

# Deep Learning-based Human MRI Reconstruction and Preprocessing with Artificial Fourier Transform Network (AFT-Net)

Yanting Yang<sup>1</sup>, Jeffery Siyuan Tian<sup>2</sup>, and Jia Guo<sup>1</sup>

<sup>1</sup>Columbia University, New York, NY, United States, <sup>2</sup>Computer Science, University of Maryland, College Park, Clarksville, MD, United States

## Synopsis

**Motivation:** Complex-valued deep learning framework has not been fully investigated in human normal-field and low-field MRI reconstruction and preprocessing.

**Goal(s):** We aim to replace conventional numerical methods with deep learning network, which reconstruct and preprocess the k-space data in parallel.

**Approach:** An artificial Fourier transform network (AFT-Net) is proposed to directly processes the complex-valued raw data in the sensor domain.

**Results:** An evaluation of accelerated reconstruction and denoised reconstruction shows that AFT-Net demonstrated the ability to reconstruct the data with significantly accelerate acquisition and random Gaussian noise. The proposed AFT-Net is an efficient and accurate approach for MRI reconstruction and preprocessing from raw data.

**Impact:** MRI reconstruction and preprocessing with AFT-Net should be able to determine the domain-manifold mapping and process k-space data directly, which shows superior performance and can be served as an efficient and accurate approach for human high-field and low-field MRI acquisition.

## Introduction

The domain shift from real to complex domain in the context of deep neural networks has uncovered the potential of utilizing the rich representational capacity of complex numbers and boosted the development of complex-valued architectures[1,2,3]. A similar but inverse domain shift is mirrored in MRI reconstruction, where raw data is acquired in complex-valued k-space. It has been shown that k-space, as a low-dimensional feature space, can be leveraged in deep neural networks to determine the between-manifold mapping of domain transforms in low signal-to-noise settings[4]. However, most complex networks do not include end-to-end domain transformation. Here, we propose a complex-valued MRI reconstruction approach, which aims to get rid of conventional numerical methods in the workflow and preprocess the data in parallel. The framework described here is the artificial Fourier transform (AFT). We utilize AFT combined with complex-valued UNet to design our AFT-Net as shown in Figure 1.

## Methods

### Complex-valued neural network

The real-valued network is extended to complex-valued network by defining the complex operator as  $\mathbf{W} = \mathbf{W}_{real} + i\mathbf{W}_{imag}$ , where  $\mathbf{W}_{real}$  and  $\mathbf{W}_{imag}$  are real-valued operators (linear or convolution operator). The output of  $\mathbf{W}$  acting on  $x$  can then denoted as complex matrix multiplications:

$$y = \mathbf{W} * x = (\mathbf{W}_{real} * x_{real} - \mathbf{W}_{imag} * x_{imag}) + i(\mathbf{W}_{imag} * x_{real} + \mathbf{W}_{real} * x_{imag}).$$

We extend the ReLU in the same way, which preserves consistency and shows comparable performance[3]. Normalization is used in deep learning to accelerate training and reduce statistical covariance shift[5]. This is mirrored in the complex-valued neural network, where we want to ensure that both real and imaginary parts have equal variance. Extending the normalization equation to matrix notation we have:

$$\tilde{z} = V^{-\frac{1}{2}}(z - \mathbb{E}(z)) = \begin{pmatrix} \text{Cov}(\Re(z), \Re(z)) & \text{Cov}(\Re(z), \Im(z)) \\ \text{Cov}(\Im(z), \Re(z)) & \text{Cov}(\Im(z), \Im(z)) \end{pmatrix} + \epsilon I)^{-\frac{1}{2}} \cdot \begin{bmatrix} \Re(z) - \text{Mean}(\Re(z)) \\ \Im(z) - \text{Mean}(\Im(z)) \end{bmatrix},$$

where  $x - \mathbb{E}(x)$  zero centers real and imaginary parts separately,  $V$  is the covariance matrix and  $\epsilon I$  is added to guarantee the existence of the inverse square root.

### Implementation detail

The general workflows are shown in Figure 2. We apply AFT-Net to the normal-field and high-field MRI of humans. Multiple network architectures are evaluated for accelerated and denoised reconstruction to verify the effectiveness of both AFT and CUNet in different domains. We refer to each of them as **AFT**, **AFT-Net (I)**, **AFT-Net (K)**, and **AFT-Net (KI)**.

## Results

The reconstruction results using fully-sampled k-space data are shown in Figure 3. It can be seen that the ground truth image obtained from FT is identical to the AFT prediction, which human observers can not distinguish. The residual map (pixel-wise difference between the ground truth image and the AFT prediction) shows that no brain structural information is presented. The grid-like remaining error is mainly caused by precision loss

during floating-point calculations in matrix multiplication.

In Figure 4, we show the results of accelerated reconstruction using under-sampled k-space data. In the first row, we see the reconstructions from 1D 4x equal-spaced sampling. Different AFT-Net structures were compared with zero-filling method. AFT-Net (KI) performs outstanding reconstruction, where less structural difference can be seen from the residual map in the second row. The third row shows zoomed-in areas of both images and residual maps. AFT-Net (I) produces more blurry reconstruction that loses the structural details. Reconstruction through AFT-Net (K) induces foggy artifacts, which is reflected in terms of SSIM.

Next, we illustrate the results of denoised reconstruction using k-space data with added Gaussian noise in Figure 5. AFT-Net (I) performs the best across all structures. The second row shows the pixel-wise difference between AFT-Net output and noiseless ground truth. It can be indicated that the noise in the background is attenuated significantly. Although the brain structure can be seen from the residual map, the zoomed-in version of the image shows that the AFT-Net reconstruction preserves the anatomy.

## Discussion and Conclusion

A novel artificial Fourier transform framework is proposed that determines the mapping between k-space and image domain as conventional DFT while having the ability to be fine-tuned/optimized with further training. The flexibility of AFT allows it to be easily incorporated into existing deep learning networks as learnable or static blocks. AFT is then utilized to design our AFT-Net, which implements complex-valued UNet to extract higher features in k-space and/or image domain. We aim to combine reconstruction and acceleration/denoising tasks into a unified network that enhances the image quality by removing artifacts directly from the k-space. Our AFT-Net achieves competitive results and proves to be more robust to noise and contrast differences. One remaining limitation is that we only implement a complex-valued network with linear layers and CNNs, which are less effective than some advanced architectures. In our future work, we will aim to replace multi-layer perceptrons and CNNs with transformer-based and diffusion-based models while extending the concept of AFT to more medical imaging tasks.

## Code Availability

The code is available at <https://github.com/yangyangting233/AFT-Net>.

## Acknowledgements

This study was supported by the Zuckerman Mind Brain Behavior Institute at Columbia University and Columbia MR Research Center site.

## References

- [1] Georgiou GM, Koutsougeras C. Complex domain backpropagation. IEEE transactions on Circuits and systems II: analog and digital signal processing. 1992 May;39(5):330-4.
- [2] Guberman N. On complex valued convolutional neural networks. arXiv preprint arXiv:1602.09046. 2016 Feb 29.
- [3] Trabelsi C, Bilaniuk O, Zhang Y, Serdyuk D, Subramanian S, Santos JF, et al. Deep Complex Networks. 2018 ICLR (Poster)
- [4] Zhu B, Liu JZ, Cauley SF, Rosen BR, Rosen MS. Image reconstruction by domain-transform manifold learning. Nature. 2018 Mar 22;555(7697):487-92.
- [5] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In International conference on machine learning 2015 Jun 1 (pp. 448-456). pmlr.

## Figures

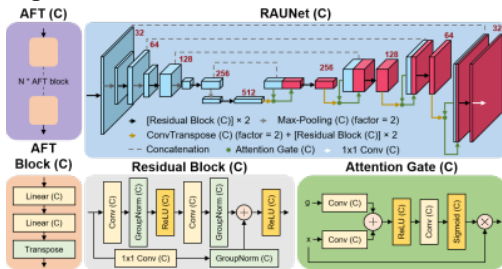


Figure 1. Structure of an N-dimensional AFT-Net ( $N = 2$  for 2D k-space input). Components include the complex-valued AFT block, the complex-valued residual attention UNet, the complex-valued residual block, and the complex-valued attention gate. All convolution layers have a kernel size of 3, except those pointed out specifically. The parameters of each convolution layer are fine-tuned so that the input and output sizes are the same. C: complex-valued. Red numbers indicate the number of channels produced by each layer.

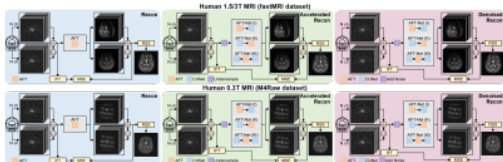


Figure 2. Workflows of experiments on each dataset. Blue block: Workflow of the reconstruction. The ground truth is derived by applying inversed Fourier transform to the input k-space data. Green block: Workflow of the accelerated reconstruction. The input k-space data is under-sampled by setting lines to zeros in the phase-encoding direction. Red block: Workflow of the denoised reconstruction. Random Gaussian noise is added to both the real and imaginary part of the input k-space data. IFT: Inverse Fourier transform. MSE: Mean square error. RSS: Root sum of square.

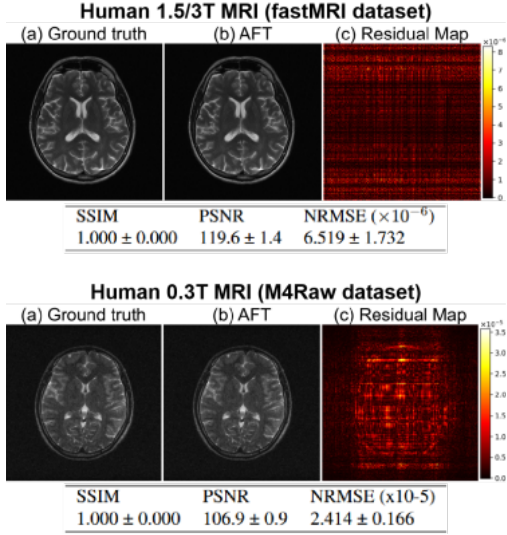


Figure 3. Human normal-field and low-field MRI reconstruction results. (a) Ground truth, (b) proposed method, (c) difference magnitude of (a) and (b) (in Hot colormap). Numbers are presented as mean value  $\pm$  standard deviation.

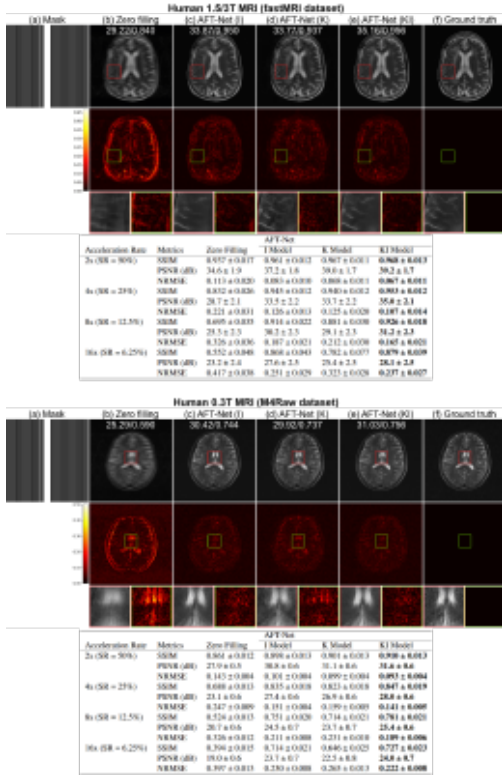
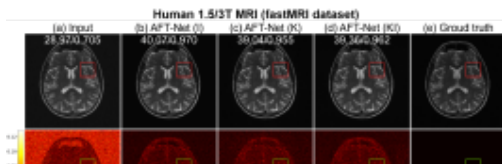


Figure 4. Human normal-field and low-field MRI accelerated reconstruction results. (a) The sub-sampling mask, (b) zero filling, (c)-(e) proposed methods, and (f) ground truth. 1st row: 1D 4x equal-spaced sampling (8% of low-frequency columns are retained), 2nd row: difference magnitude against (f) (in Hot colormap), 3rd row: zoomed-in version of the indicated box. White numbers in the upper center location indicate PSNR (db) and SSIM. Numbers in the table are presented as mean value  $\pm$  standard deviation. Numbers in boldface indicate the best metric out of all the methods.



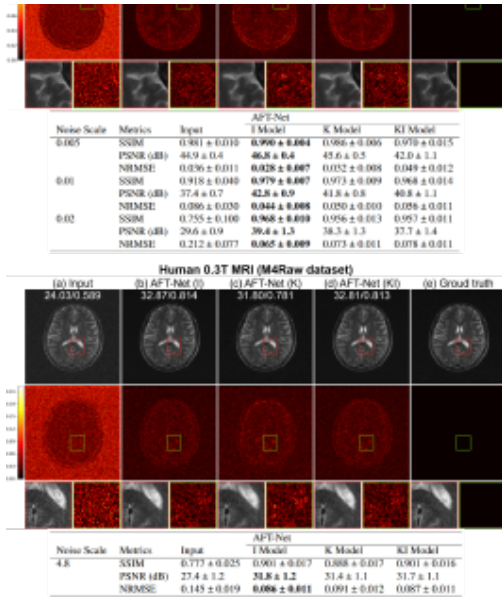


Figure 5. Human normal-field and low-field MRI denoised reconstruction results. (a) Input, (b)-(d) proposed methods, and (e) ground truth. 1st row: Randomly added Gaussian noise in k-space (scale = 0.02 for normal-field MRI and scale = 4.8 for low-field MRI), 2nd row: difference magnitude against (e) (in Hot colormap), 3rd row: zoomed-in version of the indicated box. White numbers indicate PSNR (db) and SSIM. Numbers in the table are presented as mean value  $\pm$  standard deviation. Numbers in boldface indicate the best metric out of all the methods.