

# End-to-End Variational Networks for Accelerated MRI Reconstruction

Anuroop Sriram<sup>1\*</sup>, Jure Zbontar<sup>1\*</sup>, Tullie Murrell<sup>1</sup>, Aaron Defazio<sup>1</sup>, C. Lawrence Zitnick<sup>1</sup>, Nafissa Yakubova<sup>1</sup>, Florian Knoll<sup>2</sup>, and Patricia Johnson<sup>2</sup>

<sup>1</sup> Facebook AI Research (FAIR)

<sup>2</sup> NYU School of Medicine

**Abstract.** The slow acquisition speed of magnetic resonance imaging (MRI) has led to the development of two complementary methods: acquiring multiple views of the anatomy simultaneously (parallel imaging) and acquiring fewer samples than necessary for traditional signal processing methods (compressed sensing). While the combination of these methods has the potential to allow much faster scan times, reconstruction from such undersampled multi-coil data has remained an open problem. In this paper, we present a new approach to this problem that extends previously proposed variational methods by learning fully end-to-end. Our method obtains new state-of-the-art results on the fastMRI dataset [18] for both brain and knee MRIs.<sup>3</sup>

**Keywords:** MRI Acceleration · End-to-end learning · Deep learning.

## 1 Introduction

Magnetic Resonance Imaging (MRI) is a powerful diagnostic tool for a variety of disorders, but its utility is often limited by its slow speed compared to competing modalities like CT or X-Rays. Reducing the time required for a scan would decrease the cost of MR imaging, and allow for acquiring scans in situations where a patient cannot stay still for the current minimum scan duration.

One approach to accelerating MRI acquisition, called Parallel Imaging (PI) [3, 8, 13], utilizes multiple receiver coils to simultaneously acquire multiple views of the underlying anatomy, which are then combined in software. Multi-coil imaging is widely used in current clinical practice. A complementary approach to accelerating MRIs acquires only a subset of measurements and utilizes Compressed Sensing (CS) [1, 7] methods to reconstruct the final image from these undersampled measurements. The combination of PI and CS, which involves acquiring undersampled measurements from multiple views of the anatomy, has the potential to allow faster scans than is possible by either method alone. Reconstructing MRIs from such undersampled multi-coil measurements has remained an active area of research.

MRI reconstruction can be viewed as an inverse problem and previous research has proposed neural networks whose design is inspired by the optimization procedure to solve such a problem [4, 6, 9, 10]. A limitation of such an approach is that it assumes the forward process is completely known, which is an unrealistic assumption for the multi-coil reconstruction problem. In this paper, we present a novel technique for reconstructing MRI images from undersampled multi-coil data that does not make this assumption. We extend previously proposed variational methods by learning the forward process in conjunction with reconstruction, alleviating this limitation. We show through experiments on the fastMRI dataset that such an approach yields higher fidelity reconstructions.

Our contributions are as follows: 1) we extend the previously proposed variational network model by learning completely end-to-end; 2) we explore the design space for the variational networks to determine the optimal intermediate representations and neural network architectures for better reconstruction quality; and 3) we perform extensive experiments using our model on the fastMRI dataset and obtain new state-of-the-art results for both the knee and the brain MRIs.

\* Equal contribution

<sup>3</sup> Code is available at <https://github.com/facebookresearch/fastMRI>.

## 2 Background and Related Work

### 2.1 Accelerated MRI acquisition

An MR scanner images a patient's anatomy by acquiring measurements in the frequency domain, called *k-space*, using a measuring instrument called a receiver coil. The image can then be obtained by applying an inverse multidimensional Fourier transform  $\mathcal{F}^{-1}$  to the measured k-space samples. The underlying image  $\mathbf{x} \in \mathbb{C}^M$  is related to the measured k-space samples  $\mathbf{k} \in \mathbb{C}^M$  as

$$\mathbf{k} = \mathcal{F}(\mathbf{x}) + \epsilon, \quad (1)$$

where  $\epsilon$  is the measurement noise and  $\mathcal{F}$  is the fourier transform operator.

Most modern scanners contain multiple receiver coils. Each coil acquires k-space samples that are modulated by the sensitivity of the coil to the MR signal arising from different regions of the anatomy. Thus, the  $i$ -th coil measures:

$$\mathbf{k}_i = \mathcal{F}(S_i \mathbf{x}) + \epsilon_i, i = 1, 2, \dots, N, \quad (2)$$

where  $S_i$  is a complex-valued diagonal matrix encoding the position dependent sensitivity map of the  $i$ -th coil and  $N$  is the number of coils. The sensitivity maps are normalized to satisfy [15]:

$$\sum_{i=1}^N S_i^* S_i = 1 \quad (3)$$

The speed of MRI acquisition is limited by the number of k-space samples obtained. This acquisition process can be accelerated by obtaining undersampled k-space data,  $\tilde{\mathbf{k}}_i = M \mathbf{k}_i$ , where  $M$  is a binary mask operator that selects a subset of the k-space points and  $\mathbf{k}_i$  denotes the measured k-space data. The same mask is used for all coils. Applying an inverse Fourier transform naively to this under-sampled k-space data results in aliasing artifacts.

Parallel Imaging can be used to accelerate imaging by exploiting redundancies in k-space samples measured by different coils. The sensitivity maps  $S_i$  can be estimated using the central region of k-space corresponding to low frequencies, called the *Auto-Calibration Signal (ACS)*, which is typically fully sampled. To accurately estimate these sensitivity maps, the ACS must be sufficiently large, which limits the maximum possible acceleration.

### 2.2 Compressed Sensing for Parallel MRI Reconstruction

Compressed Sensing [2] enables reconstruction of images by using fewer k-space measurements than is possible with classical signal processing methods by enforcing suitable priors. Classical compressed sensing methods solve the following optimization problem:

$$\hat{\mathbf{x}} = \operatorname{argmin}_{\mathbf{x}} \frac{1}{2} \sum_i \left\| M \mathcal{F}(S_i \mathbf{x}) - \tilde{\mathbf{k}}_i \right\|^2 + \lambda \Psi(\mathbf{x}) \quad (4)$$

$$= \operatorname{argmin}_{\mathbf{x}} \frac{1}{2} \left\| A(\mathbf{x}) - \tilde{\mathbf{k}} \right\|^2 + \lambda \Psi(\mathbf{x}), \quad (5)$$

where  $\Psi$  is a regularization function that enforces a sparsity constraint,  $A$  is the linear forward operator that multiplies by the sensitivity maps, applies 2D fourier transform and then under-samples the data, and  $\tilde{\mathbf{k}}$  is the vector of masked k-space data from all coils. This problem can be solved by iterative gradient descent methods. In the  $t$ -th step the image is updated from  $\mathbf{x}^t$  to  $\mathbf{x}^{t+1}$  using:

$$\mathbf{x}^{t+1} = \mathbf{x}^t - \eta^t \left( A^*(A(\mathbf{x}) - \tilde{\mathbf{k}}) + \lambda \Phi(\mathbf{x}^t) \right), \quad (6)$$

where  $\eta^t$  is the learning rate,  $\Phi(\mathbf{x})$  is the gradient of  $\Psi$  with respect to  $\mathbf{x}$ , and  $A^*$  is the hermitian of the forward operator  $A$ .

### 2.3 Deep Learning for Parallel MRI Reconstruction

In the past few years, there has been rapid development of deep learning based approaches to MRI reconstruction [4–6, 9, 10, 12]. A comprehensive survey of recent developments in using deep learning for parallel MRI reconstruction can be found in [5]. Our work builds upon the Variational Network (VarNet) [4], which consists of multiple layers, each modeled after a single gradient update step in equation 6. Thus, the  $t$ -th layer of the VarNet takes  $\mathbf{x}^t$  as input and computes  $\mathbf{x}^{t+1}$  using:

$$\mathbf{x}^{t+1} = \mathbf{x}^t - \eta^t A^*(A(\mathbf{x}^t) - \tilde{\mathbf{k}}) + \text{CNN}(\mathbf{x}^t), \quad (7)$$

where CNN is a small convolutional neural network that maps complex-valued images to complex-valued images of the same shape. The  $\eta^t$  values as well as the parameters of the CNNs are learned from data.

The  $A$  and  $A^*$  operators involve the use of sensitivity maps which are computed using a traditional PI method and fed in as additional inputs. As noted in section 2.1, these sensitivity maps cannot be estimated accurately when the number of auto-calibration lines is small, which is necessary to achieve higher acceleration factors. As a result, the performance of the VarNet degrades significantly at higher accelerations. We alleviate this problem in our model by learning to predict the sensitivity maps from data as part of the network.

## 3 End-to-End Variational Network

Let  $\mathbf{k}_0 = \tilde{\mathbf{k}}$  be the vector of masked multi-coil k-space data. Similar to the VarNet, our model takes this masked k-space data  $\mathbf{k}_0$  as input and applies a number of refinement steps of the same form. We refer to each of these steps as a *cascade* (following [12]), to avoid overloading the term "layer" which is already heavily used. Unlike the VN, however, our model uses **k-space intermediate quantities rather than image-space quantities**. We call the resulting method the End-to-End Variational Network or E2E-VarNet.

### 3.1 Preliminaries

To simplify notation, we first define two operators: the expand operator ( $\mathcal{E}$ ) and the reduce operator ( $\mathcal{R}$ ). The *expand* operator ( $\mathcal{E}$ ) takes the image  $\mathbf{x}$  and sensitivity maps as input and computes the corresponding image seen by each coil in the idealized noise-free case:

$$\mathcal{E}(\mathbf{x}) = (\mathbf{x}_1, \dots, \mathbf{x}_N) = (S_1 \mathbf{x}, \dots, S_N \mathbf{x}). \quad (8)$$

where  $S_i$  is the sensitivity map of coil  $i$ . We do not explicitly represent the sensitivity maps as inputs for the sake of readability. The inverse operator, called the *reduce* operator ( $\mathcal{R}$ ) combines the individual coil images:

$$\mathcal{R}(\mathbf{x}_1, \dots, \mathbf{x}_N) = \sum_{i=1}^N S_i^* \mathbf{x}_i. \quad (9)$$

Using the expand and reduce operators,  $A$  and  $A^*$  can be written succinctly as  $A = M \circ \mathcal{F} \circ \mathcal{E}$  and  $A^* = \mathcal{R} \circ \mathcal{F}^{-1} \circ M$ .

### 3.2 Cascades

Each cascade in our model applies a refinement step similar to the gradient descent step in equation 7, except that the intermediate quantities are in k-space. Applying  $\mathcal{F} \circ \mathcal{E}$  to both sides of 7 gives the corresponding update equation in k-space:

$$\mathbf{k}^{t+1} = \mathbf{k}^t - \eta^t M(\mathbf{k}^t - \tilde{\mathbf{k}}) + G(\mathbf{k}^t) \quad (10)$$

where  $G$  is the *refinement module* given by:

$$G(\mathbf{k}^t) = \mathcal{F} \circ \mathcal{E} \circ \text{CNN}(\mathcal{R} \circ \mathcal{F}^{-1}(\mathbf{k}^t)). \quad (11)$$

Here, we use the fact that  $\mathbf{x}^t = \mathcal{R} \circ \mathcal{F}^{-1}(\mathbf{k}^t)$ . CNN can be any parametric function that takes a complex image as input and maps it to another complex image. Since the CNN is applied after combining all coils into a single complex image, the same network can be used for MRIs with different number of coils.

Each cascade applies the function represented by equation 10 to refine the k-space. In our experiments, we use a U-Net [11] for the CNN.

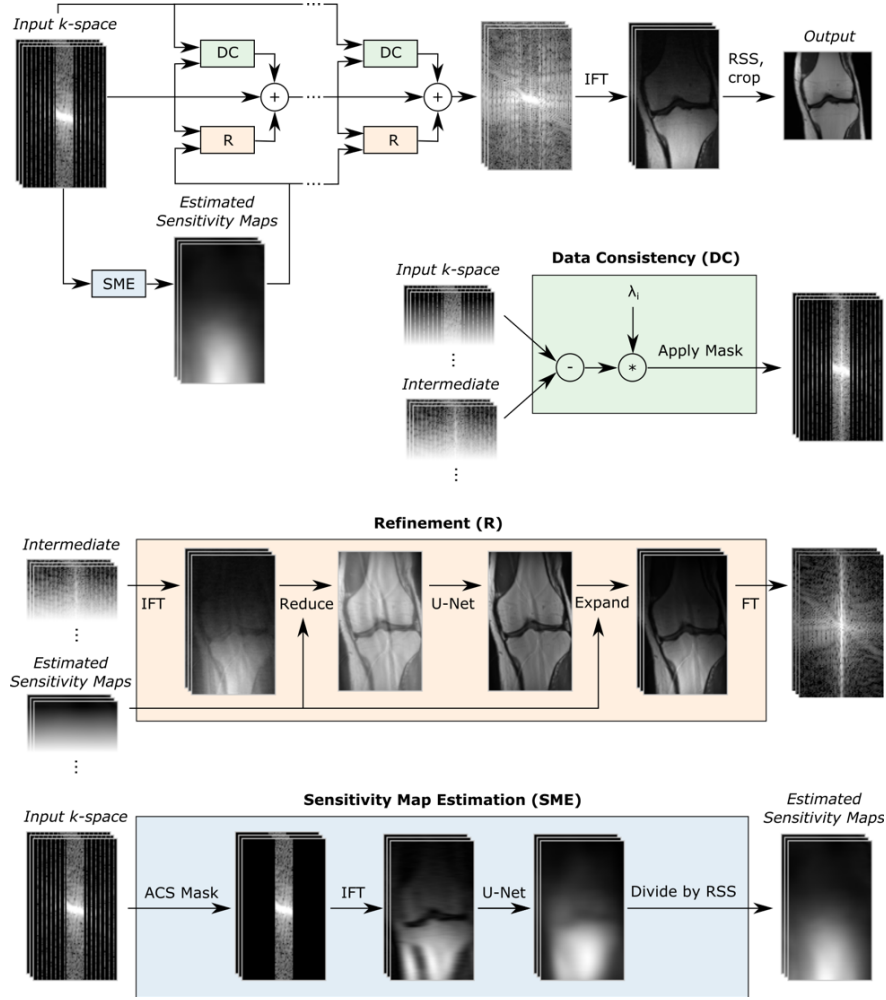


Fig. 1: **Top:** Block diagram of our model which takes under-sampled k-space as input and applies several cascades, followed by an inverse Fourier transform (IFT) and an RSS transform. The **Data Consistency (DC)** module computes a correction map that brings the intermediate k-space closer to the measured k-space values. The **Refinement (R)** module maps multi-coil k-space data into one image, applies a U-Net, and then back to multi-coil k-space data. The **Sensitivity Map Estimation (SME)** module estimates the sensitivity maps used in the Refinement module.

### 3.3 Learned sensitivity maps

The expand and reduce operators in equation 11 take sensitivity maps  $(S_1, \dots, S_N)$  as inputs. In the original VarNet model, these sensitivity maps are computed using the ESPIRiT algorithm [15] and fed in to the model as additional inputs. In our model, however, we estimate the sensitivity maps as part of the reconstruction network using a *Sensitivity Map Estimation (SME)* module:

$$H = \text{dSS} \circ \text{CNN} \circ \mathcal{F}^{-1} \circ M_{\text{center}}. \quad (12)$$

The  $M_{\text{center}}$  operator zeros out all lines except for the autocalibration or ACS lines (described in Section 2.1). This is similar to classical parallel imaging approaches which estimate sensitivity maps from the ACS lines. The CNN follows the same architecture as the CNN in the cascades, except with fewer channels and thus fewer parameters in intermediate layers. This CNN is applied to each coil image independently. Finally, the dSS operator normalizes the estimated sensitivity maps to ensure that the property in equation 3 is satisfied.

### 3.4 E2E-VarNet model architecture

As previously described, our model takes the masked multi-coil k-space  $\mathbf{k}_0$  as input. First, we apply the SME module to  $\mathbf{k}_0$  to compute the sensitivity maps. Next we apply a series of cascades, each of which applies the function in equation 10, to the input k-space to obtain the final k-space representation  $K^T$ . This final k-space representation is converted to image space by applying an inverse Fourier transform followed by a root-sum-squares (RSS) reduction for each pixel:

$$\hat{\mathbf{x}} = \text{RSS}(\mathbf{x}_1, \dots, \mathbf{x}_N) = \sqrt{\sum_{i=1}^N |\mathbf{x}_i|^2} \quad (13)$$

where  $\mathbf{x}_i = \mathcal{F}^{-1}(K_i^T)$  and  $K_i^T$  is the k-space representation for coil  $i$ . The model is illustrated in figure 1.

All of the parameters of the network, including the parameters of the CNN model in SME, the parameters of the CNN in each cascade along with the  $\eta^t$ s, are estimated from the training data by minimizing the structural similarity loss,  $J(\hat{\mathbf{x}}, \mathbf{x}^*) = -\text{SSIM}(\hat{\mathbf{x}}, \mathbf{x}^*)$ , where SSIM is the Structural Similarity index [16] and  $\hat{\mathbf{x}}, \mathbf{x}^*$  are the reconstruction and ground truth images, respectively.

## 4 Experiments

### 4.1 Experimental setup

We designed and validated our method using the multicoil track of the fastMRI dataset [18] which is a large and open dataset of knee and brain MRIs. To validate the various design choices we made, we evaluated the following models on the knee dataset:

1. Variational network [4] (VN)
2. Variational network with the shallow CNNs replaced with U-Nets (VNU)
3. Similar to VNU, but with k-space intermediate quantities (VNU-K)
4. Our proposed end-to-end variational network model (E2E-VN)

The VN model employs shallow convolutional networks with RBF kernels that have about 150K parameters in total. VNU replaces these shallow networks with U-Nets to ensure a fair comparison with our model. VNU-K is similar to our proposed model but uses fixed sensitivity maps computed using classical parallel imaging methods. The difference in reconstruction quality between VNU and VNU-K shows the value of using k-space intermediate quantities for reconstruction, while the difference between VNU-K and E2E-VN shows the importance of learning sensitivity maps as part of the network.

We used the same model architecture and training procedure for the VN model as in the original VarNet [4] paper. For each of the other models, we used  $T = 12$  cascades, containing a total of about 29.5M parameters. The E2E-VN model contained an additional 0.5M parameters in the SME module, taking the total number

of parameters to 30M. We trained these models using the Adam optimizer with a learning rate of 0.0003 for 50 epochs, without using any regularization or data augmentation techniques.

We used two types of under-sampling masks: *equispaced masks*  $M_e(r, l)$ , which sample  $l$  low-frequency lines from the center of k-space and every  $r$ -th line from the remaining k-space; and *random masks*  $M_r(a, f)$ , which sample a fraction  $f$  of the full width of k-space for the ACS lines in addition to a subset of higher frequency lines, selected uniformly at random, to make the overall acceleration equal to  $a$ . These random masks are identical to those used in [18]. We also use equispaced masks as they are easier to implement in MRI machines.

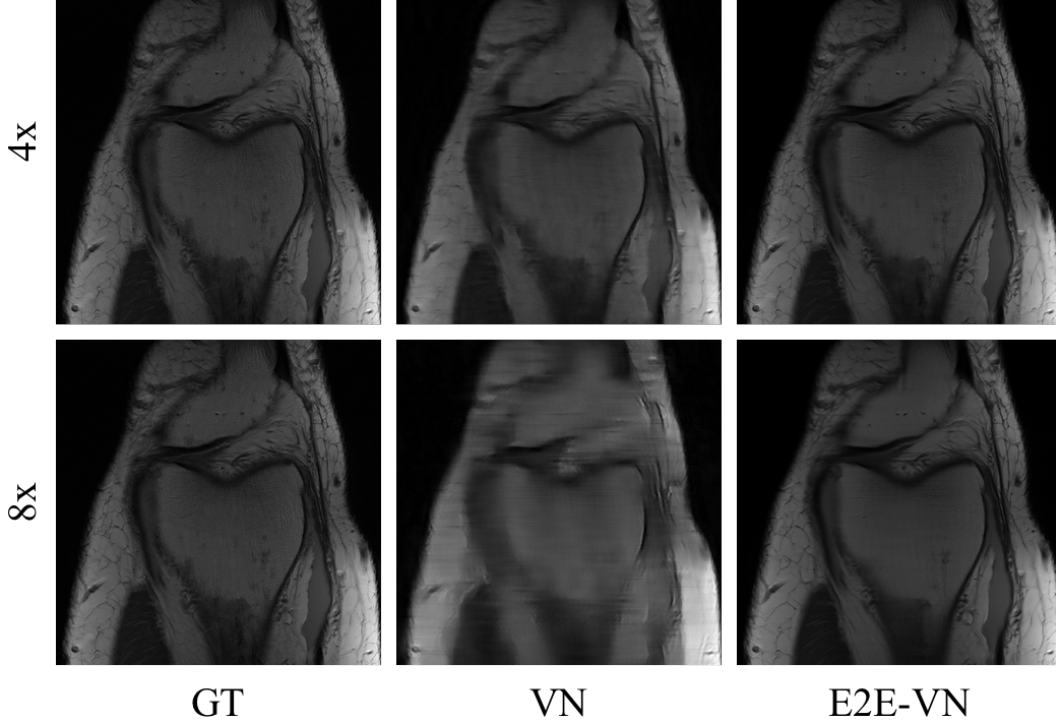


Fig. 2: Examples comparing the ground truth (GT) to VarNet (VN) and E2E-VarNet (E2E-VN). At  $8\times$  acceleration, the VN images contain severe artifacts.

## 4.2 Results

Tables 1 and 2 show the results of our experiments for equispaced and random masks respectively, over a range of down-sampling mask parameters. The VNU model outperforms the baseline VN model by a large margin due to its larger capacity and the multi-scale modeling ability of the U-Nets. VNU-K outperforms VNU demonstrating the value of using k-space intermediate quantities. E2E-VN outperforms VNU-K showing the importance of learning sensitivity maps as part of the network. It is worth noting that the relative performance does not depend on the type of mask or the mask parameters. Some example reconstructions are shown in figure 2.

**Significance of learning sensitivity maps** Figure 3 shows the SSIM values for each model with various equispaced mask parameters. In all cases, learning the sensitivity maps improves the SSIM score. Notably, this improvement in SSIM is larger when the number of low frequency lines is smaller. As previously stated, the quality of the estimated sensitivity maps tends to be poor when there are few ACS lines, which leads to a

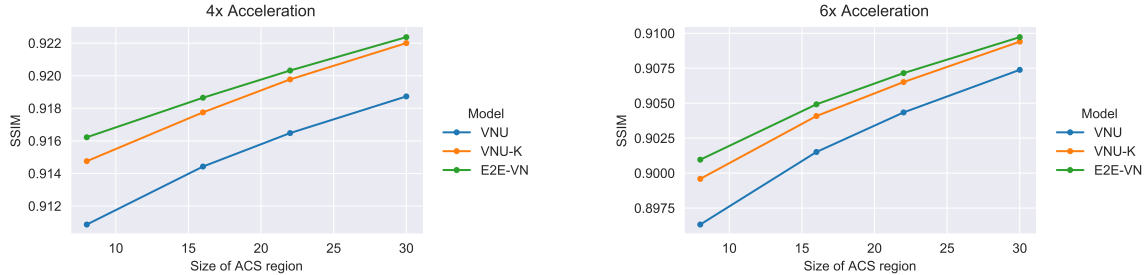
Accel( $r$ )	Num ACS( $l$ )	Model	SSIM
4	30	VN	0.818
		VNU	0.919
		VNU-K	0.922
		E2E-VN	<b>0.923</b>
6	22	VN	0.795
		VNU	0.904
		VNU-K	0.907
		E2E-VN	<b>0.907</b>
8	16	VN	0.782
		VNU	0.889
		VNU-K	0.891
		E2E-VN	<b>0.893</b>

Table 1: Experimental results using equispaced masks on the knee MRIs

Accel( $a$ )	Frac ACS( $f$ )	Model	SSIM
4	0.08	VN	0.813
		VNU	0.906
		VNU-K	0.907
		E2E-VN	<b>0.910</b>
6	0.06	VN	0.767
		VNU	0.886
		VNU-K	0.887
		E2E-VN	<b>0.892</b>
8	0.04	VN	0.762
		VNU	0.870
		VNU-K	0.871
		E2E-VN	<b>0.878</b>

Table 2: Experimental results using random masks on the knee MRIs

degradation in the final reconstruction quality. The E2E-VN model is able to overcome this limitation and generate good reconstructions even with a small number of ACS lines.

Fig. 3: Effect of different equispaced mask parameters on reconstruction quality at  $4\times$  and  $6\times$  acceleration.

**Experiments on test data** Table 3 shows our results on the test datasets for both the brain and knee MRIs compared with the best models on the fastMRI leaderboard<sup>4</sup>. To obtain these results, we used the same training procedure as our previous experiments, except that we trained on both the training and validation sets for 100 epochs. We used the same type of masks that are used for the fastMRI paper [4]. Our model outperforms all other models published on the fastMRI leaderboard for both anatomies.

## 5 Conclusion

In this paper, we introduced End-to-End Variational Networks for multi-coil MRI reconstruction. While MRI reconstruction can be posed as an inverse problem, multi-coil MRI reconstruction is particularly challenging because the forward process (which is determined by the sensitivity maps) is not completely known. We alleviate this problem by estimating the sensitivity maps within the network, and learning fully end-to-end. Further, we explored the architecture space to identify the best neural network layers and intermediate representation for this problem, which allowed our model to obtain new state-of-the-art results on both brain and knee MRIs.

<sup>4</sup> <http://fastmri.org/leaderboards>

Dataset	Model	4× Acceleration			8× Acceleration		
		SSIM	NMSE	PSNR	SSIM	NMSE	PSNR
Knee	E2E-VN	<b>0.930</b>	0.005	40	<b>0.890</b>	0.009	37
	SubtleMR	0.929	0.005	40	0.888	0.009	37
	AIRS Medical	0.929	0.005	40	0.888	0.009	37
	SigmaNet	0.928	0.005	40	0.888	0.009	37
	i-RIM	0.928	0.005	40	0.888	0.009	37
Brain	E2E-VN	<b>0.959</b>	0.004	41	<b>0.943</b>	0.008	38
	U-Net	0.945	0.011	38	0.915	0.023	35

Table 3: Results on the test data compared with the best models on the fastMRI leaderboard

The quantitative measures we have used only provide a rough estimate for the quality of the reconstructions. Many clinically important details tend to be subtle and limited to small regions of the MR image. Rigorous clinical validation needs to be performed before such methods can be used in clinical practice to ensure that there is no degradation in the quality of diagnosis.



## Bibliography

- [1] Emmanuel J Candès et al. Compressive sampling. In *Proceedings of the international congress of mathematicians*, volume 3, pages 1433–1452. Madrid, Spain, 2006.
- [2] David Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, 2006.
- [3] Mark A. Griswold, Peter M. Jakob, Robin M. Heidemann, Mathias Nittka, Vladimir Jellus, Jianmin Wang, Berthold Kiefer, and Axel Haase. Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magnetic Resonance in Medicine*, 47(6):1202–1210, 2002.
- [4] Kerstin Hammernik, Teresa Klatzer, Erich Kobler, Michael P. Recht, Daniel K. Sodickson, Thomas Pock, and Florian Knoll. Learning a variational network for reconstruction of accelerated MRI data. *Magnetic Resonance in Medicine*, 79(6):3055–3071, 2018.
- [5] Florian Knoll, Kerstin Hammernik, Chi Zhang, S. Möller, Thomas Pock, Daniel K. Sodickson, and Mehmet Akçakaya. Deep learning methods for parallel magnetic resonance image reconstruction. *CoRR*, abs/1904.01112, 2019.
- [6] Dong Liang, Jing Cheng, Ziwen Ke, and Leslie Ying. Deep mri reconstruction: Unrolled optimization algorithms meet neural networks. *arXiv preprint arXiv:1907.11711*, 2019.
- [7] Michael Lustig, David Donoho, and John M Pauly. Sparse MRI: The Application of Compressed Sensing for Rapid MR Imaging. *Magnetic Resonance in Medicine*, 58(6), 2007.
- [8] Klaas P Pruessmann, Markus Weiger, Markus B Scheidegger, and Peter Boesiger. SENSE: sensitivity encoding for fast MRI. *Magnetic resonance in medicine*, 42(5), 1999.
- [9] Patrick Putzky and Max Welling. Invert to learn to invert. In *Advances in Neural Information Processing Systems*, pages 444–454, 2019.
- [10] Patrick Putzky, Dimitrios Karkaloulos, Jonas Teuwen, Nikita Miriakov, Bart Bakker, Matthan Caan, and Max Welling. i-rim applied to the fastmri challenge. *arXiv preprint arXiv:1910.08952*, 2019.
- [11] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional networks for biomedical image segmentation. In *MICCAI 2015: Medical Image Computing and Computer-Assisted Intervention*, volume 9351 of *Lecture Notes in Computer Science*, pages 234–241. Springer, 2015.
- [12] Jo Schlemper, Jose Caballero, Joseph V. Hajnal, Anthony N. Price, and Daniel Rueckert. A deep cascade of convolutional neural networks for dynamic MR image reconstruction. *IEEE Transactions on medical imaging*, 37(2), 2018.
- [13] Daniel K Sodickson and Warren J Manning. Simultaneous acquisition of spatial harmonics (SMASH): fast imaging with radiofrequency coil arrays. *Magnetic resonance in medicine*, 38(4), 1997.
- [14] Anuroop Sriram, Jure Zbontar, Tullie Murrell, C Lawrence Zitnick, Aaron Defazio, and Daniel K Sodickson. Grappanet: Combining parallel imaging with deep learning for multi-coil mri reconstruction. *arXiv preprint arXiv:1910.12325*, 2019.
- [15] Martin Uecker, Peng Lai, Mark J Murphy, Patrick Virtue, Michael Elad, John M Pauly, Shreyas S Vasanawala, and Michael Lustig. ESPIRiT -an eigenvalue approach to autocalibrating parallel MRI: where SENSE meets GRAPPA. *Magnetic resonance in medicine*, 71(3), 2014.
- [16] Zhou Wang, Eero P Simoncelli, and Alan C Bovik. Multiscale structural similarity for image quality assessment. In *Asilomar Conference on Signals, Systems & Computers*, 2003.
- [17] Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, and Eero P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.
- [18] Jure Zbontar, Florian Knoll, Anuroop Sriram, Tullie Murrell, Zhengnan Huang, Matthew J. Muckley, Aaron Defazio, Ruben Stern, Patricia Johnson, Mary Bruno, Marc Parente, Krzysztof J. Geras, Joe Katsnelson, Hersh Chandarana, Zizhao Zhang, Michal Drozdal, Adriana Romero, Michael Rabbat, Pascal Vincent, Nafissa Yakubova, James Pinkerton, Duo Wang, Erich Owens, C. Lawrence Zitnick, Michael P. Recht, Daniel K. Sodickson, and Yvonne W. Lui. fastmri: An open dataset and benchmarks for accelerated MRI. *CoRR*, abs/1811.08839, 2018. URL <http://arxiv.org/abs/1811.08839>.

## 6 Supplementary Materials

### 6.1 Dithering as post-processing

The Structural Similarity (SSIM) loss [17] we used to train our models has a tendency to produce overly smooth reconstructions even when all of the diagnostic content is preserved. We noticed a similar behavior with other frequently used loss functions like mean squared error, mean absolute error, etc. Sriram et al. [14] found that dithering the image by adding a small amount of random gaussian noise helped enhance the perceived sharpness of their reconstructions. We found that the same kind of dithering helped improve the sharpness of our reconstructions, but we tuned the scale of the noise by manual inspection.

Similar to [14], we adjusted the scale of the added noise to the brightness of the image around each pixel to avoid obscuring dark areas of the reconstruction. Specifically, we first normalize the image by dividing each pixel by the maximum pixel intensity. Then we add zero-mean random gaussian noise to each pixel. The standard deviation of the noise at a given pixel location is equal to  $\sigma$  times the square root of the local median computed over a patch of  $11 \times 11$  pixels around that pixel location. We set  $\sigma = 0.02$  for the brain images and non fat suppressed knee images, and  $\sigma = 0.03$  for the fat suppressed knee images.

Example reconstructions with and without noise are shown in figs. 4 to 7. The dithered images look more natural, especially at higher accelerations.

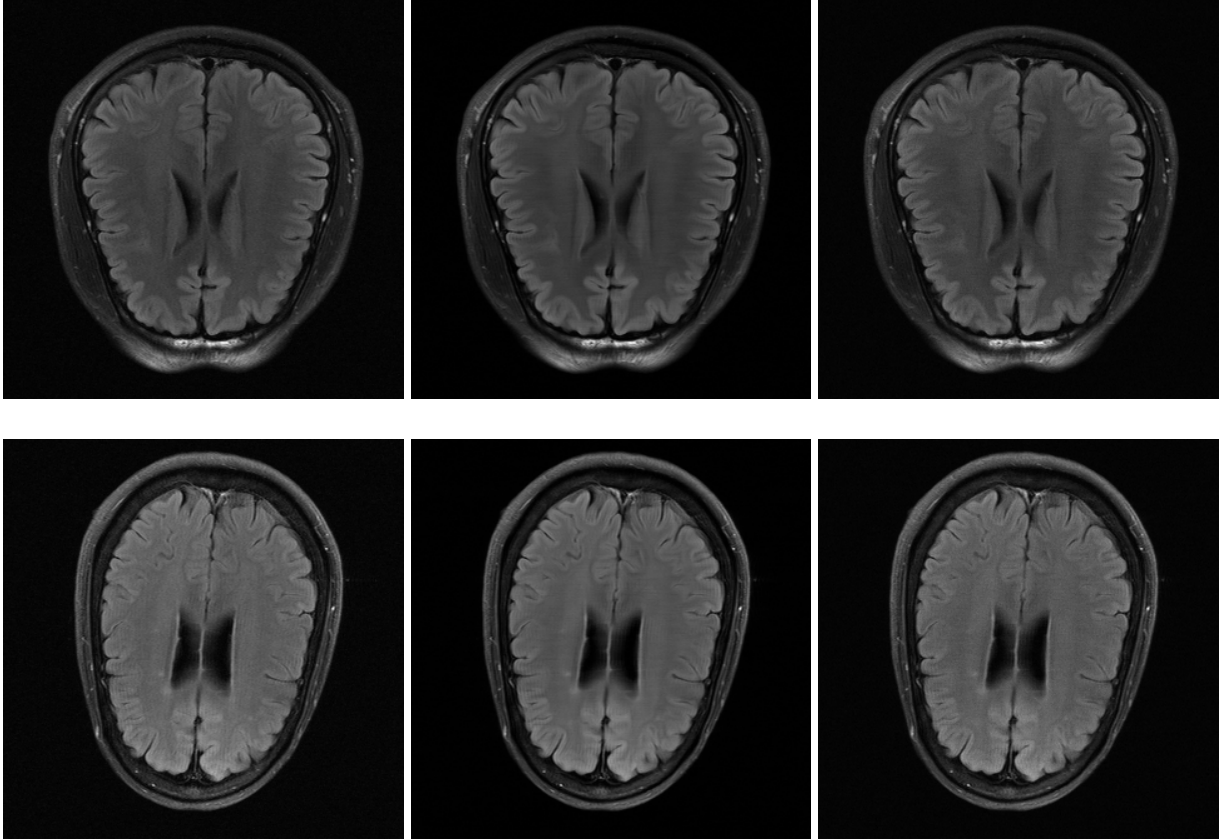


Fig. 4: Brain images showing ground truth (left), reconstruction (middle) and reconstruction with added noise (right) at  $4\times$  acceleration

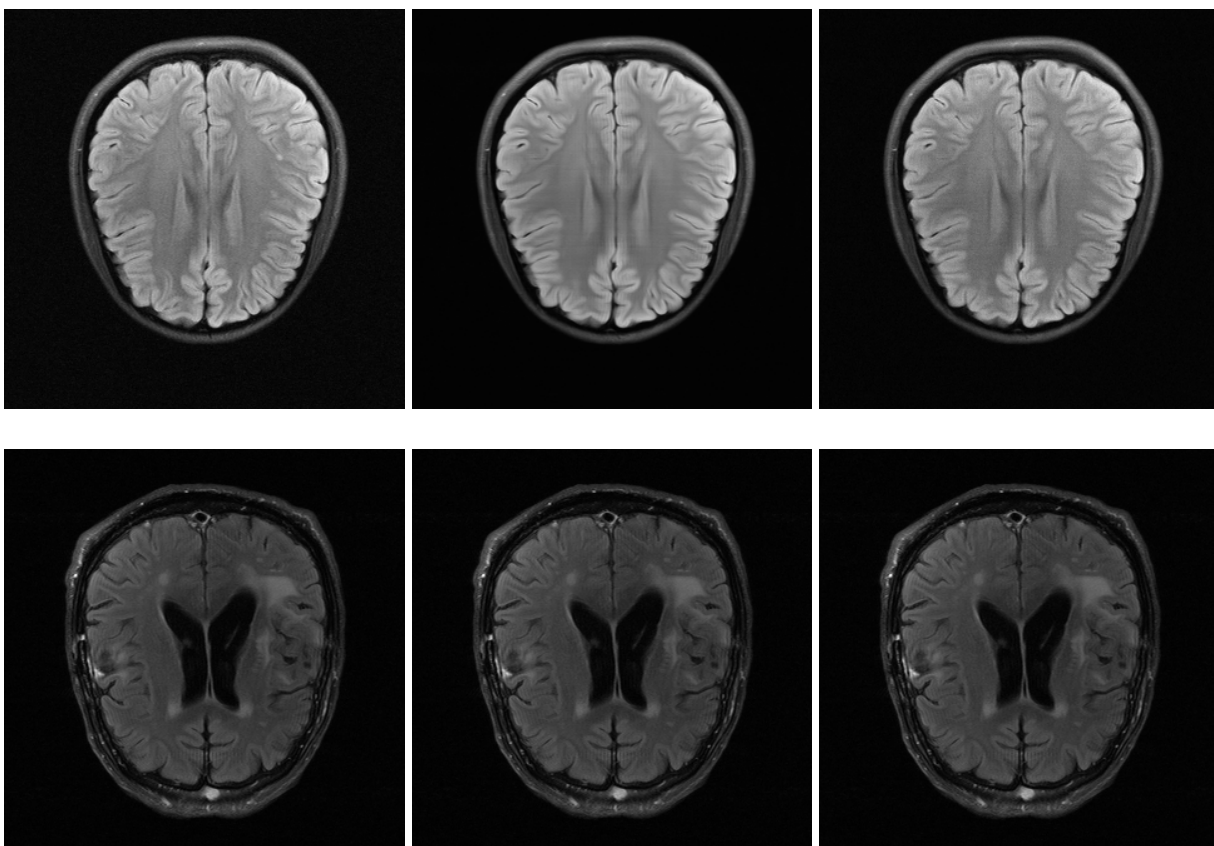


Fig. 5: Brain images showing ground truth (left), reconstruction (middle) and reconstruction with added noise (right) at  $8\times$  acceleration

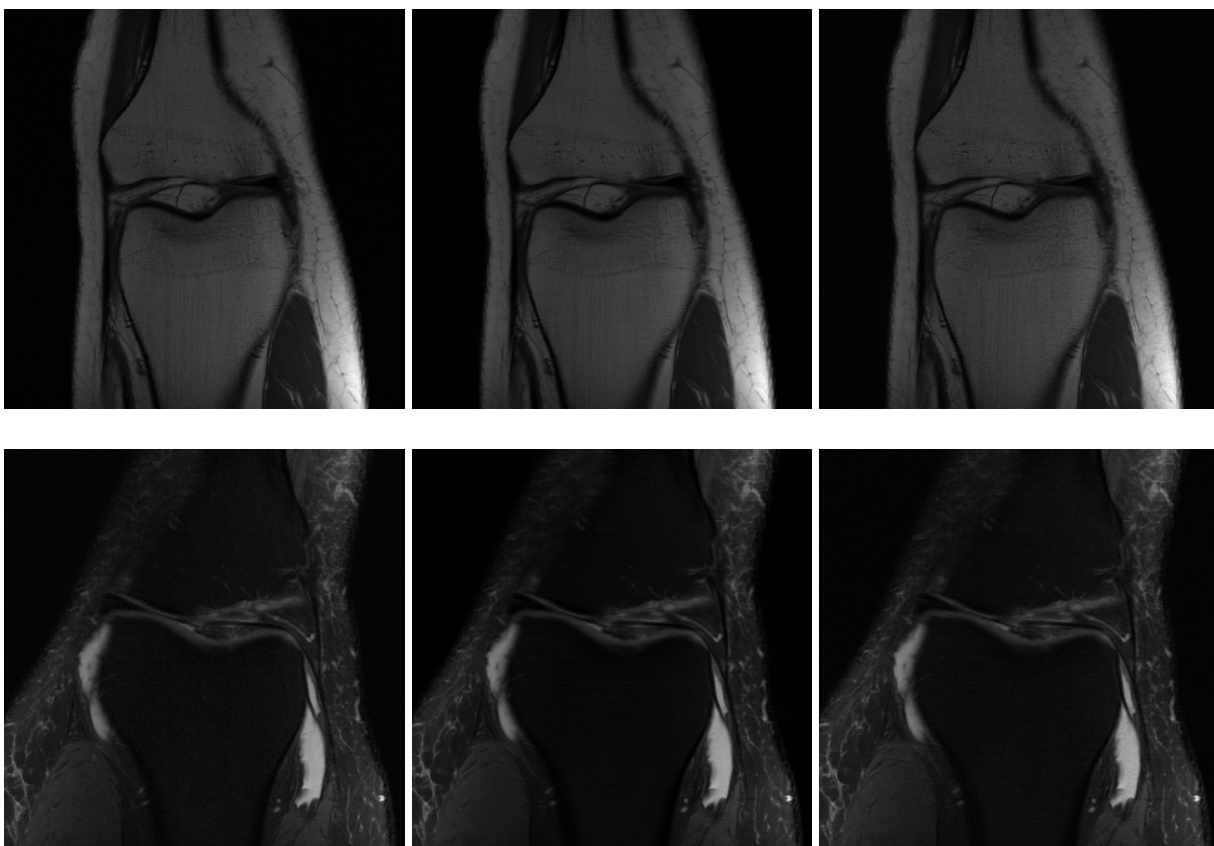


Fig. 6: Knee images showing ground truth (left), reconstruction (middle) and reconstruction with added noise (right) at  $4\times$  acceleration

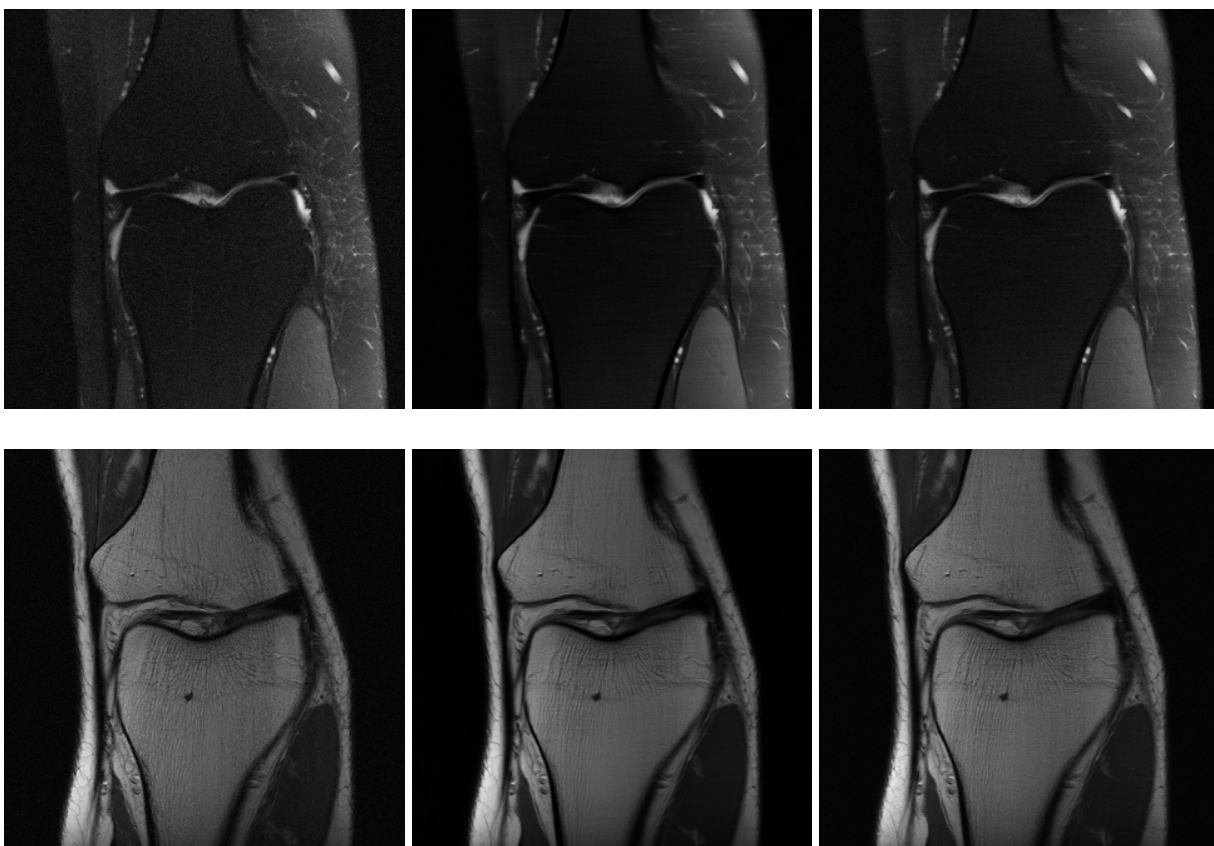


Fig. 7: Knee images showing ground truth (left), reconstruction (middle) and reconstruction with added noise (right) at  $8\times$  acceleration