Biomedical Signal Processing and Control 96 (2024) 106560

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Compressed sensing using a deep adaptive perceptual generative adversarial network for MRI reconstruction from undersampled K-space data

Kun Wu [a,b,*], Yan Xia [a,b], Nishant Ravikumar [a,b], Alejandro F. Frangi [c,d]

[a] *Centre for Computational Imaging and Simulation Technologies in Biomedicine (CISTIB), School of Computing, University of Leeds, Leeds, LS2 9JT, United Kingdom*
[b] *Leeds Institute for Cardiovascular and Metabolic Medicine (LICAMM), School of Medicine, University of Leeds, Leeds, LS2 9JT, United Kingdom*
[c] *Medical Imaging Research Center (MIRC), Cardiovascular Science and Electronic Engineering Departments, KU Leuven, B-3000 Leuven, Belgium*
[d] *Department of Computer Science, School of Health Sciences, University of Manchester, Manchester, M13 9PL, United Kingdom*

ARTICLE INFO

Keywords:
MRI reconstruction
Undersampling
Compressed sensing
Perceptual feature guidance
Error-correction

ABSTRACT

Objective: Magnetic resonance imaging (MRI) reconstruction from undersampled k-space data has received great interest due to its capability in reducing physical scan time. Meanwhile, the reconstruction problem is challenging because of its ill-posed and inverse nature. Nevertheless, existing compressed sensing (CS) and deep learning-based methods still need improvement as they suffer from limited generalisation ability, especially when higher undersampling factors are applied.
Methods: To recover high-quality images with reliable fine anatomical structures, we propose DAPGAN - a deep adaptive perceptual generative adversarial network which reconstructs high-quality MR images from undersampled k-space data. In particular, a novel perceptual feature guidance (PFG) mechanism is proposed which has the capability to retrieve effective features from each level that is useful in emphasising underlying anatomical structures. In addition, the model explores information in a dual-domain style.
Results: Experimental results show that the proposed method outperforms state-of-the-art baselines in terms of quantitative and qualitative evaluations. Our method improves the average SSIM (structure similarity index measurement) from 0.81 to 0.93 at a low CS ratio of 10%, compared to the average performance of competing methods on cardiac datasets using Cartesian sampling.
Conclusion: An innovative mechanism was proposed for accurate and perceptual feature guidance. It is an adaptive error-correction-based mechanism during multi-level feature reconstruction. The effectiveness was proved by the superior performance in extracting reliable anatomical details.
Significance: The architecture of our proposed model offers a new solution for accurate feature guidance, considering enhancing conventional optimisation-based problems. In particular, it is a robust mechanism in aggressive undersampling scenarios.

1. Introduction

Magnetic resonance imaging (MRI) is a widely used noninvasive medical imaging modality for clinical diagnosis and treatment planning. MRI has excellent strengths to provide outstanding contrast, especially in the field of assessing brain diseases, cardiovascular diseases, and corresponding treatment monitoring. Nevertheless, the long acquisition time caused by the hardware or patients nature always lead to discomfort to patients, which may hinders the related time-critical issues. Additionally, due to the limitations of MRI scanning, the image quality may be influenced heavily by aliasing like motion artefacts. As a result, an efficient MR image reconstruction approach that recovers high-quality images with faithful anatomical structures from undersampled data is a challenging problem.

To accelerate MR imaging, the efforts can be categorised into two directions: hardware-based parallel imaging techniques [1] and MR image reconstruction algorithms from accelerated k-space signal. As a mainstream direction, most researches focus on reconstructing MR images with undersampled data in k-space. The violation of the Nyquist sampling theorem results in aliasing artefacts when undersampled images are acquired. To solve this, the MR image reconstruction process can be formulated as an optimisation problem composed of two elements: data fidelity which ensures the consistency between reconstructed and original items in k-space, and prior regularisation, which

---

* Correspondence to: School of Computing, University of Leeds, Leeds, LS2 9JT, United Kingdom.
E-mail address: sckw@leeds.ac.uk (K. Wu).

https://doi.org/10.1016/j.bspc.2024.106560
Received 17 October 2023; Received in revised form 22 April 2024; Accepted 7 June 2024

Table 1
The configuration of investigated methods in terms of the algorithms and experiments.

	Ours	Ran et al. [8]	Wang et al. [9]	Sun et al. [10]	Yang et al. [11]	Ravishankar et al. [12]
K-space loss	✓	✓	✓	✗	✓	✗
Image-space loss	✓	✓	✓	✓	✓	✓
Parallel domain learning	✓	✓	✗	✗	✗	✗
Active adaptation-based guidance	✓	✗	✗	✗	✗	✗
Multi-medical datasets	✓	✓	✗	✗	✗	✓
Irregular/aggressive undersampling patterns	✓	✗	✗	✗	✗	✗
Ablation study	✓	✗	✓	✓	✓	✗
3D downstream analysis (clinical relevance)	✓	✗	✗	✗	✗	✗

usually includes sparsity, smoothness or spatio-temporal redundancy. This technique is well-known as compressed sensing MRI [2,3] (CS-MRI). Conventional CS methods are essentially regularised convex optimisation algorithms, most commonly based on Total Variation (TV) regularisation [4–6] and low-rank constraints [6,7]. Great progress has been made since CS-MRI-based conventional methods were proposed, nevertheless, these methods are iterative and they are heavily limited by the reconstruction speed which is time-consuming.

Recently, deep learning (DL) approaches have gained popularity for CS-MRI, as they are more efficient (i.e. faster execution/reconstruction times) and yield higher-quality image reconstructions than their conventional CS-MRI counterparts. There have been quite a few newly proposed methods, which can be roughly categorised into two types, (a) single domain-based methods and (b) dual-domain-based methods. The models performed in a single domain usually leverage the inverse Fourier transform (IFT) to obtain the original input from initial frequency domain data, then feed it to an image domain network or transfer the outputs from a network in k-space to obtain the final image domain results. The former one is the most adopted workflow for most current single-domain convolutional neural network(CNN)-based methods. For instance, Liu et al. [13] exploited the shareable features from potentially misaligned target images of an arbitrary contrast to reconstruct accelerated images based on an end-to-end CNN.Dar et al. [14] addressed the problem of data scarcity in network training for accelerated MRI by utilising transfer-learning based approach. The methods above addressed high computational cost relatively compared with conventional CS-based algorithms. Further, generative adversarial network (GAN) [15] has been proven to be effective in MRI reconstruction of image space [11,16–21]. De-aliasing generative adversarial network (DAGAN) [11] was the first study to explore using GANs for MRI reconstruction. To improve the recovery of high-frequency data, the adversarial loss was combined with a mean squared error (MSE) loss, mean absolute error (MAE) loss, perceptual loss and other non-adversarial losses. Zhou et al. [17] proposed an approach which introduced structurally strengthened connections for enhancing feature propagation, further more the in-between concatenated autoencoders were utilised except residual layers. Considering widening the channels of feature propagation in image space, Liu et al. [18] used a cross-stage skip connection between two end-to-end cascaded U-Net to form the generator. For learning the long-range dependency of feature maps, Xu et al. [20] proposed a dilated depthwise separable convolution dense block and a self-attention module which is squeeze-and-excitation lightweight. In order to consider the data consistency (DC) with k-space information for the correction of reconstructed results, several works [10,22–28] proposed to incorporate the data fidelity in the network architecture by applying linear combination between the network prediction and the known undersampled k-space measurements for the known samples, and replacing the unknown samples by the prediction of the CNN. For example, Wang et al. [9] proposed a deep cascade-based convolutional neural network to simulate the iterative reconstruction of dictionary learning-based approaches by embedding the data consistency layer into the unrolling iteration. By adding a data-sharing layer, the temporal correlation for dynamic reconstruction can be learned as an enhanced regularisation for better performance. The dual-domain approaches [28–30] adopt learning feature information in both k-space and image domain. To simultaneously reconstruct information in k-space and image space, Zhou et al. [29] proposed a recurrent network with deep T1 prior embedded for accelerating the acquisition of MRI.

We compared the proposed method with the state-of-the-art in Table 1 from the following aspects: Firstly, most comparison methods consider the reconstruction of the k-space data to aid the spatial domain through data consistency. For example, DIMENSION designs k-space networks to learn the domain information in particular. Both of our methods and MD-Recon Net addressed the sequential learning of splitting the k-space and image domains, fully exploring the internal and intersection information of the dual domains. Our method facilitates an adaptive error-correction-based feature guidance mechanism. From the aspect of the configuration for experiments, several methods utilise one type of dataset. Our method verified the superior performance on both brain and cardiac datasets. Further, we demonstrated the robustness of our model with simulated irregular undersampling patterns based on compressed sensing. Considering the clinical relevance of investigated methods, we conducted 3D downstream analysis with volumetric quantifications of key indices for both brain and cardiac MRI.

Previous works on dual-domain MRI reconstruction have proved the superiority compared with single-domain approaches. Nevertheless, the reconstruction quality is still limited, and the performance can be improved based on two observations:

1. From the view of network design, the conventional optimisation strategies adopted the combination of different loss functions which only consider the fidelity between the final output and the ground truth. Therefore, the performance is easy to suffer from blurry results or natural-looking structures which are hallucinated in the reconstructed results, because the prediction is essentially an average of all likely solutions to satisfy the multiple criteria. There are still more valuable intermediate predictions regarding different levels of reconstruction to be utilised for much more accurate MR image reconstruction.
2. From the perspective of data to be reconstructed, the increased acceleration factors of CS, and changing the target organs for reconstruction are two main terms that affect the robustness of the network when facilitating reasonable structural anatomy recovery.

To essentially address the above limitations, we propose the deep adaptive perceptual generative adversarial network (DAPGAN), a novel DL-based method for CS-MRI reconstruction from under-sampled k-space data, it performs in an adaptive supervision style to provide aliasing-robust and supervision-efficient reconstruction. Our network simultaneously leverages feature learning in both frequency and spatial domains. The main contributions of our approach are listed below:

1. We propose a perceptual feature guidance at multi-level for our network. It can be embedded in the original output of the previous multi-level CNN layers to build the adaptive mechanism which iterates between intermediate reconstructions. Such a mechanism can adaptively constrain the spatial domain information as an additional and efficient supervision which enables emphasising underlying and missed anatomical structures of multi-levels based on the previous iterations.

2. De-aliasing robustness with multi-datasets and multi-undersampling conditions. We perform extensive experiments on brain and cardiac MRI datasets. Results show the DAPGAN outperforms several state-of-the-art methods, particularly under high undersampling acceleration factors and irregular undersampling masks.
3. Clinical relevance. We address the clinical relevance of this approach by demonstrating that subsequent image analysis steps – viz. segmentation – can be performed using the reconstructed images from undersampled data, without significant influence on the biomarkers/anatomical measurements derived thereof, relative to those extracted from the original images (i.e. reconstructed from fully sampled data).

2. Problem formulation for CS–MRI

The observation and data acquisition model for image reconstruction can be regarded as a discrete linear model:

$$\boldsymbol{F}\boldsymbol{X} + \varepsilon = \boldsymbol{Y}, \tag{1}$$

where $\boldsymbol{X} \in \mathbb{C}^{\mathrm{N}}$ represents the complex-valued MR image which is desired, $\boldsymbol{Y} \in \mathbb{C}^{M}$ denotes an observation matrix that represents undersampled k-space measurements, ε is acquisition noise and $\boldsymbol{F}$ represents an undersampled Fourier coding matrix. From the CS perspective, the CS-MRI reconstruction task can be expressed as an optimisation problem:

$$\min_{x} \zeta R(\boldsymbol{x}) + \lambda \left\| \boldsymbol{F}_u \boldsymbol{x} - \boldsymbol{y} \right\|_2^2, \tag{2}$$

where $\boldsymbol{F}_u \in \mathbb{C}^{M\times N}$ denotes the undersampled Fourier encoding matrix, R denotes a regularisation term based on input vector $\boldsymbol{x}$, λ and ζ are regularisation parameters.

The deep learning-based CS-MRI reconstruction model can be formulated as follows:

$$\min_{x} \zeta \left\| \boldsymbol{x} - f_{cnn}\left(\boldsymbol{x}_u \mid \boldsymbol{\vartheta}\right) \right\|_2^2 + \lambda \left\| \boldsymbol{F}_u \boldsymbol{x} - \boldsymbol{y} \right\|_2^2, \tag{3}$$

where f_{cnn} represents a CNN parameterised by $\boldsymbol{\vartheta}$, and ζ is a regularisation parameter, $\boldsymbol{x}_u$ denotes the zero-filled reconstruction utilised as the initial input of the network. Here, the reconstruction process applied by CNN methods (i.e. $f_{cnn}\left(\boldsymbol{x}_u \mid \boldsymbol{\vartheta}\right)$) are used as alternatives for the regularisation term $R(x)$ by optimising the parameters of the model.

Inspired by the conventional regularisation term in CNN-based methods (Eq. (3)), the proposed perceptual feature guidance can be leveraged as an additional constrain to formulate the new optimisation problem:

$$\min_{x} \zeta \left\| \boldsymbol{x} - f_{cnn}\left(c\left(f_{corr}\left(\left(\boldsymbol{x} - \boldsymbol{x}_i\right) \mid \hat{\boldsymbol{\vartheta}}\right), \boldsymbol{x}_u\right) \mid \boldsymbol{\vartheta}\right) \right\|_2^2 + \lambda \left\| \boldsymbol{F}_u \boldsymbol{x} - \boldsymbol{y} \right\|_2^2, \tag{4}$$

where $\boldsymbol{x}_i$ denotes the intermediate prediction of each iteration, f_{corr} represent the correlation layers parameterised by $\hat{\boldsymbol{\vartheta}}$, which can acquire and learning the difference feature in each training iteration when we force $\boldsymbol{x}$ to be well-approximated by the reconstructed results. Then the correlation features are concatenated with the original input as a constrain for next iteration of $f_{cnn}\left(\boldsymbol{x}_u \mid \boldsymbol{\vartheta}\right)$. How to provide accurate guidance for training process to find the optimal parameters is what we focused on in this paper.

3. Methodology

In this section, we introduce the details of our proposed method – DAPGAN. Inspired by the concept of dual-domain learning (MRI Dual-domain Reconstruction Network: MD-Recon-Net) [8], we developed a novel deep learning-based framework DAPGAN as illustrated in Fig. 1. Most approaches so far have only focused on fidelity between the output and the referenced ground truth based on minimising a cost function. This mechanism conventionally derived a tendency for model training to fall into the local minima. In contrast, our approach handles image reconstruction by proposing a self-adaptive guidance mechanism to address the problem relatively. The generator comprises two parallel convolutional neural networks: a multi-level perceptual feature guidance network (PFG-Net) based on U-Net architecture for real-space (image-domain) data and a standard U-Net for processing k-space data. The structure of PFG-Net is shown in Fig. 2, and PFG-Net is driven by the mechanism of perceptual feature guidance (PFG), which is presented in the following subsections.

3.1. *Perceptual Feature Guidance mechanism of PFG-Net*

We construct the pyramid Perceptual Feature Guidance (PFG) as an active-tuning-based deep conditioning mechanism, which consists of three steps (as Fig. 2.A shows): PFE (a VGG-based extension) for multi-scale feature extraction with shared weights, in which two parallel branches take reference signals (fully sampled or multi-accelerated data) and intermediate prediction as input respectively, instead of using the VGG-Net as the loss network only [31], we employ the proposed PFE as a pre-trained and pure multi-level feature extractor as the initial step of PFG; a pyramidal correlation layer between the corresponding two branches; a pyramidal generation with correlation features from pairs of frames. Thus, the adaptive tuning function of our generator is facilitated by learning the dynamic correlation guidance alongside training iterations. For model testing and validation, we employ the architecture with trained parameters in the training phase and leave out PFG (see Fig. 2.B).

3.1.1. *Perceptual feature extractor*

As the first stage of the PFG mechanism, PFE includes the following process: the reconstruction results of the last iteration are processed as the intermediate prediction IP and then fed with the ground truth GT jointly to the PFE to obtain the paired output $\boldsymbol{PFE}_{\boldsymbol{k}}^{\boldsymbol{GT}}$ and $\boldsymbol{PFE}_{\boldsymbol{k}}^{\boldsymbol{IP}}$ as the inputs of the next stage (see Fig. 2.A). PFE is an extension of the VGG-Net [32] (Fig. 3). Previous approaches [33] have shown that preserving the negative features with reference to the inputs is beneficial for MRI reconstruction. Motivated by this, we employed recursive residual learning to enhance the capacity of the VGG in the multi-level feature extraction task. Each layer of PFE contains three pre-activation residual blocks (PRB), and each PRB comprises batch normalisation (BN), a pre-activation unit (ReLU) and convolutional layers (in sequence) as sub-blocks with skip connections.

Considering $\boldsymbol{R}_{\boldsymbol{i-1}}$ and $\boldsymbol{R}_{\boldsymbol{i}}$ as the input and output of the ith PRB module, respectively, the PRB can be expressed as

$$\boldsymbol{R}_{\boldsymbol{i}} = f_{PRB}\left(\boldsymbol{R}_{\boldsymbol{i-1}}\right) + \boldsymbol{R}_{\boldsymbol{i-1}} = \boldsymbol{W}_{\boldsymbol{RB}}\left(\boldsymbol{BN}\left(r_a\left(\boldsymbol{R}_{\boldsymbol{i-1}}\right)\right)\right) + \boldsymbol{R}_{\boldsymbol{i-1}}. \tag{5}$$

where $f_{PRB}(\bullet)$ is a function of the residual network, $\boldsymbol{W}_{\boldsymbol{RB}}$ represents the weights of the convolutional network, r_a and $\boldsymbol{BN}$ denote ReLU and batch normalisation which included in the pre-activation function.

We extract the convolution and pooling layers from the standard VGG-Net and concatenate them with the PRB modules. We can represent the expression as follows:

$$SP_k = \boldsymbol{W}_{\boldsymbol{k}} \otimes \left[f_c\left(\boldsymbol{R}_{\boldsymbol{k}}^{\boldsymbol{i}} + \boldsymbol{X}_{\boldsymbol{k}}\right)\right] + b_k, \tag{6}$$

where SP_k represents a sub-unit of PFE, k denotes the sequential level of feature extraction that corresponds to multi-scale features in the encoder, $\boldsymbol{W}_{\boldsymbol{k}}$ represents the convolution network weight of each block in standard VGG, b_k is a bias, f_c denotes the concatenation and fusion of the PRB $\boldsymbol{R}_{\boldsymbol{k}}^{\boldsymbol{i}}$ and initial input $\boldsymbol{X}_{\boldsymbol{k}}$, $\otimes$ denotes the convolution operation.

The input for the PFE is the pairs of ground truth $\boldsymbol{X}_{\boldsymbol{GT}}$ and the intermediate predicted $\boldsymbol{X}_{\boldsymbol{IP}}$ over the iterations during training; we instantiate Eq. (6) as

$$\begin{aligned} PFE_k &= SP_k\left(\boldsymbol{X}_{\boldsymbol{GT}}, \boldsymbol{X}_{\boldsymbol{IP}}\right) \\ &= \boldsymbol{W}_{\boldsymbol{k}} \otimes \left[f_c\left(\boldsymbol{R}_{\boldsymbol{k}}^{\boldsymbol{i}} + \boldsymbol{X}_{\boldsymbol{k}}\right)\right] + b_k \mid \boldsymbol{X}_{\boldsymbol{GT}}, \boldsymbol{X}_{\boldsymbol{IP}}. \end{aligned} \tag{7}$$

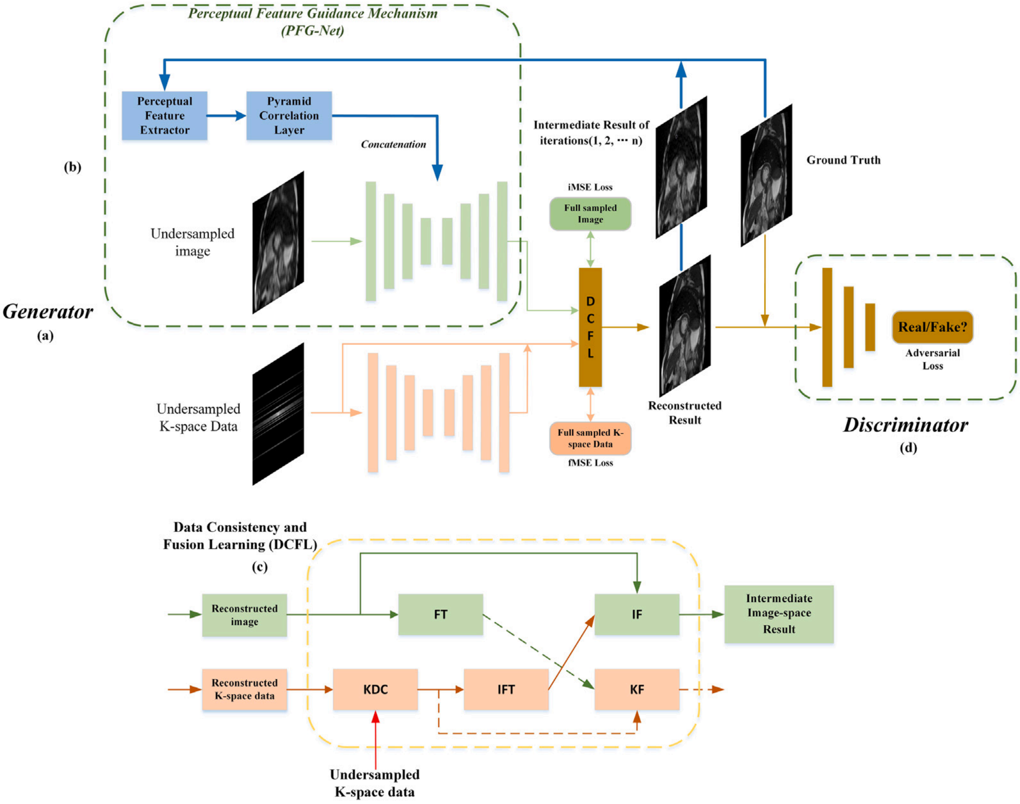
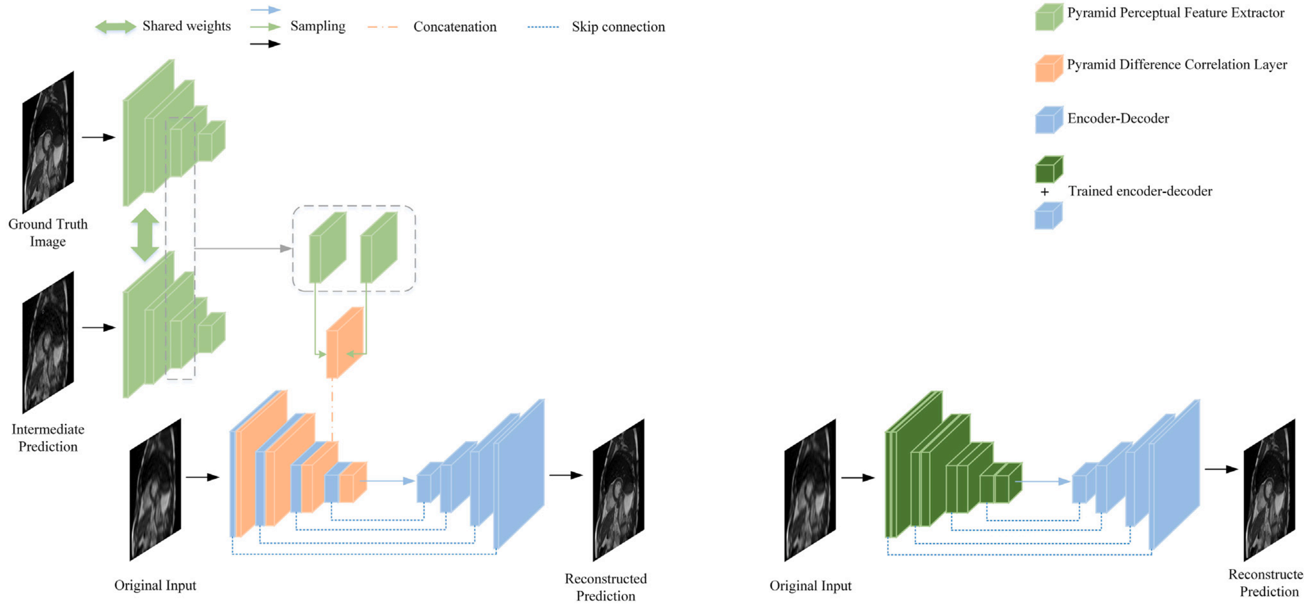


Fig. 1. Network architecture of DAPGAN for high-quality MRI reconstruction. (a) The generative network takes the undersampled k-space data and the zero-filling images (with inverse Fourier transform of undersampled k-space data) as inputs in a parallel learning fashion, and outputs reconstructed images. (b) For the image domain, the proposed PFG-Net is a U-Net-based network driven by the PFG mechanism; for the k-space domain, we utilise a basic U-Net to reconstruct the frequency information. (c) Data consistency module and fusion learning layers are utilised in the cross-domain for information sharing. (d) The discriminator distinguishes the reconstructed images from the fully sampled images. (KDC: K-space Data Consistency, FT: Fourier Transform, IFT: Inverse Fourier Transform, IF: Image-space Fusion, KF: K-space Fusion).



(A) Pyramid Perceptual Feature Guidance for training phase

(B) Trained model for testing and validation phase

Fig. 2. The overall schematic architecture of proposed pyramid Perceptual Feature Guidance (PFG) mechanism for deep conditioning reconstruction and PFG-Net. (A) The pipeline includes PFE, Pyramid multi-level correlation layers and concatenation-generation layers for the encoder at different scales. (B) The testing and validation phase of the generator.

3.1.2. Pyramidal Correlation and Generation (PCG)

For PCG, as Fig. 2.A shows, there are two stream branches utilised as inputs: $\boldsymbol{PFE_k^{GT}}$ and $\boldsymbol{PFE_k^{IP}}$ represent the pyramidal outputs from the PFE, the pairwise features are inputs of pyramidal correlation layer, the original features denoted by $\boldsymbol{X_u}$ from the convolutional block as the backbone of the generator. Along the iteration axis, we match the

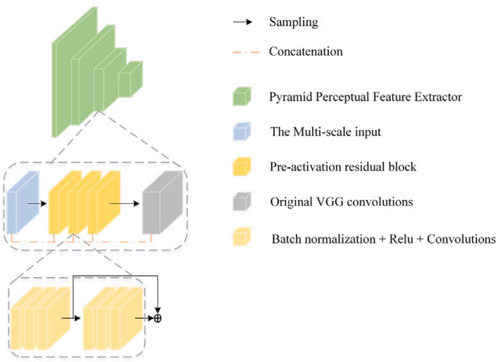


Fig. 3. Overview of the proposed pyramid perceptual feature extractor, for exploring multi-level features without information redundancy, a pre-activation module with residual option is employed, inheriting the VGG-16 architecture before fully connected layer.

correlation features to the primary encoder at four distinct scales to form the generation layers, which are 256 × 256, 128 × 128, 64 × 64, 32 × 32, respectively. The outputs from convolution and residual blocks are concatenated with the correlation feature maps in the form of $\left[f_{\mathrm{Conv}\ +RB}\left(\boldsymbol{X}_{\boldsymbol{u}}\right), f_{\mathrm{Corr}}\ \right]$. The correlation layer comprises convolutional layers and activation options (ReLU). We formulate the PFG mechanism as follows:

$$\begin{aligned}\boldsymbol{F}_{PFG} = \boldsymbol{W}_{\boldsymbol{k}}^{\boldsymbol{DAP}} \otimes \Big(f_c \Big(\boldsymbol{W}_{\boldsymbol{k}}^{\boldsymbol{Corr}} \otimes \left(\boldsymbol{PFE}_{\boldsymbol{k}}^{\boldsymbol{GT}} - \boldsymbol{PFE}_{\boldsymbol{k}}^{\boldsymbol{IP}} \right), \\ f_{\mathrm{Conv}\ +RB}^{k} \left(\boldsymbol{X}_{\boldsymbol{u}} \right) \Big) \Big) \\ +\ b_k^{DAP},\end{aligned} \tag{8}$$

where, $\boldsymbol{W}_{\boldsymbol{k}}^{\boldsymbol{DAP}}$ represents the weight and activation function of the spatial-domain reconstruction network PFG-Net at the convolutional scale k, $\boldsymbol{W}_{\boldsymbol{k}}^{\boldsymbol{Corr}}$ represents the weights of the convolutional block and activation function in correlation layers, from where we obtain the correlation features f_{Corr} of level k in the mechanism, f_c represents the concatenation function, b_k^{DAP} denotes the bias of layer k and ⊗ represents a convolution operation.

3.2. *Data consistency and dual-domain fusion*

Motivated by the concepts of CDNN [22] and MD-Recon Net [8], we use the data consistency of the k-space to incorporate data fidelity into the network architecture and employ a fusion block to enhance dual-domain learning.

A closed-form solution of Eq. (3) is applied for k-space data consistency (KDC module in Fig. 1)

$$DC(j) = \begin{cases} \hat{\boldsymbol{Y}} & \text{if} \quad \boldsymbol{j} \notin \Omega \\ \frac{\hat{\boldsymbol{Y}} + \lambda_0 \boldsymbol{U}(\boldsymbol{j})}{1+\lambda_0} & \text{if} \quad \boldsymbol{j} \in \Omega \end{cases}, \tag{9}$$

where, $\hat{\boldsymbol{Y}}$ denotes the predicted value learned from the k-space network. $\boldsymbol{j}$ denotes the index of k-space data, $\boldsymbol{U}(\boldsymbol{j})$ represents the undersampling value in the frequency domain. The DC block is used to perform linear fitting between the network prediction and original measurements. The hyper-parameter λ_0 is related to the imaging noise level, which is determined during training and it controls the constraints of the original k-space measurements, Ω represents the index set of the obtained k-space samples. If the k-space coefficient is initially not sampled ($\boldsymbol{j} \notin \Omega$), implying that we use the network prediction results as the final results. For the sampled k-space entries, we calculate it linearly and weighted by λ_0 thus we combine the original sampled data with the unconditionally embedded network prediction. The forward and backward passes of the k-space data consistency can be derived easily.

As the merit of dual-domain reconstruction is taking advantage of the reconstructed details that only k-space domain network can extract, to further improve the ability of PFG-Net in image-domain, the iterative outputs of k-space network and KDC module are interactive via the image domain fusion (IF) layers, thus, it is utilised as one of the input for reconstructed intermediate results (see Fig. 1). We employ fusion blocks as the last step of each DC block. For k-space fusion (KF), the prediction from the branch network of the k-space is fed into the k-space data consistency (KDC) module with the original undersampled k-space data. Then, KF takes the output of KDC and the output of FT which is the transformed prediction of the image-domain network, as inputs.

3.3. *Optimisation*

The proposed approach is trained as a conditional GAN, where the perceptual feature guidance maps are used as the conditioning variables to drive image reconstruction. The network is trained end-to-end by minimising the adversarial loss and normalising MSE between the reconstructed data and fully sampled data. In particular, the MSE loss is defined in both the k-space and image domains to enforce data consistency across both domains.

3.3.1. *Multi-scale perceptual loss*

Let Θ be the set of all the learnable parameters in the model, which includes the perceptual feature pyramid extractor and the encoder at different pyramid levels. Let n denote the iteration number, and $\boldsymbol{W}_{\boldsymbol{n-1}}^{\boldsymbol{l}}$ denote the intermediate prediction at the lth pyramid level reconstructed by our model, and $\boldsymbol{W}_{\boldsymbol{GT}}^{\boldsymbol{l}}$ the corresponding reference signal. The basic training process can be formulated as:

$$\mathcal{L}(\Theta_n) = \sum_{l=l_0}^{L} \alpha_l \sum_{\boldsymbol{X}} \left| \boldsymbol{W}_{n-1}^{l}(\boldsymbol{X}) - \boldsymbol{W}_{\boldsymbol{GT}}^{l}(\boldsymbol{X}) \right|_2 + \gamma |\Theta|_2, \tag{10}$$

where we compute the $\boldsymbol{L_2}$ norm $|\cdot|_2$ between the intermediate results and the ground truth in a multi-scale style. The second term denotes the regularisation of the model. We designed robust training loss for fine-tuning:

$$\mathcal{L}(\Theta_n) = \sum_{l=l_0}^{L} \alpha_l \sum_{\boldsymbol{X}} \left(\left| \boldsymbol{W}_{n-1}^{l}(\boldsymbol{X}) - \boldsymbol{W}_{\mathrm{GT}}^{l}(\boldsymbol{X}) \right| + \epsilon \right)^q + \gamma |\Theta|_2, \tag{11}$$

where $|\cdot|$ denotes the MAE norm, and we set further $q < 1$ to give less penalty to the short-range difference of perceptual fields, further, ϵ denote a constant.

3.3.2. *Conditional adversarial loss*

We formulate the conditional GAN as follows:

$$\begin{aligned}\min_{\theta_G} \max_{\theta_D} \mathcal{L}_{cGAN}\left(\theta_D, \theta_G\right) = \mathrm{E}_{X_g \sim p_{\mathrm{train}\ (X_g)}} \left[\log D_{\theta_D} \left(\boldsymbol{X}_{\boldsymbol{g}} \mid \boldsymbol{Y} \right) \right] \\ + \mathrm{E}_{X_u \sim p_{G(X_u)}} \left[\log \left(1 - \left(D_{\theta_D} \left(G_{\theta_G} \left(\boldsymbol{X}_{\boldsymbol{u}} \mid \boldsymbol{Y} \right) \right) \right) \right) \right],\end{aligned} \tag{12}$$

where, $\boldsymbol{X}_{\boldsymbol{u}}$ denotes the undersampled data and $\boldsymbol{X}_{\boldsymbol{g}}$ represents the fully-sampled ground-truth data; Y denotes the guidance information that acts as conditioning variables, with G_{θ_G} and D_{θ_D} representing the generator and discriminator networks parameterised by θ_G and θ_D respectively; and $p_{\mathrm{train}\ (X_g)}$ and $p_{G(X_u)}$ denote the data distribution induced by the discriminator and generator respectively.

3.3.3. *Dual-domain MSE loss*

As the aforementioned PFG is additional supervision when considering the MSE loss as the basic content loss to improve the quality of our reconstruction results, here we define MSE loss across the dual domain. This helps overcome the non-smooth inspection, which lacks coherent image details, based solely on the optimisation of normalised MSE which is defined by pixel-wise difference.

The MSE loss in the image-space domain is formulated as follows:

$$\min_{\theta_G} \mathcal{L}_{iMSE}\left(\theta_G\right) = \frac{1}{2} \left\| \boldsymbol{X}_{\boldsymbol{g}} - \hat{\boldsymbol{X}}_{\boldsymbol{u}} \right\|_2^2, \tag{13}$$

where X_g represents the ground truth, and $\hat{\boldsymbol{X}}_{\boldsymbol{u}}$ represents the reconstructed image.

Alternatively, the MSE formulation for the frequency domain is expressed as:

$$\min_{\theta_G} \mathcal{L}_{fMSE}\left(\theta_G\right) = \sum_{i=1}^{m} \frac{1}{2} \left\| \boldsymbol{K}_{fs} - \boldsymbol{K}^{i}_{\boldsymbol{DC}} \right\|_2^2, \tag{14}$$

where $\mathcal{L}_{fMSE}$ denotes the MSE loss in the frequency domain, $\boldsymbol{K}_{fs}$ represents the full-sampled k-space data and $\boldsymbol{K}^{i}_{\boldsymbol{DC}}$ denotes the output of the ith generation sub-net with the data-consistency branch.

Consider that K^i denotes the generation networks with convolution blocks in the k-space domain. The forward-pass process can be expressed as

$$\begin{cases} \boldsymbol{K}^1 = \sigma\left(\boldsymbol{W}^1 \otimes \boldsymbol{K}_{\boldsymbol{u}} + \boldsymbol{b}^1\right) \\ \quad\vdots \\ \boldsymbol{K}^i = \sigma\left(\boldsymbol{W}^i \otimes \boldsymbol{K}^{i-1} + \boldsymbol{b}^i\right) \\ \boldsymbol{K}^{i}_{\boldsymbol{DC}} = KDC\left(\boldsymbol{K}^i\right) \end{cases} \tag{15}$$

where $\boldsymbol{K}^{i}_{\boldsymbol{DC}}$ represents the results from the KDC (k-space-domain data consistency). $\boldsymbol{W}^i$ and $\boldsymbol{b}^i$ denote the weight and bias of the ith filter, respectively, and $\otimes$ denotes a convolution operation.

We express the total loss function as

$$\mathcal{L}_{\text{TOTAL}} = a\mathcal{L}(\Theta_n) + b\mathcal{L}_{cGAN} + c\mathcal{L}_{iMSE} + d\mathcal{L}_{fMSE}. \tag{16}$$

where a, b, c and d are weighting factors that control the trade-off between the different terms in the overall loss function.

4. Experimental setup

4.1. Datasets

We trained and evaluated the proposed DAPGAN using a publicly available brain MRI dataset, namely, Calgary Campinas 359[1] (CC-359), and cardiac MRI data from the UK Biobank [34] (UKBB). We applied different undersampling schemes/patterns to the datasets and reconstruct the undersampled data using the proposed method.

In the CC-359 brain dataset, images were acquired from a clinically used MR scanner (Discovery MR750, GE Healthcare, Waukesha, WI). Thirty-five subjects were used, and the data were divided into training (70%), validation (15%) and test (15%) sets. The dataset contains sequential two-dimensional slices stacked as volumetric images, with a size of 170 × 256 × 256.

In the UKBB dataset, cardiac MRI data were acquired using a clinical wide bore 1.5T MR system (MAGNETOM Aera, Syngo Platform VD13 A, Siemens Healthcare, Erlangen, Germany) equipped with an 18-channel anterior body surface coil (45 mT/m and 200 T/m/s gradient system). Further, 2D cine b-SSFP SAX (short axis) image stacks were acquired with the following acquisition protocol: in-plane spatial resolution of 1.8×1.8 mm, slice thickness of 8 mm, slice gap of 2 mm and image size of 198 × 208. We used 250 subjects to train the model. For each subject's cine-MR sequence, we selected two time points, namely, at end diastolic (ED) and end systolic (ES) frames respectively, resulting in 5000 slices (10 slices per time point); 1000 slices at ED and ES of another 50 subjects were used as validation and test data respectively.

4.2. Implementation details

We propose to use a U-Net-based architecture to construct the generators for image-space and k-space. For image space, the PFG-Net consists of 5 convolutional layers as encoding layers with 32 filters in each layer, followed by leaky ReLU activation functions with negative slope 1×10^{-2}, and 7 decoding layers with 32, 32, 32, 32, 32, 2 and 2 filters followed by leaky ReLU layers, respectively. Skip connections are introduced to connect mirrored layers between the encoder and decoder of matched scales. All kernels of the encoder are set to 3 × 3, the stride is set to 2 except for the first layer (set to 1), and padding is set to 1. Pyramidal correlation features are concatenated between convolutional layers and residual blocks in each encoder block to form the final generation layer. For the decoder of PFG-Net, all the kernels size are set to 4 × 4, and we set stride to 2, while padding is set to 1. As for the generator of k-space, it consists of 5 convolutional layers for encoder and decoder respectively, followed by batch normalisation and leaky ReLU layers. On the other hand, the discriminator D is formed using a standard CNN with 9 convolutional layers, followed by batch normalisation and leaky ReLU layers.

The proposed approach was trained using the Adam optimiser, with an empirical learning rate of 5×10^{-5}. The specific settings of the optimiser are $\alpha = 5 \times 10^{-5}$, $\beta_1 = 0.9$ and $\beta_2 = 0.999$. For the training procedure, we performed 500 iterations. In Eqs. (2)–(4), λ and ζ are determined as learning parameters during network training. The hyper-parameters a, b, c and d (weighting factors in the total loss function, refer to Eq. (16)) were set to 0.5, 0.5, 0.05 and 0.05 respectively. It is worth noting that, from the perspective of practical experience, these hyperparameters were determined empirically with the average value between the brain and cardiac datasets during fine-tuned training. From the view of theoretical design of feature guidance, the proposed method is robust with minimal parameter tuning when different anatomical structures are applied.

Our training strategy is to pre-train the PFE using public brain and cardiac MRI datasets as the first step, then utilise the pre-trained PFE as an additional feature extractor during PFG to leverage correlation feature guidance in the training and fine-tuning processes of our model with CC-359 and UKBB. The PFG plays a semi-supervision role in the training process and we used a trained model to infer the independent testing and validation dataset.

4.3. Competing methods

Competing methods that the proposed approach, DAPGAN, is compared against in this study can be divided into three categories: (i) DLMRI [12] and DAGAN [11] representing conventional methods without k-space data consistency; (ii) DIMENSION [9] and ADMM-CS Net [10] involving reconstruction that leverages data consistency in the frequency and spatial domains and (iii) MD-Recon Net [8] that is most related to the proposed method, and was the first study to employ dual-domain learning. The proposed DAPGAN inherits the concept of dual-domain learning with data consistency, and implements a mechanism that leverages multi-level PFG in the spatial domain. This is demonstrated later in the ablation study, in which we evaluate the influence of different network components at retaining key structural and morphological characteristics of the anatomical regions of interest in the case of the cardiac MRI data.

4.4. Evaluation design

We evaluated the performance of the proposed DAPGAN and investigated competing methods to reconstruct high-resolution imaging data, from various versions of undersampled imaging data, generated using different types of undersampling patterns/schemes. Two different imaging data sets (i.e. brain and cardiac MRI) were used across all experiments conducted in this study. The performance of the proposed

---

[1] https://sites.google.com/view/calgary-campinas-dataset/home/mr-reconstruction-challenge.

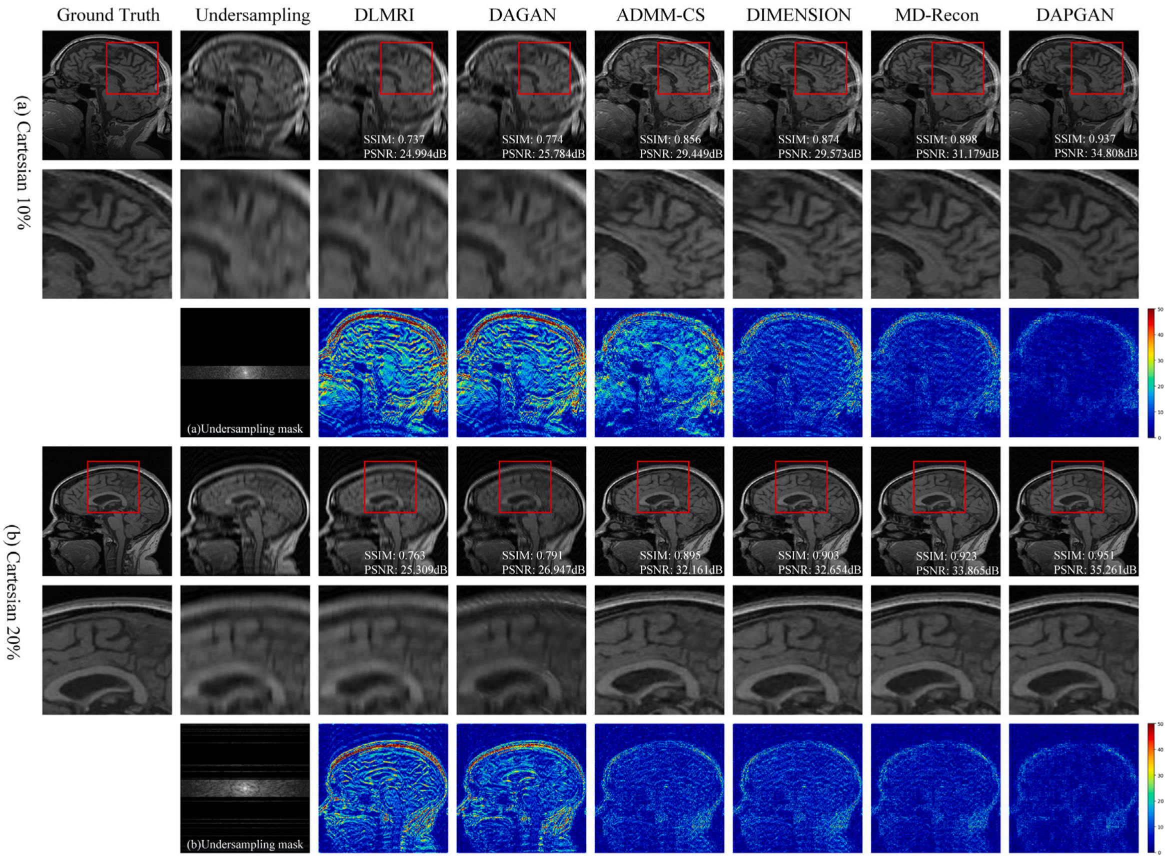


Fig. 4. Visual inspection results of CS-MRI reconstruction on brain images using regular Cartesian sampling at rates 10% (from first row to third row) and 20% (from fourth row to last row). Here we show the comparison between the ground truth, undersampled results, and the investigated methods. The third column to last column are the results of different methods. The second and fifth rows represent the magnifying views of the red boxes in the original images. The third row and last row show the undersampling masks and the absolute difference maps (error maps) indicating the regions with multi-degrees of distinctions. The quantitative evaluation metrics (SSIM/PSNR) of the reconstructed images are indicated in the first and fourth rows.

approach was evaluated and compared against competing methods in terms of image reconstruction quality, and the quality of downstream analyses (e.g. image segmentation, biomarker quantification) afforded by images reconstructed from undersampled data.

To demonstrate the flexibility of the proposed approach, we simulate two types of k-space masks for generating undersampled data: variable-density Cartesian and random at sampling rates of 10%, 20%, equal to acceleration factors of ×10 and ×5. To evaluate the robustness of the proposed model under aggressive undersampling conditions, we design two challenging sampling patterns with sampling rates of 10% and 20% based on the Cartesian sampling rules.

Two quantitative evaluation metrics are considered to assess the reconstructed results in terms of structure similarity index measurement (SSIM) and peak signal-to-noise ratio (PSNR); these metrics are used to evaluate the structural similarity and information-to-noise ratio, respectively. The fully sampled data was used for the ground-truth in training and the reference in evaluation. Higher SSIM and PSNR scores mean better reconstruction, which can be verified by the retention of perceptual features and artefact removal indicated by visual inspection simultaneously.

We address several aspects of clinical relevance by demonstrating that subsequent image analysis steps, such as segmentation, can be performed using images reconstructed by the proposed method, producing results comparable to those derived using the original/fully sampled data on both brain and cardiac MRI datasets. Then we compared volumetric indices/biomarkers extracted from segmented images, in addition to comparing the quality of segmentations achieved by 3 state-of-the-art methods.

5. Results

5.1. Brain MRI reconstruction

We trained and evaluated DAPGAN across a series of experiments using brain MRI data, where images were undersampled to different degrees using both regular and irregular undersampling strategies, i.e. based on Cartesian grid and random undersampling patterns. DAPGAN was trained to reconstruct high-resolution images from their undersampled counterparts, and was compared against several state-of-the-art image reconstruction methods.

5.1.1. Regular undersampling reconstruction

As shown in Table 2, when using the Cartesian mask with a sampling rate of 20%, DLMRI and DAGAN achieved SSIM values in the range of 0.750–0.800. ADMM-CS Net and DIMENSION achieved a better SSIM value of ∼0.90, while MD-Recon Net yielded an SSIM value of 0.923 and the highest PSNR value of 33.865 dB among all competing methods. Our DAPGAN achieved the best SSIM value of 0.951 and reached a PSNR value of 35.261 dB. At a sampling rate of 10% that results in a more aggressive reconstruction problem, the performance of all state-of-the-art methods and that of our proposed method decreased, as expected. Whilst DIMENSION, ADMM-CS Net and MD-Recon Net achieved SSIM values in the range of 0.850–0.900, our DAPGAN still achieved an SSIM value of >0.900, reaching 0.937. The decrease in the SSIM values of these three competing methods was ∼3%, whilst that of our method was <2%. The PSNR values of the state-of-the-art methods reduced by ∼3 dB, whereas that of our DAPGAN was <1 dB.

Table 2

Quantitative comparison of CS-MRI reconstruction on brain data using Cartesian and random sampling at different rates in terms of SSIM and PSNR (dB). To further verify the robustness of investigated methods, Type I and Type II of Cartesian sampling are conducted as challenging masks, with multi-irregular phase-encoding gradients.

Methods	Cartesian (Mean ± Std.)								Random (Mean ± Std.)			
	10%		Type I of irregular 10%		Type II of irregular 10%		20%		10%		20%	
	SSIM	PSNR (dB)	SSIM	PSNR (dB)	SSIM	PSNR (dB)	SSIM	PSNR (dB)	SSIM	PSNR (dB)	SSIM	PSNR (dB)
DLMRI	0.737 ± 0.087	24.994 ± 2.718	0.701 ± 0.084	24.260 ± 2.681	0.673 ± 0.082	24.288 ± 2.443	0.763 ± 0.079	25.309 ± 2.995	0.663 ± 0.071	24.906 ± 3.104	0.757 ± 0.049	28.150 ± 2.988
DAGAN	0.774 ± 0.076	25.784 ± 2.927	0.765 ± 0.060	26.174 ± 2.774	0.719 ± 0.063	25.649 ± 2.528	0.791 ± 0.067	26.947 ± 2.952	0.737 ± 0.049	27.364 ± 2.863	0.785 ± 0.042	29.132 ± 2.726
ADMM-CS Net	0.856 ± 0.034	29.449 ± 2.370	0.805 ± 0.047	27.265 ± 2.651	0.818 ± 0.040	28.947 ± 2.564	0.895 ± 0.026	32.161 ± 2.411	0.824 ± 0.030	30.098 ± 2.348	0.853 ± 0.027	31.683 ± 2.301
DIMENSION	0.874 ± 0.037	29.573 ± 2.870	0.835 ± 0.038	28.594 ± 2.466	0.841 ± 0.038	29.598 ± 2.579	0.903 ± 0.024	32.654 ± 2.344	0.844 ± 0.027	30.865 ± 2.237	0.864 ± 0.025	32.145 ± 2.258
MD-Recon	0.898 ± 0.031	31.179 ± 2.554	0.855 ± 0.040	29.259 ± 2.758	0.868 ± 0.032	30.839 ± 2.438	0.923 ± 0.020	33.865 ± 2.236	0.869 ± 0.024	31.859 ± 2.188	0.880 ± 0.023	32.838 ± 2.228
DAPGAN	**0.937 ± 0.017**	**34.808 ± 2.188**	**0.923 ± 0.023**	**32.861 ± 2.126**	**0.925 ± 0.015**	**34.092 ± 2.033**	**0.951 ± 0.018**	**35.261 ± 2.192**	**0.911 ± 0.019**	**33.368 ± 2.375**	**0.915 ± 0.025**	**34.222 ± 2.180**

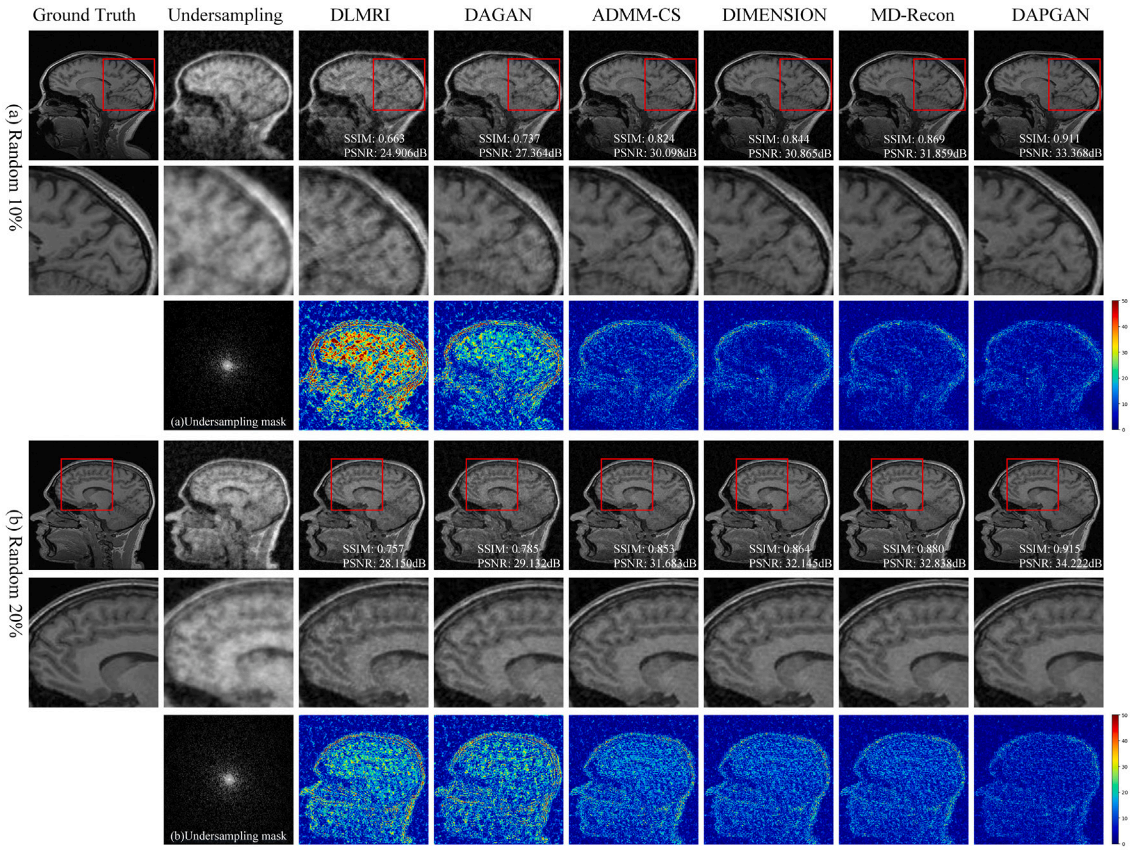


Fig. 5. Visual inspection results of CS-MRI reconstruction on brain images using regular random sampling at rates 10% (from first row to third row)and 20% (from fourth row to last row), here we show the comparison between the ground truth, undersampled results and investigated methods. The second column (Undersampling) of the first and fourth rows represent the random sampling pattern in image space. The third column to the last column are the results of different methods. The second and fifth rows represent the magnifying views of the red boxes in the original images. The third row and last row show the undersampling masks and the absolute difference maps (error maps) indicating the regions with multi-degrees of distinctions. The quantitative evaluation metrics (SSIM/PSNR) of the reconstructed images are indicated in the first and fourth rows.

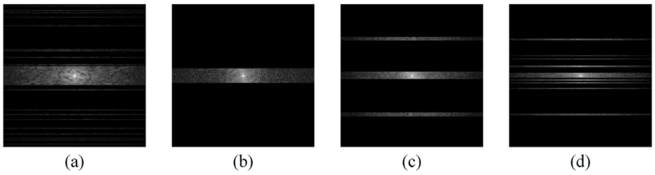

(a) (b) (c) (d)

Fig. 6. Examples of regular and irregular Cartesian undersampling masks. (a) Regular Cartesian mask for 20% sampling, (b) Regular Cartesian mask for 10% sampling, (c) Type I of irregular Cartesian mask for 10% sampling, (d) Type II of irregular Cartesian mask for 10% sampling.

When a random mask was used, our DAPGAN achieved the highest SSIM values of 0.911 and 0.915 at sampling rates of 10% and 20%, respectively. Only our method could reach SSIM values of >0.900. Moreover, our method achieved the highest PSNR value at the aggressive sampling rate of 10% and maintained competitive performance at 20%.

The reconstruction results of five competing methods and our proposed method using different sampling strategies are shown in Figs. 4 and 5, respectively. Zoomed-in images indicated by red boxes and reconstruction error maps presented below the reconstructed images clearly highlight the improvements in image quality achieved by DAPGAN relative to all other competing methods.

When the Cartesian sampling rate is modified to 10%, our approach reconstructed the most accurate perceptual textures of underlying images with the smallest differences observed from the error maps (see Fig. 4). Images reconstructed using DLMRI and DAGAN contained significant artefacts, with poor resolution and blurring of structural and textural details throughout the brain. While ADMM-CS suppressed a large proportion of such artefacts, recovering structural information more accurately than DLMRI and DAGAN throughout the brain, the error map indicates it still failed to preserve realistic and detailed textural information. DIMENSION and MD-Recon Net were the competitive approaches in terms of visual inspection, which reduced most of the artefacts, but still suffered defects of unrealistic textures and loss of structural details in some specific regions, compared with our method. When we set the sampling rate to 20%, our method again achieved the best reconstruction results based on visual inspection of the reconstructed images and corresponding error maps (shown in Fig. 4), with all other approaches achieving higher quality image reconstructions than when a sampling rate of 10% was used.

With regards to random undersampling at different sampling rates (see Fig. 5), again, the proposed method recovered high-frequency content better than the competing methods and captured detailed structural information more accurately than the rest (as evidenced by the error maps presented in Fig. 5). As seen before in the Cartesian undersampling scenario, images reconstructed using DLMRI and DAGAN contained severe structural and textural artefacts. The other approaches, namely ADMM-CS Net, DIMENSION and MD-Recon, were more successful in suppressing artefacts, but still exhibited higher errors than DAPGAN.

5.1.2. *Irregular Cartesian undersampling reconstruction*

Among types of k-space imaging trajectories, Cartesian sampling is the most common type, where each line represents a phase-encoding gradient of frequency-encoding readout. All lines in the raster are

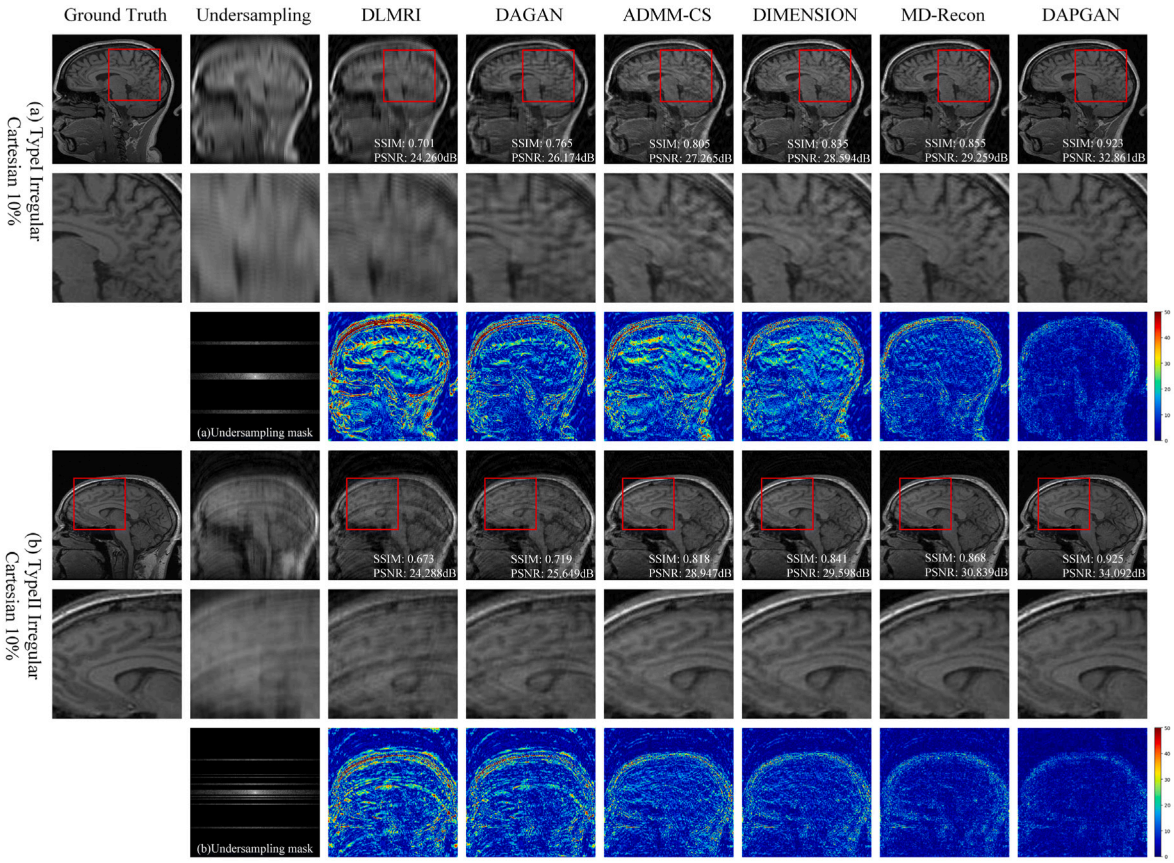


Fig. 7. Visual inspection results of CS-MRI reconstruction on brain images using two types of irregular Cartesian sampling at rates 10%, here we show the comparison between the ground truth, undersampled results and investigated methods. The third column to the last column are the results of different methods. The second and fifth rows represent the magnifying views of the red boxes in the original images. The third row and last row show the irregular undersampling masks and the absolute difference maps (error maps) indicating the regions with multi-degrees of distinctions. To further verify the robustness of investigated methods, Type I and Type II of Cartesian sampling are conducted as challenging masks, with multi-irregular phase-encoding gradients. The lines in the raster of Type I reduce the density of central region compared with regular masks, while the space between lines in the raster of Type II is provided with random values, and the lines are randomly located. The quantitative evaluation metrics (SSIM/PSNR) of the reconstructed images are indicated in the first and fourth rows.

provided in parallel as a sampling mask in k-space. Typical Cartesian sampling employs equally spaced lines in the raster, which can be reconstructed consistently after IFT. Usually, partial or parallel acquisition is adopted to save scan time. As illustrated in Fig. 6, to further verify the robustness of investigated methods, two types of Cartesian sampling are investigated (see Fig. 6c, d), namely, Type I and Type II, to generate undersampling masks under a fixed sampling rate of 10%, with multi-irregular phase-encoding directions. The lines in the raster of Type I reduce the density of the central region compared with regular masks (instead of only acquiring data at the central region, we also acquire the k-space data towards the top and bottom directions), while the space between lines in the raster of Type II is provided with random values, and the lines are randomly located.

Quantitative image reconstruction results for the undersampling with the Type I and Type II irregular Cartesian undersampling masks are presented in Table 2. The SSIM yielded by DLMRI were both ~ 0.70, DAGAN achieved SSIM values of 0.70–0.80 for both types, outperforming DLMRI marginally. DIMENSION and ADMM-CS Net meanwhile, improved image reconstruction quality for both Type I and II undersampled data, relative to DAGAN and DLMRI, achieving SSIM values in the range of 0.80–0.85. Similarly, MD-Recon Net afforded further improvement, achieving SSIM values in the range of 0.85–0.90. The proposed approach outperformed all competing methods for both types of irregular undersampled data, consistently achieving SSIM values >0.90. The performance of all competing methods in terms of PSNR, for Type I undersampled data, was <30 dB. DAPGAN consistently outperformed all methods in this regard, achieving PSNR of >30 dB. For the Type II undersampled data, both MD-Recon Net and DAPGAN achieved PSNR values >30 dB, while the PSNR values for all other methods were <30 dB.

Fig. 7 provides a visual comparison of image reconstruction results achieved using several state-of-the-art approaches and the proposed approach, for two types (i.e. Type I and Type II) of Cartesian undersampling patterns. DLMRI and DAGAN failed to preserve detailed structural and textural information, and the reconstructed images contain significant visible artefacts. ADMM-CS and DIMENSION also performed poorly in this regard, although their reconstructions were of higher quality than DLMRI and DAGAN, as evidenced by the error maps depicted in Fig. 7. Images reconstructed by ADMM-CS and DIMENSION appear to be affected by streak-artefacts, as illustrated in the enlarged regions in Fig. 7 (refer to red boxes in Fig. 7). Visually, images reconstructed by MD-Recon Net appeared similar to those reconstructed using our approach. The corresponding error maps, however, suggest that our approach, DAPGAN, achieved lower magnitudes of reconstruction errors throughout the brain than MD-Recon Net, as illustrated in Fig. 7.

5.1.3. *Clinical brain quantification*

To demonstrate the clinical feasibility of the proposed image reconstruction approach, we also conducted a volumetric analysis in terms of the selected key brain indices based on segmentations derived from images reconstructed using undersampled data. Volumetric measurements of key anatomical structures are often used as biomarkers for

Table 3
Quantification of brain voxel volume (mm^3) of multiple volumetric indices extracted from segmented 3D volumes with 2D reconstructed results derived from Ground Truth, our method and competing methods. (LV/RV: left/right ventricle, LWM/RWM: left/right white matter, LT/RT: left/right thalamus, LP/RP: left/right putamen, BS: brain stem, CCM: corpus callosum of middle-anterior).

Parameters	GT	**DAPGAN**		MD-Recon		DIMENSION		ADMM		DAGAN	
	Mean ± Std	**Mean ± Std**	**p-value**	Mean ± Std	p-value	Mean ± Std	p-value	Mean ± Std	p-value	Mean ± Std	p-value
LV	3302.67 ± 188.23	**3519 ± 224.55**	**0.0002**	4798.17 ± 749.63	<0.0002	4712.17 ± 1210.88	<0.0002	5034.83 ± 671.01	<0.0002	5716.33 ± 837.57	<0.0002
LWM	8722.33 ± 478.60	**8406.17 ± 509.41**	**0.0179**	7176.33 ± 1106.57	<0.0002	5600.67 ± 579.03	<0.0002	6597.33 ± 689.54	<0.0002	4792.17 ± 967.35	<0.0002
LT	5873.5 ± 487.71	**5619 ± 455.72**	**0.0446**	5218.83 ± 1433.46	0.0234	5539.33 ± 1379.54	0.0237	4840.67 ± 1375.83	0.0003	4527.24 ± 1484.65	0.0256
LP	1669.67 ± 67.89	**1735.83 ± 67.91**	**0.0005**	2860.83 ± 1012.60	<0.0002	4433.83 ± 645.89	<0.0002	4193.67 ± 826.52	<0.0002	4819.43 ± 891.38	<0.0002
BS	14 415.83 ± 1232.73	**13 109.5 ± 1469.73**	**0.0005**	7909.33 ± 948.93	<0.0002	8441.67 ± 364.38	<0.0002	8447.67 ± 1059.51	<0.0002	7233.17 ± 918.24	<0.0002
RP	2747.83 ± 117.60	**2647.5 ± 148.03**	**0.0059**	2130.33 ± 160.67	<0.0002	1834.83 ± 226.38	<0.0002	1998.17 ± 89.010	<0.0002	1761.63 ± 288.29	<0.0002
RT	5050.83 ± 459.48	**4700 ± 461.32**	**0.0052**	4194 ± 750.15	<0.0002	3913.17 ± 641.05	<0.0002	4210.33 ± 375.90	<0.0002	3792.19 ± 259.26	<0.0002
RWM	10 335.17 ± 817.08	**9856 ± 802.90**	**0.0281**	7349.83 ± 704.03	<0.0002	7941.17 ± 1049.23	<0.0002	6425 ± 759.40	<0.0002	6193.21 ± 785.73	<0.0002
RV	2731.33 ± 422.33	**2869.5 ± 397.27**	**0.1523**	3005.67 ± 402.98	0.0091	1751.83 ± 351.69	<0.0002	1830.33 ± 490.11	<0.0002	1585.45 ± 493.17	<0.0002
CCM	1562.67 ± 127.64	**1413.17 ± 218.09**	**0.0023**	934 ± 53.53	<0.0002	785.17 ± 120.88	<0.0002	959.67 ± 279.73	<0.0002	717.06 ± 138.83	<0.0002

Table 4
Quantitative comparison of cardiac image reconstruction performance, using Cartesian and random undersampling patterns at different rates in terms of SSIM and PSNR (dB). To further verify the robustness of investigated methods, Type II of Cartesian sampling is conducted as a challenging mask, with irregular phase-encoding gradient.

Methods	Cartesian (Mean ± Std.)				Random (Mean ± Std.)			
	Type II of irregular 10%		20%		10%		20%	
	SSIM	PSNR (dB)	SSIM	PSNR (dB)	SSIM	PSNR (dB)	SSIM	PSNR (dB)
DLMRI	0.710 ± 0.040	26.400 ± 1.369	0.803 ± 0.036	25.758 ± 2.338	0.689 ± 0.031	29.146 ± 1.604	0.767 ± 0.030	32.843 ± 1.437
DAGAN	0.801 ± 0.027	29.375 ± 1.332	0.849 ± 0.024	30.152 ± 1.359	0.743 ± 0.033	29.892 ± 1.604	0.815 ± 0.030	33.762 ± 1.502
ADMM-CS Net	0.826 ± 0.024	29.913 ± 1.343	0.894 ± 0.018	31.859 ± 1.481	0.832 ± 0.024	31.823 ± 1.687	0.873 ± 0.019	34.960 ± 1.558
DIMENSION	0.855 ± 0.021	31.376 ± 1.412	0.907 ± 0.016	32.383 ± 1.545	0.869 ± 0.018	32.941 ± 1.705	0.919 ± 0.016	36.681 ± 1.477
MD-Recon	0.859 ± 0.021	31.193 ± 1.441	0.925 ± 0.013	33.291 ± 1.576	0.887 ± 0.017	33.639 ± 1.669	0.938 ± 0.017	37.347 ± 1.446
DAPGAN	**0.930 ± 0.011**	**34.512 ± 1.472**	**0.959 ± 0.007**	**35.657 ± 1.579**	**0.918 ± 0.018**	**35.081 ± 1.546**	**0.953 ± 0.015**	**37.877 ± 1.413**

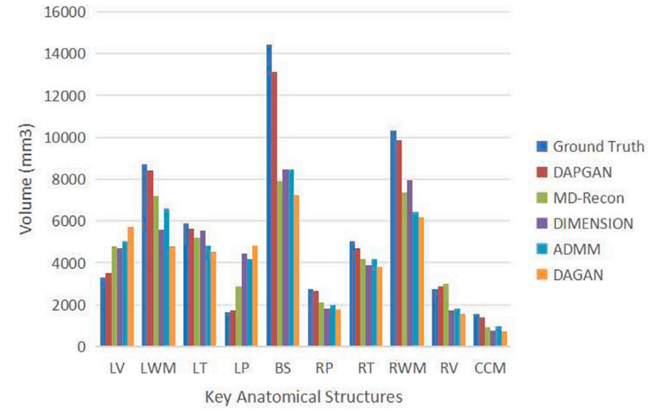


Fig. 8. Quantification of brain volumetric indices based on segmentations derived from the original images (Ground Truth) and from images reconstructed from undersampled data using DAPGAN and each competing method of interest. (LV/RV: left/right ventricle, LWM/RWM: left/right white matter, LT/RT: left/right thalamus, LP/RP: left/right putamen, BS: brain stem, CCM: corpus callosum of middle-anterior).

disease diagnosis and monitoring. The rationale here is that for the proposed approach to be clinically useful, volumetric biomarkers derived from reconstructed images (from undersampled data) should show no/limited statistically significant differences to values derived from the original images. To this end, volumetric analysis was conducted using reconstructions of structural brain MR images, by quantifying volumes for the following structures using FreeSurfer [35] — lateral ventricles, white matter, thalamus, and putamen for both left and right brain hemispheres, and the brain stem and middle-anterior corpus callosum.

Brain volumetric analysis following image reconstruction using DAPGAN and all other state-of-the-art methods of interest in this study is summarised in Fig. 8, alongside the ground truth structural brain volumes (i.e. volumes derived from the original images). For fair comparison, all evaluations were conducted identically, via automatic brain segmentation and volumetric quantification using FreeSurfer. Our method, DAPGAN, consistently outperformed all competing methods and showed no statistically significant differences to the ground truth volumes for most brain structures analysed (see Table 3).

5.2. *Cardiac MRI reconstruction*

For the cardiac dataset, we evaluated the performance of our method on the reconstruction of undersampled data generated using Cartesian and Gaussian random undersampling patterns, using different sampling ratios (regular and irregular masks with 10% and 20%). Furthermore, we conducted a set of ablation experiments to evaluate the key components of the proposed approach.

5.2.1. *Regular Cartesian undersampling*

Table 4 shows results from an experiment with an undersampling rate of 20% in the regular Cartesian pattern. DLMRI and DAGAN achieve SSIM values of $0.80 - 0.85$, with the latter providing marginal improvements over the former. DIMENSION, ADMM-CS Net and MD-Recon Net all outperform DLMRI and DAGAN, achieving SSIM values of 0.894, 0.907 and 0.925, respectively. As with the brain MR data set, our approach significantly outperformed all competing methods, at reconstructing images undersampled using a regular Cartesian pattern at a rate of 20%, achieving an average SSIM value of 0.959. Our method achieved the best performance with an SSIM value of 0.959. Moreover, our method reached the highest PSNR value of 35.657 dB, more than 2 dB higher than that achieved using MD-Recon Net. In the random sampling mask with a sampling rate of 10%, DLMRI yielded an SSIM lower than 0.7, and DAGAN achieved an SSIM value of only 0.743. Both ADMM-CS Net and DIMENSION achieved scores ~0.85. MD-Recon Net achieved an SSIM value of almost 0.900, while our method achieved an SSIM value of >0.900. When the sampling rate was increased to 20%, the SSIM value increased by 4% and 5% in the cases of ADMM-CS Net and DIMENSION, respectively. Our method and MD-Recon Net, DIMENSION, achieved SSIM values of >0.900.

Fig. 9 provides results of visual inspection between the proposed approach and other approaches on the cardiac dataset. When we set the 20% Cartesian strategy, as shown in the enlarged images and error maps, DLMRI and DAGAN show severe blurriness, and the remaining three approaches produced the images with artefacts corrupted near the myocardium (MYO) wall (as indicated more clearly in error maps

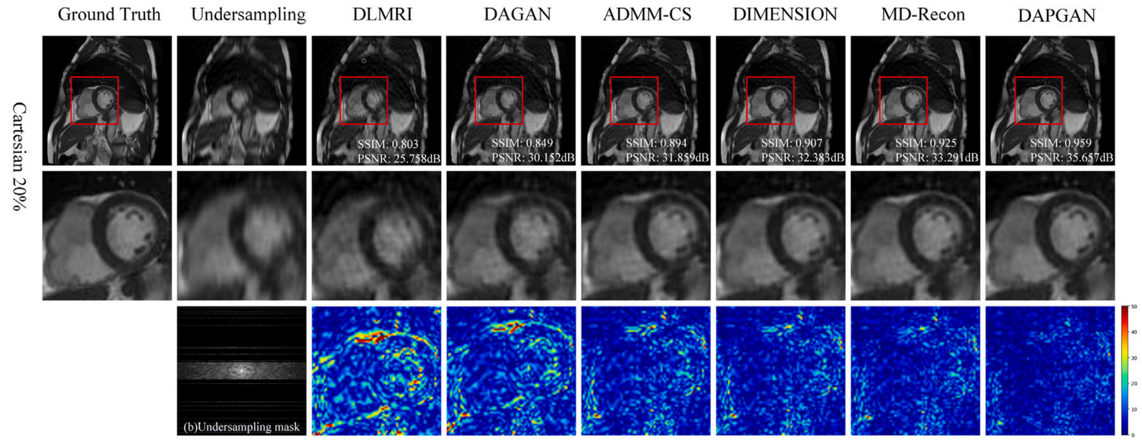


Fig. 9. Visual inspection results of CS-MRI reconstruction on cardiac images using regular Cartesian sampling at rates 20%, here we show the comparison between the ground truth, undersampled results and investigated methods. The third column to the last column are the results of different methods. The second row represents the magnifying views of the red boxes in the original images. The third row shows the undersampling mask and the absolute difference maps (error maps) indicating the regions with multi-degrees of distinctions. The quantitative evaluation metrics (SSIM/PSNR) of the reconstructed images are indicated in the first row.

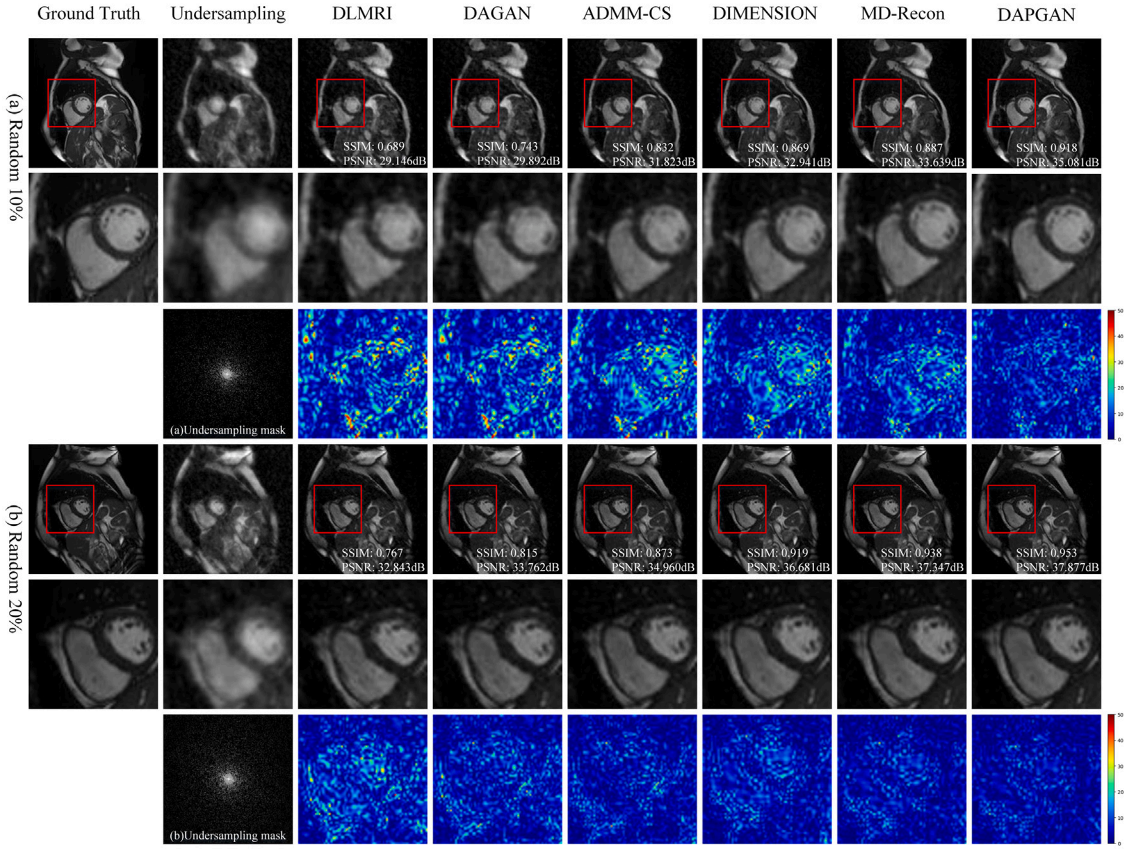


Fig. 10. Visual inspection results of CS-MRI reconstruction on cardiac images using regular random sampling at rates 10% (from first row to third row) and 20% (from fourth row to last row), here we show the comparison between the ground truth, undersampled results and investigated methods. The second column (Undersampling) of the first and fourth rows represent the random sampling pattern in image space. The third column to the last column are the results of different methods. The second and fifth rows represent the magnifying views of the red boxes in the original images. The third row and last row show the undersampling masks and the absolute difference maps (error maps) indicating the regions with multi-degrees of distinctions. The quantitative evaluation metrics (SSIM/PSNR) of the reconstructed images are indicated in the first and fourth rows.

compared with magnifying views). The DAPGAN consistently outperformed the other methods in terms of the removal of the artefacts and recovering high-resolution details.

For Gaussian random results with a sampling rate of 10% in Fig. 10, the five compared methods failed to remove the severe noise signals and did not preserve realistic textures in the ventricle regions. However, the proposed method yielded the most comparable results to the GT with fine details.

5.2.2. *Irregular Cartesian undersampling*

Table 4 evaluates the de-aliasing robustness of the model, regarding Type II of the irregular mask (see Fig. 6) performing qualitative and quantitative experiments under aggressive sampling conditions. We alter the regular mask with a challenging pattern designed based on the Cartesian sampling rule. The SSIM and PSNR values obtained using DLMRI were ~0.7 and 26 dB, respectively. DAGAN achieved an SSIM value of ~0.8. The performance of ADMM-CS Net was most affected

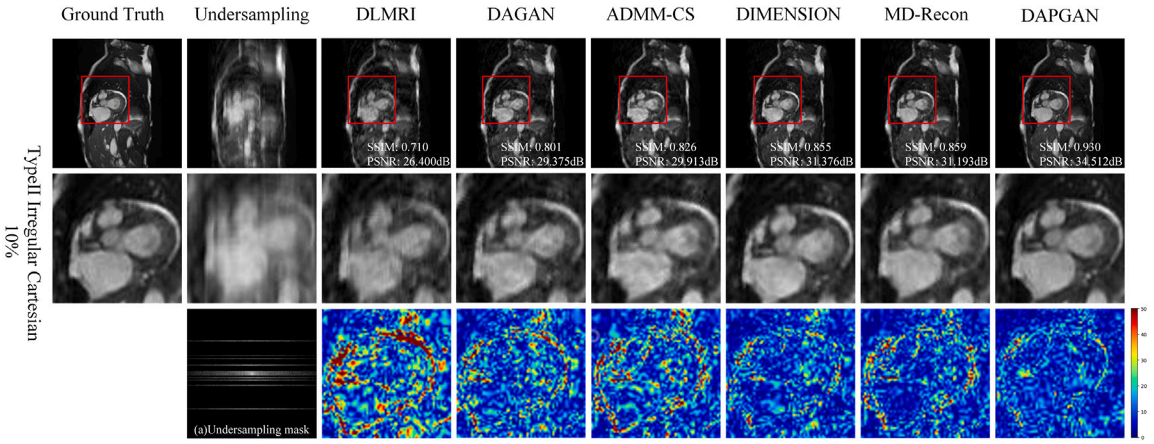


Fig. 11. Visual inspection results of CS-MRI reconstruction on cardiac images using irregular Cartesian sampling at rates 10%, here we show the comparison between the ground truth, undersampled results and investigated methods. The third column to the last column are the results of different methods. The second row represents the magnifying views of the red boxes in the original images. The third row shows the irregular undersampling mask and the absolute difference maps (error maps) indicating the regions with multi-degrees of distinctions. To further verify the robustness of investigated methods, Type II of Cartesian sampling is conducted as a challenging mask, with irregular phase-encoding gradient. The space between lines in the raster of Type II is provided with random values and the lines are randomly located. The quantitative evaluation metrics (SSIM/PSNR) of the reconstructed images are indicated in the first row.

Table 5
3D cardiac segmentation indices with 2D reconstruction results.

Parameters	GT	**DAPGAN**		MD-Recon Net		DIMENSION		ADMM-CS Net		DAGAN	
	Mean ± Std.	**Mean ± Std.**	**p-value**	Mean ± Std.	p-value	Mean ± Std.	p-value	Mean ± Std.	p-value	Mean ± Std.	p-value
LVEDV (mL)	154.31 ± 32.17	**153.11 ± 31.87**	**0.676**	149.79 ± 32.79	0.121	152.18 ± 31.87	0.461	150.57 ± 31.58	0.192	142.17 ± 31.37	0.116
LVESV (mL)	69.93 ± 22.18	**67.38 ± 22.28**	**0.202**	62.75 ± 22.81	<0.005	65.66 ± 22.18	0.034	62.64 ± 22.44	<0.005	56.52 ± 22.57	<0.005
LVSV (mL)	84.38 ± 16.74	**85.74 ± 16.01**	**0.369**	87.05 ± 17.88	0.090	86.53 ± 16.24	0.161	87.92 ± 16.71	0.020	89.81 ± 16.93	0.036
LVEF (%)	55.09 ± 6.46	**56.51 ± 6.34**	**0.015**	58.64 ± 7.43	<0.005	57.39 ± 6.41	<0.005	58.97 ± 6.84	<0.005	60.89 ± 7.83	<0.005
LVM (g)	87.97 ± 16.93	**84.94 ± 16.79**	**0.049**	75.10 ± 16.53	<0.005	80.92 ± 16.38	<0.005	78.25 ± 15.97	<0.005	67.36 ± 16.19	<0.005
RVEDV (mL)	165.48 ± 31.11	**164.78 ± 31.22**	**0.813**	160.50 ± 28.99	0.067	162.02 ± 30.41	0.221	159.62 ± 29.82	0.032	156.24 ± 30.65	0.022
RVESV (mL)	76.82 ± 19.63	**76.34 ± 19.15**	**0.756**	73.28 ± 19.88	0.044	74.98 ± 18.79	0.286	73.47 ± 18.95	0.053	70.37 ± 18.68	0.043
RVSV (mL)	88.65 ± 15.91	**88.44 ± 15.93**	**0.878**	87.22 ± 14.34	0.265	87.04 ± 15.46	0.251	86.14 ± 15.50	0.076	84.58 ± 15.69	0.032
RVEF (%)	53.92 ± 5.32	**53.99 ± 5.15**	**0.799**	54.84 ± 6.26	0.067	54.06 ± 5.19	0.610	54.33 ± 5.78	0.335	55.62 ± 5.76	0.117

under the aforementioned sampling conditions. Compared with the 20% sampling rate from a longitudinal perspective, the SSIM value of ADMM-CS Net decreased by 7% to ~0.83 under this pattern, while the SSIM values obtained using DIMENSION and MD-Recon Net decreased by 5% (0.855) and ~7% (0.859), respectively. Our method maintained an SSIM value of >0.900, and the decrease was only <3%.

Fig. 11 shows that DAPGAN yields superior and consistent performance in preserving accurate perceptual details with better contrast and higher resolution, as depicted by the magnified cardiac regions, which reflected more underlying anatomy. We faithfully reconstructed the images with the smallest difference illustrated by the error maps, especially the improved performances in the region of the MYO wall. In contrast, the comparison methods failed to recover the high-frequency details, and produced relatively large differences, as shown in the error maps and enlarged images.

5.2.3. *Clinical cardiac quantification*

In this section, we demonstrate the impact of our method on various cardiac functional indexes based on the two most representative time points: ED and ES. Table 5 and Fig. 12 demonstrate the performance of the competing methods in terms of several key cardiac clinical analysis metrics, including the LV ED volume (LVEDV), ES volume (LVESV), LV stroke volume (LVSV), LV ejection fraction (LVEF), LV myocardial mass (LVM), RV ED volume (RVEDV), ES volume (RVESV), RV stroke volume (RVSV), and RV ejection fraction (RVEF). Our proposed DAPGAN outperformed all the competing methods in all clinical parameters and yielded the most comparable results to the reference ranges derived from the fully-sampled data.

5.2.4. *Ablation study*

To fully evaluate the benefit of our algorithm, we perform an ablation experiment. The results were analysed under the challenging 10% Cartesian undersampling conditions.

Table 6
Quantitative comparison of CS-MRI reconstruction on cardiac data using irregular Cartesian sampling at rates 10% for ablation study in terms of SSIM and PSNR (dB). IS: image-space Net only, KS + IS: dual-domain learning with k-space, KS + IS + PL: adding perceptual loss into the loss functions, and KS + IS + PFG: a complete structure for DAPGAN with PFG.

Methods	Type II of irregular 10% Cartesian (Mean ± Std.)	
	SSIM	PSNR
IS	0.816 ± 0.030	29.780 ± 1.469
KS + IS	0.846 ± 0.027	31.644 ± 1.457
KS + IS + PFL	0.875 ± 0.023	32.427 ± 1.491
KS + IS + PFG (DAPGAN)	**0.930** ± **0.011**	**34.512** ± **1.472**

We use three structural variants of our approach: image-space only (IS), dual-domain learning, which is image-space joint with k-space (KS + IS), adding perceptual feature loss into the loss functions (KS + IS + PFL) and a complete structure for DAPGAN with PFG (KS + IS + PFG). We first evaluate the importance of incorporating the k-space domain into the reconstruction process. In Table 6, compared with single-domain learning, better performance can be achieved using dual-domain learning, with SSIM and PSNR value improvements of 3% and 2 dB, respectively. In Fig. 13, when we performed the dual-domain network, an improved performance with smaller errors in the region of MYO wall could be observed. Furthermore, we evaluated the variant with perceptual loss added to the linear loss function. The model achieved an increased SSIM value of ~3% compared with “KS+IS”. However, from the view of error maps, ‘KS + IS + PFL’ appeared smaller differences slightly than the former two variants, which indicated the limitation of utilising the linear combination of multi-loss functions as the only supervision. DAPGAN considerably outperformed the other variants in terms of all metrics and visual inspection based on the

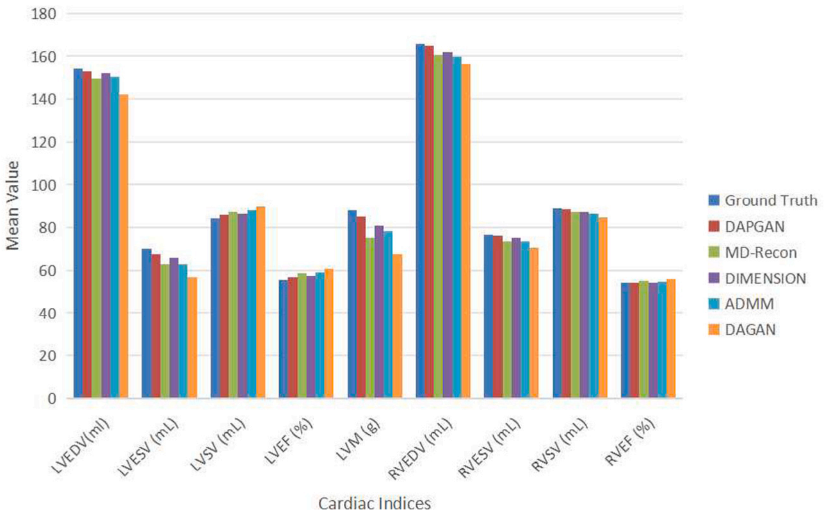


Fig. 12. Mean value of 3D cardiac functional indices for the ground truth and investigated methods.

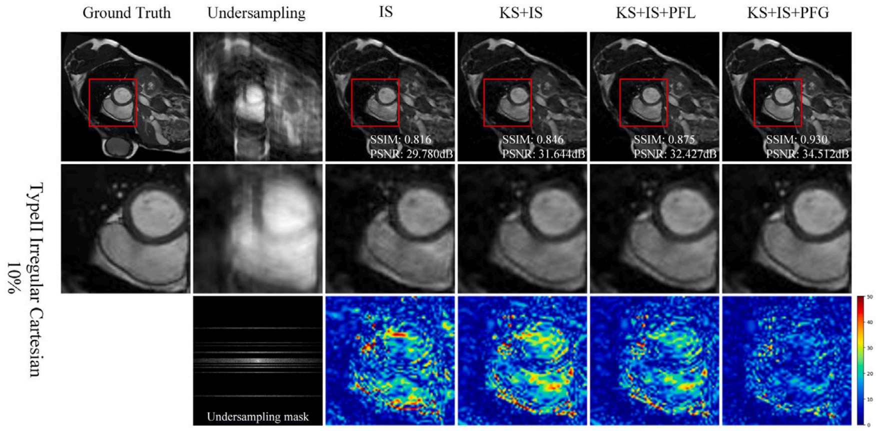


Fig. 13. Visual inspection results of CS-MRI reconstruction on cardiac images using irregular Cartesian sampling at rates 10% for ablation study, here we show the comparison between the ground truth, undersampled results and investigated variants. IS: image-space Net only, KS + IS: dual-domain learning with k-space, KS + IS + PL: adding perceptual loss into the loss functions, KS + IS + PFG: a complete structure for DAPGAN with PFG. The third column to the last column are the results of different variants of our method. The second represents the magnifying views of the red boxes in the original images. The third row shows the undersampling mask and the absolute difference maps (error maps) indicating the regions with multi-degrees of distinctions. The quantitative evaluation metrics (SSIM/PSNR) of the reconstructed images are indicated in the first row.

challenge mask, significantly decreasing differences in the regions of MYO wall and ventricle. This consistently confirmed the merits of exploiting multi-level perceptual feature guidance (PFG) as an effective supervision mechanism, instead of optimising or simply combining loss functions linearly in a conventional way.

6. Discussion

Accurate assessment and high acceleration of the MR image reconstruction from undersampled k-space data is essential. The relationship between convolutional networks and conventional methods has already been justified by the theoretical analysis, which solidly supports the development of the CNN-based method. Herein, we focus on exploring an innovative approach for MRI reconstruction suitable for practical undersampling conditions, saving the physical sampling time, and enabling the reconstruction of undersampled data with high quality and comparable to reconstruction results of fully-sampled data.

Exploring the information in image-space and k-space in parallel improved the average performance relatively compared with single domain-based approaches. Nevertheless, we should pay more attention to the control of the optimisation solution of a CNN-based method during training, in other words, the “black box”, if we desire an effective mechanism for avoiding overfitting and achieving superior performance. In summary, recent methods are inadequate for two main terms: clinical practice and improvements for regularisation constraints. In addition, an efficient architecture enables the network to have a good potential for application extension.

Firstly, when the amount of training data based on fullysampling as references is not sufficient for the task, the supervised learning paradigms may not be able to train a sufficiently robust model. Moreover, for accelerated MR image acquisition, higher undersamping factors add severe aliasing to the initial zero-filled inputs, making the reconstruction more challenging. In addition, the generalisation ability across multiple conditions is significantly reduced when training the model with certain datasets. For instance, the reconstruction tasks for medical imaging with multiple organs. Compensating the improvement of network generalisation capability with existing transfer learning offers a better solution.

For another aspect, it is defined as the inverse problem which learns the end-to-end mapping from aliased undersampled data to reconstructed high-quality prediction. Nevertheless, as the regularisation term utilised in the optimisation problem, recent deep learning-based methods have adopted optimisation strategies based on the fusion of multiple typical loss functions. Therefore, it will make them still suffer from blurry results, because the output is essentially an averaged optimal solution to satisfy the multiple criteria based on corresponding weights. Meanwhile, the intermediate information of multi-level features which perceive more detailed structural information is less considered for optimising specially. The effectiveness is also confirmed by our simulation experiments, especially under aggressive sampling strategies. Inspired by the VGG network for deep feature extraction, we proposed multi-level feature extraction and difference correlation layer as an iterate-specific error-correction unit, the outputs are inserted into embedding and reconstruction nets. By retaining the useful information, this unit can adaptively correct the intermediate prediction of each iteration. Overall, for reflecting more accurate tissue structures, the feature guidance adaptively constrains the feature learning of multi-level in the solution space. For relative tasks like segmentation or synthesis, it can be a flexible mechanism that enables the network to have a good potential for application extension. Furthermore, our work proved another option for controlling the “black box” of deep learning-based optimisation, which is the effective feature guidance that enables avoiding gradient vanishing or overfitting.

Considering that the performances of most deep learning-based CS approaches are limited by higher undersampling factors, in addition, it would also be influenced when exploring the application on multiple organs with different structural complexity. We evaluate our method on brain data and cardiac data. For imaging acceleration simulation, we considered CS-MRI reconstruction at sample rates lower than 20% which are relatively hard to deal with utilising optimisation-based approaches. In view of simulating more undersampling patterns for improving the robustness of our model, we conducted irregular masks at a 10% rate for simulating aggressive CS-MRI situations with significantly increased aliasing. Our experimental results show that the proposed DAPGAN can achieve better promising results both in terms of visual images and quantitative indicators compared with conventional and competing approaches.

Currently, we only focus on the reconstruction in static 2D cardiac cine MR imaging, while clinical indices such as myocardium strain are in demand to be computed to assess the dynamic deformation functionality of the heart, which indicates the potential correlation and redundancy exist along the spatio-temporal dimension. Additionally, as estimating the deformation field of multiple cardiac frames in a complete cardiac cycle is a crucial step for dynamic reconstruction with cine imaging, it is still a challenge to estimate accurate respiratory motion fields directly from undersampled images. It would be interesting to investigate how motion compensation improves dynamic CMR reconstruction and research its generalisability for different downstream tasks.

7. Conclusion

We proposed a novel deep learning-based approach, namely DAPGAN, for the reconstruction of undersampled k-space data with high quality. In particular, an active adaptation-based guidance mechanism for multi-level feature reconstruction is proposed. Also, the dual-domain learning with k-space further improved the generalisation ability. DAPGAN is robust to de-aliasing and other artefacts removal in accelerated k-space datasets. It produces perceptual features-focused results, which is confirmed by superior performance on evaluated metrics.

CRediT authorship contribution statement

Kun Wu: Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Yan Xia:** Writing – review & editing, Visualization, Validation, Formal analysis. **Nishant Ravikumar:** Writing – review & editing, Supervision. **Alejandro F. Frangi:** Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This research was conducted using the UK Biobank Resource under Application 11350. The authors are grateful to all UK Biobank participants and staff. AFF acknowledges support from the Royal Academy of Engineering Chair in Emerging Technologies Scheme (CiET1819/19), EPSRC-funded Grow MedTech CardioX (POC041), and the MedIAN Network (EP/N026993/1) funded by the Engineering and Physical Sciences Research Council (EPSRC), United Kingdom.

References

[1] Daniel K. Sodickson, Warren J. Manning, Simultaneous acquisition of spatial harmonics (SMASH): fast imaging with radiofrequency coil arrays, Magn. Reson. Med. 38 (4) (1997) 591–603.

[2] Emmanuel J. Candès, Justin Romberg, Terence Tao, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information, IEEE Trans. Inform. Theory 52 (2) (2006) 489–509.

[3] David L. Donoho, Compressed sensing, IEEE Trans. Inform. Theory 52 (4) (2006) 1289–1306.

[4] Kai Tobias Block, Martin Uecker, Jens Frahm, Undersampled radial MRI with multiple coils. Iterative image reconstruction using a total variation constraint, Magn. Reson. Med.: Off. J. Int. Soc. Magn. Reson. Med. 57 (6) (2007) 1086–1098.

[5] Xiaoyu Fan, Qiusheng Lian, Baoshun Shi, Compressed sensing MRI with phase noise disturbance based on adaptive tight frame and total variation, IEEE Access 5 (2017) 19311–19321.

[6] Feng Shi, Jian Cheng, Li Wang, Pew-Thian Yap, Dinggang Shen, LRTV: MR image super-resolution with low-rank and total variation regularizations, IEEE Trans. Med. Imaging 34 (12) (2015) 2459–2466.

[7] Benjamin Trémoulhéac, Nikolaos Dikaios, David Atkinson, Simon R Arridge, Dynamic MR image reconstruction-separation from undersampled (k,t)-space via low-rank plus sparse prior, IEEE Trans. Med. Imaging 33 (8) (2014) 1689–1701.

[8] Maosong Ran, Wenjun Xia, Yongqiang Huang, Zexin Lu, Peng Bao, Yan Liu, Huaiqiang Sun, Jiliu Zhou, Yi Zhang, Md-recon-net: A parallel dual-domain convolutional neural network for compressed sensing MRI, IEEE Trans. Radiat. Plasma Med. Sci. 5 (1) (2020) 120–135.

[9] Shanshan Wang, Ziwen Ke, Huitao Cheng, Sen Jia, Leslie Ying, Hairong Zheng, Dong Liang, DIMENSION: dynamic MR imaging with both k-space and spatial prior knowledge obtained via multi-supervised network training, NMR Biomed. 35 (4) (2022) e4131.

[10] Jian Sun, Huibin Li, Zongben Xu, et al., Deep ADMM-net for compressive sensing MRI, Adv. Neural Inf. Process. Syst. 29 (2016).

[11] Guang Yang, Simiao Yu, Hao Dong, Greg Slabaugh, Pier Luigi Dragotti, Xujiong Ye, Fangde Liu, Simon Arridge, Jennifer Keegan, Yike Guo, et al., DAGAN: deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction, IEEE Trans. Med. Imaging 37 (6) (2017) 1310–1321.

[12] Saiprasad Ravishankar, Yoram Bresler, MR image reconstruction from highly undersampled k-space data by dictionary learning, IEEE Trans. Med. Imaging 30 (5) (2010) 1028–1041.

[13] Xinwen Liu, Jing Wang, Jin Jin, Mingyan Li, Fangfang Tang, Stuart Crozier, Feng Liu, Deep unregistered multi-contrast MRI reconstruction, Magn. Reson. Imaging 81 (2021) 33–41.

[14] Salman Ul Hassan Dar, Muzaffer Özbey, Ahmet Burak Çatlı, Tolga Çukur, A transfer-learning approach for accelerated MRI using deep neural networks, Magn. Reson. Med. 84 (2) (2020) 663–685.

[15] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio, Generative adversarial nets, Adv. Neural Inf. Process. Syst. 27 (2014) 2672–2680.

[16] Tran Minh Quan, Thanh Nguyen-Duc, Won-Ki Jeong, Compressed sensing MRI reconstruction using a generative adversarial network with a cyclic loss, IEEE Trans. Med. Imaging 37 (6) (2018) 1488–1497.

[17] Wenzhong Zhou, Huiqian Du, Wenbo Mei, Liping Fang, Efficient structurally-strengthened generative adversarial network for MRI reconstruction, Neurocomputing 422 (2021) 51–61.

[18] Xianzhe Liu, Hongwei Du, Jinzhang Xu, Bensheng Qiu, DBGAN: A dual-branch generative adversarial network for undersampled MRI reconstruction, Magn. Reson. Imaging 89 (2022) 77–91.

[19] Bhavya Vasudeva, Puneesh Deora, Saumik Bhattacharya, Pyari Mohan Pradhan, Compressed sensing mri reconstruction with co-vegan: Complex-valued generative adversarial network, in: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2022, pp. 672–681.

[20] Jianan Xu, Wanqing Bi, Lier Yan, Hongwei Du, Bensheng Qiu, An efficient lightweight generative adversarial network for compressed sensing magnetic resonance imaging reconstruction, IEEE Access 11 (2023) 24604–24614.

[21] Xia Li, Hui Zhang, Hao Yang, Tie-Qiang Li, CS-MRI reconstruction using an improved GAN with dilated residual networks and channel attention mechanism, Sensors 23 (18) (2023) 7685.

[22] Jo Schlemper, Jose Caballero, Joseph V Hajnal, Anthony N Price, Daniel Rueckert, A deep cascade of convolutional neural networks for dynamic MR image reconstruction, IEEE Trans. Med. Imaging 37 (2) (2017) 491–503.

[23] Taejoon Eo, Yohan Jun, Taeseong Kim, Jinseong Jang, Ho-Joon Lee, Dosik Hwang, KIKI-net: cross-domain convolutional neural networks for reconstructing undersampled magnetic resonance images, Magn. Reson. Med. 80 (5) (2018) 2188–2201.

[24] Liyan Sun, Zhiwen Fan, Yue Huang, Xinghao Ding, John Paisley, Compressed sensing MRI using a recursive dilated network, in: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 32, 2018.

[25] Zhiwen Fan, Liyan Sun, Xinghao Ding, Yue Huang, Congbo Cai, John Paisley, A segmentation-aware deep fusion network for compressed sensing mri, in: Proceedings of the European Conference on Computer Vision, ECCV, 2018, pp. 55–70.

[26] Qiaoying Huang, Dong Yang, Pengxiang Wu, Hui Qu, Jingru Yi, Dimitris Metaxas, MRI reconstruction via cascaded channel-wise attention network, in: 2019 IEEE 16th International Symposium on Biomedical Imaging, ISBI 2019, IEEE, 2019, pp. 1622–1626.

[27] Wenyuan Qiu, Dongxiao Li, Xinyu Jin, Fan Liu, Bin Sun, Deep neural network inspired by iterative shrinkage-thresholding algorithm with data consistency (NISTAD) for fast undersampled MRI reconstruction, Magn. Reson. Imaging 70 (2020) 134–144.

[28] Roberto Souza, Mariana Bento, Nikita Nogovitsyn, Kevin J Chung, Wallace Loos, R Marc Lebel, Richard Frayne, Dual-domain cascade of U-nets for multi-channel magnetic resonance image reconstruction, Magn. Reson. Imaging 71 (2020) 140–153.

[29] Bo Zhou, S. Kevin Zhou, DuDoRNet: learning a dual-domain recurrent network for fast MRI reconstruction with deep T1 prior, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 4273–4282.

[30] Bo Zhou, Jo Schlemper, Neel Dey, Seyed Sadegh Mohseni Salehi, Kevin Sheth, Chi Liu, James S Duncan, Michal Sofka, Dual-domain self-supervised learning for accelerated non-cartesian mri reconstruction, Med. Image Anal. 81 (2022) 102538.

[31] Johnson Justin, Alahi Alexandre, Fei-Fei Li, Perceptual losses for real-time style transfer and super-resolution, in: European Conference on Computer Vision, Springer, 2016, pp. 694–711.

[32] Karen Simonyan, Andrew Zisserman, Very deep convolutional networks for large-scale image recognition, 2014, arXiv preprint arXiv:1409.1556.

[33] Ying Tai, Jian Yang, Xiaoming Liu, Image super-resolution via deep recursive residual network, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 3147–3155.

[34] Steffen E Petersen, Paul M Matthews, Jane M Francis, Matthew D Robson, Filip Zemrak, Redha Boubertakh, Alistair A Young, Sarah Hudson, Peter Weale, Steve Garratt, et al., UK biobanks cardiovascular magnetic resonance protocol, J. Cardiovasc. Magn. Reson. 18 (1) (2015) 8.

[35] Bruce Fischl, Freesurfer, Neuroimage 62 (2) (2012) 774–781.