



# HHS Public Access

Author manuscript

*IEEE Trans Med Imaging*. Author manuscript; available in PMC 2019 October 04.

Published in final edited form as:

*IEEE Trans Med Imaging*. 2019 October ; 38(10): 2364–2374. doi:10.1109/TMI.2019.2899328.

## Deep Learning for Fast and Spatially-Constrained Tissue Quantification from Highly-Accelerated Data in Magnetic Resonance Fingerprinting

**Zhengan Fang,**

Department of Radiology, BRIC, and the Department of Biomedical Engineering, University of North Carolina at Chapel Hill, NC, USA.

**Yong Chen,**

Department of Radiology and BRIC, University of North Carolina at Chapel Hill, NC, USA.

**Mingxia Liu,**

Department of Radiology and BRIC, University of North Carolina at Chapel Hill, NC, USA.

**Lei Xiang,**

Institute for Medical Imaging Technology, School of Biomedical Engineering, Shanghai Jiao Tong University, China.

**Qian Zhang,**

Department of Radiology and BRIC, University of North Carolina at Chapel Hill, NC, USA.

**Qian Wang,**

Institute for Medical Imaging Technology, School of Biomedical Engineering, Shanghai Jiao Tong University, China.

**Weili Lin,**

Department of Radiology and BRIC, University of North Carolina at Chapel Hill, NC, USA.

**Dinggang Shen [Fellow, IEEE]**

Department of Radiology and BRIC, University of North Carolina at Chapel Hill, NC, USA, and also with the Department of Brain and Cognitive Engineering, Korea University, Seoul 02841, Republic of Korea.

### Abstract

Magnetic resonance fingerprinting (MRF) is a quantitative imaging technique that can simultaneously measure multiple important tissue properties of human body. Although MRF has demonstrated improved scan efficiency as compared to conventional techniques, further acceleration is still desired for translation into routine clinical practice. The purpose of this work is to accelerate MRF acquisition by developing a new tissue quantification method for MRF that allows accurate quantification with *fewer* sampling data. Most of existing approaches use the MRF signal evolution at each individual pixel to estimate tissue properties, without considering the spatial association among neighboring pixels. In this work, we propose a spatially-constrained

---

\*Corresponding author: D. Shen (dgshen@med.unc.edu).

quantification method that uses the signals at multiple neighboring pixels to better estimate tissue properties at the central pixel. Specifically, we design a unique two-step deep learning model that learns the mapping from the observed signals to the desired properties for tissue quantification, i.e., 1) with a *feature extraction module* for reducing the dimension of signals by extracting a low-dimensional feature vector from the high-dimensional signal evolution and 2) a *spatially-constrained quantification module* for exploiting the spatial information from the extracted feature maps to generate the final tissue property map. A corresponding two-step training strategy is developed for the network training. The proposed method is tested on highly undersampled MRF data acquired from human brains. Experimental results demonstrate that our method can achieve accurate quantification for T1 and T2 relaxation times by using only 1/4 time points of the original sequence (i.e., *four* times of acceleration for MRF acquisition).

## Keywords

Magnetic resonance fingerprinting; neural network; quantitative magnetic resonance imaging; tissue quantification

## I. Introduction

QUANTITATIVE imaging, i.e., quantification of important tissue properties in human body such as the T1 and T2 relaxation times, is desired for both clinical and research purposes. Compared with conventional qualitative imaging techniques, e.g., T1- and T2-weighted imaging, quantitative imaging can provide more accurate and unbiased information of tissue properties and enable objective comparison of scans acquired in longitudinal studies [1]–[3]. However, one of the major barriers to translate conventional quantitative imaging techniques [4]–[7] for clinical applications is the prohibitively long time for data acquisition.

Recently, magnetic resonance fingerprinting (MRF) [8], a new approach for MR image acquisition and post-processing, has been introduced, which can significantly reduce the acquisition time for quantitative imaging. Compared with conventional MR imaging methods, the MRF framework uses pseudo-randomized imaging parameters to generate a unique signal evolution for each tissue type. Multiple tissue properties can be estimated by matching the observed signal evolution to a precomputed MRF dictionary using a template matching algorithm. The parallel measurement of multiple tissue properties using highly undersampled data acquisition enables fast quantitative imaging within the MRF framework. The efficacy of MRF for quantification of various tissue properties in multiple organs has been demonstrated in recent studies [9]–[13].

Although MRF has demonstrated better scan efficiency than conventional quantitative imaging techniques, further acceleration is still needed to improve spatial resolution/coverage and translate MRF into routine clinical practice, especially for pediatric and body imaging. Further reduction of acquisition time is of great benefit to clinical practice as it allows reduction of motion artifact, improvement of efficiency and throughput, and alleviation of patient discomfort during the MRF scan. In this work, we aim to shorten the acquisition time of MRF by acquiring fewer time points for each scan. However, acquiring fewer time points will result in fewer sampled data available for tissue quantification and

thus degraded quantification accuracy. Thus, a robust tissue quantification method that can accurately estimate tissue properties from fewer time points is needed for further acceleration of MRF acquisition.

In the literature, various methods have been proposed for tissue quantification in MRF. The original framework [8] uses a template matching method, where the signal evolution at each pixel is matched to a precomputed dictionary containing signal evolutions corresponding to a wide range of tissue types. The tissue properties used to generate the best matching entry in the dictionary are then assigned to the corresponding pixel. McGivney et al. [14] propose to use singular value decomposition (SVD) to reduce the dictionary dimension, thus improving the computational efficiency. Cauley et al. [15] has introduced a fast group matching algorithm that assigns dictionary entries to several highly-clustered groups to further accelerate the matching process. Yang et al. [16] has developed randomized SVD and dictionary polynomial fitting to largely reduce the memory requirement of calculating low-rank approximations for large-scale dictionaries. In [17], the authors combine the dictionary matching technique with the compressed sensing (CS) framework by iteratively performing MR image reconstruction and tissue quantification to improve the quantification accuracy. Furthermore, the CS-based method is improved by adding certain constraints (i.e., low-rank constraint [18] and wavelet domain sparsity [19]) on the reconstructed MR images and reconstructing the MR images at different resolution levels to suppress the noise in the reconstructed images [20]. However, these methods only use the signal at a single pixel to estimate tissue properties, without considering spatial information of the whole image. To incorporate spatial information, [21] introduces a spatiotemporal dictionary, with each entry containing signals at multiple neighboring pixels. However, this method dramatically increases the dimension of each dictionary entry and thus could only exploit spatial information from a small neighborhood (e.g.,  $7 \times 7$  pixels) due to its high computation load.

Besides the aforementioned dictionary-based methods, other methods that do not involve dictionary matching have also been developed for the tissue quantification process. These methods can be roughly divided into two categories: 1) model-based and 2) learning-based methods. The model-based methods [22], [23] usually employ mathematical models to simulate the signal generation and acquisition process of MRF, and then obtain an statistically optimal estimation of the tissue property maps from the observed signals. Although the model-based methods provide new theoretical insights into the MRF framework, they are subject to unavoidable mismatches between mathematical models and the real-world non-ideal imaging systems [24], and also need complex algorithms to solve the statistical optimization problem. In the second category, the learning-based methods [25]–[28] use deep learning models to approximate a direct mapping from MRF signals to the underlying tissue properties. These methods have demonstrated improved computation efficiency since they perform tissue quantification with feed-forward neural networks without iterative computations. The performance of some learning-based methods has been evaluated on *in vivo* human brain data [27], [28]. However, previous learning-based methods rely on signal evolution acquired from single pixel for tissue quantification, without exploiting spatial context information of images. Moreover, these methods aim at improving computation efficiency, while the potential for further acceleration of MRF acquisition has not been evaluated in these approaches.

In this work, we propose a learning-based method for tissue quantification in MRF. Our method exploits spatial context information by using a deep learning model to learn the mapping from the signals at multiple neighboring pixels to the tissue properties at the central pixel. We hypothesize that the spatial context information is helpful for accurate quantification for two reasons. *First*, there is correlation between tissue properties at neighboring pixels. For example, the tissue properties at adjacent pixels in a homogeneous region are likely to be similar. Therefore, information from neighboring pixels could be utilized as spatial constraint to regularize the estimation at the central target pixel for improving the accuracy of tissue quantification. *Second*, signals at the target pixel are distributed to its neighboring pixels due to the aliasing effect caused by undersampling in k-space during MRF acquisition. Therefore, using spatial information can help recover the scattered signals and achieve a better quantification with MRF.

A major challenge here is the high dimension of the observed signal evolution at each pixel due to the large number of acquired time points. To this end, we develop a unique two-step deep learning model for spatially-constrained tissue quantification, including 1) a *feature extraction module* in the first step used to reduce the dimension of signals by extracting a low-dimensional feature vector from the high-dimensional signal evolution, and 2) a *spatially-constrained quantification module* used to exploit the spatial information from the extracted feature maps to generate the final tissue property map. We further design a two-step training strategy for learning this two-step model. Moreover, a relative-difference-based loss function is adopted to tackle with the large range of the tissue property values to be estimated. Experiments on 6 subjects demonstrate that our method is superior to several state-of-the-art methods.

The rest of this paper is organized as follows. We first present materials used in this work and our proposed method in Section II, and then introduce experimental results and related analysis in Section III. In Section IV, we compare quantification results for T1 and T2, compare our method with previous studies, and present the limitations of the current work and future research direction. We finally conclude this paper in Section V.

## II. Materials and Method

In this section, we first describe the data acquisition and preprocessing approach used in this work. We then present our proposed tissue quantification method in detail, including our proposed two-step deep learning model, two-step training strategy, and implementation details.

### A. Data Acquisition and Preprocessing

All imaging was performed on a Siemens 3T Prisma scanner. A 32-channel head coil was used for signal reception. MRF data of axial human brain slices were acquired using the fast imaging with steady state precession (FISP) sequence [29]. A total of 2,304 time points were acquired for each scan, and data from only one spiral readout was acquired for each time point (reduction factor = 48). Other imaging parameters used in this study included: field of view (FOV) = 30 cm; matrix size =  $256 \times 256$ ; slice thickness = 5 mm; flip angle =  $5^\circ \sim 12^\circ$ .

A constant TR of 7.0 msec was used and the acquisition time for each 2D MRF dataset was ~23 seconds.

The MRF dictionary used in this study contains 13,123 combinations of T1 (60 ~ 5000 ms, with an increment of 10 ms below 2000 ms, an increment of 20 ms between 2000 ms and 3000 ms, an increment of 50 ms between 3000 ms and 3500 ms, and an increment of 500 ms above 3500 ms) and T2 (10 ~ 500 ms, with an increment of 5 ms below 200 ms, an increment of 10 ms between 200 ms and 300 ms, and an increment of 50 ms above 300 ms). The signal evolution corresponding to each combination was simulated using Bloch equations.

The ground-truth tissue property maps were obtained from the acquired MRF data of all 2,304 time points by using the dictionary matching method as introduced in the original MRF framework [8]. Specifically, MR images are first reconstructed using non-uniform Fast Fourier Transform (NUFFT) [30]. Next, the signal evolution in the dictionary that best matches the observed signal evolution at each pixel is selected by using the cross correlation as similarity metric. The T1 and T2 values corresponding to the best-matching entry are assigned to that pixel. Repeated over the entire image, this process yields quantitative T1 and T2 maps simultaneously. The obtained tissue property maps are then used as the ground truth in the following experiments. Since the magnitudes of MRF signal evolutions vary largely across different subjects, the MRF signal is normalized to a common range for better generalization of the deep learning model. Specifically, we normalize the energy (i.e., the sum of squared magnitudes) of each signal evolution to one, which is similar to the normalization performed when calculating cross correlation in the dictionary matching method [8].

For simplicity, denote the normalized MRF signals of an axial slice as  $X \in \mathbb{C}^{M \times N \times T}$ , where  $M \times N$  is the size of the imaging matrix, i.e.,  $256 \times 256$  in this study, and  $T$  is the number of time points used for tissue quantification. Denote the signal evolution at the pixel  $(m, n)$  as  $x_{m,n} \in \mathbb{C}^T$ . Denote the ground-truth tissue property (T1 or T2) map of that axial slice as  $\Theta \in \mathbb{R}^{M \times N}$ .

## B. Proposed Model

We design a two-step deep learning model to learn the mapping from the MRF signals  $X$  to the tissue property map  $\Theta$ . Our model contains two sequential components, i.e., 1) a feature extraction (**FE**) module for reducing the dimension of signal evolutions, and 2) a spatially-constrained quantification (**SQ**) module for estimating the tissue property maps from the extracted feature map. The schematic overview of our proposed model is shown in Fig. 1, with Figs. 2 and 3 showing the detailed network structure for each of the two modules. The structures of FE and SQ modules are described in details in the following sections.

**1) Feature Extraction Module:** In the feature extraction (FE) module, fully-connected neural networks (FNNs) are used to convert the input high-dimensional signal evolution into a low-dimensional feature vector which contains useful information for tissue property estimation. One FNN is used to extract the features useful for estimation of one specific tissue property. In total, our FE module contains two FNNs for estimation of two tissue

properties, i.e., T1 and T2 relaxation times. Specifically, each FNN learns a nonlinear mapping  $f$  from the signal evolution at a certain pixel  $x_{m,n} \in \mathbb{C}^T$ , to the feature vector at that pixel  $y_{m,n} \in \mathbb{R}^D$ :

$$y_{m,n} = f(x_{m,n}) \quad (1)$$

where  $T$  is the number of time points used for tissue quantification and  $D$  ( $D=46$  in this study) is the number of features extracted. Applying the FE module to all the pixels in the axial slice yields a low-dimensional feature map  $Y \in \mathbb{R}^{M \times N \times D}$  corresponding to the high-dimensional MRF signals  $X$ :

$$Y = f(X) \quad (2)$$

The FE module is needed for the following reasons. First, MRF implementation usually acquires a large number of time points, such as 2,304 in the FISP sequence and 576 after 4 times of acceleration. In this case, directly feeding such high-dimensional data into the subsequent spatially-constrained quantification module will lead to large network size and challenges in efficient network training. Besides, the FE module can provide a better representation of the original signal for tissue quantification, by extracting only the essential information for estimation of the target tissue property while filtering out the noise and unrelated information in the original signal.

Previous studies [14], [18], [23], [31] used singular value decomposition (SVD) to reduce the dimension of signal evolutions. The advantage of using deep neural networks, i.e., FNN in this study, instead of SVD is twofold. *First*, the deep neural network learns a multilayer nonlinear mapping from the original signal to the extracted features, whereas SVD learns only a single-layer linear mapping. Therefore, the deep neural network can extract more abstract and higher-level information from the input signal, which can be used to improve the robustness and accuracy of tissue quantification. *Second*, the combination of neural networks for both FE and SQ modules allows end-to-end training of the entire two-step model through back propagation, which further improves the compatibility between two modules and thus the performance of our method.

There are various choices for the structure of FNN in the FE module. The particular structure adopted in this study is composed of 4 fully-connected (FC) layers (as shown in Fig. 2), where each FC layer includes a full connection followed by batch normalization and ReLU activation. The output dimensions of all the FC layers are the same (i.e.,  $D=46$ ). The input of FNN, i.e., a signal evolution  $x_{m,n} \in \mathbb{C}^T$ , is transformed to a real vector by splitting the real and imaginary parts, thus the input dimension of FNN is  $2T$ .

**2) Spatially-Constrained Quantification Module:** In the spatially-constrained quantification (SQ) module, convolutional neural networks (CNNs) are used to capture spatial information of the feature map  $Y$  and finally generate the estimated tissue property map  $\hat{\Theta} \in \mathbb{R}^{M \times N}$ . One CNN is used for the estimation of one tissue property. In total, our SQ

module contains two CNNs for estimation of two tissue properties, i.e., T1 and T2 relaxation times. Specifically, each CNN learns a nonlinear mapping  $s$  from the feature map  $Y$  to the estimated tissue property map  $\hat{\Theta}$ :

$$\hat{\Theta} = s(Y) \quad (3)$$

There are also various choices for the structure of CNN in the SQ module. In this study, we use U-Net [32], [33] (as shown in Fig. 3) to capture both local and global spatial information of feature map  $Y$ . This network includes an encoder sub-network (i.e., left part of Fig. 3) to extract spatial features at multiple resolution levels from the input feature map, and a subsequent decoder sub-network (i.e., right part of Fig. 3) to estimate the output tissue property (T1 or T2) map based on the extracted spatial features from encoder subnetwork.

In the encoder sub-network, the spatial context information is incorporated into the extracted feature maps by feature extraction (with convolution) and  $2\times$  down-sampling (with max pooling). Specifically, each convolution layer incorporates spatial information from  $3\times 3$  neighboring pixels by applying  $3\times 3$  convolution kernels to the input feature map. Moreover, each max pooling layer incorporates information from  $2\times 2$  pixels by retaining the max value of  $2\times 2$  subregion in the input. Spatial sharing further increases with the concatenation of convolution layers and pooling layers.

In the subsequent decoder sub-network, the feature maps at different spatial scales are up-sampled (with transpose convolution), concatenated, and merged (with convolution), to combine global context knowledge with complementary local details and generate the final tissue property maps with spatial context information from different scales. The U-Net shown in Fig. 3 has a receptive field size of  $54\times 54$  pixels, which means that the signals from  $54\times 54$  neighboring pixels in the input MRF images are used for estimation of tissue property at one pixel in the output tissue property map.

### C. Two-Step Training Strategy

To better train the proposed two-step model, we design a two-step training strategy, including 1) pretraining of the FE module by signals and tissue properties at individual pixels, and 2) end-to-end training of the entire model by signals and tissue property maps of whole axial slices.

**1) Pretraining of the Feature Extraction (FE) Module:** As there is no ground truth for the output features of FE module, we extend each FNN with one FC layer and trains the extended FNN to estimate the desired tissue property (corresponding to the input signal evolution). Therefore, the features extracted by the original FNN will capture useful information for estimation of the desired tissue property. Denote the mapping learned by the added FC layer as  $f_a$  and the mapping learned by the extended FNN as  $f_a \circ f$ . The objective function of pretraining of the FE module can be formulated as the following:

$$\xi_f, \xi_{f_a} = \arg \min_{\xi_f, \xi_{f_a}} \mathbb{E}_{x, \theta} \left\| \frac{\theta_{m,n} - f_a \circ f(x_{m,n})}{\theta_{m,n}} \right\| \quad (4)$$

where  $\theta_{m,n} \in \mathbb{R}$  is the ground-truth tissue property at the pixel  $(m, n)$ ,  $f_a \circ f(x_{m,n}) \in \mathbb{R}$  is the output of extended FNN for input signal evolution  $x_{m,n}$ ,  $\xi_f$  and  $\xi_{f_a}$  are the network parameters of the original FNN and the added FC layer respectively, and  $\mathbb{E}[\cdot]$  represents mathematical expectation.

Note that we use *relative difference* between the ground truth  $\theta_{m,n}$  and the network output  $f_a \circ f(x_{m,n})$  as the loss function, instead of the absolute difference which is commonly used for regression problems. The main reason is that T1 and T2 measures in human body span a large range, and thus the loss function based on the conventional absolute difference will be dominated by the pixels with high tissue properties. Therefore, we use relative difference as the loss function to balance across tissues with different ranges in our proposed method.

The pretraining of FE module is helpful in providing better initial parameters of FE module for the subsequent end-to-end training of the entire model. *Besides*, during pretraining, more data can be used to better train the FE module since the signal and tissue property at each individual pixel can be used as training data. However, during the subsequent end-to-end training, only the whole images or image patches can be used as training data, to provide the spatial context information needed for spatially-constrained tissue quantification.

**2) End-to-End Training of the Entire Model:** After pretraining of the FE module, the end-to-end training is performed to train the SQ module and fine-tune the FE module. During the end-to-end training, the parameters in both the FE and SQ modules are tuned together, so that the two modules are more compatible and the performance of the entire model can be improved. The objective function of end-to-end training can be formulated as the following:

$$\xi_s, \xi_f = \arg \min_{\xi_s, \xi_f} \mathbb{E}_{X, \Theta} \left[ \left\| \frac{\Theta - s \circ f(X)}{\Theta} \right\|_1 \right] \quad (5)$$

where  $s$  and  $f$  are the mappings learned by SQ and FE modules, respectively, while  $\xi_s$  and  $\xi_f$  are the network parameters of SQ and FE modules, respectively. Also,  $s \circ f(X) \in \mathbb{R}^{M \times N}$  is the output of entire model for the input MRF signals  $X$ , and  $\|\cdot\|_1$  stands for the entry-wise L<sub>1</sub>-norm.

**3) Training Details:** For both training stages, we used ADAM optimizer [34] with a batch size of 32 and an initial learning rate of 0.0002. At each stage, we trained the networks for 400 epochs, by keeping the learning rate the same for the first 200 epochs and linearly decaying it to zero over the next 200 epochs. For end-to-end training, we extracted image patches of 64×64 pixels as training data. Network weights were initialized from a Gaussian



distribution with mean 0 and standard deviation 0.02. We implemented the algorithm in Python with PyTorch library and trained the models on a GeForce GTX TITAN XP GPU.

#### D. Implementation

When the training is completed, the model can be applied to new testing data for tissue quantification. Specifically, the model calculates the desired tissue property map  $\hat{\Theta} \in \mathbb{R}^{M \times N}$  for the input MRF signals of an axial slice  $Z \in \mathbb{C}^{M \times N \times T}$  by:

$$\hat{\Theta} = s \circ f(Z) \quad (6)$$

Note that the tissue quantification is performed by a direct mapping from the observed signals to the tissue property map, i.e.,  $s \circ f$ , thus our method is much more computationally efficient than the dictionary-based or model-based methods that require iterative computations.

### III. Experiments

In this section, we first present experimental settings and competing methods, then report experimental results achieved by our method and competing methods, and finally investigate the influence of several key components of our method.

#### A. Experimental Settings

Our dataset contains axial brain slices from 6 human subjects. Among the 6 subjects, 5 subjects were acquired with 12 slices per subject and one subject was acquired with 10 slices. The MRF image series and ground-truth T1 and T2 maps of each slice were obtained through the method described in Section II.A.

In this work, we accelerate MRF acquisition by using fewer time points for tissue quantification. Specifically, for the acceleration rate  $ar$ , we only use the first  $1/ar \cdot T_a$  of all  $T_a$  ( $T_a = 2,304$  in this study) time points. For example, when  $ar = 4$ ,  $T$  (i.e., the number of time points used for tissue quantification) =  $1/4 \times 2,304 = 576$ .

In our experiments, we use the relative error to measure the quantification accuracy achieved by different methods:  $e_{m,n} = |(\theta_{m,n} - \hat{\theta}_{m,n})/\theta_{m,n}|$ , where  $e_{m,n} \in \mathbb{R}$  is the quantification error at the pixel  $(m, n)$ , and  $\theta_{m,n} \in \mathbb{R}$  and  $\hat{\theta}_{m,n} \in \mathbb{R}$  are the ground-truth and estimated tissue properties at that pixel, respectively. The relative error of an axial slice is calculated by averaging the relative errors across all brain tissues (excluding background) in the slice. The mean and standard deviation of relative errors of all testing slices are calculated for quantitative comparison between different methods.

#### B. Competing Methods

We first compare our proposed spatially-constrained tissue quantification (**SCQ**) method with the baseline dictionary matching (**DM**) method proposed in the original MRF framework [8]. In the DM method, we use the same dictionary as that used to calculate the ground

truth (as described in Section II.A). Specifically, the dictionary is generated using the Bloch equations and contains 13,123 entries corresponding to different combinations of T1 and T2 values.

Besides, we further compare our SCQ method with three state-of-the-art methods, including 1) SVD-compressed dictionary matching (**SDM**) [14], 2) compressed sensing multiresolution reconstruction (**CSMR**) [20], and 3) conventional deep learning (**DL**) method [28]. The details of these state-of-the-art methods are listed as follows.

**1) SVD-Compressed Dictionary Matching (SDM) [14]:** A variant of DM that uses SVD to compress the dictionary, which is reported to have better computation efficiency than DM. Specifically, in SDM, SVD is applied to all the entries in the dictionary used in DM. Then, each entry is projected to a lower dimensional subspace spanned by the singular vectors. The dimension of the subspace is set as 17, which corresponds to 99.9% of the total energy of all singular values.

**2) Compressed Sensing Multi-resolution Reconstruction (CSMR) [20]:** A compressed-sensing-based method that uses multi-resolution reconstruction for MR images, which is reported to have good quantification accuracy for accelerated data with fewer time points. Specifically, in the CSMR method, MR image reconstruction is performed iteratively at different resolution levels, and tissue quantification is performed at each iteration through dictionary matching. The number of iterations is set as 14, and the tissue property maps obtained from the last (i.e., 14th) iteration are used as the final quantification results.

**3) Conventional Deep Learning (DL) Method [28]:** A non-spatially-constrained deep-learning-based method, which is reported to have better computation efficiency than DM. Specifically, the DL method treats each observed signal evolution as a time sequence and feeds it into a CNN with six 1-D temporal convolution layers followed by six fully-connected layers. The network then outputs the desired tissue properties.

It is worth noting that DM, SDM, and CSMR are dictionary-matching-based methods, while DL and our SCQ method are deep-learning-based methods.

### C. Comparison with Baseline Method

In this group of experiments, we compare our SCQ method with the baseline DM method with 3 acceleration rates: 1)  $ar=2$ ,  $T=1152$ ; 2)  $ar=4$ ,  $T=576$ ; and 3)  $ar=8$ ,  $T=288$ . Here, we use slices of 5 randomly selected subjects as the training data, and slices of the remaining one subject as the test data. The quantification results for a representative slice in the test dataset yielded by the two methods are shown in Fig. 4. In general, our method achieves more accurate quantification results than the baseline DM method. Notably, with  $ar=8$  for T2 quantification, our method yields a reasonable result (error = 8.0%), while DM nearly fails to estimate the T2 map (error = 60.9%).

The means and standard deviations of the relative errors calculated from all slices in the test data are summarized in Fig. 5. As shown in Fig. 5, our method yields lower error than the baseline method for T2 quantification when  $ar=8$ , 4, and 2, and also for T1 quantification

when  $ar = 8$  and 4. In addition, the improvement is more evident when the acceleration rate is higher, i.e., the acquisition time is shorter.

Besides, we also compared the SCQ method with DM for the data with no acceleration. Note that the DM results with no acceleration are used as the ground-truth tissue property maps in this study. The quantification errors yielded by SCQ are  $1.00\% \pm 0.19\%$  for T1 and  $2.82\% \pm 0.19\%$  for T2, which suggests a good agreement between the quantification results of SCQ and DM with no data acceleration.

#### D. Comparison with State-of-the-art Methods

In this group of experiments, we compare our method with three state-of-the-art methods in terms of quantification accuracy and processing time, with  $ar = 4$ . Here, we performed subject-level leave-one-out cross validation. Specifically, each time we used the slices of one subject as test data and slices of all the remaining 5 subjects as training data. This process was repeated 6 times until all subjects were alternatively used as test data. The quantification errors yielded by different methods for each test subject are summarized in Table I. For comparison, we also show the errors of DM in Table I.

From Table I, we can observe that our SCQ method consistently achieves superior accuracy for T1 and T2 quantification compared to the four competing methods. Also, compared with the conventional deep-learning-based method (i.e., DL), our method achieves higher T2 quantification accuracy because our method can exploit spatial context information through CNNs in its SQ module, while the DL method cannot exploit that information through its network. On the other hand, compared with the dictionary-based methods (i.e., DM, SDM, and CSMR), our method learns a multilayer nonlinear mapping from the signal to the underlying tissue properties, which is potentially more noise-robust than the simple cross-correlation-based template matching procedure performed in the dictionary-based methods. Also, the dictionary-based methods perform tissue quantification for each pixel separately, without considering the spatial context information.

In addition, we also report the average testing times for one axial slice required by different methods in Table II. The testing time includes all computations performed using the reconstructed MRF image series for the quantification of T1 and T2 maps. Specifically, both signal compression time and matching time are included for SDM. From Table II, we can clearly see that our SCQ method achieves the shortest testing time, i.e., 2.3 seconds for quantification of T1 and T2 for an axial slice with a matrix size of  $256 \times 256$ , while the second fastest method (i.e., SDM) needs 14.2s. These results suggest that, compared with the competing methods, our method is more efficient in tissue quantification, which is particularly useful in practice.

#### E. Effects of Important Components in Our Method

In this subsection, we study the effect of three essential components in our method for tissue quantification. All the experiments in this subsection were performed with  $ar = 4$  and  $D = 46$  unless stated otherwise. In the Supplementary Materials, we further analyze the influence of another two components (i.e., network architecture and data normalization) in SCQ, with results shown in Fig. S1 and Tables SI–SIII.

**1) Influence of Feature Number, and Comparison between FNN and SVD:** We evaluate the influence of the number of features extracted by the FE module, i.e.,  $D$ . To estimate possible values for  $D$ , we first apply SVD to the MRF dictionary using the method in [14]. Multiple numbers of singular values corresponding to certain energy thresholds are then selected as the  $D$  values. Specifically, we select  $D = 8, 10, 22, 46, 76, 124$ , and  $170$ , corresponding to the energy thresholds of 90%, 95%, 99%, 99.9%, 99.99%, 99.999%, and 99.9999%, respectively.

We also compare our method (SCQ) with its variant **SCQ\_S** which uses SVD (instead of FNN) for feature extraction in the FE module. We use the method proposed in [14] to apply SVD to the MRF signal evolutions.

The quantification errors yielded by SCQ and SCQ\_S with different numbers of extracted features are summarized in Fig. 6. As shown in Fig. 6, SCQ generally yields lower quantification error than SCQ\_S. Besides, the quantification error of SCQ first drops then rises when the number of extracted features increases. Specifically, the error of SCQ reaches the minimum for T1 when  $D = 124$ , and for T2 when  $D = 46$ . It is worth noting that the quantification error is lower from SCQ\_S than SCQ when the number of extracted features is around 10. To further study this case, we examine the visual results for  $D = 10$  and  $D = 46$  in Section I.C and Fig. S2 in the Supplementary Materials.

**2) Influence of the FE and SQ Modules:** In this group of experiments, we study the importance of the FE module and the SQ module, respectively, by removing each of them from the original SCQ model.

Specifically, we compare SCQ with its two degraded variants. In the first variant (denoted as **SQ-only**), we feed the high-dimensional MRF signals directly into the SQ module, without using the FE module for dimension reduction. In SQ-only, the number of input channels for the CNN in SQ module increases from  $D$  to  $2T$ , with each channel corresponding to the real or imaginary part of the signal at one time point. In the second variant (called **FE-only**), we use the extended FNN obtained from the pretraining of FE module to estimate the tissue property from individual signal evolutions, without using the SQ module to exploit spatial context information.

The results yielded by our SCQ method and its two variants (i.e., SQ-only and FE-only) are summarized in Table III. As shown in Table III, removing the FE module (in SQ-only) increases the quantification errors, which is likely due to the over-fitting problem caused by the increased network size in SQ module. Removing the SQ module (in FE-only) increases quantification errors, demonstrating the importance of the proposed SQ module and the spatial context information it extracts for accurate tissue quantification.

**3) Importance of the Phase Information in MRF Signals:** In our SCQ method, both the real and imaginary parts of the complex MRF signals are inputted to our model for tissue quantification. In the previous work that applies deep learning to MRF [26], [28], only the magnitude of signals (without any phase information) is inputted. In this group of experiments, we study whether the phase in MRF signals provides valuable information for

tissue quantification, by comparing our method (SCQ) with its another variant, called **SCQ\_M**, which uses only the magnitude of signals as input. The results are summarized in Table IV. As shown in Table IV, SCQ\_M yields much higher quantification errors than SCQ, suggesting that both magnitude and phase values in MRF signals contain useful information for tissue quantification. It is worth noting that the error increase due to removing phase information is much higher than that due to removing FE or SQ module (as shown in Table III), which further demonstrates the importance of phase information in MRF signals for accurate tissue quantification. This result is consistent with recent findings in the literature [35].

## IV. Discussion

### A. Difference between T1 and T2

It has been observed in Figs. 4, 5, and 6 and Tables I, III, and IV that both our SCQ method and the existing methods yield higher accuracy for T1 quantification than for T2 quantification. Previous studies have also observed this difference for longer acquisitions [22], which indicates that the FISP sequence is not as sensitive to T2 as to T1. Moreover, the FISP sequence applies an inversion preparation pulse at the beginning of the acquisition, stimulating T1-recovery-like dynamics that dominate the early portion of signal evolutions [22]. Therefore, the early portion of signal evolutions, which is used for tissue quantification in this study, contains more information of T1 than of T2, resulting in higher accuracy in T1 quantification. This also explains why the optimal number of extracted features for T1 (i.e., 124) is greater than that for T2 (i.e., 46), as shown in Section III.E.1, as more features are needed to encode the information of T1 contained in the MRF signals used for tissue quantification.

### B. Flexibility of SCQ Framework

As demonstrated by the experiments with different acceleration rates, our framework is flexible with different acquisition lengths. Moreover, the framework can also be applied to various other acquisition protocols, including different MRF pulse sequences, under-sampling designs, and 3D imaging. When applying our framework to a new acquisition protocol, a new set of training dataset is likely needed to train the deep neural networks (i.e., FNN and CNN in our framework) and the model can then be applied to the data acquired with the corresponding protocol.

Compared to the standard dictionary matching approach, our framework could potentially provide more flexibility in the estimation of tissue properties. Specifically, the MRF dictionary only covers a limited range of discrete tissue property values, which limits the precision of quantification and leads to partial volume problems. In comparison, our framework learns an unbounded and continuous mapping from MRF signals to underlying tissue properties, which could potentially improve the performance of tissue quantification.

### C. Comparison with Previous Studies

Compared with the baseline DM method [8], our method can significantly improve the quantification accuracy for highly-accelerated data, as shown in Figs. 4 and 5. However, it is

worth noting that some blurring effects can be observed in the quantitative maps estimated by our method with 8 times of acceleration (see Fig. 4). In the Supplementary Materials, we further study this blurriness through frequency-domain representations of the quantitative maps. The blurring effect of our method potentially degrades spatial resolution and hinders application for high-resolution imaging [36]–[39]. To tackle with this problem, a loss function can be added to enhance the high frequency details of our output tissue property map, such as a discriminative loss within generative adversarial networks (GAN) [40]–[42], or a gradient loss that minimizes the difference between the gradient maps of network output and the ground-truth tissue property map [42], [43].

Compared with three state-of-the-art methods (i.e., SDM, CSMR, and DL), our method achieves higher quantification accuracy with a shorter processing time, as shown in Tables I and II. Among the state-of-the-art methods, SDM reduces the processing time of the baseline DM method by 46% but achieves poorer quantification accuracy than DM. DL shortens the processing time of DM by 31% while simultaneously achieving higher accuracy. CSMR can achieve higher quantification accuracy than the baseline DM method in certain cases for T2, but at a cost of longer processing time. On the other hand, our method can improve both the quantification accuracy and the processing speed of the baseline DM method. Furthermore, the dictionary-based methods, i.e., DM, SDM, and CSMR, all have a testing time that grows exponentially with the number of tissue properties to be estimated, since the dictionary dimension grows exponentially with the number of tissue properties [22], [25]. However, the testing time of our method only grows linearly with the number of tissue properties, as the number of networks used in our method grows linearly with the number of tissue properties. Given these relationships, our method is potentially superior to the dictionary-based methods in computation efficiency when more tissue properties are quantified with MRF in future work.

#### D. Limitations and Future Work

There are several limitations in the current method. *First*, our method, like other deep-learning-based methods, needs large amounts of training data for network learning. All the training data used in this study are acquired from *in vivo* subjects at our institution, while the number of subjects (i.e., 6) is not large enough. In the future work, simulation data, obtained from computer simulation based on the mathematical models of the data generation and acquisition process, will also be used for training to increase the training data size and further optimize our model. *Second*, our method only exploits the spatial context information inside an axial slice, without considering the spatial information from neighboring slices. Inspired by [44]–[47] that extend the MRF framework from 2D to 3D, we will extend our 2D model for 3D application to exploit the spatial information from all 3 spatial dimensions. *Third*, our method estimates different tissue properties separately, ignoring the inherent relationship among different tissue properties, e.g., T1 and T2. We plan to explicitly exploit this correlation through a multi-target joint estimation network [48] to improve the quantification accuracy. *Finally*, other advanced convolutional neural networks besides the U-Net [32], e.g., ResNet [49], DenseNet [50] and others [51]–[53], can also be used in SQ module to capture spatial context information in MRF signals in future work.

Besides the aforementioned improvements in methodology, more experiments can be also conducted to better evaluate our method. *First*, in this work, we only test our method with the FISP sequence. The performance of our method on other MRF sequences, e.g., bSSFP [8], [54], is worth evaluating. *Second*, we only examine two tissue properties, i.e., T1 and T2 relaxation times, while other properties, e.g., spin density and T2\* relaxation time, are also worth evaluating. *Third*, our dataset contains only healthy subjects with no brain lesions. In the future work, data from subjects with brain lesions or other brain diseases can be also used for testing, to better evaluate our method's potential for clinical usage.

## V. Conclusion

In this paper, we have introduced a deep-learning-based spatially-constrained tissue quantification method for magnetic resonance fingerprinting (MRF). A unique two-step deep learning model is proposed to tackle with the high dimensionality of MRF data, including a feature extraction module and a spatially-constrained quantification module. A two-step training strategy and a relative-difference-based loss function are used to optimize the performance of our model. Based on the experimental results, our method can 1) achieve accurate quantification of T1 and T2 relaxation times from highly undersampled *in vivo* brain data, and 2) allow at least *four* times of acceleration for MRF data acquisition without significant loss of accuracy in tissue quantification.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## References

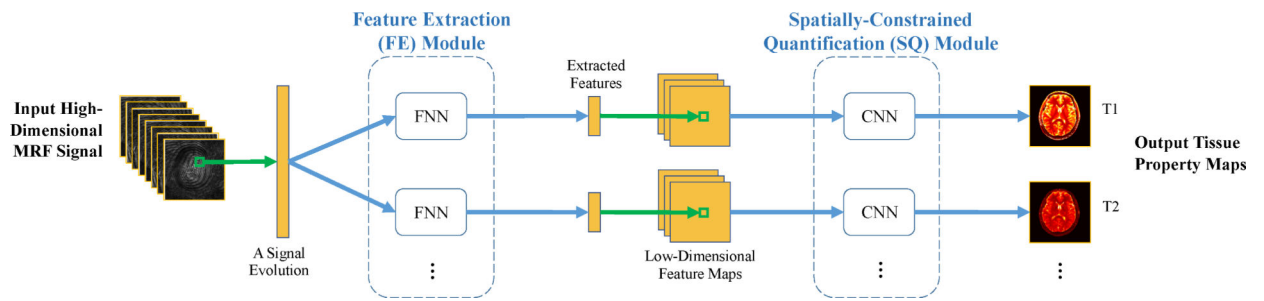
- [1]. Panda A, Mehta BB, Coppo S, Jiang Y, Ma D, Seiberlich N, Griswold MA, and Gulani V, "Magnetic resonance fingerprinting-an overview," *Current Opinion in Biomedical Engineering*, 2017.
- [2]. Coppo S, Mehta BB, McGivney D, Ma D, Chen Y, Jiang Y, Hamilton J, Pahwa S, Badve C, Seiberlich N, Griswold M, and Gulani V, "Overview of magnetic resonance fingerprinting," *Magnetom Flash*, vol. 1, no. 65, pp. 12–21, 2016.
- [3]. Hernando D, Levin YS, Sirlin CB, and Reeder SB, "Quantification of liver iron with MRI: State of the art and remaining challenges," *Journal of Magnetic Resonance Imaging*, vol. 40, no. 5, pp. 1003–1021, 2014. [PubMed: 24585403]
- [4]. Ehses P, Seiberlich N, Ma D, Breuer FA, Jakob PM, Griswold MA, and Gulani V, "IR TrueFISP with a golden-ratio-based radial readout: Fast quantification of T1, T2, and proton density," *Magnetic Resonance in Medicine*, vol. 69, no. 1, pp. 71–81, 2013. [PubMed: 22378141]
- [5]. Warntjes JBM, Dahlqvist Leinhard O, West J, and Lundberg P, "Rapid magnetic resonance quantification on the brain: Optimization for clinical usage," *Magnetic Resonance in Medicine*, vol. 60, no. 2, pp. 320–329, 2008. [PubMed: 18666127]
- [6]. Schmitt P, Griswold MA, Jakob PM, Kotas M, Gulani V, Flentje M, and Haase A, "Inversion Recovery TrueFISP: Quantification of T1, T2, and Spin Density," *Magnetic Resonance in Medicine*, vol. 51, no. 4, pp. 661–667, 2004. [PubMed: 15065237]
- [7]. Warntjes JB, Dahlqvist O, and Lundberg P, "Novel method for rapid, simultaneous T1, T2\*, and proton density quantification," *Magnetic Resonance in Medicine*, vol. 57, no. 3, pp. 528–537, 2007. [PubMed: 17326183]
- [8]. Ma D, Gulani V, Seiberlich N, Liu K, Sunshine JL, Duerk JL, and Griswold MA, "Magnetic resonance fingerprinting," *Nature*, vol. 495, no. 7440, pp. 187–192, 3 2013. [PubMed: 23486058]

- [9]. Jiang Y, Ma D, Keenan KE, Stupic KF, Gulani V, and Griswold MA, "Repeatability of magnetic resonance fingerprinting T1 and T2 estimates assessed using the ISMRM/NIST MRI system phantom," *Magnetic Resonance in Medicine*, vol. 78, no. 4, pp. 1452–1457, 2017. [PubMed: 27790751]
- [10]. Badve C, Yu A, Rogers M, Liu Y, Schluchter M, Sunshine J, Griswold M, Gulani V, and Core B, "Simultaneous T1 and T2 Brain Relaxometry in Asymptomatic Volunteers Using Magnetic Resonance Fingerprinting," *Tomography*, vol. 1, no. 2, pp. 136–144, 2015. [PubMed: 26824078]
- [11]. Chen Y, Jiang Y, Ma D, Wright KL, Seiberlich N, Griswold MA, and Gulani V, "Magnetic Resonance Fingerprinting (MRF) for Rapid Quantitative Abdominal Imaging," *Proceedings of the 22nd Annual Meeting of ISMRM, Milan, Italy*, vol. 22, p. 561, 2014.
- [12]. Chen Y, Jiang Y, Pahwa S, Ma D, Lu L, Twieg MD, Wright KL, Seiberlich N, Griswold MA, and Gulani V, "Mr fingerprinting for rapid quantitative abdominal imaging," *Radiology*, vol. 279, no. 1, pp. 278–286, 2016, PMID: 26794935. [PubMed: 26794935]
- [13]. Hamilton JI, Jiang Y, Chen Y, Ma D, Lo W-C, Griswold M, and Seiberlich N, "MR fingerprinting for rapid quantification of myocardial T1, T2, and proton spin density," *Magnetic Resonance in Medicine*, vol. 77, no. 4, pp. 1446–1458, 2017. [PubMed: 27038043]
- [14]. McGivney DF, Pierre E, Ma D, Jiang Y, Saybasili H, Gulani V, and Griswold MA, "SVD Compression for Magnetic Resonance Fingerprinting in the Time Domain," *IEEE Transactions on Medical Imaging*, vol. 33, no. 12, pp. 2311–2322, 12 2014. [PubMed: 25029380]
- [15]. Cauley SF, Setsompop K, Ma D, Jiang Y, Ye H, Adalsteinsson E, Griswold MA, and Wald LL, "Fast group matching for mr fingerprinting reconstruction," *Magnetic Resonance in Medicine*, vol. 74, no. 2, pp. 523–528, 2015. [PubMed: 25168690]
- [16]. Yang M, Ma D, Jiang Y, Hamilton J, Seiberlich N, Griswold MA, and McGivney D, "Low rank approximation methods for mr finger-printing with large scale dictionaries," *Magnetic Resonance in Medicine*, vol. 79, no. 4, pp. 2392–2400, 2018. [PubMed: 28804918]
- [17]. Davies M, Puy G, Vanderghenst P, and Wiaux Y, "A Compressed Sensing Framework for Magnetic Resonance Fingerprinting," *SIAM Journal on Imaging Sciences*, vol. 7, no. 4, pp. 2623–2656, 1 2014.
- [18]. Mazor G, Weizman L, Tal A, and Eldar YC, "Low rank magnetic resonance fingerprinting," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) IEEE*, 8 2016, pp. 439–442.
- [19]. Wang Z, Li H, Zhang Q, Yuan J, and Wang X, "Magnetic Resonance Fingerprinting with compressed sensing and distance metric learning," *Neurocomputing*, vol. 174, pp. 560–570, 2016.
- [20]. Pierre EY, Ma D, Chen Y, Badve C, and Griswold MA, "Multiscale reconstruction for MR fingerprinting," *Magnetic Resonance in Medicine*, vol. 75, no. 6, pp. 2481–2492, 6 2016. [PubMed: 26132462]
- [21]. Gómez PA, Ulas C, Sperl JI, Sprenger T, Molina-Romero M, Menzel MI, and Menze BH, "Learning a spatiotemporal dictionary for magnetic resonance fingerprinting with compressed sensing," *Patch-MI*, vol. 9467, pp. 112–119, 2015.
- [22]. Zhao B, Setsompop K, Ye H, Cauley SF, and Wald LL, "Maximum Likelihood Reconstruction for Magnetic Resonance Fingerprinting," *IEEE Transactions on Medical Imaging*, vol. 35, no. 8, pp. 1812–1823, 8 2016. [PubMed: 26915119]
- [23]. Zhao B, "Model-based iterative reconstruction for magnetic resonance fingerprinting," *Proceedings - International Conference on Image Processing, ICIP*, vol. 2015-Decem, pp. 3392–3396, 2015.
- [24]. Wang G, Ye JC, Mueller K, and Fessler JA, "Image Reconstruction is a New Frontier of Machine Learning," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1289–1296, 6 2018. [PubMed: 29870359]
- [25]. Cohen O, Zhu B, and Rosen M, "Deep learning for fast mr fingerprinting reconstruction," in *2017 Scientific Meeting Proceedings. International Society for Magnetic Resonance in Medicine*, 2017, p. 688.
- [26]. Hoppe E, Kördörfer G, Würfl T, Wetzl J, Lugauer F, Pfeuffer J, and Maier A, "Deep Learning for Magnetic Resonance Fingerprinting: A New Approach for Predicting Quantitative Parameter



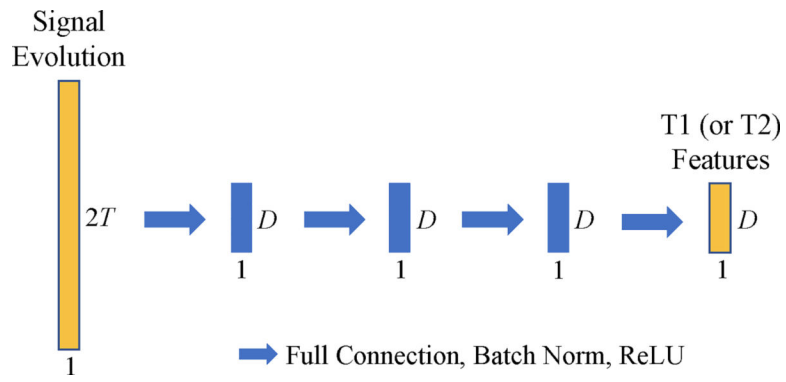
- Values from Time Series.” *Studies in health technology and informatics*, vol. 243, pp. 202–206, 2017. [PubMed: 28883201]
- [27]. Cohen O, Zhu B, and Rosen MS, “Mr fingerprinting deep reconstruction network (drone),” *Magnetic Resonance in Medicine*, vol. 80, no. 3, pp. 885–894, 2018. [PubMed: 29624736]
- [28]. Hoppe E, Kördörfer G, Nittka M, Wür T, Wetzl J, Lugauer F, and Schneider M, “Deep learning for magnetic resonance fingerprinting: Accelerating the reconstruction of quantitative relaxation maps,” in *Proceedings of the 26th Annual Meeting of ISMRM, Paris, France, 2018*.
- [29]. Jiang Y, Ma D, Seiberlich N, Gulani V, and Griswold MA, “MR fingerprinting using fast imaging with steady state precession (FISP) with spiral readout,” *Magnetic Resonance in Medicine*, vol. 74, no. 6, pp. 1621–1631, 2015. [PubMed: 25491018]
- [30]. Fessler JA, “On nufft-based gridding for non-cartesian mri,” *Journal of Magnetic Resonance*, vol. 188, no. 2, pp. 191–195, 2007. [PubMed: 17689121]
- [31]. Zhao B, Setsompop K, Adalsteinsson E, Gagoski B, Ye H, Ma D, Jiang Y, Ellen Grant P, Griswold MA, and Wald LL, “Improved magnetic resonance fingerprinting reconstruction with low-rank and subspace modeling,” *Magnetic Resonance in Medicine*, vol. 79, no. 2, pp. 933–942, 2018. [PubMed: 28411394]
- [32]. Ronneberger O, Fischer P, and Brox T, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” in *Medical Image Computing and Computer-Assisted Intervention*, vol. 9351, 2015, pp. 234–241.
- [33]. Lian C, Zhang J, Liu M, Zong X, Hung SC, Lin W, and Shen D, “Multi-channel multi-scale fully convolutional network for 3D perivascular spaces segmentation in 7T MR images,” *Medical Image Analysis*, vol. 46, pp. 106–117, 2018. [PubMed: 29518675]
- [34]. Kingma DP and Ba JL, “Adam: a Method for Stochastic Optimization,” *International Conference on Learning Representations 2015*, pp. 1–15, 2015.
- [35]. Wang CY, Coppo S, Mehta BB, Seiberlich N, Yu X, and Griswold MA, “Magnetic resonance fingerprinting with quadratic rf phase for measurement of  $t_2^*$  simultaneously with  $\delta_f$   $t_1$ , and  $t_2$ ,” *Magnetic resonance in medicine*, 2018.
- [36]. Chen X, Usman M, Baumgartner CF, Balfour DR, Marsden PK, Reader AJ, Prieto C, and King AP, “High-resolution self-gated dynamic abdominal mri using manifold alignment.” *IEEE Trans. Med. Imaging*, vol. 36, no. 4, pp. 960–971, 2017. [PubMed: 28113339]
- [37]. Kashyap S, Zhang H, Rao K, and Sonka M, “Learning-based cost functions for 3-d and 4-d multi-surface multi-object segmentation of knee mri: Data from the osteoarthritis initiative,” *IEEE transactions on medical imaging*, vol. 37, no. 5, pp. 1103–1113, 2018. [PubMed: 29727274]
- [38]. Miao Z, Wang X-D, Mao L-X, Xia Y-H, Yuan L-J, Cai M, Liu J-Q, Wang B, Yang X, Zhu L, Yu H-B, and Fang B, “Influence of temporomandibular joint disc displacement on mandibular advancement in patients without pre-treatment condylar resorption,” *International journal of oral and maxillofacial surgery*, vol. 46, 09 2016.
- [39]. Miao Z, Wang B, Wu D, Zhang S, Wong S, Shi O, Hu A, Mao L, and Fang B, “Temporomandibular joint positional change accompanies post-surgical mandibular relapse-a long-term retrospective study among patients who underwent mandibular advancement,” *Orthodontics and Craniofacial Research*, vol. 21, 12 2017. [PubMed: 29265578]
- [40]. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, and Bengio Y, “Generative adversarial nets,” in *Advances in neural information processing systems*, 2014, pp. 2672–2680.
- [41]. Isola P, Zhu J-Y, Zhou T, and Efros AA, “Image-to-Image Translation with Conditional Adversarial Networks,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) IEEE*, 7 2017, pp. 5967–5976.
- [42]. Nie D, Trullo R, Petitjean C, Ruan S, and Shen D, “Medical Image Synthesis with Context-Aware Generative Adversarial Networks,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2017, pp. 417–425.
- [43]. Liu H, Xiong R, Zhang X, Zhang Y, Ma S, and Gao W, “Nonlocal Gradient Sparsity Regularization for Image Restoration,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 27, no. 9, pp. 1909–1921, 2017.

- [44]. Liao C, Bilgic B, Manhard MK, Zhao B, Cao X, Zhong J, Wald LL, and Setsompop K, "3D MR fingerprinting with accelerated stack-of-spirals and hybrid sliding-window and GRAPPA reconstruction," *NeuroImage*, vol. 162, pp. 13–22, 11 2017. [PubMed: 28842384]
- [45]. Gómez PA, Buonincontri G, Molina-Romero M, Ulas C, Sperl JI, Menzel MI, and Menze BH, "3d magnetic resonance fingerprinting with a clustered spatiotemporal dictionary," in *Proceedings of the 24th Annual Meeting of ISMRM, Singapore*, 2016.
- [46]. Ma D, Jiang Y, Chen Y, McGivney D, Mehta B, Gulani V, and Griswold M, "Fast 3D magnetic resonance fingerprinting for a whole-brain coverage," *Magnetic Resonance in Medicine*, vol. 79, no. 4, pp. 2190–2197, 4 2018. [PubMed: 28833436]
- [47]. Chen Y, Panda A, Pahwa S, Hamilton JI, Dastmalchian S, McGivney DF, Ma D, Batesole J, Seiberlich N, Griswold MA, Plecha D, and Gulani V, "Three-dimensional mr fingerprinting for quantitative breast imaging," *Radiology*, vol. 290, no. 1, pp. 33–40, 2019, pMID: 30375925. [PubMed: 30375925]
- [48]. Bahrami K, Rekik I, Shi F, and Shen D, "Joint Reconstruction and Segmentation of 7T-like MR Images from 3T MRI Based on Cascaded Convolutional Neural Networks," in *Medical Image Computing and Computer-Assisted Intervention*, 2017, pp. 764–772. [PubMed: 30009281]
- [49]. He K, Zhang X, Ren S, and Sun J, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.
- [50]. Huang G, Liu Z, Maaten L. v. d., and Weinberger KQ, "Densely Connected Convolutional Networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2261–2269.
- [51]. Zhang J, Liu M, and Shen D, "Detecting anatomical landmarks from limited medical imaging data using two-stage task-oriented deep neural networks," *IEEE Transactions on Image Processing*, vol. 26, no. 10, pp. 4753–4764, 2017. [PubMed: 28678706]
- [52]. Lian C, Liu M, Zhang J, and Shen D, "Hierarchical fully convolutional network for joint atrophy localization and Alzheimer's disease diagnosis using structural MRI," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018.
- [53]. Zhou Z, Wang Y, Yu J, Guo W, and Fang Z, "Super-Resolution Reconstruction of Plane-Wave Ultrasound Imaging Based on the Improved CNN Method," in *Lecture Notes in Computational Vision and Biomechanics*, 2018, vol. 27, pp. 111–120.
- [54]. Zhang X, Zhou Z, Chen S, Chen S, Li R, and Hu X, "MR fingerprinting reconstruction with Kalman filter," *Magnetic Resonance Imaging*, vol. 41, pp. 53–62, 2017. [PubMed: 28433503]

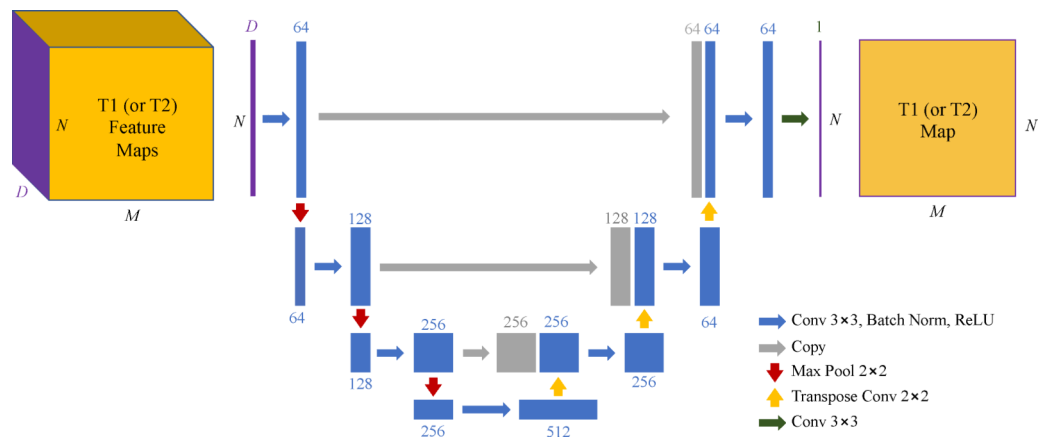


**Fig. 1.**

Schematic overview of our proposed two-step deep learning model for spatially-constrained tissue quantification in MRF. FNN: fully-connected neural network. CNN: convolutional neural network. The number of channels (each with paired FNN and CNN) is equal to the number of tissue properties to be estimated (i.e., 2 in this study).

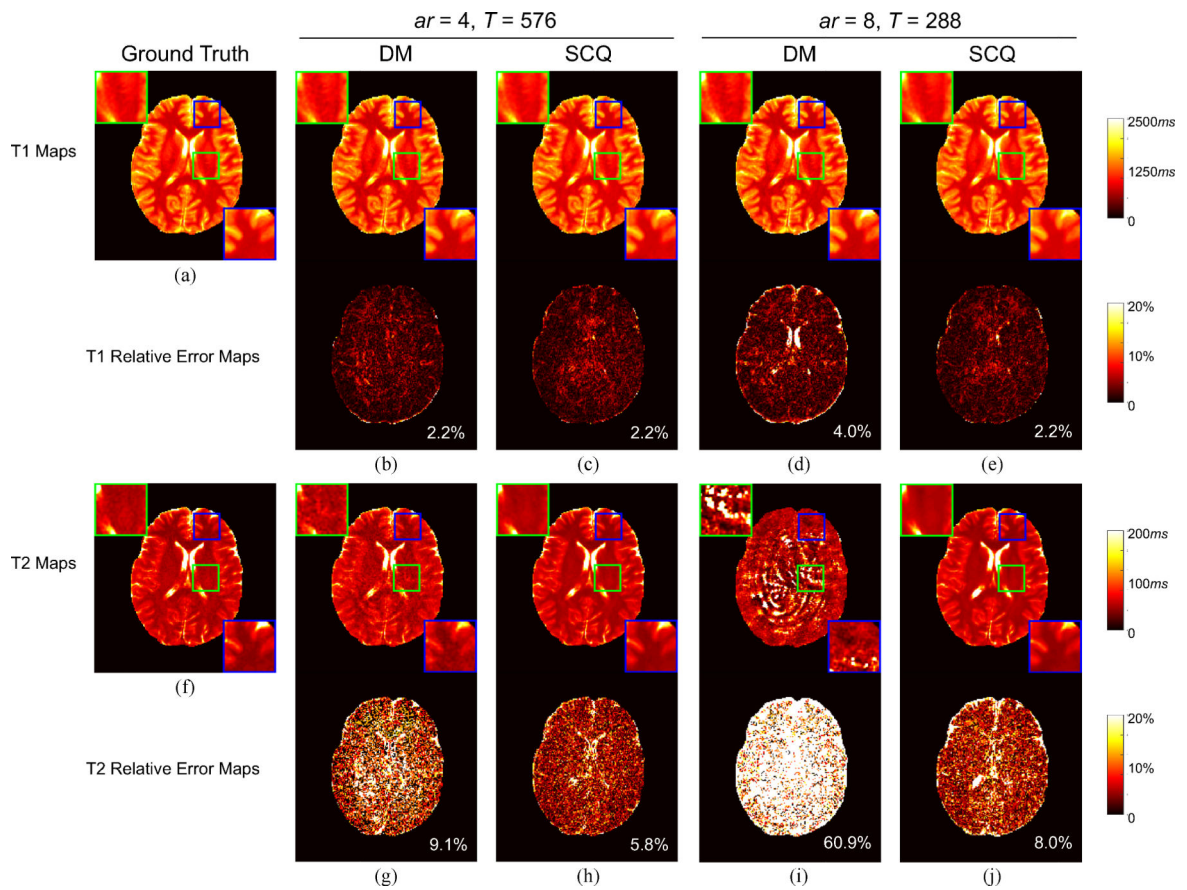


**Fig. 2.** Network structure of the fully-connected neural network (FNN) in our proposed feature extraction (FE) module. Each blue arrow represents a fully-connected (FC) layer, and each blue block represents the output of an FC layer.

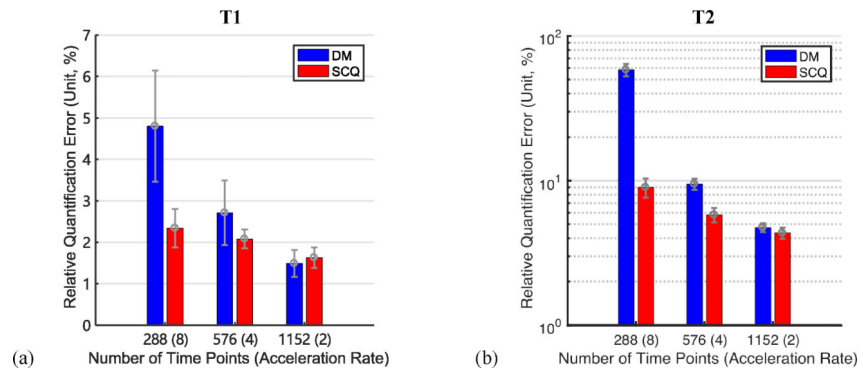


**Fig. 3.**

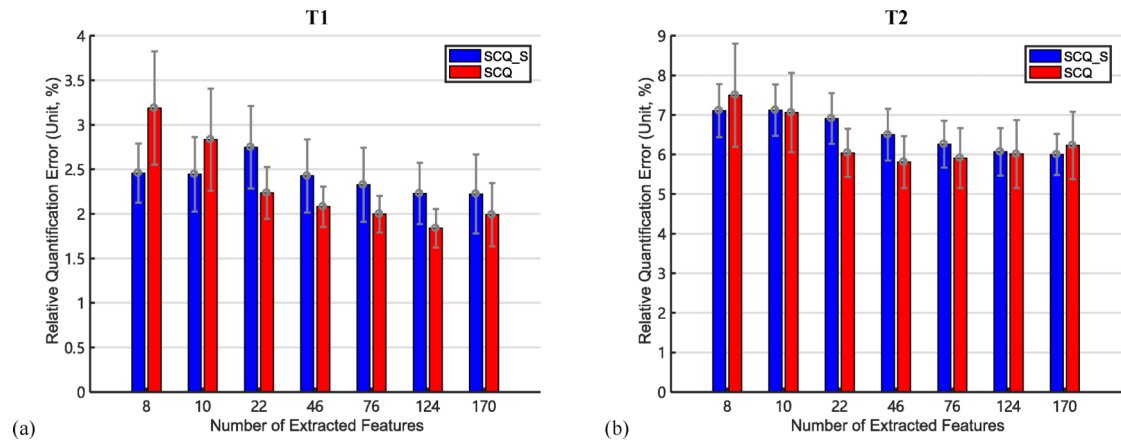
Network structure of the convolutional neural network (CNN) in our proposed spatially-constrained quantification (SQ) module, i.e., U-Net. The blue blocks represent spatial feature maps obtained from spatial convolutions, and gray blocks represent feature maps copied from previous layers. Each feature map has one feature dimension and two spatial dimensions, while only one feature dimension and one spatial dimension are shown for simplicity. The feature dimension is labeled on the top of each block.



**Fig. 4.** Visual results of T1 and T2 quantification by our method (SCQ) and the baseline method (DM). (a) Ground-truth T1 map. (b)-(e): Estimated T1 maps and associated relative error maps. (f) Ground-truth T2 map. (g)-(j): Estimated T2 maps and associated relative error maps. The overall error is labeled at the lower right corner of each error map.  $ar$ : acceleration rate;  $T$ : number of time points used for tissue quantification.



**Fig. 5.** Quantification errors of (a) T1 and (b) T2 (mean  $\pm$  standard deviation, unit: %) yielded by our method (SCQ) and the baseline method (DM) for different acceleration rates.



**Fig. 6.** Quantification errors of (a) T1 and (b) T2 (mean  $\pm$  standard deviation, unit: %) yielded by our method (SCQ) and its variant (SCQ S) which uses SVD (instead of FNN) for feature extraction in the FE module, with different numbers of extracted features. Here,  $ar = 4$ .



Quantification errors of T1 and T2 (mean  $\pm$  standard deviation, unit: %) yielded by our method (SCQ), the baseline method (DM) and the state-of-the-art methods (SDM, CSMR, AND DL) for each test subject. Here,  $nR=4$ .

**TABLE I:**

Subject	DM	SDM	CSMR	DL	SCQ (ours)
1	2.39 $\pm$ 0.24	2.78 $\pm$ 0.29	2.61 $\pm$ 0.29	<b>1.86 <math>\pm</math> 0.18</b>	1.87 $\pm$ 0.23
2	2.40 $\pm$ 0.80	2.89 $\pm$ 0.93	2.65 $\pm$ 0.67	<b>1.78 <math>\pm</math> 0.36</b>	1.83 $\pm$ 0.27
3	2.75 $\pm$ 0.77	3.17 $\pm$ 0.86	2.91 $\pm$ 0.79	<b>2.04 <math>\pm</math> 0.30</b>	2.07 $\pm$ 0.27
4	2.90 $\pm$ 0.74	3.51 $\pm$ 0.97	4.38 $\pm$ 0.97	2.17 $\pm$ 0.46	<b>1.85 <math>\pm</math> 0.28</b>
5	2.71 $\pm$ 0.78	3.15 $\pm$ 0.96	3.34 $\pm$ 0.83	2.12 $\pm$ 0.46	<b>2.08 <math>\pm</math> 0.23</b>
6	2.13 $\pm$ 0.39	2.53 $\pm$ 0.58	2.29 $\pm$ 0.47	1.72 $\pm$ 0.20	<b>1.64 <math>\pm</math> 0.15</b>
Overall	2.55 $\pm$ 0.62	3.00 $\pm$ 0.76	3.03 $\pm$ 0.67	1.95 $\pm$ 0.33	<b>1.89 <math>\pm</math> 0.24</b>
<hr/>					
1	10.06 $\pm$ 0.75	13.99 $\pm$ 1.52	10.40 $\pm$ 0.93	6.46 $\pm$ 0.95	<b>5.73 <math>\pm</math> 0.71</b>
2	8.69 $\pm$ 0.68	11.87 $\pm$ 1.44	8.57 $\pm$ 0.59	6.35 $\pm$ 1.04	<b>5.47 <math>\pm</math> 0.60</b>
3	9.96 $\pm$ 1.52	13.96 $\pm$ 2.23	9.60 $\pm$ 1.43	6.68 $\pm$ 1.13	<b>6.29 <math>\pm</math> 0.89</b>
4	9.81 $\pm$ 0.58	14.16 $\pm$ 1.37	14.89 $\pm$ 4.91	6.65 $\pm$ 0.85	<b>5.88 <math>\pm</math> 0.56</b>
5	9.48 $\pm$ 0.84	13.46 $\pm$ 1.38	9.36 $\pm$ 0.65	7.00 $\pm$ 1.51	<b>5.81 <math>\pm</math> 0.66</b>
6	8.82 $\pm$ 0.87	13.34 $\pm$ 1.41	8.78 $\pm$ 1.57	5.69 $\pm$ 0.77	<b>5.45 <math>\pm</math> 0.70</b>
overall	9.47 $\pm$ 0.87	13.46 $\pm$ 1.56	10.27 $\pm$ 1.68	6.47 $\pm$ 1.04	<b>5.77 <math>\pm</math> 0.69</b>

**TABLE II:**

Average testing time required by our method (SCQ), the baseline method (DM) and the state-of-the-art methods (SDM, CSMR, AND DL) for T1 and T2 quantification for an axial slice with a matrix size of  $256 \times 256$ . Units: second (S) and hour (H). Here,  $ar = 4$ .

DM	SDM	CSMR	DL	SCQ (ours)
26.2s	14.2s	~2h	18.0s	2.3s

Experiments were performed on an iMac (2017) desktop, with CPU: 4.2 GHz Intel Core i7, and memory: 16 GB 2400 MHz DDR4. DM, SDM, DL, and SCQ were implemented in Python, and CSMR was implemented in Matlab R2017a. Testing time is defined as the time for all computations performed using the reconstructed MRF image series for the quantification of both T1 and T2 maps. Specifically, the testing time of SDM includes both signal compression time and matching time.

**TABLE III:**

Quantification errors of T1 and T2 (mean  $\pm$  standard deviation, unit: %) yielded by our method (SCQ) and its variants (SQ-only and FE-only) where only one of the modules exists. Here,  $ar = 4$ .

	<b>SQ-only</b>	<b>FE-only</b>	<b>SCQ (ours)</b>
T1	2.34 $\pm$ 0.40	2.41 $\pm$ 0.49	2.08 $\pm$ 0.23
T2	6.42 $\pm$ 0.64	7.02 $\pm$ 0.78	5.81 $\pm$ 0.66

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**TABLE IV:**

Quantification errors of T1 and T2 (mean  $\pm$  standard deviation, unit: %) yielded by our method (SCQ) which uses both the real and imaginary parts of mrf signals as input, and its variant (SCQ\_M) which uses only the magnitude of signals as input. Here,  $ar = 4$ .

	SCQ_M	SCQ (ours)
T1	10.82 $\pm$ 2.12	2.08 $\pm$ 0.23
T2	13.71 $\pm$ 2.11	5.81 $\pm$ 0.66