Semi-Supervised Learning of MRI Synthesis without Fully-Sampled Ground Truths

Mahmut Yurt, Onat Dalmaz, Salman Dar, Muzaffer Ozbey, Berk Tınaz, Kader Oguz, Tolga Çukur

***Abstract*— Learning-based translation between MRI contrasts involves supervised deep models trained using high-quality source- and target-contrast images derived from fully-sampled acquisitions, which might be difficult to collect under limitations on scan costs or time. To facilitate curation of training sets, here we introduce the first semi-supervised model for MRI contrast translation (ssGAN) that can be trained directly using undersampled k-space data. To enable semi-supervised learning on undersampled data, ssGAN introduces novel multi-coil losses in image, k-space, and adversarial domains. The multi-coil losses are selectively enforced on acquired k-space samples unlike traditional losses in single-coil synthesis models. Comprehensive experiments on retrospectively undersampled multi-contrast brain MRI datasets are provided. Our results demonstrate that ssGAN yields on par performance to a supervised model, while outperforming single-coil models trained on coil-combined magnitude images. It also outperforms cascaded reconstruction-synthesis models where a supervised synthesis model is trained following self-supervised reconstruction of undersampled data. Thus, ssGAN holds great promise to improve the feasibility of learning-based multi-contrast MRI synthesis.**

***Index Terms*— magnetic resonance imaging, image synthesis, semi-supervised, adversarial, undersampled**

I. Introduction

MRI is a clinical powerhouse in neuroimaging due to its noninvasiveness and excellent soft-tissue discrimination. Its unique ability to image the same anatomy under a diverse set of tissue contrasts empowers it to accumulate complementary diagnostic information on a single scanner [1], [2]. However, economic and time costs often limit the number of distinct contrasts that can be captured within an MRI exam [3], [4]. A promising solution is to synthesize missing images (i.e., target-contrast images) within the protocol via translation from available images (i.e., source-contrast images) [5]. Through imputation of target images, multi-contrast MRI synthesis can enhance radiological assessments as well as image analysis tasks such as registration, segmentation, or detection [6]–[8]. From a scan-time perspective, synthesis of an unacquired target image from an acquired source image, and reconstruction from mutually undersampled acquisitions of source-target contrasts could yield similar outcomes [9]. Yet, synthesis offers several key benefits. While reconstruction may not be able to suppress artifacts that corrupt target-contrast acquisitions (e.g., motion artifacts), successful synthesis can be performed given a high-quality source image [4]. Reconstructions might be unreliable for target-contrast acquisitions with intrinsically low signal-to-noise ratio (SNR), yet successful synthesis can be performed given a high-SNR source image [10]. Finally, the reconstruction framework requires prolonged scan times as the number of distinct contrasts grows, albeit synthesis of multiple target contrasts can be performed efficiently based on a single source contrast [5].

In recent years, there has been emerging interest in learning-based MRI synthesis based on deep neural networks, given their state-of-the-art performance in other computer vision [11], [12] and medical imaging tasks [13]–[17]. An earlier group of studies proposed convolutional neural networks (CNNs) to learn nonlinear latent representations that mediate conversion from source to target images [18]–[22]. These studies typically involved encoder-decoder architectures, where the encoder embeds hierarchical image features onto a latent space that is later used by the decoder to recover the target image. For improved capture of structural details, a second group has proposed deep architectures based on conditional generative adversarial networks (GAN) [23]–[34], where the generator that performs the source-to-target mapping benefits from the game-theoretic interplay with the discriminator [12]. Pioneering studies have exploited pixel- or feature-wise correspondence between source-target images in an adversarial setup [23], [31], [32], [35]. Later studies have proposed unified models capable of multiple types of contrast translation [24], [27], [28], [36], or multi-tasking frameworks [24], [27], [33] to reduce computational complexity. These previous studies have collectively highlighted the immense potential of learning-based synthesis in multi-contrast MRI.

That said, existing MRI synthesis models are trained to translate between coil-combined images pertaining to source and target contrasts [18], [23]. These training data are canonically collected in a matching set of subjects, and derived from Nyquist-sampled acquisitions [18], [23]. Resultant models leverage supervision regarding pairing of source-target images across subjects, and regarding the use of ground-truth images obtained from fully-sampled k-space data [23], [24], [35]. However, compilation of training datasets with paired and

This study was supported in part by a TUBITAK 1001 Research Grant (118E256), a TUBA GEBIP 2015 fellowship, and a BAGEP 2017 fellowship (Corresponding author: Tolga Çukur).

M. Yurt, O. Dalmaz, S. Dar, M. Ozbey, and T. Çukur are with the Department of Electrical and Electronics Engineering, Bilkent University, Ankara, Turkey (e-mails: {mahmut, onat, salman, muzaffer, cukur}@ee.bilkent.edu.tr).

B. Tınaz is with the Department of Electrical and Computer Engineering, University of Southern California, Los Angeles, California, USA (email: tinaz@usc.edu).

K. Oguz is with the Department of Radiology, Hacettepe University, Ankara, Turkey (email: kkarlioguz@gmail.com).

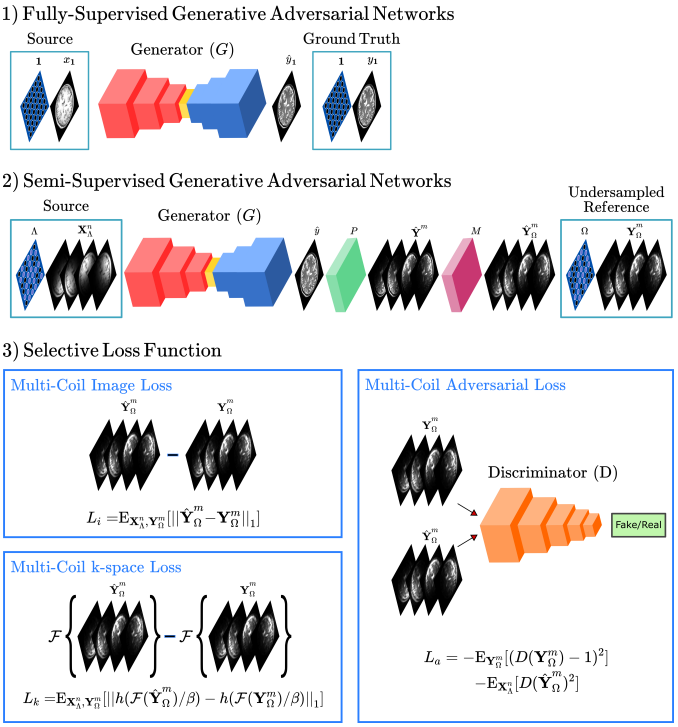

Fig. 1: Illustration of the proposed ssGAN model. 1) Fully-supervised synthesis models demand a training dataset of high-quality images derived from Nyquist-sampled acquisitions of the source and target contrasts. 2) Instead, ssGAN uses a training dataset containing only undersampled acquisitions of the source and target contrasts. To enable training on undersampled data, ssGAN synthesizes a coil-combined target image that is backprojected onto individual coils via an operator P. These multi-coil images are subjected to the sampling mask (Ω) of the reference target-contrast acquisition via operator M. 3) Afterwards, selective multi-coil losses are defined between undersampled synthetic and reference target images in image, k-space and adversarial domains.

fully-sampled acquisitions might prove impractical due to limitations on scan time or cost [3], [4]. As such, there is a dire need to lower reliance on these supervision factors in training of synthesis models to improve practicality.

Recent efforts to lower supervision requirements in MRI synthesis have primarily focused on model training with unpaired images across subjects. A successful approach has been to replace pixel-wise losses in GAN models with cycle-consistency, shape-consistency or mutual information losses [23], [27], [37]–[40]. Similar to supervised models, unpaired models that unify multiple contrast translation tasks have also been introduced to reduce computational complexity [38], [41]. As an alternative, [42], [43] have proposed a hybrid method where the model is trained on a composite dataset with both paired and unpaired samples. These previous methods have allowed synthesis models to learn from unpaired data, but they still leverage explicit supervision on ground-truth images derived from fully-sampled acquisitions. While unsupervised training of MRI reconstruction models on undersampled data has received recent interest [44]–[48], to the best of our knowledge, no prior study has considered training of MRI synthesis models using undersampled k-space data.

To avoid reliance on ground-truth images from fully-sampled acquisitions, here we propose a novel semi-supervised generative model for source-to-target contrast translation (Fig. 1). The proposed model, ssGAN, is learned using a training dataset of undersampled source- and target-contrast acquisitions from accelerated MRI scans. Unlike prior methods that use single-coil loss terms on coil-combined images, ssGAN enables learning of MRI synthesis based on undersampled data by introducing novel multi-coil losses in image, k-space, and adversarial domains. These losses are selectively enforced on acquired k-space samples in target-contrast acquisitions, and the sampling masks are randomized across training subjects to promote homogeneous learning across k-space.

Comprehensive experiments on retrospectively undersampled brain MRI datasets clearly demonstrate that ssGAN achieves on par performance to a benchmark supervised model trained with ground truth derived from fully-sampled acquisitions. Meanwhile, ssGAN outperforms single-coil synthesis models trained on inverse Fourier transform of undersampled data, and cascaded reconstruction-synthesis models that involve a reconstruct step to recover images from undersampled acquisitions and a synthesis step trained on the reconstructed images. Our results suggest that ssGAN can facilitate curation of training datasets for learning-based synthesis by enabling the use of undersampled MRI acquisitions. Code for ssGAN is publicly available at https://github.com/icon-lab/ssGAN.

Contributions:

- To our knowledge, ssGAN is the first semi-supervised method for MRI contrast translation that learns from a training dataset of undersampled source- and target-contrast acquisitions.
- For training, ssGAN introduces novel multi-coil losses expressed only on acquired k-space samples of undersampled target acquisitions.
- During inference, ssGAN synthesizes target images directly from undersampled multi-coil source acquisitions without the need for intermediate reconstruction.

II. Methods

In this section, we first overview basics of generative adversarial networks, and the foundation of the proposed architecture for semi-supervised multi-contrast MRI synthesis. We then describe in detail the datasets and experiments conducted to evaluate the proposed methodology.

A. Generative Adversarial Networks

Generative adversarial networks (GANs) [12] are deep generative models comprising a pair of competing subnetworks: a generator (G) and a discriminator (D). G aims to map a random noise vector z to a sample resembling a target-domain distribution, whereas D aims to distinguish between real and fake samples from the target domain. These two subnetworks are alternately trained via an adversarial loss function, formulated as follows:

$$L_{GAN} = \mathrm{E}_y[\log D(y)] + \mathrm{E}_z[\log(1 - D(G(z)))] \quad (1)$$

where E denotes expectation, and y is an arbitrary real sample

in the target domain. A squared loss can be adopted in place of log-likelihood loss to stabilize the training process [49]:

$$L_{GAN} = -\mathrm{E}_y[(D(y)-1)^2] - E_z[D(G(z))^2] \quad (2)$$

Upon convergence, G is expected to generate realistic target-domain samples that D cannot tell apart from the actual ones. While the initial GAN models generated target samples from a random noise vector, later studies have demonstrated success in image-to-image translation with conditional GAN (cGAN) models. These models receive as input a separate source-domain image x to capture conditional dependencies between the source and target domains [12]. The adversarial loss function is then modified by conditioning G on x:

$$L_{cGAN} = -\mathrm{E}_{x,y}[(D(y)-1)^2] - \mathrm{E}_x[D(G(x))^2] \quad (3)$$

As the contribution of latent variables is relatively limited in cGANs [12], z is typically removed from the formulation. When spatially aligned source-target images are available, a pixel-wise loss can be further included [12]:

$$\begin{aligned} L_{cGAN} = &- \mathrm{E}_{x,y}[(D(y)-1)^2] - E_x[D(G(x))^2] \\ &+ \mathrm{E}_{x,y}[||y - G(x)||_1] \end{aligned} \quad (4)$$

Several studies have demonstrated cGANs on multi-contrast MRI that synthesize target-contrast images from source-contrast images of the same underlying anatomy [23]–[34]. These models typically learn the source-to-target mapping in a fully-supervised setup. A comprehensive training set is needed containing high-quality source and target images derived from fully-sampled k-space acquisitions ($x_\mathbf{1}$, $y_\mathbf{1}$), where $x_\mathbf{1}$ is an arbitrary source, $y_\mathbf{1}$ is an arbitrary target image in the training set, and $\mathbf{1}$ denotes the sampling mask for Nyquist-sampled acquisitions. Such supervised models have demonstrated state-of-the-art performance for synthetic multi-contrast MRI. However, they rely on the availability of a training dataset of images obtained from fully-sampled acquisitions that might prove impractical to collect. Therefore, methods that can directly learn from undersampled k-space data are direly needed.

B. Semi-Supervised Generative Adversarial Networks

The proposed semi-supervised model (ssGAN) mitigates reliance on training datasets composed of fully-sampled MRI acquisitions (Fig. 1). Instead, ssGAN is trained using undersampled acquisitions of source and target contrasts (Fig. 2). To do this, ssGAN introduces multi-coil loss functions enforced selectively on the acquired k-space samples of the target-contrast acquisitions. During inference, ssGAN synthesizes target-contrast images given as input only multi-coil data from undersampled source-contrast acquisitions (Fig. 2). The optimization objectives are detailed in this section.

Receiving as input multi-coil images computed as the inverse Fourier transform of the undersampled source-contrast acquisition, ssGAN learns to estimate the corresponding coil-combined image of the target contrast. To do this, the generator G in ssGAN implements a forward mapping:

$$G(\mathbf{X}_\Lambda^n) = \hat{y}, \quad \text{with} \quad \mathbf{X}_\Lambda^n = \{x_\Lambda^1, \ldots, x_\Lambda^n\} \quad (5)$$

where $\mathbf{X}_\Lambda^n$ denotes multi-coil source-contrast images acquired with a k-space sampling mask Λ, n denotes the number of receive coils with sensitivity maps $\hat{\mathbf{C}}_\mathbf{X}^n$ computed via ESPIRiT [50], and $\hat{y}$ denotes the coil-combined target-contrast image. Since supervision based on fully-sampled data is excluded, no ground truth for the target-contrast image is available. As reference for the network output, ssGAN instead leverages multi-coil images computed as the inverse Fourier transform of the undersampled target-contrast acquisition: $\mathbf{Y}_\Omega^m = \{y_\Omega^1, \ldots, y_\Omega^m\}$ collected with a sampling mask Ω and m receive coils of coil sensitivities $\mathbf{C}_\mathbf{Y}^m$. These reference images are corrupted with aliasing artifacts, so canonical single-coil loss terms on coil-combined images cannot be employed. Instead, ssGAN uses multi-coil loss functions expressed selectively on the acquired subset of k-space samples. To do this, the synthesized coil-combined image is first projected onto individual coils:

$$\hat{\mathbf{Y}}^m = P(\hat{y}, \hat{\mathbf{C}}_\mathbf{Y}^m) = \hat{y} \cdot \hat{\mathbf{C}}_\mathbf{Y}^m \quad (6)$$

where $\hat{\mathbf{Y}}^m$ denotes the synthesized multi-coil target images, $\hat{\mathbf{C}}_\mathbf{Y}^m$ denotes coil sensitivity maps estimated via ESPIRiT [50], and P performs the coil projection in the image domain as element-wise multiplication between the input image and coil sensitivity maps. The multi-coil target image projections are then subjected to the binary sampling mask in Fourier domain:

$$\begin{aligned} \hat{k}_\Omega^m &= M(\mathcal{F}(\hat{\mathbf{Y}}^m), \Omega) = \mathcal{F}(\hat{\mathbf{Y}}^m) \cdot \Omega \\ \hat{\mathbf{Y}}_\Omega^m &= \mathcal{F}^{-1}(\hat{k}_\Omega^m) \end{aligned} \quad (7)$$

where $\mathcal{F}$ denotes the forward and $\mathcal{F}^{-1}$ denotes the inverse Fourier transform, M is the operator that performs binary masking in k-space with a given sampling mask. In Eq. (7), $\hat{k}_\Omega^m$ and $\hat{Y}_\Omega^m$ denote undersampled multi-coil data respectively in k-space and image domain for the synthesized target-contrast image. Based on undersampled multi-coil data for the synthesized versus reference target-contrast images, the overall loss function containing three components is calculated: multi-coil image, k-space and adversarial losses.

1) Multi-Coil Image Loss: The first component is an image-domain loss between synthesized versus reference multi-coil images of the target contrast. The image loss is based on the L_1-norm to reduce sensitivity to outliers:

$$L_i = \mathrm{E}_{\mathbf{X}_\Lambda^n, \mathbf{Y}_\Omega^m}[||\hat{\mathbf{Y}}_\Omega^m - \mathbf{Y}_\Omega^m||_1] \quad (8)$$

where $\mathbf{Y}_\Omega^m$ are reference images, and $\hat{\mathbf{Y}}_\Omega^m$ are synthesized images, both subjected to the same undersampling mask, Ω.

2) Multi-Coil k-space Loss: Because the training images for ssGAN are corrupted with aliasing artifacts, a Fourier-domain loss is used between synthesized and reference multi-coil

k-space data for the target contrast for improved performance.

$$L_k = \mathrm{E}_{\mathbf{X}_\Lambda^n, \mathbf{Y}_\Omega^m}[||h(\mathcal{F}(\hat{\mathbf{Y}}_\Omega^m)/\beta) - h(\mathcal{F}(\mathbf{Y}_\Omega^m)/\beta)||_1] \quad (9)$$

where $\mathcal{F}(\mathbf{Y}_\Omega^m)$ and $\mathcal{F}(\hat{\mathbf{Y}}_\Omega^m)$ denote undersampled k-space data for the reference and synthesized images, respectively. To provide comparable signal intensities across different spatial frequencies in k-space, both reference and synthesized k-space data are processed with h, a $tanh$ function with a shape parameter β controlling its slope [51]. The k-space loss is based on the L_1-norm to reduce sensitivity to outliers [44].

3) Multi-Coil Adversarial Loss: To improve the realism of synthesis, the third component is an adversarial loss based on synthesized and reference multi-coil target images:

$$L_a = -\mathrm{E}_{\mathbf{Y}_\Omega^m}[(D(\mathbf{Y}_\Omega^m) - 1)^2] - \mathrm{E}_{\mathbf{X}_\Lambda^n}[D(\hat{\mathbf{Y}}_\Omega^m)^2] \quad (10)$$

where D denotes the discriminator that distinguishes between undersampled images of the actual and synthetic target contrast. To compute Eq. 10, images for individual coils can be sequentially provided to the discriminator.

The overall loss for ssGAN is constructed as a weighted combination of the three components as $L_{ssGAN} = \lambda_k L_k + \lambda_i L_i + \lambda_a L_a$, where λ_k, λ_i, and λ_a denote the relative weighting of the k-space, image, and adversarial terms. All loss terms are expressed selectively on the subset of acquired k-space samples in the undersampled target-contrast acquisition. Using the same sampling mask across all training subjects would focus exclusively on a specific k-space subregion, thereby yield suboptimal learning. Instead, a different random sampling mask is prescribed for each subject in the training set. Note that within a given subject, the sampling mask is fixed across all cross-sections, following the common procedure for implementing accelerated MRI acquisitions [52].

C. Datasets

The IXI dataset (https://brain-development.org/ixi-dataset/) with single-coil magnitude brain images and an in-house dataset with multi-coil complex brain images were used.

1) The IXI Dataset: T_1- and T_2-weighted images from 94 subjects were used, with 64 reserved for training, 10 for validation, 20 for testing. Images were collected at axial orientation with matrix size = 256×256×130, spatial resolution = 0.94×0.94×1.2 mm^3. For T_1-weighted scans, a 3D MPRAGE sequence was used with TR = 9.81 ms, TE = 4.603 ms, flip angle = 8°, scan time = 5:26. For T_2-weighted scans, a 2D fast spin-echo (FSE) sequence was used with TR = 8178.34 ms, TE = 100 ms, flip angle = 90°, turbo factor of 16, scan time = 6:33. T_2-weighted images were spatially registered onto T_1-weighted images via FSL [53], using an affine transformation based on mutual information.

2) In vivo Brain Dataset: T_2- and PD-weighted images of 10 healthy volunteers were used, with 7 reserved for training, 1 for validation, 2 for testing. Images were collected using a 3D FSE sequence at coronal orientation with the following parameters: matrix size = 256×192×88, spatial resolution = 1×1×2 mm^3, flip angle = 90°, a turbo factor of 16. For T_2-weighted scans, TR = 1000 ms, TE = 118 ms, scan time = 17:39 were used. For PD-weighted scans, TR = 750 ms, TE = 12 ms, scan time = 13:14 were used.

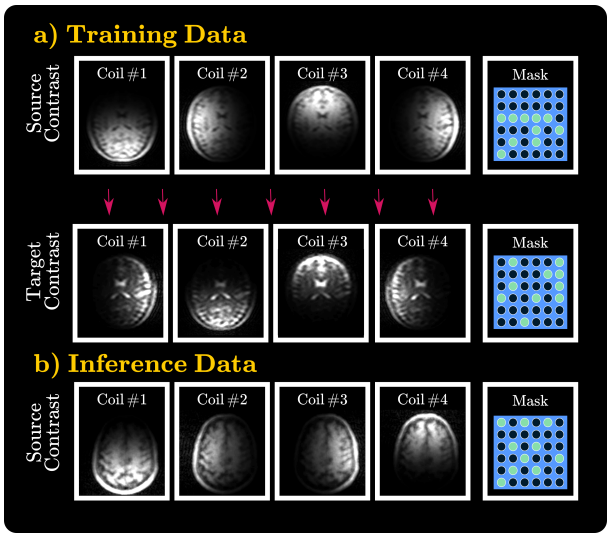

Fig. 2: a) The proposed ssGAN model is trained using undersampled multi-coil acquisitions, so no high-quality ground truth is available. The source and target contrasts are acquired with different sampling masks Λ and Ω, respectively. b) During inference in a test subject, ssGAN receives as input the undersampled acquisition of the source contrast, and it generates the coil-combined target-contrast image that is missing from the subject's MRI exam.

Data were collected on a 3T Siemens Magnetom scanner using a 32-channel receive-only head coil at Bilkent University, Ankara, Turkey. Imaging protocols were approved by the local ethics committee at Bilkent University, and all participants provided written informed consent. No spatial registration was performed as analyses in FSL suggested negligible interscan motion for multi-contrast images within subjects (less than 1 mm in translation and 0.6^o in rotation corresponding to the voxel size in the minimum dimension). To lower computational complexity, geometric coil compression was performed to reduce the number of coils from 32 to 5 [54].

3) Retrospective Undersampling: To obtain training datasets with undersampled source and target acquisitions, data from the IXI and in vivo datasets were retrospectively undersampled. Undersampling was performed across the two phase-encoding dimensions in three-dimensional k-space. Undersampling factors and sampling masks were independently selected for the source and target contrasts. For fair comparison, all competing methods were implemented based on the same sampling masks. For IXI, k-space data were obtained as Fourier transform of coil-combined magnitude images. Data in axial cross-sections were then randomly undersampled to achieve undersampling factors $R = [2 - 10]$. Uniform-density sampling was performed with a 10×10 fully-sampled calibration region. For the in vivo dataset, k-space data for coronal cross-sections were randomly undersampled to achieve $R = [2-4]$. Uniform-density sampling was used with a 16×16 calibration region to permit coil-sensitivity estimation [50]. A different random sampling mask was generated for each

contrast within each subject. The mask was kept identical across separate cross-sections of a given contrast, and was kept fixed for a given subject across the modeling procedures. These sampling procedures accurately emulate the way undersampled acquisitions would be performed in practice, as pervasively utilized in MRI studies [10], [52]. Note that the fully-sampled acquisitions in IXI or in vivo datasets were not used for training ssGAN, instead fully-sampled data were only utilized to measure model performance on test subjects.

D. Implementation Details for ssGAN

The generator and discriminator in ssGAN were adopted from a previous study on multi-contrast MRI synthesis [23]. The generator contained an encoder of 3 convolutional layers, a residual network of 9 ResNet blocks, and a decoder of 3 convolutional layers. The discriminator contained 5 convolutional layers. An unlearned coil-combination block was placed at the input of the generator [55], which recovered real and imaginary parts of the coil-combined target image given real and imaginary parts of the coil-combined source image. The coil-combined target image was backprojected onto individual coils, and the complex target images from each coil were sequentially fed to a patch discriminator.

Cross-validation was used to select the relative weighting of the selective loss components in image, k-space and adversarial domains, as well as the slope of the tanh function in the k-space loss (λ_i, λ_k, λ_a, β). This selection aimed to minimize the network loss L_{ssGAN} on the validation set. A common set of parameters (λ_i=100, λ_k=3000, λ_a=1, β=5000) observed to yield near-optimal performance consistently across datasets were used in all experiments, including ssGAN models trained for varying undersampling factors for the source and target contrasts (R_S and R_T respectively).

Ablation studies were conducted to demonstrate the influence of the main building blocks in ssGAN to synthesis performance. For this purpose, variant ssGAN models were trained by removing particular loss terms from L_{ssGAN}, by replacing multi-coil loss terms with single-coil versions following compression of multi-coil data onto a single coil [54], and by prescribing a fixed sampling mask across training subjects. All variant models were trained with the same loss function and hyperparameters as in ssGAN, except for variants with ablated losses that omitted specific loss components.

E. Competing Methods

We compared ssGAN against a benchmark fully-supervised model, state-of-the-art single-coil synthesis models, and cascaded reconstruction-synthesis models.

1) Benchmark Supervised Model: A supervised model was trained using fully-sampled multi-coil acquisitions of the source and target contrasts. This supervised model serves as an upper performance limit for ssGAN.

fsGAN: A GAN model was implemented with matching architecture, loss terms and hyperparameters to ssGAN, albeit each loss term was non-selectively expressed on the entire k-space to leverage information on fully-sampled acquisitions.

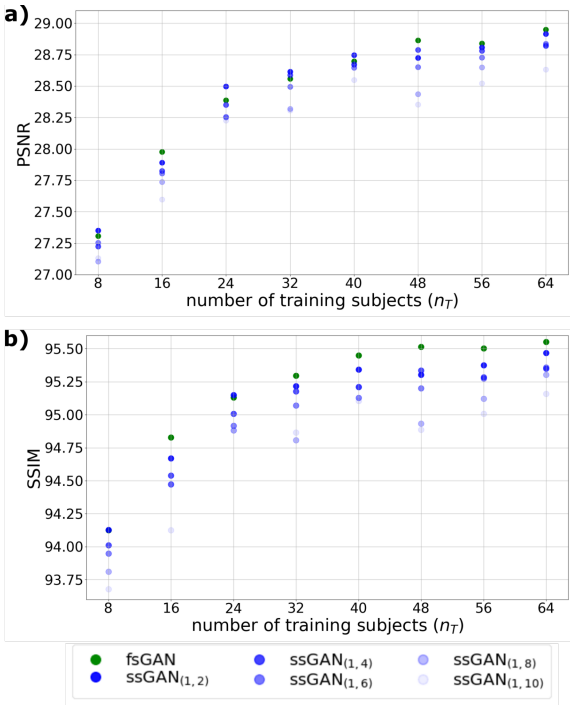

Fig. 3: Reliability against the undersampling factor of target-contrast acquisitions in the training set ($R_T \in [2:2:10]$) and the number of training subjects ($n_T = [8:8:64]$). The supervised fsGAN model was trained with fully-sampled source and target acquisitions ($R_S = 1$, $R_T = 1$); ssGAN was trained with $R_S = 1$ and varying R_T. PSNR and SSIM are reported in the test set.

2) Single-coil Synthesis Models: We considered two state-of-the-art single-coil synthesis models designed to operate on coil-combined magnitude images. To learn these models using the same training set as ssGAN, training images were formed via inverse Fourier transform of undersampled acquisitions followed by coil combination [50]. Models were trained to translate between the resultant coil-combined magnitude images of source and target contrasts.

pix2pix: A GAN model was implemented with the architecture, loss terms and hyperparameters adopted from [12].

pGAN: A GAN model was implemented with the architecture, loss terms and hyperparameters adopted from [23].

3) Cascaded Reconstruction-Synthesis Models: Given a training set of undersampled data, a cascaded reconstruction-synthesis approach is an alternative to ssGAN. In the absence of fully-sampled ground truth, an unsupervised model must be used to reconstruct undersampled data from the source and target contrasts. A supervised synthesis model can then be trained on the reconstructed coil-combined source and target images. Five cascaded models were considered based on different reconstruction methods. In all cases, synthesis was implemented based on the pGAN method, with the loss terms and hyperparameters adopted from [23].

CasCS: For single-coil data in IXI, reconstruction was implemented using SparseMRI [10]. Hyperparameters were selected via cross-validation: 4 iterations, total variation weight of 10^{-4}, and wavelet-domain L_1 weight of 10^{-1}.

CasSPIRiT: For multi-coil data from the in vivo dataset, reconstruction was implemented using L_1-SPIRiT [52]. Hyper-

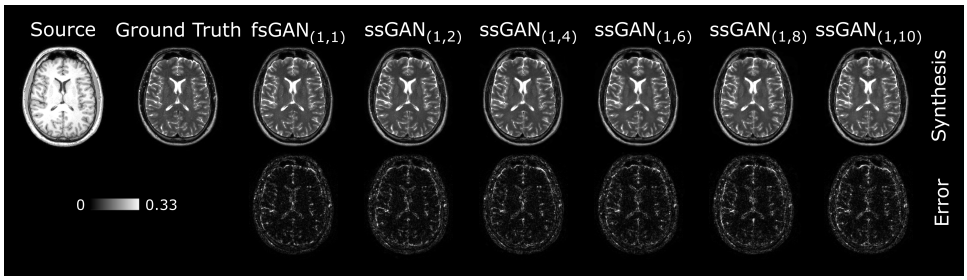

Fig. 4: Independent ssGAN$_{(R_S,R_T)}$ models with $R_S = 1$, $R_T = [2 - 10]$ were demonstrated on IXI for $T_1 \rightarrow T_2$ mapping against fsGAN$_{(R_S,R_T)}$ with $R_S = 1$, $R_T = 1$. Synthesized images are displayed along with error images underneath (see colorbar). The source image and the ground-truth target image derived from fully-sampled acquisitions are also shown.

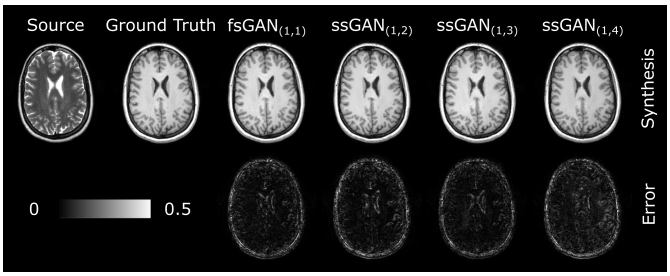

Fig. 5: ssGAN$_{(R_S,R_T)}$ models with $R_S = 1$, $R_T = [2 : 1 : 4]$ were demonstrated on IXI for $T_2 \rightarrow T_1$ mapping against fsGAN with $R_S = 1$, $R_T = 1$. Synthesized images are displayed along with error images. The source image and the ground-truth target image derived from fully-sampled acquisitions are also shown.

parameters were selected via cross-validation: a 5×5 kernel, wavelet-domain L_1 weight of 10^{-1}, Tikhonov weight of 10^{-3}, 10 iterations for PD and 20 iterations for T_2 images.

CasRAKI: Reconstruction was implemented using RAKI with the architecture and loss terms adopted from [56]. Hyperparameters reported in [56] were selected that also yielded near-optimal performance during cross-validation.

CasUnsup: Reconstruction was implemented using an unsupervised GAN method with the architecture and loss terms adopted from [45]. Hyperparameters reported in [45] were selected that also yielded near-optimal performance during cross-validation.

CasSSDU: Reconstruction was implemented using a self-supervised approach, SSDU [44]. As SSDU is a model-agnostic training strategy, the network architecture and loss terms were adopted from ssGAN for fair comparison. Undersampled k-space data were split into two nonoverlapping subsets, where 40% were used for defining losses, and 60% were used to estimate model weights as proposed in [44]. The relative weighting of image, k-space, and adversarial losses was [100, 3000, 1] as determined via cross-validation.

F. Modeling Procedures

All models were trained for 100 epochs using the Adam optimizer with parameters β_1=0.5 and β_2=0.999. The learning rate was set to 0.0002 in the first 50 epochs and linearly decayed to 0 in the last 50 epochs. Expectation values of the loss components were estimated via Monte Carlo sampling over images sampled from the training set. Models were run on Nvidia 2080 Ti GPUs in Python2.7 using PyTorch. Training ssGAN on 32 subjects with 100 cross-sections required nearly 8 hrs. Meanwhile, it required nearly 8 sec of inference for a single test subject with 100 cross-sections.

TABLE I: Image quality in IXI for $R_S = 1$, $R_T = [1 - 4]$

	$T_1 \rightarrow T_2$		$T_2 \rightarrow T_1$	
	PSNR	SSIM	PSNR	SSIM
fsGAN$_{(1,1)}$	28.57±1.39	95.33±1.35	28.62±1.45	95.95±1.38
ssGAN$_{(1,2)}$	28.63±1.42	95.25±1.33	28.52±1.61	95.79±1.46
ssGAN$_{(1,3)}$	28.56±1.39	95.24±1.32	28.44±1.55	95.72±1.41
ssGAN$_{(1,4)}$	28.60±1.43	95.21±1.34	28.42±1.56	95.70±1.42

Performance was evaluated using peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) on coil-combined magnitude images derived from synthesized and ground-truth target-contrast images. The ground-truth image was based on inverse Fourier transform of the fully-sampled target acquisition. In Tables, summary statistics of quantitative metrics were provided as mean ± std across test subjects, except for ablation analyses where metrics were provided as mean. Significance of PSNR, SSIM, and radiological opinion scores was assessed via Kruskal Wallis H-test ($p < 0.05$) to collectively compare ssGAN models with fsGAN, and via Wilcoxon signed-rank test ($p < 0.05$) to individually compare ssGAN against competing methods trained under matched $R_{S,T}$.

III. Results

A. Robustness against training set deficiencies

We first examined the reliability of ssGAN against the degree of undersampling and the number of subjects in the training dataset. Models were learned for $T_1 \rightarrow T_2$ mapping in the IXI dataset (network input: T_1, network output: T_2), while the undersampling factor of target-contrast acquisitions in the training set ranged in $R_T = [2 : 2 : 10]$ and the number of training subjects ranged in $n_T = [8 : 8 : 64]$. As a benchmark, the fully-supervised fsGAN model was also trained for matching n_T but with ground-truth images derived from fully-sampled target-contrast acquisitions $R_T = 1$. All models received as input source-contrast images derived from fully-sampled acquisitions ($R_S = 1$). Note that all models were trained on the same selected set of subjects for a given n_T, and they were tested on the same set of test subjects that did not overlap with training subjects. Synthesis

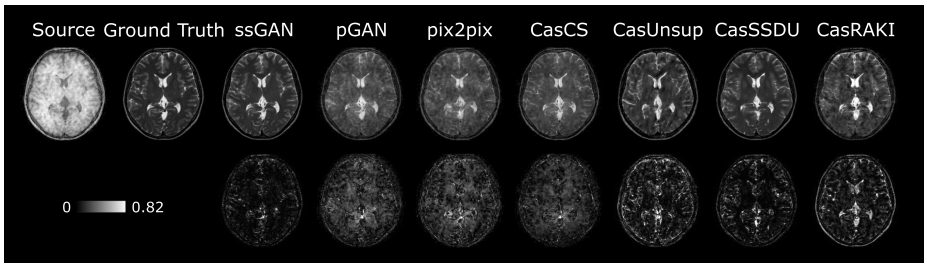

Fig. 6: ssGAN was demonstrated on the IXI dataset for the $T_1 \rightarrow T_2$ synthesis task. Synthesized images from all competing methods trained at (R_S, R_T)=(2, 2) are displayed along with error images underneath (see colorbar). The undersampled source image and the ground-truth target image are also shown.

performance is displayed in Fig. 3 as a function of R_T and n_T. Synthesis quality of ssGAN is on par with the supervised fsGAN model ($p > 0.05$), where ssGAN performance is within [-0.51,0.13]dB PSNR, [-0.70,0.02]% SSIM of fsGAN. Synthesized images displayed in Fig. 4 further indicate that ssGAN models trained with varying R_T yield highly similar performance, and near-optimal synthesis quality can be maintained for a broad range of undersampling rates for target-contrast acquisitions in the training set.

Synthesis performance naturally improves towards larger n_T for both fsGAN and ssGAN as depicted in Fig. 3. Comparing $n_T = 16$ against $n_T = 32$, similar average improvements in (PSNR, SSIM) are observed with (0.68dB,0.57%) for ssGAN, and (0.58dB,0.47%) for fsGAN. Yet, because ssGAN permits training with undersampled acquisitions (i.e., $R_T > 1$), a larger group of training subjects can be recruited under the same scan budget compared to fsGAN that requires $R_T = 1$. For instance, T_1- and T_2-weighted images required to train fsGAN can be collected from 16 subjects in a total scan time of 192 min. In the same time, training data for ssGAN can be captured from 32 subjects with this two-contrast protocol at $R_T = 10$. Therefore, ssGAN facilitates curation of training sets to improve practicality of learning-based MRI synthesis.

We then extended comparison of ssGAN against fsGAN to the $T_2 \rightarrow T_1$ task in the IXI dataset with a fixed number of training subjects $n_T = 32$ (used hereafter in all evaluations in IXI). Measurements of synthesis quality are reported in Table I. The reported measurements reveal that ssGAN models trained with $R_T = [2 : 1 : 4]$ maintain near-optimal synthesis quality on par with fsGAN. Representative results displayed in Fig. 5 corroborate the quantitative findings by showing that ssGAN offers a similar level of accuracy in tissue depiction to the benchmark supervised model, fsGAN.

B. Single-coil MRI Synthesis

Next, we comparatively demonstrated ssGAN for $T_1 \rightarrow T_2$ and $T_2 \rightarrow T_1$ on the single-coil IXI dataset. Competing methods included single-coil synthesis models trained on magnitude images (pix2pix, pGAN), and cascaded reconstruction-synthesis models (CasCS, CasUnsup, CasSSDU, CasRAKI). For all models, the training dataset comprised undersampled acquisitions of both source and target contrasts. Models were built for $R_S = \{2, 4\}$, $R_T = \{2, 4\}$. Synthesis performance is reported in Table II for $R_S = 2$, $R_T = \{2, 4\}$ and in Table III for $R_S = 4$, $R_T = \{2, 4\}$. ssGAN outperforms all competing models consistently across tasks ($p < 0.05$). On average, ssGAN improves PSNR by 7.49dB, SSIM by 14.49% over single-coil models, and PSNR by 4.86dB, SSIM by 10.09% over cascaded models. Note that incremental steps from $R_T = 2$ to $R_T = 4$ result in modest performance losses of 0.16dB PSNR, 0.34% SSIM for

TABLE II: Image quality in IXI for $R_S = 2$, $R_T = \{2, 4\}$

	$R_S = 2, R_T = 2$				$R_S = 2, R_T = 4$			
	$T_1 \rightarrow T_2$		$T_2 \rightarrow T_1$		$T_1 \rightarrow T_2$		$T_2 \rightarrow T_1$	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
ssGAN	**26.92** ±**1.37**	**93.48** ±**1.54**	**27.47** ±**1.35**	**94.16** ±**1.51**	**26.81** ±**1.34**	**93.22** ±**1.50**	**27.30** ±**1.38**	**93.92** ±**1.55**
pGAN	22.80 ±0.70	82.34 ±2.03	19.38 ±0.85	82.77 ±2.29	18.83 ±0.56	75.83 ±2.26	16.31 ±0.66	77.31 ±2.71
pix2pix	22.24 ±0.79	80.39 ±2.39	18.86 ±0.99	81.41 ±2.55	18.78 ±0.63	74.83 ±2.58	16.08 ±0.66	76.52 ±2.76
CasCS	24.04 ±0.81	85.22 ±2.12	21.25 ±1.05	85.56 ±2.31	20.06 ±0.6	79.03 ±2.35	16.81 ±0.69	79.29 ±2.68
CasSSDU	25.04 ±1.34	91.00 ±1.73	25.79 ±1.36	91.80 ±1.69	24.83 ±1.28	90.52 ±1.69	25.13 ±1.36	90.33 ±1.85
CasUnsup	22.55 ±1.25	84.67 ±2.73	23.34 ±0.93	86.05 ±2.37	22.86 ±1.36	84.84 ±2.77	22.04 ±0.59	82.81 ±2.40
CasRAKI	22.51 ±2.05	83.87 ±3.13	16.72 ±0.49	68.32 ±2.80	21.62 ±1.84	80.9 ±3.15	16.81 ±0.45	69.16 ±2.71

TABLE III: Image quality in IXI for $R_S = 4$, $R_T = \{2, 4\}$

	$R_S = 4, R_T = 2$				$R_S = 4, R_T = 4$			
	$T_1 \rightarrow T_2$		$T_2 \rightarrow T_1$		$T_1 \rightarrow T_2$		$T_2 \rightarrow T_1$	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
ssGAN	**25.49** ±**1.39**	**91.39** ±**1.73**	**26.64** ±**1.29**	**92.84** ±**1.62**	**25.28** ±**1.43**	**90.90** ±**1.83**	**26.51** ±**1.28**	**92.46** ±**1.68**
pGAN	22.32 ±0.76	80.55 ±2.25	19.06 ±0.63	81.74 ±2.28	18.93 ±0.65	75.07 ±2.47	16.23 ±0.58	76.73 ±2.78
pix2pix	21.66 ±0.84	78.25 ±2.76	18.89 ±0.87	80.09 ±2.55	18.54 ±0.66	73.42 ±2.78	16.07 ±0.66	75.72 ±2.77
CasCS	23.49 ±0.90	83.70 ±2.31	21.04 ±1.19	84.42 ±2.39	19.85 ±0.66	77.72 ±2.52	16.87 ±0.68	78.55 ±2.73
CasSSDU	23.38 ±1.30	88.39 ±2.03	25.15 ±1.32	90.58 ±1.87	23.37 ±1.28	88.20 ±2.03	24.72 ±1.32	89.45 ±2.01
CasUnsup	21.63 ±1.18	82.57 ±2.89	23.55 ±0.91	86.28 ±2.38	21.80 ±1.29	82.39 ±3.07	22.10 ±0.56	82.82 ±2.33
CasRAKI	21.47 ±1.92	81.57 ±3.17	16.70 ±0.53	68.21 ±2.82	20.90 ±1.77	79.26 ±3.13	16.82 ±0.46	69.07 ±2.64

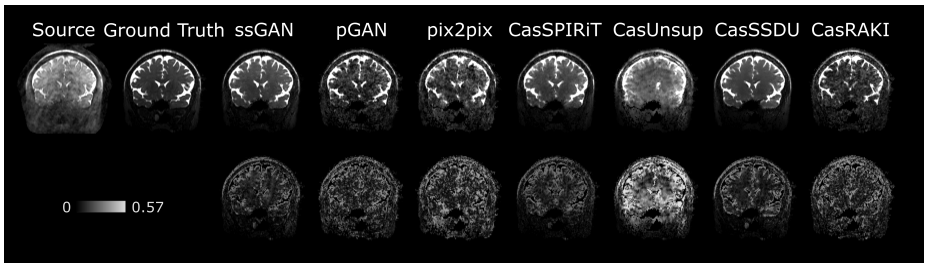

Fig. 7: ssGAN was demonstrated on the in vivo dataset for the PD $\rightarrow$ T_2 synthesis task. Synthesized images from all competing methods trained at (R_S, R_T)=(2, 2) are displayed along with error images underneath (see colorbar). The undersampled source image and the ground-truth target images are also shown.

ssGAN. In contrast, performance losses are 3.18dB PSNR, 5.26% SSIM for single-coil models and 1.32dB PSNR, 2.37% SSIM for cascaded models. Representative synthesis results are shown in Fig. 6. Single-coil models generate images with native aliasing artifacts; and cascaded models suffer from residual errors from the reconstruction step. Meanwhile, ssGAN yields superior synthesis quality with the sharpest tissue depiction and highest artifact suppression.

C. Multi-coil MRI Synthesis

We also conducted experiments to demonstrate ssGAN for $T_2 \rightarrow$ PD and PD $\rightarrow T_2$ tasks on the multi-coil in vivo dataset. Competing methods included single-coil synthesis models and cascaded reconstruction-synthesis models. For all methods, the training dataset comprised undersampled acquisitions of source and target contrasts ($R_S = \{2, 4\}$, $R_T = \{2, 4\}$). Performance is reported in Table IV for $R_S = 2$ and $R_T = \{2, 4\}$, and in Table V for $R_S = 4$ and $R_T = \{2, 4\}$. Overall, ssGAN achieves the highest performance among competing methods in all examined tasks ($p < 0.05$), except for PSNR against CasSPIRiT PD $\rightarrow T_2$ with ($R_S = 4, R_T = 2$). On average, ssGAN improves PSNR by 2.30dB, SSIM by 17.71% over single-coil models, and PSNR by 2.70dB, SSIM by 12.30% over cascaded models. Representative results displayed in Fig. 7 corroborate the superior synthesis quality of ssGAN over single-coil models that suffer from aliasing artifacts, and cascaded models that suffer from residual reconstruction errors. Note that increasing the target-contrast undersampling factor from $R_T = 2$ to 4 results in a 0.30dB PSNR, 0.96% SSIM performance loss for ssGAN, and instead 1.72dB PSNR, 9.52% SSIM loss for single-coil models, and 0.75dB PSNR, 4.60% SSIM loss for cascaded models.

D. Radiological Evaluations

Radiological evaluations were performed to assess synthesized images in terms of their general visual similarity to the ground-truth target image derived from fully-sampled acquisitions (Fig. 8). On a 5-point Likert scale, we examined the comparative performance of ssGAN against three top-contending competing methods, pGAN, CasSPIRiT, CasSSDU. This evaluation was conducted for PD $\rightarrow T_2$ and $T_2 \rightarrow$ PD tasks on the in vivo dataset, where independent models were trained at (R_S, R_T)=(2,2) and (2,4). ssGAN outperforms all competing methods in opinion scores ($p < 0.05$). On average, the proposed method achieves a 3.72 opinion score across tasks, whereas the opinion score is 1.03 for pGAN, 1.94 for CasSSDU, and 2.28 for CasSPIRiT.

E. Ablation Studies

We first conducted ablation studies to examine the contribution of individual loss components to ssGAN's perfor-

TABLE IV: Image quality in the in vivo dataset for $R_S = 2$

	$R_S = 2, R_T = 2$				$R_S = 2, R_T = 4$			
	$T_2 \rightarrow$ PD		PD $\rightarrow T_2$		$T_2 \rightarrow$ PD		PD $\rightarrow T_2$	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
ssGAN	**25.75** **±0.01**	**88.14** **±1.83**	**24.59** **±0.21**	**86.56** **±0.52**	**25.47** **±0.15**	**87.72** **±1.92**	**24.46** **±0.20**	**85.97** **±0.30**
pGAN	23.46 ±0.04	73.92 ±2.13	24.50 ±1.20	79.18 ±4.29	20.41 ±0.01	63.01 ±1.67	23.47 ±1.33	69.10 ±5.01
pix2pix	22.19 ±0.29	69.38 ±2.12	24.08 ±0.95	75.97 ±3.56	20.72 ±0.06	61.26 ±1.80	22.44 ±0.80	65.59 ±4.44
CasSPIRiT	25.01 ±0.03	86.96 ±1.61	24.44 ±0.43	85.02 ±2.45	22.95 ±0.59	82.22 ±2.32	23.55 ±0.24	81.53 ±3.03
CasSSDU	23.22 ±0.1	84.69 ±1.28	23.74 ±0.14	84.97 ±0.38	23.61 ±0.01	85.79 ±1.03	23.69 ±0.13	85.25 ±0.32
CasUnsup	18.02 ±0.03	62.40 ±0.38	21.09 ±0.36	67.48 ±1.52	19.37 ±0.32	65.12 ±0.44	20.13 ±0.40	60.16 ±1.70
CasRAKI	21.55 ±0.41	70.53 ±0.86	23.93 ±0.78	77.16 ±2.99	19.74 ±0.16	57.56 ±1.76	21.77 ±0.16	62.92 ±1.36

TABLE V: Image quality in the in vivo dataset for $R_S = 4$

	$R_S = 4, R_T = 2$				$R_S = 4, R_T = 4$			
	$T_2 \rightarrow$ PD		PD $\rightarrow T_2$		$T_2 \rightarrow$ PD		PD $\rightarrow T_2$	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
ssGAN	**24.85** **±0.00**	**86.64** **±1.47**	24.70 ±0.35	**87.25** **±0.20**	**24.73** **±0.12**	**86.22** **±1.83**	**24.05** **±0.38**	**84.84** **±0.39**
pGAN	22.88 ±0.10	72.21 ±2.25	24.45 ±1.09	77.77 ±4.17	20.79 ±0.11	61.32 ±2.50	23.05 ±0.95	68.5 ±4.89
pix2pix	21.78 ±0.16	68.33 ±2.96	23.30 ±0.96	72.93 ±5.31	20.19 ±0.08	59.57 ±2.62	22.23 ±0.83	65.22 ±5.78
CasSPIRiT	23.72 ±0.48	84.78 ±2.24	**25.12** **±1.33**	85.91 ±3.81	22.31 ±0.23	80.32 ±2.50	23.51 ±1.17	81.68 ±3.79
CasSSDU	22.42 ±0.18	82.13 ±2.12	23.47 ±0.19	83.87 ±0.54	23.05 ±0.06	83.63 ±1.50	23.61 ±0.34	84.29 ±0.86
CasUnsup	17.94 ±0.07	61.51 ±0.27	21.01 ±0.52	65.88 ±3.03	19.34 ±0.17	63.82 ±0.88	19.90 ±0.86	58.95 ±3.31
CasRAKI	22.18 ±0.43	69.34 ±0.78	23.26 ±0.54	74.27 ±3.38	19.97 ±0.11	58.10 ±1.47	21.50 ±0.74	61.46 ±3.75

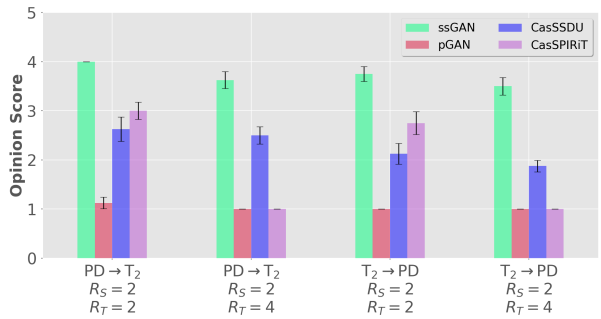

Fig. 8: Radiological opinion scores for ssGAN, pGAN, CasSPIRiT, and CasSSDU for the PD $\rightarrow$ T_2 and T_2 $\rightarrow$ PD tasks on the in vivo dataset, with independent models trained at (R_S, R_T)=(2,2) and (2,4) are displayed. Error bars denote standard error.

mance by forming three variants: ssGAN(w/o image) without image loss, ssGAN(w/o k-space) without k-space loss, and ssGAN(w/o adv) without adversarial loss. Models were trained for T_1 $\rightarrow$ T_2 and T_2 $\rightarrow$ T_1 in IXI. The effects of image and k-space losses were evaluated using PSNR and SSIM, whereas the effect of adversarial loss was assessed using Frechlet Inception Distance (FID) [57] and visual inspection as common in literature [58]. Performance metrics in Table VI indicate that image and k-space losses serve to improve synthesis quality in both tasks. Meanwhile, adversarial loss increases realism with decreased FID scores. Representative results in Fig. 9 corroborate the quantitative findings by showing that ssGAN(w/o image) and ssGAN(w/o k-space) suffer from residual artifacts and ssGAN(w/o adv) suffers from smoothing. In contrast, ssGAN with all loss components yields enhanced artifact suppression and visual acuity.

Next, we compared ssGAN against a variant model ssGAN(single) with traditional single-coil losses. Models were trained for T_2 $\rightarrow$ PD and PD $\rightarrow$ T_2 tasks on the multi-coil in vivo dataset. Performance metrics are listed for $R_S = \{2, 4\}$ and $R_T = \{2, 4\}$ in Table VII. ssGAN model consistently yields enhanced synthesis performance. On average, it improves PSNR by 1.46dB, SSIM by 6.32% over ssGAN(single). These results demonstrate the value of the multi-coil loss terms in improving quality of synthetic images.

Finally, we compared ssGAN against a variant model ssGAN(w/o rand) with a common sampling mask across training subjects. Models were trained for T_1 $\rightarrow$ T_2 and T_2 $\rightarrow$ T_1 tasks on the IXI dataset. Performance metrics for $R_S = 2$ and $R_T = [2 : 2 : 10]$ are listed in Table VIII. ssGAN yields consistently higher performance compared to the non-randomized variant, with 0.15dB PSNR, 0.25% SSIM improvement over ssGAN(w/o rand). These results indicate that leveraging randomized sampling masks across training subjects helps improve synthesis quality.

IV. Discussion

A novel semi-supervised deep generative model was introduced for image synthesis in multi-contrast MRI. Unlike supervised models, ssGAN learns to synthesize high-quality target-contrast images in the absence of training sets composed of costly Nyquist-sampled acquisitions. Unlike prior synthesis

TABLE VI: Effects of losses on synthesis performance

	$T_1 \rightarrow T_2$		$T_2 \rightarrow T_1$	
	PSNR	SSIM	PSNR	SSIM
ssGAN	**25.25**	**90.75**	**26.47**	**92.27**
ssGAN (w/o image)	25.12	90.36	26.12	91.78
ssGAN (w/o k-space)	24.96	90.30	26.41	92.19
	$T_1 \rightarrow T_2$		$T_2 \rightarrow T_1$	
	FID		FID	
ssGAN	**22.89**		**16.80**	
ssGAN (w/o adv)	23.13		24.35	

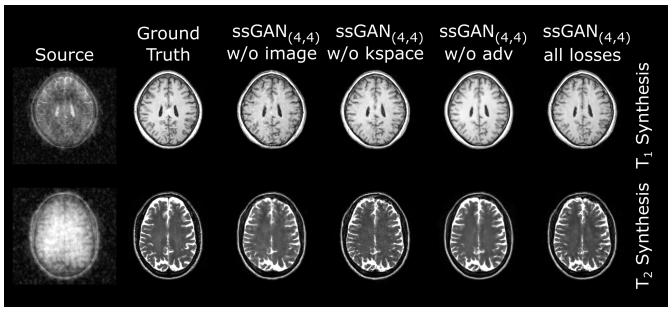

Fig. 9: Variant $ssGAN_{(4,4)}$ models with $R_S = 4$ and $R_T = 4$ were trained on IXI for T_2 $\rightarrow$ T_1 (upper row) and T_1 $\rightarrow$ T_2 (lower row) tasks. Synthesized images are shown for variant models with ablated image, k-space or adversarial loss along with ssGAN. The source image derived via inverse Fourier transform of the undersampled acquisition, and the ground-truth target image derived from a fully-sampled acquisition are also shown.

models trained on coil-combined magnitude images, ssGAN is trained on multi-coil complex MRI data from undersampled acquisitions with selective multi-coil losses. It holds promise in facilitating curation of large training sets required for deep-learning models, thereby advancing the practicality of multi-contrast MRI synthesis. Through imputation of missing target contrasts, ssGAN can aid in radiological assessment and image analysis tasks such as registration, segmentation or detection. Note that ssGAN can receive as input multi-coil data from undersampled source-contrast acquisitions during inference without an intermediate reconstruction step. Thus, a side benefit of ssGAN is to shorten the scan time for the source-contrast acquisition, albeit obtaining high-quality source images requires a separate reconstruction.

We find superior synthesis performance with ssGAN over cascaded models, best attributed to differences in learning strategy. ssGAN performs end-to-end mapping between undersampled source-target acquisitions to directly maximize synthesis quality. In contrast, cascaded models recover images from undersampled acquisitions to maximize reconstruction quality, and then map between reconstructed images for synthesis. This step-wise learning causes error propagation in cascaded models, where the synthesis step supervised by reconstructed images will learn to reproduce reconstruction errors. Note that reconstruction performance is primarily dominated by low-spatial-frequency information given k-space spectra of MR images [10], so elevated reconstruction errors are typically evident at high frequencies. This can undermine synthesis performance as contrast translation critically relies

on high-frequency information regarding shared tissue boundaries between contrasts [23].

Several lines of development can be pursued for ssGAN. Cycle-consistency loss was proposed to improve performance and enable unsupervised training of GAN-based MRI reconstruction [45], [59] and synthesis models [16], [23], [60]. Incorporation of cycle-consistency in ssGAN can also permit training on unpaired source-target contrast acquisitions. A GAN model proposed for retrospective motion correction of MRI scans employed a generator with cascaded U-Net backbones for improved performance [61]. Synthesis performance for ssGAN might also benefit from cascading multiple residual backbones in its generator. Third, many-to-one synthesis was recently demonstrated with a multi-stream model that adaptively fuses one-to-one mapping streams between each individual source and the target, and a many-to-one mapping stream between all sources and the target [35]. Unlike the semi-supervised ssGAN model trained on undersampled acquisitions with multi-coil loss functions, this prior method used a supervised model trained on fully-sampled acquisitions with single-coil loss functions. Yet, ssGAN might be generalized to perform many-to-one synthesis by adopting a multi-stream approach. Lastly, recent studies demonstrated improvements in image quality and resolution during joint reconstruction of multi-contrast MRI scans [9], [17], [62]. Accordingly, recovery of target-contrast images might be enhanced by modifying ssGAN to receive as input not only source-contrast but also undersampled target-contrast acquisitions, and adopting a self-supervised training procedure [44].

Here, we demonstrated all methods using uniform-density random sampling patterns that improve peripheral k-space coverage for a given undersampling factor. Our motivation was to aid recovery of high-frequency information during MRI synthesis. An alternative would be to collect low-resolution images by Nyquist-sampling in a central k-space region to achieve similar acceleration. A superresolution task on coil-combined images could then be performed [63]–[65]. Since superresolution is challenging when both source and target images are low resolution, external priors might be required to assist recovery. Another alternative would be to adopt variable-density sampling that trades-off high-frequency coverage for increased SNR [10]. While aggregate performance metrics (e.g., PSNR) that are dominated by low-frequency errors might be improved, variable-density masks would have relatively sparse coverage of peripheral k-space compared to uniform-density masks, potentially compromising recovery of detailed structure. The influence of sampling patterns on the relative performance of ssGAN versus cascaded models remains an important topic for future research.

Training data for ssGAN were obtained via retrospective undersampling of the IXI and in vivo datasets. When undersampled data are derived by masking fully-sampled data, a paired ground-truth image is available that allows reference-based performance assessments. That said, our analyses might be insensitive to certain imperfections in MRI acquisitions. First, IXI contained coil-combined magnitude images. While assessment of influences from coil-sensitivity encoding and image phase are precluded, analyses on IXI still demonstrate the benefits of selective expression of loss terms on acquired k-space coefficients. The in vivo dataset contained multi-coil complex images, albeit no significant phase perturbations or swaps were observed within brain tissue, likely due to adequate shimming. Under large field inhomogeneity, a field map can be collected or phase unwrapping can be used against perturbations. Reliability against phase-related effects remains an important topic for future research. Second, we estimated coil sensitivities via ESPIRiT that has been reported to work reliably with undersampled acquisitions [50]. ssGAN uses sensitivity estimates to combine images across coils, to backproject the combined image onto individual coils, and thereby to define its loss function. Thus, significant errors in sensitivity estimates might affect signal homogeneity in the combined image, elicit suboptimal weighting across coils that can degrade SNR, and introduce biases that reduce accuracy of the loss function. Lastly, prospectively undersampled experiments are warranted to demonstrate the reliability of ssGAN in practice [61], [66]. Prospective undersampling can be achieved by modifying stock pulse sequences for acceleration, where the sampling mask for a scan is either selected from a stored set of a priori determined random masks, or generated on the fly as the sequence is initiated.

Here, we followed an empirical risk minimization strategy for model training, as many other studies in the domain of multi-contrast MRI synthesis [12], [18], [23], [24], [26], [27]. While this strategy performs desirably when training sets are sufficiently diverse, the resultant models might suboptimally generalize to atypical cases scarcely represented in the training set. Note that the IXI and in vivo datasets only contained healthy subjects, so atypicaly anatomy was not observed. It

TABLE VII: Effects of multi-/single-coil loss terms

	$T_2 \rightarrow PD$		$PD \rightarrow T_2$	
	PSNR	SSIM	PSNR	SSIM
$ssGAN_{(2,2)}$	**25.75**	**88.14**	**24.59**	**86.56**
$ssGAN_{(2,2)}$(single)	23.13	80.77	24.51	83.18
$ssGAN_{(2,4)}$	**25.47**	**87.82**	**24.46**	**85.97**
$ssGAN_{(2,4)}$(single)	22.91	80.77	24.07	81.00
$ssGAN_{(4,2)}$	**24.85**	**86.64**	**24.70**	**87.25**
$ssGAN_{(4,2)}$(single)	22.61	78.95	23.69	79.90
$ssGAN_{(4,4)}$	**24.73**	**86.22**	**24.05**	**84.84**
$ssGAN_{(4,4)}$(single)	22.49	78.85	23.55	79.45

TABLE VIII: Effects of randomized/non-randomized sampling masks

	$T_1 \rightarrow T_2$		$T_2 \rightarrow T_1$	
	PSNR	SSIM	PSNR	SSIM
$ssGAN_{(2,2)}$	**26.89**	**93.32**	**27.43**	**94.05**
$ssGAN_{(2,2)}$(w/o rand)	26.82	93.19	27.27	93.75
$ssGAN_{(2,4)}$	**26.77**	**93.10**	**27.25**	**93.80**
$ssGAN_{(2,4)}$(w/o rand)	26.59	92.74	27.18	93.72
$ssGAN_{(2,6)}$	**26.70**	**92.88**	**27.10**	**93.67**
$ssGAN_{(2,6)}$(w/o rand)	26.50	92.63	27.07	93.61
$ssGAN_{(2,8)}$	**26.68**	**92.83**	**27.14**	**93.79**
$ssGAN_{(2,8)}$(w/o rand)	26.44	92.40	26.96	93.42
$ssGAN_{(2,10)}$	**26.50**	**92.43**	**26.95**	**93.35**
$ssGAN_{(2,10)}$(w/o rand)	26.26	92.08	26.86	93.22

remains important future work to assess the generalization performance of ssGAN on patients with brain pathology. When needed, reliability might be improved by adopting distributionally robust optimization strategies with asymmetric risk measures that give higher weight to atypical cases [67].

We examined one-to-one synthesis tasks to translate between single source and target MRI contrasts. A preliminary radiological evaluation was conducted where synthesized images were rated according to their general visual similarity to the ground-truth images. Since the examined datasets involved only healthy subjects, contrast transfer in pathological tissue was not evaluated. Future studies are required to assess the quantitative and radiological performance of ssGAN on patient cohorts with pathology. Furthermore, it is likely that desired information in the target contrast may not always be sufficiently encoded in a single source contrast, particularly for pathological tissue [27]. In those cases, complementary information from multiple distinct source-contrast acquisitions might be required for successful estimation of target-contrast images. To enable many-to-one synthesis, the ssGAN model could be modified to receive information from multiple source contrasts as separate input channels [24], and dedicated fusion modules can be introduced across higher stages of the network model for improved performance [25], [35].

The ssGAN implementation considered here leverages a multi-coil loss between undersampled versions of the synthesized and reference target images. This image-domain loss implicitly assumes that the source- and target-contrast acquisitions are spatially registered. The datasets examined here were either aligned or a registration step was performed during preprocessing. Alternatively, a deep-learning model can be cascaded to the input of ssGAN for spatial registration. It remains important future work to explore this integrated registration and synthesis approach.

In summary, we proposed a semi-supervised learning framework based on a GAN architecture to train synthesis models in the absence of fully-sampled ground truth. While the data-efficiency of ssGAN was primarily demonstrated for MRI contrast translation in the brain, it can also be adopted to other anatomies, other recovery tasks including multi-parametric MRI, or cross-modality mapping [39], [42], [68].

REFERENCES

[1] B. Moraal *et al.*, "Multi-contrast, isotropic, single-slab 3D MR imaging in multiple sclerosis," *Neuroradiol J*, vol. 22, pp. 33–42, 2009.

[2] S. Bakas *et al.*, "Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features," *Sci Data*, vol. 4, 2017.

[3] B. Thukral, "Problems and preferences in pediatric imaging," *Indian J Radiol Imaging*, vol. 25, pp. 359–364, 2015.

[4] K. Krupa and M. Bekiesińska-Figatowska, "Artifacts in magnetic resonance imaging," *Pol J Radiol*, vol. 80, pp. 93–106, 2015.

[5] J. E. Iglesias, E. Konukoglu, D. Zikic, B. Glocker, K. Van Leemput, and B. Fischl, "Is synthesizing MRI contrast useful for inter-modality analysis?" in *Med Image Comput Comput Assist Interv*, 2013, pp. 631–638.

[6] D. Lee, W.-J. Moon, and J. C. Ye, "Assessing the importance of magnetic resonance contrasts using collaborative generative adversarial networks," *Nat Mach Intell*, vol. 2, p. 34–42, 2020.

[7] Y. Huo, Z. Xu, S. Bao, A. Assad, R. G. Abramson, and B. A. Landman, "Adversarial synthesis learning enables segmentation without target modality ground truth," in *IEEE Int Symp Biomed Imaging*, 2018, pp. 1217–1220.

[8] S. Roy, A. Carass, and J. Prince, "Magnetic resonance image example-based contrast synthesis," *IEEE Trans Med Imaging*, vol. 32, pp. 2348–2363, 2013.

[9] S. U. H. Dar, M. Yurt, M. Shahdloo, M. E. Ildız, B. Tinaz, and T. Çukur, "Prior-guided image reconstruction for accelerated multi-contrast MRI via generative adversarial networks," *IEEE J Sel Top Sign Process*, vol. 14, pp. 1072–1087, 2020.

[10] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magn Reson Med*, vol. 58, no. 6, pp. 1182–1195, 2007.

[11] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *IEEE International Conference on Computer Vision*, 2017, pp. 2242–2251.

[12] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *IEEE Conf Comput Vis Pattern Recognit*, 2017, pp. 5967–5976.

[13] H.-C. Shin *et al.*, "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE Trans Med Imaging*, vol. 35, no. 5, pp. 1285–1298, 2016.

[14] J. Schlemper, J. Caballero, J. V. Hajnal, A. N. Price, and D. Rueckert, "A deep cascade of convolutional neural networks for dynamic MR image reconstruction," *IEEE Trans Med Imaging*, vol. 37, no. 2, pp. 491–503, 2018.

[15] H. Gupta, K. H. Jin, H. Q. Nguyen, M. T. McCann, and M. Unser, "CNN-based projected gradient descent for consistent CT image reconstruction," *IEEE Trans Med Imaging*, vol. 37, no. 6, pp. 1440–1453, 2018.

[16] K. Armanious, C. Jiang, S. Abdulatif, T. Küstner, S. Gatidis, and B. Yang, "Unsupervised medical image translation using Cycle-MedGAN," in *European Signal Processing Conference*, 2019, pp. 1–5.

[17] K. H. Kim, W.-J. Do, and S.-H. Park, "Improving resolution of MR images with an adversarial network incorporating images with different contrast," *Med Phys*, vol. 45, no. 7, pp. 3120–3131, 2018.

[18] A. Chartsias, T. Joyce, M. V. Giuffrida, and S. A. Tsaftaris, "Multimodal MR synthesis via modality-invariant latent representation," *IEEE Trans Med Imaging*, vol. 37, no. 3, pp. 803–814, 2018.

[19] V. Sevetlidis, M. Giuffrida, and S. Tsaftaris, "Whole image synthesis using a deep encoder-decoder network," in *Simul Synth Med Imaging*, 2016, pp. 127–137.

[20] T. Joyce, A. Chartsias, and S. Tsaftaris, "Robust multi-modal MR image synthesis," in *Med Image Comput Comput Assist Interv*, 2017, pp. 347–355.

[21] W. Wei *et al.*, "Fluid-attenuated inversion recovery MRI synthesis from multisequence MRI using three-dimensional fully convolutional networks for multiple sclerosis," *J Med Imaging*, vol. 6, p. 1, 2019.

[22] C. Bowles *et al.*, "Pseudo-healthy image synthesis for white matter lesion segmentation," in *Simul Synth Med Imaging*, Cham, 2016, pp. 87–96.

[23] S. U. H. Dar, M. Yurt, L. Karacan, A. Erdem, E. Erdem, and T. Çukur, "Image synthesis in multi-contrast MRI with conditional generative adversarial networks," *IEEE Trans Med Imaging*, vol. 38, no. 10, pp. 2375–2388, 2019.

[24] A. Sharma and G. Hamarneh, "Missing MRI pulse sequence synthesis using multi-modal generative adversarial network," *IEEE Trans Med Imaging*, vol. 39, no. 4, pp. 1170–1183, 2020.

[25] T. Zhou, H. Fu, G. Chen, J. Shen, and L. Shao, "Hi-Net: hybrid-fusion network for multi-modal MR image synthesis," *IEEE Trans Med Imaging*, vol. 39, no. 9, pp. 2772–2781, 2020.

[26] D. Nie *et al.*, "Medical image synthesis with deep convolutional adversarial networks," *IEEE Trans Biomed Eng*, vol. 65, no. 12, pp. 2720–2730, 2018.

[27] D. Lee, J. Kim, W.-J. Moon, and J. C. Ye, "CollaGAN: Collaborative GAN for missing image data imputation," in *IEEE Conf Comput Vis Pattern Recognit*, 2019, pp. 2482–2491.

[28] H. Li *et al.*, "DiamondGAN: Unified multi-modal generative adversarial networks for MRI sequences synthesis," in *Medical Image Computing and Computer Assisted Intervention*, Cham, 2019, pp. 795–803.

[29] B. Yu, L. Zhou, L. Wang, Y. Shi, J. Fripp, and P. Bourgeat, "Ea-GANs: Edge-aware generative adversarial networks for cross-modality MR image synthesis," *IEEE Trans Med Imaging*, vol. 38, no. 7, pp. 1750–1762, 2019.

[30] K. Armanious *et al.*, "MedGAN: Medical image translation using GANs," *Comput Med Imaging Graph*, vol. 79, p. 101684, 2020.

[31] A. Beers *et al.*, "High-resolution medical image synthesis using progressively grown generative adversarial networks," *arXiv:1805.03144*, 2018.

[32] A. Hagiwara *et al.*, "Improving the quality of synthetic FLAIR images with deep learning using a conditional generative adversarial network for pixel-by-pixel image translation," *Am J Neuroradiol*, vol. 40, pp. 224–230, 2019.

[33] G. Wang *et al.*, "Synthesize high-quality multi-contrast magnetic resonance imaging from multi-echo acquisition using multi-task deep generative model," *IEEE Trans Med Imaging*, vol. 39, no. 10, pp. 3089–3099, 2020.

[34] B. Yu, L. Zhou, L. Wang, J. Fripp, and P. Bourgeat, "3D cGAN based cross-modality MR image synthesis for brain tumor segmentation," in *IEEE Int Symp Biomed Imaging*, 2018, pp. 626–630.

[35] M. Yurt, S. U. Dar, A. Erdem, E. Erdem, K. K. Oğuz, and T. Çukur, "mustGAN: multi-stream generative adversarial networks for MR image synthesis," *Med Image Anal*, vol. 70, p. 101944, 2021.

[36] O. Dalmaz, M. Yurt, and T. Çukur, "ResViT: Residual vision transformers for multi-modal medical image synthesis," *IEEE Trans Med Imaging*, pp. 1–1, 2022.

[37] S. Olut, Y. H. Sahin, U. Demir, and G. Unal, "Generative adversarial training for MRA image synthesis using multi-contrast MRI," in *Predict Intell Med*, Cham, 2018, pp. 147–154.

[38] M. Sohail, M. N. Riaz, J. Wu, C. Long, and S. Li, "Unpaired multi-contrast MR image synthesis using generative adversarial networks," in *Simul Synth Med Imaging*, Cham, 2019, pp. 22–31.

[39] J. M. Wolterink, A. M. Dinkla, M. H. F. Savenije, P. R. Seevinck, C. A. T. van den Berg, and I. Išgum, "Deep MR to CT synthesis using unpaired data," in *Simul Synth Med Imaging*, Cham, 2017, pp. 14–23.

[40] Y. Ge *et al.*, "Unpaired MR to CT synthesis with explicit structural constrained adversarial learning," in *IEEE Int Symp Biomed Imaging*, 2019, pp. 1096–1099.

[41] Y. Choi, M. Choi, M. Kim, J.-W. Ha, S. Kim, and J. Choo, "StarGAN: unified generative adversarial networks for multi-domain image-to-image translation," in *IEEE Conf Comput Vis Pattern Recognit*, 2018, pp. 8789–8797.

[42] C.-B. Jin *et al.*, "Deep CT to MR synthesis using paired and unpaired data," *Sensors*, vol. 19, no. 10, 2019.

[43] Z. Wang, Y. Lin, K.-T. T. Cheng, and X. Yang, "Semi-supervised mp-MRI data synthesis with StitchLayer and auxiliary distance maximization," *Med Image Anal*, vol. 59, p. 101565, 2020.

[44] B. Yaman, S. A. H. Hosseini, S. Moeller, J. Ellermann, K. Uğurbil, and M. Akçakaya, "Self-supervised learning of physics-guided reconstruction neural networks without fully sampled reference data," *Magn Reson Med*, vol. 84, no. 6, pp. 3172–3191, 2020.

[45] E. K. Cole, J. M. Pauly, S. S. Vasanawala, and F. Ong, "Unsupervised MRI reconstruction with generative adversarial networks," *arXiv:2008.13065*, 2020.

[46] Y. Korkmaz, S. U. Dar, M. Yurt, M. Özbey, and T. Çukur, "Unsupervised MRI reconstruction via zero-shot learned adversarial transformers," *IEEE Trans Med Imaging*, pp. 1–1, 2022.

[47] H. K. Aggarwal, A. Pramanik, and M. Jacob, "ENSURE: A general approach for unsupervised training of deep image reconstruction algorithms," *arXiv:2010.10631*, 2020.

[48] G. Elmas *et al.*, "Federated learning of generative image priors for MRI reconstruction," *arXiv:2202.04175*, 2022.

[49] X. Mao, Q. Li, H. Xie, R. Y. Lau, Z. Wang, and S. Paul Smolley, "Least squares generative adversarial networks," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, Oct 2017, pp. 2794–2802.

[50] M. Uecker *et al.*, "ESPIRiT—an eigenvalue approach to autocalibrating parallel MRI: Where SENSE meets GRAPPA," *Magn Reson Med*, vol. 71, no. 3, pp. 990–1001, 2014.

[51] Y. Han, L. Sunwoo, and J. C. Ye, "k-Space Deep Learning for Accelerated MRI," *IEEE Trans Med Imaging*, vol. 39, no. 2, pp. 377–386, 2020.

[52] M. Lustig and J. M. Pauly, "SPIRiT: iterative self-consistent parallel imaging reconstruction from arbitrary k-space," *Magn Reson Med*, vol. 64, no. 2, pp. 457–471, 2010.

[53] S. Smith *et al.*, "Advances in functional and structural MR image analysis and implementation as FSL," *NeuroImage*, vol. 23 Suppl 1, pp. S208–19, 2004.

[54] T. Zhang, J. M. Pauly, S. S. Vasanawala, and M. Lustig, "Coil compression for accelerated imaging with Cartesian sampling," *Magn Reson Med*, vol. 69, no. 2, pp. 571–582, 2013.

[55] M. Bydder, D. Larkman, and J. Hajnal, "Combination of signals from array coils using image-based estimation of coil sensitivity profiles," *Magn Reson Med*, vol. 47, no. 3, pp. 539–548, 2002.

[56] M. Akçakaya, S. Moeller, S. Weingärtner, and K. Uğurbil, "Scan-specific robust artificial-neural-networks for k-space interpolation (RAKI) reconstruction: Database-free deep learning for fast imaging," *Magn Reson Med*, vol. 81, no. 1, pp. 439–453, 2019.

[57] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, "GANs trained by a two time-scale update rule converge to a local nash equilibrium," in *Adv Neural Inf Process Syst*, 2017, p. 6629–6640.

[58] T. DeVries, A. Romero, L. Pineda, G. W. Taylor, and M. Drozdzal, "On the evaluation of conditional GANs," *arXiv:1907.08175*, 2019.

[59] T. M. Quan, T. Nguyen-Duc, and W.-K. Jeong, "Compressed sensing MRI reconstruction using a generative adversarial network with a cyclic loss," *IEEE Trans Med Imaging*, vol. 37, no. 6, pp. 1488–1497, 2018.

[60] E. Kerfoot *et al.*, "Synthesising images and labels between MR sequence types with CycleGAN," in *Dom Adapt Rep Transfer and Med Image Learn Less Labels Imperfect Data*. Cham: Springer, 2019, pp. 45–53.

[61] T. Küstner, K. Armanious, J. Yang, B. Yang, F. Schick, and S. Gatidis, "Retrospective correction of motion-affected MR images using deep learning frameworks," *Magn Reson Med*, vol. 82, no. 4, pp. 1527–1540, 2019.

[62] W.-J. Do, S. Seo, Y. Han, J. C. Ye, S. H. Choi, and S.-H. Park, "Reconstruction of multicontrast MR images through deep learning," *Medical Physics*, vol. 47, no. 3, pp. 983–997, 2020.

[63] C. Zhao, B. E. Dewey, D. L. Pham, P. A. Calabresi, D. S. Reich, and J. L. Prince, "SMORE: A Self-Supervised Anti-Aliasing and Super-Resolution Algorithm for MRI Using Deep Learning," *IEEE Trans Med Imaging*, vol. 40, no. 3, pp. 805–817, 2021.

[64] Q. Lyu *et al.*, "Multi-contrast super-resolution MRI through a progressive network," *IEEE Trans Med Imaging*, vol. 39, no. 9, pp. 2738–2749, 2020.

[65] A. Gungor, B. Askin, D. A. Soydan, E. U. Saritas, C. B. Top, and T. Çukur, "TranSMS: Transformers for super-resolution calibration in magnetic particle imaging," *arXiv:2111.02163*, 2021.

[66] K. H. Kim, S. Seo, W.-J. Do, H. M. Luu, and S.-H. Park, "Varying undersampling directions for accelerating multiple acquisition magnetic resonance imaging," *NMR in Biomedicine*, vol. 35, no. 4, p. e4572, 2022.

[67] H. Rahimian and S. Mehrotra, "Distributionally robust optimization: A review," *arXiv:1908.05659*, 2019.

[68] B. Dewey *et al.*, "DeepHarmony: a deep learning approach to contrast harmonization across scanner changes," *Magn Reson Imaging*, vol. 64, 2019.