




# Edge-enhanced dual discriminator generative adversarial network for fast MRI with parallel imaging using multi-view information

Jiahao Huang<sup>1,2</sup> · Weiping Ding<sup>3</sup> · Jun Lv<sup>4</sup> · Jingwen Yang<sup>5</sup> · Hao Dong<sup>6</sup> · Javier Del Ser<sup>7,8</sup> · Jun Xia<sup>9</sup> · Tiaojuan Ren<sup>1</sup> · Stephen T. Wong<sup>10</sup> · Guang Yang<sup>2,11</sup> 

Accepted: 9 December 2021 / Published online: 28 January 2022  
© The Author(s) 2021

## Abstract

In clinical medicine, magnetic resonance imaging (MRI) is one of the most important tools for diagnosis, triage, prognosis, and treatment planning. However, MRI suffers from an inherent slow data acquisition process because data is collected sequentially in  $k$ -space. In recent years, most MRI reconstruction methods proposed in the literature focus on holistic image reconstruction rather than enhancing the edge information. This work steps aside this general trend by elaborating on the enhancement of edge information. Specifically, we introduce a novel parallel imaging coupled dual discriminator generative adversarial network (PIDD-GAN) for fast multi-channel MRI reconstruction by incorporating multi-view information. The dual discriminator design aims to improve the edge information in MRI reconstruction. One discriminator is used for holistic image reconstruction, whereas the other one is responsible for enhancing edge information. An improved U-Net with local and global residual learning is proposed for the generator. Frequency channel attention blocks (FCA Blocks) are embedded in the generator for incorporating attention mechanisms. Content loss is introduced to train the generator for better reconstruction quality. We performed comprehensive experiments on Calgary-Campinas public brain MR dataset and compared our method with state-of-the-art MRI reconstruction methods. Ablation studies of residual learning were conducted on the MICCAI13 dataset to validate the proposed modules. Results show that our PIDD-GAN provides high-quality reconstructed MR images, with well-preserved edge information. The time of single-image reconstruction is below 5ms, which meets the demand of faster processing.

**Keywords** Fast MRI · Parallel imaging · Multi-view learning · Generative adversarial networks · Edge enhancement

✉ Tiaojuan Ren  
rentj@zjsru.edu.cn

✉ Guang Yang  
g.yang@imperial.ac.uk

<sup>1</sup> College of Information Science and Technology, Zhejiang Shuren University, 310015 Hangzhou, China

<sup>2</sup> National Heart and Lung Institute, Imperial College London, London, UK

<sup>3</sup> School of Information Science and Technology, Nantong University, 226019 Nantong, China

<sup>4</sup> School of Computer and Control Engineering, Yantai University, 264005 Yantai, China

<sup>5</sup> Department of Prosthodontics, Peking University School and Hospital of Stomatology, Beijing, China

<sup>6</sup> Center on Frontiers of Computing Studies, Peking University, Beijing, China

<sup>7</sup> TECNALIA, Basque Research and Technology Alliance (BRTA), 48160 Derio, Spain

<sup>8</sup> University of the Basque Country (UPV/EHU), 48013 Bilbao, Spain

<sup>9</sup> Department of Radiology, Shenzhen Second People's Hospital, The First Affiliated Hospital of Shenzhen University Health Science Center, Shenzhen, China

<sup>10</sup> Systems Medicine and Bioengineering Department, Departments of Radiology and Pathology, Houston Methodist Cancer Center, Houston Methodist Hospital, Weill Cornell Medicine, 77030 Houston, TX, USA

<sup>11</sup> Cardiovascular Research Centre, Royal Brompton Hospital, London, UK

## 1 Introduction

Magnetic resonance imaging (MRI) is one of the most important clinical tools for diagnosis, triage, prognosis, and treatment planning. MRI produces accurate, potentially high-resolution, and reproducible images with various contrast and functional information. Furthermore, it is non-invasive and harmless to the human body. However, MRI has an inherently slow data acquisition process since data is collected sequentially in  $k$ -space, where speed is limited by physiological and hardware constraints, rather than directly in image space [1]. Prolonged acquisition time can lead to severe motion artefacts due to patient movement and physiological motion. Early approaches to accelerate acquisition followed Nyquist-Shannon sampling criteria, e.g. implementing multiple radio frequency [2] or gradient refocusing [3], with limited speed improvement.

Undersampling in  $k$ -space can improve acquisition speed, but causes aliasing artefacts and blur. Several studies have attempted to reduce aliasing artefacts, such as parallel imaging (PI) and compressed sensing (CS). Parallel imaging was first introduced in 1997 as the simultaneous acquisition of spatial harmonic (SMASH) [4] to reduce scan time using multi-channel  $k$ -space data and includes sensitivity encoding (SENSE) [5] and generalized auto-calibrating partially parallel acquisition (GRAPPA) [6]. The acceleration factor and geometry factor influence signal noise ratio for the reconstructed image, where geometry factor depends on the receiver coil distribution. Parallel imaging requires phased array coils where each coil receives data at the same time (i.e., parallel). If the local sensitivity for each coil is already known, then the field of view can be made arbitrarily small in the phase-encoding direction, enabling aliasing to be unwrapped using this information.

Compressed sensing [7] reconstructs signals from significantly fewer measurements with higher sampling efficiency than traditional Nyquist-Shannon sampling. There are three requirements for successful CS [8], which MRI naturally meets except for the third one:

- (1) *Transform sparsity*: MRI is compressible by transform coding and has a sparse representation in an appropriate transform domain.
- (2) *Incoherence*: It can be achieved by random undersampling MRI data in  $k$ -space.
- (3) *Nonlinear reconstruction*: MRI should be reconstructed using a nonlinear method. The main task of CS is to find an appropriate nonlinear method to reconstruct MRI.

Several methods applying fixed sparsifying transforms for reconstruction have been proposed, e.g., total variation (TV) [9], curvelets [10], and double-density complex

wavelet [11], while adaptive sparse model represented by dictionary learning [12] was also developed as an extension. In addition, CS-MRI worked well with parallel imaging data jointly. Aelterman et al. [13] proposed a joint algorithm, i.e., COMPASS, of parallel imaging and compressed sensing for MRI reconstruction. Trzasko et al. [14] designed an offline, sparsity-driven reconstruction framework for Cartesian sampling time-series acquisitions. However, the iterative computation significantly limits the speed of the reconstruction, leading to the fact that the traditional CS-MRI is hard to be applied in the clinical environment.

Deep learning has been developed enormously recently. Convolutional neural networks (CNNs) have made a significant impact on many computer vision learning tasks, such as classification [15], segmentation [16], object detection [17], and image reconstruction [18]. CNNs offer better feature extraction performance compared with traditional machine learning algorithms due to their deeper structure of hierarchically stacked neural layers. Consequently, several CNNs have been proposed for medical imaging to solve traditional limitations. CNN was first introduced for MRI reconstruction in [19], where it was used to identify mapping relationships between MR images obtained from zero-filled and fully-sampled  $k$ -space data. Yang et al. [20] proposed a deep architecture, i.e., ADMM-Net, inspired by the alternating direction method of multipliers (ADMM) algorithm to optimizing CS based MRI models. Schlemper et al. [21] developed a deep cascade CNN to reconstruct dynamic sequences for 2D cardiac MR images from undersampled data. Zhu et al. [22] proposed a novel framework for MRI reconstruction, i.e., automated transform by manifold approximation (AUTOMAP), allowing a mapping between the sensor and the image domain to emerge from an appropriate corpus of training data. Transfer learning has been also proposed to solve the data scarcity problem when training deep networks for accelerated MRI [23].

In 2014, Goodfellow et al. [24] designed generative adversarial networks (GANs). Then, various improved GAN models were presented. Wasserstein GAN (WGAN) [25] was proposed for improving the training stability of GAN and optimise the learning curves. Radford et al. [26] introduced DCGAN, first applied CNNs to GANs, bridging the gap between supervised learning and unsupervised learning. GAN-based models has been widely used in image-to-image translation [27–30], super-resolution [31–33], as well as MRI reconstruction [34–36].

Several studies have reported that the deep learning based GAN models perform well in MRI reconstruction. DAGAN [34, 37] applied the modified U-Net [38] architecture as the generator and incorporated additional perceptual loss by pre-trained VGG [39] networks. Shaul et al. [35, 37]

introduced a dual-generator GAN, i.e., KIGAN, for both estimating the missing  $k$ -space samples and refinement of the image space data. Quan et al. [36, 37] presented RefineGAN, which was a variant of a fully-residual convolutional autoencoder and GAN with cyclic data consistency loss. Lv et al. [40], proposed a deep GAN model, i.e., PICGAN, with parallel imaging for accelerated multi-channel MRI reconstruction, where data fidelity and regularisation terms were integrated into the generator. Nader et al. [41] combined the traditional MR reconstruction algorithm GRAPPA with a conditional GAN to build the GRAPPA-GAN, which was developed and tested on the multi-coil brain data. Guo et al. [42] proposed DAWAGAN, which coupled WGAN with Recurrent Neural Networks to adopt the relationship among MRI slices for fast MRI reconstruction. Hu et al. [43] designed a general texture-based GAN for MR image synthesis, where a multi-scale mechanism for the texture transfer between source and target domain was adopted. Ma et al. [44] introduced a novel GAN-based model, i.e., CSI-GAN, for medical image enhancement, where illumination regularisation and structure loss were used as constraints of training. Zhang et al. [45] combined a Retinex model with the reverse edge attention network for cerebrovascular segmentation. The utilisation of reverse edge attention module significantly improved the performance of segmentation. Yuan et al. [46] incorporated the self-attention mechanism into the generator for a global understanding of images and improved the discriminator for the utilisation of prior knowledge. Li et al. [47] proposed RSCA-GAN for fast MRI reconstruction, where both spatial and channel-wise attention mechanisms were applied in the generator. This team also applied the GAN-based model with attention mechanisms in the super-resolution task [48]. Chen et al. [49], incorporated wavelet packet decomposition into the de-aliasing GAN [34] for the texture feature. Lv et al. [50] applied transfer learning to a parallel image coupled GAN model, improving the generalisability of networks based on small samples. Jiang et al. [51] proposed FA-GAN for the super-resolution task, where both global and local feature fusion were utilised in the generator for better performance. Zhou et al. [52] designed a GAN-based model, i.e., ESSGAN, with a structurally strengthened generator, which consisted of the strengthened connection and the residual in the residual block. This team also introduced a novel spatial orthogonal attention GAN model [53], where the computational complexity was significantly decreased.

Most proposed MRI reconstruction approaches focused on integral MR image property. The design of GAN-approaches relied on loss function definitions that do not consider structural characteristics of practical value that are present in the image such as edge information. However, edge information can be crucially conclusive for clinical diagnosis. Accordingly, to solve this problem, studies related

to edge information preservation in MRI reconstruction has been reported. Yu et al. [54] proposed Ea-GANs, in which the edge information was utilised via the Sobel operator. Ea-GANs contain a generator-induced gEa-GAN, and a discriminator-induced dEa-GAN, for enriching the reconstruction images with more details. Chai et al. [55] designed an edge-guided GAN (EG-GAN), to restore brain MRI images which decoupled reconstruction into edge connection and contrast completion. Li et al. [56], proposed a dual-discriminator GAN, i.e., EDDGAN, of which one discriminator was used for holistic image reconstruction and the other one was for edge information preservation.

Recently, a large number of multi-view data based methods by considering the diversity of various views have been proposed [57, 58]. The main task of multi-view learning is to find a function to model each view and jointly optimises all the functions to improve the generalisation performance [59, 60]. In this work, the idea of multi-view learning was adopted. The multi-channel MR data we used in training was the multi-view information of the MRI raw data, by multi-coil parallel imaging. In addition, information from different views including image-space information,  $k$ -space information and edge details were all included and used as constraints on the training process.

The methods mentioned above focused on the preservation of edge information from various aspects. However, all were based on single-channel MR data. In fact, complementary multi-view information can be provided by multi-channel MR data. This work takes advantage of parallel imaging technology by using multi-channel data rather than single-channel data. In particular, we propose a novel parallel imaging coupled dual discriminator generative adversarial network (PIDD-GAN) for fast multi-channel MRI reconstruction. The main contributions can be summarised as follows.

- We introduce a dual discriminator GAN architecture, where two discriminators are used for holistic image reconstruction and edge information enhancement respectively.
- The GAN generator is an improved U-Net with local and global residual learning that stabilises and accelerates the training process. Frequency channel attention blocks (FCA Blocks) are embedded in the improved U-Net to incorporate attention mechanisms.
- Although the proposed network is designed for multi-channel MR image reconstruction, single-channel image reconstruction can also be accomplished.
- Comprehensive experiments on the Calgary-Campinas public brain MR dataset<sup>1</sup> (359 subjects) compared the proposed method with current state-of-the-art MR recon-

<sup>1</sup> <https://sites.google.com/view/calgary-campinas-dataset/>

struction methods. Ablation studies for residual learning were conducted on the MICCAI13 dataset<sup>2</sup> to validate the proposed modules. Experimental results confirm that the proposed PIDD-GAN achieves high reconstruction quality with faster processing time.

## 2 Method

### 2.1 Traditional MRI

The problem of traditional CS-MRI is to recover a vector  $x \in \mathbb{C}^N$  in image space from an undersampled vector  $y \in \mathbb{C}^M$  ( $M \ll N$ ) in  $k$ -space, which can be expressed as follows

$$y = \mathcal{M}\mathcal{F}x + n, \quad (1)$$

where  $\mathcal{M}$  denotes the undersampling trajectory,  $\mathcal{F}$  denotes the Fourier transform, and  $n$  denotes noise.

For parallel imaging, coil sensitivity encoding is incorporated in the reconstruction, i.e.,

$$y = \mathcal{M}\mathcal{F}\mathcal{C}x + n, \quad (2)$$

where  $\mathcal{C}$  is the coil sensitivity.

Let  $\mathcal{E}$  be an operator including undersampling trajectory  $\mathcal{M}$ , Fourier transform  $\mathcal{F}$  and coil sensitivity  $\mathcal{C}$ , the reconstruction problem can now be expressed as follows

$$\min_x \frac{1}{2} \|y - \mathcal{E}x\|_2^2 + \lambda R(x), \quad (3)$$

where  $R(x)$  denotes regularisation terms on  $x$ , and  $\lambda$  is a regularisation coefficient, controlling the degree of regularisation.  $\|\cdot\|_2$  denotes the  $l_2$  norm.

### 2.2 Deep learning-based fast MRI

#### 2.2.1 CNN-based fast MRI

A deep network can be incorporated into fast MRI reconstruction to generate image  $f_{\text{CNN}}(x_u | \theta)$  from the zero-filled image  $x_u$ , where  $\theta$  are the optimised parameters of the deep network. The problem can be expressed as follows

$$\min_x \frac{1}{2} \|y - \mathcal{E}x\|_2^2 + \lambda R(x) + \zeta \|x - f_{\text{CNN}}(x_u | \theta)\|_2^2, \quad (4)$$

where  $\lambda$  and  $\zeta$  are coefficients that balance each term.

#### 2.2.2 GAN-based fast MRI

In general, a GAN model consists of two parts: a generator and a discriminator. The generator aims to produce fake data

$G_{\theta_G}(z)$  that is as real as possible by modelling and sampling the distribution of the ground truth  $x$ , so that samples drawn from the modelled distribution succeed at deceiving the discriminator. The goal of the discriminator is to distinguish the fake data generated by the generator from the ground truth.

Ideally, the best discriminator can be represented as

$$D_{\theta_D}(x) = 1, \quad D_{\theta_D}(G_{\theta_G}(z)) = 0. \quad (5)$$

In this way, the generator and the discriminator form a min-max game. The training process of GAN can be described as follows

$$\min_{\theta_G} \max_{\theta_D} \mathcal{L}(\theta_G, \theta_D) = \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log D_{\theta_D}(x)] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D_{\theta_D}(G_{\theta_G}(z)))], \quad (6)$$

where  $p_{\text{data}}(x)$  is the distribution of real data, and  $p_z(z)$  is the latent variables distribution. During the training process, both sides constantly optimise themselves until the balance is reached — neither side can get better, that is, the fake samples are completely indistinguishable from the true samples.

However, the discriminator may be very confident and almost always output 0 initially. To circumvent this problem practically, the loss function is replaced by:

$$\min_{\theta_G} \max_{\theta_D} \mathcal{L}(\theta_G, \theta_D) = \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log D_{\theta_D}(x)] - \mathbb{E}_{z \sim p_z(z)} [\log(D_{\theta_D}(G_{\theta_G}(z)))]. \quad (7)$$

When the GAN model is used for the reconstruction task, the generator is trained to generate reconstructed MR images  $G_{\theta_G}(x_u)$  from zero-filled undersampled MR images  $x_u$ . The discriminator is trained to distinguish  $G_{\theta_G}(x_u)$  from the ground truth MR image  $x_t$  by maximizing the log-likelihood for estimating the conditional probability, which can be represented as

$$D_{\theta_D}(x_t) = 1, \quad D_{\theta_D}(G_{\theta_G}(x_u)) = 0. \quad (8)$$

The adversarial loss  $\mathcal{L}_{\text{adv}}$  that drives the training process can be parameterised by  $\theta_G$  and  $\theta_D$  as follows

$$\min_{\theta_G} \max_{\theta_D} \mathcal{L}_{\text{adv}}(\theta_G, \theta_D) = \mathbb{E}_{x_t \sim p_{\text{train}}(x_t)} [\log D_{\theta_D}(x_t)] - \mathbb{E}_{x_u \sim p_G(x_u)} [\log(D_{\theta_D}(G_{\theta_G}(x_u)))]. \quad (9)$$

### 2.3 Proposed dual discriminator GAN for fast MRI reconstruction

#### 2.3.1 Formulation

In this work, a first discriminator  $D_1$  is used to distinguish the reconstructed MR images  $\hat{x}_u$  from the ground truth MR images  $x_t$ , whereas an additional discriminator  $D_2$  is designed to assist the reconstruction of the edge information. Edge information is usually extracted by means of a Sobel operator  $\mathcal{S}(\cdot)$ . The edge information of the reconstructed MR image  $\mathcal{S}(\hat{x}_u)$  and the edge information of the MR ground truth  $\mathcal{S}(x_t)$  are fed to  $D_2$ , so that the result is counted into the adversarial loss  $\mathcal{L}_{\text{adv}}$ . The new adversarial loss  $\mathcal{L}_{\text{adv}}$  can be defined as follows

<sup>2</sup> <https://mrbrains13.isi.uu.nl/>

$$\begin{aligned}
& \min_{\theta_G} \max_{\theta_{D_1}} \max_{\theta_{D_2}} \mathcal{L}_{\text{adv}}(\theta_G, \theta_{D_1}, \theta_{D_2}) \\
& = \mu \{ \mathbb{E}_{x_r \sim p_{\text{train}}(x_r)} [\log D_{\theta_{D_1}}(x_r)] - \mathbb{E}_{x_u \sim p_G(x_u)} [\log D_{\theta_{D_1}}(\hat{x}_u)] \} \\
& + \nu \{ \mathbb{E}_{x_r \sim p_{\text{train}}(x_r)} [\log D_{\theta_{D_2}}(\mathcal{S}(x_r))] - \mathbb{E}_{x_u \sim p_G(x_u)} [\log D_{\theta_{D_2}}(\mathcal{S}(\hat{x}_u))] \}, \quad (10)
\end{aligned}$$

where  $\mu$  and  $\nu$  denote the weights of the discriminator for the holistic image and the edge information correspondingly, and  $\hat{x}_u = G_{\theta_G}(x_u) + x_u$  as it will be later discussed.

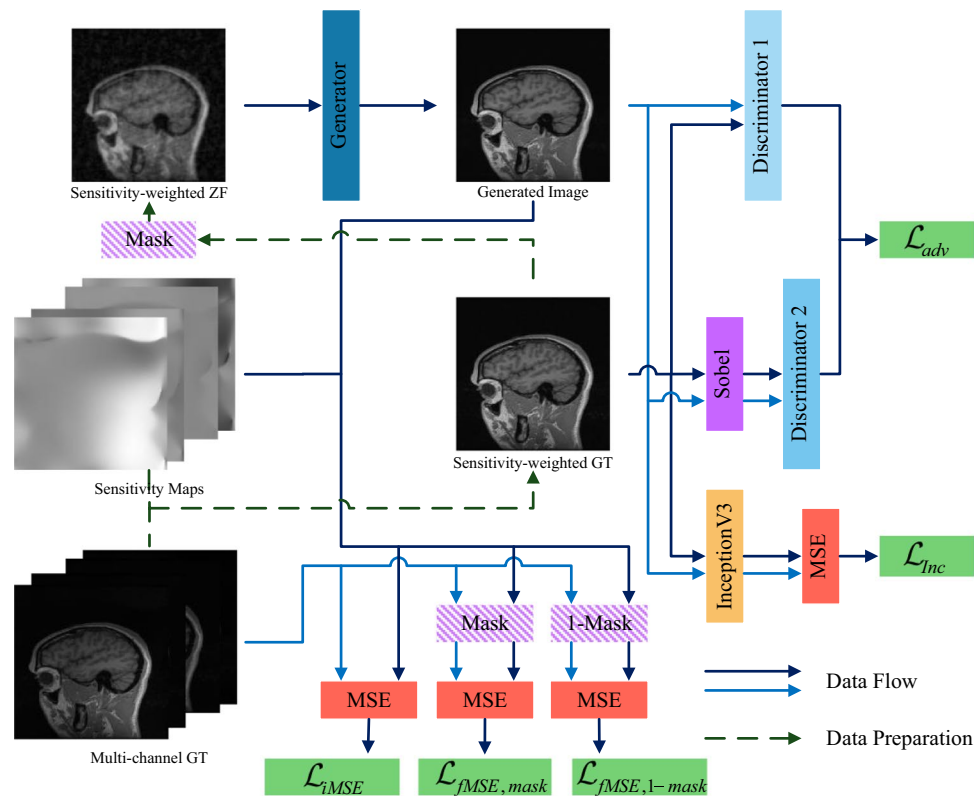
## 2.4 Network architecture

The overall architecture of the proposed PIDD-GAN for MR image reconstruction is shown in Fig. 1. Sensitivity-weighted ground truth  $x_r$  is derived from the multi-channel ground truth  $x_i^q$ , corresponding sensitivity map  $\mathcal{C}^q$  ( $q$  denotes the coil number) and sensitivity-weighted zero-filled image  $x_u$  by undersampling  $x_r$ . The generator produces the reconstructed MR image  $\hat{x}_u$  from  $x_u$ . Two

discriminators are used for the holistic image and the edge information, respectively.

Ronneberger et al. [38] proposed U-Net for semantic segmentation of medical images. This deep architecture comprises an encoder path to capture context and a symmetric decoder path that enables precise localisation, i.e., a U-shaped model. Skip connections are applied between corresponding layers in encoder and decoder paths, passing features directly from undersampling path to upsampling path. U-Net can be trained in an end-to-end manner and performs well in medical image segmentation [34, 36, 56].

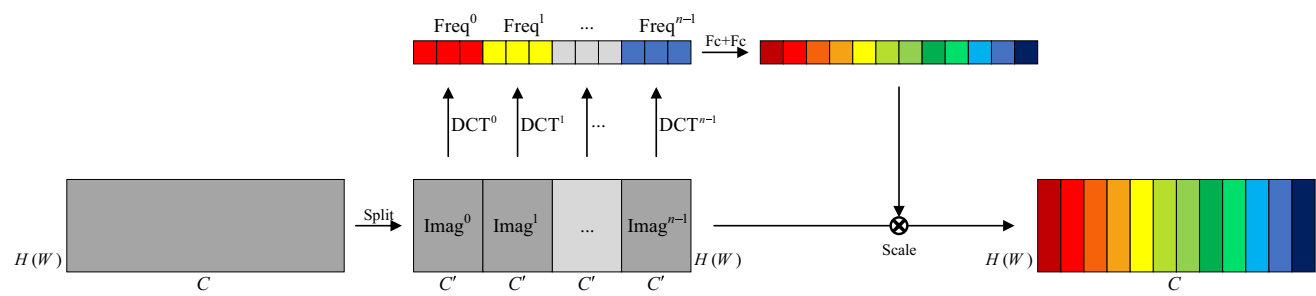
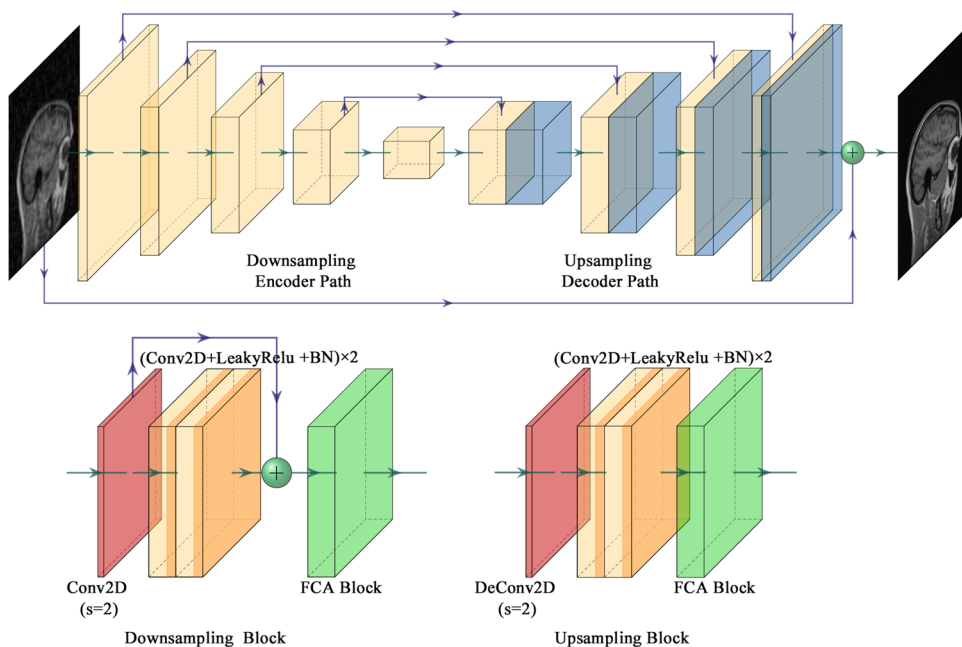
An improved U-Net is proposed as the generator in our proposed PIDD-GAN for higher reconstruction performance, where the generator input is sensitivity-weighted zero-filled image  $x_u$ . As shown in Fig. 2, in our improved U-Net, cascaded downsampling blocks are placed in the encoder path (left), and corresponding cascaded upsampling blocks are placed in the decoder path (right). Skip connection and concatenation are applied between the downsampling blocks and the symmetrical upsampling blocks with



**Fig. 1** The architecture of our PIDD-GAN. The generator produces generated MR images from sensitivity-weighted zero-filled MR images. Sensitivity maps and multi-channel ground truth images are used to produce sensitivity-weighted ground truth images. Two discriminators are used for the reconstruction of holistic images and edge information respectively. A pre-trained Inception V3 model is applied for the perceptual loss. Dark blue and sky blue lines are

both used to represent the data flow for easier identification. Green dash lines are used to represent the data preparation steps before the training. (MASK: undersampling trajectory mask, I-MASK: inverse undersampling trajectory mask, InceptionV3: a pre-trained InceptionV3 model, MSE: mean squared error, Sobel: Sobel operator, Discriminator 1: the discriminator for holistic images, Discriminator 2: the discriminator for edge information)

**Fig. 2** The structure of the generator in PIDD-GAN. Four cascaded downsampling blocks and four cascaded upsampling blocks are placed in encoder and decoder paths respectively. Skip connection and concatenation are used between layers with the same scales. A shortcut connection is applied between the input and output of the generator for further refined learning. (Conv2D: 2D convolutional layer, DeConv2D: 2D deconvolution layer, BN: batch normalisation, LeakyRelu: Leaky ReLU layer)



**Fig. 3** The structure of the FCA Block. An  $H \times W \times C$  feature map is divided into  $n H \times W \times C'$  parts, i.e.,  $\text{Imag}^i$  ( $i = 0, 1, \dots, n - 1$ ). The  $i^{\text{th}}$  frequency component  $\text{Freq}^i$  is obtained by corresponding  $\text{Imag}^i$  using

$\text{DCT}^i$ . The weight of channels can be counted by all the frequency components using two full connection layers (FC in the Fig)

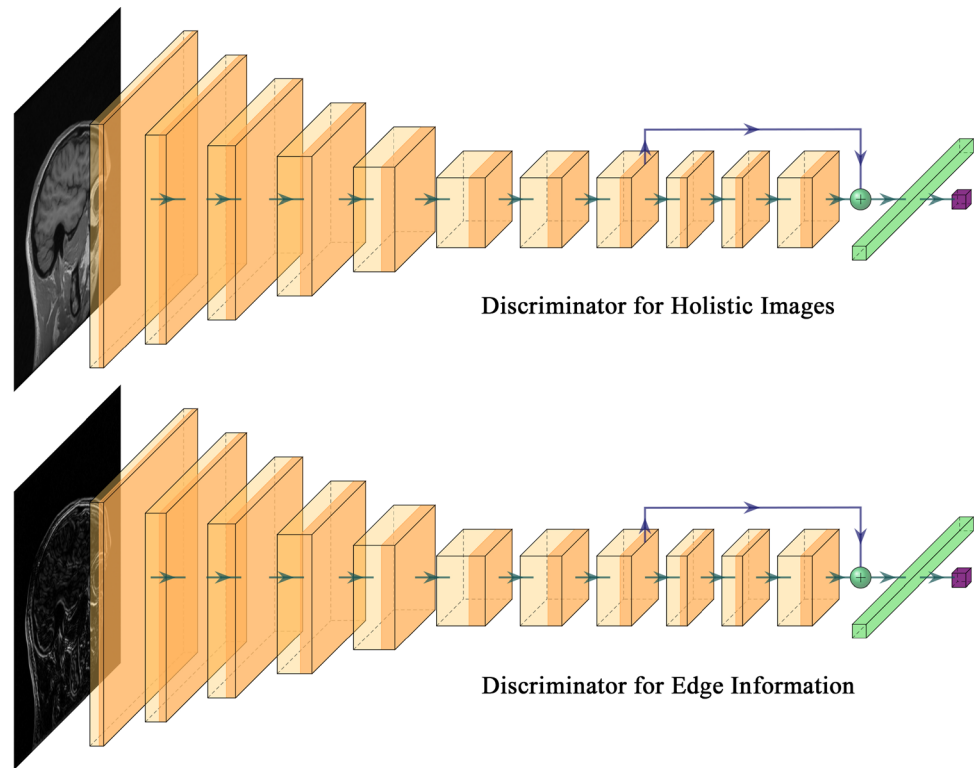
the same scale to preserve the feature from different levels and, ultimately, yield better reconstruction details.

Figure 2 shows the  $i^{\text{th}}$  downsampling block structure. First, a  $3 \times 3$  convolution layer with stride = 2 is applied to downsample the  $(i - 1)^{\text{th}}$  output. Then, a residual block extracts further features and avoids gradient vanishing and exploding problems [61]. There are two  $3 \times 3$  convolution layers in the backbone and a  $1 \times 1$  convolution layer adjusting the channel and fusing feature maps at different scales. Leaky ReLU layers and Batch normalisation (BN) are applied after each convolution layer except the final one. Finally, a FCA Block [62] learns the different channel weights with attention. The structure of upsampling blocks and downsampling blocks are similar. A deconvolution layer with stride = 2 is applied to upsample the  $(i - 1)^{\text{th}}$  output. The shortcut is removed to reduce computation cost but maintain high reconstruction quality here.

Traditional CNNs treat all channels in a feature map with the same importance, ignoring the importance differences. Therefore, the attention mechanism is adopted to make use of importance difference information by learning the different channel weights, i.e., effective channels have high weights and ineffective channels have small weights, which helps to train the model and enhance the results.

FCA Block [62] is a novel attention mechanism based on the squeeze and excitation block (SE Block) [63], as shown in Fig. 3. First, an  $H \times W \times C$  feature map is squeezed into a  $1 \times C$  vector. Then channel weights are extracted by two fully-connected layers, and channel weights multiply the original feature maps. A two-dimensional discrete cosine transform (DCT) is applied in FCA Blocks to squeeze the feature map, rather than global average pooling employed in SE Blocks, since this latter operation is equivalent to the lowest DCT frequency. Hence using only GAP leads to loss of other frequency components in the feature channel

**Fig. 4** The structure of the discriminator in PIDD-GAN. A standard 11-layer CNN is used as the discriminator, where each convolutional layer is followed by a BN layer and a Leaky ReLU layer. A full connection layer and a Sigmoid layer are cascaded at the end of discriminators to output the result of classification. The upper discriminator is the discriminator for holistic image and the lower one is for edge information



containing useful information. In the squeezing step, the  $H \times W \times C$  feature map is divided evenly into  $n$  parts (each size is  $H \times W \times C'$ ). The squeezing step can be represented as

$$\text{Freq}^i = \text{DCT}^i(\text{Imag}^i), \quad (i = 0, 1 \dots n - 1), \quad (11)$$

where  $\text{DCT}^i$  denotes the  $i^{\text{th}}$  preset DCT template and  $\text{Freq}^i$  denotes the corresponding frequency component. In this work, FCA Blocks are utilised after every residual blocks in the generator.

As the network architecture goes deeper, more granular information can be extracted, but leading to gradient vanishing and exploding problems, making the network converge slowly. The introduction of residual learning [61] solves this problem effectively. The main idea of residual learning is the utilisation of shortcut connections between the convolution layer. It makes the deep network easier to be trained and converge faster. Global residual learning (GR) is applied in our improved U-Net. The output of the generator adopts  $\hat{x}_u = G_{\theta_G}(x_u) + x_u$  instead of  $\hat{x}_u = G_{\theta_G}(x_u)$  in the original U-Net. This change transfers the generator from a conditional generative function to a refinement function. This work also applies local residual learning (LR) by the shortcut connection in each residual block in the downsampling path. The utilisation of LR aimed to stabilise the training and accelerate the model convergence.

Traditional GAN trains a single discriminator to compete against the generator. Although it improves reconstruction

quality compared with other methods, only integral MR image properties are considered, without enhancing edge details. The current study proposes a dual discriminator GAN for edge information enhancement. The generated MR image  $\hat{x}_u$  and sensitivity-weighted MR ground truth  $x_t$  are fed into discriminator  $D_1$  for holistic image reconstruction. We use the Sobel operator  $\mathcal{S}(\cdot)$  to extract the edge information from MR images, and input edge information for the reconstructed image  $\mathcal{S}(\hat{x}_u)$  and ground truth  $\mathcal{S}(x_t)$  into  $D_2$ . Thus, both holistic image information and edge details can be simultaneously reconstructed.

Figure 4 shows the common network structure used for both discriminators. First, a cascade of  $3 \times 3$  convolution layers with stride = 2 downsamples and extracts MR image features. Two  $1 \times 1$  convolution layers and a residual block follow the final  $3 \times 3$  convolution layer. The residual block consists of three  $1 \times 1$  convolution layers, and input and output are connected by a shortcut. All convolution layers above are followed by a BN layer and Leaky ReLU layer. Finally a full connection layer and a Sigmoid layer output the prediction results. Results of both discriminators are incorporated into the adversarial loss  $\mathcal{L}_{\text{adv}}$ .

## 2.5 Loss function

This study introduces content loss to train the generator for better reconstruction quality. Content loss comprises

pixel-wise mean square error (MSE) loss  $\mathcal{L}_{\text{iMSE}}$ ; the frequency MSE loss  $\mathcal{L}_{\text{fMSE,mask}}$  and  $\mathcal{L}_{\text{fMSE,1-mask}}$ ; and the perceptual InceptionV3 loss  $\mathcal{L}_{\text{Inc}}$ . Pixel-wise MSE loss can be defined as

$$\min_{\theta_G} \mathcal{L}_{\text{iMSE}}(\theta_G) = \sum_q \frac{1}{2} \|x_t^q - \mathcal{C}^q \hat{x}_u\|_2^2, \quad (12)$$

where  $\mathcal{C}^q$  is the sensitivity map of the  $q^{\text{th}}$  coil.  $\mathcal{L}_{\text{iMSE}}$  is used to reduce the artefact between the generated image and the ground truth. However, optimisation with only  $\mathcal{L}_{\text{iMSE}}$  would make the reconstructed image lack coherent image details. Therefore, frequency MSE loss is applied to train the generator in  $k$ -space.

The frequency MSE loss can be defined as

$$\min_{\theta_G} \mathcal{L}_{\text{fMSE,mask}}(\theta_G) = \sum_q \frac{1}{2} \|y_{\mathcal{M}}^q - \mathcal{M}\mathcal{F}\mathcal{C}^q \hat{x}_u\|_2^2, \quad (13)$$

$$\min_{\theta_G} \mathcal{L}_{\text{fMSE,1-mask}}(\theta_G) = \sum_q \frac{1}{2} \|y_{1-\mathcal{M}}^q - (1 - \mathcal{M})\mathcal{F}\mathcal{C}^q \hat{x}_u\|_2^2, \quad (14)$$

where  $\mathcal{L}_{\text{fMSE,mask}}$  eliminates differences between undersampled generated images  $\mathcal{M}\mathcal{F}\mathcal{C}^q \hat{x}_u$  and undersampled  $k$ -space measurements  $y_{\mathcal{M}}^q$ .  $\mathcal{L}_{\text{fMSE,1-mask}}$  is used to minimise the differences between the interpolated data based on the generated image  $(1 - \mathcal{M})\mathcal{F}\mathcal{C}^q \hat{x}_u$  and the unacquired  $k$ -space data  $y_{1-\mathcal{M}}^q$ .

In addition, the perceptual Inception V3 loss can be defined as

$$\min_{\theta_G} \mathcal{L}_{\text{Inc}}(\theta_G) = \frac{1}{2} \|f_{\text{Inc}}(x_t) - f_{\text{Inc}}(\hat{x}_u)\|_2^2, \quad (15)$$

where  $f_{\text{Inc}}(\cdot)$  denotes the Inception V3 network [15].  $\mathcal{L}_{\text{Inc}}$  is used to optimise the perceptual quality of reconstructed results.

The adversarial loss is defined as

$$\begin{aligned} & \min_{\theta_G} \max_{\theta_{D_1}} \max_{\theta_{D_2}} \mathcal{L}_{\text{adv}}(\theta_G, \theta_{D_1}, \theta_{D_2}) \\ &= \mu \{ \mathbb{E}_{x_t \sim p_{\text{train}}(x_t)} [\log D_{\theta_{D_1}}(x_t)] - \mathbb{E}_{x_u \sim p_G(x_u)} [\log D_{\theta_{D_1}}(\hat{x}_u)] \} \\ &+ \nu \{ \mathbb{E}_{x_t \sim p_{\text{train}}(x_t)} [\log D_{\theta_{D_2}}(S(x_t))] - \mathbb{E}_{x_u \sim p_G(x_u)} [\log D_{\theta_{D_2}}(S(\hat{x}_u))] \}, \end{aligned} \quad (16)$$

Hence, the total loss can be described as

$$\mathcal{L}_{\text{TOTAL}} = \alpha \mathcal{L}_{\text{iMSE}} + \beta (\mathcal{L}_{\text{fMSE,mask}} + \mathcal{L}_{\text{fMSE,1-mask}}) + \gamma \mathcal{L}_{\text{Inc}} + \mathcal{L}_{\text{adv}}, \quad (17)$$

where  $\alpha$ ,  $\beta$  and  $\gamma$  are coefficients balancing each term in the loss function.

## 3 Experiments and results

### 3.1 Datasets

This work used the Calgary Campinas multi-channel dataset [64] and the MICCAI 2013 grand challenge single-channel dataset [65], which are both publicly available for the experiment section.

The Calgary Campinas dataset was used to train and validate our proposed method and compare with other methods. The MR images were acquired with a 12-channel coil. We randomly chose 15360 12-channel T1-weighted brain MR images. The dataset was divided into training, validation, and testing sets (7680, 3072, and 4608 slices respectively), according to the ratio of 5:2:3.

The MICCAI 2013 grand challenge dataset was used for ablation studies. We randomly chose 18850 single-channel T1-weighted brain MRI images, and divided these into training, validation, and testing sets (9935, 3974, and 5961 slices, i.e., ratio of 5:2:3, respectively).

### 3.2 Implementation detail

The proposed PIDD-GAN was implemented using PyTorch, and trained and tested on an NVIDIA TITAN RTX GPU with 24GB GPU memory. We set the same hyperparameters for all experiments. Adam optimiser with a momentum of 0.5 was adopted during training. Empirically, we set  $\alpha = 15$ ,  $\beta = 0.1$ ,  $\gamma = 10$  (for Inception V3) or 0.0025 (for VGG) in the total loss function, and  $\mu = 0.6$ ,  $\nu = 0.4$  for discriminator weights. Initial and minimal learning rates were 0.001 and 0.00001, decayed by 50% every 5 epochs. The batch size was set to 12. An early stopping mechanism was adopted to halt training and prevent overfitting: training was stopped if there were no normalised mean square error (NMSE) reduction on the validation set for 8 epochs.

To evaluate the proposed method, we compared the following variations: (1) PIDD: dual discriminator GAN model trained with parallel imaging dataset; (2) PISD: single discriminator GAN trained with parallel imaging dataset; (3) nPIDD: dual discriminator GAN model trained with single-channel imaging dataset. Thus, the role for each component in the proposed method can be compared more fairly and clearly.

### 3.3 Evaluation methods

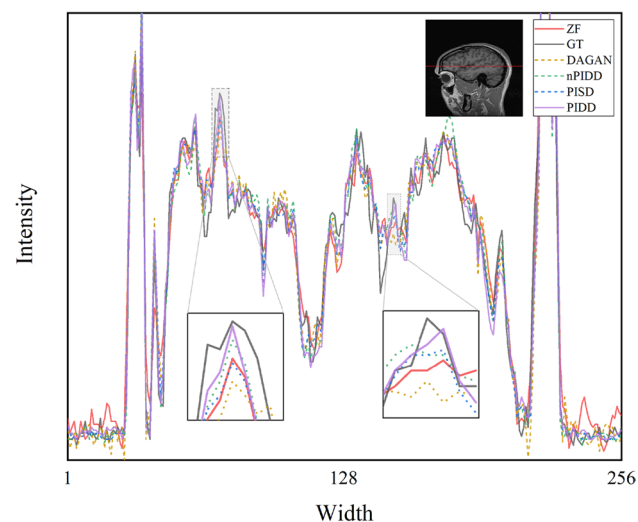
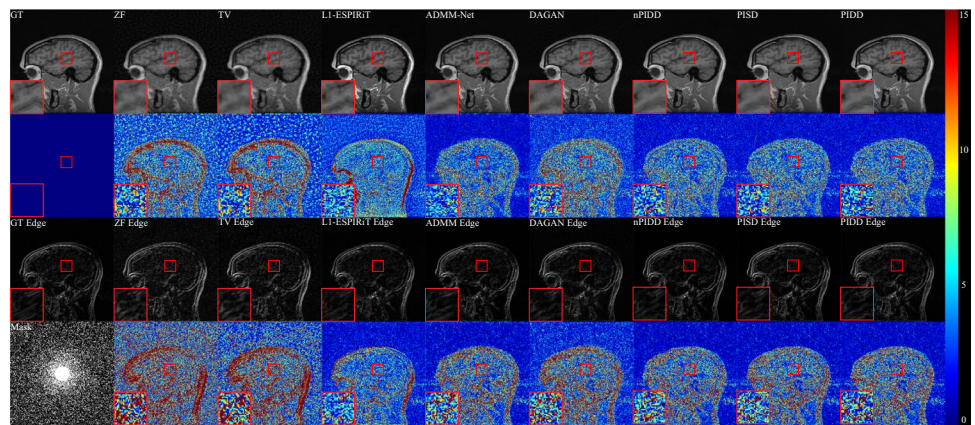
Various assessment methods were applied to evaluate reconstruction quality. Normalised mean square error measures average squared difference between generated images and ground truth. Structural similarity (SSIM) was used to



**Table 1** Testing results of different method. Results are obtained by Gaussian 2D masks with 30%. (Bold values indicate the best performed method)

Method	Gaussian 30%				MOS		Time	
	NMSE	PSNR	SSIM	FID	Holistic	Edge	GPU Time (ms)	CPU Time (s)
GT					3.83±0.52	3.70±0.46		
ZF	0.0247±0.0013	28.4701±0.3806	0.8756±0.0060	103.33	1.03±0.18	1.03±0.18		
TV	0.0254±0.0023	28.5259±0.6997	0.8840±0.0117	106.04				0.64±0.02
L1-ESPIRiT	<b>0.0071±0.0008</b>	32.4709±0.8978	0.9056±0.0122	21.21				53.24±0.82
ADMM-Net	0.0129±0.0022	31.5171±0.9744	0.9368±0.0115	67.97	1.10±0.30	1.13±0.34		1.19±0.12
DAGAN	0.0169±0.0011	30.1123±0.3710	0.9072±0.0053	53.67	2.60±0.66	2.43±0.56	2.99±0.01	0.21±0.01
nPIDD	0.0099±0.0009	<b>32.4809±0.4680</b>	<b>0.9438±0.0044</b>	29.42			4.95±0.01	0.47±0.00
PISD	0.0103±0.0009	32.2685±0.4808	0.9398±0.0050	24.12	3.70±0.59	3.26±0.51	4.95±0.01	0.47±0.00
PIDD	0.0101±0.0009	32.3694±0.4443	0.9425±0.0046	<b>15.38</b>	<b>3.80±0.54</b>	<b>3.40±0.66</b>	<b>4.95±0.01</b>	<b>0.47±0.00</b>

Bold values indicate the best performed method

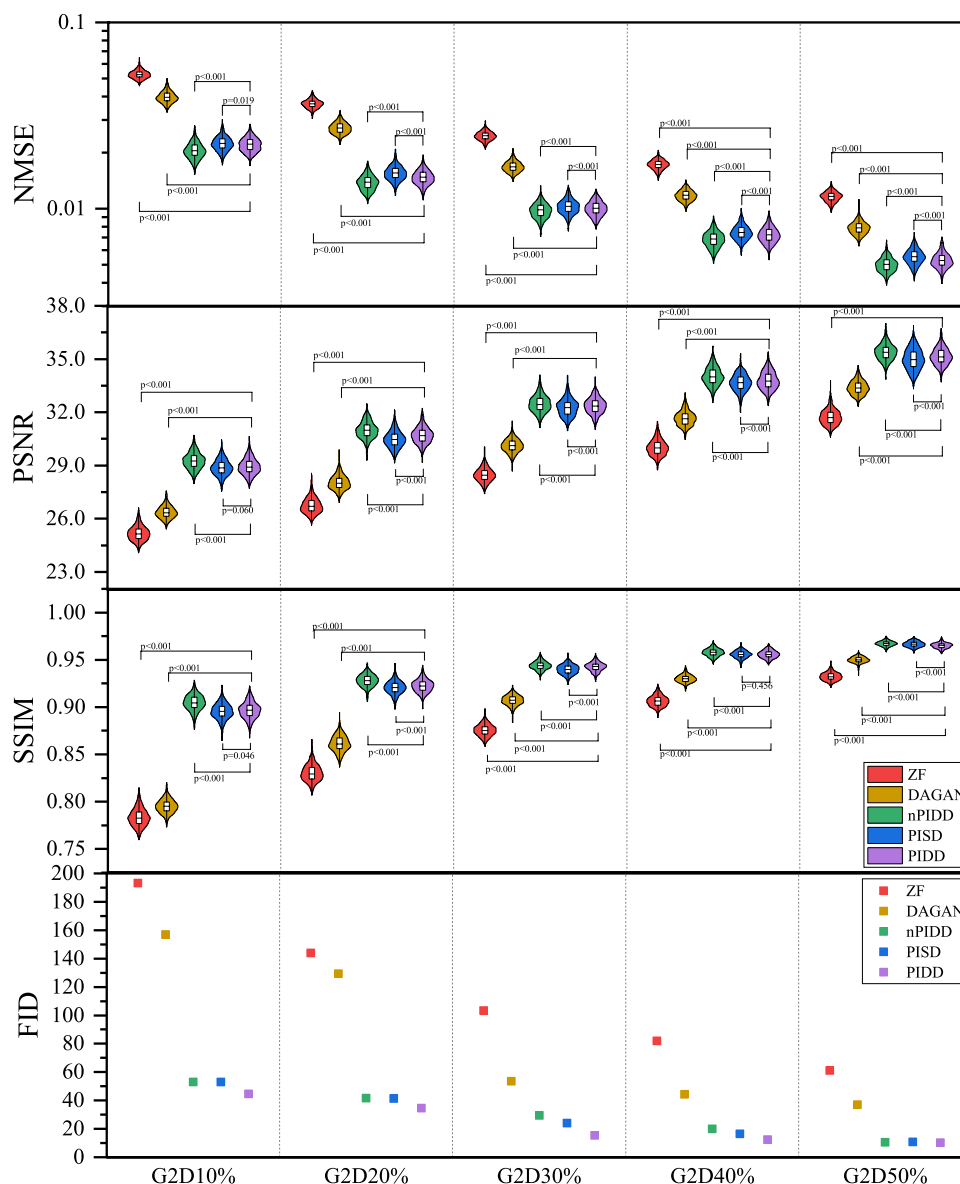
**Fig. 5** Testing examples of different methods. Results are obtained by Gaussian 2D masks with 30%. Line1: MR images; Line2: Differences of MR images ( $\times 15$ ); Line3: Edge Information of MR images; Line4: Differences of Edge Information ( $\times 15$ )**Fig. 6** Horizontal line profiles of different methods. Results are obtained by Gaussian 2D masks with 30%

measure similarity between two images and hence predict perceived quality. Peak signal-to-noise ratio (PSNR) is the ratio between maximum signal power and corrupting noise power, which quantifies the representations fidelity.

However, PSNR and SSIM do not necessarily correspond with visual quality for human observers. Therefore, we adopted two different metrics to evaluate reconstruction quality. Frchet inception distance (FID) [66] measures similarity between two sets of images, calculated by computing the Frchet distance between two Gaussian fitted feature representations for the inception network. FID correlates well with human derived visual quality and it is widely used to evaluate GAN sample quality.

Meanwhile, the judgement of domain experts is considered. Mean opinion score (MOS) from expert observers was used to evaluate the holistic image and edge information for the reconstructed images. Likert scales [67] from 1 (poor), 2 (fair), 3 (good) to 4 (very good) were used that were based on the holistic image quality and edge information

**Fig. 7** Testing results of the experiment on masks using different Gaussian 2D downsampling percentage



quality, the visibility of the atrial scar and occurrence of the artefacts. All compared reconstruction results of different methods were randomly shuffled and blinded for the expert observers' scrutiny.

### 3.4 Comparisons with other methods

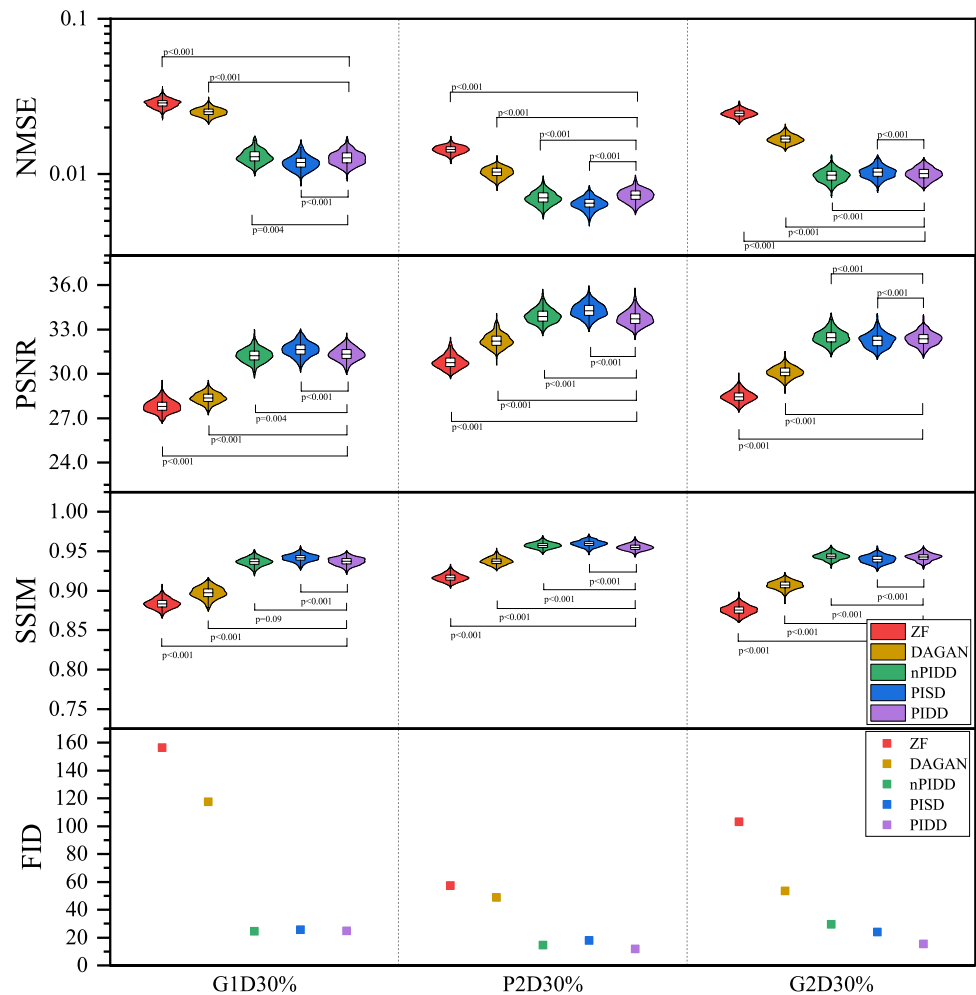
To better evaluate the reconstruction performance of the PIDD-GAN, we compare it with other traditional and MR reconstruction algorithms, including TV [9], L1-EPsRiT [68], ADMM-Net [20], DAGAN [34], as well as PISD and nPIDD. Among them, TV, ADMM-Net, DAGAN, nPIDD were implemented based on the single-channel MRI data, whereas L1-EPsRiT, PIDD, PISD were implemented based on the multi-channel MRI data. The Calgary Campinas multi-channel MRI data was used in this experiment.

The testing results are shown in Table 1. Zero-filled image (ZF) is undersampled by a Gaussian 30% downsampling trajectory. According to the comparison study results, the proposed method shows significant improvement for all test indicators compared with other methods. The FID of PIDD is significantly lower compared to PISD and nPIDD.

Reconstruction samples of six methods, including PIDD, PISD, DAGAN, ADMM-Net, together with GT and ZF, were chosen for unbiased rating by domain experts. PIDD achieved the highest MOS (except GT) both in holistic image and edge information reconstruction.

Testing examples of different methods are shown in Fig. 5. Our method is superior to other methods in terms of overall reconstruction quality and edge information reconstruction. We can see that ZF is quite blurred, mixed with many artefacts. Compared with traditional TV and L1-EPsRiT

**Fig. 8** Testing results of the experiment on masks using different downsampling trajectories



methods, the noise reduction of ADMM and DAGAN are greatly improved, but the detailed information is still lost significantly. The reconstruction result of L1-ESPIRiT shows rich details, but it is poor in de-noising and time-consuming. Compared with PIDD and PISD, nPIDD lacks multi-channel information. Although the basic structure can be completely restored, and most noise is reduced, the excessive smoothing phenomenon is severe compared to PIDD and PISD. From the zoom-in area, it can be seen that PIDD clearly reconstructs the edge information of the brain, but this structure is very shallow in the results of PISD.

Horizontal line profiles of the samples in Fig. 5 are shown in Fig. 6. ZF and DAGAN still contain lots of noise, while our proposed methods preserve more detail information. Zoom-in areas clearly show that the line profile of PIDD-DAGAN achieve more accurate than other methods.

### 3.5 Experiments on mask

In this experiment, different downsampling trajectories were adopted to evaluate the robustness of the proposed

method. G2D10%, G2D20%, G2D30%, G2D40%, G2D50%, G1D30% and P2D30% indicate Gaussian 2D 10%, Gaussian 2D 20%, Gaussian 2D 30%, Gaussian 2D 40%, Gaussian 2D 50%, Gaussian 1D 30% and Poisson 2D 30% downsampling trajectories, respectively. This experiment was tested on the Calgary Campinas multi-channel MRI dataset.

The testing result are shown in Figs. 7 and 8. Experimental results exhibit the same trend under all the different downsampling trajectories. The proposed method provides significant advantages for low downsampling percentage (high acceleration factor), with correspondingly significantly improved reconstruction quality.

The testing examples of the reconstruction are shown in Fig. 9. It can be seen from the results that PIDD has a significant recovery effect on edge information in the case of a low downsampling percentage (10–30%). The edge information of PIDD, particularly in sulci and cerebellum areas, is greatly preserved, compared to PISD. The texture details of PIDD are also richer than nPIDD. In the case of a high downsampling percentage (40–50%), the

**Fig. 9** Testing examples of the experiment on masks using different downsampling trajectories and percentage



reconstruction problem becomes simpler, and the advantages of PIDD with enhanced reconstruction of edge information and the advantages of multi-channel data are less obvious. In the experiment of Gaussian 2D 50% undersampling, PIDD, PISD and nPIDD basically have the same quality of the reconstruction.

### 3.6 Experiments on noise

In this experiment, the same downsampling trajectory and different noise levels were used to test the reconstruction performance of the model under the influence of noise. This experiment was tested on the Calgary Campinas multi-channel MRI dataset. Here the noise level (NL) is defined as follows

$$NL = \frac{N}{N + S}, \quad (18)$$

where  $S$  and  $N$  denotes the power of signal and noise respectively.

The testing results are shown in Fig. 10, and testing samples are shown in the Fig. 11.

All considered methods can restore image structure and edge information for low and medium noise levels (20–50%), with PIDD having strong advantages over PISD. This advantage weakens as noise level increases, and PISD, which focuses on overall information recovery, performance slightly surpasses PIDD, which focuses more on edge information preservation, for high noise (70–80%).

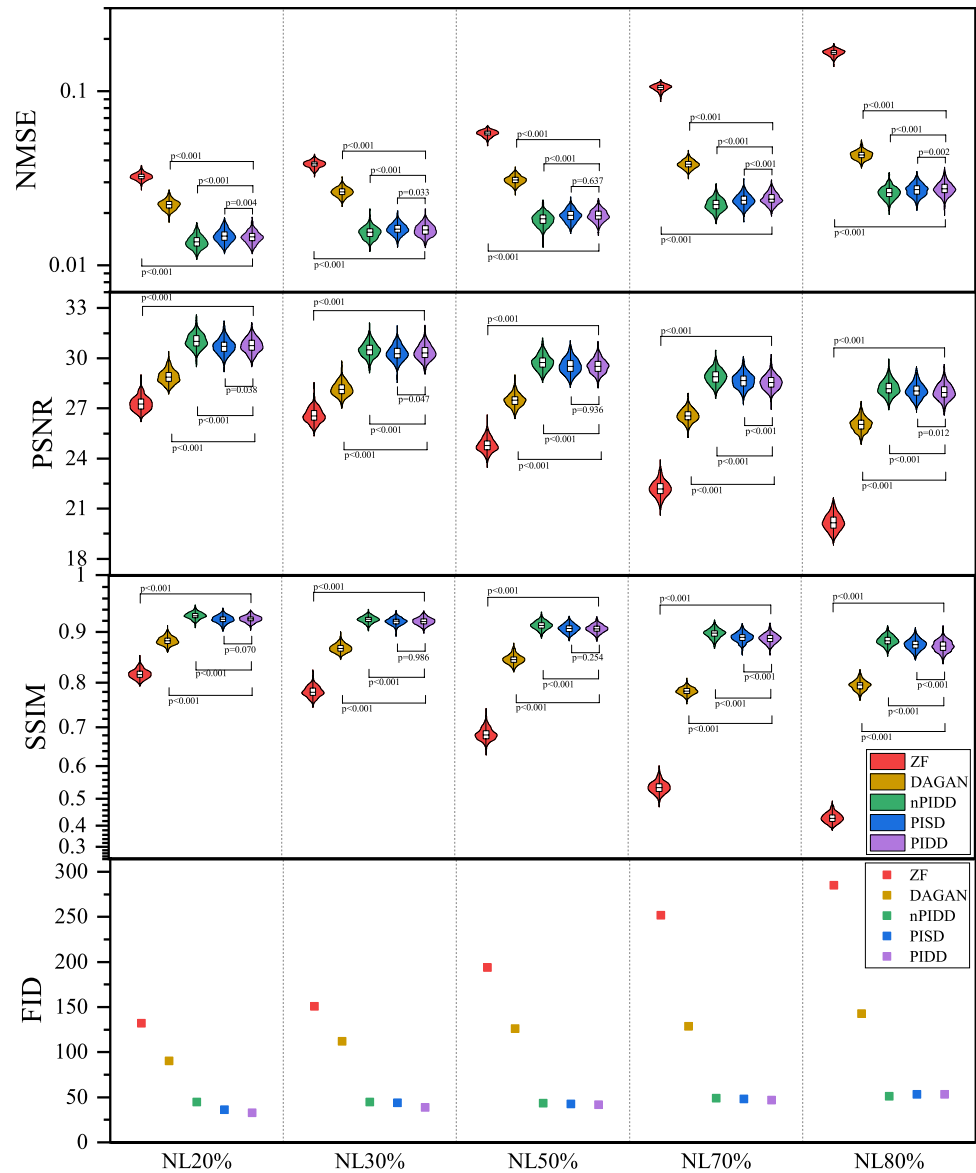
### 3.7 Ablation experiments on residual learning

In this experiment, the effect of residual learning in the network was discussed. Our proposed model was tested on MICCAI 2013 grand challenge dataset, using Gaussian 1D 30% downsampling.

The experiment was divided into four groups: (1) GRLR (model with GR and LR), (2) GRnLR (model with GR without LR), (3) nGRLR (model with LR without GR), (4) nGRnLR (model without LR and GR). Early stopping strategy was turned off in this experiment to prolong the training process for a better and more distinguishable comparison for the training step.

Figure 12 shows NMSE, SSIM, PSNR and generator loss (G Loss) of the four groups changing with the training process, and Fig. 13 shows testing examples with respect to different epoch weights.

**Fig. 10** Testing results of the experiment on noise using different noise level



Models with GR (GRLR, GRnLR) have faster convergence and better final results compared with those without GR (nGRLR, nGRnLR). If the model applies GR, then using LR has little further impact effect on the results. For non-GR models, nGRLR converges significantly slower than nGRnLR but final results are superior. Therefore, we chose GRLR as the generator for subsequent study.

## 4 Discussion

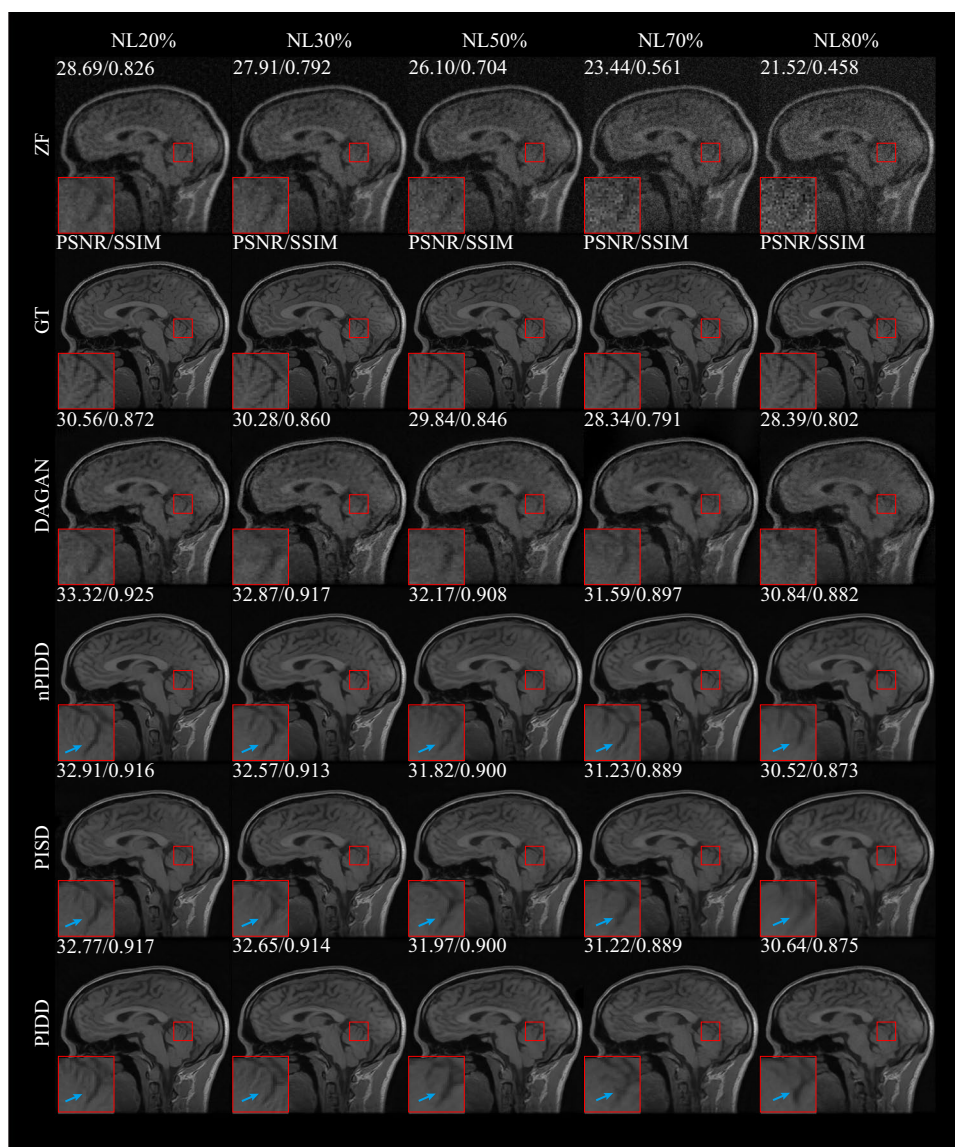
In this work, we introduce the PIDD-GAN for multi-channel MRI reconstruction using multi-view parallel imaging information, focusing on the enhancement of

edge information and the utilisation of the multi-channel MR data.

Experimental results verified that the proposed dual discriminator design does greatly improve reconstruction quality of edge information, particularly for sulci and cerebellum areas with rich edge information. The utilisation of multi-channel data also reduces the excessive smoothing phenomenon to a certain extent, and the texture of the tissue can be also better preserved.

During experiments, we found that SSIM and PSNR do not reflect the reconstruction quality very well, since these two indicators have better tolerance for over-smoothing and are relatively insensitive to the edge information details. Therefore, we adopted FID using a deep network

**Fig. 11** Testing examples of the experiment on noise using different noise level



and MOS based on subjective scoring by experts to assess our results more comprehensively.

Meanwhile, we did a series of experiments to test the role of the attention mechanism (FCA Block and SE Block) in the entire model. In most cases, the utilisation of attention mechanisms slightly improved the convergence speed and final results of the model. However, incorporating the attention mechanism degraded the final result for a few cases. The specific mechanism requires further study.

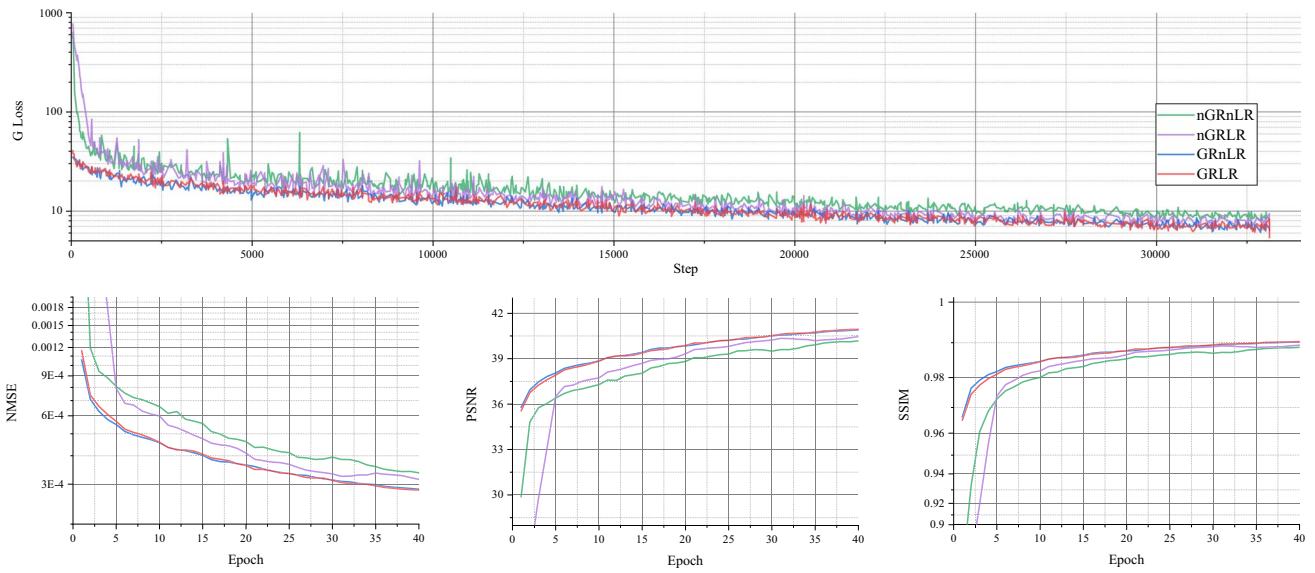
The proposed model still has some remaining limitations. Dual discriminator structure superiority reduces for high downsampling percentage, and single discriminator structure, focusing on overall recovery, may offer better performance. Noise can affect edge information extraction, particularly for higher noise levels. This effect is very marked when using the Sobel operator, and hence may

jeopardise  $D_2$  performance, reducing image reconstruction quality.

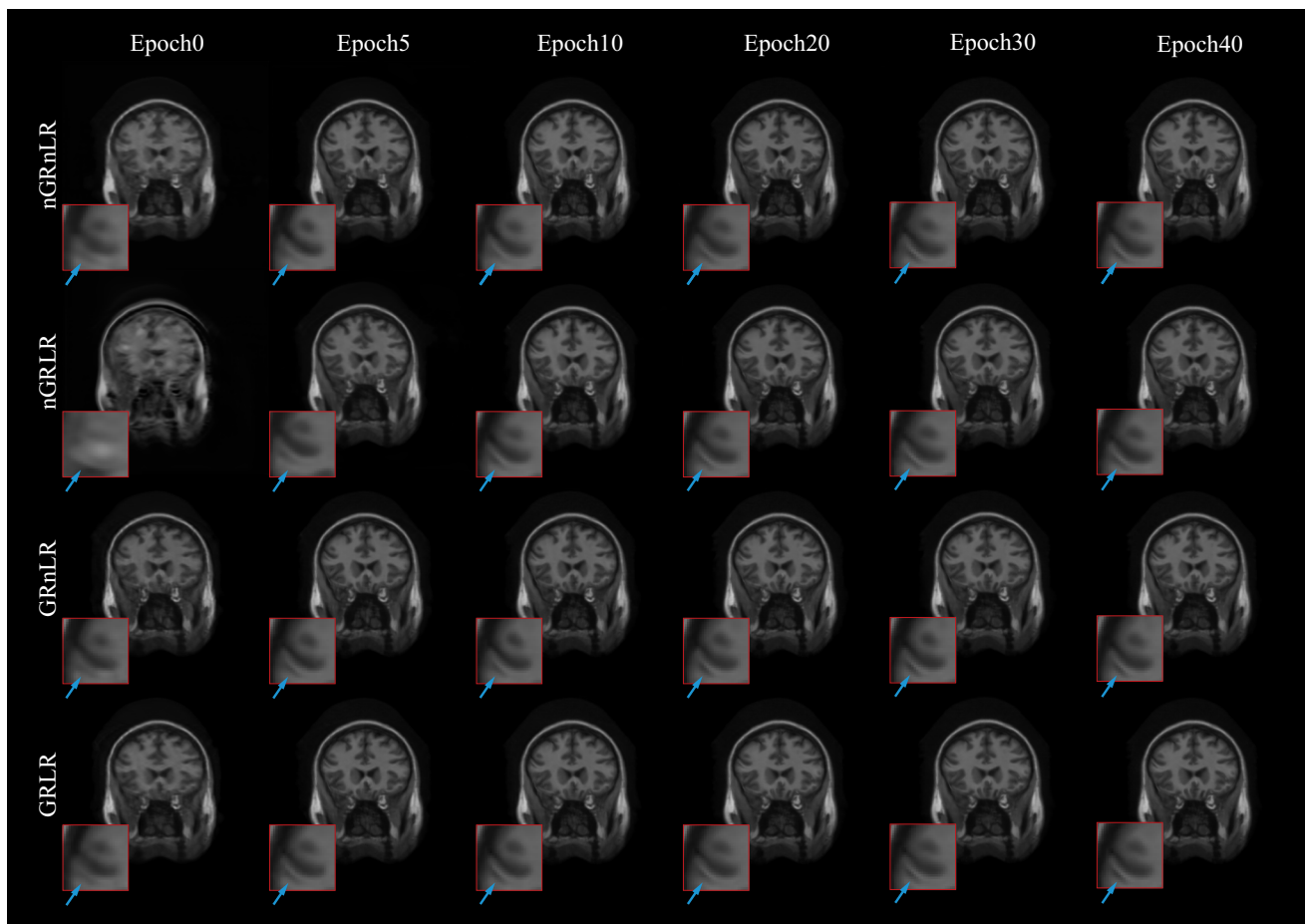
Further studies will continue to focus using multi-channel information to help edge information reconstruction and restoration. We will also consider how to reduce network calculations and improve network efficiency.

## 5 Conclusion

This work proposed a parallel imaging based dual discriminator generative adversarial network for multi-channel MRI reconstruction, enhancing edge information and multi-channel MR data utilisation using multi-view information. Experiment results verified that the proposed method offers significantly better performance preserving edge information for MRI reconstruction.



**Fig. 12** Testing results of the ablation studies with residual learning. GRLR: model with GR and LR; GRnLR: model with GR without LR; nGRLR: model with LR without GR; nGRnLR: model without LR and GR



**Fig. 13** Testing examples of the ablation studies with residual learning. GRLR: model with GR and LR; GRnLR: model with GR without LR; nGRLR: model with LR without GR; nGRnLR: model without LR and GR

**Acknowledgements** This work was supported in part by the Zhejiang Shuren University Basic Scientific Research Special Funds, in part by the European Research Council Innovative Medicines Initiative (DRAGON, H2020-JTI-IMI2 101005122), in part by the AI for Health Imaging Award (CHAIMELEON, H2020-SC1-FA-DTS-2019-1 952172), in part by the UK Research and Innovation Future Leaders Fellowship (MR/V023799/1), in part by the British Heart Foundation (Project Number: TG/18/5/34111, PG/16/78/32402), in part by the Foundation of Peking University School and Hospital of Stomatology [KUSSNT-19B11], in part by the Peking University Health Science Center Youth Science and Technology Innovation Cultivation Fund [BMU2021PYB017], in part by the National Natural Science Foundation of China [61976120], in part by the Natural Science Foundation of Jiangsu Province [BK20191445], in part by the Qing Lan Project of Jiangsu Province, in part by National Natural Science Foundation of China [61902338], in part by the Project of Shenzhen International Cooperation Foundation [GJHZ20180926165402083], in part by the Basque Government through the ELKARTEK funding program [KK-2020/00049], and in part by the consolidated research group MATHMODE [IT1294-19].

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Lustig M, Donoho DL, Santos JM, Pauly JM (2008) Compressed sensing MRI. *IEEE Signal Process Mag* 25(2):72–82
- Hennig J, Nauerth A, Friedburg H (1986) RARE imaging: a fast imaging method for clinical MR. *Magnetic Resonance in Medicine* 3(6):823–833
- Stehling MK, Turner R, Mansfield P (1991) Echo-planar imaging: magnetic resonance imaging in a fraction of a second. *Science* 254(5028):43–50
- Sodickson Daniel K, Manning Warren J (1997) Simultaneous acquisition of spatial harmonics (SMASH): fast imaging with radio frequency coil arrays. *Magnetic Resonance in Medicine* 38(4):591–603
- Pruessmann Klaas P, Weiger Markus, Scheidegger Markus B, Boesiger Peter (1999) SENSE: sensitivity encoding for fast MRI. *Magnetic Resonance in Medicine* 42(5):952–962
- Griswold Mark A, Jakob Peter M, Heidemann Robin M, Nittka Mathias, Jellus Vladimir, Wang Jianmin, Kiefer Berthold, Haase Axel (2002) Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magnetic Resonance in Medicine* 47(6):1202–1210
- Donoho L, Compressed D (2006) sensing. *IEEE Trans Inform Theory* 52(4):1289–1306
- Ye JC (2019) Compressed sensing MRI: a review from signal processing perspective. *BMC Biomed Eng* 1:8
- Block KT, Uecker M, Frahm J (2007) Undersampled radial MRI with multiple coils. iterative image reconstruction using a total variation constraint. *Magnetic Resonance in Medicine* 57(6):1086–1098
- Beladgham M, Boucli Hacene I, TalebAhmed A, Khelif M (2008) MRI images compression using curvelets transforms. *AIP Conference Proceedings* 1019(1):249–253
- Zhu Zangen, Wahid Khan, Babyn Paul, Yang Ran, Wong Koon-Pong (2013) Compressed sensing-based MRI reconstruction using complex double-density dual-tree DWT. *Int J Biomed Imag* 2013:907501
- Ravishankar S, Bresler Y (2011) MR image reconstruction from highly undersampled k-space data by dictionary learning. *IEEE Trans Med Imag* 30(5):1028–1041
- Aelterman J, Luong HQ, Goossens B, Pižurica A, Philips W (2010) COMPASS: a joint framework for parallel imaging and compressive sensing in MRI. In: *Proceedings - International conference on image processing, ICIP*, pp 1653–1656
- Trzasko Joshua D, Haider Clifton R, Borisch Eric A, Campeau Norbert G, Glockner James F, Riederer Stephen J, Manduca Armando (2011) Sparse-CAPR: highly accelerated 4D CE-MRA with parallel imaging and nonconvex compressive sensing. *Magnetic Resonance in Medicine* 66(4):1019–1032
- Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*
- Long J, Shelhamer E, Darrell T (2015) Fully convolutional networks for semantic segmentation. In: *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*
- Girshick R, Donahue J, Darrell T, Malik J (2014) Rich feature hierarchies for accurate object detection and semantic segmentation. In: *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*
- Dong C, Loy C, He K, Tang X (2016) Image super-resolution using deep convolutional networks. *IEEE Trans Pattern Anal Mach Intell* 38(2):295–307
- Wang S, Su Z, Ying L, Peng X, Zhu S, Liang F, Feng D, Liang D (2016) Accelerating magnetic resonance imaging via deep learning. In: *2016 IEEE 13th international symposium on bio-medical imaging (ISBI)*, pp 514–517
- Yang Y, Sun J, Li H, Xu Z (2016) Deep ADMM-Net for compressive sensing MRI. In: *Proceedings of the 30th international conference on neural information processing systems*, p 1018
- Schlemper J, Caballero J, Hajnal VJ, Price NA, Rueckert D (2018) A deep cascade of convolutional neural networks for dynamic MR image reconstruction. *IEEE Trans Med Imag* 37(2):491–503. *IEEE Transactions on Medical Imaging*
- Zhu Bo, Liu Jeremiah Z, Cauley Stephen F, Rosen Bruce R, Rosen Matthew S (2018) Image reconstruction by domain-transform manifold learning. *Nature* 555(7697):487–492
- Dar SUH, zbey M, iatl AB, ukur T (2020) A transfer-learning approach for accelerated MRI using deep neural networks. *Magnetic Resonance in Medicine* 84(2):685
- Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y (2014) Generative adversarial networks. *arXiv e-prints*, page arXiv:1406.2661
- Arjovsky M, Chintala S, Bottou L (2017) Wasserstein generative adversarial networks. In: *Proceedings of the 34th international conference on machine learning*, vol 70 of *Proceedings of machine learning research*, PMLR, 06–11, pp 214–223
- Radford A, Metz L, Chintala S (2015) Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv e-prints*, page arXiv:1511.06434
- Zhu J-Y, Park T, Isola P, Efros AA (2017) Unpaired image-to-image translation using cycle-consistent adversarial networks. In: *Proceedings of the IEEE international conference on computer vision (ICCV)*



28. Choi Y, Choi M, Kim M, Ha J-W, Kim S, Choo J (2018) StarGAN: unified generative adversarial networks for multi-domain image-to-image translation. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)
29. Xia W, Yang Y, Xue J-H, Wu B (2021) TediGAN: text-guided diverse face image generation and manipulation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (CVPR), pp 2256–2265
30. Pizzati F, Cerri P, de Charette R (2021) CoMoGAN: continuous model-guided image-to-image translation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (CVPR), pp 14288–14298
31. Ledig C, Theis L, Huszar F, Caballero J, Cunningham A, Acosta A, Aitken A, Tejani A, Totz J, Wang Z, Shi W (2017) Photo-realistic single image super-resolution using a generative adversarial network. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)
32. Chen T, Zhai X, Ritter M, Lucic M, Houlsby N (2019) Self-supervised GANs via auxiliary rotation loss. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (CVPR)
33. Jiang Y, Chang S, Wang Z (2021) TransGAN: two transformers can make one strong GAN. *arXiv e-prints*, page arXiv:2102.07074
34. Yang G, Yu S, Slabaugh G, Dragotti PL, Ye X, Liu F, Arridge S, Keegan J, Guo Y, Firmin D (2018) DAGAN: deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction. *IEEE Trans Med Imag* 37(6):1310–1321
35. Shaul R, David I, Shitrit O, Raviv TR (2020) Subsampled brain MRI reconstruction by generative adversarial neural networks. *Med Image Anal* 65:101747
36. Quan MT, Nguyen-Duc T, Jeong KW (2018) Compressed sensing MRI reconstruction using a generative adversarial network with a cyclic loss. *IEEE Tran Med Imag* 37(6):1488–1497
37. (2021) Which GAN? a comparative study of generative adversarial network-based fast MRI reconstruction. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 379(2200):20200203
38. Ronneberger O, Fischer P, Brox T (2015) U-Net: convolutional networks for biomedical image segmentation. In: Medical image computing and computer-assisted intervention – MICCAI 2015, Cham, Springer International Publishing, pp 234–241
39. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. *arXiv e-prints*, page arXiv:1409.1556
40. Lv Jun, Wang Chengyan, Yang Guang (2021) PIC-GAN: a parallel imaging coupled generative adversarial network for accelerated multi-channel MRI reconstruction. *Diagnostics* 11(1):61
41. Tavaf N, Torfi A, Ugurbil K, Van de Moortele P-F (2021) GRAPPA-GANs for parallel MRI reconstruction. *arXiv e-prints*, page arXiv:2101.03135
42. Guo Yifeng, Wang Chengjia, Zhang Heye, Yang Guang (2020) Deep attentive wasserstein generative adversarial networks for MRI reconstruction with recurrent context-awareness. In: Medical Image Computing and Computer Assisted Intervention - MICCAI 2020:167–177
43. Hu X (2021) Multi-texture GAN: exploring the multi-scale texture translation for brain MR images. *arXiv e-prints*, page arXiv:2102.07225
44. Ma Y, Liu J, Liu Y, Fu H, Hu Y, Cheng J, Qi H, Wu Y, Zhang J, Zhao Y (2021) Structure and illumination constrained GAN for medical image enhancement. *IEEE Transactions on Medical Imaging*, pp 1–1
45. Zhang H, Xia L, Song R, Yang J, Hao H, Liu J, Zhao Y (2020) Cerebrovascular segmentation in MRA via reverse edge attention network. In: Medical image computing and computer assisted intervention – MICCAI 2020, Cham, Springer International Publishing, pp 66–75
46. Yuan Z, Jiang M, Wang Y, Wei B, Li Y, Wang P, Menpes-Smith W, Niu Z, Yang G (2020) SARA-GAN: self-attention and relative average discriminator based generative adversarial networks for fast compressed sensing MRI reconstruction. *Frontiers in Neuroinformatics*, 14(November)
47. Li Guangyuan, Lv Jun, Wang Chengyan (2021) A modified generative adversarial network using spatial and channel-wise attention for CS-MRI reconstruction. *IEEE Access* 9:83185–83198
48. Li Guangyuan, Lv Jun, Tong Xiangrong, Wang Chengyan, Yang Guang (2021) High-resolution pelvic MRI reconstruction using a generative adversarial network with attention and cyclic loss. *IEEE Access* 9:105951–105964
49. Chen Y, Firmin D, Yang G (2021) Wavelet improved GAN for MRI reconstruction. In: Medical imaging 2021: Physics of medical imaging, vol 11595. SPIE, pp 285 – 295
50. Lv Jun, Li Guangyuan, Tong Xiangrong, Chen Weibo, Huang Jiahao, Wang Chengyan, Yang Guang (2021) Transfer learning enhanced generative adversarial networks for multi-channel MRI reconstruction. *Computers in Biology and Medicine* 134:104504
51. Jiang Mingfeng, Zhi Minghao, Wei Liying, Yang Xiaocheng, Zhang Jucheng, Li Yongming, Wang Pin, Huang Jiahao, Yang Guang (2021) FA-GAN: fused attentive generative adversarial networks for MRI image super-resolution. *Computerized Medical Imaging and Graphics* 92:101969
52. Zhou Wenzhong, Huiqian Du, Mei Wenbo, Fang Liping (2021) Efficient structurally-strengthened generative adversarial network for MRI reconstruction. *Neurocomputing* 422:51–61
53. Zhou Wenzhong, Huiqian Du, Mei Wenbo, Fang Liping (2021) Spatial orthogonal attention generative adversarial network for MRI reconstruction. *Med Phys* 48(2):627–639
54. Biting Yu, Zhou Luping, Wang Lei, Shi Yinghuan, Fripp Jürgen, Bourgeat Pierrick (2019) Ea-GANs: edge-aware generative adversarial networks for cross-modality MR image synthesis. *IEEE Trans Med Imag* 38(7):1750–1762
55. Chai Yaqiong, Botian Xu, Zhang Kangning, Lepore Natasha, Wood John C (2020) MRI restoration using edge-guided adversarial learning. *IEEE Access* 8:83858–83870
56. Li Yixuan, Li Jie, Ma Fengfei, Shuangli Du, Liu Yiguang (2021) High quality and fast compressed sensing MRI reconstruction via edge-enhanced dual discriminator generative adversarial network. *Magnetic Resonance Imaging* 77:124–136
57. Wang Hao, Yang Yan, Liu Bing, Fujita Hamido (2019) A study of graph-based system for multi-view clustering. *Knowledge-Based Systems* 163:1009–1019
58. Zhang Yiling, Yang Yan, Li Tianrui, Fujita Hamido (2019) A multitask multiview clustering algorithm in heterogeneous situations based on LLE and LE. *Knowledge-Based Systems* 163:776–786
59. Chao Guoqing, Sun Shiliang, Bi Jinbo (2021) A survey on multi-view clustering. *IEEE Trans Artif Intell* 2(2):146–168
60. Zhao Jing, Xie Xijiong, Xin Xu, Sun Shiliang (2017) Multi-view learning overview: recent progress and new challenges. *Information Fusion* 38:43–54
61. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)
62. Qin Z, Zhang P, Wu F, Li X (2021) FcaNet: frequency channel attention networks. In: Proceedings of the IEEE/CVF international conference on computer vision (ICCV), pp 783–792

63. Hu J, Shen L, Sun G (2018) Squeeze-and-excitation networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)
64. (2018) An open, multi-vendor, multi-field-strength brain MR dataset and analysis of publicly available skull stripping methods agreement. *NeuroImage* 170:482–494
65. Mendrik AM, Vincken KL, Kuijf HJ, Breeuwer M, Bouvy WH, De Bresser J, Alansary A, De Bruijne M, Carass A, El-Baz A, et al (2015) MRBrainS challenge: Online evaluation framework for brain image segmentation in 3T MRI scans. *Computational Intelligence and Neuroscience*
66. (2017) GANs trained by a two time-scale update rule converge to a local nash equilibrium. *Advances in Neural Information Processing Systems 2017-Decem(Nips)*:6627–6638
67. Yang Guang, Zhuang Xiahai, Khan Habib, Haldar Shouvik, Nyktari Eva, Li Lei, Wage Ricardo, Ye Xujiang, Slabaugh Greg, Mohiaddin Raad, Wong Tom, Keegan Jennifer, Firmin David (2018) Fully automatic segmentation and objective assessment of atrial scars for long-standing persistent atrial fibrillation patients using late gadolinium-enhanced MRI. *Med Phys* 45(4):1562–1576
68. Uecker Martin, Lai Peng, Murphy Mark J, Virtue Patrick, Elad Michael, Pauly John M, Vasanawala Shreyas S, Lustig Michael (2014) ESPIRiT - an eigenvalue approach to autocalibrating parallel MRI: where SENSE meets GRAPPA. *Magnetic Resonance in Medicine* 71(3):990–1001