

# PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

## Wavelet improved GAN for MRI reconstruction

Chen, Yutong, Firmin, David, Yang, Guang

Yutong Chen, David Firmin, Guang Yang, "Wavelet improved GAN for MRI reconstruction," Proc. SPIE 11595, Medical Imaging 2021: Physics of Medical Imaging, 1159513 (15 February 2021); doi: 10.1117/12.2581004

**SPIE.**

Event: SPIE Medical Imaging, 2021, Online Only

# Wavelet Improved GAN for MRI reconstruction

Yutong Chen <sup>a,b</sup>, David Firmin <sup>a</sup>, and Guang Yang <sup>a</sup>

<sup>a</sup> National Heart & Lung Institute, Imperial College London, London SW3 6NP, U.K.

<sup>b</sup> Faculty of Biology, Cambridge University, Cambridge CB2 1TN, U.K.

## ABSTRACT

**Background:** Compressed sensing magnetic resonance imaging (CS-MRI) is an important technique of accelerating the acquisition process of magnetic resonance (MR) images by undersampling. It has the potential of reducing MR scanning time and costs, thus minimising patient discomfort.

**Motivation:** One of the successful CS-MRI techniques to recover the original image from undersampled images is generative adversarial network (GAN). However, GAN-based techniques suffer from three key limitations: training instability, slow convergence and input size constraints.

**Method and Result:** In this study, we propose a novel GAN-based CS-MRI technique: WPD-DAGAN (Wavelet Packet Decomposition Improved de-aliaising GAN). We incorporate Wasserstein loss function and a novel structure based on wavelet packet decomposition (WPD) into the de-aliaising GAN (DAGAN) architecture, which is a well established GAN-based CS-MRI technique. We show that the proposed network architecture achieves a significant performance improvement over the state-of-the-art CS-MRI techniques.

**Keywords:** Compressed sensing, Generative adversarial network, Wavelet packet decomposition, MRI

## 1. INTRODUCTION

Magnetic resonance imaging (MRI) is a revolutionary non-invasive medical imaging technique, offering high resolution definition of the structure and function of most body tissues and organs. However, a major limitation of MRI is its slow rate of image acquisition,<sup>1,2</sup> resulting in a prolonged scanning time compared to other imaging modality. This limitation not only increases its scanning cost, but restricts its application in emergency settings. Furthermore, patients have to ‘lie still throughout an acquisition in order not to degrade the quality of the images’.<sup>3</sup> Patients must hold their breath for abdominal/thoracic imaging, which may be difficult for ‘children, obese individuals and those with respiratory compromise’.<sup>3</sup> Hence, the MR scanning process can bring feelings of discomfort and claustrophobia to many patients. Therefore, to reduce the scanning cost and improve patient experience, it is necessary to accelerate the MR image acquisition process.

The reason why MR imaging process is slow is that unlike other imaging techniques, MR images are obtained from the k-space, i.e. the spatial frequency domain. The image domain information is recovered from the k-space data by inverse Fourier transform.<sup>4</sup> The measurements in the k-space are acquired sequentially, i.e. one line by one line per repetition time across the phase encoding direction(s).<sup>3,4</sup> Various MR acceleration techniques focus on taking measurements of multiple lines per repetition time, thus transversing through the k-space at a faster rate.<sup>3</sup> Examples include echo planar imaging,<sup>5</sup> rapid acquisition with relaxation enhancement,<sup>6</sup> and fast low angle shot imaging.<sup>7</sup>

To achieve a higher degree of acceleration, it is possible to sample the k-space only partially, i.e. not obtaining all lines of measurements across the phase encoding direction(s). The original k-space information can be inferred from the undersampled measurement. The result is an acceleration factor inversely proportional to the undersampling ratio. For instance, if 50% of the k-space is sampled, the acceleration factor is 2-fold. Undersampling techniques therefore attempt to bypass the Nyquist-Shannon sampling criteria.<sup>1</sup>

A promising undersampling method is compressed sensing (CS),<sup>1</sup> potentially allowing for more aggressive undersampling and acceleration. The theory of CS is related to the idea of compressing signals for transmission, and

---

Further author information: (Send correspondence to Yutong Chen and Guang Yang)

Y Chen: stcyt1024@outlook.com and G Yang: g.yang@imperial.ac.uk

decompressing them afterwards,<sup>8</sup> as applied to JPEG, JPEG2000, MPEG, and MP3 standards.<sup>3</sup> CS assumes that if undersampled signals or images can be compressed accurately, they could be decompressed or reconstructed accurately.<sup>9</sup> Hence, CS imposes 3 criteria to the reconstruction process: 1) The signal or the image must be compressible. In other words, the image must be sparse, i.e. the majority of its pixel values being zeros, either in itself or in a suitable transformation domain, e.g. wavelet or frequency domain. 2) The undersampling pattern should be incoherent, such as being random, to prevent aliasing artifacts. 3) The reconstruction algorithm should be non-linear. Under these 3 criteria, it is possible to reconstruct the original image from its undersampled measurements.

Earlier work on applying CS as an MR acceleration method involves iterative non-linear optimisation algorithms, which enforce sparsity and reconstruction fidelity. Examples include total variation (TV),<sup>1</sup> dictionary learning such as DLMRI,<sup>10</sup> RecPF,<sup>11</sup> and BM3D.<sup>12</sup> There are however 4 main problems associated with these techniques:

1. Iterative optimisation algorithm is time consuming<sup>3,13</sup>
2. These algorithms tend to cause over-generalisation, leading to artificially smoothed appearance.<sup>3</sup>
3. The reconstruction results show blocky artifacts.<sup>14–16</sup>
4. They reconstruct each image as an isolated case, and thus fail to take into account of expected anatomical features in MR images to enhance performance.<sup>17</sup>

Most of the above limitations of the earlier CS methods are however effectively addressed by deep learning based methods. A deep learning model consists of multiple layers of nodes. The weights of the node connections between layers are optimised to learn the mapping from the undersampled MR images to their corresponding fully sampled ones. This process of weight optimisation is referred to as training the model. Once trained, the model can reconstruct original images from their undersampled measurements. Deep learning based CS methods have been shown to consistently outperform non-deep learning based ones, in terms of reconstruction accuracy, speed and visual quality.<sup>2,18–23</sup>

An important type of deep learning based CS-MRI reconstruction method is generative adversarial network (GAN). In the context of CS-MRI, GAN involves training a generator to reconstruct the original image from its undersampled measurements and a discriminator to output the probability of whether the generated image resembles the original, i.e. fully sampled ones. The discriminator output in turn modifies the learning of the generator.<sup>24</sup> This results in the generator producing photo-realistic images.<sup>25</sup> GAN-based methods<sup>2,18</sup> outperform the non-GAN based deep learning method—deep ADMM-net—both in reconstruction accuracy and speed. One GAN-based method<sup>19</sup> also claims to produce fewer blurry and aliasing artifacts than the non-deep learning-based ones. Hence, GAN-based methods can potentially achieve state-of-the-art CS-MRI reconstruction result.

However, there are several limitations associated with GAN-based methods:

1. The training history is unstable with the possibility of mode collapse.<sup>26</sup> GAN also suffers from vanishing gradient<sup>27</sup> and the problem of balancing the training between the discriminator and generator.<sup>28</sup>
2. GAN suffers from local minima and non-convergence.<sup>28</sup>
3. The GAN-based CS-MRI methods proposed thus far in the literature<sup>2,18–20</sup> can only reconstruct MR images of a particular size, e.g. 256x256 pixels. To reconstruct MR images of a different size, it would be necessary to resize them by interpolation, which can introduce deformity to the images.<sup>29</sup> This input size constraint may limit the applicability of GAN-based CS-MRI methods and the realm of training datasets.

In order to address the above limitations of GAN, we propose a novel GAN-based CS-MRI technique: WPD-DAGAN. This model is developed upon de-aliasing GAN (DAGAN),<sup>11</sup> with the incorporation of wavelet packet decomposition (WPD) into the DAGAN skeleton and with the removal of the input size constraint. In the following section, we will demonstrate how and why this model achieves a superior reconstruction performance.

## 2. METHOD

### 2.1 GAN

Before describing our proposed method, it is necessary to outline the principles of GAN, in order to understand the success of GAN-based CS-MRI techniques and the foundation of our work. GAN consists of a generator  $G$ , whose parameters or weights are denoted as  $\theta_G$ , and a discriminator  $D$ , with parameters  $\theta_D$ . The discriminator is trained to label the original, fully sampled MR images ( $x$ ) as real, and the reconstruction from the undersampled images ( $z$ ) by the generator as fake, i.e. not being real. Conversely, the generator is trained to achieve the opposite ends: the discriminator cannot tell whether the reconstructed images are real or fake, that is, it fails to distinguish the reconstructed images from the fully sampled ones. The entire training process can be described mathematically as a minimax game, using the following loss function:

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

in which  $p_{data}$  is the distribution of the fully sampled images and  $p_z$  is that of the undersampled images. It can be shown that with an optimised discriminator, the generator minimises the Jensen-Shannon divergence between the distribution of the original and reconstructed images. In other words, GAN models can be seen as minimising the difference in distributions between the fully sampled and reconstructed images.

### 2.2 Proposed loss function

While there is a theoretical justification of the effectiveness of GAN, in practice, GAN suffers from non-convergence and training fluctuations as discussed in the introduction section. These 2 problems are claimed to be resolved by training the model with Wasserstein loss function,<sup>28</sup> which we will use in our model. Instead of minimising JSD, Arjovsky<sup>28</sup> proposed to minimise the Wasserstein distance between the generator and discriminator. It is claimed to achieve a more stable training history:

$$\mathcal{L}(\theta_D, \theta_G) = \min_{\theta_G} \max_{\theta_D} \mathbb{E}_{x \sim p_{data}(x)} [D(x)] - \mathbb{E}_{z \sim p_z(z)} [D(G(z))] \quad (2)$$

This loss function minimises the Wasserstein distance provided that the discriminator function  $D$  is 1-Lipschitz:

$$\|D\|_L \leq 1$$

This 1-Lipschitz constraint can be enforced by weight clipping.<sup>28</sup> To simplify Equation 2, we can write separate loss functions for the discriminator and generator, which we name as  $L_D$  and  $L_{adv}$  respectively:

$$\mathcal{L}_D(\theta_D) = \max_{\theta_D} \mathbb{E}_{x \sim p_{data}(x)} [D(x)] - \mathbb{E}_{z \sim p_z(z)} [D(G(z))] \quad (3)$$

$$\mathcal{L}_{adv}(\theta_G) = \min_{\theta_G} - \mathbb{E}_{z \sim p_z(z)} [D(G(z))] \quad (4)$$

In addition to  $\mathcal{L}_{adv}$ , we add the normalised root mean squared error (NMSE) loss  $\mathcal{L}_{NMSE}$  and frequency domain loss ( $\mathcal{L}_{freq}$ ) to the total generator loss.  $\mathcal{L}_{NMSE}$  and  $\mathcal{L}_{freq}$  are expected to minimise the difference between fully sampled and reconstructed images in the spatial and frequency domains respectively. They are defined as:

$$\mathcal{L}_{NMSE}(\theta_G) = \sqrt{\frac{\|x - G(z)\|_2^2}{\|x\|_2^2}} \quad (5)$$

$$\mathcal{L}_{freq}(\theta_G) = \|\mathcal{F}\{x\} - \mathcal{F}\{G(z)\}\|_2^2 \quad (6)$$

in which  $\mathcal{F}$  is the Fourier transform operator to obtain the frequency, or the k-space, representation of an image. The final loss function for the generator in our model is:

$$\mathcal{L}_G(\theta_G) = \mathcal{L}_{adv} + \alpha \mathcal{L}_{NMSE} + \beta \mathcal{L}_{freq} \quad (7)$$

in which  $\alpha$  and  $\beta$  are weights to adjust the relative contributions of the three loss components to the total loss.

### 2.3 Proposed discriminator architecture

In the previous section, we described our rationale of reducing the non-convergence and instability of GAN-based CS-MRI method by using Wasserstein loss function. In this section, we will outline how we modify the architecture of GAN to address the problem of input size constraint.

The main factor imposing the input size constraint is the flattening layer of the discriminator used in DAGAN. Consider the DAGAN discriminator in Figure 1(a) processing an MR image of  $256 \times 256$  pixels. The image, after passing 6 convolutional operations, is reduced in size by  $2^6 = 64$  times, becoming  $4 \times 4$  pixels in size. When it reaches the ‘concat, 1024’ layer, it has 1024 channels. Its dimension at this layer can be described as a  $4 \times 4 \times 1024$  dimension tensor. The next flattening layer unravels the tensor of the previous layer ‘concat, 1024’ into a column vector of 8192 units. This can be accepted by the subsequent dense layer, which possess 8192 input nodes. However, if the same model is supplied with an image of  $128 \times 128$  pixels. This image will be processed into a  $2 \times 2 \times 1024$  dimension tensor, unravelled into a 4096-unit column vector, which contains too few input nodes to be accepted by the 8192-unit dense layer. This example illustrates that the input size constraint of the DAGAN discriminator.

To remove such a constraint, we replace the flattening operation with a global average pooling (GAP) layer.<sup>29</sup> GAP computes the mean value of all pixels per channel of the input, regardless of the pixel number of image size. For instance, the input to the GAP layer in our model has 1024 channels. The output will be a vector of 1024 units regardless of the dimension of the input tensor from ‘concat, 1024’. Thence, we place a dense layer of 1024 units followed by sigmoid activation. Our design of the discriminator thus removes the input size constraint significantly. It is to our knowledge the first attempt to do so for GAN-based CS-MRI techniques.

### 2.4 Proposed generator structure

To improve training stability and facilitate model convergence, we introduce a novel generator architecture. The skeleton is based on Unet,<sup>30</sup> but we incorporate wavelet package decomposition operations (WPD), a form of wavelet transform. Before describing our motive of using WPD and our proposed architecture in details, we will consider the success of wavelet representation as a feature extractor in general.

Wavelet transform is a useful feature extractor,<sup>31,32</sup> as this representation contains both spatial and frequency features.<sup>32</sup> There has been growing research efforts to investigate how wavelet transform operations may be incorporated into deep learning models, for image classification,<sup>33</sup> image super-resolution,<sup>25,34,35</sup> and image dehazing.<sup>36</sup> A possible reason for this increased attention is suggested by Fujieda (2018). ‘Convolutional operations are good at capturing spatial features while wavelet transform can capture scale invariant features based on spectral information. It is thus preferable to consider both the spatial and spectral information within a single model’.<sup>33</sup>

However, there are different types of wavelet transform. Most of the aforementioned studies used discrete wavelet transform (DWT), except Deng (2019).<sup>25</sup> None of the studies used WPD, which we propose in the model. Both DWT and WPD involves the convolution of the image with different wavelet filters. The key difference between them is that in WPD, wavelet transform is performed on every subband from the previous level. In contrast, in DWT, only the LL (low frequency) subband is transformed. The consequence of this difference is that low level DWT coefficients contain only high frequency features whereas each level of WPD coefficients contain both high and low frequency features. We hypothesise this property of WPD can potentially accelerate feature learning. Therefore, it may offer faster convergence and higher training stability. To our knowledge, this is the first attempt to incorporate WPD-based structures into a deep learning model.

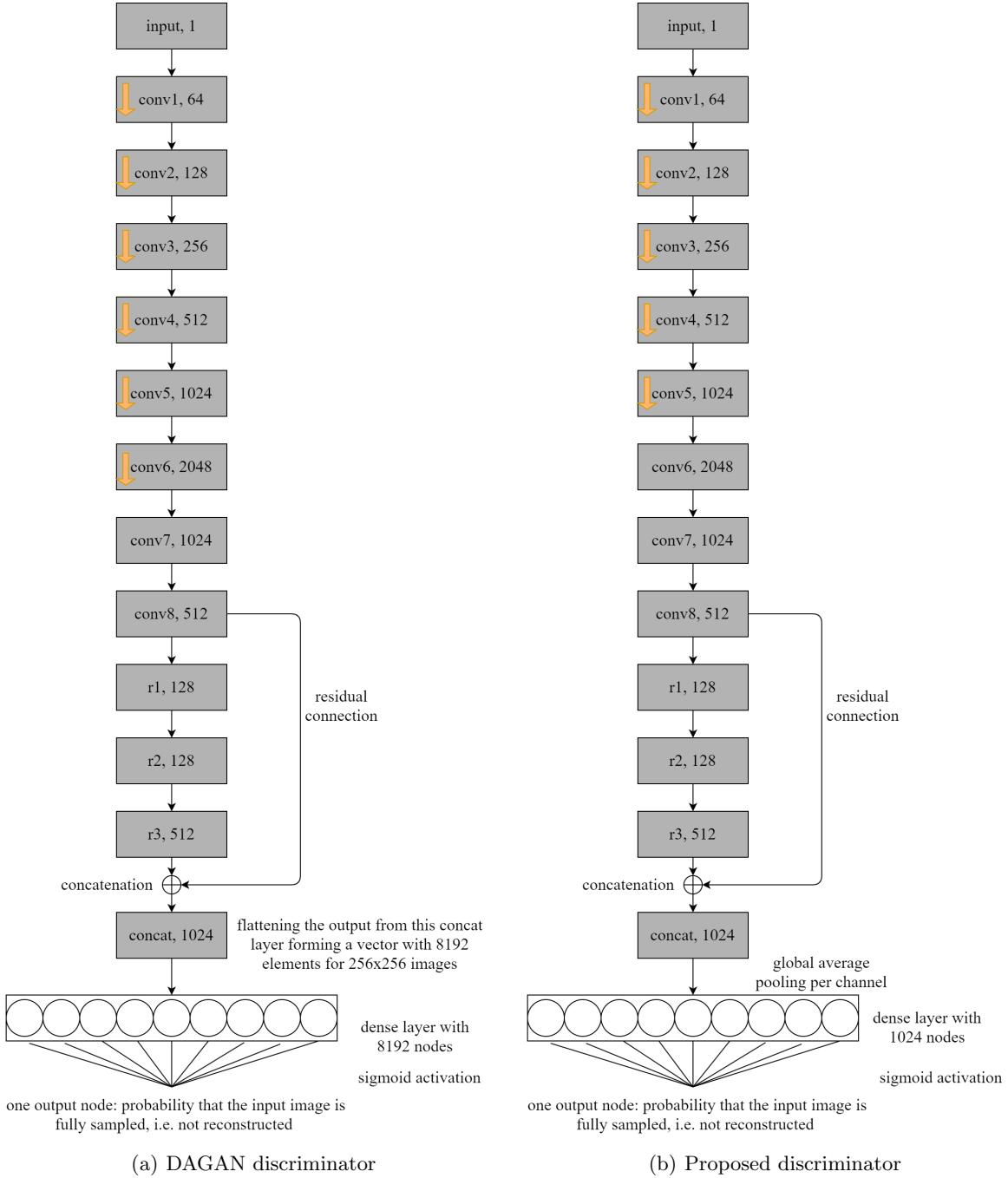


Figure 1. The architecture of the discriminators used in DAGAN<sup>2</sup> 1(a) and in this study 1(b). In both figures, the output of each layer of the discriminator is denoted as ‘name, channel number’. For example, ‘conv1, 64’ means that the output of this layer has 64 channels. The yellow downward pointing arrow denotes the presence of a convolutional operation with a stride of two. No arrows indicate convolutional operation with a stride of one. A black arrow connecting two successive convolutional layers indicates a convolutional operation followed by a batch normalisation layer and a leaky relu activation with a negative slope of 0.2.

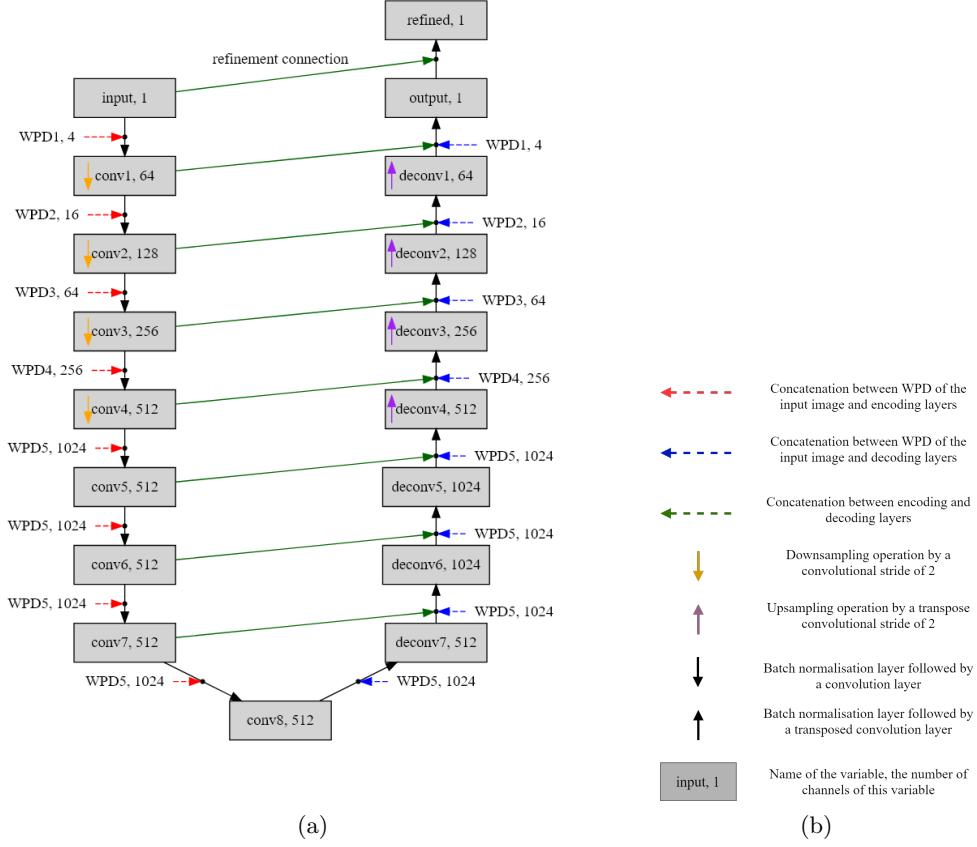


Figure 2. (a): Proposed generator structure. WPD1 means the first level WPD coefficients, i.e. WPD on the input image. WPD2 means the second level coefficients, i.e. WPD on WPD1 coefficients. (b): Legend for (a)

## 2.5 Evaluation

### 2.5.1 Dataset

To evaluate the performance of WPD-DAGAN, we use the MICCAI 2013 grand challenge dataset \*. The dataset consists of  $T_1$  weighted coronal brain sections. We use 14963 images for training, 6277 for validation, and 9901 for testing. Unless otherwise stated, we present the performance on the test set in this paper. All images are  $256 \times 256$  pixels in sizes and pixel values range from 0 to 1.

To test the model on reconstructing images of different sizes— $64 \times 128$ ,  $128 \times 256$ ,  $256 \times 256$ —in our experiment, we crop the central location of the original images, if the required size is smaller than  $256 \times 256$ . We only crop one patch per image, so that the number of images for training, validation and testing is the same for images of different sizes. This ensures the fairness of comparing performance on images of different sizes. Furthermore, if a model is trained with images of a particular size, it is tested with images of the same size.

### 2.5.2 Undersampling mask

Random undersampling masks are generated using the source code <sup>†</sup> from Lustig (2007).<sup>1</sup> We use Gaussian masks with 1D undersampling pattern, which are typically used for CS of 2D MRI slices. This is because 2D MRI involves sampling the frequency encoding lines across one phase encoding axis. Acquisition of samples along the frequency encoding lines is rapid, and thus does not require undersampling.<sup>3</sup> Only the phase encoding direction requires sampling, hence a 1D undersampling pattern is applied. We test 3 undersampling ratios: 10%, 30%, and 50%, corresponding to acceleration factors of 10X, 3.33X and 2X respectively.

\*<https://my.vanderbilt.edu/masi/workshops/>

<sup>†</sup><https://people.eecs.berkeley.edu/~mlustig/Software.html>

### 2.5.3 Comparison study

To establish the validity of our approach, we compare it with zero filling (ZF) method, i.e. filling the k-space values that are not sampled with zeros. We also run non-deep learning-based CS techniques: total variation (TV)<sup>1</sup> and DLMRI.<sup>10</sup> Furthermore, we compare our results to DAGAN.<sup>2</sup> To investigate whether it is our WPD architecture or the Wasserstein loss function or both that affect model performance, we run DAGAN with discriminator loss ( $\mathcal{L}_D$ ) and the adversarial loss ( $\mathcal{L}_{adv}$ ) replaced with Wasserstein loss, as in Equation 3 and 4 respectively. We call this model WDAGAN.

All deep learning models in this study are implemented in pytorch <sup>‡</sup> in python. They are trained on one Nvidia Titan RTX GPU. All non-deep learning models, i.e. TV and DLMRI, are trained on 6 Intel(R) Core(TM) i5-8305G CPUs. TV <sup>§</sup> and DLMRI <sup>¶</sup> are implemented by directly calling their respective matlab source code. The evaluation metrics for model performance on reconstruction of undersampled images are structural similarity index measure (SSIM) and peak signal to noise ratio (PSNR).

## 3. RESULTS

### 3.1 Reconstruction Accuracy

Method	10%		30%		50%	
	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR
ZF	0.620±0.099	25.220±2.874	0.752±0.073	32.141±3.761	0.835±0.058	37.470±4.134
TV	0.696±0.082	25.711±2.983	0.908±0.033	34.499±3.822	0.984±0.009	41.938±3.705
DLMRI	0.644±0.102	23.265±4.233	0.896±0.044	32.796±5.854	0.953±0.021	39.978±6.203
DAGAN	0.891±0.051	30.271±3.461	0.968±0.017	37.841±3.768	0.991±0.006	44.616±4.450
WDAGAN	0.920±0.043	31.688±3.959	0.974±0.016	38.836±4.149	0.992±0.006	45.042±4.714
<b>WPD-DAGAN</b>	<b>0.949±0.025</b>	<b>34.276±3.525</b>	<b>0.987±0.009</b>	<b>42.000±3.980</b>	<b>0.996±0.003</b>	<b>48.611±4.475</b>

Table 1. Reconstruction results of different models for image size  $256 \times 256$  pixels. ‘WDAGAN’ means the original DAGAN model with the incorporation of Wasserstein loss function. This set of experiments shows that while adding the Wasserstein loss function only slightly improves model performance, the WPD-based generator further enhances the model performance.

From table 1, it is clear that our WPD-DAGAN consistently outperforms other methods in terms of reconstruction SSIM and PSNR. The qualitative improvement is also evident, as shown in Figure 3. In this section, we will present the analysis of the possible reasons of this performance improvement

Comparing DAGAN and WDAGAN, the improvement of reconstruction accuracy by Wasserstein loss is insignificant, especially at high sampling ratio (50%). However, comparing WPD-DAGAN with WDAGAN, WPD-based architecture appears to significantly improve the reconstruction accuracy. This improvement is also evident qualitatively. For instance, consider the orange boxes in Figure 3(e) and 3(f). Comparing with the corresponding region in the reconstructed image, the bright vertically oriented line just inferior to the brain<sup>37,38</sup> is the internal carotid artery. In the ‘Difference Image’, the reconstruction error of this artery is high in WDAGAN (3 (e)), but slightly reduced in WPD-DAGAN (3 (f)). The error in 3(f) appears less structured and more noise-like. This indicates a better qualitative performance of WPD-DAGAN. From the magnified view of the red box (corpus callosum and white matter tract) and the green box (medial temporal cortex), the reconstruction error of WPD-DAGAN is almost nil. In contrast, a higher error is noted in the image reconstructed by WDAGAN.

Take together, WPD-DAGAN can reconstruct MR images more accurately than other CS-MRI techniques tested in this study. The main contributor to this performance improvement appears to be the WPD architecture, while the Wasserstein loss appears less important

<sup>‡</sup><https://pytorch.org/>

<sup>§</sup><https://people.eecs.berkeley.edu/~mlustig/Software.html>

<sup>¶</sup><http://www.ifp.illinois.edu/~yoram/DLMRI-Lab/DLMRI.html>

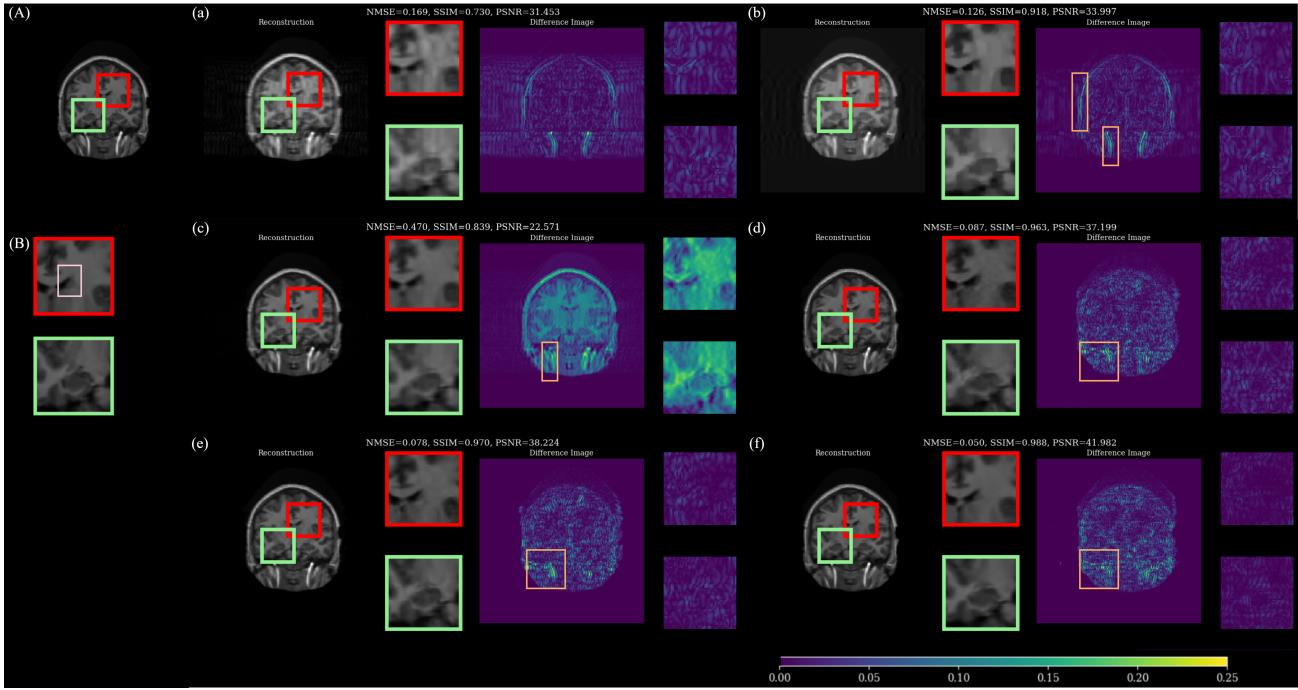


Figure 3. Typical example reconstruction: (A): fully sampled ground truth; (B): magnified regions in (A); (a): ZF; (b): total variation (TV); (c): DLRMI; (d): DAGAN; (e): WDAGAN; (f): WPD-DAGAN. In the sub-figures, we show the reconstructed image at 30% undersampling ratio on the left hand side. On the right hand side, we show the absolute difference between the the reconstructed and original image as ‘Difference Image’. This absolute difference is shown with a pseudocoloring scheme, according to the colorbar at the bottom of the figure. In each sub-figure, the **red** boxes contain a magnified view of the corpus callosum and adjacent white matter region, which in general are well reconstructed across different methods. The **green** boxes contain a magnified view of the medial temporal cortex including the hippocampus. There are discrepancies in the reconstruction performance between different algorithms. Our WPD-DAGAN reconstructs this region with higher fidelity and perceptual quality than DAGAN and TV. The **pink** box in (B) encompass the caudate nucleus. Reconstruction of this structure has low perceptual quality across all the methods. The **orange** boxes denote regions where reconstruction error is high. In (b), the upper orange box is part of the cranium, the lower box is the right internal carotid artery. In (d), (e) and (f), the orange boxes encompass the right internal carotid artery.

### 3.2 Performance on other image sizes

A key breakthrough of WPD-DAGAN over DAGAN is the removal of the constraint of the input image sizes using global average pooling without altering the underlying model architecture. This feature allows the same WPD-DAGAN model to process a wide range of image sizes that may be encountered in MR reconstruction, broadening the range of the training samples. For a fair comparison, we have implemented global average pooling in the original DAGAN model to remove its input size constraint. With this improvement, DAGAN still displays significantly worse performance compared to WPD-DAGAN in images of  $64 \times 128$  and  $128 \times 256$  pixels, as shown in Table 2 and Table 3. Even though the performance improvement at 10% undersampling for  $64 \times 128$  is smaller, the improvement overall is comparable to that in  $256 \times 256$  images. This suggests that not only does WPD-DAGAN have no input size constraint, but also that the WPD operations and Wasserstein loss functions in WPD-DAGAN underlie the performance enhancement over a range of image sizes.

## 4. CONCLUSION

We propose a novel GAN-based CS-MRI reconstruction method—WPD-DAGAN—and effectively remove the input size constraints of GAN. To accelerate feature learning, we incorporate Wasserstein loss function and WPD connections. The model has demonstrated superior reconstruction accuracy to other CS-MRI techniques, while maintaining its performance on images of different sizes. Therefore, our contribution is to introduce a potentially promising GAN-based CS-MRI reconstruction method.

Method	10%		30%		50%	
	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR
ZF	0.500±0.082	19.074±2.544	0.811±0.058	27.300±3.642	0.883±0.053	30.694±3.674
TV	0.514±0.082	19.374±2.609	0.880±0.041	29.208±3.714	0.955±0.018	34.067±3.461
DAGAN	0.733±0.080	24.202±3.068	0.887±0.035	29.910±3.201	0.975±0.011	37.340±3.539
WDAGAN	0.763±0.079	24.905±3.082	0.940±0.023	32.843±3.125	0.976±0.012	37.567±3.365
<b>WPD-DAGAN</b>	<b>0.778±0.068</b>	<b>25.586±3.089</b>	<b>0.950±0.021</b>	<b>33.846±3.365</b>	<b>0.981±0.010</b>	<b>38.707±3.577</b>

Table 2. Reconstruction results of different models for image size  $64 \times 128$  pixels.

Method	10%		30%		50%	
	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR
ZF	0.394±0.092	22.293±2.394	0.607±0.090	29.792±3.738	0.709±0.087	33.498±4.049
TV	0.555±0.087	23.089±2.598	0.889±0.040	33.727±3.804	0.969±0.015	38.508±3.621
DAGAN	0.886±0.046	30.468±3.487	0.963±0.018	37.178±3.864	0.979±0.010	40.587±3.913
WDAGAN	<b>0.915±0.036</b>	31.586±3.590	0.978±0.011	39.727±3.961	0.988±0.006	42.522±4.064
<b>WPD-DAGAN</b>	0.913±0.039	<b>32.136±3.660</b>	<b>0.981±0.011</b>	<b>40.668±4.065</b>	<b>0.991±0.005</b>	<b>44.138±4.246</b>

Table 3. Reconstruction results of different models for image size  $128 \times 256$  pixels.

## 5. ACKNOWLEDGEMENT

This work was supported in part by the British Heart Foundation (Project Number: PG/16/78/32402), in part by the Hangzhou Economic and Technological Development Area Strategical Grant [Imperial Institute of Advanced Technology], in part by the European Research Council Innovative Medicines Initiative on Development of Therapeutics and Diagnostics Combatting Coronavirus Infections Award 'DRAGON: rapiD and secuRe AI imaging based diaGnosis, stratification, fOLLOW-up, and preparedness for coronavirus paNdemics' [H2020-JTI-IMI2 101005122], and in part by the AI for Health Imaging Award 'CHAIMELEON: Accelerating the Lab to Market Transition of AI Tools for Cancer Management' [H2020-SC1-FA-DTS-2019-1 952172].

## REFERENCES

- [1] Lustig, M., Donoho, D., and Pauly, J. M., "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine* **58**, 1182–1195 (dec 2007).
- [2] Yang, G., Yu, S., Dong, H., Slabaugh, G., Dragotti, P. L., Ye, X., Liu, F., Arridge, S., Keegan, J., Guo, Y., and Firmin, D., "DAGAN: Deep De-Aliasing Generative Adversarial Networks for Fast Compressed Sensing MRI Reconstruction," *IEEE Transactions on Medical Imaging* **37**, 1310–1321 (jun 2018).
- [3] Hollingsworth, K. G., "Reducing acquisition time in clinical MRI by data undersampling and compressed sensing reconstruction," *Phys. Med. Biol.* , R297–322 (2015).
- [4] Suetens, P., [Fundamentals of Medical Imaging], Cambridge University Press, 2nd ed. (2009).
- [5] Mansfield, P., "Multi-planar image formation using NMR spin echoes," *Journal of Physics C: Solid State Physics* (1977).
- [6] Hennig, J., Nauerth, A., and Friedburg, H., "RARE imaging: A fast imaging method for clinical MR," *Magnetic Resonance in Medicine* **3**(6), 823–833 (1986).
- [7] Haase, A., Frahm, J., Matthaei, D., and Merboldt, K.-D., "FLASH Imaging. Rapid NMR Imaging Using Low Flip-Angle Pulses," tech. rep. (1986).
- [8] Zisselman, E., Adler, A., and Elad, M., "Compressed Learning for Image Classification: A Deep Neural Network Approach," in [Handbook of Numerical Analysis], **19**, 3–17, Elsevier B.V. (jan 2018).
- [9] Fair, M. J., Gatehouse, P. D., DiBella, E. V., and Firmin, D. N., "A review of 3D first-pass, whole-heart, myocardial perfusion cardiovascular magnetic resonance," (aug 2015).
- [10] Ravishankar, S. and Bresler, Y., "MR image reconstruction from highly undersampled k-space data by dictionary learning," *IEEE Transactions on Medical Imaging* **30**, 1028–1041 (may 2011).
- [11] Yang, J., Zhang, Y., and Yin, W., "A fast alternating direction method for TVL1-L2 signal reconstruction from partial Fourier data," *IEEE Journal on Selected Topics in Signal Processing* **4**, 288–297 (apr 2010).
- [12] Eksioglu, E. M., "Decoupled Algorithm for MRI Reconstruction Using Nonlocal Block Matching Model: BM3D-MRI," *Journal of Mathematical Imaging and Vision* **56**, 430–440 (nov 2016).
- [13] Hu, Y., Ongie, G., Ramani, S., and Jacob, M., "Generalized higher degree total variation (HDTV) regularization," *IEEE Transactions on Image Processing* **23**(6), 2423–2435 (2014).
- [14] Liu, Y., Cai, J.-F., Zhan, Z., Guo, D., Ye, J., Chen, Z., and Qu, X., "Balanced Sparse Model for Tight Frames in Compressed Sensing Magnetic Resonance Imaging," *PLOS ONE* **10**, e0119584 (apr 2015).
- [15] Kayvanrad, M. H., McLeod, A. J., Baxter, J. S., McKenzie, C. A., and Peters, T. M., "Stationary wavelet transform for under-sampled MRI reconstruction," *Magnetic Resonance Imaging* **32**, 1353–1364 (dec 2014).
- [16] Guerquin-Kern, M., Haberlin, M., Pruessmann, K. P., and Unser, M., "A fast wavelet-based reconstruction method for magnetic resonance imaging," *IEEE Transactions on Medical Imaging* **30**, 1649–1660 (sep 2011).
- [17] Hammernik, K., Klatzer, T., Kobler, E., Recht, M. P., Sodickson, D. K., Pock, T., and Knoll, F., "Learning a variational network for reconstruction of accelerated MRI data," *Magnetic Resonance in Medicine* **79**, 3055–3071 (jun 2018).
- [18] Quan, T. M., Nguyen-Duc, T., and Jeong, W. K., "Compressed Sensing MRI Reconstruction Using a Generative Adversarial Network With a Cyclic Loss," *IEEE Transactions on Medical Imaging* **37**, 1488–1497 (jun 2018).
- [19] Mardani, M., Gong, E., Cheng, J. Y., Vasanawala, S. S., Zaharchuk, G., Xing, L., and Pauly, J. M., "Deep generative adversarial neural networks for compressive sensing MRI," *IEEE Transactions on Medical Imaging* **38**, 167–179 (jan 2019).
- [20] An, H. and Zhang, Y. J., "A Structural Oriented Training Method for GAN Based Fast Compressed Sensing MRI," in [Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)], **11902 LNCS**, 483–494, Springer (aug 2019).
- [21] Schlemper, J., Caballero, J., Hajnal, J. V., Price, A. N., and Rueckert, D., "A Deep Cascade of Convolutional Neural Networks for Dynamic MR Image Reconstruction," *IEEE Transactions on Medical Imaging* **37**, 491–503 (feb 2018).
- [22] Huang, Q., Yang, D., Wu, P., Qu, H., Yi, J., and Metaxas, D., "MRI reconstruction via cascaded channel-wise attention network," in [Proceedings - International Symposium on Biomedical Imaging], **2019-April**, 1622–1626, IEEE Computer Society (apr 2019).

- [23] Eo, T., Jun, Y., Kim, T., Jang, J., Lee, H., and Hwang, D., "KIKI-net: cross-domain convolutional neural networks for reconstructing undersampled magnetic resonance images," *Magnetic Resonance in Medicine* **80**, 2188–2201 (nov 2018).
- [24] Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y., "Generative Adversarial Nets," tech. rep. (2014).
- [25] Deng, X., Yang, R., Zurich, E., Xu, M., and Dragotti, P. L., "Wavelet Domain Style Transfer for an Effective Perception-distortion Tradeoff in Single Image Super-Resolution," tech. rep. (2019).
- [26] Chen, Z. and Tong, Y., "Face Super-Resolution Through Wasserstein GANs," tech. rep. (2017).
- [27] Wiatrak, M., Albrecht, S. V., and Nystrom, A., "Stabilizing GANs: A Survey Stabilizing Generative Adversarial Networks: A Survey," tech. rep. (2020).
- [28] Arjovsky, M., Chintala, S., and Bottou, L., "Wasserstein GAN," tech. rep. (2017).
- [29] Kendrick, C., Gillespie, D., and Yap, M. H., "Anysize GAN: A solution to the image-warping problem," (mar 2020).
- [30] Ronneberger, O., Fischer, P., and 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)], **9351**, 234–241, Springer Verlag (may 2015).
- [31] Unser, M., "Texture Classification and Segmentation Using Wavelet Frames," *IEEE Transactions on Image Processing* **4**(11), 1549–1560 (1995).
- [32] Mallat, S. G., "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11**(7), 674–693 (1989).
- [33] Fujieda, S., Takayama, K., and Hachisuka, T., "Wavelet Convolutional Neural Networks," tech. rep. (2018).
- [34] Huang, H., He, R., Sun, Z., and Tan, T., "Wavelet-SRNet: A Wavelet-based CNN for Multi-scale Face Super Resolution," tech. rep. (2017).
- [35] Liu, P., Zhang, H., Zhang, K., Lin, L., and Zuo, W., "Multi-level Wavelet-CNN for Image Restoration," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops* **2018-June**, 886–895 (may 2018).
- [36] Yang, H.-H., Huck Yang, C.-H., and James Tsai, Y.-C., "Y-NET: Multi-scale feature aggregation network with wavelet structure similarity loss function for single image dehazing," tech. rep. (2020).
- [37] Drexler, I., Traenka, C., Von Hessling, A., and Gensicke, H., "Internal carotid artery dissection and asymmetrical facial flushing: The harlequin sign," *Stroke* **45** (may 2014).
- [38] Ben Hassen, W., MacHet, A., Edjlali-Goujon, M., Legrand, L., Ladoux, A., Mellerio, C., Bodiguel, E., Gobin-Metteil, M. P., Trystram, D., Rodriguez-Regent, C., Mas, J. L., Plat, M., Oppenheim, C., Meder, J. F., and Naggara, O., "Imaging of cervical artery dissection," (dec 2014).