

FAST. ACCURATE. TRUSTED.



Everything you would expect from Medical Physics' leading QA software for over 25 years, plus so much more.

LET RIT'S NEW SOFTWARE UPGRADE EXCEED YOUR EXPECTATIONS:

## THE RIT FAMILY OF PRODUCTS **VERSION 6.7.X**



Achieve sub-pixel accuracy with the enhanced 3D Winston-Lutz (Isocenter Optimization) routine.



Perform efficient and accurate Elekta Leaf Speed Analysis, with minimal setup.



Cerberus 2.0: Completely streamline your imaging QA workflow with hands-free automation.



Set custom tolerance values for every measurement used in automated phantom analyses with RIT's new Tolerance Manager.



Easily perform exact dose comparisons with TomoTherapy® Registration for Patient QA.



Manage software licenses at your convenience with RIT's new, cloud-based system.



Fully-customize your software experience with RIT's updated, dynamic interface.

**DEMO VERSION 6.7.X  
AT ASTRO 2018:  
BOOTH #2025**

RADIOLOGICAL IMAGING TECHNOLOGY, INC.  
**RADIMAGE.COM**

(+1) 719.590.1077 x211 // [sales@radimage.com](mailto:sales@radimage.com)

Connect with RIT  
**@RIT4QA**



©2018, Radiological Imaging Technology, Inc.  
TomoTherapy® is a registered trademark of Accuray, Inc.

# HYDRA: Hybrid Deep Magnetic Resonance Fingerprinting

Pingfan Song<sup>1</sup> Yonina C. Eldar<sup>2</sup> Gal Mazor<sup>3</sup> Miguel R. D. Rodrigues<sup>4</sup>

## Abstract

**Purpose:** Magnetic resonance fingerprinting (MRF) methods typically rely on dictionary matching to map the temporal MRF signals to quantitative tissue parameters. Such approaches suffer from inherent discretization errors, as well as high computational complexity as the dictionary size grows. To alleviate these issues, we propose a HYbrid Deep magnetic Resonance fingerprinting approach, referred to as HYDRA.

**Methods:** HYDRA involves two stages: a model-based signature restoration phase and a learning-based parameter restoration phase. Signal restoration is implemented using low-rank based de-aliasing techniques while parameter restoration is performed using a deep nonlocal residual convolutional neural network. The designed network is trained on synthesized MRF data simulated with the Bloch equations and fast imaging with steady state precession (FISP) sequences. In test mode, it takes a temporal MRF signal as input and produces the corresponding tissue parameters.

**Results:** We validated our approach on both synthetic data and anatomical data generated from a healthy subject. The results demonstrate that, in contrast to conventional dictionary-matching based MRF techniques, our approach significantly improves inference speed by eliminating the time-consuming dictionary matching operation, and alleviates discretization errors by outputting continuous-valued parameters. We further avoid the need to store a large dictionary, thus reducing memory requirements.

**Conclusions:** Our approach demonstrates advantages in terms of inference speed, accuracy and storage requirements over competing MRF methods.

**Key words:** Magnetic Resonance Fingerprinting, Quantitative Magnetic Resonance Imaging, Deep Learning, Nonlocal Residual Convolutional Neural Network, Self-attention

---

*This work was supported by the Royal Society International Exchange Scheme IE160348, by the European Union's Horizon 2020 grant ERC-BNYQ, by the Israel Science Foundation grant no. 335/14, by ICore: the Israeli Excellence Center 'Circle of Light', by the Ministry of Science and Technology, Israel, by UCL Overseas Research Scholarship (UCL-ORS) and by China Scholarship Council (CSC) and by EPSRC grant EP/K033166/1.*

<sup>1</sup> *Department of Electronic and Electrical Engineering, Imperial College London, UK*

<sup>2</sup> *Faculty of Mathematics and Computer Science, Weizmann Institute of Science, Israel*

<sup>3</sup> *Department of Electrical Engineering, Technion – Israel Institute of Technology, Israel*

<sup>4</sup> *Department of Electronic and Electrical Engineering, University College London, UK*

*p.song@imperial.ac.uk, yonina.eldar@weizmann.ac.il, galmazor@technion.ac.il, m.rodrigues@ucl.ac.uk*

---

## 1. Introduction

Magnetic Resonance Fingerprinting (MRF)<sup>1,2,3,4,5,6,7,8</sup> has emerged as a promising Quantitative Magnetic Resonance Imaging (QMRI) approach, with the capability of providing multiple tissue's intrinsic spin parameters simultaneously, such as the spin-lattice magnetic relaxation time (T1) and the spin-spin magnetic relaxation time (T2). Based on the fact that the response from each tissue with respect to a given pseudo-random pulse sequence is unique, MRF exploits pseudo-randomized acquisition parameters to create unique temporal signal signatures, analogous to a "fingerprint", for different tissues. A dictionary matching operation is then performed to map an inquiry temporal signature to the best matching entry in a precomputed dictionary, leading to multiple tissue parameters directly.

The temporal signatures are generated by varying the acquisition parameters of a pseudo-random excitation pulse sequence, such as repetition time (TR), time of echo (TE), and radio frequency flip angle (FA) over time. The dictionary is composed of a large number of entries that are usually simulated by the Bloch equations given pseudo-random pulse sequences. Each entry represents a unique temporal signature associated with a specific tissue and its quantitative parameters, such as the T1 and T2 relaxation times. Thus, once the best matching (i.e. most correlated) entry is found, it directly leads to multiple quantitative parameters simultaneously via a lookup-table operation.

MRI physics and physiological constraints make the MR scanning procedure time-consuming. To shorten acquisition time, subsampling is commonly performed in k-space (a.k.a conjugate Fourier transform domain) in order to reduce the number of samples and accelerate imaging speed. However, such k-space subsampling results in temporal signatures that are corrupted by aliasing, blurring and noise. This hampers the accuracy associated with estimation of the tissue parameters using a dictionary matching procedure. In order to alleviate the impact of such distortion and corruption, de-aliasing operations are often exploited to restore cleaner signatures before performing signature-to-parameter mapping. Therefore, MRF reconstruction usually involves two operations: signature restoration and parameter restoration.

Inspired by the successful application of sparsity-driven image processing approaches in MRI reconstruction<sup>9,10,11,12</sup>, several works<sup>3,4,5,6</sup> suggest to incorporate prior knowledge such as sparsity and low-rank to attenuate distortion and corruption, improving the signature

71 restoration performance, during the initial MRF reconstruction stage. This is then followed  
72 by a dictionary matching operation, performing mapping from purified temporal signatures  
73 to tissue’s quantitative parameters. However, such dictionary matching based signature-  
74 to-parameter mapping exhibits several drawbacks<sup>13,14</sup>. Since the simulated dictionary and  
75 lookup-table contain a finite number of elements, they can only cover a limited number  
76 of discrete values for each type of tissue parameter. We refer to the difference between a  
77 continuous-valued tissue parameter and its closest available discrete value on a lattice as the  
78 discretization error. For example, a pair of dictionary and lookup-table that contain 101  
79 elements will lead to a discretization error of maximum 25 ms if they cover the range of 0 ms  
80 - 5000 ms with a fixed interval of 50 ms for a specific tissue parameter, e.g. T1. To reduce  
81 the discretization error, a huge dictionary that is composed of a large number of entries is  
82 needed to represent tissues with fine granularity over the entire value range of target tissue  
83 parameters. However, storing a large dictionary becomes prohibitively memory-consuming,  
84 as the dictionary size and density often increase exponentially with the number of tissue  
85 parameters. Specifically, the number of entries in a dictionary will be  $P^s$  for  $s$  parameters  
86 each containing  $P$  values, since every combination of these  $s$  parameters determines a specific  
87 tissue which is characterized by a specific signature. For example, given T1, T2 relaxations,  
88 i.e.  $s = 2$ , if each of them contains 1000 values, the dictionary will have  $1000^2$  entries.  
89 In addition, finding the best matching entry becomes computationally intense for a large  
90 dictionary, considerably limiting the inference speed.

91 In this paper, we propose an alternative approach to dictionary matching based on deep  
92 neural networks (a.k.a. deep learning)<sup>15,16</sup>, which we refer to as HYDRA: HYbrid Deep mag-  
93 netic Resonance fingerprinting. The motivation derives from the fact that a well designed  
94 and tuned deep neural network is capable of approximating complex functions, leading to  
95 state-of-the-art results in a number of tasks such as image classification, super-resolution,  
96 speech recognition, and more<sup>17,18,19,20,21,22,23</sup>. Recent work<sup>13,14</sup> proposed to exploit neural  
97 networks to replace the dictionary and lookup-table used in conventional MRF reconstruc-  
98 tion approaches. These proposed neural networks suffer from two limitations: First, these  
99 approaches are based on neural network models containing only 3-layers, thus suffer from  
100 limited capacity of capturing complex mapping functions. Second, these methods focused  
101 exclusively on parameter restoration stage (the second stage in MRF reconstruction), but not  
102 on signature restoration (the first stage in MRF reconstruction). Therefore, these techniques

103 rely on fully-sampled data instead of typically available sub-sampled k-space data.

104 Different from Cohen et al.’s fully-connected feed-forward neural network<sup>13</sup>, and Hoppe  
105 et al.’s vanilla convolutional neural network (CNN)<sup>14</sup>, the proposed HYDRA involves both  
106 a signature restoration and a parameter restoration phase. Signature restoration is im-  
107 plemented using a low-rank based de-aliasing method adapted from Mazor et al.<sup>6</sup> while  
108 parameter restoration is implemented using a deep nonlocal residual convolutional neural  
109 network developed for this purpose. Our key contributions with respect to prior work are:

- 110 • HYDRA is, to the best of our knowledge, the first deep network approach to combine  
111 model-based de-aliasing and learning-based parameter mapping. HYDRA eliminates  
112 the requirement for the memory and time-consuming dictionary matching operation,  
113 thus significantly improving inference speed without compromising on reconstruction  
114 performance.
- 115 • A 1D nonlocal residual convolutional neural network is designed to capture the map-  
116 pings from temporal MRF signals to tissue parameters. Owing to residual learning  
117 and a self-attention mechanism, our network is deeper and more sophisticated than  
118 competing network models. This allows to capture complex parameter mappings more  
119 effectively, and output continuous parameters to alleviate discretization issues.
- 120 • The designed network is trained on synthesized MRF data simulated with the Bloch  
121 equations, but is still applicable to anatomical data. This contributes to eliminating  
122 the requirement for a large amount of real MRF data.
- 123 • A low-rank based de-aliasing technique is developed in order to take advantage of  
124 temporal similarity for signature restoration.
- 125 • The low-rank based signature restoration is organically combined with the learning-  
126 based parameter restoration to achieve fast and accurate MRF reconstruction. Such  
127 strategy enables HYDRA to handle both fully-sampled k-space data and more impor-  
128 tantly sub-sampled k-space data.
- 129 • A series of numerical experiments are conducted to evaluate the proposed approach  
130 on both synthetic and anatomical data. The results demonstrate improved inference  
131 speed, accuracy and discretization errors over competing methods<sup>1,2,3,4,5,6,13,14</sup>.

132 The rest of the paper is organized as follows. In Section II., we formulate the MRF  
133 reconstruction problem, introduce related methods, and present our approach, involving the

134 use of a low-rank based signature restoration procedure together with a deep network for  
 135 parameter restoration. Section III. is devoted to experimental results, followed by a discussion  
 136 in Section IV. and a conclusion in Section V..

## 137 II. Materials and Methods

### 138 II.A. The MRF Problem Formulation

139 MRF data is composed of multiple frames sampled in k-space over time. A series of such  
 140 frames are vectorized and then stacked together along the temporal dimension to construct  
 141 a measurement matrix  $\mathbf{Y} \in \mathcal{C}^{Q \times L}$ , where  $Q$  is the number of k-space samples in each frame,  
 142 and  $L$  is the number of frames. Due to k-space subsampling, every column vector  $\mathbf{Y}_{:,i}$   
 143 represents a subsampled Fourier transform of a vectorized image frame  $\mathbf{X}_{:,i}$ :

$$144 \quad \mathbf{Y} = [\mathbf{Y}_{:,1}, \dots, \mathbf{Y}_{:,L}] = [F_u\{\mathbf{X}_{:,1}\}, \dots, F_u\{\mathbf{X}_{:,L}\}], \quad (1)$$

145 where  $F_u\{\cdot\}$  denotes a subsampled 2D Fourier transform.

146 Each column  $\mathbf{X}_{:,i}$  represents a MR contrast acquired with RF sequence parameters:

$$147 \quad \Theta_{:,i}^{TRE} = [TR^i, TE^i, FA^i]^T, \quad i \in [1, L] \quad (2)$$

148 where  $TR^i$  and  $TE^i$  denote the repetition time and echo time, respectively, and  $FA^i$  denotes  
 149 the flip angle of the RF pulse during sampling the  $i$ -th contrast. Every row  $\mathbf{X}_{j,:}$  represents  
 150 a temporal signature, i.e. temporal signal evolution of a specific tissue at the  $j$ -th image  
 151 pixel. The signature depends on the tissue's relaxation times, such as T1 and T2, grouped  
 152 as a row vector:

$$153 \quad \Theta_{j,:}^{T12} = [T1^j, T2^j], \quad j \in [1, N] \quad (3)$$

154 where,  $N$  denotes the number of pixels in each image frame. Note that,  $j$  is the spatial  
 155 index while  $i$  is the temporal index throughout. Given RF sequence parameters  $\Theta^{TRE}$ , and  
 156 parameters  $\Theta_{j,:}^{T12}$  of a specific tissue, its temporal signature  $\mathbf{X}_{j,:}$  can be derived as:

$$157 \quad \mathbf{X}_{j,:} = f(\Theta_{j,:}^{T12}, \Theta^{TRE}) \quad (4)$$

158 where  $f(\cdot)$  denotes the Bloch equations. This MR contrast matrix  $\mathbf{X}$  is associated with the  
 159 k-space measurements  $\mathbf{Y}$  per column by the subsampled Fourier transform, and it is related  
 160 to tissue parameters  $\Theta^{T12}$  per row by the Bloch equations, as illustrated in Fig. 1.

161 Given RF sequence parameters  $\Theta^{TRE}$  and k-space measurements  $\mathbf{Y}$ , the goal of MRF  
 162 reconstruction is to estimate the tissue parameters  $\Theta^{T12}$ . Typically, the image stack  $\mathbf{X}$  is  
 163 first reconstructed from  $\mathbf{Y}$ , referred to as signature restoration, and then mapped to tissue  
 164 parameters  $\Theta^{T12}$  via dictionary matching, referred to as parameter restoration<sup>1,2,3,4,5,6</sup>. This  
 165 process is illustrated in Fig. 2.

166 The dictionary is a collection of temporal signatures that are usually simulated by the  
 167 Bloch equations for various typical tissues, given the pseudo-random RF pulse sequences and  
 168 tissue parameters. Given an inquiry temporal signature, dictionary matching computes the  
 169 inner product between the temporal signature with each dictionary entry, selecting the entry  
 170 in the dictionary exhibiting the highest correlation with the inquiry one as the best matching  
 171 signature. Once the best entry is found, it directly leads to multiple tissue parameters, such  
 172 as T1, T2, simultaneously, via searching a lookup-table.

173 Let  $\mathbf{LUT} \in \mathcal{R}^{K \times 2}$  denote a lookup-table composed of  $K$  tissues, each containing 2  
 174 parameters, i.e., T1 and T2 relaxation times<sup>1</sup>. Let  $\mathbf{D} \in \mathcal{C}^{K \times L}$  denote the corresponding  
 175 dictionary simulated using Bloch equations given the RF sequence parameters  $\Theta^{TRE}$ , for-  
 176 mulated as  $\mathbf{D}_{k,:} = f(\mathbf{LUT}_{k,:}, \Theta^{TRE})$ . Since each temporal signature  $\mathbf{D}_{k,:}$  is linked with the  
 177  $k$ -th tissue's parameters  $\mathbf{LUT}_{k,:}$ , the choice of a large dictionary size  $K$  can in principle  
 178 provide enough granularity to capture a range of possible tissue values.

179 In conclusion, existing MRF reconstruction approaches involve two stages: signature  
 180 restoration and parameter restoration, that can be succinctly written as

$$181 \quad \Theta_{j,:}^{T12} = g(h(\mathbf{Y})_{j,:} | \Theta^{TRE}), \quad j \in [1, N], \quad (5)$$

182 where the function  $\mathbf{X} = h(\mathbf{Y})$  represents the signature restoration operation such as sparsity  
 183 or low-rank based de-aliasing and denoising methods, whereas  $g(\mathbf{X}_{j,:} | \Theta^{TRE})$  denotes the  
 184 parameter restoration operation, such as dictionary matching based methods<sup>1,2,3,4,5,6</sup>.

185 Our approach aims to perform signature restoration via low-rank based de-aliasing and  
 186 parameter restoration via a neural network in order to achieve improved MRF reconstruction  
 187 performance. We highlight that our method only requires a simulated dictionary during  
 188 network training. Once the network is trained, the dictionary is not needed anymore. In

---

<sup>1</sup>Note that the off resonance parameter, which appeared in the original MRF paper<sup>1</sup>, has been omitted here, since the sequence used in our experiments is derived from the FISP sequence, which is insensitive to off resonance effects<sup>2,6</sup>.

---

**Algorithm 1** Original MRF method<sup>1</sup>


---

**Input:**A set of subsampled k-space images:  $\mathbf{Y}$ A pre-simulated dictionary:  $\mathbf{D}$ An appropriate lookup-table:  $\mathbf{LUT}$ **Output:**Magnetic parameter maps:  $\hat{T}_1, \hat{T}_2$ **Step 1. Restore signatures:**

$$\hat{\mathbf{X}}_{:,i} = F_u^H \{\mathbf{Y}_{:,i}\}, \forall i$$

**Step 2. Restore parameters for every  $j$  via dictionary matching:**

$$\hat{k}_j = \arg \max_k \frac{\mathbf{Re} \langle \mathbf{D}_{k,:}, \hat{\mathbf{X}}_{j,:} \rangle}{\|\mathbf{D}_{k,:}\|_2^2}, \quad \hat{T}_1^j, \hat{T}_2^j = \mathbf{LUT}(\hat{k}_j)$$


---

189 addition, our approach also eliminates a simulated dictionary for signature restoration, which  
 190 is a key difference from FLOR<sup>5,6</sup> during signature restoration.

## 191 II.B. Previous Methods

192 **Dictionary Matching based MRF approaches.** The original MRF reconstruction al-  
 193 gorithm<sup>1</sup> is based on dictionary matching, as presented in Algorithm 1. It finds the best  
 194 matching dictionary entry for the acquired temporal signature according to their inner prod-  
 195 uct and then searches the lookup-table to obtain corresponding tissue parameters. Here,  
 196  $F_u^H \{\cdot\}$  denotes the inverse Fourier transform operating on the zero filled k-space data where  
 197 zeros are filled at the unknown frequencies and symbol  $\mathbf{Re} \langle \mathbf{a}, \mathbf{b} \rangle$  represents the real part of  
 198 the inner product of two vectors  $\mathbf{a}$  and  $\mathbf{b}$ .

199 Exploiting the nature of signals, by using appropriate prior knowledge, can often con-  
 200 tribute to improved signal processing performance. In this spirit, later works suggested  
 201 to incorporate sparsity in MRF reconstruction to further improve performance, inspired  
 202 by successful applications of sparsity in MRI reconstruction<sup>9,10,11</sup>. Davies et al.<sup>3</sup> proposed  
 203 Bloch response recovery via Iterative Projection (BLIP) which exploits sparsity in the dic-  
 204 tionary domain. BLIP consists of iterating between two main steps: (a) a gradient step  
 205 which enforces consistency with the measurements, based on the Projected Landweber Al-  
 206 gorithm (PLA) generalized from the iterative hard thresholding method; (b) a projection  
 207 which matches each row of  $\mathbf{X}$  to a single dictionary atom. Instead of exploiting sparsity in

208 the dictionary domain, Wang et al.<sup>4</sup> suggested to leverage sparsity in the wavelet domain of  
 209 each imaging frame,  $\mathbf{X}_{:,i}$ . They further replaced the Euclidean norm with the Mahalanobis  
 210 distance for dictionary matching. Considering that adjacent MR image frames along the  
 211 temporal dimension should exhibit high resemblance, Mazor et al.<sup>5,6</sup> proposed a magnetic  
 212 resonance Fingerprint with LOW-Rank prior for reconstructing the image stack and quanti-  
 213 tative parameters, referred to as FLOR, which achieved state-of-the-art performance. The  
 214 algorithm, described in Algorithm 2, relies on two priors: a low rank prior on the the matrix  
 215  $\mathbf{X}$ , and the fact that the rows of  $\mathbf{X}$  lie in the column space of the dictionary  $\mathbf{D}$ .

---

**Algorithm 2** FLOR<sup>6</sup>


---

**Input:**

A set of subsampled k-space images:  $\mathbf{Y}$ ; A pre-simulated dictionary:  $\mathbf{D}$ ; An appropriate lookup-  
 table:  $\mathbf{LUT}$ ; Parameters  $\mu$  for gradient step and  $\lambda$  for regularization

**Output:**

Magnetic parameter maps:  $\hat{T}_1, \hat{T}_2$

**Initialization:**

$\hat{\mathbf{X}}^0 = 0$ ,  $\mathbf{P} = \mathbf{D}^\dagger \mathbf{D}$ , where  $\mathbf{D}^\dagger$  is the pseudo-inverse of  $\mathbf{D}$ .

**Step 1. Restore signatures via iterating until convergence:**

- Gradient step for every  $i$ :

$$\hat{\mathbf{Z}}_{:,i}^{t+1} = \hat{\mathbf{X}}_{:,i}^t - \mu F_u^H \{F_u \{\hat{\mathbf{X}}_{:,i}^t\} - \mathbf{Y}_{:,i}\} \quad (6)$$

where the superscript  $t$  represents the index of iterations.

- Project onto the dictionary subspace:

$$[\mathbf{U}, \mathbf{S}, \mathbf{V}] = \text{svd}(\hat{\mathbf{Z}}^{t+1} \mathbf{P}) \quad (7)$$

where svd denotes the singular-value decomposition operation, and  $\mathbf{S} = \text{diag}(\{\sigma_j\})$  is a rectangular diagonal matrix with singular values  $\{\sigma_j\}$  on its diagonal.

- Soft-threshold the non-zero singular values with  $\lambda\mu$  and reconstruct signatures  $\hat{\mathbf{X}}^{t+1}$ :

$$\sigma'_j = \max\{\sigma_j - \lambda\mu, 0\}, \quad \hat{\mathbf{X}}^{t+1} = \mathbf{U} \mathbf{S}' \mathbf{V}^H \quad (8)$$

where  $\mathbf{S}' = \text{diag}(\{\sigma'_j\})$ .

**Step 2. Restore parameters for every  $j$  via dictionary matching:**

$$\hat{k}_j = \arg \max_k \frac{\text{Re} \langle \mathbf{D}_{k,:}, \hat{\mathbf{X}}_{j,:} \rangle}{\|\mathbf{D}_{k,:}\|_2^2}, \quad \hat{T}_1^j, \hat{T}_2^j = \mathbf{LUT}(\hat{k}_j)$$

---

216 **Learning-based MRF approaches.** The above techniques all use dictionary matching

217 to perform mapping from temporal signatures to tissue parameters. Therefore, these meth-  
 218 ods suffer from drawbacks such as discretization error, slow inference speed and memory-  
 219 consuming storage. In order to alleviate these issues, recent works<sup>13,14</sup> propose to exploit  
 220 neural networks to replace dictionaries and lookup-tables used in conventional MRF recon-  
 221 struction approaches. Cohen et al. suggest a fully-connected feed-forward neural network  
 222 (FNN)<sup>13</sup>. Since the input layer of the FNN is fully connected with the input temporal  
 223 signature, the number of neurons in the input layer corresponds to the length of the input  
 224 temporal signature. This makes the network structure less flexible, as a FNN network trained  
 225 on temporal signatures with a certain length is not applicable to temporal signatures with  
 226 a different length. In addition, the fully-connected structure results in rapid increase in the  
 227 number of parameters along with the growth of depth, making the network more suscepti-  
 228 ble to overfitting. Hoppe et al.<sup>14</sup> propose a 3-layer vanilla CNN for parameter restoration.  
 229 Both<sup>13</sup> and<sup>14</sup> focus exclusively on learning the signature-to-parameter mapping from a pair  
 230 of dictionary and lookup-table simulated using the Bloch equations. During the validation,  
 231 they assume that clean temporal signatures are available as input into the trained networks.  
 232 However, since temporal signatures obtained from k-space subsampled MRF data are al-  
 233 ways contaminated by aliasing and noise, their approaches, when applied directly in such  
 234 k-space subsampling situations, suffer from heavy artifacts introduced during the signature  
 235 restoration phase, leading to poor performance.

## 236 II.C. Proposed Methods

237 The proposed hybrid deep magnetic resonance fingerprinting (HYDRA) approach, summa-  
 238 rized in Algorithm 3, consists of two stages: signature restoration and parameter restoration,  
 239 (see also (5)). As illustrated in Fig. 3, a low-rank based de-aliasing method is used to restore  
 240 signatures, and then a 1D nonlocal residual convolutional neural network is used to map  
 241 each restored signature to corresponding tissue parameters.

242 In particular, given  $\Theta^{TRE}$  and k-space samples  $\mathbf{Y}$ , in our proposed approach, the func-  
 243 tion  $\mathbf{X} = h(\mathbf{Y})$  in (5) represents a signature restoration operation using low-rank based  
 244 de-aliasing techniques without requiring a dictionary. The function  $\Theta_{j,:}^{T12} = g(\mathbf{X}_{j,:}|\Theta^{TRE})$   
 245 in (5) represents a parameter restoration operation that exploits a trained neural network to  
 246 map each restored signature  $\mathbf{X}_{j,:}$  to corresponding tissue parameters  $\Theta_{j,:}^{T12}$  directly. In the  
 247 subsequent sections, we provide a detailed description of both stages of our technique.

### 248 II.C.1. Low-rank based signature restoration

249 Since MRF data consists of multiple frames exhibiting temporal similarity, the imaging  
 250 contrasts matrix  $\mathbf{X}$  is typically a low-rank matrix<sup>6</sup>. Therefore,  $h(\cdot)$  leverages a low-rank  
 251 prior for denoising and de-aliasing, formulated as

$$252 \quad h(\mathbf{Y}) = \arg \min_{\mathbf{X}} \frac{1}{2} \sum_i \|\mathbf{Y}_{:,i} - F_u\{\mathbf{X}_{:,i}\}\|_2^2 \quad (9)$$

s.t.  $\text{rank}(\mathbf{X}) < r$

253 where the parameter  $r$  is the rank of the matrix, a fixed pre-chosen parameter. Since typically  
 254  $r$  is not known in advance, we consider a relaxed regularized version:

$$255 \quad h(\mathbf{Y}) = \arg \min_{\mathbf{X}} \frac{1}{2} \sum_i \|\mathbf{Y}_{:,i} - F_u\{\mathbf{X}_{:,i}\}\|_2^2 + \lambda \|\mathbf{X}\|_* \quad (10)$$

256 where  $\|\mathbf{X}\|_*$  denotes the nuclear norm<sup>24</sup> of  $\mathbf{X}$ , defined as the sum of the singular values of  
 257  $\mathbf{X}$ , and  $\lambda$  is the Lagrangian multiplier manually selected for balancing data fidelity and the  
 258 rank. Problem (10) can be solved using the incremental subgradient proximal method<sup>25</sup>,  
 259 similar as to FLOR<sup>6</sup>. The procedure for solving (5) is shown in Algorithm 3.

260 One of differences from FLOR<sup>6</sup> is the fact that we removed the operation of project-  
 261 ing the temporal signal onto a dictionary. This allows to eliminate the requirement for a  
 262 simulated dictionary in the signature restoration stage, which also alleviates the memory  
 263 consumption issue. In addition, the computational complexity is reduced by  $N \cdot L^2$  floating-  
 264 point operations in each iteration, where  $L$  is the dimension of a dictionary element, and  $N$   
 265 is the number of pixels in each image frame. On the other hand, the gained benefits are at  
 266 the price of requiring more iterations to converge. Another difference from FLOR<sup>6</sup> is that  
 267 we exploit a network, instead of dictionary matching, for signature-to-parameter mapping.

### 268 II.C.2. Learning-based parameter restoration

269 Once the imaging contrasts matrix  $\mathbf{X}$  is recovered from the k-space samples  $\mathbf{Y}$ , each temporal  
 270 signature  $\mathbf{X}_{j,:}$  is input into the trained network for parameter restoration, formulated as:

$$271 \quad \Theta_{j,:}^{T12} = g(\mathbf{X}_{j,:} | \Theta^{TRE}), \quad j \in [1, N] \quad (11)$$

272 where  $g(\cdot)$  denotes the trained network,  $\Theta^{TRE}$  denotes the fixed RF sequence parameters.

273 We next describe the network structure, training and testing procedures.

#### 274 Network structure.

---

**Algorithm 3** Proposed MRF reconstruction approach: HYDRA
 

---

**Input:**A set of subsampled k-space images:  $\mathbf{Y}$ The trained network:  $g$ Parameters  $\mu$  for gradient step and  $\lambda$  for regularization**Output:** Magnetic parameter maps  $\hat{T}_1, \hat{T}_2$ **Initialization:**  $\hat{\mathbf{X}}^0 = 0$ **Step 1. Restore signatures via iterating until convergence:**

- Gradient step for every  $i$ , the same as (6).
- Perform SVD:

$$[\mathbf{U}, \mathbf{S}, \mathbf{V}] = \text{svd}(\hat{\mathbf{Z}}^{t+1})$$

- Soft-threshold the non-zero singular values  $\{\sigma_j\}$  of  $\mathbf{S}$  with parameter  $\lambda\mu$  and reconstruct signatures  $\hat{\mathbf{X}}^{t+1}$ , the same as (8).

**Step 2. Restore parameters for every  $j$  via the trained network:**

$$\hat{T}_1^j, \hat{T}_2^j = g(\hat{\mathbf{X}}_{j,:})$$


---

275 The proposed network has a 1D nonlocal residual CNN architecture with short-cuts  
 276 for residual learning and nonlocal operations for achieving a self-attention mechanism. As  
 277 illustrated in Fig. 3, it starts with two 1D convolutional layers before connecting with 4  
 278 residual / non-local operation blocks, and finally ends with a global-average-pooling layer  
 279 followed by a fully-connected layer. Every residual block is followed by a non-local operation  
 280 block. Four such blocks are interspersed with each other.

281 Each residual block contains a max-pooling layer with stride 2, two convolution layers  
 282 and a shortcut that enforces the network to learn the residual content. The filter size is set  
 283 to be equal to 21 throughout convolutional layers. The number of channels, a.k.a feature  
 284 maps, in the first two convolutional layers is set to 16 and then is doubled in subsequent four  
 285 residual blocks until 128 in the final residual block. The size of feature maps in the next  
 286 block halves in contrast with the previous one due to max-pooling. In this way, we gradually  
 287 reduce temporal resolution while extract more features to increase content information.

288 Inspired by the self-attention scheme and nonlocal neural networks<sup>26,27</sup>, non-local op-  
 289 erations are incorporated into the designed network to achieve the attention mechanism, in  
 290 order to capture long-range dependencies with fewer layers. In contrast to the progressive  
 291 behavior of convolutional operations that process one local neighborhood at a time, the non-

292 local operations compute the response at a position as a weighted sum of the features at all  
 293 positions in the feature maps. Formally, the nonlocal operation is formulated as<sup>27</sup>:

$$294 \quad \mathbf{y}_i = \frac{1}{C(\mathbf{x})} \sum_{\forall j} f(\mathbf{x}_i, \mathbf{x}_j) g(\mathbf{x}_j). \quad (12)$$

295 Here,  $i$  is the index of an output position and  $j$  is the index that enumerates all possible  
 296 temporal positions,  $\mathbf{x}$  is the input temporal signal or its features and  $\mathbf{y}$  is the output signal  
 297 of the same size as  $\mathbf{x}$ . A pairwise function  $f$  computes a scalar between  $i$  and all  $j$  to  
 298 represent the affinity relationship of these two positions. The unary function  $g$  computes a  
 299 representation of the input signal at position  $j$ . The response is normalized by a factor  $C(\mathbf{x})$ .  
 300 There exists a few instantiations for function  $f$  and  $g$ . For simplicity, the unary function  
 301  $g$  is chosen as a linear embedding:  $g(\mathbf{x}_j) = \mathbf{W}_g(\mathbf{x}_j)$ , where  $\mathbf{W}_g$  is a weight matrix to be  
 302 learned. Regarding the affinity matrix  $f$ , we adopt the embedded Gaussian to compute  
 303 similarity in an embedding space, which is formulated as:  $f(\mathbf{x}_i, \mathbf{x}_j) = e^{\theta(\mathbf{x}_i)^\top \phi(\mathbf{x}_j)}$ . Here,  
 304  $\theta(\mathbf{x}_i) = \mathbf{W}_\theta \mathbf{x}_i$  and  $\phi(\mathbf{x}_j) = \mathbf{W}_\phi \mathbf{x}_j$  are two learned embeddings. The normalization factor is  
 305 set as  $C(\mathbf{x}) = \sum_{\forall j} f(\mathbf{x}_i, \mathbf{x}_j)$ . For a given  $i$ ,  $\frac{1}{C(\mathbf{x})} f(\mathbf{x}_i, \mathbf{x}_j)$  becomes the softmax computation  
 306 along the dimension  $j$ , which leads to  $\mathbf{y} = \text{softmax}(\mathbf{x}^\top \mathbf{W}_\theta^\top \mathbf{W}_\phi \mathbf{x}) g(\mathbf{x})$ , which is the self-  
 307 attention form<sup>26</sup>.

308 The non-local behavior in (12) is due to the fact that all positions ( $j$ ) are considered in  
 309 the operation. As a comparison, a convolutional operation sums up the weighted input in  
 310 a local neighborhood<sup>27</sup>. It implies that the non-local operation directly captures long-range  
 311 dependencies in the temporal dimension via computing interactions between any two points,  
 312 regardless of their positional distance. In this way, the network is able to extract global  
 313 features and take advantage of the full receptive field in each layer.

314 The global-average-pooling layer is used to average each feature map in order to integrate  
 315 information in every channel for improved robustness to corrupted input data. This global-  
 316 average-pooling layer also reduces the number of parameters significantly, thus lessening the  
 317 computation cost as well as preventing over-fitting. The last fully-connected layer outputs  
 318 estimated parameters – T1 and T2 relaxation times. The designed network contains around  
 319 0.27 million parameters. The weights are initialized using He-normal-distribution<sup>28</sup>. The  
 320 max-norm kernel constraint<sup>29</sup> is exploited to regularize the weight matrix directly in order to  
 321 prevent over-fitting. The designed network can also be adapted for various MRF sequences,  
 322 such as the original MRF sequence – inversion-recovery balanced steady state free-precession

323 (IR-bSSFP) sequence, that depends also on the intrinsic  $T_2$  parameter. It is possible to adjust  
 324 the number of outputs to adapt to more parameters, such as proton density,  $B_0$ .

325 To summarize, our network is motivated and inspired by recent successful applications  
 326 of convolutional neural networks and variants. Convolutional neural networks have been  
 327 proved to be a powerful model to capture useful features from signals and images. By  
 328 introducing convolution, local receptive field and weight sharing design, a CNN is capable  
 329 of taking advantage of local spatial coherence and translation invariance characteristics in  
 330 the input signal, thus become especially well suited to extract relevant information at a low  
 331 computational cost<sup>17,18,19,20,21,22</sup>. On the other hand, the residual network architecture<sup>18,19</sup>  
 332 provides an effective way to design and train a deeper model, since it alleviates the gradient  
 333 vanishing or exploding problems by propagating gradients throughout the model via short-  
 334 cuts, a.k.a skip connections. By leveraging non-local operation based attention mechanism,  
 335 neural networks are endowed with capability of extracting global features and capturing  
 336 long-range dependencies.

### 337 **Network training.**

338 The designed network is trained on a synthesized dictionary  $\mathbf{D}$  and corresponding  
 339 lookup-table  $\mathbf{LUT}$  to learn the signature-to-parameter mappings  $\mathbf{LUT}_{k,:} = g(\mathbf{D}_{k,:} | \Theta^{TRE})$ .

340 The training dataset is synthesized as follows. First, we determine the range of tissue  
 341 parameters. For example, one may set  $T_1$  relaxation times to cover a range of [1, 5000] ms  
 342 and  $T_2$  relaxation times to cover a range of [1, 2000] ms with an increment of 10 ms for both.  
 343 Thus, the  $T_1$  and  $T_2$  values constitute a grid with dimension  $500 \times 200$ , in which each point  
 344 represents a specific combination of  $T_1$  and  $T_2$  values, and hence characterizes a specific  
 345 tissue. Points corresponding to  $T_1 < T_2$  have been excluded as such combinations have no  
 346 physical meaning. All the valid points are stacked together to generate a lookup-table. For  
 347 instance, the above setting for  $T_1$  and  $T_2$  leads to a lookup-table of dimension  $80100 \times 2$ .  
 348 The RF pulse sequences used in our work are fast imaging with steady state precession  
 349 (FISP) pulse sequences with parameters that have been used in previous publications in the  
 350 field of MRF<sup>2,6,8</sup>. Given the lookup-table and RF pulse sequences, dictionary entries can be  
 351 synthesized by solving the Bloch equations using the extended phase graph formalism<sup>30,31</sup>.

352 When the training dataset is ready, the dictionary entries are used as input signals and

353 corresponding lookup-table entries serve as the groundtruth. All the dictionary entries are  
 354 input into the designed network batch by batch which outputs estimated parameters. The  
 355 root mean square errors (RMSE) of the outputs are calculated with respect to corresponding  
 356 groundtruth. The resulting RMSE loss is then backpropagated from the output layer to the  
 357 first layer to update the weights and bias by using Adam<sup>32</sup> as the optimization algorithm.  
 358 More training details are provided in the subsequent experiment section. Once the training  
 359 procedure is completed, given an inquiry signal evolution  $\mathbf{X}_{j,:}$ , it is able to map such a time  
 360 sequence directly to corresponding tissue parameters, as formulated in (11), implying that  
 361 no dictionary or lookup-table are required during the inference. Since we only need to store  
 362 the trained network which is a compact model, it consumes less memory than storing the  
 363 dictionary and lookup-table.

364 We emphasize that even though the network is trained on a grid of tissue values, it is  
 365 expected to capture the mapping function from temporal signatures to tissue parameters.  
 366 Thus the trained network is capable of outputting tissue values not existing in the grid  
 367 of training values. Detailed results can be found in Fig. 7 and Table 2. This feature is  
 368 favorable, as it implies that well designed and trained networks have an ability to overcome  
 369 discretization issues. The overall procedures for solving (10) and (11) are shown in Algorithm  
 370 3.

### 371 III. Experimental Results

372 In this section, we conduct a series of experiments to evaluate our approach, comparing it  
 373 with other state-of-the-art MRF methods<sup>1,3,6,13,14</sup>.

374 The experiments are categorized into a few types: training, testing on synthetic data,  
 375 testing on anatomical data using variable density Gaussian sampling patterns and spiral  
 376 sampling patterns at different sampling ratios and number of time frames, as described in  
 377 Table 1. For the network training, synthesized temporal signatures, i.e. simulated dictionary  
 378 entries of  $\mathbf{D}$  shown as Fig. 4, are used as input signals and corresponding parameter values  
 379 in the lookup-table  $\mathbf{LUT}$  serve as the groundtruth. The proposed network is trained to  
 380 capture the signature-to-parameter mappings. For testing on synthetic data, synthesized  
 381 temporal signatures in  $\mathbf{X}$  are used as input signals and corresponding parameter values in  
 382  $\Theta^{T12}$  serve as the groundtruth. The aim is to test the parameter restoration performance

383 only. For testing on anatomical data, the k-space measurements  $\mathbf{Y}$  which are derived from  
 384 the Fourier transform (for Gaussian patterns) or non-uniform FFT (for spiral trajectories)<sup>33</sup>  
 385 of  $\mathbf{X}$ , are used as input and corresponding parameter values in  $\Theta^{T12}$  serve as reference. When  
 386 there is no k-space subsampling, the aim is to test the parameter restoration performance  
 387 only. When there exists k-space subsampling, the aim is to test the overall performance,  
 388 including both signature restoration and parameter restoration. More detailed descriptions  
 389 are provided in each subsection.

### 390 III.A. Training

391 As mentioned in Section II.C.2., the designed network is trained on a pair of synthesized  
 392 dictionary  $\mathbf{D}$  and lookup-table  $\mathbf{LUT}$ , simulated using Bloch equations and FISP pulse se-  
 393 quences<sup>2,6</sup>.

394 The FISP pulse sequence used in our experiments was designed with parameters  $\Theta_{:,i}^{TRE} =$   
 395  $[TR^i, TE^i, FA^i]^T, i \in [1, L]$  that have been used in previous publications in the field of  
 396 MRF<sup>2,6,8</sup>. The echo time  $TE^i$  was constant of 2ms. The repetition time  $TR^i$  was randomly  
 397 varied in the range of 11.5 - 14.5 ms with a Perlin noise pattern. All the flip angles  $FA^i, i \in$   
 398  $[1, L]$  constituted a sinusoidal variation in the range of 0 - 70 degrees to ensure smoothly  
 399 varying transient state of the magnetization, as shown in Figure 5.

400 For the range of tissue parameters, T1 relaxation times are set to cover a range of  
 401  $[1, 5000]$  ms and T2 relaxation times to cover a range of  $[1, 2000]$  ms with an increment  
 402 of 10 ms for both. Such parameter ranges cover the relaxation time values that can be  
 403 commonly found in a brain scan<sup>34</sup>. All the valid combinations of T1 and T2 values are  
 404 stacked together, generating a lookup-table  $\mathbf{LUT}$  of dimension  $K \times 2$  where  $K = 80100$ .  
 405 Given the lookup-table and RF pulse sequences, dictionary entries are synthesized by solving  
 406 the Bloch equations using the extended phase graph formalism, leading to a dictionary of  
 407 dimension  $K \times L$  where  $L = 200$  or 1000 is the number of time frames.

408 When the training dataset is ready, the dictionary entries are used as input signals and  
 409 corresponding lookup-table entries serve as the groundtruth to train the designed network,  
 410 as mentioned in Section II.C.2.. The model was trained for 50 epochs. It takes around 30  
 411 seconds for running one epoch on average, thus around 25 minutes for completing 50 epochs,  
 412 on a NVIDIA GeForce GTX 1080 Ti GPU. In each training epoch, 20% of the training

413 samples are separated aside for validation dataset. The learning rate decays from 1e-2 to  
 414 1e-6 every 10 epochs. Each batch was experimentally set to contain 256 time-sequences  
 415 in order to balance the convergence rate and weights updating rate well. For comparison  
 416 purposes, we also implemented Hoppe et al.’s CNN referring to<sup>14</sup>, and Cohen et al.’s FNN  
 417 referring to<sup>13</sup> with the same structure and parameters as specified in their papers. Then we  
 418 use the same GPU and training dataset to train their networks with specified learning rate  
 419 and number of epochs until convergence.

We adopt a few widely used metrics, such as root mean square error (RMSE), signal-to-noise ratio (SNR) and peak signal-to-noise ratio (PSNR) to evaluate the image quality quantitatively. The definitions of RMSE, SNR and PSNR are given as follows:

$$\text{RMSE} = \sqrt{\frac{\|\mathbf{X} - \hat{\mathbf{X}}\|_F^2}{N}}, \quad (13)$$

$$\text{SNR} = 20 \log_{10} \frac{\|\mathbf{X}\|_F^2}{\text{RMSE}}, \quad (14)$$

$$\text{PSNR} = 20 \log_{10} \frac{\text{PeakVal}}{\text{RMSE}}, \quad (15)$$

420 where matrices  $\mathbf{X}$  and  $\hat{\mathbf{X}}$  denote the ground truth signal and its reconstructed version,  
 421 respectively,  $N$  denotes the total number of elements in the signal and  $\|\cdot\|_F$  denotes the  
 422 Frobenius norm. PeakVal stands for the pixel peak value in an image, e.g., 1 for a normalized  
 423 signal.

### 424 III.B. Testing on synthetic dataset

425 In this subsection, we evaluate the performance of HYDRA on a synthetic testing dataset.  
 426 The procedures of constructing a synthetic testing dataset is similar to the construction of  
 427 the training dataset: 500 different T1 values are randomly selected from 1 - 5000 ms, while  
 428 200 different T2 values are randomly selected from 1 - 2000 ms, using random permutation  
 429 based on uniformly distributed pseudorandom numbers. All the valid combinations from  
 430 the selected T1 and T2 values are stacked together, generating a parameter matrix  $\Theta^{T12}$   
 431 of dimension  $80000 \times 2$  with  $N = 80000$ . The RF pulse sequences are the same as in the  
 432 training stage. Given the parameter matrix and RF pulse sequences, input signal signatures  
 433 are synthesized by solving the Bloch equations using the extended phase graph formalism,  
 434 leading to a signature matrix  $\mathbf{X}$  of dimension  $N \times L = 80000 \times 200$ , with each row representing

435 a temporal signature corresponding to a specific combination of T1 and T2 values. The  
 436 signature matrix  $\mathbf{X}$  and parameter matrix  $\Theta^{T12}$  constitute the synthetic testing dataset,  
 437 with  $\mathbf{X}$  as input and  $\Theta^{T12}$  as the groundtruth.

438 We input the synthetic testing signatures  $\mathbf{X}$  into Hoppe et al.’s CNN<sup>14</sup>, Cohen et al.’s  
 439 FNN<sup>13</sup>, and the network of HYDRA to compare the outputs with groundtruth T1 and T2  
 440 values in  $\Theta^{T12}$ . We also compare with dictionary matching methods<sup>1,2,3,4,5,6</sup> which exploit  
 441 the same dictionary  $\mathbf{D}$  and lookup-table  $\mathbf{LUT}$  to find the best matching entry for each sig-  
 442 nature in  $\mathbf{X}$  and then estimate parameter values by searching the lookup-table. As shown  
 443 in Table 2, Table 3, Fig. 6 and Fig. 7, the estimated parameter values using the proposed  
 444 network obtained outstanding agreement with the groundtruth, yielding higher PSNR, SNR  
 445 and smaller RMSE than the dictionary matching method<sup>1,2,3,4,5,6</sup>, as well as competing net-  
 446 works<sup>13,14</sup>.

447 In particular, to illustrate in detail how well neural networks tackle the discretization  
 448 issue inherent to dictionary matching, we show the testing performance on continuous-valued  
 449 T1, T2 parameters which have small intervals, e.g. 0.5ms, that is 20 times smaller than the  
 450 training grid intervals 10ms, between neighboring values in Table 3. Since these values and  
 451 their corresponding MRF signatures do not exist in the training dictionary and lookup-table<sup>2</sup>,  
 452 the dictionary matching methods report a T1 and T2 value – the closest discretized value  
 453 present in the dictionary – that can be quite distinct from the groundtruth. In contrast, the  
 454 various neural network approaches can potentially learn an underlying mapping from the  
 455 temporal signatures to the respective T1 and T2 values, leading to estimates that are much  
 456 closer to the groundtruth. Interestingly, our approach outperforms previous networks<sup>13,14</sup>  
 457 as shown in Table 3 and Fig. 6. Evidently, neural networks demonstrate much better ro-  
 458 bustness to discretization issues, leading to improved parameter restoration in comparison  
 459 to dictionary based methods.

460 Another impressive advantage of HYDRA is the fast inference speed. HYDRA takes only  
 461 8.2 s to complete the mapping operation for eighty thousand temporal signatures, that is,  
 462 53× faster than dictionary matching. Furthermore, the inference speed of HYDRA is subject  
 463 to the network topology. That is, once the network structure is fixed, the complexity is fixed.

---

<sup>2</sup> As mentioned in the experiment setting in section 5.1, in the training dataset, T1 relaxation times are set to cover a range of [1, 5000] ms and T2 relaxation times to cover a range of [1, 2000] ms with an increment of 10 ms for both, that is, T1 values = {1, 11, 21, ..., 4991}, and T2 values = {1, 11, 21, ..., 1991}.

464 In contrast, the complexity of dictionary matching is limited by the dictionary density. This  
 465 implies that our advantage will be more prominent in comparison with competing techniques  
 466 using a dictionary with higher density.

### 467 III.C. Testing on anatomical dataset

468 In this subsection, we evaluate our approach on an anatomical testing dataset. We construct  
 469 the dataset from brain scans that were acquired with GE Signa 3T HDXT scanner from a  
 470 healthy subject.<sup>3</sup> Since there are no groundtruth parameter values for the T1 and T2 param-  
 471 eter maps, we obtain gold standard data by acquiring Fast Imaging Employing Steady-state  
 472 Acquisition (FIESTA) and Spoiled Gradient Recalled Acquisition in Steady State (SPGR)  
 473 images, at 4 different flip angles ( $3^\circ, 5^\circ, 12^\circ$  and  $20^\circ$ ), and implementing corrections<sup>35</sup> fol-  
 474 lowed by DESPOT1 and DESPOT2<sup>36</sup> algorithms. The constructed gold standard T1, T2  
 475 parameter maps have a dimension of  $128 \times 128$  for each map, accordingly leading to a pa-  
 476 rameter matrix  $\Theta^{T12}$  of size  $16384 \times 2$  by stacking vectorized T1, T2 maps together. Based  
 477 on the parameter matrix  $\Theta^{T12}$  and pre-defined RF pulse sequences, we generate temporal  
 478 signatures using Bloch equations, the same mechanism as generating the synthetic testing  
 479 dataset, leading to a signature matrix  $\mathbf{X}$  of dimension  $N \times L = 16384 \times 200$ . The signature  
 480 matrix  $\mathbf{X}$  and parameter matrix  $\Theta^{T12}$  constitute the anatomical testing dataset, with  $\mathbf{X}$  as  
 481 input and  $\Theta^{T12}$  as the gold standard reference.

482 Note that, since the gold standard T1, T2 maps exhibit spatial structures in the image  
 483 domain, the resulting signature matrix  $\mathbf{X}$  can be regarded as a stack of  $L = 200$  vectorized  
 484 image frames, where each frame exhibits specific spatial structures. Therefore, it makes  
 485 sense to perform Fourier transform and k-space subsampling for each column of  $\mathbf{X}$  to get  
 486 k-space measurements  $\mathbf{Y}$ . This is the key difference between the anatomical dataset and the  
 487 synthetic dataset.

488 We first explore the case with full k-space sampling in order to evaluate the parameter  
 489 restoration performance of HYDRA. Then, we consider situations with k-space subsampling  
 490 in order to evaluate both the signature restoration and the parameter restoration performance  
 491 of HYDRA.

---

<sup>3</sup>The experiment procedures involving human subjects described in this paper were approved by the Institutional Review Board of Tel-Aviv Sourasky Medical Center, Israel.

### 492 III.C.1. Full k-space sampling

493 In the first case, the fully-sampled k-space measurements  $\mathbf{Y}$ , derived from the Fourier trans-  
 494 form of  $\mathbf{X}$ , are used as input to obtain the estimated  $\Theta^{T12}$ . This is equivalent to inputting  $\mathbf{X}$   
 495 into the network of HYDRA, or performing dictionary matching based on  $\mathbf{X}$  directly, since  
 496 the inverse Fourier transform of the fully-sampled measurements  $\mathbf{Y}$  is exactly the same as  
 497  $\mathbf{X}$ . The aim is to test the parameter restoration performance only. In the experiment, cor-  
 498 responding parameter values in  $\Theta^{T12}$  serve as the gold standard reference. For comparison,  
 499 dictionary matching methods<sup>1,2,3,4,5,6</sup> exploit the same dictionary  $\mathbf{D}$  and lookup-table  $\mathbf{LUT}$   
 500 as in our training stage to find the best matching entry and estimate parameter values for  
 501 each signature in  $\mathbf{X}$ .

502 Visual and quantitative results are shown in Fig. 8, Fig. 9 and Table 4. It can be  
 503 seen that our basic version of HYDRA outperforms dictionary matching<sup>1,2,3,4,5,6</sup>, yielding  
 504 better visual and quantitative performance, e.g., 7.9 dB SNR gains for T2 mapping. The  
 505 RMSE of T2 mapping is also reduced to 2.498 from 6.252, accordingly. Our nonlocal version  
 506 of HYDRA achieves even better performance, leading to 10 dB SNR gains with RMSE  
 507 as small as 1.86. This is owing to the advantage that the trained network is a powerful  
 508 function approximator, which is able to output well-estimated parameter values based on  
 509 learnt mapping functions, even though these values do not exist in the training dictionary  
 510 and lookup-table. In contrast, dictionary matching only matches signatures to discrete  
 511 parameters existing in the training dataset. In other words, if there are no exact matching  
 512 dictionary element and parameter values for an inquiry MRF signature, it will find adjacent  
 513 values as approximations, thus introducing discretization error. On the other hand, the  
 514 advantage of HYDRA over dictionary matching on T1 mapping is not as significant as on  
 515 T2 mapping quantitatively. But the visual improvements are evident. A similar trend  
 516 is observed when comparing our network with competing networks such as Hoppe et al.'s  
 517 CNN<sup>14</sup> and Cohen et al.'s FNN<sup>13</sup>. In addition, HYDRA takes around 2 s to accomplish the  
 518 mapping for 16384 signatures, 40× faster than dictionary matching<sup>1,2,3,4,5,6</sup>.

### 519 III.C.2. k-space subsampling using Gaussian patterns

520 In k-space subsampling situations, the developed low-rank based de-aliasing method is ap-  
 521 plied to restore the signature matrix  $\mathbf{X}$  from the measurements matrix  $\mathbf{Y}$ . Then, the re-

522 constructed  $\mathbf{X}$  is used as input into the network for parameter mapping to obtain the cor-  
 523 responding tissue parameter values. In the experiments, the sub-sampling factor  $\beta$  is set to  
 524 be 70% and 15%. For  $\beta = 15\%$ , 15% k-space data is acquired by a series of 2D random  
 525 Gaussian sampling patterns, shown in Fig. 10, leading to a k-space measurement matrix  $\mathbf{Y}$   
 526 of size  $Q \times L = 2458 \times 200$ . Similarly,  $\beta = 70\%$  gives rise to a k-space measurement matrix  
 527  $\mathbf{Y}$  of size  $Q \times L = 11469 \times 200$ . A larger  $\lambda$  enforces lower rank for the restored signature  
 528 matrix  $\mathbf{X}$  to strengthen the de-aliasing effect, while a smaller  $\lambda$  encourages  $\mathbf{X}$  to have a sub-  
 529 sampled Fourier transform that approximates the k-space measurements matrix  $\mathbf{Y}$  better.  
 530 Therefore, we tried a range of values from 1 to 20 for  $\lambda$  and experimentally select the best  
 531 one  $\lambda = 5$ . Since the low-rank based signature restoration involves gradient descent steps, a  
 532 larger step size  $\mu$  accelerates gradient descent speed, but tends to result in oscillation or even  
 533 divergence, while a smaller  $\mu$  leads to a slower convergence. We experimentally find that  
 534  $\mu = 1$  gives a good balance. The same k-space measurements  $\mathbf{Y}$  are also used by dictionary-  
 535 matching based methods<sup>1,3,6</sup> for comparison, and the same signature restoration approach  
 536 is used to convert  $\mathbf{Y}$  onto  $\mathbf{X}$  for learning based methods<sup>13,14</sup>. The aim is to evaluate the  
 537 overall performance on both signature restoration and parameter restoration.

538 Quantitative performance is shown in Table 5. Note that the advantage of learning-  
 539 based methods over dictionary matching degrades when the subsampling factor increases.  
 540 This is due to the fact that the restored signatures from highly subsampled k-space data  
 541 exhibit deviations and distortions, thus leading to poorer input for the trained networks. In  
 542 spite of this, the proposed approach outperforms the dictionary matching based methods<sup>1,3</sup>  
 543 with significant gains, and also yields better or comparable performance as the state-of-the-  
 544 art methods FLOR<sup>6</sup>, CNN<sup>14</sup> and FNN<sup>13</sup>. In addition, it takes around 23s for low-rank based  
 545 signature restoration and less than 3s for network based parameter restoration. Thus, the  
 546 total time cost is around 26s, almost  $4.8\times$  faster than FLOR<sup>6</sup>. Furthermore, the speed of  
 547 our method is  $60\times$  faster than FLOR<sup>6</sup> for parameter restoration.

548 We compared the performance with/without nonlocal operations in our developed net-  
 549 work. The results in Table 4 and 5 show that the proposed network with nonlocal operations  
 550 based self-attention scheme outperforms the basic counterpart. In particular, the nonlocal  
 551 version achieves 6 dB gains in terms of SNR over the basic version for T2 mapping. Such  
 552 significant improvement demonstrates the benefits of capturing long-range dependencies and  
 553 global features using the nonlocal operation based attention scheme.

554 We also investigated the performance with respect to the number of time frames. In  
555 particular, we increased  $L$  from 200 to 1000 and kept other experiment settings the same as  
556 before. The quantitative results are shown in Table 6. It is noticed that given more time  
557 frames, all the methods show better performance. Moreover, the performance of learning-  
558 based methods, including CNN<sup>14</sup>, FNN<sup>13</sup> and HYDRA, improve more than model-based  
559 techniques<sup>1,3,6</sup>. In particular, our approach outperforms competing algorithms quantitatively  
560 in terms of PSNR, SNR, and RMSE, as well as demonstrates visual advantage, as shown in  
561 Figure 11 and Figure 12.

### 562 III.C.3. k-space subsampling using spiral trajectories

563 We carried out additional experiments with widely used non-Cartesian sampling patterns –  
564 variable density spiral trajectories<sup>6,37</sup>. A set of spiral trajectories used in the experiments are  
565 shown in Figure 13. They have FOV of 24 and rotation angle difference of 7.5 degrees between  
566 any two adjacent spirals to spread out the alias artifacts. Given such spiral trajectories,  
567 data were subsampled to acquire 1488 k-space samples in each time frame, leading to a  
568 subsampling ratio of 9% which is defined by the number of acquired samples in the k-space  
569 domain divided by the number of pixels in a frame. This setting closely matches the original  
570 MRF paper<sup>1</sup> where each single spiral trajectory samples 1450 k-space points (leading to a  
571 subsampling ratio around 9%) and any two adjacent spiral trajectories have a rotation angle  
572 of 7.5 degrees.

573 In the case of spiral subsampling, during the signature restoration, SParse Uniform Re-  
574 Sampling (SPURS) algorithm<sup>38</sup> was exploited to implement nonuniform Fourier transform  
575 between k-space domain and image domain, as SPURS has proved to achieve smaller approx-  
576 imation errors while maintaining low computational cost comparing with other resampling  
577 methods, such as nonuniform-FFT algorithm<sup>33</sup> and regularized Block Uniform ReSampling  
578 (rBURS)<sup>39</sup>. In the experiments, 1000 density variable spiral trajectories were used for k-space  
579 subsampling, leading to 1000 time frames. The quantitative and qualitative reconstruction  
580 results demonstrate that HYDRA outperforms competing methods with smaller estimation  
581 errors, as shown in Table 7, Figure 15 and Figure 14.

## 582 IV. Discussion

### 583 IV.A. Relation to previous works

584 Our low-rank based signature restoration method is adapted from FLOR<sup>6</sup> by removing the  
585 operation of projecting the temporal signal onto a dictionary. Thus, the signature restoration  
586 does not require a simulated dictionary, and saves computational cost. Although recent  
587 works<sup>13,14</sup> exploit neural networks to perform parameter mapping, replacing dictionaries and  
588 lookup-tables used in conventional MRF reconstruction approaches, our technique is different  
589 from these methods<sup>13,14</sup>. We design a deep nonlocal residual CNN for capturing signature-to-  
590 parameter mapping which is organically combined with low-rank based de-aliasing techniques  
591 for signature restoration. In this way, our algorithm can bypass some of the issues associated  
592 with other techniques: (1) The input dimension issue. The proposed approach can ingest  
593 temporal signatures with different lengths without the need to change the structure of the  
594 network. This is due to the fact that we rely on convolutional neural networks (CNNs)  
595 rather than fully-connected neural networks (FNNs) such as the model used in<sup>13</sup>. (2) The  
596 k-space subsampling issue. The proposed approach involves a hybrid of a neural network  
597 with a low-rank based de-aliasing approach. Thus it is able to deal with correlations both  
598 over time and space via exploiting low-rank regularization and convolution operation. This  
599 enables our work to handle k-space subsampling situations. (3) The complex mappings  
600 issue. By exploiting a residual network structure, our method can be successfully extended  
601 to deeper levels and thus obtain a better capacity to learn complex signature-to-parameter  
602 mapping functions. (4) Distortion and corruption issue. Due to the subsampling in k-space,  
603 the restored temporal signatures suffer from local distortion and corruption. Such deviation  
604 may lead to performance degeneration in the second stage. By incorporating non-local  
605 operations in the network design, our method is able to capture global features and find  
606 most relevant components for inference, thereby reducing interference of local distortion and  
607 corruption.

### 608 IV.B. Computational complexity

609 HYDRA involves two main stages: the low-rank based signature restoration stage and the  
610 network based parameter restoration stage. Even though the time cost for parameter restora-  
611 tion is longer than previous methods<sup>13,14</sup>, the time cost in the this stage is only a small  
612 fraction of the total time consumption, as the computational complexity is dominated by

613 the signature restoration stage. In other words, the computational burden of HYDRA lies  
614 in the SVD calculation in the first stage. Hence, fast SVD methods can be employed to  
615 dramatically improve the efficiency of signature restoration.

#### 616 IV.C. Model storage requirements

617 Regarding the storage requirement (in double precision), HYDRA needs only 2.1 megabytes  
618 to store the network with 0.5 million parameters, while it requires 108 megabytes to store  
619 a simulated dictionary of size  $80100 \times 200$  and 551 megabytes for size  $80100 \times 1000$ . Note  
620 that the dictionary volume will grow exponentially with the number of parameters, but the  
621 space required for storing a network is not strictly limited by the dictionary density once the  
622 topology of the network is fixed, thus significantly alleviating the storage burden inherent to  
623 the exponential growth of multi-dimensional dictionaries.

#### 624 IV.D. Impact of providing continuous T1/T2 values

625 Providing continuous T1/T2 values is an advantage of neural network based parameter  
626 mapping over dictionary matching. This property may find promising applications in some  
627 practical scenarios, for exempling, monitoring sensitive changing of pathology condition  
628 over time, such as multiple sclerosis<sup>40,41</sup>, stroke<sup>42</sup>, and treatment responses<sup>43,44</sup>, where the  
629 differences in T1 and T2 values between healthy and diseased tissues or between disease  
630 stages could be very small<sup>45</sup>. On the other hand, to fulfil this potential of network based  
631 MRF techniques, prerequisites on the accuracy and precision of MRI measurements are  
632 needed. Taking T1/T2 quantification as an example, even for the inversion recovery spin  
633 echo (IR-SE) / multiple single-echo spin echo MRI sequences which are considered as the  
634 gold standard for T1/T2 quantification, there exist variations of 2% - 9% on the measured  
635 relaxation times<sup>45</sup>. Such anatomical measurement uncertainties and model imperfections  
636 may weaken the advantage and clinical impact of providing continuous T1/T2 values using  
637 network based MRF techniques to some extent. Therefore, improving the accuracy of gold  
638 standard approaches in the future would contribute to making the most of the potential of  
639 neural networks in the MRF domain.

## V. Conclusion

We proposed a hybrid deep MRF approach which combines low-rank based signature restoration with learning-based parameter restoration. In our approach, a low-rank based de-aliasing method is used to restore clean signatures from subsampled k-space measurements. Then, a 1D deep nonlocal residual CNN is developed for efficient signature-to-parameter mapping, replacing the time-consuming dictionary matching operation in conventional MRF techniques. Our approach demonstrates advantages in terms of inference speed, accuracy and storage requirements over competing MRF methods as no dictionary is needed for recovery.

## References

- <sup>1</sup> D. Ma, V. Gulani, N. Seiberlich, K. Liu, J. L. Sunshine, J. L. Duerk, and M. A. Griswold, Magnetic resonance fingerprinting, *Nature* **495**, 187 (2013).
- <sup>2</sup> Y. Jiang, D. Ma, N. Seiberlich, V. Gulani, and M. A. Griswold, MR fingerprinting using fast imaging with steady state precession (FISP) with spiral readout, *Magnetic resonance in medicine* **74**, 1621–1631 (2015).
- <sup>3</sup> M. Davies, G. Puy, P. Vandergheynst, and Y. Wiaux, A compressed sensing framework for magnetic resonance fingerprinting, *SIAM Journal on Imaging Sciences* **7**, 2623–2656 (2014).
- <sup>4</sup> Z. Wang, H. Li, Q. Zhang, J. Yuan, and X. Wang, Magnetic Resonance Fingerprinting with compressed sensing and distance metric learning, *Neurocomputing* **174**, 560–570 (2016).
- <sup>5</sup> G. Mazor, L. Weizman, A. Tal, and Y. C. Eldar, Low rank magnetic resonance fingerprinting, in *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the*, pages 439–442, IEEE, 2016.
- <sup>6</sup> G. Mazor, L. Weizman, A. Tal, and Y. C. Eldar, Low-rank magnetic resonance fingerprinting, *Medical physics* **45**, 4066–4084 (2018).
- <sup>7</sup> C. Liao, B. Bilgic, M. K. Manhard, B. Zhao, X. Cao, J. Zhong, L. L. Wald, and K. Set-

- 666 sompop, 3D MR fingerprinting with accelerated stack-of-spirals and hybrid sliding-  
667 window and GRAPPA reconstruction, *Neuroimage* **162**, 13–22 (2017).
- 668 <sup>8</sup> X. Cao, C. Liao, Z. Wang, Y. Chen, H. Ye, H. He, and J. Zhong, Robust sliding-window  
669 reconstruction for Accelerating the acquisition of MR fingerprinting, *Magnetic resonance*  
670 *in medicine* **78**, 1579–1588 (2017).
- 671 <sup>9</sup> M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, Compressed sensing MRI,  
672 *IEEE signal processing magazine* **25**, 72–82 (2008).
- 673 <sup>10</sup> L. Weizman, Y. C. Eldar, and D. Ben Bashat, Compressed sensing for longitudinal MRI:  
674 An adaptive-weighted approach, *Medical physics* **42**, 5195–5208 (2015).
- 675 <sup>11</sup> L. Weizman, Y. C. Eldar, and D. Ben Bashat, Reference-based MRI, *Medical physics*  
676 **43**, 5357–5369 (2016).
- 677 <sup>12</sup> Y. C. Eldar, *Sampling Theory: Beyond Bandlimited Systems*, Cambridge University  
678 Press, 2015.
- 679 <sup>13</sup> O. Cohen, B. Zhu, and M. S. Rosen, MR fingerprinting deep reconstruction network  
680 (DRONE), *Magnetic resonance in medicine* **80**, 885–894 (2018).
- 681 <sup>14</sup> E. Hoppe, G. Körzdörfer, T. Würfl, J. Wetzl, F. Lugauer, J. Pfeuffer, and A. Maier,  
682 Deep Learning for Magnetic Resonance Fingerprinting: A New Approach for Predicting  
683 Quantitative Parameter Values from Time Series, *Stud Health Technol Inform* **243**,  
684 202–206 (2017).
- 685 <sup>15</sup> Y. LeCun, Y. Bengio, and G. Hinton, Deep learning, *Nature* **521**, 436 (2015).
- 686 <sup>16</sup> I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep learning*, volume 1, MIT  
687 press Cambridge, 2016.
- 688 <sup>17</sup> A. Krizhevsky, I. Sutskever, and G. E. Hinton, Imagenet classification with deep convo-  
689 lutional neural networks, in *Advances in neural information processing systems*, pages  
690 1097–1105, 2012.
- 691 <sup>18</sup> K. He, X. Zhang, S. Ren, and J. Sun, Deep residual learning for image recognition, in  
692 *Proc. IEEE Conf. Comput. Vision Pattern Recog*, pages 770–778, 2016.

- 693 19 K. He, X. Zhang, S. Ren, and J. Sun, Identity mappings in deep residual networks, in  
694 *Proc. Eur. Conf. Comput. Vision*, pages 630–645, Springer, 2016.
- 695 20 C. Dong, C. C. Loy, K. He, and X. Tang, Image super-resolution using deep convolutional  
696 networks, *IEEE Trans. Pattern Anal. Mach. Intell.* **38**, 295–307 (2016).
- 697 21 J. Kim, J. Kwon Lee, and K. Mu Lee, Deeply-recursive convolutional network for image  
698 super-resolution, in *Proc. IEEE Conf. Comput. Vision Pattern Recog*, pages 1637–1645,  
699 2016.
- 700 22 J. Gehring, M. Auli, D. Grangier, D. Yarats, and Y. N. Dauphin, Convolutional sequence  
701 to sequence learning, arXiv preprint arXiv:1705.03122 (2017).
- 702 23 G. Hinton et al., Deep neural networks for acoustic modeling in speech recognition:  
703 The shared views of four research groups, *IEEE Signal processing magazine* **29**, 82–97  
704 (2012).
- 705 24 J.-F. Cai, E. J. Candès, and Z. Shen, A singular value thresholding algorithm for matrix  
706 completion, *SIAM Journal on Optimization* **20**, 1956–1982 (2010).
- 707 25 S. Sra, S. Nowozin, and S. J. Wright, *Optimization for machine learning*, Mit Press,  
708 2012.
- 709 26 A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and  
710 I. Polosukhin, Attention is all you need, in *Advances in neural information processing*  
711 *systems*, pages 5998–6008, 2017.
- 712 27 X. Wang, R. Girshick, A. Gupta, and K. He, Non-local neural networks, in *Proceedings*  
713 *of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 7794–7803,  
714 2018.
- 715 28 K. He, X. Zhang, S. Ren, and J. Sun, Delving deep into rectifiers: Surpassing human-  
716 level performance on imagenet classification, in *Proc. IEEE Int. Conf. Comput. Vision*,  
717 pages 1026–1034, 2015.
- 718 29 N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, Dropout:  
719 a simple way to prevent neural networks from overfitting, *The Journal of Machine*  
720 *Learning Research* **15**, 1929–1958 (2014).

- 721 <sup>30</sup> J. Hennig, Echoes – how to generate, recognize, use or avoid them in MR-imaging se-  
722 quences. Part I: Fundamental and not so fundamental properties of spin echoes, Concepts  
723 in Magnetic Resonance **3**, 125–143 (1991).
- 724 <sup>31</sup> M. Weigel, Extended phase graphs: dephasing, RF pulses, and echoes-pure and simple,  
725 Journal of Magnetic Resonance Imaging **41**, 266–295 (2015).
- 726 <sup>32</sup> D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, arXiv preprint  
727 arXiv:1412.6980 (2014).
- 728 <sup>33</sup> J. A. Fessler and B. P. Sutton, Nonuniform fast Fourier transforms using min-max  
729 interpolation, IEEE Trans. Sig. Proc. **51**, 560–574 (2003).
- 730 <sup>34</sup> J. Vymazal, A. Righini, R. A. Brooks, M. Canesi, C. Mariani, M. Leonardi, and G. Pez-  
731 zoli, T1 and T2 in the brain of healthy subjects, patients with Parkinson disease, and  
732 patients with multiple system atrophy: relation to iron content, Radiology **211**, 489–495  
733 (1999).
- 734 <sup>35</sup> G. Liberman, Y. Louzoun, and D. Ben Bashat, T1 mapping using variable flip angle  
735 SPGR data with flip angle correction, Journal of Magnetic Resonance Imaging **40**,  
736 171–180 (2014).
- 737 <sup>36</sup> S. C. Deoni, T. M. Peters, and B. K. Rutt, High-resolution T1 and T2 mapping of  
738 the brain in a clinically acceptable time with DESPOT1 and DESPOT2, Magnetic  
739 Resonance in Medicine: An Official Journal of the International Society for Magnetic  
740 Resonance in Medicine **53**, 237–241 (2005).
- 741 <sup>37</sup> J. H. Lee, B. A. Hargreaves, B. S. Hu, and D. G. Nishimura, Fast 3D imaging us-  
742 ing variable-density spiral trajectories with applications to limb perfusion, Magnetic  
743 Resonance in Medicine: An Official Journal of the International Society for Magnetic  
744 Resonance in Medicine **50**, 1276–1285 (2003).
- 745 <sup>38</sup> A. Kiperwas, D. Rosenfeld, and Y. C. Eldar, The SPURS algorithm for resampling an  
746 irregularly sampled signal onto a cartesian grid, IEEE transactions on medical imaging  
747 **36**, 628–640 (2017).

- 748 <sup>39</sup> D. Rosenfeld, New approach to gridding using regularization and estimation theory,  
749 *Magnetic Resonance in Medicine: An Official Journal of the International Society for*  
750 *Magnetic Resonance in Medicine* **48**, 193–202 (2002).
- 751 <sup>40</sup> F. Manfredonia, O. Ciccarelli, Z. Khaleeli, D. J. Tozer, J. Sastre-Garriga, D. H. Miller,  
752 and A. J. Thompson, Normal-appearing brain t1 relaxation time predicts disability in  
753 early primary progressive multiple sclerosis, *Archives of neurology* **64**, 411–415 (2007).
- 754 <sup>41</sup> K. Papadopoulos, D. J. Tozer, L. Fisniku, D. R. Altmann, G. Davies, W. Rashid, A. J.  
755 Thompson, D. H. Miller, and D. T. Chard, T1-relaxation time changes over five years  
756 in relapsing-remitting multiple sclerosis, *Multiple Sclerosis Journal* **16**, 427–433 (2010).
- 757 <sup>42</sup> J. Bernarding, J. Braun, J. Hohmann, U. Mansmann, M. Hoehn-Berlage, C. Stapf, K.-  
758 J. Wolf, and T. Tolxdorff, Histogram-based characterization of healthy and ischemic  
759 brain tissues using multiparametric MR imaging including apparent diffusion coefficient  
760 maps and relaxometry, *Magnetic Resonance in Medicine: An Official Journal of the*  
761 *International Society for Magnetic Resonance in Medicine* **43**, 52–61 (2000).
- 762 <sup>43</sup> P. M. McSheehy, C. Weidensteiner, C. Cannet, S. Ferretti, D. Laurent, S. Ruetz,  
763 M. Stumm, and P. R. Allegrini, Quantified tumor T1 is a generic early-response imag-  
764 ing biomarker for chemotherapy reflecting cell viability, *Clinical Cancer Research* **16**,  
765 212–225 (2010).
- 766 <sup>44</sup> C. Weidensteiner, P. R. Allegrini, M. Sticker-Jantscheff, V. Romanet, S. Ferretti, and  
767 P. M. McSheehy, Tumour T 1 changes in vivo are highly predictive of response to  
768 chemotherapy and reflect the number of viable tumour cells—a preclinical MR study in  
769 mice, *BMC cancer* **14**, 88 (2014).
- 770 <sup>45</sup> Y. Jiang, D. Ma, K. E. Keenan, K. F. Stupic, V. Gulani, and M. A. Griswold, Re-  
771 peatability of magnetic resonance fingerprinting T1 and T2 estimates assessed using the  
772 ISMRM/NIST MRI system phantom, *Magnetic resonance in medicine* **78**, 1452–1457  
773 (2017).

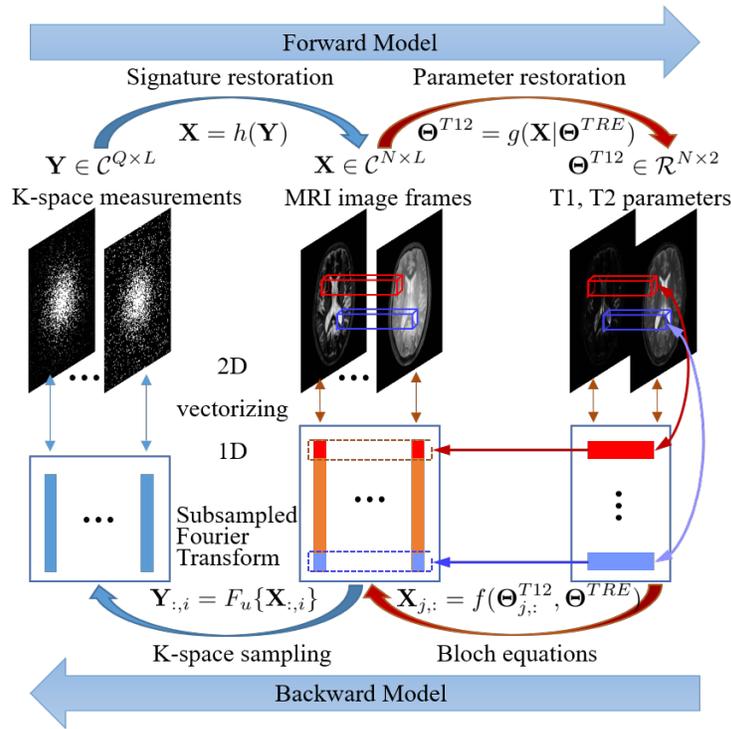


Figure 1: Relationship between key variables. The MR contrast matrix  $\mathbf{X}$  is associated with the k-space measurements  $\mathbf{Y}$  per column by the subsampled Fourier transform. It is related to tissue parameters  $\Theta^{T12}$  per row by the Bloch equations. Given  $\Theta^{T12}$  and  $\mathbf{Y}$ , the image stack  $\mathbf{X}$  is commonly first reconstructed from  $\mathbf{Y}$ , referred to as signature restoration, and then mapped to tissue parameters  $\Theta^{T12}$  via dictionary matching, referred to as parameter restoration.

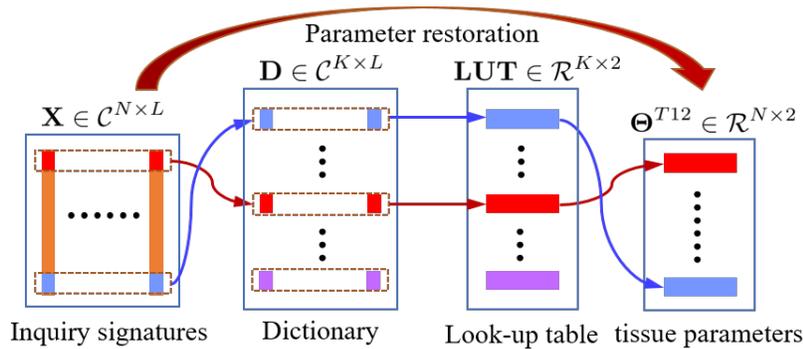


Figure 2: Parameter restoration using dictionary matching. Given an inquiry temporal signature, dictionary matching computes its inner product with each dictionary entry, and selects the most correlated one with the highest inner product as the best matching signature. Once the best matching entry is found, it directly leads to multiple tissue parameters, such as T1, T2, simultaneously, via searching a lookup-table.

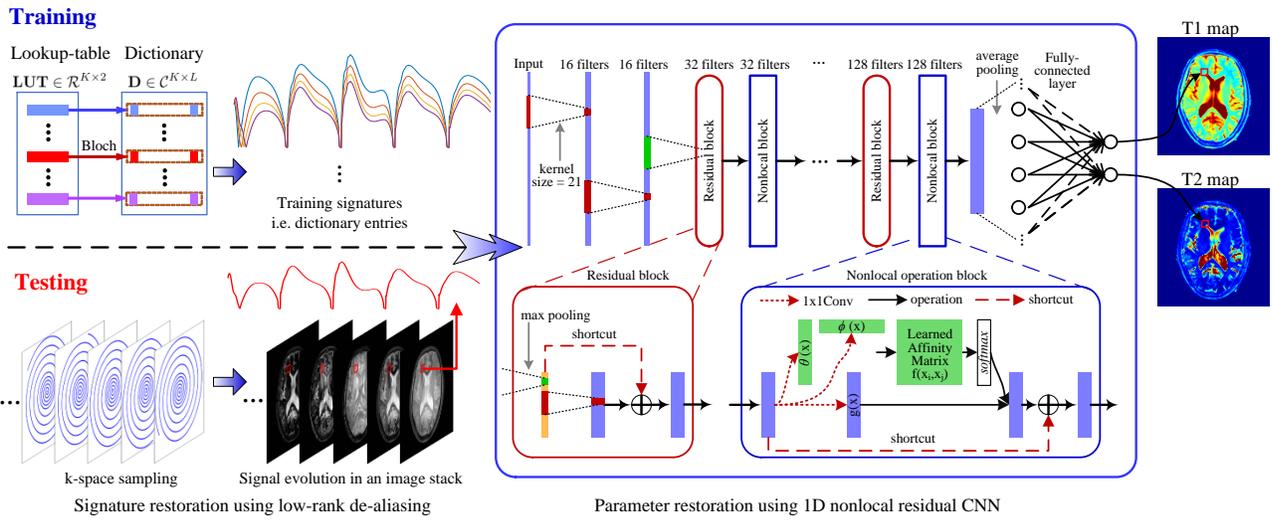


Figure 3: Diagram of the proposed MRF reconstruction approach. During the training stage, synthesized dictionary entries are used as training signatures to train the designed 1D nonlocal residual CNN until the outputs approximate parameter values in **LUT** well. In this way, the network captures the signature-to-parameter mapping. During the testing stage, a low-rank based algorithm is used to restore the image stack, a matrix containing signatures in rows, from k-space measurements. Then the restored signatures are input into the trained network to obtain corresponding tissue parameters directly.

Table 1: Brief description of experiment types and settings.

Experiment	Settings
Training	Input: <b>D</b> , size $K \times L = 80100 \times 200$ . Groundtruth: <b>LUT</b> , size $80100 \times 2$ . k-space subsampling factor $\beta$ : not available.
Testing on synthetic data	Input: <b>X</b> , size $N \times L = 80000 \times 200$ . Groundtruth: $\Theta^{T12}$ , size $80000 \times 2$ . k-space subsampling factor $\beta$ : not available.
Testing on anatomical data	Input: <b>Y</b> , size $Q \times L = 16384\beta \times 200$ or $16384\beta \times 1000$ . Reference: $\Theta^{T12}$ , size $N \times 2 = 16384 \times 2$ . k-space subsampling factor $\beta$ : 70%, 15% using Gaussian patterns.
Testing on anatomical data	Input: <b>Y</b> , size $Q \times L = 16384\beta \times 1000$ . Reference: $\Theta^{T12}$ , size $N \times 2 = 16384 \times 2$ . k-space subsampling factor $\beta$ : 9% using spiral trajectories.

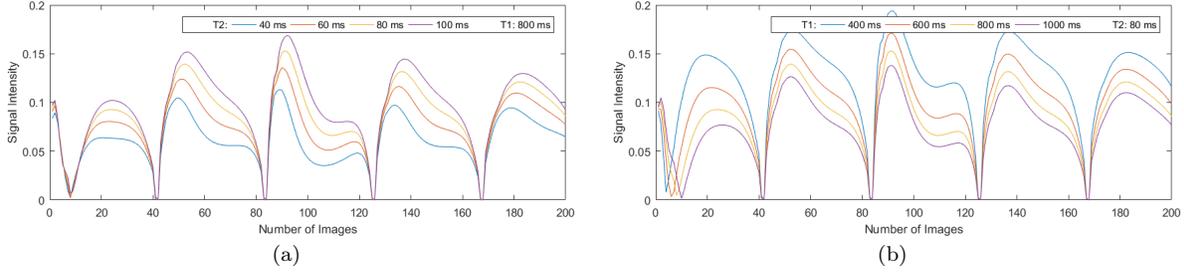


Figure 4: Synthetic MRF temporal signatures with 200 time frames. (a) Temporal signatures corresponding to parameter values  $\{(T1, T2)\}$  ms =  $\{(800,40),(800,60),(800,80),(800,100)\}$  ms. (2) Temporal signatures corresponding to parameter values  $\{(T1, T2)\}$  ms =  $\{(400,80),(600,80),(800,80),(1000,80)\}$  ms.

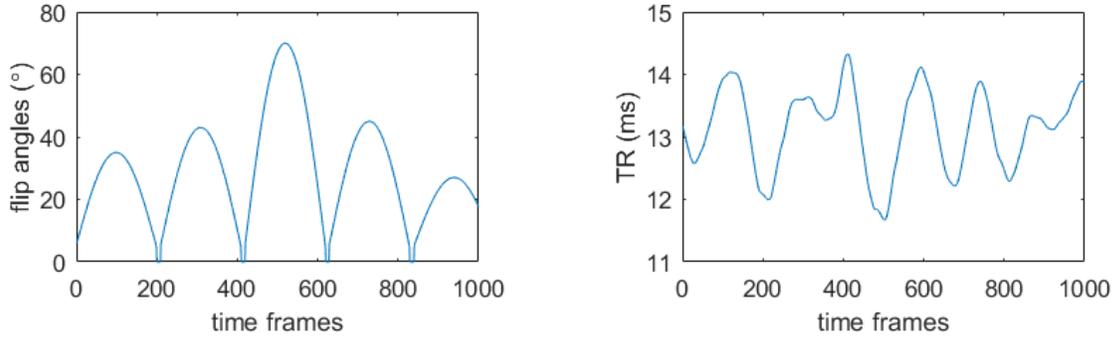


Figure 5: FISP pulse sequence parameters. All the flip angles (FA) constituted a sinusoidal variation in the range of 0 - 70 degrees to ensure smoothly varying transient state of the magnetization. The repetition time (TR) was randomly varied in the range of 11.5 - 14.5 ms with a Perlin noise pattern.

Table 2: Testing on synthetic dataset. Comparing parameter restoration performance, in terms of PSNR, SNR, RMSE and correlation coefficient.

	Dict. Match. T1 / T2	CNN <sup>14</sup> T1 / T2	FNN <sup>13</sup> T1 / T2	Proposed T1 / T2
PSNR (dB)	59.15 / 52.31	62.96 / 49.64	58.97 / 54.96	<b>79.30 / 72.99</b>
SNR (dB)	55.23 / 47.15	59.05 / 44.49	55.06 / 49.81	<b>75.38 / 67.83</b>
RMSE (ms)	5.515 / 4.847	3.554 / 6.591	5.63 / 3.57	<b>0.542 / 0.448</b>
CorrCoef	1.00 / 1.00	1.00 / 1.00	1.00 / 1.00	1.00 / 1.00
time cost (s)	464.10	2.87	<b>1.58</b>	8.2

Table 3: Testing on synthetic dataset involving detailed T1 / T2 examples that are not on the training grid and their intervals are much smaller than the training grid intervals. D.M. denotes dictionary matching. T1 and T2 errors are defined as the difference between estimated values and groundtruth values.

Truth	T1 Estimation				T1 Errors			
	D.M.	<sup>14</sup>	<sup>13</sup>	Ours	D.M.	<sup>14</sup>	<sup>13</sup>	Ours
1005.0	1001.0	1002.9	1009.3	1004.8	-4.0	-2.1	4.3	<b>-0.2</b>
1005.5	1001.0	1003.3	1010.0	1005.3	-4.5	-2.3	4.5	<b>-0.2</b>
1006.0	1011.0	1003.6	1010.6	1005.8	5.0	-2.4	4.6	<b>-0.2</b>
1006.5	1011.0	1004.1	1011.2	1006.3	4.5	-2.5	4.7	<b>-0.2</b>
1007.0	1011.0	1004.5	1011.8	1006.8	4.0	-2.5	4.8	<b>-0.3</b>
RMSE	-	-	-	-	4.4	2.3	4.6	<b>0.2</b>

Truth	T2 Estimation				T2 Errors			
	D.M.	<sup>14</sup>	<sup>13</sup>	Ours	D.M.	<sup>14</sup>	<sup>13</sup>	Ours
505.0	501.0	513.6	504.3	505.2	-4.0	8.6	-0.7	<b>0.2</b>
505.5	511.0	514.1	504.8	505.7	5.5	8.6	-0.7	<b>0.2</b>
506.0	501.0	514.7	505.3	506.2	-5.0	8.6	-0.7	<b>0.2</b>
506.5	511.0	515.2	505.9	506.8	4.5	8.7	-0.6	<b>0.3</b>
507.0	511.0	515.7	506.4	507.3	4.0	8.7	-0.6	<b>0.3</b>
RMSE	-	-	-	-	4.6	8.6	0.7	<b>0.2</b>

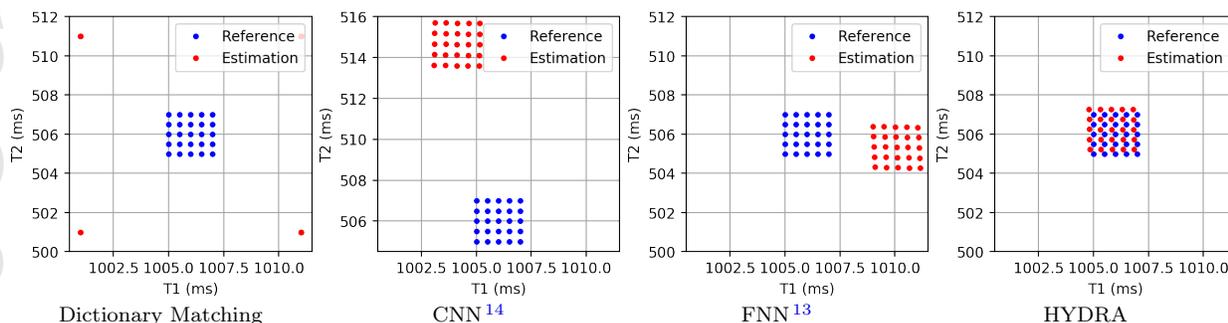


Figure 6: Testing on synthetic dataset involving detailed T1 / T2 examples that are not on the training grid and their intervals are much smaller than the training grid intervals. Dictionary matching finds best adjacent values from the dictionary, i.e. 1001, 1011 for T1, and 501, 511 for T2. In contrast, owing to the captured mapping functions, neural networks output continuous values. Proposed HYDRA leads to the smallest deviations and bias.

Table 4: Testing on anatomical dataset with full k-space sampling. Comparing parameter restoration performance, in terms of PSNR, SNR, RMSE and correlation coefficient.

	Dict. Match.	CNN <sup>14</sup>	FNN <sup>13</sup>	Proposed basic	Proposed nonlocal
	T1 / T2	T1 / T2	T1 / T2	T1 / T2	T1 / T2
PSNR (dB)	56.64 / 52.04	54.06 / 49.88	54.53 / 54.36	56.59 / 60.01	56.47 / 62.56
SNR (dB)	42.20 / 27.81	39.63 / 25.66	40.09 / 30.07	42.15 / 35.76	42.03 / 38.32
RMSE (ms)	6.623 / 6.252	8.912 / 8.015	8.45 / 4.78	6.661 / 2.498	6.76 / 1.86
CorrCoef	1.00 / 1.00	1.00 / 1.00	1.00 / 1.00	1.00 / 1.00	1.00 / 1.00
time cost (s)	84.56	0.69	0.41	1.6	2.1

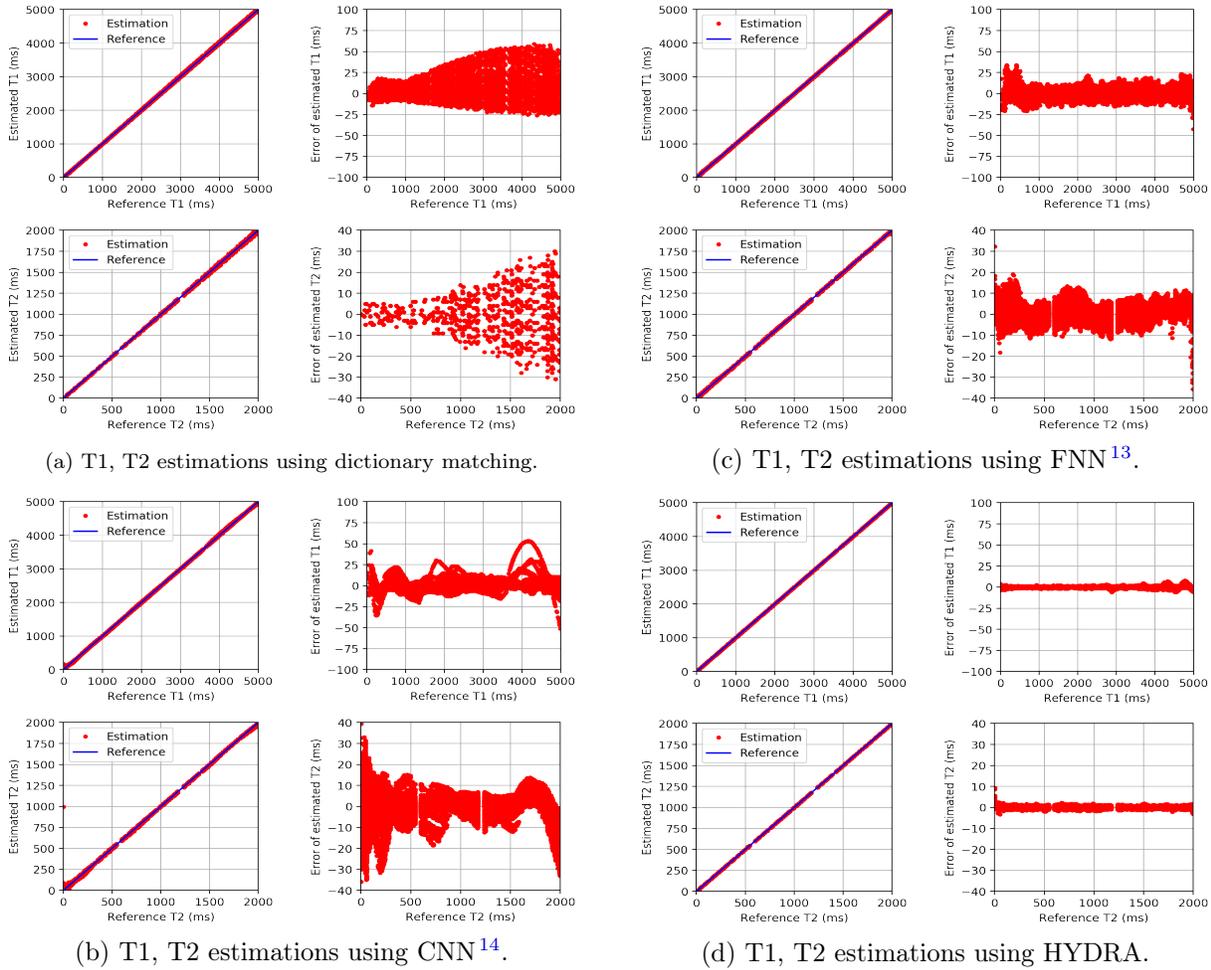


Figure 7: Testing on the synthetic dataset for comparing parameter restoration performance. Subfig. (a) - (d) show the results using dictionary matching<sup>1,2,3,4,5,6</sup>, FNN<sup>13</sup>, CNN<sup>14</sup> and HYDRA. In each subfigure, the left figure compares the estimated T1 or T2 values (marked with red dot) with groundtruth values (marked with blue line), and the right figure shows the deviations of the estimation from the groundtruth. Parameter mapping performance of HYDRA is much better than competing methods, in the entire value range of T1 and T2 parameters, resulting in smaller deviations.

Table 5: Testing on anatomical dataset with k-space subsampling ratio 70% and 15% using Gaussian patterns and 200 time frames.

	k-space subsampling factor $\beta = 70\%$						Proposed	
	Ma et al. <sup>1</sup>	BLIP <sup>3</sup>	FLOR <sup>6</sup>	CNN <sup>14</sup>	FNN <sup>13</sup>	basic	nonlocal	
PSNR (dB)	23.69 / 38.17	45.67 / 47.84	50.11 / 50.85	49.71 / 45.48	50.15 / 51.08	50.79 / 51.59	49.87 / 57.57	
SNR (dB)	8.73 / 13.84	31.28 / 23.49	35.67 / 26.48	35.26 / 21.19	35.70 / 26.67	36.34 / 27.19	35.42 / 33.30	
RMSE (ms)	294.32 / 30.87	23.42 / 10.14	14.01 / 7.17	14.71 / 13.31	13.99 / 6.98	12.99 / 6.57	14.44 / 3.31	
time cost (s)	72.88	75.70	85.35	23.72	23.53	24.85	26.3	
	k-space subsampling factor $\beta = 15\%$						Proposed	
	Ma et al. <sup>1</sup>	BLIP <sup>3</sup>	FLOR <sup>6</sup>	CNN <sup>14</sup>	FNN <sup>13</sup>	basic	nonlocal	
PSNR (dB)	27.94 / 32.84	35.45 / 39.25	44.95 / 46.11	43.74 / 35.98	45.03 / 45.90	45.23 / 44.44	45.39 / 51.32	
SNR (dB)	13.50 / 8.61	20.99 / 14.58	30.51 / 21.89	29.23 / 12.26	30.58 / 21.32	30.76 / 19.78	30.91 / 26.99	
RMSE (ms)	180.3 / 57.03	76.01 / 27.25	25.46 / 12.37	29.27 / 39.73	25.21 / 12.68	24.65 / 15.00	24.20 / 6.79	
time cost (s)	106	112.8	121.7	24.54	24.36	25.67	27.31	

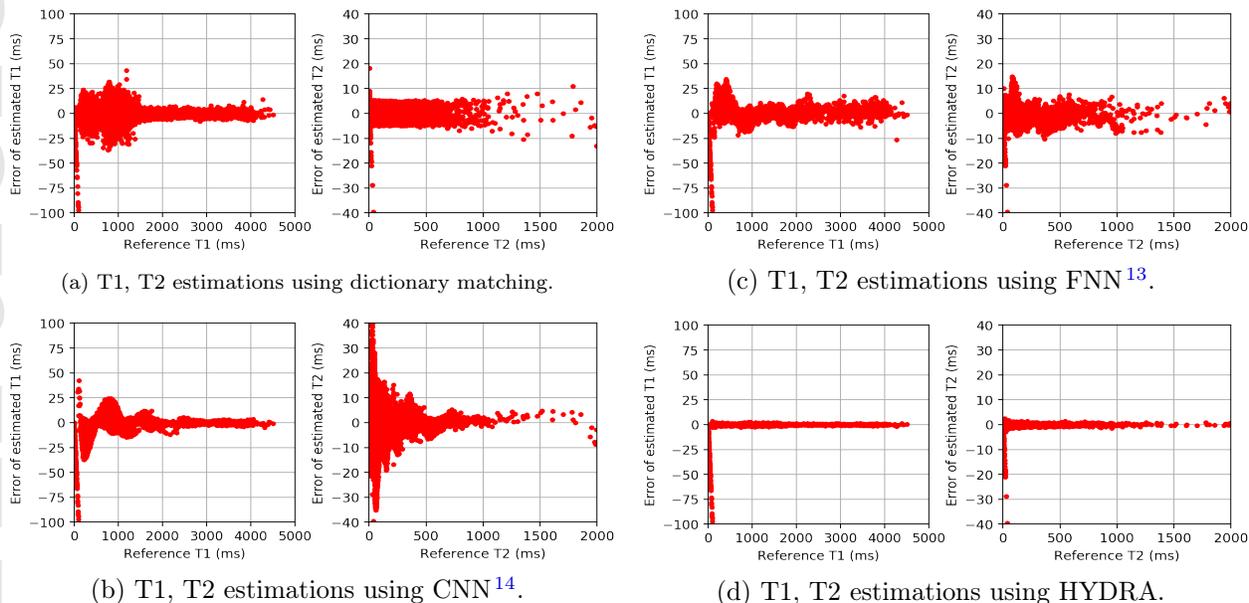


Figure 8: Testing on the anatomical dataset with full k-space sampling for comparing parameter restoration performance. Subfig. (a) - (d) show the results using dictionary matching<sup>1,2,3,4,5,6</sup>, FNN<sup>13</sup>, CNN<sup>14</sup> and HYDRA. Each subfigure shows the deviations of the estimation from the reference. Parameter mapping performance using HYDRA outperforms competing methods significantly, resulting in smaller deviations. The performance is also verified by quantitative metrics, as shown in Table 4.

Table 6: Testing on anatomical dataset with k-space subsampling ratio 15% using Gaussian patterns and 1000 time frames.

	Ma et al. <sup>1</sup>	BLIP <sup>3</sup>	FLOR <sup>6</sup>	CNN <sup>14</sup>	FNN <sup>13</sup>	Proposed
PSNR (dB)	27.53 / 33.28	35.50 / 39.10	50.90 / 50.04	41.96 / 39.21	52.62 / 49.86	52.32 / 52.79
SNR (dB)	13.09 / 9.05	21.06 / 14.87	36.44 / 25.65	27.44 / 15.05	38.17 / 25.43	37.86 / 28.35
RMSE (ms)	189.09 / 54.21	75.53 / 27.74	12.83 / 7.87	35.91 / 27.37	10.52 / 8.04	10.89 / 5.74

Table 7: Testing on anatomical dataset with k-space subsampling ratio 9% using spiral trajectories and 1000 time frames.

	Ma et al. <sup>1</sup>	BLIP <sup>3</sup>	FLOR <sup>6</sup>	CNN <sup>14</sup>	FNN <sup>13</sup>	Proposed
PSNR (dB)	26.66 / 30.44	29.35 / 39.47	39.32 / 44.60	35.68 / 27.74	40.26 / 44.70	41.45 / 45.41
SNR (dB)	12.22 / 6.21	15.03 / 15.22	24.88 / 20.38	21.45 / 4.48	25.84 / 20.30	27.02 / 21.04
RMSE (ms)	209.01 / 75.18	153.37 / 26.57	48.67 / 14.72	73.96 / 102.57	43.65 / 14.55	38.08 / 13.41

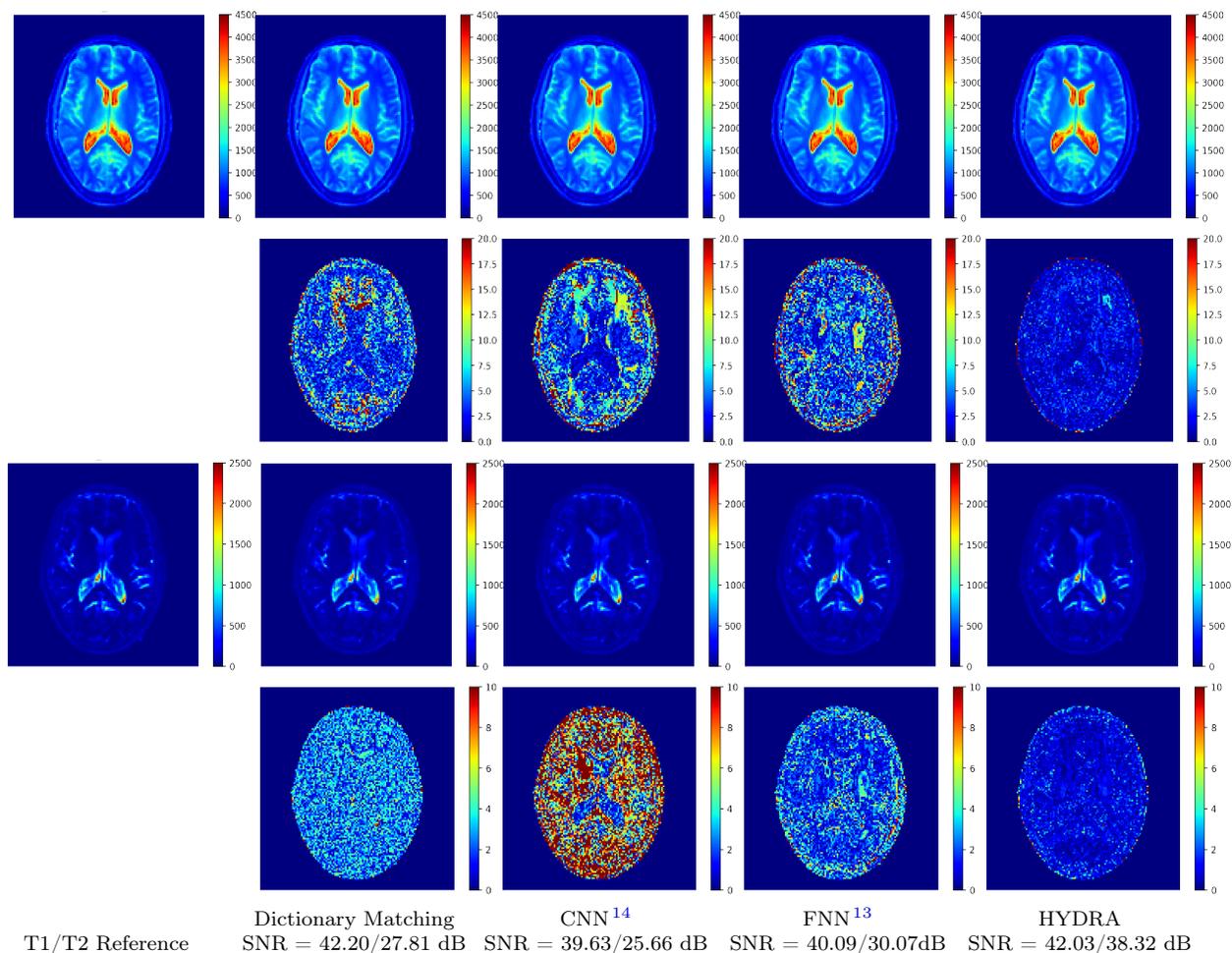


Figure 9: Visual results of testing on anatomical dataset with full k-space sampling for comparing parameter restoration performance. Top two rows correspond to T1 maps and residual errors while bottom two rows correspond to T2 maps and residual errors. Proposed HYDRA results in comparable performance for T1 mapping and yields much better performance for T2 mapping, obtaining 10dB higher SNR gains than competing dictionary-matching based methods<sup>1,2,3,4,5,6</sup>. HYDRA also outperforms previous networks, such as CNN by Hoppe et al.<sup>14</sup> and FNN by Cohen et al.<sup>13</sup>.



Figure 10: A series of Gaussian patterns used for k-space subsampling.

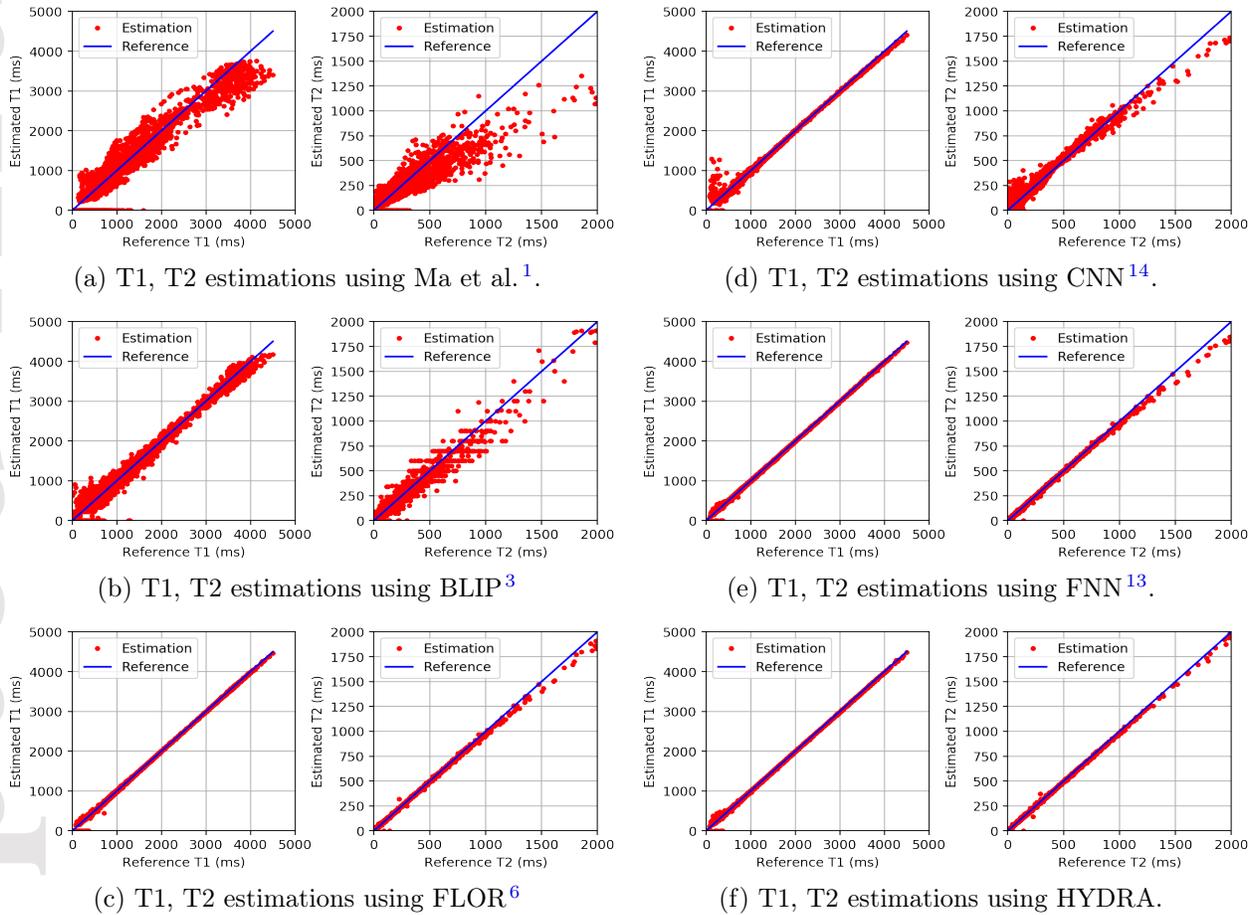
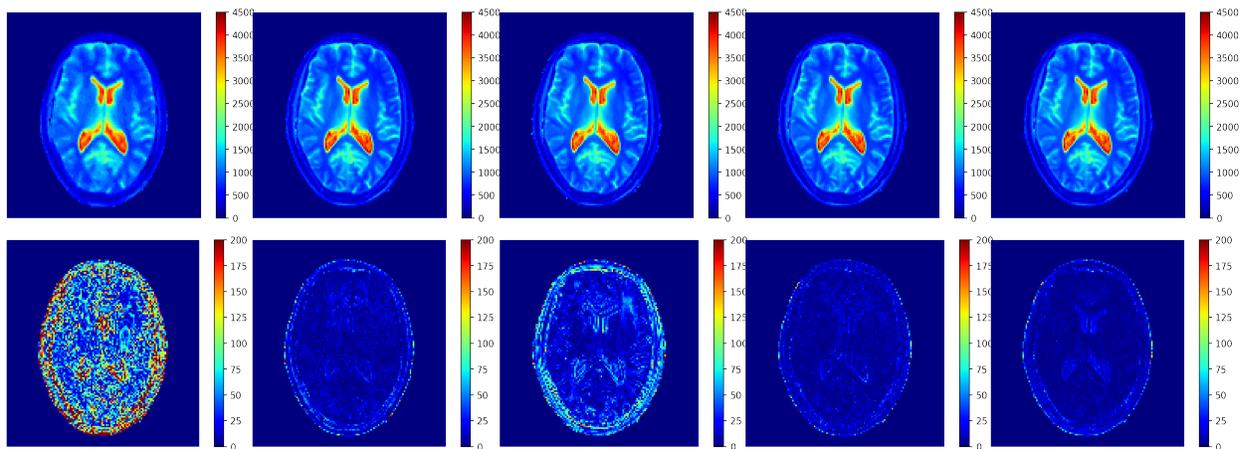
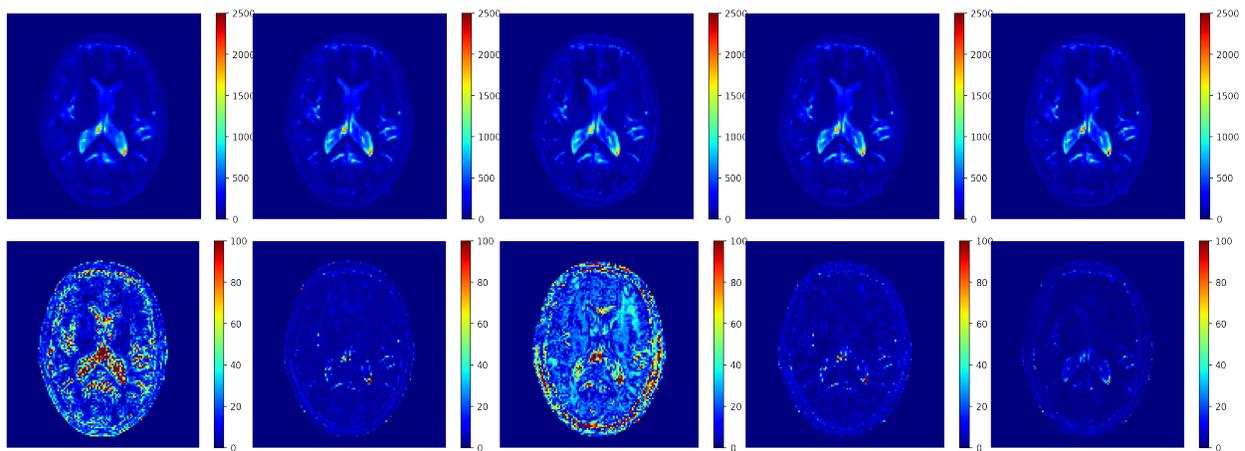


Figure 11: Testing on the anatomical dataset with k-space subsampling factor 15% using Gaussian patterns and 1000 time frames. Subfig. (a) - (f) show the results using Ma et al.<sup>1</sup>, BLIP<sup>3</sup>, FLOR<sup>6</sup>, CNN by Hoppe et al.<sup>14</sup>, FNN by Cohen et al.<sup>13</sup> and HYDRA.



(a) T1 estimation (top row) and residual errors (bottom row).



(b) T2 estimation (top row) and residual errors (bottom row).

BLIP<sup>3</sup> SNR = 21.06/14.87 dB   
 FLOR<sup>6</sup> SNR = 36.44/25.65 dB   
 CNN<sup>14</sup> SNR = 27.44/15.05 dB   
 FNN<sup>13</sup> SNR = 38.17/25.43 dB   
 HYDRA SNR = 37.86/28.35 dB

Figure 12: Visual results of testing on anatomical dataset with k-space subsampling factor 15% using Gaussian pattern with  $L = 1000$ . Comparison between BLIP<sup>3</sup>, FLOR<sup>6</sup>, CNN by Hoppe et al.<sup>14</sup>, FNN by Cohen et al.<sup>13</sup> and HYDRA.

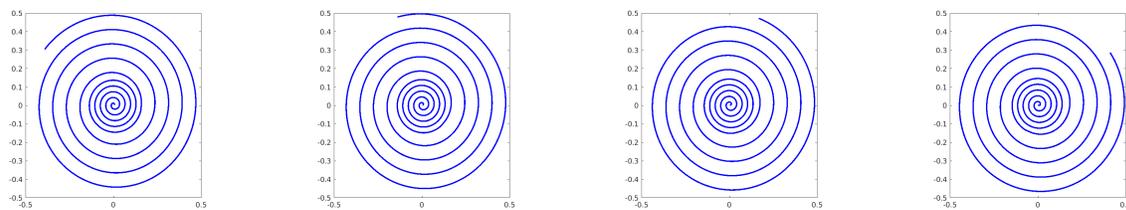


Figure 13: A series of spiral trajectories for k-space subsampling.

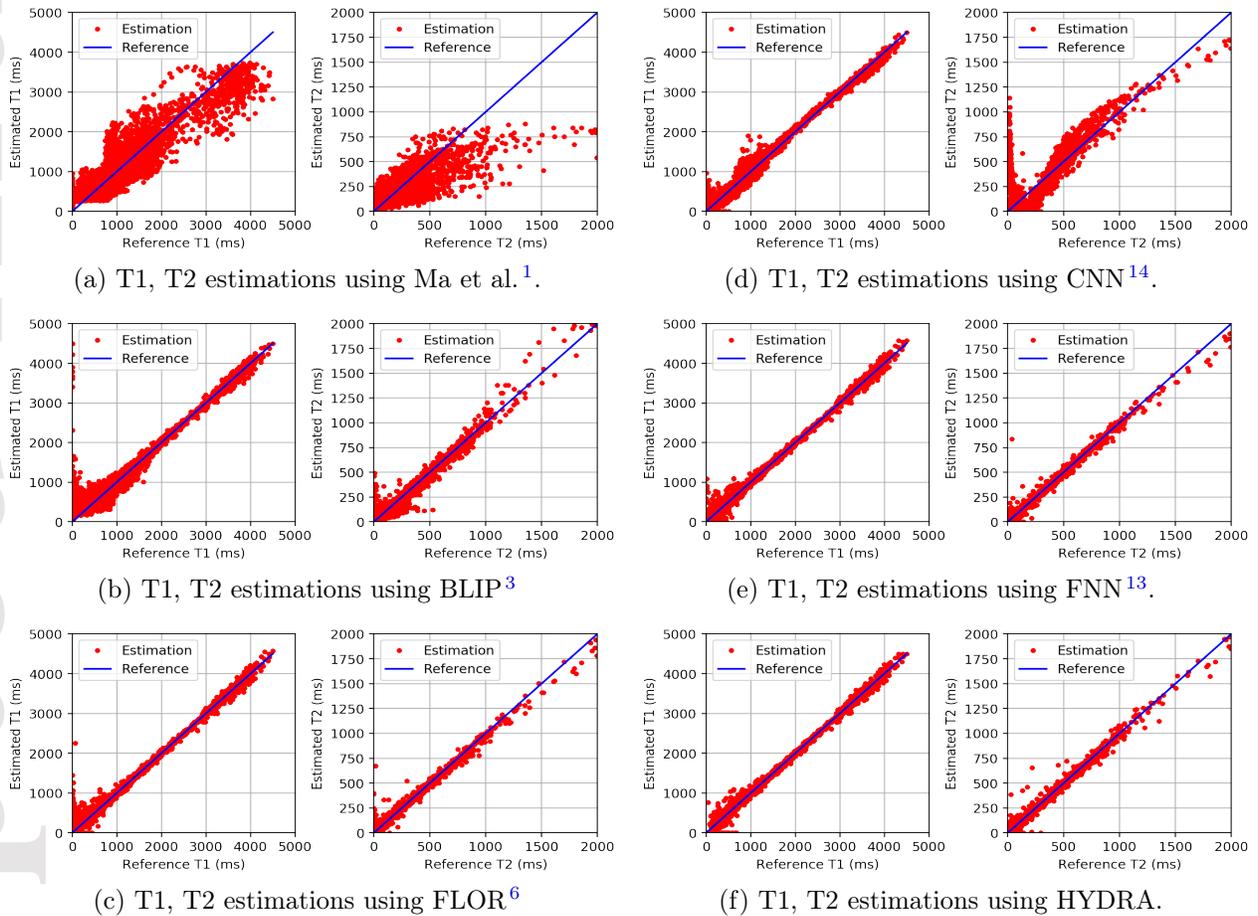
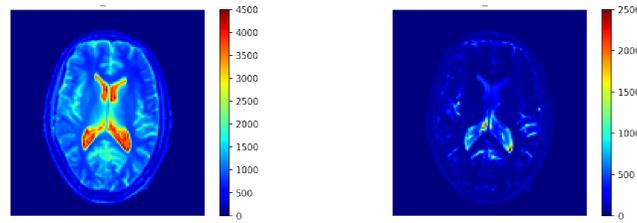
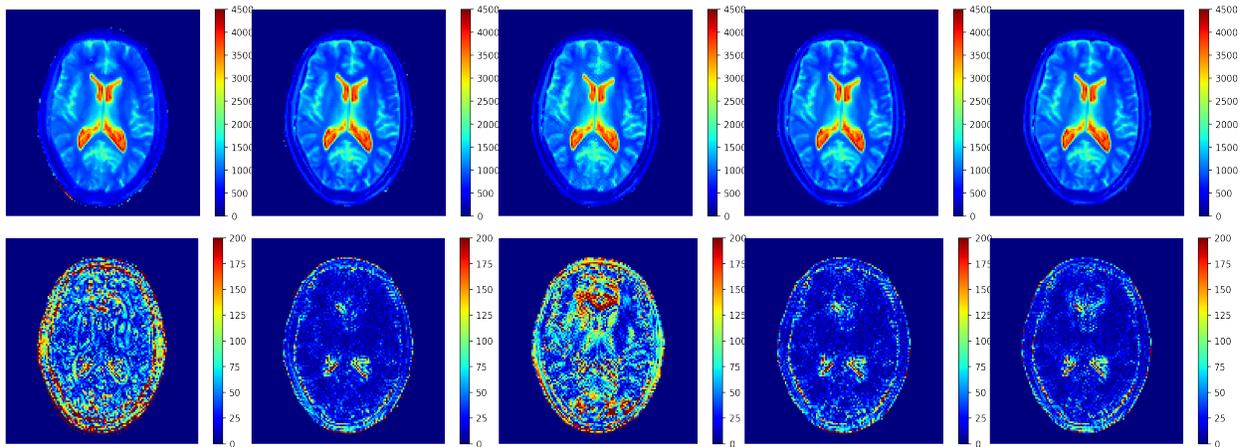


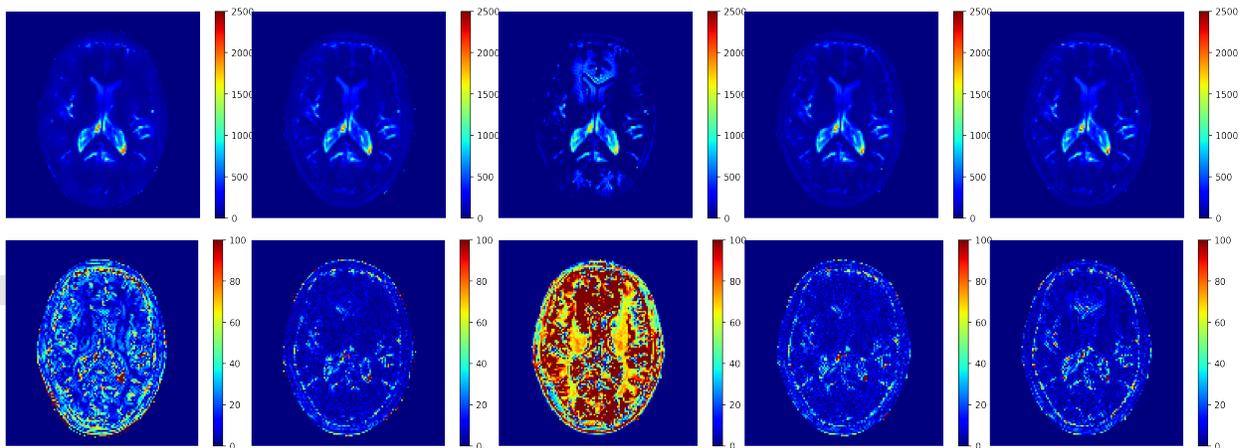
Figure 14: Testing on the anatomical dataset with k-space subsampling factor 9% using spiral trajectories and 1000 time frames. Subfig. (a) - (f) show the results using Ma et al.<sup>1</sup>, BLIP<sup>3</sup>, FLOR<sup>6</sup>, CNN by Hoppe et al.<sup>14</sup>, FNN by Cohen et al.<sup>13</sup> and HYDRA.



(a) T1 Reference (left) and T2 Reference (right).



(b) T1 estimation (top row) and residual errors (bottom row).



(c) T2 estimation (top row) and residual errors (bottom row).

BLIP <sup>3</sup>	FLOR <sup>6</sup>	CNN <sup>14</sup>	FNN <sup>13</sup>	HYDRA
SNR = 15.03/15.22 dB	SNR = 24.88/20.38 dB	SNR = 21.45/4.48 dB	SNR = 25.84/20.30 dB	SNR = 27.02/21.04 dB

Figure 15: Visual results of testing on anatomical dataset with k-space subsampling factor 9% using spiral trajectories with  $L = 1000$ . Comparison between BLIP<sup>3</sup>, FLOR<sup>6</sup>, CNN by Hoppe et al.<sup>14</sup>, FNN by Cohen et al.<sup>13</sup> and HYDRA.



## The closest approximation to water *for all applications*

Plastic Water® is a solid material that simulates water and is a convenient and reliable water alternative for photon and electron beam QA. Mimicking true water in a solid material is one of the more challenging tasks in tissue simulation but CIRS achieves a 1% and better accuracy. Plastic Water® is available in 3 energy ranges.



### **Plastic Water® Original – High energy range 50 keV-100 MeV**

Best for calibration of photon and electron beams within +/- 0.5% of true water dose according to AAPM TG51, and is faithful from 150 keV to 100 MeV



### **Plastic Water® LR – Low Energy Range 5 keV-8 MeV**

Originally formulated for water equivalency at lower photon energies and is useful in dosimetry of low energy brachytherapy sources. For CT and X-Ray imaging and all RT applications up to 20 MV



### **Plastic Water® DT – Diagnostic Therapy Range 50 keV-25 MeV**

Designed to meet demands of VMAT/IMRT verification techniques where it is desirable to match attenuation and absorption properties in both the diagnostic CT and RT energy ranges up to 30 MV

Visit us at ASTRO, **Booth 4215** to learn more.

**CIRS**

Tissue Simulation & Phantom Technology

[cirsinc.com](http://cirsinc.com)

900 Asbury Ave, Norfolk, VA 23513, USA • (800) 617-1177