



Magnetic Resonance Fingerprinting Reconstruction via Spatiotemporal Convolutional Neural Networks

Fabian Balsiger¹(✉) , Amaresha Shridhar Konar², Shivaprasad Chikop², Vimal Chandran¹, Olivier Scheidegger³, Sairam Geethanath², and Mauricio Reyes¹

¹ Institute for Surgical Technology and Biomechanics, University of Bern, Bern, Switzerland

`fabian.balsiger@istb.unibe.ch`

² Medical Imaging Research Center, Dayananda Sagar Institutions, Bangalore, India

³ Support Center for Advanced Neuroimaging (SCAN),

Institute for Diagnostic and Interventional Neuroradiology, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland

Abstract. Magnetic resonance fingerprinting (MRF) quantifies multiple nuclear magnetic resonance parameters in a single and fast acquisition. Standard MRF reconstructs parametric maps using dictionary matching, which lacks scalability due to computational inefficiency. We propose to perform MRF map reconstruction using a spatiotemporal convolutional neural network, which exploits the relationship between neighboring MRF signal evolutions to replace the dictionary matching. We evaluate our method on multiparametric brain scans and compare it to three recent MRF reconstruction approaches. Our method achieves state-of-the-art reconstruction accuracy and yields qualitatively more appealing maps compared to other reconstruction methods. In addition, the reconstruction time is significantly reduced compared to a dictionary-based approach.

Keywords: Magnetic resonance fingerprinting · Parameter mapping
Image reconstruction · Convolutional neural network

1 Introduction

Magnetic resonance imaging (MRI) is widely used in healthcare centers for the diagnosis of pathologies. The diagnosis from MRI relies mostly on weighted images, where the contrast between tissues is used to identify pathologies rather than the absolute intensities in the images. This qualitative approach limits the objective evaluation and reproducibility of MRI in the clinics. Although significant effort has been made for quantitative MRI, a clinical relevant solution for nuclear magnetic resonance (NMR) parameter mapping has not been achieved so far. Mainly time-inefficiency and the limitation to one NMR parameter at

interest (e.g. T1 and T2 relaxation times) make quantitative MRI inappropriate for clinical use. To overcome the drawbacks of quantitative MRI, magnetic resonance fingerprinting (MRF) has been proposed recently as a novel quantitative MRI technique [6]. MRF quantifies multiple NMR parameters in a single and fast acquisition. The acquisition relies on a MR sequence with pseudo-randomly varying parameters to obtain a unique signal evolution, i.e. fingerprint, per tissue and voxel. After the acquisition, a dictionary matching algorithm assigns the voxel’s signal evolutions to an entry of a dictionary of simulated and pre-computed signal evolutions, which allows reconstructing quantitative maps of NMR parameters of interest. However, this dictionary matching is time-consuming, lacks scalability, and can introduce artefacts due to the under-sampled k -space during the acquisition [8].

Recently, three approaches have been proposed aiming to overcome the issues associated with dictionary matching during the MRF reconstruction. Gómez et al. [4] proposed a spatiotemporal dictionary matching that matches a spatial neighborhood of fingerprints instead of using a fingerprint-wise approach. They additionally improve the computational efficiency by limiting the matching to a local search window. However, the search window comes at the cost of requiring spatially aligned MRF scans, and ultimately only alleviates the problem of scalability of dictionary-based MRF reconstruction methods. Therefore, approaches replacing the dictionary matching using deep learning have been proposed to overcome the bottleneck of scalability. Cohen et al. [3] proposed a fully-connected neural network and Hoppe et al. [5] proposed a convolutional neural network (CNN) to learn the matching of a MRF signal evolution to NMR properties. Both approaches show promising results regarding reconstruction accuracy and speed, and their concepts might be a feasible way to replace the dictionary matching involved in MRF reconstruction. However, they use a fingerprint-wise approach, i.e. do not consider any spatial characteristics during the reconstruction, which might result in noisy reconstructions. Moreover, all three approaches use maps reconstructed by the standard dictionary matching with simulated entries as ground truth to compare their reconstructed maps. This ultimately adds a bias to the methods, which resemble the dictionary matching instead of learning the underlying relation of the fingerprints to the NMR parameter maps.

We propose a MRF reconstruction approach that exploits the spatiotemporal relationship between neighboring signal evolutions motivated by noisy reconstructions of fingerprint-wise approaches and the findings of [4]. Our approach bases on CNNs and yields fast and more accurate reconstructions than recently proposed methods on six healthy brain MRF images with three NMR maps: proton density (PD), T1 relaxation time (T1), and T2 relaxation time (T2). Unlike previously published methods, we rely on parametric maps acquired through MR parameter mapping as ground truth instead of reconstructed maps by dictionary matching. We compare our performance to the aforementioned spatiotemporal dictionary- and deep learning-based methods. We report quantitative and qualitative results and discuss open issues and challenges towards a relevant solution for accurate and fast MRF reconstruction.

2 Materials and Methods

We consider a four-dimensional (4-D) MRF image $I \in \mathbb{C}^{X \times Y \times Z \times T}$, where each voxel $I(\mathbf{v}) = \{t_1, t_2, \dots, t_T\}$ at location $\mathbf{v} = (x, y, z)$ contains a MRF signal evolution, or fingerprint, with T temporal signal intensities t_i . For each MRF image I , a set $Q = \{q_1, q_2, \dots, q_M\} \in \mathbb{R}^{X \times Y \times Z \times M}$ with M parametric maps are available as ground truth for the reconstruction. In this work, six brain MRF images with $M = 3$ parametric maps $Q = \{\text{PD}, \text{T1}, \text{T2}\}$ were used.

2.1 MRF and Parametric Map Acquisition

We acquired brain scans from six healthy male volunteers (21 to 43 years) using a tailored MRF sequence [7] on a 1.5 Tesla GE SIGNA Artist scanner (GE Medical Systems, Milwaukee, WI, U.S.) with a 16-channel head coil as part of an institution approved study. Each scan consisted of $Z = 16$ axial-oriented slices with a matrix size of $X \times Y = 256 \times 256$, field of view (FOV) of $256 \times 256 \text{ mm}^2$, voxel size of $1.0 \times 1.0 \times 5.0 \text{ mm}^3$, and a total of 720 temporal images per slice. After the acquisition, the images were pre-processed using a sliding-window reconstruction [1] with a window size of 48 resulting in $T = 673$ temporal images.

The parametric maps serving as ground truth for the MRF reconstruction were acquired with the same number of slices, matrix size, FOV, and voxel size. The T1 and T2 maps were generated using curve fitting of the MR signal of multi-FA and multi-echo sequences, respectively. Seven T1-weighted images were acquired with a gradient recalled echo pulse sequence with FAs of $1^\circ, 2^\circ, 5^\circ, 8^\circ, 11^\circ, 14^\circ$, and 25° , and constant TR/TE = 5.85/1.77 ms. A fast spin echo sequence with eight TEs starting from 20 ms at an interval of 20 ms was used to generate the T2 map (FA = 90° and TR = 1626 ms). By using a signal intensity equation, the PD maps were generated from the T1-weighted images acquired for the T1 mapping.

2.2 Spatiotemporal CNN MRF Reconstruction

We propose a CNN to learn spatiotemporal features to reconstruct the maps Q from a MRF image I . Input to the CNN are MRF image patches $I_P(\mathbf{v}) \in \mathbb{C}^{5 \times 5 \times T} \subset I$, centered at location \mathbf{v} . Output of the CNN are the values of the estimated maps $\hat{Q}(\mathbf{v}) \in \mathbb{R}^M$ at location \mathbf{v} . The CNN is trained to learn the mapping $\mathcal{M} : I_P(\mathbf{v}) \rightarrow Q(\mathbf{v})$. We remark that the reconstruction was performed slice-wise due to the large slice spacing of 5.0 mm of our data. Figure 1 provides an overview of how the input and output data are defined for the proposed multiparametric spatiotemporal MRF reconstruction.

Pre-processing. We first apply a brain mask to the MRF images and the corresponding maps to exclude the background in all experiments. The masks were manually segmented using the T1 map with the polygon tool in ITK-SNAP

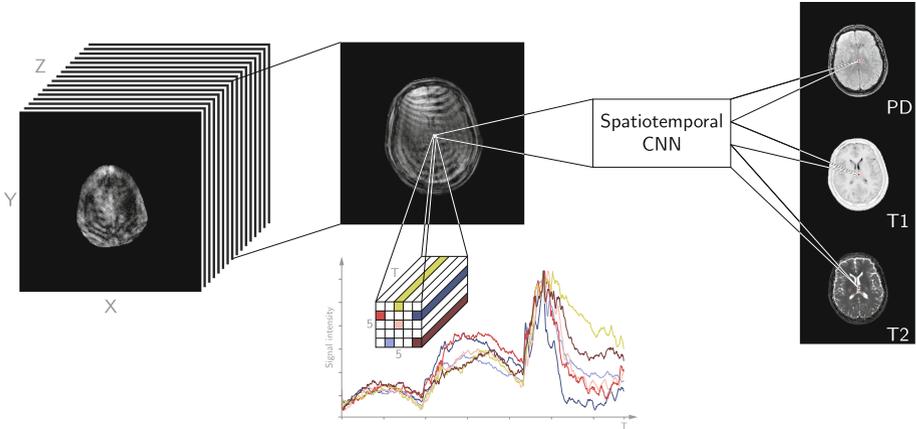


Fig. 1. Overview of the proposed spatiotemporal MRF reconstruction. Note that the signal evolutions are complex-valued but the absolute numbers are plotted for simplicity.

(www.itksnap.org). Outliers from each map in Q are removed by clipping the values to the percentiles $[0.1, 99.9]$. Finally, we normalize I along the temporal axis T to have zero mean and unit variance, and Q along the temporal axis M to the range $[0, 1]$. Note that within each subject the maps were spatially aligned and therefore no registration was applied.

Architecture. Our network consists of five convolutional layers, which learn the mapping \mathcal{M} , i.e. we predict the M map values at location \mathbf{v} from an MRF image patch I_P (Fig. 2). We first concatenate the real and imaginary part of the complex-valued input $I_P(\mathbf{v}) \in \mathbb{C}^{5 \times 5 \times T}$ to a real-valued input $I_P(\mathbf{v}) \in \mathbb{R}^{5 \times 5 \times 2T}$ and consider the temporal dimension ($2T$) as the channels in our network. Second, we apply two 1×1 convolutional blocks to reduce the number of channels to 256. Subsequently, we apply two convolutional blocks in parallel with different receptive fields of 5×5 and 3×3 motivated by [2]. The output channels of these two convolutional blocks are concatenated and fed into the last convolutional layer with M output channels corresponding to the values of the estimated maps \hat{Q} . A convolutional block in our network consists of a sequence of 2-D convolutional layer (valid padding and stride one), dropout, batch normalization, and rectified linear unit (ReLU) activation function. We maintain a linear activation at the last convolutional layer (valid padding and stride one). The estimated maps are denormalized to get quantitative values in the range prior to the pre-processing. We implemented our network using the open source machine learning library TensorFlow 1.8.0 (Google, Mountain View, CA, U.S.) with Python 3.6 (Python Software Foundation, Wilmington, DE, U.S.).

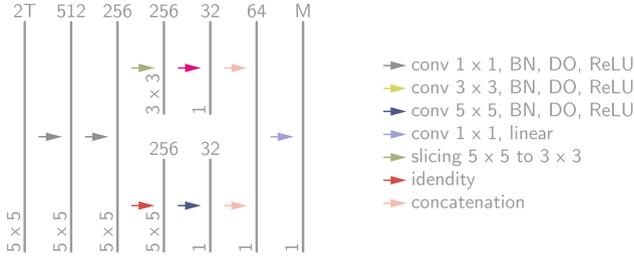


Fig. 2. The architecture of our spatiotemporal CNN. To perform a convolution with a filter size of 3×3 , we extract a 3×3 patch from a 5×5 patch denoted as slicing operation. The number of channels are denoted on the top of the bars and the $x \times y$ size is provided at the lower left edge of the bars. BN: batch normalization, DO: dropout.

Training. The network was trained using an Adam optimizer with a learning rate of 0.001, which minimized a mean squared error (MSE) loss with a batch size of 600 randomly sampled patches I_P . The dropout rate was set to 0.2 and the training was stopped after 50 epochs, which we empirically found to be sufficient.

2.3 Evaluation

We evaluate the performance our model and the baselines using a leave-one-out cross-validation, i.e. we train the model on five brain scans and test it on the left-out brain scan. Note that we tuned the architecture on one randomly chosen cross-validation split and did not use the other splits to develop and tune the architecture.

Baselines. We compare our method to recent approaches for MRF reconstruction using the fully-connected neural network [3], the CNN [5], and the spatiotemporal dictionary matching [4]. For the deep learning-based methods, we performed the same leave-one-out cross-validation and the data underwent the same pre-processing as for our method. The approaches were implemented as proposed in the papers. For [4], we also perform a leave-one-out cross-validation, i.e. construct a dictionary using five brain scans and reconstruct the left-out brain scan with following parameters: $W_n = 11 \times 11 \times 3$, $P = 3 \times 3 \times 3$, $C = 5$, and $\alpha = 0.5$ with two iterations.

Metrics. Quantitatively, we report the mean and standard deviation of the mean absolute difference (MAE) and the root mean square error (RMSE) for the leave-one-out cross-validation. The metrics are reported separately for the three brain tissues white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF). The brain tissue masks were obtained from the T1 maps using thresholding according to literature values [6].

3 Results

Mean and standard deviation of the MAE and RMSE for the PD, T1, and T2 map reconstructions are given in Table 1. The proposed method outperforms the other methods for most brain tissues and maps. Reconstructed maps of a mid-brain slice are shown in Fig. 3. Qualitatively, our maps show a good reconstruction with visible brain structures like the ventricles. It is noticeable that our spatiotemporal approach yields a less noisy reconstruction than the fingerprint-wise approach of Cohen et al. [3] (similar noisy reconstructions were obtained for Hoppe et al. [5] but not reported in Fig. 3). The dictionary-based approach [4] yields a qualitatively coarser reconstruction than our method. Overall, large reconstruction errors are mainly present at the skull, meninges, ventricles as well as at the boundary of the brain mask (rightmost column in Fig. 3), which could be consistently observed for all methods.

Table 1. Mean absolute error (MAE) and root mean square error (RMSE) for the PD, T1, and T2 map reconstructions separated by the brain tissues white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF).

Tissue	Method	PD		T1 (ms)		T2 (ms)	
		MAE	RMSE	MAE	RMSE	MAE	RMSE
WM	Cohen	0.084 ± 0.030	0.107 ± 0.032	209.0 ± 28.3	267.5 ± 21.5	43.6 ± 23.9	77.3 ± 54.7
	Hoppe	0.080 ± 0.030	0.101 ± 0.031	253.9 ± 64.3	317.7 ± 67.2	61.3 ± 36.3	92.0 ± 53.8
	Gómez	0.058 ± 0.015	0.074 ± 0.021	258.6 ± 61.0	327.2 ± 68.5	33.4 ± 20.4	73.4 ± 50.6
	Proposed	0.055 ± 0.015	0.072 ± 0.016	159.4 ± 36.3	242.7 ± 54.7	28.0 ± 17.6	71.1 ± 62.0
GM	Cohen	0.094 ± 0.028	0.121 ± 0.026	197.0 ± 28.4	258.0 ± 42.0	57.6 ± 35.4	97.3 ± 65.9
	Hoppe	0.092 ± 0.030	0.119 ± 0.029	218.5 ± 37.6	287.5 ± 42.5	70.2 ± 41.0	105.1 ± 61.7
	Gómez	0.060 ± 0.017	0.081 ± 0.021	190.8 ± 24.7	269.1 ± 43.1	45.8 ± 26.9	90.5 ± 59.6
	Proposed	0.061 ± 0.017	0.077 ± 0.020	208.2 ± 34.1	286.6 ± 46.5	43.2 ± 31.2	93.6 ± 76.3
CSF	Cohen	0.126 ± 0.024	0.152 ± 0.025	1162.6 ± 256.2	1364.1 ± 265.1	183.3 ± 54.1	237.0 ± 58.0
	Hoppe	0.128 ± 0.013	0.156 ± 0.014	1013.3 ± 236.0	1219.4 ± 251.5	174.7 ± 49.3	227.8 ± 55.5
	Gómez	0.102 ± 0.020	0.129 ± 0.019	1072.5 ± 164.5	1268.8 ± 172.6	228.6 ± 85.7	286.8 ± 86.0
	Proposed	0.093 ± 0.013	0.113 ± 0.009	989.2 ± 254.7	1181.5 ± 288.6	181.6 ± 48.6	240.2 ± 48.7

4 Discussion and Conclusion

We presented a deep learning-based, dictionary-free approach to reconstruct parametric maps from MRF images that exploits the spatiotemporal relationship between neighboring fingerprints. The approach is designed as CNN that yields a reconstruction of parametric maps in a more accurate way than previously proposed dictionary-free methods and competes with a dictionary-based method.

In general, the results show that a spatiotemporal reconstruction is favorable to a fingerprint-wise reconstruction for almost all brain tissues and parametric maps (Table 1). Out of the three brain tissues, the GM yielded the most inconsistent results among the different methods. We think that this might arise due to partial volume effects at the interface to WM and CSF. A spatial analysis reveals high reconstruction errors in the skull, meningeal layers, and ventricles for all

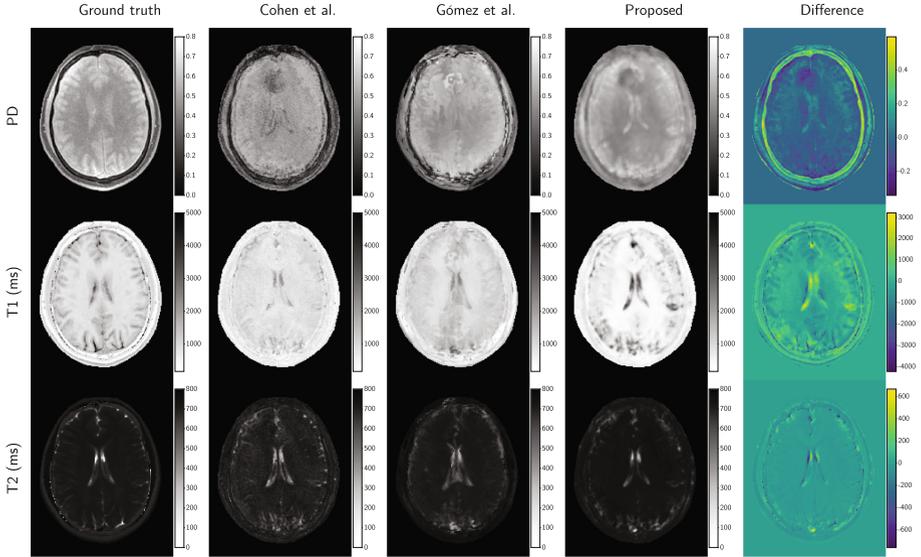


Fig. 3. Exemplary map reconstructions of one axial brain slice. The rows represent the three maps PD, T1, and T2. The columns represent from left to right: the ground truth map, results of Cohen et al. [3], Gómez et al. [4], and our proposed method. The rightmost column shows the difference $\hat{Q} - Q$ between our estimated map \hat{Q} and the ground truth map Q .

methods (rightmost column in Fig. 3). These reconstruction errors could originate from a partial volume effect or an apparent lack of training examples. Reconstruction artifacts are only present in the method of Gómez et al. [4], confirming the findings of [8]. In regards to the computational costs, all deep learning-based approaches yield reconstructed maps within few seconds. Conversely, the dictionary-based approach is computationally intensive, with calculations in the order of several minutes per reconstruction. For clinically used MRF reconstruction, we therefore think that machine learning-based approaches are favorable to dictionary-based approaches in the long term.

Our study design included parametric maps acquired through MR parameter mapping as ground truth. The reconstruction errors of all methods are large compared to the errors reported in the studies of the baselines [3–5]. Such large errors are especially surprising for the dictionary-based method, which can be interpreted as a k-nearest neighbor search. Unfortunately, all baselines compared their performance with a ground truth obtained by dictionary matching as proposed in the original MRF paper [6]. Therefore, the methods resembled the dictionary matching instead of learning the underlying relation between fingerprints and NMR maps. We think that a comparison to acquired NMR maps is ultimately more meaningful than a comparison with maps reconstructed from simulated dictionaries. Our results suggest that the direct learning of fingerprints to acquired NMR maps is possible, although additional investigations and work are needed to lower the reconstruction errors.

In conclusion, we demonstrated that a spatiotemporal MRF reconstruction is favorable to a fingerprint-wise MRF reconstruction designed within a CNN by achieving quantitatively and qualitatively better parametric map reconstructions.

Acknowledgements. This research was supported by the Swiss Foundation for Research on Muscle Diseases (ssem), grant attributed to author OS. The authors thank the NVIDIA Corporation for their GPU donation.

References

1. Cao, X., et al.: Robust sliding-window reconstruction for accelerating the acquisition of MR fingerprinting. *Magn. Reson. Med.* **78**(4), 1579–1588 (2017). <https://doi.org/10.1002/mrm.26521>
2. Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L.: DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Trans. Pattern Anal. Mach. Intell.* **40**(4), 834–848 (2016). <https://doi.org/10.1109/TPAMI.2017.2699184>
3. Cohen, O., Zhu, B., Rosen, M.S.: MR fingerprinting deep reconstruction network (DRONE). *Magnetic Reson. Med.* **80**(3), 885–894 (2018). <https://doi.org/10.1002/mrm.27198>
4. Gómez, P.A., et al.: Simultaneous parameter mapping, modality synthesis, and anatomical labeling of the brain with MR fingerprinting. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (eds.) *MICCAI 2016*. LNCS, vol. 9902, pp. 579–586. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46726-9_67
5. Hoppe, E., et al.: Deep learning for magnetic resonance fingerprinting: a new approach for predicting quantitative parameter values from time series. In: Röhrig, R., Timmer, A., Binder, H., Sax, U. (eds.) *German Medical Data Sciences: Visions and Bridges*, Oldenburg, vol. 243, pp. 202–206 (2017). <https://doi.org/10.3233/978-1-61499-808-2-202>
6. Ma, D., et al.: Magnetic resonance fingerprinting. *Nature* **495**(7440), 187–192 (2013). <https://doi.org/10.1038/nature11971>
7. Shaik, I., et al.: Tailored magnetic resonance fingerprinting: optimizing acquisition schedule and intelligent reconstruction using a block approach. In: *ISMRM 2018* (2018)
8. Wang, Z., Zhang, Q., Yuan, J., Wang, X.: MRF denoising with compressed sensing and adaptive filtering. In: *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*, pp. 870–873. IEEE, Beijing (2014). <https://doi.org/10.1109/ISBI.2014.6868009>