Deep learning for Magnetic Resonance Fingerprinting

Mohammad Golbabaee*, Dondong Chen†, Mike Davies†, Carolin M. Pirkl*, Marion I. Menzel† and Pedro Gómez*<br><br>Abstract—Current popular methods for Magnetic Resonance Fingerprint (MRF) recovery are bottlenecked by the heavy storage and computation requirements of a dictionary-matching step due to the growing size and complexity of the fingerprint dictionaries in multi-parametric quantitative MRI applications. We study a deep learning approach to address these shortcomings. The MRF-Net provides a piecewise affine approximation to the (temporal) Bloch response manifold projection. Fed with non-iterated back-projected images, the network alone is unable to fully resolve spatially-correlated artefacts which appear in highly undersampling regimes. We propose an accelerated iterative reconstruction to minimize these artefacts before feeding into the network. This is done through a convex regularization that jointly promotes spatio-temporal regularities of the MRF time-series.

I. INTRODUCTION

Magnetic Resonance Fingerprinting (MRF) [1] recently emerged to accelerate the acquisition of quantitative NMR characteristics based on using i) short excitation pulses which simultaneously encode multitude of NMR parameters, and ii) significantly undersampled k-space data. To overcome the lack of sufficient spatio-temporal information MRF incorporates a physical model based on exhaustively simulating a large dictionary of magnetic responses (fingerprints) for all combinations of the quantized parameters. As occurs to any multi-parametric manifold enumeration, the main drawback of such an approach is the size of this dictionary which grows exponentially in terms of the number of parameters. Dictionary-matching (DM) approaches therefore do not scale well to the complexity of the emerging multi-parametric quantitative MRI problems.

II. DEEP LEARNING FOR BLOCH MANIFOLD PROJECTION

Deep learning (DL) methodologies have recently been introduced to overcome this problem [2, 3, 4]. Time-series of Back-Projected Images (BPI) are fed into a compact neural network which temporally processes voxel sequences and approximates the DM step to output the parametric maps. For instance our proposed MRF-Net (Figure 2) is able to accurately approximate the DM step by saving more than 60 times in memory and computations [4]. The MRF dictionary is only used for training and not during parameter recovery. Figure 1 shows that the network provides a piecewise affine approximation to the Bloch response manifold projection and that rather than memorizing the dictionary, it efficiently clusters this manifold and learns a set of hierarchical matched-filters for affine regression of the NMR characteristics in each segment (for more details see [4]).

III. OUR PARAMETER ESTIMATION PIPELINE

Trained by independently corrupted (i.i.d. Gaussian) noisy fingerprints, the MRF-Net acts only along temporal domain and is unable to correct for dominant spatially-correlated (alising) artefacts appearing in BPIs in highly undersampled regimes. Also larger DL models aiming to learn spatio-temporal data structures are prone to overfitting due to the limited access to properly large ground-truth parametric maps in practice. Further, such approaches build customized denoisers which require expensive re-training by changing sampling parameters i.e. the forward model. We address these shortcomings by taking a dictionary-free compressed sensing approach to spatio-temporally process data before feeding into the compact and easily-trained MRF-Net.

The undersampled k-space measurements $Y \in \mathbb{C}^{n \times L}$ acquired across $L$ timeframes are first processed by solving the following convex and dictionary-free regularized problem [5]:

$$\hat{X} = \arg\min_X \|Y - A(XV^H)\|_2^2 + \lambda \sum_{i=1}^{S} \|X_i\|_{TV} \quad (P1)$$

in order to find $S \ll L$ principal/subspace images $X \in \mathbb{C}^{n \times S}$. The subspace bases $V_i \in \mathbb{C}^{L \times S}$ are the $S$ leading (left) SVD components of a large-size MRF dictionary $D \in \mathbb{C}^{L \times d}$ ($L \ll d$) used here for (unsupervised) dimensionality reduction. The forward operator $A$ models the multi-coil sensitivities and the per-frame subsampled 2D Fourier Transforms. The low-rank subspace model is a convex (in fact linear) relaxed representation of the temporal dictionary responses and when accurate enough, it is computationally advantageous over the full image representation $X^{Full} \approx X^S V^H \in \mathbb{C}^{n \times L}$ because it reconstructs smaller objects and promotes temporally low-rank structures [6]. This prior alone is, however, insufficient to obtain artefact-free solutions e.g. when using spiral readouts [7]. We additionally use the Total Variation (TV) regularization to promote spatial smoothness across recovered subspace images. (P1) can be efficiently solved using FISTA algorithm [8]. Fed with the iteratively reconstructed images $X$, the MRF-Net processes each (normalized) voxel sequence and outputs per-voxel quantitative parameters.

IV. NUMERICAL RESULTS

Methods are tested on a healthy human brain acquired using the Steady State Precession (FISP) sequence in [9] and spiral readouts which sample $m = 732$ k-space locations in each of the $L = 1000$ time-frames in order to reconstruct $n = 256 \times 256$ resolution parametric T1 and T2 maps. We simulate a dictionary of $d = 1137460$ fingerprints for combinations of $T1 \in [0.1, 4]$ sec, $T2 \in [0.02, 0.6]$ sec. Clean fingerprints are used for unsupervised subspace model learning of sufficiently low-rank ($S = 10$). Further, fingerprints corrupted by additive white Gaussian noise (data augmentation by factor 100) supervisedly train the dimension-reduced MRF-Net on a standard CPU desktop.

We compare three methods for reconstructing subspace images before feeding to the MRF-Net: non-iterative BPIs i.e. $\hat{X} := A^H(Y)V\hat{w}$, and iterative reconstructions incorporating i) only the low-rank (LR) subspace prior by solving (P1) with $\lambda = 0$, and iii) joint TV and subspace spatio-temporal priors (LRTV) by solving P1 with an experimentally tuned $\lambda = 2 \times 10^{-5}$. Note that the BPIs are the first iteration of the LR. Figure 3 shows the reconstructed maps for scanner data. Undersampling artefacts are visible in BPI+MRF-Net. The subspace iterations of LR+MRF-Net also admit undesirable solutions with high-frequency artefacts due to the insufficient measurements collected from the k-space corners in spiral readouts (for details see [7]). By adding sufficient spatial regularization, the proposed LRTV+MRF-Net outputs artefact-free maps within 8-12 iterations.
Fig. 1: MRF-Net hierarchically segments the input space and learns a piece-wise affine mapping between input-outputs for each segment [4]. Segments only created by the layers \( h^{(2)} \) and \( h^{(3)} \) are mapped to the \( T_1, T_2 \) grid used for generating the inputs (fingerprints) and are visualised in the first two figures from the left. The deep hidden layer \( h^{(3)} \) creates more abstract (coarse) segments. Finally the end-to-end segments which are the intersection of all layers’ segments are visualised in the third figure from the left. The end-to-end segments on the manifold of Bloch responses are also depicted across the three dominant principal components (right image).

Fig. 2: Illustration of the MRF-Net: Inputs \( h^{(1)} \) are the voxel sequences of the subspace image reconstructed by (P1) and outputs \( h^{(4)} \) are the per-voxel \( T_1 \) and \( T_2 \) parameters. The MRF-Net has implicitly 4 layers by including the unsupervisedly learned subspace projection (first layer in grey) incorporated in solving (P1). Thanks to this dimensionality-reduction, MRF-Net requires less units and training resources compared to the uncompressed DL approaches [2, 3]. Three last layers use nonlinear ReLU activations (orange) and are supervisedly trained by standard backpropagation to approximate subspace dictionary matching.

Fig. 3: Reconstructed \( T_1 \) (top row) and \( T_2 \) (bottom row) maps for healthy volunteer data using MRF-Net fed with the non-iterated BPIs (left column), iteratively reconstructed images with only low-rank (LR) subspace prior (middle column), and iteratively reconstructed images with joint TV and low-rank subspace (LRTV) priors (right column).

REFERENCES


