

---

# CS-MRI Reconstruction Using an Improved GAN with Dilated-Residual Networks and Channel Attention Mechanism

---

Xia Li , Hui Zhang , Hao Yang , [Tie-Qiang Li](#) \*

Posted Date: 11 July 2023

doi: 10.20944/preprints202307.0666.v1

Keywords: Compressed sensing MRI; GAN; U-net; dilated-residual blocks; channel attention mechanism



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

# CS-MRI Reconstruction Using an Improved GAN with Dilated-Residual Networks and Channel Attention Mechanism

Xia Li <sup>1</sup>, Hui Zhang <sup>1</sup>, Hao Yang <sup>1</sup> and Tie-Qiang Li <sup>1,2,3\*</sup>

<sup>1</sup> College of Information Engineering, China Jiliang University, Hangzhou, China

<sup>2</sup> Department of Clinical Science, Intervention, and Technology, Karolinska Institute, Sweden

<sup>3</sup> Department of Medical Radiation and Nuclear Medicine, Karolinska University Hospital, Sweden

\* tie-qiang.li@ki.se.

**Abstract:** Compressed Sensing (CS) MRI has shown great potential in enhancing time efficiency. Deep learning techniques, specifically Generative Adversarial Networks (GANs), have emerged as potent tools for speedy CS-MRI reconstruction. Yet, as the complexity of deep learning reconstruction models increases, this can lead to prolonged reconstruction time and challenges in achieving convergence. In this study we present a novel GAN-based model that delivers superior performance without escalating model complexity. Our generator module, built on the U-net architecture, incorporates dilated residual (DR) networks, thus expanding the network's receptive field without increasing parameters or computational load. At every step of the downsampling path, this revamped generator module includes a DR network, with the dilation rates adjusted according to the depth of the network layer. Moreover, we have introduced a channel attention mechanism (CAM) to distinguish between channels and reduce background noise, thereby focusing on key information. This mechanism adeptly combines global maximum and average pooling approaches to refine channel attention. We conducted comprehensive experiments with the designed model using public domain MRI datasets of the human brain. Ablation studies affirmed the efficacy of the modified modules within the network. Compared to other relevant models, our proposed model exhibits exceptional performance, achieving not only excellent stability but also outperforming other networks in terms of Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM). The model presents a promising pathway for enhancing the efficiency and quality of CS-MRI reconstruction.

**Keywords:** Compressed Sensing MRI; GAN; U-net; dilated-residual blocks; channel attention mechanism

---

## 1. Introduction

Compressed Sensing (CS) is a promising technique that capitalizes on the sparsity property for signal recovery [1]. CS-based methods have been increasingly employed to enhance MRI time efficiency. Introduced by Donoho in 2006 [2], the CS method, also known as compressed or sparse sampling, was first applied to fast MRI by Lustig et al. [3]. CS-MRI leverages sparse sampling and convex optimization algorithms to improve clinical MRI time efficiency [4]. Although CS-MRI surpasses the Nyquist–Shannon sampling barrier, its acceleration rate remains limited. Current sparse transforms in CS-MRI struggle to capture intricate image details [5], and iterative calculations in the nonlinear optimization solver prolong reconstruction time [6]. Inappropriate hyperparameter prediction may result in overly smooth, unnatural images [7]. However, the widespread influence of artificial intelligence (AI) and deep learning (DL) advancements [8] have enhanced various aspects of medical image reconstruction, including speed, accuracy, and robustness [9], making them increasingly important for CS-MRI fast reconstruction.

Deep learning techniques, such as generative adversarial networks (GANs) [10, 11], have emerged as a leading approach to CS-MRI reconstruction. These techniques have enabled high-quality reconstruction without increasing model complexity. Convolutional Neural Networks (CNNs) are one of the most prominent DL models used for CS-MRI reconstruction. They excel in learning the non-linear mapping between under-sampled and fully sampled MRI images. Models like U-net and its variants [12] have been particularly successful due to their ability to extract and combine features at different levels. Recurrent Neural Networks (RNNs) and specifically their Long Short-Term Memory (LSTM) variant have also been used in CS-MRI reconstruction [13]. They take advantage of the temporal correlation in dynamic MRI data, thus improving the reconstruction quality of dynamic sequences. The GAN-based models have also shown potential in CS-MRI reconstruction [11, 14-17]. The competitive learning process between the generator and discriminator networks results in images with high perceptual quality. Some of the mainstream GAN-based CS-MRI reconstruction models are summarized below: CycleGAN [18] model has been used for unsupervised image translation, where the goal is to transform images from one domain to another. In CS-MRI, CycleGAN can be used to map under-sampled k-space data to fully sampled k-space data. The inherent cycle-consistency loss helps to preserve the important structures in the MRI images, leading to better reconstruction. The pix2pix GAN model [19], also known as the Image-to-Image Translation GAN, has been used in CS-MRI for mapping the under-sampled MRI images to their fully sampled counterparts. This model has shown promising results in generating high-quality reconstructions, although the quality of the output largely depends on the paired data quality. Deep Convolutional GAN (DCGAN) model [20] has been used in CS-MRI to learn the mapping between the under-sampled and fully sampled MRI data. The deep convolutional layers in this model can better capture the complex patterns in the MRI data, thereby producing high-quality reconstructed images. Wasserstein GAN (WGAN) model [21] uses the Wasserstein loss instead of the traditional GAN loss to stabilize the training process and improve the quality of the generated images. WGAN has been used in CS-MRI to generate more realistic reconstructions and better preserve the details of the original images. Progressive GAN model [22] adopts a new training methodology that grows both the generator and discriminator progressively, adding layers that model increasingly fine details as training progresses. This method allows the model to effectively generate high-resolution details, leading to improved CS-MRI reconstruction.

Deep learning-based methods have shown impressive results in image and signal reconstruction. They often outperform conventional algorithms in terms of both speed and quality [23]. However, many models require increased network depth and width, complicating training, and prolonging reconstruction time [21, 24]. In this study, we improved a GAN-based model for construction of CS-MRI. The generator module of the model was derived from the U-Net architecture by integrating dilated residual (DR) structures to expand the network's receptive field, avoiding an increase in parameters or computation. At each step of the downsampling path the revamped generator module incorporates three such structures with varying dilation rates depending on the depth of the network layer [25, 26]. To concentrate on essential information, we also introduced a channel attention mechanism (CAM) that distinguishes between channels and reduces background noise [27]. This mechanism integrates global maximum and average pooling for a more precise channel attention. We've named this GAN-based CS-MRI reconstruction model, equipped with DR networks and CAM, as DR-CAM-GAN.

Using public domain MRI datasets for the human brain, we conducted extensive reconstruction experimentation with the designed DR-CAM-GAN model at different undersampling levels. To assess the performance of the model, we compared the reconstructed image qualities with the results for a few reference models, including CRNN [13], DAGAN [15] and RefineGAN [28]. The image quality was evaluated using peak signal-to-noise ratio (PSNR), structural similarity (SSIM), and mean square error (MSE) [29]. The highlights of the study include the following: (1) Dilated-residual networks with varying dilation rates to fully satisfy the receptive field [30], a channel attention mechanism for refining network resource allocation [31], and a multi-scale information fusion module for feature fusion and improved MRI reconstruction quality [32]. (2) A discriminator

structure design that avoids max-pooling layers and implements feature downsampling through varied convolution strides, with batch normalization to address gradient vanishing. (3) A combination of four loss functions—pixel-wise image-domain mean square error loss, frequency domain MSE loss, perceptual VGG loss, and adversarial loss—assigned different weights to enhance reconstruction quality and visual perception.

## 2. Materials and Methods

### 2.1. Datasets and data processing:

To demonstrate the effectiveness of the proposed model, we utilized publicly available T<sub>1</sub>-weighted magnetization prepared rapid gradient echo (MPRAGE) datasets of the human brain for testing, including the diencephalon challenge dataset of the MICCAI 2013 grand challenge (<https://www.synapse.org/>) and the Open Access Series of Imaging Studies (OASIS) dataset (<https://www.oasis-brains.org/>), which contains neuroimaging data for 1378 participants collected over a 30-year period. There are 176 image slices in each 3D T<sub>1</sub>-weighted MPRAGE file. We randomly shuffled these files, allocating 70% for training, 10% for validation, and 20% for testing. The training set updates the network parameters through gradient descent approach to best fit the real data distribution, while the validation set helps hyperparameter fine-tuning, identify overfitting, and select higher accuracy models.

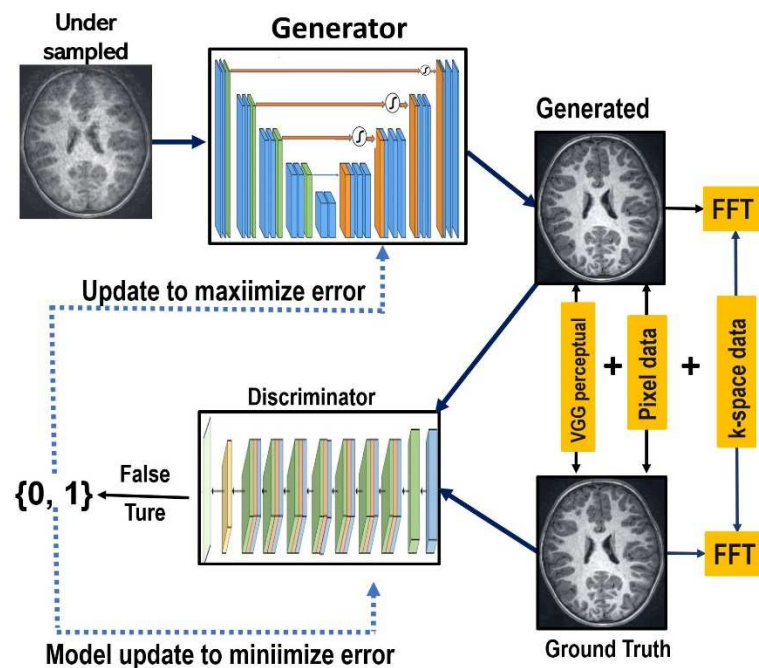
We utilized a fast Fourier transform procedure to convert the images into fully sampled k-space data within the complex domain. To generate undersampled images for CS-MRI reconstruction at different sampling rates, we extracted the k-space data at sub-Nyquist rates of 10%, 20%, 30%, and 50%, followed by zero-filling of the undersampled k-space data and applying an inverse Fourier transform. This process aimed to demonstrate the adaptability of our proposed model to various CS-MRI sampling rates. During the reconstruction model training, the subsampled k-space data was used as input to generate output images that matched the Fourier transforms of fully sampled k-space data. Sparsity plays a crucial role in CS-MRI, particularly in the wavelet or Fourier domain. It is essential for the CS-MRI data sampling pattern to be incoherent with the sparse basis. Typically, variable-density random undersampling in the k-space is employed to achieve this. Since the central k-space region contains important high-energy information for image quality and tissue contrast, and higher k-space data contributes to image details and fine structures [12], we always include the central 4% low-frequency information to ensure good image quality during undersampling. To simulate incoherent k-space data undersampling, we adopted a Monte Carlo's random undersampling strategy based on Gaussian sampling distribution. This approach exploits the high energy at the center of the k-space while also including some sparse-distributed high-frequency k-space data points needed to preserve structure fine details and correct for aliasing artifacts.

Given the importance of extensive datasets for deep learning models, and the limited availability of medical images, we utilized data augmentation techniques to expand our datasets and bolster model resilience. To achieve this, we implemented online stochastic augmentation, generating random augmentations for a data batch [33], which not only enhanced model stability but also alleviated storage constraints as compared to offline augmentation methods. We employed four different random enhancement methods with equal probability to ensure low data repeatability. We divided the data into four equal parts, augmenting each by flipping up and down, translating, mirroring horizontally, or rotating 90 degrees. These methods reduce model sensitivity to target position, accelerate model convergence, and improve performance.

### 2.2. The Proposed Reconstruction Model:

The GAN-based models are innovative deep generative models that leverage game theory and competitive learning between a generator and a discriminator to enhance the network's fitting ability. The proposed DR-CAM-GAN framework consists of generator and discriminator modules. The generator module takes undersampled MR images and outputs reconstructed images after multiple levels of convolutional operations, while the discriminator classifies de-aliased reconstructed images

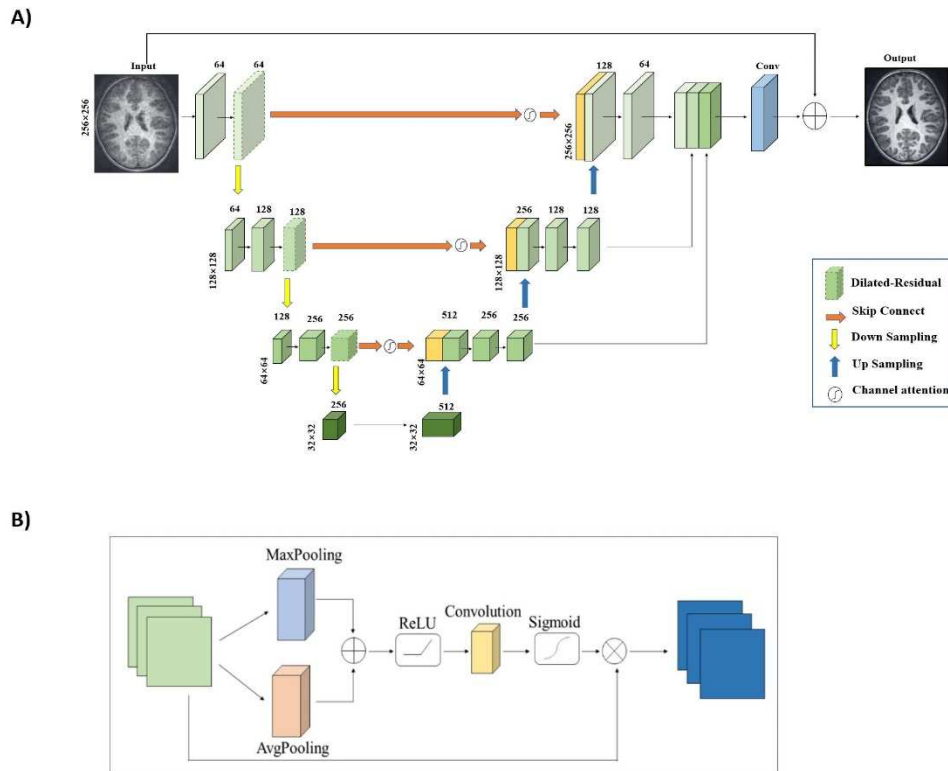
from fully sampled ground truth images. The outline of the proposed GAN framework is displayed in Figure 1.



**Figure 1.** Schematically illustration of the improved GAN model with dilated residual networks and channel attention mechanism (DR-CAM-GAN) aimed for CS-MRI reconstruction.

In the DR-CAM-GAN framework, we use a U-net base structure for the generator, with an encoding layer for feature extraction and a decoding layer for feature amplification in each stage. Skip connections transfer information from the encoding layer to the corresponding decoding layer, enabling feature fusion and providing more accurate detailed features for reconstruction.

To alleviate gradient disappearance or network degradation as network depth increases, we modify the residual structures by replacing the second standard convolution with a dilated convolution. This expands the network's receptive field without increasing parameters or computation. Our generator structure employs three dilated residual blocks with different dilation rates (1, 2, and 3), which vary depending on the network layer's depth. We use skip connections in the U-net structure [34] to fuse deep and shallow features, improving MR image reconstruction quality. To better focus on key information, we propose a new channel attention mechanism that establishes dependencies between channels, suppressing background information. This mechanism combines global maximum pooling and global average pooling to obtain more accurate channel attention. Despite the dilated-residual structure and channel attention mechanism's strengths, the decoding layer structure is a bottom-up process, which can lead to information loss or corruption. We employ a multi-scale information fusion module to aggregate features from multiple levels, using linear interpolation techniques for upsampling to avoid midway detail information loss and fully utilize the information. Figure 2 depicts the generator with integrated DR blocks and CAM.



**Figure 2.** The U-net based generator architecture (A) with integrated DR blocks and CAM (B).

The discriminator's primary function is to determine the input MR images' source. It consists of a series of convolutions with batch-normalization after each convolution to normalize input and avoid gradient disappearance. The discriminator structure follows Radford et al.'s [35] architectural guidelines, using convolution stride variations for feature downsampling instead of max-pooling layers. We utilize LeakyReLU for enhanced nonlinearity, and a sigmoid function to determine whether an MR image is fully sampled or reconstructed.

### 2.3. The Loss Function:

The GAN model is trained in alternating fashion: the discriminator trains for one or more epochs followed by the training of the generator. This procedure is repeated until the targeted loss or the ultimate iterations is reached. To enhance the reconstruction quality and visual perception of our model, we used a combination of pixel-wise image domain mean square error (MSE) loss, frequency domain MSE loss, perceptual VGG loss and adversarial loss [36], assigning them different weights to form the generator's loss function. The combined loss function for the generator is given by:

$$L_{combine} = \alpha L_{iMSE} + \beta L_{fMSE} + \delta L_{VGG} + L_{GEN} \quad (1)$$

where the hyperparameters  $\alpha, \beta, \delta$  are the weights associated with different loss terms. Balancing these weights is achieved by setting them for different undersampling ratios. We train the network with  $\alpha=15$ ,  $\beta=0.1$ , and  $\delta=0.0025$  as used previously in GAN-related model training. Pixel loss ensures positive similarity between the reconstructed and original images by calculating the MSE between each pixel point of the reconstructed MRI and the fully sampled image:

$$L_{iMSE} = \frac{1}{2} \|X_t - X_u\|_2^2 \quad (2)$$

Here,  $X_t$  represents the fully sampled image and  $X_u$  is the reconstructed image generated by a cascaded convolutional neural network. Frequency domain MSE loss enhances the similarity between a fully sampled image and a reconstructed image, defined as:

$$L_{MSE} = \frac{1}{2} \|Y_t - Y_u\|_2^2 \quad (3)$$

In this expression,  $Y_t$  and  $Y_u$  are the corresponding frequency domain data of  $X_t$  and  $X_u$ . Pixel and frequency domain losses yield reconstructed MRIs with higher PSNR and lower MSE. However, optimizing MSE may result in the loss of high-frequency information, leading to over-smoothed images that affect human visual perception. To address this issue, we introduce perceptual loss from Gatys et al. [30] to reduce the gap between reconstructed MRI and fully sampled image in feature space, obtaining higher texture similarity:

$$L_{VGG} = \frac{1}{2} \|f_{VGG}(X_t) - f_{VGG}(X_u)\|_2^2 \quad (4)$$

where  $f_{VGG}$  refers to the VGG16 network.

The adversarial loss in the GAN network represents the difference between the predicted probability distribution produced by the discriminator and the actual probability distribution of real samples. It is expressed as the log of discriminator probability distribution for the generated image data:

$$L_{GEN} = -\log\left(D_{\theta_d}\left(G_{\theta_g}(X_u)\right)\right) \quad (5)$$

While the generator aims to produce a better image from the undersampled data by focusing on improving image quality, the discriminator's primary objective is to maximize the probability assigned to real and fake images. That is to distinguish the image produced by the generator from the fully sampled ground truth image. The two modules are engaged in min-max game, where simultaneous improvements are achieved for both modules through competition. Mathematically, the model is fit seeking to minimize the average binary cross entropy. This can be expressed as follows [36]:

$$L_D = -\log\left(D_{\theta_d}(X_t)\right) - \log\left(1 - D_{\theta_d}(X_u)\right) \quad (6)$$

where  $\log\left(D_{\theta_d}(X_t)\right)$  represents the log of discriminator probability distribution for the fully sampled image data, and  $\log\left(1 - D_{\theta_d}(X_u)\right)$  is the log of invert probability distribution for the generated images from the undersampled data. For the discriminator, minimizing the  $L_D$  loss function is equivalent to maximizing the judgement accuracy of the fully sampled image and the network's reconstructed image using under-sampled data.

#### 2.4. Model Training:

We employed an NVIDIA Geforce 3060 GPU for both training and testing phases, utilizing the PyTorch development environment and a model with a total of 41.84 MB in parameters. Our model was trained using the ADAM optimizer with set parameters:  $\beta_1=0.9$ ,  $\beta_2=0.999$ , an initial learning rate of 0.0001, a learning rate decay factor of 0.5, and an update interval of 10 epochs for the learning rate. To mitigate overfitting, we used MSE as the evaluation metric to optimize our model. Training was halted and the current model saved if the observed MSE was lower than any MSE from the subsequent 20 epochs. Each epoch, which included the computation of PSNR, SSIM, and MSE for the validation sets, was completed within approximately 20 minutes.

We conducted ablation studies to ascertain the contribution of individual elements to our proposed reconstruction model. By systematically altering the model, we gauged the performance of

these variants in comparison to the original comprehensive model. We specifically scrutinized the effects of the dilated-residual structure, channel attention mechanism, and multi-scale information fusion module on the model's performance. This was accomplished by modifying or removing each component and evaluating the resultant model's performance.

In addition, as detailed in Table 1, we compared the performance of four renowned CS-MRI reconstruction models: U-net [12], CRNN [13], DAGAN [15], and RefineGAN [28] using the same training datasets and similar training procedures as described above. We used three metrics (PSNR, SSIM, and MSE) to assess and compare the performance of the different models under different levels of under-sampling.

**Table 1.** Summary of the different DL models used to compare their performance in CS-MRI reconstruction.

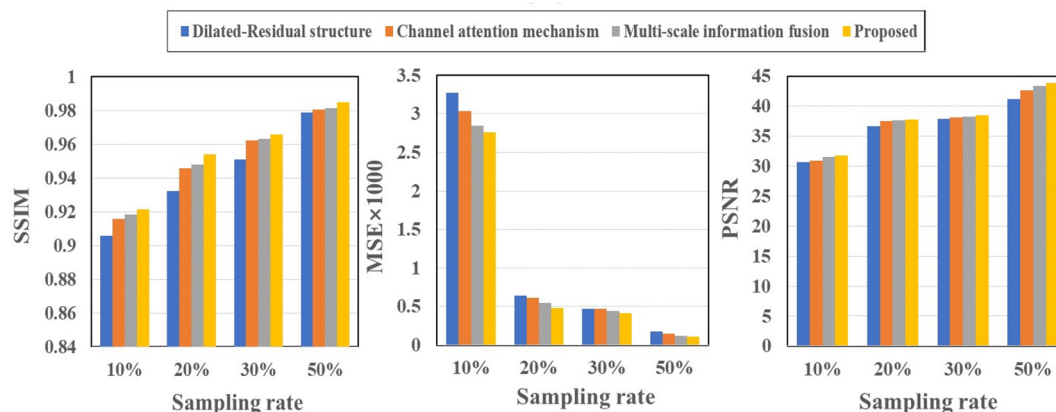
Model	Reference	Parameters	Model Size	Batch	Epoch	Training
U-net	[12]	2.04M	5.12MB	16	200	3.6 h
CRNN	[13]	0.32M	1.13MB	16	100	2.4 h
DAGAN	[15]	98.60M	376.16MB	16	60	10.6 h
RefineGAN	[28]	40.91M	105.34MB	16	50	18.4 h
DR-CAM-GAN		15.77M	41.84MB	16	50	14.2 h

### 3. Results

Table 2 and Figure 3 present the quantitative outcomes of our ablation studies, encompassing evaluation metrics such as SSIM, MSE, and PSNR across four different undersampling levels. The full model we propose, DR-CAM-GAN, outperforms in all three metrics. The DR networks elevate PSNR, SSIM, and MSE by 1.6-6.3%, 0.6-2.3%, and 14-64%, respectively, contingent on the undersampling levels. Introducing the CAM results in improvements of 0.9-3.0% for PSNR, 0.4-0.9 for SSIM, and 10-26% for MSE. Multi-scale information fusion modestly boosts PSNR, SSIM, and MSE by 0.4-1.3%, 0.3-0.6, and 3-15%, respectively. Among these different innovative modifications, the DR networks contribute most significantly to performance enhancement, followed by the improvement from the CAM.

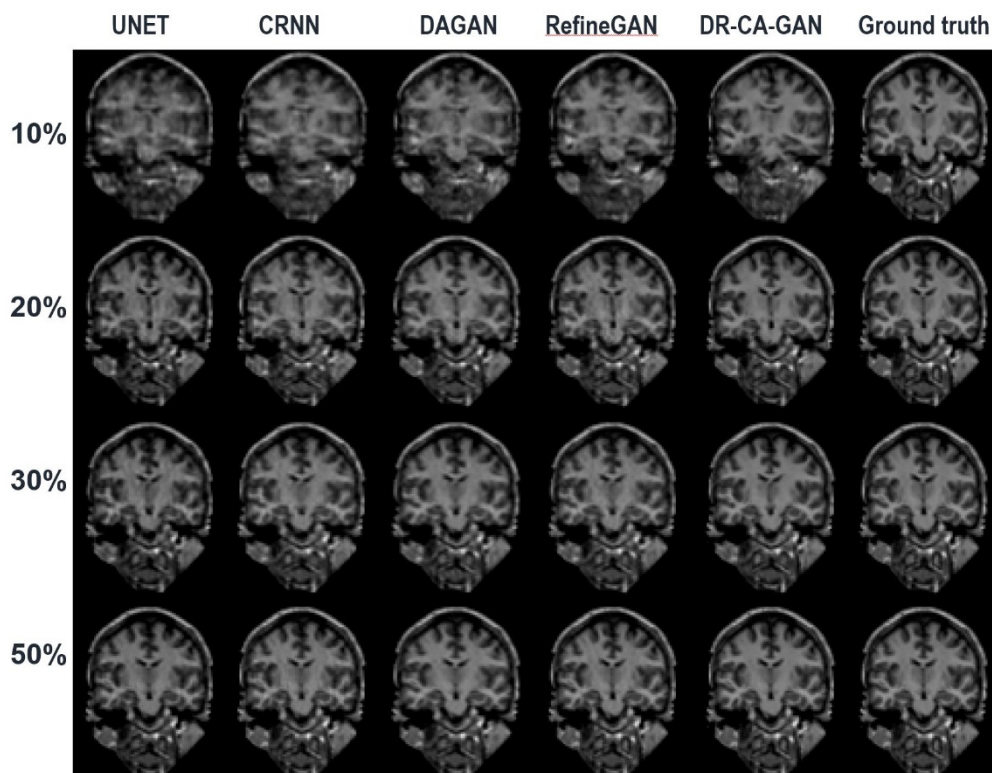
**Table 2.** Results of the ablation experiments with the DR-CAM-GAN framework.

Sampling	Ablation comparison	SSIM	MSE( $\times 10^{-3}$ )	PSNR
10%	Dilated-Residual structure	0.9059	3.27	30.62
	Channel attention mechanism	0.9157	3.04	30.91
	Multi-scale information fusion	0.9184	2.85	31.49
	DR-CAM-GAN full model	0.9213	2.76	31.76
20%	Dilated-Residual structure	0.9325	0.64	36.68
	Channel attention mechanism	0.9456	0.61	37.45
	Multi-scale information fusion	0.9479	0.55	37.63
	DR-CAM-GAN full model	0.9541	0.48	37.79
30%	Dilated-Residual structure	0.9508	0.47	37.82
	Channel attention mechanism	0.9622	0.47	38.05
	Multi-scale information fusion	0.9631	0.44	38.21
	DR-CAM-GAN full model	0.9656	0.41	38.43
50%	Dilated-Residual structure	0.9786	0.18	41.13
	Channel attention mechanism	0.9804	0.15	42.62
	Multi-scale information fusion	0.9812	0.12	43.35
	DR-CAM-GAN full model	0.9847	0.11	43.90

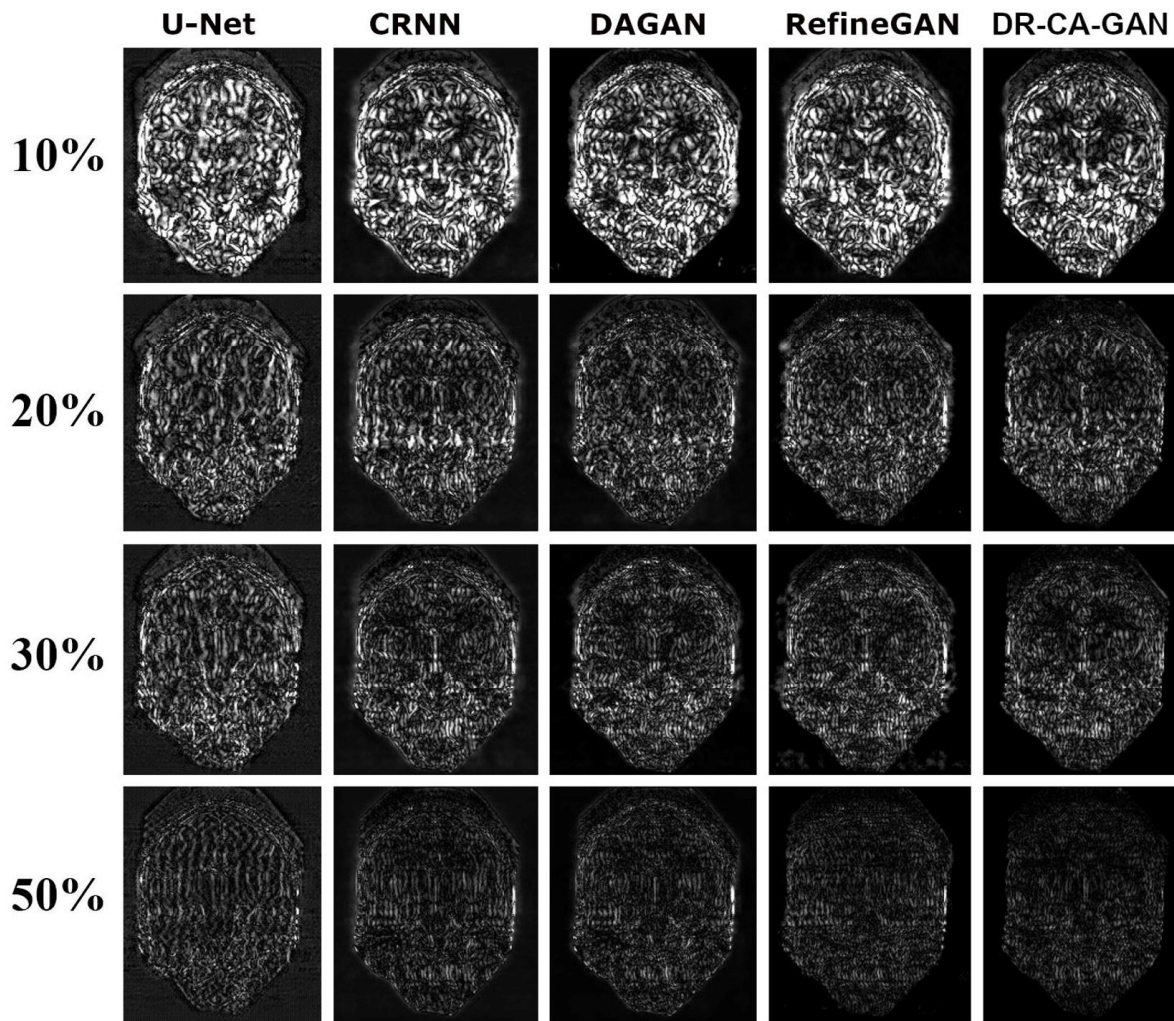


**Figure 3.** Bar graph of the SSIM, MSE, and PSNR as a function of the CS-undersampling rates for the different models.

Figure 4 depicts a representative coronal slice of images reconstructed by various models, including our proposed DR-CAM-GAN framework, at four distinct undersampling levels. Figure 5 demonstrates the MSE outcomes for the same slice. At lower sampling rates (10% and 20%), models such as U-net, CRNN, and DAGAN grapple with signal recovery, leading to blurring in the reconstructed images. In stark contrast, both the RefineGAN and proposed DR-CAM-GAN models display superior performance, with less blur and enhanced detail recovery in brain structure. The DR-CAM-GAN framework excels, even surpassing the highly effective RefineGAN model. At a 20% sampling rate, the network succeeds in reconstructing largely the brain structure, although edge details remain slightly blurred with minor artifacts visible. On the other hand, reconstructions from the CRNN and DAGAN models appear noisy and present more significant loss of brain anatomy details.



**Figure 4.** A coronal slice reconstructed from a T1-weighted MPRAGE volume by five distinct DL models (column-wise) at four different CS undersampling rates (row-wise).



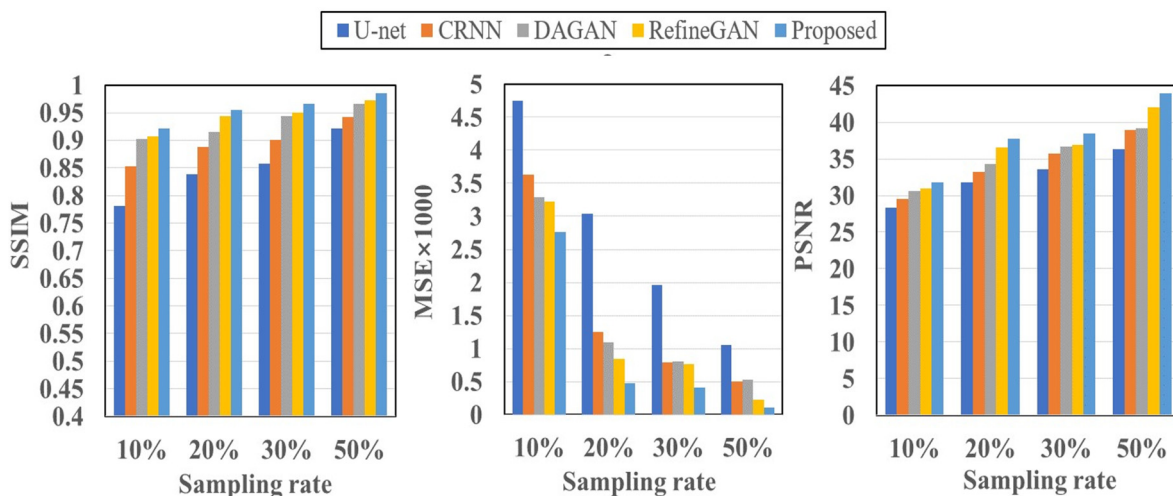
**Figure 5.** The MSE maps corresponding to the coronal slice depicted in Figure 4, illustrating results from five DL models (column-wise) at four different CS undersampling rates (row-wise).

Table 3 and Figure 6 encapsulate the quantitative evaluations of the reconstructed image quality. DL models showcase robust performance, even at elevated undersampling rates, with various quality metrics (PSNR, SSIM, and MSE) following a similar trajectory. Our proposed DR-CAM-GAN model stands out, demonstrating remarkable performance and stability. Regardless of the sampling rate, DR-CAM-GAN consistently outperforms other networks, achieving substantial improvements in PSNR and SSIM. When pitted against U-net, enhancements in SSIM and PSNR range from 6-15% and 11-17% respectively. Even compared to RefineGAN, improvements in SSIM and PSNR consistently surpass 1-3% respectively, depending on the undersampling levels. Particularly at higher sampling rates, the reduction in MSE is noteworthy, with improvements ranging from 2 to 9 times.

**Table 3.** Performance comparison for the different models at various Cs under sampling rates.

Sampling	Model	SSIM	MSE( $\times 10^{-3}$ )	PSNR
10%	U-net	0.7812	4.75	28.34
	CRNN	0.8531	3.63	29.53
	DAGAN	0.9026	3.29	30.59
	RefineGAN	0.9064	3.22	30.97
	DR-CAM-GAN	0.9213	2.76	31.76
20%	U-net	0.8381	3.04	31.75
	CRNN	0.8874	1.25	33.16

	DAGAN	0.9149	1.09	34.32
	RefineGAN	0.9437	0.85	36.50
	DR-CAM-GAN	0.9541	0.48	37.79
30%	U-net	0.8569	1.96	33.54
	CRNN	0.9003	0.79	35.67
	DAGAN	0.9444	0.81	36.69
	RefineGAN	0.9506	0.77	36.95
	DR-CAM-GAN	0.9656	0.41	38.43
	50%	U-net	0.9209	1.05
CRNN		0.9416	0.50	38.93
DAGAN		0.9665	0.53	39.18
RefineGAN		0.9718	0.23	42.05
DR-CAM-GAN		0.9847	0.11	43.90



**Figure 6.** Bar graph of the SSIM, MSE, and PSNR as a function of the CS-undersampling rates for the different DL models.

#### 4. Discussion

To summarize, our proposed DR-CAM-GAN model exhibits superior performance in image reconstruction compared to other DL models, even at elevated undersampling rates. Quantitative assessments indicated substantial improvements in PSNR, SSIM, and MSE metrics. The DR networks contribute the most significantly to this performance enhancement. The DR-CAM-GAN model consistently outperforms U-net, CRNN, DAGAN, and even the efficient RefineGAN, with notable improvements in image quality and noise reduction. In essence, the DR-CAM-GAN model demonstrates exceptional performance and stability, effectively recovering detailed brain structures from undersampled data.

The ablation study results presented in Table 2 and Figure 3 offer valuable insights into the contributions of each component of our proposed DR-CAM-GAN model. By evaluating the model's performance using SSIM, MSE, and PSNR metrics across four distinct under-sampling levels, we can better understand how each element impacts the overall performance. In summary, the ablation study highlights the importance of each component in the DR-CAM-GAN model. The DR networks are the primary driver of performance improvement, while the CA mechanism and multi-scale information fusion provide valuable support. These combined elements allow the DR-CAM-GAN model to achieve superior performance across all evaluation metrics, demonstrating its effectiveness in reconstructing high-quality images from undersampled data. The DR networks play a crucial role in enhancing the model's performance. The DR networks demonstrate their effectiveness in capturing multi-scale information and preserving the spatial structure of the images. This leads to better image quality and improved reconstruction, particularly in capturing fine details and maintaining the

overall structure of the brain. The CA mechanism, though providing a moderate contribution, is still an essential component of the model. It focuses on important features by adaptively weighing channel-wise information, thereby improving the model's ability to recover specific details and suppress less relevant information. This results in better image quality and reduced noise. The multiscale information fusion further refines the model's performance by moderately increasing PSNR, SSIM, and MSE. This component enables the model to incorporate information from different scales, ensuring that both global and local features are well-represented in the final re-construction. This fusion process contributes to a more accurate and detailed image representation.

As suggested by the ablation study results, the superior performance can be attributed to the combination of DR networks, CA mechanism, and multi-scale information fusion. This combination allows the DR-CAM-GAN model to outperform other DL models and achieve exceptional performance and stability in CS image reconstruction tasks. The model comparison results showcased in Figures 4–6 provide a comprehensive understanding of how the proposed DR-CAM-GAN model performs against other deep-learning models in image reconstruction tasks. By examining the reconstruction quality of various models, such as U-net, CRNN, DAGAN, RefineGAN, and our proposed DR-CAM-GAN, we can evaluate their strengths and weaknesses and identify the factors contributing to the superior performance of the DR-CAM-GAN model.

At low sampling rates (10 and 20%), the limited amount of sampled data poses a challenge for most models. U-net, CRNN, and DAGAN struggle to recover lost signals, resulting in blurry reconstructed images with less detail. This limitation can be attributed to these models' inability to capture multi-scale information effectively, which is crucial for preserving the spatial structure and finer details in the images.

In contrast, RefineGAN and our DR-CAM-GAN model demonstrate better performance in reconstructing images with less blur and more detailed brain structures. The DR-CAM-GAN model, in particular, leverages Dilated-Residual networks, which enable the model to capture multi-scale information and preserve the spatial structure more effectively. Additionally, the channel attention mechanism focuses on relevant features, improving the model's ability to recover specific details and suppress less important information.

The DR-CAM-GAN model outperforms even the highly efficient RefineGAN model. At a 20% sampling rate, the network recovers much of the brain structure. However, edge information remains somewhat blurred, and minor artifacts are present. Conversely, reconstruction results from CRNN and DAGAN models are noisy, with some loss of brain anatomy details.

Quantitative assessments of image quality, including PSNR, SSIM, and MSE metrics, further confirm the outstanding performance of the DR-CAM-GAN model. As shown in Table 3 and Figure 6, deep-learning models demonstrate good performance even at high undersampling rates, but the DR-CAM-GAN model consistently achieves significant improvements in PSNR and SSIM compared to other networks. This superior performance can be attributed to the model's ability to effectively capture multi-scale information and focus on relevant features through the DR networks and the CA mechanism.

## 5. Conclusion

In conclusion, the DR-CAM-GAN model outperforms other deep-learning models in image reconstruction tasks, even at high undersampling rates. Its superior performance and stability in recovering detailed brain structures stem from the integration of Dilated-Residual networks, channel attention mechanism, and multi-scale information fusion. The ablation study highlights the importance of DR networks, with the CA mechanism and multi-scale fusion providing valuable support. Consequently, the DR-CAM-GAN model excels across all evaluation metrics, proving its effectiveness in reconstructing high-quality images from undersampled data. The model surpasses deep-learning models like U-net, CRNN, DAGAN, and even RefineGAN, with quantitative assessments confirming its outstanding performance. The DR-CAM-GAN model's success lies in its ability to capture multi-scale information and focus on relevant features. Ultimately, this model

presents a promising solution for reconstructing high-quality images from undersampled data, with significant potential for various medical imaging and computer vision applications.

**Funding Acknowledgements:** This research was supported by a grant from Zhejiang Natural Science Foundation of China (No. LY23F010005), the ALF foundation in Stockholm Region, and the Joint China-Sweden Mobility program from STINT (Dnr: CH2019-8397).

## References

1. Montefusco, L. B.; Lazzaro, D.; Papi, S.; Guerrini, C., A fast compressed sensing approach to 3D MR image reconstruction. *IEEE Trans Med Imaging* **2011**, 30, (5), 1064-75.
2. Donoho, D. L., Compressed Sensing. *IEEE TRANSACTIONS ON INFORMATION THEORY* **2006**, 52, (4), 1289-1306.
3. Lustig, M.; Donoho, D.; Pauly, J. M., Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magn Reson Med* **2007**, 58, (6), 1182-95.
4. Jin, K. H.; Lee, D.; Ye, J. C., A General Framework for Compressed Sensing and Parallel MRI Using Annihilating Filter Based Low-Rank Hankel Matrix. *IEEE Transactions on Computational Imaging* **2016**, 2, (4), 480-495.
5. Huang, Y.; Paisley, J.; Lin, Q.; Ding, X.; Fu, X.; Zhang, X. P., Bayesian nonparametric dictionary learning for compressed sensing MRI. *IEEE Trans Image Process* **2014**, 23, (12), 5007-19.
6. Qu, X.; Zhang, W.; Guo, D.; Cai, C.; Cai, S.; Chen, Z., Iterative thresholding compressed sensing MRI based on contourlet transform. *Inverse Problems in Science and Engineering* **2010**, 18, (6), 737-758.
7. Hollingsworth, K. G., Reducing acquisition time in clinical MRI by data undersampling and compressed sensing reconstruction. *Phys Med Biol* **2015**, 60, (21), R297-322.
8. Wang, J.; Zhu, H.; Wang, S.-H.; Zhang, Y.-D., A Review of Deep Learning on Medical Image Analysis. *Mobile Networks and Applications* **2020**, 26, (1), 351-380.
9. Wang, G.; Ye, J. C.; Mueller, K.; Fessler, J. A., Image Reconstruction is a New Frontier of Machine Learning. *IEEE Trans Med Imaging* **2018**, 37, (6), 1289-1296.
10. Michael Lustig; David L. Donoho; Juan M. Santos, a; Pauly], J. M., Compressed Sensing MRI [A look at how CS can improve on current imaging techniques]. *IEEE SIGNAL PROCESSING MAGAZINE* **2008**, 73-82.
11. Zhang, H.; Goodfellow, I.; Metaxas, D.; Odena, A. In *Self-attention generative adversarial networks*, International conference on machine learning, 2019; PMLR: 2019; pp 7354-7363.
12. Hyun, C. M.; Kim, H. P.; Lee, S. M.; Lee, S.; Seo, J. K., Deep learning for undersampled MRI reconstruction. *Phys Med Biol* **2018**, 63, (13), 135007.
13. 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 Transactions on Medical Imaging* **2018**, 37, (2), 491-503.
14. Sun, L.; Fan, Z.; Huang, Y.; Ding, X.; Paisley, J. In *Compressed sensing MRI using a recursive dilated network*, Proceedings of the AAAI Conference on Artificial Intelligence, 2018; 2018.
15. Yang, G.; Yu, S.; Dong, H.; Slabaugh, G.; Dragotti, P. L.; Ye, X.; Liu, F.; Arridge, S.; Keegan, J.; Guo, Y.; Firmin, D.; Keegan, J.; Slabaugh, G.; Arridge, S.; Ye, X.; Guo, Y.; Yu, S.; Liu, F.; Firmin, D.; Dragotti, P. L.; Yang, G.; Dong, H., DAGAN: Deep De-Aliasing Generative Adversarial Networks for Fast Compressed Sensing MRI Reconstruction. *IEEE Trans Med Imaging* **2018**, 37, (6), 1310-1321.
16. 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* **2019**, 38, (1), 167-179.
17. Guo, Y.; Wang, C.; Zhang, H.; Yang, G. In *Deep attentive wasserstein generative adversarial networks for mri reconstruction with recurrent context-awareness*, International Conference on Medical Image Computing and Computer-Assisted Intervention, 2020; Springer: 2020; pp 167-177.
18. Zhu, J. Y.; Park, T.; Isola, P.; Efros, A. A. In *Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks*, 2017 IEEE International Conference on Computer Vision (ICCV), 22-29 Oct. 2017, 2017; 2017; pp 2242-2251.
19. Isola, P.; Zhu, J. Y.; Zhou, T.; Efros, A. A. In *Image-to-Image Translation with Conditional Adversarial Networks*, 2017 IEEE International Conference on Computer Vision (ICCV), 22-29 Oct. 2017, 2017; 2017; pp 1125-1134.
20. Radford, A.; Metz, L.; Chintala, S., Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. *CoRR* **2015**, abs/1511.06434.
21. Jiang, M.; Yuan, Z.; Yang, X.; Zhang, J.; Gong, Y.; Xia, L.; Li, T., Accelerating CS-MRI reconstruction with fine-tuning Wasserstein generative adversarial network. *IEEE Access* **2019**, 7, 152347-152357.

22. Karras, T.; Aila, T.; Laine, S.; Lehtinen, J., Progressive Growing of GANs for Improved Quality, Stability, and Variation. In *International Conference on Learning Representations*, 2018.
23. Tanuj Kumar Jham, V. R., V.K. Govinda, A Review on Image Reconstruction through MRI
24. k-Space Data. *I.J. Image, Graphics and Signal Processing* **2015**, *7*, 42-59.
25. Chandra, S. S.; Bran Lorenzana, M.; Liu, X.; Liu, S.; Bollmann, S.; Crozier, S., Deep learning in magnetic resonance image reconstruction. *J Med Imaging Radiat Oncol* **2021**, *65*, (5), 564-577.
26. Dai, Y.; Zhuang, P., Compressed sensing MRI via a multi-scale dilated residual convolution network. *Magn Reson Imaging* **2019**, *63*, 93-104.
27. Yu, F.; Funkhouser, V.; Funkhouser, T. In *Dilated residual networks*, Proceedings of the IEEE conference on computer vision and pattern recognition, 2017; 2017; pp 472-480.
28. Huang, G.; Zhu, J.; Li, J.; Wang, Z.; Cheng, L.; Liu, L.; Li, H.; Zhou, J., Channel-Attention U-Net: Channel Attention Mechanism for Semantic Segmentation of Esophagus and Esophageal Cancer. *IEEE Access* **2020**, *8*, 122798-122810.
29. 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* **2018**, *37*, (6), 1488-1497.
30. Horé, A.; Ziou, D. In *Image Quality Metrics: PSNR vs. SSIM*, 2010 20th International Conference on Pattern Recognition, 23-26 Aug. 2010, 2010; 2010; pp 2366-2369.
31. Yu, F.; Koltun, V., Multi-scale context aggregation by dilated convolutions. *arXiv preprint arXiv:1511.07122* **2015**.
32. Hu, J.; Shen, L.; Albanie, S.; Sun, G.; Wu, E., Squeeze-and-Excitation Networks. *IEEE Trans Pattern Anal Mach Intell* **2020**, *42*, (8), 2011-2023.
33. Chen, L.-C.; Papandreou, G.; Schroff, F.; Adam, H., Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587* **2017**.
34. Girshick, R. In *Fast r-cnn*, Proceedings of the IEEE international conference on computer vision, 2015; 2015; pp 1440-1448.
35. Ronneberger, O.; Fischer, P.; Brox, T. In *U-Net: Convolutional Networks for Biomedical Image Segmentation*, Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Cham, 2015//, 2015; Navab, N.; Hornegger, J.; Wells, W. M.; Frangi, A. F., Eds. Springer International Publishing: Cham, 2015; pp 234-241.
36. Radford, A.; Metz, L.; Chintala, S., Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. *CoRR* **2016**, abs/1511.06434.