City Research Online

City St George's, University of London

Citation: Yu, S., Dong, H., Yang, G., Slabaugh, G. G., Dragotti, P. L., Ye, X., Liu, F., Arridge, S., Keegan, J., Firmin, D. & et al (2017). Deep De-Aliasing for Fast Compressive Sensing MRI. IEEE Transactions on Medical Imaging,

This is the submitted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/17941/

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to City Research Online policy.

City Research Online: http://openaccess.city.ac.uk/ publications@citystgeorges.ac.uk

arXiv:1705.07137v1 [cs.CV] 19 May 2017

Deep De-Aliasing for Fast Compressive Sensing MRI

Simiao Yu[†1], Hao Dong[†1], Guang Yang[2], Greg Slabaugh[3], Pier Luigi Dragotti[4], Xujiong Ye[5], Fangde Liu[1], Simon Arridge[6], Jennifer Keegan[2], David Firmin[2], and Yike Guo[1‡]

[1] Data Science Institute, Imperial College London, SW7 2AZ, London, UK
[2] National Heart & Lung Institute, Imperial College London, SW3 6NP, London, UK
[3] Dept. of Computer Science, City University London, EC1V 0HB, London, UK
[4] EEE Department, Imperial College London, SW3 6NP, London, UK
[5] School of Computer Science, University of Lincoln, LN6 7TS, Lincoln, UK
[6] CMIC, University College London, WC1E 6BT, London, UK

Abstract. Fast Magnetic Resonance Imaging (MRI) is highly in demand for many clinical applications in order to reduce the scanning cost and improve the patient experience. This can also potentially increase the image quality by reducing the motion artefacts and contrast washout. However, once an image field of view and the desired resolution are chosen, the minimum scanning time is normally determined by the requirement of acquiring sufficient raw data to meet the Nyquist–Shannon sampling criteria. Compressive Sensing (CS) theory has been perfectly matched to the MRI scanning sequence design with much less required raw data for the image reconstruction. Inspired by recent advances in deep learning for solving various inverse problems, we propose a conditional Generative Adversarial Networks-based deep learning framework for de-aliasing and reconstructing MRI images from highly undersampled data with great promise to accelerate the data acquisition process. By coupling an innovative content loss with the adversarial loss our de-aliasing results are more realistic. Furthermore, we propose a refinement learning procedure for training the generator network, which can stabilise the training with fast convergence and less parameter tuning. We demonstrate that the proposed framework outperforms state-of-the-art CS-MRI methods, in terms of reconstruction error and perceptual image quality. In addition, our method can reconstruct each image in 0.22ms–0.37ms, which is promising for real-time applications.

1 Introduction

Magnetic Resonance Imaging (MRI) is an invaluable technique for clinical medical imaging in that it provides non-invasive, reproducible and quantitative measurements of tissue structural, anatomical and functional information. Despite its

[†] Co-first authors
[‡] Corresponding author. Emails: {simiao.yu13,hao.dong11,g.yang}@imperial.ac.uk {p.dragotti,fangde.liu,j.keegan,d.firmin,y.guo}@imperial.ac.uk gregory.slabaugh.1@city.ac.uk, xye@lincoln.ac.uk, s.arridge@cs.ucl.ac.uk

unique flexibility for imaging different tissue types and organs of the human body (because the sensitivity of the image to tissue characteristics can be extensively tuned), prolonged acquisition times limit its usage due to expensive cost and considerations of patient comfort and compliance [1].

MRI is associated with an inherently slow acquisition speed that is due to data samples not being collected directly in the image space but rather in k-space, which contains spatial-frequency information. Here k-space and the image space are inversely related: resolution in one domain determines extent in the other. The raw data samples are acquired sequentially in k-space and the speed at which k-space can be traversed is limited by physiological and hardware constraints [2]. Once the desired field-of-view and spatial resolution of the MRI images are prescribed, the k-space raw data we need to acquire is conventionally determined by the Nyquist–Shannon sampling criteria [3].

One possible fast MRI approach is to undersample k-space, to which can provide an acceleration rate proportional to the undersampling ratio. However, this undersampling in k-space would violate the Nyquist–Shannon sampling criteria, and thus generates aliasing artefacts once the images have been reconstructed. Therefore, the main challenge for fast MRI is to find an algorithm that can reconstruct an uncorrupted or de-aliased image from the undersampled k-space.

The mathematical framework of Compressive Sensing (CS) has been intensively investigated from about a decade ago [4], and was almost immediately considered for fast MRI applications due to the inherent suitability of the MRI data. Firstly, in general, the medical imagery acquired by MRI is naturally compressible. CS utilises the implicit sparsity of MRI images to reconstruct accelerated acquisitions [5]. Here the term 'sparsity' describes a matrix of image pixels or raw data points, which is predominately zero valued or namely compressible. Such sparseness may exist either in the image domain or more commonly via a suitable mathematical representation in a transform domain of the images due to redundancy in a single image or over a series of related images. Using this property, CS allows accurate reconstruction from undersampled raw data, with the proviso that the sampling pattern is 'random' to create incoherent undersampling artefacts that a proper 'nonlinear reconstruction' can be applied to suppress noise-like artefacts without degrading image quality of the reconstruction [2]. Secondly, the MRI scanners acquire raw data samples in spatial-frequency encoded k-space. This allows aforementioned random undersampling to be implemented on the scanner, and the fast MRI can be achieved by acquiring less data in the first place, although this violates the requirements of the Nyquist–Shannon sampling criteria.

Sparse regularisation, which is a key component for successful fast CS based MRI (CS-MRI), can be explored in a specific transform domain or generally in a dictionary-based subspace [6]. Classic fast CS-MRI used predefined and fixed sparsifying transforms, e.g., total variation (TV) [7, 8, 9], discrete cosine transforms [10, 11, 12] and discrete wavelet transforms [13, 14, 15]. In addition, this has been extended to more flexible sparse representation learnt directly from data using dictionary learning [16, 17, 18].

Despite the promise of using fast CS-MRI, most routine clinical MRI scannings are still based on standard Cartesian sequences or accelerated only using parallel imaging. The main challenges are: (1) it can be challenging to satisfy the incoherence criteria required by CS [1]; (2) the sparsifying transforms used in current CS methods might be too simple to capture complex image details associated with subtle differences of the biological tissues, e.g., TV based sparsifying transform can introduce staircase artefacts in the reconstructed image [19]; (3) nonlinear optimisation solvers usually involve iterative computing and updating that may result in relatively long reconstruction time [1]; (4) inappropriate hyper-parameters predicted in current CS methods can cause over-regularisation that will yield overly smooth and unnatural looking reconstructions or images with residual undersampling artefacts [1] and (5) the acceleration rate is still limited ($2\times$ to $6\times$ acceleration).

Recently, deep learning has been widely and successfully applied in many computer vision problems. Essentially, CS-MRI reconstruction is solving a generalised inverse problem that is analogous to image super-resolution (SR), de-noising and inpainting that have been well investigated in computer vision. Deep neural network architectures, especially convolutional neural networks (CNNs), are becoming the state-of-the-art technique for tackling such inverse problems, e.g., image SR [20], de-noising and inpainting [21, 22]. Comprehensive reviews on classic CS methods and clinical applications can be found elsewhere [1, 23], and here we briefly review the most relevant publications using deep learning models for the CS-MRI. It is of note that despite the popularity of deep learning in computer vision applications, there has only been preliminary research on deep learning based CS-MRI. From our literature search, we found only two formal publications on this topic [19, 24], and the other three preprints on the arXiv [25, 26, 27]. In general these methods leveraged deep learning (e.g., CNNs) to derive an optimal mapping between the undersampled k-space (aliased reconstruction) and the desired uncorrupted image (de-aliased reconstruction). This has been done in either a sequential manner with classic CS-MRI or in an integrated manner that considered the CNNs based training as an additional regularisation term [24]. The experimental results have shown some promise; however, the improvement was not significantly different from what classic CS-MRI can achieve although the reconstruction speed has been dramatically improved [19]. Moreover, only up to $6\times$ acceleration could be achieved by these methods [19, 24, 25, 26, 27].

In this study, we proposed a novel conditional Generative Adversarial Networks (GAN) based deep learning architecture for fast CS-MRI. Our main contributions are: (1) we used the U-Net architecture [28] with skip connections to achieve better reconstruction details; (2) for more realistic reconstruction, we proposed to combine the adversarial loss with a novel content loss considering both pixel-wise mean square error (MSE) and perceptual loss defined by pretrained VGG networks and (3) we proposed a refinement learning for training the generator that can stabilise the training with fast convergence and less parameter tuning. Compared to other state-of-the-art CS-MRI methods, we can achieve up to $10\times$ acceleration with superior results and faster processing time using the GPU.

2 Method

2.1 General CS-MRI

MRI reconstruction naturally deals with complex numbers. Let $\mathrm{x} \in \mathbb{C}^N$ represent a complex-valued MRI image, which consists of $\sqrt{N} \times \sqrt{N}$ pixels formatted as a column vector. The aim is to reconstruct x from the undersampled k-space measurements $\mathrm{y} \in \mathbb{C}^M$ ($M << N$), such that $\mathrm{y} = \mathrm{F}_u\mathrm{x}$, in which $\mathrm{F}_u \in \mathbb{C}^{M \times N}$ is the undersampled Fourier encoding matrix. In order to solve this underdetermined and ill-posed system, one must exploit a-priori knowledge of x that can be formulated as an unconstrained optimisation problem, that is

$$\min_{\mathrm{x}} \ \frac{1}{2}||\mathrm{F}_u\mathrm{x} - \mathrm{y}||_2^2 + \lambda\mathcal{R}(\mathrm{x}), \tag{1}$$

where the least squares part represents the data fidelity term. $\mathcal{R}$ expresses regularisation terms on x and λ is a regularisation parameter. The regularisation terms $\mathcal{R}$ typically involve l_q-norms ($0 \leq q \leq 1$) in the sparsifying domain of x [2].

Previous deep learning based fast CS-MRI studies [24, 25] integrated CNNs into CS-MRI, that is

$$\min_{\mathrm{x}} \ \frac{1}{2}||\mathrm{F}_u\mathrm{x} - \mathrm{y}||_2^2 + \lambda\mathcal{R}(\mathrm{x}) + \zeta||\mathrm{x} - f_{\mathrm{cnn}}(\mathrm{x}_u|\hat{\theta})||_2^2, \tag{2}$$

in which f_{cnn} is the forward propagation of the CNNs parametrised by θ, and ζ is another regularisation parameter. The image generated by the CNNs (i.e., $f_{\mathrm{cnn}}(\mathrm{x}_u|\hat{\theta})$) was used as a reference image and as an additional regularisation term, in which $\hat{\theta}$ represents the optimised parameters of the trained CNNs. In addition, $\mathrm{x}_u = \mathrm{F}_u^H\mathrm{y}$ is the reconstruction from the zero-filled undersampled k-space measurements, where H represents the Hermitian transpose operation.

2.2 General GAN

Generative Adversarial Networks (GAN) [29] consist of a generator network G and a discriminator network D. The goal of the generator G is to map latent variable $\boldsymbol{z}$ to the distribution of the given true data in order to fool a discriminator D, while the discriminator aims to distinguish the true data $\boldsymbol{x}$ from the synthesised fake data $G_{\theta_G}(\boldsymbol{z})$. Mathematically, the training process can be represented by a minimax function with network parameters θ_G and θ_D as following

$$\min_{\theta_G}\max_{\theta_D} \ \mathbb{E}_{\boldsymbol{x}\sim p_{\mathrm{data}}(\boldsymbol{x})}[\log D_{\theta_D}(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D_{\theta_D}(G_{\theta_G}(\boldsymbol{z})))], \tag{3}$$

where latent variable $\boldsymbol{z}$ is sampled from a fixed latent distribution $p_{\boldsymbol{z}}(\boldsymbol{z})$ and real samples $\boldsymbol{x}$ come from a real data distribution $p_{\mathrm{data}}(\boldsymbol{x})$. Extra prior information can constrain the generator G to learn to create samples conditioned on such information, which is known as conditional GAN [30].

2.3 Proposed Method

Figure 1 shows the overall framework of our GAN-based de-aliasing architecture. We used GAN conditioned on images. We fed the zero-filling reconstruction x_u (contains aliasing artefacts) to the generator and would yield the corresponding de-aliased reconstruction $\hat{\mathrm{x}}_u$ in order to stop the discriminator from recognising it from fully-sampled ground truth reconstruction x_t. In other words, we have paired x_t and x_u as the input training data, we output $\hat{\mathrm{x}}_u$.

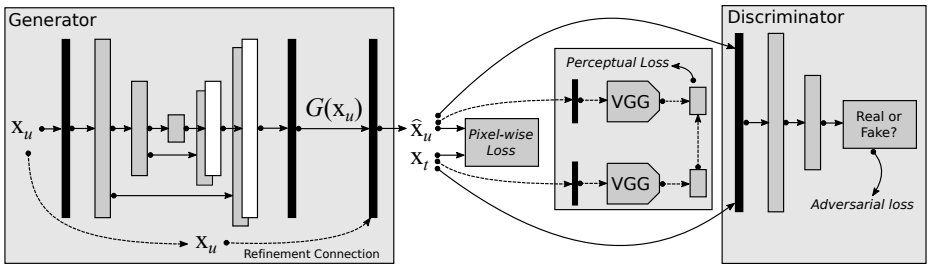

Fig. 1: Schema for our proposed GAN-based de-aliasing for fast CS-MRI.

Adversarial Loss First, instead of using CNNs, we incorporated adversarial loss into our de-aliasing process for the fast CS-MRI that can be expressed as

$$\min_{\theta_G}\max_{\theta_D}\ \mathbb{E}_{\mathrm{x}_t\sim p_{\mathrm{train}}(\mathrm{x}_t)}[\log D_{\theta_D}(\mathrm{x}_t)] + \mathbb{E}_{\mathrm{x}_u\sim p_G(\mathrm{x}_u)}[\log(1 - D_{\theta_D}(G_{\theta_G}(\mathrm{x}_u)))]. \quad (4)$$

Adversarial Learning For the generator G, we used the U-Net architecture [28], which applied skip connections between mirrored layers in the encoder and decoder paths. These skip connections can pass different levels of features to the decoder and gain better reconstruction details. For the output activation function of the generator, we used the hyperbolic tangent function. This adversarial learning process can be considered as using an adaptive loss function to iteratively shift the distribution of the de-aliased reconstruction towards the ground truth distribution.

Content Loss In order to make the generated images more realistic, in addition to the adversarial loss, we also designed a content loss for training the generator, which is formed by coupling a pixel-wise MSE loss and a perceptual VGG loss, that is

$$\min_{\theta_G}\ \alpha\frac{1}{2}||\mathrm{x}_t - \hat{\mathrm{x}}_u||_2^2 + \beta\frac{1}{2}||f_{\mathrm{vgg}}(\mathrm{x}_t) - f_{\mathrm{vgg}}(\hat{\mathrm{x}}_u)||_2^2 + \log(1 - D_{\theta_D}(\hat{\mathrm{x}}_u)) \quad (5)$$
$$\text{s.t.}\ \ \hat{\mathrm{x}}_u = G_{\theta_G}(\mathrm{x}_u),$$

in which the first term describes the MSE loss that is a common choice of the optimisation cost function for the deep learning based fast CS-MRI [19, 26]. In

our study, we used the normalised MSE (NMSE). However, the solution solely based on the optimisation of the MSE loss, which is defined on pixel-wise image difference, could result in perceptually overly smooth reconstructions that often lack high-frequency image details. We therefore defined an additional VGG loss (second term of Eq. 5) to take the perceptual similarity into account [31]. In particular, we used the conv4 output of the VGG as the encoded embedding of the de-aliased output and the ground truth, and computed the MSE between them. By optimising this combined loss, the aim is to train a generator network successfully that can yield realistic de-aliased reconstruction that can fool the discriminator network. Once the generator network has been trained, we can apply it to any new inputs (i.e., initial aliased zero-filled reconstructions), and it will result in the de-aliased reconstruction.

Refinement Learning Another main innovation of our method is that we added the undersampled reconstruction x_u to the generator output to model the final de-aliased reconstruction, i.e., instead of using $\hat{\mathrm{x}}_u = G_{\theta_G}(\mathrm{x}_u)$ we proposed to use $\hat{\mathrm{x}}_u = G_{\theta_G}(\mathrm{x}_u) + \mathrm{x}_u$. In so doing, we transferred the generator from a conditional generative function to a refinement function, i.e., only generate the missing information. This can dramatically reduce the complexity of the learning, make the model more stable with faster convergence. In order to ensure that the de-aliased reconstruction $\hat{\mathrm{x}}_u$ is in a proper intensity scale as the ground truth, we applied a simple ramp function to rescale the image.

Our generator can learn to perform de-aliasing of the zero-filled reconstruction, and create solutions that highly resemble the fully sampled ground truth. We named our method (using GAN architecture, pixel-wise MSE loss, VGG loss, and refinement learning) as Pixel-Perceptual-GAN-Refinement (PPGR). For comparison purpose, we also tested the method without refinement learning that is named as Pixel-Perceptual-GAN (PPG), and the method with pixel-wise MSE only and GAN architecture is denoted as Pixel-GAN (PG).

3 Experimental Settings and Results

3.1 Experimental Settings

Datasets First, we trained and validated our GAN-based deep de-aliasing model using a subset of the IXI dataset [1]. We randomly selected 1605 T_1-weighted MRI images acquired for healthy volunteers. For this subset of the IXI dataset, we demonstrated the robustness of our model using 5-fold cross-validation. Second, we also tested our model using a MICCAI 2013 grand challenge dataset [2]. We randomly included 100 T_1-weighted MRI data for training and 50 MRI data for testing as described by [19] for a comparison study. We simulated both 1D and 2D Gaussian distribution based undersampling masks for our experiments. For each mask, 10%, 20%, 30%, 40% and 50% remaining raw k-space data were simulated representing 10×, 5×, 3.3×, 2.5× and 2× accelerations.

[1] http://brain-development.org/ixi-dataset

[2] http://masiweb.vuse.vanderbilt.edu/workshop2013/index.php/Segmentation_Challenge_Details

Evaluation Methods For both datasets, we reported the NMSE, the Peak Signal-to-Noise Ratio (PSNR in dB) and the Structural Similarity Index (SSIM) [31]. The reconstructed fully sampled k-space data was used as ground truth (GT) for validation. In addition to quantitative metrics, we also evaluated our method using qualitative visualisation of the reconstructed MRI images and the error with respect to the GT (e.g., using absolute difference image amplified $10\times$).

Networks and Training Settings Our GAN architecture was inspired by [32, 33] that consists of multiple convolutional and deconvolutional layers with batch normalisation [34] and leaky ReLU layers. The VGG network [35] we used was pretrained on ImageNet [36]. Detailed architecture settings can be found in the Supplementary Material. For both datasets, we trained separate networks for different sampling ratios with the following mutual hyperparameters: $\alpha = 15$, $\beta = 0.0025$, initial learning rate of 0.0001, batch size of 25. We adopted Adam optimisation [37] with momentum of 0.5. For the IXI dataset, each model was trained by fixed 60 epochs and the learning rate was halved every 30 epochs. For the MICCAI dataset, each model was learnt by employing early stopping and the learning rate was halved every 5 epochs.

Data Augmentation The purpose of data augmentation is to improve the network performance by intentionally producing more training data from the original one. Conventional data augmentation (e.g., image flipping, rotation, shift, brightness adjustment and zoom) can result in displacement fields change but can not create training samples with diverse shapes. The shape of the organs imaged by MRI could be diverse but the variation is limited; therefore, in addition to conventional data augmentation, we also applied elastic distortion [38] that can generate more training data with arbitrary but reasonable shape variations.

Implementation The implementation of our GAN-based de-aliasing model has been done using a high-level Python wrapper (TensorLayer [3]) of the TensorFlow [4] library. We will publish our open-source implementation on the Github.

3.2 Results

Figure 2 shows comparison results of the SSIM between our proposed methods (PG, PPG and PPGR) and the baseline zero-filling (ZF) reconstruction using the IXI dataset (higher SSIM indicates better results). In general, all versions of our GAN-based methods outperformed the baseline reconstruction significantly. PPGR with refinement learning obtained the best SSIM with less variation than PPG and PG methods. We also calculated the NMSE and PSNR (refer to the Supplementary Material), both of which gained significant improvement compared to the baseline reconstruction regardless the random sampling distribution and undersampling ratio.

Table 1 tabulates the NMSE and PSNR results for our MICCAI dataset. Compared to the ZF baseline, our methods also performed much better. When

[3] http://tensorlayer.readthedocs.io
[4] https://www.tensorflow.org

only 10% of k-space remained using a 2D undersampling mask, we could still obtain $>$ 38dB PSNR. Compared to [19] study that used a similar experimental setting and data, we achieved better reconstruction in terms of NMSE and PSNR improvements. Also, we demonstrated that our method could work when the k-space is highly undersampled (e.g., when only 10% of raw data remains we can still obtain SSIM > 0.97). Figure 4 (a) shows the line profile comparison using different reconstruction methods. Compared to the line profile of the GT image, ZF results were clearly over-smooth. Although better details can be observed in PG and PPG results compared to ZF, the PPGR results were the best that were closer to the GT. Figure 4 (b) and (c) demonstrate that after adding the refinement connection, our PPGR method had a much faster convergence and a more stable improvement over the PSNR than the PPG.

As the MICCAI datasets we used and the experimental settings (2D undersampling masks) were similar to [19], we also compared with their quantitative results. Compared to the conventional CS-MRI methods using predefined and fixed sparsifying transforms (e.g., TV [6], RecPF [8] and PBDW [13]), our PPGR method achieved lower NMSE and higher PSNR (e.g., NMSE ≥ 0.062 using conventional CS-MRI v.s. NMSE $= 0.059$ using PPGR, and PSNR $<$ 39dB v.s. PSNR $=$ 43.56dB when the undersampling ratio is 30%). Compared to state-of-the-art CS-MRI methods using non-local sparsity operator, dictionary learning and deep learning (e.g., PANO [39], FDLCP [18], BM3D-MRI [40] and ADMM-Net [19]), we obtained comparable NMSE and improved PSNR, and more importantly our GPU implementation could process each reconstruction in only 0.22ms–0.37ms.

Figure 3 shows examples of reconstructed images using IXI dataset (upper two rows) under 1D Gaussian random sampling and MICCAI datasets (bottom two rows) under 2D Gaussian random sampling with various undersampling ratios, respectively. For each example case, we showed the GT and the reconstructions with various undersampling ratios (Figure 3 1st and 3rd rows), and the difference images between the GT and each reconstruction (Figure 3 2nd and 4th rows). We obtained compelling de-aliasing results compared to the initial difference images that contained significant aliasing artefacts. We can hardly observe any qualitative differences between the reconstructed images and the GT when the undersampling ratio $\geq$30%. However, when the undersampling ratio $\leq$20%, we might start noticing the loss of structural information (e.g., organ edges); however, most of the aliasing artefacts have still been suppressed effectively. Also, in general results obtained from 2D undersampling masks were better than the ones recovered using 1D masks.

4 Discussion

Classic fast CS-MRI treats the image reconstruction task as a nonlinear optimisation problem without considering prior information of the expected appearance of the anatomy or the possible structure of the undersampling artefacts. This is significantly different from how human radiologists read images. Radiologists

have been trained to read MRI images and scrutinise for certain reproducible anatomical and contextual patterns [26]. By reading thousands of MRI images over the course of their career, they can obtain remarkable skills to understand images with known artefacts presented [1]. Therefore, imitating this human learning experience using GAN based deep learning model can change the conventional online nonlinear optimisation task into an offline training procedure. In other words, compared to classic CS methods that solved the inverse problem for each new input dataset, our GAN based deep learning model can learn the complex nonlinear mapping between the undersampled k-space (aliased reconstruction) and the desired uncorrupted image (de-aliased reconstruction) offline. Once such optimal mapping is learnt, it can be applied to any new input dataset. Although the training procedure can take a long time to finish (depending on the size of the training dataset and the desired quality of the reconstructed image), the reconstruction for the new input dataset is really fast (average 0.22ms per image on a Nvidia TitanX Pascal GPU) that is suitable for real time applications.

Compared to previous studies based on deep learning (CNNs) for fast CS-MRI, our conditional GAN-based method also incorporated content loss and refinement learning. The refinement learning can stabilise and speed up our training procedure (Figure 4 (b) and (c)), while the perceptual loss can make our reconstruction more realistic. Interestingly, we can observe that our PG method without perceptual loss may yield smaller NMSE and higher PSNR than the PPG and PPGR methods that utilised the perceptual loss when the undersampling rate is $\leq$30% (Table 1). This means quantitatively the PG method has outperformed the PPG and PPGR methods; however, the qualitative visualisation demonstrated that results from both PPG and PPGR are superior to PG results in terms of finer perceptual details and less jagged artefacts (Figure 5). This can be attributed to the fact that when reconstructing a highly undersampled k-space, the PG method without perceptual loss can only find an optimal solution to satisfy the MSE criteria but may not perceptually resemble the real data. More importantly, conventional or recently proposed deep learning-based fast CS-MRI methods solve the inverse problem by assuming the correctness of the 'forward model'. In contrast, our GAN-based method could still perform well without such assumptions imparted by the forward model.

Our method has a similar architecture as [31], which also contains a perceptual loss for solving SR. Essentially, CS-MRI is a more general inverse problem to recover data from undersampled measurements, in which the undersampling pattern is random and noise and artefacts propagation is global due to frequency domain operation (compared to the regular downsampling pattern and local artefacts in SR). Therefore, the CS-MRI is a more challenging problem to solve. On the other hand, our model can be generalised to solve SR. Compared to [31], our PPGR method applied U-Net architecture to reconstruct better image details. In addition, our model is reliable when the raw data is randomly and highly undersampled using the proposed refinement learning. We notice that our current study may have two limitations, but will not influence the final conclusion: (1) we compared to the results in [19] without implementation. This is because the

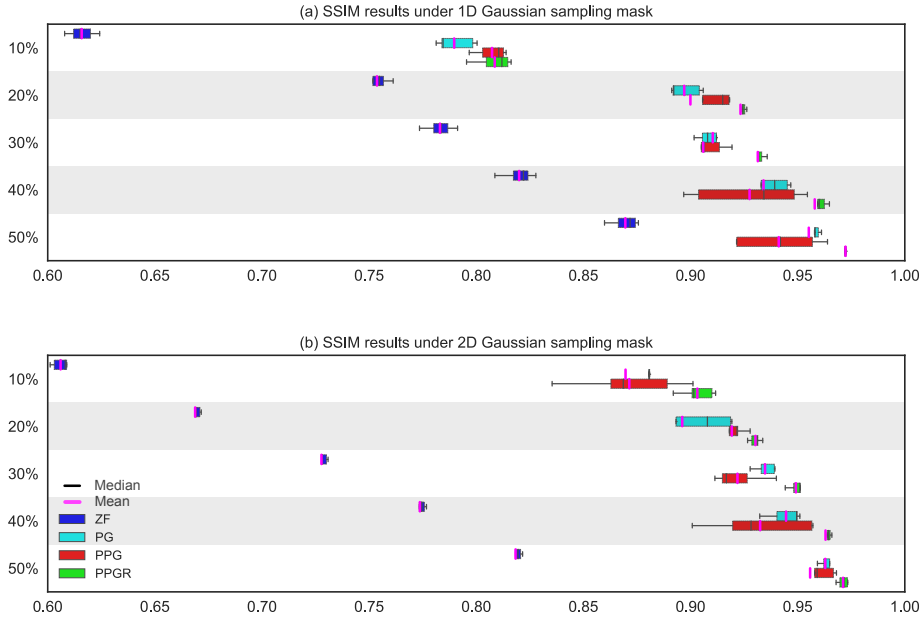

Fig. 2: *Box plots of the SSIM for the reconstruction of the IXI dataset (a) using 1D Gaussian random sampling mask (for only the phase encoding direction) and (b) using 2D Gaussian random sampling mask.*

implementation of those methods is difficult as the fine-tuned hyper-parameters used are not always clearly stated and the methodologies cannot be reproduced exactly. However, we validated our methods using the test cases randomly selected from the same datasets as [19] that can achieve a relatively fair comparison; (2) the frequency domain information, which may provide useful constraints for reconstruction fidelity, is not explicitly considered in our current model. This is now considered as a future working direction.

5 Conclusion

In this study, we proposed a conditional GAN-based deep learning method to solve the de-aliasing for fast CS-MRI. Remarkably, by incorporating the adversarial loss with a content loss that consists of pixel-wise MSE and perceptual VGG loss, our method can achieve promising and realistic MRI reconstruction results. By using the refinement learning, our method is fast and robust even when the k-space raw data is highly undersampled. Convincing simulation based results show promise of our technique to be translational for real MRI acquisition applications.

Table 1: *NMSE and PSNR of the MICCAI dataset using different random undersampling masks and ratios.*

Mask	Method	10%		20%		30%		40%		50%	
		NMSE	PSNR	NMSE	PSNR	NMSE	PSNR	NMSE	PSNR	NMSE	PSNR
Gaussian 1D	ZF	0.3297	27.26	0.1776	33.86	0.1513	35.32	0.1186	37.49	0.0828	40.65
	PG	0.1911	33.16	**0.0949**	**39.29**	**0.0875**	**40.01**	0.0599	43.33	0.0530	44.41
	PPG	0.2321	31.51	0.1004	38.91	**0.0875**	39.99	0.0695	42.31	0.0461	45.66
	PPGR	**0.1901**	**33.24**	0.0958	39.22	0.0906	39.73	**0.0528**	**44.49**	**0.0385**	**47.30**
Gaussian 2D	ZF	0.2374	31.25	0.1797	33.78	0.1315	36.57	0.0981	39.16	0.0708	41.99
	PG	**0.0890**	**39.91**	**0.0781**	**41.03**	**0.0569**	**43.89**	0.0519	44.68	0.0371	47.64
	PPG	0.0925	39.53	0.0878	40.12	0.0612	43.16	0.0555	44.12	0.0403	46.85
	PPGR	0.1029	38.61	0.0804	40.82	0.0586	43.56	**0.0440**	**46.20**	**0.0347**	**48.32**

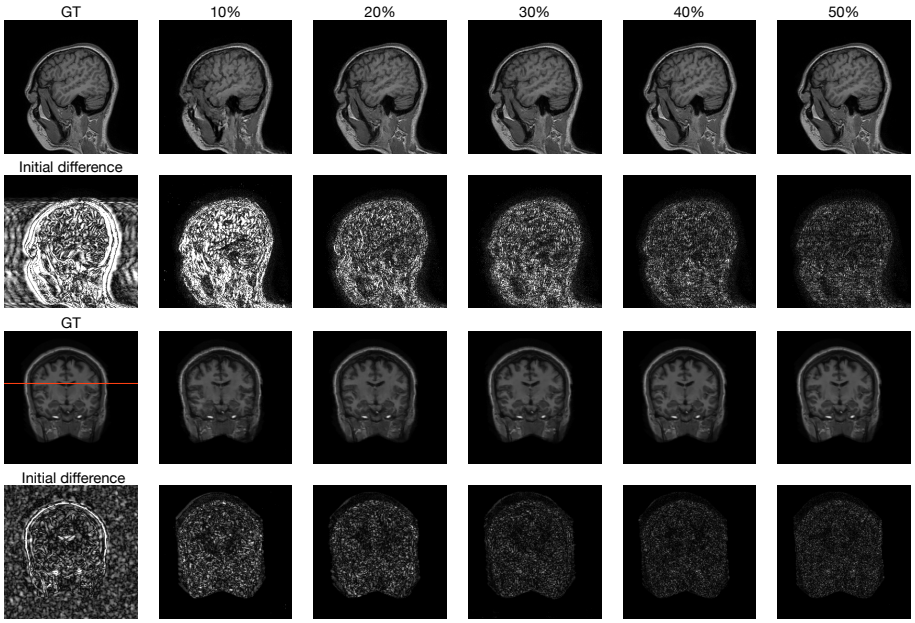

Fig. 3: *Qualitative visualisation of our PPGR de-aliasing for fast CS-MRI.*

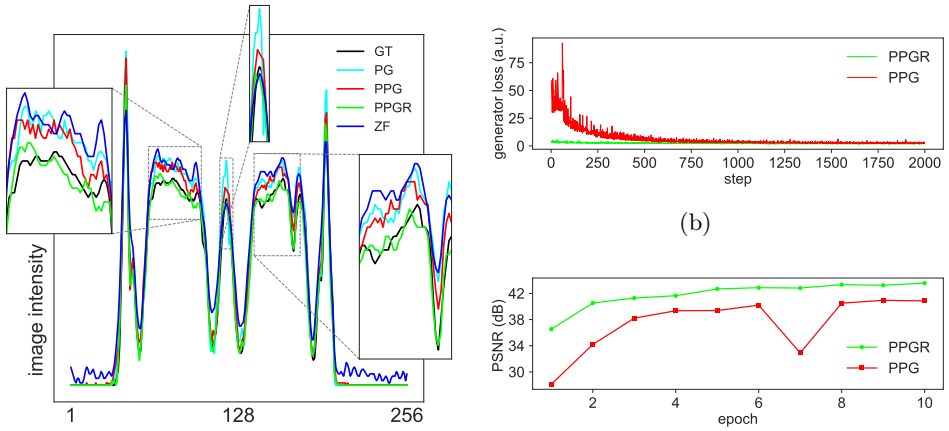

Fig. 4: *(a) Line profile comparison (red horizontal line in the GT image as shown in Figure 3), (b) Convergence speed comparison w.r.t. the generator loss and (c) Stability comparison w.r.t. the PSNR of the MICCAI dataset.*

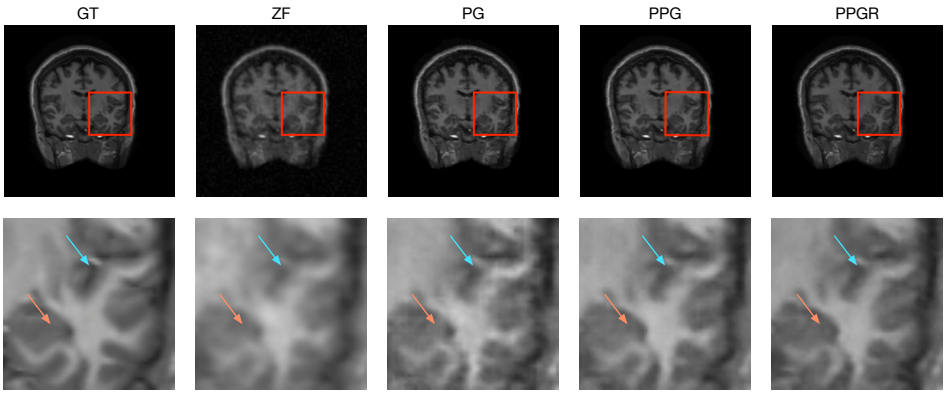

Fig. 5: *Qualitative comparison for our proposed models using 20% 2D Gaussian random undersampling.*

References

[1] K. G. Hollingsworth, "Reducing acquisition time in clinical MRI by data undersampling and compressed sensing reconstruction," *Phys. Med. Biol.*, vol. 60, no. 21, pp. 297–322, 2015.

[2] M. Lustig, D. Donoho, J. Santos, and J. Pauly, "Compressed sensing MRI," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 72–82, 2008.

[3] H. Nyquist, "Certain topics in telegraph transmission theory," *Trans. Am. Inst. Elect. Eng.*, vol. 47, no. 2, pp. 617–644, 1928.

[4] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[5] M. J. Fair, P. D. Gatehouse, E. V. R. DiBella, and D. N. Firmin, "A review of 3D first-pass, whole-heart, myocardial perfusion cardiovascular magnetic resonance," *J. Cardiovasc. Magn. Reson.*, vol. 17, no. 1, p. 68, 2015.

[6] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: the application of compressed sensing for rapid MR imaging," *Magn. Reson. Med.*, vol. 58, no. 6, pp. 1182–1195, 2007.

[7] K. T. Block, M. Uecker, and J. Frahm, "Undersampled radial MRI with multiple coils. Iterative image reconstruction using a total variation constraint.," *Magn. Reson. Med.*, vol. 57, no. 6, pp. 1086–1098, 2007.

[8] J. Yang, Y. Zhang, and W. Yin, "A fast alternating direction method for TVL1-L2 signal reconstruction from partial fourier data," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 288–297, 2010.

[9] F. Knoll, K. Bredies, T. Pock, and R. Stollberger, "Second order total generalized variation (TGV) for MRI," *Magn. Reson. Med.*, vol. 65, no. 2, pp. 480–491, 2011.

[10] M. Hong, Y. Yu, H. Wang, F. Liu, and S. Crozier, "Compressed sensing MRI with singular value decomposition-based sparsity basis," *Phys. Med. Biol.*, vol. 56, no. 19, pp. 6311–6325, 2011.

[11] S. G. Lingala and M. Jacob, "Blind compressive sensing dynamic MRI.," *IEEE Trans. Med. Imag.*, vol. 32, no. 6, pp. 1132–1145, 2013.

[12] Y. Wang and L. Ying, "Undersampled dynamic magnetic resonance imaging using kernel principal component analysis," in *EMBC*, 2014.

[13] X. Qu, D. Guo, B. Ning, Y. Hou, Y. Lin, S. Cai, and Z. Chen, "Undersampled MRI reconstruction with patch-based directional wavelets," *Magn. Reson. Imag.*, vol. 30, no. 7, pp. 964–977, 2012.

[14] Z. Zhu, K. Wahid, P. Babyn, and R. Yang, "Compressed sensing-based MRI reconstruction using complex double-density dual-tree DWT," *Int. J. Biomed. Imaging*, vol. 2013, p. 10, 2013.

[15] Z. Lai, X. Qu, Y. Liu, D. Guo, J. Ye, Z. Zhan, and Z. Chen, "Image reconstruction of compressed sensing MRI using graph-based redundant wavelet transform," *Med. Image Anal.*, vol. 27, pp. 93–104, 2016.

[16] S. Ravishankar and Y. Bresler, "MR image reconstruction from highly undersampled k-space data by dictionary learning," *IEEE Trans. Med. Imag.*, vol. 30, no. 5, pp. 1028–1041, 2011.

[17] J. Caballero, A. N. Price, D. Rueckert, and J. V. Hajnal, "Dictionary learning and time sparsity for dynamic MR data reconstruction," *IEEE Trans. Med. Imag.*, vol. 33, no. 4, pp. 979–994, 2014.

[18] Z. Zhan, J.-F. Cai, D. Guo, Y. Liu, Z. Chen, and X. Qu, "Fast multiclass dictionaries learning with geometrical directions in MRI reconstruction," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 9, pp. 1850–1861, 2016.

[19] Y. Yang, J. Sun, H. Li, and Z. Xu, "Deep ADMM-Net for compressive sensing MRI," in *NIPS*, 2016.

[20] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, 2016.

[21] J. Xie, L. Xu, and E. Chen, "Image denoising and inpainting with deep neural networks," in *NIPS*, 2012.

[22] F. Agostinelli, M. R. Anderson, and H. Lee, "Adaptive multi-column deep neural networks with application to robust image denoising," in *NIPS*, 2013.

[23] O. N. Jaspan, R. Fleysher, and M. L. Lipton, "Compressed sensing MRI: a review of the clinical literature," *Br. J. Radiol.*, vol. 88, no. 1056, pp. 1–12, 2015.

[24] S. Wang, Z. Su, L. Ying, X. Peng, S. Zhu, F. Liang, D. Feng, and D. Liang, "Accelerating magnetic resonance imaging via deep learning," in *ISBI*, 2016.

[25] J. Schlemper, J. Caballero, J. V. Hajnal, A. Price, and D. Rueckert, "A deep cascade of convolutional neural networks for dynamic MR image reconstruction," *arXiv preprint arXiv:1704.02422*, 2017.

[26] K. Hammernik, T. Klatzer, E. Kobler, M. P. Recht, D. K. Sodickson, T. Pock, and F. Knoll, "Learning a variational network for reconstruction of accelerated MRI data," *arXiv preprint arXiv:1704.00447*, 2017.

[27] D. Lee, J. Yoo, and J. C. Ye, "Deep artifact learning for compressed sensing and parallel MRI," *arXiv preprint arXiv:1703.01120*, 2017.

[28] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," in *MICCAI*, 2015.

[29] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *NIPS*, 2014.

[30] M. Mirza and S. Osindero, "Conditional generative adversarial nets," *arXiv preprint arXiv:1411.1784*, 2014.

[31] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network," in *CVPR*, 2017.

[32] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," in *ICLR*, 2016.

[33] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *CVPR*, 2017.

[34] S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," in *ICML*, 2015.

[35] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *ICLR*, 2015.

[36] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, and M. Bernstein, "Imagenet large scale visual recognition challenge," *Int. J. Comput. Vision*, vol. 115, no. 3, pp. 211–252, 2015.

[37] D. Kingma and J. Ba, "Adam: a method for stochastic optimization," in *ICLR*, 2014.

[38] P. Simard, D. Steinkraus, and J. Platt, "Best practices for convolutional neural networks applied to visual document analysis," in *ICDAR*, 2003.

[39] X. Qu, Y. Hou, F. Lam, D. Guo, J. Zhong, and Z. Chen, "Magnetic resonance image reconstruction from undersampled measurements using a patch-based nonlocal operator," *Med. Image Anal.*, vol. 18, no. 6, pp. 843–856, 2014.

[40] E. M. Eksioglu, "Decoupled algorithm for MRI reconstruction using nonlocal block matching model: BM3D-MRI," *J. of Math. Imaging Vis.*, vol. 56, no. 3, pp. 430–440, 2016.