DR. JESSE IAN HAMILTON (Orcid ID : 0000-0002-4463-481X)

Article type : Note

Deep Learning Reconstruction for Cardiac Magnetic Resonance Fingerprinting T_1 and T_2 Mapping

** Jesse I. Hamilton[1], Danielle Currey[2], Sanjay Rajagopalan[3,4], Nicole Seiberlich[1]

Author Affiliations:

[1]Department of Radiology, University of Michigan, Ann Arbor, MI, USA

[2]Department of Computer Science, Johns Hopkins University, Baltimore, MD, USA

[3]Department of Radiology, University Hospitals Cleveland Medical Center, Cleveland, OH, USA

[4]Division of Cardiovascular Medicine, University Hospitals Cleveland Medical Center, Cleveland, OH, USA

Contact Email: JI Hamilton – hamiljes@med.umich.edu; D Currey – dcurrey1@jhu.edu; S Rajagopalan – Sanjay.Rajagopalan@uhhospitals.org; N Seiberlich – nse@med.umich.edu

** Corresponding author at MSRB-2 Room #1590, 1150 West Medical Center Drive, Ann Arbor, MI 48109-1590

Running Head: Deep Learning Cardiac MRF

Word Count: 3103

Keywords: Magnetic Resonance Fingerprinting; deep learning; tissue characterization; T_1 mapping; T_2 mapping; neural network

Pages: 8

Figures: 5

Tables: 0

Supplementary Files: 1

Abstract

Purpose: To develop a deep learning method for rapidly reconstructing T_1 and T_2 maps from undersampled electrocardiogram (ECG) triggered cardiac Magnetic Resonance Fingerprinting (cMRF) images.

Methods: A neural network was developed that outputs T_1 and T_2 values when given a measured cMRF signal timecourse and cardiac RR interval times recorded by an ECG. Over 8 million cMRF signals, corresponding to 4000 random cardiac rhythms, were simulated for training. The training signals were corrupted by simulated k-space undersampling artifacts and random phase shifts to promote robust learning. The deep learning reconstruction was evaluated in Monte Carlo simulations for a variety of cardiac rhythms and compared with dictionary-based pattern matching in 58 healthy subjects at 1.5T.

Results: In simulations, the normalized root-mean-square-error (nRMSE) for T_1 was below 1% in myocardium, blood, and liver for all tested heart rates. For T_2, the nRMSE was below 4% for myocardium and liver and below 6% for blood for all heart rates. The difference in the mean myocardial T_1 or T_2 observed *in vivo* between dictionary matching and deep learning was 3.6ms for T_1 and –0.2ms for T_2. Whereas dictionary generation and pattern matching required more than 4 minutes per slice, the deep learning reconstruction only required 336ms.

Conclusion: A neural network is introduced for reconstructing cMRF T_1 and T_2 maps directly from undersampled spiral images in under 400ms and is robust to arbitrary cardiac rhythms, which paves the way for rapid online display of cMRF maps.

Keywords: Magnetic Resonance Fingerprinting; deep learning; tissue characterization; T_1 mapping; T_2 mapping; neural network

Introduction

Quantitative MRI is a powerful tool for assessing cardiac health. Two clinically measured tissue properties are T_1 and T_2, which can be used for early detection and monitoring of fibrosis,[1] inflammation,[2] and edema,[3] among other conditions. Cardiac Magnetic Resonance Fingerprinting (cMRF) is one technique for simultaneous T_1-T_2 mapping,[4,5] which uses a time-varying sequence, an undersampled spiral k-space trajectory, and pattern matching with a dictionary of simulated signals to estimate quantitative maps.

Although cMRF is efficient, as data are collected during one breathhold, the reconstruction time is long and prohibits real-time display of the maps. The major hurdle is that the subject's cardiac rhythm dictates the sequence timings because the scan is electrocardiogram (ECG) triggered, and thus a new dictionary must be simulated after every acquisition. The dictionary simulation time increases if slice profile imperfections or other effects are modeled; both dictionary simulation and pattern matching take longer if additional properties (e.g., B_1^+) beyond T_1 and T_2 are quantified.[6–8] A typical cMRF reconstruction for T_1-T_2 mapping requires 4 minutes for dictionary simulation (including corrections for slice profile and preparation pulse efficiency) and 10 seconds for pattern matching.

The combination of deep learning and MRF is gaining interest because of the potential for orders of magnitude reductions in computation time.[9–11] Previously, a neural network was proposed that reduces cMRF dictionary simulation time to one second and generalizes to arbitrary cardiac rhythms, which eliminates the need for time-consuming and scan-specific Bloch equation simulations.[12] However, this approach still generates a scan-specific dictionary that occupies memory (220MB). Measuring additional properties beyond T_1 and T_2 would require exponentially more memory and time and quickly become infeasible. In addition, the maps have quantization errors due to the discrete step sizes in the dictionary.

Neural networks have been proposed to directly quantify T_1 and T_2 from MRF images in non-cardiac applications, thereby bypassing dictionary simulation and pattern matching to reduce computation time and memory requirements. However, existing methods are not directly applicable to cMRF. Previous approaches have only considered scans with fixed sequence timings, whereas the cMRF sequence timings are determined by the subject's cardiac rhythm.[11] Some existing neural network approaches cannot reconstruct maps from undersampled non-Cartesian data and require additional reconstruction steps.[9] Other approaches require *in vivo* MRF datasets for training,[10] which may be time-consuming and expensive to collect, and may not generalize to scenarios that are underrepresented in the training set.

In this work, a deep learning reconstruction is proposed for cMRF that directly outputs T_1 and T_2 maps from undersampled spiral images in under 400ms per slice without using a dictionary. The network is robust to arbitrary cardiac rhythms and eliminates the need for scan-specific Bloch equation simulations and pattern matching. The cMRF deep learning reconstruction is evaluated in simulations and compared with dictionary-based pattern matching using *in vivo* data acquired in 58 healthy subjects at 1.5T.

Methods

cMRF Sequence Parameters

The cMRF sequence has been described in previous work,[13] although the breathhold duration here was reduced from 15 to 10 heartbeats. A FISP readout is used that is relatively insensitive to off-resonance due to the unbalanced gradient moment on the slice-select axis.[14] Multiple preparation pulses are applied with the following pattern (which repeats twice): inversion (TI=21ms), no preparation, T_2-prep (30ms), T_2-prep (50ms), T_2-prep (80ms). The acquisition is ECG-triggered with a 250ms diastolic readout with 50 TRs collected each heartbeat and 500 TRs collected during the entire scan. Data are acquired using an undersampled spiral k-space trajectory with golden angle rotation[15] that requires 48 interleaves to fully sample k-space.[16] Other parameters include a 192x192 matrix, 300mm^2 field-of-view, 1.6x1.6x8.0mm^3 resolution, and constant TR/TE 5.1/1.4ms.

Neural Network Architecture

Figure 1 shows a diagram of the proposed network. The network takes two inputs—the measured signal timecourse from one voxel and the cardiac RR interval times from the ECG. The timecourse is split into real and imaginary parts and concatenated with the RR interval

times, resulting in a vector of length $2N + M$, where N is the number of TRs and M is the number of heartbeats. This study uses $N = 500$ and $M = 10$. The input is normalized by dividing the RR intervals (in milliseconds) by 1000 and dividing the signal by its l_2-norm. The network has 18 hidden layers with 300 nodes per layer. Skip connections are used every 4 layers beginning after the first, which avoids problems with vanishing gradients during training. Supporting Information Figures S1 and S2 provide justification for the number of hidden layers and use of skip connections. The final outputs are the T_1 and T_2 estimates for the given voxel.

Neural Network Training

The training dataset consists of cMRF signals simulated using the Bloch equations, corresponding to 4000 randomly generated cardiac rhythms. Each cardiac rhythm has an average heart rate (HR) between 40-120 beats per minute (bpm), and random Gaussian noise with a standard deviation (SD) between 0-100% of the mean RR interval is added to the RR interval times to introduce variability. Adding noise with a large SD (i.e., near 100%) mimics ECG mis-triggering because it results in large timing variations between heartbeats. For each cardiac rhythm, 2000 signals were generated with T_1 and T_2 values selected from a uniform random distribution between 50-3000ms and 5-700ms, respectively. In total, 8 million (4000x2000) training signals were simulated including corrections for slice profile (assuming a 0.8ms duration sinc RF pulse with time-bandwidth product 2) and preparation pulse efficiency.[6,7]

Although adding Gaussian noise to training data is common in machine learning to promote robust learning, non-Cartesian undersampling artifacts do not fall along a Gaussian distribution. Therefore, the network is trained using simulated cMRF signals corrupted with noise that mimics non-Cartesian aliasing, hereafter called “pseudo-noise” (Supporting Information Figure S3 compares neural networks trained with pseudo-noise versus Gaussian noise). A repository of pseudo-noise is generated before training (Supporting Information Figure S4). The pseudo-noise is meant to be agnostic to cardiac rhythm and image content. To create the repository, random T_1, T_2, and M_0 maps are synthesized where each voxel has a random value between 50-3000ms for T_1, 5-700ms for T_2, and 0-1 for M_0. A random cardiac rhythm is also generated with an average HR between 40-120bpm, with Gaussian noise having SD between 0-100% of the mean RR interval added to the RR interval times. Signals are simulated using the Bloch equations to yield a time series of reference images. Data acquisition is simulated using the spiral k-space sampling pattern, and undersampled images are gridded using the non-uniform fast Fourier Transform (NUFFT).[17] The fully-sampled reference images are subtracted from the undersampled images. Each voxel in the resulting difference images is treated as an

independent pseudo-noise sample and saved in the repository. For a 192x192 matrix, these steps result in 36,864 (192^2) pseudo-noise samples. The complete repository contains 1.8 million pseudo-noise samples generated by repeating this process 50 times using random parameter maps and cardiac rhythms.

When training the network (Figure 2B), pseudo-noise samples are randomly selected from the repository every epoch and added to the simulated cMRF signals, similar to an approach described for contrast synthesis by Virtue, et al.[18] Let $s(t)$ denote an arbitrary cMRF signal and $n(t)$ denote an arbitrary pseudo-noise sample. The pseudo-noise is randomly scaled by a factor C so the SNR is between 0.2 and 1.0, which was empirically determined to be appropriate for the k-space trajectory employed in this study (Supporting Information Figure S5) and would need to be tuned for other trajectories. The SNR is defined as follows:

$$SNR = \frac{\|s(t)\|_2}{\|n(t)\|_2} \quad \text{[Eq. 1]}$$

Each cMRF signal is also multiplied by a random phase shift ϕ_1, and the pseudo-noise is multiplied by a different random phase shift ϕ_2. Phase shifts are performed because *in vivo* datasets have arbitrary phase due to factors such as receiver coil sensitivity profiles and off-resonance. Supporting Information Figure S6 compares the performance of networks trained with and without random phase shifts. The final cMRF signal used for training is denoted by $\tilde{s}(t)$.

$$\tilde{s}(t) = s(t) \cdot e^{i\phi_1} + C \cdot n(t) \cdot e^{i\phi_2} \quad \text{[Eq. 2]}$$

A separate validation dataset was created by generating 400 random cardiac rhythms and simulating 500 cMRF signals for each rhythm corrupted by pseudo-noise and phase shifts. The neural network was implemented in PyTorch and trained for 5 epochs using an Adam optimizer with learning rate 10^{-4} and batch size 128. The network parameters with the smallest validation loss were saved. A normalized l_1 loss function (Equation 3) was used that was the sum of the relative errors in T_1 and T_2, where B is the batch size, $T_{1,i}^{net}$ and $T_{2,i}^{net}$ are the network estimates for T_1 and T_2, and $T_{1,i}^{ref}$ and $T_{2,i}^{ref}$ are the reference T_1 and T_2 values.

$$loss = \frac{1}{B}\sum_{i=1}^{B}\left(\frac{\left|T_{1,i}^{net} - T_{1,i}^{ref}\right|}{T_{1,i}^{ref}} + \frac{\left|T_{2,i}^{net} - T_{2,i}^{ref}\right|}{T_{1,i}^{ref}}\right) \quad \text{[Eq. 3]}$$

Simulation Experiments

Monte Carlo simulations were performed using a digital cardiac phantom (MRXCAT)[19] to evaluate the accuracy of the deep learning reconstruction. The phantom used myocardial T_1/T_2=1400/50ms, blood T_1/T_2=1950/280ms, and liver T_1/T_2=800/40ms. Datasets with different cardiac rhythms were simulated where the average HR was swept from 40 to 120bpm (step size 10bpm), and Gaussian noise was added to the RR interval times with SD 0%, 10%, 20%, 50%, 75%, and 100% of the mean RR interval to introduce heart rate variability. For each combination of average HR and noise level, 50 cMRF datasets with different cardiac rhythms were simulated by performing Bloch equation simulations, spiral k-space sampling, and gridding. The undersampled images and RR interval times were input to the neural network to reconstruct T_1 and T_2 maps. The mean T_1 and T_2 values were computed in the myocardial wall, left ventricular blood pool, and liver and are reported using normalized root mean square error (nRMSE).

In Vivo Experiments

cMRF scans from 58 healthy adult subjects were retrospectively collected in a HIPAA-compliant, IRB-approved study. The scans were performed on a 1.5T MRI scanner (MAGNETOM Aera, Siemens Healthineers, Germany) at a medial short-axis slice position during an end-expiration breathhold with a 192x192 matrix size, 300mm^2 field-of-view, and 1.6x1.6x8.0mm^3 resolution. T_1 and T_2 maps were reconstructed in two ways: 1) using the Bloch equations to simulate a scan-specific dictionary and performing pattern matching as in previous cMRF work,[13] hereafter called “dictionary matching”, and 2) using the deep learning reconstruction. The dictionary contained 23,345 entries with T_1 [10:10:2000, 2020:20:3000]ms and T_2 [4:2:100, 105:5:300, 320:20:500]ms. The mean T_1 and T_2 over the entire myocardial wall were compared between both reconstructions using a two-tailed Student’s t-test for pairwise comparisons, with $p<0.05$ considered statistically significant. The mean T_1 and T_2 were also compared using linear regression and Bland-Altman analyses.[20] Intrasubject variability for dictionary matching and deep learning were assessed by computing the SD in T_1 and T_2 over the myocardium for each subject. Intersubject variability was assessed by computing the coefficient of variation (CV), obtained by calculating the SD of the mean T_1 and T_2 measured for each subject and dividing by the group-averaged T_1 and T_2.

Results

Computation Time

Gridding required 30s and was required for both deep learning and dictionary matching reconstructions. The average time to quantify T_1 and T_2 maps from the gridded images using deep learning was 336ms. For comparison, simulating a scan-specific dictionary required 4 minutes, and pattern matching required an additional 10s. Each dictionary occupied 220MB of memory. The deep learning reconstruction does not utilize a dictionary, and the network parameters only occupied 7MB.

Simulation Experiments

Figure 3 shows results from the Monte Carlo simulations. The deep learning reconstruction was more accurate at estimating T_1 than T_2. The nRMSE for T_1 was generally below 1% for all tissue types (myocardium, blood, and liver). Note that a 1% error corresponds to a 14ms difference from the true T_1 of 1400ms in myocardium. The T_2 nRMSE was below 4% for myocardium and liver, and below 6% for blood. A 4% error corresponds to a 2ms difference from the true T_2 of 50ms in myocardium. The quantification accuracy for T_1 and T_2 was similar regardless of average HR or the variability of the cardiac rhythm.

In Vivo Imaging

Maps from two representative subjects are shown in Figure 4. Subject A had a steady cardiac rhythm (mean RR 775 ± 28ms), while Subject B had a variable cardiac rhythm (mean RR 770 $\pm$ 215ms) with one missed ECG trigger during heartbeat 10. The maps in the myocardium were visually similar between the deep learning and dictionary matching reconstructions. There were differences in some areas, such as subcutaneous fat. Figure 5A shows the linear regression analysis between the mean myocardial T_1 and T_2 values from deep learning and dictionary matching. The measurements were strongly correlated, with R^2=0.93 for T_1 and R^2=0.95 for T_2. As seen in the Bland-Altman analysis (Figure 5B), the mean T_1 bias was 3.6ms with 95% limits of agreement (-18.9, 26.1)ms, and the mean T_2 bias was -0.2ms with 95% limits of agreement (-1.9, 1.5)ms. Using a paired t-test, the differences in the mean myocardial values between deep learning and dictionary matching were statistically significant for T_1 (p=0.019) and T_2 (p=0.038). Figure 5C compares the intrasubject standard deviations. The SD for T_1 was 106.9ms for dictionary matching and 110.2ms for deep learning, and the difference was statistically significant (p=0.013). The SD for T_2 was 6.8ms for dictionary matching and 7.3ms for deep learning, and the difference was statistically significant ($p < 0.0001$). The intersubject variability

was similar for both reconstructions. For T_1, the CV was 4.4% for dictionary matching and 4.5% for deep learning; for T_2, the CV was 9.1% for dictionary matching and 8.9% for deep learning.

Discussion

This study introduces a deep learning method for rapidly performing cMRF T_1 and T_2 quantification that is robust to arbitrary cardiac rhythms. A neural network was trained to directly output T_1 and T_2 maps from undersampled spiral cMRF images and cardiac RR interval timings. The deep learning reconstruction does not require Bloch equation simulations to create a dictionary or use pattern matching. The main advantage is the large reduction in computation time, which could enable real-time display of cMRF maps. The deep learning method takes less than 400ms per slice to reconstruct T_1/T_2 maps from cMRF images, which is more than a 700-fold speedup compared to dictionary matching. The deep learning reconstruction also requires less memory than dictionary matching. Whereas the dictionary occupies 220MB of memory, the network coefficients only occupy 7 MB. Although this study focuses on T_1 and T_2 quantification, the savings in computation time and memory may be more pronounced for applications seeking to measure additional tissue properties.

The deep learning reconstruction yielded accurate T_1 and T_2 estimates in simulations, with T_1 errors below 1% and T_2 errors below 6% regardless of the variability in the cardiac rhythms. *In vivo*, the deep learning reconstruction had similar accuracy and precision as dictionary matching. Although a statistically significant bias was observed in the mean and SD of the myocardial T_1 and T_2 values compared to dictionary matching, their magnitude was small (3.6ms difference in mean T_1 and -0.2ms difference in mean T_2).

There are several interesting features of the cMRF deep learning reconstruction. First, whereas dictionary matching leads to quantization errors because the T_1 and T_2 estimates are restricted to discrete values, the neural network produces continuous outputs. Supporting Information Figure S7 compares dictionary matching and deep learning in an example where the T_1 and T_2 values of a ground truth signal do not lie exactly on the T_1-T_2 grid used to populate the dictionary. Second, the network is trained for a fixed k-space undersampling pattern. To achieve the best performance, the network should be retrained if data are acquired with a different sampling pattern, as the distribution of aliasing artifacts would change (Supporting Information Figure S8).

Recently, other neural network approaches have been described for MRF and for cardiac parameter mapping. DRONE uses a 2-layer fully-connected network for MRF T_1 and T_2

quantification, although the sequence timings are fixed and non-Cartesian k-space undersampling is not taken into consideration.[9] Cao, et al. have proposed a 4-layer fully-connected network and developed a method for simulating training data with non-Cartesian undersampling artifacts, although limited to MRF sequences with fixed timings.[11] Fang, et al. have developed a U-net for high-resolution spiral MRF in the brain. However, the network uses *in vivo* training data, which may be time-consuming and expensive to collect, and may not generalize to pathological scenarios underrepresented in the training set. In this study, a neural network is trained using simulated cMRF signals, which has the advantage that an arbitrarily large training set can be generated to improve performance. Also, whereas a U-net may introduce blurring, the network used here operates voxelwise and therefore does not induce spatial smoothing. Another recent technique is DeepBLESS, which is a deep learning reconstruction for simultaneous cardiac T_1-T_2 mapping using a non-fingerprinting sequence.[21] Similar to this study, it is trained to be robust to arbitrary cardiac rhythms. However, highly undersampled radial images are first reconstructed using compressed sensing before being input to the network, which requires three minutes of additional computation time.

This study has several limitations. First, it is still necessary to grid the spiral k-space data, which requires 30s on a standard workstation using a CPU; thus, the computation bottleneck is now gridding rather than dictionary simulation. Gridding could be accelerated using parallel GPUs[22] or by applying GRAPPA operator gridding (GROG) to shift the k-space data points onto a Cartesian grid,[23] although these approaches were not investigated here. Second, although the cMRF T_1 and T_2 estimation is robust to field inhomogeneities (Supporting Information Figure S9), no corrections were made for off-resonance blurring during the spiral readout, which can degrade spatial resolution and lead to fat signal contamination, especially near epicardial fat or in regions with intramyocardial fat (Supporting Information Figure S10). Third, both dictionary-based and deep learning cMRF reconstructions can be affected by partial volume artifacts (Supporting Information Figure S11). Fourth, B_1^+ corrections were not considered. Fifth, no attempt was made to model the complicated spin history of flowing blood, and thus the blood T_1/T_2 estimates may not be reliable. Sixth, no post-contrast T_1/T_2 mapping was performed, although simulations suggest the network could be used for post-contrast data (Supporting Information Figure S12). Seventh, no comparison was made between deep learning cMRF and conventional T_1/T_2 mapping techniques, although prior work has compared dictionary-based cMRF with conventional mapping.[13] Finally, the in vivo results were limited to healthy subjects, and additional validation of the deep learning cMRF reconstruction should be performed in patients with known cardiac pathologies.

In conclusion, this work introduces a deep learning method for reconstructing T_1 and T_2 maps from undersampled spiral cMRF images in less than 400ms per slice with similar accuracy and precision *in vivo* as dictionary matching. By eliminating the need for scan-specific dictionary generation and pattern matching, this approach may enable rapid at-the-scanner reconstructions and facilitate the clinical translation of cMRF.

Funding Sources

This work was funded by the National Institutes of Health (NIH/NHLBI R01HL094557), National Science Foundation (NSF CBET 1553441), and Siemens Healthineers. The funding sources were not involved in the study design; collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

References

1. Okur A, Kantarcı M, Kızrak Y, et al. Quantitative evaluation of ischemic myocardial scar tissue by unenhanced T1 mapping using 3.0 Tesla MR scanner. *Diagn Interv Radiol*. 20(5):407-413.

2. Hinojar R, Nagel E, Puntmann VO. T1 mapping in myocarditis - headway to a new era for cardiovascular magnetic resonance. *Expert Rev Cardiovasc Ther*. 2015;13(8):871-874.

3. Giri S, Chung Y-C, Merchant A, et al. T2 quantification for improved detection of myocardial edema. *J Cardiovasc Magn Reson*. 2009;11(1):56.

4. Ma D, Gulani V, Seiberlich N, et al. Magnetic resonance fingerprinting. *Nature*. 2013;495(7440):187-192.

5. Hamilton JI, Jiang Y, Chen Y, et al. MR fingerprinting for rapid quantification of myocardial T1, T2, and proton spin density. *Magn Reson Med*. 2017;77(4):1446-1458.

6. Hamilton JI, Jiang Y, Ma D, et al. Investigating and reducing the effects of confounding factors for robust T1 and T2 mapping with cardiac MR fingerprinting. *Magn Reson Imaging*. 2018;53:40-51.

7. Ma D, Coppo S, Chen Y, et al. Slice profile and B1 corrections in 2D magnetic resonance fingerprinting. *Magn Reson Med*. 2017;78(5):1781-1789.

8. Buonincontri G, Schulte RF, Cosottini M, Tosetti M. Spiral MR fingerprinting at 7 T with

simultaneous B1 estimation. *Magn Reson Imaging*. 2017;41:1-6.

9. Cohen O, Zhu B, Rosen MS. MR fingerprinting Deep RecOnstruction NEtwork (DRONE). *Magn Reson Med*. 2018;80(3):885-894.

10. Fang Z, Chen Y, Hung S, Zhang X, Lin W, Shen D. Submillimeter MR fingerprinting using deep learning–based tissue quantification. *Magn Reson Med*. 2020;84(2):579-591.

11. Cao P, Cui D, Vardhanabhuti V, Hui ES. Development of fast deep learning quantification for magnetic resonance fingerprinting in vivo. *Magn Reson Imaging*. 2020;70:81-90.

12. Hamilton JI, Seiberlich N. Machine Learning for Rapid Magnetic Resonance Fingerprinting Tissue Property Quantification. *Proc IEEE*. 2019;PP:1-17.

13. Hamilton JI, Pahwa S, Adedigba J, et al. Simultaneous Mapping of T1 and T2 Using Cardiac Magnetic Resonance Fingerprinting in a Cohort of Healthy Subjects at 1.5T. *J Magn Reson Imaging*. March 2020.

14. 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*. 2015;74(6):1621-1631.

15. Winkelmann S, Schaeffter T, Koehler T, Eggers H, Doessel O. An optimal radial profile order based on the Golden Ratio for time-resolved MRI. *IEEE Trans Med Imaging*. 2007;26(1):68-76.

16. Hargreaves B. Variable-Density Spiral Design Functions. http://mrsrl.stanford.edu/~brian/vdspiral/. Published 2005. Accessed June 1, 2017.

17. Fessler J, Sutton B. Nonuniform fast Fourier transforms using min-max interpolation. *IEEE Trans Signal Process*. 2003;51(2):560-574.

18. Virtue P, Tamir JI, Doneva M, Yu SX, Lustig M. Learning Contrast Synthesis from MR Fingerprinting. In: *Proc. 26th Annu. Meet. ISMRM*. Paris, France; 2018:676.

19. Wissmann L, Santelli C, Segars WP, Kozerke S. MRXCAT: Realistic numerical phantoms for cardiovascular magnetic resonance. *J Cardiovasc Magn Reson*. 2014;16(1):63.

20. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1(8476):307-310.

21. Shao J, Ghodrati V, Nguyen K, Hu P. Fast and accurate calculation of myocardial T1 and

T2 values using deep learning Bloch equation simulations (DeepBLESS). *Magn Reson Med*. May 2020.

22. Knoll F, Schwarzl A, Diwoky C, Sodickson DK. gpuNUFFT - An open source GPU library for 3D regridding with direct Matlab interface. In: *Proceedings of the ISMRM*. ; 2014:4297.

23. Seiberlich N, Breuer FA, Blaimer M, Barkauskas K, Jakob PM, Griswold MA. Non-Cartesian data reconstruction using GRAPPA operator gridding (GROG). *Magn Reson Med*. 2007;58(6):1257-1265.

Figure Captions

Figure 1. *Neural network for cMRF T_1 and T_2 map reconstruction.* A neural network is used with 18 hidden layers with 300 nodes per layer (blue rectangles) and rectified linear unit (ReLU) activation functions (yellow arrows). Skip connections (black lines) are used every 4 layers. The inputs to the network are a measured cMRF signal timecourse concatenated with the cardiac RR interval times, and the outputs are the estimated T_1 and T_2 values. The network operations are performed independently for each voxel.

Figure 2. *Generation of training data*. The network is trained using simulated cMRF signal timecourses. A pseudo-noise sample is randomly drawn from the repository (generation of pseudo-noise is described in Supporting Information Figure S4). The amplitude of the pseudo-noise is scaled by a factor *C* so the SNR of the noisy signal is between specific bounds (0.2-1.0 for this study). Random phase shifts ϕ_1 and ϕ_2 are applied to the cMRF signal and pseudo-noise, respectively. The pseudo-noise is added to the cMRF signal to yield the noisy signal that will be used for training.

Figure 3. *Monte Carlo simulation results for (A) T_1 and (B) T_2 in myocardium, liver, and blood.* The different color lines refer to the standard deviation (SD) of Gaussian noise that is added to the RR interval times, with the SD given as a percentage of the mean RR interval. SD 0% refers to a constant heart rate, while SD 100% refers to a highly variable heart rate.

Figure 4. *cMRF T_1 and T_2 maps in two healthy subjects at 1.5T.* T_1 and T_2 maps are shown corresponding to dictionary-based pattern matching and the deep learning reconstruction, along

with difference maps. Subject A had a steady cardiac rhythm, while Subject B had a variable cardiac rhythm with one missed ECG trigger. The mean and standard deviation in T_1 and T_2 over the entire myocardium are displayed as insets.

Figure 5. *Analysis of the in vivo data*. Linear regression plots are shown comparing the mean myocardial (A) T_1 and (B) T_2 values using dictionary matching and deep learning. Bland-Altman plots are shown comparing the mean myocardial (C) T_1 and (D) T_2. The solid line indicates the bias, and the dotted lines indicate the 95% limits of agreement. Boxplots comparing the intrasubject standard deviation (SD) for (E) T_1 and (F) T_2 in the myocardium are also presented.

Supporting Information Figure S1. Diagram of a neural network with 18 hidden layers and skip connections every 2 layers.

Supporting Information Figure S2. Validation loss as a function of (A) the number of hidden layers and (B) the number of layers between skip connections.

Supporting Information Figure S3. Maps from a healthy subject using a network trained using pseudo-noise, a network trained using Gaussian noise, and dictionary matching. The mean and standard deviation of T_1 and T_2 in the myocardial wall are reported in the insets.

Supporting Information Figure S4. A repository of noise representative of k-space undersampling artifacts ("pseudo-noise") is generated before training the network. Simulated parameter maps are generated where each voxel has a random parameter value. The cMRF data acquisition and spiral k-space sampling are simulated to yield undersampled images. The reference images are subtracted from the undersampled images, and each voxel in the resulting difference images is treated as an independent sample of pseudo-noise.

Supporting Information Figure S5. Distribution of SNR values in a simulated cardiac phantom, used to inform the choice of SNR scaling factor when training the network.

Supporting Information Figure S6. Maps from a healthy subject using a network trained with random phase shifts, a network trained without random phase shifts, and dictionary matching.

Supporting Information Figure S7. Monte Carlo simulation results demonstrating quantization errors using dictionary matching. Histograms are displayed showing the distribution of T_1 values estimated with (A) dictionary matching and (B) the neural network. In both (A) and (B), the width of the histogram bins is 1ms. The red line depicts the ground truth T_1 of 1397ms. Similar histograms are shown for T_2 using (C) dictionary matching and (D) the neural network. Here the width of the histogram bins is 0.5ms in both (C) and (D), and the ground truth T_2 is 41ms.

Supporting Information Figure S8. Monte Carlo simulation results for myocardial T_1 and T_2 using different spiral k-space interleaf orderings. The different color lines refer to the standard deviation (SD) of Gaussian noise added to the RR interval times, with SD given as a percentage of the mean RR interval. (A) Both training and testing datasets employed spiral golden angle sampling. (B) The training data employed spiral golden angle sampling, while the testing data used spiral sampling with incremental rotation. (C) Both training and testing datasets employed spiral sampling with incremental rotation.

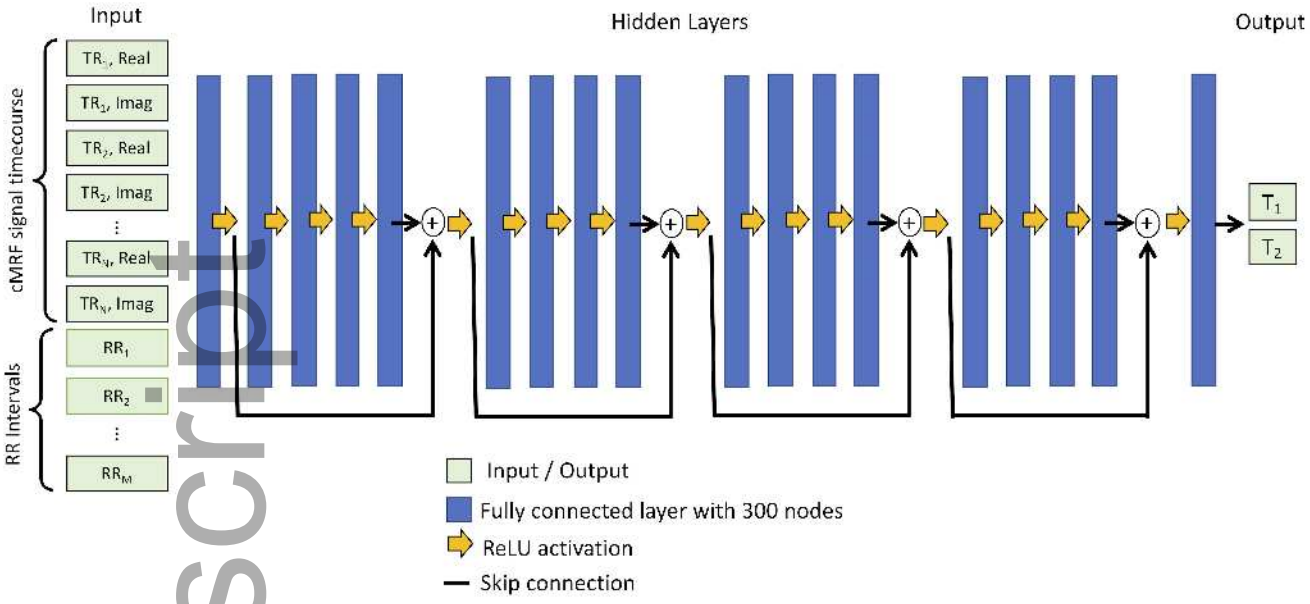
Supporting Information Figure S9. Monte Carlo simulation results for the neural network (A) T_1 and (B) T_2 estimates as a function of off-resonance frequency, plotted in blue. The red line indicates the ground truth T_1 and T_2 values.

Supporting Information Figure S10. Simulations of partial volume effects between myocardium and fat on network (A) T_1 and (B) T_2 estimates. Results are shown for dictionary matching (blue) and the neural network (red), with the vertical error bars indicating the standard deviation over 5000 Monte Carlo repetitions. The dotted lines indicate the ground truth values for pure myocardium and pure fat.

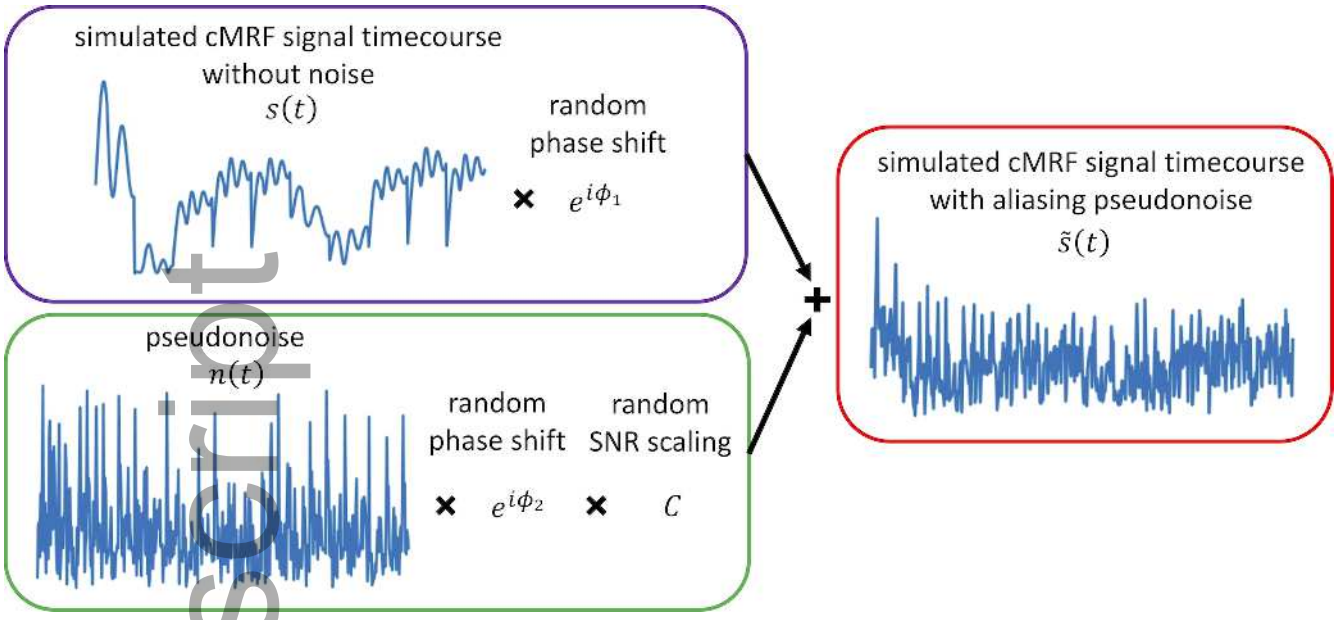
Supporting Information Figure S11. Simulations of partial volume effects between myocardium and blood on network (A) T_1 and (B) T_2 estimates. Results are shown for dictionary matching (blue) and the neural network (red), with the vertical error bars indicating the standard deviation over 5000 Monte Carlo repetitions. The dotted lines indicate the ground truth values for pure myocardium and pure blood.

Supporting Information Figure S12. Monte Carlo simulation using a post-contrast cardiac phantom. Results are shown for (A) T_1 and (B) T_2 in myocardium, liver, and blood. The different color lines refer to the standard deviation (SD) of Gaussian noise that is added to the RR interval times, with the SD given as a percentage of the mean RR interval. SD 0% refers to a constant heart rate, while SD 100% refers to a highly variable heart rate.

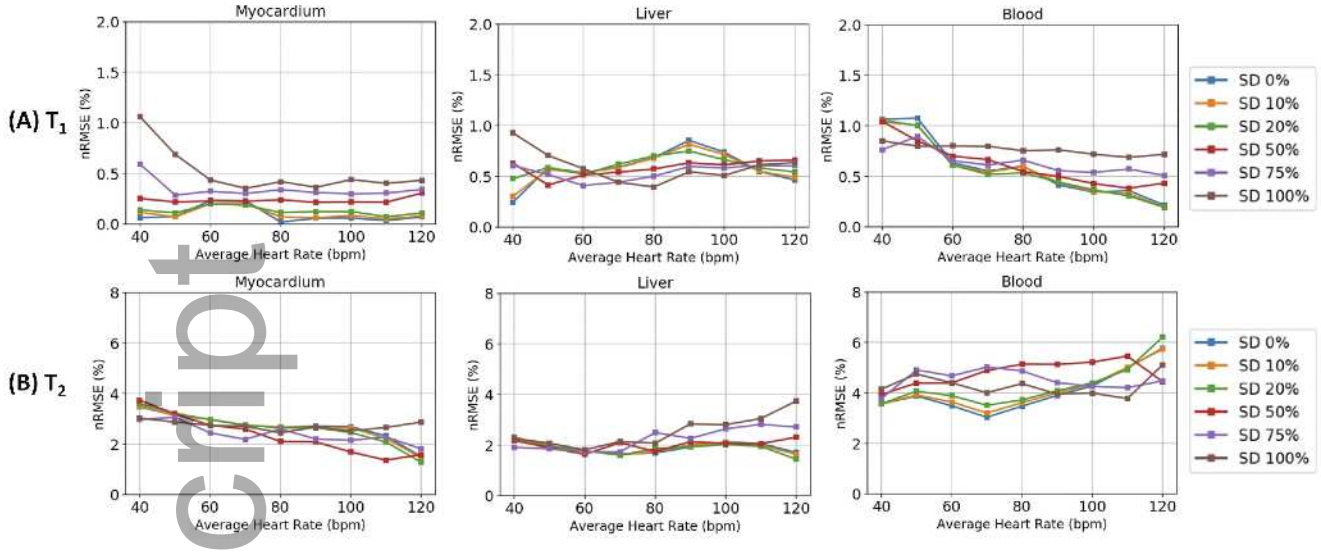
Supporting Information Table S1. Spiral interleaf orderings used during training and testing.



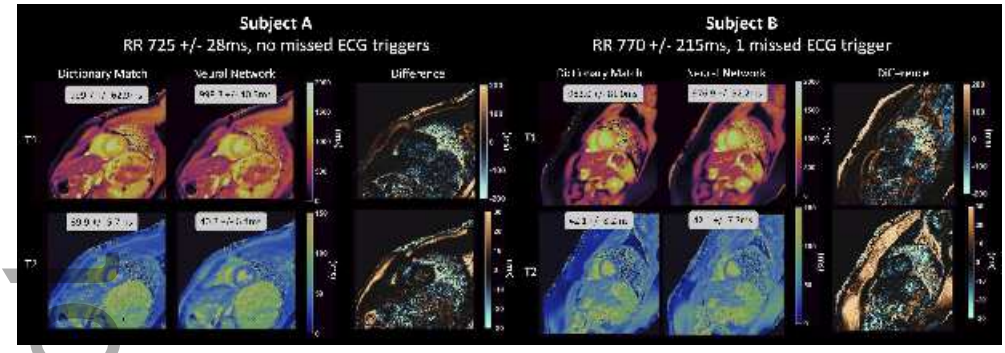
mrm_28568_f1.tiff



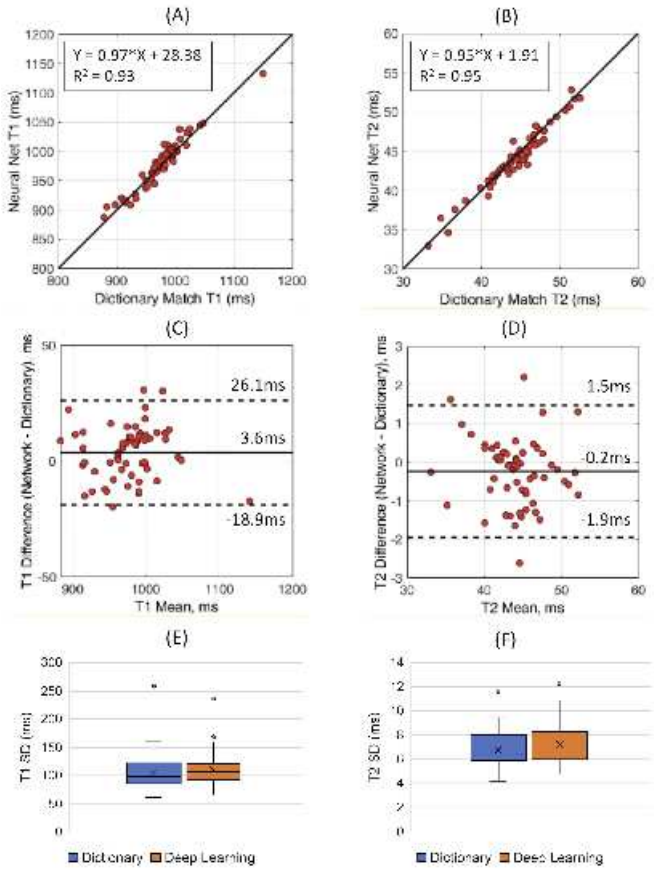
mrm_28568_f2.tiff



mrm_28568_f3.tiff



mrm_28568_f4.tiff



mrm_28568_f5.tiff