

# Deep Learning-based MRS Reconstruction with Artificial Fourier Transform Network (AFT-Net)

Yanting Yang<sup>1</sup>, Matthieu Dagonmmer<sup>1</sup>, and Jia Guo<sup>1</sup>  
<sup>1</sup>Columbia University, New York, NY, United States

## Synopsis

**Motivation:** Complex-valued deep neural network has not been fully investigated in MRS reconstruction and preprocessing.

**Goal(s):** We aim to solve the spectroscopy inverse problems in domain transform from FIDs to spectra, especially for accelerated MRS reconstruction.

**Approach:** A complex-valued deep learning framework artificial Fourier transform network (AFT-Net) is proposed to directly reconstruct and process the complex-valued raw data in the sensor domain.

**Results:** Evaluation of different acceleration rates was performed on the in vivo dataset. AFT-Net demonstrated the ability to reconstruct the data under up to 80 times acceleration rate. The proposed AFT-Net is an efficient and accurate approach for MEGA-PRESS MRS accelerated reconstruction.

**Impact:** MRS reconstruction and preprocessing with AFT-Net should be able to determine the domain-manifold mapping and process FID data directly, which shows superior performance compared with numerical method and can be served as an efficient and accurate approach for MRS acquisition.

## Introduction

MRS is widely used to quantify metabolic chemical changes in brains, which provides crucial information on brain health. However, due to scanner variability and subject motion, frequency and phase shifts may arise, affecting data quality. The vast majority of present deep learning-based methods are based on real-valued operations, and few complex-valued methods[1] have fully leveraged the rich representational capacity of complex numbers to develop complex-valued architectures. Here, we propose a complex-valued MRS reconstruction approach, which aims to reconstruct and denoise the FID 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, convolution and non-linear[2] operator). The output of  $\mathbf{W}$  acting on  $x$  can then denoted as complex matrix multiplications:

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

The normalization[3] equation can also be extended as:

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

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 workflow is shown in Figure 2. We first trained AFT-Net on the simulated MEGA-PRESS dataset and then on the in vivo Big GABA dataset[4]. A total number of 101 subjects acquired by the Philips scanners were used in the training. For each subject, a standard GABA+-edited MRS acquisition was run, where ON editing pulses were placed at 1.9 ppm and OFF editing pulses were placed at 7.46 ppm. The acquisition number is 320 (160 ON and 160 OFF transients) per subject. The DIFF spectra are denoted as the subtraction of the ON and OFF spectra.

The ground truth of the ON/OFF/DIFF spectra is derived by taking the average over 160 acquisitions. For the training, we combined randomly sampled acquisitions of each subject to retrieve a signal with noise. Decreasing the number of samples results in higher noise and acceleration rate, which is defined as the ratio of the total acquisition number and the number of acquisitions sampled. This quantity is very handy to assess the power of denoising methods in practical terms. Retrieving accurate denoised signals at a high acceleration rate has implications for the potential reduction of total experimental time.

## Results

The results of the AFT-Net approach and conventional numerical methods with Gaussian line broadening are illustrated in Figure 3. The first row shows the reconstructed spectrum from the numerical methods and the proposed AFT-Net. The second row indicates the reconstructed spectrum overlaid with the ground truth. The third row plots the difference between the reconstructed spectrum and the ground truth. Under an acceleration rate of 80, where only 2 acquisitions were used over all 160 acquisitions, the AFT-Net shows excellent performance at high acceleration factors. The AFT-Net outperforms other methods for the DIFF spectra, indicating that the AFT-Net removes the noise in the FIDs while preserving the subject-level features.

We used PSNR (dB), PCC, and SCC to measure the similarity between the reconstructed spectra and the ground truth, as shown in Figure 4. Take PSNR as an example, the metric value increases as the acceleration rate decreases, but the absolute difference between high and low acceleration rates is tiny ( $31.1 \pm 2.1$  for ON spectra under an acceleration rate of 160 vs.  $32.7 \pm 1.9$  for ON spectra under an acceleration rate of 10). In addition, AFT-Net outperforms the DFT+GLB (Gaussian Line Broadening) method across all metrics (Figure 5).

## Discussion and Conclusion

A novel artificial Fourier transform framework is proposed that determines the mapping between FIDs and spectra 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 any existing deep learning network as learnable or static blocks. We then utilized AFT to design our AFT-Net, which implements complex-valued UNet to extract higher features in the temporal domain. We aim to combine reconstruction and denoising into a unified network that simultaneously enhances spectrum quality by removing artifacts directly from the FIDs. Our AFT-Net achieves competitive results and proves to be more robust to noise. 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 spectroscopy tasks.

## Code Availability

The code is available at <https://github.com/yangyanting233/AFT-Net>.

## Acknowledgements

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

## References

- [1] Ma DJ, Yang Y, Harguindeguy N, Tian Y, Small SA, Liu F, Rothman DL, Guo J. Magnetic Resonance Spectroscopy Spectral Registration Using Deep Learning. *Journal of Magnetic Resonance Imaging*. 2022 Sep 3.
- [2] Trabelsi C, Bilaniuk O, Zhang Y, Serdyuk D, Subramanian S, Santos JF, et al. Deep Complex Networks. 2018 ICLR (Poster)
- [3] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *International conference on machine learning* 2015 Jun 1 (pp. 448-456). pmlr.
- [4] Mikkelsen M, Barker PB, Bhattacharyya PK, Brix MK, Buur PF, Cecil KM, Chan KL, Chen DY, Craven AR, Cuypers K, Dacko M. Big GABA: Edited MR spectroscopy at 24 research sites. *Neuroimage*. 2017 Oct 1;159:32-45.

## Figures

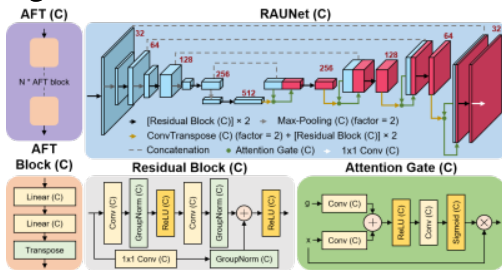


Figure 1. Structure of an N-dimensional AFT-Net ( $N = 1$  for 1D FID input). Components include the complex-valued AFT block, the complex-valued residual attention UNet, the complex-valued residual block, and the complex-valued attention gate. All convolution layers have a kernel size of 3, except those 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.

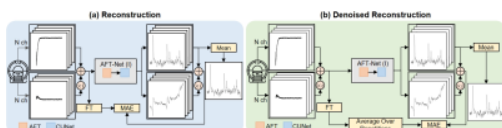


Figure 2. Workflows of experiments on the MRS dataset. (a) Workflow of MRS reconstruction. The inputs are randomly sampled FIDs from a single acquisition, and the outputs are reconstructed spectra. The ground truth is derived by applying the Fourier transform to the input FIDs. (b) Workflow of denoised MRS reconstruction. Here, we illustrate the case of an acceleration rate of 160, where the inputs are randomly sampled FIDs from a single acquisition and the ground truth is derived by averaging over the total 160 acquisitions. FT: Fourier transform. MAE: Mean absolute error.

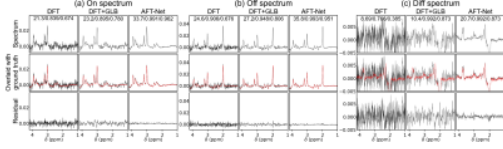


Figure 3. Qualitative results of human 3T MRS denoised reconstruction. The acceleration rate is 80 for each spectrum. (a) Reconstruction results for the ON spectrum; (b) Reconstruction results for the off spectrum; (c) Reconstruction results for the DIFF spectrum derived from (a) and (b). 1st row: reconstructed spectra; 2nd row: reconstructed spectra overlaid with ground truth (in red line); 3rd row: difference of reconstructed spectra against ground truth. Black numbers in the upper center location indicate PSNR (dB), PCC, and SCC, respectively.

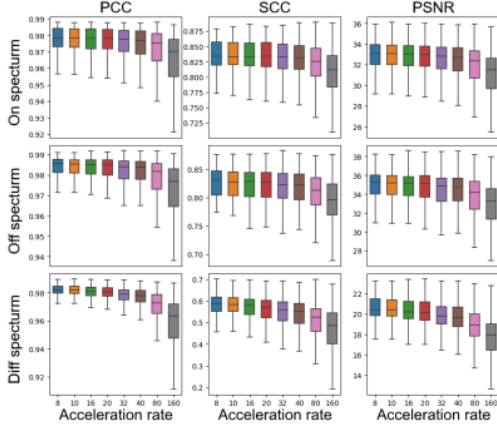


Figure 4. Metrics results of human 3T MRS denoised reconstruction for ON, OFF, and DIFF spectra respectively. Metrics are compared under acceleration rates 8, 10, 16, 20, 32, 40, 80, and 160 for each spectrum. PCC: Pearson correlation coefficient. SCC: Spearman's rank correlation coefficient. PSNR: Peak signal-to-noise ratio.

Acceleration Rate	Spectrum	Metric	DFT	DFT+GLB	AFT
10	On	PSNR	28.5 ± 1.1	29.4 ± 1.0	<b>32.7 ± 1.9</b>
		PCC	0.941 ± 0.012	0.951 ± 0.011	<b>0.975 ± 0.012</b>
		SCC	0.711 ± 0.045	0.751 ± 0.051	<b>0.836 ± 0.023</b>
	Off	PSNR	32.0 ± 1.1	31.4 ± 0.8	<b>34.8 ± 1.9</b>
		PCC	0.968 ± 0.006	0.962 ± 0.008	<b>0.982 ± 0.009</b>
		SCC	0.726 ± 0.033	0.771 ± 0.035	<b>0.826 ± 0.025</b>
	On/Off Mismatch	PSNR	15.2 ± 1.4	15.3 ± 1.1	<b>20.6 ± 1.1</b>
		PCC	0.902 ± 0.019	0.927 ± 0.013	<b>0.982 ± 0.004</b>
		SCC	0.497 ± 0.044	0.485 ± 0.056	<b>0.580 ± 0.059</b>
20	On	PSNR	25.5 ± 1.1	27.5 ± 1.2	<b>32.6 ± 1.9</b>
		PCC	0.891 ± 0.020	0.925 ± 0.018	<b>0.974 ± 0.013</b>
		SCC	0.615 ± 0.050	0.677 ± 0.057	<b>0.834 ± 0.025</b>
	Off	PSNR	29.0 ± 1.1	30.2 ± 1.0	<b>34.6 ± 2.0</b>
		PCC	0.939 ± 0.011	0.951 ± 0.011	<b>0.981 ± 0.009</b>
		SCC	0.635 ± 0.037	0.706 ± 0.042	<b>0.824 ± 0.027</b>
	On/Off Mismatch	PSNR	13.9 ± 1.6	13.8 ± 1.3	<b>20.2 ± 1.2</b>
		PCC	0.828 ± 0.029	0.889 ± 0.022	<b>0.980 ± 0.005</b>
		SCC	0.393 ± 0.049	0.406 ± 0.066	<b>0.560 ± 0.069</b>
40	On	PSNR	22.7 ± 1.1	25.3 ± 1.2	<b>32.3 ± 1.9</b>
		PCC	0.813 ± 0.030	0.883 ± 0.029	<b>0.973 ± 0.014</b>
		SCC	0.517 ± 0.054	0.598 ± 0.067	<b>0.830 ± 0.028</b>
	Off	PSNR	26.1 ± 1.1	28.5 ± 1.0	<b>34.3 ± 2.0</b>
		PCC	0.890 ± 0.017	0.930 ± 0.016	<b>0.980 ± 0.010</b>
		SCC	0.542 ± 0.037	0.633 ± 0.046	<b>0.820 ± 0.028</b>
	On/Off Mismatch	PSNR	13.1 ± 1.7	12.8 ± 1.4	<b>19.7 ± 1.4</b>
		PCC	0.725 ± 0.038	0.830 ± 0.033	<b>0.977 ± 0.007</b>
		SCC	0.309 ± 0.050	0.336 ± 0.070	<b>0.541 ± 0.073</b>
80	On	PSNR	19.9 ± 1.1	22.6 ± 1.3	<b>31.9 ± 2.0</b>
		PCC	0.702 ± 0.041	0.809 ± 0.044	<b>0.970 ± 0.015</b>
		SCC	0.417 ± 0.057	0.506 ± 0.073	<b>0.821 ± 0.035</b>
	Off	PSNR	23.2 ± 1.1	26.4 ± 1.1	<b>33.8 ± 2.1</b>
		PCC	0.889 ± 0.027	0.893 ± 0.022	<b>0.977 ± 0.012</b>
		SCC	0.447 ± 0.037	0.552 ± 0.049	<b>0.810 ± 0.033</b>
	On/Off Mismatch	PSNR	12.7 ± 1.8	12.2 ± 1.6	<b>18.9 ± 1.6</b>
		PCC	0.595 ± 0.044	0.736 ± 0.045	<b>0.970 ± 0.014</b>
		SCC	0.238 ± 0.051	0.270 ± 0.074	<b>0.510 ± 0.083</b>
160	On	PSNR	17.3 ± 1.2	19.9 ± 1.2	<b>31.1 ± 2.1</b>
		PCC	0.572 ± 0.049	0.706 ± 0.058	<b>0.964 ± 0.019</b>
		SCC	0.329 ± 0.057	0.417 ± 0.076	<b>0.804 ± 0.045</b>
	Off	PSNR	20.4 ± 1.1	23.9 ± 1.1	<b>32.9 ± 2.3</b>
		PCC	0.696 ± 0.034	0.827 ± 0.032	<b>0.972 ± 0.016</b>
		SCC	0.359 ± 0.039	0.464 ± 0.053	<b>0.793 ± 0.040</b>
	On/Off Mismatch	PSNR	12.5 ± 1.9	11.8 ± 1.7	<b>17.7 ± 1.9</b>
		PCC	0.464 ± 0.048	0.615 ± 0.058	<b>0.957 ± 0.023</b>
		SCC	0.178 ± 0.052	0.207 ± 0.078	<b>0.469 ± 0.101</b>

Figure 5. Quantitative results of human 3T MRS denoised reconstruction. Numbers are presented as mean value ± standard deviation. Numbers in boldface indicate the best metric out of all the methods.