

# BETTER THAN REAL: COMPLEX-VALUED NEURAL NETS FOR MRI FINGERPRINTING

Patrick Virtue<sup>\*†</sup> Stella X. Yu<sup>\*†</sup> Michael Lustig<sup>\*</sup>

<sup>\*</sup> University of California, Berkeley / <sup>†</sup>International Computer Science Institute

## ABSTRACT

The task of MRI fingerprinting is to identify tissue parameters from complex-valued MRI signals. The prevalent approach is dictionary based, where a test MRI signal is compared to stored MRI signals with known tissue parameters and the most similar signals and tissue parameters retrieved. Such an approach does not scale with the number of parameters and is rather slow when the tissue parameter space is large.

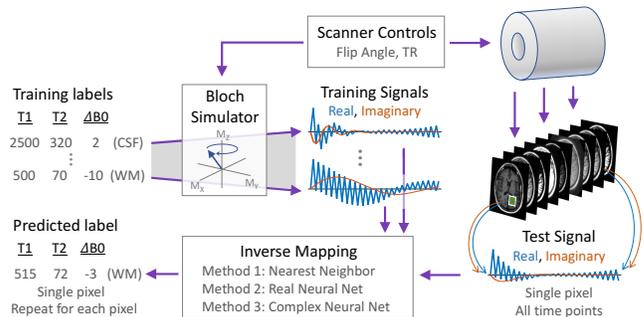
Our first novel contribution is to use deep learning as an efficient nonlinear inverse mapping approach. We generate synthetic (tissue, MRI) data from an MRI simulator, and use them to train a deep net to map the MRI signal to the tissue parameters directly. Our second novel contribution is to develop a complex-valued neural network with new cardioid activation functions. Our results demonstrate that complex-valued neural nets could be much more accurate than real-valued neural nets at complex-valued MRI fingerprinting.

**Index Terms**— Magnetic Resonance Fingerprinting, Parameter Mapping, Complex-valued Neural Networks

## 1. INTRODUCTION

Fingerprinting in the magnetic resonance imaging (MRI) domain [1] quantifies tissue parameters from complex-valued MRI signals. Tissue in the body may be characterized by how it interacts with the magnetic field during a MRI scan. Two tissue parameters, T1 and T2, are exponential time constants, e.g.  $e^{-t/T_2}$ , that describe how fast hydrogen protons in different tissues react to the applied magnetic field. For example, T1 and T2 values allow us to discern the boundary between gray matter (T1= 830, T2= 80) and white matter (T1= 500, T2= 70) in MRI brain images. These parameters also enable radiologists to differentiate between benign and malignant tissues.

Traditional MRI requires *many different scans* that each accentuates one of the desired parameters. Additionally, those scans only provide a *qualitative* visual contrast between tissues, e.g. *in this image, tissues with high T1 are brighter than other tissues*. MRI fingerprinting as proposed in [1], however, simultaneously produces *quantitative* values for T1, T2, and proton density in *one single scan*. It can also provide information about imperfections in the parameters for the applied magnetic field, i.e. B0 and B1.



**Fig. 1.** MRI fingerprinting is an inverse mapping problem that infers the tissue parameters from MRI signals. MRI simulator turns a ground-truth (T1, T2, B0) parameter tuple into an observed MRI temporal signal. The inverse mapping, by either nearest neighbor search or a neural network, solves for the tissue parameters given the simulated signal. At test time, the MRI signal will arrive from the scanner rather than the simulator. Example complex-valued signals are shown for cerebral spinal fluid (CSF) and white matter (WM) parameters.

MRI fingerprinting works by scanning the subject using a predetermined progression of scanner controls, e.g. flip angles and repetition time (TR) values of the pulse sequence. The various tissues in the body will react to this pulse sequence, producing measurable signals that have unique signatures depending on their specific tissue parameter (T1, T2, proton density) and applied magnetic field parameters (B0, B1). Just like a fingerprint pattern can identify a specific person, these measured signals may be decoded to determine the tissue and magnetic field parameters at each pixel location in the image. Figure 1 shows how MRI fingerprinting uses a numerical simulator to convert parameters into MRI signals, which are then used to train an algorithm to solve the inverse problem of mapping the signal back to the original parameters. When scanning a patient, the unlabeled MRI signals will arrive from the scanner to be decoded by the inverse mapping algorithm.

From a machine learning perspective, the MRI simulator provides a potentially unlimited number of complex-valued MRI signals for training, each of these temporal sequences is paired with a tuple of real-valued labels (T1, T2, B0). At the test time, the MRI scanner acquires a temporal signal at each pixel location, which must be decoded into to the tissue

and MRI field parameters. Prior MRI fingerprinting works [1, 2, 3] have used a nearest neighbor search based approach to match the measured signal to a dictionary of simulated training signals. Due to the non-parametric nature of nearest neighbor methods, the computation time scales linearly with the size of the dictionary, quickly becoming infeasible with a finer parameter resolution or when more tissue parameters are required. Additional research has improved the nearest neighbor matching efficiency by incorporating SVD [4] and group matching [5].

Rather than non-parametric nearest neighbor-based methods, we propose to learn a parameterized model for solving the MRI fingerprinting inverse mapping problem. Specifically, we demonstrate that feedforward neural networks can accurately model the complex non-linear MRI fingerprinting inverse mapping function with a computational efficiency that does not scale with the number of training examples.

We also investigate using complex-valued neural networks for MRI fingerprinting, since the MRI signals are inherently complex-valued. While complex-valued signals can be represented by 2-channel real signals, each channel containing real and imaginary components respectively, such a representation does not respect the phase information that is captured by complex algebra. Indeed, by introducing a new complex activation function for complex neural networks, we demonstrate that complex-valued neural nets are more effective than real-valued networks at MRI fingerprinting.

## 2. COMPLEX-VALUED NEURAL NETWORKS

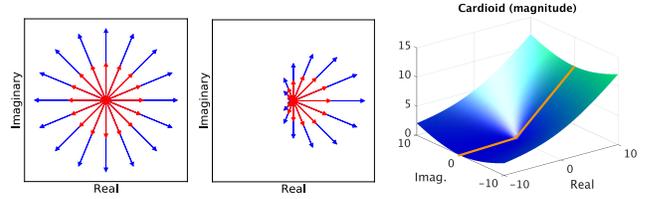
A significant facet of complex network research since the early 1990's has been to overcome the issue that standard real-valued non-linear layers do not transfer well to complex-valued networks. We tackle several aspects here.

### 2.1. Complex Cardioid Activation Function

Standard non-linear functions are either unbounded, e.g. sigmoid( $i\pi$ ), or undefined, e.g. the max operator in max pooling and ReLU. With complex outputs, we have also lost the probabilistic interpretations that functions like sigmoid and softmax provide. Past research has explored a range of potential solutions, for example, limiting the range of the activation input to avoid unbounded regions [6], or applying non-linearities to real and imaginary components separately [7, 8]. In their 1992 paper [9], Georgiou and Koutsougeras presented an activation that attenuates the magnitude of the signal while preserving the phase. In our experiments, we refer to this activation function as **siglog** as it modifies the magnitude by applying the sigmoid of the log of the magnitude:

$$\text{siglog}(z) = \frac{z}{1 + |z|} = g(\log(|z|))e^{-i\angle z} \quad (1)$$

$$g(z) = 1/(1 + e^{-z}). \quad (2)$$



**Fig. 2.** Our new cardioid activation function is a phase sensitive complex extension of ReLU. Left / Center: Each arrow indicates a sample input/output of our cardioid function on the complex plane. Right: The magnitude transformation of the cardioid function shows that it is reduced to ReLU on the real axis (orange line).

We propose a new complex activation function, **complex cardioid**, which is sensitive to the input phase rather than the input magnitude. The output magnitude is attenuated based on the input phase, while the output phase remains equal to the input phase. The complex cardioid is defined as:

$$f(z) = \frac{1}{2}(1 + \cos(\angle z))z \quad (3)$$

With this activation, input values that lie on the positive real axis are scaled by one, input values on the negative real axis are scaled by zero, and input values with nonzero imaginary components are gradually scaled from one to zero as the complex number rotates in phase from positive real axis towards the negative real axis. When the input values are restricted to real values, the complex cardioid function is simply the ReLU activation function. The  $\mathbb{C}\mathbb{R}$  [10] derivatives are as follows:

$$\frac{\partial f}{\partial z} = \frac{1}{2} + \frac{1}{2} \cos(\angle z) + \frac{i}{4} \sin(\angle z) \quad (4)$$

$$\frac{\partial f}{\partial \bar{z}} = \frac{-i}{4} \sin(\angle z) \frac{z}{\bar{z}} \quad (5)$$

### 2.2. Complex Calculus and Optimization

We leverage Wirtinger calculus, or  $\mathbb{C}\mathbb{R}$  calculus, [11, 10] to do gradient descent on functions that are not complex differentiable as long as they are differentiable with respect to their real and imaginary components. The first of two  $\mathbb{C}\mathbb{R}$  calculus derivatives is the  $\mathbb{R}$ -derivative (or real derivative) which computes  $\partial f / \partial z$  by treating  $z$  as a real variable and holding instances of  $\bar{z}$  constant. Likewise, the other derivative is the **conjugate  $\mathbb{R}$ -derivative**,  $\partial f / \partial \bar{z}$ , where  $\bar{z}$  acts as a real variable and  $z$  is held constant.

To optimize a real-valued loss function ( $L(w)$ ) at the end of a complex feedforward neural net, we update the weights by applying the complex version of gradient descent:

$$w = w - \alpha \nabla_{\bar{w}} L, \text{ where } \nabla_{\bar{w}} L = \left[ \frac{\partial L}{\partial \bar{w}_1} \dots \frac{\partial L}{\partial \bar{w}_n} \right]^T \quad (6)$$

This is the same as real-valued gradient descent with careful attention paid to the gradient operator. As shown in [12], the direction of steepest descent is the complex cogradient,  $\nabla_{\bar{w}} L$ .

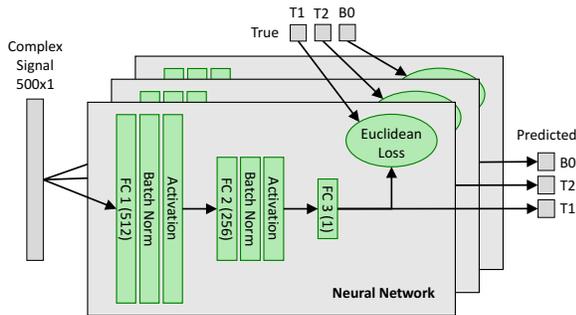
### 3. MRI FINGERPRINTING EXPERIMENTS

**Training Data.** We simulate the MRI signal with the Bloch equations and the first of the two pulse sequence parameters from [1] with signal length 500. We use 100,000 simulated points for training, randomly sampled with the same T1, T2, B0 density as used in the baseline nearest neighbor dictionary.

**Testing Data.** Following [2], we test our methods with a numerical MRI phantom [13, 14] with the T1, T2, and proton density values specified for each tissue type in [14]. We add a liner ramp in the B0 field across the image from -60 Hz to 60 Hz. We compute proton density from the norm of the test signal as in [1]. Although we do not include any B1 inhomogeneity in our experiments, a fourth neural network could easily be added to incorporate this or any other parameter(s). In addition to testing with a clean signal from the phantom, we also tested phantom signals with complex-valued Gaussian noise added to produce a peak signal-to-noise ratio (pSNR) of 40.

**Methods.** Fig.3 shows our 3-layer neural network architecture. We compare six MRI fingerprinting methods.

1. Baseline inner product nearest neighbor and T1, T2, B0 dictionary setup used in [1].
2. Real-valued neural nets with 2-channel real/imaginary inputs representing complex MRI signals, using the ReLU activation function.
3. Real-valued neural nets that are twice as wide as the second model, with 1024 and 512 feature channels in the two hidden layers.
4. Complex-valued neural nets with 1-channel complex MRI signals, using our new cardioid activation function.
5. Complex-valued neural nets using separable sigmoid activation functions (i.e. sigmoid applied to real and imaginary independently) [7].
6. Complex-valued neural nets using the siglog activation function [9].



**Fig. 3.** Fully connected neural network architecture, repeated for each desired output label (T1, T2, B0).

**Table 1.** NRMSE results: fingerprinting from clean signals.

Network	T1	T2	$\Delta$ B0
Nearest neighbor	10.63	39.78	<b>1.02</b>
2-ch real/imaginary network	2.71	8.21	2.11
2-ch real/imaginary network 2x	2.21	8.04	2.44
<b>Complex (cardioid)</b>	<b>1.42</b>	<b>4.34</b>	1.32
Complex (separable sigmoid)	4.72	9.24	3.33
Complex (siglog)	2.99	12.04	3.05

**Table 2.** NRMSE results: fingerprinting from noisy signals.

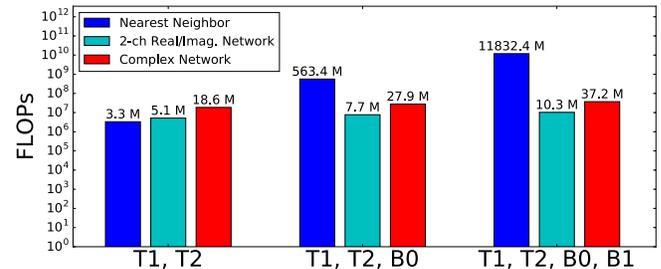
Network	T1	T2	$\Delta$ B0
Nearest neighbor	12.21	40.38	<b>1.08</b>
2-ch real/imaginary network	11.15	<b>17.96</b>	5.23
2-ch real/imaginary network 2x	11.08	22.15	7.08
<b>Complex (cardioid)</b>	<b>9.40</b>	20.98	4.43
Complex (separable sigmoid)	17.31	33.09	18.83
Complex (siglog)	102.22	237.88	266.33

Here we focus on pixel-wise fingerprinting reconstruction. We plan to extend our approach to full image predictions for under-sampled MRI fingerprinting in the future.

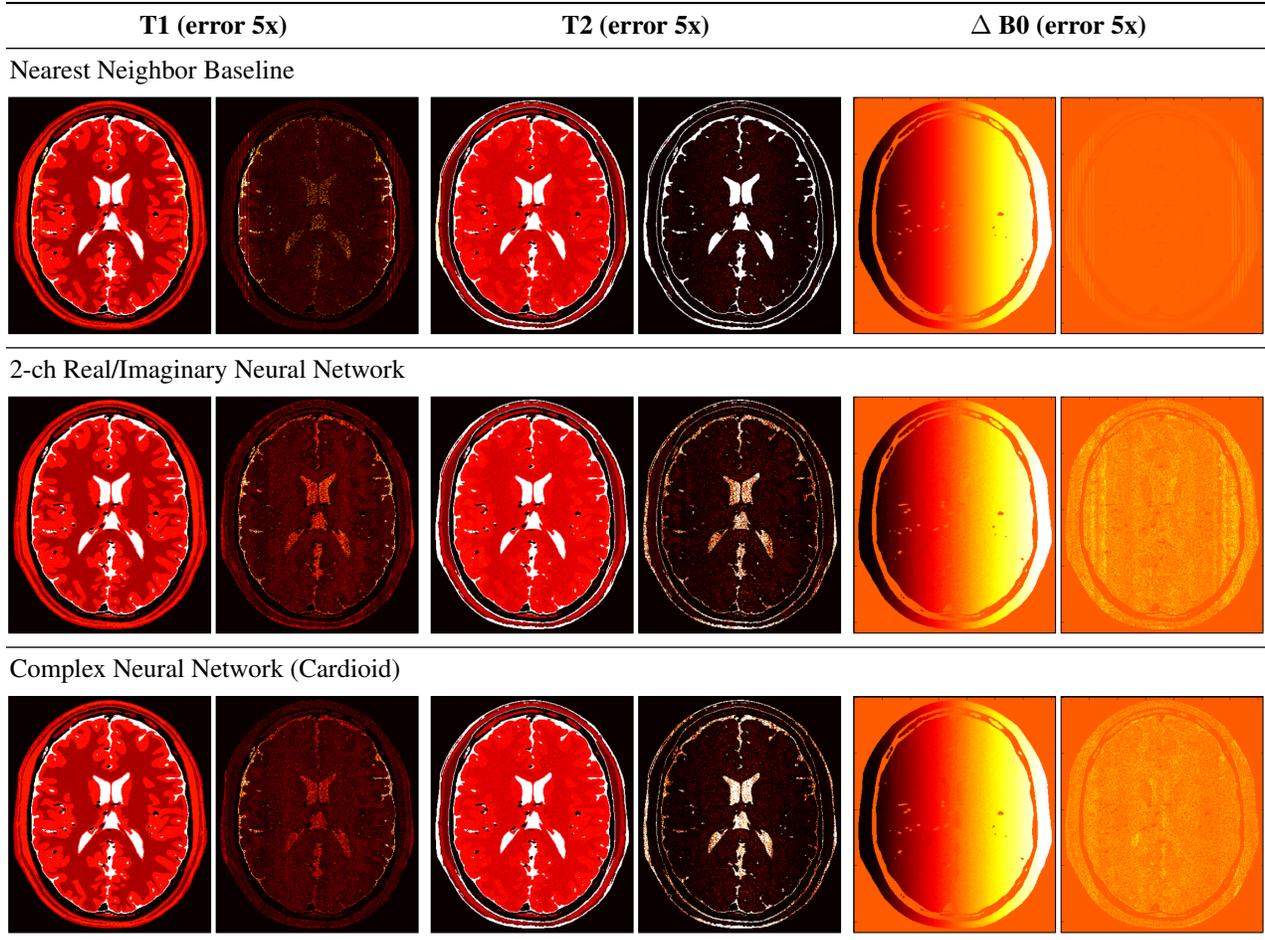
**Deep Learning Implementation.** We implement all our networks in Caffe [15]. For the complex-valued neural nets, we extend the Caffe platform with complex versions of the fully connected layer, batch normalization layer, and complex activation layers, including the  $\mathbb{C}\mathbb{R}$  calculus back propagation for all the layer functions.

**Results.** Tables 1 and 2 compare the prediction accuracy at no noise and pSNR=40 noise level, respectively. Fig. 5 shows sample reconstruction results. Fig. 4 compares the computational efficiency in terms of the number of floating point operations (FLOPs). We observe the following:

1. A dictionary based approach explodes exponentially with more outputs and becomes infeasible. Compared to the two outputs (T1,T2), the #FLOPs increases by  $171\times$  for the three outputs (T1,T2,B0), and by  $3,585\times$  for the four outputs (T1,T2,B0,B1).



**Fig. 4.** Comparison of floating point operations required to compute the parameters for a single pixel. Note the log scale.



**Fig. 5.** Numerical phantom with added noise (pSNR=40). Predicted quantitative parameters maps images are shown adjacent to the error image. For visualization purposes, the error images are displayed at 5x the scale of the images.

2. Inverse mapping by neural nets outperforms the traditional nearest neighbor baseline on T1 and T2 values, whereas the nearest neighbor approach predicts B0 values more accurately.
3. Complex-valued neural networks outperform 2-channel real-valued networks for almost all of our experiments, and this advantage cannot be explained away by the twice large model capacity, suggesting that complex-valued networks can bring out information in the complex data more effectively than treating them as arbitrary two-channel real data.
4. The complex cardioid activation significantly outperformed both the separable sigmoid and siglog activation functions, allowing complex networks to not only compete with, but surpass, real-valued networks.

**Summary.** For the complex-valued MRI fingerprinting problem, we propose a deep learning approach that implements an efficient nonlinear inverse mapping function that turns MRI

signals to tissue parameters directly<sup>1</sup>. We generate synthetic (tissue, MRI) data from an MRI simulator, and use them to train a neural network. We develop a novel cardioid activation function that enables the successful real-world application of complex-valued neural networks. Our results demonstrate that complex-valued nets can be more accurate than real-valued nets at complex-valued MRI fingerprinting.

#### 4. ACKNOWLEDGEMENTS

This research was supported in part by the National Institutes of Health R01EB009690 grant and a Sloan Research Fellowship. We thank Michael Kellman, Frank Ong, Jonathan Tamir, and Hong Shang for great discussions about complex calculus, fingerprinting, pulse sequences, and simulator software.

<sup>1</sup>A conference abstract exploring neural networks for MRI fingerprinting [16] was concurrently published with this work.

## 5. REFERENCES

- [1] Dan Ma, Vikas Gulani, Nicole Seiberlich, Kecheng Liu, Jeffrey L Sunshine, Jeffrey L Duerk, and Mark A Griswold, "Magnetic resonance fingerprinting," *Nature*, vol. 495, no. 7440, pp. 187–192, 2013.
- [2] Eric Y Pierre, Dan Ma, Yong Chen, Chaitra Badve, and Mark A Griswold, "Multiscale reconstruction for MR fingerprinting," *Magnetic resonance in medicine*, 2015.
- [3] Jakob Assländer, Steffen J Glaser, and Jürgen Hennig, "Pseudo steady-state free precession for MR-fingerprinting," *Magnetic resonance in medicine*, 2016.
- [4] Debra F McGivney, Eric Pierre, Dan Ma, Yun Jiang, Haris Saybasili, Vikas Gulani, and Mark A Griswold, "SVD compression for magnetic resonance fingerprinting in the time domain," *IEEE transactions on medical imaging*, vol. 33, no. 12, pp. 2311–2322, 2014.
- [5] Stephen F Cauley, Kawin Setsompop, Dan Ma, Yun Jiang, Huihui Ye, Elfar Adalsteinsson, Mark A Griswold, and Lawrence L Wald, "Fast group matching for MR fingerprinting reconstruction," *Magnetic resonance in medicine*, vol. 74, no. 2, pp. 523–528, 2015.
- [6] H. Leung and S. Haykin, "The complex backpropagation algorithm," *IEEE Transactions on Signal Processing*, vol. 39, no. 9, pp. 2101–2104, Sep 1991.
- [7] Tohru Nitta, "An extension of the back-propagation algorithm to complex numbers," *Neural Netw.*, vol. 10, no. 9, pp. 1391–1415, Nov. 1997.
- [8] Md. Faijul Amin and Kazuyuki Murase, "Single-layered complex-valued neural network for real-valued classification problems," *Neurocomput.*, vol. 72, no. 4-6, pp. 945–955, Jan. 2009.
- [9] G. M. Georgiou and C. Koutsougeras, "Complex domain backpropagation," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 39, no. 5, pp. 330–334, May 1992.
- [10] K. Kreutz-Delgado, "The Complex Gradient Operator and the CR-Calculus," *ArXiv e-prints*, June 2009.
- [11] W. Wirtinger, "Zur formalen theorie der funktionen von mehr komplexen veränderlichen," *Mathematische Annalen*, vol. 97, no. 1, pp. 357–375, 1927.
- [12] DH Brandwood, "A complex gradient operator and its application in adaptive array theory," in *IEE Proceedings F-Communications, Radar and Signal Processing*. IET, 1983, vol. 130, pp. 11–16.
- [13] D Louis Collins, Alex P Zijdenbos, Vasken Kollokian, John G Sled, Noor J Kabani, Colin J Holmes, and Alan C Evans, "Design and construction of a realistic digital brain phantom," *IEEE transactions on medical imaging*, vol. 17, no. 3, pp. 463–468, 1998.
- [14] Berengere Aubert-Broche, Alan C Evans, and Louis Collins, "A new improved version of the realistic digital brain phantom," *NeuroImage*, vol. 32, no. 1, pp. 138–145, 2006.
- [15] Yangqing Jia, Evan Shelhamer, Jeff Donahue, Sergey Karayev, Jonathan Long, Ross Girshick, Sergio Guadarrama, and Trevor Darrell, "Caffe: Convolutional architecture for fast feature embedding," *arXiv preprint arXiv:1408.5093*, 2014.
- [16] Ouri Cohen, Bo Zhu, and Matthew Rosen, "Deep learning for fast MR fingerprinting reconstruction," in *2017 Scientific Meeting Proceedings*. International Society for Magnetic Resonance in Medicine, 2017, p. 688.