German Medical Data Sciences: Visions and Bridges
R. Röhrig et al. (Eds.)

doi:10.3233/978-1-61499-808-2-202

Deep Learning for Magnetic Resonance Fingerprinting: A New Approach for Predicting Quantitative Parameter Values from Time Series

Elisabeth HOPPE[a,b,1], Gregor KÖRZDÖRFER[a,c], Tobias WÜRFL[b], Jens WETZL[b], Felix LUGAUER[b], Josef PFEUFFER[a] and Andreas MAIER[b]
[a] *MR Application Development, Siemens Healthcare, Erlangen, Germany*
[b] *Pattern Recognition Lab, Department of Computer Science, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany*
[c] *Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany*

Abstract. The purpose of this work is to evaluate methods from deep learning for application to Magnetic Resonance Fingerprinting (MRF). MRF is a recently proposed measurement technique for generating quantitative parameter maps. In MRF a non-steady state signal is generated by a pseudo-random excitation pattern. A comparison of the measured signal in each voxel with the physical model yields quantitative parameter maps. Currently, the comparison is done by matching a dictionary of simulated signals to the acquired signals. To accelerate the computation of quantitative maps we train a Convolutional Neural Network (CNN) on simulated dictionary data. As a proof of principle we show that the neural network implicitly encodes the dictionary and can replace the matching process.

Keywords. Convolutional Neural Networks, Deep Learning, Machine Learning, Magnetic Resonance Fingerprinting, Supervised Machine Learning

1. Introduction

Previously presented methods for generating parameter maps in MRF are time-consuming and require a dictionary of time series for every possible combination of parameters like T1 and T2 relaxation times [1, 2]. Furthermore, such a dictionary will only have discrete entries for reasons of efficiency. This can lead to errors in MRF parameter maps [3]. To overcome these time and storage limitations, we train a Convolutional Neural Network (CNN) to predict quantitative T1 and T2 values from MRF time series. Deep leaning has recently been shown to be a promising technique for many applications in medical imaging, e.g. reconstruction in X-ray computed tomography [4, 5]. Two advantages of this method are i) fast computation of the quantitative parameter maps and ii) a better representation of the dictionary data (our trained model requires about 2 MB compared to 210 MB for a dictionary with e.g. 8750 T1/T2 combinations). Moreover, the CNN can predict values for time series

---

[1] Corresponding author, Pattern Recognition Lab, Department of Computer Science, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany, E-Mail: elisabethhoppe@web.de

continuously. This study investigates the execution time and accuracy of predicted parameters using a CNN compared to the conventional dictionary matching approach using simulated data from a FISP MRF implementation [6].

2. Methods

Experiments were performed on a head-shaped gel phantom, a NIST phantom [7] and healthy subjects on a MAGNETOM Skyra 3T (Siemens Healthcare, Erlangen, Germany). 2D MRF-FISP [6] data served as the experimental basis of the dictionary simulation. It was acquired using a prototype sequence with the following parameters: Field-of-view (FOV) 300 mm, resolution 1.17x1.17x5 mm^3, variable repetition time (TR, 12-15 ms), flip angle (FA, 5-74°), number of repetitions (Nrep) 3000. A dictionary with high resolution was simulated to obtain a large amount of training and testing data. Relaxation parameters present in normal human tissues [8, 9, 10] and in the NIST phantom [7] were selected for the simulation (T1: 50 to 4500 ms, T2: 20 to 800 ms, with steps from 2 to 50 ms, with relative B1+ magnitude values ranging from 0.7 to 1.3, step: 0.05, overall about 120,000 time series). Implementation and testing were run on GPU using the machine learning library TensorFlow [11]. The architecture of the network (Figure 1) was inspired by neural networks used in the domain of speech processing [12], as these problems are similar to our problem. We tested different architectures using different numbers of convolutional and fully connected layers. We found that the network model with smallest average error for validation data consists of 3 convolutional layers (kernel size = 3, stride size = 2), each followed by a rectified linear unit (ReLU) activation function. The number of the feature maps per convolutional layer is increasing, from 32 in the first to 128 in the last. After convolution an average pooling layer follows with the same size as the stride size. The last layer is fully connected, with 2 outputs. The simulated time series data was randomly partitioned into disjunct sets for training, validation and testing (80/10/10 %). The weights were initialized uniformly randomly distributed. Training was done with the ADAM [13] optimization method with an initial learning rate of 5*10^-4, by minimizing the Mean Squared Error. Batch size was set to 5 time series. The model was trained for maximal 200 epochs. Early stopping was performed using the validation data. The training was stopped when the validation average error of current epochs increased in comparison to the past epochs.

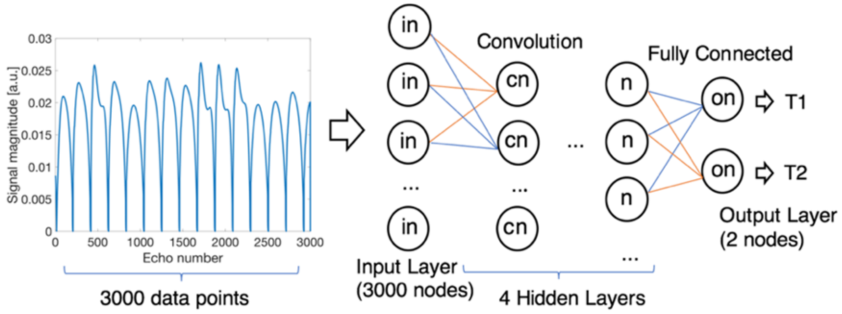

Figure 1. Schema of the CNN. The input is one simulated time series, the outputs are the estimated quantitative values for T1 and T2.

3. Results

The range of T1 and T2 values was chosen to be close to values present in the NIST phantom [7] and in human tissues, especially brain regions [8, 9, 10]. Estimated values of our proposed method show only small average deviations from ground truth values (Figures 2 and 3): The mean absolute deviations (standard deviation, SD) are 0.29% (0.44%) for T1 and 1.22% (2.04%) for T2 for relevant NIST values (T1: 80-1500, T2: 20-450), and the mean absolute deviations (SD) for complete test data (T1: 50-4500, T2: 20-800) are 0.18% (0.44%) for T1 and 1.3% (2.23%) for T2. The comparison between estimated parameters from the network and the simulated dictionary time series for some T1 and T2 values and their deviations are shown in Figures 4 and 5. The comparison of execution time was carried out on a 2.7 GHz Intel Core i5. While the estimation for one time series using the matching method proposed by [1] took about 100 ms, the CNN approach vastly improves this by a factor of 30 (or about a factor of 100 on a GPU).

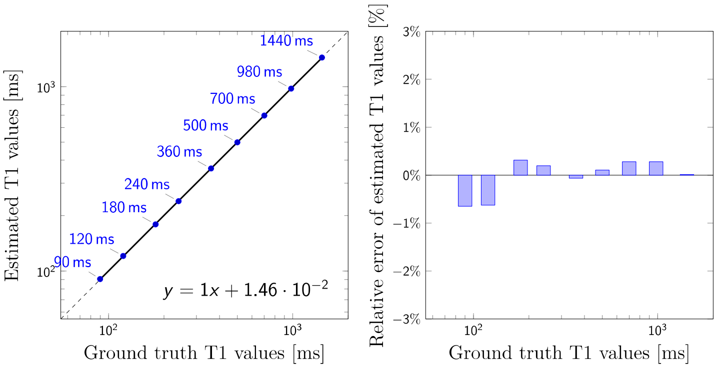

Figure 2. Left: Estimated T1 values for some typical NIST values (90, 120, 180, 240, 360, 500, 700, 980 and 1440 ms). The dashed line is the x = y line, the solid line is the linear regression, with its formula in the bottom right corner. Right: The relative deviations of the estimated values from ground truth values.

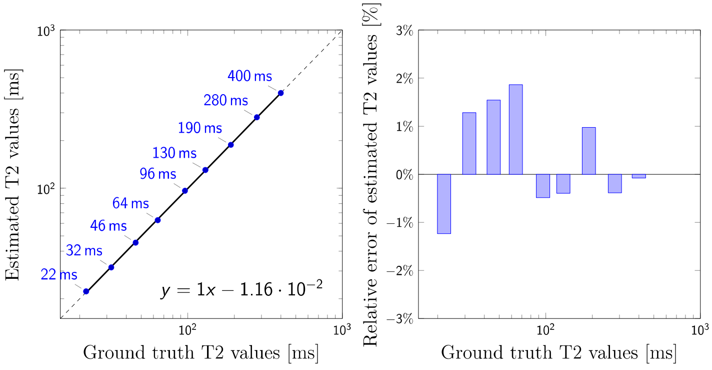

Figure 3. Left: Estimated T2 values for some typical NIST values (22, 32, 46, 64, 96, 130, 190, 280 and 400 ms). The dashed line is the x = y line, the solid line is the linear regression, with its formula in the bottom right corner. Right: The relative deviations of the estimated values from ground truth values.

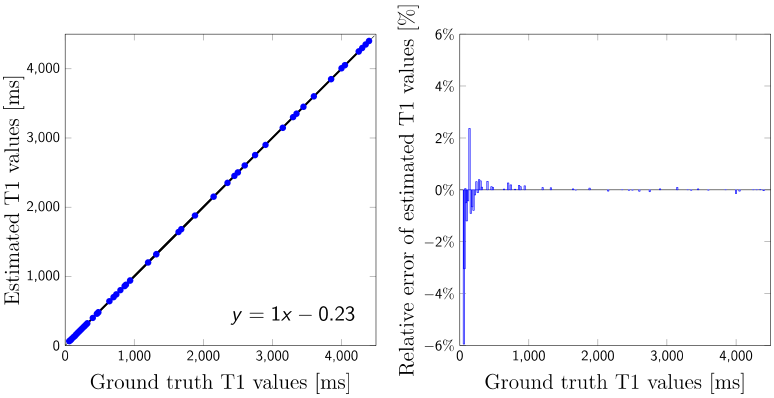

Figure 4. Left: Quantitative examples for estimated T1 values. The solid line is the linear regression, with its formula in the bottom right corner. Right: The relative deviations of the estimated values from ground truth values.

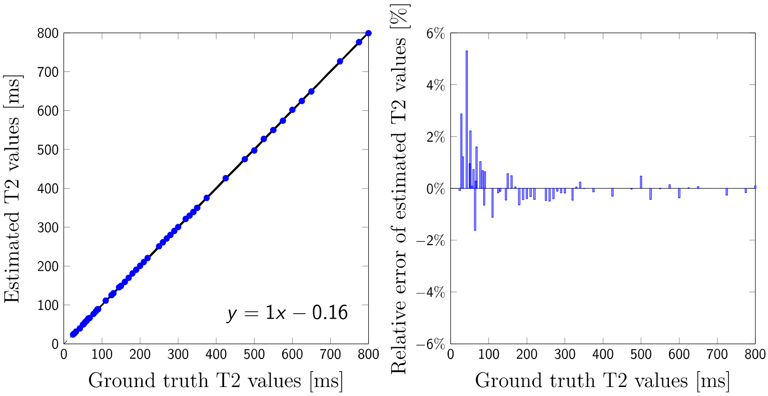

Figure 5. Left: Quantitative examples for estimated T2 values. The solid line is the linear regression, with its formula in the bottom right corner. Right: The relative deviations of the estimated values from ground truth values.

4. Discussion and conclusion

A CNN, trained with simulated time series from a MRF dictionary, is shown to produce accurate predictive results for quantitative parameters like T1 and T2 times. Thus the CNN is able to detect the relevant features and differences between time series for the different T1 and T2 times and can learn these features by itself during training. A CNN model works well as an alternative approach for prediction of quantitative values from time series in MRF. Besides providing accurate results for simulated time series, it also has the following two advantages compared to state-of-the-art matching methods: Firstly, the computation time can be greatly reduced. Secondly, the neural network also provides a very efficient representation of the model when compared to a dictionary. Future work will include the adaption of the proposed concept to real measured time series for generating quantitative parameter maps of human tissues.

References

[1] Ma D, Gulani V, Seiberlich N, Liu K, Sunshine JL, Duerk JL, et al. Magnetic Resonance Fingerprinting. Nature. 2013;495(7440):187-92.

[2] Cauley SF, Setsompop K, Ma D, Jiang Y, Ye H, Adalsteinsson E, et al. Fast group matching for MR fingerprinting reconstruction. Magn Reson Med. 2015;74(2):523-8.

[3] 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 2014): Proceedings of the 11th International Symposium on Biomedical Imaging; 2014 Apr 29-May 2; Beijing, China. IEEE; 2014. p. 870-3.

[4] Würfl T, Ghesu FC, Christlein V, Maier A. Deep Learning Computed Tomography. In: Ourselin S, Joskowicz L, Sabuncu M, Unal G, Wells W, editors. Medical Image Computing and Computer-Assisted Intervention -- MICCAI 2016. MICCAI 2016. Lecture Notes in Computer Science. Vol 9902. Cham: Springer; 2016. p. 432-40.

[5] Hammernik K, Würfl T, Pock T, Maier A. A Deep Learning Architecture for Limited-Angle Computed Tomography Reconstruction. In: Maier-Hein K, Deserno T, Handels H, Tolxdorff T, editors. Bildverarbeitung für die Medizin 2017. Informatik aktuell. Berlin, Heidelberg: Springer Vieweg; 2017. p. 92-7.

[6] Jiang Y, Ma D, Seiberlich N, Gulani V, Griswold MA. MR fingerprinting using fast imaging with steady state precession (FISP) with spiral readout. Magn Reson Med. 2014;74(6):1621-31.

[7] Calibrate MRI Scanners with NIST Referenced Quantitative MRI (qMRI) Phantoms (QIBA DWI and ISMRM) [Internet]. Boulder(CO): High Precision Devices, Inc.; c2017 [cited 2017 Feb 16]. Available from: http://www.hpd-online.com/MRI-phantoms.php.

[8] Bojorquez JZ, Bricq S, Acquitter C, Brunotte F, Walker PM, Lalande A. What are normal relaxation times of tissues at 3 T?. Magn Reson Imaging. 2017;35:69-80.

[9] Stanisz GJ, Odrobina EE, Pun J, Escaravage M, Graham SJ, Bronskill MJ. T1, T2 relaxation and magnetization transfer in tissue at 3T. Magn Reson Med. 2005;54(3):507-12.

[10] Li Y, Xu D, Ozturk-Isik E, Lupo JM, Chen AP, Vigneron DB, et al. T1 and T2 metabolite relaxation times in normal brain at 3T and 7T. J Mol Imaging Dynam S. 2012;1:2.

[11] Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, et al. TensorFlow: A system for large-scale machine learning. In: 12th USENIX Symposium on Operating Systems Design and Implementation: Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation; 2016 Nov 2-4; Savannah(GA), USA. USENIX Association; 2016 [cited 2017 Feb 16]. p. 265-83. Available from: https://www.usenix.org/sites/default/files/osdi16_full_proceedings.pdf.

[12] Zhang Y, Pezeshki M, Brakel P, Zhang S, Bengio CLY, Courville A. Towards end-to-end speech recognition with deep convolutional neural networks. In: Interspeech 2016. Proceedings of Interspeech 2016; 2016 Sep 8-12; San Francisco(CA), USA. 2016 [cited 2017 Feb 16]. p. 214-16. Available from: http://dx.doi.org/10.21437/Interspeech.2016-1446. doi: 10.21437/Interspeech.2016-1446.

[13] Kingma D, Ba J. Adam: A method for stochastic optimization. arXiv preprint [Internet]. 2014 [cited 2017 Feb 16]. Available from: https://arxiv.org/abs/1412.6980.