Supplementary Material for Compressed Sensing MRI Reconstruction with Co-VeGAN: Complex-Valued Generative Adversarial Network

Bhavya Vasudeva[1]*, Puneesh Deora[1]*, Saumik Bhattacharya[2], Pyari Mohan Pradhan[3]
[1] ISI Kolkata [2] IIT Kharagpur [3] IIT Roorkee

1. Introduction

This document is organized as follows: Section 2 contains the details of complex-valued operations, Section 3 presents the visual schematic of the network architecture, Section 4 contains the details of the training settings, Section 5 compares the number of trainable parameters in real-valued and equivalent complex-valued networks, Section 6 contains the qualitative results of the ablation study and comparison of various activation functions, Section 7 contains the learning curves, Section 8 compares the reconstruction time of the proposed approach with other deep learning (DL) based methods, Section 9 presents the ablation study of the loss function, Section 10 shows the high-resolution versions of the qualitative results for MICCAI 2013 and fastMRI datasets and the reconstruction results for the MRNet dataset, Section 11 shows the images of zero-filled reconstruction, Section 12 presents the performance of our method for various sampling ratios and patterns, and Section 13 contains the qualitative results of zero-shot inference.

2. Complex-valued Operations

Convolution: The complex-valued equivalent of real-valued 2D convolution for a complex-valued kernel $\mathbf{W} = \mathbf{W_R} + i\mathbf{W_I}$ with complex-valued feature maps $\mathbf{F} = \mathbf{F_R} + i\mathbf{F_I}$, is given by $\mathbf{A} = \mathbf{W} * \mathbf{F} = \mathbf{A_R} + i\mathbf{A_I}$, where $*$ denotes convolution operation, i denotes the imaginary unit, and

$$\begin{aligned} \mathbf{A_R} &= \mathbf{W_R} * \mathbf{F_R} - \mathbf{W_I} * \mathbf{F_I}, \\ \mathbf{A_I} &= \mathbf{W_R} * \mathbf{F_I} + \mathbf{W_I} * \mathbf{F_R}, \end{aligned} \tag{1}$$

similar to complex-valued multiplication. The subscripts $\mathbf{R}$ and $\mathbf{I}$ denote the real and imaginary parts of the complex-valued entities, respectively. To implement this, we make use of real-valued tensors, where $\mathbf{W}$ ($\mathbf{F}$) is stored by concatenating $\mathbf{W_I}$ ($\mathbf{F_I}$) to $\mathbf{W_R}$ ($\mathbf{F_R}$). The resultant includes four real-valued 2D convolutions as defined in (1), and is stored in a similar manner by concatenating $\mathbf{A_I}$ to $\mathbf{A_R}$.

*Equal contribution; Work done while at IIT Roorkee.

Backpropagation: Backpropagation can be performed on a function f that is non-holomorphic as long as it is differentiable with respect to its real and imaginary parts [3]. Since all the loss functions considered in this work are real-valued, we consider f to be a real-valued function of l-dimensional weight vector $\mathbf{w}$. The update rule of $\mathbf{w}$ using gradient descent is written as:

$$\mathbf{w} = \mathbf{w} - \rho \nabla_{\bar{\mathbf{w}}} f(\mathbf{w}), \tag{2}$$

where ρ is the learning rate (LR), $\bar{\mathbf{w}}$ denotes the complex conjugate of $\mathbf{w}$, and the gradient of f is calculated as:

$$\begin{aligned} \nabla_{\bar{\mathbf{w}}} f(\mathbf{w}) &= \Big[\frac{\partial f}{\partial \bar{w_1}} \cdots \frac{\partial f}{\partial \bar{w_l}}\Big]^T \\ \frac{\partial f}{\partial \bar{w}} &= \frac{1}{2}\Big(\frac{\partial f}{\partial w_R} + i\frac{\partial f}{\partial w_I}\Big). \end{aligned} \tag{3}$$

Batch Normalization: We make use of the complex batch normalization (CBN), proposed in [4]. To ensure that the complex data is scaled in such a way that the distribution of real and imaginary components is circular, the 2D complex vector can be whitened as shown below:

$$\mathbf{x}_{std} = \mathbf{B}^{-\frac{1}{2}}(\mathbf{x} - \mathbb{E}[\mathbf{x}]), \tag{4}$$

where $\mathbf{B}$ denotes the covariance matrix, and $\mathbb{E}$ denotes expectation operator. $\mathbf{B}$ can be represented as:

$$\mathbf{B} = \begin{bmatrix} \mathrm{Cov}(\mathbf{x_R}, \mathbf{x_R}) & \mathrm{Cov}(\mathbf{x_R}, \mathbf{x_I}) \\ \mathrm{Cov}(\mathbf{x_I}, \mathbf{x_R}) & \mathrm{Cov}(\mathbf{x_I}, \mathbf{x_I}) \end{bmatrix}. \tag{5}$$

Learnable parameters $\boldsymbol{\gamma}$, $\boldsymbol{\beta}$ are used to scale and shift the aforementioned standardized vector as follows:

$$\mathbf{x}_{BN} = \boldsymbol{\gamma}\mathbf{x}_{std} + \boldsymbol{\beta}, \tag{6}$$

where $\boldsymbol{\gamma}$ is a 2×2 matrix, and $\boldsymbol{\beta}$ is a complex number.

3. Network Architecture

Fig. 1 shows the visual diagram of the network architecture of the proposed framework.

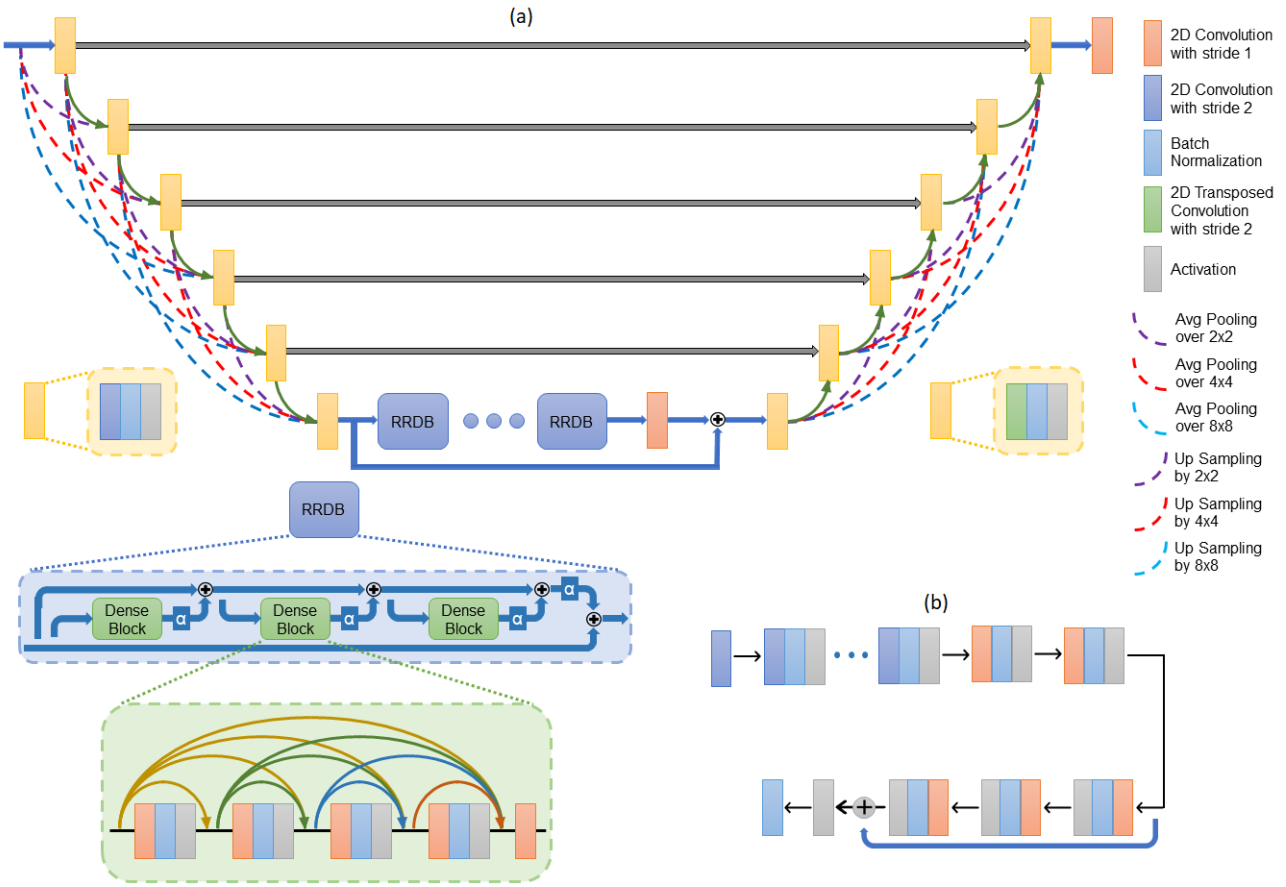

Figure 1: (a) Proposed dense U-net generator architecture with RRDBs and (b) discriminator architecture.

4. Training settings

For implementing the models, the Keras framework with TensorFlow backend is used. The models are trained using 4 NVIDIA GeForce GTX 1080 Ti GPUs. In this work, the batch size is set as 16. In the generator, each layer produces 32 feature maps. 4 residual-in-residual dense blocks (RRDBs) are present in the bottleneck layer, the number of dense blocks is set as 8, and α is 0.2. The absolute value of the discriminator weights is clipped at 0.05, and n_D is set as 3. For training the models, we use Adam optimizer [2], with $\beta_1 = 0.5$ and $\beta_2 = 0.999$. The initial learning rate is set as 10^{-4}, with a decay of 1.39×10^{-3}, so that it becomes $1/10^{th}$ of the initial value after five epochs.

5. Trainable Parameters in Real- and Complex-valued Networks

To consider the real-valued version of the fifth case as the sixth in the ablation study, each convolutional layer is converted from complex to real by doubling the number of output channels present in each layer.

For complex-valued networks, trainable parameters include both real and imaginary components of the network weights. The number of parameters of a convolutional layer is given by $W \times H \times I \times C$, where W denotes the width of the convolutional filter, H denotes the height, I denotes the number of input channels, and C denotes the number of output channels. For the complex-valued case, the number of trainable parameters would be $2 \times W \times H \times I \times C$, as each filter weight has a real and imaginary component. As mentioned in the manuscript, we doubled the number of channels in each layer in the 6^{th} case, leading to $W \times H \times 2I \times 2C = 4 \times W \times H \times I \times C$ parameters, which is twice the number of trainable parameters as compared to the 5^{th} case. These figures account for the trainable parameters due to the convolutional layers. The actual number is not exactly twice due to the trainable parameters used by the activation and batch normalization layers.

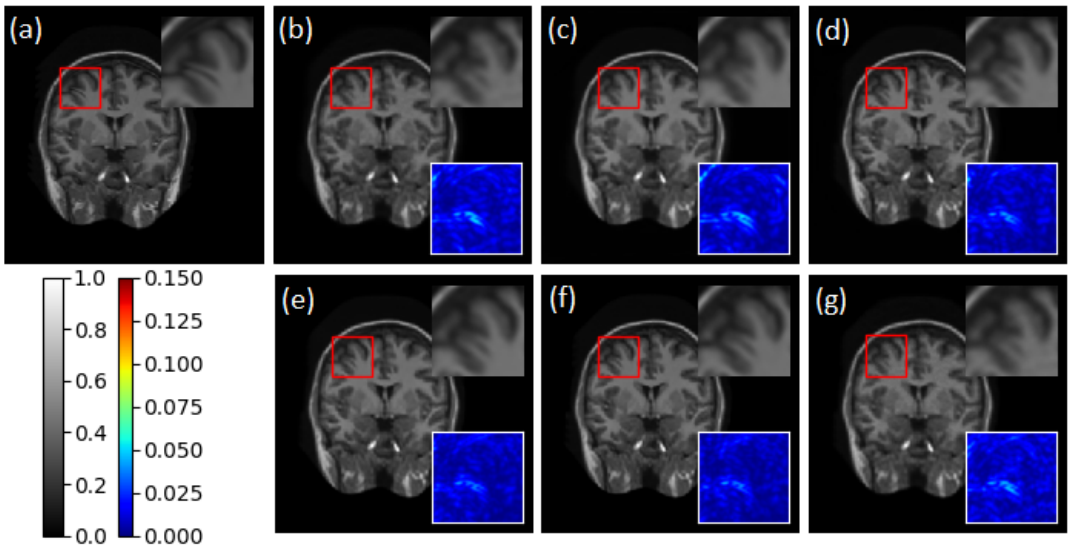

Figure 2: Qualitative results for ablation study of the proposed model. (a) GT, reconstruction results for (b) 1^{st}, (c) 2^{nd}, (d) 3^{rd}, (e) 4^{th}, (f) 5^{th}, and (g) 6^{th} network settings. Inset- top right: the ZiR enclosed by the red box, bottom right: the absolute difference between the ZiR and its corresponding GT.

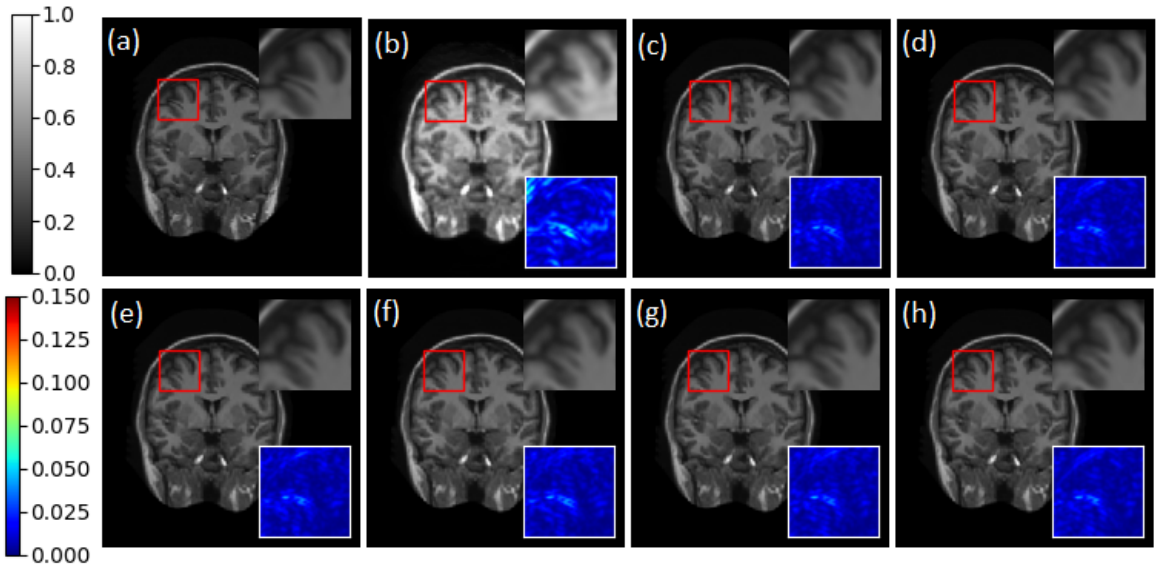

Figure 3: Qualitative comparison of various activation functions. (a) GT, reconstruction results using (b) zReLU, (c) $\mathbb{C}$ReLU, (d) $\mathbb{C}$PReLU, (e) Cardioid, (f) PP-WSS, (g) TIP-WSS, and (h) PC-WSS activation functions. Inset- top right: the ZiR enclosed by the red box, bottom right: the absolute difference between the ZiR and its corresponding GT.

6. Qualitative Results (Ablation Study and Activations)

Figs. 2 and 3 show the qualitative results for the ablation study of the proposed model and the comparison of various activation functions, respectively.

7. Learning Curves

The learning curves (loss value vs. iteration) for the proposed method are shown in Fig. 4. These values are obtained when the proposed method is trained on 30% 1D Gaussian (1D-G) undersampled images from the MICCAI 2013 dataset. It can be seen from the plots that all the losses decrease and converge within 5000 iterations.

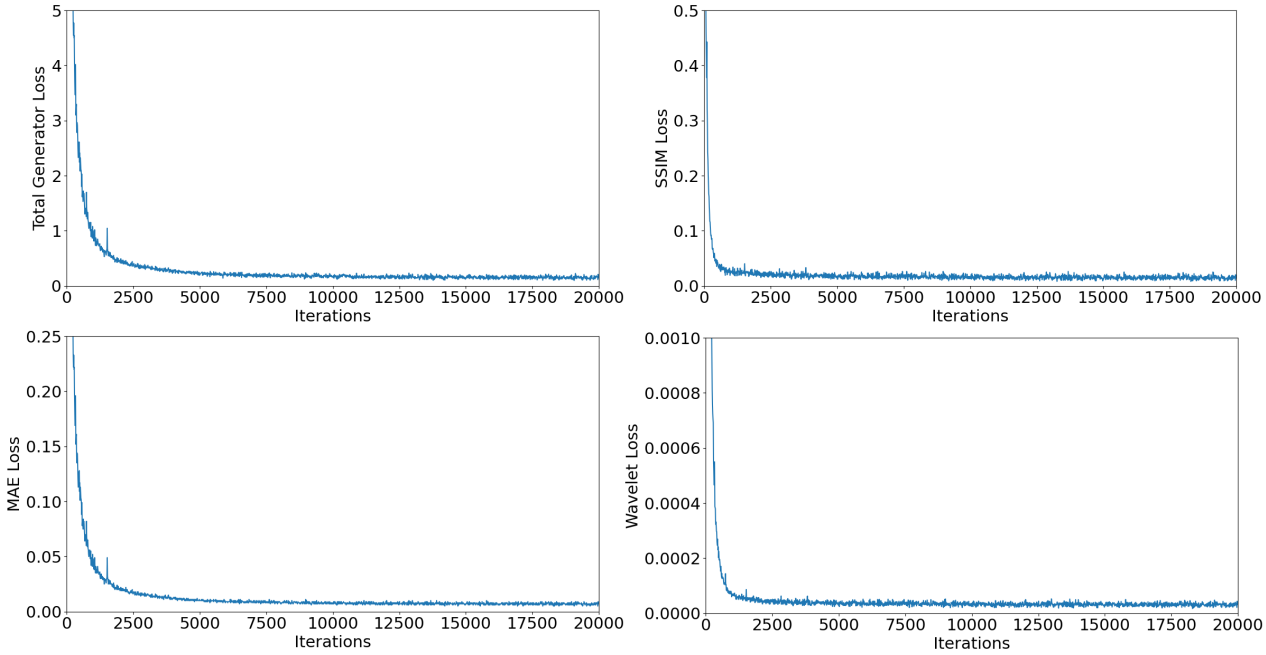

Figure 4: Learning curves for the proposed method.

Table 1: Runtimes of SOTA deep learning-based methods for CS-MRI reconstruction.

Method	DeepADMM [6]	MoDL [1]	DAGAN [5]	Co-VeGAN (Proposed)
Runtime (ms)	313.50	82.33	6.27	34.14

8. Reconstruction Time

The runtimes (time taken to reconstruct an MR image) of state-of-the-art (SOTA) DL-based methods are shown in Table 1. These values are obtained for 30% 1D-G undersampled images from the MICCAI 2013 dataset. The runtime of the proposed approach is in milliseconds, which is suitable for real-time reconstruction of MR images. Although it is second-best in this context, the proposed approach obtains superior reconstruction with significantly fewer trainable parameters, thereby lowering the storage requirement.

9. Ablation Study of the Loss Function

Table 2 shows the results of the ablation study of the loss function. These results were obtained by using a lighter version of the final model in the manuscript which has fewer trainable parameters. We use the fastMRI dataset with batch size 10 to train the model.

We see that in the absence of MAE loss, the PSNR is significantly low. This is because both SSIM and wavelet loss focus only on the reconstruction of the magnitude content. In the second case, when we add MAE loss, the PSNR value improves but the PSNR and mSSIM values are lower than the next two cases because the SSIM loss, which focuses on the reconstruction of the high-frequency content, is not used. In the third case, we observe that using the SSIM leads to an increase in both PSNR and mSSIM values. Finally, adding wavelet loss improves the reconstruction of mid-frequency content, leading to better PSNR and mSSIM values.

Table 2: Results for ablation study of the loss function.

Loss	Model 1	Model 2	Model 3	Model 4
MAE	✗	✓	✓	✓
SSIM	✓	✗	✓	✓
Wavelet	✓	✓	✗	✓
PSNR (dB)	19.446	32.864	33.747	33.887
mSSIM	0.7993	0.7240	0.7689	0.7710

10. Qualitative Results and Comparison

Figs. 5 and 6 show the high-resolution versions of the figures showing qualitative results of the proposed method and comparisons with SOTA methods using images from MICCAI 2013 and fastMRI datasets, respectively.

Fig. 7 shows the qualitative results and comparison using the MRNet dataset. The qualitative results and the differ-

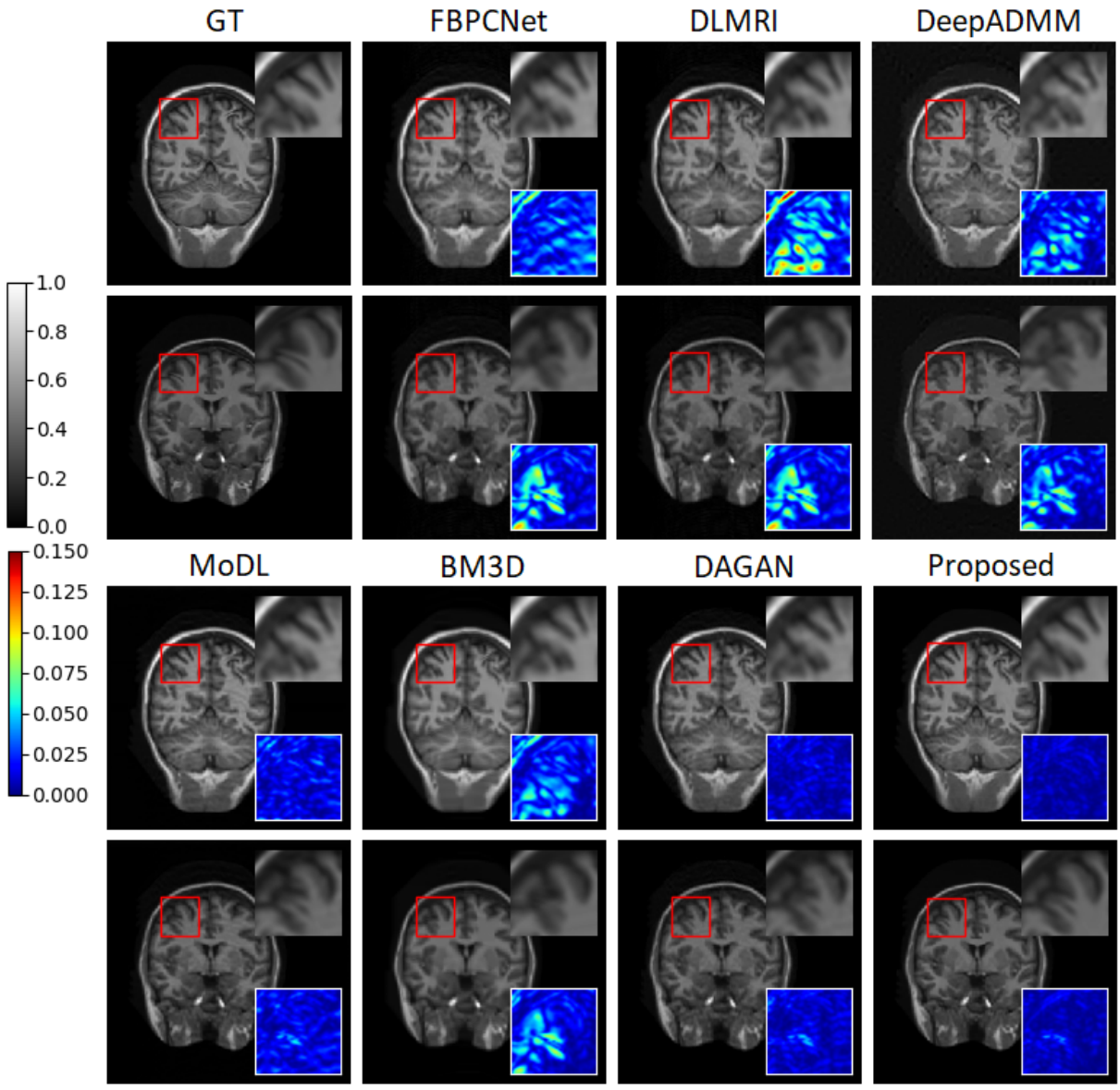

Figure 5: (Figure 3 in manuscript) Qualitative results and comparison of the proposed method for two images taken from the MICCAI 2013 dataset. Inset- top right: the ZiR enclosed by the red box, bottom right: the absolute difference between the ZiR and its corresponding GT.

ence between the reconstructed output and the ground truth (GT) demonstrate that the proposed method obtains superior reconstruction results.

11. Zero-filled Reconstruction

Figs. 8 and 9 show the qualitative results of zero-filled reconstruction (ZFR) of 30% 1D-G undersampled images from MRNet and fastMRI datasets, respectively.

12. Effect of Sampling Masks

12.1. Sampling Ratio

Figs. 10 and 11 show the qualitative results of the proposed method for various sampling ratios for 1D-G undersampled images, taken from the MICCAI 2013 and the fastMRI dataset, respectively. It is observed that the proposed approach can obtain high-quality reconstructions for 20% and 30% undersampled images for both datasets. Table 3 shows the quantitative results for this set of experi-

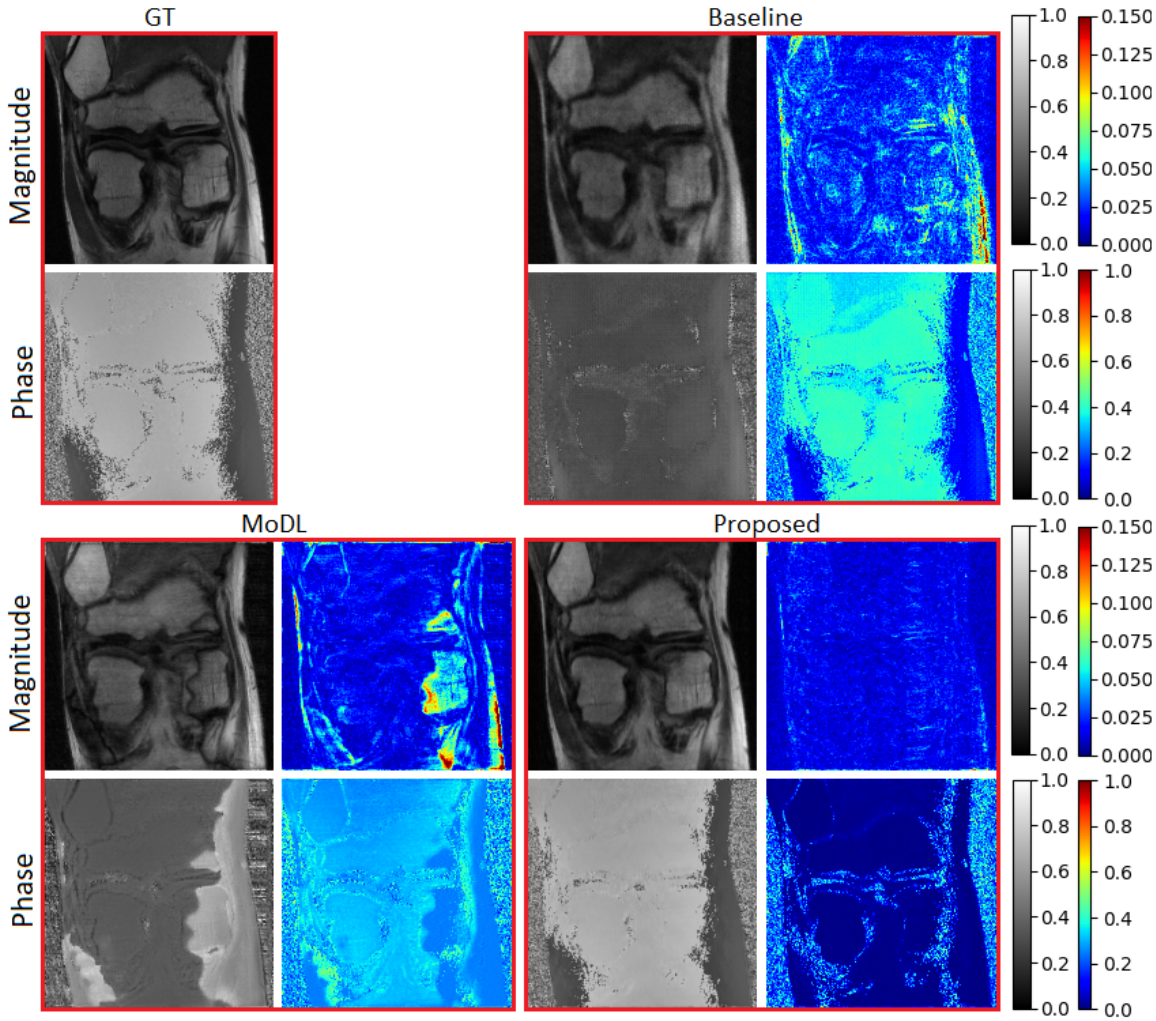

Figure 6: (Figure 4 in manuscript) Qualitative results and comparison of our method for complex-valued images from fastMRI dataset. For each method, the two columns show the reconstructed outputs, and their absolute difference with the GT.

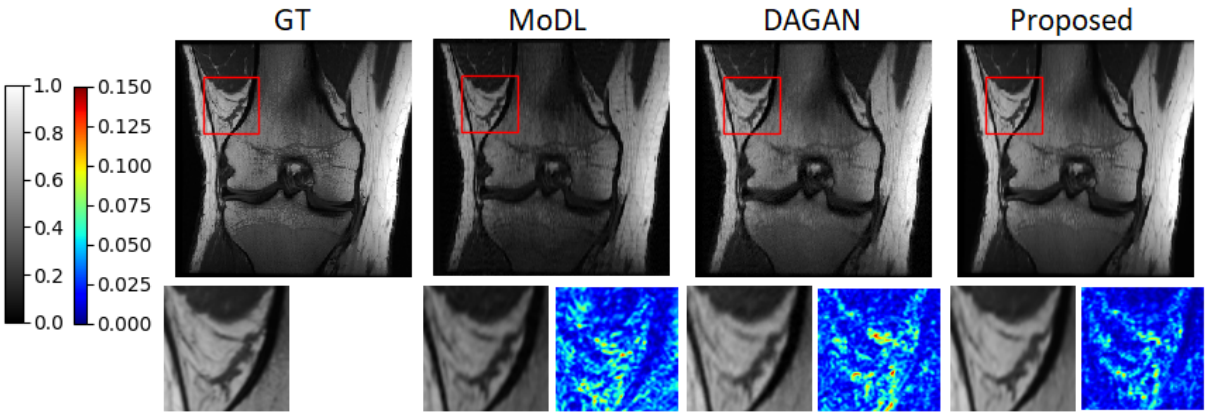

Figure 7: Qualitative results and comparison of the proposed method on images from the MRNet dataset. Inset- bottom left: the ZiR enclosed by the red box, bottom right: the absolute difference between the ZiR and its corresponding GT.

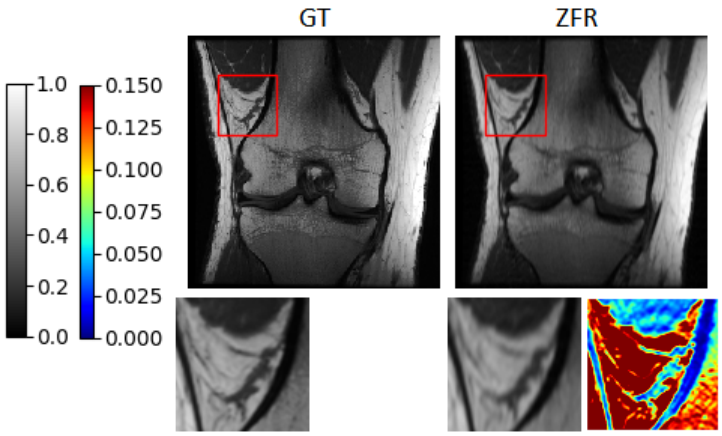

Figure 8: Qualitative results of ZFR for 30% 1D-G undersampled images from the MRNet dataset. Inset- bottom left: the ZiR enclosed by the red box, bottom right: the absolute difference between the ZiR and its corresponding GT.

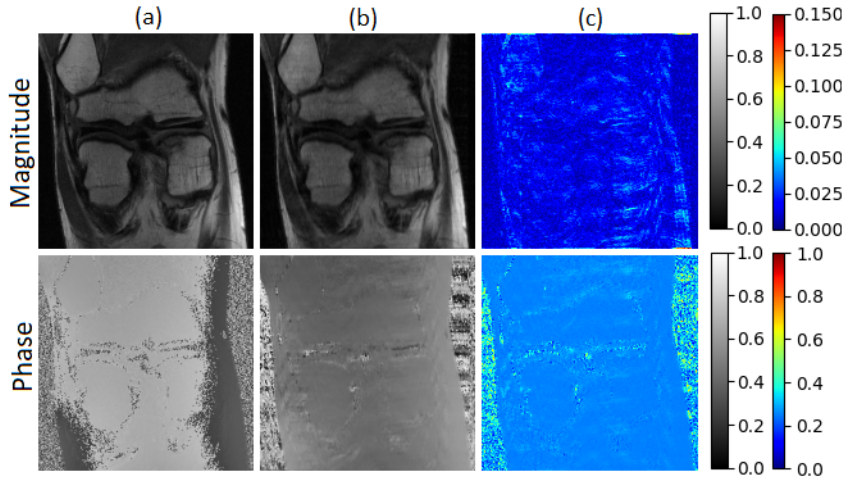

Figure 9: Qualitative results of ZFR for 30% 1D-G undersampled images from the fastMRI dataset. (a) GT, (b) ZFR, (c) the absolute difference between ZFR and GT.

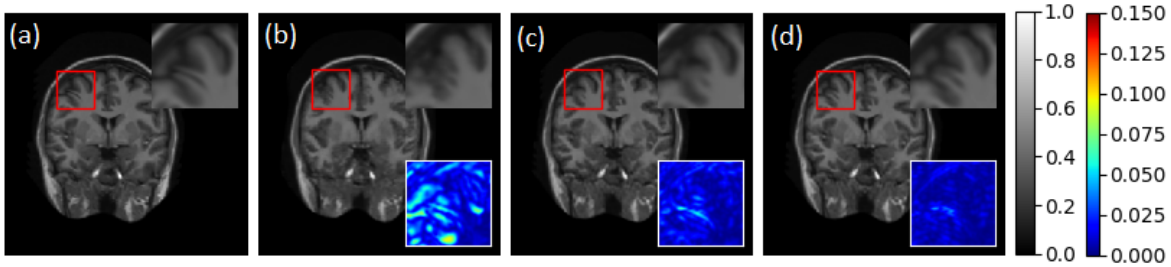

Figure 10: Qualitative results for various sampling ratios using 1D-G undersampled images from the MICCAI 2013 dataset. (a) GT, reconstruction results of the proposed approach for (b) 10%, (c) 20%, and (d) 30% sampling ratios. Inset- top right: the ZiR enclosed by the red box, bottom right: the absolute difference between the ZiR and its corresponding GT.

ments. For the MICCAI 2013 dataset, the results obtained for 20% undersampled images are quantitatively as well as qualitatively better than the results obtained by some of the SOTA methods (Table 3 in manuscript) for 30% under-

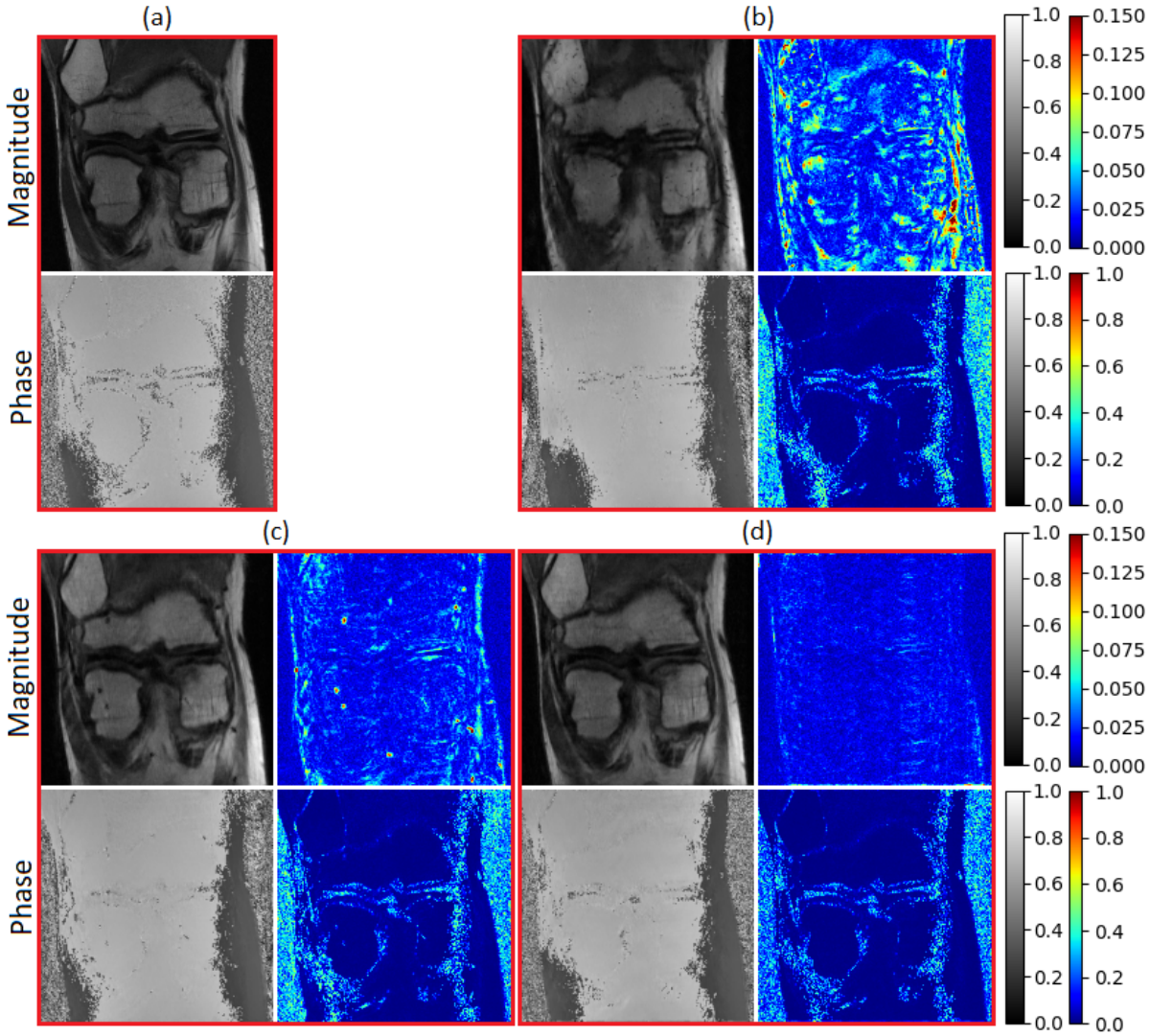

Figure 11: Qualitative results for various sampling ratios using 1D-G undersampled complex-valued images from the fastMRI dataset. (a) GT, reconstruction results of the proposed approach for (b) 10%, (c) 20%, and (d) 30% sampling ratios. For all the ratios, the two columns show the reconstructed outputs, and their absolute difference with the GT.

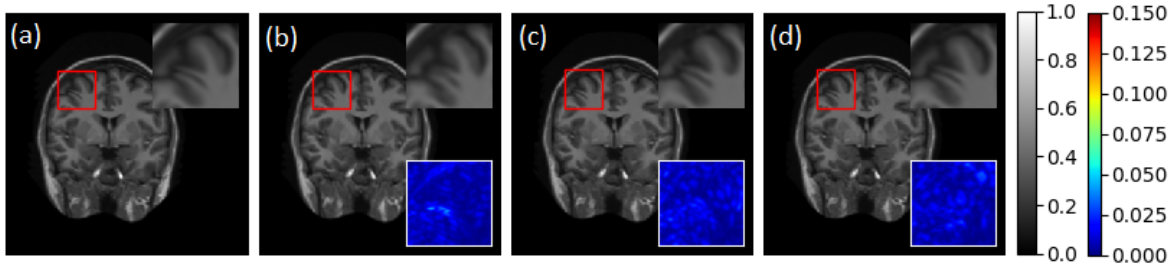

Figure 12: Qualitative results for various sampling patterns using 30% undersampled images from the MICCAI 2013 dataset. (a) GT, reconstruction results of the proposed approach for (b) 1D-G, (c) radial, and (d) spiral sampling patterns. Inset- top right: the ZiR enclosed by the red box, bottom right: the absolute difference between the ZiR and its corresponding GT.

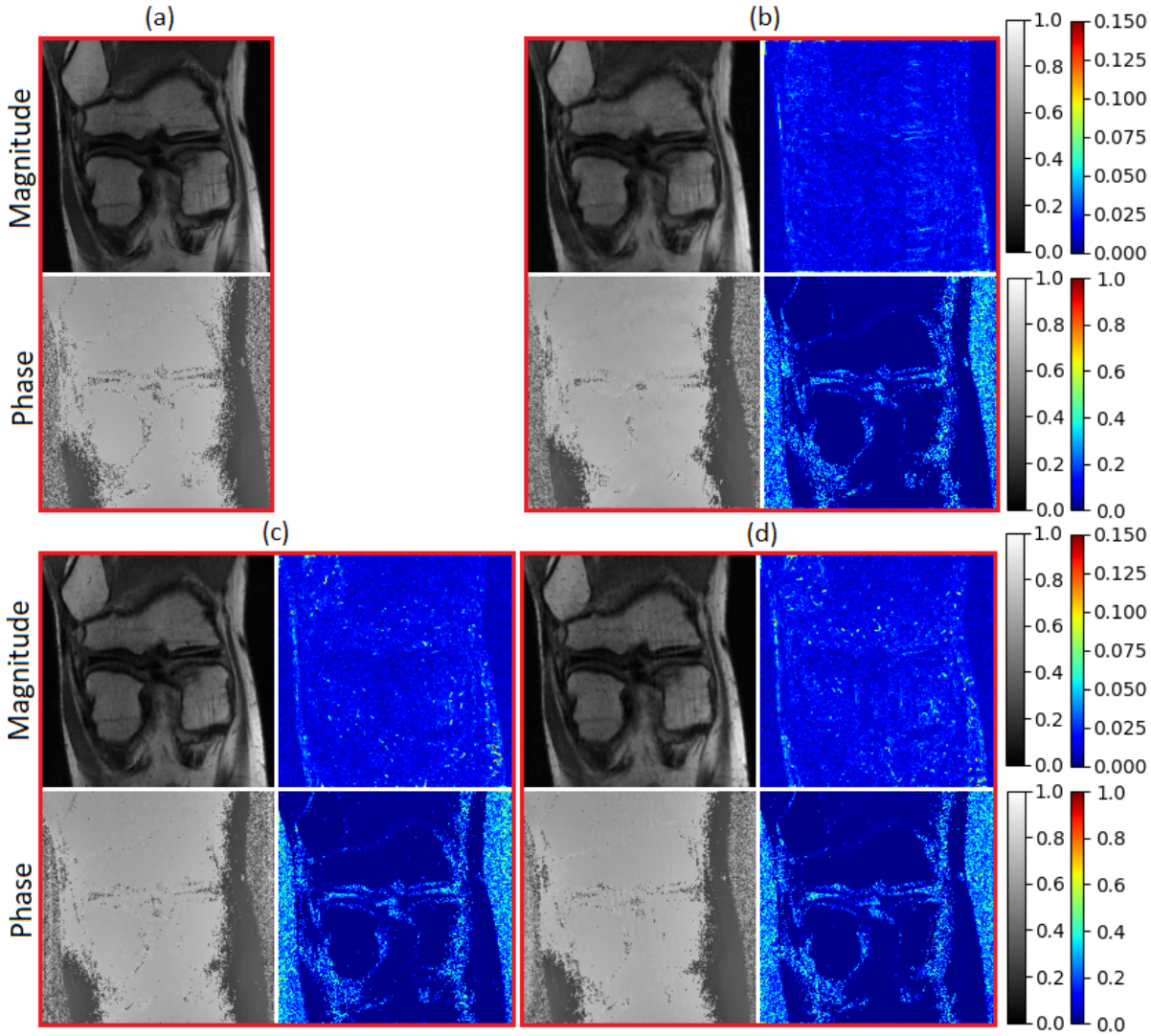

Figure 13: Qualitative results for various sampling patterns using 30% undersampled complex-valued images from the fastMRI dataset. (a) GT, reconstruction results of the proposed approach for (b) 1D-G, (c) radial, and (d) spiral sampling patterns. For all the patterns, the two columns show the reconstructed outputs, and their absolute difference with the GT.

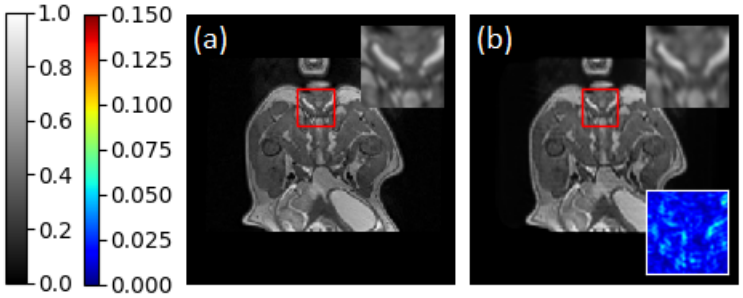

Figure 14: Qualitative results of zero-shot inference experiment. (a) GT, (b) reconstruction results of the proposed model trained on images from the MICCAI 2013 dataset. Inset- top right: the ZiR enclosed by the red box, bottom right: the absolute difference between the ZiR and its corresponding GT.

sampled images. For the fastMRI dataset, the results obtained for even 10% undersampled images are better than the results obtained by the baseline (Table 5 in manuscript) for 30% undersampled images. These observations further highlight the superior quality reconstructions achieved by the proposed approach as compared to SOTA methods. For 10% undersampled images from both the aforementioned datasets, the proposed approach is able to obtain a de-aliased output, where the contrast, as well as a significant portion of structural content, has been preserved. However, it is evident that a highly faithful reconstruction may not be achieved for this ratio in both cases. This is because the k-space has been highly undersampled, and only 10% of the data has been retained.

Table 3: Quantitative results (PSNR (dB) / mSSIM) for various sampling ratios

Dataset	Sampling Ratio		
	10%	20%	30%
MICCAI 2013	35.799 / 0.9485	41.396 / 0.9817	45.678 / 0.9927
fastMRI	30.596 / 0.6673	32.512 / 0.7356	34.538 / 0.7893

Table 4: Quantitative results (PSNR (dB) / mSSIM) for various sampling patterns

Dataset	Sampling Pattern		
	1D-G	Radial	Spiral
MICCAI 2013	45.678 / 0.9927	46.629 / 0.9922	46.747 / 0.9929
fastMRI	34.538 / 0.7893	34.173 / 0.7819	33.918 / 0.7745

12.2. Sampling Pattern

Figs. 12 and 13 demonstrate the visual outputs generated by the proposed approach for comparing various sampling patterns, using 30% undersampled images, from the MICCAI 2013 dataset and the fastMRI dataset, respectively. For both the datasets, highly accurate recoveries are achieved for all three sampling masks, as evident from the difference between the generated output and the GT. For the MICCAI 2013 dataset, this difference is close to zero, signifying the proposed method's ability to reconstruct the finest details as seen in the zoomed in region (ZiR). For the fastMRI dataset, the proposed method obtains high-quality reconstruction of both magnitude and phase images. These experiments show that our method generalizes well to various sampling patterns. This is also supported by the quantitative results for this set of experiments shown in Table 4.

13. Zero-shot Inference

The qualitative results of zero-shot inference are shown in Fig. 14. We can see that although the images of canine legs (used to test the model) are of a completely different anatomy compared to brain (used for training), our approach can obtain high-quality reconstruction.

References

[1] H. K. Aggarwal, M. P. Mani, and M. Jacob. MoDL: Model-based deep learning architecture for inverse problems. *IEEE Transactions on Medical Imaging*, 38(2):394–405, 2019.

[2] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*, 2015.

[3] K. Kreutz-Delgado. The complex gradient operator and the CR-calculus. *ArXiv*, abs/0906.4835, 2009.

[4] C. Trabelsi et al. Deep complex networks. In *International Conference on Learning Representations (ICLR)*, 2018.

[5] G. Yang et al. DAGAN: Deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction. *IEEE Transactions on Medical Imaging*, 37(6):1310–1321, 2018.

[6] Y. Yang, J. Sun, H. Li, and Z. Xu. Deep ADMM-Net for compressive sensing MRI. In *Advances in Neural Information Processing Systems*, pages 10–18. 2016.