 1.7**
	SSIM	83.8 ± 4.8	94.4 ± 1.2	95.7 ± 1.2	74.6 ± 4.3	92.5 ± 7.6	94.1 ± 1.8	**95.9 ± 1.4**
T_2	PSNR	34.8 ± 1.0	35.3 ± 0.7	36.7 ± 0.8	30.4 ± 0.8	37.5 ± 0.6	**39.1 ± 0.7**	37.7 ± 0.8
	SSIM	91.5 ± 1.5	94.7 ± 0.6	96.0 ± 0.5	79.6 ± 2.2	95.7 ± 0.5	**96.8 ± 0.4**	96.2 ± 0.4
FL	PSNR	28.6 ± 3.2	34.6 ± 1.9	36.0 ± 2.2	28.2 ± 2.6	34.1 ± 2.9	35.4 ± 2.3	**36.2 ± 2.6**
	SSIM	76.8 ± 9.6	91.0 ± 4.0	**92.7 ± 4.0**	72.3 ± 8.7	88.3 ± 6.9	91.7 ± 4.7	92.5 ± 4.8
		R = 8x						
T_1	PSNR	34.1 ± 2.0	35.6 ± 0.9	37.1 ± 1.1	26.7 ± 1.2	36.2 ± 1.3	34.7 ± 1.0	**37.2 ± 1.5**
	SSIM	85.3 ± 4.5	92.2 ± 1.6	**93.5 ± 1.6**	57.1 ± 4.5	91.1 ± 2.1	89.4 ± 1.8	**93.5 ± 2.2**
T_2	PSNR	33.9 ± 0.7	33.0 ± 0.7	33.9 ± 0.8	25.9 ± 0.8	34.8 ± 0.6	35.2 ± 0.5	**35.3 ± 0.8**
	SSIM	92.6 ± 1.0	92.8 ± 0.8	93.9 ± 0.7	66.0 ± 2.9	94.2 ± 0.6	93.9 ± 0.7	**94.4 ± 0.7**
FL	PSNR	28.9 ± 3.2	33.0 ± 1.9	**33.7 ± 1.9**	24.0 ± 1.9	32.8 ± 2.2	32.8 ± 1.6	**33.7 ± 2.4**
	SSIM	77.5 ± 10.2	88.3 ± 5.1	**88.9 ± 4.8**	57.1 ± 7.7	86.5 ± 6.6	87.3 ± 4.8	88.8 ± 6.2

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: (T_1,T_2,PD) in IXI and (T_1,T_2,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 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 LPIPS 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 (GAN_{prior}), 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 T_1 in general and on T_2 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 T_1, FLAIR in general, and DiffRecon that yields higher performance on T_2 at R = 4x and similar PSNR on T_2 at R = 8x. On average, AdaDiff outperforms the traditional method by 6.8 dB PSNR and 15.8% SSIM, conditional models by 2.0 dB PSNR and 2.5% SSIM, the adaptive GAN by 6.6 dB PSNR and 15.0% SSIM, and static diffusion models by 1.3 dB PSNR and 1.4% SSIM.
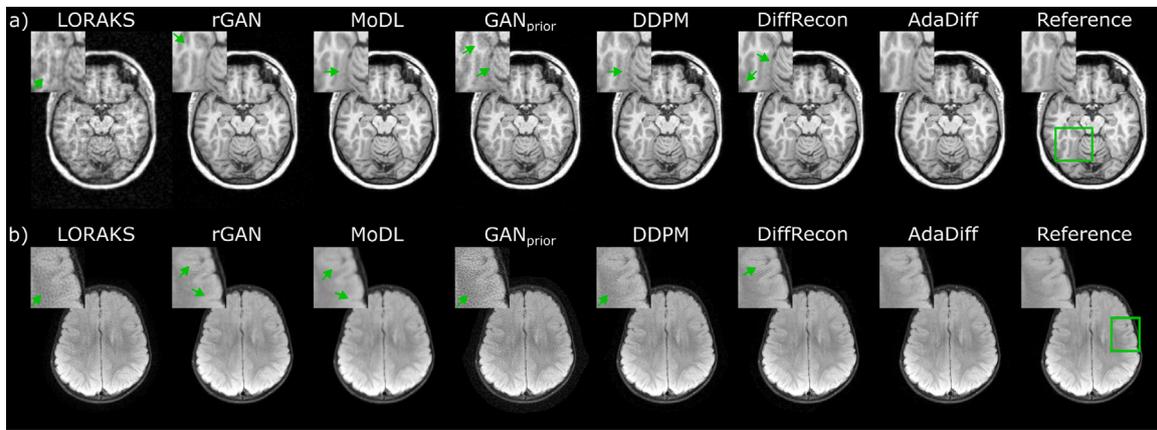
Fig. 4. Within-domain reconstructions at R = 4x. Results are shown for **(a)** T_1-weighted acquisitions in IXI, and **(b)** FLAIR-weighted acquisitions in fastMRI. Reconstructed images are given along with the reference image derived from fully-sampled acquisitions, and zoom-in windows and arrows are included to highlight differences among methods. LORAKS and GAN_{prior} 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.

Table 3
Cross-domain performance for T_1-, T_2-, PD-weighted contrasts in IXI. Results listed for training at R = 4x, testing at R = 8x and R = 12x.

		LORAKS	rGAN	MoDL	GAN_{prior}	DDPM	DiffRecon	AdaDiff
		R = 8x						
T_1	PSNR	28.0 ± 1.7	31.6 ± 1.0	34.7 ± 1.3	32.8 ± 1.3	35.3 ± 1.2	36.0 ± 0.8	**36.3 ± 1.5**
	SSIM	77.4 ± 4.5	90.7 ± 1.6	96.7 ± 0.8	93.2 ± 2.4	96.8 ± 0.7	97.5 ± 0.4	**97.6 ± 0.6**
T_2	PSNR	28.8 ± 1.7	31.1 ± 0.8	34.2 ± 1.4	32.1 ± 1.6	34.6 ± 1.0	34.9 ± 0.4	**36.0 ± 1.5**
	SSIM	72.9 ± 3.6	86.2 ± 1.1	95.4 ± 0.8	91.1 ± 4.0	95.8 ± 0.6	95.5 ± 0.4	**97.3 ± 0.6**
PD	PSNR	28.0 ± 2.2	31.6 ± 1.2	34.9 ± 1.7	32.9 ± 1.8	35.2 ± 1.2	35.4 ± 0.5	**36.6 ± 1.8**
	SSIM	66.6 ± 4.7	85.6 ± 1.9	94.9 ± 0.8	92.7 ± 2.9	96.5 ± 0.6	96.4 ± 0.3	**97.6 ± 0.6**
		R = 12x						
T_1	PSNR	26.5 ± 1.6	29.2 ± 1.0	31.6 ± 1.3	31.1 ± 1.3	32.7 ± 1.1	32.9 ± 0.9	**33.4 ± 1.3**
	SSIM	73.7 ± 5.3	86.7 ± 2.4	92.7 ± 1.5	92.1 ± 1.9	95.2 ± 1.0	95.3 ± 0.8	**96.2 ± 1.0**
T_2	PSNR	27.3 ± 1.6	28.9 ± 0.9	31.1 ± 1.3	29.5 ± 1.5	32.0 ± 1.0	31.9 ± 0.8	**33.1 ± 1.5**
	SSIM	69.2 ± 4.1	81.4 ± 1.7	89.2 ± 1.6	86.3 ± 4.8	94.0 ± 1.0	93.1 ± 0.7	**95.7 ± 1.0**
PD	PSNR	26.4 ± 2.0	29.3 ± 1.2	31.7 ± 1.6	31.1 ± 1.6	32.6 ± 1.2	32.7 ± 1.0	**33.9 ± 1.8**
	SSIM	62.1 ± 5.2	80.2 ± 2.3	85.5 ± 2.4	90.4 ± 3.3	94.8 ± 1.0	94.5 ± 0.7	**96.2 ± 1.0**

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 GAN_{prior} 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_1 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 4.9 dB PSNR and 16.0% SSIM, conditional models by 2.3 dB PSNR and 4.4% SSIM, the adaptive GAN by 6.8 dB PSNR and 20.6% SSIM, and static diffusion models by 0.9 dB PSNR and 1.5% SSIM. Note that, at R = 8x, the performance benefit of AdaDiff over MoDL is 1.5 dB PSNR, 2.3% SSIM under domain shift in acceleration rate, versus 0.3 dB 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 shows 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

Table 4
Cross-domain performance for T_1-, T_2-, and FLAIR- (FL.) weighted contrasts in fastMRI. Results for training at R = 4x, testing at R = 8x and R = 12x.

		LORAKS	rGAN	MoDL	GAN$_{prior}$	DDPM	DiffRecon	AdaDiff
		R = 8x						
T_1	PSNR	34.1 ± 2.0	36.0 ± 0.9	36.4 ± 1.0	26.7 ± 1.2	36.2 ± 1.3	34.7 ± 1.0	**37.2 ± 1.5**
	SSIM	85.3 ± 4.5	92.4 ± 1.5	93.3 ± 1.6	57.1 ± 4.5	91.1 ± 2.1	89.4 ± 1.8	**93.5 ± 2.2**
T_2	PSNR	33.9 ± 0.7	33.0 ± 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	66.0 ± 2.9	94.2 ± 0.6	93.9 ± 0.7	**94.4 ± 0.6**
FL	PSNR	28.9 ± 3.2	32.7 ± 1.8	33.2 ± 1.9	24.0 ± 1.9	32.8 ± 2.2	32.8 ± 1.6	**33.7 ± 2.4**
	SSIM	77.5 ± 10.2	87.7 ± 4.9	**88.8 ± 4.8**	57.1 ± 7.7	86.5 ± 6.6	87.3 ± 4.8	**88.8 ± 6.2**
		R = 12x						
T_1	PSNR	34.1 ± 1.6	34.6 ± 0.9	34.7 ± 1.0	24.4 ± 1.1	35.9 ± 0.9	33.0 ± 0.8	**35.2 ± 1.6**
	SSIM	86.9 ± 4.0	91.1 ± 1.6	**91.7 ± 1.8**	48.8 ± 5.0	91.2 ± 1.5	86.4 ± 1.9	91.2 ± 2.7
T_2	PSNR	32.9 ± 0.7	31.5 ± 0.8	31.2 ± 0.8	23.0 ± 0.8	33.2 ± 0.5	33.2 ± 0.5	**33.9 ± 0.8**
	SSIM	92.2 ± 0.9	90.5 ± 1.0	91.4 ± 1.0	56.1 ± 3.4	92.8 ± 0.8	91.3 ± 1.0	**93.2 ± 0.8**
FL	PSNR	29.5 ± 3.0	31.6 ± 1.7	31.8 ± 1.7	22.1 ± 1.6	32.2 ± 1.7	31.4 ± 1.4	**32.3 ± 2.3**
	SSIM	79.2 ± 9.9	85.5 ± 5.3	86.2 ± 5.3	50.2 ± 7.3	86.3 ± 5.7	84.1 ± 4.8	**86.4 ± 6.9**

Table 5
Cross-domain performance for T_1-, T_2-, and FL.-weighted contrasts in fastMRI at R = 4x. Training under 2D undersampling with 5 coils, testing under 1D undersampling with 5 coils and 2D undersampling with 10 coils.

		LORAKS	rGAN	MoDL	GAN$_{prior}$	DDPM	DiffRecon	AdaDiff
		1D, 5 coils						
T_1	PSNR	33.4 ± 1.6	34.6 ± 1.0	35.3 ± 1.1	28.7 ± 1.3	35.7 ± 1.1	32.1 ± 0.9	**36.5 ± 1.8**
	SSIM	84.9 ± 4.2	91.7 ± 1.3	92.8 ± 1.3	66.8 ± 4.2	91.2 ± 1.3	85.1 ± 2.4	**93.3 ± 2.0**
T_2	PSNR	32.2 ± 0.8	30.7 ± 0.8	31.1 ± 0.9	28.4 ± 0.7	33.5 ± 0.6	33.0 ± 0.5	**34.0 ± 0.8**
	SSIM	91.9 ± 1.5	90.2 ± 1.2	90.9 ± 1.2	77.1 ± 2.5	93.4 ± 0.8	92.7 ± 1.0	**94.3 ± 0.8**
FL	PSNR	28.1 ± 2.5	31.7 ± 1.5	32.3 ± 1.5	26.1 ± 2.0	32.7 ± 1.6	31.5 ± 1.3	**32.8 ± 2.0**
	SSIM	78.2 ± 6.3	88.3 ± 2.7	**89.7 ± 2.5**	66.6 ± 6.5	88.6 ± 3.5	86.5 ± 3.6	89.3 ± 4.3
		2D, 10 coils						
T_1	PSNR	33.2 ± 1.9	38.1 ± 1.0	39.7 ± 1.3	32.1 ± 1.7	38.4 ± 1.5	38.8 ± 1.5	**40.2 ± 1.7**
	SSIM	82.3 ± 4.2	94.6 ± 1.3	95.7 ± 1.3	76.4 ± 4.4	93.3 ± 1.8	94.5 ± 1.8	**96.1 ± 1.5**
T_2	PSNR	33.6 ± 1.0	35.4 ± 0.7	36.7 ± 0.7	30.5 ± 0.7	37.4 ± 0.6	**39.2 ± 0.7**	37.7 ± 0.8
	SSIM	90.3 ± 1.8	95.0 ± 0.5	96.1 ± 0.5	80.6 ± 2.0	95.9 ± 0.4	**97.1 ± 0.3**	96.4 ± 0.4
FL	PSNR	27.2 ± 3.0	34.6 ± 2.3	36.0 ± 2.5	28.4 ± 3.0	34.6 ± 2.9	35.5 ± 2.6	**36.3 ± 3.0**
	SSIM	74.5 ± 9.5	91.1 ± 5.3	**92.5 ± 5.3**	73.1 ± 10.2	89.6 ± 7.5	91.6 ± 6.0	92.4 ± 6.4
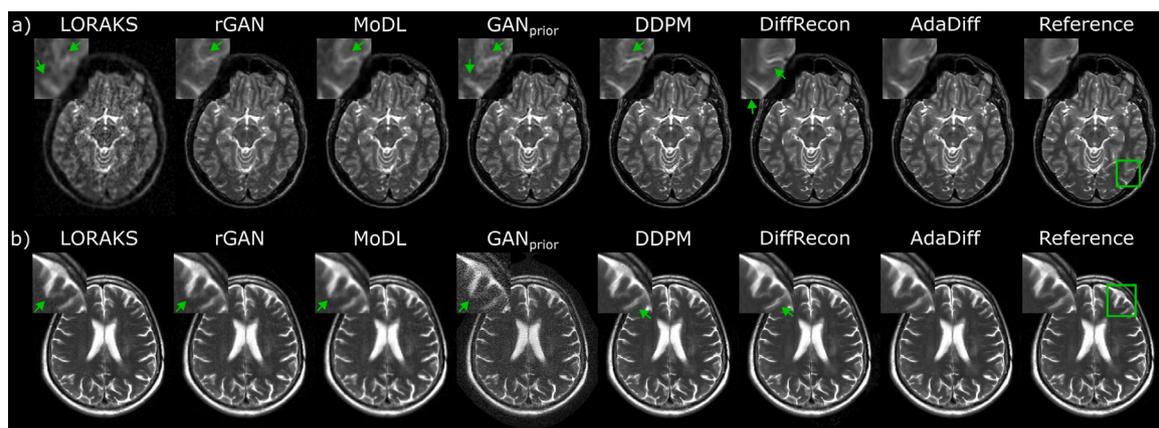
Fig. 5. Cross-domain reconstructions under domain shifts in the acceleration rate. Results are shown for **(a)** T_2-weighted acquisitions at R = 8x in IXI, and **(b)** T_2-weighted acquisitions at R = 12x in fastMRI. Reconstructed images are given along with the reference image derived from fully-sampled acquisitions, and zoom-in windows and arrows are included to highlight differences among methods. Conditional models were trained at R = 4x. LORAKS and GAN$_{prior}$ show high noise amplification, rGAN and MoDL show some residual reconstruction artifacts and spatial blurring. Among diffusion models, DDPM has relatively high noise and DiffRecon has local ringing artifacts. AdaDiff reconstructs images with low artifacts/noise and a closer appearance to the reference images.

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 SSIM 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
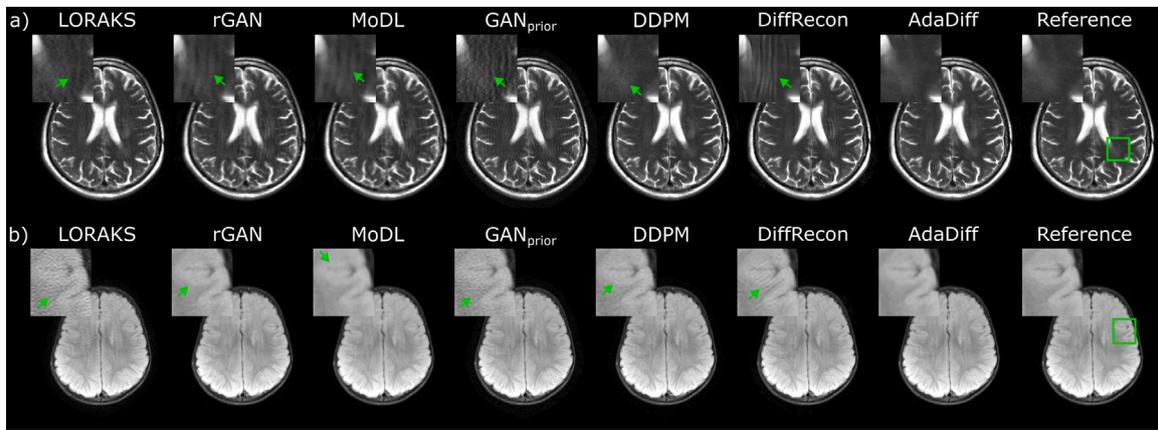
Fig. 6. Cross-domain reconstructions at R = 4x under domain shifts in the sampling trajectory and number of coils. Results are shown for **(a)** T_2-weighted acquisitions with 1D undersampling, and **(b)** PD-weighted acquisitions with 10 virtual coils in fastMRI. Reconstructed images are given along 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.

Table 6

Cross-domain performance for T_1-, T_2-, PD-weighted contrasts at R = 4x-8x. Results listed for training on fastMRI, testing on IXI.

		LORAKS	rGAN	MoDL	GAN_{prior}	DDPM	DiffRecon	AdaDiff
		R = 4x						
T_1	PSNR	31.8 ± 3.4	30.3 ± 2.1	34.1 ± 1.8	28.2 ± 1.8	35.3 ± 1.7	38.8 ± 0.9	**41.0 ± 2.1**
	SSIM	83.5 ± 5.4	79.4 ± 5.4	89.3 ± 2.4	73.3 ± 5.3	90.1 ± 2.3	**98.6 ± 0.2**	**98.6 ± 0.5**
T_2	PSNR	32.6 ± 3.6	30.1 ± 1.3	32.8 ± 1.7	27.2 ± 1.7	35.4 ± 0.8	38.6 ± 0.5	**40.5 ± 1.9**
	SSIM	80.8 ± 6.4	72.9 ± 3.6	81.3 ± 3.0	59.3 ± 6.1	87.8 ± 1.1	**98.3 ± 0.2**	98.1 ± 0.5
PD	PSNR	32.0 ± 3.9	29.6 ± 1.9	32.9 ± 2.0	28.4 ± 1.9	34.5 ± 0.9	39.1 ± 0.7	**40.8 ± 2.0**
	SSIM	75.2 ± 7.9	68.7 ± 5.3	79.3 ± 4.3	64.4 ± 6.2	83.6 ± 1.7	**98.7 ± 0.2**	98.0 ± 0.6
		R = 8x						
T_1	PSNR	28.4 ± 2.6	27.0 ± 1.9	30.8 ± 1.6	22.9 ± 1.1	32.1 ± 1.7	34.4 ± 0.8	**35.6 ± 2.0**
	SSIM	78.0 ± 6.6	71.1 ± 6.9	85.5 ± 3.1	55.9 ± 6.4	86.5 ± 2.9	**96.9 ± 0.5**	96.1 ± 1.3
T_2	PSNR	29.1 ± 2.8	27.4 ± 0.9	30.0 ± 1.5	22.8 ± 1.4	32.0 ± 0.9	34.2 ± 0.4	**35.1 ± 1.9**
	SSIM	74.5 ± 7.8	66.1 ± 3.2	76.9 ± 3.3	41.5 ± 7.1	82.9 ± 1.6	**96.5 ± 0.5**	95.8 ± 1.3
PD	PSNR	28.6 ± 3.2	26.1 ± 1.5	30.0 ± 1.8	22.7 ± 1.8	31.4 ± 1.0	34.5 ± 0.6	**35.6 ± 2.0**
	SSIM	68.0 ± 9.4	58.1 ± 5.4	73.9 ± 4.6	43.1 ± 8.1	78.6 ± 2.1	**96.9 ± 0.5**	95.2 ± 1.6

SSIM on FLAIR, and DiffRecon that yields higher performance on T_2. On average, AdaDiff outperforms the traditional method by 5.0 dB PSNR and 10.0% SSIM, conditional models by 1.6 dB PSNR and 1.2% SSIM, the adaptive GAN by 7.2 dB PSNR and 20.2% SSIM, and static diffusion models by 1.0 dB PSNR and 2.0% SSIM. Note that, at R = 4x in fastMRI, the performance benefit of AdaDiff over MoDL is 1.6 dB PSNR, 1.2% SSIM under domain shift in sampling trajectory, 0.6 dB PSNR, 0.2% SSIM under domain shift in number of coils, versus 0.5 dB 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 mis-matched 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 7.7 dB PSNR and 20.3% SSIM, conditional models by 8.0 dB PSNR and 21.8% SSIM, the adaptive GAN by 12.7 dB PSNR and 40.7% SSIM, and static diffusion models by 3.1 dB 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 LO-RAKS and GAN_{prior}, 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
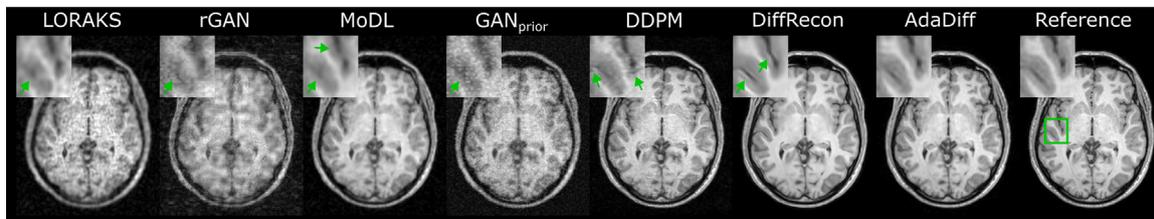
Fig. 7. Cross-domain reconstructions under domain shifts in the MR image distribution. Results are shown for T_1-weighted acquisitions at R = 8x. All models were trained on fastMRI and tested on IXI. The imaging operator matched between training and testing for conditional models. Reconstructed images 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 GAN$_{prior}$, 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.

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 = 4x in IXI.

		$T/k = 4$	$T/k = \mathbf{8}$	$T/k = 16$	$T/k = 32$
T_1	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
T_2	FID	37.75	30.75	30.18	28.92
	LPIPS	12.80	12.09	11.81	11.82
	PSNR	40.44	40.93	40.97	40.97
	SSIM	98.20	98.57	98.70	98.68
PD	FID	38.41	32.04	32.92	30.67
	LPIPS	12.92	12.63	12.59	12.72
	PSNR	40.79	41.23	41.29	41.21
	SSIM	98.41	98.75	98.85	98.82

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.

		$N_e = \mathbf{500}$	$N_e = 1000$	$N_e = 1500$	$N_e = 2000$
T_1	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
T_2	FID	30.75	29.65	30.46	29.96
	LPIPS	12.09	12.09	12.18	12.16
	PSNR	40.93	40.84	40.68	40.65
	SSIM	98.57	98.55	98.45	98.42
PD	FID	32.04	30.41	30.0	29.97
	LPIPS	12.63	12.59	12.54	12.51
	PSNR	41.23	41.19	41.16	41.12
	SSIM	98.75	98.73	98.68	98.67

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

		w/o z	AdaDiff
T_1	FID	34.05	31.99
	LPIPS	4.47	4.27
	PSNR	40.84	40.89
	SSIM	98.87	98.88
T_2	FID	31.39	30.75
	LPIPS	12.16	12.09
	PSNR	40.83	40.93
	SSIM	98.52	98.57
PD	FID	32.34	32.04
	LPIPS	12.72	12.63
	PSNR	41.17	41.23
	SSIM	98.69	98.75

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 adversarial loss with a pixel-wise ℓ_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 10.72 dB, 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 0.94 dB, 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 0.42 dB, 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>

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

Table 10
Performance of AdaDiff and ablated variants. A variant omitting the prior adaptation phase (w/o adapt.), 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.

		ZF	w/o adapt.	J = 500 iterations			J = 1000 iterations		
				w/o train.	w/o adv.	AdaDiff	w/o train.	w/o adv.	AdaDiff
T_1	FID	273.00	406.78	64.26	64.70	44.59	38.88	36.35	31.99
	LPIPS	40.27	49.87	7.64	8.74	5.45	5.27	4.89	4.27
	PSNR	32.55	30.85	39.03	39.65	40.23	40.33	40.90	40.89
	SSIM	88.80	84.99	98.06	98.53	98.71	98.62	98.90	98.88
T_2	FID	280.57	375.95	65.25	62.48	46.64	48.24	34.14	30.75
	LPIPS	53.64	76.31	15.07	15.75	13.49	13.32	12.64	12.09
	PSNR	31.64	29.04	39.08	39.59	40.20	40.15	40.86	40.93
	SSIM	81.16	73.87	97.41	97.54	98.05	98.21	98.46	98.57
PD	FID	298.05	423.69	73.84	63.63	45.60	54.81	36.40	32.04
	LPIPS	53.64	70.92	15.59	16.85	13.69	14.27	13.23	12.63
	PSNR	32.12	29.89	39.33	39.54	40.45	40.48	40.93	41.23
	SSIM	81.94	76.18	97.67	97.51	98.27	98.48	98.54	98.75

Table 11
Training and inference times in seconds per cross-section for reconstructions at R = 4x in IXI.

	LORAKS	rGAN	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	131.3
Inference	3.0	0.03	0.05	129.6	57.5	12.0	0.4	131.0	131.4	131.4

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 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.

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.

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 Narnhofer 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.

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., 2020a). 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., 2020b; Polak et al., 2020; Gaillochet 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; Oh et al., 2020; Ozbey 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., 2022b; Güngör 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; Küstner et al., 2020; Cole et al., 2021; Wang et al., 2020a). 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.

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.

Data availability

The url for the code used for analyses in the study is provided in the manuscript. The public repositories containing data have been listed in the manuscript.

Acknowledgments

References

Adler, J., Oktem, O., 2018. Learned primal-dual reconstruction. IEEE Trans. Med. Imaging 37 (6), 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 Trans. Med. Imaging 38 (2), 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., Reehorst, E.T., Schniter, P., 2020. Plug-and-play methods for magnetic resonance imaging: Using denoisers for image recovery. IEEE Signal Process. Mag. 37 (1), 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. Magn. Reson. Med. 82 (1), 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. Med. Image Anal. 71, 102047.

Chung, H., Sim, B., Ye, J.C., 2022. Come-closer-diffuse-faster: Accelerating conditional diffusion models for inverse problems through stochastic contraction. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. CVPR, pp. 12413–12422.

Chung, H., Ye, J.C., 2022. Score-based diffusion models for accelerated MRI. Med. Image Anal. 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. Magn. Reson. Med. 86 (2), 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., Avestimehr, 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 Trans. Med. Imaging 41 (10), 2598–2614.

Dar, S.U.H., Özbey, M., Çatli, A.B., Çukur, T., 2020a. A transfer-learning approach for accelerated MRI using deep neural networks. Magn. Reson. Med. 84 (2), 663–685.

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.H., Yurt, M., Karacan, L., Erdem, A., Erdem, E., Çukur, T., 2019. Image synthesis in multi-contrast MRI with conditional generative adversarial networks. IEEE Trans. Med. Imaging 38 (10), 2375–2388.

Dar, S.U., Yurt, M., Shahdloo, M., Ildız, M.E., Tınaz, B., Çukur, T., 2020b. Prior-guided image reconstruction for accelerated multi-contrast MRI via generative adversarial networks. IEEE J. Sel. Top. Sign. Proces. 14 (6), 1072–1087.

Darestani, M.Z., Heckel, R., 2021. Accelerated MRI with un-trained neural networks. IEEE Trans. Comput. Imag. 7, 724–733.

Dedmari, M.A., Conjeti, S., Estrada, S., Ehses, P., Stöcker, T., Reuter, M., 2018. Complex fully convolutional neural networks for MR 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. Vol. 34. Curran Associates, Inc, pp. 8780–8794.

Elmas, G., Dar, S.U., Korkmaz, Y., Ceyani, E., Susam, B., Ozbey, M., Avestimehr, S., Çukur, T., 2022. Federated learning of generative image priors for MRI reconstruction. IEEE Trans. Med. Imaging http://dx.doi.org/10.1109/TMI.2022.3220757.

Eo, T., Jun, Y., Kim, T., Jang, J., Lee, H.-J., Hwang, D., 2018. KIKI-net: cross-domain convolutional neural networks for reconstructing undersampled magnetic resonance images. Magn. Reson. Med. 80 (5), 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 Trans. Neural Netw. Learn. Syst. 1–11.

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

Gu, H., Yaman, B., Ugurbil, K., Moeller, S., Akçakaya, M., 2021. Compressed sensing MRI with ℓ1-wavelet reconstruction revisited using modern data science tools. In: International Conference of the IEEE Engineering in Medicine & Biology Society. pp. 3596–3600.

Güngör, A., Askin, B., Soydan, D.A., Saritas, E.U., Top, C.B., Çukur, T., 2022. TranSMS: Transformers for super-resolution calibration in magnetic particle imaging. IEEE Trans. Med. Imaging 41 (12), 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 RNN for MRI reconstruction. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 13–23.

Haldar, J.P., Zhuo, J., 2016. P-LORAKS: Low-rank modeling of local k-space neighborhoods with parallel imaging data. Magn. Reson. Med. 75 (4), 1499.

Hammernik, K., Klatzer, T., Kobler, E., Recht, M.P., Sodickson, D.K., Pock, T., Knoll, F., 2017. Learning a variational network for reconstruction of accelerated MRI data. Magn. Reson. Med. 79 (6), 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. Magn. Reson. Med. 80 (3), 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. Vol. 30.

Ho, J., Jain, A., Abbeel, P., 2020. Denoising diffusion probabilistic models. In: Advances in Neural Information Processing Systems. Vol. 33. pp. 6840–6851.

Hosseini, S.A.H., Yaman, B., Moeller, S., Hong, M., Akcakaya, M., 2020. Dense recurrent neural networks for accelerated MRI: History-cognizant unrolling of optimization algorithms. IEEE J. Sel. Top. Sign. Proces. 14 (6), 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. Phys. Med. Biol. 63 (13), 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. Vol. 34. pp. 14938–14954.

Jin, K.H., Gupta, H., Yerly, 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: IEEE CVPR. 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.

Knoll, F., Hammernik, K., Kobler, E., Pock, T., Recht, M.P., Sodickson, D.K., 2019. Assessment of the generalization of learned image reconstruction and the potential for transfer learning. Magn. Reson. Med. 81 (1), 116–128.

Knoll, F., Zbontar, J., Sriram, A., Muckley, M.J., Bruno, M., Defazio, A., Parente, M., Geras, K.J., Katsnelson, J., Chandarana, H., Zhang, Z., Drozdzalv, M., Romero, A., Rabbat, M., Vincent, P., Pinkerton, J., Wang, D., Yakubova, N., Owens, E., Zitnick, C.L., Recht, M.P., Sodickson, D.K., Lui, Y.W., 2020. fastMRI: A publicly available raw k-space and DICOM dataset of knee images for accelerated MR image reconstruction using machine learning. Radiol. Artif. Intell. 2 (1), e190007.

Korkmaz, Y., Dar, S.U.H., Yurt, M., Ozbey, M., Çukur, T., 2022. Unsupervised MRI reconstruction via zero-shot learned adversarial transformers. IEEE Trans. Med. Imaging 41 (7), 1747–1763.

Küstner, T., Fuin, N., Hammernik, K., Bustin, A., Qi, H., Hajhosseiny, R., Masci, P.G., Neji, R., Rueckert, D., Botnar, R.M., Prieto, C., 2020. CINENet: deep learning-based 3D cardiac CINE MRI reconstruction with multi-coil complex-valued 4D spatio-temporal convolutions. Sci. Rep. 10 (1), 1–13.

Kwon, K., Kim, D., Park, H., 2017. A parallel MR imaging method using multilayer perceptron. Med. Phys. 44 (12), 6209–6224.

Lee, D., Yoo, J., Tak, S., Ye, J.C., 2018. Deep residual learning for accelerated MRI using magnitude and phase networks. IEEE Trans. Biomed. Eng. 65 (9), 1985–1995.

Lei, K., Mardani, M., Pauly, J.M., Vasanawala, S.S., 2021. Wasserstein GANs for MR imaging: From paired to unpaired training. IEEE Trans. Med. Imaging 40 (1), 105–115.

Liang, D., Cheng, J., Ke, Z., Ying, L., 2020. Deep magnetic resonance image reconstruction: Inverse problems meet neural networks. IEEE Signal Process. Mag. 37 (1), 141–151.

Liu, J., Sun, Y., Eldeniz, C., Gan, W., An, H., Kamilov, U.S., 2020. RARE: Image reconstruction using deep priors learned without groundtruth. IEEE J. Sel. Top. Sign. Proces. 14 (6), 1088–1099.

Liu, X., Wang, J., Liu, F., Zhou, S.K., 2021. Universal undersampled MRI reconstruction. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 211–221.

Liu, X., Wang, J., Peng, C., Chandra, S.S., Liu, F., Zhou, S.K., 2022. Undersampled MRI reconstruction with side information-guided normalisation. arXiv preprint arXiv:2203.03196.

Liu, Q., Yang, Q., Cheng, H., Wang, S., Zhang, M., Liang, D., 2020. Highly undersampled magnetic resonance imaging reconstruction using autoencoding priors. Magn. Reson. Med. 83 (1), 322–336.

Lu, Z., Li, Z., Wang, J., Shi, J., Shen, D., 2021. Two-stage self-supervised cycle-consistency network for reconstruction of thin-slice MR images. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 3–12.

Luo, G., Heide, M., Uecker, M., 2022. MRI reconstruction via data driven Markov chain with joint uncertainty estimation. arXiv:2202.01479.

Luo, G., Zhao, N., Jiang, W., Hui, E.S., Cao, P., 2020. MRI reconstruction using deep Bayesian estimation. Magn. Reson. Med. 84 (4), 2246–2261.

Lustig, M., Donoho, D., Pauly, J.M., 2007. Sparse MRI: The application of compressed sensing for rapid MR imaging. Magn. Reson. Med. 58 (6), 1182–1195.

Mardani, M., Gong, E., Cheng, J.Y., Vasanawala, S., Zaharchuk, G., Xing, L., Pauly, J.M., 2019. Deep generative adversarial neural networks for compressive sensing MRI. IEEE Trans. Med. Imaging 38 (1), 167–179.

Mescheder, L., Geiger, A., Nowozin, S., 2018. Which training methods for GANs do actually converge? In: Proceedings of ICML. Vol. 80. pp. 3481–3490.

Narnhofer, D., Hammernik, K., Knoll, F., Pock, T., 2019. Inverse GANs for accelerated MRI reconstruction. In: Proceedings of SPIE. Vol. 11138. pp. 381–392.

Oh, G., Sim, B., Chung, H., Sunwoo, L., Ye, J.C., 2020. Unpaired deep learning for accelerated MRI using optimal transport driven cycleGAN. IEEE Trans. Comput. Imag. 6, 1285–1296.

Ozbey, M., Dalmaz, O., Dar, S.U., Bedel, H.A., Ozturk, S., Güngör, A., Çukur, T., 2022. Unsupervised medical image translation with adversarial diffusion models. arXiv preprint arXiv:2207.08208.

Peng, C., Guo, P., Zhou, S.K., Patel, V., Chellappa, R., 2022. Towards performant and reliable undersampled MR reconstruction via diffusion model sampling. arXiv:2203.04292.

Pock, T., Sabach, S., 2016. Inertial proximal alternating linearized minimization (iPALM) for nonconvex and nonsmooth problems. SIAM J. Imaging Sci. 9 (4), 1756–1787.

Polak, D., Cauley, S., Bilgic, B., Gong, E., Bachert, P., Adalsteinsson, E., Setsompop, K., 2020. Joint multi-contrast variational network reconstruction (jVN) with application to rapid 2D and 3D imaging. Magn. Reson. Med. 84 (3), 1456–1469.

Qin, C., Schlemper, J., Caballero, J., Price, A.N., Hajnal, J.V., Rueckert, D., 2019. Convolutional recurrent neural networks for dynamic MR image reconstruction. IEEE Trans. Med. Imaging 38 (1), 280–290.

Quan, T.M., Nguyen-Duc, T., Jeong, W.-K., 2018. Compressed sensing MRI reconstruction with cyclic loss in generative adversarial networks. IEEE Trans. Med. Imaging 37 (6), 1488–1497.

Schlemper, J., Caballero, J., Hajnal, J.V., Price, A., Rueckert, D., 2017. A deep cascade of convolutional neural networks for MR image reconstruction. In: Proceedings of IPMI. pp. 647–658.

Song, Y., Shen, L., Xing, L., Ermon, S., 2022. Solving inverse problems in medical imaging with score-based generative models. In: International Conference on Learning Representations.

Song, Y., Sohl-Dickstein, J., Kingma, D.P., Kumar, A., Ermon, S., Poole, B., 2020. Score-based generative modeling through stochastic differential equations. arXiv:2011.13456.

Sriram, A., Zbontar, J., Murrell, T., Defazio, A., Zitnick, C.L., Yakubova, N., Knoll, F., Johnson, P., 2020. End-to-end variational networks for accelerated MRI reconstruction. In: Proceedings of MICCAI. pp. 64–73.

Tamir, J.I., Yu, S.X., Lustig, M., 2019. Unsupervised deep basis pursuit: Learning reconstruction without ground-truth data. In: Proceedings of ISMRM. p. 0660.

Tezcan, K.C., Baumgartner, C.F., Luechinger, R., Pruessmann, K.P., Konukoglu, E., 2019. MR image reconstruction using deep density priors. IEEE Trans. Med. Imaging 38 (7), 1633–1642.

Tezcan, K.C., Karani, N., Baumgartner, C.F., Konukoglu, E., 2022. Sampling possible reconstructions of undersampled acquisitions in MR imaging with a deep learned prior. IEEE Trans. Med. Imaging.

Uecker, M., Lai, P., Murphy, M.J., Virtue, P., Elad, M., Pauly, J.M., Vasanawala, S.S., Lustig, M., 2014. ESPIRiT-an eigenvalue approach to autocalibrating parallel MRI: Where SENSE meets GRAPPA. Magn. Reson. Med. 71 (3), 990–1001.

Ulyanov, D., Vedaldi, A., Lempitsky, V., 2018. Deep image prior. In: IEEE CVPR. pp. 9446–9454.

Wang, S., Cheng, H., Ying, L., Xiao, T., Ke, Z., Zheng, H., Liang, D., 2020a. DeepcomplexMRI: Exploiting deep residual network for fast parallel MR imaging with complex convolution. Magn. Reson. Imag. 68, 136–147.

Wang, A.Q., Dalca, A.V., Sabuncu, M.R., 2020b. Neural network-based reconstruction in compressed sensing MRI without fully-sampled training data. In: MLMIR. pp. 27–37.

Wang, S., Ke, Z., Cheng, H., Jia, S., Ying, L., Zheng, H., Liang, D., 2019. DIMENSION: Dynamic MR imaging with both k-space and spatial prior knowledge obtained via multi-supervised network training. NMR Biomed. e4131.

Wang, S., Su, Z., Ying, L., Peng, X., Zhu, S., Liang, F., Feng, D., Liang, D., 2016. Accelerating magnetic resonance imaging via deep learning. In: IEEE ISBI. pp. 514–517.

Xiao, Z., Kreis, K., Vahdat, A., 2022. Tackling the generative learning trilemma with denoising diffusion GANs. In: International Conference on Learning Representations. ICLR.

Xie, Y., Li, Q., 2022. Measurement-conditioned denoising diffusion probabilistic model for under-sampled medical image reconstruction. arXiv:2203.03623.

Xuan, K., Xiang, L., Huang, X., Zhang, L., Liao, S., Shen, D., Wang, Q., 2022. Multi-modal MRI reconstruction assisted with spatial alignment network. IEEE Trans. Med. Imaging.

Yaman, B., Hosseini, S.A.H., Akcakaya, M., 2021. Zero-shot physics-guided deep learning for subject-specific MRI reconstruction. In: NeurIPS 2021 Workshop on Deep Learning and Inverse Problems.

Yaman, B., Hosseini, S.A.H., Moeller, S., Ellermann, J., Uğurbil, K., Akçakaya, M., 2020. Self-supervised learning of physics-guided reconstruction neural networks without fully sampled reference data. Magn. Reson. Med. 84 (6), 3172–3191.

Yang, Y., Sun, J., Li, H., Xu, Z., 2016. Deep ADMM-Net for compressive sensing MRI. In: Advances in Neural Information Processing Systems. Vol. 29.

Yang, Y., Sun, J., Li, H., Xu, Z., 2020. ADMM-CSNet: A deep learning approach for image compressive sensing. IEEE Trans. Pattern Anal. Mach. Intell. 42 (3), 521–538.

Ye, J.C., Han, Y., Cha, E., 2018. Deep convolutional framelets: A general deep learning framework for inverse problems. SIAM J. Imaging Sci. 11 (2), 991–1048.

Yoon, J., Gong, E., Chatnuntawech, I., Bilgic, B., Lee, J., Jung, W., Ko, J., Jung, H., Setsompop, 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 Trans. Med. Imaging 37 (6), 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. Magn. Reson. Med. 69 (2), 571–582.

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. Vol. 33. pp. 21853–21864.

Zhu, B., Liu, J.Z., Rosen, B.R., Rosen, M.S., 2018. Image reconstruction by domain transform manifold learning. Nature 555 (7697), 487–492.

Zou, Q., Ahmed, A.H., Nagpal, P., Kruger, S., Jacob, M., 2021. Deep generative SToRM model for dynamic imaging. arXiv:2101.12366.