Chapter 13
Generative Adversarial Network Powered Fast Magnetic Resonance Imaging—Comparative Study and New Perspectives

Guang Yang, Jun Lv, Yutong Chen, Jiahao Huang, and Jin Zhu

Abstract Magnetic Resonance Imaging (MRI) is a vital component of medical imaging. When compared to other image modalities, it has advantages such as the absence of radiation, superior soft tissue contrast, and complementary multiple sequence information. However, one drawback of MRI is its comparatively slow scanning and reconstruction compared to other image modalities, limiting its usage in some clinical applications when imaging time is critical. Traditional compressive sensing based MRI (CS-MRI) reconstruction can speed up MRI acquisition, but suffers from a long iterative process and noise-induced artefacts. Recently, Deep Neural Networks (DNNs) have been used in sparse MRI reconstruction models to recreate relatively high-quality images from heavily undersampled k-space data, allowing for much faster MRI scanning. However, there are still some hurdles to tackle. For example, directly training DNNs based on L1/L2 distance to the target fully sampled images could result in blurry reconstruction because L1/L2 loss can only enforce overall image or patch similarity and does not take into account local information such as anatomical sharpness. It is also hard to preserve fine image details while

G. Yang (✉) · Y. Chen
National Heart & Lung Institute, Imperial College London, London SW7 2AZ, UK
e-mail: g.yang@imperial.ac.uk

G. Yang
Cardiovascular Research Centre, Royal Brompton Hospital, London SW3 6NP, UK

J. Lv
School of Computer and Control Engineering, Yantai University, Yantai, China

Y. Chen
Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge CB2 3EG, UK

J. Huang
School of Optics and Photonics, Beijing Institute of Technology, Beijing, China

J. Zhu
Department of Computer Science and Technology, University of Cambridge, Cambridge CB3 0FD, UK

maintaining a natural appearance. More recently, Generative Adversarial Networks (GAN) based methods are proposed to solve fast MRI with enhanced image perceptual quality. The encoder obtains a latent space for the undersampling image, and the image is reconstructed by the decoder using the GAN loss. In this chapter, we review the GAN powered fast MRI methods with a comparative study on various anatomical datasets to demonstrate the generalisability and robustness of this kind of fast MRI while providing future perspectives.

Keywords Generative adversarial networks (gan) · Fast magnetic resonance imaging (mri) · Compressive sensing · Deep learning

13.1 Introduction

13.1.1 Magnetic Resonance Imaging

Through offering repeatable, non-invasive measures of tissue structure and function, Magnetic resonance imaging (MRI) has transformed clinical medical imaging and medical science. The sensitivity of the image to tissue properties can be greatly varied with MRI, either by changing the timing with which MR signals are obtained (e.g., the echo time—TE and the repetition time—TR), or by using magnetisation preparation or contrast agents. These so-called multimodal or multi-sequence MRI methods can not only provide a comparison in traditional anatomical or structural MRI but can also quantify the function of most tissues and organs of the human body in clinical and pre-clinical laboratory environments. Invasive methods, such as tissue biopsy or radionuclide tests, have therefore become less needed as a result of the prosperity of MRI and other non-invasive medical imaging technologies [1]. Since MRI is non-invasive, it can be used to provide longitudinal and quantitative imaging biomarkers in the therapy trials. MR methods have begun to represent gold standard measurement for clinical research, despite the fact that the modality is still relatively underutilised [1]. Why is this so? 'MRI is complex and pricey', which is a commonly heard melancholy.

13.1.2 Limitations of Magnetic Resonance Imaging

Although MRI is a revolutionary non-invasive diagnostic imaging technique that provides high-resolution definition of the structural and functional information of most body tissues and organs, one significant drawback of MRI is its slow rate of image acquisition, which results in a longer scanning period as compared to other imaging modalities [2, 3].

In MRI, raw data is acquired in the *k*-space, which includes information about spatial frequencies within the image, rather than collected directly in the image space [4]. The Fourier transformation links image space and the *k*-space. The Nyquist criterion defines the *k*-space information that must be satisfied conventionally after we have defined the field-of-view (FOV) and spatial resolution of the image that we want to obtain. The distance between *k*-space neighbours is inversely proportional to the field of view in each direction. The highest frequency obtained in each direction is inversely proportional to the desired resolution. Data encoded with pulsed magnetic field gradients are acquired to fill the *k*-space. We may obtain a line of *k*-space points very quickly in one direction, known as the read direction, using either a spin or gradient echo in one repetition time. Further directions, on the other hand, must be phase-encoded, which takes one repetition time to encrypt one line of *k*-space lines [1, 4]. This must then be rerun for all possible combinations of the number of phase encoding steps needed in the anterior-posterior and foot-head directions. As a result, MRI acquisitions can be time-consuming, particularly when a high resolution or large FOV is needed.

This drawback not only raises the cost of imaging but also limits its use in emergency situations. Furthermore, in order to maintain image consistency, patients must lie still during the acquisition. For abdominal/thoracic imaging, patients must hold their breath, which can be problematic for paediatric, obese patients, and those with respiratory compromise [1]. As a result, many patients can experience anxiety and claustrophobia during the MR scanning procedure [1]. In order to minimise scanning costs and increase patient throughput, the MR image acquisition process must be sped up. Various MR acceleration methods rely on taking measurements of several lines per repetition time, allowing for quicker traversing of the *k*-space. Examples include echo planar imaging [5], rapid acquisition with relaxation enhancement [6], and fast low angle shot imaging [7].

13.1.3 Conventional Acceleration Using Compressive Sensing

It is possible to attain a higher degree of acceleration by sampling the *k*-space only partially, i.e., not collecting all lines of measurements in the phase encoding path(s). The undersampled calculation can be used to infer the original *k*-space details. As a consequence, the acceleration metric is equal to the undersampling ratio. For example, if half of the *k*-space is sampled, the acceleration factor is doubled. As a result, undersampling methods aim to circumvent the Nyquist-Shannon sampling criterion [2].

Compressed sensing (CS) is a promising undersampling approach that may allow for more aggressive undersampling and acceleration [2]. CS principle is similar to the concept of compressing signals for transmission and then decompressing them [8], as seen in the JPEG, JPEG2000, MPEG, and MP3 standards [1]. If undersampled

signals or images can be compressed correctly, they can also be decompressed or recovered accurately, according to CS [9]. Hereby, CS sets three conditions on the MRI reconstruction:

1. The image or signal must be compressible. In other words, the MRI images must be sparse, with the bulk of its pixel values being zeros, either in its native domain or in an appropriate transformation domain, such as in the wavelet or frequency domain.
2. To avoid aliasing artefacts, the undersampling patterns should be incoherent using random undersampling.
3. A non-linear reconstruction algorithm must be used.

It is possible to recover the original MRI images from their undersampled measurements following these three criteria.

Previously published research on using CS as an MR acceleration approach employs iterative non-linear optimisation algorithms that implement sparsity and reconstruction fidelity. Total variation (TV) [2], dictionary learning (DLMRI [10], RecPF [11], and BM3D [12]) are typical examples. However, there are four major issues with these approaches:

1. Iterative optimisation can be time-consuming [1, 13].
2. These algorithms tend to generate artificially smoothed image appearance [1].
3. The reconstruction results can have blocky artefacts [14–16].
4. They reconstruct each image as an individual event, failing to account for the expected anatomical features in MR images that may be used to improve the reconstruction accuracy [17].

13.1.4 Deep Learning Based Fast MRI

Deep learning based approaches have recently achieved performance dividends in a variety of medical image processing problems by using 'big data' and advancements in computational power. To date, however, the majority of research studies have concentrated on downstream medical image interpretation and post-processing activities, such as anatomical segmentation [18–28], lesion segmentation [29–34], co-registration [35–37], synthesis [38, 39], and multimodal data detection [40–46], for disease identification [47–49], prognosis [50, 51], and treatment prediction [52, 53]. To increase the precision of these post-processing operations, imaging methods must be improved, which can also be aided by deep learning [54–60].

Since its principle was developed in 2006, CS has had a long history for fast imaging applications, including the embodiment of MRI reconstruction [61]. However, the related less efficient iterative optimisation can stymie further implementation. Although deep learning based tomographic reconstruction technology has only been around for a few years, there is a lot of interest in this area, and there are many ongoing advances and exciting applications, including MRI.

Deep learning based approaches can successfully overcome the majority of the aforementioned shortcomings of earlier CS methods. A deep learning algorithm, e.g., convolutional neural networks (CNN), is made up of many layers of nodes. To learn the mapping from undersampled MR images to their corresponding fully sampled ones, the weights of the node relations between layers are optimised. The method of optimising weights is known as training the model. Once trained, the model is capable of reconstructing original images from undersampled measurements. In terms of reconstruction accuracy, speed, and visual consistency, deep learning based methods have been shown to consistently outperform non-deep learning based ones [3, 62–66].

13.1.5   GAN Powered Fast MRI

Generative Adversarial Networks, or GAN for short, represent a type of generative modelling technique that employs neural network and deep learning methods, e.g., CNN. Generative modelling is an unsupervised learning task in machine learning that entails automatically finding and learning the regularities or patterns in input data such that the model can be used to produce or output new examples that could have been plausibly drawn from the original dataset.

GAN employs a clever method for training a generative model by posing the problem as a supervised learning problem with two sub-models: the generator model, which we train to produce new examples, and the discriminator model, which attempts to identify examples as either true (from the original domain) or false (generated). The two models are trained in an adversarial zero-sum game before the discriminator model is tricked about half of the time, indicating that the generator model is producing plausible instances.

GAN is a fascinating and quickly evolving area that delivers on the promise of generative models by producing plausible instances in a variety of problems, most notably in image-to-image conversion tasks such as translating image styles, and in generating photo-realistic images of objects, scenes, and individuals that even humans cannot recognise which ones are fake.

GAN is an important type of deep learning based CS-MRI reconstruction method, which was proposed first by Yang et al. in [3, 67]. In the context of CS-MRI, GAN entails training a generator to recreate the original image from undersampled measurements and a discriminator to produce the likelihood of whether the generated image matches the original, i.e., fully sampled measurements. The discriminator, in turn, modifies the generator's learning [68]. As a consequence, the generator generates photo-realistic images [69]. In terms of reconstruction accuracy and efficiency, GAN based methods [3, 62] outperform the non-GAN based deep learning method, e.g., deep ADMM-net. One GAN based approach [63] also claims to generate less fuzzy and aliasing artefacts than non-deep learning based methods. As a result, GAN based approaches have the capability to produce state-of-the-art CS-MRI reconstruction results.

In this book chapter,

- we will perform a mini topical review on GAN powered fast MRI, including the original Deep De-Aliasing Generative Adversarial Networks (DAGAN) method and other more advanced and recently proposed GAN based models;
- we will analyse and explain different GAN models, and compare the results obtained by different GAN based models;
- we will provide a comparison study on different datasets, e.g., MRI for various anatomical scans.
- we will highlight the recent development and discuss future directions.

13.2 Methods

13.2.1 Fundamentals of MRI Reconstruction

The basis of undersampled MR reconstruction is inferring the missing k-space data from the already sampled k-space values. For example, in partial Fourier imaging (PFI), 50% of the k-space is acquired. Because the k-space is conjugate symmetric, the other missing 50% is obtained via complex conjugation of the existing 50%. However, the maximum acceleration is only 2-fold and the reconstructed image exhibits a lower signal-to-noise ratio (SNR) [1]. Another key undersampling technique is parallel imaging (PI). In PI, multiple receiver coils simultaneously collect the k-space information of the tissues closest to each coil. However, the acceleration factor is limited by the number and the configuration of the receiver coils [1, 70]. Therefore, limits of accelerating MR image acquisition exist for both PI and PFI, bound by the Nyquist-Shannon sampling criteria [2].

Compressed sensing (CS)-based reconstruction circumvents the Nyquist-Shannon sampling criteria and achieves a higher acceleration ratio. CS establishes the model of MR image acquisition as (notations in Table 13.1):

$$y_u = Ax_t, \tag{13.1}$$

where y_u is the undersampled k-space signal and x_t is the original fully sampled image. A is an operator defined as:

$$A = \Psi \mathcal{F}, \tag{13.2}$$

where Ψ is the undersampling mask, a binary matrix denoting which k-space locations are sampled and which are not, and $\mathcal{F}$ is the Fourier transform operator. Hence, $\Psi \mathcal{F} x_t$, or Ax_t is equivalent to undersampling the k-space of the putative reconstructed MR image.

Table 13.1 Mathematical notations in this book chapter

Symbol	Definition
x_t	Ground truth MR image
x_u	Undersampled MR image
$\hat{x}_u$	Reconstructed MR image
y_t	Ground truth k-space signals
y_u	Undersampled k-space signals
$\mathcal{F}$	Fourier transform operator
Ψ	Binary matrix denoting undersampled k-space positions
A	$\Psi\mathcal{F}$
R	Regularisation function
λ	A parameter adjusting the contribution of data fidelity term to the reconstruction loss function
θ	Parameters within a convolutional neural network

A CS-model infers the underlying full resolution MR image x_t from the collected incomplete k-space samples y_u by solving the following optimisation problem [2]:

$$\arg\min_{\hat{x}_u} \frac{\lambda}{2}||y_u - A\hat{x}_u||_2^2 + R(\hat{x}_u), \tag{13.3}$$

where λ adjusts the contribution of the first term to the optimisation objective and $R(\hat{x}_u)$ is an image regulariser function. If this reconstructed image $\hat{x}_u$ is to match the original image, their undersampled k-space results ought to match, as reflected in the first term of Eq. 13.3. This reinforces the k-space data fidelity of reconstructing the undersampled data.

The second term in Eq. 13.3 ensures the reconstructed image possesses certain attributes such as smoothness or sparsity, which is required by the theory of CS. For example, total variation—an early CS technique—uses the following regulariser term to ensure the underlying image is smooth [2, 71]:

$$\arg\min_{\hat{x}_u} \frac{\lambda}{2}||y_u - A\hat{x}_u||_2^2 + ||\nabla\hat{x}_u||_1, \tag{13.4}$$

where ∇ is the gradient operator to minimise the difference between adjacent pixels in the final reconstruction to ensure a smooth texture. To solve Eq. 13.4, conjugate gradient descent is employed [2], by updating the reconstructed image with the gradient of Eq. 13.4 iteratively until convergence.

Another example of CS algorithm is dictionary learning, in which the regulariser function enforces the sparsity of the dictionary representation of each image patch [10]:

$$\underset{\hat{x}_u, D, \Gamma}{\arg\min} \frac{\lambda}{2} ||y_u - A\hat{x}_u||_2^2 + \sum_{ij} ||R_{ij}\hat{x}_u - Da_{ij}||_2^2, \quad (13.5)$$

where R_{ij} is the image patch extractor, D is the dictionary used to transform the dictionary representation a_{ij} into the image domain, and Γ is the set of all dictionary representations $\{a_{ij}\}_{ij}$. This optimisation problem in Eq. 13.5 is solved by alternating minimisation of the 3 parameter sets iteratively: $\hat{x}_u$, D and Γ collectively. CS algorithms such as TV and dictionary learning repeatedly update the reconstruction to ensure it fulfils k-space data fidelity and certain regularisation properties such as sparsity.

From a Bayesian perspective, data fidelity is seen as maximising the conditional probability of observing the undersampled k-space data given the reconstructed image. The regulariser term represents the prior expectation of the statistical properties of the final reconstructed image [72]. By jointly enforcing data fidelity and regulariser constraint, CS reconstruction is equivalent to maximising the posterior probability that the reconstructed image matches the original one given the existing k-space samples. This justifies the effectiveness of the CS in reconstructing undersampled MR images.

13.2.2 CNN Based MRI Reconstruction

CNN excels in computer vision applications compared to traditional machine learning methods in general, and has been increasingly incorporated into CS-based MRI reconstruction models. It achieves higher accuracy and acceleration of MRI acquisition compared to traditional CS models [3, 64, 66, 73]. A CNN consists of a series of convolutional layers connected by non-linear activation functions [74]. In each convolutional layer, multiple filters are convolved with the input image or the output from a previous layer to extract image specific features [74]. The weight of each filter is optimised such that the final reconstructed image matches the original fully sampled image. By stacking multiple convolutional layers, the large number of weight parameters in a CNN endows it with the potential to model complex functions, including recovering the fully sampled images from the undersampled ones.

The process of optimising the weights of CNN is known as training the network and is governed by an optimisation objective called the loss function. The loss functions differ between supervised and unsupervised settings, depending on whether fully sampled ground truth images are available. In a supervised learning setting, fully sampled images are used to "teach" the CNN model to minimise the loss, or difference, between the reconstructed results and the fully sampled ground truth. However, in an unsupervised environment, ground truth images are not available, meaning the loss function can be, for example, optimised based on the Deep Image Prior framework and uses a high-resolution reference MR image as the input of the CNN to induce the structural prior in the learning procedure [75]. Unsupervised

CNN-based CS-MRI models are beyond the scope of this chapter, which will instead focus on the supervised learning models.

Supervised CNN-based CS-MRI methods can be broadly divided into two categories: end-to-end and unrolled optimisation [76–78]. End-to-end methods model the inverse of the acquisition process (Eq. 13.1) by mapping the undersampled input to the reconstructed output directly, hence the name "end-to-end":

$$\hat{x}_u = f_{\text{CNN}}(y_u|\theta), \tag{13.6}$$

where f_{CNN} is the operation performed by the CNN and θ represents all the parameters within the CNN. The θ is optimised by minimising the difference between $\hat{x}_u$ and the ground truth x_t. Such end-to-end techniques are exemplified by U-Net [79] and generative adversarial network (GAN) [3, 62, 63].

In contrast, CNN-based unrolled optimisation methods perform iterative image update based upon the general CS reconstruction model (Eq. 13.3). Unlike non-CNN based CS techniques, where the term $R(\hat{x}_u)$ is an expert-designed regularisation function, the CNN-based unrolled methods apply CNN to learn the optimal way of regularising an image. This is exemplified by the deep cascade CNN (DC-CNN) network [64], whose regulariser term penalises the deviation of the reconstructed image from the CNN output:

$$\arg\min_{\hat{x}_u,\theta} \frac{\lambda}{2}||y_u - A\hat{x}_u||_2^2 + ||\hat{x}_u - f_{\text{CNN}}(y_u|\theta)||. \tag{13.7}$$

Another example is the variational network [17], which uses a Field-of-Expert regularisation function:

$$\arg\min_{\hat{x}_u,\theta} \frac{\lambda}{2}||y_u - A\hat{x}_u||_2^2 + \sum_i f_i(k_i\hat{x}_u), \tag{13.8}$$

where f_i is a learnable activation function and k_i represents a convolutional filter. From a Bayesian point of view, these CNN-based regularisers form the prior expectation of the reconstructed images by learning from the ground truth MR images.

While both end-to-end and unrolled methods can incorporate CNN to learn the image reconstruction or regularisation process, a key difference between them is that unrolled methods need to iteratively update the images during the reconstruction process. In contrast, end-to-end ones compute the output directly. Hence, it is more time consuming to train unrolled CNN-based models and to apply them in reconstruction, compared with end-to-end methods, with other comparisons summarised in Table 13.2. This suggests an advantage of developing end-to-end methods, such as GAN-based models, for CS-MRI reconstruction.

Table 13.2 Comparison between end-to-end and unrolled optimisation methods in CNN-based CS-MRI models [76–78]

Category	End-to-end	Unrolled
Reconstruction time	Short	Long
Data fidelity	No	Yes
Sample size	Larger	Smaller
Performance	Lower	Higher
Weight update	Yes	Yes
Image update	No	Yes
Parameter number	Larger	Smaller

13.2.3 GAN Based MRI Reconstruction

13.2.3.1 General GAN

Inspired by the two-player zero-sum game in game theory, GAN [68] consists of two players: a generator and a discriminator. Traditionally, the generator G captures the distribution of sample data, and uses noise z that follows a certain distribution to generate a sample $G(z)$ similar to the real training data x_t. The discriminator D is a binary classifier, which aims to distinguish fake data generated by G from the ground truth. If the sample comes from the real training data, D outputs a large probability, otherwise, D outputs a small probability. Ideally, the best D can be represented as $D(x_t) = 1$ and $D(G(z)) = 0$. In this way, the generator and the discriminator form a min-max game. The training process of GAN can be described as follows:

$$\min_{\theta_G} \max_{\theta_D} L(\theta_G, \theta_D) = \mathbb{E}_{\boldsymbol{x} \sim p_{\text{data}}(\boldsymbol{x})}[\log D_{\theta_D}(\boldsymbol{x})]$$
$$+ \mathbb{E}_{\boldsymbol{z} \sim p_z(\boldsymbol{z})}[\log(1 - D_{\theta_D}(G_{\theta_G}(\boldsymbol{z})))], \qquad (13.9)$$

where $p_{\text{data}}(\boldsymbol{x})$ is the distribution of the training dataset, and the $p_z(\boldsymbol{z})$ is the distribution of the latent variables.

Alternating gradient optimisation between G and D is used to train the GAN. In this process, both models try their best to optimise their networks to form a competitive confrontation until the two models reach a dynamic balance. Ideally, the final result is that the image generated by G is very similar to the real image, and it is difficult for the D network to distinguish between the real image and the image generated by G, i.e., $D(G(Z)) = 0.5$.
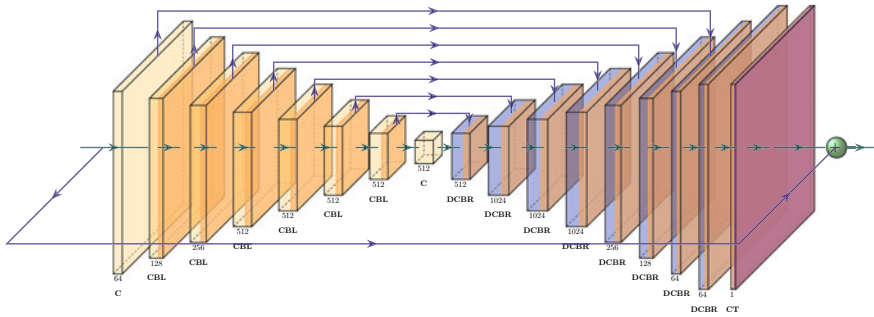
13.2.3.2   DAGAN

Deep De-Aliasing Generative Adversarial Networks (DAGAN) was proposed in 2017 by Yang et al. [3, 67] for fast compressed sensing MRI reconstruction. Figure 13.1 shows the architecture of DAGAN.

As shown in Fig. 13.2, a modified U-Net [80] was used as the generator G, which consisted of 8 convolutional layers in the encoding path and 8 deconvolutional layers in the decoding path. The stride of all convolutional and deconvolutional layers was set to 2 for downsampling and upsampling the feature maps. Each convolutional layer was followed by a Batch Normalisation (BN) layer and a Leaky ReLU (LReLU) layer. Skip connection was applied between corresponding layers in the encoding and decoding paths in order to pass the features in the encoding layer to the decoding layer for better reconstruction details. The modified U-Net ended up with a hyperbolic tangent function as the activation function. Due to the alternating training strategy in the adversarial components, the original GAN model is hard to train. DAGAN applied $\hat{x}_u = G(x_u) + x_u$ instead of $\hat{x}_u = G(x_u)$ as the output of the generator for better stability and faster convergence speed. In this way, DAGAN turned the generator into a refinement function, which means that it only generated the missing information, and the complexity of the model was reduced.

As shown in Fig. 13.3, an 11-layer CNN architecture was used as the discriminator D. A BN layer and a LReLU layer were followed with each convolutional layer. Finally, a full connection (FC) layer was cascaded, and the classification result was output through the Sigmoid activation function.
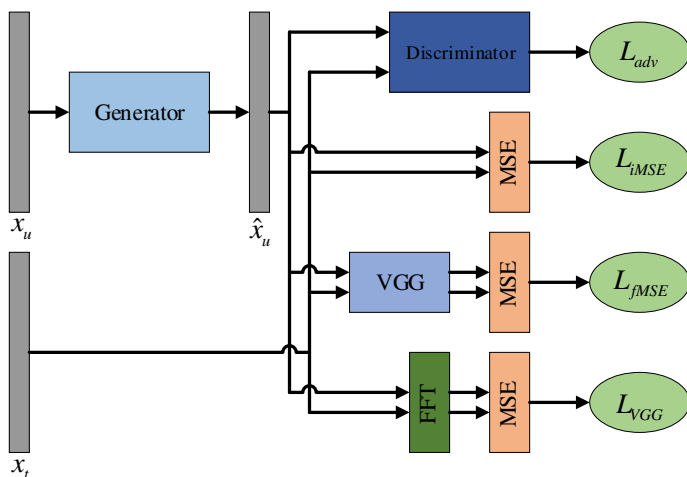
Fig. 13.1 The architecture of DAGAN. The generator produced the MR image $\hat{x}_u$ from the under-sampled MR image. The generated MR image $\hat{x}_u$ and the ground truth MR image were input into the discriminator for adversarial loss L_{adv} calculation. The data consistency loss consisted of image domain MSE loss L_{iMSE}, frequency domain MSE loss L_{fMSE}, and VGG perceptual loss L_{VGG}

Fig. 13.2 Structure of the generator in DAGAN. The generator was based on a modified U-Net, which consisted of 8 convolutional layers and 8 transposed convolutional layers (de-convolutional layers), with a BN layer and a LReLU or ReLu layer followed with each layer. Skip connection between layers that were the same scale and shortcut connections between the input and output of the generator were applied. A hyperbolic tangent function was used as an output activation function (C: Convolutional layer; DC: De-Convolutional layer; B: Batch Normalization layer; L: Leaky ReLU layer; R: ReLU layer; T: hyperbolic tangent function.)

Fig. 13.3 Structure of the discriminator in DAGAN. Essentially, the discriminator was an 11-layer CNN classifier. Each convolutional layer was cascaded by a BN layer and a LReLu Layer. A full connection layer and a Sigmoid function were applied to output the result of the classification (C: Convolutional layer; B: Batch Normalization layer; L: Leaky ReLU layer; F: Full Connection layer; S: Sigmoid function.)

In DAGAN, normalised MSE was used as the optimisation cost function. However, only optimising the pixel-wise MSE loss in the image domain could result in non-smooth reconstructions, which might lack coherent image details. To solve this problem, a data consistency loss was designed for training the generator in both frequency and image domains to help the optimisation and to exploit the complementary properties of the two domains. Besides, perceptual similarity [81] was also incorporated. In so doing, the data consistency loss consists of three parts: pixel-wise image domain MSE loss L_{iMSE}, frequency domain MSE loss L_{fMSE}, and the pre-trained VGG perceptual loss L_{VGG} [82]. They can be defined as:

$$L_{\text{iMSE}}(\theta_G) = \frac{1}{2} \parallel x_t - \hat{x}_u \parallel_2^2, \tag{13.10}$$

$$L_{\text{fMSE}}(\theta_G) = \frac{1}{2} \parallel \mathcal{F}x_t - \mathcal{F}\hat{x}_u \parallel_2^2, \tag{13.11}$$

$$L_{\text{VGG}}(\theta_G) = \frac{1}{2} \parallel f_{\text{VGG}}(x_t) - f_{\text{VGG}}(\hat{x}_u) \parallel_2^2. \tag{13.12}$$

Here, $\parallel \cdot \parallel_2^2$ denotes the L2 norm, $f_{\text{VGG}}(\cdot)$ denotes VGG network (the Conv4 output of the VGG16), and was pretrained on the ImageNet [83].

The adversarial loss L can be defined as:

$$\begin{aligned} L_{\text{adv}}(\theta_G, \theta_D) = &\ \mathbb{E}_{x_t \sim P_{\text{train}}(x_t)}[\log D_{\theta_D}(x_t)] \\ &+ \mathbb{E}_{x_u \sim p_G(x_u)}[\log(1 - D_{\theta_D}(G_{\theta_G}(x_u)))], \end{aligned} \tag{13.13}$$

where $P_{\text{train}}(x_t)$ denotes the collection of MR image ground truth, and $p_G(x_u)$ denotes the collection of MR image reconstruction.

Therefore the total loss function can be represented as:

$$\begin{aligned} L_{\text{TOTAL}}(\theta_G, \theta_D) = &\ \alpha L_{\text{iMSE}}(\theta_G) + \beta L_{\text{fMSE}}(\theta_G) + \gamma L_{\text{VGG}}(\theta_G) \\ &+ L_{\text{adv}}(\theta_G, \theta_D), \end{aligned} \tag{13.14}$$

where α, β and γ were the weights of different components of the loss function, which balanced different loss terms into similar scales according to previous study. Here, $\alpha = 15$, $\beta = 0.1$ and $\gamma = 0.0025$ were set empirically, according to the original paper.

13.2.3.3 KIGAN

KIGAN was introduced by Shaul et al. [84]. The overall structure of KIGAN is shown in Fig. 13.4. A k-space generator G_K and an image space generator G_{IM} were cascaded in KIGAN. Adjacent undersampled k-space slices (along the third dimension or the temporal dimension) y_u^{l-1}, y_u^l and y_u^{l+1} were the input of G_K. A data consistency step, i.e., $\widetilde{y}_u = \bar{\Psi} G_K(y_u^{l-1}, y_u^l, y_u^{l+1}) + y_u^l$ was applied for merging the output with the reconstructed missing data. G_{IM} was able to generate the output of G_K in image space, i.e., $\widetilde{x}_u$, to reconstructed image $\hat{x}_u$. Refinement learning $\hat{x}_u = G_{\text{IM}}(\widetilde{x}_u) + \widetilde{x}_u$ was also used as the output of G_{IM} instead of $\hat{x}_u = G_{\text{IM}}(\widetilde{x}_u)$. The reconstruction MR image $\hat{x}_u$, together with MR ground truth image x_t were sent to the discriminator for the adversarial loss.

As shown in Figs. 13.5 and 13.6, the main structure of the generators G_K and G_{IM} were based on a modified U-Net, which consisted of 5 convolutional layers in the encoding path and 5 deconvolutional layers in the decoding path, where each

Fig. 13.4 The Architecture of KIGAN. Two generators and one discriminator were used in KIGAN. k-space generator was able to reconstruct k-space MR data $G_K(y_u^l)$ from a cascade of undersampled k-space MR data y_u^{l-1}, y_u^l and y_u^{l+1}, and image space generator was able to reconstruct the MR image $\hat{x}_u$ (final result). The discriminator was to output the classification result of ground truth MR image x_t and generated MR image $\hat{x}_u$ for adversarial loss L_{loss}. Image domain MSE loss L_{iMSE} and frequency domain MSE loss were added into the total loss function

Fig. 13.5 The Structure of k-space Generator in KIGAN. Five convolutional layers and five deconvolutional layers were applied as encoding path and decoding path in the generator respectively. A convolutional layer was used as the output layer to adjust the channel of the reconstructed MR image. Corresponding layers of the same scales were linked by the skip connection (C: Convolutional layer; DC: De-Convolutional layer; B: Batch Normalization layer; L: Leaky ReLU layer; R: ReLU layer.)

Fig. 13.6 The Structure of image space Generator in KIGAN. 5 convolutional layers and 5 de-convolutional layers were applied as encoding path and decoding path in the generator respectively. A convolutional layer and a hyperbolic tangent function were used as the output layer. Skip connection and shortcut connection were adopted in the image space generator (C: Convolutional layer; DC: De-Convolutional layer; B: Batch Normalization layer; L: Leaky ReLU layer; R: ReLU layer; T: hyperbolic tangent function.)

convolutional and deconvolutional layer were followed by a BN layer and a LReLU layer. Skip connection was applied between corresponding layers in the encoding and decoding paths, in order to pass the feature of different scales to the decoding layer for better reconstruction details. Additionally, in G_{IM}, a shortcut connection between the input and output was applied for turning G_{IM} into a refinement function.

As shown in Fig. 13.7, the discriminator was a standard 9-layer CNN structure with an FC layer and a Sigmoid activation function connected for the result of the classification.

The loss function of KIGAN consisted of: an image domain MSE loss L_{iMSE}, a frequency domain MSE loss L_{fMSE}, and an adversarial loss L_{adv}, which can be defined as:

$$L_{\text{iMSE}}(\theta_{\text{K}}, \theta_{\text{IM}}) = \frac{1}{2} \| x_t - \hat{x}_u \|_2^2, \tag{13.15}$$

$$L_{\text{fMSE}}(\theta_{\text{K}}) = \frac{1}{2} \| y_t^l - \tilde{y}_u \|_2^2, \tag{13.16}$$

$$L_{\text{adv}}(\theta_{\text{K}}, \theta_{\text{IM}}, \theta_D) = \mathbb{E}_{x_t \sim M}[\log D_{\theta_D}(x_t)] + \mathbb{E}_{\hat{x}_u \sim S}[\log(1 - D_{\theta_D}(\hat{x}_u))], \tag{13.17}$$

where M denotes the collection of image space ground truth, and S denotes the collection of image space reconstruction.
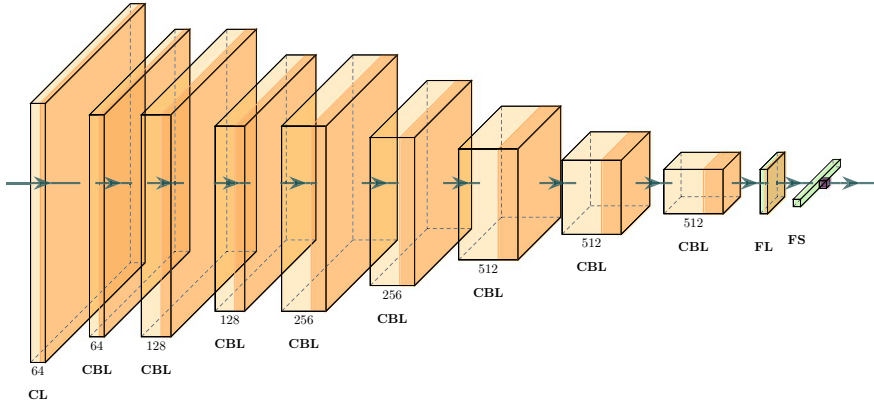
Fig. 13.7 The Structure of Discriminator in KIGAN. The discriminator was a standard 9-layer CNN structure with an FC layer and a Sigmoid activation function connected for the result of the classification (C: Convolutional layer; B: Batch Normalization layer; L: Leaky ReLU layer; F: Full Connection layer; S: Sigmoid function.)

The whole loss function can be represented as:

$$L_{\text{TOTAL}}(\theta_{\text{K}}, \theta_{\text{IM}}, \theta_D) = \alpha L_{\text{iMSE}}(\theta_{\text{K}}, \theta_{\text{IM}}) + \beta L_{\text{fMSE}}(\theta_{\text{K}})$$
$$+ L_{\text{adv}}(\theta_{\text{K}}, \theta_{\text{IM}}, \theta_D), \tag{13.18}$$

where α, β are the hyperparameters that control the balance of different components in the loss function.

13.2.3.4 ReconGAN/RefineGAN

ReconGAN/RefineGAN was proposed by Quan et al. [62] for compressed sensing MRI reconstruction. Figure 13.8 shows the architecture of ReconGAN/RefineGAN. The two-fold chained network that consisted of 2 U-shaped generators (G_1 and G_2 respectively) was able to generate a 2-channel zero-filling MR image (real part and imaginary part) directly into a 2-channel reconstructed MR image. The checkpoints after G_1 and G_2 were defined as ReconGAN and RefineGAN respectively.

As shown in Fig. 13.9, the generator consisted of 4 encoder blocks in the encoding path and 4 decoder blocks in the decoding path. Skip connection between corresponding blocks that had the same scale was adopted to pass the feature from the encoding path to the decoding path. $\bar{x}_u = G_1(x_u) + x_u$ and $\hat{x}_u = G_2(\bar{x}_u) + \bar{x}_u$ were used as the output of the generator instead of $\bar{x}_u = G_1(x_u)$ and $\hat{x}_u = G_2(\bar{x}_u)$ for better reconstruction details and faster convergence speed. Figures 13.10 and 13.11 show the structure of the encoder block and decoder block. The encoder block consisted of 2 convolutional layers and a residual block inserted between them. The stride of the first convolutional layer was set to 2 for the downsampling. The decoder block
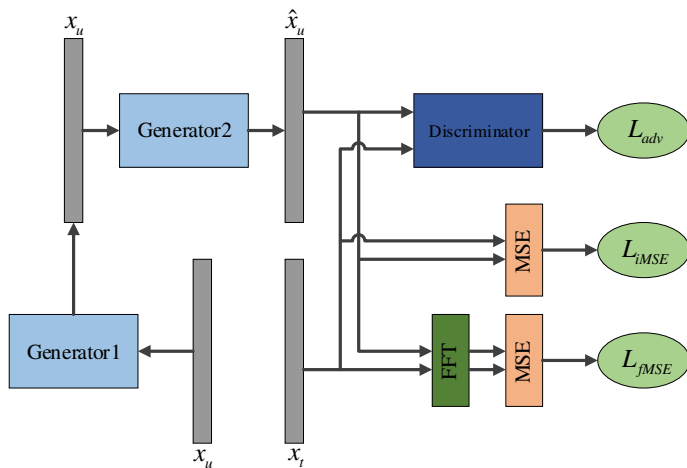
Fig. 13.8 The architecture of ReconGAN/RefineGAN. A two-fold chained network with 2 generators cascaded was used as the generator to produce reconstructed MR images $\hat{x}_u$ from undersampled MR images x_u. Reconstructed MR images $\hat{x}_u$ together with ground truth MR images x_t were sent to the discriminator for the calculation of the adversarial loss L_{adv}. Besides, image domain MSE loss L_{iMSE} and frequency domain MSE loss L_{fMSE} were added to the total loss function for training the generator
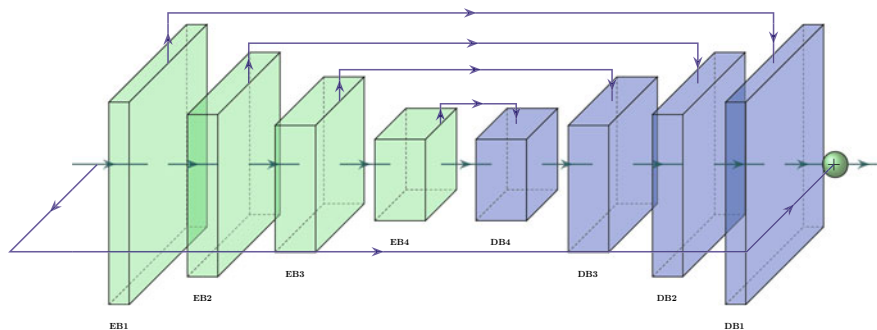
Fig. 13.9 Structure of the generator in ReconGAN/RefineGAN. The generator consisted of 4 encoder blocks in the encoding path and 4 decoder blocks in the decoding path. Skip connection was linked between blocks of the same scale in different paths. A shortcut connection was applied between the input and output of the generator, which turns the generator into a refinement function (EB: Encoder Block; DB: Decoder Block.)

Fig. 13.10 Structure of the encoder block in the generator of ReconGAN/RefineGAN (C: Convolutional layer; B: Batch Normalization layer; L: Leaky ReLU layer.)

CBL(s2)   CBL*2   C   CBL
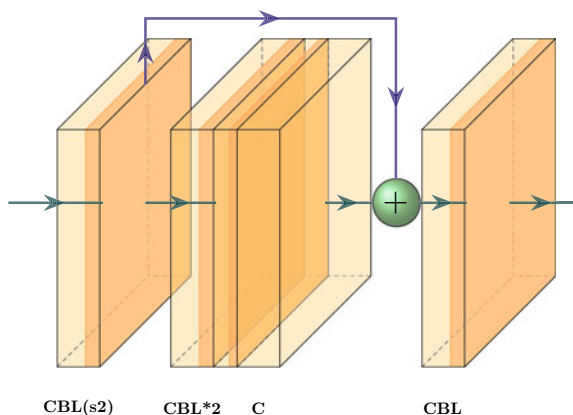
Fig. 13.11 Structure of the decoder block in the generator of ReconGAN/RefineGAN (C: Convolutional layer; DC: De-Convolutional layer; B: Batch Normalization layer; L: Leaky ReLU layer.)

DCBL   CBL*2   C

DCBL(s2)
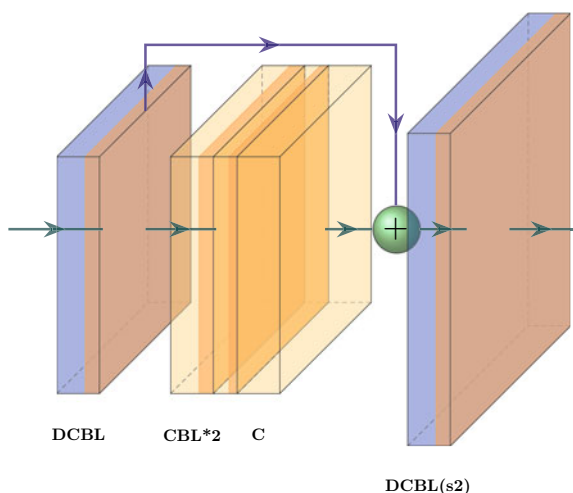
consisted of 2 transposed convolutional layers and a residual block inserted between them. The stride of the second transposed convolutional layer was set to 2 for upsampling. The residual block consisted of three convolutional layers, the kernel number in the second convolutional layer was half of the kernel number in the first and third convolutional layer. A shortcut connection was linked between the input and output of the residual block.

As shown in Fig. 13.12, the encoding path in the generator was adopted as the main structure of the discriminator. An FC layer and Sigmoid function were cascaded after the encoding path for the classification result.

The loss function of ReconGAN/RefineGAN was consisted of: image domain MSE loss L_{iMSE}, frequency domain MSE loss L_{fMSE}, and the adversarial loss L_{adv}, which can be defined as:

Fig. 13.12 Structure of the discriminator in ReconGAN/RefineGAN. The main structure adopted the encoding path of the generator. A full connection layer and a Sigmoid activation function were cascaded at the end of the discriminator (C: Convolutional layer; L: Leaky ReLU layer; F: Full Connection layer; S: Sigmoid function; EB: Encoder Block.)
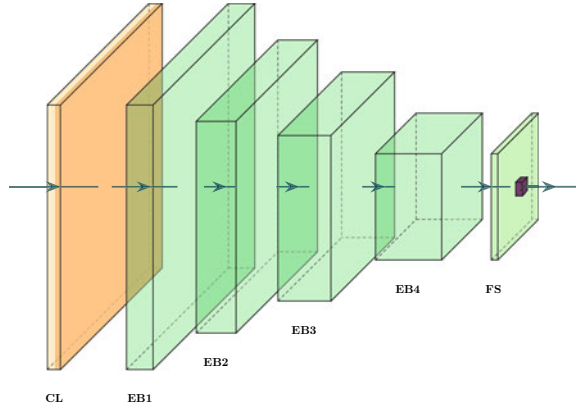
$$L_{\text{iMSE}}(\theta_{G_1}, \theta_{G_2}) = \frac{1}{2} \parallel x_t - \hat{x}_u \parallel_2^2, \tag{13.19}$$

$$L_{\text{fMSE}}(\theta_{G_1}, \theta_{G_2}) = \frac{1}{2} \parallel \mathcal{F}x_t - \mathcal{F}\hat{x}_u \parallel_2^2, \tag{13.20}$$

$$L_{\text{adv}}(\theta_{G_1}, \theta_{G_2}, \theta_D) = \mathbb{E}_{x_t \sim P_{train}(x_t)}[\log D_{\theta_D}(x_t)]$$
$$+ \mathbb{E}_{x_u \sim p_{G_2 G_1}(x_u)}[\log(1 - D_{\theta_D}(G_{\theta_{G_2}}(G_{\theta_{G_1}}(x_u))))]. \tag{13.21}$$

The total loss function can be represented as:

$$L_{\text{TOTAL}}(\theta_{G_1}, \theta_{G_2}, \theta_D) = \alpha L_{\text{iMSE}}(\theta_{G_1}, \theta_{G_2}) + \beta L_{\text{fMSE}}(\theta_{G_1}, \theta_{G_2})$$
$$+ L_{\text{adv}}(\theta_{G_1}, \theta_{G_2}, \theta_D), \tag{13.22}$$

where $\alpha = 10$, $\beta = 0.1$ were the hyperparameters that controlled the balance of different components in the loss function.

13.2.4 Evaluation Methods

Generally, the evaluation methods include objective methods for fidelity quality assessment and subjective methods for perceptual quality assessment. In this section, we review the most popular metrics for fast MRI quality evaluation.

13.2.4.1 Fidelity Quality Assessment

First, we introduce the Peak Signal-to-Noise Ratio (PSNR), which is the most commonly used evaluation criteria for image transformation tasks (e.g., reconstruction,

super-resolution, de-noising). It involves the data range to measure the pixel-level Mean Squared Error (MSE):

$$\text{PSNR}(I_{\text{rec}}, I_{\text{gt}}) = 10 \cdot \log_{10} \left(\frac{L^2}{\frac{1}{N} \sum_{i=1}^{N} (I_{\text{rec}}(i) - I_{\text{gt}}(i))^2} \right), \qquad (13.23)$$

where L denotes the data range (generally $L = 1.0$ in MRI reconstruction tasks), and N is the number of all the pixels in I_{rec} and I_{gt}. PSNR represents the pixelwise accuracy of the reconstruction regardless of the acquisition sequences of the multimodal MRI.

Besides, considering the importance of image structural information, such as brightness, contrast and structures, Structural SIMilarity index (SSIM) is formed as:

$$\text{SSIM}(x, y) = \frac{2\mu_x \mu_y + \kappa_1}{\mu_x^2 + \mu_y^2 + \kappa_1} \cdot \frac{\sigma_{xy} + \kappa_2}{\sigma_x^2 + \sigma_y^2 + \kappa_2}, \qquad (13.24)$$

where x, y denote two images, μ and σ^2 are the mean and variance, σ_{xy} is the covariance between x and y, and κ_1, κ_2 are constant relaxation terms.

13.2.4.2 Perceptual Quality Assessment

The perceptual quality of an image represents how realistic it looks. In MRI images reconstruction tasks, the most reliable perceptual quality assessment is the mean opinion score (MOS), which asks experienced radiologists to rate the reconstructed images. Typically, the images are rated from 0 to 4 depending on the reconstructed image quality (i.e., non-diagnostic, poor, fair, good, and excellent), and the final MOS is calculated as the arithmetic mean of the scores of all raters. In some cases, the rater may also mark the low perceptual quality features such as low SNR and motion artefacts. Although the MOS seems to be faithful, it has limitations such as inter-/inner-raters bias and variance of rating criteria and the scoring might be timeconsuming. Thus, the Frechet Inception Distance (FID) [85], as a learning based perceptual quality assessment, is becoming more commonly used for evaluation in GAN based image reconstruction tasks. It considers the high-level global features of a group of images (e.g., the reconstructed images) as a multidimensional Gaussian distribution $\mathcal{N}(\mu, \Sigma)$, and measures the differences between the two distributions of the reconstructed images $\mathbb{I}_{\text{rec}}$ and the ground truth images $\mathbb{I}_{\text{gt}}$. It first converts each group of images into a distribution of 2048 features in the latent space of a pre-trained image classification model Inception-V3 [86]. Then, the FID between these two distributions is calculated as:

$$\text{FID}(\mathbb{I}_{\text{rec}}, \mathbb{I}_{\text{gt}}) = \left\| \mu_{\text{gt}} - \mu_{\text{rec}} \right\|^2 + \text{Tr}(\Sigma_{\text{gt}} + \Sigma_{\text{rec}} - 2(\Sigma_{\text{gt}} \Sigma_{\text{rec}})^{1/2}). \qquad (13.25)$$

The FID becomes a popular metric for image perceptual quality assessment in GAN based image generation tasks because it is fully automatic and the features extracted from Inception-V3 are close to real-world object classification problems, which tend to mimic human perception similarity for images.

13.3 Benchmarking

In this chapter, we benchmark four GAN based algorithms, i.e., DAGAN, KIGAN, ReconGAN and RefineGAN, for fast MRI. Figure 13.13 shows the brain reconstruction results using different acceleration factors ($2\times$, $4\times$, $6\times$). It is obvious that the zero-filled (ZF) image has strong artefacts inside the brain tissue. From the entire image, the DAGAN effectively removes the artefacts in the ZF. However, in terms of the zoomed-in areas, there still exists some residual artefacts. For the reconstructions produced by KIGAN, blurring artefacts still exist. Although ReconGAN shows a significant reduction of aliasing artefacts, the edge details are not reconstructed clearly enough. It can be seen that the reconstructed details of RefineGAN are relatively fine, and the reconstruction quality is close to that of the ground truth.
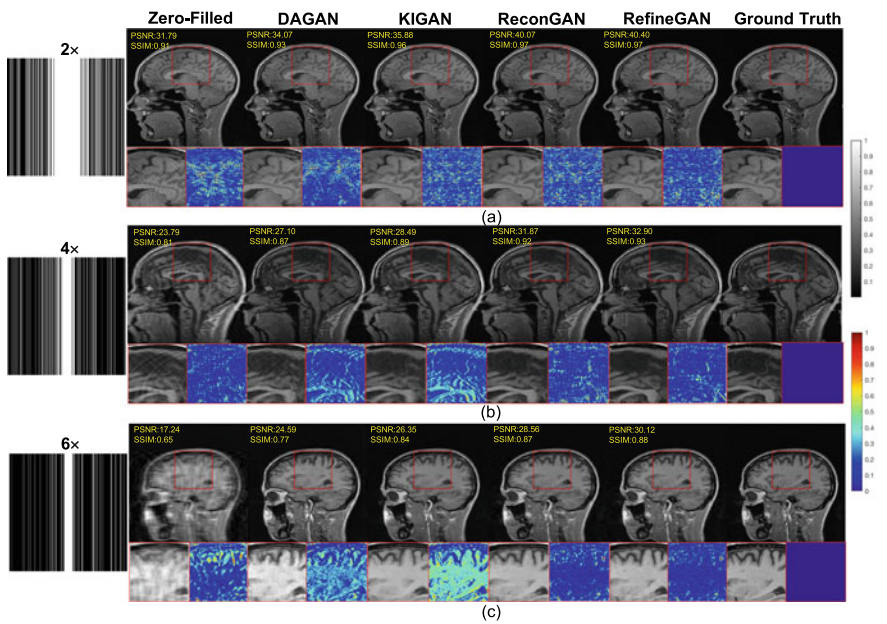
Fig. 13.13 Brain reconstruction result using the Cartesian mask. From top to bottom $2\times$, $4\times$ and $6\times$ acceleration, respectively. From left to right are Zero-Filled (ZF), DAGAN, KIGAN, ReconGAN, RefineGAN, Ground Truth (GT)
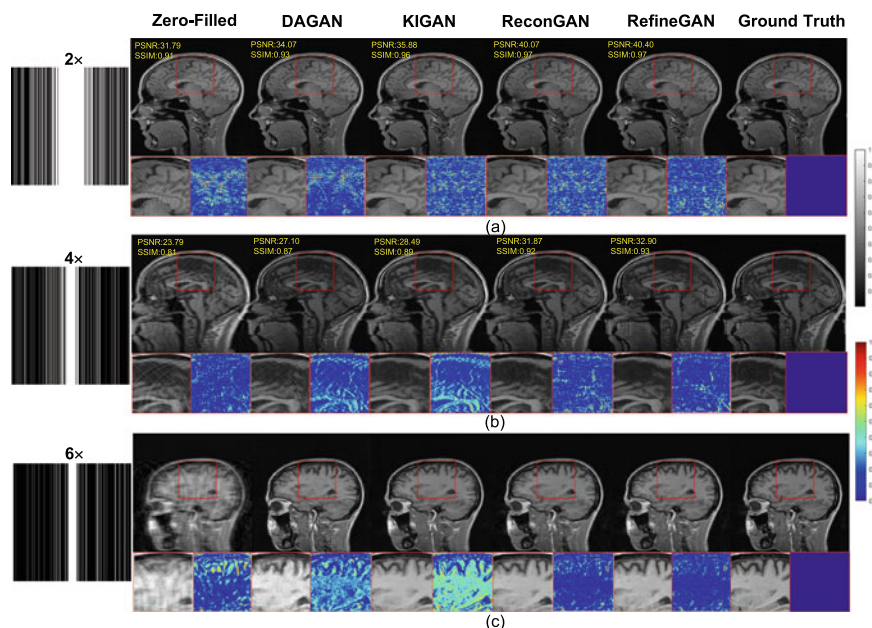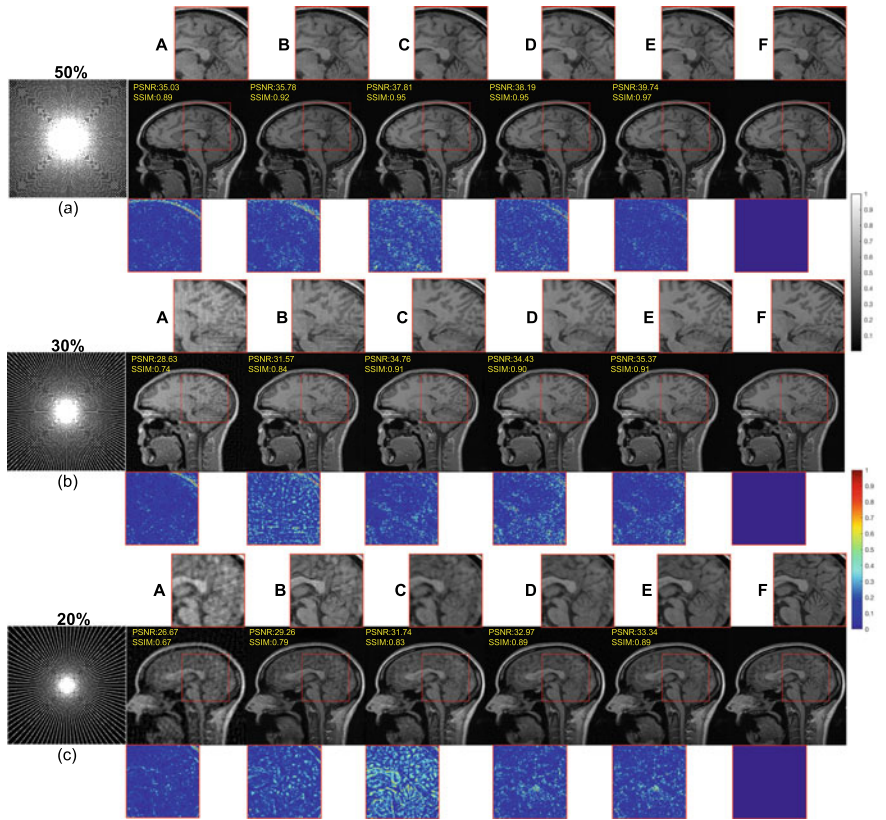
Fig. 13.14 Knee reconstruction result using the Cartesian mask. From top to bottom 2×, 4× and 6× acceleration, respectively. From left to right are Zero-Filled (ZF), DAGAN, KIGAN, ReconGAN, RefineGAN, Ground Truth (GT)

Besides, knee reconstruction results using different Cartesian masks are shown in Fig. 13.14. It can be seen that except for the ZF images, all methods can reconstruct acceptable MR images. As the acceleration factor goes high, obvious aliasing artefacts are produced in DAGAN images. As can be observed in the zoomed-in areas and the corresponding error maps, KIGAN can not restore clear vessels. ReconGAN and RefineGAN show better reconstruction results with higher PSNR and SSIM. In addition, the quantitative values of the RefineGAN are superior to those of the other methods.

Furthermore, we also used radial and spiral masks for training and testing each GAN based method. The sampling rate (SR) of each mask is 50, 30 and 20%. Figures 13.15 and 13.16 show the brain and knee reconstruction results using radial masks. We can see that the image reconstructed by the ZF method under the radial mask has strong blurring artefacts, and the details in the brain and knee cannot be distinguished clearly. When SR=20%, from the error maps, we can see that there are still obvious blurring artefacts and obscure blood vessels in the results of DAGAN and KIGAN. However, both ReconGAN and RefineGAN can restore sharper vessel edges and finer textures compared to other methods. Besides, RefineGAN has better PSNR and SSIM quantification.

The results are similar using spiral masks. Figures 13.17 and 13.18 show the brain and knee reconstruction results of each method using different spiral masks. Through
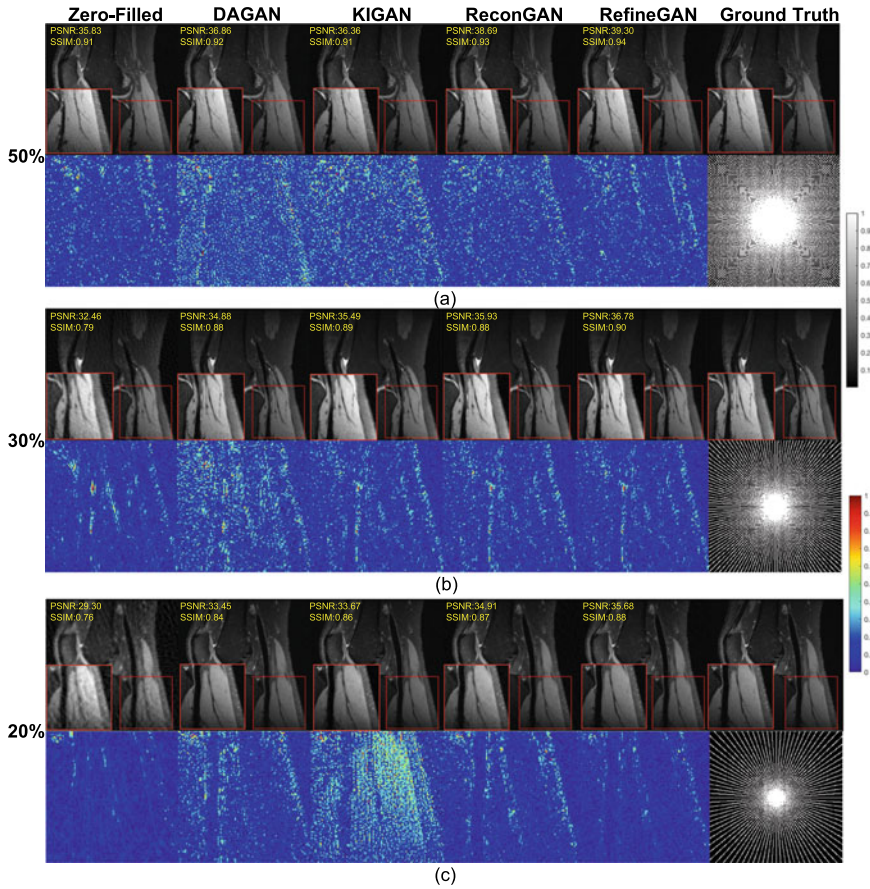
Fig. 13.15 Brain reconstruction result using the radial mask. From top to bottom, the sampling rate (SR) is 50, 30, and 20%, respectively. From left to right are Zero-Filled (ZF), DAGAN, KIGAN, ReconGAN, RefineGAN, Ground Truth (GT)

the error map, we can intuitively see that the reconstructed image of RefineGAN has fewer errors, and the reconstruction details are also better. The image reconstructed by RefineGAN can clearly show the details of the gray matter in the brain and the details of the blood vessels in the knee.

The quantitative metrics (PSNR, SSIM and RMSE) of each GAN based method using different under-sampling masks are shown in Tables 13.3 and 13.4. We can draw similar conclusions as the qualitative visualisation results that the image quality after RefineGAN reconstruction is better than other methods. Even at a high acceleration factor or high under-sampling rate, the reconstructed images of RefineGAN still have high SNR.

Fig. 13.16 Knee reconstruction result using the radial mask. From top to bottom 2×, 4× and 6× acceleration, respectively. From left to right are Zero-Filled (ZF), DAGAN, KIGAN, ReconGAN, RefineGAN, Ground Truth (GT)

Tables 13.3 and 13.4 show the quantitative metrics, including PSNR, SSIM and RMSE for all compared methods. The numbers in Tables 13.3 and 13.4 represent the mean and standard deviation values of the corresponding metrics (bold numbers indicate the best performance). Compared to DAGAN, KIGAN and ReconGAN, the RefineGAN framework has outperformed them remarkably at different acceleration factors.
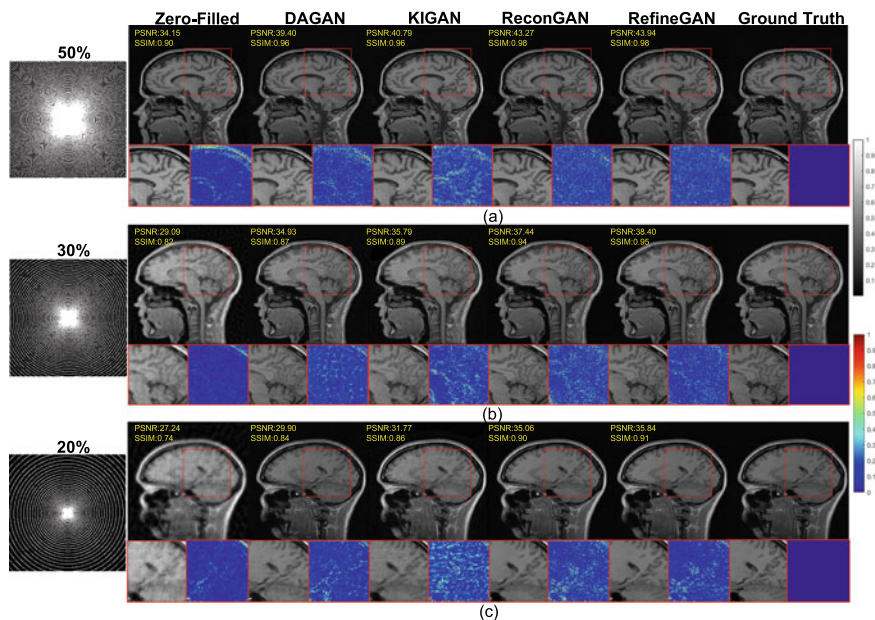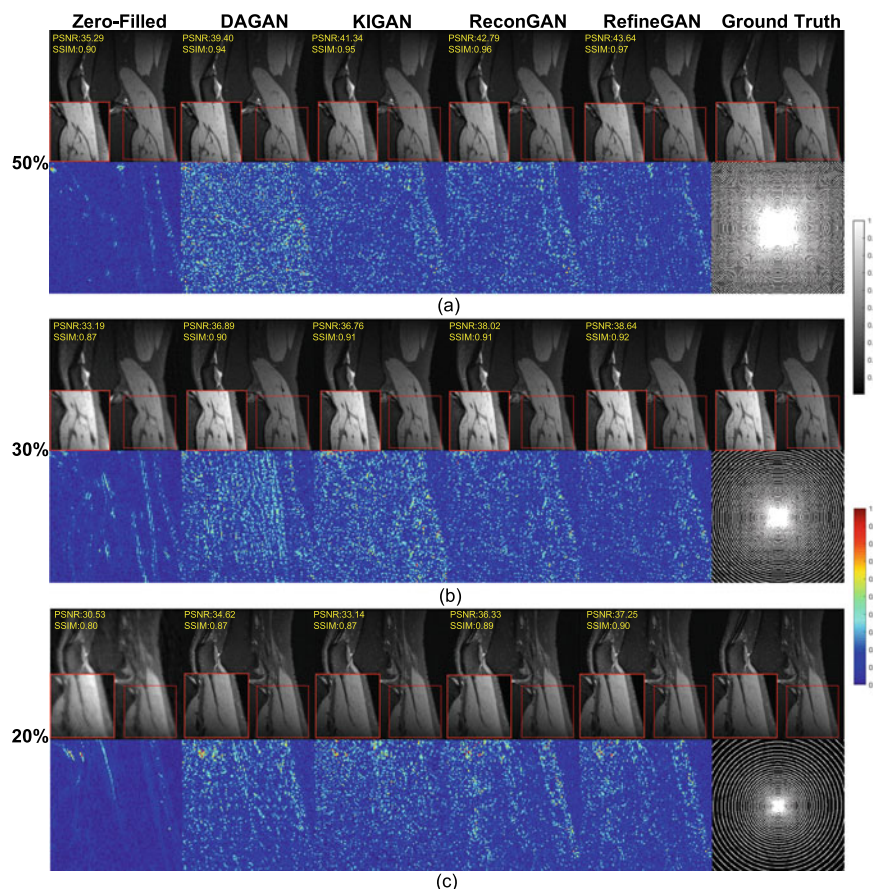
Fig. 13.17 Brain reconstruction result using the radial mask. From top to bottom, the sampling rate (SR) is 50, 30, and 20%, respectively. From left to right are Zero-Filled (ZF), DAGAN, KIGAN, ReconGAN, RefineGAN, Ground Truth (GT)

13.4   Discussion

We have established that GAN based methods such as DAGAN, KIGAN, Recon-GAN, and RefineGAN excel in generating faithful, photo-realistic reconstruction of undersampled MR images and in removing the undersampling artefacts. Despite their representing a successful category of CS-MRI techniques, GAN based methods suffer from training instability and slow convergence to a global minimum [87, 88]. This problem can be alleviated by Wasserstein GAN (WGAN) [89]. Instead of minimising Jensen-Shannon divergence between the reconstructed image and ground truth in the training phase, WGAN aims to reduce the Wasserstein distance. WGAN based CS-MRI models have achieved superior reconstruction performance compared with GAN based methods e.g., DAGAN [90, 91]. However, to minimise the Wasserstein distance, the discriminator needs to satisfy the 1-Lipschitz constraint. The original WGAN paper adopted a weight clipping approach to enforce this constraint but this can itself cause training instability. The potential improvement includes a gradient clipping approach [92] or dividing the weights of neural networks by their spectral norm [93]. Altogether, modifying the loss function of GAN is key to addressing its training instability issue.

Another problem with the GAN loss function is that it may affect the high frequency texture [63] and smoothing the reconstructed image [94]. GAN may also

Fig. 13.18 Knee reconstruction result using the radial mask. From top to bottom 2×, 4× and 6× acceleration, respectively. From left to right are Zero-Filled (ZF), DAGAN, KIGAN, ReconGAN, RefineGAN, Ground Truth (GT)

generate high frequency noise [63]. Mardani et al. (2019) suggested solving the high frequency texture issue with the least square GAN [63]. L1 loss between reconstructed images and ground truth was also added as an auxiliary function, to serve as a low-pass filter to remove high frequency noises.

Apart from the issues with the GAN loss function during training, GAN based methods suffer from reconstruction instability [95]. This means that a small perturbation to the input image leads to large scale changes in the reconstruction output. Antun et al. [95] evaluated the instabilities of DAGAN and other deep learning based CS-MRI models and showed that the performance of DAGAN deteriorated as the undersampling ratio increased during testing. That is if more *k*-space pixels were undersampled more than the undersampling ratio in the images used to train DAGAN, its performance decreases. In contrast, variational network and deep cas-

Table 13.3 The quantitative metrics (PSNR, SSIM and RMSE) of the brain MRI data using different GAN based methods. The bold numbers indicate the best results

Mask	AF/SR	Metric	ZF	DAGAN	KIGAN	ReconGAN	RefineGAN
Cartesian	2X	PSNR	30.94±2.75	33.79±1.88	33.90±2.55	39.08±1.34	39.40±1.33
		SSIM	0.92±0.02	0.93±0.01	0.96±0.01	**0.97±0.00**	0.97±0.00
		RMSE	1.57±0.01	0.72±0.43	0.78±0.53	0.20±0.09	0.19±0.08
	4X	PSNR	23.69±3.02	28.76±1.95	28.14±1.84	32.07±1.65	32.67±1.56
		SSIM	0.79±0.03	0.86±0.02	0.88±0.02	0.92±0.01	0.93±0.01
		RMSE	8.86±7.01	2.35±1.50	2.67±1.40	1.05±0.54	0.91±0.08
	6X	PSNR	19.47±2.31	25.4±1.57	27.91±1.57	29.23±1.68	29.95±1.61
		SSIM	0.66±0.04	0.77±0.03	0.86±0.02	0.88±0.02	0.89±0.02
		RMSE	21.3±14.1	4.89±2.33	2.75±1.31	2.03±1.00	1.71±0.83
Radial	50%	PSNR	34.28±1.03	35.24±1.14	38.93±1.10	37.89±0.93	39.38±0.88
		SSIM	0.87±0.02	0.91±0.01	0.96±0.01	0.95±0.01	0.97±0.00
		RMSE	1.94±0.22	1.74±0.22	1.01±0.12	1.28±0.13	0.88±0.11
	30%	PSNR	28.83±0.72	30.99±2.33	33.97±1.05	34.35±0.99	35.21±1.05
		SSIM	0.73±0.02	0.82±0.02	0.91±0.01	0.91±0.01	0.92±0.01
		RMSE	3.10±0.30	2.01±0.27	1.60±0.18	1.93±0.21	1.02±0.17
	20%	PSNR	26.47±0.64	28.96±1.90	31.87±0.94	31.96±0.87	32.83±0.94
		SSIM	0.65±0.02	0.78±0.02	0.84±0.03	0.87±0.02	0.89±0.02
		RMSE	6.08±0.36	2.47±0.45	2.56±0.27	2.54±0.25	1.56±0.32
Spiral	50%	PSNR	34.87±1.01	38.08±0.99	41.56±1.00	43.69±0.58	44.36±0.57
		SSIM	0.90±0.01	0.95±0.01	0.96±0.01	0.97±0.01	0.98±0.00
		RMSE	1.62±0.18	1.25±0.01	0.84±0.10	0.72±0.66	0.59±0.61
	30%	PSNR	29.55±0.71	34.87±2.79	36.71±1.09	37.93±0.78	38.61±0.82
		SSIM	0.83±0.01	0.87±0.01	0.91±0.02	0.95±0.01	0.95±0.00
		RMSE	3.93±0.28	1.72±0.41	1.65±0.20	1.27±0.11	0.94±0.19
	20%	PSNR	26.62±0.63	28.87±2.29	32.43±0.72	35.02±0.81	35.11±0.85
		SSIM	0.73±0.01	0.80±0.04	0.88±0.03	0.91±0.01	0.92±0.01
		RMSE	7.12±0.36	2.29±0.31	2.40±0.20	1.78±0.16	1.52±0.27

caded convolutional neural networks (DC-CNN), two non-GAN based techniques, did not experience the same issue. This indicates the incorporation of GAN may reduce the generalisability of the model to images collected under different undersampling ratios and contribute to its reconstruction instability.

Another instability issue with DAGAN is that if small structural details were added to an MRI image, such as random letters, and this image was undersampled retrospectively, DAGAN was unable to reconstruct the structural details. In contrast, variational network and DC-CNN were superior in recovering these structural details [95]. Even though these two instability tests revealed a critical weakness of DAGAN, DAGAN, being one of the earliest GAN based methods, may not capture the improved reconstruction accuracy and network architectural complexity of the

Table 13.4 The quantitative metrics (PSNR, SSIM, and RMSE) of the knee MRI data using different GAN based methods. The bold numbers indicate the best results

Mask	AF/SR	Metric	ZF	DAGAN	KIGAN	ReconGAN	RefineGAN
Cartesian	2X	PSNR	34.66±2.98	38.91±1.59	38.53±2.51	42.37±1.60	42.41±1.98
		SSIM	0.95±0.01	0.94±0.01	0.96±0.01	0.97±0.00	0.98±0.00
		RMSE	1.64±1.32	0.52±0.31	0.83±1.03	0.23±0.09	0.24±0.16
	4X	PSNR	27.31±3.23	34.35±1.77	34.70±1.74	**34.88±1.96**	35.58±1.74
		SSIM	0.84±0.02	0.86±0.02	0.89±0.02	0.90±0.02	0.91±0.02
		RMSE	10.30±11.00	1.55±1.20	1.49±1.27	1.37±0.93	1.14±0.66
	6X	PSNR	25.15±3.37	32.67±1.89	30.83±2.09	32.34±2.16	33.36±1.81
		SSIM	0.79±0.03	0.82±0.03	0.84±0.02	0.86±0.02	0.87±0.02
		RMSE	18.2±22.3	2.36±2.09	4.08±4.29	2.62±2.19	2.00±1.52
Radial	50%	PSNR	35.17±1.37	36.70±1.37	37.17±1.37	38.38±1.22	38.91±1.21
		SSIM	0.90±0.02	0.91±0.01	0.92±0.01	**0.93±0.01**	0.93±0.01
		RMSE	2.37±0.50	1.97±0.59	1.53±0.18	1.22±0.16	1.14±0.15
	30%	PSNR	32.69±1.13	34.02±1.39	35.27±1.40	35.96±1.27	36.21±1.29
		SSIM	0.80±0.03	0.87±0.02	0.88±0.01	0.88±0.02	0.89±0.02
		RMSE	4.87±0.89	2.23±0.76	1.39±0.67	1.61±0.22	1.36±0.32
	20%	PSNR	30.77±1.15	33.69±1.51	33.49±1.54	34.48±1.26	34.72±1.30
		SSIM	0.75±0.03	0.84±0.02	0.85±0.02	**0.86±0.02**	0.86±0.02
		RMSE	7.13±1.24	2.76±0.87	2.68±0.73	1.91±0.27	1.61±0.34
Spiral	50%	PSNR	35.62±1.27	39.37±1.39	41.39±1.37	42.53±0.75	43.04±0.77
		SSIM	0.91±0.02	0.94±0.01	0.94±0.02	0.96±0.01	0.97±0.02
		RMSE	1.78±0.54	1.06±0.10	0.97±0.14	0.73±0.15	0.61±0.07
	30%	PSNR	33.48±1.12	36.19±1.35	36.26±1.18	38.20±1.31	38.49±1.33
		SSIM	0.86±0.02	0.90±0.01	0.90±0.02	**0.92±0.01**	0.92±0.01
		RMSE	3.71±0.91	1.67±0.72	1.64±0.35	1.24±0.18	0.98±0.23
	20%	PSNR	30.97±1.14	34.88±1.39	33.83±1.50	36.51±1.23	36.75±1.27
		SSIM	0.81±0.03	0.88±0.02	0.88±0.01	**0.90±0.02**	0.90±0.02
		RMSE	5.39±1.26	2.29±0.56	1.95±0.25	1.51±0.20	1.32±0.14

more recent methods [90, 91]. Hence, more instability testing is required on more GAN based methods to conclude whether this category of CS-MRI techniques suffers from reconstruction instability.

These results on instability testing still question the generalisability of GAN based models. In other words, it is unclear whether after being trained with images from one source, e.g., one organ, under one undersampling ratio etc., a model will achieve equivalent performance on images from another source. This is often referred to in the literature as a zero-shot inference test. However, GAN based methods appear to excel in these inference tests. To illustrate, DAGAN, having been trained with MR images of the brain of a healthy volunteer, successfully reconstructed the MR images from a patient with the brain tumour, faithfully preserving the tumour structure [3]. Another example is that GANCS, another GAN based method, was trained with MR

images of the abdomen and any images containing the enlarged adrenal gland were removed from the training set. In testing, GANCS was capable of recovering enlarged adrenal gland in patients with adrenal hypertrophy. No evidence of hallucination of the reconstructed images was found [63]. These successes in the zero-shot inference of pathological structures challenged the instability test results from Antun et al. One interpretation is that the additional structural details added to the MR images by Antun et al., such as letters, may not represent real-life physiological or pathological deviations of the MR images from the training set. It was also possible that the zero-shot inference tests on DAGAN and GANCS concerned only a small sample size, limited to a few pathological conditions. Larger scale zero-shot inference tests are necessary to evaluate the generalisability of GAN based CS-MRI methods.

In summary, future research directions on GAN based fast MRI may include more robust training strategies, e.g., combining GAN with genetic algorithm [96], incorporating edge and texture enhancement [54], reducing possible hallucination [97, 98], coupling with explainable AI (XAI) modules [99, 100], and further consideration with MR physics [55].

13.5 Conclusion

We carried out a mini review, benchmarked and compared four different GAN-based network architectures for fast MRI reconstruction in this chapter. Our comparison used the various sampling patterns, different masks on corresponding datasets that have covered commonly used clinical MRI scenarios. For our systematic research and measurement, we used both traditional and newly proposed quantitative tools. The outcomes of qualitative visualization were also examined and compared. To summarise, our mini review and benchmarking have revealed that all GAN-based approaches could obtain promising results at lower acceleration factors. However, when the acceleration factors are high, some GAN-based architectures, such as DAGAN and KIGAN, may not be sufficient for MRI reconstruction applications. Furthermore, as compared to other GAN-based methods, the RefineGAN has improved reconstruction accuracy and perceptual efficiency. Future development incorporating MR physics and XAI into GAN based models will provide promising pathways for clinical deployment with more transparency of the reconstruction algorithms.

References

1. Kieren, G.H.: Reducing acquisition time in clinical MRI by data undersampling and compressed sensing reconstruction. Phys. Med. Biol. R297–322 (2015)
2. Lustig, M., Donoho, D., Pauly, J.M.: Sparse MRI: the application of compressed sensing for rapid MR imaging. Magn. Reson. Med. **58**(6), 1182–1195 (2007)
3. Yang, G., Yu, S., Dong, H., Slabaugh, G., Pier, L.D., Ye, X., Liu, F., Arridge, S., Keegan, J., Guo, Y., Firmin, D.: DAGAN: deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction. IEEE Trans. Med. Imaging **37**(6), 1310–1321 (2018)
4. Suetens, P.: Fundamentals of Medical Imaging, 2nd edn. Cambridge University Press (2009)
5. Mansfield, P.: Multi-planar image formation using NMR spin echoes. J. Phys. C: Solid State Phys. (1977)
6. Hennig, J., Nauerth, A., Friedburg, H.: RARE imaging: a fast imaging method for clinical MR. Magn. Reson. Med. **3**(6), 823–833 (1986)
7. Haase, A., Frahm, J., Matthaei, D., Merboldt, K.-D.: FLASH imaging. Rapid NMR Imaging Using Low Flip-Angle Pulses. Technical report (1986)
8. Zisselman, E., Adler, A., Elad, M.: Compressed learning for image classification: a deep neural network approach. In: Handbook of Numerical Analysis, vol. 19, pp. 3–17. Elsevier B.V., Jan 2018
9. Fair, M.J., Gatehouse, P.D., DiBella, E.V.R., Firmin, D.N.: A review of 3D first-pass, whole-heart, myocardial perfusion cardiovascular magnetic resonance, Aug 2015
10. Ravishankar, S., Bresler, Y.: MR image reconstruction from highly undersampled k-space data by dictionary learning. IEEE Trans. Med. Imaging **30**(5), 1028–1041 (2011)
11. Yang, J., Zhang, Y., Yin, W.: A fast alternating direction method for TVL1-L2 signal reconstruction from partial Fourier data. IEEE J. Sel. Top. Signal Process. **4**(2), 288–297 (2010)
12. Eksioglu, E.M.: Decoupled algorithm for MRI reconstruction using nonlocal block matching model: BM3D-MRI. J. Math. Imaging Vis. **56**(3), 430–440 (2016)
13. Yue, H., Ongie, G., Ramani, S., Jacob, M.: Generalized higher degree total variation (HDTV) regularization. IEEE Trans. Image Process. **23**(6), 2423–2435 (2014)
14. Liu, Y., Cai, J.-F., Zhan, Z., Guo, D., Ye, J., Chen, Z., Xiaobo, Q.: Balanced sparse model for tight frames in compressed sensing magnetic resonance imaging. PLOS ONE **10**(4), e0119584 (2015)
15. Mohammad H. Kayvanrad, A. Jonathan McLeod, John S.H. Baxter, Charles A. McKenzie, and Terry M. Peters. Stationary wavelet transform for under-sampled MRI reconstruction. *Magnetic Resonance Imaging*, 32(10):1353–1364, dec 2014
16. M. Guerquin-Kern, M. Haberlin, K. P. Pruessmann, and M. Unser. A fast wavelet-based reconstruction method for magnetic resonance imaging. *IEEE Transactions on Medical Imaging*, 30(9):1649–1660, sep 2011
17. Hammernik, K., Klatzer, T., Kobler, E., Recht, M.P., Sodickson, D.K., Pock, T., Knoll, F.: Learning a variational network for reconstruction of accelerated MRI data. Mag. Reson. Med. **79**(6), 3055–3071 (2018)
18. Chen, J., Yang, G., Khan, H., Zhang, H., Zhang, Y., Zhao, S., Mohiaddin, R., Wong, T., Firmin, D., Keegan, J.: JAS-GAN: generative adversarial network based joint atrium and scar segmentation on unbalanced atrial targets. IEEE J. Biomed. Health Inf. (2021)
19. Yinzhe, W., Hatipoglu, S., Alonso-Álvarez, D., Gatehouse, P., Li, B., Gao, Y., Firmin, D., Keegan, J., Yang, G.: Fast and automated segmentation for the three-directional multi-slice cine myocardial velocity mapping. Diagnostics **11**(2), 346 (2021)
20. Wu, Y., Hatipoglu, S., Alonso-Álvarez, D., Gatehouse, P., Firmin, D., Keegan, J., Yang, G.: Automated multi-channel segmentation for the 4D myocardial velocity mapping cardiac MR. In: Medical Imaging 2021: Computer-Aided Diagnosis, vol. 11597, pp. 115970P. International Society for Optics and Photonics (2021)
21. Jin, Yao, Y., G., Fang, Y., Li, R., Xiaomei, X., Liu, Y., Lai, X.: 3D PBV-Net: an automated prostate MRI data segmentation method. Comput. Biol. Med. **128**, 104160 (2021)

22. Zhou, X., Ye, Q., Jiang, Y., Wang, M., Niu, Z., Menpes-Smith, W., Fang, E.F., Liu, Z., Xia, J., Yang, G.: Systematic and comprehensive automated ventricle segmentation on ventricle images of the elderly patients: a retrospective study. Frontiers Aging Neurosci. **12** (2020)

23. Liu, Y., Yang, G., Hosseiny, M., Azadikhah, A., Mirak, S.A., Miao, Q., Raman, S.S., Sung, K.: Exploring uncertainty measures in bayesian deep attentive neural networks for prostate zonal segmentation. IEEE Access **8**, 151817–151828 (2020)

24. Ferreira, P.F., Martin, R.R., Scott, A.D., Khalique, Z., Yang, G., Nielles-Vallespin, S., Pennell, D.J., Firmin, D.N.: Automating in vivo cardiac diffusion tensor postprocessing with deep learning–based segmentation. Magn. Reson. Med. **84**(5), 2801–2814 (2020)

25. Li, M., Wang, C., Zhang, H., Yang, G.: Mv-ran: multiview recurrent aggregation network for echocardiographic sequences segmentation and full cardiac cycle analysis. Comput. Biol. Med. **120**, 103728 (2020)

26. Liu, Y., Yang, G., Mirak, S.A., Hosseiny, M., Azadikhah, A., Zhong, X., Reiter, R.E., Lee, Y., Raman, S.S., Sung, K.: Automatic prostate zonal segmentation using fully convolutional network with feature pyramid attention. IEEE Access **7**, 163626–163632 (2019)

27. Zhuang, X., Li, L., Payer, C., Štern, D., Urschler, M., Heinrich, M.P., Oster, J., Wang, C., Smedby, Ö., Bian, C., et al.: Evaluation of algorithms for multi-modality whole heart segmentation: an open-access grand challenge. Med. Image Anal. **58**, 101537 (2019)

28. Mo, Y., Liu, F., McIlwraith, D., Yang, G., Zhang, J., He, T., Guo, Y. (2018) The deep poincaré map: A novel approach for left ventricle segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 561–568. Springer

29. Zhang, W., Yang, G., Huang, H., Yang, W., Xu, X., Liu, Y., Lai, X.: ME-Net: multi-encoder net framework for brain tumor segmentation. Int. J. Imaging Syst. Technol. (2021)

30. Yang, G., Chen, J., Gao, Z., Li, S., Ni, H., Angelini, E., Wong, T., Mohiaddin, R., Nyktari, E., Wage, R., et al.: Simultaneous left atrium anatomy and scar segmentations via deep learning in multiview information with attention. Fut. Gen. Comput. Syst. **107**, 215–228 (2020)

31. Li, L., Fuping, W., Yang, G., Lingchao, X., Wong, T., Mohiaddin, R., Firmin, D., Keegan, J., Zhuang, X.: Atrial scar quantification via multi-scale CNN in the graph-cuts framework. Med. Image Anal. **60**, 101595 (2020)

32. Zhang, L., Yang, G., Ye, X.: Automatic skin lesion segmentation by coupling deep fully convolutional networks and shallow network with textons. J. Med. Imaging **6**(2), 024001 (2019)

33. Yang, G., Zhuang, X., Khan, H., Nyktari, E., Haldar, S., Li, L., Wage, R., Ye, X., Slabaugh, G., Mohiaddin, R. et al.: Left atrial scarring segmentation from delayed-enhancement cardiac MRI images: a deep learning approach. Cardiovasc. Imaging Image Anal. 109 (2018)

34. Bakas, S., Reyes, M., Jakab, A., Bauer, S., Rempfler, M., Crimi, A., Shinohara, R.T., Berger, C., Ha, S.M., Rozycki, M., et al.: Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the brats challenge (2018). arXiv:1811.02629

35. Mok, T.C.W., Chung, A.: Fast symmetric diffeomorphic image registration with convolutional neural networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 4644–4653 (2020)

36. de Vos, B.D., Berendsen, F.F., Viergever, M.A., Staring, M., Išgum, I.: End-to-end unsupervised deformable image registration with a convolutional neural network. In: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, pp. 204–212. Springer (2017)

37. Wu, G., Kim, M., Wang, Q., Munsell, B.C., Shen, D.: Scalable high-performance image registration framework by unsupervised deep feature representations learning. IEEE Trans. Biomed. Eng. **63**(7), 1505–1516 (2015)

38. Chenchu, X., Zhang, D., Chong, J., Chen, B., Li, S.: Synthesis of gadolinium-enhanced liver tumors on nonenhanced liver MR images using pixel-level graph reinforcement learning. Med. Image Anal. **69**, 101976 (2021)

39. Wang, C., Yang, G., Papanastasiou, G., Tsaftaris, S.A., Newby, D.E., Gray, C., Macnaught, G., MacGillivray, T.J.: DiCyc: GAN-based deformation invariant cross-domain information fusion for medical image synthesis. Inf. Fusion **67**, 147–160 (2021)

40. Gao, Z., Zhang, H., Dong, S., Sun, S., Wang, X., Yang, G., Wu, W., Li, S., de Albuquerque, V.H.C.: Salient object detection in the distributed cloud-edge intelligent network. IEEE Netw. **34**(2), 216–224 (2020)

41. Wang, C., Dong, S., Zhao, X., Papanastasiou, G., Zhang, H., Yang, G.: Saliencygan: deep learning semisupervised salient object detection in the fog of IOT. IEEE Trans. Ind. Inf. **16**(4), 2667–2676 (2019)

42. Ali, A.-R., Li, J., Kanwal, S., Yang, G., Hussain, A., O'Shea, S.J.: A novel fuzzy multilayer perceptron (F-MLP) for the detection of irregularity in skin lesion border using dermoscopic images. Frontiers Med. **7** (2020)

43. Yang, M., Xiao, X., Liu, Z., Sun, L., Guo, W., Cui, L., Sun, D., Zhang, P., Yang, G.: Deep retinanet for dynamic left ventricle detection in multiview echocardiography classification. Sci. Program. (2020)

44. Li, M., Dong, S., Zhang, K., Gao, Z., Wu, X., Zhang, H., Yang, G., Li, S.: Deep learning intra-image and inter-images features for co-saliency detection. In: BMVC, vol. 291 (2018)

45. Dong, S., Sun, G.S., Wang, X., Li, M., Zhang, H., Yang, G., Liu, H., Li, S., et al.: Holistic and deep feature pyramids for saliency detection. In: BMVC, p. 67 (2018)

46. Dong, H., Yang, G., Liu, F., Mo, Y., Guo, Y.: Automatic brain tumor detection and segmentation using U-Net based fully convolutional networks. In: Annual Conference on Medical Image Understanding and Analysis, pp. 506–517. Springer (2017)

47. Hu, S., Gao, Y., Niu, Z., Jiang, Y., Li, L., Xiao, X., Wang, M., Fang, E.F., Menpes-Smith, W., Xia, J., et al.: Weakly supervised deep learning for COVID-19 infection detection and classification from CT images. IEEE Access **8**, 118869–118883 (2020)

48. Cao, Y., Wang, Z., Liu, Z., Li, Y., Xiao, X., Sun, L., Zhang, Y., Hou, H., Zhang, P., Yang, G.: Multiparameter synchronous measurement with IVUS images for intelligently diagnosing coronary cardiac disease. IEEE Trans. Instrum. Meas. (2020)

49. Zhang, N., Yang, G., Gao, Z., Chenchu, X., Zhang, Y., Shi, R., Keegan, J., Lei, X., Zhang, H., Fan, Z., et al.: Deep learning for diagnosis of chronic myocardial infarction on nonenhanced cardiac cine MRI. Radiology **291**(3), 606–617 (2019)

50. Roberts, M., Driggs, D., Thorpe, M., Gilbey, J., Yeung, M., Ursprung, S., Aviles-Rivero, A.I., Etmann, C., McCague, C., Beer, L., et al.: Machine learning for COVID-19 detection and prognostication using chest radiographs and CT scans: a systematic methodological review (2020). arXiv:2008.06388

51. Soltaninejad, M., Zhang, L., Lambrou, T., Yang, G., Allinson, N., Ye, X.: MRI brain tumor segmentation and patient survival prediction using random forests and fully convolutional networks. In: International MICCAI Brainlesion Workshop, pp. 204–215. Springer (2017)

52. Jin, C., Yu, H., Ke, J., Ding, P., Yi, Y., Jiang, X., Duan, X., Tang, J., Chang, D.T., Wu, X., et al.: Predicting treatment response from longitudinal images using multi-task deep learning. Nat. Commun. **12**(1), 1–11 (2021)

53. Nielsen, A., Hansen, M.B., Tietze, A., Mouridsen, K.: Prediction of tissue outcome and assessment of treatment effect in acute ischemic stroke using deep learning. Stroke **49**(6), 1394–1401 (2018)

54. Chen, Y., Firmin, D., Yang, G.: Wavelet improved GAN for MRI reconstruction. In: Medical Imaging 2021: Physics of Medical Imaging, vol. 11595, p. 1159513. International Society for Optics and Photonics (2021)

55. Lv, J., Wang, C., Yang, G.: PIC-GAN: a parallel imaging coupled generative adversarial network for accelerated multi-channel MRI reconstruction. Diagnostics **11**(1), 61 (2021)

56. Lv, J., Zhu, J., Yang, G.: Which GAN? a comparative study of generative adversarial network (GAN) based fast MRI reconstruction. Philos. Trans. R. Soc. A

57. Yuan, Z., Jiang, M., Wang, Y., Wei, B., Li, Y., Wang, P., Menpes-Smith, W., Niu, Z., Yang, G.: SARA-GAN: Self-attention and relative average discriminator based generative adversarial networks for fast compressed sensing MRI reconstruction. Frontiers Neuroinformatics **14** (2020)

58. Guo, Y., Wang, C., Zhang, H., Yang, G.: Deep attentive wasserstein generative adversarial networks for mri reconstruction with recurrent context-awareness. In: International Conference

on Medical Image Computing and Computer-Assisted Intervention, pp. 167–177. Springer (2020)

59. Schlemper, J., Yang, G., Ferreira, P., Scott, A., McGill, L.-A., Khalique, Z., Gorodezky, M., Roehl, M., Keegan, Pennell, J.D., et al.: Stochastic deep compressive sensing for the reconstruction of diffusion tensor cardiac MRI. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 295–303. Springer (2018)

60. Seitzer, M., Yang, G., Schlemper, J., Oktay, O., Würfl, T., Christlein, V., Wong, T., Mohiaddin, R., Firmin, D., Keegan, J., et al.: Adversarial and perceptual refinement for compressed sensing mri reconstruction. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 232–240. Springer (2018)

61. Donoho, D.L., et al.: Compressed sensing. IEEE Trans. Inf. Theory **52**(4), 1289–1306 (2006)

62. Quan, T.M., Nguyen-Duc, T., Jeong, W.K.: Compressed sensing MRI reconstruction using a generative adversarial network with a cyclic loss. IEEE Trans. Med. Imaging **37**(6), 1488–1497 (2018)

63. Mardani, M., Gong, E., Cheng, J.Y., Vasanawala, S.S., Zaharchuk, G., Xing, L., Pauly, J.M.: Deep generative adversarial neural networks for compressive sensing MRI. IEEE Trans. Med. Imaging **38**(1), 167–179 (2019)

64. Schlemper, J., Caballero, J, Hajnal, J.V., Price, A.N., Rueckert, D.: A deep cascade of convolutional neural networks for dynamic MR image reconstruction. IEEE Trans. Med. Imaging **37**(2), 491–503 (2018)

65. Huang, Q., Yang, D., Wu, P., Qu, H., Yi, J., Metaxas, D.: MRI reconstruction via cascaded channel-wise attention network. In: Proceedings—International Symposium on Biomedical Imaging, Apr. 2019, pp. 1622–1626. IEEE Computer Society, Apr. 2019

66. Eo, Jun, Y., Kim, T., Jang, J., Lee, H.-J., Hwang, D.: KIKI-net: cross-domain convolutional neural networks for reconstructing undersampled magnetic resonance images. Magn. Reson. Med. **80**(5), 2188–2201 (2018)

67. Yu, S., Dong, H., Yang, G., Slabaugh, G., Dragotti, P.L., Ye, X., Liu, F., Arridge, S., Keegan, J., Firmin, D., et al.: Deep de-aliasing for fast compressive sensing MRI (2017). arXiv:1705.07137

68. Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative Adversarial Nets. Technical report (2014)

69. Deng, X., Yang, R., Zurich, E., Xu, M., Dragotti, P.L.: Wavelet Domain Style Transfer for an Effective Perception-distortion Tradeoff in Single Image Super-Resolution. Technical report (2019)

70. Deshmane, A., Gulani, V., Griswold, M.A., Seiberlich, N.: Parallel MR imaging. J. Magn. Reson. Imaging **36**(1), 55–72 (2012). https://onlinelibrary.wiley.com/doi/pdf/10.1002/jmri.23639

71. Zhang, X., Lian, Q., Yang, Y., Su, Y.: A deep unrolling network inspired by total variation for compressed sensing MRI. Digit. Signal Process.: Rev. J. **107** (2020). Publisher: Elsevier

72. Diamond, S., Sitzmann, V., Heide, F., Wetzstein, G.: Unrolled Optimization with Deep Priors (2017). Publisher: arxiv.org

73. Qin, C., Schlemper, J., Caballero, J., Price, A.N., Hajnal, J.V., Rueckert, D.: Convolutional recurrent neural networks for dynamic MR image reconstruction. IEEE Trans. Med. Imaging **38**(1), 280–290 (2019). Publisher: ieeexplore.ieee.org

74. Mohana, M.J., Madhulika, M.S., Divya, G.D., Meghana, R.K., Apoorva, S.: Feature extraction using convolution neural networks (CNN) and deep learning. In: 2018 3rd IEEE International Conference on Recent Trends in Electronics, Information Communication Technology (RTE-ICT), pp. 2319–2323 (2018)

75. Zhao, D., Zhao, F., Gan, Y.: Reference-driven compressed sensing MR image reconstruction using deep convolutional neural networks without pre-training. Sensors (Switzerland) **20**(1) (2020)

76. Liang, D., Cheng, J., Ke, Z., Ying, L.: Deep MRI Reconstruction: Unrolled Optimization Algorithms Meet Neural Networks (2019). arXiv:1907.11711. Publisher: arxiv.org

77. Liang, D., Cheng, J., Ke, Z., Ying, L.: Deep magnetic resonance image reconstruction: inverse problems meet neural networks. IEEE Signal Process. Mag. **37**(1), 141–151 (2020). Publisher: ieeexplore.ieee.org

78. Zhang, H.-M., Dong, B.: A review on deep learning in medical image reconstruction. J. Oper. Res. Soc. China **8**(2), 311–340 (2020)

79. Lee, D., Yoo, J., Tak, S., Ye, J.C.: Deep residual learning for accelerated MRI using magnitude and phase networks. IEEE Trans. Biomed. Eng. **65**(9), 1985–1995 (2018)

80. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 9351, pp. 234–241. Springer, May 2015

81. Ledig, C., Theis, L., Huszar, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., Shi, W.: Photo-realistic single image super-resolution using a generative adversarial network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 2017

82. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition (2015)

83. Russakovsky, O., Deng, J., Hao, S., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C., Fei-Fei, L.: Imagenet large scale visual recognition challenge. Int. J. Comput. Vis. **115**(3), 211–252 (2015)

84. Shaul, R., David, I., Shitrit, O., Raviv, R.: Subsampled brain MRI reconstruction by generative adversarial neural networks. Med. Image Anal. **65**, 101747 (2020)

85. Heusel, M., Ramsauer, M., Unterthiner, T., Nessler, B., Hochreiter, S.: GANs trained by a two time-scale update rule converge to a local nash equilibrium (2018). arXiv:1706.08500

86. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z.: Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2818–2826 (2016)

87. Chen, Z., Tong, Y.: Face Super-Resolution Through Wasserstein GANs. Technical report (2017)

88. Wiatrak, M., Albrecht, S.V., Nystrom, A.: Stabilizing GANs: A Survey Stabilizing Generative Adversarial Networks: A Survey. Technical report (2020)

89. Arjovsky, S.C., Bottou, L.: Wasserstein GAN. Technical report (2017)

90. Jiang, M., Yuan, Yang, Z.X., Zhang, J., Gong, Y., Xia, L., Li, T.: Accelerating CS-MRI reconstruction with fine-tuning wasserstein generative adversarial network. IEEE Access **7**, 152347–152357 (2019). Publisher: ieeexplore.ieee.org

91. Oh, G., Sim, B., Chung, H.J., Sunwoo, L., Ye, J.C.: Unpaired deep learning for accelerated MRI using optimal transport driven CycleGAN. IEEE Trans. Comput. Imaging **6**, 1285–1296 (2020). Publisher: ieeexplore.ieee.org

92. Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., Courville, A.: Improved Training of Wasserstein GANs Montreal Institute for Learning Algorithms. Technical report (2017)

93. Miyato, T., Kataoka, T., Koyama, M., Yoshida, Y.: Spectral Normalization for Generative Adversarial Networks (2018). [cs, stat], Feb. 2018. arXiv: 1802.05957

94. Zhao, H., Gallo, O., Frosio, I., Kautz, J.: Loss functions for image restoration with neural networks. IEEE Trans. Comput. Imaging **3**(1), 47–57 (2017). Conference Name: IEEE Transactions on Computational Imaging

95. Antun, V., Renna, F., Poon, C., Adcock, B., Hansen, A.C.: On instabilities of deep learning in image reconstruction and the potential costs of AI. In: Proceedings of the National Academy of Sciences, vol. 117(48), pp. 30088–30095, Dec. 2020. Publisher: National Academy of Sciences Section: Colloquium on the Science of Deep Learning (2020)

96. Hao, J., Wang, C., Zhang, H., Yang, G.: Annealing genetic GAN for minority oversampling (2020). arXiv:2008.01967

97. Zhu, J., Yang, G., Lio, P.: How can we make GAN perform better in single medical image super-resolution? A lesion focused multi-scale approach. In: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), pp. 1669–1673. IEEE (2019)

98. Zhu, J., Yang, G., Lio, P.: Lesion focused super-resolution. In: Medical Imaging 2019: Image Processing, vol. 10949, p. 109491L. International Society for Optics and Photonics (2019)
99. Ye, Q., Xia, J., Yang, G.: Explainable AI for covid-19 CT classifiers: an initial comparison study (2021). arXiv:2104.14506
100. Yang, G., Ye, Q., Xia, J.: Unbox the black-box for the medical explainable AI via multi-modal and multi-centre data fusion: a mini-review, two showcases and beyond (2021). arXiv:2102.01998