



OPEN ACCESS

EDITED BY

Qiang Zhang,
University of Oxford, United Kingdom

REVIEWED BY

Ye Tian,
University of Southern California,
United States
Jesse Hamilton,
University of Michigan, United States

*CORRESPONDENCE

Chenxi Hu
✉ chenxi.hu@sjtu.edu.cn

RECEIVED 28 March 2024

ACCEPTED 27 August 2024

PUBLISHED 12 September 2024

CITATION

Xue Z, Zhu S, Yang F, Gao J, Peng H, Zou C,
Jin H and Hu C (2024) A hybrid deep image
prior and compressed sensing reconstruction
method for highly accelerated 3D coronary
magnetic resonance angiography.
Front. Cardiovasc. Med. 11:1408351.
doi: 10.3389/fcvm.2024.1408351

COPYRIGHT

© 2024 Xue, Zhu, Yang, Gao, Peng, Zou, Jin
and Hu. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

A hybrid deep image prior and compressed sensing reconstruction method for highly accelerated 3D coronary magnetic resonance angiography

Zhihao Xue¹, Sicheng Zhu¹, Fan Yang¹, Juan Gao¹, Hao Peng²,
Chao Zou², Hang Jin³ and Chenxi Hu^{1*}

¹National Engineering Research Center of Advanced Magnetic Resonance Technologies for Diagnosis and Therapy, School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China,

²Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, Guangdong, China, ³Department of Radiology, Zhongshan Hospital, Fudan University and Shanghai Medical Imaging Institute, Shanghai, China

Introduction: High-resolution whole-heart coronary magnetic resonance angiography (CMRA) often suffers from unreasonably long scan times, rendering imaging acceleration highly desirable. Traditional reconstruction methods used in CMRA rely on either hand-crafted priors or supervised learning models. Although the latter often yield superior reconstruction quality, they require a large amount of training data and memory resources, and may encounter generalization issues when dealing with out-of-distribution datasets.

Methods: To address these challenges, we introduce an unsupervised reconstruction method that combines deep image prior (DIP) with compressed sensing (CS) to accelerate 3D CMRA. This method incorporates a slice-by-slice DIP reconstruction and 3D total variation (TV) regularization, enabling high-quality reconstruction under a significant acceleration while enforcing continuity in the slice direction. We evaluated our method by comparing it to iterative SENSE, CS-TV, CS-wavelet, and other DIP-based variants, using both retrospectively and prospectively undersampled datasets.

Results: The results demonstrate the superiority of our 3D DIP-CS approach, which improved the reconstruction accuracy relative to the other approaches across both datasets. Ablation studies further reveal the benefits of combining DIP with 3D TV regularization, which leads to significant improvements of image quality over pure DIP-based methods. Evaluation of vessel sharpness and image quality scores shows that DIP-CS improves the quality of reformatted coronary arteries.

Discussion: The proposed method enables scan-specific reconstruction of high-quality 3D CMRA from a five-minute acquisition, without relying on fully-sampled training data or placing a heavy burden on memory resources.

KEYWORDS

coronary magnetic resonance angiography, deep image prior, compressed sensing, unsupervised learning, image reconstruction

1 Introduction

Three-dimensional coronary magnetic resonance angiography (CMRA) is a promising imaging modality for the assessment of coronary artery disease (CAD) due to its non-invasiveness and freedom from ionizing radiation (1–3). Despite these benefits, CMRA still suffers from a long scan time, because of the need to achieve millimeter-level 3D

whole-heart imaging. To address this issue, highly undersampled acquisition is often employed to accelerate CMRA, with acceleration factors ranging from 7 to 10 (4). These high acceleration rates can cause serious degradations of image quality due to the incurrance of aliasing artifacts and noise amplification in the image. Therefore, reconstruction algorithms that minimize these artifacts and noise at high acceleration rates are highly desirable.

Various reconstruction methods have been developed to improve image quality for accelerated MRI, and some have been applied to the CMRA reconstruction. Early methods are often based on the compressed sensing theory. These methods leverage general image properties to accelerate imaging, such as sparsity in the wavelet domain (5), finite total variation (6), or low-rankness of signals extracted from non-local patches (7). Some of these methods have achieved reconstructions of adequate quality with a scan time of 5–7 min (4, 8, 9). Recently, learning-based methods have emerged as a new genre of reconstruction methods for fast MRI. These methods can learn image priors from existing data and have shown the state-of-the-art performance in various reconstruction tasks. Among these methods, some are purely data-driven, such as AUTOMAP (10) and U-Net (11), while others are partly model-driven, which include deep unrolling (12–15), plug-and-play (16, 17), and methods based on learned explicit priors (18, 19). Usually, model-driven methods are more robust and generalizable than purely data-driven methods. However, despite an improved generalizability for the model-driven methods, a major challenge remains: they often require a substantial amount of training data, which can be scarce or unavailable in various scenarios, including 3D CMRA (20). Additionally, performance degradation due to distribution shifts between training and testing datasets is not negligible. For instance, several studies have found performance degradation of deep learning-based methods in reconstructions of out-of-distribution data (21, 22). Finally, certain methods, such as deep unrolling, are challenging to use in 3D imaging due to the high memory demand of network training. Given these limitations of supervised methods, there is a growing demand for database-free, self-supervised, or unsupervised reconstruction approaches.

Deep imaging prior (DIP) (23) has emerged as a powerful unsupervised technique for solving inverse problems in image processing, including denoising, super-resolution, and inpainting. DIP does not need external training data; instead, DIP optimizes the parameters of a randomly initialized deep neural network from scratch for each reconstruction. During training, the network learns to map a vector of noise to the reconstructed image, whose regularization is attributed to the inductive bias captured by the network architecture. Recently, several DIP-based methods have been proposed for MRI reconstruction, exhibiting an improvement of image quality when compared with other unsupervised methods (24–29). Compared with supervised methods, DIP does not require any training data, reduces the memory requirement compared with deep unrolling, and induces less concern on its generalizability under different undersampling patterns or anatomies (24, 25). Recently, the application of DIP

has been extended to cardiac imaging, including cardiac magnetic resonance fingerprinting (MRF) (27), functional CMR (28), and 2D + t dynamic MRI (30). Some research has also investigated the combination of DIP and classical regularization methods within an iterative optimization framework (31, 32).

Although DIP has been used in 2D imaging, no prior study has utilized DIP for 3D CMRA reconstruction. Thus, it remains unclear whether 2D DIP can generalize well to 3D reconstruction in terms of reconstruction quality and memory usage. In this work, we fill this gap by proposing a novel hybrid DIP-CS method for 3D CMRA reconstruction. This approach combines the unsupervised DIP model and the total variation (TV) regularization within a framework that uses the alternating direction method of multipliers (ADMM). We compared the proposed 3D DIP-CS method with several alternative methods, including iterative SENSE, CS-wavelet, CS-TV, 2D slice-by-slice DIP, and 2D slice-by-slice DIP equipped with 2D TV regularization with both retrospective and prospective undersampling experiments.

2 Theory and methods

2.1 Compressed sensing MRI and deep image prior

The reconstruction of 3D MRI can be considered as an inverse problem in the following format:

$$\mathbf{m} = \mathbf{A}\mathbf{f} + \boldsymbol{\epsilon} = \mathbf{DFS}\mathbf{f} + \boldsymbol{\epsilon} \quad (1)$$

Where $\mathbf{f} \in \mathbb{C}^N$ is the 3D image, $\mathbf{m} \in \mathbb{C}^M$ is the 3D k-space measurement, $\boldsymbol{\epsilon}$ is the noise, and \mathbf{A} is the matrix representation of the forward imaging model. In parallel imaging, the forward model consists of three linear operators: coil sensitivity encoding \mathbf{S} , the Fourier transform operator \mathbf{F} , and k-space undersampling operator \mathbf{D} . Due to the undersampling, the inverse problem is ill-conditioned. Consequently, regularization is usually required to solve the problem. Many approaches based on compressed sensing theory leverage the sparsity of images in a certain transform domain to improve the conditioning. The objective function of reconstruction is described as follows in Equation 2:

$$\underset{\mathbf{f}}{\operatorname{argmin}} \|\mathbf{m} - \mathbf{A}\mathbf{f}\|_2^2 + \lambda\Phi(\mathbf{f}) \quad (2)$$

where $\|\mathbf{m} - \mathbf{A}\mathbf{f}\|_2^2$ is the data-fidelity term, $\Phi(\mathbf{f})$ the regularization term, and λ the regularization weight. Several algorithms like ADMM can be used to solve Equation 2.

Deep image prior (23) is an approach proposed to solve image recovery problems, such as image inpainting, denoising, and reconstruction. In general DIP schemes, a random but fixed latent noise variable $\mathbf{z} \in \mathbb{R}^L$ serves as the input to an un-trained

deep convolutional neural network (CNN) $G(\mathbf{z}; \boldsymbol{\theta})$, where $\boldsymbol{\theta} \in \mathbb{R}^d$ are the parameters of the network, which are either randomly initialized or tuned with pre-training. During training, $\boldsymbol{\theta}$ is optimized to make the network output $G(\mathbf{z}; \boldsymbol{\theta})$ minimize the reconstruction loss with respect to the measurement \mathbf{m} through the model A :

$$\boldsymbol{\theta}^* = \operatorname{argmin}_{\boldsymbol{\theta}} \|\mathbf{m} - A G(\mathbf{z}; \boldsymbol{\theta})\|_2^2 \quad (3)$$

Where $\boldsymbol{\theta}^*$ is the optimized network weights. No prior information is needed to solve Equation 3; instead, the inductive bias of the convolutional neural network can implicitly enforce a natural appearance in the reconstructed images (30). Typically, an over-parameterized network G is used (33), which means that the number of network parameters d is much larger than the output dimension N .

2.2 Proposed framework

In this work, we propose a hybrid DIP-CS framework for 3D image reconstruction. Because k-space is fully sampled in the readout direction of MR acquisition, we firstly perform 1D inverse Fourier transform in this direction to partition k-space into a sequence $\{\mathbf{m}^{(i)}\}$, where $\mathbf{m}^{(i)}$ is the k-space data for the i th slice along the readout direction. We then solve the 3D reconstruction problem by a slice-by-slice application of regular 2D DIP. However, a slice-by-slice reconstruction not only may induce inter-slice inconsistency of image intensities, but also lacks regularization power over the inter-slice direction. To

address this issue, we combine 2D DIP with a 3D TV regularization, which is separately applied along the intra-slice dimension and the inter-slice dimension. The objective function is shown below:

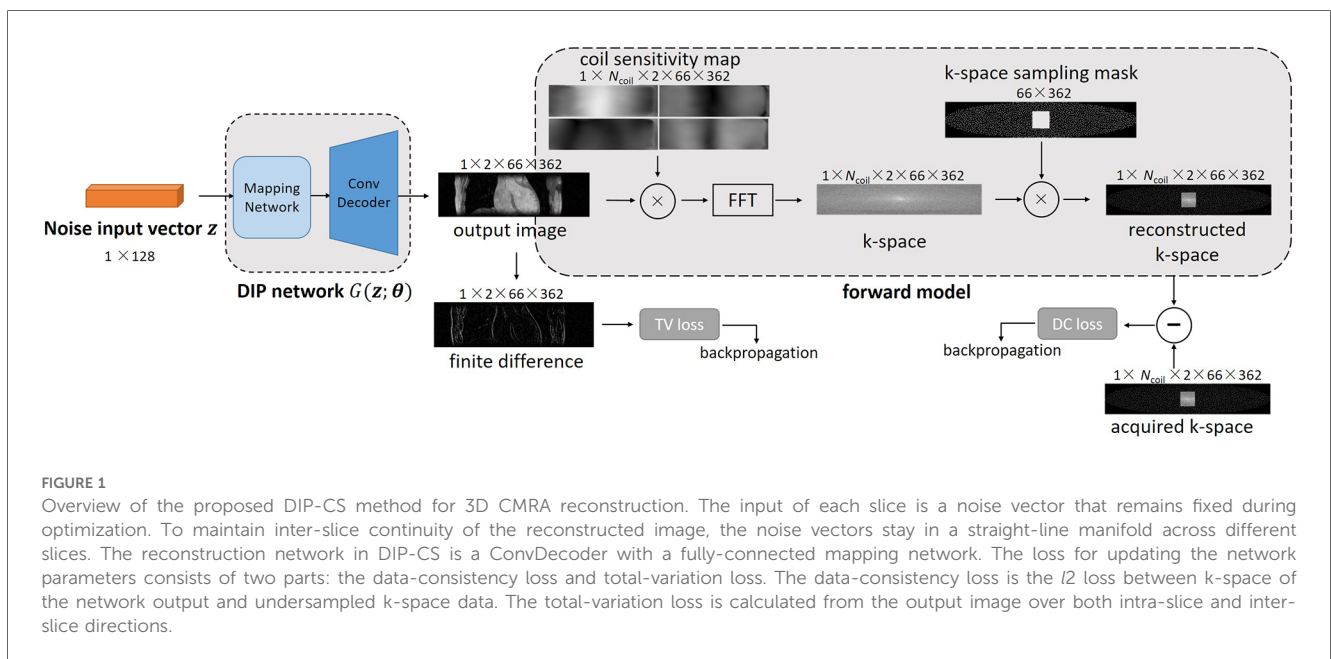
$$\operatorname{argmin}_{\boldsymbol{\theta}} \sum_{i=1}^{N_{\text{slice}}} (\|\mathbf{m}^{(i)} - A G(\mathbf{z}^{(i)}; \boldsymbol{\theta}^{(i)})\|_2^2 + \lambda \|G(\mathbf{z}^{(i)}; \boldsymbol{\theta}^{(i)})\|_{\text{TV}_{yz}}) + \lambda \|G(\mathbf{z}; \boldsymbol{\theta})\|_{\text{TV}_x} \quad (4)$$

where $\mathbf{z}^{(i)}$ and $\boldsymbol{\theta}^{(i)}$ are the corresponding latent representation and network weights of each slice, $G(\mathbf{z}; \boldsymbol{\theta})$ is the 3D image obtained by concatenating all the network output slices, and $\|\cdot\|_{\text{TV}_{yz}}$ and $\|\cdot\|_{\text{TV}_x}$ are the intra-slice and inter-slice TV regularization, respectively. To minimize this function, we combine the ADMM with the proposed DIP-CS formulation by using variable splitting. The problem is transformed into a constrained optimization as in Equation 5:

$$\operatorname{argmin}_{\boldsymbol{\theta}} \sum_{i=1}^{N_{\text{slice}}} (\|\mathbf{m}^{(i)} - A G(\mathbf{z}^{(i)}; \boldsymbol{\theta}^{(i)})\|_2^2 + \lambda \|G(\mathbf{z}^{(i)}; \boldsymbol{\theta}^{(i)})\|_{\text{TV}_{yz}}) + \lambda \|\mathbf{v}\|_{\text{TV}_x} \quad (5)$$

s.t. $\mathbf{v}^{(i)} = G(\mathbf{z}^{(i)}; \boldsymbol{\theta}^{(i)})$

Where \mathbf{v} is the auxiliary variable used to replace $G(\mathbf{z}; \boldsymbol{\theta})$ in the inter-slice regularization term. Figure 1 shows an overview of 3D DIP-CS. The numbers denote the size of each object in this framework. \mathbf{z} is a vector of length 128. We solve



Equation 5 based on the ADMM algorithm, which recursively executes Equations 6a–c:

$$\begin{aligned} \theta_{t+1} = \operatorname{argmin}_{\theta} & \sum_{i=1}^{N_{\text{slice}}} (\|m^{(i)} - AG(z^{(i)}; \theta^{(i)})\|_2^2 \\ & + \lambda \|G(z^{(i)}; \theta^{(i)})\|_{\text{TV}_{yz}} + \frac{\rho}{2} \|G(z^{(i)}; \theta^{(i)}) - v_t^{(i)} \\ & + u_t^{(i)}\|_2^2) \end{aligned} \quad (6a)$$

$$v_{t+1} = \operatorname{argmin}_v \lambda \|v\|_{\text{TV}_x} + \frac{\rho}{2} \|G(z; \theta_{t+1}) - v + u_t\|_2^2 \quad (6b)$$

$$u_{t+1} = u_t + G(z; \theta_{t+1}) - v_{t+1} \quad (6c)$$

Note that the first subproblem in Equation 6a is solved using DIP with a 2D TV regularization slice-by-slice, whereas the second subproblem in Equation 6b is solved along the inter-slice direction only, which can be computed in parallel. Furthermore, since adjacent slices should have similar appearances, we hypothesize that $\{z^{(i)}\}$ across different slices should reside in a connected manifold in the latent space. We thus add an additional constraint on $\{z^{(i)}\}$, as described in Equation 7, which asks $\{z^{(i)}\}$ to reside in a learnable straight-line manifold:

$$z^{(i)} = (1 - \alpha^{(i)})z^{(0)} + \alpha^{(i)}z^{(N_{\text{slice}}-1)}, \quad \alpha^{(i)} = i/(N_{\text{slice}} - 1) \quad (7)$$

Where $z^{(0)}$ and $z^{(N_{\text{slice}}-1)}$ are sampled from the standard uniform distribution $\mathcal{U}(0, 1)$, and $\alpha^{(i)}$ controls the position on this straight line. This constraint was inspired by a previous work of DIP (30), which constrains the latent representation of dynamic MRI to a learnable straight-line manifold.

To accelerate the reconstruction of each slice, we propose initializing the network weights of the i th slice based on those of

the $(i-1)$ th slice, as shown in Figure 2A. The whole algorithm is summarized in Algorithm 1 of Figure 3. Optimization in Equation 6a was performed for N_{DIP} times and the whole iterative process of Equations 6a–c was T iterations.

2.3 Network architecture

Many network architectures can be selected as the backbone of the un-trained model, such as CNN in the original DIP work (23), ConvDecoder (25), MLP in DeepDecoder (34, 35), skip-connected U-Net (36), and even Transformer (26, 37). In this paper, we modified the architecture in ConvDecoder (25) to construct the neural network $G(z^{(i)}; \theta^{(i)})$ in DIP-CS. The network architecture to model each image slice is illustrated in Figure 2B. The input to the network is a fixed vector with length L , containing uniformly distributed noise. Similar to the MapNet in 2D + t time-dependent DIP (30) and the “Mapper” in zero-Shot Learned Adversarial Transformer (37), we firstly used a mapping network consisting of two fully-connected layers to map the noise vector into the input space of the modified ConvDecoder. The first layer in the mapping network outputs a 128-channel vector, and the second transforms this vector into a 128-channel 16×16 feature map. The ConvDecoder module then produces the output image by progressively increasing the image resolution across layers. The input and output of ConvDecoder have two channels for real and imaginary parts because of the complex-valued MR images. Each block in this module comprises of a nearest-neighbor up-sampling layer, a 3×3 128-channel convolutional layer, ReLU function and a batch normalization layer as in the original work. The final layer only uses a 1×1 convolution to linearly combine the channels.

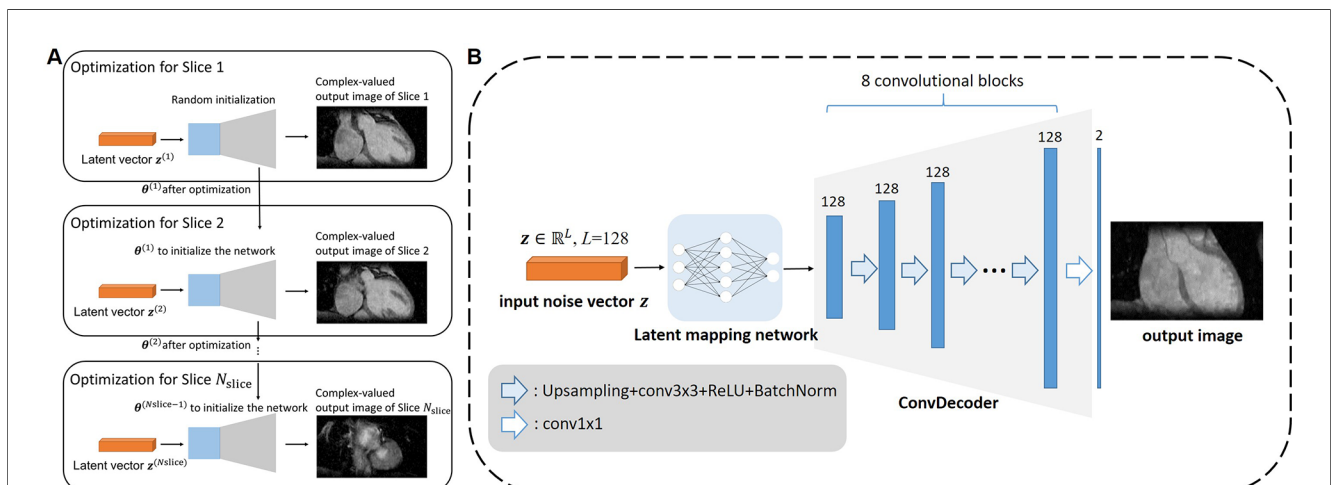


FIGURE 2 Architecture of the proposed framework. (A) The diagram for how the multi-slice reconstruction was performed using the network and multi-slice latent vectors. For each slice except the first, the network parameters were initialized with those obtained from the previous slice to expedite the reconstruction process. (B) Network architecture adopted in the proposed 3D DIP-CS framework. The latent mapping network comprises of fully-connect layers that map the random noise to the input of the ConvDecoder. The ConvDecoder comprises of 8 network blocks, each formed by an up-sampling layer, a convolutional layer, and a batch normalization layer. The output is an image with two channels holding the real and imaginary parts.

Algorithm 1 DIP-CS for 3D Reconstruction**Input:** image \mathbf{x} , k-space measurements \mathbf{y} , noise vector \mathbf{z} , network weights $\boldsymbol{\theta}$ **Output:** \mathbf{x}

```

1: for  $t = 1 : T$  do
2:   Optimization 1
3:   for  $i = 1 : N_{slice}$  do
4:     // Initialize the network parameter
5:      $\boldsymbol{\theta}^{(i)} \leftarrow \boldsymbol{\theta}^{(i-1)}$ 
6:     for  $k = 1 : N_{DIP}$  do
7:       // Compute the image at  $i$ th slice
8:        $\mathbf{f}^{(i)} \leftarrow G(\mathbf{z}^{(i)}; \boldsymbol{\theta}^{(i)})$ 
9:       // Compute the loss
10:       $L \leftarrow \frac{1}{2} \|\mathbf{m}^{(i)} - \mathbf{A}\mathbf{f}^{(i)}\|_2^2 + \lambda \|\mathbf{f}^{(i)}\|_{TV_{yz}} + \frac{\rho}{2} \|\mathbf{f}^{(i)} - \mathbf{v}^{(i)} + \mathbf{u}^{(i)}\|_2^2$ 
11:      // update the weights with gradient
12:       $\boldsymbol{\theta}^{(i)} \leftarrow \text{Adam}(\boldsymbol{\theta}^{(i)}, L)$ 
13:    end for
14:  end for
15:  Optimization 2
16:   $\mathbf{v} \leftarrow \arg \min_{\mathbf{v}} \lambda \|\mathbf{v}\|_{TV_x} + \frac{\rho}{2} \|\mathbf{f} - \mathbf{v} + \mathbf{u}\|_2^2$ 
17:   $\mathbf{u} \leftarrow \mathbf{u} + \mathbf{f} - \mathbf{v}$ 
18: end for

```

FIGURE 3
The DIP-CS algorithm for 3D reconstruction.

3 Experiments

3.1 Datasets preparation

We conducted both retrospective and prospective experiments to investigate the effectiveness of the proposed DIP-CS method. For the retrospective experiment, 10 healthy volunteers were recruited and underwent CMRA imaging in a 3.0 T MRI scanner (uMR 790, United Imaging Healthcare Co. Ltd., Shanghai, China) with a 12-channel torso coil and 32-channel spine coil. Written informed consent was obtained from all subjects before the scan. No contrast agent was used in this study. A navigator-gated, electrocardiogram-triggered, 2-point Dixon T2-prepared spoiled gradient echo sequence was used for 3D CMRA (8, 38), which generated two images of different echo times for each subject. The acquisition window for each cardiac cycle was 164 ms, which was placed in the quiescent diastolic phase to reduce cardiac motion. A four-chamber cardiac cine imaging was performed prior to the CMRA scan to determine the subject-specific optimal trigger delay to minimize cardiac motion during the acquisition. Data were acquired with image readout in the anterior-posterior (AP) direction, and phase encoding in the left-right (LR) and feet-head (FH) directions. Scan parameters of the 3D GRE sequence with Cartesian sampling include: TE1/TE2 = 2.24/3.17 ms, TR = 5.21 ms, duration of T2 preparation = 32 ms, field of view (FOV) = 400 × 300 × 90 mm³ (LR × AP × FH), acquisition matrix = 362 × 272 × 60 (10% oversampling in FH direction), acquired resolution = 1.10 × 1.10 × 1.50 mm³, flip

angle = 10°, bandwidth = 1,070 Hz/pixel, and T2-prep TE = 32 ms. Two-fold undersampling in the phase encoding direction k_y was performed on the basis of elliptical scanning to shorten the scan time. [Supplementary Figure 1A](#) gives a demonstration of the undersampling trajectory. Coil compression was used to compress the multi-coil data into 12 channels (39) using the BART toolbox (40). The 24 × 24 central reference lines were fully-sampled and used for ESPIRiT (41) coil sensitivity estimation. The mean scan time was 23 min 36 s. The 3D image was then reconstructed using iterative SENSE with Tikhonov regularization (regularization weight = 0.1). The result was treated as the ground truth and used to generate retrospectively undersampled data. Eight-fold undersampling was used based on a pseudo-random poisson-disc mask in the k_y - k_z plane (as in [Supplementary Figure 1B](#)). From the 2× accelerated acquisition to the 8× undersampling, the scan time can be reduced to approximately 5–6 min. Different reconstruction methods were then used to reconstruct the images from the retrospectively undersampled data.

For the prospective experiment, we scanned another 13 volunteers with a self-navigated, electrocardiogram-triggered, 2-point Dixon T2-prepared spoiled gradient echo sequence (38), which is different from the navigator-gated sequence used in the retrospective experiment. The sequence uses a variable-density Cartesian undersampling trajectory in the k_y - k_z plane, where the acquisition in each heartbeat passes the k-space center. The k-space center sampling is then reconstructed by 1D inverse Fourier transform to provide a self-navigating signal for retrospective

binning of the data. All data were binned into 4 respiratory phases, and data from the end-expiration bin were used for reconstruction. Data were acquired with image readout in the FH direction, and phase encoding in the LR and AP directions in this sequence. The sequence had the following imaging parameters: TE1/TE2 = 2.24/3.17 ms, TR = 5.21 ms, FOV = 400 × 300 × 132–150 mm³(RL × FH × AP), acquisition matrix = 362 × 272 × (88–100), flip angle = 10°, bandwidth = 1,070 Hz/pixel, and T2-prep TE = 32 ms. We scanned each volunteer for 5 min to obtain the undersampled data. To provide reference images as an evaluation of reconstruction quality, we continued the scan for another 15 min to obtain nearly fully-sampled data ($R \approx 2.5$ for the end-expiration bin). The reference images were then reconstructed using iterative SENSE with Tikhonov regularization (regularization weight = 0.1).

3.2 Ablation study

We performed the ablation study to evaluate the effectiveness of 3D DIP-CS by comparing the method to (1) standard 2D DIP applied slice-by-slice, where the TV regularization was completely removed, and (2) 2D DIP-CS applied slice-by-slice, where only the inter-slice TV regularization was removed. We then assessed the performance of the three methods on the retrospective dataset.

3.3 Comparison methods

We applied four baseline methods for comparison to show the advantage of the proposed method. These methods included (1) zero-filling, which directly fills the unsampled k-space points with zero followed by the inverse Fourier transform and coil combination, (2) iterative SENSE, which is a classical reconstruction method for parallel imaging, (3) CS-TV, which uses total-variation regularization in the classic iterative CS-based reconstruction, and (4) CS-wavelet, which is a CS-based algorithm based on a l_1 -norm constraint on the wavelet transform of the image. The last 3 methods were also compared in multiple undersampling rates and patterns to investigate the generalizability.

3.4 Image and statistical analysis

In our experiments, the two-echo acquisitions were reconstructed separately as two 3D images. To quantitatively evaluate the performance of the reconstruction methods, the peak signal-to-noise ratio (PSNR), structural similarity index (SSIM), and normalized mean-squared error (NMSE) were calculated on each 3D image volume with respect to the reference image.

In addition to the generic reconstruction quality comparison, we also compared the coronary vessel sharpness and conspicuity between different methods. To evaluate vessel sharpness, we firstly performed the two-point Dixon water-fat separation (43) based on the method described in (44) to generate the water

image. We then reformatted the water image using curved planar reformation based on 3D Slicer (Version 5.2.2) (42) to visualize the coronary arteries. After that, we computed sharpness of the left anterior descending artery (LAD) and the right coronary artery (RCA) based on the intensity profile of vessel edges, using the method previously described in Ref (45, 46). To evaluate the coronary artery conspicuity, we invited a clinician (with more than 5 years of experience in cardiology) and a MR physicist (with more than 10 years of experience) to rate the images of reformatted coronary arteries. The images of different comparison methods were randomized, and their identity was blinded to the readers. For each image, the visualized vessel was given a rank using a five-point scoring system: score 1, uninterpretable (vessels that were almost indiscernible); score 2, poor (vessels that were visible but highly degraded by noise or artifacts); score 3, fair (vessels with moderately blurred borders); score 4, good (clear vessels but with slight blurring); score 5, excellent (vessels with sharply defined borders and structure). The average score from the two readers was used as the final score.

In tables, statistics for quantitative metrics were provided as mean ± standard deviation across all the subjects. Statistical significance of the difference between two methods was assessed by using the Wilcoxon signed-rank test (for expert scoring) and paired *t*-test (for others). Statistical significance of the difference between multiple methods was assessed using repeated-measures one-way analysis of variance (ANOVA) with Bonferroni correction. A *P* value less than 0.05 was considered statistically significant.

3.5 Implementation details

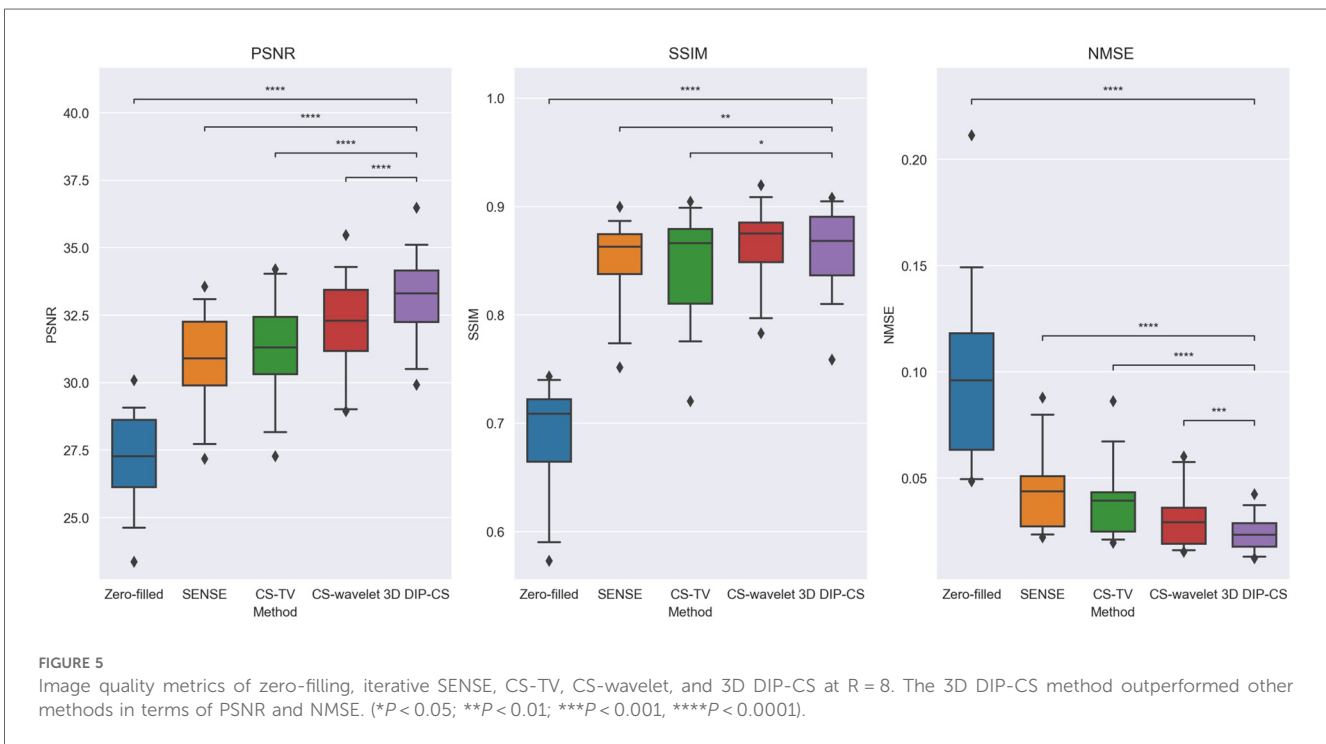
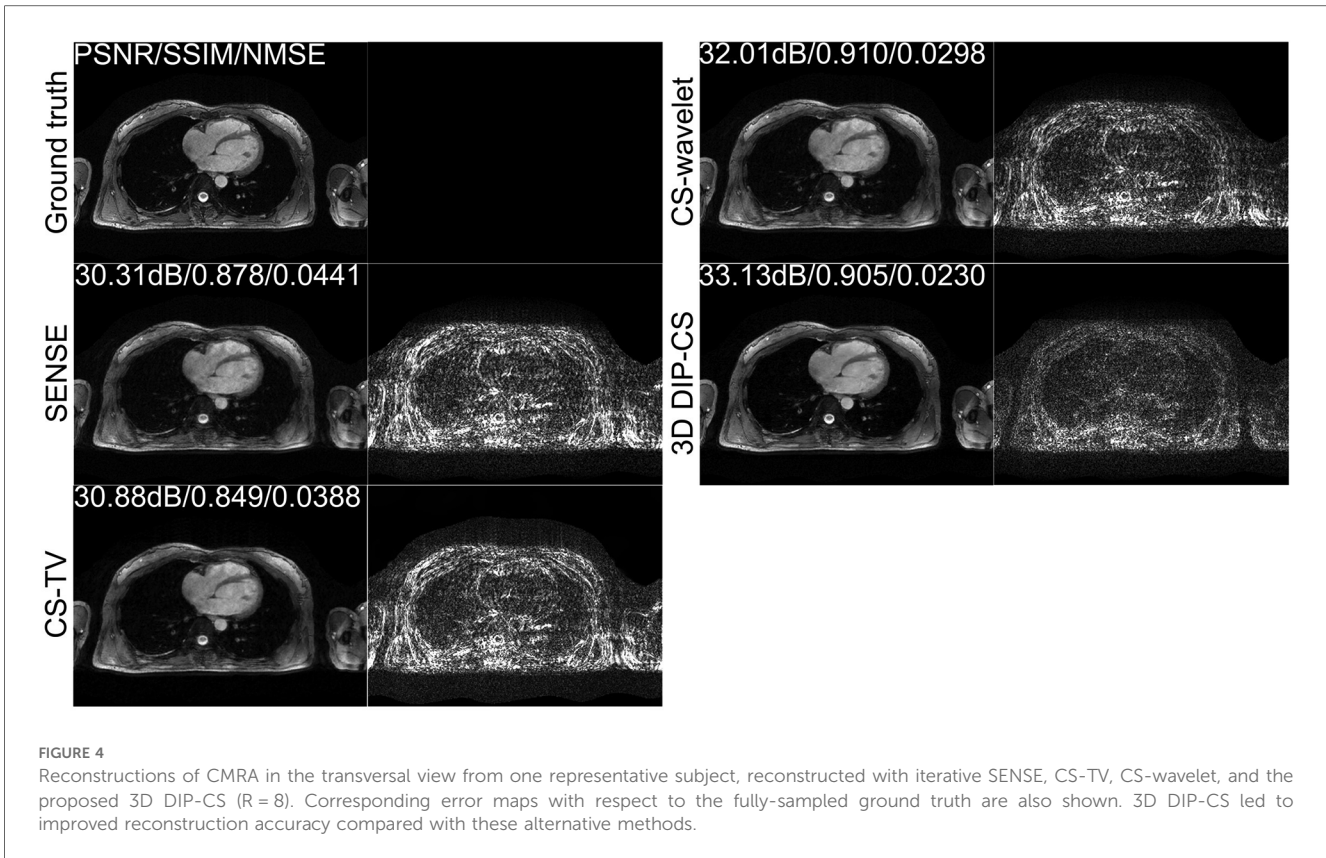
A platform equipped with 4 Intel Xeon Platinum 8260 (2.4 GHz) CPU and two NVIDIA A100 GPU was used for the experiments. Our framework was implemented with Pytorch 1.9.1 on Python 3.7.7. The proximal operator of total variation was solved using proxTV toolbox (47, 48). For each slice, the network of DIP was optimized for 1,000 iterations (N_{DIP} in Algorithm 1) using the Adam (49) optimizer with a learning rate of 5×10^{-5} . After parameter tuning, the parameters in Algorithm 1 were set to be $\rho = 0.001$ and $\lambda = 0.006$. The number of outer loops was set to $T = 3$ for faster reconstruction. The results of optimizing the parameters were demonstrated in Supplementary Figure 2. The baseline CS-TV and CS-wavelet were implemented with BART (40). The regularization parameter was optimized on one sample and was finally set to be 0.01.

4 Results

4.1 Imaging with retrospective accelerations

4.1.1 Comparison with other methods

Figure 4 shows the comparison of iterative SENSE, CS-TV, CS-wavelet, and the proposed method on $R = 8$. Compared with the other two methods, DIP-CS produced results with fewer artifacts in the image and lower reconstruction errors. The



quantitative values calculated from the entire volume were also in accordance with the visual perception. Figure 5 shows the statistical comparison of these quantitative metrics across 10 subjects. The 3D DIP-CS method achieved higher PSNR and

lower NMSE (PSNR: 33.20 ± 1.63 dB; NMSE: 0.0236 ± 0.0087) compared with SENSE (PSNR: 30.75 ± 1.75 dB, $P = 5 \times 10^{-10}$; NMSE: 0.0424 ± 0.0186 , $P = 2 \times 10^{-6}$), CS-TV (PSNR: 31.22 ± 1.87 dB, $P = 3 \times 10^{-6}$; NMSE: 0.0382 ± 0.0174 , $P = 6 \times 10^{-5}$), and

CS-wavelet (PSNR: 32.25 ± 1.76 dB, $P = 2 \times 10^{-5}$; NMSE: 0.0299 ± 0.0128 , $P = 4 \times 10^{-4}$). Differences in SSIM between DIP-CS and CS-wavelet were not statistically significant (SSIM: 0.862 ± 0.038 vs. 0.864 ± 0.038 , $P = 0.57$) in our experiments.

4.1.2 Performance on multiple acceleration factors

Figure 6 shows performance of the methods under different acceleration factors. We observed that under all the undersampling rates, 3D DIP-CS provided overall better reconstructions than the comparison methods. The generalizability to different acceleration factors is useful because the actual acceleration factor can be variant for a practical five-minute acquisition. Table 1 presents the quantitative results for reconstruction quality metrics using different undersampling rates. 3D DIP-CS led to overall better PSNR, SSIM, and NMSE than other methods.

4.1.3 Performance on the changed sampling pattern

To show the feasibility of the proposed method for uniform sampling, which is more commonly used in practice, we compared the reconstruction of each method with one-dimensional uniform undersampling. Figure 7A shows a representative result. In this example, only 3D DIP-CS successfully removed the aliasing artifacts of uniform undersampling, which can be found in the reconstruction of other methods (blue arrows). Figure 7B shows the results of quantitative comparisons. 3D DIP-CS led to higher image

TABLE 1 Average PSNR, SSIM, and NMSE for different methods across a variety of undersampling rates. Numbers indicate mean \pm standard deviation.

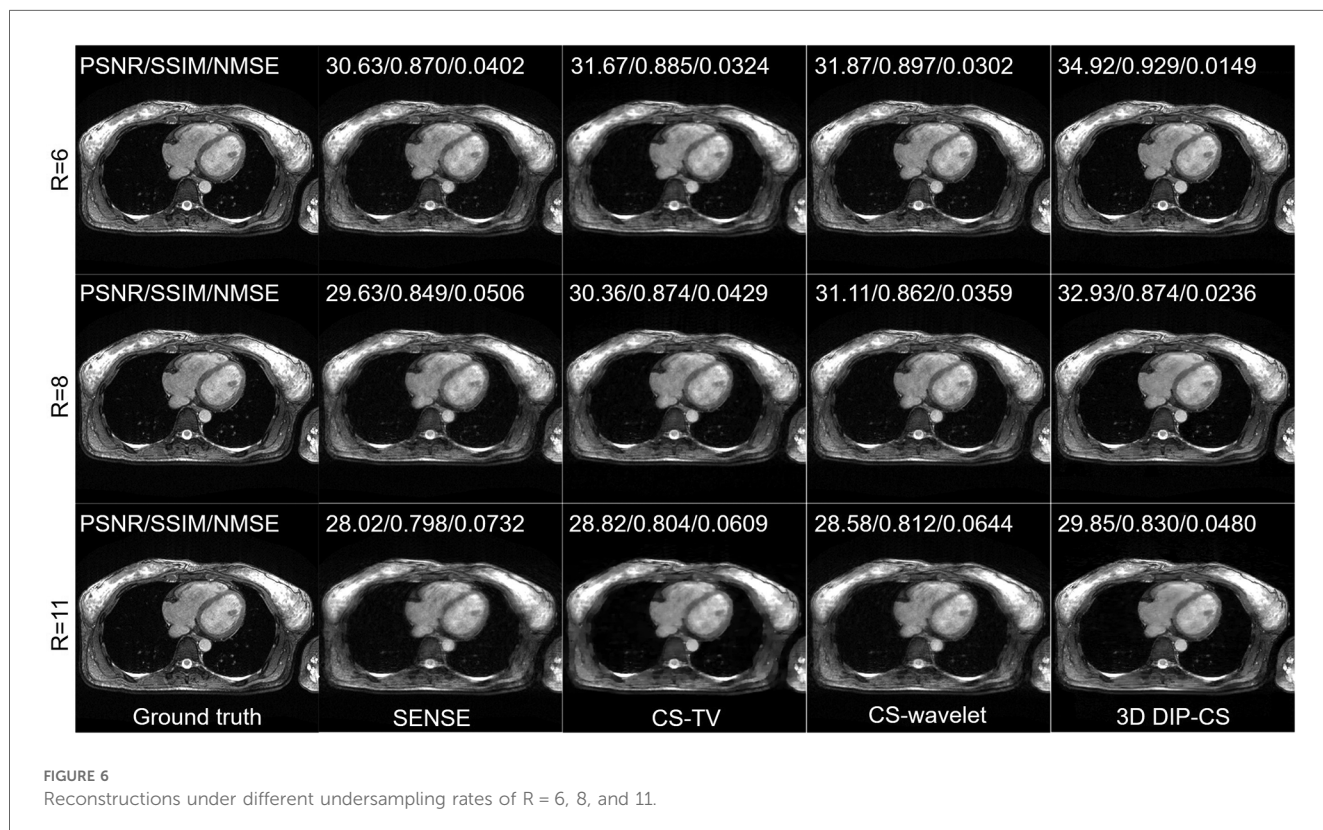
Sampling rate	Method	PSNR(dB)	SSIM	NMSE
R = 6	SENSE	31.86 \pm 1.88	0.872 \pm 0.044	0.0331 \pm 0.0155
	CS-TV	32.05 \pm 1.92	0.881 \pm 0.038	0.0317 \pm 0.0142
	CS-wavelet	33.32 \pm 2.01	0.901 \pm 0.041	0.0238 \pm 0.0116
	3D DIP-CS	35.98 \pm 1.89	0.928 \pm 0.032	0.0127 \pm 0.0054
R = 8	SENSE	30.75 \pm 1.75	0.848 \pm 0.041	0.0424 \pm 0.0186
	CS-TV	31.22 \pm 1.87	0.847 \pm 0.049	0.0382 \pm 0.0174
	CS-wavelet	32.25 \pm 1.76	0.864 \pm 0.038	0.0299 \pm 0.0128
	3D DIP-CS	33.20 \pm 1.63	0.862 \pm 0.038	0.0236 \pm 0.0087
R = 11	SENSE	28.98 \pm 1.78	0.789 \pm 0.045	0.0639 \pm 0.0287
	CS-TV	29.72 \pm 1.61	0.0795 \pm 0.051	0.0546 \pm 0.0234
	CS-wavelet	29.58 \pm 1.78	0.808 \pm 0.044	0.0556 \pm 0.0252
	3D DIP-CS	30.83 \pm 1.69	0.0824 \pm 0.030	0.0408 \pm 0.0149

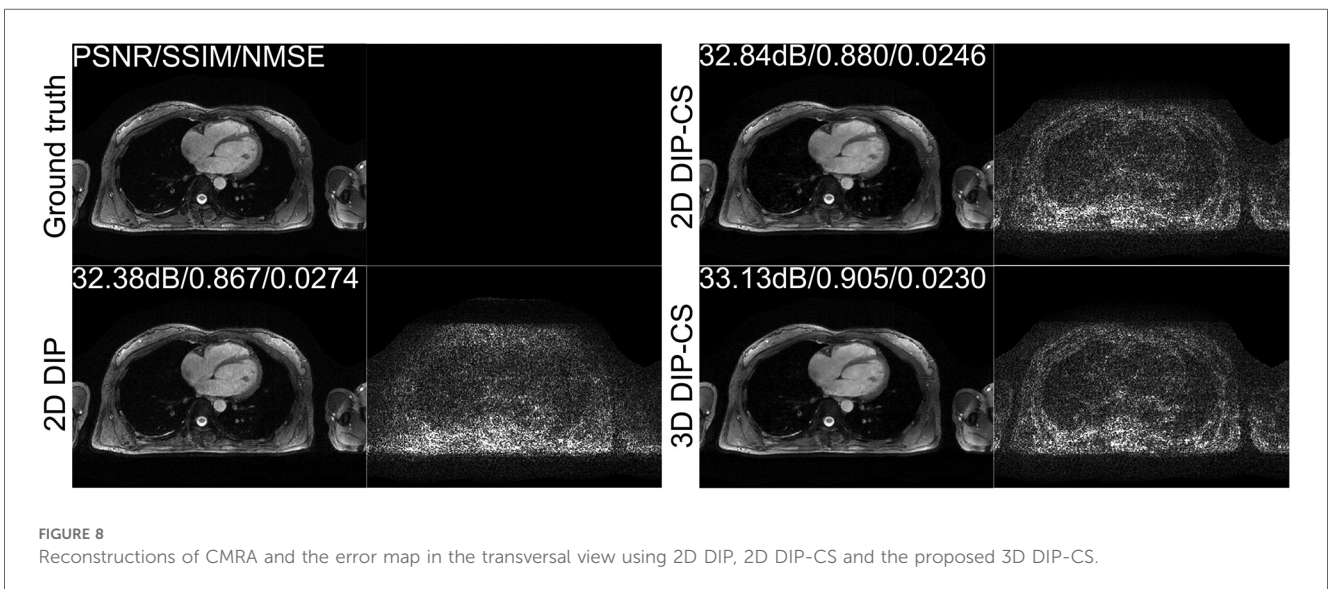
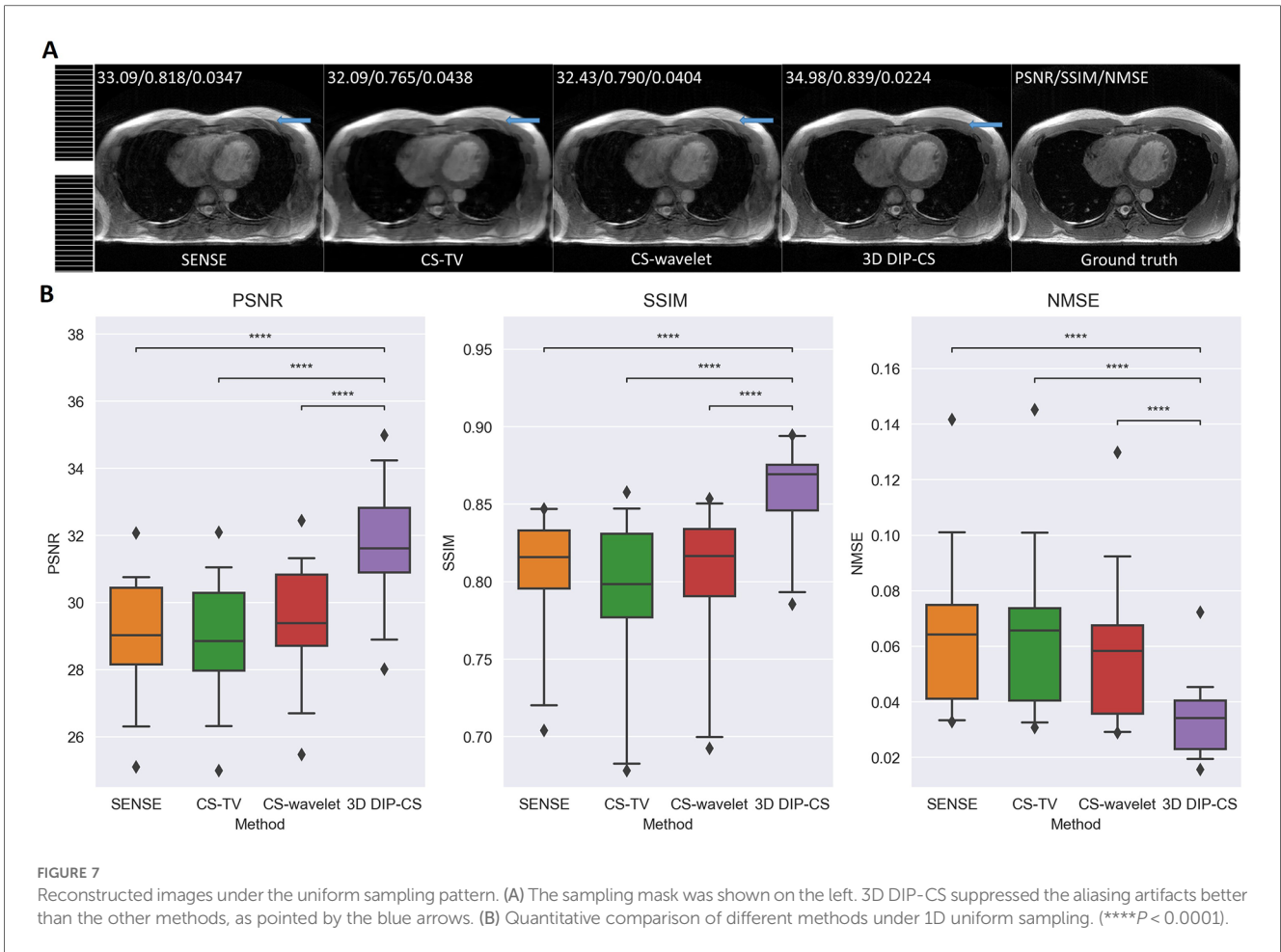
Bold value indicates the highest mean value of the metric among the comparison methods.

qualities compared with other methods when uniform sampling was employed.

4.2 Ablation study

We employed ablation studies to investigate the effect of the intra-slice and inter-slice total variation regularization in the proposed 3D DIP-CS method. To do that, we compared the method with 2D DIP, which employed no total variation, and 2D DIP-CS, which employed only intra-slice total variation. Figure 8 shows the results. The proposed DIP-CS method





achieved a better image quality compared with both 2D DIP and 2D DIP-CS. The 2D DIP method exhibited stronger noise in the heart region. While both 2D and 3D DIP-CS methods reduced noise, 3D DIP-CS yielded slightly higher values in image quality

metrics, suggesting the addition of inter-slice CS further improved the reconstruction.

Table 2 shows the statistical comparison of different quantitative metrics between 2D DIP, 2D DIP-CS, and 3D DIP-CS. The proposed

TABLE 2 Comparisons of PSNR/SSIM/NMSE on 2D DIP, 2D DIP-CS, and 3D DIP-CS.

Method	PSNR (dB)	SSIM	NMSE
2D DIP	31.92 ± 1.97	0.850 ± 0.034	0.0344 ± 0.0248
2D DIP-CS	32.96 ± 1.63	0.852 ± 0.034	0.0249 ± 0.0091
3D DIP-CS	33.20 ± 1.63	0.862 ± 0.038	0.0236 ± 0.0087

Bold value indicates the highest mean value of the metric among the comparison methods.

method achieved a higher PSNR/SSIM and lower NMSE (3D DIP-CS vs. 2D DIP: $P = 5 \times 10^{-4}/0.004/0.019$ for PSNR/SSIM/NMSE; 3D DIP-CS vs. 2D DIP-CS: $P = 2 \times 10^{-7}/3 \times 10^{-4}/1 \times 10^{-6}$ for PSNR/SSIM/NMSE) compared to the other methods. 2D DIP-CS also achieved a better PSNR and NMSE than 2D DIP ($P = 0.003/0.037$ for PSNR and NMSE), although SSIM was not significantly different between the two methods ($P = 0.71$).

4.3 Imaging with prospective accelerations

We assessed our method using real CMRA data acquired within 5 min, which had an acceleration factor of 7–11. An image reconstructed from a 15 min scan was used as the reference. Figure 9 shows a representative result. SENSE and CS-based

methods were able to give reconstructions with reduced undersampling artifacts compared with zero-filling. However, noise appeared stronger in the SENSE reconstruction compared to the other methods. CS-TV significantly suppressed the noise but caused visible blurring in the images. CS-wavelet reconstructions mitigated some noise and undersampling artifacts; however, the images still showed some unsmooth features compared with the reference image (indicated by the white arrows in Figure 9). The 3D DIP-CS method removed most of the artifacts and well suppressed the noise, generating an image quality more similar to the reference image. Table 3 quantitatively compared the image quality metrics among these methods with respect to the 15 min reference images. DIP-CS achieved significantly higher PSNR, SSIM, and lower NMSE compared to classical CS-wavelet (PSNR: $P = 0.001$; SSIM: $P = 0.015$; NMSE: $P = 0.001$).

4.4 Evaluation of the coronary arteries

Figure 10A compares reformatted coronary arteries of different methods using curved planar reformation. CMRA from SENSE reconstruction was noisy compared to that from CS-TV,

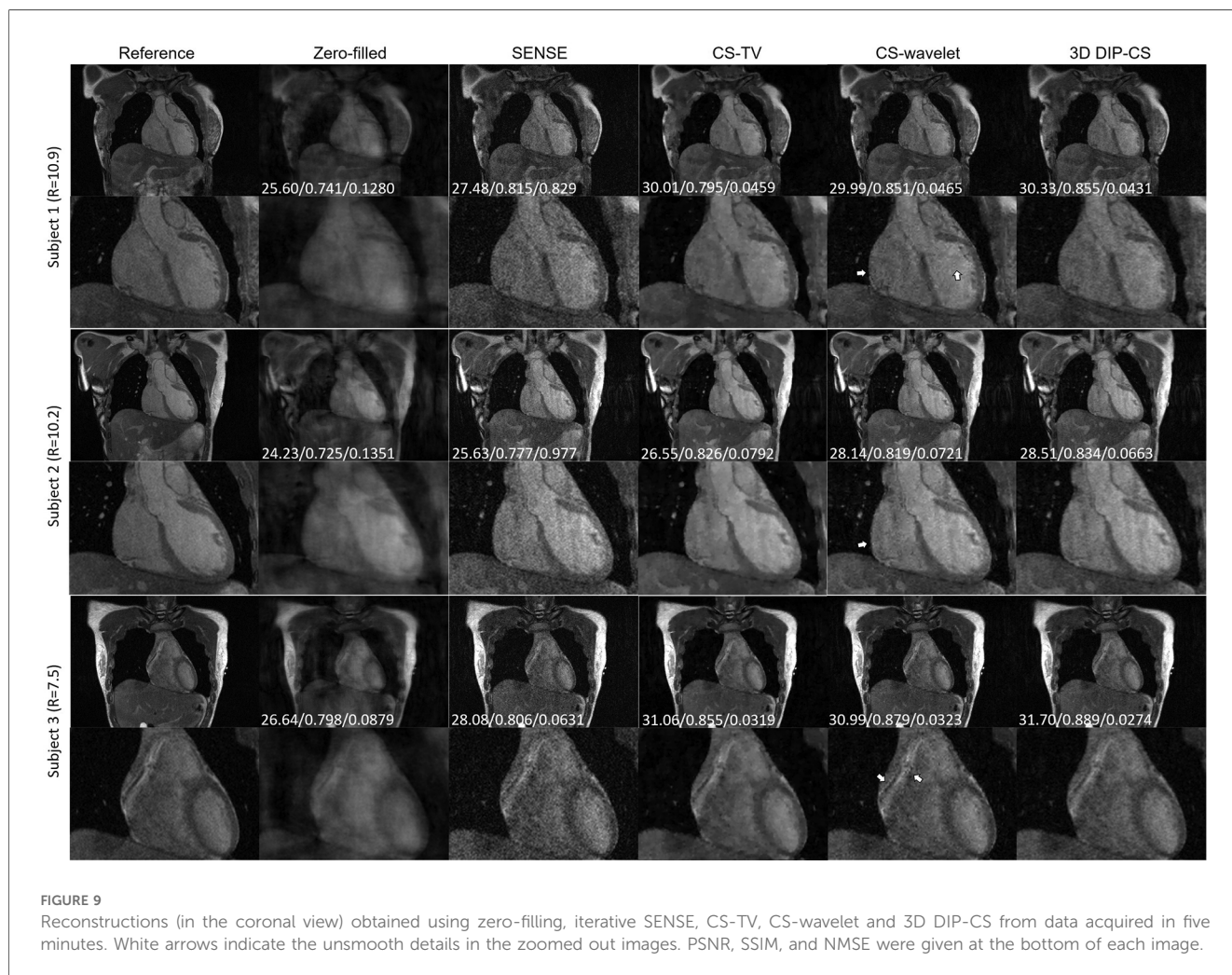


TABLE 3 Quantitative comparison of the reconstruction quality with respect to reference images among all the methods on the dataset.

Method	PSNR (dB)	SSIM	NMSE
Zero-filled	25.80 ± 0.96	0.738 ± 0.075	0.1161 ± 0.0315
SENSE	26.89 ± 0.97	0.768 ± 0.067	0.0889 ± 0.0194
CS-TV	29.16 ± 1.28	0.819 ± 0.044	0.0530 ± 0.0161
CS-wavelet	29.09 ± 1.10	0.826 ± 0.061	0.0564 ± 0.0166
DIP-CS	29.82 ± 0.94	0.841 ± 0.050	0.0472 ± 0.0130

Bold value indicates the highest mean value of the metric among the comparison methods.

CS-wavelet and DIP-CS. The boundaries of vessels in CS-TV reconstruction were slightly blurred. CS-wavelet reconstruction better mitigated the noise, yet residual artifacts and unsmooth edges of the coronary arteries were still visible. DIP-CS achieved the best coronary visualization among the compared methods.

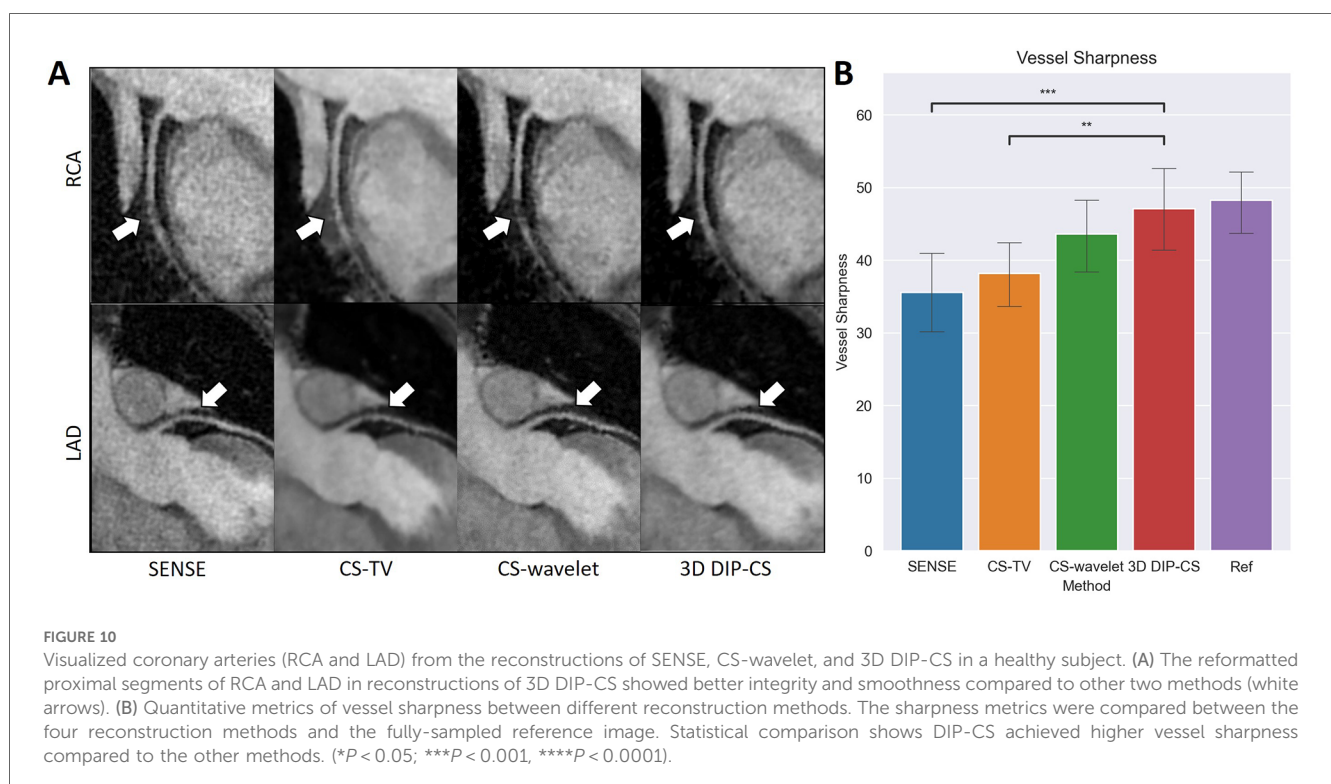
Since reconstruction error metrics like NMSE and SSIM do not directly reflect the clinical quality of imaging, we compared the sharpness and conspicuity of the coronary arteries between different reconstruction methods. **Figure 10B** shows the results of quantitative vessel sharpness evaluation. 3D DIP-CS (47.2 ± 12.8%) significantly improved the vessel sharpness compared with CS-TV (38.1 ± 10.2%, $P = 0.005 < 0.01$), and SENSE (35.53 ± 12.5%, $P = 9 \times 10^{-3} < 0.001$). 3D DIP-CS also led to higher vessel sharpness compared with CS-wavelet (43.6 ± 10.9%, $P = 0.20 > 0.05$), although lacking statistical significance after Bonferroni correction. No significant difference was found between DIP-CS and the mildly undersampled reference images. **Figure 11** shows the results of conspicuity scores rated by two readers. 3D DIP-CS resulted in significantly higher average score (RCA: 4.0 ± 0.6; LAD: 4.3 ± 0.5) than SENSE (RCA: 2.3 ± 0.9,

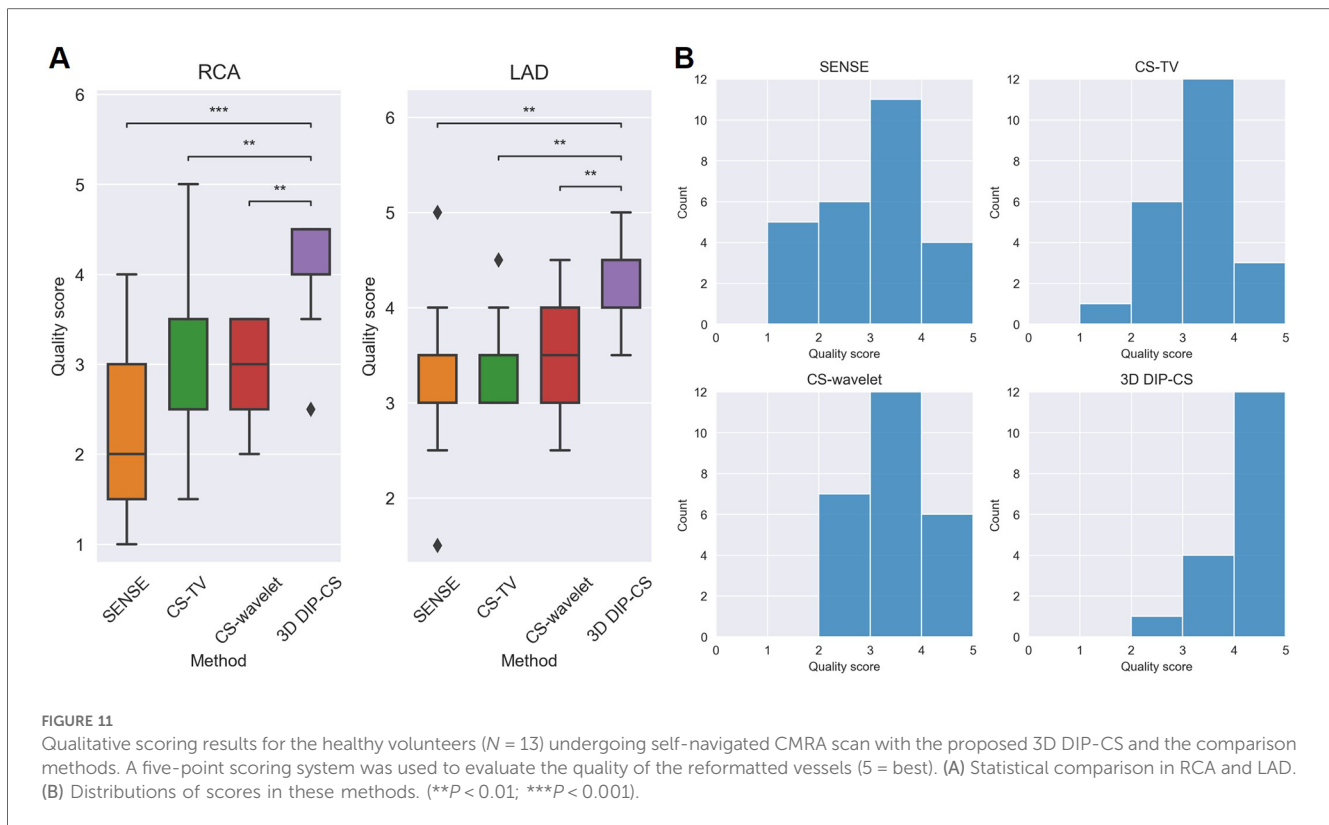
$P = 2 \times 10^{-4}$; LAD: 3.3 ± 0.8; $P = 0.007$), CS-TV (RCA: 2.9 ± 0.9, $P = 0.003$; LAD: 3.4 ± 0.4, $P = 0.003$), and CS-wavelet (RCA: 3.0 ± 0.5, $P = 0.002$; LAD: 3.5 ± 0.6; $P = 0.002$).

5 Discussion

In this work, we introduce a scan-specific 3D DIP-CS method for accelerated CMRA reconstruction. The method demonstrates sufficient accuracy in reconstructing the 3D image from a 7-to-11-fold undersampling, which shortens the scan time of 3D CMRA from more than 20 min–5 min. Compared to the baseline methods such as iterative SENSE, CS-TV and CS-wavelet, our method exhibits superior mitigation of artifacts and noise, resulting in improved image quality for multiple k-space sampling rates and sampling patterns, and improved coronary artery sharpness and conspicuity. The short scan time underscores the potential of our method for clinical application.

Many existing learning-based reconstruction methods rely on large datasets for training. For example (50), trained a variational neural network on 2.7k images to perform nine-fold accelerated motion-compensated reconstruction (51). used data acquired from 45 subjects with an extensive data augmentation to train an unrolled MoCo-MoDL network for nine-fold accelerated CMRA. A recent study (52) used an unrolled CS-AI framework trained on 740k images of various anatomies and contrasts to accelerate data acquisition with an acceleration factor of 6 (the mean scan time was 8.1 min). Compared to these methods, our method has a key advantage in that it does not need external training data. Furthermore, since it is scan-specific, less concerns are present





regarding its generalizability compared to fully-supervised methods (29). The method can thus have the potential to be used with different CMRA sequences and scan parameters, as shown in our experiments. Due to the same reason, it also holds promise for extension to solving other 3D reconstruction problems in MRI and other imaging modalities.

Compared with deep learning methods such as deep unrolling, our method requires significantly less GPU memory during optimization. In our experiment, 3D DIP-CS used less than 5 GB of GPU memory, making it a more practical and memory-efficient option. Owing to its efficient memory usage, the DIP-CS method could potentially be expanded with a 3D DIP network architecture, which has already been explored in a previous study for 3D PET imaging (53). This 3D DIP architecture is particularly useful for those 3D sampling trajectories that cannot be transformed into 2D slice-by-slice sampling, such as 3D radial sampling. In our initial validation, we have implemented such a 3D network architecture by changing the 2D convolutional layers to 3D layers. However, reconstruction results from the 3D DIP method were quite blurred, suggesting the presence of model underfitting. We opine that this lack of fitting may be due to the limited network capacity of a 3D DIP encoder compared to the large number of voxels, as previous studies have found that 2D DIP networks are often over-parameterized (20, 33, 54). Therefore, more sophisticated modifications may be needed to develop a 3D DIP method with high reconstruction quality.

In this work, we compared the performance of 3D DIP-CS with several other unsupervised methods, including iterative SENSE, CS-TV and CS-wavelet. We found that 3D DIP-CS led to

improved reconstruction accuracy and image quality. In addition to these benefits, we also found that the incorporation of TV regularization often resulted in faster convergence, with an example shown in Figure 12. This may be an additional advantage for a combination of DIP and the CS-based regularization.

The study design and the method have limitations. Firstly, due to the absence of a large CMRA dataset, our experiments were conducted on a relatively small dataset, which necessitates further studies on larger datasets. A larger number of subjects can enhance the statistical power of the analysis and increase reliability of the conclusions. Furthermore, a larger, multicenter dataset also contains more diverse data, providing a better vehicle for evaluating generalizability of the proposed method. Secondly, the acquisition window of our CMRA sequence was 164 ms, which was mildly longer than acquisition windows in previous studies [90–130 ms in (55)]. We will investigate the use of shorter acquisition window in future to reduce the potential adverse effect of cardiac motion. However, this limitation should not affect the relative differences between different reconstruction methods, as the same sequence was used for all of them. Thirdly, due to the limitation of computing resources and time, we only explored one type of network architecture. However, many other architectures such as Transformer or U-Net have been proposed as the backbone of DIP and may bring improvement to the results (26, 29, 37, 56). Finally, our method needs approximately 3 h to reconstruct a single volume, which needs acceleration before clinical translation. Long reconstruction time is a common issue for DIP. For example, in (24), the un-trained DIP-based method took 6 min for reconstruction of a single slice even after acceleration. Furthermore,

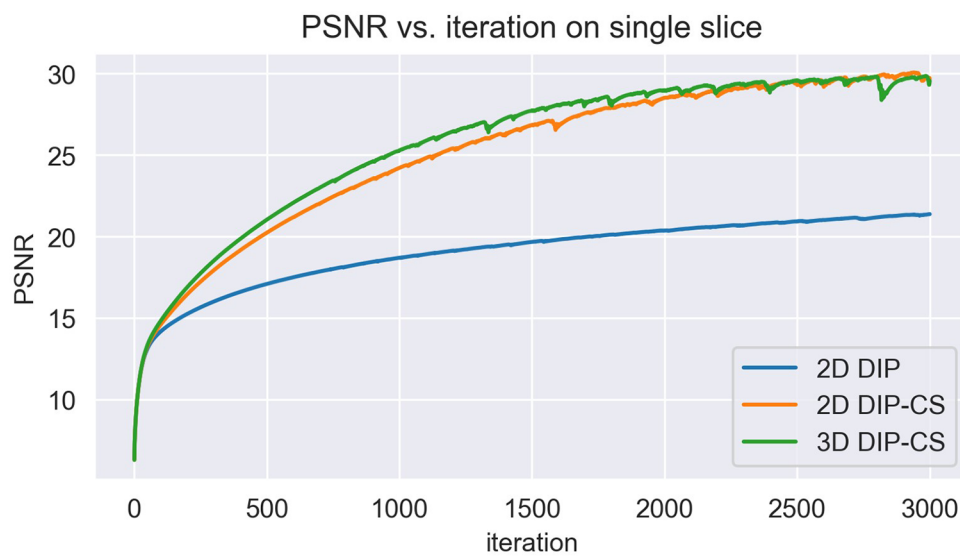


FIGURE 12

The variation of PSNR over each iteration of training. Initializations of network parameters were the same among the three methods for a fair comparison.

due to the large data size, even classical CMRA reconstruction methods like non-rigid motion-compensated PROST need ~50 min for reconstruction (9). Parallel computing and the use of 3D networks may be feasible approaches for further exploration.

6 Conclusion

In conclusion, we propose a hybrid DIP-CS method for the acceleration of 3D CMRA. By leveraging the implicit prior from the CNN architecture and the sparsifying prior from total variation, the proposed 3D DIP-CS method demonstrates the capability to recover images from 7 to 11 folds of acceleration without training on any external datasets. The superior image quality obtained by the proposed method renders it a useful method for accelerating 3D CMRA.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Institutional Review Board for Human Research Protections, Shanghai Jiao Tong University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

ZX: Conceptualization, Formal Analysis, Investigation, Methodology, Software, Validation, Writing – original draft, Writing – review & editing. SZ: Formal Analysis, Methodology, Software, Writing – review & editing. FY: Data curation, Resources, Writing – review & editing. JG: Resources, Writing – review & editing. HP: Software, Writing – review & editing. CZ: Software, Writing – review & editing. HJ: Conceptualization, Writing – review & editing. CH: Conceptualization, Funding acquisition, Methodology, Supervision, Visualization, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was partially supported by the National Natural Science Foundation of China (No. 62001288) and the Shanghai Science and Technology Commission (No. 22TS1400200).

Acknowledgments

The authors gratefully acknowledge Dr. Hongfei Lu from Department of Radiology at Zhongshan Hospital, Fudan University, for her help in the clinical scoring of images and the revision of the manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcvm.2024.1408351/full#supplementary-material>

References

- Hajhosseiny R, Bustin A, Munoz C, Rashid I, Cruz G, Manning WJ, et al. Coronary magnetic resonance angiography: technical innovations leading US to the promised land? *JACC Cardiovasc Imaging*. (2020) 13(12):2653–72. doi: 10.1016/j.jcmg.2020.01.006
- Hajhosseiny R, Munoz C, Cruz G, Khamis R, Kim WY, Prieto C, et al. Coronary magnetic resonance angiography in chronic coronary syndromes. *Front Cardiovasc Med*. (2021) 8:682924. doi: 10.3389/fcvm.2021.682924
- Sakuma H. Coronary CT versus mr angiography: the role of mr angiography. *Radiology*. (2011) 258(2):340–9. doi: 10.1148/radiol.10100116
- Bustin A, Ginami G, Cruz G, Correia T, Ismail TF, Rashid I, et al. Five-minute whole-heart coronary mra with sub-millimeter isotropic resolution, 100% respiratory scan efficiency, and 3d-prost reconstruction. *Magn Reson Med*. (2019) 81(1):102–15. doi: 10.1002/mrm.27354
- Lustig M, Donoho D, Pauly JM. Sparse MRI: the application of compressed sensing for rapid mr imaging. *Magn Reson Med*. (2007) 58(6):1182–95. doi: 10.1002/mrm.21391
- Liang D, Liu B, Wang J, Ying L. Accelerating sense using compressed sensing. *Magn Reson Med*. (2009) 62(6):1574–84. doi: 10.1002/mrm.22161
- Bustin A, Lima da Cruz G, Jaubert O, Lopez K, Botnar RM, Prieto C. High-dimensionality undersampled patch-based reconstruction (hd-prost) for accelerated multi-contrast MRI. *Magn Reson Med*. (2019) 81(6):3705–19. doi: 10.1002/mrm.27694
- Zhang Y, Zhang X, Jiang Y, Yang P, Hu X, Peng B, et al. 3D whole-heart noncontrast coronary mr angiography based on compressed sense technology: a comparative study of conventional sense sequence and coronary computed tomography angiography. *Insights Imaging*. (2023) 14(1):35. doi: 10.1186/s13244-023-01378-w
- Bustin A, Rashid I, Cruz G, Hajhosseiny R, Correia T, Neji R, et al. 3D whole-heart isotropic sub-millimeter resolution coronary magnetic resonance angiography with non-rigid motion-compensated prost. *J Cardiovasc Magn Reson*. (2020) 22(1):24. doi: 10.1186/s12968-020-00611-5
- Zhu B, Liu JZ, Cauley SF, Rosen BR, Rosen MS. Image reconstruction by domain-transform manifold learning. *Nature*. (2018) 555(7697):487–92. doi: 10.1038/nature25988
- Hyun CM, Kim HP, Lee SM, Lee S, Seo JK. Deep learning for undersampled MRI reconstruction. *Phys Med Biol*. (2018) 63(13):135007. doi: 10.1088/1361-6560/aac71a
- Yang Y, Sun J, Li H, Xu Z. Admm-net: a deep learning approach for compressive sensing MRI. *arXiv [Preprint]*. *arXiv.1705.06869* (2017) 1705.06869:1–14. doi: 10.48550/arXiv.1705.06869
- Aggarwal HK, Mani MP, Jacob M. Modl: model-based deep learning architecture for inverse problems. *IEEE Trans Med Imaging*. (2019) 38(2):394–405. doi: 10.1109/TMI.2018.2865356
- Zhang J, Ghanem B. *Ista-net: interpretable optimization-inspired deep network for image compressive sensing*. *IEEE Conference on Computer Vision and Pattern Recognition* (2018). p. 1828–37
- Sriram A, Zbontar J, Murrell T, Defazio A, Zitnick CL, Yakubova N, et al. End-to-end variational networks for accelerated MRI reconstruction. Medical image computing and computer assisted intervention—miccai 2020. *Lect Notes Comput Sci*. (2020) 12262:64–73. doi: 10.1007/978-3-030-59713-9_7
- Venkatakrisnan SV, Bouman CA, Wohlberg B. *Plug-and-play priors for model based reconstruction*. *IEEE Global Conference on Signal and Information Processing* (2013). p. 945–8
- Ahmad R, Bouman CA, Buzzard GT, Chan S, Liu S, Reehorst ET, et al. Plug-and-play methods for magnetic resonance imaging: using denoisers for image recovery. *IEEE Signal Process Mag*. (2020) 37(1):105–16. doi: 10.1109/msp.2019.2949470
- Tezcan KC, Baumgartner CF, Luechinger R, Pruessmann KP, Konukoglu E. Mr image reconstruction using deep density priors. *IEEE Trans Med Imaging*. (2019) 38(7):1633–42. doi: 10.1109/TMI.2018.2887072
- Luo G, Zhao N, Jiang W, Hui ES, Cao P. MRI reconstruction using deep Bayesian estimation. *Magn Reson Med*. (2020) 84(4):2246–61. doi: 10.1002/mrm.28274
- Zou Q, Ahmed AH, Nagpal P, Priya S, Schulte RF, Jacob M. Variational manifold learning from incomplete data: application to multislice dynamic MRI. *IEEE Trans Med Imaging*. (2022) 41(12):3552–61. doi: 10.1109/tmi.2022.3189905
- Antun V, Renna F, Poon C, Adcock B, Hansen AC. On instabilities of deep learning in image reconstruction and the potential costs of AI. *Proc Natl Acad Sci USA*. (2020) 117(48):30088–95. doi: 10.1073/pnas.1907377117
- 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*. (2018) 81(1):116–28. doi: 10.1002/mrm.27355
- Ulyanov D, Vedaldi A, Lempitsky V. Deep image prior. *Int J Comput Vis*. (2020) 128(7):1867–88. doi: 10.1007/s11263-020-01303-4
- Zalbagi Darestani M, Heckel R. Accelerated MRI with un-trained neural networks. *IEEE Trans Comput Imaging*. (2021) 7:724–33. doi: 10.1109/tci.2021.3097596
- Darestani MZ, Heckel R. *Can Un-Trained Networks Compete with Trained Ones for Accelerated MRI?* Concord, NH: ISMRM & SMRT Annual Meeting & Exhibition (2021). p. 271.
- Korkmaz Y, Yurt M, Dar SUH, Özbey M, Cukur T. Deep MRI reconstruction with generative vision transformers. Machine learning for medical image reconstruction. *Lect Notes Comput Sci*. (2021) 12964:54–64. doi: 10.1007/978-3-030-88552-6_6
- Hamilton JI. A self-supervised deep learning reconstruction for shortening the breathhold and acquisition window in cardiac magnetic resonance fingerprinting. *Front Cardiovasc Med*. (2022) 9:928546. doi: 10.3389/fcvm.2022.928546
- Hamilton JI, Truesdell W, Galizia M, Burris N, Agarwal P, Seiberlich N. A low-rank deep image prior reconstruction for free-breathing ungated spiral functional CMR at 0.55 t and 1.5 t. *Magn Reson Mater Phys Biol Med*. (2023) 36(3):451–64. doi: 10.1007/s10334-023-01088-w
- Leynes AP, Deveshwar N, Nagarajan SS, Larson PEZ. Scan-specific self-supervised Bayesian deep non-linear inversion for undersampled MRI reconstruction. *IEEE Trans Med Imaging*. (2024) 43:1. doi: 10.1109/tmi.2024.3364911
- Yoo J, Jin KH, Gupta H, Yerly J, Stuber M, Unser M. Time-dependent deep image prior for dynamic MRI. *IEEE Trans Med Imaging*. (2021) 40(12):3337–48. doi: 10.1109/TMI.2021.3084288
- Sun Z, Latorre F, Sanchez T, Cevher V. *A plug-and-play deep image prior*. *ICASSP 2021–2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (2021). p. 8103–7
- Cascarano P, Sebastiani A, Comes MC, Franchini G, Porta F. *Combining weighted total variation and deep image prior for natural and medical image restoration via admm*. *2021 21st International Conference on Computational Science and Its Applications (ICCSA)* (2021).
- Van Veen D, Jalal A, Soltanolkotabi M, Price E, Vishwanath S, Dimakis AG. Compressed sensing with deep image prior and learned regularization. *arXiv [Preprint]*. *arxiv.1806.06438* (2018). doi: 10.48550/arxiv.1806.06438
- Heckel R, Hand P. Deep decoder: concise image representations from untrained non-convolutional networks. *arXiv [Preprint]*. *arxiv.1810.03982* (2018). doi: 10.48550/arxiv.1810.03982

35. Arora S, Roeloffs V, Lustig M. Untrained modified deep decoder for joint denoising and parallel imaging reconstruction. *Proc Intl Soc Mag Reson Med.* (2020). p. 3585. Available online at: <https://archive.ismrm.org/2020/3585.html> (Accessed September 1, 2024).
36. Ronneberger O, Fischer P, Brox T. *U-Net: Convolutional Networks for Biomedical Image Segmentation*. Cham: Springer International Publishing (2015).
37. Korkmaz Y, Dar SUH, Yurt M, Ozbey M, Cukur T. Unsupervised MRI reconstruction via zero-shot learned adversarial transformers. *IEEE Trans Med Imaging.* (2022) 41(7):1747–63. doi: 10.1109/tmi.2022.3147426
38. Tian D, Zhao S-h, Wang Y, Lu H-f, Chen Y-y, Guo J-j, et al. Unenhanced whole-heart coronary MRA: prospective intraindividual comparison of 1.5-T ssfp and 3-T dixon water-fat separation gre methods using coronary angiography as reference. *Am J Roentgenol.* (2022) 219(2):199–211. doi: 10.2214/ajr.21.27292
39. Zhang T, Pauly JM, Vasanaawala SS, Lustig M. Coil compression for accelerated imaging with cartesian sampling. *Magn Reson Med.* (2013) 69(2):571–82. doi: 10.1002/mrm.24267
40. Berkeley Advanced reconstruction toolbox. In: Uecker M, Ong F, Tamir JI, Bahri D, Virtue P, Cheng JY, et al. editors. *Proceedings of the 23rd Annual Meeting of ISMRM*. Toronto, Canada: International Society for Magnetic Resonance in Medicine (ISMRM) (2015). p. 2486.
41. Uecker M, Lai P, Murphy MJ, Virtue P, Elad M, Pauly JM, et al. Esprit—an eigenvalue approach to autocalibrating parallel MRI: where sense meets grappa. *Magn Reson Med.* (2014) 71(3):990–1001. doi: 10.1002/mrm.24751
42. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin J-C, Pujol S, et al. 3D slicer as an image computing platform for the quantitative imaging network. *Magn Reson Imaging.* (2012) 30(9):1323–41. doi: 10.1016/j.mri.2012.05.001
43. Berglund J, Ahlstrom H, Johansson L, Kullberg J. Two-point dixon method with flexible echo times. *Magn Reson Med.* (2011) 65(4):994–1004. doi: 10.1002/mrm.22679
44. Peng H, Cheng C, Liu X, Zheng H, Zou C. *Solving Fat-Water Separation with Arbitrary Echo Combination by Phase Unwrapping*. Toronto, Canada: ISMRM & SMRT Annual Meeting & Exhibition (2023).
45. Etienne A, Botnar RM, Van Muiswinkel AM, Boesiger P, Manning WJ, Stuber M. “Soap-bubble” visualization and quantitative analysis of 3D coronary magnetic resonance angiograms. *Magn Reson Med.* (2002) 48(4):658–66. doi: 10.1002/mrm.10253
46. Botnar RM, Stuber M, Dianas PG, Kissinger KV, Manning WJ. Improved coronary artery definition with T2-weighted, free-breathing, three-dimensional coronary MRA. *Circulation.* (1999) 99(24):3139–48. doi: 10.1161/01.CIR.99.24.3139
47. Jiménez ÁB, Sra S. *Fast newton-type methods for total variation regularization*. *International Conference on Machine Learning* (2011).
48. Barbero Á, Sra S. Modular proximal optimization for multidimensional total-variation regularization. *arXiv [Preprint]. arxiv.1411.0589* (2014). doi: 10.48550/arxiv.1411.0589
49. Kingma DP, Ba J. *Adam: a method for stochastic optimization*. *The 3rd International Conference for Learning Representations*; San Diego (2014).
50. Fuin N, Bustin A, Kustner T, Oksuz I, Clough J, King AP, et al. A multi-scale variational neural network for accelerating motion-compensated whole-heart 3D coronary mr angiography. *Magn Reson Imaging.* (2020) 70:155–67. doi: 10.1016/j.mri.2020.04.007
51. Qi H, Hajhosseiny R, Cruz G, Kuestner T, Kunze K, Neji R, et al. End-to-end deep learning nonrigid motion-corrected reconstruction for highly accelerated free-breathing coronary MRA. *Magn Reson Med.* (2021) 86(4):1983–96. doi: 10.1002/mrm.28851
52. Wu X, Deng L, Li W, Peng P, Yue X, Tang L, et al. Deep learning-based acceleration of compressed sensing for noncontrast-enhanced coronary magnetic resonance angiography in patients with suspected coronary artery disease. *J Magn Reson Imaging.* (2023) 1:1–10. doi: 10.1002/jmri.28653
53. Hashimoto F, Onishi Y, Ote K, Tashima H, Yamaya T. Fully 3D implementation of the end-to-end deep image prior-based pet image reconstruction using block iterative algorithm. *PMB.* (2023) 68(15):155009. doi: 10.1088/1361-6560/ace49c
54. Heckel R, Soltanolkotabi M. *Compressive sensing with un-trained neural networks: gradient descent finds the smoothest approximation*. *Proceedings of the 37th International Conference on Machine Learning*. JMLR.org (2020). p. 388
55. Kustner T, Munoz C, Psenicny A, Bustin A, Fuin N, Qi H, et al. Deep-learning based super-resolution for 3D isotropic coronary mr angiography in less than a minute. *Magn Reson Med.* (2021) 86(5):2837–52. doi: 10.1002/mrm.28911
56. Siedler TM, Jakob PM, Herold V. Enhancing quality and speed in database-free neural network reconstructions of undersampled MRI with SCAMPI. *Magn Reson Med Sci.* (2024) 92(3):1232–47. doi: 10.1002/mrm.30114