Cardiac MRI reconstruction from undersampled k-space using double-stream IFFT and a denoising GNA-UNET pipeline

Julia Dietlmeier[1 [0000−0001−9980−0910]],
Carles Garcia-Cabrera[2 [0000−0001−8139−9647]],
Anam Hashmi[2 [0009−0007−0887−246X]], Kathleen M. Curran[3 [0000−0003−0095−9337]],
and Noel E. O'Connor[1 [0000−0002−4033−9135]]

[1] Insight SFI Research Centre for Data Analytics, Glasnevin, Dublin 9, Ireland
julia.dietlmeier,noel.oconnor@insight-centre.org
http://www.insight-centre.org
[2] ML-Labs, Dublin City University, Glasnevin, Dublin 9, Ireland
carles.garciacabrera6,anam.hashmi2@mail.dcu.ie
http://www.ml-labs.ie/
[3] University College Dublin, School of Medicine, Dublin, Ireland
kathleen.curran@ucd.ie
http://www.ucd.ie/

Abstract. In this work, we approach the problem of cardiac Magnetic Resonance Imaging (MRI) image reconstruction from undersampled k-space. This is an inherently ill-posed problem leading to a variety of noise and aliasing artifacts if not appropriately addressed. We propose a two-step double-stream processing pipeline that first reconstructs a noisy sample from the undersampled k-space (frequency domain) using the inverse Fourier transform. Second, in the spatial domain we train a denoising GNA-UNET (enhanced by Group Normalization and Attention layers) on the noisy aliased and fully sampled image data using the Mean Square Error loss function. We achieve competitive results on the leaderboard and show that the algorithmic combination proposed is effective in high-quality MRI reconstruction from undersampled cardiac long-axis and short-axis complex k-space data.

Keywords: Cardiac MRI · Undersampled k-space · Deep Learning · Denoising UNET.

1 Introduction

In clinical diagnostics, Magnetic Resonance Imaging (MRI) is a widely used non-invasive medical imaging modality providing superior soft tissue contrast. Cardiac MRI, in particular, offers the opportunity for the diagnosis of cardiovascular disease. The long acquisition time in MRI remains a major weakness of the approach. The undersampling of k-space by different factors (also known as the compressed sensing (CS) approach [10]) offers accelerated data collection providing discomfort relief for paediatric patients. In addition to the motion artifacts

inherently present in cardiac MRI, the process of frequency domain undersampling results in the loss of high-frequency information [11], which translates to a variety of noise and structural aliasing artifacts in the spatial domain, and thus further deteriorates the reconstructed image quality.

A simple and a straightforward approach to denoising and artifact removal is to use a deep learning (DL) model, e.g. a Convolutional Neural Network (CNN), in the spatial domain to map a noisy aliased MR image to a higher-quality aliasing-free image. In fact, deep learning (DL) became an irreplaceable tool in MRI reconstruction and most DL approaches treat it as a denoising problem while some propose to directly complete the missing and corrupted k-space data using specially designed interpolation CNNs [7] or Transformers [8]. Some challenges still remain such as degraded performance of such models due to the high level of noise or limited training data and limited computational resources.

The work reported in this paper was conducted as part of the CMRxRecon challenge, Task 1: Cine reconstruction, and specifically we target high-quality MR image reconstruction in spatial domain. Our contribution is primarily in designing an effective double-stream pipeline to process the undersampled complex k-space data and to reconstruct the CMR images.

2 Related Works

Recently, end-to-end DL approaches have tried to solve the inverse problem of reconstructing MR images from sub-Nyquist sampled data [12]. For example, Wang et al. [13] trained a CNN based model using a large number of existing high quality MR images from downsampled reconstruction images as either an initialization or regularization term in classical CS approaches, to learn fully sampled reconstruction. Kwon et al. (2017), utilized a multilayer perceptron for fast reconstruction of MR images [14]. Lee et al. (2018), trained a residual network for phase and magnitude information separately to perform MR image reconstruction [15]. Numerous studies have incorporated the MRI domain knowledge, such as data consistency in the space to improve the standard deep learning algorithms for CS-MRI [16]. Some research used generative adversarial networks (GANs) [17]. Others used k-space reconstruction neural networks trained on autocalibration signal (ACS) data [18]. Typically, DL based methods require a large amount of training data, so several studies have tried to employ untrained CNNs such as Deep Image prior and Deep encoder and their variants [19–21]. More recently the MR image reconstruction research has shifted to using an *unrolled* framework with DL models because of better reconstruction performance and faster reconstruction times compared to iterative methods [22–24]. However, these approaches still rely on fully sampled scans as the ground truth data while overall it is preferable to reduce reliance on data.

In order to increase data efficiency, recent work has proposed designing data augmentation pipelines specifically suited to accelerated MRI reconstruction with appropriate image-based [25] or acquisition-based, physics-driven transformations [26]. These pipelines would supplement prior proposals that utilize

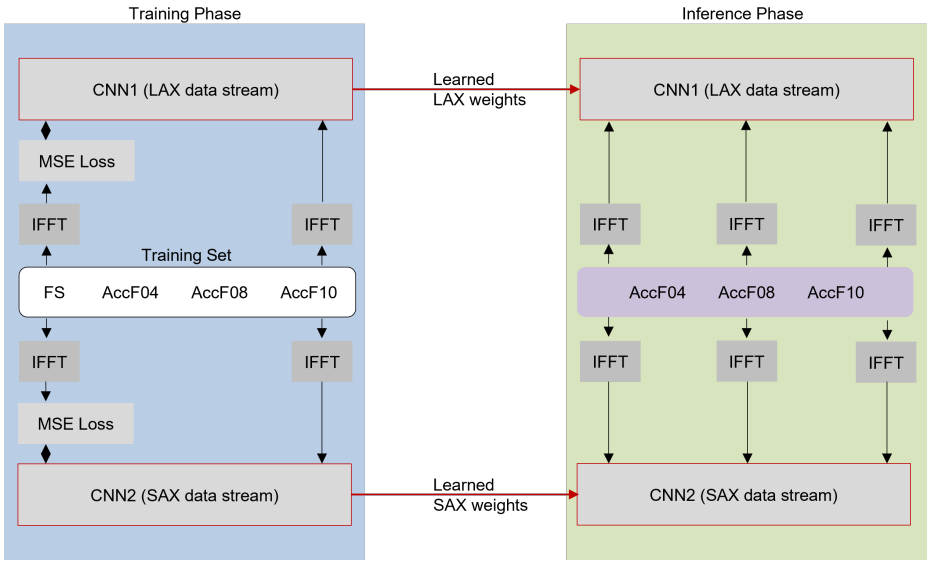

Fig. 1. Architecture of the proposed double-stream cardiac MRI reconstruction pipeline. Our proposed design involves processing the long-axis (LAX) and short-axis (SAX) data streams separately. Also, in the training phase we train the CNNs using the MSE loss function only on the AccFactor10 undersampled data. Inspired by the results from [9], we hypothesize that because the $\times 10$ acceleration factor is the most sparse one it results in strongest structural artifacts. Therefore a model trained on the fully sampled (FS) and AccF10 data can also be applied to the factor $\times 4$ (AccF04) and factor $\times 8$ (AccF08) undersampled data during the inference phase.

prospectively undersampled (unsupervised) data that currently lag in reconstruction performance [21]. However, data augmentation places an additional burden on the neural network by requiring it to learn every conceivable scale of every feature separately, even though it just attempts to approximate equivariance. The outcomes of such acquired equivariance are frequently worse than those of assured equivariance [27]. Therefore, scale-equivariant CNNs have been gaining traction in improving the data efficiency [28, 29].

3 Architecture

The original data are a complex k-space with different subsampling (acceleration) factors such as $\times 4$, $\times 8$, and $\times 10$. Subsampling masks and fully sampled k-space are also provided in the training set. A diagram of our processing pipeline is illustrated in Fig. 1. The pipeline operates directly on k-space data by converting frequency domain into spatial domain using the inverse Fourier transform and uses aliased undersampled images and the fully sampled images from the spatial domain to train a CNN.

We process the long-axis (LAX) and short-axis (SAX) k-space data in two separate but identical data streams due to our empirical observation during

experimentation that training two models on separate LAX and SAX data yields better results.

3.1 GNA-UNET with Group Normalization and GCT Attention

In the medical imaging domain, a UNET model is probably the most well known supervised deep learning architecture that was initially introduced by Ronneberger et al. in 2015 [3] and since then has found its applications in many downstream tasks such as image segmentation and image to image translation [6]. Furthermore, the concept of UNET is essential to recent Transformer architectures [5] and Denoising Diffusion Probabilistic Models (DDPM) [4]. In a nutshell, it is a symmetric encoder-decoder architecture consisting of contracting and an expanding branches and skip connections that enable sharing of information. Another remarkable property of UNET is that this model is known to be able to perform well with limited training data [31]. Here, we work with 2D slices and build a 2D GNA-UNET with the configuration outlined in Table 1. In particular, we design a GNA-UNET with five downsampling/upsampling stages.

Table 1. GNA-UNET configuration developed in this work.

Module	**Shape (in_channels, out_channels)**
encoder block	(1,64)
encoder block	(64,128)
encoder block	(128,256)
encoder block	(256,512)
encoder block	(512,1024)
conv_block	(1024,2048)
decoder block	(2048,1024)
decoder block	(1024,512)
decoder block	(512,256)
decoder block	(256,128)
decoder block	(128,64)
output	(64,1)

Our GNA-UNET model is first enhanced by the Group Normalization (GN) layers [30]. This is motivated by the fact that we are using a small batch_size=2 in our experiments. As has been shown in [30], GN outperforms commonly used Batch Normalization (BN) for small batch sizes. In particular, GN divides the channels into groups and computes within each group the mean and variance for normalization. Our Ablation study in Section 4.6 demonstrates the effectiveness of GN as compared to BN.

Second, we introduce Gated Channel Transformation (GCT) attention layers into GNA-UNET. GCT attention has been shown to improve the discriminability of deep CNNs by leveraging the relationship among channels

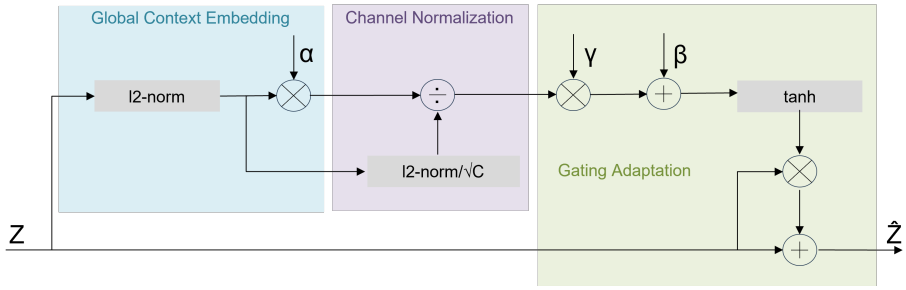

Fig. 2. Simplified diagram of the Gated Channel Transformation (GCT) attention layer. The weight parameter α controls the weight of each channel before the channel normalization. The gating weight and bias, γ and β, are responsible for adjusting the scale of the input feature Z channel-wisely. Image source: adapted from [32].

[32]. In particular, we add GCT layers after each convolutional **conv2d** layer and in the skip connections. This configuration has been determined by our Ablation study in Section 4.6. The overview diagram of GCT is given in Fig.2 where α, β and γ are trainable parameters. As shown by the authors, GCT layer inserted far away from the network output, reduces the variance of input features and thus encourages cooperation among channels and avoids excessive activation values or loss of useful features. The GCT consists of *Global Context Embedding* module that exploits global contextual information ooutside the small receptive fields of convolutional layers. The trainable embedding weight α controls the weight of each channel. The *Channel Normalization* module uses l_2-based channel normalization. The scalar $\sqrt{C}$ is used to normalize the scale, with C being the number of channels. The *Gating Adaptation* module is used to adapt the original feature. GCT can facilitate competition and cooperation during the training process by introducing the gating mechanism with *tanh* activation function. Gating weight γ and the gating bias β are trainable parameters.

To regularize our model, we further add **Dropout** layers with Dropout probability $p = 0.25$ in each of five encoder blocks. Without the **Dropout** layers, the model starts overfitting from the very beginning of the training phase.

Thus, our encoder block consists of **conv_block**(in_channels, out_channels) $\rightarrow$ **Dropout**($p = 0.25$) $\rightarrow$ **MaxPooling2d**(2,2) layers.

A convolutional block **conv_block** consists of the following layers: **conv2d** (kernel_size=3, padding=1) $\rightarrow$ **GCT** $\rightarrow$ **GN**(ng) $\rightarrow$ **conv2d** (kernel_size=3, padding=1) $\rightarrow$ **GCT** $\rightarrow$ **GN**(ng) $\rightarrow$ **ReLU**. Where **GN**(ng) is a group normalization layer with the number of groups hyperparameter $ng = 8$ determined in Section 4.6, and **ReLU** is the Rectified Linear Unit activation function.

The decoder block mirrors the encoder block in reverse and consists of the combination of the transposed convolution layer **ConvTranspose2d** (kernel_size=2, stride=2, padding=0) $\rightarrow$ **conv_block**.

The total number of learnable GNA-UNET parameters is 124,427,137.

3.2 Loss function

We use the Mean Square Error (MSE) as the objective function to update the model parameters:

$$Loss_{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{1}$$

where N is the number of pixels in the image, and y_i and $\hat{y}_i$ represent the target and reconstructed images.

4 Methodology

In order to be able to obtain quantitative validation metrics, we perform our experiments on the training and validation sets of CMRxRecon challenge data which contain the long-axis and short-axis fully sampled and undersampled images of 120 and 60 patients, respectively. We do not use the sampling masks.

4.1 Preprocessing

We process 1-channel 2D greyscale images in the spatial domain first and after performing the IFFT, we rescale by the maximum image intensity and then apply the linear scaling transform to map the pixel intensity I into $\overline{I} \in [0, 1]$:

$$\overline{I} = \frac{(I - I_{min})}{(I_{max} - I_{min})} \tag{2}$$

While the IFFT operates on the original k-space resolutions, we resize all reconstructed images to 512×512 pixels as an input to the GNA-UNET model.

4.2 Dataset

A total of 300 healthy volunteers from a single center were included in this study. The released dataset [34] includes 120 volunteers for training data, 60 for validation data and 120 for test data. Training data includes fully sampled k-space data, auto-calibration lines (ACS, 24 lines) and undersampled k-space with acceleration factors $\times 4$, $\times 8$ and $\times 10$.

4.3 Implementation and training details

The processing pipeline was implemented in Python 3.9 and the DL open source library Pytorch 2.0.1. All experiments were performed on a desktop computer with the Ubuntu operating system 18.04.3 LTS with the Intel(R) Core(TM) i9-9900K CPU, Nvidia GeForce RTX 2080 Ti GPU, and a total of 62GB RAM.

We train the GNA-UNET for 300 epochs using an AdamW optimizer [33] with a learning rate of $lr = 0.001$. We use the batch size of 2. No data augmentation was used in the training phase.

4.4 Performance evaluation metrics

Mean Square Error (MSE) is the most widely used image quality assessment metric with better values closer to zero. The MSE and the normalized MSE (NMSE) between two images $\hat{y}$ and y are defined as follows:

$$MSE = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}[\hat{y}(n,m) - y(n,m)]^2 \tag{3}$$

$$NMSE = 1 - \frac{||y - \hat{y}||_2}{||y - \overline{y}||}, \overline{y} = \frac{1}{N}\sum_{i} y_i \tag{4}$$

Peak Signal to Noise Ratio (PSNR) is defined as the ratio between the maximum possible signal power and the power of the distorting noise. This ratio between two images is computed in the decibel form as [2]:

$$PSNR = 10\log_{10}(peakval^2)/MSE \tag{5}$$

where *peakval* is the maximum possible intensity value in an image.

The Structural Similarity Index Measure (SSIM) is a perception-based model that captures the mutual dependencies among adjacent pixels to assess the similarity of two images, such as brightness, contrast and structural properties [1]:

$$SSIM(y,\hat{y}) = \frac{(2\mu_y\mu_{\hat{y}} + c_1)(2\sigma_{y\hat{y}} + c_2)}{(\mu_y^2 + \mu_{\hat{y}}^2 + c_1)(\sigma_y^2 + \sigma_{\hat{y}}^2 + c_2)} \tag{6}$$

where μ_y and $\mu_{\hat{y}}$ represent the mean values of the model output $\hat{y}$ and the target output y, σ_y and $\sigma_{\hat{y}}$ denote the corresponding pixel variance values and $\sigma_{y\hat{y}}$ is the covariance value. In order to stabilize the division, and with $P = \max(y) - \min(y)$, the constants c_1 and c_2 are defined as follows:

$$c_1 = (0.01P)^2, c_2 = (0.03P^2) \tag{7}$$

4.5 Experimental results

To allow fast experimentation, we do not use all training data provided but construct two training subsets as follows. For subset $\mathbb{S}_1$ (single coil) and subset $\mathbb{M}_1$ (multi coil), we process all timeframes and sz slices in the supplied `cine_lax.mat` and `cine_sax.mat` files. We further randomly subsample each full training set and select 1,000 LAX and 1,000 SAX images for training.

Table 2 shows our best results on the validation set provided by the challenge platform while Table 3 shows validation results obtained by using a simple denoising Autoencoder model (DAE) with the encoder-bottleneck-decoder architecture, i.e. without skip connections. Two large DAE models each with 131,282,057 learnable parameters were trained on 128×128 resized LAX and SAX images. The comparison between Tables 2 and 3 shows that GNA-UNET based processing outperforms the DAE model by a large margin. We think that

Table 2. Our best results on the CMRxRecon validation set using GNA-UNET.

Modality/Metric	SSIM ↑	NMSE ↓	PSNR ↑	Modality/Metric	SSIM ↑	NMSE ↓	PSNR ↑
Single_LAX_04	0.5505	0.1514	21.9632	Multi_LAX_04	0.5933	0.1358	22.5673
Single_LAX_08	0.5683	0.1506	22.1463	Multi_LAX_08	0.6121	0.1285	23.3891
Single_LAX_10	0.5677	0.1536	22.0728	Multi_LAX_10	0.6323	0.1399	21.8900
Single_SAX_04	0.6003	0.2004	22.1731	Multi_SAX_04	0.5903	0.2017	22.0045
Single_SAX_08	0.6116	0.1995	22.3188	Multi_SAX_08	0.6520	0.1541	22.9344
Single_SAX_10	0.6153	0.1991	22.3178	Multi_SAX_10	0.6699	0.2087	23.3594
Cine average	0.5856	0.1757	22.1653	Cine average	0.6250	0.1614	22.6908

Table 3. Results on the CMRxRecon validation set (single coil) by using the DAE model (trained for 100 epochs) as a denoising CNN.

Modality/Metric	SSIM ↑	NMSE ↓	PSNR ↑
Single_LAX_04	0.3047	0.7520	15.6542
Single_LAX_08	0.3054	0.7422	15.6861
Single_LAX_10	0.3051	0.7448	15.6745
Single_SAX_04	0.3218	1.1617	14.6089
Single_SAX_08	0.3227	1.1517	14.6522
Single_SAX_10	0.3228	1.1466	14.6672
Cine average	0.3137	0.9498	15.1572

this is due to the absence of encoder-decoder skip connections. The results for the multi coil data are some way off the top of the leaderboard, and we are currently trying to better understand why this might be the case with a view to solving this problem in future work.

Selected qualitative results are shown in Fig. 3. While working with the single coil data on a subset of the training set, we obtained high $PSNR = 35.8596dB$ however when evaluating on the validation set provided by the challenge platform, our best $PSNR = 22.3188dB$ was much lower. This shows that despite the regularization, our model does not yet generalize well to the new unseen data.

Our results on the test set are: $PSNR = 34.4493dB$, $SSIM = 0.921$ and $NMSE = 0.0563$.

4.6 Ablation studies

We perform two ablations studies to first demonstrate the effectiveness of the Group Normalization (GN) and second the effectiveness of the Gated Channel Transformation (GCT) attention unit. Both ablation studies were performed on a random subset of the TrainingSet of the CMRxRecon dataset with the 70%-10%-20% non-overlapping training-validation-test split.

We tune the hyperparameters of the Group Normalization layers and report the best result based on five training runs. The base model in Table 4 refers to the encoder-decoder structure with skip connections without the Batch Normalization (BN) layers. It can be seen that the model performs poorly without any

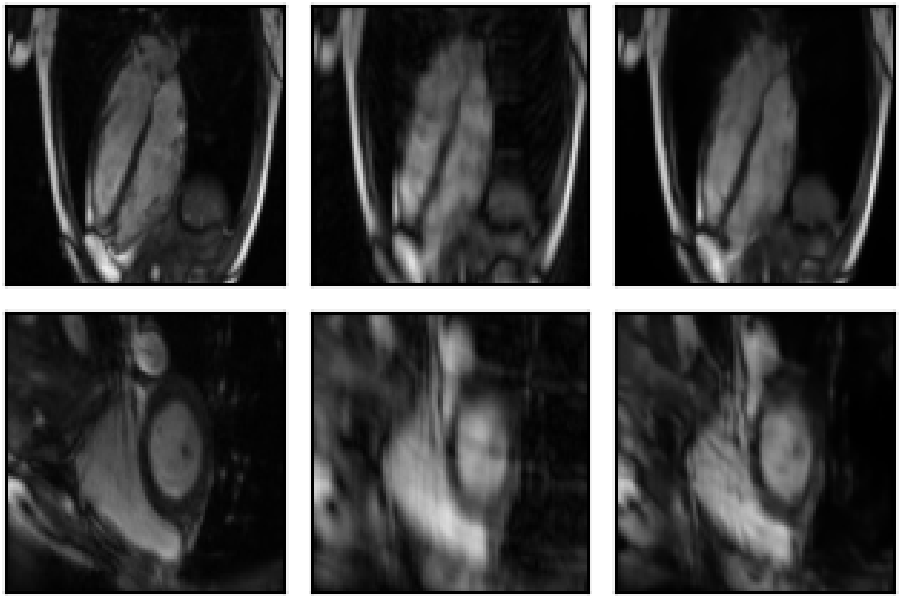

Fig. 3. Selected qualitative results (single coil, patient P002). First row: long-axis images. Second row: short-axis images. **Left**: ground truth; **Middle**: undersampled AccFactor10; **Right**: reconstructed images using our GNA-UNET pipeline.

type of normalization. The addition of BN layers results in $12.66dB\ PSNR$ gain while the addition of GN layers results in $14.56dB\ PSNR$ gain (with $ng = 8$) relative to the base model. Interestingly, the best performing hyperparameter $ng = 8$ is fixed for all convolutional blocks. We have also investigated such configuration as $2-4-8-16-32-64$ in the encoder with the reverse configuration $64-32-16-8-4-2$ in the decoder. This type of configuration resulted in a slightly worse performance. The addition of GCT layers results in a marginal improvement over the use of GN layers as can be seen in Table 5. The position of GCT layers, as well as the normalization norm and the gating function influence performance.

5 Conclusion

In this paper, we report the research carried out as part of our participation in the exciting CMRxRecon challenge, Task 1: Cine reconstruction. While not reaching the top of the leaderboard, we achieved competitive results on the validation set using a double-stream processing pipeline including the IFFT and a denoising GNA-UNET model with Group Normalization and GCT attention layers. We also verified that GNA-UNET outperforms the denoising Autoencoder by a large margin ($7.5336dB\ PSNR$ gain). Our ongoing work is centered around the use

Table 4. Ablation study to demonstrate the effectiveness of Group Normalization (GN) layers instead of Batch Normalization (BN) layers. All investigated fixed number of groups (ng) values consistently outperform BN. The best value $ng = 8$ corresponds to $1.9dB$ $PSNR$ gain relative to the use of BN.

Configuration	SSIM ↑	MSE ↓	PSNR ↑
base model	0.3759	0.00781	21.2789
base model with Batch Normalization	0.9152	0.0004722	33.9389
base model with Group Normalization			
$ng = 2$	0.9322	0.0003059	35.7680
$ng = 4$	0.9333	0.0003118	35.6037
$ng = 8$	**0.9335**	**0.0002984**	**35.8393**
$ng = 16$	0.9244	0.0003304	35.2098
$ng = 32$	0.9302	0.0003147	35.4652
$ng = 64$	0.9285	0.0003631	34.8299
$ng = 2 - 4 - 8 - 16 - 32 - 64$	0.9257	0.0003041	35.6154
$ng = 64 - 32 - 16 - 8 - 4 - 2$	0.9285	0.0003816	34.7310

Table 5. Ablation study to demonstrate the effectiveness of GCT layers.

Position	SSIM ↑	MSE ↓	PSNR ↑
after each **conv2d** layer	0.9333	0.0002983	35.8570
only in the skip connections	0.9326	0.0003062	35.8225
only in the bottleneck	0.9318	0.0003412	35.4324
after each **conv2d** layer and in the skip connections	**0.9346**	**0.0002969**	**35.8596**

Table 6. Our model information. Performance on a training set shows PSNR, SSIM and MSE values while performance on a validation set shows PSNR, SSIM and NMSE values. Inference time is given per patient and includes LAX and SAX processing time.

Task of participation	Task1: Cine	Use of pre-training	No
University/organization	Insight SFI Centre	Data augmentation	No
Single- or multi-channel	Single	Data standardization	No
Hardware configuration	RTX 2080 Ti GPU	Model parameters	124,427,137
Training time	9 hours	Loss function	MSE
Inference time	13.27 sec	Physical model	No
Performance on a train. set	(35.86, 0.93, 0.00029)	Use of unrolling	No
Performance on a val. set	(22.17, 0.59, 0.17)	k-space fidelity	No
Docker submitted?	Yes	Model backbone	UNET
Use of segmentation labels	No	Operations	Amplitude

of generative denoising diffusion probabilistic models and k-space interpolation methods for the cine reconstruction on the dataset provided.

Acknowledgements This publication has emanated from research conducted with the financial support of Science Foundation Ireland under Grant numbers 18/CRT/6183 and 12/RC/2289_P2.

References

1. Wang Z., Bovik A.C., Sheikh H.R., Simoncelli E.P. Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing; 13:600–612. doi: 10.1109/TIP.2003.819861 (2004)
2. Sara U., Akter M., Uddin M.S. Image Quality Assessment through FSIM, SSIM, MSE and PSNR—A Comparative Study. Journal of Computer and Communications; 7(3) (2019)
3. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. MICCAI (2015)
4. Yuan X., Li L., Wang J., Yang Z., Lin K., Liu Z., Wang L. Spatial-Frequency U-Net for Denoising Diffusion Probabilistic Models. arXiv (27 July 2023)
5. Hatamizadeh A., Tang Y., Nath Y., Yang D., Myronenko A., Landman B., Roth H., Xu D. UNETR: Transformers for 3D Medical Image Segmentation. IEEE WACV Conference (2022)
6. Kalantar R., Messiou C., Winfield J. M., Renn A., Latifoltojar A., Downey K., Sohaib A., Lalondrelle S., Koh D-M., Blackledge M. D. CT-Based Pelvic T1-Weighted MR Image Synthesis Using UNet, UNet++ and Cycle-Consistent Generative Adversarial Network (Cycle-GAN). Frontiers in Oncology Front 11: 665807. (2021)
7. Ding Q., Zhang, X. MRI Reconstruction by Completing Under-sampled K-space Data with Learnable Fourier Interpolation. MICCAI (2022)
8. Zhao Z., Zhang T., Xie W., Wang Y., Zhang Y. K-Space Transformer for Undersampled MRI Reconstruction. BMVC (2022)
9. Versteeg E., Klomp D. W. J., Siero J. C. W. Accelerating Brain Imaging Using a Silent Spatial Encoding Axis. Magnetic Resonance in Medicine 88(4) (2022)
10. Kojima S., Shinohara H., Hashimoto T., Suzuki S. Undersampling patterns in k-space for compressed sensing MRI using two-dimensional Cartesian sampling. Radiological Physics and Technology 11(3):303-319 (2018)
11. McGibney G., Smith M. R., Nichols S. T., Crawley A. Quantitative evaluation of several partial Fourier reconstruction algorithms used in MRI. Magnetic Resonance Medicine 30(1):51-9 (1993)
12. Sriram A., Zbontar J., Murrell T., Defazio A., Zitnick C. L., Yakubova N., Knoll F., Johnson P. End-to-End Variational Networks for Accelerated MRI Reconstruction. MICCAI (2020) pp. 64-73.
13. Wang S., Su Z., Ying L., Peng X., Zhu S., Liang F., Feng D., Liang D. Accelerating magnetic resonance imaging via deep learning. IEEE 13th International Symposium on Biomedical Imaging (ISBI), pp. 514-517, doi: 10.1109/ISBI.2016.7493320. (2016)
14. Kwon K., Kim D., Park H. A parallel MR imaging method using multilayer perceptron. Medical Physics 44(12):6209-6224 doi: 10.1002/mp.12600 (2017)
15. Lee D., Yoo J., Tak S., Ye J.C. Deep residual learning for accelerated MRI using magnitude and phase networks. IEEE Transactions on Biomedical Engineering 65(9):1985-1995 (2018)
16. Hyun C. M., Kim H. P., Lee S. M., Lee S., Seo J. K. Deep learning for undersampled MRI reconstruction. Physics in Medicine and Biology 63(13) (2018)
17. Yang G., Yu S., Dong H., Slabaugh G., Dragotti P. L., Ye X., Liu F., Arridge S., Keegan J., Guo Y., Firmin D. DAGAN: Deep De-Aliasing Generative Adversarial Networks for Fast Compressed Sensing MRI Reconstruction. IEEE Transactions on Medical Imaging 37(6):1310-1321 (2018)
18. Akçakaya M., Moeller S., Weingärtner S., Uğurbil K., Scan-specific robust artificial-neural-networks for k-space interpolation (RAKI) reconstruction: Database-free

deep learning for fast imaging. Magnetic Resonance in Medicine 81(1):439-453 (2019)
19. Arora S., Roeloffs V., Lustig M. Untrained modified deep decoder for joint denoising parallel imaging reconstruction. ISMRM and SMRT Virtual Conference and Exhibition (2020)
20. Yoo J., Jin K. H., Gupta H., Yerly J., Stuber M., Unser M. Time-dependent deep image prior for dynamic MRI. IEEE Transactions on Medical Imaging 40(12):3337-3348 (2021)
21. Darestani M. Z., Heckel R. Accelerated MRI With Un-Trained Neural Networks. IEEE Transactions on Computational Imaging 7:724-733 (2021)
22. 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 37(2):491-503 (2017)
23. Muckley M. J., Riemenschneider B., Radmanesh A., Kim S., Jeong G., Ko J., Jun Y., Shin H., Hwang D., Mostapha M., Arberet S., Nickel D., Ramzi Z., Ciuciu P., Starck J.-L., Teuwen J., Karkalousos D., Zhang C., Sriram A., Huang Z., Yakubova N., Lui Y. W., Knoll F. Results of the 2020 fastMRI Challenge for Machine Learning MR Image Reconstruction. IEEE Transactions on Medical Imaging 40(9):2306-2317 (2021)
24. Ramzi Z., G. R. C, Starck J.-L., Ciuciu P. NC-PDNet: A Density-Compensated Unrolled Network for 2D and 3D Non-Cartesian MRI Reconstruction. IEEE Transactions on Medical Imaging 41(7):1625-1638 (2022)
25. Fabian Z., Heckel R., Soltanolkotabi, M. Data augmentation for deep learning based accelerated MRI reconstruction with limited data. Proceedings of the 38th International Conference on Machine Learning (ICML) pp. 3057-3067 (2021)
26. Desai A. D., Gunel B., Ozturkler B. M., Beg H., Vasanawala S., Hargreaves B., Ré C., Pauly J., Chaudhari A. Vortex: Physics-driven data augmentations for consistency training for robust accelerated MRI reconstruction. International Conference on Medical Imaging with Deep Learning (MIDL) (2021)
27. E. J. Bekkers. B-spline CNNs on lie groups. arXiv preprint arXiv:1909.12057 (2019)
28. Gunel B., Sahiner Arda., Desai A. D., Chaudhari A. S., Vasanawala S., Pilanci M., Pauly John. Scale-equivariant unrolled neural networks for data-efficient accelerated MRI reconstruction. MICCAI. pp. 737-747 (2022)
29. Wimmer T., Golkov V., Dang H. N., Zaiss M., Maier A., Cremers D. Scale-Equivariant Deep Learning for 3D Data. arXiv preprint arXiv:2304.5864 (2023)
30. Wu Y. and He K. Group Normalization. ECCV (2018)
31. Ali O., Ali H., Ayaz Ali Shah S., Shahzad A. Implementation of a Modified U-Net for Medical Image Segmentation on Edge Devices. IEEE Transactions on Circuits and Systems II: Express Brief (2022)
32. Yang Z., Zhu L., Wu Y., Yang Y. Gated Channel Transformation for Visual Recognition. CVPR (2020)
33. Loshchilov I., Hutter F. Decoupled Weight Decay Regularization. ICLR (2019)
34. Wang C. et al. CMRxRecon: An open cardiac MRI dataset for the competition of accelerated image reconstruction. https://arxiv.org/abs/2309.10836 (2023)