



Published in final edited form as:

Magn Reson Med. 2019 November ; 82(5): 1890–1904. doi:10.1002/mrm.27827.

SANTIS: Sampling-Augmented Neural neTwork with Incoherent Structure for MR image reconstruction

Fang Liu^{1,4}, Alexey Samsonov¹, Lihua Chen², Richard Kijowski¹, Li Feng³

¹Department of Radiology, University of Wisconsin-Madison, Madison, Wisconsin, USA

²Department of Radiology, Southwest Hospital, Chongqing, China

³Department of Medical Physics, Memorial Sloan Kettering Cancer Center, New York, USA

Abstract

Purpose: To develop and evaluate a novel deep learning-based reconstruction framework called SANTIS (Sampling-Augmented Neural neTwork with Incoherent Structure) for efficient MR image reconstruction with improved robustness against sampling pattern discrepancy.

Methods: With a combination of data cycle-consistent adversarial network, end-to-end convolutional neural network mapping, and data fidelity enforcement for reconstructing undersampled MR data, SANTIS additionally employs a sampling-augmented training strategy by extensively varying undersampling patterns during training, so that the network is capable of learning various aliasing structures and thereby removing undersampling artifacts more effectively and robustly. The performance of SANTIS was demonstrated for accelerated knee imaging and liver imaging using a Cartesian trajectory and a golden-angle radial trajectory, respectively. Quantitative metrics were used to assess its performance against different references. The feasibility of SANTIS in reconstructing dynamic contrast-enhanced images was also demonstrated using transfer learning.

Results: Compared to conventional reconstruction that exploits image sparsity, SANTIS achieved consistently improved reconstruction performance (lower errors and greater image sharpness). Compared to standard learning-based methods without sampling augmentation (e.g., training with a fixed undersampling pattern), SANTIS provides comparable reconstruction performance but significantly improved robustness against sampling pattern discrepancy. SANTIS also achieved encouraging results for reconstructing liver images acquired at different contrast phases.

Conclusion: By extensively varying undersampling patterns, the sampling-augmented training strategy in SANTIS can remove undersampling artifacts more robustly. The novel concept behind SANTIS can particularly be useful for improving the robustness of deep learning-based image reconstruction against discrepancy between training and inference, an important but currently less explored topic.

⁴Corresponding Author: Fang Liu, Ph.D., Department of Radiology, University of Wisconsin-Madison, 1111 Highland Avenue, Madison, Wisconsin USA, 53705-2275, Phone: 612-222-9728, FAX: 608-263-5112, leoliuf@gmail.com.

Keywords

Deep Learning; Image Reconstruction; Data Augmentation; Sampling Discrepancy; Reconstruction Robustness; Data Cycle-Consistent Adversarial Network

INTRODUCTION

There has been much recent interest in applying deep learning to a wide range of medical imaging applications, including disease classification (1,2), tissue segmentation (3–8), and lesion detection and recognition (9–12). Deep learning also holds great promise for reconstructing undersampled MRI data, providing new opportunities to further improve the performance of rapid MRI. Preliminary studies performed over the past few years (13–22) have demonstrated the ability of deep learning-based reconstruction methods to provide improved image quality with reduced reconstruction time. Compared to conventional rapid imaging techniques, such as parallel imaging (23–25) and compressed sensing (26–28), deep learning-based reconstruction uses a data-driven approach to characterize image features and removes undersampling artifacts by inferring features from a large image database. For example, a straightforward end-to-end mapping using convolutional neural networks (CNNs) has been shown to efficiently remove artifacts from undersampled images (16–18), translate k -space information into images (19), estimate missing k -space data (20,21), and reconstruct MR parameter maps(22). More importantly, upon completion of the CNN training for a particular application, a deep learning system can reconstruct newly-acquired undersampled images efficiently, generally in order of seconds (13–22).

In existing deep learning-based reconstruction methods, supervised training is performed using high-quality reference images and their corresponding undersampled pairs (13–22). The undersampled images are typically generated by a fixed undersampling pattern for network training, and the trained network is then applied to reconstruct new images acquired with the same undersampling pattern. While such a training strategy can achieve a favorable reconstruction performance for a pre-selected undersampling pattern, the robustness of the trained network against any discrepancy of undersampling schemes is typically poor. For example, a recent study has shown that a deviation of undersampling patterns between the training and inference can degrade image quality (29). This observation emphasizes the need for better network training strategies that can more robustly remove image artifacts from a wide variety of undersampling patterns.

The main contribution of this work is to propose a new training strategy for improving the robustness of a trained network against discrepancies of undersampling schemes that may occur between training and inference. The hypothesis was that deep learning-based reconstruction could benefit from extensively varying undersampling patterns during the training process, leading to better generalization and performance of the trained network for removing undersampling artifacts. Such a training strategy enforces sampling or k -space trajectory augmentation, providing an optimal method for highly efficient and robust image reconstruction. The technique may also be better suited for applications that employ non-repeating continuous k -space undersampling. This new framework, termed as Sampling-

Augmented Neural network with Incoherent Structure (SANTIS), was implemented using a novel data cycle-consistent adversarial network(30) and was evaluated in retrospectively undersampled knee and liver MR image datasets using variable-density Cartesian sampling and golden-angle radial sampling, respectively.

THEORY

CNN-Regularized Image Reconstruction

In MR image reconstruction, an image is estimated from a discrete set of k -space measurements that can be expressed as:

$$d = Ex + \varepsilon \quad (1)$$

Here, d is a vector of length N for measured k -space data and $N = N_x \times N_y$ is the total number of k -space samples. x is the image to be reconstructed, and ε represents complex noise in the measurements. In a general case, the encoding matrix E can be expanded as:

$$E = U\mathfrak{F}C \quad (2)$$

where \mathfrak{F} is the Fourier transform operator and C represents coil sensitivities. In the case of undersampling, U is an undersampling mask to select desired k -space data points. An SNR-optimized reconstruction can be accomplished in the least squares sense as:

$$\hat{x} = \arg \min_x \frac{1}{2} \|Ex - d\|_2^2 \quad (3)$$

where $\|\cdot\|_2$ is the l_2 norm. Eq.[3] can be poorly-conditioned with a high undersampling factor, and thus, a regularization term can be additionally incorporated as:

$$\hat{x} = \arg \min_x \left(\frac{1}{2} \|Ex - d\|_2^2 + \lambda \mathfrak{R}(x) \right) \quad (4)$$

The design of the penalty function \mathfrak{R} is governed by prior knowledge about the image x , with a regularization parameter λ controlling the trade-off between the data fidelity (the left term) and prior information assumptions (the right term). In conventional MR reconstruction, \mathfrak{R} is often chosen as

$$\mathfrak{R}(x) = \|Tx\|_p \quad (5)$$

where $\|\cdot\|_p$ represents a p -norm and T is a specific transform operator, which, for example, is often selected as the wavelet transform or finite differences in the context of compressed sensing (26).

In contrast to conventional regularizations, the penalty function can be implemented by learning an end-to-end deep CNN mapping from an undersampled image to its corresponding fully sampled pair, similar to many recent proposals (13–15,17). In other words, CNN can serve as a spatial or spatial-temporal regularizer in constrained image reconstruction, and Eq.[4] can be rewritten as:

$$\hat{x} = \arg \min_x \left(\frac{1}{2} \|Ex - d\|_2^2 + \lambda \|F(x_u) - x\|_p \right) \quad (6)$$

Here, $F: f_{gen}(x_u | \theta) \rightarrow \hat{x}$ is a generator function using a deep CNN conditioned on network parameters θ , and the x_u is an undersampled image to be reconstructed.

In a practical implementation from recent CNN reconstruction studies (14,31), a data fidelity loss can be incorporated to enforce data fidelity in combination with the CNN-regularization as shown in Figure 1. The image reconstruction, expressed in Eq.[6] can then be approximated as:

$$\hat{F} = \arg \min_F \left(\lambda_{loss2} \mathbb{E}_{x_u \sim P(x_u)} [\| \Phi_u F(x_u) - x_u \|_p] + \lambda_{loss1} \mathbb{E}_{x_u \sim P(x_u)} [\| F(x_u) - x \|_p] \right) \quad (7)$$

where two loss functions were generated including a CNN generator term for CNN-regularization (pixel-wise loss 1: the right term) and a data fidelity loss term for data consistency (pixel-wise loss 2: the left term). $\Phi_u = C^H \mathfrak{S}^H U \mathfrak{S} C$ is the undersampling encoding operator and $\Phi_u \hat{x} = C^H \mathfrak{S}^H U \mathfrak{S} C \hat{x} \rightarrow \hat{x}_u$ generates undersampled images using the undersampling mask and coil sensitivities. $\mathbb{E}_{x_u \sim P(x_u)}[\cdot]$ is the expectation of a probability function given x_u belongs to the data distribution of the undersampled image domain $P(x_u)$. λ_{loss1} and λ_{loss2} are corresponding weighting parameters controlling the balance between the two loss terms. More specifically, the first loss term (pixel-wise loss 1) removes aliasing artifacts such that the reconstructed images look like their reference, while the second loss term (pixel-wise loss 2) ensures that the reconstructed images from the CNN generator produce undersampled images matching the acquired k -space measurements.

Adversarial Loss

A l_1 norm or a l_2 norm is often chosen for the pixel-wise loss terms (loss 1 and 2) of Eq.[7] in most studies. While either of these two norms can ensure good reconstruction performance by suppressing noise and artifacts, such a reconstruction scheme has been demonstrated to be suboptimal for restoring undersampled images with a ‘natural’ appearance as in corresponding fully sampled images (14,15). Several studies have shown that using CNN mapping only with pixel-wise losses can result in image blurring, detail loss, and degraded perceptual quality (14,15). Therefore, Generative Adversarial Network (GAN) has been recently proposed to overcome the drawbacks of l_1 or l_2 norm, where an adversarial loss is used to promote image texture preservation in image restoration (32). In

the context of image reconstruction as shown in Eq.[7], an adversarial loss can be incorporated into the CNN generator as an additional loss term to promote texture preservation as shown in Figure 1. In other words, in addition to the CNN generator function F , another multiple layer CNN discriminator $D: f_{dis}(x|\delta) \rightarrow 1$ conditioned on network parameters δ is designed to distinguish reconstructed images versus corresponding reference images in the fully sampled image domain $P(x)$. Mathematically, this discriminator $D(x)$ outputs a scalar representing the probability that x comes from the fully sampled image domain rather than the CNN generator output $F(x_u)$. The adversarial loss is formatted as:

$$L_{gan}(F, D) = \mathbf{E}_{x \sim P(x)}[\log D(x)] + \mathbf{E}_{x_u \sim P(x_u)}[\log(1 - D(F(x_u)))] \quad (8)$$

Data Cycle-Consistent Adversarial Network

Following all loss terms mentioned above, a Data Cycle-Consistent Adversarial Network (30) can be formulated for reconstructing undersampled MR images using a multi-loss joint training model as shown in Figure 1. This model consists of a CNN generator that generates a pixel-wise loss term (loss 1) and an adversarial loss term as CNN regularization. An additional data fidelity loss path creates another pixel-wise loss term (loss 2) for enforcing data fidelity. Therefore, a full objective function can be written as:

$$L_{full} = L_{cyc} + L_{gan} = \lambda_{loss2} \mathbf{E}_{x_u \sim P(x_u)}[\|\Phi_u F(x_u) - x_u\|_p] + \lambda_{loss1} \mathbf{E}_{x_u \sim P(x_u)}[\|F(x_u) - x\|_p] + \lambda_{gan} \left(\mathbf{E}_{x \sim P(x)}[\log D(x)] + \mathbf{E}_{x_u \sim P(x_u)}[\log(1 - D(F(x_u)))] \right) \quad (9)$$

Here, λ_{gan} is a weighting factor to control the contribution of the GAN regularization to the reconstruction. The overall aim of the training is to solve the full objective function in a two-player minimax fashion (32) as:

$$\hat{F}, \hat{D} = \arg \min_F \max_D L_{full}(F, D) \quad (10)$$

Namely, F aims to minimize this objective function against the adversary D that seeks to maximize it. In other words, F tries to reconstruct undersampled images to look similar to corresponding fully sampled images, while D aims to distinguish reconstructed images from real fully sampled images. As demonstrated in the studies for natural image-to-image translation (32–34), successful training using such a competing scheme can result in an optimal CNN generator capable of generating artificial images indistinguishable from the original images. It should be noted that optimizing Eq.[10] aims to search for optimal network parameters $\hat{\theta}$ for $F: f_{gen}(x_u | \theta) \rightarrow \hat{x}$ during the training phase, so that once the training is completed and satisfied, $\hat{\theta}$ can be fixed for the CNN generator and directly

applied to reconstruct a new undersampled image by using the CNN mapping F to the undersampled image as:

$$x_{cnn} = f_{gen}(x_u | \hat{\theta}), x_u \in P(x_u) \quad (11)$$

Sampling-Augmented Data Cycle-Consistent Adversarial Network

To strengthen the robustness of the trained network towards sampling pattern discrepancy, the training of Eq.[10] can be further augmented in a manner where x_u is generated to have different artifact patterns at each training iteration. Once the training is successful, the trained network is capable of learning a variety of aliasing artifact patterns and can then be generalized towards removing a broad range of artifacts. This can be achieved by varying the undersampling mask for generating each training data pair throughout the entire training iterations. Eq.[9] is then updated as:

$$\begin{aligned} L_{full} = L_{cyc} + L_{gan} = & \lambda_{loss1} \mathbf{E}_{x_{u,ij} \sim P(x_u)} [\|\Phi_{u,ij} F(x_{u,ij}) - x_{u,ij}\|_P] \\ & + \lambda_{loss2} \mathbf{E}_{x_{u,ij} \sim P(x_u)} [\|F(x_{u,ij}) - x\|_P] \\ & + \lambda_{gan} \left(\mathbf{E}_{x \sim P(x)} [\log D(x)] + \mathbf{E}_{x_{u,ij} \sim P(x_u)} [\log(1 - D(F(x_{u,ij})))] \right) \end{aligned} \quad (12)$$

where i denotes the i^{th} image in the training dataset and j denotes the j^{th} training iteration. In other words, x_u has a different aliasing artifact pattern in each image of the dataset ($x_{u,i}$), and the artifact is also changing for different iterations ($x_{u,ij}$). Accordingly, the undersampling operator $\Phi_{u,ij}$ is rewritten as:

$$\begin{aligned} \Phi_{u,ij} &= C^H \mathfrak{S}^H U_{ij} \mathfrak{S} C \quad (13) \\ \Phi_{u,ij} \hat{x} &= C^H \mathfrak{S}^H U_{ij} \mathfrak{S} C \hat{x} \rightarrow \hat{x}_{u,ij} \end{aligned}$$

to reflect the variation of undersampling patterns U_{ij} during training.

METHODS

In-Vivo Image Datasets

The study was performed in compliance with the Health Insurance Portability and Accountability Act (HIPPA) regulations and was approved by the local Institutional Review Boards (IRB). Knee datasets were acquired on 30 patients undergoing a clinical knee MRI examination, with a waiver of written informed consent. The datasets were acquired on a 3.0T scanner (Signa Premier, GE Healthcare, USA) equipped with an 18-element knee coil array using a clinical two-dimensional (2D) coronal proton density-weighted fast spin-echo (Cor-PD-FSE) sequence. Relevant imaging parameters included: matrix size = 420×448, FOV = 15×15 cm², TR/TE = 2500/20ms, flip angle = 111°, echo train length = 5, slice thickness = 3mm, number of slices = 30-38, and scan time of 210 seconds

Liver datasets were acquired on 63 patients undergoing a clinical liver MRI examination, with written informed consent. The datasets were acquired approximately 20 minutes after the injection of Gd-EOB-DTPA contrast (Eovist) on a 3.0T scanner (TimTrio, Siemens Medical Systems, Germany) equipped with a 12-element body matrix coil using a prototype fat-saturated stack-of-stars golden-angle radial sequence (35,36). Relevant imaging parameters included: matrix size = 256×256 , FOV = $33 \times 33 \text{cm}^2$, TR/TE = 3.40/1.68ms, flip angle = 10° , slice thickness = 5mm, number of slices = 44, 80% partial Fourier along the slice dimension, 1000 spokes acquired in each partition, and scan time of 178 seconds.

In addition, a dynamic contrast-enhanced liver image dataset was also acquired on one healthy volunteer, with written informed consent. This dataset was acquired during contrast injection using the same golden angle radial sequence with a similar protocol as previously described except 1222 spokes for each partition and longer scan time of 190 seconds to ensure capture of contrast dynamics from a pre-contrast phase to a delayed phase.

Implementation of Neural Network

A U-Net architecture (Supporting Information Figure S1) with residual learning was adapted from previous image-to-image translation studies (5,34) (<https://github.com/phillipi/pix2pix>) for the CNN generator (F) to map undersampled image domain into corresponding fully sample image domain. This U-Net structure is composed of an encoder network and a decoder network with multiple shortcut connections between them (37). The encoder is used to achieve efficient data compression while probing robust and spatial invariant image features of input images. A decoder network with a mirrored structure of the encoder is applied for restoring image features and increasing image resolution using the output of the encoder network. Multiple shortcut connections are incorporated to concatenate entire feature maps from encoder to decoder to enhance mapping performance. Residual learning is implemented to further boost learning robustness and accuracy (38,39). Such a network structure and its variants have shown promising results for image-to-image translation in many recent deep learning studies (5,37–39). For the adversarial configuration, a network architecture developed in PatchGAN (34) was used for the discriminator network (D), which aims to differentiate real from artificial images in the adversarial process using image patch-based assessment on input images. An illustration of the applied PatchGAN is shown in Supporting Information Figure S1. Such a patch-level discriminator architecture provides efficient network training with fewer network parameters and has demonstrated favorable results at differentiating real from artificial images in many recent GAN studies (5,34,40,41).

Network Training

In our framework, the complex MR image $x \in \mathbb{C}^M$ is converted into a new representation $x \in \mathbb{R}^{2M}$ so that the real and imaginary components are treated as two individual image channels for the network input (14,17). Twenty-two of the 30 knee datasets were randomly selected and used for training the deep learning network, while the remaining eight datasets were used for evaluation. Similarly, 55 of the 63 liver datasets were randomly selected and used for training, while the remaining eight datasets were used for evaluation.

The network weights for the training were initialized using the initialization scheme as described in (42) and were updated using Adam algorithm (43) with a fixed learning rate of 0.0002. The training was performed in a mini-batch manner with three image slices in a single mini-batch (5). The l_1 norm was applied for the pixel-wise loss terms in Eq. [12] to promote image sharpness following recent GAN studies (15). Total iteration steps corresponding to 200 epochs were carried out to ensure convergence of the training (see Supporting Information Figure S2). During each iteration, a two-step training strategy was applied where the CNN generator (F) and adversarial discriminator (D) were updated alternatively (5). After qualitative evaluation with a broad range of combinations, the parameters for the weight factors in the full objective function of Eq. [12] were empirically selected as $\lambda_{loss1} = 10$, $\lambda_{loss2} = 10$ and $\lambda_{gan} = 0.1$ for our study.

All the proposed algorithms were implemented in Python (v2.7, Python Software Foundation, Wilmington, DE). The CNNs were designed using the Keras package (44) running Tensorflow computing backend (45) on a 64-bit Ubuntu Linux system. All the training and evaluation were performed on a computer server with an Intel Xeon W3520 quad-core CPU, 32GB RAM, and an Nvidia GeForce GTX 1080Ti graphic card with a total of 3584 CUDA cores and 11GB GDDR5 RAM.

Fixed Training vs. Sampling-Augmented Training

Two training strategies were investigated in the study for both knee and liver imaging, including **1**) the fixed training strategy (referred to as **CNN-Fix** hereafter) and **2**) the sampling-augmented training strategy (**SANTIS**).

A library of one-dimensional (1D) variable-density Cartesian random undersampling patterns (26), consisting of 3000 different undersampling masks, were generated for the knee datasets as shown in Figure 2a. Each random sampling pattern was generated using Monte-Carlo simulation with a fixed probability density function, and corresponding sampling pattern lines were selected randomly according to this density. For each undersampling mask, the k -space center (5% of the total measurements) was consistently fully sampled. For the training of CNN-Fix, only the first undersampling mask was used throughout the entire training process as outlined in Eq. [9]. For the training of SANTIS, in contrast, an undersampling mask was randomly selected from this library for each image in each training iteration, as shown in Eq. [12]. This operation ensures that different artifacts can be created for each image at each training iteration so that the network can learn a wide range of undersampling artifact structures during training. As a pre-processing step, all the knee datasets were compressed into eight virtual coils to fit into GPU memory, and corresponding coil sensitivity maps were estimated using the adaptive combination approach (46). To test the acceleration performance, two acceleration rates ($R=3$ and $R=5$) were investigated for knee imaging. In addition, experiments for the uniform Cartesian undersampling patterns were also performed and described in the Supporting Information.

The training for the liver datasets was performed using the following steps. First, all the datasets were compressed into four virtual coils, and a multi-coil three-dimensional (3D) liver image (e.g., without coil combination) was reconstructed using a standard gridding algorithm for each dataset. Corresponding coil sensitivity maps were estimated using the

adaptive combination approach (46). Second, a library of radial trajectories consisting of 3000 consecutive golden-angle rotations was calculated, from which different undersampled golden-angle patterns could then be generated by combining a subset of consecutive radial trajectories (89 consecutive spokes in this study). Third, using these radial undersampling patterns, corresponding undersampled radial spokes and radial images could then be generated using non-uniform fast Fourier transform (NUFFT) and inverse NUFFT, respectively. Fourth, similar to knee imaging, the training of CNN-Fix was performed with only the first undersampling pattern (e.g., the 1st to the 89th spokes), while the training of SANTIS was performed by random selection of a radial undersampling pattern from the trajectory library. Due to the GPU memory limit, the 3D liver datasets were treated as multi-slice 2D images, so that both the training and reconstruction were performed slice-by-slice. As a pre-processing step, self-calibrating GRAPPA Operator Gridding (GROG) (47,48) was applied to all the generated reference-undersampled radial data pairs, so that the subsequent training could be performed entirely on a Cartesian grid.

Evaluation of Reconstruction Methods

In the first experiment, the robustness of the different training strategies was compared. Specifically, the trained networks from CNN-Fix and SANTIS were applied to reconstruct (the inference step) undersampled images generated with two undersampling masks/patterns, including the one used for CNN-Fix training (denoted as MaskC1 for knee imaging and MaskR1 for liver imaging hereafter) and a newly-generated one not used for training (denoted as MaskC2 for knee imaging and MaskR2 for liver imaging hereafter). It was ensured that both undersampling patterns were not included in the training of SANTIS to fully assess its robustness.

In the second experiment, both CNN-Fix and SANTIS were compared with conventional reconstruction methods, including a combination of compressed sensing and parallel imaging (CS-PI) (49) and a recently proposed Variational Network (VN) (13,29) that represents the current state-of-the-art deep learning-based reconstruction approach. For the CS-PI method, reconstruction was performed slice by slice, and coil sensitivity maps were generated as previously described. A Daubechies-4 wavelet transform was selected as the sparsifying transform. The VN method was implemented using the source code provided by the original developers (<https://github.com/VLOGroup/mri-variationalnetwork>) using reconstruction parameters previously configured for accelerated knee imaging (13,29). The VN approach was used as a reference for knee imaging as it has been previously described for this application. The undersampling patterns MaskC1 and MaskR1 were used for the comparison for knee and liver imaging, respectively.

Reconstruction performance was evaluated using quantitative metrics focusing on different aspects of the reconstruction quality. The normalized Root Mean Squared Error (nRMSE) was used to assess the overall reconstructed image errors. The Structural Similarity Index (SSIM) was used to estimate the overall image similarity with respect to the reference. The image sharpness was evaluated by calculating the relative reduction of Tenengrad measure (50) between the reconstructed and reference images, where local characteristics of pronounced image edges were explored. Differences were evaluated using a paired non-

parametric Wilcoxon signed-rank test with statistical significance defined as p -value less than 0.05.

In the third experiment, the trained network was evaluated on the dynamic contrast-enhanced liver dataset with retrospective undersampling to test the robustness of SANTIS for reconstructing images with different image contrasts. Specifically, the liver dataset was reconstructed using RACER-GRASP (51) with a temporal resolution of 15 seconds, resulting in a pre-contrast phase, three arterial phases, two venous phases, and several delayed phases. Undersampled radial images were then generated from these reconstructed contrast phases using different newly-generated golden-angle radial sampling patterns with 89 consecutive spokes each.

RESULTS

Figure 3 shows a comparison of different reconstruction methods for a representative knee dataset at $R=3$ with MaskC1. Compared to the CS-PI, learning-based reconstruction methods (VN, CNN-Fix-MaskC1, and SANTIS) provided better removal of aliasing artifacts in the bone and cartilage, as illustrated in the zoomed images. SANTIS showed the best reconstruction performance, creating a high-quality reconstructed image. A sharp texture, as indicated by the green arrows, was better preserved in SANTIS compared to the other methods. Similar result for uniform undersampling patterns is shown in Supporting Information Figure S4.

Figure 4 shows a comparison of different reconstruction methods for another knee dataset at $R=5$. Similar improvement in image quality was observed for SANTIS, with better representation of the layered structure of the femoral and tibial cartilage (green arrows). At this acceleration, CS-PI failed to effectively reconstruct the undersampled image, and both VN and CNN-Fix suffered from notable residual blurring and loss of image details, as shown by the red arrows in the zoomed bone and subcutaneous fat regions.

Figure 5 shows knee images comparing CNN-Fix with SANTIS using both MaskC1 (top row) and MaskC2 (bottom row) at $R=3$. As described in the Methods, MaskC1 was only used for the training of CNN-Fix training, while neither MaskC1 or MaskC2 were used for the SANTIS training. As expected, MaskC1 and MaskC2 created different aliasing artifacts, as highlighted by the green arrows in the zero-filling images. While CNN-Fix-MaskC1 showed decent image quality, CNN-Fix-MaskC2 failed to effectively reconstruct the undersampled image generated with MaskC2, with noticeable residual aliasing artifacts in the zoomed image, as indicated by the green arrows. In contrast, SANTIS demonstrated consistent reconstruction performance for both MaskC1 and MaskC2. Similar result for uniform undersampling patterns is shown in Supporting Information Figure S5.

Figure 6 shows a comparison of different reconstruction methods for radial liver imaging with MaskR1. Similar to knee imaging, SANTIS achieved the best image quality with better removal of streaking artifacts and preservation of tissue texture. CS-PI generated pseudo residual structures (hollow green arrows), which were potentially caused by the use of wavelet constraint. Meanwhile, CNN-Fix-MaskR1 suffered from noticeable residual

blurring as indicated by the green arrows. For both CS-PI and CNN-Fix-MaskR1, residual streaking artifacts were present, as illustrated by the red arrows in the zoomed region.

Figure 7 shows zoomed liver images with a tumor (green box in the reference image) comparing CNN-Fix and SANTIS using both MaskR1 and MaskR2. CNN-Fix-MaskR1 and CNN-Fix-MaskR2 suffered from residual streaking artifacts. SANTIS provided improved image quality for both sampling patterns, with a reduced level of residual streaking artifacts and a better delineation of the tumor. Corresponding full FOV images of this comparison are shown in Supporting Information Figure S3.

The group-wise quantitative analyses further confirmed the qualitative observations in the exemplary figures. The comparisons of reconstruction methods using different metrics, including the nRMSE, SSIM and Tenengrad, are summarized in Table 1 and Table 2 for all the testing knee and liver datasets. For knee imaging at both R=3 and R=5 (Table 1), all the deep learning-based reconstruction methods (VN, CNN-Fix-MaskC1, and SANTIS) achieved significantly better image quality ($p < 0.05$) than CS-PI. Among different deep learning-based methods, SANTIS achieved significantly better image quality than CNN-Fix-MaskC1 for nRMSE ($p = 0.003$) and Tenengrad ($p = 0.001$) at R=3 and significantly better image quality ($p < 0.003$) than CNN-Fix-MaskC1 at R=5. Image quality was significantly degraded ($p < 0.003$) from CNN-Fix-MaskC1 to CNN-Fix-MaskC2 when an unmatched undersampling mask (MaskC2) was used for inference, while SANTIS achieved more consistent performance for both MaskC1 and MaskC2 ($p > 0.05$).

For liver imaging (Table 2), CNN-Fix-MaskR1 and SANTIS also achieved significantly better image quality ($p < 0.001$) than CS-PI for nRMSE and SSIM. However, different from Cartesian imaging, the results of CNN-Fix-MaskR2 was only marginally ($p = 0.01$ for nRMSE and SSIM; $p = 0.03$ for Tenengrad) degraded compared to CNN-Fix-MaskR1, which was likely due to the more incoherent undersampling behavior of radial sampling. Nevertheless, SANTIS produced significantly better image quality than CNN-Fix-MaskR1 for SSIM ($p = 0.01$) and Tenengrad ($p = 0.001$).

For both knee and liver imaging, although SANTIS was better than CNN-Fix in general, there was no significant difference between SANTIS and CNN-Fix-MaskC1 for SSIM ($p = 0.1$) in knee datasets at R=3 and there was no significant difference between SANTIS and CNN-Fix-MaskR1 for nRMSE ($p = 0.07$) in liver imaging. These observations indicated that CNN-Fix can provide reconstruction quality that is statistically comparable to SANTIS at certain aspects. Regarding reconstruction time, deep learning-based methods provided faster reconstruction speed compared to CS-PI. Both SANTIS and CNN-Fix reconstruction could be completed in approximately 60ms per image slice.

Figure 8 demonstrates the performance of SANTIS for reconstructing accelerated dynamic contrast-enhanced liver images at different contrast phases. Although the SANTIS training was performed using images acquired in a delayed contrast phase, the method was able to effectively reconstruct images in a different contrast phases (e.g., pre-contrast phase, arterial phase, or venous phase) without affecting image contrast. The recovery of image quality from the zero-filling results could be clearly noted.

DISCUSSION

The main contribution of this work is a new training strategy for improving the robustness of a trained network against discrepancies of undersampling schemes that may occur between training and inference. The performance of SANTIS was demonstrated for accelerating Cartesian-based knee imaging and radial-based liver imaging by comparing with fully sampled references images and different reference reconstruction approaches. Utilizing a network architecture combining end-to-end CNN mapping, data fidelity enforcement, and adversarial training using GAN, the SANTIS approach additionally performs an extensive variation of undersampling patterns during training, referred to as sampling or k -space trajectory augmentation. Such a training strategy allows the network to explore a variety of aliasing artifact structures during the training process, thus providing more effective and robust removal of undersampling artifacts during reconstruction compared to network trained with a fixed undersampling pattern, as demonstrated for the accelerated knee and liver imaging in this work. Moreover, the SANTIS framework might be particularly useful in imaging applications where a non-repeating sampling scheme is performed. An example of this application scenario is the combination of continuous golden-angle radial sampling scheme(35,52) with SANTIS, in which a trained network can be employed to reconstruct undersampled images acquired with varying patterns in different time points, as demonstrated for the accelerated liver imaging in this study. Meanwhile, the SANTIS framework may also be useful to better handle trajectory variations caused by system imperfections (e.g. eddy current and field inhomogeneity) that are often different from scanner to scanner.

Despite the recent interest in deep learning-based reconstruction, most existing work has focused on improving reconstruction accuracy, applicability, and efficiency (13–21). The robustness of the reconstruction methods, despite its importance, has been less studied but remains an essential component in the emerging era of deep learning-based reconstruction. Here, a general open question is whether a trained network that performs well on one image dataset can be effectively used for reconstructing different image datasets with varying image contrasts, noise levels, imaging hardware, or acquisition protocols. A pilot study by Knoll et al. has investigated this question using a variational network (29). Although their findings were insightful, a practical solution was not provided. The SANTIS framework represents a first attempt to solve this challenging question in deep learning-based reconstruction. Although our current study focused on the robustness of acquisition trajectories, the framework could be potentially extended to handle different types of image contrasts and even different noise levels.

The performance of CNN-Fix and SANTIS was compared for accelerated knee imaging using Cartesian trajectories and accelerated liver imaging using radial trajectories. While SANTIS achieved relatively consistent performance, our results found that the differences between CNN-Fix-MaskC1 and CNN-Fix-MaskC2 were more pronounced than the differences between CNN-Fix-MaskR1 and CNN-Fix-MaskR2. This can be attributed to the fact that radial sampling is more incoherent, resulting in more benign undersampling artifacts. As a result, the change of undersampling mask has less of an influence on radial sampling than Cartesian sampling. Our results also showed that the performance of SANTIS

was better than CNN-Fix for both Cartesian and radial sampling in general. The superior performance of SANTIS may be due to two factors. On the one hand, when the number of training datasets is relatively small, successful training does not ensure good reconstruction performance in the testing datasets. Although datasets were acquired on the same organ, corresponding images likely show substantial variations between patients. Thus, when image features are limited, the network may be over-adapted to the training datasets and may not be well-generalized to the testing datasets. The variation of undersampling patterns in SANTIS creates abundant image features, which allow the trained network to better characterize undersampling artifacts for improved image reconstruction. On the other hand, the variation of the undersampling pattern can introduce additional incoherence, thereby facilitating better recognition of undersampled artifacts from underlying image structure. This is similar to the effect of time-varying incoherent sampling in temporal compressed sensing methods (26) and randomized acquisitions in MR fingerprinting (53). However, it should be noted that CNN-Fix also yielded comparable results to SANTIS at certain aspects, indicating that the sampling augmentation might have preferably more contribution to certain image features than others. It is also anticipated that the effectiveness of the data augmentation in image reconstruction is dependent on the deep learning model complexity, the training accuracy and the data size. Further investigations to better understand the comprehensive relationship of the reconstruction improvement via sampling augmentation and penitential affecting factors are needed in future work.

This study also investigated whether the SANTIS framework could be used for reconstructing liver images acquired in different contrast phases during dynamic contrast-enhanced imaging. Our proof-of-concept investigation on a single patient showed encouraging results for reconstructing a rapid contrast wash-in phase (e.g., arterial phase) image using image features learned from a delayed phase image. This suggests the potential to reconstruct accelerated dynamic contrast-enhanced images using transfer learning. Our example showed that SANTIS is capable of learning essential image features that are uniquely dependent on the undersampling pattern with the reconstruction less affected by the change in image contrast caused by contrast injection. In addition to the sampling-augmented training, such performance is also partly due to the use of a combination of residual learning and data fidelity loss. The data fidelity consistency ensures that the reconstructed image matches the acquired undersampled measurement. The residual learning strategy aims to directly subtract image artifacts from undersampled input images, which is advantageous in identifying artifacts without affecting image contrast and anatomy (38,39). However, additional studies on larger image datasets are needed to more thoroughly investigate the applicability and utility of this intriguing feature of SANTIS.

The application of adversarial training has proven to be quite successful in preserving perceptual image structure and texture in reconstructed images, as shown in (14,15). In contrast to the reconstruction methods using solely pixel-wise losses that may alter underlying image texture and noise spectrum, the reconstructed images using adversarial loss is more advantageous for maintaining a visually plausible image appearance. The adversarial learning ensures that the restored images fall into the same data distribution as the fully sampled reference images in a transformed high dimensional image manifold (32,54). The data cycle consistency enforces data fidelity and further prevents degradation of

the adversarial process from generating hallucinated image features (33). The combination of these two components in SANTIS imposes both high-quality image feature learning and information consistency to provide efficient and accurate CNN end-to-end mapping. Although the data cycle-consistent network implemented in the current study is similar to that in our previous report (30), the intriguing contribution of this work is the idea of SANTIS, which is expected to be a generalizable framework that can be combined with other network structures. Moreover, our experiments were performed on multi-coil k -space data in both Cartesian and golden-angle radial trajectories in the context of parallel imaging, which is also different from prior works that have used adversarial training on simulated single-coil k -space datasets. Further investigation will also focus on incorporating dedicated network structures (such as deep cascade network(17), KIKI-net(20) and RAKI(21)) into the data cycle-consistent network to maximize the benefits of both the database feature learning and case-specific feature estimation with multi-coil data.

Our study has several limitations. First, although the SANTIS training was performed using a library of different undersampling patterns, they were generated to have the same variable-density distribution in Cartesian imaging. The robustness of the trained network for reconstructing a completely different sampling scheme remains to be explored. However, it is straightforward to extend the SANTIS training to incorporate different undersampling schemes with varying distributions of sampling and different acceleration rates, so that the generalization of the network can be further extended. However, such a training scheme may require an increased number of iterations to achieve convergence. Second, our study did not analyze the speed of training convergence between CNN-Fix and SANTIS. For both networks, the training was performed with a sufficient number of epochs to ensure training convergence. Thus, further studies are needed to evaluate the influence of sampling-augmented training and training parameters on convergence speed. Third, the reconstruction parameters, as shown in Eq. [12], were empirically determined in our study. The weight of the GAN regularization, in particular, is vital to ensure good reconstruction performance (14,15). However, the selection of reconstruction parameters is not only a challenge for our work but remains an open question in constrained image reconstruction in general. Finally, due to the need for large training datasets, the SANTIS approach was evaluated on a relatively small number of knee and liver image testing datasets and only a single dynamic contrast-enhanced liver testing dataset. Additional studies with larger testing datasets are needed to further evaluate SANTIS before the approach can be used in clinical practice to accelerate knee and liver imaging.

CONCLUSIONS

SANTIS represents a novel concept for deep learning-based image reconstruction. With sampling or k -space trajectory-augmented training by extensively varying undersampling pattern, SANTIS can recognize more aliasing artifact structures and thereby provide more robust and effective removal of image artifacts. The SANTIS framework can be particularly useful for improving the robustness of deep learning-based image reconstruction against discrepancies of undersampling schemes that may occur between training and inference.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENT

This work was supported by National Institutes of Health (grant R01AR068373 and R01EB027087). The authors thank the technologists at the University of Wisconsin Hospital in Madison, USA for acquiring the knee image datasets and the technologists at the Southwest Hospital in Chongqing, China for acquiring the liver image datasets with institutional IRB approval. The authors thank Dr. Julia Velikina for discussion about compressed sensing and parallel imaging reconstruction.

REFERENCES

1. Lakhani P, Sundaram B. Deep Learning at Chest Radiography: Automated Classification of Pulmonary Tuberculosis by Using Convolutional Neural Networks. *Radiology* [Internet] 2017;284:574–582. doi: 10.1148/radiol.2017162326.
2. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* [Internet] 2017;542:115–118. doi: 10.1038/nature21056.
3. Liu F, Zhou Z, Jang H, Samsonov A, Zhao G, Kijowski R. Deep Convolutional Neural Network and 3D Deformable Approach for Tissue Segmentation in Musculoskeletal Magnetic Resonance Imaging. *Magn. Reson. Med.* [Internet] 2017;DOI: 10.1002/mrm.26841. doi: 10.1002/mrm.26841.
4. Zhou Z, Zhao G, Kijowski R, Liu F. Deep Convolutional Neural Network for Segmentation of Knee Joint Anatomy. *Magn. Reson. Med.* 2018;doi:10.1002/mrm.27229.
5. Liu F SUSAN: Segment Unannotated image Structure using Adversarial Network. *Magn Reson Med* [Internet] 2018;DOI: 10.1002/mrm.27627. doi: 10.1002/mrm.27627.
6. Norman B, Padoia V, Majumdar S. Use of 2D U-Net Convolutional Neural Networks for Automated Cartilage and Meniscus Segmentation of Knee MR Imaging Data to Determine Relaxometry and Morphometry. *Radiology* [Internet] 2018;172322. doi: 10.1148/radiol.2018172322.
7. Zhao G, Liu F, Oler JA, Meyerand ME, Kalin NH, Birn RM. Bayesian convolutional neural network based MRI brain extraction on nonhuman primates. *Neuroimage* [Internet] 2018;175:32–44. doi: 10.1016/j.neuroimage.2018.03.065.
8. Zha W, Fain SB, Schiebler ML, Evans MD, Nagle SK, Liu F. Deep Convolutional Neural Networks With Multiplane Consensus Labeling for Lung Function Quantification Using UTE Proton MRI. *J. Magn. Reson. Imaging* [Internet] 2019;jmri.26734. doi: 10.1002/jmri.26734.
9. Liu F, Zhou Z, Samsonov A, Blankenbaker D, Larison W, Kanarek A, Lian K, Kambhampati S, Kijowski R. Deep Learning Approach for Evaluating Knee MR Images: Achieving High Diagnostic Performance for Cartilage Lesion Detection. *Radiology* [Internet] 2018;172986. doi: 10.1148/radiol.2018172986.
10. Liu F, Guan B, Zhou Z, et al. Fully-Automated Diagnosis of Anterior Cruciate Ligament Tears on Knee MR Images Using Deep Learning. *Radiol. Artif. Intell.* 2019;In-press.
11. Anthimopoulos M, Christodoulidis S, Ebner L, Christe A, Mougiakakou S. Lung Pattern Classification for Interstitial Lung Diseases Using a Deep Convolutional Neural Network. *IEEE Trans. Med. Imaging* [Internet] 2016;35:1207–1216. doi: 10.1109/TMI.2016.2535865.
12. Kooi T, Litjens G, van Ginneken B, Gubern-Mérida A, Sánchez CI, Mann R, den Heeten A, Karssemeijer N. Large scale deep learning for computer aided detection of mammographic lesions. *Med. Image Anal.* [Internet] 2017;35:303–312. doi: 10.1016/j.media.2016.07.007.
13. Hammernik K, Klatzer T, Kobler E, Recht MP, Sodickson DK, Pock T, Knoll F. Learning a Variational Network for Reconstruction of Accelerated MRI Data. *Magn. Reson. Med.* [Internet] 2017;79:3055–3071. doi: 10.1002/mrm.26977.
14. Quan TM, Nguyen-Duc T, Jeong W-K. Compressed Sensing MRI Reconstruction using a Generative Adversarial Network with a Cyclic Loss. *IEEE Trans. Med. Imaging* [Internet] 2017;37:1488–1497. doi: 10.1109/TMI.2018.2820120.

15. Mardani M, Gong E, Cheng JY, Vasanawala SS, Zaharchuk G, Xing L, Pauly JM. Deep Generative Adversarial Neural Networks for Compressive Sensing (GANCS) MRI. *IEEE Trans. Med. Imaging* [Internet] 2018;1–1. doi: 10.1109/TMI.2018.2858752.
16. Wang S, Su Z, Ying L, Peng X, Zhu S, Liang F, Feng D, Liang D, Technologies I. ACCELERATING MAGNETIC RESONANCE IMAGING VIA DEEP LEARNING. In: 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI) IEEE; 2016 pp. 514–517. doi: 10.1109/ISBI.2016.7493320.
17. Schlemper J, Caballero J, Hajnal JV., Price A, Rueckert D A Deep Cascade of Convolutional Neural Networks for Dynamic MR Image Reconstruction. *IEEE Trans. Med. Imaging* [Internet] 2017;1–1. doi: 10.1007/978-3-319-59050-9_51.
18. Han Y, Yoo J, Kim HH, Shin HJ, Sung K, Ye JC. Deep learning with domain adaptation for accelerated projection-reconstruction MR. *Magn. Reson. Med.* [Internet] 2018;80:1189–1205. doi: 10.1002/mrm.27106.
19. Zhu B, Liu JZ, Cauley SF, Rosen BR, Rosen MS. Image reconstruction by domain-transform manifold learning. *Nature* [Internet] 2018;555:487–492. doi: 10.1038/nature25988.
20. Eo T, Jun Y, Kim T, Jang J, Lee HJ, Hwang D. KIKI-net: Cross-domain convolutional neural networks for reconstructing undersampled magnetic resonance images. *Magn. Reson. Med.* [Internet] 2018. doi: 10.1002/mrm.27201.
21. Akçakaya M, Moeller S, Weingärtner S, Urbil K. Scan-specific robust artificial-neural-networks for k-space interpolation (RAKI) reconstruction: Database-free deep learning for fast imaging. *Magn. Reson. Med.* [Internet] 2019;81:439–453. doi: 10.1002/mrm.27420.
22. Liu F, Feng L, Kijowski R. MANTIS: Model-Augmented Neural network with Incoherent k-space Sampling for efficient MR parameter mapping. *Magn. Reson. Med.* [Internet] 2019. doi: 10.1002/mrm.27707.
23. Sodickson DK, Manning WJ. Simultaneous acquisition of spatial harmonics (SMASH): fast imaging with radiofrequency coil arrays. *Magn. Reson. Med.* [Internet] 1997;38:591–603.
24. Griswold MA, Jakob PM, Heidemann RM, Nittka M, Jellus V, Wang J, Kiefer B, Haase A. Generalized Autocalibrating Partially Parallel Acquisitions (GRAPPA). *Magn. Reson. Med.* 2002;47:1202–1210. doi: 10.1002/mrm.10171. [PubMed: 12111967]
25. Pruessmann KP, Weiger M, Scheidegger MB, Boesiger P. SENSE: Sensitivity Encoding for Fast MRI. *Magn Reson Med* [Internet] 1999;42:952–962.
26. Lustig M, Donoho D, Pauly JM. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magn Reson Med* [Internet] 2007;58:1182–1195. doi: 10.1002/mrm.21391.
27. Lustig M, Pauly JM. SPIRiT: Iterative self-consistent parallel imaging reconstruction from arbitrary k-space. *Magn. Reson. Med.* [Internet] 2010;64:457–471. doi: 10.1002/mrm.22428.
28. Otazo R, Kim D, Axel L, Sodickson DK. Combination of compressed sensing and parallel imaging for highly accelerated first-pass cardiac perfusion MRI. *Magn. Reson. Med.* [Internet] 2010;64:767–776. doi: 10.1002/mrm.22463.
29. Knoll F, Hammernik K, Kobler E, Pock T, Recht MP, Sodickson DK. Assessment of the generalization of learned image reconstruction and the potential for transfer learning. *Magn. Reson. Med.* [Internet] 2018. doi: 10.1002/mrm.27355.
30. Liu F, Samsonov A. Data-Cycle-Consistent Adversarial Networks for High-Quality Reconstruction of Undersampled MRI Data. In: the ISMRM Machine Learning Workshop ; 2018.
31. Yang G, Yu S, Dong H, et al. DAGAN: Deep De-Aliasing Generative Adversarial Networks for Fast Compressed Sensing MRI Reconstruction. *IEEE Trans. Med. Imaging* [Internet] 2018;37:1310–1321. doi: 10.1109/TMI.2017.2785879.
32. Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Generative Adversarial Networks. *arXiv Prepr. arXiv ...* [Internet] 2014:1–9. doi: 10.1001/jamainternmed.2016.8245.
33. Zhu J-Y, Park T, Isola P, Efros AA. Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. In: 2017 IEEE International Conference on Computer Vision (ICCV) Vol. 2017–Octob. IEEE; 2017 pp. 2242–2251. doi: 10.1109/ICCV.2017.244.
34. Isola P, Zhu J-Y, Zhou T, Efros AA. Image-to-Image Translation with Conditional Adversarial Networks. *ArXiv e-prints* [Internet] 2016.

35. Feng L, Grimm R, Block KT, Chandarana H, Kim S, Xu J, Axel L, Sodickson DK, Otazo R. Golden-angle radial sparse parallel MRI: Combination of compressed sensing, parallel imaging, and golden-angle radial sampling for fast and flexible dynamic volumetric MRI. *Magn. Reson. Med.* [Internet] 2014;72:707–717. doi: 10.1002/mrm.24980.
36. Chandarana H, Feng L, Block TK, Rosenkrantz AB, Lim RP, Babb JS, Sodickson DK, Otazo R. Free-Breathing Contrast-Enhanced Multiphase MRI of the Liver Using a Combination of Compressed Sensing, Parallel Imaging, and Golden-Angle Radial Sampling. *Invest. Radiol.* [Internet] 2013;48:10–16. doi: 10.1097/RLI.0b013e318271869c.
37. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab N, Hornegger J, Wells WM, Frangi AF, editors. *Medical Image Computing and Computer-Assisted Intervention -- MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III*. Cham: Springer International Publishing; 2015 pp. 234–241. doi: 10.1007/978-3-319-24574-4_28.
38. Gong E, Pauly JM, Wintermark M, Zaharchuk G. Deep learning enables reduced gadolinium dose for contrast-enhanced brain MRI. *J. Magn. Reson. Imaging* [Internet] 2018. doi: 10.1002/jmri.25970.
39. Han YS, Yoo J, Ye JC. Deep Residual Learning for Compressed Sensing CT Reconstruction via Persistent Homology Analysis. *arXiv* [Internet] 2016.
40. Ledig C, Theis L, Huszar F, et al. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. *ArXiv e-prints* [Internet] 2016.
41. Li C, Wand M. Precomputed Real-Time Texture Synthesis with Markovian Generative Adversarial Networks. *ArXiv e-prints* [Internet] 2016.
42. He K, Zhang X, Ren S, Sun J. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. *ArXiv e-prints* [Internet] 2015;1502.
43. Kingma DP, Ba J. Adam: A Method for Stochastic Optimization. *ArXiv e-prints* [Internet] 2014.
44. Chollet François. Keras. GitHub 2015:<https://github.com/fchollet/keras>.
45. Abadi M, Agarwal A, Barham P, et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. *ArXiv e-prints* [Internet] 2016. doi: 10.1109/TIP.2003.819861.
46. Walsh DO, Gmitro AF, Marcellin MW. Adaptive reconstruction of phased array MR imagery. *Magn. Reson. Med.* [Internet] 2000;43:682–690. doi: 10.1002/(SICI)1522-2594(200005)43:5<682::AID-MRM10>3.0.CO;2-G.
47. Seiberlich N, Breuer F, Blaimer M, Jakob P, Griswold M. Self-calibrating GRAPPA operator gridding for radial and spiral trajectories. *Magn. Reson. Med.* [Internet] 2008;59:930–935. doi: 10.1002/mrm.21565.
48. Benkert T, Tian Y, Huang C, DiBella EVR, Chandarana H, Feng L. Optimization and validation of accelerated golden-angle radial sparse MRI reconstruction with self-calibrating GRAPPA operator gridding. *Magn. Reson. Med.* [Internet] 2018;80:286–293. doi: 10.1002/mrm.27030.
49. Liu Bo, Yi Ming Zou, Leslie Ying. Sparsesense: Application of compressed sensing in parallel MRI. In: *2008 International Conference on Technology and Applications in Biomedicine IEEE*; 2008 pp. 127–130. doi: 10.1109/ITAB.2008.4570588.
50. Krotkov EP. *Active Computer Vision by Cooperative Focus and Stereo*. New York, NY: Springer New York; 1989. doi: 10.1007/978-1-4613-9663-5.
51. Feng L, Huang C, Shanbhogue K, Sodickson DK, Chandarana H, Otazo R. RACER-GRASP: Respiratory-weighted, aortic contrast enhancement-guided and coil-unstreaking golden-angle radial sparse MRI. *Magn. Reson. Med.* [Internet] 2018;80:77–89. doi: 10.1002/mrm.27002.
52. Feng L, Axel L, Chandarana H, Block KT, Sodickson DK, Otazo R. XD-GRASP: Golden-angle radial MRI with reconstruction of extra motion-state dimensions using compressed sensing. *Magn. Reson. Med.* [Internet] 2016;75:775–788. doi: 10.1002/mrm.25665.
53. Ma D, Gulani V, Seiberlich N, Liu K, Sunshine JL, Duerk JL, Griswold MA. Magnetic resonance fingerprinting. *Nature* [Internet] 2013;495:187–192. doi: 10.1038/nature11971.
54. Ye JC, Han Y, Cha E. Deep Convolutional Framelets: A General Deep Learning Framework for Inverse Problems. *ArXiv e-prints* [Internet] 2017.

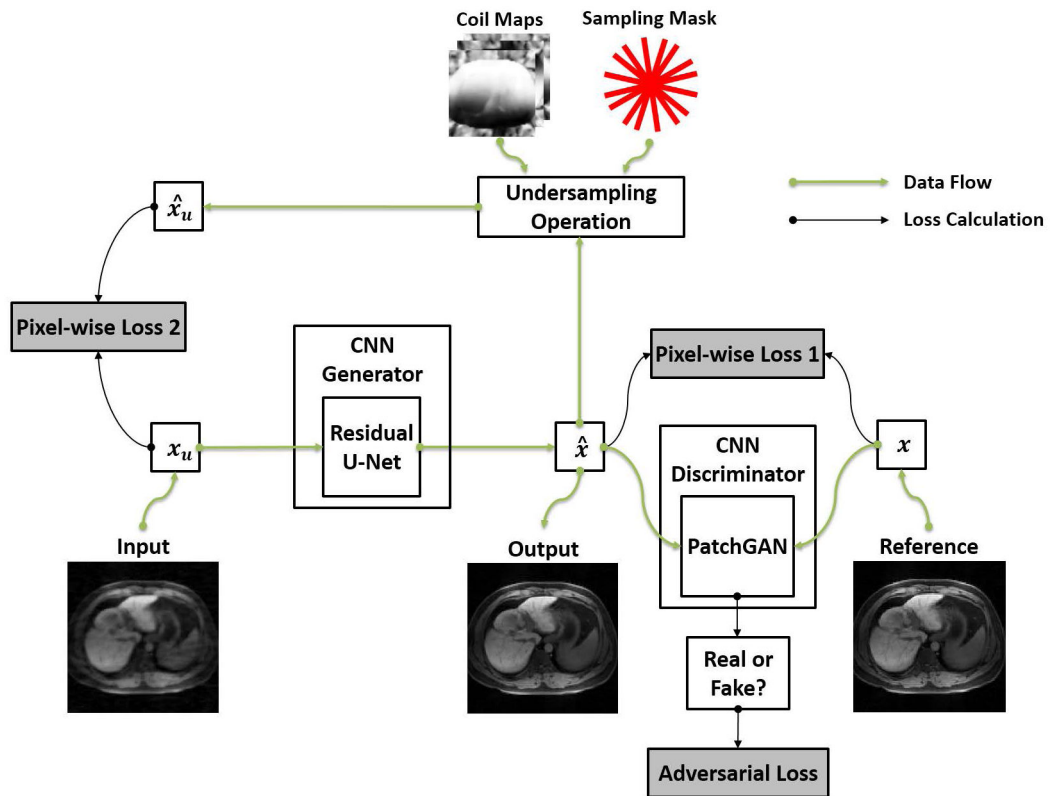


Figure 1.

Illustration of the SANTIS reconstruction using a data cycle-consistent adversarial network, which features three loss components. The first loss term (pixel-wise loss 1) removes aliasing artifacts such that the reconstructed image \hat{x} looks like their reference (x). The second loss term (pixel-wise loss 2) ensures that the reconstructed images from CNN generator (e.g., Residual U-Net) produce undersampled images \hat{x}_u matching the acquired k -space measurements (x_u). The adversarial loss term enforces a high perceptual quality for the reconstructed image to maintain image details and texture using a CNN discriminator (e.g., PatchGAN). The notation in this figure follows the main text description.

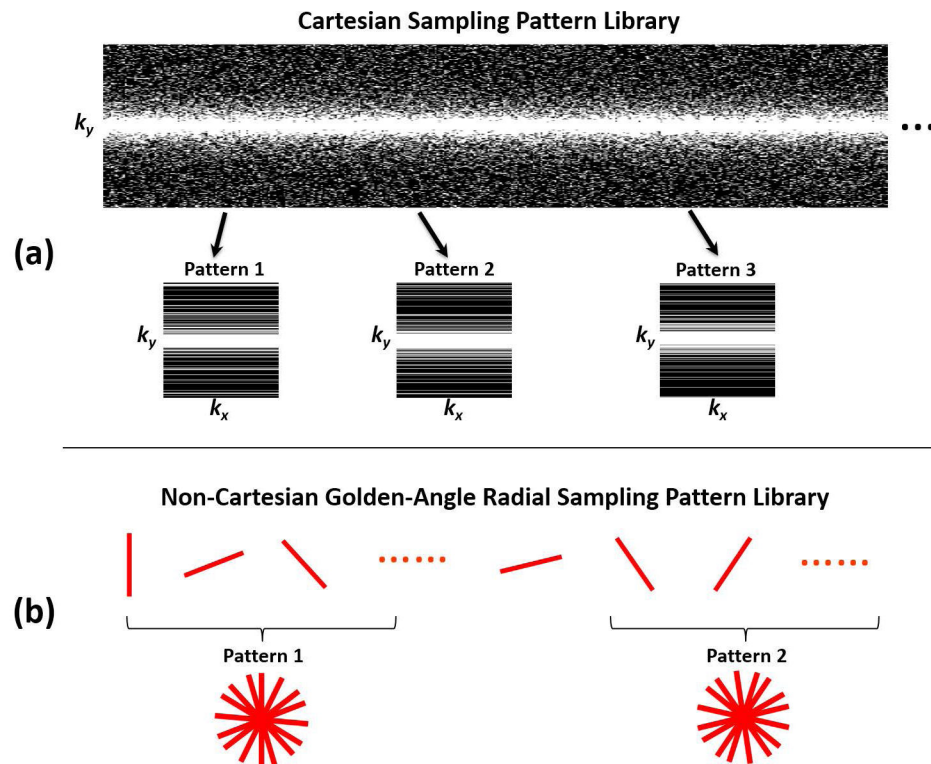


Figure 2. Schematic demonstration of the undersampling patterns used in the study. **(a)** Examples of the 1D variable-density Cartesian random undersampling patterns used for knee imaging. **(b)** Example of the gold-angle radial undersampling masks used for liver imaging. The undersampling mask was varying for each iteration during the network training to augment the training data for SANTIS framework.

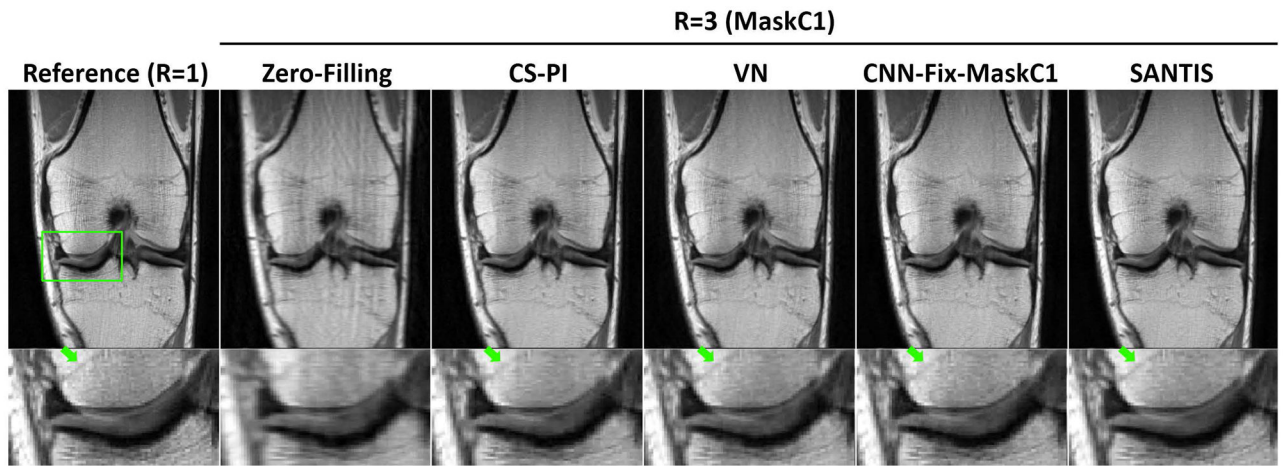


Figure 3. Representative examples of knee images obtained using the different reconstruction methods at R=3. SANTIS showed the highest image quality with favorable preservation of tissue sharpness (green arrows) and texture comparable to the reference.

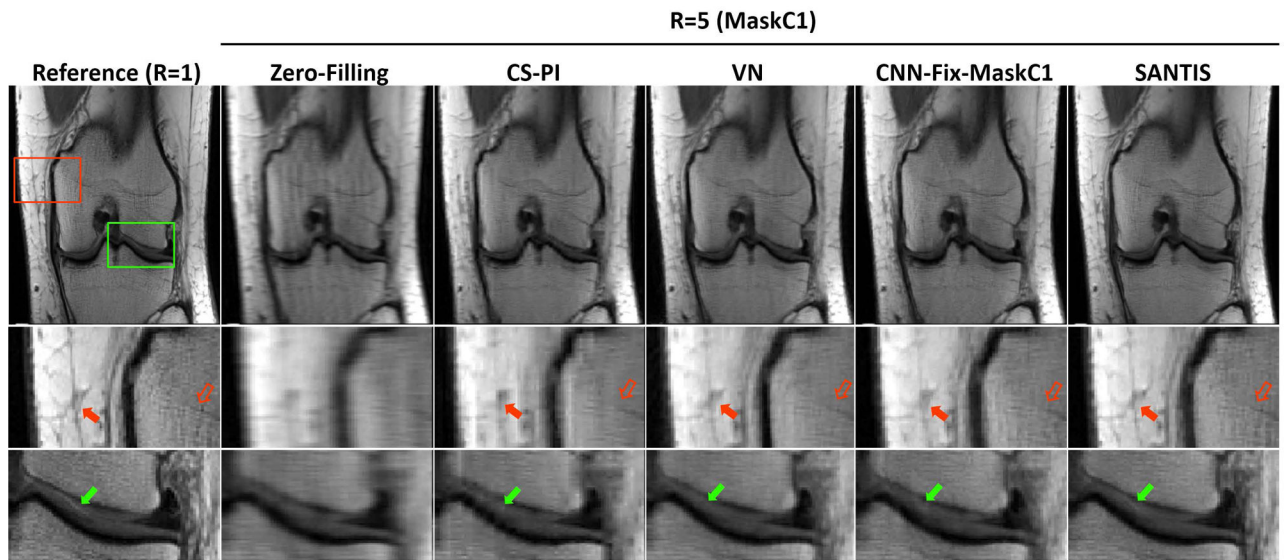


Figure 4. Representative examples of knee images obtained using the different reconstruction methods at R=5. SANTIS showed the highest image quality with a better representation of the layered structure of the femoral and tibial cartilage (green arrows) with favorable preservation of tissue sharpness and texture (red arrows) comparable to the reference.

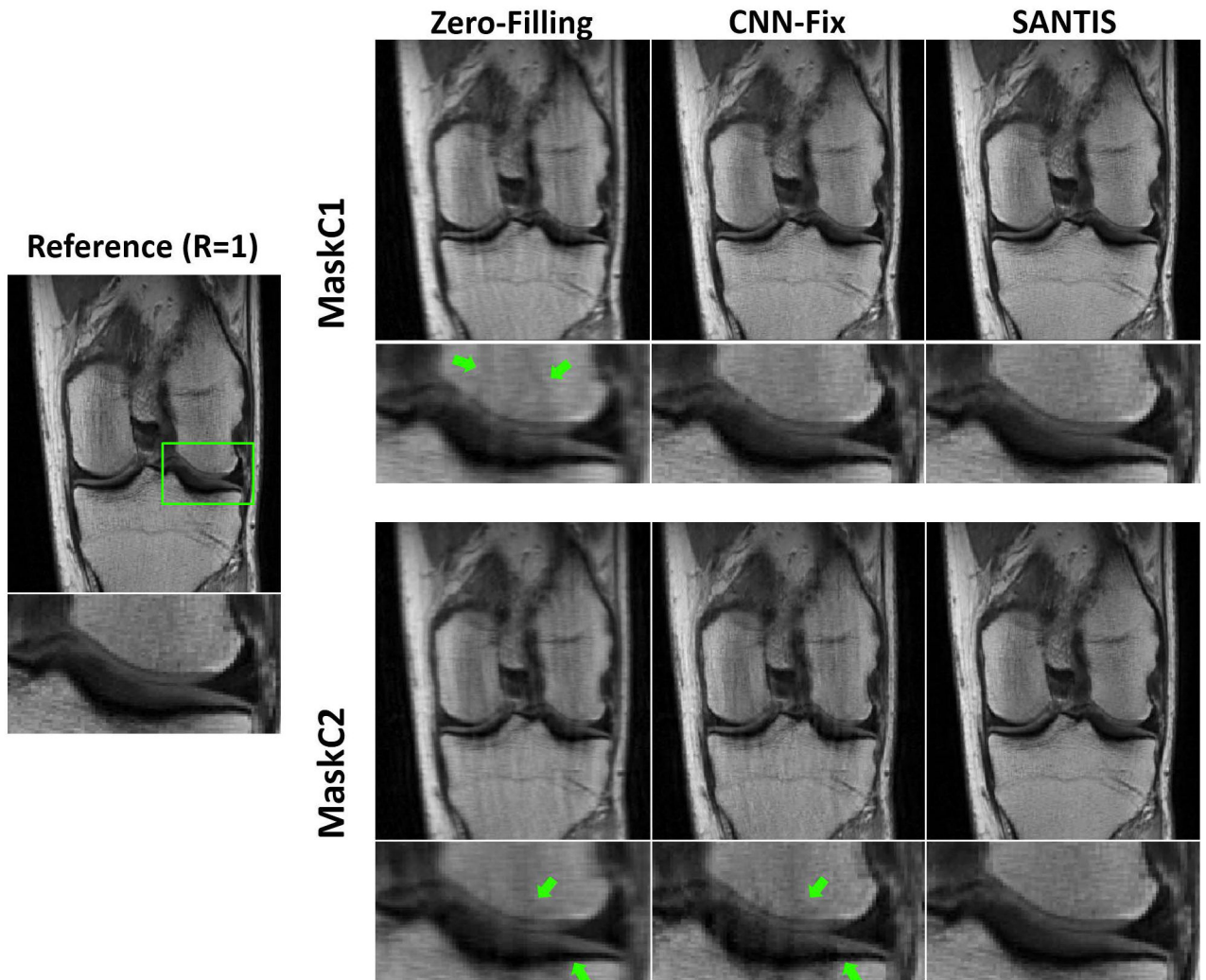


Figure 5. Evaluation of reconstruction robustness at different undersampling masks for MaskC1 and MaskC2. The fixed training CNN-Fix failed to reconstruct MaskC2 with noticeable residual artifacts, as indicated by the green arrows. SANTIS provided a consistent high-quality reconstruction for both masks.

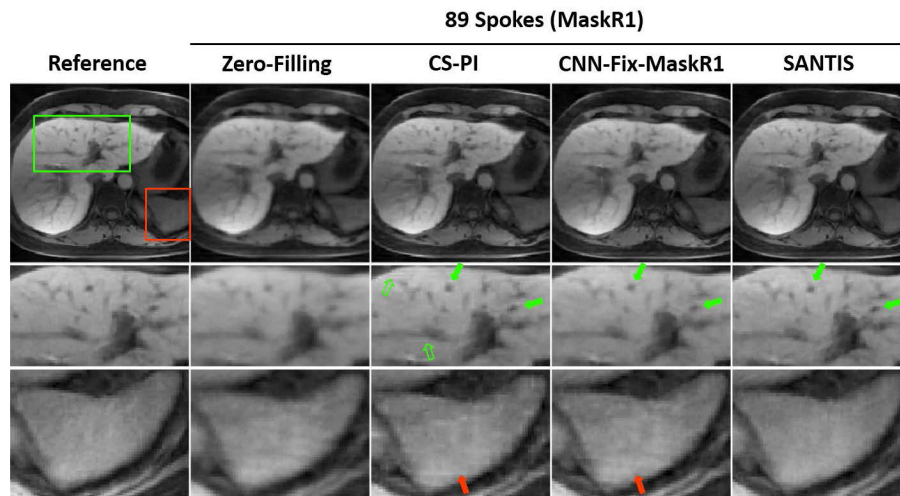


Figure 6.

A representative comparison of different reconstruction methods in radial liver imaging with MaskR1. SANTIS showed the highest image quality with successful removal of streaking artifacts and with favorable preservation of tissue sharpness and texture. CS-PI showed pseudo residual structures (hollow green arrows). Meanwhile, CNN-Fix-MaskR1 suffered from noticeable residual blurring, as indicated by the green arrows. For both CS-PI and CNN-Fix-MaskR1, residual streaking artifacts could be noted, as highlighted by the red arrows in the zoomed region.

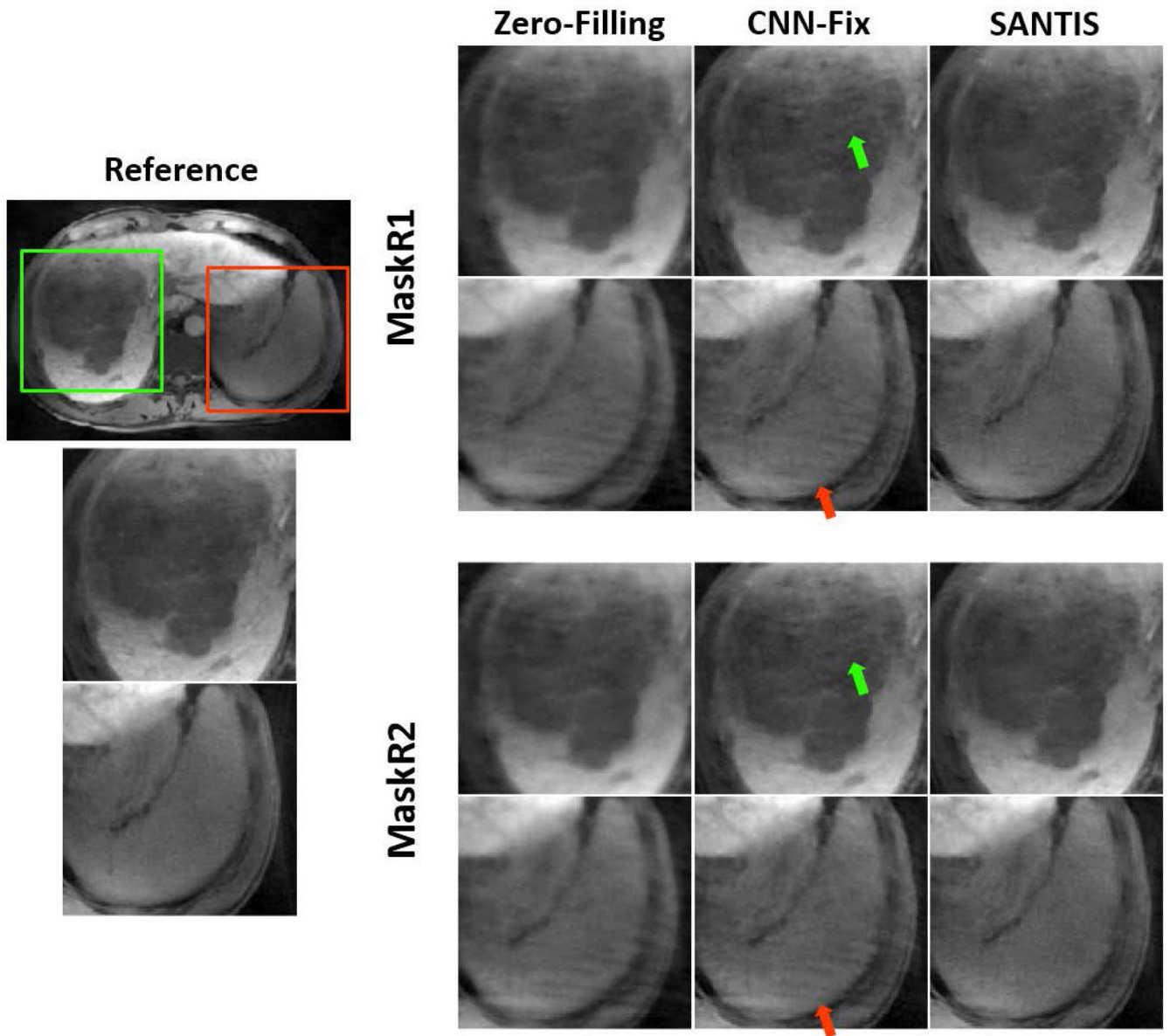


Figure 7. Example of zoomed liver images with a tumor comparing CNN-Fix and SANTIS using both MaskR1 and MaskR2 for inference. Both CNN-Fix-MaskR1 and CNN-Fix-MaskR2 suffered from residual artifacts (red arrows), while SANTIS consistently showed improved image quality for both MaskR1 and MaskR2, with a reduced level of residual streaking artifacts and a delineation of the tumor that is closer to the reference.

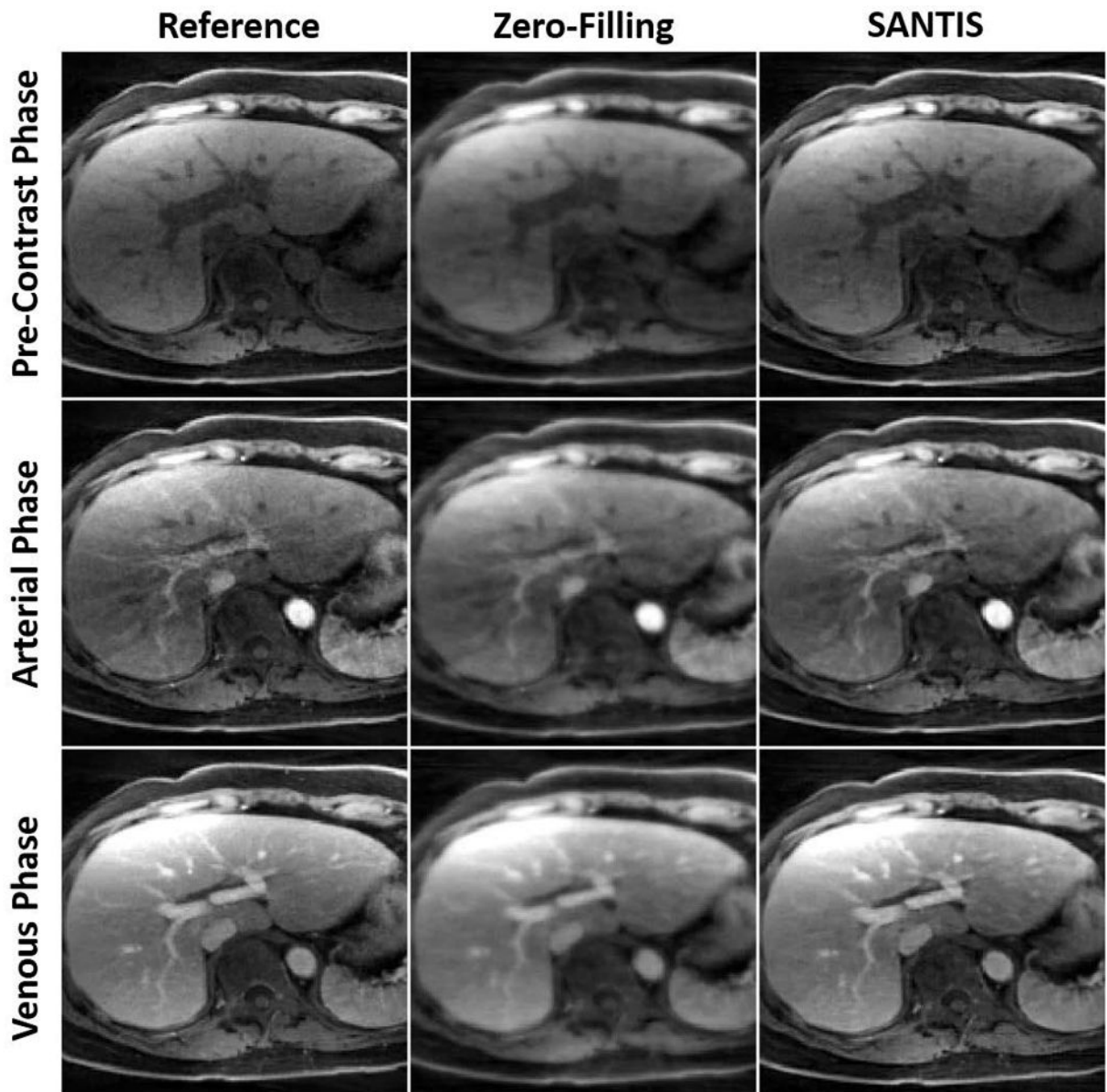


Figure 8.

Examples demonstrating the performance of SANTIS in reconstructing accelerated dynamic contrast-enhanced liver images at different contrast phases. SANTIS effectively reconstructed images at a different contrast phase (e.g., pre-contrast phase, arterial phase, or venous phase) without affecting image contrast, and the recovery of image quality from the zero-filling results could be clearly noted.

Table 1

Quantitative metrics comparing the reference fully sampled images and the reconstructed images from different reconstruction methods for Cartesian knee imaging. Results were averaged over the eight test datasets and represent mean value \pm standard deviation. SANTIS achieved the highest and most robust reconstruction performance at both R=3 and 5.

Method		R=3				R=5			
		nRMSE [%]	SSIM [%]	Tenengrad [%]	Average Recon Time [s]	nRMSE [%]	SSIM [%]	Tenengrad [%]	Average Recon Time [s]
ZF		9.87 \pm 1.16	82.25 \pm 3.77	43.75 \pm 2.69	0.001	13.58 \pm 1.75	75.38 \pm 4.92	55.19 \pm 3.52	0.001
CS-PI		6.74 \pm 0.63	87.64 \pm 2.22	15.66 \pm 1.99	2.12	10.91 \pm 1.13	83.78 \pm 3.80	33.01 \pm 2.99	2.21
VN		6.01 \pm 1.12	90.21 \pm 2.54	10.11 \pm 2.06	0.11	10.61 \pm 1.67	85.11 \pm 3.96	20.13 \pm 3.10	0.12
CNN-Fix	MaskC1 (match training)	5.67 \pm 0.94	90.78 \pm 2.89	10.30 \pm 2.34	0.06	10.48 \pm 1.41	85.06 \pm 4.79	19.63 \pm 4.10	0.06
	MaskC2 (unmatch training)	6.92 \pm 1.09	86.33 \pm 2.91	10.67 \pm 2.86	0.06	11.56 \pm 1.94	82.98 \pm 4.84	20.59 \pm 4.95	0.06
SANTIS	MaskC1	5.15 \pm 1.07	91.41 \pm 2.82	8.67 \pm 2.45	0.06	9.60 \pm 1.81	89.06 \pm 4.97	17.29 \pm 3.79	0.06
	MaskC2	5.10 \pm 1.02	91.96 \pm 2.90	8.51 \pm 2.38	0.06	9.66 \pm 1.40	88.66 \pm 4.62	17.11 \pm 3.84	0.06

Table 2

Quantitative metrics comparing the reference fully sampled images and the reconstructed images from different reconstruction methods for golden-angle radial liver imaging. Results were averaged over the eight test datasets and represent mean value \pm standard deviation. SANTIS achieved the highest and most robust reconstruction performance.

Method		89 spokes			
		nRMSE [%]	SSIM [%]	Tenengrad [%]	Average Recon time [s]
ZF		11.69 \pm 1.39	85.80 \pm 2.38	37.98 \pm 2.92	0.001
CS-PI		9.50 \pm 1.56	87.42 \pm 3.14	3.55 \pm 2.75	2.32
CNN-Fix	MaskR1 (match training)	9.00 \pm 1.32	91.07 \pm 2.27	6.16 \pm 2.33	0.06
	MaskR2 (unmatch training)	9.29 \pm 1.32	89.80 \pm 2.24	6.76 \pm 2.43	0.06
SANTIS	MaskR1	8.82 \pm 1.31	92.66 \pm 2.00	2.00 \pm 2.72	0.06
	MaskR2	8.78 \pm 1.29	92.17 \pm 2.00	1.77 \pm 3.19	0.06