Published in final edited form as:

Mach Learn Med Imaging. 2018 September ; 11046: 398–405. doi:10.1007/978-3-030-00919-9_46.

Deep Learning for Fast and Spatially-Constrained Tissue Quantification from Highly-Undersampled Data in Magnetic Resonance Fingerprinting (MRF)

Zhenghan Fang¹, Yong Chen¹, Mingxia Liu¹, Yiqiang Zhan², Weili Lin¹, and Dinggang Shen¹

¹Department of Radiology and BRIC, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

²Institute for Medical Imaging Technology, School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

Abstract

Magnetic resonance fingerprinting (MRF) is a novel quantitative imaging technique that allows simultaneous measurements of multiple important tissue properties in human body, e.g., T1 and T2 relaxation times. While MRF has demonstrated better scan efficiency as compared to conventional quantitative imaging techniques, further acceleration is desired, especially for certain subjects such as infants and young children. However, the conventional MRF framework only uses a simple template matching algorithm to quantify tissue properties, without considering the underlying spatial association among pixels in MRF signals. In this work, we aim to accelerate MRF acquisition by developing a new post-processing method that allows accurate quantification of tissue properties with *fewer* sampling data. Moreover, to improve the accuracy in quantification, the MRF signals from multiple surrounding pixels are used together to better estimate tissue properties at the central target pixel, which was simply done with the signal only from the target pixel in the original template matching method. In particular, a deep learning model, i.e., U-Net, is used to learn the mapping from the MRF signal evolutions to the tissue property map. To further reduce the network size of U-Net, principal component analysis (PCA) is used to reduce the dimensionality of the input signals. Based on *in vivo* brain data, our method can achieve accurate quantification for both T1 and T2 by using only 25% time points, which are *four times* of acceleration in data acquisition compared to the original template matching method.

Keywords

Magnetic resonance fingerprinting; Relaxation times Deep learning

1 Introduction

Quantitative imaging, i.e., quantification of tissue properties in human body, is desired in both clinics and research. The quantified tissue properties allow physicians to better distinguish between healthy and pathological tissues [1] with certain quantitative metrics,

thus making it easier to objectively compare different examinations in longitudinal studies [2]. Also, these quantitative measurements could be more representative of the underlying changes at the cellular level [3, 4], compared to qualitative results obtained from standard MR imaging data. [5]

One of the major barriers of translating conventional quantitative imaging techniques for clinical use is the prohibitively long data acquisition time. Recently, a new framework for MR image acquisition and post-processing, termed as Magnetic Resonance Fingerprinting (MRF) [6], has been introduced, which can significantly reduce the acquisition time needed for quantitative measurement. The MRF framework first acquires a series of highly-undersampled MR images with different contrast weightings using pseudo-randomized acquisition parameters, such as repetition times and flip angles. Note that different tissues or materials have unique signal evolutions that depend on multiple tissue properties. Then, the signal evolution from each pixel is matched to a precomputed dictionary, containing the signal evolutions of a wide range of tissue types. The entry with the best matching signal evolution in the dictionary is selected, with its tissue properties finally assigned for this pixel. The simultaneous measurement of multiple tissue properties and high undersampling rate in the acquisition enable fast quantitative imaging using MRF.

While MRF has demonstrated higher scan efficiency as compared to conventional quantitative imaging techniques, further acceleration is needed as the current scan time is still too long for certain subjects such as infants and young children. In the MRF framework, the acquisition time is approximately proportional to the number of time points used in the imaging sequence, i.e., the number of acquired images with different contrast weightings. In this study, we propose to reduce the acquisition time by reducing the number of time points acquired in each scan. However, the reduction of time points will result in a shorter signal evolution and therefore the loss of useful information at each pixel, which influences the accuracy of tissue quantification.

To compensate for the loss of information at each pixel and improve the quantification accuracy, we propose to use additional information from the space domain, i.e., signal evolutions at neighboring pixels, to assist the estimation at each target pixel. We believe that spatial context information is critical for the following two reasons. First, the tissue properties at different pixels are not independent, but actually correlated. For example, the adjacent pixels of one tissue are likely to have similar tissue properties. Therefore, neighboring pixels could be used together as spatial constraint to regulate the estimation and correct errors at the central target pixel. *Second*, the undersampling in k-space in MRF acquisition results in aliasing in the image space, due to distribution of the target pixel signal to neighboring pixels. Therefore, using spatial information may help retrieve the scattered signals and finally provide a better quantification with MRF.

To achieve the aforementioned spatially-constrained quantification, we resort to a deep learning model, i.e., U-Net [7], to learn the mapping from the MRF signals of a cross-section slice to its tissue property map. Because of using the convolution and down- and up-sampling operations in the U-Net, quantitative tissue properties from one pixel in the output space are spatially correlated with signal evolutions from multiple neighboring pixels in the

input space. In addition, a principal component analysis (PCA) is further performed to reduce the dimensionality of the MRF signals before feeding them into the U-Net, thus significantly reducing network size as well as facilitating network training and finally improving accuracy in tissue quantification.

Note that deep learning has been used for MRF post-processing in previous studies. For example, [8] used a neural network to map the signal evolution at each pixel to the underlying tissue properties. For the same mapping, [9] used a convolutional neural network, where each signal evolution was treated as a time sequence and the convolution was done in time domain. The results of these studies demonstrated potential of deep learning methods for tissue quantification in MRF. However, both studies focused only on improving the quantification speed, instead of accelerating MRF scanning. In addition, these previous studies used *either* fully-sampled MRF images *or* phantom data for method validation. The performance using the actual highly-undersampled *in vivo* data as proposed in the MRF framework needs to be evaluated.

2 Materials and Method

In our spatially-constrained tissue quantification method, the MRF images are fed into a U-Net to first extract spatial features and then estimate the tissue property map. One network is trained for each tissue property under estimation, i.e., T1 or T2. To deal with the large quantitative range of tissue properties in human body, we develop a relative difference based loss function to balance the contributions of tissues with different property ranges. Also, to address the across-subject variation and high-dimensionality of MRF signals, we propose two data pre-processing strategies, i.e., (1) energy-based normalization, and (2) PCA-based compression. In the following, we first introduce the data acquisition and pre-processing methods in Sect. 2.1, and then present our proposed deep learning based method in Sect. 2.2.

2.1 Data Acquisition and Pre-processing

Data Acquisition.—We acquired MRF data of cross-section slices of human brains on a Siemens 3T Prisma scanner using a 32-channel head coil. Highly-undersampled 2D MR images were acquired using fast imaging with steady state precession (FISP) sequence. For each slice, 2,304 time points were acquired and each time point consists of data from only one spiral readout (reduction factor = 48). Other imaging parameters included: field of view (FOV): 30 cm; matrix size: 256 256; slice thickness: 5 mm; flip angle: 5°–12°.

The MRF dictionary was simulated with 13,123 combinations of T1 (60–5000 ms) and T2 (10–500 ms). The reference tissue property maps were obtained by dictionary matching from all 2,304 time points. These maps are considered as “ground truth” in this work.

Energy-based Data Normalization.—Since the magnitude of MRF signals varies largely across different subjects, it is critical to normalize these signals to a common magnitude range in MRF-based learning framework. In this work, we propose to normalize the energy (i.e., sum of squared magnitude) of the acquired signal evolution at each pixel to 1.

PCA-based Compression.—MRF implementation typically acquires a large number of images with different contrast weightings for one subject, i.e., 2,304 in total and 576 after undersampling. In this case, it is unreasonable to feed such high-dimensional data directly into U-Net, as it results in a prohibitively large network size, which is challenging for efficient training and good generalization. Inspired by a previous work [10], we propose to perform PCA for dimensionality reduction of the signal evolutions after normalization. Specifically, the bases of the principal component subspace are obtained from signal evolutions in the dictionary. The k largest eigenvalues are retained with a threshold energy E (i.e., $E = 99.9\%$ and $k = 17$ in our study). As a result, the input channels of U-Net are compressed to $2k$, with each channel representing the real or imaginary part of the coefficient along one basis.

2.2 Proposed U-Net Model

In conventional template matching method, only the signal evolution from a particular local pixel is used to estimate tissue properties at the corresponding pixel in the tissue property map, without considering the global context information (e.g., spatial association among pixels) of the input MRF images. Actually, for each pixel, signals from its neighboring pixels may also provide important *spatial constraints* for estimating the tissue properties. Therefore, in this work, we resort to U-Net to capture both the local and global information of MRF images, with the architecture shown in Fig. 1.

As shown in Fig. 1, this network consists of an encoder sub-network (i.e., left part of Fig. 1) that extracts multi-scale spatial features from the input MRF images, and a successive decoder sub-network (i.e., right part of Fig. 1) that uses the extracted features to generate the output tissue property (T1 or T2) map. During alternate feature extraction (3 3 convolution followed by ReLU activation) and down-sampling (2 2 max pooling) operations in the encoder sub-network, the information from distributed signals due to aliasing in MRF images is retrieved and summarized. Also, the spatial constraints among different pixels are now implicitly incorporated into the extracted feature maps. Then, the decoder sub-network expands (with the transpose convolution 2 2) the down-sampled feature maps and combines feature maps at different scales by copying and concatenating, to fuse global context knowledge with complementary local details, for accurate and spatially-consistent estimation of the tissue property map.

It is worth noting that T1 and T2 measures in human body have very large quantitative ranges. As a result, the loss function based on the conventional absolute difference between network estimation and ground-truth will be dominated by the tissues with high T1 or T2 values. To address this issue, we propose to use the *relative difference*, rather than absolute difference, in our loss function, as defined below:

$$L = \sum_{x \in \Omega} \left| \frac{\theta_x - \hat{\theta}_x}{\theta_x} \right| \quad (1)$$

where Ω is the set of all pixels inside the brain region of a training slice, $\hat{\theta}_x$ is the network output at pixel x , and θ_x is the corresponding ground-truth property. With Eq. (1), the loss function is balanced over tissues with different property ranges.

3 Experiments

3.1 Experimental Settings

Our dataset includes axial cross-section brain slices collected from 5 subjects, with 12 slices for each subject. In our experiments, the slices of 4 randomly selected subjects were used as the training data, and the slices of the remaining 1 subject were used as the testing data to validate effectiveness of the proposed method. To assess the estimation accuracy, similar to Eq. (1), the pixel-wise relative error at pixel x was computed as $e_x = |\theta_x - \hat{\theta}_x|/\theta_x$, where θ_x and $\hat{\theta}_x$ denote the reference and estimated tissue properties at pixel x . The average relative error over all pixels in each testing slice and the mean and standard deviation of the errors over all testing slices were also computed.

We first compared our proposed Spatially-constrained Tissue Quantification method (denoted as **STQ**) with the standard Dictionary Matching approach (denoted as **DM**). It is worth noting that there are three components in our method, including (1) energy-based data normalization, (2) relative difference based loss in Eq. (1), and (3) PCA-based compression. To evaluate the influence of each component, we further compared our proposed STQ method with its three variants: (1) **STQ** without data normalization (denoted as **STQ_N**), where the acquired MRF signals were directly used for the following PCA-based compression without energy normalization; (2) **STQ** using conventional absolute difference as loss function (denoted as **STQ_A**); and (3) **STQ** without using PCA-based compression (denoted as **STQ_P**), where the high-dimensional MRF signals were directly fed into U-Net without dimensionality reduction. For fair comparison, all these five competing methods were used to estimate tissue properties from the first 576 time points among all 2,304 time points (i.e., with the undersampling rate of 25%, or reducing the MRF acquisition time by 4 times).

3.2 Results

Comparison with Dictionary Matching.—The tissue property maps of a testing slice obtained by DM and STQ, along with their error maps, are shown in Fig. 2. The estimation errors over the testing set achieved by DM are 19.4% (T1) and 12.3% (T2), and 5.6% (T1) and 8.8% (T2) by STQ, as shown in Table 1. We can see that our STQ method achieves 71.1% (T1) and 28.5% (T2) reductions in estimation errors compared to DM. Moreover, as shown by the visual results in Fig. 2, the quantitative maps of STQ have much less noise than those of DM, especially for T2 estimation, suggesting that the spatial constraint from neighboring pixels can help correct estimation errors and achieve more spatially-consistent quantification results. It is worth noting that the spatial resolution is not deteriorated by the denoising effect of STQ, i.e., the boundaries between different tissues are well-preserved (see the blue zoom-in boxes in Fig. 2). This advantage of our method is particularly important for the subsequent statistical analysis procedures such as segmentation of lesions.

Comparison with Three STQ Variants.—We further compare our proposed STQ method with its three variants, with numerical results shown in Table 1. We can observe that our STQ method obtains the best results for both T1 and T2 estimations, compared with its three variants (i.e., STQ_N, STQ_A and STQ_P). These results clearly demonstrate the merits of our three proposed strategies, i.e., energy-based data normalization, relative difference based loss function, and PCA-based compression. It is worth noting that the training time of STQ is much shorter than that of STQ_P, due to the significantly reduced network size by PCA-based representation of MRF signals. This suggests that our proposed PCA-based compression strategy can improve *not only* the generalization of the network *but also* the efficiency of network training.

Influence of Parameters.—There are two important parameters in our proposed STQ method, i.e., (1) the energy threshold for PCA eigenvector selection, and (2) data acquisition time (corresponding to the undersampling rate). Now we investigate the influences of these two parameters on the performance of our method, with results reported in Fig. 3(a) and (b), respectively. From Fig. 3(a), we can see that the estimation error yielded by STQ decreases quickly when the energy threshold is increased from 90 to 99.9%, but does not change much when it is further increased to 99.99%. Also, we change the acquisition time by extracting tissue properties from different numbers of time points, i.e., 288 (12.5%), 576 (25%), 1152 (50%), 2304 (100%). As shown in Fig. 3(b), there is a trade-off between acquisition time and quantification accuracy, and STQ can achieve good estimation accuracy using 100%, 50% and 25% acquisition time. When the acquisition time is further reduced to 12.5%, the estimation error of STQ for T1 is still acceptable, but becomes much worse for T2 (i.e., relative error > 13%).

4 Conclusion

In this study, we have proposed a new MRF post-processing method for accurate tissue quantification from highly-undersampled data. Unlike the conventional method that performs estimation for each pixel separately, our method uses the signals at multiple pixels around the target pixel to jointly estimate its tissue property, resulting in more accurate and spatially-consistent quantification results. A deep learning model, i.e., U-Net, is used to learn the mapping from MRF signals to the desired tissue property map. The experimental results from *in vivo* brain data show that our method achieves good accuracy in T1 and T2 quantification using only 25% of time points (i.e., four times of reduction in acquisition time), and 71.1% (T1) and 28.5% (T2) reductions in estimation errors compared to the conventional dictionary matching method.

References

1. Larsson HBW, et al.: Assessment of demyelination, edema, and gliosis by *in vivo* determination of T1 and T2 in the brain of patients with acute attack of multiple sclerosis. *Magn. Reson. Med* 11(3), 337–348 (1989) [PubMed: 2779421]
2. Usman AA, et al.: Cardiac magnetic resonance T2 mapping in the monitoring and follow-up of acute cardiac transplant rejection clinical perspective: a pilot study. *Circ. Cardiovasc. Imaging* 5(6), 782–790 (2012) [PubMed: 23071145]

3. Payne AR, et al.: Bright-blood T2-weighted MRI has high diagnostic accuracy for myocardial hemorrhage in myocardial infarction clinical perspective: a preclinical validation study in swine. *Circ. Cardiovasc. Imaging* 4(6), 738–745 (2011) [PubMed: 21930836]
4. Van Heeswijk RB, et al.: Free-breathing 3T magnetic resonance T2-mapping of the heart. *JACC: Cardiovasc. Imaging* 5(12), 1231–1239 (2012)
5. Coppo S, et al.: Overview of magnetic resonance fingerprinting. In: *MAGNETOM Flash*, 65 (2016)
6. Ma D, et al.: Magnetic resonance fingerprinting. *Nature* 495, 187–192 (2013) [PubMed: 23486058]
7. Ronneberger O, et al.: U-net: convolutional networks for biomedical image segmentation In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, Cham (2015)
8. Cohen O, et al.: Deep learning for rapid sparse MR fingerprinting reconstruction. *arXiv preprint [arXiv:1710.05267](https://arxiv.org/abs/1710.05267)* (2017)
9. Hoppe E, et al.: Deep learning for magnetic resonance fingerprinting: a new approach for predicting quantitative parameter values from time series. *Stud. Health Technol. Inf* 243, 202–206 (2017)
10. McGivney DF, et al.: SVD compression for magnetic resonance fingerprinting in the time domain. *IEEE Trans. Med. Imaging* 33(12), 2311 (2014) [PubMed: 25029380]

Fig. 1. U-net architecture. Each blue block represents a multi-channel feature map, and gray blocks represent copied feature maps. The number of channels is denoted at the top or bottom of each block. The arrows denote different operations.

Fig. 2.

T1 and T2 estimation results for a testing slice obtained by DM and STQ (our proposed method). (a) Reference T1 map. (b), (c) Estimated T1 maps and their corresponding relative error maps. (d) Reference T2 map. (e), (f) Estimated T2 maps and their corresponding relative error maps. For each error map, the mean relative error over brain area is shown at the lower-right corner.

Fig. 3.

The performance of the proposed method using (a) different energy thresholds for PCA eigenvector selection (with the x-axis representing the energy threshold), and (b) different acquisition times (with the x-axis representing the number of time points).

Table 1.

The means and standard deviations of T1 and T2 estimation errors achieved by DM, STQ (our proposed method) and three variants of STQ (unit: %).

	DM	STQ	STQ_N	STQ_A	STQ_P
T1	19.4 ± 8.8	5.6 ± 1.8	7.7 ± 1.9	7.9 ± 3.7	6.9 ± 2.0
T2	12.3 ± 3.0	8.8 ± 1.9	10.8 ± 1.7	10.6 ± 3.2	9.0 ± 2.3