

# Stage U-Net Framework: Streamlining MRI Reconstruction from Under-sampled K-Space

Aya Mohamed ElBehairy<sup>1</sup>, Inas A. Yassine<sup>2</sup>, and Mustafa Elattar<sup>1</sup>

<sup>1</sup>Medical Imaging and Image Processing Research Group, Centre of Informatics Science, Nile University, Egypt

ayhassan@nu.edu.eg; melattar@nu.edu.eg

<sup>2</sup> Systems and Biomedical Engineering Department, Faculty of Engineering, Cairo University, Egypt  
[Iyassine@eng.cu.edu.eg](mailto:Iyassine@eng.cu.edu.eg)

**Abstract** - Magnetic resonance Image MRI with various protocols is extensively used for diagnosis because it gives detailed images. Still, its acquisition time is excessively long, which causes motion artifacts due to patient impatience. To accelerate, two key strategies were utilized to decrease acquisition time parallel imaging and compressed sensing. Parallel imaging reduces time by simultaneously capturing subsampled MRI data acquired from multiple receiver coils. Compressed sensing acquires partially observed k-space data using regularized iterative optimization techniques. Both methods provide complementary possibilities for speeding up MRI acquisition. Recently Our proposed architecture uses an exclusive two-phase technique, Image2Image Understanding and Reconstruction, to rebuild MR images from under-sampled k-space data. In the first phase, a U-Net is trained on images reconstructed from full sampled k-space, laying the groundwork for future reconstruction. In the second phase to subsample images we apply the radial mask to k-space: a Fourier transform. The reconstructed images from subsampled k-space are used to train the trained U-Net from the first phase to improve image details as if it is reconstructed from a fully sampled k-space. This dual-phase technique improves performance by modifying the U-Net to learn image structure from fully sampled k-space first, establishing a solid foundation for future high-quality image reconstruction from subsampled k-space. The stability created in the first step saves time and minimizes processing power needs in the second phase.

**Keywords:** MRI reconstruction, k-space, deep learning, U-Net, fastMRI.

## 1. Introduction

Magnetic Resonance Imaging (MRI) is the leading diagnostic modality for a wide range of illnesses, including musculoskeletal, neurological, and oncological conditions. MRI does not employ harmful ionizing radiation. However, the inherent sluggishness of the MRI data collecting process, caused by its underlying physics, distinguishes it from speedier alternatives such as CT or X-ray [1]. Consequently, the persistent focus of research over the decades has been to expedite the MRI acquisition speed. Reducing MRI scan times has benefits like improving patient satisfaction, cutting motion artifacts, and lowering medical costs. Scan duration is linked to the number of phase-encoding steps in k-space, and efforts to expedite MRI involve omitting specific lines while addressing aliasing issues from Nyquist criterion violations. Compressed Sensing and Parallel Imaging have achieved significant progress in fasten MRI acquisition [2], [3]. To reduce or decrease aliasing artifacts, compressed sensing MRI can be viewed as sub-Nyquist sampling approach that compensates for under-sampled data by gathering fewer samples than typical signal processing methods. Previous knowledge of MR images is acquired from unmeasured k-space data. Parallel MRI, on the other hand, uses numerous receiver coils with space-dependent properties to reduce aliasing effects [4]. A potential for fast MRI strategy involves under-sampling k-space, leading to an acceleration rate that is proportionate to the degree of under-sampling. This paper focuses solely on single-channel MRI so Parallel MRI will not be discussed. Deep learning, a method for representing complicated abstractions in data via numerous layers of processing, has lately gained prominence, owing to the availability of powerful GPUs. Convolutional neural networks (CNNs) have shown useful in solving complex vision problems such as action detection, super-resolution, and denoising. CNNs have various benefits, including independence from past information, no requirement for hand-engineered feature design, and a strong capacity to capture visual structures. These characteristics influenced our decision to use CNNs for reconstructing MR images from under-sampled k-space data [5].

## 2. Background and Related Work

### 2.1. MRI Acquisition

MR scanners image a patient’s anatomy by acquiring measurements in the frequency domain using a measuring instrument called a receiver coil. In clinical settings that frequency domain is known as k-space where  $K$  refers to the spatial wave number and images can be obtained by applying two-dimensional Fourier transform  $F$ .

$$Images = F^{-1}(K) \quad (1)$$

As our aim is to accelerate MRI acquisition, the k-space is subsampled with acceleration factor represented as  $R$  is used to define the degree of acceleration, also known as  $R$ -fold acceleration.  $R$  is calculated by dividing the total number of k-space measurements necessary for a full scan by the number of measurements obtained. The sub-sampled k-space measurements can be defined in terms of the fully-sampled k-space measurements [6] where  $M$  denotes sampling mask.

$$k = M \odot K \quad (2)$$

### 2.2. Deep Learning Accelerating MRI Reconstruction

The latest developments in Deep Learning (DL), particularly in Convolutional Neural Networks (CNNs), have facilitated their utilization in various domains, including addressing inverse problems in imaging, notably tasks related to Accelerated MRI Reconstruction. When integrated with PI-CS, DL-based techniques demonstrate superior performance compared to conventional PI and CS methods. This results in reconstructed MR images that exhibit fewer artifacts[1], [7] . All these deep learning-based algorithms are designed to take an image reconstructed from under-sampled k-space as input and output an image reconstructed from fully sampled k-space; however, detailed structures in the input images to CNNs are likely to be distorted or even disappear. While there have been recent attempts to restore properly sampled images from under-sampled k-space [8]. Our Method get a high resolution even with U-Net which is a very basic architecture.

## 3. Method

In our architecture, to allow the network to fully utilize all of the information needed from the images to get high-resolution image from subsampled k-space, we divided our architecture into two stages. The first phase, Image2Image Understanding, consists of training a U-Net that seeks to recreate high-resolution images from images reconstructed from full sampled k-space . Initially, Images are translated into k-space using the Fourier transform and multiplied by the subsampling mask to get subsampled k-space. The subsampled k-space is then translated back into the image domain via the inverse Fourier transform. U-Net is trained using reconstructed images from subsampled k-spaces. This dual-phase technique improves image reconstruction performance by modifying U-Net to be trained on image structure first from images collected from fully sampled k-space, laying a solid foundation for the next phase, which involves reconstructing an image from subsampled k-space with high detail as if it were reconstructed from fully sampled k-space.

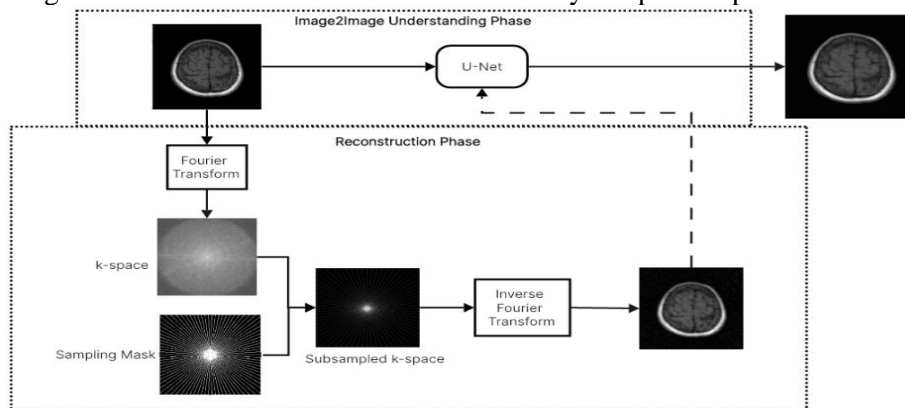


Fig. 1: Our two-phase architecture begins with Image2Image Understanding then the obtained weights are then employed in the Reconstruction phase.

### 3.1. Image2Image Understanding Phase

The initial model acts as a fundamental model for obtaining launching weights by utilizing previous information encoded in images that reflect high-quality, detailed. Using U-Net to encode the image features to be used in the reconstruction phase. This collection of images serves as a standard of excellence before the subsampling process, introducing stability in the second phase. In other words, it acts as a pre-trained model that needs to be fine-tuned later. This stabilization not only saves time but also decreases the processing power necessary for the next phase.

### 3.2. Reconstruction Phase

The Reconstruction Phase begins with translating images  $I$  to  $k$ -space  $K$  using a two-dimensional Fourier transform and adding measurement noise to them. To mimic the  $k$ -space transformation, which is a complex data type with both input and output being actual data, each image in the spatial domain with one channel is transformed to two channels in  $k$ -space that represent the  $k$ -space phase and magnitude respectively.

$$K = F(I) + e \quad (3)$$

Then applying sampling mask to each channel in  $k$ -space using (2) to get subsampled  $k$  and using  $k$  in (1) get noisy images served as input for U-Net. U-Net initially began with weights created by Image2Image Understanding network. With this prior information, U-Net can rebuild high-quality images with fine details from noisy images without the requirement for  $k$ -space training and simply on the spatial domain.

### 3.3. U-Net

Main block in our architecture is U-Net which is useful for image-to-image mapping tasks. As it encode the most important features in the image and then decoded it back. The use of pooling and up-sampling layers enables the learning of usable feature maps across many sizes and abstraction levels. The U-Net uses a multi-resolution feature representation to forecast higher-level information at the decoder's lowest level, gradually adding finer, higher-frequency details through up-sampling layers.

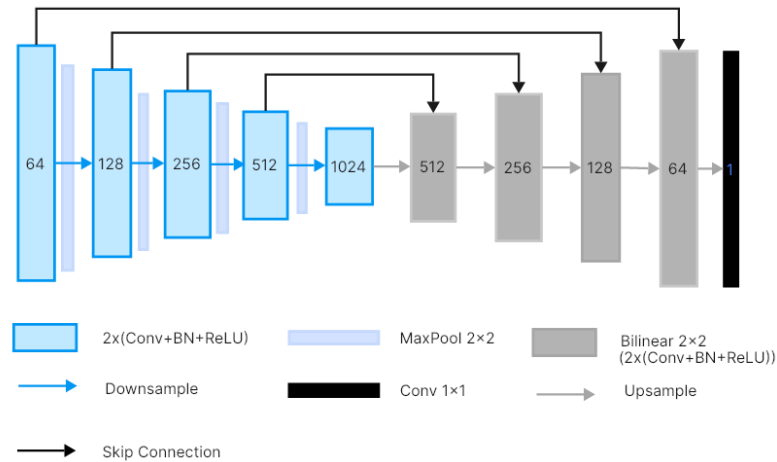


Fig. 2: Our U-Net architecture

### 3.4. Loss Function

For image reconstruction problems, the most used loss function is MSE, which is L2 loss, and MAE, which is L1. Both have advantages and disadvantages. For L2, the main disadvantage is that it produces blurring images, and in MR imaging, details are very important. For that point, L1 is much better because it preserves edges, but L1 has a drawback, which may produce images with less contrast, and L2 does not have that drawback, so we need to use both with different ratios so our loss  $L$ . Furthermore, in the context of Compressed Sensing (CS) and speeding up the model conversion process, using a linear mix of L2 and L1 loss functions is critical. This smart combination optimizes the reconstruction process by reducing the blurring effects of L2 while keeping important features using L1's edge-preserving properties. A well-balanced and efficient model conversion for rapid image reconstruction relies heavily on the linear combination of these loss functions,

which are tuned with optimal ratios. We used grid search to get optimum  $\alpha$  and  $1-\alpha$  for L1 and L2 respectively. Given a ground truth image  $g$  and a test image  $f$ , both of size  $M \times N$ , our customized loss between  $g$  and  $f$  is defined by:

$$Loss(g, f) = \alpha MSE(g, f) + (1 - \alpha) MAE(g, f) \quad (4)$$

## 4. Experiments

### 4.1. Dataset

The dataset utilized in that work is from the fastMRI dataset [9], which is the result of a collaboration between the Center for Advanced Imaging Innovation and Research (CAI2R) and Facebook AI Research (FAIR) on fastMRI - a joint research initiative to investigate the use of AI to speed up MRI scans by up to 10 times [4]. The dataset contains two forms of raw data, unprocessed complex-valued multi-coil MR readings and DICOM images. The data was intended to assist two types of challenges. There are two reconstruction tasks: single coils and multi-coils. We worked on the single-coil reconstruction assignment. The fastMRI dataset provides two types of DICOM data: Brain DICOM and Knee DICOM. Our results rely on Brain DICOM data that are axial 2D image volumes with different pulse sequence T1, T2, and T2 FLAIR. We trained our architecture on two experiments using T1 and T2 separately. Our dataset is divided into training, testing and validation sets with 70%, 20% and 10% respectively.

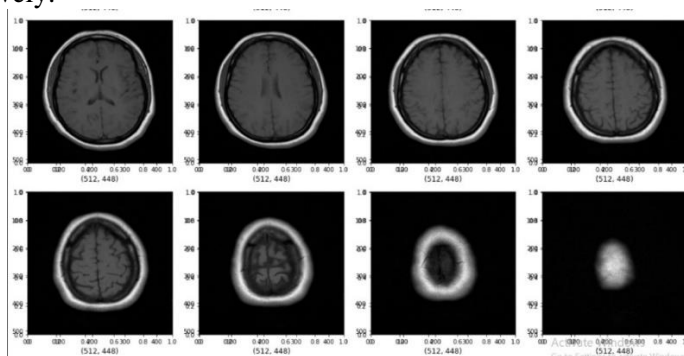


Fig. 3: Samples from T1 DICOM data

### 4.2. Mask

To sample k-space, we implemented a radial subsampling mask with varying acceleration rates, specifically  $R = 10$  and  $R = 5$ . As the acceleration rate increased, the reconstruction process became more challenging. The intuition for using radial mask as k-space can be filled with radial sampling lines.

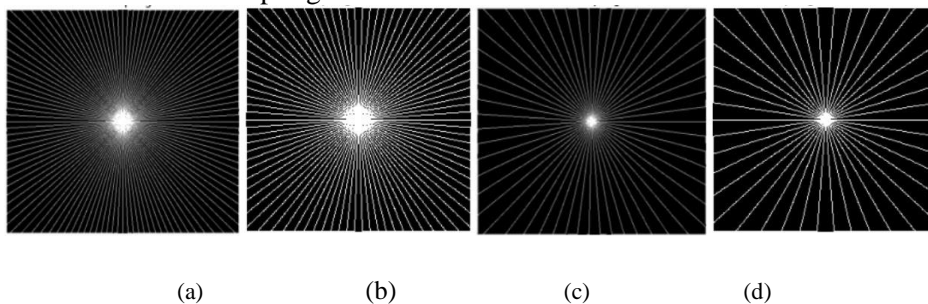


Fig. 4(a). Radial Mask with  $R=4$ . Fig. 4(b). Radial Mask with  $R=5$ . Fig. 4(c). Radial Mask with  $R=8$ . Fig. 4(d). Radial Mask with  $R=10$ .

### 4.3. Evaluation Metrics

A traditional approach for evaluating image reconstruction employs quantitative measurements, namely the Peak Signal-to-Noise Ratio (PSNR) and the Structural Similarity (SSIM) index. PSNR is defined as the ratio of the power of the highest potential image intensity inside a volume to the power of distortion noise and other faults. This metric gives a quantifiable

assessment of the reconstructed magnitude image's faithfulness to the ground truth. Given a ground truth image  $g$  and a test image  $f$ , both of size  $M \times N$ , the PSNR between  $g$  and  $f$  is defined by:

$$PSNR(g, f) = 10 \log_{10} \frac{255^2}{MSE(g, f)} \quad (5)$$

The SSIM index, on the other hand, uses interdependencies among neighbouring pixels to determine the similarity between two images. SSIM examines the structural attributes of objects in images and computes at various image regions using a sliding window technique, allowing for a nuanced analysis of the structural congruence between the reconstructed image and ground truth image[5]. SSIM is designed by modeling any image distortion as a combination of three factors that are loss of correlation, luminance distortion and contrast distortion. The SSIM is defined as:

$$SSIM(g, f) = l(g, f)c(g, f)s(g, f) \quad (6)$$

The first is term the luminance comparison function which measures the closeness of the two images. The second term is the contrast comparison function which measures the closeness of the contrast of the two images. The third term is the structure comparison function which measures the correlation coefficient between the two images.

#### 4.4. Training Details

Our U-Net model was implemented with PyTorch and five scales: 64, 128, 256, 512, and 1024 filters for convolutional layers with kernel sizes of  $3 \times 3$ , followed by a batch normalization layer. The input and output images are  $256 \times 256$  in size in some experiments and in other  $320 \times 320$ , and the ReLU activation function is employed. Same U-Net used in both phases. Model was optimized using AdamW with weight decay 0.00001 and fixed learning rate 0.001. Each phase trained for 200 epoch with 16 batch size without adding any data augmentation. The approximate trainable parameters of our model is 31.4M parameter.

### 5. Experimental Results

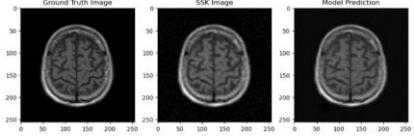
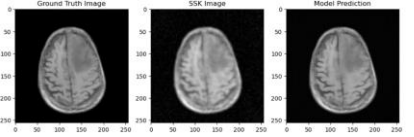
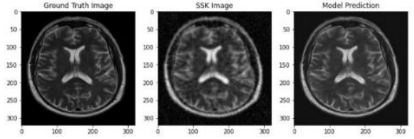
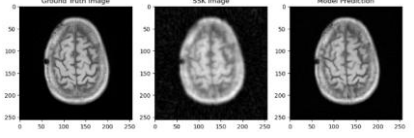
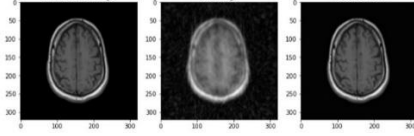
To provide a different point of view of our results, we studied several acceleration rates and time sequences to ensure our technique's competitiveness. In our search for similar models, We uncovered the basic model of the fastMRI dataset, which is a U-Net architecture on the T1 brain dataset. As a consequence, we looked at other models with similar qualities, such as architecture, subsampling mask, and time sequence. Table (1) presents a thorough comparison of our data with these selected models., allowing us to assess their efficacy and distinctiveness.

As DuDoRNet [8] employed various masks with varying acceleration rates, one of which is a radial mask with acceleration factor 5 so it used in 5 acceleration factor only. Even though the datasets acquired an in-house, but it still have some thing in common that we could compare with. U-Net stated in [9] employs a mask linked to the [4] dataset with acceleration factors 5 and 10 The U-Net[7] has essentially the same convolution filter, and 64, 128, 256, and 512 will be included in the comparison.

Table 1: Other Approaches Details

Model	Dataset	Mask Type	R	Sequence
fastMRI BaseModel[4]	fastMRI	Cartesian	4,8	T1
U-Net [7]	Calgary-Campinas	Provided in dataset	5,10	T1
DuDoRNet [8]	In-house	Radial	5	T2

Table 2: Experimental results

R	Sequence	Comparative Results			
		Model	Metric		Our Approach Output Sample
			SSIM	PSNR	
4	T1	fastMRI BaseModel [4]	0.9275	36.24	
		Ours	1.0	69.56	
5	T1	U-Net [6]	0.8682	29.25	
		Ours	0.999	71.67	
	T2	DuDoRNet [8]	0.981	40.815	
		Ours	1.0	63.55	
8	T1	fastMRI BaseModel [4]	0.8837	32.02	
		Ours	1.0	66.534	
10	T1	U-Net [6]	0.8177	27.23	
		Ours	0.9999	60.8	

## 6. Conclusion

Our findings indicate that our two-phase approach works well for MR image reconstruction, even with varying time sequences and acceleration rates, producing competitive results even though all training is done in the image domain. The Image2Image Understanding step guarantees that the U-Net catches critical information from fully sampled k-space, establishing the groundwork for future reconstruction. The Reconstruction phase successfully converts images into k-space, eliminating the requirement for k-space training and allowing the U-Net to reconstruct high-quality images directly in the spatial domain. The strategic combination of L2 and L1 loss functions, tweaked via grid search for optimal ratios, improves the model's performance by decreasing blurring effects while keeping key characteristics. The positive results demonstrate the potential of our approach in enhancing rapid MR imaging techniques and provide useful insights for future study in this area.

## 7. Future Work

Our future attempts will entail investigating different masks, such as the spiral mask, with varying acceleration speeds. To offer a more impartial comparison of models, we intend to train their architectures on the public CalgaryCampinas dataset while using the same mask across all instances.

## References

- [1] A. Sriram, J. Zbontar, T. Murrell, C. L. Zitnick, A. Defazio, and D. K. Sodickson, ‘GrappaNet: Combining Parallel Imaging with Deep Learning for Multi-Coil MRI Reconstruction’.
- [2] C. M. Hyun, H. P. Kim, S. M. Lee, S. Lee, and J. K. Seo, ‘Deep learning for undersampled MRI reconstruction’, *Phys Med Biol*, vol. 63, no. 13, Jun. 2018, doi: 10.1088/1361-6560/aac71a.
- [3] G. Yiasemis, J.-J. Sonke, C. Sánchez, and J. Teuwen, ‘Recurrent Variational Network: A Deep Learning Inverse Problem Solver applied to the task of Accelerated MRI Reconstruction’. [Online]. Available: <https://github.com/NKI-AI/direct>.
- [4] Zbontar, J., Knoll, F., Sriram, A., Murrell, T., Huang, Z., Muckley, M. J., Defazio, A., Stern, R., Johnson, P., Bruno, M., Parente, M., Geras, K. J., Katsnelson, J., Chandarana, H., Zhang, Z., Drozdal, M., Romero, A., Rabbat, M., Vincent, P., ... Lui, Y. W. (2018). *fastMRI: An Open Dataset and Benchmarks for Accelerated MRI*. <http://arxiv.org/abs/1811.08839>
- [5] Wang, S., Su, Z., Ying, L., Peng, X., Zhu, S., Liang, F., Feng, D., & Liang, D. (2016). Accelerating magnetic resonance imaging via deep learning. *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, 514–517. <https://doi.org/10.1109/ISBI.2016.7493320>
- [6] G. Yiasemis, J.-J. Sonke, C. Sánchez, and J. Teuwen, ‘Recurrent Variational Network: A Deep Learning Inverse Problem Solver applied to the task of Accelerated MRI Reconstruction’. [Online]. Available: <https://github.com/NKI-AI/direct>.
- [7] Souza, R., Lucena, O., Garrafa, J., Gobbi, D., Saluzzi, M., Appenzeller, S., Rittner, L., Frayne, R., & Lotufo, R. (2018). An open, multi-vendor, multi-field-strength brain MR dataset and analysis of publicly available skull stripping methods agreement. *NeuroImage*, 170, 482–494. <https://doi.org/https://doi.org/10.1016/j.neuroimage.2017.08.021>
- [8] B. Zhou and S. K. Zhou, ‘DuDoRNet: Learning a Dual-Domain Recurrent Network for Fast MRI Reconstruction with Deep T1 Prior’.
- [9] Sherry, F., Benning, M., Reyes, J. C. D. los, Graves, M. J., Maierhofer, G., Williams, G., Schönlieb, C.-B., & Ehrhardt, M. J. (2019). Learning the Sampling Pattern for MRI. <http://arxiv.org/abs/1906.08754>