NONLINEAR EQUIVARIANT IMAGING: LEARNING MULTI-PARAMETRIC TISSUE MAPPING WITHOUT GROUND TRUTH FOR COMPRESSIVE QUANTITATIVE MRI

Ketan Fatania, Kwai Y. Chau, Carolin M. Pirkl, Marion I. Menzel and Mohammad Golbabaee

ABSTRACT

Current state-of-the-art reconstruction for quantitative tissue maps from fast, compressive, Magnetic Resonance Fingerprinting (MRF), use supervised deep learning, with the drawback of requiring high-fidelity ground truth tissue map training data which is limited. This paper proposes NonLinear Equivariant Imaging for MRF (NLEI-MRF), a self-supervised learning approach to eliminate the need for ground truth for deep MRF image reconstruction. NLEI-MRF extends the recent Equivariant Imaging framework to the MRF non-linear inverse problem. Only compressed-sampled MRF scans are used for training. NLEI-MRF learns tissue mapping using spatiotemporal priors: spatial priors are obtained from the invariance of MRF data to a group of geometric image transformations, while temporal priors are obtained from a nonlinear Bloch response model approximated by a pre-trained neural network. Tested retrospectively on two acquisition settings, we observe that NLEI-MRF closely approaches the performance of supervised learning.

Index Terms— Quantitative MRI, Magnetic Resonance Fingerprinting, Compressed Sensing, Inverse Problems, Self-Supervised Deep Learning, Equivariant Imaging

1. INTRODUCTION

Magnetic Resonance Fingerprinting (MRF) [1], is an accelerated Quantitative MRI (QMRI) method, for the acquisition of multi-parametric quantitative bio-property maps (QMaps) of the tissues, in a single time-efficient scan. The reduced acquisition times are due to aggressive spatiotemporal subsampling, which leads to aliasing artefacts in the MRF image time-series data, and as a result QMaps. Current state-of-the-art for MRF image reconstruction use supervised deep learning e.g. [2, 3], for which training requires pairs of under-sampled, aliasing-contaminated MRF data, and their corresponding alias-free QMaps as ground truth. However, relying on ground truth is challenging, as: i) obtaining them requires long, clinically-infeasible scans, ii) long acquisitions are susceptible to motion artefacts (and correcting these may introduce interpolation artefacts), and iii) there is no real ground truth: each method for estimating ground truth QMaps depends on its own measurement effects and reconstruction artefacts, hence can only be considered a reference rather than real ground truth.

Therefore, an alternative approach to supervised learning which does not rely on ground truth during training, would be highly beneficial. This work proposes a self-supervised deep learning approach, NonLinear Equivariant Imaging for MRF (NLEI-MRF), to enable MRF quantitative mapping (reconstruction of QMaps), using only fast compressive MRF scans as training data, without requiring ground truth QMaps. We also apply linear Equivariant Imaging (EI) [4], the foundation for NLEI-MRF, for the first time to the MRF reconstruction problem. For competing algorithms, see [5, 6, 7]. NLEI-MRF learns a reconstruction mapping for the MRF nonlinear inverse problem by incorporating spatiotemporal priors from the invariance of spatial transformations (e.g. rotations, flips) on estimated QMaps, and additionally (unlike EI) a differentiable model for the nonlinear Bloch response temporal dynamics approximated by a pre-trained neural network, BlochNet [8]. Tested retrospectively on two distinct MRF acquisitions, we observed that NLEI-MRF (self-supervised learning) closely approached the performance of supervised learning, despite not using ground truth during training.

2. THE MRF INVERSE IMAGING PROBLEM

MRF adopts a spatiotemporal compressed sensing acquisition:

\[ y \approx Ax(q) \]

where \( y \in \mathbb{C}^{m \times T} \) are \( m \) k-space measurements taken at \( T \) timeframes, and \( q = \{T1, T2, PD\} \) are the unknown QMaps i.e. \( n \times 3 \) images of the tissues’ \( T1 \) and \( T2 \) relaxation times and Proton Density (PD) across \( n > m \) voxels. The linear acquisition operator \( A : \mathbb{C}^{n \times T} \rightarrow \mathbb{C}^{m \times T} \) models Fourier subsampling according to a set of temporally-varying k-space locations in each timeframe, combined with a temporal-domain SVD dimensionality reduction scheme [9, 3] i.e., \( 3 < t < T \). The Time-Series of Magnetisation Images (TSMI) for \( n \) voxels and \( t \) dimension-reduced timeframes are denoted by \( x \in \mathbb{C}^{n \times t} \). The TSMIs’ magnetisation responses (fingerprints) per voxel \( x_v \) are nonlinearly related to the tissue properties \( T1 \) and \( T2 \) relaxation times by the solutions of the Bloch differential equations, \( B \), scaled by the Proton Density, PD [1, 10]:

\[ x_v \approx PD_v B(T1_v, T2_v) \]

The compressive nature of the acquisitions makes the estimation of QMaps \( q \) from the undersampled MRF measurements \( y \) a non-linear ill-posed inverse problem (1).

The Bloch model can temporally constrain (1), but alone is absent of spatially-constraining priors to make the inverse problem well-posed. While supervised deep image reconstruction models can learn effective spatial priors from ground truth QMaps, i.e. inter-dependencies across image voxels, they would impose a significant scan-time challenge, as mentioned. We therefore build on the EI self-supervised learning framework to obtain a set of more generic (but still effective) spatially-constraining image priors, with the advantage of using only fast compressed MRF scans as training data.

3. EQUIVARIANT IMAGING

Equivariant Imaging [4] exploits the assumption that an image (to be reconstructed) is invariant to certain types of transformations, e.g. reflections and rotations, in order to train a deep image reconstruction model \( f \) in a self-supervised fashion. This is done by applying

---

KF, KYC and MG are with the Department of Computer Science at the University of Bath, UK. (KF432@bath.ac.uk). CMP and MIM are with GE Healthcare, Germany. MIM is also with the Department of Physics at the Technical University of Munich, Germany and AImotion Bavaria at the Technische Hochschule Ingolstadt, Germany. MG is also with the Department of Engineering Mathematics at the University of Bristol, UK.
the acquisition operator $A$, on the reconstructed transformed images obtained from $f$, to yield new observations (k-space measurements) with information outside the range of the original observations. The new observations, once reconstructed by the same model $f$, must result in a transformed image compared to the original reconstruction.

While the EI idea [4] works for linear inverse problems, i.e. estimating TSMIs ($x$) and not QMaps ($q$) in (1), it can be suboptimal for MRF by neglecting temporally-constraining Bloch response priors that nonlinearly relate the TSMIs to lower-dimensional QMaps. Our NLEI-MRF algorithm builds on EI to additionally incorporate the Bloch priors (2) to estimate the QMaps in the nonlinear problem (1). Fig.1 shows the NLEI-MRF training pipeline.

**NonLinear Equivariant Imaging:** The reconstruction model $f(y) : A^H y \rightarrow q$ is a U-Net CNN following [4], which learns a spatiotemporal mapping from aliased TSMIs, obtained from back-projected k-space measurements ($A^H y$), to the artefact-free QMaps. $A^H$ is the adjoint of $A$, and our experiments used the temporal dimension $t = 10$ for the complex-valued TSMIs, leading to 20ch stacked real and imaginary parts for the inputs of $f$, whereas outputs are 4ch QMaps of T1, T2, PD$_{real}$ and PD$_{imag}$. Training uses a weighted sum of two MSE losses $L_{MC} + \alpha L_{EI}$ ($\alpha > 0$).

- **Measurement Consistency (MC) loss, $L_{MC}(y_{MC}, y)$:** is a routinely-used loss in compressed sensing literature, first applied to deep MRF in [8] for minimising discrepancies between scanner k-space measurements $y$, and those obtained from the reconstructed QMaps $q := f(y)$, following the nonlinear forward model (1). To be specific, a pre-trained BlochNet model $B$, which approximates (2), was used to map $q$ to TSMIs $x(q) \approx B(q)$ with PD$_{real}$ and PD$_{imag}$ used to obtain complex-valued TSMIs, followed then by the compressed subsampling operator $A$, to obtain k-space data $y_{MC} := A \circ B(q)$. Minimising $L_{MC}$ enables $f$ to find an inverse mapping for $A \circ B$, where the reconstructed QMaps respect the forward model physics.

- **EI loss $L_{EI}(q_{EI}, q_{T})$:** minimises discrepancies between reconstructed $q$, and the spatially-transformed reconstructed QMaps $q_{EI}$, (Fig.1). To be specific, spatial transformations $T$, are applied to the QMaps $q := f(y)$, reconstructed from the original (scanner) k-space data. An approximate $A \circ B$ of the nonlinear forward model (1) is applied to $q_T := T(q)$, to obtain new k-space measurements, which were then reconstructed by $f$ into $q_{EI} := f \circ A \circ B \circ T(q)$. The EI loss enables $f$ to learn a reconstruction mapping in a self-supervised manner that respects the image invariance properties, i.e. $f$ should learn that $q$ and $q_{EI}$ are only different by a transformation: $q_T \approx q_{EI}$.

**Linear Equivariant Imaging:** The linear EI algorithm [4] for MRF can be reduced from the NLEI-MRF pipeline in Fig.1: (i) let $f(y) : A^H y \rightarrow x$ reconstruct a TSMI, and (ii) remove the BlochNet $B$ (diamond-shapes in Fig.1) responsible for the forward model’s non-linearity. The result is the EI algorithm to reconstruct an artefact-reduced TSMI $x$, albeit uninformed/unconstrained by the Bloch response priors. An MRF dictionary-matching step [1] then can be used to estimate QMaps from the EI-reconstructed TSMI.

**BlochNet, $B$ [8]:** approximates (2) by a differentiable neural network model, that is kept frozen and used within the NLEI-MRF’s training pipeline (Fig.1) to add temporal Bloch response priors. Implemented by a CNN of 2 hidden layers (each with 300 filters, ReLU activations), BlochNet uses $1 \times 1$ filters to process QMaps (input) in a voxel-wise manner and output the corresponding TSMI i.e. $x(q) \approx B(q)$. This network is trained offline from NLEI-MRF. Training data uses an SUV dimension-reduced ($t = 10$) FISP-MRF dictionary [9] with 94,777 fingerprints, that are simulated Bloch responses (using EPG simulator [11]) for combinations of T1/T2 values in a logarithmically-sampled grid ($T1, T2) \in [0.01, 6] \times [0.004, 4]$ (sec).

**Transformations, $T$:** An important component of EI is the selection of appropriate transformations to learn image invariances. For compressed sensing MRI, randomly selected rotations have been successfully used for EI [7], while shift transformations have no benefit for Fourier based acquisitions [4]. For our work, NLEI-MRF and EI use a random selection from 7 transformations defined by combinations of 90° rotations and vertical flips, accounting for all orientations: 1) Vertical Flip, 2) 90° Rotation, 3) Vertical Flip with 90° Rotation, 4) 180° Rotation, 5) Vertical Flip with 180° Rotation, 6) 270° Rotation and 7) Vertical Flip with 270° Rotation. We limit rotation angles to multiples of 90° to prevent interpolation artefacts.

### 4. NUMERICAL EXPERIMENTS

**Dataset:** We used a dataset of T1, T2 and PD QMaps of 2D axial brain scans from 8 subjects across 15 slices. Complex-valued TSMIs and k-space MRF data were retrospectively simulated from these QMaps via (2) and (1), respectively. The reference QMaps and TSMIs had spatial dimensions of $n = 224 \times 224$ pixels with head-masks applied, which were generated using the Proton Density. For the Bloch response model a truncated FISP-MRF protocol [10] was used with $T = 200$ repetitions i.e. 5 times less (accelerated) than the original FISP-MRF. For MRF k-space data, we simulated single-coil acquisitions using two distinct cartesian (FFT) k-space subsampling patterns evolving across each temporal frame: (i) a rotating Spiral, as in [8] and (ii) shifting horizontal lines using multi-shot Echo Planar Imaging (EPI) [12]. We sampled k-space locations in each timeframe corresponding to a spatial compression ratio of 81:1 for Spiral and 74:1 for EPI. The TSMIs were dimension-reduced ($t = 10$) following [9]. The dataset was split into 105 slices from 7 subjects for training, and 15 slices from the 8th subject for testing.
Fig. 2: Tissue Map Results using Spiral Subsampling for Slice 10 of 15, with MAPE (%) for T1 and T2, and PSNR (dB) for PD.

**Tested Algorithms:** We compared the performance of the proposed NLEI-MRF to SVD-MRF [9], EI [4] and RCA-U-Net [2] baselines. SVD-MRF is a non-data driven approach which reconstructs backprojected (aliased) TSMIs, $A^H y$, from k-space data, followed by MRF dictionary-matching to estimate QMaps. RCA-U-Net is a state-of-the-art deep supervised learning MRF model that uses pairs of aliased TSMIs, $A^H y$, and ground truth QMaps for training. Separate RCA-U-Net models were trained for T1, T2, $PD_{real}$ and $PD_{mag}$ following [2] using 1000 epochs, L1 loss, and linearly decaying learning rate from 200 to 1000 epochs. On the other hand, NLEI-MRF and EI are self-supervised learning models that only use aliased backprojected TSMIs, $A^H y$, from under-sampled MRF k-space data, and do not use ground truth for training. NLEI-MRF and EI used the U-Net as in [4], without the residual connection between the initial and final layers, and trained using 1000 epochs, batch size 2, Adam optimiser, weight decay $10^{-8}$, initial learning rate $5 \times 10^{-4}$ decreasing by factor 10 at 300 epochs. We used 3 randomly selected transformations per iteration from the 7 previously defined, and applied them across each batch to create new batch sizes of 6. Optimal values for $\alpha$ were found experimentally: $10^{-8}$ for NLEI-MRF Spiral, $10^{-4}$ for NLEI-MRF EPI, $10^{-5}$ for EI Spiral, and $10^{-2}$ for EI EPI.

**Evaluation Metrics:** We used the Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Peak Signal-to-Noise-Ratio (PSNR) and Structural Similarity Index Measure (SSIM). Head-masks were applied to all reconstructions and metrics were then calculated and averaged across 15 test slices.

**Results and Discussion:** Metrics in Table 1, for both subsampling patterns, show a clear progression in performance from SVD-MRF to EI, NLEI-MRF and RCA-U-Net, with NLEI-MRF (self-supervised model) approaching the performance of the supervised model RCA-U-Net. The results in Fig.2 show similar performance for EI and NLEI-MRF, while Fig.3 shows NLEI-MRF outperforms SVD-MRF and EI, while being close to RCA-U-Net. SVD-MRF, EI and NLEI-MRF exhibit blurring, which can also be seen in RCA-U-Net to a lesser extent. This is intrinsic to spiral subsampling due to prioritised sampling of low frequencies at the centre of k-space.
Table 1: Metrics averaged over 15 test slices, show increasing performance for each acquisition scheme, moving from SVD-MRF [9], EI [4], NLEI-MRF (ours) to RCA-U-Net [2]. NLEI-MRF (self-supervised learning) performs closest to RCA-U-Net (supervised learning).

| Nature   | SVD-MRF | EI  | NLEI-MRF | RCA-U-Net |
|----------|---------|-----|----------|-----------|
| MAE (s)  |         |     |          |           |
| T1       | 0.1371  | 0.0577 | 0.0532   | 0.0073    |
| T2       | 0.5048  | 0.0178 | 0.0162   | 0.0134    |
| MAPE (%) |         |     |          |           |
| T1       | 12.2519 | 5.0527 | 4.2236   | 3.8465    |
| T2       | 34.9255 | 9.3400 | 8.6272   | 7.3329    |
| PSNR (dB)|         |     |          |           |
| TSMI     | 11.3135 | 26.0396 | -       | -        |
| T1       | 23.9384 | 33.2161 | 33.6557 | 34.7350  |
| T2       | 25.7089 | 34.8714 | 36.7658 | 38.3916  |
| SSIM     |         |     |          |           |
| TSMI     | 0.5979  | 0.7721 | -        | -        |
| PD       | 0.8469  | 0.9154 | 0.9537   | 0.9688   |
| T2       | 0.7676  | 0.8896 | 0.8913   | 0.9224   |
| T1       | 0.1895  | 0.9409 | 0.9425   | 0.9558   |

5. CONCLUSION

A proof-of-concept for a self-supervised learning approach (NLEI-MRF) for MRF multi-parametric quantitative tissue mapping was proposed. The method was validated on two cartesian (FFT) k-space sampling patterns on retrospectively simulated MRF data. NLEI-MRF’s performance was observed to approach the state-of-the-art supervised learning method, despite not using ground truth for training. Future work will include extensions to address noisy, non-cartesian acquisitions from prospective in-vivo scans.

6. COMPLIANCE WITH ETHICAL STANDARDS

This research study was conducted retrospectively using anonymised human subject scans made available by GE Healthcare who obtained informed consent in compliance with the German Act on Medical Devices. Approval was granted by the Ethics Committee of The University of Bath (Date. Sept 2021 / No. 6568).

7. ACKNOWLEDGMENTS

CMP and MIM are supported by the EU’s Horizon 2020 (grant No. 952172). MG is supported by the EPSRC grant EP/X001091/1.

8. REFERENCES

[1] D. Ma et al., “Magnetic resonance fingerprinting,” Nature, vol. 495, no. 7440, pp. 187–192, 2013.
[2] Z. Fang, Y. Chen, D. Nie, W. Lin, and D. Shen, “Rca-u-net: Residual channel attention u-net for fast tissue quantification in magnetic resonance fingerprinting,” in MICCAI, 2019, pp. 101–109.
[3] M. Golbabaee et al., “Compressive mri quantification using convex spatiotemporal priors and deep encoder-decoder networks,” Medical Image Analysis, vol. 69, pp. 101945, 2021.
[4] D. Chen, J. Tachella, and M. E. Davies, “Equivariant imaging: Learning beyond the range space,” in 2021 IEEE/CVF Intl. Conf. on Computer Vision (ICCV), Oct 2021, pp. 4359–4368.
[5] M. Gao, H. Ye, T. H. Kim, Z. Zhang, S. So, and B. Bilgic, “Accurate parameter estimation using scan-specific unsupervised deep learning for relaxometry and mr fingerprinting,” 2021.
[6] J. I. Hamilton, “A self-supervised deep learning reconstruction for shortening the breathhold and acquisition window in cardiac magnetic resonance fingerprinting,” Frontiers in Cardiovascular Medicine, vol. 9, 2022.
[7] D. Chen, J. Tachella, and M. E. Davies, “Robust equivariant imaging: a fully unsupervised framework for learning to image from noisy and partial measurements,” in 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2022, pp. 5637–5646.
[8] D. Chen, M. E. Davies, and M. Golbabaee, “Compressive mr fingerprinting reconstruction with neural proximal gradient iterations,” in MICCAI, 2020, pp. 13–22.
[9] D. F. McGivney et al., “Svd compression for magnetic resonance fingerprinting in the time domain,” IEEE Trans. Med. Imag., vol. 33, no. 12, pp. 2311–2322, 2014.
[10] Y. Jiang, D. Ma, N. Seiberlich, V. Gulani, and M. A. Griswold, “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.
[11] M. Weigel, “Extended phase graphs: dephasing, rf pulses, and echoes-pure and simple,” Journal of Magnetic Resonance Imaging, vol. 41, no. 2, pp. 266–295, 2015.
[12] A. J. V. Benjamin et al., “Multi-shot echo planar imaging for accelerated cartesian mr fingerprinting: an alternative to conventional spiral mr fingerprinting,” Magnetic resonance imaging, vol. 61, pp. 20–32, 2019.