Data-Driven Joint Optimization of Acquisition and Reconstruction of Quantitative MRI

by

Jiaren Zou

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Biomedical Engineering)
in the University of Michigan
2024

Doctoral Committee:

Professor Yue Cao, Chair
Professor Jeffrey Fessler
Professor Douglas Noll
Professor Nicole Seiberlich

Jiaren Zou

jiarenz@umich.edu

ORCID iD: 0000-0002-8356-1967

Acknowledgements

My doctoral studies turn out to be a unique life experience that reshapes my values, my way of thinking, and my view of the world. This journey is long and challenging, yet immensely rewarding. First and foremost, I would like to express my deepest gratitude to Professor Yue Cao, who encouraged me to pursue doctoral studies in the first place and provided continuous guidance and support throughout my Ph.D. especially in the uncertain times of the pandemic. Through her, I learned how to be a researcher and how to tackle various difficulties in research with rational strategies and scientific rigor.

I would like to thank Professor Jeffrey Fessler for his insightful thoughts and help on my projects. His courses and knowledge on signal and image processing have built my mathematical foundation in MR image reconstruction.

I am also grateful to Professor Douglas Noll for his support and advice which helped shape my dissertation. His course on medical imaging was my initial exposure to the fascinating field of MRI.

I am also thankful to Professor Nicole Seiberlich for providing insights and comments and critical support on MR fingerprinting projects. I also learnt a lot on dynamic MRI and non-Cartesian image reconstruction from the discussion with her.

Additional thanks to Professor James Balter for providing continuous support and advice, Dr. Michelle Mierzwa for providing clinical data on head and neck cancer, Professor Yun Jiang, Dr. Yuchi Liu, Professor Jesse Hamilton, Sydney Kaplan for their assistance, advice, and

collaboration on MR fingerprinting projects. Particular thanks to Josiah Simeth, Yuan Li, Yuhang Zhang, Robert Jones, Daekeun You and Madhava Aryal for their support to my research as lab mates.

Most importantly, I am deeply grateful to my parents and my girlfriend for their unfailing love and support. Without them, it is impossible for me to complete this thesis.

This work is supported by NIH grant of R01 EB016079.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AIF	Arterial Input Function
BAT	Bolus Arrival Time
CA	Contrast Agent
CNN	Convolutional Neural Network
CNR	Contrast-to-Noise Ratio
CRNN	Convolutional Recurrent Neural Network
CS	Compressed Sensing
DC	Data Consistency
DCE MRI	Dynamic Contrast Enhanced Magnetic Resonance Imaging
eTofts	Extended Tofts Model
LLSQ	Linear Least-Squares
LSTM	Long Short-Term Memory
MBDL	Model-Based Deep Learning
MRF	Magnetic Resonance Fingerprinting
MSE	Mean Squared Error
NLLSQ	Nonlinear Least-Squares
NRMSE	Normalized Root Mean Squared Error
PK	Pharmacokinetic
PSF	Point Spread Function
PSNR	Peak Signal to Noise Ratio

qMRI	Quantitative Magnetic Resonance Imaging
R	Undersampling Rate
RNN	Recurrent Neural Networks
SNR	Signal-to-Noise Ratio
SP	Sampling Pattern
SSIM	Structural Similarity Index Measure
T_1	Longitudinal Relaxation Time
T_2	Transverse Relaxation Time
TE	Echo Time
TI	Inversion Time
TR	Repetition Time

Abstract

Quantitative magnetic resonance imaging (qMRI) measures physical, physiological or biological properties of tissues and thus provides reproducible imaging biomarkers for disease diagnosis and therapy response monitoring. However, long scanning and reconstruction time, low reproducibility, spatial resolution, and volume of coverage limit the clinical translation of qMRI.

The overall goal of this dissertation is to improve qMRI by exploiting its sparsity by data-driven deep learning methods. Such methods can provide more accurate and precise tissue parameters from highly undersampled or accelerated scans using a fraction of reconstruction time of conventional methods. Sampling patterns, image reconstruction and parameter estimation could be jointly optimized to directly minimize parameter quantification error under the same scan time.

Temporal sparsity in dynamic contrast enhanced magnetic resonance imaging (DCE MRI) was exploited by a long short-term memory (LSTM) neural network-based approach to provide more robust tissue parameter estimation. The network was trained on simulated DCE signals and tested on both simulated and real data. Compared to a conventional linear least squares (LLSQ) fitting method, the LSTM-based approach had higher accuracy for the data with temporally subsampling, total acquisition time truncation, or high noise level. Also, the LSTM-based method reduced the inference time by ~14 times compared to the LLSQ fitting. Validation of the method on real data demonstrated its clinical feasibility to provide high-quality tissue parameter maps.

Beyond temporal sparsity, the spatiotemporal sparsity of DCE MRI was further exploited by convolutional recurrent neural network. 2D Cartesian phase encoding k-space subsampling patterns were jointly optimized with image reconstruction to identify the most informative k-space data to acquire beyond the learned population prior knowledge. Both reconstruction image quality and parameter estimation accuracy were used to guide network training. The proposed method was trained and tested by multi-coil complex digital reference objects of DCE images. The proposed method achieved lower parameter estimation error at four temporal resolutions (2s, 3s, 4s, and 5s) compared with two benchmark methods and reduced parameter estimation bias and uncertainty in tumor regions at temporal resolution of 2s. The proposed method also showed robustness to contrast arrival timing variations across patients.

Compared with Cartesian sampling, non-Cartesian sampling provides more flexibility in trajectory design. This work also develops a deep learning framework that is able to synergistically optimize rotation angles of 3D spiral trajectories, image reconstruction, and parameter estimation of magnetic resonance fingerprinting. To counter the large problem size, an efficient model-based deep learning (MBDL) image reconstruction framework was developed. The MBDL image reconstruction provided more accurate parameter estimation than a state-of-the-art reconstruction method on both simulated and *in vivo* data. On simulated data, joint optimization of image-parameter reconstruction or sampling trajectory-image reconstruction were incorporated into the baseline MBDL framework and further improved tissue parameter estimation.

Chapter 1 Introduction

1.1 Motivation

Magnetic resonance imaging (MRI) is a non-invasive imaging modality with excellent soft-tissue contrast and no ionizing radiation. More importantly, MRI has flexibility of providing contrasts associated with various anatomic and physiological information for both disease diagnosis and therapy response monitoring. Conventional contrast-weighted MR images are routinely used in clinics. For example, with contrast agent, T_1- and T_2-weighted images are used respectively for identification of high-grade brain tumors and pathological tissues in liver and spleen [1]. However, conventional contrast-weighted MR images are sensitive to various confounding factors and lead to inter-scanner and across-center variations of images, which have an impact on diagnosis accuracy.

Quantitative MRI (qMRI) measures underlying physical, physiological or biological properties of tissues directly and thus could provide reproducible imaging biomarkers. Various qMRI techniques have been developed with a wide range of clinical applications and demonstrate high diagnostic power [2]–[4]. However, in general, qMRI requires longer scanning and reconstruction time, limiting its clinical translation. In practice, one must balance high spatiotemporal resolution, volume coverage, signal-to-noise ratio and quantitative parameter accuracy and precision against limited scanning times.

qMRI usually requires long scanning time to acquire images of the same anatomical structures with varying parameters for multiple contrasts or with the same parameters for dynamic courses of a contrast uptake. However, these acquired images contain a high degree of

redundant information due to spatial, temporal and/or contrast sparsity. This sparsity can be learned by statistical methods and utilized as prior knowledge complimentary to acquired data that can then be undersampled. The extraction of robust and accurate sparse representations of qMRI and the efficient incorporation of such prior knowledge into image reconstruction are thus promising research areas to reduce scan time.

Compressed sensing [5] is one of the ways to exploit such sparsity. It incorporates prior knowledge of the spatiotemporal sparsity of the underlying MR image-time series into the reconstruction from highly accelerated scans by using hand-crafted regularizations, for example, spatiotemporal total variation (TV) [6], low rank [7], and locally low rank [8], [9]. Combining with parallel imaging [10], [11] that exploits additional spatial information provided by multiple coils, compressed sensing has achieved promising results in accelerating MR scans. However, these regularizations are empirically designed and may not fully capture the sparsity in anatomy of interest and dynamics of the data. More recently, deep learning has shown state-of-the-art image reconstruction performance from shortened scans by learning regularizations tailored to the anatomy and dynamics in MRI signals [12]–[16]. Direct mapping from k-space data to MR images by fully connected and convolutional neural networks are explored [17], followed by methods working in image space using a fully convolutional network with one data consistency step at the end [18]. However, due to the high-dimensional nature of MRI reconstruction, large training dataset is required which is not readily available in MRI. More recently, methods based on interleaves of neural networks and data consistency steps better leverage learned priors by mimicking conventional optimization algorithm of model-based reconstruction and have achieved the state-of-the-art robustness and image quality [15] even with small training dataset.

K-space sampling patterns determining where to acquire the data in the Fourier domain have critical impacts on the quality of reconstructed MR images. While previous methods have optimized sampling patterns independent of reconstruction methods [19], [20], deep learning has recently emerged as a data-driven approach to optimize the two (sampling and reconstruction) together in a large space with high degrees of freedom and achieves state-of-the-art image reconstruction performance for static MRI [21], [22]. Deep learning-based joint optimization of sampling and reconstruction for qMRI, although holding great promise, has not been explored due to several reasons. 1) The data acquired for qMRI not only contains anatomical sparsity but also temporal and contrast sparsity, which can be explored to reduce required samplings dramatically; 2) error propagation from image reconstruction to parameter quantification can be directly minimized to improve accuracy in the final parameter estiamtion; 3) parameter estimation algorithms can be jointly optimized along with sampling and image reconstruction, which further enlarges the optimization space.

In this dissertation, we explore the concept of optimizing the entire pipeline of qMRI from k-space sampling patterns to image reconstruction to parameter estimation for reducing scanning time and increasing reproducibility and accuracy of the final estimated quantitative parameters. We utilize dynamic contrast enhanced (DCE) MRI that contains spatial and temporal sparsity and 3D magnetic resonance fingerprinting (MRF) that consists of spatial, temporal and contrast sparsity as platforms to test the proposed concept. Brief introductions of DCE MRI and MRF are provided in the following subsections.

1.2 Dynamic Contrast Enhanced Magnetic Resonance Imaging (DCE MRI)

DCE MRI is an imaging technique that acquires a time-series of T_1 weighted images before, during and after a bolus administration of a contrast agent (CA) to enable quantitative

physiological parameter extraction from a pharmacokinetic (PK) model. The extracted parameters, e.g., blood volume and blood flow, could help assess the histological grade of tumors [23], differentiate tumors from normal tissue [24], and monitor as well as predict cancer response to therapy [25]. For example, a previous work shows that the tumor subvolume with low blood volume derived from DCE MRI has been used as a radiation boost target in a clinical trial of head and neck cancers, which shows improved tumor local control [26].

The extended Tofts (eTofts) model [27] was used in this dissertation to quantify physiological parameters, including the transfer constant K^{trans} of the CA that diffuses from blood vessels to the interstitial space, the rate constant k_{ep} of the CA efflux from the interstitial space to blood plasma and the fractional volume of blood plasma v_p. Also, a voxel-wise contrast bolus arrival time (BAT), $\tau_{BAT}(\mathbf{r})$ is warranted for accuracy of PK parameter estimation [28]. The implemented eTofts model is written as:

$$C_t(\mathbf{r},t) = K^{trans}(\mathbf{r}) \int_0^t C_p\big(\tau - \tau_{BAT}(\mathbf{r})\big) e^{-K_{ep}(\mathbf{r})(t-\tau)} d\tau + v_p(\mathbf{r}) C_p\big(t - \tau_{BAT}(\mathbf{r})\big), \quad (1.1)$$

where $C_t(\mathbf{r},t)$ is the CA concentration in the tissue voxel, C_p is the CA concentration in the blood plasma, k_{ep} equals to K^{trans}/v_e, where v_e is the fractional interstitial volume.

Although a time-series of images is acquired in DCE MRI, the image frames contain similar anatomical structures, and the contrast changes are relatively smooth in time domain and happen only in part of the anatomy. The spatial and temporal sparsity can be explored jointly to improve DCE MRI reconstruction.

1.3 Magnetic Resonance Fingerprinting (MRF)

MR fingerprinting [29] is a qMRI technique that enables fast simultaneous quantification of multiple tissue parameters for potential disease diagnosis and treatment monitoring. The MRF

sequence design sensitizes MR signals to physiological parameters of interest by using preparation pulses and pseudo-random schedule of flip angles and TRs. Temporal incoherence of MR signals is enhanced by using different spiral readout trajectories across time. While MRF is initially demonstrated for mapping relaxometry parameters, extension works have shown its feasibility for quantitative diffusion [30], chemical exchange saturation transfer (CEST) [31], [32], and microvascular structure [33] imaging. The potential clinical applications of MRF have been investigated for different body sites, including brain [34]–[38], heart [39], liver [40], and abdomen [41]. For example, MRF can improve accuracy of identifying lesions suspicious for hippocampal sclerosis in patients with mesial temporal lobe epilepsy [35]. T_1 and T_2 values derived from MRF can be used to distinguish tumor from normal-appearing white matter, differentiate tumor grade, and identify abnormalities in peritumoral regions [42]. The repeatability and reproducibility of MRF have also been demonstrated [43].

Full 3D non-Cartesian sampling of k-space is highly flexible in trajectory design and can potentially be optimized for MRF to improve efficiency. However, full 3D non-Cartesian sampling based MRF requires extremely high computation power to reconstruct image series and estimate quantitative parameters. Spatiotemporal sparsity of MRF can be exploited to reduce the complexity of the problem.

1.4 Organization of the Dissertation

Chapter 2 introduces a novel neural network approach to estimate the PK parameters by extracting long and short time-dependent features in DCE MRI. A Long Short-Term Memory (LSTM) network, widely used for processing sequence data, was employed to map DCE MRI time-series accompanied with an arterial input function to parameters of the eTofts model. Our study suggests that the LSTM model can achieve improved robustness and computation speed

for PK parameter estimation compared to a conventional parameter fitting method and a convolutional neural network (CNN)-based network, particularly for suboptimal data. The results are published in a paper [44].

Chapter 3 develops and evaluates a deep learning framework that jointly optimizes k-t sampling patterns and image reconstruction for head and neck DCE MRI. The k-space subsampling patterns were jointly optimized in a deep learning-based dynamic MRI reconstruction network. Comparing with conventional iterative reconstruction methods, the proposed method achieved low PK parameter quantification errors at various temporal resolutions and reduced PK parameter estimation bias and uncertainty in tumor regions. The results are published in a paper [45].

Chapter 4 investigates the idea of jointly optimizing the k-space sampling patterns of two highly distinct contrasts to accelerate multi-contrast MRI. The sampling patterns of T_1 and T_2 weighted images were optimized with an image reconstruction framework with decoupled single- and multi-contrast learning modules. The jointly learned sampling patterns outperformed empirical patterns at various undersampling rates in both 1D and 2D imaging. The related results are described in a conference proceeding [46].

Chapter 5 describes an end-to-end framework to jointly optimize image reconstruction, parameter reconstruction and trajectory of full 3D spiral MRF with minimal computation demand. By evaluating the framework on both simulated and *in vivo* datasets, the proposed method improved parameter quantification accuracy and reduced reconstruction time compared with a state-of-the-art reconstruction method. The results are summarized in a manuscript to be submitted as a journal paper.

Chapter 6 summarizes the dissertation and outlines future research directions.

Chapter 2 Estimation of Pharmacokinetic Parameters from DCE-MRI by Extracting Long and Short Time-Dependent Features Using an LSTM Network

2.1 Introduction

The purpose of this chapter is to develop a more time-efficient and more robust method for pharmacokinetic parameter estimation from DCE-MRI compared with a convolutional neural network-based method and a conventional pharmacokinetic model fitting by linear least squares.

Quantification of PK parameter maps is often done in two steps: 1) reconstruction of time series of DCE images, and 2) fitting DCE time series to a PK model, so refereed as an indirect method. The PK model fitting is often done by nonlinear least-squares (NLLSQ) fitting, which is sensitive to DCE data sampling interval, total acquisition time, and noise [47], and requires intensive computation. However, a LLSQ fitting can be used [48] after re-formulating the optimization problem, which is more efficient in computation time and improves accuracy in the estimates, particularly for DCE MRI signals with low signal-to-noise ratio. More recently, an efficient derivative based LLSQ method with a low-pass filter in time domain [49] is introduced. However, these developed methods show degraded performance for noisy and low temporal resolution DCE data.[1]

Recently, machine learning methods have been investigated to learn mapping from fully sampled or subsampled image time-series to the parameter maps utilizing 2D or 3D CNN [50],

[1]This chapter is based on our paper published in *Medical Physics* [44].

[51]. Although having a short inference time, CNN is not designed to learn long and short-term temporal relationships in the hemodynamics of the CA from the image time-series. Recurrent neural networks (RNNs), especially LSTM [52], have been successfully applied to learn temporal relationships in sequence data, such as video description and image captioning [53]. The LSTM has been applied to explore applications in biomedical images. A modified U-net was combined with an LSTM variant for 3D biomedical volume segmentation [54], where the LSTM explored the correlation of slices in the cranial-caudal direction. A pre-trained fine-tuned Visual Geometry Group (VGG) network was used to extract feature maps from DCE MRI slices, and then the sequence of feature maps was processed by an LSTM to determine whether a breast lesion was benign or malignant [55].

It remains as a challenge of how to input an arterial input function (AIF) during training and inferencing a machine learning model for mapping time series of DCE images to PK parameter maps. In general, AIFs vary in shape, peak amplitude, and time delay from subject to subject. The machine learning model could produce biased estimations in PK parameters if a model is trained with insufficient amounts of subject AIFs [51].

Inspired by the recent successes of LSTM and the challenges faced by CNN-based approaches for PK parameter estimation, we proposed a LSTM-based approach to learn the mapping of temporal dynamics in single-voxel signals accompanied, with their corresponding AIFs, to the PK parameters in the extended Tofts model [27]. Our approach is motivated by four factors. First, the signal intensity-time curves of DCE-MRI describe temporal hemodynamics of a CA passing through microvasculature in tissue. The LSTM architecture is able to learn long-term (temporal) dependence of signals [53], [56] and thus could improve performance of PK parameter estimation compared to the CNN-based approaches. Second, in practice, the AIF

varies from patient to patient [57]. Inclusion of a subject-specific AIF in the estimation process could significantly improve estimation performance. AIFs can be readily incorporated into the input of LSTM as another input dimension, removing the bias observed in the reported CNN-based approaches [50], [51]. Third, by capturing a low-dimensional manifold where the tissue concentration-time curve and AIF reside using LSTM, more robust parameter estimation can be achieved. Fourth, the inference time can be reduced because of the small computational burden of the LSTM as compared to direct model fitting (DMF). We compared the results of LSTM with conventional direct model fitting [58] as well as a state-of-the-art CNN-based method [51], including performances on the DCE MRI at low contrast-to-noise ratio, low temporal sampling and short total acquisition time.

2.2 Materials and Methods

Our proposed LSTM-based method treats the PK parameter estimation problem as a mapping from a CA concentration-time curve accompanying an AIF to the underlying PK parameters. Here, we investigated our method in the most commonly used extended Tofts model.

2.2.1 Extended Tofts model

The eTofts model (equation 1.1) was used in this chapter. While equation 1.1 has been fitted using NLLSQ methods [48], [49] previously, we selected a LLSQ method [48] that has a better tolerance to low signal-to-noise ratio (SNR) in the DCE data and a more efficient computation speed than NLLSQ fitting, as a benchmark to compare with our proposed LSTM method. In this LLSQ method, equation 1.1 can be re-written as:

$$C_t(\boldsymbol{r},t) = (K^{trans}(\boldsymbol{r}) + K_{ep}(\boldsymbol{r})v_p(\boldsymbol{r}))\int_0^t C_p\big(\tau - \tau_{BAT}(\boldsymbol{r})\big)d\tau - K_{ep}(\boldsymbol{r})\int_0^t C_t(\boldsymbol{r},\tau)d\tau + v_p(\boldsymbol{r})C_p\big(t - \tau_{BAT}(\boldsymbol{r})\big), \quad (2.2)$$

where K^{trans}, K_{ep}, and v_p are linearly related to integrals of $C_t(\boldsymbol{r}, t)$ and $C_p(t)$, and $C_p(t)$. For given $\tau_{BAT}(\boldsymbol{r})$, K^{trans}, K_{ep}, and v_p can been rapidly estimated by LLSQ fitting. $\tau_{BAT}(\boldsymbol{r})$ can be estimated iteratively with K^{trans}, K_{ep}, and v_p. The $\tau_{BAT}(\boldsymbol{r})$ range can be determined using priori knowledge, e.g., 0-10 seconds for the tissue in head and neck regions. In our implementation, we tested the $\tau_{BAT}(\boldsymbol{r})$ values between 0 and 10 sec with an incremental step 1 sec [47]. Hereafter, we refer this implementation of the LLSQ fitting as DMF.

2.2.2 PK parameter inference via LSTM.

2.2.2.1 Formulation

We estimate the PK parameters by mapping $(C_t(\boldsymbol{r}, t), C_p(t))$ to the underlying physiological parameters $\theta = (K^{trans}(\boldsymbol{r}), v_e(\boldsymbol{r}), v_p(\boldsymbol{r}))$ using LSTM, which we denote as $\theta = f\left((C_t(\boldsymbol{r}, t), C_p(t))\right|\Theta)$, where $f(\cdot\,|\Theta)$ represents the forward mapping of the LSTM network parameterized by Θ.

2.2.2.2 Loss function

Our loss function seeks to reduce the mean squared error (MSE) between the estimated parameters θ_e and the ground truth parameters θ_g corresponding to the training signal series. Given a set of N training samples $(c_{t_i}(t), c_{p_i}(t), \theta_{g_i})$, i = 1, 2, …, N, we train the LSTM network to minimize the following loss function:

$$\mathcal{L}(\Theta) = \frac{1}{N}\sum_{i=1}^{N} \left\| \theta_{g_i} - f\left(\left(c_{t_i}(t), c_{p_i}(t)\right)\middle|\Theta\right)\right\|_2^2 \tag{2.3}$$

2.2.2.3 LSTM network architecture

The proposed network (Figure 2.1) consists of *m* LSTM layers with n sequentially connected cells in each layer. The network takes an input sequence $[C_t, C_p]^T$, where the AIF is incorporated as another input dimension. The first LSTM layer extracts lower-level temporal relationships. The output feature sequence is then passed through the remaining *m-1* LSTM layers each to extract higher-order level temporal relationships from the signal and AIF.

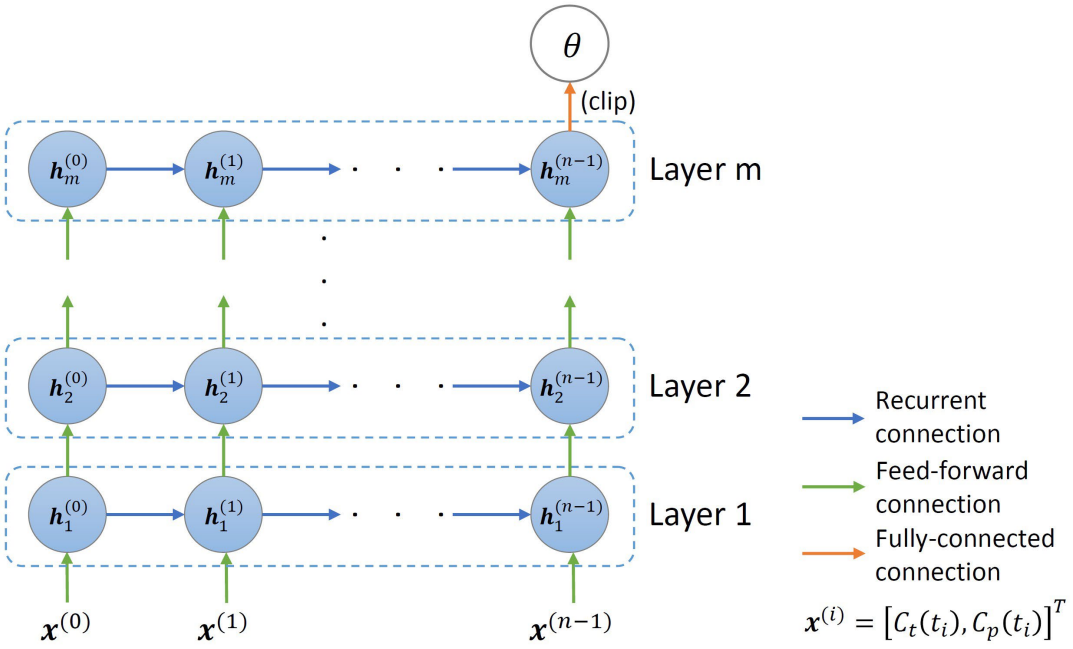

Figure 2.1: Illustration of the network architecture used for PK parameter estimation from an input of a CA concentration time-series and an AIF as two separate channels. Each layer has n sequentially connected cells.

Each LSTM layer captures changes in the input sequence by maintaining a hidden state $h^{(t)}$ and a memory cell $c^{(t)}$ by updating them using gating mechanisms when stepping through the input sequence. Specifically, the lth LSTM layer takes a sequence of hidden states $h_{l-1}^{(0)}$, $h_{l-1}^{(1)}$,…, $h_{l-1}^{(n-1)}$, $l \in \{1,2,\dots,m\}$, where the superscript and subscript stand for timestep and layer, respectively, and $h_0^{(0)}$, $h_0^{(1)}$,…, $h_0^{(n-1)}$ are defined as $x^{(0)}$, $x^{(1)}$,…, $x^{(n-1)}$. The new hidden states $h_l^{(0)}$, $h_l^{(1)}$,…, $h_l^{(n-1)}$ are then defined by the equations shown in (2.4) below

$$
\begin{aligned}
i_l^{(t)} &= \sigma(W_{in,x_l} h_{l-1}^{(t)} + W_{in,h_l} h_l^{(t-1)} + b_{i_l}) \\
f_l^{(t)} &= \sigma(W_{f,x_l} h_{l-1}^{(t)} + W_{f,h_l} h_l^{(t-1)} + b_{f_l}) \\
o_l^{(t)} &= \sigma(W_{o,x_l} h_{l-1}^{(t)} + W_{o,h_l} h_l^{(t-1)} + b_{o_l}) \\
\tilde{c}_l^{(t)} &= \tanh(W_{c,x_l} h_{l-1}^{(t)} + W_{c,h_l} h_l^{(t-1)} + b_{c_l}) \\
c_l^{(t)} &= \sigma(f_l^{(t)} \circ c_l^{(t-1)} + i_l^{(t)} \circ \tilde{c}_l^{(t)}) \\
h_l^{(t)} &= o_l^{(t)} \tanh(c_l^{(t)})
\end{aligned}
\tag{2.4}
$$

for $t \in \{0,1,\dots,n-1\}$ and $l \in \{1,2,\dots,m\}$, where $\sigma(\cdot)$ is the sigmoid function and "$\circ$" denotes the Hadamard product. $i_l^{(t)}$ and $f_l^{(t)}$ control which information to "input" to and "forget" from the memory cell, respectively. The new memory state $c_l^{(t)}$ is then obtained based on the candidate values $\tilde{c}_l^{(t)}$ and the gate values $i_l^{(t)}$ and $f_l^{(t)}$. Finally, the new hidden state $h_l^{(t)}$ is generated by the candidate values $o_l^{(t)}$ and memory state $c_l^{(t)}$.

Batch normalization is applied after each LSTM layer except for the last one. A fully connected layer with three features is then applied to find the best combination of features to generate an estimation of the PK parameters. The parameters are clipped to our targeted range, $K^{trans} \in [0,3](min^{-1})$, $v_e \in [0,0.4]$, and $v_p \in [0,0.55]$, thus confining the parameters to fall within a physiologically realistic range [47], to produce the final estimation of the PK parameters.

2.2.3 Data preparation

DCE MRI time series for training and testing were synthesized using AIFs from 103 patients with head and neck cancers. DCE MRI images were acquired using a dynamic scanning sequence (TWIST) with an injection of 0.149 cc/kg of gadobenate dimeglumine on a 3 Tesla MRI scanner (Skyra, Siemens Healthineers, Erlangen Germany). The protocol was approved by the Institutional Review Board of the University of Michigan. The scanning parameters were: flip angle = 10º, echo time (TE) = 0.97ms, repetition time (TR) = 2.73ms, 60 time frames, voxel

size = 1.5625×1.5625×2.5mm^3, matrix = 192×192. There were small variations in time step of temporal sampling of the dynamic series (median of 3.34 s) and in the numbers of the slices in z-direction between the patients (median of 72). For all cases, the subject-specific AIFs were extracted manually by averaging the signal intensity-time curves of 20 voxels from the carotid artery, which had maximum intensities at the time frame before the enhancement peak [59], and then subtracting and dividing by the average pre-contrast signal intensities of the voxels. The targeted parameter maps were estimated using DMF [47].

Of 103 patients, 78 cases were randomly selected for training, and 25 for testing. To overcome the limited size of the *in vivo* DCE MRI dataset, synthetic data were created for network training and testing. The synthesized data allow us to obtain a reliable and accurate assessment of the performance of the proposed methods by comparing the estimates to the ground truth (the parameters that created the synthetic data). Data augmentation was also applied during training data synthesis. Using equation 2.1 of the extended Tofts model, the training signal intensity time-curves were created from different combinations of the AIF, time step (Δt), bolus arrival time (τ_{BAT}), and the parameters ($\theta = (K^{trans}(\boldsymbol{r}), v_e(\boldsymbol{r}), v_p(\boldsymbol{r}))$). Particularly, AIFs and time steps of 78 training cases formed a set of 78 60-dimensional vectors and a set of 78 scalars respectively, defined as A_{train} and T_{train}. This yielded a set of 75,678,643 3-dimensional parameter vectors (K^{trans}, v_e, and v_p), which was denoted as P_{train}. For testing, we used the DCE time-series synthesized using the AIFs from the 25 cases in the testing data pool as well as acquired empirical data. The synthesized testing data were generated in the same manner as the training data. Testing with real data can assure that the model is 'realistic' enough.

2.2.4 Experimental setup

Performances of the LSTM networks on DCE MRI acquired with different total acquisition times and temporal sampling rates as well as different contrast-to-noise ratios (CNRs) were compared with the DMF method and a CNN model [51]. The LSTM networks were trained with input of fully temporal-sampled signal time-series as well as temporally subsampled time-series.

2.2.4.1 Training with synthetic data using acquired temporal sampling

The training data synthesis was executed on-the-fly during the network training and used the data from the training data generation pool consisting of the P_{train}, A_{train}, and T_{train}. Figure 2.2 shows the data generating process during network training. Specifically, at each iteration, 1000 combinations of $\{\theta_{g_i}\}_{i=1}^{1000}$, AIFs, and time steps $\{\Delta t\}_{i=1}^{1000}$, were selected randomly from the aforementioned training data generation pool. For each combination, the AIF was first randomly scaled between 70% and 130% (AIF augmentation) [60]. The concentration time-curve was then generated by randomly time-shifting the AIF (between 0 and 10 seconds to simulate the delay of CA arrival) using the extended Tofts model, and random Gaussian noise was added to the signal time-series to have contrast-to-noise ratios between 20 and 30. The resultant signal time-series and the corresponding scaled AIF (without time shifting) were concatenated as an input. As a result, this batch data consisted of a vector of dimension 1000×n×2 (where n was the number of time points in the series), and was passed to the network. Each epoch consisted of 1000 batches of training data and 200 batches of validation data. The network was trained with Adam optimizer [61] with an initial learning rate of 10^{-4}. The learning rate was reduced by a factor of 0.9 when the validation error was not improved in 30 consecutive epochs. The training was terminated when the validation error was not improved in 75 consecutive epochs.

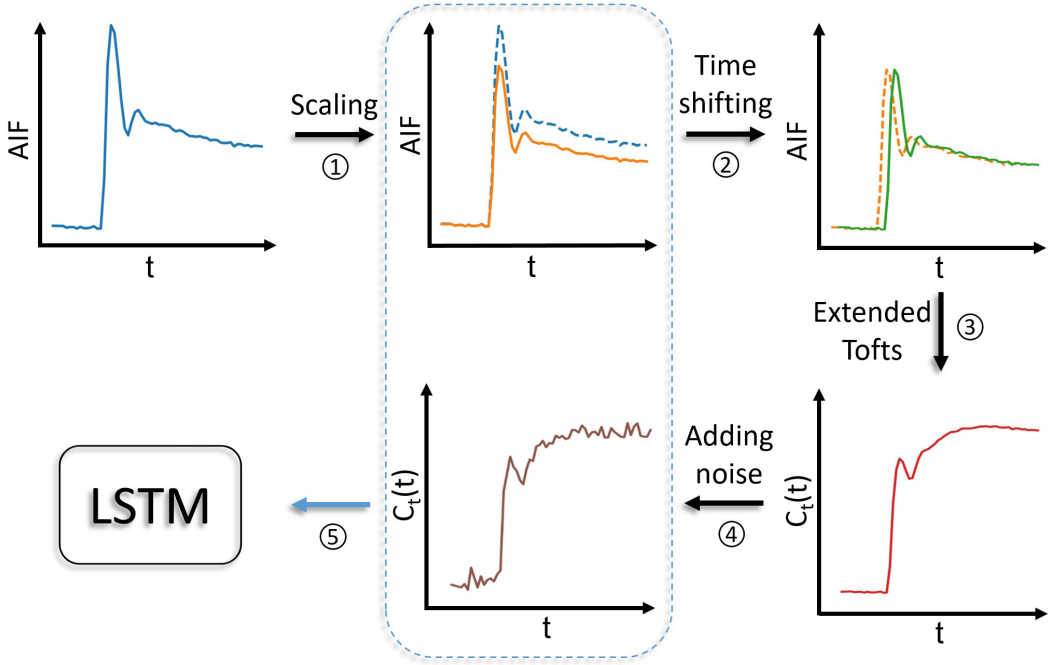

Figure 2.2: Training data generating scheme with BAT simulation and AIF augmentation by random scaling.

To evaluate sufficiency of the size of the training data generation pool on performance of the LSTM model, two other models with the same architecture as that for the initial model (LSTM3) were trained with the same scheme using 60% (LSTM1) and 80% (LSTM2) of the training data generation pool.

2.2.4.2 Training with temporally subsampled synthetic data

The proposed LSTM network was also trained and tested on temporally subsampled synthetic signal time-curves, where signal time-curves were generated with sampling time steps of Δt: 3, 4, 5 and 6 seconds. Lengths of the time-series were truncated at an integer number of time steps that was close to 168 seconds. The AIF and the PK parameters were drawn from the same data generating pool. The LSTM input size was modified accordingly to match the number of time points of the subsampled signal time-series, while the other processes and parameters were kept the same, including the targeted PK parameters.

2.2.4.3 Training with truncated data and data with different CNR levels

To examine robustness of the proposed network to signal noise and total acquisition time, the proposed LSTM network was further trained and tested on: (1) truncated synthetic data with the total acquisition times 168s, 141s, 114s, and 87s when keeping the same sampling time step 3 s, and (2) fully temporally sampled synthetic data but with CNR variations at 20-30, 10 and 5.

2.2.4.4 Performance evaluation and comparison

First, our proposed LSTM network was optimized for the numbers of LSTM layers and hidden state features using fully temporal sampled synthetic data. Particularly, 2, 4, 6 and 8 layers, and 16, 32 and 64 features were trained and tested. Then, the optimal numbers of layers and features were used to train and test the LSTM model, whose performance was compared to both the conventional DMF approach and a CNN model. In testing, both synthetic and actual patient data were used. The synthetic testing data allows us to quantitatively assess the performance of PK parameters estimation by calculating the structural similarity (SSIM) index and normalized root mean squared error (NRMSE) of the estimated parameter maps with respect to the ground truth parameter maps. The NRMSE is defined by $NRMSE_i = \sqrt{\frac{\sum_{j=1}^{N}(\tilde{\theta}_{ij}-\theta_{ij})^2}{N-1}} / (\max_j\{\theta_{ij}\} - \min_j\{\theta_{ij}\}) \times 100(\%)$ for a slice, where $NRMSE_i$, $\tilde{\theta}_{ij}$, θ_{ij}, and N are the NRMSE of the ith parameter, the estimated and ground truth ith parameter for jth voxel, and the number of voxels in the slice. To synthesize the testing signal time-curves, the AIF and the PK parameter maps were from the same patients, for which no cross-combination of the AIF and the PK parameters nor AIF augmentation were used for the training data synthesis. The SSIM and NRMSE were calculated from each of the 2D PK parameter slice maps first, and then averaged

over multiple slices and across 25 cases. The proposed LSTM model was further evaluated using the empirical data from the same 25 cases.

The CNN model proposed by Ulas and colleagues [51] was implemented for comparison with the LSTM model. The input of the CNN is 2D+time DCE image-series concatenated along time dimension across different channels. The CNN was trained on 1500 2D+time image-series with ℓ_2 loss and validated on 300 image-series generated using the same data generation pool as for LSTM training. The model loss term in the original model was dropped since BAT was not considered in the original paper, and enforcing model consistency without BAT correction produced worse results in our experiment. The size effect of the training data generation pool on the performance of the CNN-based method was also investigated by training with 60% (CNN1), 80% (CNN2), and 100% (CNN3) of the training data generation pool.

All codes were implemented using Keras library with Tensorflow backend, and the experiments were performed on an NVIDIA Tesla K40c GPU with 12 GB RAM.

2.3 Results

2.3.1 Optimization of the LSTM networks

We investigated the distributions of the PK parameters $\theta = (K^{trans}(\boldsymbol{r}), v_e(\boldsymbol{r}), v_p(\boldsymbol{r}))$ of the 78 patients, which were used as the training data generation pool. The parameter values in the pool covered the whole desired ranges of the parameters (Figure 2.3).

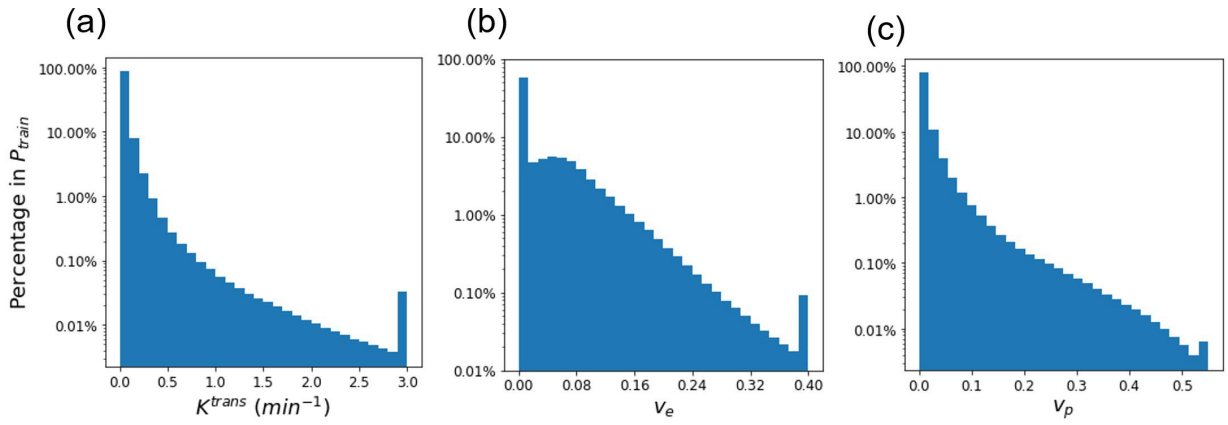

Figure 2.3: The distribution of (a) K^{trans}, (b) v_e, and (c) v_p in the training data generation pool.

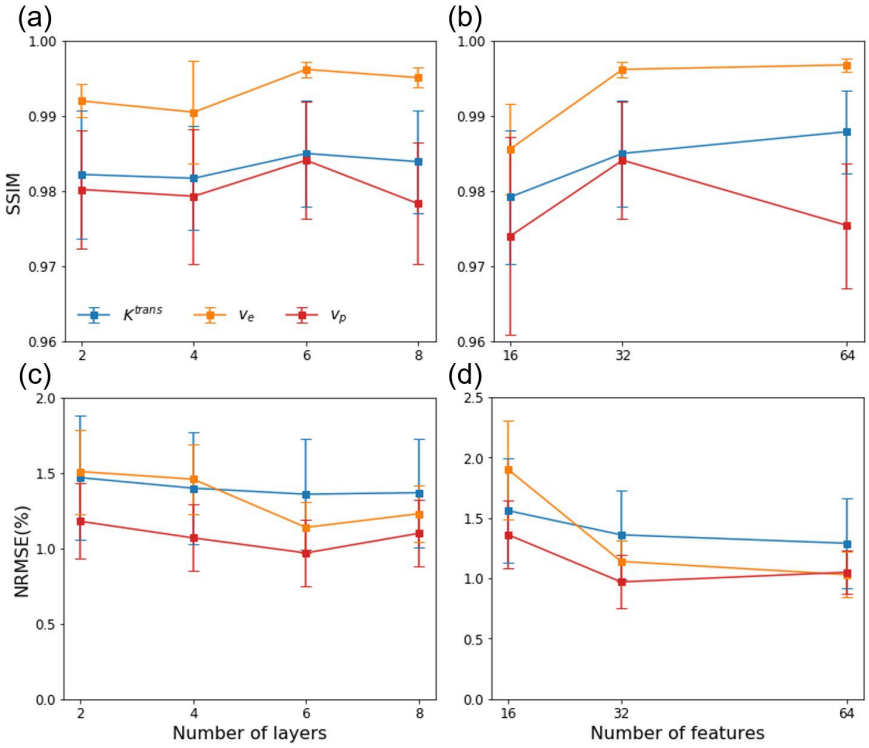

Figure 2.4: SSIM and NRMSE of the parameter map estimation by LSTM networks with (a, c) 2, 4, 6, and 8 LSTM layers with 32 features, and (b, d) 16, 32, and 64 features with 6 layers. Error bar: standard deviation.

We trained the LSTM networks with 2, 4, 6 and 8 layers and 32 features, and with 16, 32 and 64 features and 6 layers. Figure 2.4 shows the impact of varying the layers and features on NRMSE% and SSIM. The LSTM network with 6 layers and 32 features had average maximal

SSIM, presenting a balance of overfitting and underfitting of the training data and thus this network configuration was used for training under various test conditions.

2.3.2 Performance of the LSTM networks trained with original temporal-sampling data

The performance of the LSTM networks trained with synthetic data with original temporal-sampling, as well as results using the CNN and DMF models are shown in Table 2.1. The proposed LSTM network achieved high SSIM and low NRMSE% in the whole field of view as well as in the gross tumor volume (Table 2.1), comparable to the DMF approach. LSTM3 had < 1% lower SSIMs for K^{trans} and v_p, higher SSIM for v_e, and 13.4%, 19.1%, and 25.4% better NRMSEs for K^{trans}, v_e and v_p, respectively, than the DMF approach. LSTM3 outperformed the CNN-based approach by reducing the NRMSE up to 55.2%. When evaluating the size effect of the data generation pool, the performance of LSTM1, LSTM2 and LSTM3 were similar, and the size effect was insignificant, indicating the training data augmentation is effective. The CNN-based method shows inferior performance compared to the LSTM3 (Table 2.1). A visual illustration of estimated PK parameters and residuals of the testing data by different methods is given in Figure 2.5.

Table 2.1: Quantitative performance of different methods on test DCE-MRI volumes. The LSTM models were trained and tested using synthetic data with original temporal-sampling. The SSIM and NRMSE% (mean + std) with respect to the ground truth parameter maps were obtained in the whole field of view. The bold numbers indicate significant differences ($p<0.05$) between LSTM and DMF. GTV: gross tumor volume.

	Method	SSIM			NRMSE (%)		
		K^{trans}	v_e	v_p	K^{trans}	v_e	v_p
Whole Range	DMF	0.9875 ± 0.0052	0.9960 ± 0.0010	0.9880 ± 0.0046	1.57 ± 0.43	1.41 ± 0.27	1.30 ± 0.38
	LSTM1	0.9853 ± 0.0068	0.9922 ± 0.0027	0.9806 ± 0.0159	1.47 ± 0.36	1.44 ± 0.25	1.15 ± 0.24
	LSTM2	0.9840 ± 0.0078	0.9931 ± 0.0025	0.9850 ± 0.0060	1.39 ± 0.41	1.19 ± 0.14	0.97 ± 0.18
	LSTM3	0.9850 ± 0.0070	0.9962 ± 0.0010	0.9841 ± 0.0078	**1.36 ± 0.37**	**1.14 ± 0.17**	**0.97 ± 0.22**
	CNN1	0.9547 ± 0.0202	0.9534 ± 0.0180	0.8946 ± 0.0267	3.07 ± 1.32	4.04 ± 1.57	3.54 ± 1.37
	CNN2	0.9629 ± 0.0159	0.9594 ± 0.0146	0.9323 ± 0.0237	2.78 ± 0.99	4.17 ± 1.32	3.16 ± 1.44
	CNN3	0.9627 ± 0.0189	0.9556 ± 0.0192	0.9302 ± 0.0352	2.76 ± 1.10	3.95 ± 1.30	2.85 ± 0.94
p	LSTM3 vs DMF	<0.05	0.54	<0.005	<0.05	<0.005	<0.005
p	LSTM3 vs CNN3	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
GTV	DMF	0.9994 ± 0.0006	0.9997 ± 0.0004	0.9994 ± 0.0008	3.85 ± 2.64	0.85 ± 0.35	0.74 ± 0.40
	LSTM3	0.9993 ± 0.0006	0.9998 ± 0.0003	0.9994 ± 0.0007	4.45 ± 2.64	0.85 ± 0.21	0.71 ± 0.24
p		0.38	<0.05	0.63	<0.05	1.00	0.45

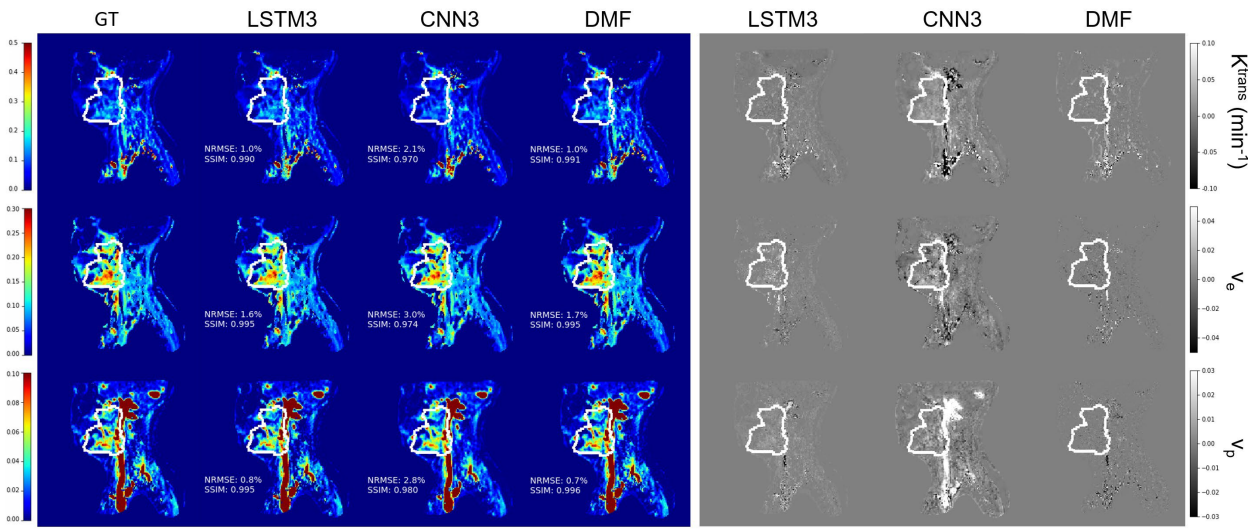

Figure 2.5: An exemplary slice of the ground truth parameter maps (column 1), the estimated maps (column 2-4), and the residual maps (last 3 columns) of estimated K^{trans} (top row) v_e (middle row) and v_p (bottom row) by the LSTM3, CNN, and DMF models from a testing case. The white contour depicts the gross tumor volume. GT: ground truth.

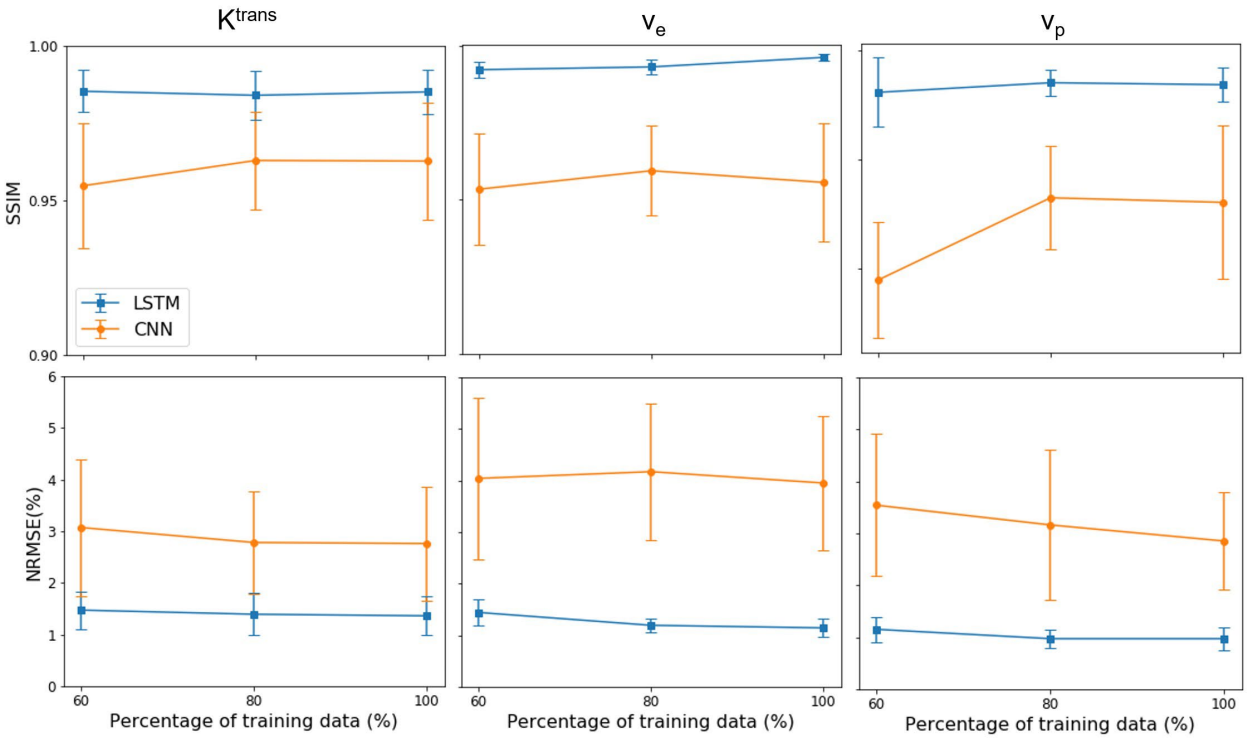

Figure 2.6: The performance of the LSTM and CNN-based methods under different amounts of training data (60%, 80%, and 100% of the data generation pool).

Figure 2.6 shows the performance of the LSTM and CNN-based methods with different amounts of training data. The performance of CNN-based method changed little with an increase in the size of the training data generation pool, indicating the sufficiency of training data. Figure 2.7 shows examples of the parameter maps generated by the DMF and LSTM methods on the acquired DCE MRI data from the testing datasets, where the estimations of DMF and LSTM approaches were highly consistent. In most cases, DMF and LSTM approaches fit the signal intensity-time curves similarly well. Examples of fitted signal intensity-time curves are shown in Figure 2.8. To quantitatively compare the fitting results by LSTM and DMF, the voxel-wise MSE was calculated for both methods. The LSTM approach yielded a lower voxel-wise MSE (0.743 ± 0.130) than the DMF (0.808 ± 0.127). The parameter inference time of LSTM for fully temporal sampled DCE MRI volumes was ~250s on CPU (~40s on GPU), while DMF approach required ~3600s to generate the PK parameter maps for the same data on the same CPU, representing approximately 14.4 times improvement in computation speed.

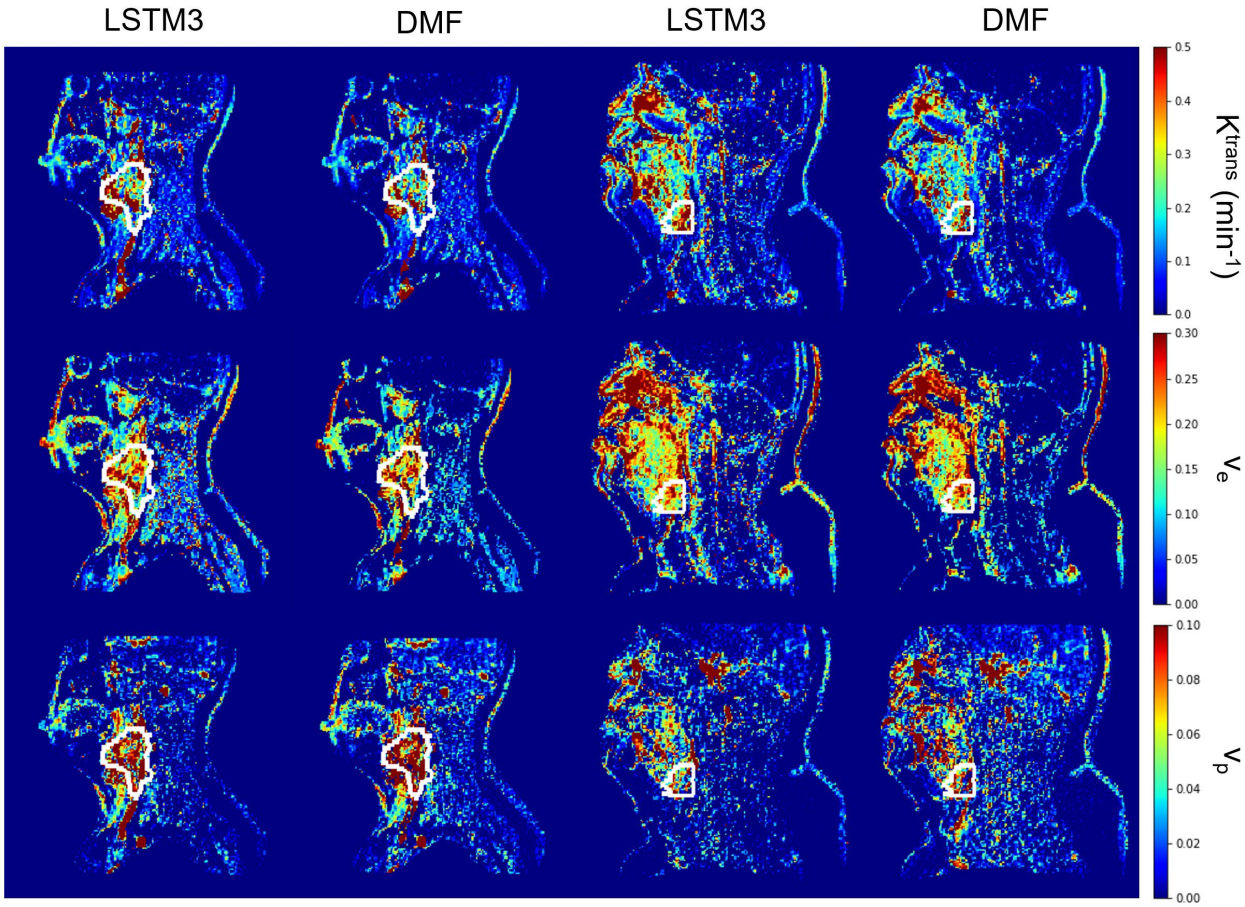

Figure 2.7: Two exemplary slices (left 2 columns for first slice, right 2 columns for second slice) of PK parameter estimation by DMF and LSTM3 on an in vivo test dataset. The results obtained from DMF and LSTM3 show high similarity in both the tumor volume (depicted by the white contour) and the full FOV.

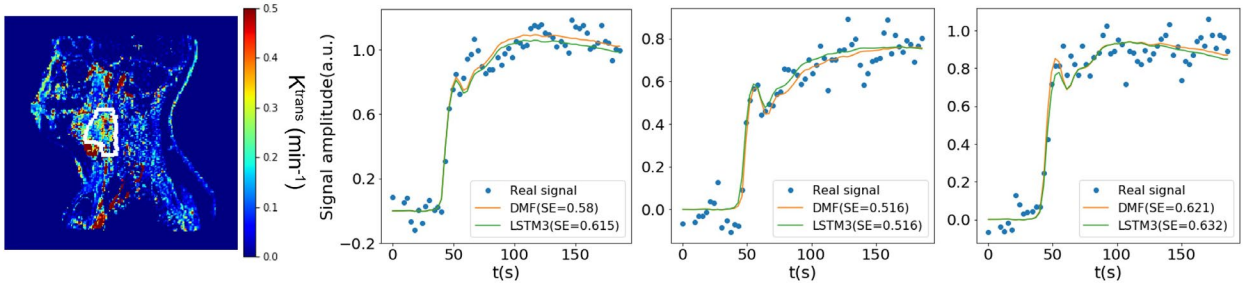

Figure 2.8: Three exemplary in vivo CA concentration-time curve fitting results by DMF and LSTM3 in the tumor region indicated by the white contour shown in the left K^{trans} parameter map estimation by DMF. Results from both methods are in reasonable alignment with observed data. (SE = squared error)

2.3.3 Performance of the LSTM networks with temporally subsampled data

Figure 2.9 and Table 2.2 show that the LSTM had significantly better performances for estimating parameters than the DMF method ($p<3.2\times10^{-6}$) when increasing temporal sampling intervals from 3s to 4s, 5s and 6s (Figure 2.10).

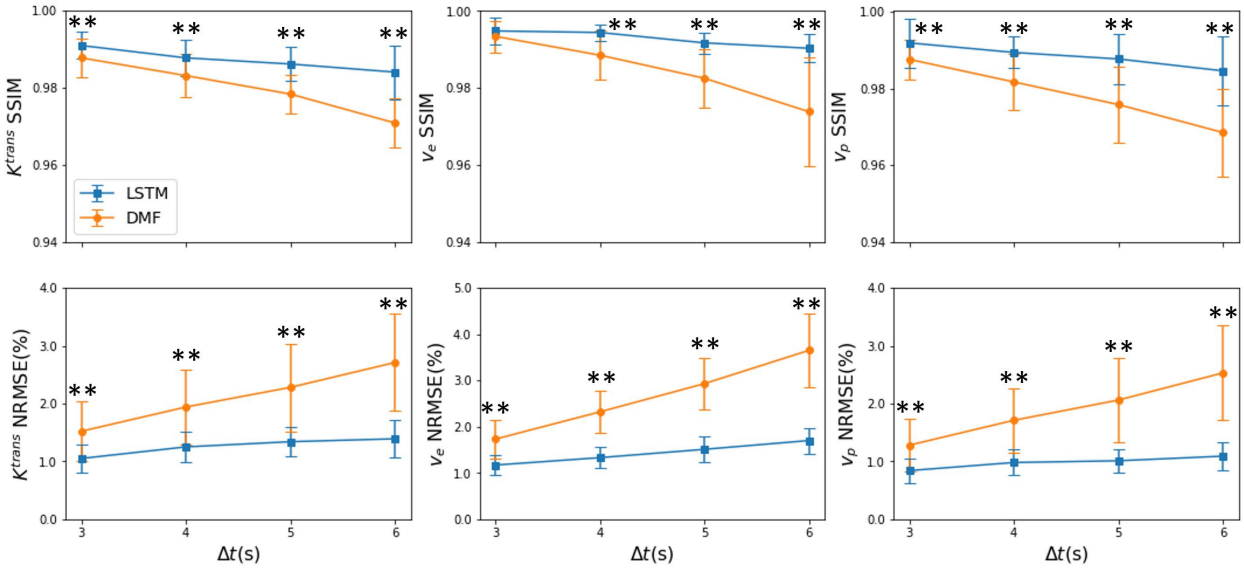

Figure 2.9: Quantitative results of the estimated parameters from the 25 synthesized testing datasets with different temporal sampling time intervals (3, 4, 5, and 6s) by the LSTM and DMF approaches. The proposed LSTM shows a more stable performance than the DMF when increasing the sampling interval. Error bar: standard deviation. *: p<0.05; **p<0.005.

Table 2.2: Quantitative performance of different methods on temporally subsampled DCE data. The SSIM and NRMSE% (mean + std) with respect to the ground truth parameter maps were obtained in the whole field of view. The bold numbers indicate significant differences (p<0.005) between the two methods.

	Method	SSIM			NRMSE (%)		
		K^{trans}	v_e	v_p	K^{trans}	v_e	v_p
Δt=3	DMF	0.9877 ± 0.0050	0.9934 ± 0.0041	0.9876 ± 0.0052	1.52 ± 0.52	1.73 ± 0.41	1.28 ± 0.46
	LSTM3	**0.9909 ± 0.0035**	0.9948 ± 0.0034	**0.9919 ± 0.0064**	**1.05 ± 0.24**	**1.17 ± 0.21**	**0.84 ± 0.21**
p		<0.005	0.12	<0.005	<0.005	<0.005	<0.005
Δt=4	DMF	0.9831 ± 0.0055	0.9885 ± 0.0064	0.9817 ± 0.0073	1.94 ± 0.64	2.32 ± 0.46	1.71 ± 0.56
	LSTM3	**0.9877 ± 0.0048**	**0.9944 ± 0.0022**	**0.9894 ± 0.0041**	**1.25 ± 0.27**	**1.33 ± 0.23**	**0.98 ± 0.22**
p		<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Δt=5	DMF	0.9783 ± 0.0050	0.9825 ± 0.0075	0.9758 ± 0.0099	2.28 ± 0.76	2.93 ± 0.56	2.06 ± 0.73
	LSTM3	**0.9861 ± 0.0044**	**0.9917 ± 0.0027**	**0.9877 ± 0.0066**	**1.34 ± 0.26**	**1.51 ± 0.27**	**1.01 ± 0.21**
p		<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Δt=6	DMF	0.9708 ± 0.0063	0.9738 ± 0.0142	0.9685 ± 0.0115	2.71 ± 0.84	3.65 ± 0.80	2.53 ± 0.82
	LSTM3	**0.9840 ± 0.0070**	**0.9903 ± 0.0037**	**0.9846 ± 0.0089**	**1.39 ± 0.33**	**1.70 ± 0.28**	**1.09 ± 0.25**
p		<0.005	<0.005	<0.005	<0.005	<0.005	<0.005

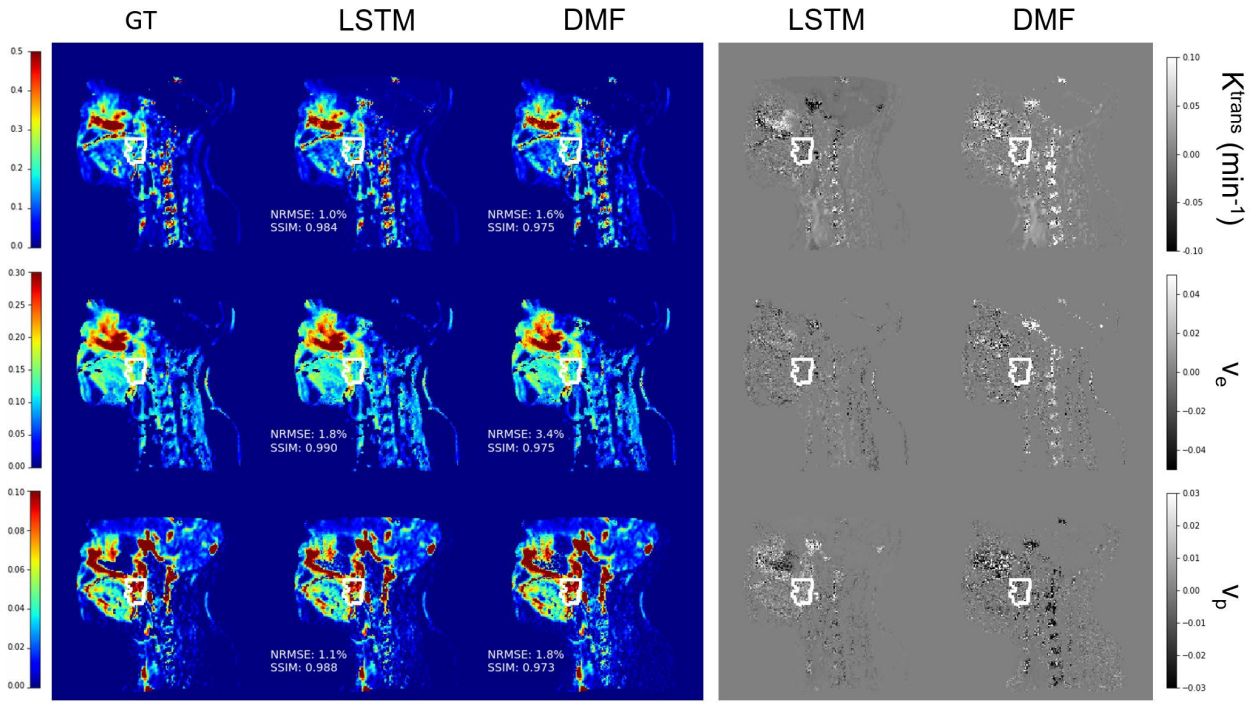

Figure 2.10: An exemplary slice of ground truth of parameter maps (column 1), estimated maps (column 2 and 3), and the residual maps (last 2 columns) by LSTM and DMF using temporal sampling interval $\Delta t = 6s$. The tumor volume is depicted by a white contour. GT: ground truth.

2.3.4 Performance of the LSTM networks with low CNR and reduced total acquisition length

The performance of LSTM on DCE MRI signals with lower CNRs and reduced total acquisition times is presented in Figure 2.11 and Table 2.3. As can be seen, LSTM consistently improved the accuracy for lower CNR levels and reduced total acquisition times from 168s to 87s. v_e estimation had the largest improvement when total acquisition time was reduced.

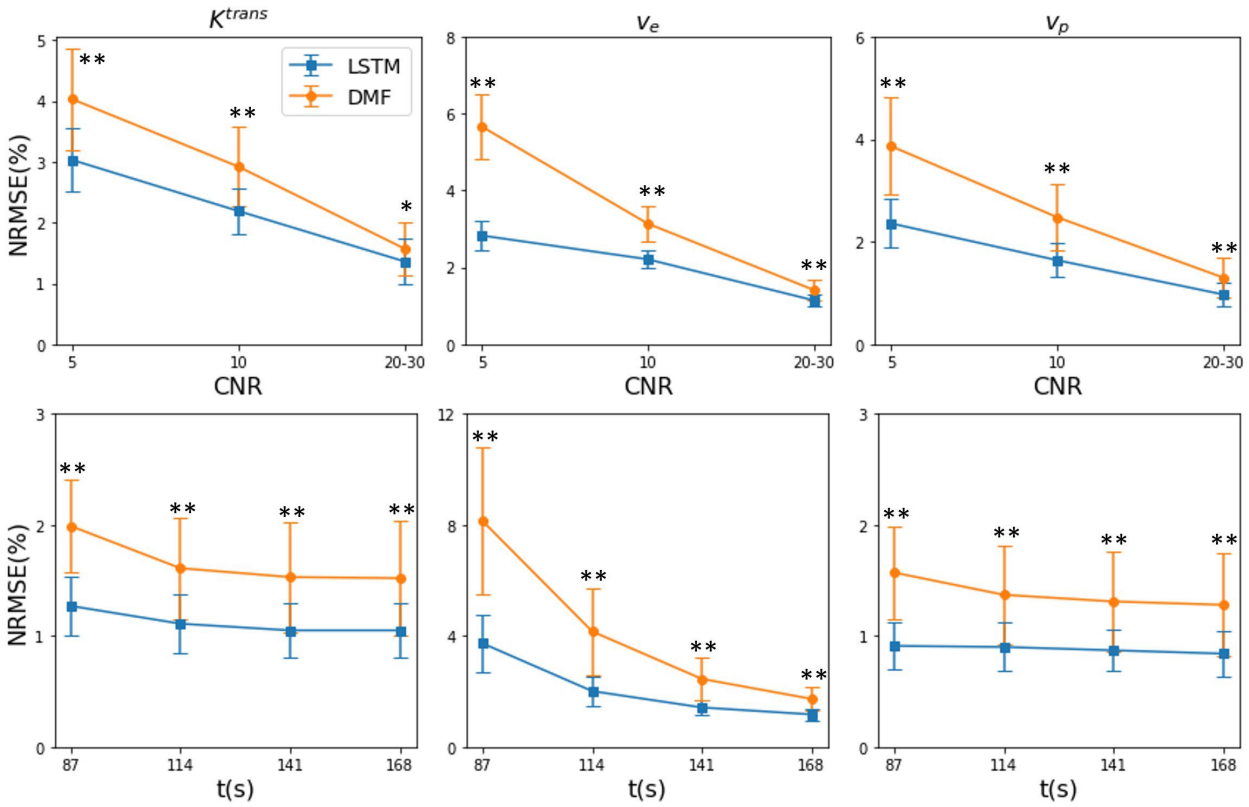

Figure 2.11: The performance (NRMSE) of the LSTM and DMF estimations under lower CNRs (first row) with full temporal sampling and reduced total acquisition times t (second row) with Δt=3s and CNR=20-30. Error bar: standard deviation. *: p<0.05; **p<0.005.

Table 2.3: NRMSE (%) of testing signals with lower CNR levels and reduced total acquisition time (t). The bold numbers are the better results of the proposed method than those of DMF.

	NRMSE (%)					
	DMF			LSTM		
	K^{trans}	v_e	v_p	K^{trans}	v_e	v_p
CNR						
5	4.03 ± 0.83	5.67 ± 0.84	3.87 ± 0.95	**3.03 ± 0.52**	**2.83 ± 0.37**	**2.36 ± 0.47**
10	2.92 ± 0.65	3.14 ± 0.46	2.48 ± 0.66	**2.19 ± 0.38**	**2.21 ± 0.24**	**1.64 ± 0.33**
t (s)						
87	1.99 ± 0.42	8.14 ± 2.66	1.57 ± 0.42	**1.27 ± 0.26**	**3.73 ± 1.02**	**0.91 ± 0.21**
114	1.61 ± 0.46	4.16 ± 1.58	1.37 ± 0.44	**1.11 ± 0.26**	**2.01 ± 0.53**	**0.90 ± 0.22**
141	1.53 ± 0.50	2.45 ± 0.78	1.31 ± 0.45	**1.05 ± 0.24**	**1.42 ± 0.29**	**0.87 ± 0.19**

2.4 Discussion

We investigated a novel and potentially powerful LSTM-based network for learning a mapping from a CA concentration-time curve including the corresponding AIF to the underlying

PK parameters. The LSTM network is capable of learning long- and short-term dependency of sequence data such as DCE MRI. We found that the performance of the LSTM on mapping DCE MRI time-curves to their corresponding PK parameters was superior to a state-of-the-art CNN-based approach [51], and better than the direct model fitting method in terms of NRMSE. The LSTM was much more robust to temporally subsampled DCE data than the direct PK model fitting, which can be utilized to increase spatial resolution of DCE images. Higher robustness of LSTM to noise and reduced acquisition time was also demonstrated compared with DMF. Our data augmentation strategies, including AIF augmentation and creation of synthetic signal time-curves from the data generation pool, overcame the limited size of the *in vivo* DCE MRI training data pool. The LSTM network trained by the synthesized data was also able to perform well on empirical DCE data. This indicates that the synthetic data simulates real signal intensity-time series well and the LSTM network is effectively trained. In addition, our proposed network enables approximately 90 times of computation time acceleration compared with the direct PK model fitting approach. The LSTM network has the potential to accelerate DCE MRI acquisition and parameter estimation.

We attribute the superior performance of the LSTM to its capability to learn long- and short-term dependency in sequence data, and to extract dynamic features and temporal correlation in the signal intensity-time series of DCE MRI. In contrast, the CNN extracts "spatial" features from the DCE MRI volumes but has a limited capability to exploit the temporal relationships in the DCE data. For example, a 2D CNN model treats the x-y-t data as a 3D volume, in which time-dependent features in the dynamic data could not be effectively extracted [62], [63]. A 3D CNN model [50] attempts to address some of the issues of the 2D CNN by leveraging more temporal correlation. Without incorporating an AIF as input in the

CNN, the PK parameters could have degeneracy and correspond to multiple signal time-curves, which can lead to a mis-mapping between the PK parameters and DCE curves. The small training and testing patch sizes limited by the high GPU memory demand of the 3D CNN training may further degrade its performance. Our proposed LSTM network structure is straightforward, and there is no PK model information required for LSTM training. Also, the AIF is incorporated into input as a second channel, which allows use of a patient-specific AIF when processing empirical patient DCE data. The proposed network can be easily trained and extended to other PK models, or even other sequence-related medical imaging data, e.g., a high-order diffusion model, with a minimum modification of the network architecture.

We used several strategies to overcome the limited size of the *in vivo* DCE MRI data. First, we did not use a fixed triad of the AIF, the PK parameters, and the signal time-curves from the *in vivo* DCE MRI dataset, which limits us to a total of 103 patient datasets. We used synthesized signal time-curves that were created by randomly selected and combined PK parameters, AIF (and augmented AIF), and time step from a data pool. In addition, we added random variations of the delay of CA bolus arrival into the signal time-curves. Our evaluation shows that the PK parameters from 78 patients sufficiently cover the parameter ranges of interest. The performance of LSTM2, and LSTM3 indicates that the PK parameters from 60% of the training data could sufficiently cover the ranges of parameter of interest. Our overall strategy of the training data synthesis seems to yield effective training of the LSTM networks, and overcomes the limited size of *in vivo* datasets. These strategies seem to reasonably mitigate small amount of variation in sampling interval in the *in vivo* data.

For extraction of PK parameters from signal time curves with different time steps (3-6s), the LSTM models show robust performance compared to the direct PK model fitting, with the

latter showing performance degradation with an increase in the time step size. This advantage of the LSTM approach can be utilized to improve the spatial resolution of DCE MRI when decreasing the temporal resolution of DCE MRI. For example, a modest increase in the spatial resolution from the currently used 1.6x1.6x2.5 mm^3 to 1.4x1.4x1.4 mm^3 would prolong the acquisition time of an image volume by a factor of 2. This increase in the temporal resolution from 3s to 6s would result in an increase in NRMSE% of the estimates by 1.7-2.1 times by the DMF but a very small increase by the LSTM model (Figure 2.9).

This work has several limitations. As we can see from Figure 2.6 and Table 2.1, the LSTM performs better than DMF in terms of NRMSE but not as good in terms of SSIM across the whole parameter range. Further analysis of the error distribution reveals that this is mainly due to the minor estimation errors in parameter combinations of zero K^{trans} and v_e but non-zero v_p (mainly in the brain region), which caused by the small portion of these parameter value combinations in the training datasets. Another concern is the accuracy of parameter estimation in the gross tumor volume, which has a different range of the parameters from normal tissue and has a small amount of the data weighting in the training dataset. We note similar SSIM values of DMF and LSTM in the gross tumor volume (Table 2.1). Further manipulation of the training dataset distribution and/or a modification of the loss function or a weighting training data in different parameter ranges could improve performance of the algorithm. Another path to improvement generalizability is to incorporate temporal sampling intervals into the network input as another channel [64]. The performance of the current vanilla LSTM architecture could be further improved by using a bidirectional LSTM with attention at the expense of longer training and inference time. A further improved model could use more realistic synthetic data

that takes motion artifacts and other factors into account to improve the robustness of performance of the LSTM on *in vivo* DCE MRI datasets.

2.5 Conclusion

In conclusion, our proposed LSTM is a promising approach to estimate PK parameters from DCE-MRI time-series. We demonstrate that the proposed approach provides more accurate PK parameter maps compared to the CNN-based approach, and is comparable to the DMF method with approximately 90 times of computation time reduction. The LSTM networks are more robust to low temporal resolution, lower CNR levels, and reduced total acquisition time than direct PK model fitting. In the future, a LSTM network with convolutional layers can be applied to image-series to leverage both temporal and spatial sparsity, and thus further improve accuracy and precision of DCE MRI.

Chapter 3 Joint Optimization of k-t Sampling Pattern and Reconstruction of DCE MRI for Pharmacokinetic Parameter Estimation

3.1 Introduction

The purpose of this chapter is to improve accuracy and precision of estimated pharmacokinetic parameters while keeping similar spatial resolution and the same scan time by jointly optimizing k-t sampling pattern and image reconstruction.

To yield accurate PK parameter estimation from DCE MRI, different k-t space sampling trajectory patterns [65]–[69] have been empirically designed to achieve relatively high temporal and spatial resolution and signal-to-noise ratio. However, the k-t sampling pattern (SP) can be systematically optimized for DCE MRI by taking advantage of the spatiotemporal sparsity of dynamic signals as well as characteristics of anatomy being imaged. A few attempts have been made to optimize the k-space SP to leverage similarity of the same body site assisted by prior knowledge for image reconstruction algorithms [19], [70]–[73]. However, these works mostly rely on pattern search algorithms that lack computational efficiency. Most importantly, currently no strategy considers the influence of underlying quantitative PK parameter estimation on the sampling optimization, even though the k-t SP influence on the variance of estimated PK parameters has been demonstrated [74]. We hypothesize that joint optimization of DCE MRI k-t SPs and PK parameter estimation has the potential to improve accuracy and reproducibility of the estimated parameters and efficiency of the raw data acquisition.[1]

[1]This chapter is based on our paper published in *IEEE Transactions on Medical Imaging* [45].

Nonlinear algorithms are needed to reconstruct high quality MR images from highly undersampled k-t space DCE data. Compressed sensing-based methods have been applied with hand-crafted constraints exploiting the spatial and temporal sparsity in DCE MRI [75]–[78]. Recently, deep learning has shown promising in MR image reconstruction by exploring data-driven constraints [12]–[16]. In deep learning-based frameworks, RNNs that processes temporal information in the dynamic data have shown superior performance in DCE MRI reconstruction [79], [80]. More recently, deep learning-based joint optimization of k-space SPs and reconstruction networks has been proposed for non-dynamic images [22], [81], [82], and shown improved quality of reconstructed MR images from optimized undersampled k-space data compared to compressed sensing-based methods with non-optimized sampling. However, these investigations of joint optimization have primarily focused on the static MR image reconstruction.

In this chapter, we extend the use of RNNs to jointly optimize k-t space SPs of DCE MRI acquisition and image reconstruction with an objective that combines image quality and parameter estimation accuracy. The dynamic MRI reconstruction network exploits the spatiotemporal sparsity of DCE MRI to optimize where in k-space to acquire the MR data across the dynamic time course. The PK parameter estimation l_2 loss was integrated into the objective of the image reconstruction network to account for PK modeling information during network training. Realistic multi-coil digital reference objects (mcDROs) were created from PK parameter maps estimated from patient scans with head and neck cancers and used for network training and testing, which provided ground truth for quantitative evaluation. The proposed method was tested in a wide range of temporal resolutions of mcDROs and showed reductions in PK parameter bias and uncertainty compared to two previous published works [78], [83].

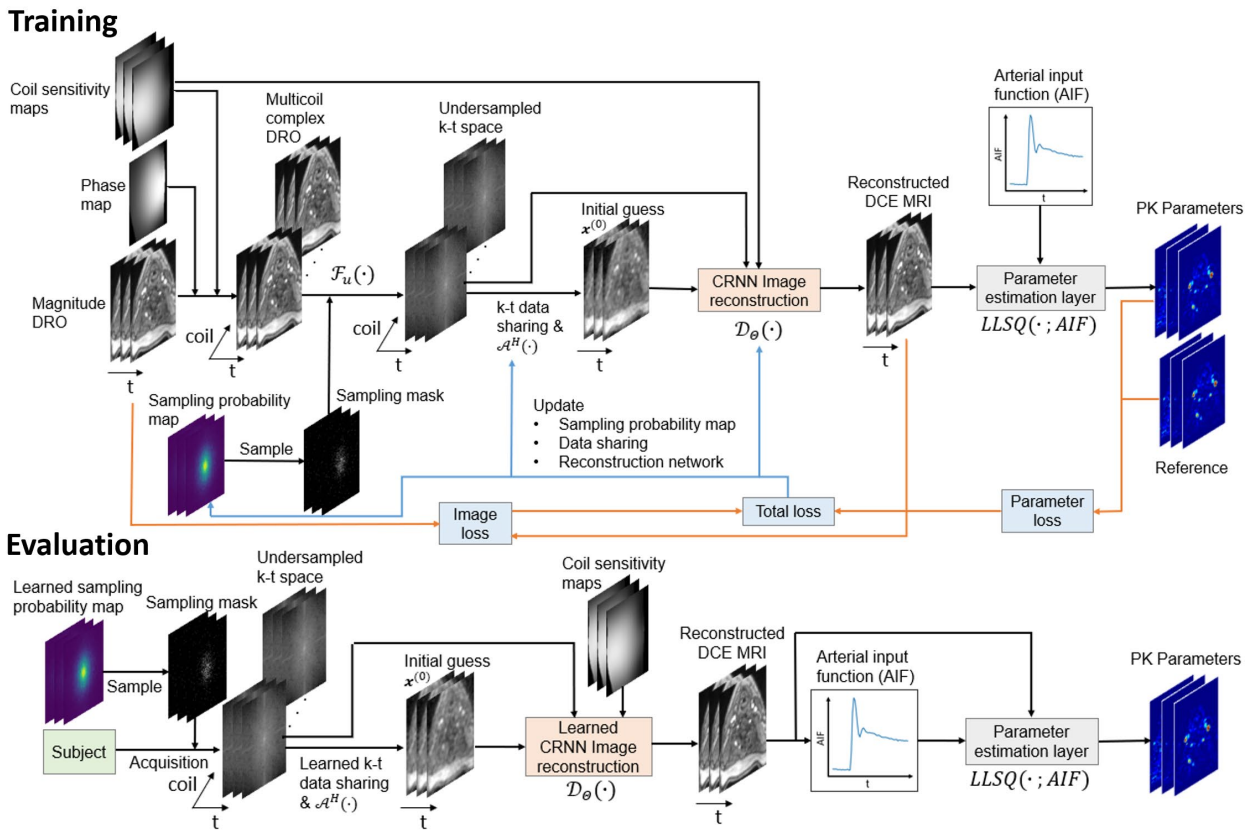

Figure 3.1: The training and evaluation workflows of the proposed method. The k-t sampling probability map, k-t data sharing, and image reconstruction network are jointly optimized with respect to both image and parameter reconstruction quality using fully sampled DCE MRI data in the training phase. In the evaluation phase, the learned sampling mask and reconstruction

3.2 Materials and Methods

Our proposed method consists of four major components: 1) k-t SP optimization, 2) k-t space data sharing optimization, 3) a dynamic MR image reconstruction network, and 4) a PK parameter estimation layer. The first three have learnable parameters and are jointly optimized during training. The last component, the PK parameter estimation layer, has no trainable parameters, but its gradient is passed on to the other components to provide feedback for network learning. The overall workflow of the proposed approach is shown in Figure 3.1. The details of each component and how they are combined to allow an end-to-end training are described in the following sub-sections.

3.2.1 k-t sampling pattern

The k-t space SP that represents 2D phase encoding locations along the time course of a DCE acquisition using a 3D T1 weighted spoiled gradient echo sequence is learned jointly with the reconstruction network. Frequency encodings are fully sampled due to their rapid sampling speed. We extended LOUPE [84] to the time domain by adding a temporal degree of freedom (TDoF) to learn dynamic SPs in the k-t space. Let $\boldsymbol{x} \in \mathbb{C}^{n_x n_y n_t}$ denote a sequence of fully sampled images of a slice with width n_x, height n_y and n_t time frames. The 2D Cartesian k-space SP in a k_x-k_y plane at time n_t is denoted as $\boldsymbol{m}^u \in \{0,1\}^{n_x n_y n_t}$ where superscript u represents "undersampling", which is formulated as a realization of random vector $\boldsymbol{M}^u$ containing independent Bernoulli random variables, i.e. $\boldsymbol{M}^u \sim \prod_{i=1}^{n_x n_y n_t} \mathcal{B}(\boldsymbol{p}_i^u)$, where $\boldsymbol{p}^u \in [0,1]^{n_x n_y n_t}$ denotes the subsampling probability to be learned. Therefore, the undersampled k-space data of the *cth* coil can be written as $\boldsymbol{F}_u \boldsymbol{C}_c \boldsymbol{x} = diag(\boldsymbol{m}^u)\boldsymbol{F}\boldsymbol{C}_c\boldsymbol{x}$, where $\boldsymbol{C}_c$ is a diagonal matrix with the coil sensitivity values of the *cth* coil as the diagonal elements, $\boldsymbol{F}$ is the forward discrete Fourier transform matrix and $\boldsymbol{F}_u = diag(\boldsymbol{m}^u)\boldsymbol{F}$ is the undersampling matrix. The first timeframe of the image slice is fully sampled to provide the anatomic baseline prior to the arrival of the injected contrast.

To implement LOUPE with a TDoF, a set of neural network weight parameters $\boldsymbol{w}^u \in \mathbb{R}^{n_x n_y n_t}$ to be learned is regularized by an element-wise sigmoid function to produce a set of sampling probability maps $\boldsymbol{p}^u \in [0,1]^{n_x n_y n_t}$. To create SPs from the sampling probability maps $\boldsymbol{p}^u$, a sigmoid function is used to approximate the operation so that $\boldsymbol{m}^u = \sigma_v(\boldsymbol{u} - \boldsymbol{p}^u)$, where $\boldsymbol{u} \in [0,1]^{n_x n_y n_t}$ is a realization of $\prod_{i=1}^{n_x n_y n_t} \mathcal{U}(0,1)$ and $\sigma_v(x) = \frac{1}{1+e^{-vx}}$ is an element-wise sigmoid function with slope *v*. The sigmoid function ensures nonzero gradient when backpropagating through the sampling of Bernoulli random variables.

3.2.2 k-t space data sharing

In DCE MRI, adjacent temporal frames share similar contrast enhancements among the same anatomy, and especially, dynamic changes are limited before starting contrast uptake and after reaching the uptake plateau. Sharing the k-space data among temporal frames within the same anatomy would reduce undersampling-caused aliasing in an initial image for reconstruction network training. Inspired by previous data sharing approaches [15], we propose a novel machine learning-based k-space data sharing scheme. A subset of frames is determined during training to share their k-space data with frame *j*, and represented by a data sharing mask $\boldsymbol{m}_j^s \in \{0,1\}^{n_t}$ where superscript *s* represents "sharing" and value 1 or 0 indicates sharing or not. The k-space data from other frames shared with frame *j* at coil *c*, $\boldsymbol{k}_{j,c}^s \in \mathbb{C}^{n_x n_y}$, is described as follows:

$$\boldsymbol{k}_{j,c}^s = \sum_{j'=1}^{n_t} \left(\boldsymbol{m}_j^s\right)_{j'} \mathrm{diag}\left(\boldsymbol{m}_{j'}^u\right) \boldsymbol{F}\boldsymbol{C}_c\boldsymbol{x}_{j'}, \tag{3.1}$$

where $\boldsymbol{x}_{j'} \in \mathbb{C}^{n_x n_y}$ denotes the j'th frame of a MR slice and $\boldsymbol{m}_j^u \in \{0,1\}^{n_x n_y}$ is the subsampling mask for frame *j*. If a k-space location in frame j has more than one frame for data sharing, an average of shared data is taken at the location. Then, the initial guess $\boldsymbol{x}_{j,c}^{(0)}$ of the image of frame *j* at coil *c* is obtained by taking an inverse discrete Fourier transform of the combined acquired and shared k-space

data as:

$$\boldsymbol{x}_{j,c}^{(0)} = \boldsymbol{F}^H\left(\mathrm{diag}(\boldsymbol{m}_j^u)\boldsymbol{F}\boldsymbol{C}_c\boldsymbol{x}_j + \mathrm{diag}(1-\boldsymbol{m}_j^u)\boldsymbol{k}_{j,c}^s\right), \tag{3.2}$$

where $\boldsymbol{F}^H$ denotes the inverse discrete Fourier transform matrix. Finally, the coil-combined initial guess of the image is $\boldsymbol{x}_j^{(0)} = \sum_{c=1}^{n_c} \boldsymbol{C}_c^H \boldsymbol{x}_{j,c}^{(0)}$. Note that if at a k-space location the data of the current frame is not acquired and no sharing data from other frames is available, zero is filled at the current frame.

Similar to the approach that generates the SP, the data sharing mask $\boldsymbol{m}_j^s$ is a realization of random vector $\boldsymbol{M}_j^s$ containing independent Bernoulli random variables of which the probability $\boldsymbol{p}_j^s \in [0,1]^{n_t}$ is parameterized by network weights that are learned during training. The same $\sigma_v(\cdot)$ is used to approximate the sampling of Bernoulli random variables.

3.2.3 Image reconstruction

To take full advantage of temporal sparsity of DCE MRI, we adopted and modified the CRNN framework [85] to utilize RNN capability of extracting temporal correlations in dynamic data and to improve computation efficiency. In the modified CRNN, connections are over both temporal and iteration dimensions, which is not used in deep-learning methods for static image reconstruction.

In the typical compressed sensing (CS) framework, the MR image reconstruction problem is usually posed as a nonlinear optimization problem in a form:

$$\widehat{\boldsymbol{x}} = \underset{\boldsymbol{x}}{\operatorname{argmin}} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_2^2 + \lambda \mathcal{R}(\boldsymbol{x}), \tag{3.3}$$

where $\boldsymbol{x} \in \mathbb{C}^{n_x n_y n_t}$ is the set of fully sampled images, $\boldsymbol{A}$ is the MRI system matrix including effects of coil sensitivity, Fourier encoding and undersampling, $\boldsymbol{y}$ is the measured k-space data and $\mathcal{R}(\cdot)$ denotes a regularization term that represents our prior knowledge of $\boldsymbol{x}$. The first l_2 norm enforces the data-consistency (DC) between the reconstructed image and the acquired k-space data. For dynamic MR reconstruction, $\mathcal{R}(\cdot)$ is often employed as spatiotemporal TV [6] or low rank [86] constraints. By applying variable splitting and alternating minimization techniques, $\widehat{\boldsymbol{x}}$ can be solved iteratively by:

$$\boldsymbol{z}^{(i+1)} = \underset{\boldsymbol{z}}{\operatorname{argmin}} \mu \left\|\boldsymbol{x}^{(i)} - \boldsymbol{z}\right\|_2^2 + \lambda \mathcal{R}(\boldsymbol{z}), \tag{3.4}$$

$$\boldsymbol{x}^{(i+1)} = \underset{\boldsymbol{x}}{\operatorname{argmin}} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_2^2 + \mu\left\|\boldsymbol{x} - \boldsymbol{z}^{(i+1)}\right\|_2^2. \tag{3.5}$$

where μ is a penalty parameter and $\boldsymbol{z}$ is an auxiliary variable. The x update is often called the DC term. We follow D-POCSENSE [87] to implement this step as

$$\boldsymbol{x}^{(i+1)} = \sum_{c=1}^{n_c} \boldsymbol{C}_c^H \boldsymbol{F}^H \left(\text{diag}(1 - \boldsymbol{m}^u)\boldsymbol{F}\boldsymbol{C}_c\boldsymbol{z}^{(i+1)} + \boldsymbol{y}_c\right), \tag{3.6}$$

assuming $\mu \to 0$. The $\boldsymbol{z}$ update is a proximal operator, which we follow previous works [15], [85] to solve with a CNN-based de-aliasing network $\boldsymbol{\mathcal{D}}_\Theta(\cdot)$ parameterized by Θ so that $\boldsymbol{z}^{(i+1)} = \boldsymbol{\mathcal{D}}_\Theta(\boldsymbol{x}^{(i)})$. We used a modified version of the CRNN framework as the de-aliasing network.

Our CRNN framework (Figure 3.2) contains 5 components: 1) one bidirectional CRNN layer over both time and iterations (BCRNN-t-i), 2) one recurrent U-net (R-U-net-i), 3) one 2D CNN layer, 4) residual connection, and 5) multi-coil DC layers. Inspired by a previous work on multi-scale image deblurring [88], we replaced the second component, a CRNN layer over iterations (CRNN-i) of the original CRNN framework, with a compact U-net structure where a recurrent connection over iterations is placed on the bottleneck of the U-net, dubbed recurrent U-net (R-U-net-i). This modification largely reduces the GPU memory and training time, allowing other components being incorporated to the framework, e.g., the PK parameter estimation layer. The CRNN was also extended to multicoil settings by using a multicoil DC layer (Equation 3.6).

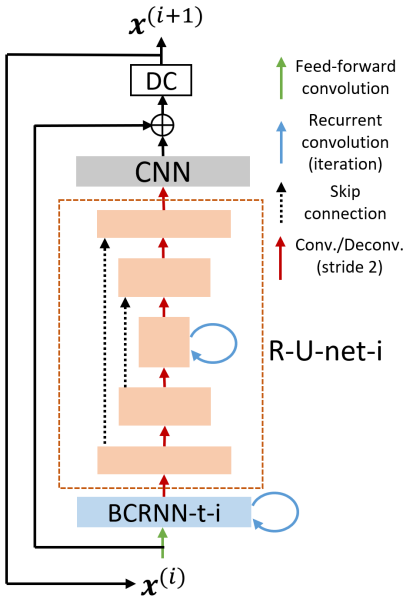

Figure 3.2: The overall structure of the image reconstruction network. The new R-U-net-i structure has 5 layers with 2 down sampling and up sampling paths. The bottleneck layer has a recurrent connection across iterations.

3.2.4 PK parameter estimation

eTofts model (equation 1.1) was implemented in this chapter. LLSQ method as described in section 2.2.1 was used for PK parameter estimation and was wrapped as a layer denoted as $LLSQ(\cdot\,; C_p)$ with the layer input as the reconstructed images $\boldsymbol{x} \in \mathbb{C}^{n_x n_y n_t}$.

3.2.5 Loss function

The learning objective is formulated to include both l_2 errors of reconstructed images and estimated PK parameters as the following:

$$\underset{\boldsymbol{p}^u, \Theta, \boldsymbol{p}^s}{\operatorname{argmin}} \mathbb{E}_{\boldsymbol{x}} \left[(1-\beta) \left\| LLSQ\left(\boldsymbol{\mathcal{D}}_\Theta\left(\boldsymbol{x}^{(0)}\right); C_p\right) - \theta \right\|_2^2 + \beta \left\| \boldsymbol{\mathcal{D}}_\Theta\left(\boldsymbol{x}^{(0)}\right) - \boldsymbol{x} \right\|_2^2 \right], \tag{3.7}$$

where $\boldsymbol{x}^{(0)} \in \mathbb{C}^{n_x n_y n_t}$ is the initial guess of an image as described by (2). $\theta \in \mathbb{R}^{3 n_{vox}}$ represents ground truth values of K^{trans}, k_{ep} and v_p. n_{vox} is the number of voxels in the anatomic region, and C_p is the ground truth CA concentration in plasma. The first term in the expectation represents the l_2 norm between reconstructed PK parameter maps and corresponding ground truth

values, called the parameter loss. The second term is the l_2 norm between reconstructed and ground truth images, called the image loss. β controls weighting between the parameter loss and the image loss.

3.3 Experiments

3.3.1 Multi-coil digital reference objects

The mcDROs of DCE MRI data used for training and testing were synthesized using patient specific AIFs and PK parameter maps estimated from real DCE MRI data of 17 patients with head and neck cancers enrolled on a protocol approved by the Institutional Review Board.

The DCE MR images of the patients were acquired using a 3D dynamic scanning sequence (TWIST) with an injection of 0.149 cc/kg of gadobenate dimeglumine on a 3 Tesla MRI scanner (Skyra, Siemens Healthineers, Erlangen Germany). The scanning parameters were: flip angle = 10º, echo time (TE) = 0.97ms, repetition time (TR) = 2.73ms, 60 time frames, voxel size = 1.56mm×1.56mm×2.5mm, and matrix = 192×192. The dynamic images were interpolated to have spatial resolution of 1.56mm×1.56mm×1.56mm. Axial slices were used due to their small dimensions and relative anatomical symmetry compared to the anatomy along the cranial-caudal direction, which might allow more aggressive k-space undersampling. For all cases, the patient specific AIFs were extracted manually by averaging the signal intensity-time curves of 20 voxels from the carotid artery, which had maximum intensities at the time frame before the enhancement peak [89], and then subtracted and divided by the average pre-contrast signal intensities of the voxels. The PK parameter maps were estimated using LLSQ with joint estimation of $\tau_{BAT}(\boldsymbol{r})$ in $\{0, 1, \ldots, 6\}$s. The voxels that had estimated parameters out of the physiologically reasonable range ($K^{trans} \in (0, 3)min^{-1}$, $k_{ep} \in (0, 6)min^{-1}$, $v_p \in (0, 0.55)$)

were considered incompatible with eTofts model and excluded from the final estimated maps. These parameter maps were regarded as ground truth values for mcDRO creation, and network training and testing.

The mcDROs with temporal resolutions of {2, 3, 4, 5}s and spatial resolution of 1.56mm×1.56mm×1.56mm were simulated. Using a TR of 2.73ms that can be achieved on this 3 T scanner yields undersampling rates (R) of {50, 34, 25, 20} for temporal resolutions of {2, 3, 4, 5}s, respectively. The undersampling rate is a reduction factor in the k-space subsampling relative to fully sampling of each frame. The mcDRO creation steps were: 1) simulate tissue concentration-time curves $C_t(\boldsymbol{r}, t)$ from ground truth PK maps of $K^{trans}(\boldsymbol{r})$, $k_{ep}(\boldsymbol{r})$, $v_p(\boldsymbol{r})$ using the eTofts model and patient specific AIF $C_p(t)$, and $\tau_{BAT}(\boldsymbol{r})$ in the voxels where the eTofts model is applicable; 2) add Gaussian noise to the simulated tissue concentration-time curves to have CNRs of 20 in each voxel to mimic the noise level present in real DCE data; 3) convert the tissue concentration-time curves to signal intensity images using baseline signal intensities $s_0(\boldsymbol{r}, 0)$ from the actual scan as: $s(\boldsymbol{r}, t) = [1 + C_t(\boldsymbol{r}, t)]s_0(\boldsymbol{r}, 0)$; 4) use the signal intensities from the real scan for the voxels where eTofts model is not applicable (e.g., vessels); 5) estimate coil sensitivity maps from fully sampled high resolution (0.875mm×0.875mm×3.3mm) post Gd T1-weighted images acquired immediately following the DCE scans by using ESPIRiT [90] and compressing coils from 30 to 8 coils [91]; 6) create complex DCE images by using the real phase variation estimated from post Gd T1-weighted images of the same patient $\phi(\boldsymbol{r})$ as $s(\boldsymbol{r}, t) = [1 + C_t(\boldsymbol{r}, t)]\mathrm{e}^{\mathrm{i}\phi(\boldsymbol{r})}s_0(\boldsymbol{r}, 0)$; and 7) create multi-coil DCE image series as well as multi-coil time-series of k-space data (see Figure 3.1). Note that the simulated mcDROs used realistic PK parameter ranges, AIFs and anatomy provided from patients.

Of the 17 patients, 8 were randomly selected for training, 3 for validation and 6 for testing. Note that during training, the ground truth patient specific AIF $C_p(t)$ was used for PK parameter estimation, while during testing, a patient specific AIF $\widehat{C_p}(t)$ was extracted from reconstructed MR images by a fully automated process that mimicked how an expert delineated an AIF [89]. As a brief, the anatomy surface was extracted by thresholding followed by closing operation. The contrast enhancement peak within the anatomy surface was detected after subtracting and dividing enhanced signals by baseline signals. The 20 voxels with the maximum enhancement in the dynamic frame of 3-10s prior to the peak of tissue enhancement was considered as an AIF. A 3×3 Gaussian filter was applied to the real scan for creation of the mcDROs used for training and validation, but not for the mcDROs set aside for testing in order to provide for better AIF characterization during testing.

3.3.2 Network training

The R-U-net-i unit of the reconstruction network used in the experiment contains 3 convolutions and 3 deconvolutions with stride 2, where the number of filters (n_f) are doubled or halved after each convolution or deconvolution. We used kernel size k=3. The number of filters of the BCRNN-t-i unit and the number of iterations were optimized and set to $n_f = 32$ and N=5, respectively. The estimated PK parameters were clipped to the physiologically reasonable range to stabilize the network training. Adam optimizer [61] was used with learning rate $1e^{-3}$. We used batch size = 1 and terminated training when the validation error was not improved in 3 consecutive epochs. All codes were implemented in PyTorch, and the experiments were performed on an NVIDIA RTX A6000 GPU with 48GB memory.

3.3.3 Evaluation and comparison

3.3.3.1 CRNN network architecture hyperparameter optimization

First, we optimized n_f and N across the combinations of $n_f \in \{8,16,32,64\}$ and $N \in \{3,5,10\}$ using the mcDRO with spatial resolution 1.56mm×1.56mm×1.56mm and temporal resolution 3s. The fixed Poisson disk SP and image loss were used. The best n_f and N were selected according to the average NRMSE across the three PK parameters and were used in experiments thereafter. The Poisson disk SP was generated using SigPy package (https://github.com/mikgroup/sigpy).

3.3.3.2 Ablation study of CRNN

To demonstrate effectiveness of CRNN for DCE MRI reconstruction, we performed an ablation study to evaluate effects of recurrent connections in temporal and iteration dimensions in CRNN by removing each or both connections. The same training data, loss function, SP, and evaluation metrics were used as in Section 3.3.3.1.

3.3.3.3 Ablation study of learnable k-t sharing

The learnable k-t sharing was compared with fixed neighbor-frame sharing strategies [15] in which {0,2,4,6,8,10,12} neighbor-frames of a center frame were shared. The same training data, loss function, SP, and evaluation metrics were used as in Section 3.3.3.1.

3.3.3.4 Ablation study of k-t sampling pattern learning module

To demonstrate the benefit of k-t SP learning module, we removed it from the proposed framework, and used the fixed Poisson disk and uniform random samplings to train the network. The same training data, loss function, and evaluation metrics were used as in Section 3.3.3.1.

3.3.3.5 Weighting parameter optimization

We optimized weighting parameter β in terms of the NRMSE of estimated PK parameters using the mcDRO with spatial resolution 1.56mm×1.56mm×1.56mm and temporal resolution 3s to search β within $\{0,0.01,0.1,0.3,0.7,1\}$. The optimized weighting parameter $\hat{\beta}$ was used in the following experiments.

3.3.3.6 The effect of contrast arrival time variation across subjects

One variation in the DCE signals across patients is the contrast arrival time (t_{CA}), which needs to be tested for its effect on the network performance. Here, t_{CA} was defined as the time when the DCE signal started rising from the baseline and calculated by the second time-derivative of the spatially averaged DCE signals, $t_{CA} = \underset{t}{argmax}\, d(\int s(\boldsymbol{r},t)d\boldsymbol{r})^2/d^2t$. We first analyzed the t_{CA} distribution in the training and testing datasets. Then, we aligned t_{CA} to 35s for all data by shifting the DCE series along the time dimension. We compared performance of the model trained by the time shifted data but tested using the data with and without time shifting. This experiment was done under spatial resolution 1.56mm×1.56mm×1.56mm and temporal resolution 3s. The performance results were used to determine how to deal with the variation of contrast arrival time variation across subjects.

3.3.3.7 Comparison with prior works

To demonstrate the advantages of joint optimization of k-t subsampling and reconstruction, we compared our method with two iterative dynamic MRI reconstruction methods. These two methods both explored spatiotemporal sparsity of dynamic MRI but did not optimize SPs, and instead used heuristic Poisson disk [92] and uniform random SPs. The first one is a dictionary learning-based indirect PK parameter estimation method [83] (DL) which was

implemented based on open-source code (https://github.com/sajanglingala/DCE_dictionary_recon). The ranges of PK parameters used for library learning were adjusted according to our data. As in the original paper, a population-based AIF was used in the test. The second method is a low-rank plus sparse model [78] (L+S) for which the open source code (https://github.com/JeffFessler/reproduce-l-s-dynamic-mri) is available. Proximal optimized gradient method (POGM) was used for optimization. We compared performances of the two methods with ours under different temporal resolutions of {2, 3, 4, 5}s and spatial resolution 1.56mm×1.56mm×1.56mm of the mcDROs of DCE MRI.

3.3.3.8 Evaluation metrics

The image reconstruction quality was measured using the SSIM and peak signal to noise ratio (PSNR) for different aspects of image similarity and quality. The PSNR is calculated for each reconstructed slice $\hat{\boldsymbol{x}} \in \mathbb{C}^{n_x \times n_y \times n_t}$ as $PSNR = 20\, log_{10}\big(max(|\boldsymbol{x}|)\,/MSE(|\boldsymbol{x}|, |\hat{\boldsymbol{x}}|)\big)$, where $\boldsymbol{x} \in \mathbb{C}^{n_x \times n_y \times n_t}$ is the ground truth image time series. The PK parameter estimation accuracy was evaluated by NRMSE, defined as $\sqrt{MSE\big(\theta_i, \widehat{\theta_i}\big)}/(max(\theta_i) - min\,(\theta_i))$, where θ_i and $\widehat{\theta_i}$ stand for the ground truth and estimation of the ith parameter, respectively, and $MSE(\cdot,\ \cdot)$ represents mean squared error. The mean and standard deviation of the parameters were calculated across testing mcDROs. Bland-Altman analysis was performed to assess bias and uncertainty in reconstructed PK maps compared with ground truth.

3.4 Results

3.4.1 CRNN network architecture hyperparameter optimization

The grid search of n_f and N of the CRNN architecture showed n_f =32 the best overall NRSMEs across three parameters, and N=5 better than 3 (Figure 3.3).

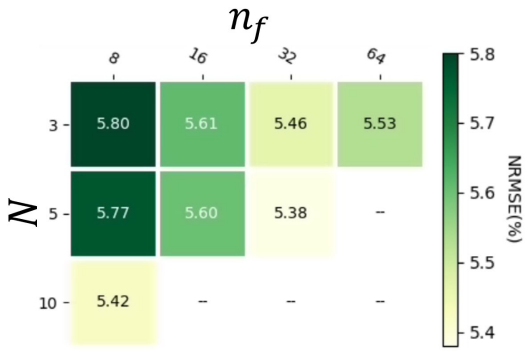

Figure 3.3: Grid search results of n_f and N of the CRNN architecture. Blank blocks are due to the combinations of n_f and N beyond the available GPU memory.

3.4.2 Ablation study of CRNN

The ablation study showed that the recurrent connections of the temporal and iteration dimensions in the CRNN reduced errors in the parameters, see NRMSEs in Table 3.1.

Table 3.1: Mean NRMSEs of estimated PK parameters with (Y) or without (N) recurrent connections in temporal and iteration dimensions.

Temporal	Iteration	NRMSE (%)
Y	Y	**5.38**
N	N	8.04
Y	N	5.61
N	Y	5.83

3.4.3 Ablation study of learnable k-t sharing

The learned k-t sharing strategy was superior to fixed neighbor-frame sharing strategies in terms of NRMSEs of PK parameters (Figure 3.4). Figure 3.5 shows the learned k-t space data sharing probability map. The data sharing is limited to fewer neighboring frames during rapid contrast uptake (wash-in) where the image contrast largely differs from the rest of the course. The k-space data across the frames before contrast uptake are shared among them with a high

probability where signal intensities change little. Similarly, it is true for frames after rapid contrast uptake.

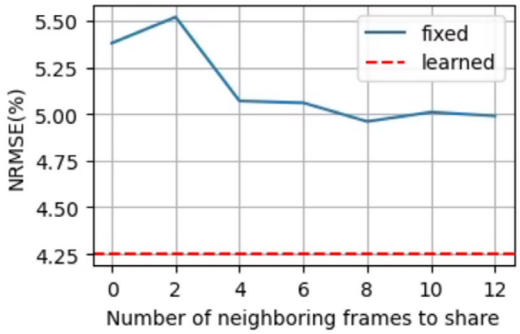

Figure 3.4: NRMSEs of PK parameter using fixed k-t sharing patterns (blue curve) over the number of shared neighboring frames. For comparison, NRMSE of learned k-t sharing is plotted as a red dashed line.

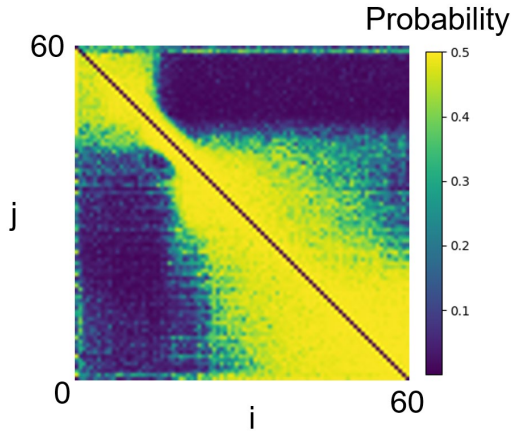

Figure 3.5: Learned k-t space data sharing probability map. For each point (i, j), the value represents the probability of sharing the k-space data of frame i with frame j.

3.4.4 Ablation study of k-t sampling pattern learning module

The learned k-t SP outperformed the fixed Poisson disk and uniform random samplings under the same CRNN reconstruction architecture in terms of NRMSEs of PK parameters (Table 3.2). Note that the CRNN was retrained for Poisson disk and uniform random sampling, respectively.

Table 3.2: Mean NRMSEs of estimated PK parameters using the Poisson disk, uniform random, and learning-based samplings.

SP	NRMSE (%)
Learned	**3.17**
Poisson	4.69
Uniform	10.17

Table 3.3: PSNR and SSIM (Mean ± σ) of reconstructed images and NRMSE of estimated PK parameters with different β values.

β	PSNR (dB)	SSIM	NRMSE (%)		
			K^{trans}	k_{ep}	v_p
0	24.66 ± 0.62	0.40 ± 0.01	3.75 ± 1.45	13.24 ± 5.76	16.76 ± 7.25
0.01	40.71 ± 0.39	0.89 ± 0.01	2.82 ± 1.50	7.72 ± 3.39	3.16 ± 1.57
0.1	41.55 ± 0.47	0.90 ± 0.00	**1.92 ± 0.96**	**6.48 ± 2.69**	**2.55 ± 1.10**
0.3	41.46 ± 0.48	0.90 ± 0.00	1.97 ± 1.02	6.82 ± 2.89	2.63 ± 1.10
0.7	41.22 ± 0.43	0.90 ± 0.01	2.49 ± 1.38	8.72 ± 4.19	3.53 ± 1.79
1	**41.65 ± 0.51**	**0.91 ± 0.00**	2.34 ± 1.50	7.46 ± 3.39	2.88 ± 1.62

Table 3.4: The PK parameter estimation NRMSEs (Mean ± σ) of the proposed model trained on time shifted DCE data and tested on the data with and without time shifting.

Testing t_{CA}	NRMSE (%)		
	K^{trans}	k_{ep}	v_p
with	2.16 ± 0.92	7.07 ± 2.88	3.08 ± 1.64
without	2.01 ± 0.91	6.94 ± 2.98	2.72 ± 1.14

3.4.5 Weighting parameter optimization

Searching the optimal weighting parameter β in {0,0.01,0.1,0.3,0.7,1} using the mcDRO yielded that β=0.1 had minimal NRMSEs of PK parameters and a near best PSNR of reconstructed images. The PSNR and SSIM of the reconstructed images and NRMSE of the estimated PK parameters of tested β values are shown in Table 3.3. Note that when $\beta = 1$, where the loss function is reduced to include the image loss l_2 only and is used commonly in deep learning-based reconstruction of MRI in the literature, the PSNR of reconstructed images was similar to one when $\beta = 0.1$ but averaged accuracy of estimated PK parameters was worse than one when $\beta = 0.1$. However, when $\beta = 0$, where the image loss was not included, both

image quality and PK parameter estimation accuracy were worse than ones with any $\beta > 0$. In the subsequent experiments, the optimal weighting parameter $\hat{\beta} = 0.1$ was used.

3.4.6 The effect of contrast arrival time variation across subjects

Variations in t_{CA} were observed in both training and testing datasets, with $t_{CA} = 38.5s \pm 3.6s$ in the training set and $t_{CA} = 40.5s \pm 4.5s$ in the testing set. The NRMSEs of PK parameters estimated from the proposed network trained using DCE data with the same t_{CA} were similarly well on testing data using a single t_{CA} value across all cases and using different subject-varied t_{CA} values (Table 3.4), indicating that the subject contrast arrival time variation within the range that we observed did not have a substantial effect on the PK parameter estimation accuracy for the proposed method.

3.4.7 Comparison with prior works under different temporal resolutions

Based upon the results in Table 3.4, we trained the model using the data without time-shifting of individual subject DCE time-series. The NRMSEs of the PK parameters estimated from the DCE data with different temporal resolutions and SPs by the proposed method compared to L+S and DL are summarized in Table 3.5.

The proposed method consistently outperformed L+S and DL that used either Poisson disk random sampling or uniform random sampling at all tested temporal resolutions or acceleration rates by approximately 2 to 20 times in NRMSE (%). The best performance achieved by the proposed method was at Δt=3s (R=34) for K^{trans} and at Δt=4s (R=25) for k_{ep} and v_p.

Examples of reconstructed PK images at temporal resolution of 2s of the proposed method, L+S with Poisson disk sampling, and DL with Poisson disk sampling as well as the

ground truth PK maps are shown in Figure 3.6. The proposed method generated the most perceptually similar PK maps to the ground truth maps, which supports the quantitative results. Note that the L+S method overestimated K^{trans} and v_p values and resulted in larger errors in k_{ep} estimation in this example slice. Also, the DL method showed overestimations in all three parameters.

The bias and uncertainty of the three PK parameters estimated from the DCE data at 2 s temporal resolution by the proposed method, L+S with Poisson disk sampling, and DL with Poisson disk sampling in gross tumor volumes of all 6 testing DROs are shown in the Bland-Altman plots in Figure 3.7. For comparison, plots of PK parameters estimated from fully sampled DCE data at the same temporal resolution were also included. The bias observed in the fully sampled data might be due to the added Gaussian noise [47], [48]. The proposed method showed comparable estimation bias and uncertainty to those from fully sampled DCE data, but reduced bias and uncertainty compared to DL and L+S methods. The proposed method reduced standard deviations of the PK parameter estimates by 0.10 (41%) and 0.19 (56%) compared to DL and L+S with Poisson disk sampling, respectively. A systematic bias on all three parameters was present in the L+S and DL methods.

The k_x-k_y plots of the learned sampling probability maps and sampling masks at frame 20 and the k_y-t plots at the central k_x for the DCE temporal resolution Δt=2s compared to the Poisson disk and uniform random sampling masks are shown in Figure 3.8. Note that the sampling probabilities near the k-space center learned by the proposed method varied over the contrast uptake time course. The learned sampling mask had a lower sampling density in the k-space center than Poisson disk sampling.

Table 3.5: PK parameter estimation NRMSEs (Mean ± σ) using the proposed method with learned sampling patterns (SP), L+S with Poisson disk and uniform random sampling, and DL with Poisson disk random sampling at $\Delta t \in \{2,3,4,5\}$s and corresponding $R \in \{50,34,25,20\}$.

Method	SP	Δt (s)	R	NRMSE (%)		
				K^{trans}	k_{ep}	v_p
Proposed	Learned	2	50	2.14 ± 1.01	6.71 ± 2.82	2.71 ± 1.21
		3	34	**1.92 ± 0.96**	6.48 ± 2.69	2.55 ± 1.10
		4	25	1.98 ± 0.90	**6.46 ± 2.61**	**2.52 ± 1.05**
		5	20	2.21 ± 1.09	7.14 ± 3.13	3.77 ± 2.45
L+S	Poisson	2	50	3.40 ± 1.87	10.73 ± 4.35	7.74 ± 5.13
		3	34	3.17 ± 1.75	10.22 ± 4.36	4.84 ± 2.15
		4	25	3.03 ± 1.46	10.32 ± 4.37	4.88 ± 2.24
		5	20	2.90 ± 1.17	10.62 ± 4.75	4.74 ± 2.12
L+S	Uniform	2	50	7.79 ± 7.00	12.62 ± 6.47	44.32 ± 20.80
		3	34	6.45 ± 4.28	12.48 ± 6.32	41.79 ± 21.57
		4	25	5.44 ± 4.08	12.76 ± 6.69	42.40 ± 19.59
		5	20	6.59 ± 3.94	13.08 ± 6.29	23.56 ± 13.04
DL	Poisson	2	50	7.30 ± 3.57	10.57 ± 4.60	6.40 ± 2.43
		3	34	7.31 ± 3.35	10.14 ± 4.30	6.51 ± 2.45
		4	25	8.05 ± 4.29	10.21 ± 4.42	6.68 ± 2.59
		5	20	8.02 ± 4.36	9.95 ± 4.25	6.90 ± 2.78
DL	Uniform	2	50	3.92 ± 1.66	13.46 ± 5.96	6.74 ± 2.83
		3	34	4.10 ± 1.94	13.41 ± 5.95	6.75 ± 2.83
		4	25	4.37 ± 2.33	13.15 ± 5.73	6.78 ± 2.86
		5	20	4.44 ± 1.99	13.09 ± 5.81	6.80 ± 2.84

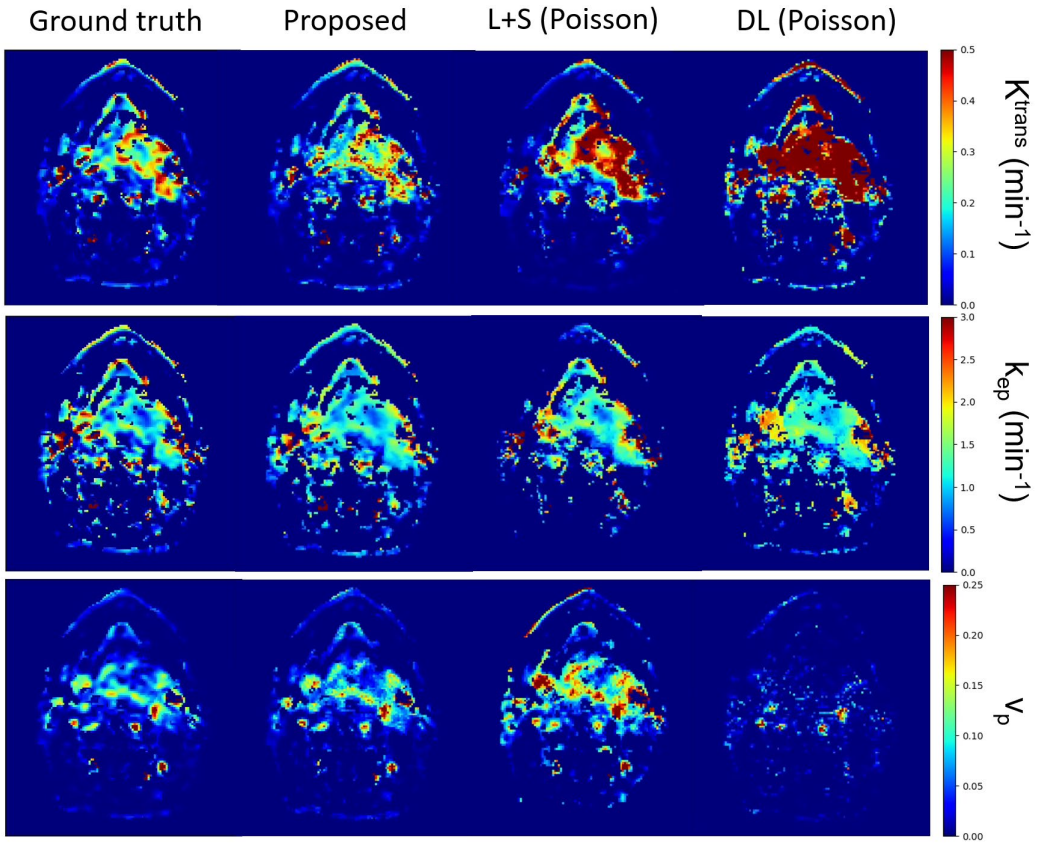

Figure 3.6: Example reconstructed PK parameter maps from one slice at temporal resolution of 2s (R =50) using the proposed method (second column), L+S with Poisson disk sampling (third column) and DL with Poisson disk sampling (forth column). The ground truth maps (first column) are also included for comparison.

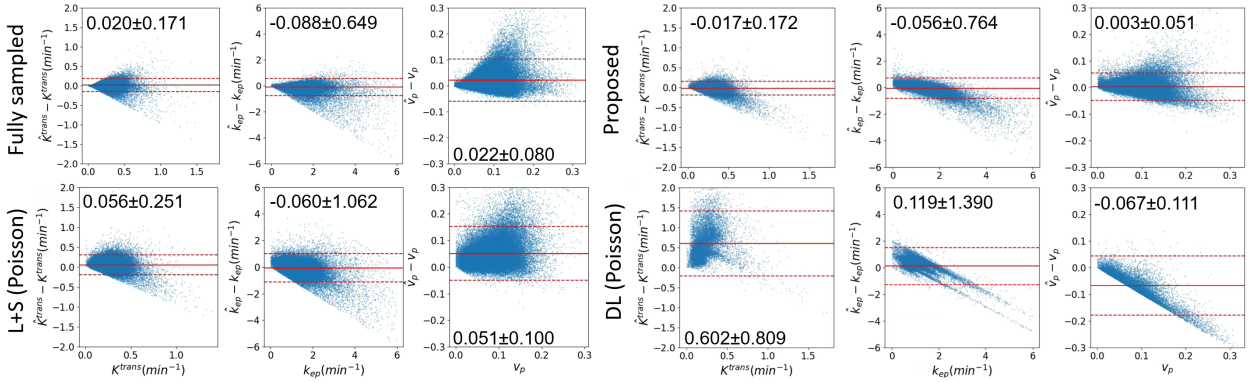

Figure 3.7: Bland-Altman plots of the difference between estimated $\widehat{\boldsymbol{K}}^{\boldsymbol{trans}}$, $\widehat{\boldsymbol{k}}_{\boldsymbol{ep}}$, and $\widehat{\boldsymbol{v}}_{\boldsymbol{p}}$ and reference values K^{trans}, k_{ep}, and v_p of fully sampled DCE data, the proposed method, L+S with Poisson disk sampling, and DL with Poisson disk sampling at temporal resolution of 2s (R=50). Each dot represents one tumor voxel in 6 mcDROs. The mean and 1.96×standard deviation were marked in each plot and represented by solid and dotted red lines, respectively.

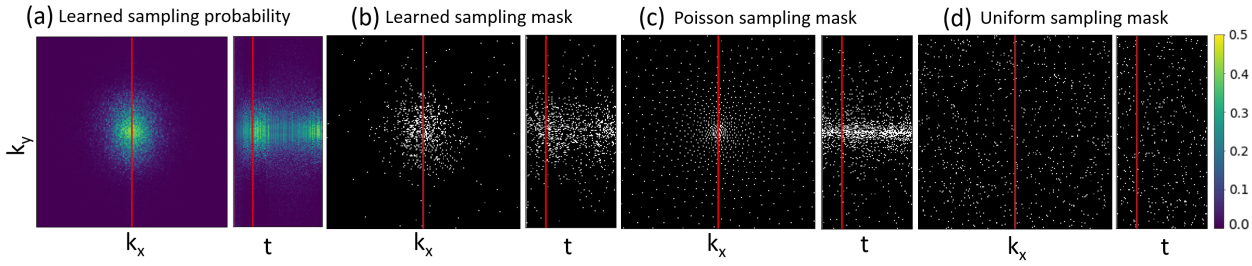

Figure 3.8: Illustration of learned phase encodings of k_x vs k_y at the 20th time frame and k_y vs t at the central k_x (respective left and right subpanels in each plot): (a) learned sampling probability maps by our method; (b) realization sampling masks by our method, (c) Poisson disk sampling masks, and (d) uniform random sampling masks. All maps were obtained at temporal resolution of 2s and R=50. White dots indicate locations of phase-encodings acquired in the k_x-k_y plane. Note that the first frame is fully sampled to provide baseline anatomy.

3.4.8 Examples of reconstructed DCE images and time courses

Examples of reconstructed images and time courses obtained by different methods are shown in Figure 3.9. Our proposed method included the optimization of the PK parameter estimates in the loss function, and resulted in more faithful reconstructed images and time courses of dynamic signals than DL and L+S methods.

3.4.9 Adaptation of the learnable SPs to specific anatomies

The learned sampling probability maps depend upon trained DCE signals and anatomies, see Figure 3.10. Note that different patterns were yielded from training on DCE time-series from different anatomic regions ((a) brain+neck, and (b) brain). Also, full width half maximums (FWHMs) of the probability density projections were different in k_x and k_y dimensions between anatomic regions. In contrast, empirical sampling patterns are pre-determined and one for all, which cannot present spatial sparsity in specific anatomy.

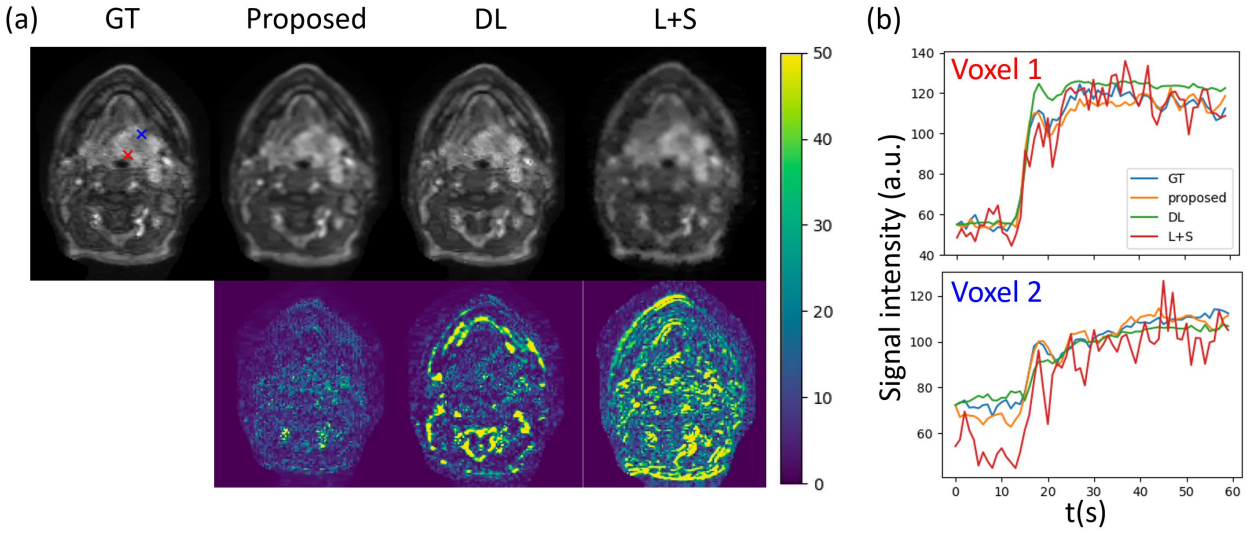

Figure 3.9: (a) Example reconstructed images at the 30th frame by different methods (first row) and their corresponding error maps (second row). (b) Time courses of dynamic signals at two example voxels within the tumor region reconstructed by different methods. The two voxels are marked by red and blue crosses in the Ground truth (GT) image in (a). In these examples, we used spatial resolution of 1.56×1.56×1.5mm and temporal resolution of 3s in mcDROs. Note large deviations and fluctuations in the DCE signals generated by DL and L+S methods, respectively.

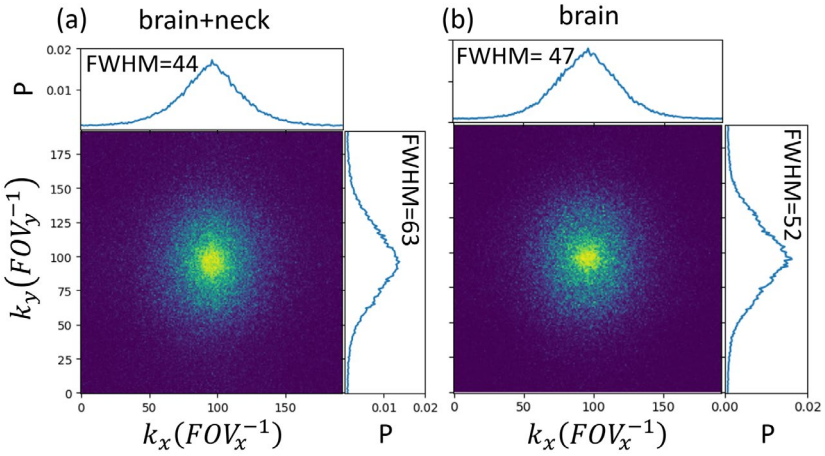

Figure 3.10: Learned sampling probability maps at the 30th frame trained by (a) a full field of view of the mcDROs including brain and neck slices and (b) brain-related slices only. Projections of the probability densities on the k_x and k_y dimensions are shown in the top and right plots of each probability map. The full width half maximum (FWHM) of each projection is shown in the top left corner of the plot. The mcDROs with spatial resolution of 1.56×1.56×1.56mm and temporal resolution of 3s were used. GT=Ground truth; P=Probability; FOV=Field of view.

3.5 Discussion

In this chapter, we presented a deep learning-based framework to jointly optimize k-t SPs and image reconstruction of PK parameters from DCE MRI by minimizing a loss function including l_2 errors of both image reconstruction and PK parameter estimation. Time series of sampling probability maps in the k-t space were optimally learned by the network to achieve rapid acquisition and accurate estimation of PK parameters. A PK parameter loss was optimally weighted into the objective function of image reconstruction, improved accuracy in the PK parameter estimation as well as quality of reconstructed images compared to using the image loss only. Overall, our proposed method performed superiorly and had reduced bias and uncertainty in the estimated PK parameters compared to two iterative dynamic MRI reconstruction methods. In addition, the proposed method was robust to patient-wise contrast arrival time variations. This method has the potential to increase spatial resolution of DCE MRI using a higher acceleration factor while providing accurate and precise PK parameter estimates. A further extension of the current approach by replacing the LLSQ layer with a neural network could simultaneously produce PK parameter estimations and corresponding uncertainties [93]. This may allow direct minimization of the uncertainty in PK parameters.

It is important to learn the optimal subsampling patten in the k-t space for dynamic image acquisition, instead of using a random SP following an assumed distribution, e.g., a Poisson distribution or uniform distribution. Similarly, temporal data sharing, although widely used in dynamic acquisitions, lacks optimization, and is often manually crafted [94], [95]. In this work, we extended LOUPE [84] to the time domain and added a TDoF to learn dynamic subsampling

and data sharing in the k-t space. In the learned sampling probability maps, the sampling density near the k-space center was high during initial contrast uptake, but decreased over time, and then increased at the end of the temporal acquisition. This could be explained as the data near the k-space center are important to capture fast contrast dynamics for accurate estimation of PK parameters, and then became less important over the time course. Also, the spread of sampling probabilities in the k-space at the frames near the time course center could be due to that the spatiotemporal sparsity is well captured by the CRNN network. The increased density near the k-space center at the end of the acquisition could be because of the zero initializations of the hidden features of the BCRNN layer which may be removed by using learnable initializations.

We included the PK parameter loss in the objective for image reconstruction. The weighting of the PK parameter loss in the objective had a nonlinear effect on NRMSEs of estimated PK parameters [50]. We found that an optimal weighting between the image loss and parameter loss improved both image and PK parameter reconstruction qualities. While the image loss provided direct guidance on image reconstruction, the parameter loss distilled PK modeling knowledge that the image l_2 loss might not be sensitive to in the model during training and in SP optimization. We observed that the sampling density was more concentrated in the k-space center when the model was trained with image loss only compared with that trained with combined image and parameter losses. This is possibly because the inclusion of the parameter loss enabled the network to directly learn hemodynamics in the signal time courses that influence the parameter estiamtion, thereby sampling the k-space center less frequently.

In our comparison with two other methods under different temporal resolutions, we observed a nonlinear dependence of NRMSE on temporal resolutions, which could be due to the interplay of temporal resolution, SNR and SPs. While both high temporal resolution and SNR are

beneficial for PK parameter estimation [47], [48], [96], the two are usually tradeoff in practice. Temporal resolutions of 3s and 4s represented a balance of these factors for the proposed method, resulting in the best NRMSEs. The AIF is another factor affecting the accuracy of PK parameter estimation [28]. Note that the proposed and L+S methods extracted patient-specific AIFs from the reconstructed images for PK parameter estimation, while the DL method used a population-based AIF that might contribute to, at least in part, the high bias observed in the Bland-Altman plot [97]. We also observed a generally higher instability in k_{ep} compared with K^{trans} and v_p as demonstrated in the corresponding NRMSEs (Table 3.4), which is consistent with prior reports [47].

In this chapter, the proposed method was validated using realistic multi-coil complex DROs to demonstrate its ability to reduce PK parameter estimation bias and uncertainty compared with iterative reconstruction methods using non-data driven SPs. One limitation of the current mcDROs simulation is that the real data was acquired using a view sharing technique at Δt=3s, which might reduce the high frequency temporal information in the subsequently derived AIFs used in the simulation of higher temporal resolution mcDROs (Δt<3s). High temporal resolution DCE MRI sequences [98], [99] could be utilized to create mcDROs at higher temporal resolutions in future works. Validating the proposed method in prospective studies in future investigations is also warranted. To facilitate these studies, realistic factors such as motion and native T_1 can be easily incorporated into training data to make the network robust to these factors present in real scans. In future prospective studies, the learned SPs can be implemented as 2D phase encoding locations in a 3D T_1 weighted spoiled gradient echo sequence without changing other sequence parameters. One of the challenges in applying the proposed method to prospective studies is that the timing of the contrast arrival may vary from patient to patient even

with the same contrast injection timing due to patient-specific factors such as cardiac output [100]. This may cause the SPs trained with a specific contrast arrival timing to fail in fully capturing the contrast dynamics information, which may result in uncertainty in PK parameter estimation. Our simulation of clinically measured variations in mcDROs demonstrated that small variations normally encountered in clinical scans had almost no impact on the PK parameter NRMSEs. This demonstrates the apparent robustness of the network to the normal variations in contrast arrival times expected in clinical DCE MRI scans of the neck region. Direct estimation of PK parameters from k-t space data has shown promise for parameter reconstruction in DCE MRI [74], [77]. An interesting direction of future works will be to incorporate data-driven priors into these approaches by extending the proposed framework to direct PK parameter reconstruction.

3.6 Conclusion

We have presented a jointly optimization framework for head and neck DCE MRI k-t SP and image reconstruction with an objective combining image reconstruction quality and PK parameter estimation accuracy. Optimization of k-t SP by learning the sparsity in the dynamic contrast enhanced images enables a dramatic reduction in the k-space sampling to achieve accurate and precise PK parameter estimations including in tumor regions while keeping similar spatial resolution. The proposed framework is general and can be applied to other quantitative MRI applications, such as T_1, T_2, and apparent diffusion coefficient quantification, to improve the scan efficiency. Future study is warranted to validate the proposed method in *in vivo* experiments.

Chapter 4 Sampling Pattern Optimization for Multi-Contrast MRI with A Fully Unrolled Reconstruction Network

4.1 Introduction

The purpose of this chapter is to accelerate conventional T_1- and T_2- weighted MR scans by jointly optimizing sampling patterns of the two scans.

MR images of different contrasts of the same anatomy, e.g., T_1 weighted and T_2 weighted images, are commonly acquired to provide complementary information and increase diagnostic power. However, the different contrasts are normally acquired and reconstructed independently, which ignores mutual information across multiple contrasts and prolong the total scanning time. Correlated features and contrast sparsity in multi-contrast MR images can be exploited to accelerate k-space acquisition. Neural networks have been used to exploit the contrast sparsity by joint multi-contrast reconstruction and outperform conventional sequential reconstruction under high undersampling rates [101]–[104]. However, sampling patterns of multi-contrast MR images can be jointly optimized with the reconstruction to further accelerate the scan. In addition, most of the previous works perform single- and multi-contrast learnings in a single network module, which may not be optimal if learning tasks of multiple contrasts are different.

In this chapter, we performed preliminary experiments to explore the idea of jointly learning k-space SPs of T_1 and T_2 weighted images with image reconstruction to improve quality of reconstructed images with undersampled acquisition.[1] In addition, we explored deep learning

[1]This chapter is based on our abstract for the 2022 AAPM annual meeting [46].

network structures that separately extract single-contrast features and cross-link the features for image reconstruction and compared it with extracting multi-contrast mutual information in a single module.

4.2 Materials and Methods

4.2.1 Problem formulation

Following a variable splitting formulation, the two-contrast image reconstruction problem can be formulated as the following optimization problem:

$$\underset{x_1,x_2}{\operatorname{argmin}} \|A_1x_1 - y_1\|_2^2 + \|A_2x_2 - y_2\|_2^2 + \lambda_1 R_1(x_1) + \lambda_2 R_2(x_2) + \lambda_3 R_3(x_1, x_2), \tag{4.1}$$

where x_i, y_i, A_i, and $R_i(\cdot)$ for $i \in \{1, 2\}$ are the image, k-space data, sampling operator, and regularization for T_i contrast, and $R_3(\cdot)$ is the joint regularization term representing sparsity in the contrast dimension. This problem can be solved with a variable splitting formulation that decouples single-contrast and multi-contrast learnings:

$$\underset{x_1,x_2,z_1,z_2,z_3,z_4}{\operatorname{argmin}} \sum_{i=1}^{2}(\|A_ix_i - y_i\|_2^2 + \lambda_i R_i(x_i) + \mu_i\|x_i - z_i\|_2^2) + \lambda_3 R_3(z_3, z_4) + \mu_3\|x_1 - z_3\|_2^2 + \mu_4\|x_2 - z_4\|_2^2, \tag{4.2}$$

which can be solved iteratively as:

$$\hat{z}_1 = \underset{z_1}{\operatorname{argmin}}\ \lambda_1 R_1(z_1) + \mu_1\|x_1 - z_1\|_2^2, \tag{4.3}$$

$$\hat{z}_3 = \underset{z_3}{\operatorname{argmin}}\ \lambda_3 R_3(z_3, z_4) + \mu_3\|x_1 - z_3\|_2^2, \tag{4.4}$$

$$\hat{x}_1 = \underset{x_1}{\operatorname{argmin}} \|A_1x_1 - y_1\|_2^2 + \mu_1\|x_1 - \hat{z}_1\|_2^2 + \mu_3\|x_1 - \hat{z}_3\|_2^2, \tag{4.5}$$

$$\hat{z}_2 = \underset{z_2}{\operatorname{argmin}}\ \lambda_2 R_2(z_2) + \mu_2\|x_2 - z_2\|_2^2, \tag{4.6}$$

$$\hat{z}_4 = \underset{z_4}{\operatorname{argmin}}\ \lambda_3 R_3(\hat{z}_3, z_4) + \mu_4\|x_2 - z_4\|_2^2, \tag{4.7}$$

$$\hat{x}_2 = \underset{x_2}{\operatorname{argmin}} \|A_2 x_2 - y_2\|_2^2 + \mu_2 \|x_2 - \hat{z}_2\|_2^2 + \mu_4 \|x_2 - \hat{z}_4\|_2^2, \tag{4.8}$$

where $\hat{z}_1$ and $\hat{z}_2$ represent the single contrast reconstruction results for T_1 and T_2 weighted images, respectively. $\hat{z}_3$ and $\hat{z}_4$ represent the refined T_1 and T_2 weighted images resulting from multi-contrast representation learning, respectively. Each z_i update for $i \in \{1,2,3,4\}$ was implemented by a U-Net in a fully unrolled neural network (FU-net) consists of four unrolled U-Nets, two for single-contrast learning and another two for multi-contrast learning. Each x_i update for $i \in \{1,2\}$ was implemented with a weighted average of $\hat{z}_i$ and $\hat{z}_{i+2}$ followed by a DC layer.

For comparison, two other network structures, one without multi-contrast mutual information learning and another with a different learning strategy, were studied. The first network structure contains two separate unrolled U-Nets (S-net) with DC that independently reconstruct T_1 and T_2 weighted images. The second one consists of a single multi-channel unrolled U-Net (MC-net) with DC with input data of the two contrasts concatenated along the channel dimension to simultaneously reconstruct them. The network structures investigated are shown in Figure 4.1. All networks have 5 unrolls.

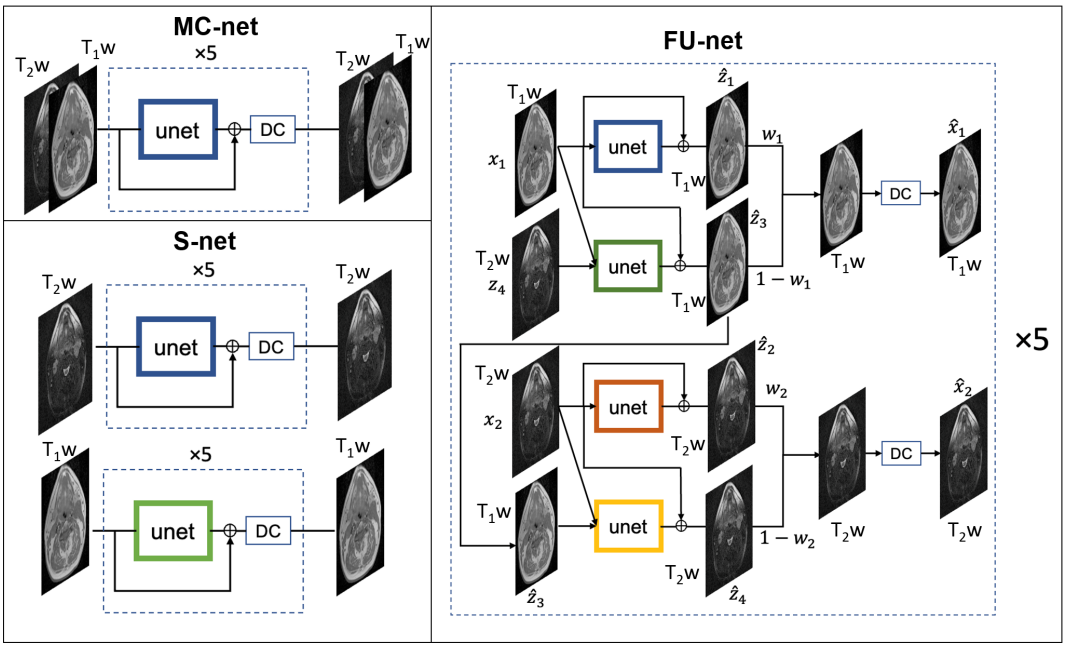

Figure 4.1: Network structures of S-net, MC-net, and FU-net.

4.2.2 Experiments

Magnitude images of fully sampled multi-slice head and neck T_1 and T_2 weighted images acquired in 16 patients who participated in an institutional review board approved protocol were used in the experiments. For both contrasts, the matrix size is 320×320, the median in-plane resolution is 0.80mm (0.78mm-0.94mm), the median number of slices is 75 (70-85), and the slice thickness is 3.3mm. The two contrast images were well aligned upon visual inspection. No image registration was performed. The k-space data was simulated from the magnitude images by inverse Fourier transform. Zero-padding in the k-space was removed to avoid subtle data crime [105], which reduced the matrix size from 320×320 to 192×192. In the first experiment, nine networks were trained for FU-net, S-net, and MC-net using retrospectively undersampled k-space data with a SP of 1D Gaussian, 2D variable density (VD), or 2D uniform random and an acceleration rate (R) of 4, 8 or 12. 615/230/360 2D slices of T_1 and T_2 weighted images from 8/3/5 patients were used for training/validation/testing. Reconstructed T_1 and T_2 weighted images from FU-net were compared with ones from S-net and MC-net. In the second experiment, 1D and 2D SPs were optimized for both contrasts jointly with FU-net by modeling them as i.i.d. samples from a learnable multivariate Bernoulli distribution [22]. R=4, 8, and 12 were tested. An image l_2 loss was used as the training loss function. The image quality was evaluated by PSNR and SSIM. To ensure fairness, the number of parameters of all networks were kept approximately the same. All training was terminated when the validation loss was not improved in 10 consecutive epochs. The code was implemented in PyTorch. The experiments were performed on an NVIDIA RTX A6000 with 48GB memory.

4.3 Results

The reconstruction performances of different methods using the three empirical SPs at the three Rs are compared in Table 4.1. The FU-net consistently outperformed MC-net and S-net in terms of PSNR and SSIM for both T_1 and T_2 contrasts. S-net and MC-net showed comparable performance, indicating inefficiency in exploiting multi-contrast sparsity by MC-net.

The reconstruction performances of FU-net jointly trained with learnable SPs are summarized in Table 4.2. Both 1D and 2D learned SPs outperformed their empirically designed counterparts under all acceleration factors. Example reconstruction results are shown in Figure 4.2, which are consistent with quantitative results. FU-net with the 2D learned SP had the best visual performance. Figure 4.3 shows the 1D and 2D learned SPs together with 1D Gaussian, 2D uniform random and VD SPs. Samplings for both 1D and 2D learned patterns were mostly concentrated near the center of k-space compared with other empirical SPs.

Table 4.1: Multi-contrast image reconstruction PSNRs and SSIMs (mean (standard deviation)) of different methods under 1D Gaussian, 2D uniform random, and 2D VD sampling and R=4, 8, and 12. SP=sampling pattern.

R	SP	Metrics	S-net	MC-net	FU-net	SP	S-net	MC-net	FU-net	SP	S-net	MC-net	FU-net
4	1D Gaussian	T1w PSNR (dB)	34.67(0.75)	35.00(0.65)	**35.40(0.73)**	2D uniform	34.36(1.06)	35.21(1.07)	**36.22(1.13)**	2D VD	39.84(0.66)	39.87(0.63)	**40.08(0.64)**
		T2w PSNR (dB)	35.57(1.19)	35.40(1.13)	**35.99(1.2)**		36.72(1.5)	36.20(1.42)	**37.80(1.69)**		38.64(1.42)	38.54(1.4)	**38.75(1.44)**
		T1w SSIM	0.88(0.02)	0.89(0.02)	**0.90(0.01)**		0.87(0.02)	0.89(0.02)	**0.90(0.02)**		0.96(0.01)	0.96(0.01)	**0.96(0.01)**
		T2w SSIM	0.82(0.03)	0.81(0.03)	**0.83(0.03)**		0.84(0.03)	0.83(0.03)	**0.86(0.03)**		0.89(0.03)	0.88(0.02)	**0.89(0.03)**
8	1D Gaussian	T1w PSNR (dB)	28.75(0.63)	29.28(0.62)	**29.84(0.46)**	2D uniform	28.17(0.85)	28.49(0.81)	**29.49(0.96)**	2D VD	34.30(0.53)	34.40(0.53)	**34.52(0.54)**
		T2w PSNR (dB)	30.9(0.99)	30.57(0.95)	**31.32(1.02)**		31.18(1.21)	30.48(1.04)	**31.99(1.27)**		34.94(1.06)	34.79(1.03)	**35.04(1.07)**
		T1w SSIM	0.70(0.03)	0.72(0.02)	**0.74(0.02)**		0.68(0.03)	0.70(0.03)	**0.73(0.03)**		0.89(0.01)	0.89(0.01)	**0.89(0.01)**
		T2w SSIM	0.61(0.03)	0.60(0.04)	**0.66(0.03)**		0.61(0.04)	0.59(0.04)	**0.66(0.04)**		0.79(0.03)	0.78(0.03)	**0.79(0.03)**
12	1D Gaussian	T1w PSNR (dB)	26.25(0.51)	26.50(0.54)	**26.51(0.57)**	2D uniform	25.97(0.74)	26.47(0.76)	**26.6(0.82)**	2D VD	32.04(0.53)	32.16(0.52)	**32.26(0.52)**
		T2w PSNR (dB)	28.56(0.92)	28.37(0.89)	**28.93(0.98)**		29.04(1.12)	28.57(1.0)	**29.75(1.16)**		33.36(0.96)	33.23(0.93)	**33.50(0.96)**
		T1w SSIM	0.56(0.03)	0.58(0.03)	**0.58(0.03)**		0.55(0.03)	0.59(0.03)	**0.60(0.04)**		0.83(0.02)	0.83(0.02)	**0.83(0.02)**
		T2w SSIM	0.46(0.03)	0.46(0.03)	**0.50(0.03)**		0.48(0.04)	0.47(0.04)	**0.53(0.04)**		0.72(0.04)	0.72(0.04)	**0.73(0.04)**

Table 4.2: Multi-contrast image reconstruction PSNRs and SSIMs (mean (standard deviation)) using learned SPs and FU-net.

Method	R	SP	T1w PSNR (dB)	T2w PSNR (dB)	T1w SSIM	T2w SSIM
FU-net	4	1D	36.90(0.74)	37.00(1.27)	0.93(0.01)	0.87(0.03)
		2D	43.41(1.01)	41.44(2.32)	0.98(0.01)	0.93(0.02)
	8	1D	31.00(0.63)	32.52(0.98)	0.80(0.02)	0.72(0.03)
		2D	36.44(0.75)	36.64(1.31)	0.92(0.01)	0.83(0.04)
	12	1D	27.79(0.47)	30.25(0.94)	0.65(0.02)	0.59(0.04)
		2D	33.67(0.68)	34.72(1.11)	0.87(0.02)	0.77(0.04)

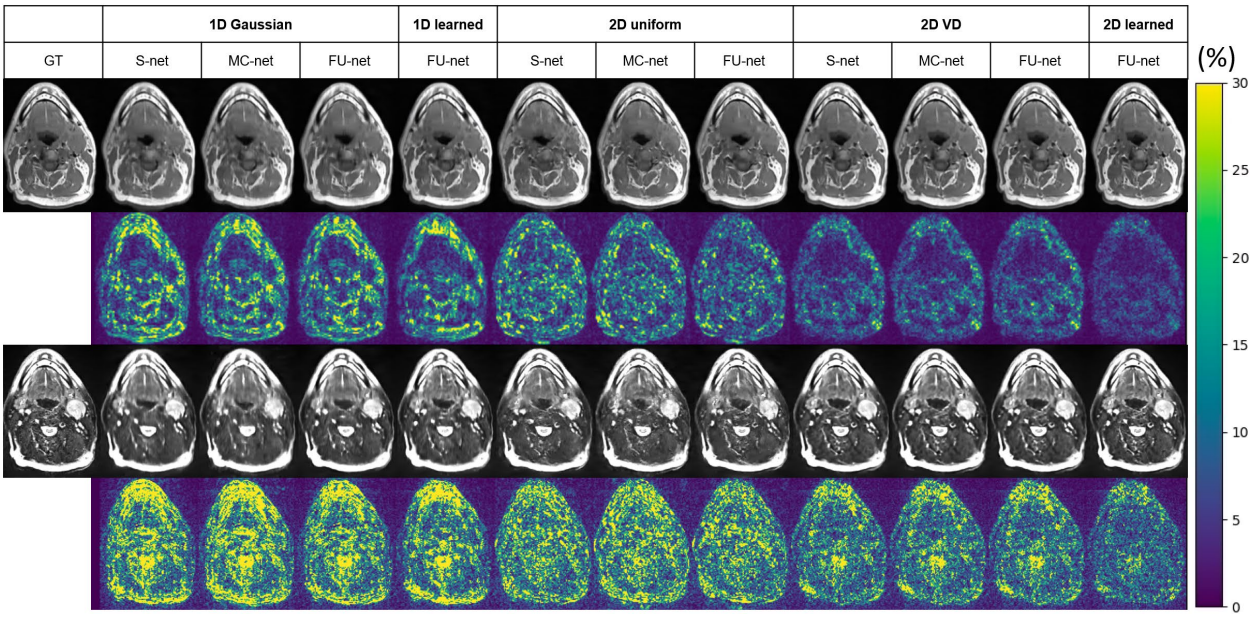

Figure 4.2: Example reconstructed T_1 (first row) and T_2 (third row) weighted images and corresponding error maps (second row and forth row, respectively.) under R=8. The error maps are shown as percentage absolute errors relative to the mean intensity of the anatomic regions of each corresponding contrast. GT=ground truth.

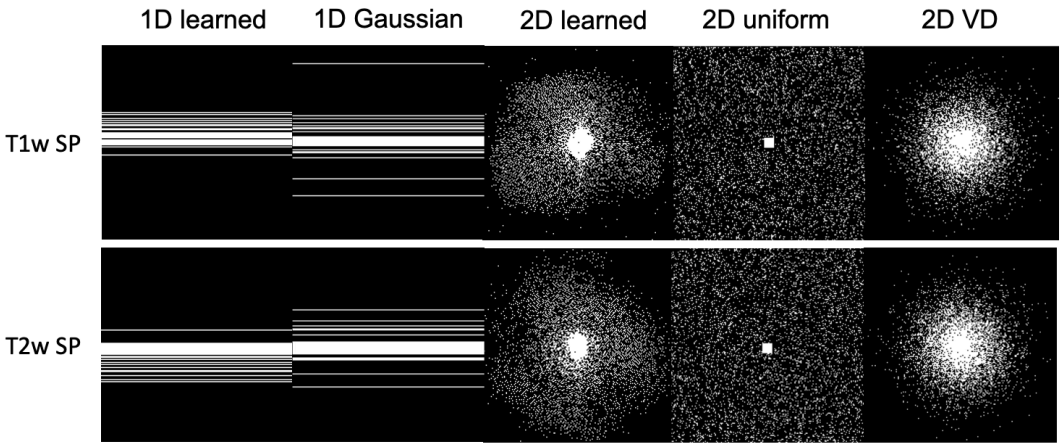

Figure 4.3: Learned SP for T_1 and T_2 weighted scans compared with 1D Gaussian and 2D uniform random and 2D variable density (VD) SPs.

4.4 Discussion

This work shows effectiveness of FU-net in multi-contrast image reconstruction and advantages of data-driven SPs over empirical ones, which may further accelerate multi-contrast imaging. Under 8-fold undersampling, FU-net reconstruction with jointly learned sampling patterns provides high fidelity images and recovers fine structures of anatomy. The FU-net decouples single- and multi-contrast learnings by using dedicated modules for each task. We compared FU-net with a S-net that processes each contrast individually and a MC-net that exploited sparsity in multi-contrast images with a single network. The performance improvement of FU-net is greater using 1D Gaussian and 2D uniform sampling strategies than 2D VD sampling. The 2D learned sampling patterns in the k-space have a sharper sampling density transition from the center to the peripheral regions than the VD sampling distribution and have a larger densely sampled k-space center than the random uniform sampling pattern. This pattern could be related to the k-space energy distribution of anatomy of interest, and is hard to guess empirically, which requires further investigation. The 1D learning sampling patterns of T_1 and T_2 contrasts complement each other by sampling opposite halves of the k-space, which could be a result of the learned correlation between the two contrasts and Hermitian symmetry of the magnitude images. The relationship between 2D learned sampling patterns of T_1 and T_2 contrasts will be investigated in the future.

Beyond this preliminary experiment, further investigation on the feature maps of different network structures may give more insights on the cause of the performance differences and the connection between single- and multi-contrast learning. Validation of the framework on real k-space data and prospectively undersampled acquisitions are also warranted. The extension of the

decoupling framework to MRI scans with more contrasts (e.g., MR fingerprinting [29]) and more advanced multi-contrast representation learning such as deep dictionary learning [106] may be investigated in the future.

Chapter 5 Improved 3D MR Fingerprinting via Memory-Efficient Synergic Optimization of 3D Spiral Trajectory, Image Reconstruction and Parameter Estimation (SOTIP)

5.1 Introduction

MRF acquisition and reconstruction involve multiple components, including k-space sampling, image reconstruction and parameter estimation, which all have critical impact on efficiency of MRF. While a number of works improve each component separately, joint optimization of these components could further improve MRF efficiency, which has not been explored.[1] The purpose of this chapter is to improve accuracy and precision of quantified MRF tissue parameters and shorten reconstruction time while keeping the same scan time and similar spatial resolution by using deep learning reconstruction to synergically optimize k-space sampling, image reconstruction and parameter estimation.

Several dynamic 3D acquisition patterns have been proposed to boost MRF scan efficiency [8], [107]–[110]. Stack-of-spirals trajectory is an efficient approach for rapid sampling and has been demonstrated to achieve whole-brain imaging with spatial resolution of $1.2\times1.2\times3.0$ mm^3 within 4.6 min [110]. Full 3D non-Cartesian sampling schemes have high flexibility in trajectory design and great sampling efficiency [8], [107]–[109]. Cao et. al. proposed a tiny-golden-angle shuffling (TGAS) sampling scheme of full 3D spiral MRF and achieved 1mm-isotropic resolution whole-brain T_1, T_2, and proton density mapping in less than 2 min [8]. However, it may be possible to optimize the sampling scheme systematically for

[1]This chapter is based on a manuscript to be submitted to *IEEE Transactions on Medical Imaging*.

specific anatomy and jointly with image reconstruction. Recently, data-driven design of k-space sampling patterns by deep learning has shown promise for accelerating MRI by exploiting spatial sparsity in anatomy of interest, for example, optimizing 2D and 3D Cartesian [19], [22], [81], [111] and non-Cartesian [21], [82], [112], [113] k-space trajectories for static MR applications. Also, benefits of optimized dynamic samplings in 3D Cartesian have been demonstrated in DCE [45] and quantitative susceptibility (QSM) MRI [114] by exploiting spatiotemporal sparsity. However, to the best of our knowledge, no extension of data-driven optimization of 3D non-Cartesian sampling to MRF has been investigated, which could improve the scan efficiency further. This is in part due to requirements of prohibitively large GPU memory footprint and long optimization time.

MRF reconstruction is done commonly by reconstructing time-series of images first and then followed by estimation of quantitative tissue parameter maps. Accuracy of parameter quantification thus depends on quality of reconstructed images. MRF data are highly redundant in the image-time domain, which is leveraged to compensate for the high degrees of undersampling in k-space. Generic spatiotemporal sparsity by low rank (LR) [7] or locally low rank (LLR) modeling [8], [9] and sequence-specific temporal sparsity by temporal subspace modeling have been utilized in iterative image reconstructions. However, when applied to full 3D non-Cartesian MRF, these methods require long reconstruction time and huge demand of random-access memory (RAM). For example, the LLR reconstruction took 4-5 hours on a Linux server and required approximately 400GB of RAM and 1 TB of swap to achieve whole-brain MRF reconstruction at 1mm-isotropic resolution [8]. A work utilized stochastic gradient descent and 9.5 GB RAM to reduce the reconstruction time to 0.4 hours on phantom data [115], which is still insufficient in clinical practice. Furthermore, generic spatiotemporal sparsity assumed by LR

or LLR does not fully capture redundancy in MRF associated with specific anatomy, which is worthwhile to exploit further. Deep learning-based MR image reconstruction methods have recently shown promising performance in various MR applications [18]. MBDL reconstruction methods [34]–[37] achieve state-of-the-art MR image reconstruction performance and increase generalizability even with a small amount of training data. MBDL that incorporates MR physics as a DC layer into deep learning-based reconstruction benefits from both learning-based and model-based image reconstruction advantages. However, to date MBDL has been applied only to 2D/3D static or 2D dynamic scans with non-Cartesian samplings [116]–[120] due to huge demands of GPU memory and training time. For example, naïve implementation of MBDL for static MRI with full 3D non-Cartesian trajectories requires over 250 GB of GPU memory [117]. For 3D MRF, spatiotemporal images of a single subject can take approximately 50 GB, which needs to be stored in RAM for the DC step. Even with several memory-efficient network training strategies [117], [119]–[122], an extension of MBDL to full 3D non-Cartesian MRF is hindered by high computation cost and GPU memory requirement. As a result, there is no MBDL framework proposed for full 3D non-Cartesian MRF.

As discussed, conventional MRF reconstruction methods construct image and quantitative parameter maps sequentially and independently. Previous works on DCE MRI [45], [123] suggest that joint image-parameter reconstruction could improve accuracy of quantitative parameters. The joint image-parameter reconstruction of MRF has been implemented in a direct iterative parameter estimation framework using maximum likelihood [124], in which solving a highly nonconvex problem was a challenge and sensitive to initialization. In a supervised deep learning framework [125], a fully connected network and a CNN were jointly trained to exploit respective temporal and spatial sparsity of 2D MRF without DC. Hamilton et. al. [126] proposed

to use an unsupervised learning approach with DC incorporation for jointly optimizing image and parameter reconstructions for 2D cardiac MRF. However, this method led to a relatively long reconstruction time, approximately 1.1h for a single subject because the network was trained de novo for every subject. Supervised joint training of image and parameter reconstructions of MRF could provide better estimation of parameters with shorter reconstruction time.

In this work, we aim to accelerate both acquisition and reconstruction of 3D MR Fingerprinting via **S**ynergic **O**ptimization of 3D Spiral **T**rajectories, **I**mage Reconstruction, and **P**arameter Estimation (**SOTIP**). To achieve this objective, we make novel contributions in three major aspects: 1) we enable MBDL for full 3D spiral MRF on a single GPU by utilizing several techniques to reduce computation and GPU memory demand; 2) we jointly optimize image and parameter reconstructions in an end-to-end fashion and reduce error propagation from image reconstruction to final parameter quantification; and 3) we synergically optimize rotation angles of full 3D spiral trajectories of MRF using MBDL to further improve acquisition efficiency. We validated our methods using both simulated and *in vivo* MRF scans with different undersampling rates and observed consistent improvement in parameter quantification accuracy and reduction in reconstruction time compared with the state-of-the-art LLR reconstruction [8], [109].

5.2 Materials and Methods

5.2.1 Problem formulation

The k-t-space data $\boldsymbol{Y}$ acquired by a MRF experiment can be formulated as follows:

$$\boldsymbol{Y} = \boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{M}}(\boldsymbol{X}); \boldsymbol{\Omega}) + \boldsymbol{N}, \tag{5.1}$$

where $\boldsymbol{X} \in \mathbb{R}^{N_v \times N_p}$ denotes a discrete approximation of an underlying continuous parameter map, $\boldsymbol{\mathcal{M}}(\cdot): \mathbb{R}^{N_v \times N_p} \rightarrow \mathbb{C}^{N_v \times T}$ represents a MRF signal model, $\boldsymbol{\mathcal{A}}(\cdot\,; \boldsymbol{\Omega}): \mathbb{C}^{N_v \times T} \rightarrow \mathbb{C}^{N_c N_k \times T}$ is a MR system operator, $\boldsymbol{N} \in \mathbb{C}^{N_c N_k \times T}$ is i.i.d. complex Gaussian noise usually assumed in MRI, and $\boldsymbol{\Omega} \in \mathbb{R}^{N_k T \times N_d}$ depicts k-t trajectories of N_d-dimensional images. N_v is the total number of voxels in a parameter map, N_p is the number of parameters of interest, N_k is the number of k-space samples acquired per TR, T is the total number of TRs, and N_c is the number of receive coils. The parameter map $\widehat{\boldsymbol{X}}$ is estimated by a reconstruction algorithm from acquired data $\boldsymbol{Y}$.

We propose a SOTIP framework to solve Equation 5.1, which is illustrated in Figure 5.1. The first step of the pipeline process is to reconstruct a low-rank representation of MRF image-time series from the acquired data. Learnable parameters $\boldsymbol{\Theta}_I$ of a MBDL image reconstruction network, $\boldsymbol{\mathcal{I}}_{\boldsymbol{\Theta}_I}(\cdot\,; \boldsymbol{\Omega}): \mathbb{C}^{N_k \times T} \rightarrow \mathbb{C}^{N_v \times K}$, are optimized to map acquired k-t-space data $\boldsymbol{Y} \in \mathbb{C}^{N_k \times T}$ to temporal subspace coefficient (TSC) images $\boldsymbol{C} \in \mathbb{C}^{N_v \times K}$ of temporal subspace bases (TSB) or principal components of a set of signal-time curves in a representative MRF dictionary, where $\boldsymbol{C} = \boldsymbol{\mathcal{M}}(\boldsymbol{X})\boldsymbol{\Phi}'$, $\boldsymbol{\Phi} \in \mathbb{C}^{K \times T}$ is a truncated set of K ($K \ll T$) TSB, and $\boldsymbol{\Phi}'$ denotes the complex conjugate transpose of $\boldsymbol{\Phi}$. The sparsity in the temporal or contrast domain of MRF data allows us to reconstruct TSC images with negligible truncation errors compared with whole image-time series and substantially reduces memory footprint. The training objective of $\boldsymbol{\mathcal{I}}_{\boldsymbol{\Theta}_I}$ is a reconstruction error of coefficient images that is a weighted l_2 distance between estimated coefficient $\widehat{\boldsymbol{C}} = \boldsymbol{\mathcal{I}}_{\boldsymbol{\Theta}_I}(\boldsymbol{Y}; \boldsymbol{\Omega})$ and ground truth $\boldsymbol{C}$ and written as:

$$\widehat{\boldsymbol{\Theta}}_I, \left(\widehat{\boldsymbol{\Omega}}\right) = \underset{\boldsymbol{\Theta}_I, (\boldsymbol{\Omega})}{\operatorname{argmin}} \mathbb{E}_{\boldsymbol{X}} \left[\left\| \boldsymbol{W}_{\boldsymbol{C}} \left(\boldsymbol{\mathcal{I}}_{\boldsymbol{\Theta}_I}(\boldsymbol{Y}; \boldsymbol{\Omega}) - \boldsymbol{\mathcal{M}}(\boldsymbol{X})\boldsymbol{\Phi}' \right) \right\|_F^2 \right], \tag{5.2}$$

where $\boldsymbol{W}_{\boldsymbol{C}} \in \mathbb{R}^{K \times K}$ is a diagonal weighting matrix on different bases and is empirically chosen for better parameter reconstruction quality. $\widehat{\boldsymbol{C}}$ is then fed into a pre-trained parameter estimation

network $\boldsymbol{\mathcal{P}}_{\boldsymbol{\Theta}_P^*}(\cdot)\colon \mathbb{C}^K \to \mathbb{R}^{N_p}$ parameterized by $\boldsymbol{\Theta}_P^*$ to estimate the parameter maps voxel-by-voxel, see Figure 5.1(a).

Given the parameters as the final estimates, an absolute percentage loss directly on T_1 and T_2 is considered as a training objective of $\boldsymbol{\mathcal{J}}_{\boldsymbol{\Theta}_I}$ in Variant 2 (Figure 5.1(b)):

$$\widehat{\boldsymbol{\Theta}}_I, \widehat{\boldsymbol{\Theta}}_P = \underset{\boldsymbol{\Theta}_I, \boldsymbol{\Theta}_P}{\operatorname{argmin}} \mathbb{E}_{\boldsymbol{X}} \left[\sum_{n_v, n_p} \left| \widehat{\boldsymbol{X}}_{n_v,n_p} - \boldsymbol{X}_{n_v,n_p} \right| / \boldsymbol{X}_{n_v,n_p} \right]$$

$$\text{s.t. } \widehat{\boldsymbol{X}}_{n_v,n_p} = \boldsymbol{\mathcal{P}}_{\boldsymbol{\Theta}_P}\left(\boldsymbol{\mathcal{J}}_{\boldsymbol{\Theta}_I}(\boldsymbol{Y}; \boldsymbol{\Omega})_{n_v,:} \right)_{n_p}, \tag{5.3}$$

where n_v and n_p denote the voxel and tissue parameter index, respectively. The weighting matrix $\boldsymbol{W_C}$ is not needed to be finetuned. This parameter loss replaces the image loss in Equation 5.2 for jointly training of TSC image reconstruction and parameter estimation in an end-to-end fashion, which may potentially improve quality of overall reconstruction.

In Variant 3, sampling trajectories $\boldsymbol{\Omega}$ of 3D MRF are jointly optimized with TSC image reconstruction, which has the potential to produce optimal sampling patterns and further improve acquisition efficiency (Figure 5.1(c)).

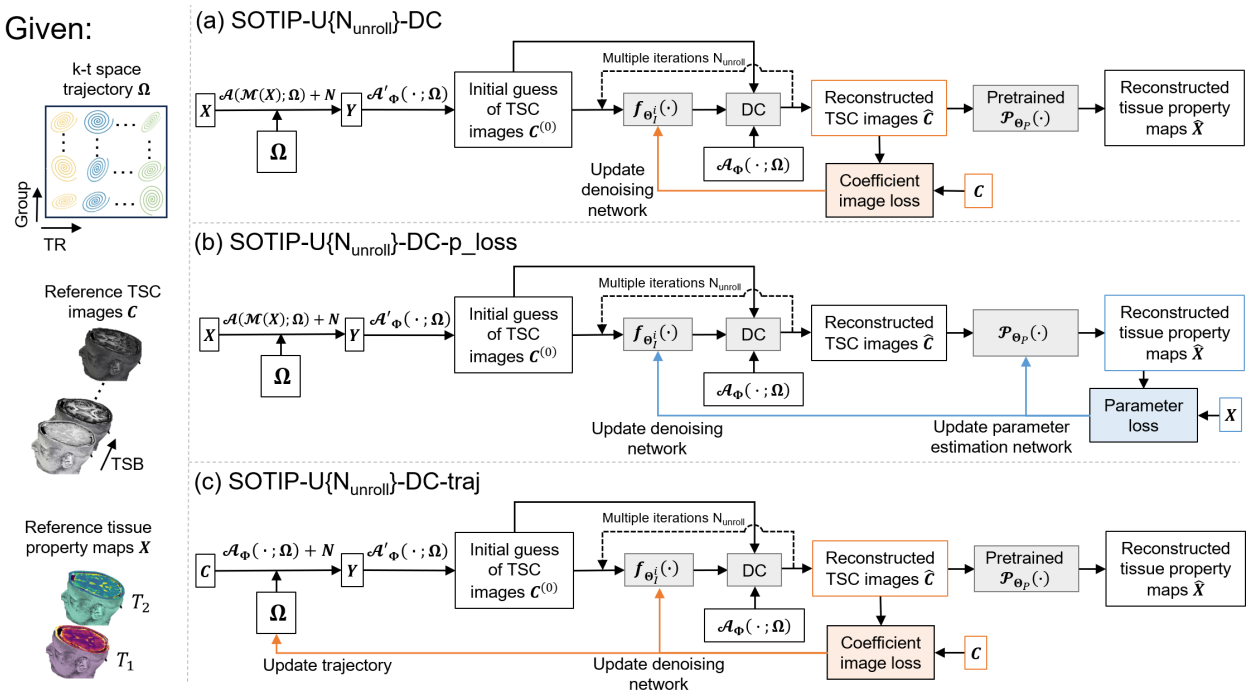

Figure 5.1: Flowcharts of three variants of a SOTIP framework. (a) SOTIP-U{N_{unroll}}-DC consists of N_{unroll} iterations of a unroll denoising network incorporated with data consistency of full 3D spiral k-space samples for temporal subspace coefficient (TSC) image reconstruction and a pretrained parameter estimation network for voxel-by-voxel parameter estimation. A l_2 image loss between reference/ground truth and reconstructed coefficient images was used to update the networks. (b) SOTIP-U{N_{unroll}}-DC-p_loss used a parameter loss that replaced the image loss for joint optimization of parameter quantification and image reconstruction. (c) SOTIP-U{N_{unroll}}-DC-traj simultaneously updated 3D sampling trajectories and coefficient image reconstruction networks using the image loss. In the training phase, the k-space data was simulated on-the-fly. The input reference TSC images were first undersampled in k-t space and then were added with complex Gaussian noise. Adjoint NUFFT was applied to the k-t space data to generate an initial guess of TSC images that input to the denoising networks.

5.2.2 MRF sequence selection and acquisition parameters

The MRF sequence was based on fast imaging with steady state precession (FISP) [127]. The sequence started with an adiabatic inversion pulse and was followed by 500 TRs (TI/TE/TR=21/0.63/9.2) with varying flip angles from 0 to 67.4 degrees, which was termed as one acquisition group. The k-space sampling trajectories were varied from TR to TR and across multiple acquisition groups. A rest time of 3s was applied between acquisition groups to allow spins to relax to equilibrium before the next acquisition group. Forty-eight acquisition groups were regarded as full sampling. An undersampling rate (R) was defined by 48 divided by the

number of acquisition groups, which followed the definition in [8]. For example, R=6 and R=3 refer to acquisition groups of 8=48/6 and 16=48/3, respectively, which led to respective total scan time of 1 min and 2 min.

5.2.3 3D spiral trajectory optimization

In Variant 3, we optimized orientations of 2D spiral readout planes in the 3D k-space during the training of the image reconstruction network to fully exploit the spatiotemporal sparsity of MRF. The 2D base trajectory is a variable density spiral trajectory [128] with FOV=230×230×230mm^3, 1mm-isotropic resolution, 29-fold in-plane undersampling rate, and variable density factor of 1.3. The maximum gradient amplitude was set as 0.04T/m and maximum slew rate as 150T/m/s. To obtain a better initialization, the trajectory was initialized using a TGAS scheme implemented in [8]. The trajectories were optimized by rotating the initial spiral planes around k_x, k_y, and k_z axes with learnable rotation angles $\boldsymbol{\delta}(i,j) = \left[\delta_x(i,j), \delta_y(i,j), \delta_z(i,j)\right]^T$, where i is the index of acquisition groups and j is the TR index.

5.2.4 MBDL TSC image reconstruction

The TSC image reconstruction can be posed as an optimization problem in a form:

$$\widehat{\boldsymbol{C}} = \underset{\boldsymbol{C}}{\operatorname{argmin}} \left\| \boldsymbol{D}^{1/2}(\boldsymbol{\mathcal{A}}_{\boldsymbol{\Phi}}(\boldsymbol{C};\boldsymbol{\Omega}) - \boldsymbol{Y}) \right\|_F^2 + \lambda \mathcal{R}(\boldsymbol{C}), \tag{5.4}$$

where $\boldsymbol{\mathcal{A}}_{\boldsymbol{\Phi}}(\cdot\,;\boldsymbol{\Omega}) = \boldsymbol{F}_{NU}^{\boldsymbol{\Omega}}\boldsymbol{S}\cdot\boldsymbol{\Phi}$, $\boldsymbol{S} \in \mathbb{C}^{N_c N_v \times N_v}$ denotes coil sensitivity maps, $\boldsymbol{F}_{NU}^{\boldsymbol{\Omega}} \in \mathbb{C}^{N_c N_k T \times N_c N_v}$ represents nonuniform discrete Fourier transform (NUDFT) operator with sampling pattern $\boldsymbol{\Omega}$, which is approximated by nonuniform fast Fourier transform (NUFFT), $\boldsymbol{D} \in \mathbb{R}^{N_c N_k T \times N_c N_k T}$ is a diagonal matrix containing the density compensation weights [129] for $\boldsymbol{\Omega}$, and $\mathcal{R}(\cdot)$ is the regularization with a weighting parameter of λ. The first term encourages DC between the

reconstructed TSC images and the acquired k-t-space data. The second term incorporates prior knowledge of the TSC images. deep learning-based data-driven regularization can be used to iteratively solve the optimization problem by:

$$\boldsymbol{Z}^{(i+1)} = \boldsymbol{f}_{\boldsymbol{\Theta}_I^{(i+1)}}\left(\boldsymbol{C}^{(i)}\right), \tag{5.5}$$

$$\boldsymbol{C}^{(i+1)} = \underset{\boldsymbol{C}}{\operatorname{argmin}} \left\|\boldsymbol{D}^{1/2}(\boldsymbol{\mathcal{A}}_{\boldsymbol{\Phi}}(\boldsymbol{C};\boldsymbol{\Omega}) - \boldsymbol{Y})\right\|_F^2 + \lambda\left\|\boldsymbol{C} - \boldsymbol{Z}^{(i+1)}\right\|_F^2, \tag{5.6}$$

where $\boldsymbol{f}_{\boldsymbol{\Theta}_I^{(i+1)}}: \mathbb{C}^{N_v \times K} \rightarrow \mathbb{C}^{N_v \times K}$ is a denoising/de-aliasing network for the (i+1)$^{\text{th}}$ iteration parameterized by $\boldsymbol{\Theta}_I^{(i+1)}$. The second step is often called a DC step, encouraging consistency of acquired k-t-space data and the reconstruction. We use one gradient decent step to approximate this step as

$$\boldsymbol{C}^{(i+1)} = \boldsymbol{Z}^{(i+1)} - \alpha\boldsymbol{\mathcal{A}}'_{\boldsymbol{\Phi}}\left(\boldsymbol{D}\left(\boldsymbol{\mathcal{A}}_{\boldsymbol{\Phi}}\left(\boldsymbol{Z}^{(i+1)};\boldsymbol{\Omega}\right) - \boldsymbol{Y}\right);\boldsymbol{\Omega}\right), \tag{5.7}$$

where α is a learnable step size of the DC update.

5.2.5 CNN network structure

We used a 3-level 2D U-Net structure for a denoising/de-aliasing network $\boldsymbol{\mathcal{D}}_{\boldsymbol{\Theta}}$ to process batched 2D axial slices of the 3D TSC images. The 2D networks could be trained by a large number of 2D slices to alleviate the need for a large dataset of high-quality 3D MRF image volumes. To account for intensity scale differences across TSC images corresponding to different TSBs, each TSC image input to the network was standardized by subtracting its mean and divided by its standard deviation. Then, the mean and standard deviation were saved to restore the intensities of each image output for the subsequent DC step. No residual connection was used, and weights were not shared across unrolls. Group norm [130] with 64 groups was added before each rectified linear unit (ReLU) non-linearity was applied after all except for the last layer.

5.2.6 Parameter estimation network structure

A three-layer fully connected neural network $\boldsymbol{\mathcal{P}}_{\boldsymbol{\Theta}_P}(\cdot)$ is used to estimate the parameters voxel-by-voxel from the TSC images $\widehat{\boldsymbol{C}}$, K=5 and $N_p = 2$ for T_1 and T_2. $\widehat{\boldsymbol{C}}$ is first normalized by its l_2 norm in the TSB dimension to eliminate the effect of proton density and other proportionality factors in the MR signals. This normalization was not used when the fully connected network was jointly trained with the TSC image reconstruction network to increase training stability. The real and imagery parts of the TSC images were concatenated so the input of the network has 10 channels. The three layers have 512, 512, and 2 nodes, respectively, where the last layer provides T_1 and T_2 estimates. ReLU nonlinearity was applied to the first two layers. The target T_1 and T_2 values were normalized by their maximum physiological values in the brain, which in this work, we used 4000ms and 1500ms, respectively. This normalization keeps the output range in [0, 1] which facilitates network training.

5.2.7 Memory- and time-efficient strategies for network training

Additional reductions of memory footprint and training time were achieved by leveraging automatic mixed precision training (AMP) [131], gradient checkpointing (GC) [132], spatial-temporal commutativity of the MR system operator and different oversampling factors and kernel sizes of NUFFT. We utilized different combinations of these techniques to allow training the network on a single GPU while maximizing the memory usage for time efficient training. The AMP training automatically assigns appropriate data type to different operations to reduce computation cost including time and memory. For example, convolution computation is much faster for data in half-precision than in default single-precision but without sacrifice of accuracy. Operations such as exponentials and summations still use data in single precision to prevent numeric overflow or underflow.

GC is a method to tradeoff memory footprint with a small increase in computation cost. GC keeps a subset of intermediate activations of an operator on memory, instead of saving all in traditional backpropagation, and recalculates others during gradient calculation. To do so, in the forward pass, 3D image volumes were decomposed into batches of 10 2D slices. Each batch was fed into the CNN denoiser $\boldsymbol{\mathcal{D}_{\theta}}$ without saving the intermediate activations, and then the output 2D slices were re-combined into 3D image volumes. In the backward pass, the intermediate activations were recalculated for each batch for gradient calculation (Figure 5.2). For more GPU memory intensive training, depending on specific dataset and reconstruction network, GC was applied to the NUFFT operator, in which the image data were decomposed in the TSB dimension and/or in the coil dimension (Figure 5.3). Note that GC only increases the computation cost during network training but not in the inference.

By using the spatial-temporal commutativity of the MR system operator [133], [134], the adjoint operator $\boldsymbol{\mathcal{A}'_{\Phi}}(\cdot\,;\boldsymbol{\Omega})$ can be implemented by first projecting the k-t space data onto a low dimensional temporal subspace by multiplying $\boldsymbol{\Phi}'$ and then applying adjoint NUFFT operator. This keeps the operations in K dimensions of the number of the significant temporal bases instead of *T* dimensions in the time domain, greatly reducing computation time and memory consumption.

To further reduce memory footprint, we investigated effects of the oversampling factor and kernel size of the NUFTT operator on MBDL. These two parameters control accuracy of the NUFFT operator but also affect memory footprint and computation time.

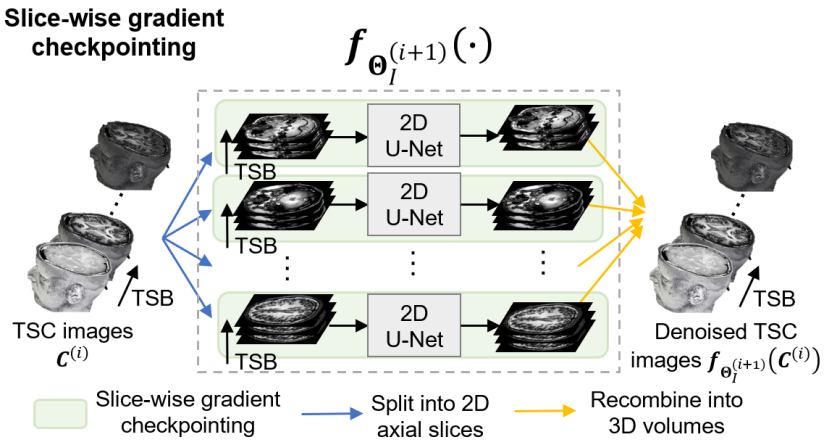

Figure 5.2: A schematic of decomposition of 3D TSC image volumes to batches of 2D axial slices for memory-efficient gradient calculation using gradient checkpointing. Outputs of the 2D slice batches are recombined into 3D volumes.

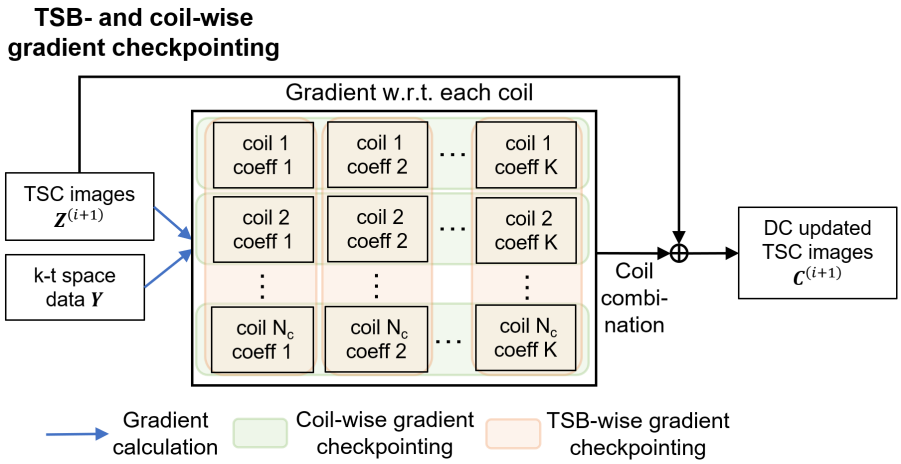

Figure 5.3: A schematic of decompositions of TSC and multi-coil image data along TSB and/or coil dimensions for memory-efficient gradient calculation required by the DC layer using gradient checkpointing.

5.2.8 Dataset

5.2.8.1 Simulation dataset

Brain T_1 and T_2 maps of 12 subjects were generated from Kirby21 public dataset [135]. T_1 mapping was conducted using a variable flip-angle spoiled 3D gradient echo sequence with TR/TE/α_1/α_2=100ms/15ms/15°/60°, FOV=212×212×143mm, and acquired

resolution=1.5×1.5×1.5mm. T_2 mapping was simulated using a 2D multi-slice spin-echo sequence to acquire dual echoes with $TR/TE_1/TE_2$=6653ms/30ms/80ms, FOV=212×212×143 mm^3 and resolution=1.5×1.5×1.5 mm^3. Both T_1 and T_2 maps were reformatted to 1mm-isotropic resolution with a field of view of 230×230×230mm^3. The MRF signals were simulated voxel-by-voxel using the extended phase graph (EPG) formalism [136] with the FISP-based MRF sequence as described in section 2.2. Coil sensitivity maps from eight virtual coils estimated from one subject in a public 3D MRF raw dataset [109], which will be described in detail in section 2.8.2, were used for simulation. The multi-coil image-time series were then retrospectively undersampled in the k-t-space using the TGAS trajectory with 48 acquisition groups. To obtain k-t-space data with high accuracy, NUFFT with oversampling factor = 2 and kernel size = 6 was used. Complex Gaussian noise was added to the simulated k-t-space data so that the SNR in the image space was 50 with respect to the mean MRF signal across time and whole brain.

5.2.8.2 In vivo dataset

Public *in vivo* brain 3D MRF raw data of 14 healthy volunteers and three patients [109] were used for training and testing of the reconstruction framework. The healthy volunteer data were acquired on a 3T Premier MRI scanner (GE Healthcare, Waukesha, WI) with 48 head receiver-coils. The patient data were acquired on two different 3T Signa Premier scanners. The MRF sequence consists of an adiabatic inversion pulse followed by a 500 TRs (TI/TE/TR = 20/0.7/12 ms) with variable flip angle from 10 to 75 degrees. Forty-eight acquisition groups with the TGAS trajectory were acquired from healthy volunteers with 6 min acquisition time. Patient data were prospectively undersampled to obtain 16 acquisition groups within 2 min and with R=3 (which was 50% TGAS undersampling of the first 32 groups and omitted the latter 16

groups, rather than 33% TGAS undersampling of 48 groups [8]). The fully acquired resolution was 1mm-isotropic and the FOV was 220×220×220 mm^3.

5.2.9 Evaluation and comparison

5.2.9.1 Simulation study

We tested SOTIP on simulation data with ground truth parameters for error calculation. A MRF dictionary containing 23538 time curves with 500 TRs in each was generated using $T_1 \in [20{:}\,20{:}\,3000] \cup [3200{:}\,200{:}\,5000]\ ms$ and $T_2 \in [10{:}\,2{:}\,200] \cup [220{:}\,20{:}\,1000] \cup [1050{:}\,50{:}\,2000] \cup [2100{:}\,100{:}\,4000]\ ms$. The dictionary was compressed to the first five principal components, $\mathbf{\Phi} \in \mathbb{C}^{5\times 500}$, by principal component analysis (PCA) to pretrain the parameter estimation network. 50%/20%/30% of the entries of the compressed dictionary were used for training/validation/testing the parameter quantification network. For all experiments involving the TSC image reconstruction networks, 1150/460/1150 2D axial slices from 5/2/5 subject heads were for training/validation/testing, respectively.

Ablation Study We conducted series of investigations of effects of 1) memory efficient training and reconstruction techniques including AMP, slice-/TSC-/coil-wise GC, 2) NUFFT with different oversampling factors and kernel sizes, 3) DC layer, and 4) number of unrolls. First, we compared different memory efficient techniques regarding training time and GPU memory by training SOTIP with 1 unroll (SOTIP-U1) and SOTIP with 1 unroll plus DC (SOTIP-U1-DC) using TGAS simulation data with $R = 6$ and using the NUFFT with oversampling factor of 1.25 and kernel size of 2. Secondly, we trained SOTIP-U1-DC on the same data using the NUFFT with {oversampling factor}×{kernel size}={1.125, 1.25, 1.5}×{2} and {1.25}×{2, 3, 4} and without TSC- or coil-wise GC. Finally, SOTIP networks with $\{1, 2, 3\}$ unrolls without DC layer (SOTIP-U$\{1, 2, 3\}$) and with $\{2, 3, 4\}$ unrolls with DC layers (SOTIP-

U$\{2, 3, 4\}$-DC) were trained by the TGAS simulation data ($R \in \{3, 6\}$). In all these experiments, no joint optimization was used, and the parameter quantification network was pre-trained (Variant 1 in Figure 5.1(a)).

Joint optimization with parameter quantification TSC Image reconstruction and parameter quantification were jointly optimized using the parameter loss (SOTIP-U3-DC-p_loss) on the TGAS simulation data ($R \in \{3, 6\}$) (Variant 2 in Figure 5.1(b)).

k-t trajectory optimization SOTIP networks were jointly trained for the learnable spiral trajectory rotation (SOTIP-U3-DC-traj) on the TSC simulation data ($R \in \{3, 6\}$), in which the parameter quantification network was pre-trained and fixed (Variant 3 in Figure 5.1(c)).

Comparison with LLR reconstruction We compared performances of SOTIP networks trained in previously described experiments with LLR reconstruction [8] on the TGAS simulation data ($R \in \{3, 6\}$). The parameter estimation accuracy and computation time were evaluated. We also utilized the spatiotemporal commutativity of $\mathbf{\Phi}$ and $\boldsymbol{F}_{NU}$ to improve efficiency of the LLR reconstruction. The LLR algorithm required the block size to be a factor of the image matrix size and found a block size of 8 optimal for the TGAS MRF reconstruction [8]. In our work, a different matrix size was used so that block sizes of 5 and 10 were tested since they were close to the optimal value of 8 reported previously. The weighting parameter of the LLR regularization was finetuned within $\{1 \times 10^{-6}, 5 \times 10^{-6}, 1 \times 10^{-5}, 5 \times 10^{-5}, 1 \times 10^{-4}\}$ for $R \in \{3, 6\}$, respectively. The LLR reconstruction was solved by Fast Iterative Shrinkage-Thresholding Algorithm (FISTA) for 100 iterations and implemented by SigPy package [137] with GPU acceleration.

5.2.9.2 in vivo study

We validated the proposed reconstruction framework on *in vivo* data. SOTIP was re-trained on the *in vivo* data with different sequence parameters than the simulation data. We trained two SOTIP-U3-DC networks on retrospectively undersampled healthy volunteer scans with $R \in \{3, 6\}$ following the TGAS scheme [8]. Coil sensitivity maps for 10 compressed coils were estimated using JSENSE [138]. Training targets of the SOTIP networks were the LLR reconstructed TSC images with matrix size = 256 × 256 × 256 and TSB=5 from the 6 min scan. The T_1 and T_2 maps were estimated using pattern matching. We used 1536/512/1536 2D axial slices from 6/2/6 healthy subjects for training/validation/testing, respectively. The SOTIP reconstructions from the 2 min (R=3) and 1 min (R=6) scans were compared with the LLR reconstructions from the 6 min (R=1), 2 min (R=3) and 1 min (R=6) scans. Prospectively undersampled 2 min scans of the 3 patients were also used to compare between SOTIP and LLR reconstructions. To ensure fairness, the public available source code and MRF dictionary provided by Iyer et. al. [109] were used for the LLR reconstruction and the pattern matching.

5.2.9.3 Evaluation Metrics

We evaluated the parameter quantification accuracy by NRMSE that is defined as $\sqrt{MSE(\boldsymbol{X}_i, \widehat{\boldsymbol{X}}_i)}/\left(\max(\boldsymbol{X}_i) - \min\left(\widehat{\boldsymbol{X}}_i\right)\right) \times 100\%$, where $\boldsymbol{X}_i$ and $\widehat{\boldsymbol{X}}_i$ stand for the ground truth and estimate of the *ith* parameter, respectively, and $MSE(\cdot,\ \cdot)$ represents the mean squared error.

5.2.10 Implementation details

The weighting matrix $\boldsymbol{W_C}$ on different bases in the TSC image loss was empirically chosen to be the reciprocal of the Frobenius norm of each reference TSC image so that different TSC images were on a similar scale to facilitate network training. SOTIP was implemented in

PyTorch with MIRTorch [139] and TorchKbNufft [140] packages on an NVIDIA A100 GPU with 40GB memory. Adam optimizer [61] was used with 1×10^{-4} learning rate which was reduced by a half when the validation error was not improved in 20 consecutive epochs. We used batch size = 1 and terminated training when the validation error was not improved in 40 consecutive epochs. GPU memory footprint was measured by peak reserved GPU memory. Paired t-test was used for all statistical comparisons between the reconstructions. A difference was considered as significant when $p<0.05$.

5.3 Results

5.3.1 Simulation study

5.3.1.1 Ablation Study

The effects of different efficient training techniques on GPU memory footprint and computation time in training SOTIP-U1 and SOTIP-U1-DC on the TGAS simulation data with R=6 were investigated (Table 5.1). The network even without the DC layer was unable to fit onto the GPU without efficient training techniques. Using GC in the slice dimension enabled the network to fit on the GPU, and adding AMP further cut down the GPU memory usage. However, incorporating a DC layer in the network increased the GPU memory by nearly 11 GB. To further reduce the GPU memory usage, GC was applied to the TSC and coil dimensions but with a cost of training time.

Table 5.1: Effects of efficient training techniques on GPU memory and training time per epoch

	w/o DC			w/ DC			
Technique	None	Slice-wise GC	Slice-wise GC +AMP	Slice-wise GC +AMP	Slice-wise GC +AMP +TSC-wise GC	Slice-wise GC +AMP +coil-wise GC	Slice-wise GC +AMP +coil-wise GC +TSC-wise GC
GPU memory (GB)	>40 (O.O.M.)	10.7	8.4	19.7	18.7	17.5	12.9

Training time per epoch (min)	-	0.4	0.4	1.6	1.9	2.5	1.7

O.O.M., out of memory.

The effects of over-sampling and interpolation kernel size of NUFFT on validation NRMSEs of T_1 and T_2, memory footprint and computation time of SOTIP-U1-DC using the TGAS simulated data with R=6 are summarized in Table 5.2. All hyperparameters of NUFFT showed comparable NRMSEs, but a large kernel size increased training and inference times while a high oversampling factor increased GPU memory footprint. We selected oversampling factor of 1.25 and kernel size of 2 for small memory usage and short computation time in the following experiments.

Table 5.2: Effects of NUFFT oversampling factor and kernel size on NRMSEs of validation T1 and T2, training and testing memory and time.

Over-sampling factor	Kernel size	T_1 NRMSE (%)	T_2 NRMSE (%)	Training memory (GB)	Testing memory (GB)	Training time per epoch (min)	Testing time per subject (s)
1.25	2	8.2	3.8	19.7	10.9	1.5	7.6
1.25	3	7.8	3.8	19.7	10.9	1.6	8.2
1.25	4	7.9	3.7	19.7	10.9	2.4	11.8
1.125	2	8.3	3.9	18.9	10.9	1.5	7.4
1.5	2	8.5	4.0	21.0	11.1	1.5	7.2

NRMSEs of T_1 and T_2 values quantified by SOTIP with different numbers of unrolls and with or without DC layers are shown in Figure 5.4. DC significantly improved both T_1 and T_2 NRMSEs ($p<0.005$) when the number of unrolls was greater than 1 for the reconstructions using either 1-min or 2-min scans. Increasing the number of unrolls from one to two improved the NRMSEs regardless of using a DC layer, but the improvement was greater with DC. However, the improvement was diminished after three unrolls for both 2 min (R=3) and 1 min (R=6) scans. Considering increases in training time and memory with the number of unrolls (Figure 5.5), 3 unrolls were used in the following experiments.

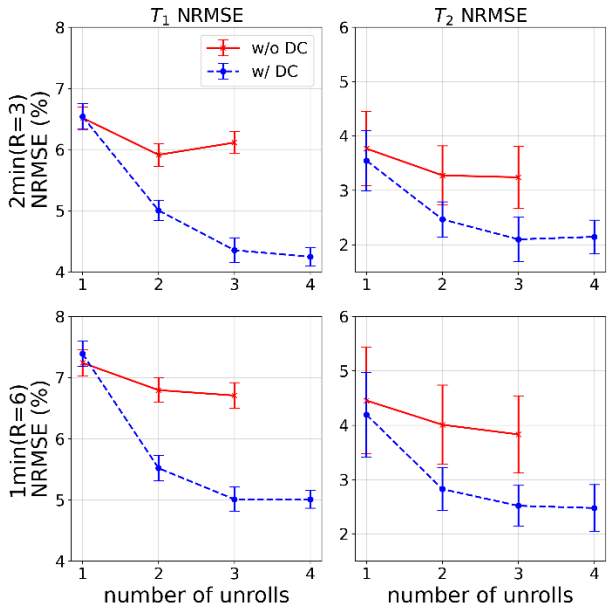

Figure 5.4: NRMSEs of T_1 (left column) and T_2 (right column) vs. the number of unrolls of SOTIP with or without DC for the 2 min (first row) and 1 min (second row) scans.

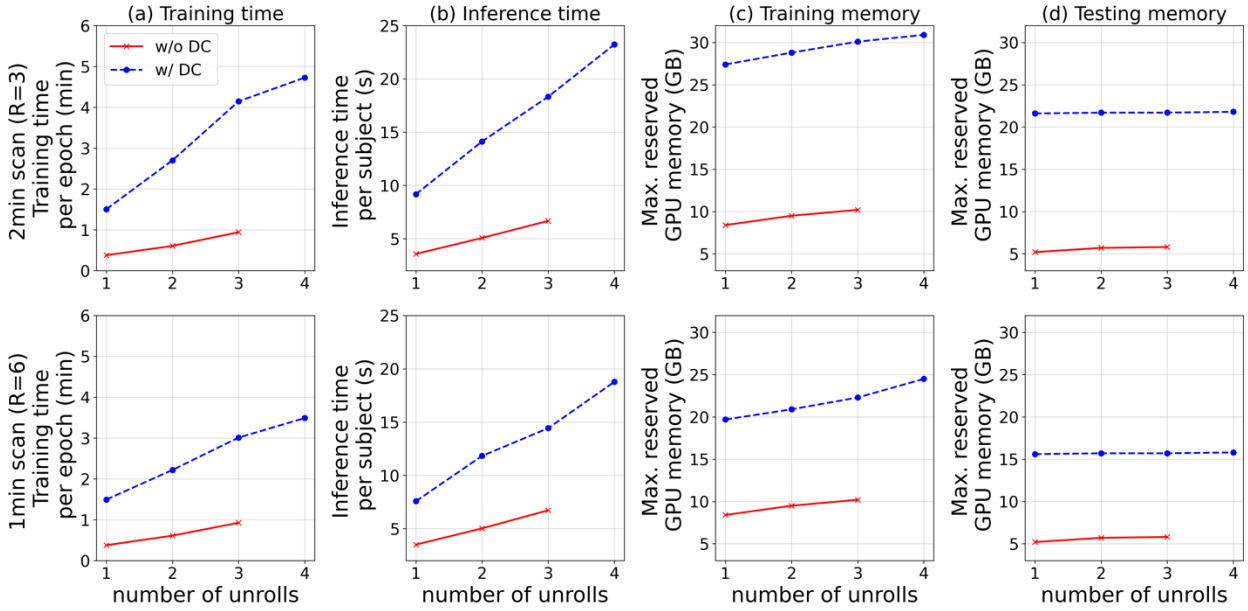

Figure 5.5: (a) Training time per epoch, (b) inference time per subject, (c) training and (d) testing GPU memory of SOTIP with respect to different numbers of unrolls and with or without DC. The training and inference times increased with the number of unrolls but more rapidly for SOTIP with DC than without DC. Although the training memory increased with the number of unrolls, the testing memory remained almost the same.

5.3.1.2 Performance of T_1 and T_2 mapping

T_1 and T_2 maps reconstructed using SOTIP-U3, SOTIP-U3-DC, SOTIP-U3-DC-p_loss, and SOTIP-U3-DC-traj were compared to the LLR reconstruction using the 2 min (R=3) and 1 min (R=6) simulated data. Examples of the reconstructed maps are shown in Figure 5.6 and Figure 5.7. Compared to the LLR reconstructions, SOTIP-U3 (without DC) did not improve the quality of T_1 and T_2 maps by qualitive review and quantitative assessment (Figure 5.8). However, adding DC layers in SOTIP-U3-DC sharply reduced percentage errors of T_1 and T_2 maps. Using the parameter loss (SOTIP-U3-DC-p_loss) improved accuracy of T_1 quantifications further, and jointly optimizing the k-t space sampling trajectory in SOTIP yielded the lowest percentage errors in the reconstructed maps. Quantitative analysis of the NRMSEs of T_1 and T_2 showed that SOTIP-U3-DC was significantly better than LLR ($p<0.005$) but SOTIP-U3 was not (Figure 5.8). Similar analyses showed that SOTIP-U3-DC-p_loss and SOTIP-U3-DC-traj significantly reduced the NRMSEs of T_1 and T_2 compared with SOTIP-U3-DC ($p<0.05$), except for T_2 quantification from the 2 min scan, suggesting effectiveness of direct minimization of parameter loss and joint optimization of k-t sampling trajectory in SOTIP. Inspection of input and output TSC image slices from LLR and SOTIP-U3-DC indicates that SOTIP-U3-DC reduced noise and errors in the output TSC images more effectively than LLR, see examples of the TSC images of the 2-min scan in Figure 5.9.

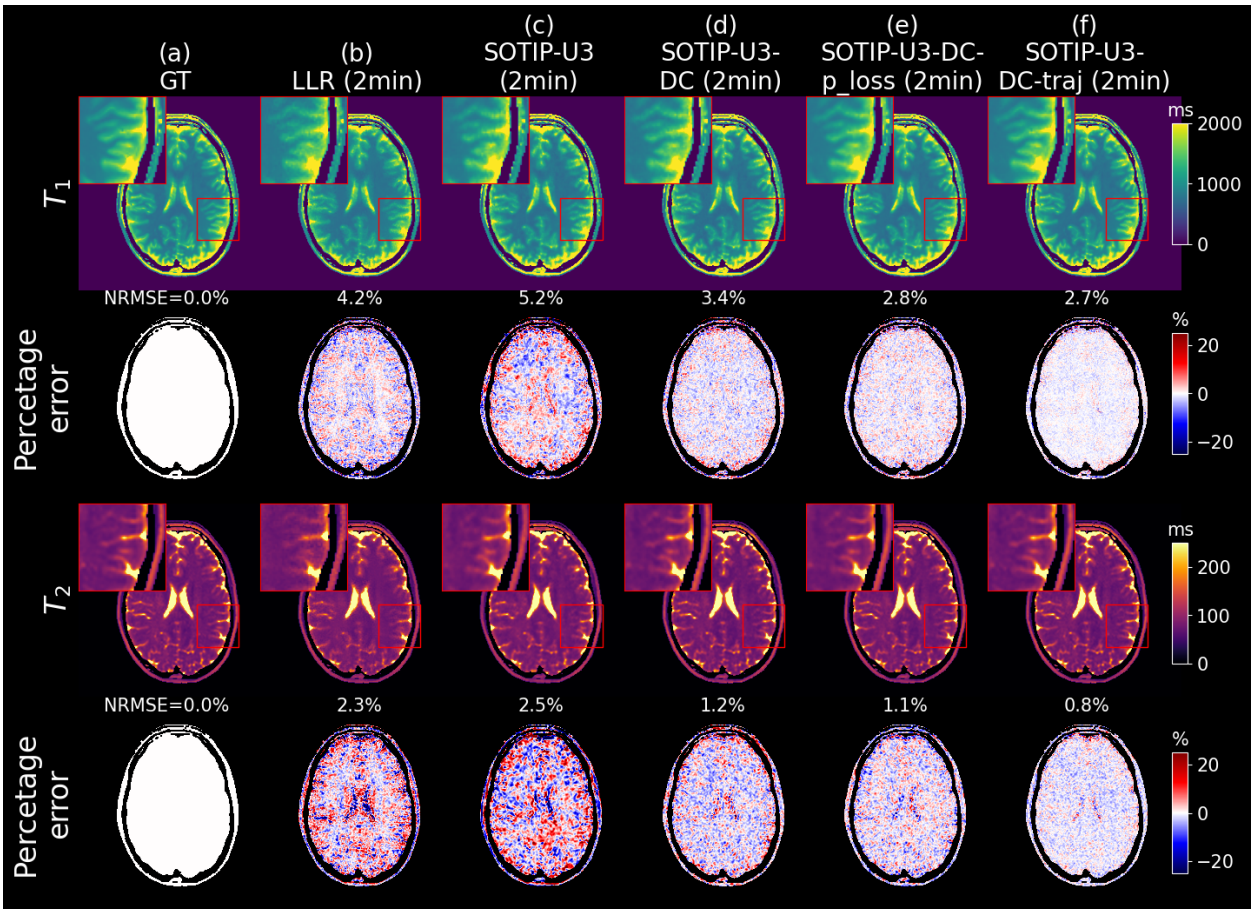

Figure 5.6: Example reconstruction slices of T_1 (first row) and T_2 (third row) and their corresponding percentage error maps (second and fourth rows) from the testing simulated data of 2-min scan by (b) the locally low rank method (LLR), (c) SOTIP with 3 unrolls (SOTIP-U3), (d) SOTIP with 3 unrolls and DC layers, (e) SOTIP-U3-DC with the parameter loss (SOTIP-U3-DC-p_loss), and (f) SOTIP-U3-DC with joint optimization of k-t space trajectory (SOTIP-U3-DC-traj). Ground truth (GT) T_1 and T_2 maps are shown in (a). Zoom-in plots of a cortical region show details of image sharpness and noise.

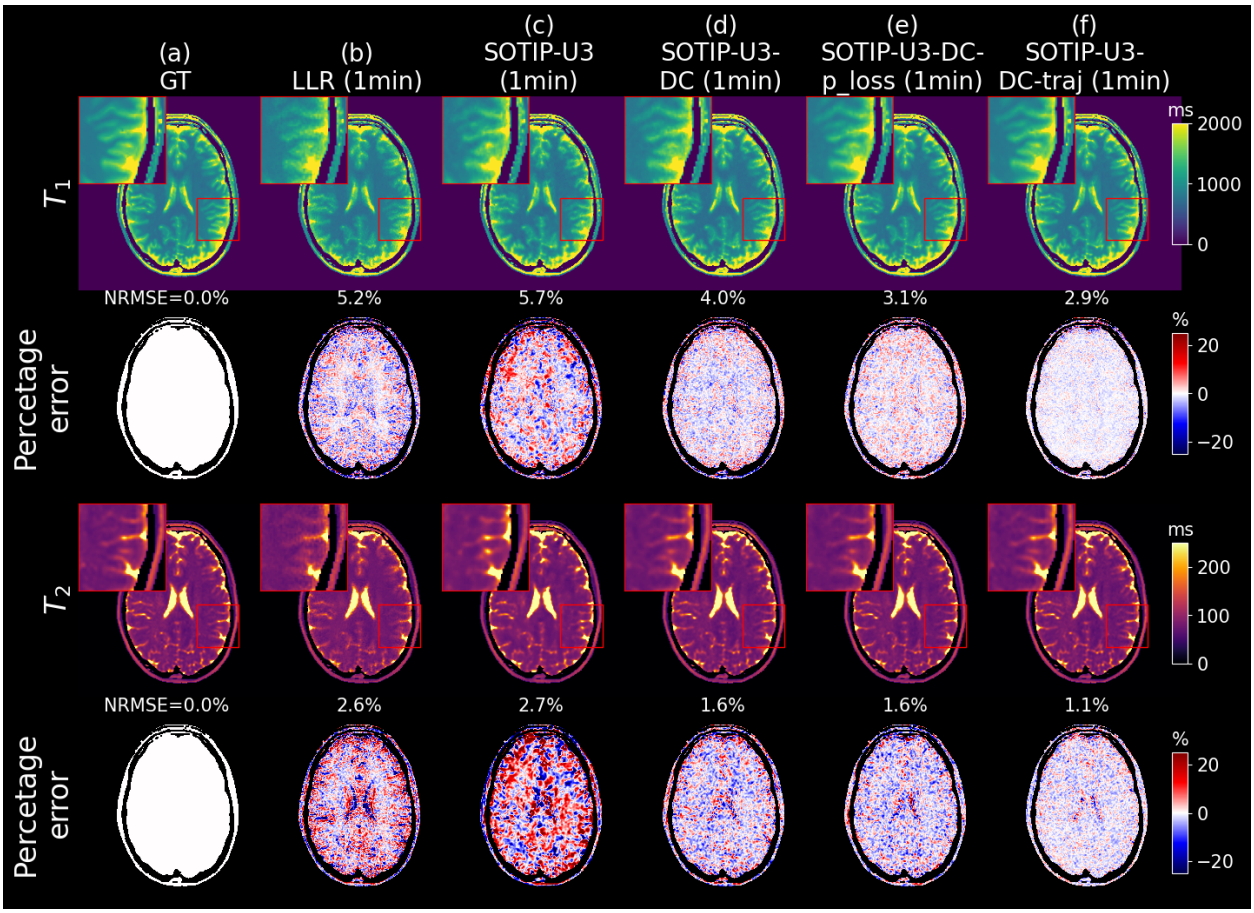

Figure 5.7: Example reconstruction slices of T_1 (first row) and T_2 (third row) and their corresponding percentage error maps (second and fourth rows) from the testing simulated data of 1-min scan by (b) the locally low rank method (LLR), (c) SOTIP with 3 unrolls (SOTIP-U3), (d) SOTIP with 3 unrolls and DC layers, (e) SOTIP-U3-DC with the parameter loss (SOTIP-U3-DC-p_loss), and (f) SOTIP-U3-DC with joint optimization of k-t space trajectory (SOTIP-U3-DC-traj). Ground truth (GT) T_1 and T_2 maps are shown in (a). Zoom-in plots of a cortical region show details of image sharpness and noise.

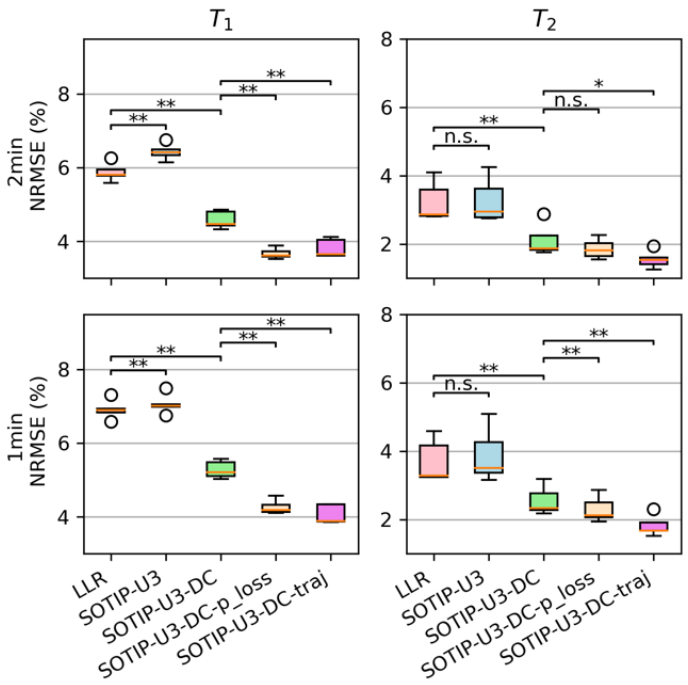

Figure 5.8: NRMSEs of quantified T_1 and T_2 by different reconstruction methods on all testing simulated data from 2-min (first row) and 1-min (second row) scans. n.s.: not significant. *: $p<0.05$. **: $p<0.005$.

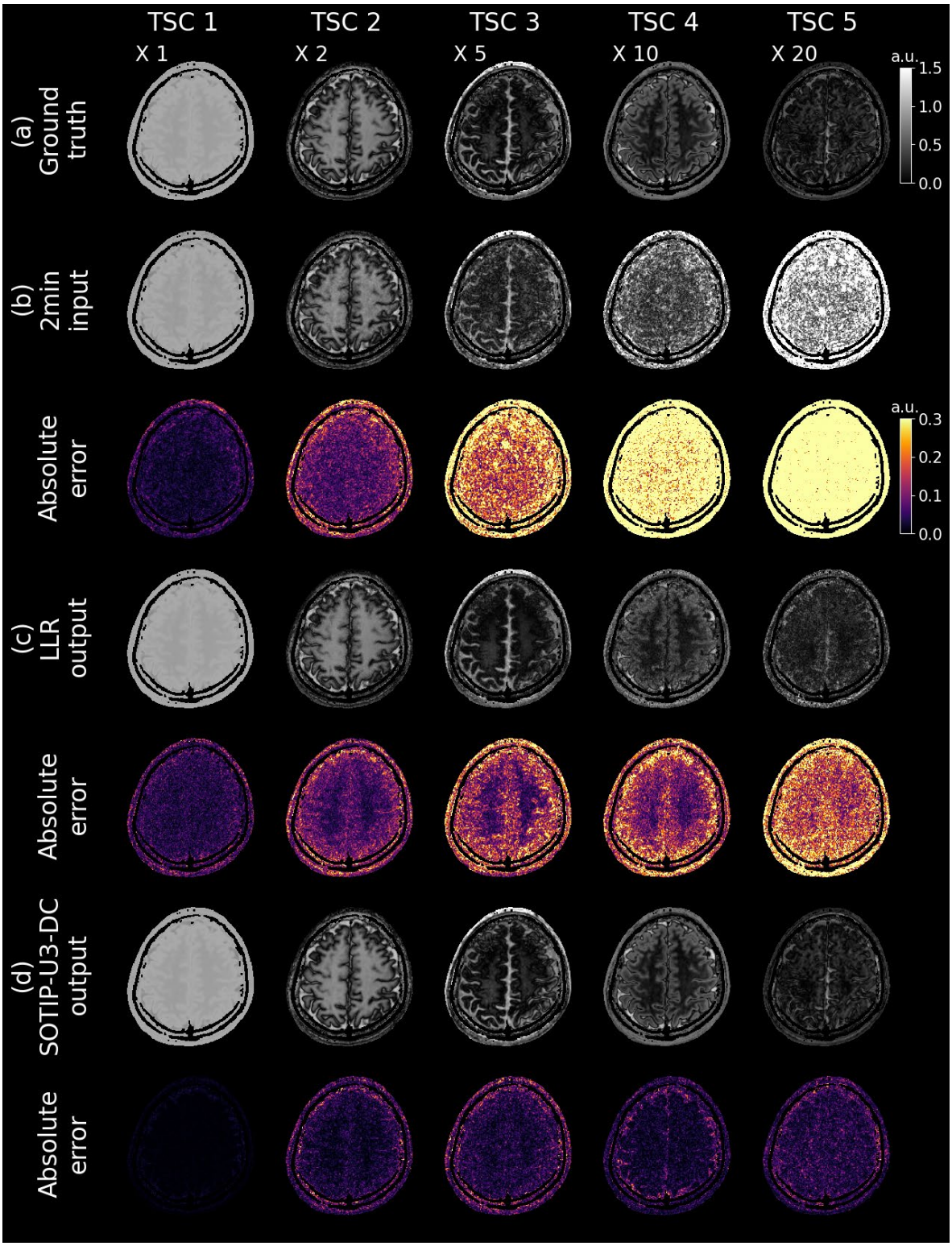

Figure 5.9: Example (a) ground truth, (b) input and output TSC image slices for (c) LLR and (d) SOTIP-U3-DC reconstructions for simulation data at 2-min scan time. The absolute error maps with respect to the ground truth for the input and output are included below each TSC image. The TSC images and their corresponding absolute error

are scaled differently across TSC 1-5 for better visualization. The TSC images are normalized across TSB dimension.

5.3.1.3 Reconstruction time comparison

T_1 and T_2 NRMSEs against wall time by different reconstruction methods of one testing case are plotted in Figure 5.10. While LLR reconstruction took ~10^3 s to converge, SOTIP-U3-DC-p_loss and SOTIP-U3-DC-traj only required ~20s to reconstruct 3D TSC image volumes of one subject, which indicates 50-fold reduction of reconstruction time by SOTIP.

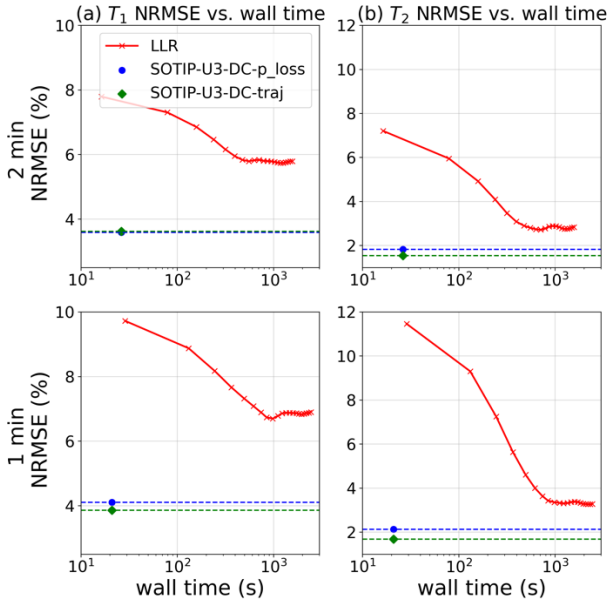

Figure 5.10: (a) T_1 and (b) T_2 NRMSEs of the reconstruction for the 2-min (first row) and 1-min (second row) scans from one testing case vs. wall time. The LLR wall time was plotted for every 5 iterations.

5.3.1.4 Optimized k-t trajectory

Point spread functions (PSFs) and line profiles of the SOTIP optimized trajectory were compared to the ones of the TGAS sampling scheme (Figure 5.11). The SOTIP optimized trajectory had less streaking artifact in the PSF and a narrower central lobe and lower side lobes in the line profile than those in TGAS sampling. The streaking artifact in PSF of TGAS

trajectory is likely caused by the limited number and distributions of different through-plane rotation angles of the trajectory that forms radial lines in certain projections of 3D k-space.

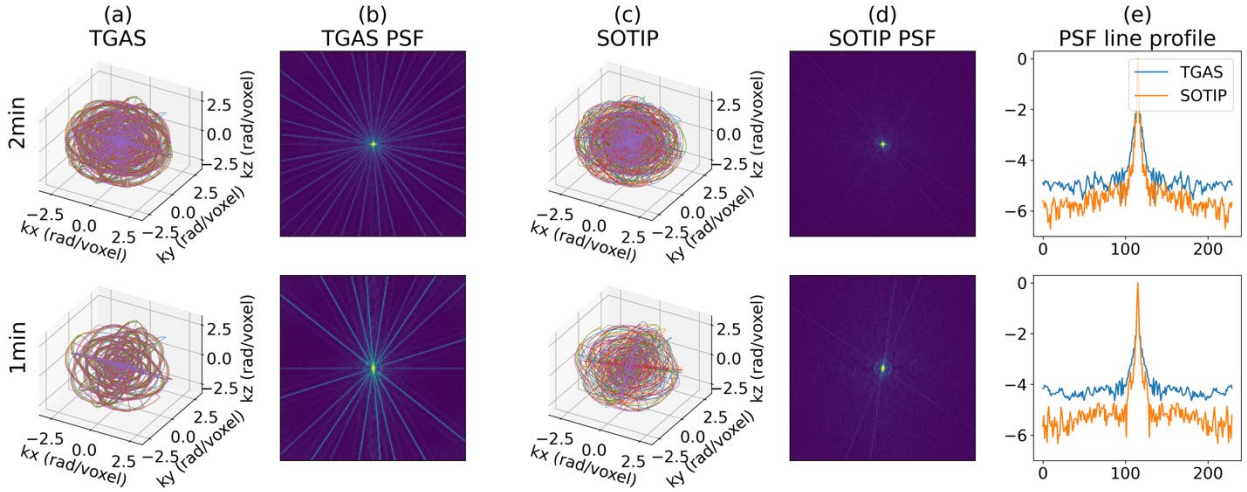

Figure 5.11: (a) TGAS samplings and (c) SOTIP optimized trajectories of the first 5 TRs, (b) and (d) their respective point spread functions (PSFs) averaged across k_x-k_y, k_x-k_z, and k_y-k_z planes, and (e) line profiles (log scale) of the PSFs averaged across x-, y-, and z-axes. Results from the 2-min and 1-min scans are shown in top and bottom two panels, respectively.

5.3.2 In vivo study

The SOTIP-U3-DC was trained and tested to reconstruct T_1 and T_2 maps using the data acquired in 2 min and 1 min in healthy subjects and patients, which were compared to the LLR constructions from the data acquired in 6 min (as reference), 2 min and 1 min. Compared to the LLR reconstructions from the 2- and 1-min scans of the healthy subjects, SOTIP reduced noise in the T_1 and T_2 maps (Figure 5.12), which is consistent with the results on the simulation data. The quantitative results showed that SOTIP-U3-DC significantly improved T_1 and T_2 NRMSEs for both 2-min ($p<0.05$) and 1-min ($p<0.005$) scans (Table 5.3). SOTIP-U3-DC also achieved 53- and 65-fold reductions in reconstruction time for the 2-min and 1-min scan, respectively. Finally, SOTIP-U3-DC was applied to reconstruct T_1 and T_2 maps from prospectively undersampled data of three patients, which again showed less noisy parameter maps and finer

structures than ones by LLR (Figure 5.13). Note that white matter abnormality of a patient was reconstructed successfully by both methods.

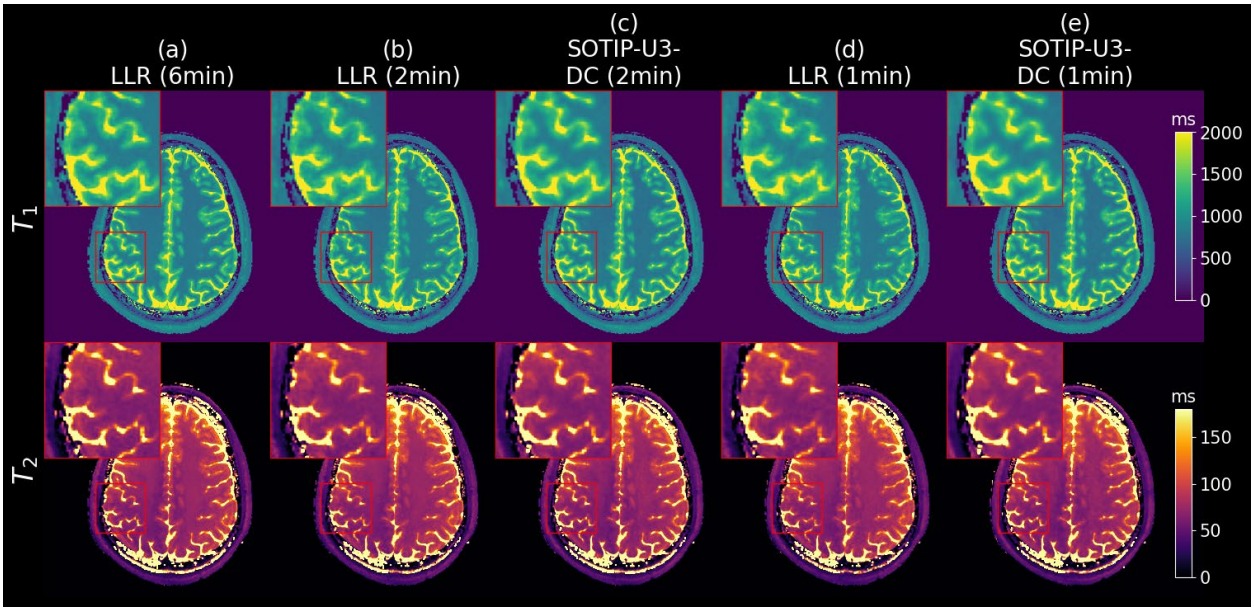

Figure 5.12: Comparison of LLR and SOTIP-U3-DC reconstructions of T_1 and T_2 of in vivo acquired 2-min and 1-min scans of one testing healthy subject. (a) LLR reconstruction of the 6-min scan as reference, (b) and (d) LLR reconstructions of respective 2-min and 1-min scans, (c) and (e) SOTIP-U3-DC reconstructions of respective 2-min and 1-min scans.

Table 5.3: NRMSEs and reconstruction times of *in vivo* data from 6 testing healthy subjects by LLR and SOTIP-U3-DC.

Scan time of data (min)	Method	T_1 NRMSE (%)	T_2 NRMSE (%)	Recon. time (s)
2	LLR	2.9 ± 0.1	5.9 ± 0.7	1600
	SOTIP-U3-DC	2.7 ± 0.1*	5.4 ± 0.4*	30
1	LLR	3.7 ± 0.2	7.2 ± 0.8	1300
	SOTIP-U3-DC	3.3 ± 0.1**	6.5 ± 0.5**	20

*: $p<0.05$. **: $p<0.005$

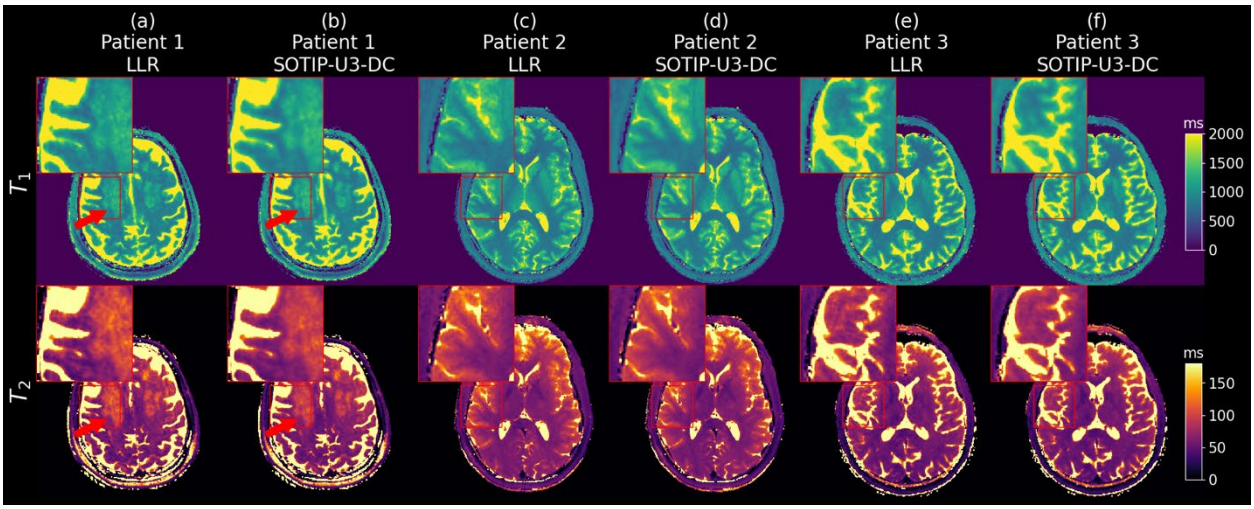

Figure 5.13: Comparison of example reconstruction slices by LLR and SOTIP-U3-DC for (a), (b) patient 1, (c), (d) patient 2, and (e), (f) patient 3. Red arrows indicate white matter abnormality of patient 1, which was recovered by both methods. Patient 1 (75 y/o male) had chronic small vessel disease while there was no significant finding in the other two patients (Patient 2, 28 y/o male and Patient 3, 49 y/o male).

5.4 Discussion

In this work, we developed a MBDL framework for quantification of T_1 and T_2 from full 3D spiral MRF to improve parameter quantification accuracy and shorten reconstruction time, evaluated it using both simulation and *in vivo* data and compared with the state-of-the-art LLR reconstruction. In this framework, we showed the benefits of synergetic optimization of the TSC image reconstruction with the parameter quantification and data-driven joint optimization of the k-t space sampling trajectory with the image reconstruction on realistic digital phantoms. The data-driven optimization of the k-t space samplings could further increase the acquisition efficiency of 3D spiral MRF. Our framework could be extended for other high-dimensional quantitative MRI acquisition and reconstruction.

One of the challenges of clinical adaptation of highly undersampled full 3D non-Cartesian scans is heavy computation burden and long reconstruction time caused by NUFFT and iterative reconstruction. MBDL holds the promise of reducing the number of iterations of

reconstruction by capturing a non-linear representation specific to MRI images in a data-driven fashion. However, the MBDL application to high-dimensional MRI reconstruction is hindered by large GPU memory requirement and/or long training time. Although future hardware development may address these issues, we demonstrated that the size and complexity of the problem can be largely reduced by utilizing the sparsity of the data and the memory footprint can be decreased by time- and memory- efficient training techniques to fit the network training onto a single modern GPU with a reasonable training time. We showed that the downgraded NUFFT operator, which reduces computation time and memory requirement, achieved a similar performance in the parameter estimation to a default optimal operator. We conjecture that the nonlinear representation learned by the neural network may compensate for the error in the data consistency step caused by the downgraded NUFFT operator. We also utilized the sequence-specific sparsity in the MRF data by temporal subspace modeling, and spatial-temporal commutativity to solve the reconstruction problem entirely in a low dimensional subspace. Mixed data precision training and GC further contained time and memory requirements. Other efficient network structures (e.g., deep equilibrium model [141], [142]) can be combined with efficient training techniques to allow further improvement in performance and computation efficiency. More efficient operators approximating NUDFT can also be combined with GC, for example, pre-calculating Toeplitz kernels of a downgraded NUFFT operator [112], [143], GRAPPA Operator Gridding (GROG) [126], [144], and FINUFFT [145] in future works.

The full 3D non-Cartesian k-t trajectory has large degrees of freedom for optimization, which presents as a challenge as well. In this work, we optimized the rotation angles of 2D spiral planes in the 3D k-space. Our optimized 3D spiral trajectories provided more focused PSFs with lower side lobes, and improved parameter quantification accuracy compared with the empirically

designed TGAS trajectory. Further validation of the optimized trajectory needs to be carried out in *in vivo* experiments. Also, future work may consider parameterizing the density of each spiral or each readout by B-spline kernels [21], [112]. With a differentiable Bloch simulator [146], the sequence parameters may also be jointly optimized with the sampling trajectory and image reconstruction.

One common challenge for developing data-driven reconstruction methods for high-dimensional quantitative MRI is the data limitation including limited availability of high-quality data for training and lack of *in vivo* ground truth for performance validation. In this work, we generated realistic 3D+time digital phantoms with ground truth tissue parameters to test the proposed framework. Effects of B_0 and B_1 inhomogeneity and eddy current will be incorporated and investigated in future works. In our *in vivo* study, since there was no ground truth available, we used the LLR reconstruction of 6 min scans as a reference even though the reference itself could be subject to artifacts imposed by the regularization. We showed that T_1 and T_2 parameters produced by SOTIP were not inferior to those by LLR. The profiles of average NRMSEs from the simulation data were slightly different from the *in vivo* study, which may be partly attributed to the different biases in the ground truth/reference tissue parameters introduced by different sequences and hardware, and to different subject demography [147]. Considering lack of ground truth in the *in vivo* evaluation, SOTIP could be tested using task-specific metrics such as diagnostic accuracy or precision in the future.

Another concern is the generalizability of SOTIP trained on a small number of subjects. In this work, we did not observe overfitting during training and demonstrated excellent parameter quantification results on testing subjects. In addition, we showed that SOTIP could also generalize well on *in vivo* patient data acquired on different scanners and trajectories. This

generalization capability may result from the fact that 2D U-Nets were trained on 2D slices containing different anatomies along the axial dimension, which increases robustness of the network. In addition, there is clear spatial-sparsity or similarity across the five TSC images (Figure 5.9). The first two TSC images with the high-quality initial guess could behave like an individual prior to facilitate reconstructions of the third, fourth and fifth TSC images, which could reduce the requirement on the training data. Validations on scanners of different vendors, with different system imperfections, are warranted. Future works may also address the issue of limited dataset by learning anatomical population prior from static MRI images with different contrasts or resolutions which are more widely available and then combining it with unsupervised learning for robust patient-specific reconstruction. In addition, training objectives other than l_2 norm will be investigated to reduce possible spatial resolution degradation.

5.5 Conclusion

In conclusion, this chapter enabled model-based deep learning for full 3D spiral MRF on a single GPU by exploiting spatiotemporal redundancy and time and memory saving techniques. We demonstrated improvement in T_1 and T_2 quantification with shortened reconstruction time while keeping the same scan time and similar spatial resolution compared with a state-of-the-art LLR method on both simulation and *in vivo* data. Also, we demonstrated advantages of synergetic optimization of image reconstruction, parameter reconstruction and k-t space trajectory by improved NRMSEs of T_1 and T_2. *In vivo* implementation of the optimized trajectory and more comprehensive clinical validation on a larger patient cohort of the proposed framework will be carried out in the future.

Chapter 6 Discussion and Conclusion

This dissertation introduces several techniques for improving the efficiency of acquisition and computation of qMRI by data-driven optimization of k-space sampling patterns, image reconstruction, and tissue parameter estimation.

The high representation power of deep neural networks and large-scale optimization algorithms enable robust learning of sparsity in existing qMRI datasets which improves the accuracy and precision of quantified parameters from new data. In addition, neural networks are highly flexible and can be tailored to a variety of applications. Firstly, different forms of sparsity can be exploited by specific network architectures with dedicated modules and connections (Chapter 2, 3, and 4). Secondly, by parameterizing different parts of image processing (optimal acquisition, image reconstruction and parameter quantification), the whole pipeline can be jointly optimized (Chapter 3, 4, and 5). Thirdly, the final objective of a specific application can be directly optimized, reducing error propagation along the processing pipeline (Chapter 3 and 5). Since a more accurate representation is learned by deep learning, image reconstruction often requires fewer numbers of iterations than conventional iterative reconstruction, thereby reducing reconstruction time. Owing to the high flexibility of neural networks, the pursuit of the optimal robust representations of MR images and the best method to incorporate this knowledge into reconstruction continues to be a vibrant research area.

The generalizability of deep learning-based image reconstruction is a major concern in its clinical translation. Given the high degrees of freedom in the network optimization, the deep

learning reconstruction algorithm is sensitive to training and testing domain changes [148]. The problem is more challenging in qMRI with limited availability in high-quality data for training and validation. It is thus crucial to incorporate other prior knowledge to constrain the reconstruction. The incorporation of high-quality individual information, as fully sampled baseline images of DCE series in Chapter 3 and linear subspace modelling of MRF time signals in Chapter 5, alleviate the generalization problem. However, in the absence of such high-quality prior knowledge, it is possible to learn an individual representation from the acquired data using self-supervised learning and then find a balance in reconstruction among data fidelity, population prior knowledge and individual representation. It is also desirable to develop a fast and robust framework to quantify the bias and uncertainty of image reconstruction or parameter estimation. We have developed a such framework based on a generative model that provides tractable distributions for MRF [115]. The framework can then be used to detect domain changes and adapt image reconstruction. Online prediction of such bias and uncertainty may also inform active sampling of k-space during dynamic scans with guidance from both population prior and continuously acquired individual data.

One limitation of this dissertation is its retrospective nature. Prospective studies will be conducted in the future to further validate the proposed frameworks. The sampling trajectories or patterns developed in Chapter 3, 4, and 5 do not involve changes in sequence parameters (e.g., TR/TE) but only require changes in phase encoding steps and thus can be readily implemented by modifying the sampling schemes of existing MR sequences. The experiments will first be conducted on phantoms to identify possible domain shift between training dataset and real data based on the reconstruction results. The source of domain shift will be characterized, and several iterations of network finetuning may be required to achieve the best outcome. For example, MR

system imperfections including B_0 and B_1 inhomogeneities may be characterized and incorporated into the training dataset for network finetuning. After phantom studies, the MR sequences with modified sampling schemes will be incorporated into scan protocols for cancer patients to validate the proposed methods on in vivo data and assess any potential improvement in clinical utility.

Direct extensions of this dissertation include optimization of sequence parameters together with sampling patterns, image reconstruction, and parameter estimation to allow higher degrees of freedom for optimization. The proposed methods can also be applied to other qMRI applications including diffusion, susceptibility, and chemical exchange saturation transfer quantification.

Several open questions remain to be addressed before qMRI can achieve broader clinical application. How to reduce scan time while providing better or the same accuracy and precision of parameter quantification is one of them and is the focus of this dissertation. The lack of realistic, specialized phantoms designed for diverse qMRI techniques poses challenges for validation of these techniques not only on a single scanner but more importantly across scanners and centers. The performances of qMRI sequences and reconstruction algorithms need to be validated on dedicated qMRI phantoms with ground truth parameters available and system imperfections measured. The development of such phantoms will facilitate clinical translation of qMRI. Biophysical models linking MR signals with underlying biophysical processes determine the accuracy and precision of the quantified parameters and their interpretations. In some qMRI applications, absence of consensus on model selection hinders its broader adoption. While a more accurate and complex model can provide more accurate parameters than a simplified model, it requires a longer scan time to achieve comparable reproducibility. The model selection

may require comprehensive evaluation of reproducibility and clinical values of the parameters. The prediction of the bias and uncertainty of the parameters under different models may be also informative for model selection. Given the versatility of modern MRI pulse sequences, exploring new MRI biomarkers for tumor hypoxia, microscopic tumor spread, and immune response presents a promising research avenue. This direction of research necessitates new insights of biology and physics and development of new biophysical models. This could significantly enhance radiotherapy adaptation and advance cancer research as a whole.

Bibliography

[1] M. A. Schmidt and G. S. Payne, "Radiotherapy planning using MRI," *Phys Med Biol*, vol. 60, no. 22, pp. R323–R361, 2015, doi: 10.1088/0031-9155/60/22/R323.

[2] D. Wu *et al.*, "Time-Dependent Diffusion MRI for Quantitative Microstructural Mapping of Prostate Cancer," *Radiology*, Mar. 2022, doi: 10.1148/radiol.211180.

[3] C. M. Pirkl *et al.*, "Accelerated 3D whole-brain T1, T2, and proton density mapping: feasibility for clinical glioma MR imaging," *Neuroradiology*, vol. 63, no. 11, pp. 1831–1851, Nov. 2021, doi: 10.1007/s00234-021-02703-0.

[4] Y. Cao, "The Promise of Dynamic Contrast-Enhanced Imaging in Radiation Therapy," *Semin Radiat Oncol*, vol. 21, no. 2, pp. 147–156, 2011, doi: 10.1016/j.semradonc.2010.11.001.

[5] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, "Compressed Sensing MRI," *IEEE Signal Process Mag*, vol. 25, no. 2, pp. 72–82, Mar. 2008, doi: 10.1109/msp.2007.914728.

[6] C. Wang, F. F. Yin, J. P. Kirkpatrick, and Z. Chang, "Accelerated Brain DCE-MRI Using Iterative Reconstruction With Total Generalized Variation Penalty for Quantitative Pharmacokinetic Analysis: A Feasibility Study," *Technol Cancer Res Treat*, vol. 16, no. 4, pp. 446–460, 2017, doi: 10.1177/1533034616649294.

[7] B. Zhao *et al.*, "Improved magnetic resonance fingerprinting reconstruction with low-rank and subspace modeling," *Magn Reson Med*, vol. 79, no. 2, pp. 933–942, Feb. 2018, doi: 10.1002/mrm.26701.

[8] X. Cao *et al.*, "Optimized multi-axis spiral projection <scp>MR</scp> fingerprinting with subspace reconstruction for rapid whole-brain high-isotropic-resolution quantitative imaging," *Magn Reson Med*, Feb. 2022, doi: 10.1002/mrm.29194.

[9] G. Lima da Cruz, A. Bustin, O. Jaubert, T. Schneider, R. M. Botnar, and C. Prieto, "Sparsity and locally low rank regularization for MR fingerprinting," *Magn Reson Med*, vol. 81, no. 6, pp. 3530–3543, Jun. 2019, doi: 10.1002/mrm.27665.

[10] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, "SENSE: Sensitivity encoding for fast MRI," *Magn Reson Med*, vol. 42, no. 5, pp. 952–962, Nov. 1999, doi: 10.1002/(SICI)1522-2594(199911)42:5<952::AID-MRM16>3.0.CO;2-S.

[11] M. A. Griswold *et al.*, “Generalized Autocalibrating Partially Parallel Acquisitions (GRAPPA),” *Magn Reson Med*, vol. 47, no. 6, pp. 1202–1210, 2002, doi: 10.1002/mrm.10171.

[12] C. Qin, J. Schlemper, K. Hammernik, J. Duan, R. M. Summers, and D. Rueckert, “Deep Network Interpolation for Accelerated Parallel MR Image Reconstruction,” vol. 2, pp. 1–6, 2020.

[13] H. K. Aggarwal, M. P. Mani, and M. Jacob, “MoDL: Model-Based Deep Learning Architecture for Inverse Problems,” *IEEE Trans Med Imaging*, vol. 38, no. 2, pp. 394–405, 2019, doi: 10.1109/TMI.2018.2865356.

[14] H. K. Aggarwal and M. Jacob, “J-MoDL: Joint Model-Based Deep Learning for Optimized Sampling and Reconstruction,” *IEEE J Sel Top Signal Process*, vol. 14, no. 6, pp. 1151–1162, Oct. 2020, doi: 10.1109/JSTSP.2020.3004094.

[15] J. Schlemper, J. Caballero, J. V. Hajnal, A. N. Price, and D. Rueckert, “A Deep Cascade of Convolutional Neural Networks for Dynamic MR Image Reconstruction,” *IEEE Trans Med Imaging*, vol. 37, no. 2, pp. 491–503, 2018, doi: 10.1109/TMI.2017.2760978.

[16] K. Hammernik *et al.*, “Learning a variational network for reconstruction of accelerated MRI data,” *Magn Reson Med*, vol. 79, no. 6, pp. 3055–3071, 2018, doi: 10.1002/mrm.26977.

[17] B. Zhu, J. Z. Liu, S. F. Cauley, B. R. Rosen, and M. S. Rosen, “Image reconstruction by domain-transform manifold learning,” *Nature*, vol. 555, no. 7697, pp. 487–492, Mar. 2018, doi: 10.1038/nature25988.

[18] C. M. Hyun, H. P. Kim, S. M. Lee, S. Lee, and J. K. Seo, “Deep learning for undersampled MRI reconstruction,” *Phys Med Biol*, vol. 63, no. 13, Jun. 2018, doi: 10.1088/1361-6560/aac71a.

[19] F. Knoll, C. Clason, C. Diwoky, and R. Stollberger, “Adapted random sampling patterns for accelerated MRI,” *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 24, no. 1, pp. 43–50, 2011, doi: 10.1007/s10334-010-0234-7.

[20] M. Seeger, H. Nickisch, R. Pohmann, and B. Schölkopf, “Optimization of k-space trajectories for compressed sensing by Bayesian experimental design,” *Magn Reson Med*, vol. 63, no. 1, pp. 116–126, 2010, doi: 10.1002/mrm.22180.

[21] G. Wang, T. Luo, J. F. Nielsen, D. C. Noll, and J. A. Fessler, “B-Spline Parameterized Joint Optimization of Reconstruction and K-Space Trajectories (BJORK) for Accelerated 2D MRI,” *IEEE Trans Med Imaging*, vol. 41, no. 9, pp. 2318–2330, Sep. 2022, doi: 10.1109/TMI.2022.3161875.

[22] C. D. Bahadir, A. Q. Wang, A. V. Dalca, and M. R. Sabuncu, "Deep-Learning-Based Optimization of the Under-Sampling Pattern in MRI," *IEEE Trans Comput Imaging*, vol. 6, pp. 1139–1152, 2020, doi: 10.1109/TCI.2020.3006727.

[23] V. Vishwanath, S. Jafarieh, and A. Rembielak, "The role of imaging in head and neck cancer: An overview of different imaging modalities in primary diagnosis and staging of the disease," *J Contemp Brachytherapy*, vol. 12, no. 5, pp. 512–518, 2020, doi: 10.5114/jcb.2020.100386.

[24] S. Bisdas, M. Baghi, J. Wagenblast, T. J. Vogl, H. T. Choon, and S. K. Tong, "Gadolinium-enhanced echo-planar T2-weighted MRI of tumors in the extracranial head and neck: Feasibility study and preliminary results using a distributed-parameter tracer kinetic analysis," *Journal of Magnetic Resonance Imaging*, vol. 27, no. 5, pp. 963–969, 2008, doi: 10.1002/jmri.21311.

[25] Y. Cao *et al.*, "Early Prediction of Outcome in Advanced Head-and-Neck Cancer Based on Tumor Blood Volume Alterations During Therapy: A Prospective Study," *Int J Radiat Oncol Biol Phys*, vol. 72, no. 5, pp. 1287–1290, Dec. 2008, doi: 10.1016/j.ijrobp.2008.08.024.

[26] M. L. Mierzwa *et al.*, "Randomized Phase II Study of Physiologic MRI-Directed Adaptive Radiation Boost in Poor Prognosis Head and Neck Cancer," *Clinical Cancer Research*, vol. 28, no. 23, pp. 5049–5057, Dec. 2022, doi: 10.1158/1078-0432.CCR-22-1522.

[27] P. S. Tofts, "Modeling tracer kinetics in dynamic Gd-DTPA MR imaging," *Journal of Magnetic Resonance Imaging*, vol. 7, no. 1, pp. 91–101, Jan. 1997, doi: 10.1002/jmri.1880070113.

[28] A. Singh, R. K. S. Rathore, M. Haris, S. K. Verma, N. Husain, and R. K. Gupta, "Improved bolus arrival time and arterial input function estimation for tracer kinetic analysis in DCE-MRI," *Journal of Magnetic Resonance Imaging*, vol. 29, no. 1, pp. 166–176, Jan. 2009, doi: 10.1002/jmri.21624.

[29] D. Ma *et al.*, "Magnetic resonance fingerprinting," *Nature*, vol. 495, no. 7440, pp. 187–192, 2013, doi: 10.1038/nature11971.

[30] M. Afzali *et al.*, "MR Fingerprinting with b-Tensor Encoding for Simultaneous Quantification of Relaxation and Diffusion in a Single Scan," *Magn Reson Med*, vol. 88, no. 5, pp. 2043–2057, Nov. 2022, doi: 10.1002/mrm.29352.

[31] O. Cohen, S. Huang, M. T. McMahon, M. S. Rosen, and C. T. Farrar, "Rapid and quantitative chemical exchange saturation transfer (CEST) imaging with magnetic resonance fingerprinting (MRF)," *Magn Reson Med*, vol. 80, no. 6, pp. 2449–2463, Dec. 2018, doi: 10.1002/mrm.27221.

[32] O. Perlman *et al.*, “Quantitative imaging of apoptosis following oncolytic virotherapy by magnetic resonance fingerprinting aided by deep learning,” *Nat Biomed Eng*, vol. 6, no. 5, pp. 648–657, May 2022, doi: 10.1038/s41551-021-00809-7.

[33] T. Christen *et al.*, “MR vascular fingerprinting: A new approach to compute cerebral blood volume, mean vessel radius, and oxygenation maps in the human brain,” *Neuroimage*, vol. 89, pp. 262–270, Apr. 2014, doi: 10.1016/j.neuroimage.2013.11.052.

[34] L. Lu *et al.*, “Initial assessment of 3D magnetic resonance fingerprinting (MRF) towards quantitative brain imaging for radiation therapy,” *Med Phys*, vol. 47, no. 3, pp. 1199–1214, 2020, doi: 10.1002/mp.13967.

[35] C. Liao *et al.*, “Detection of lesions in mesial temporal lobe epilepsy by using MR fingerprinting,” *Radiology*, vol. 288, no. 3, pp. 804–812, Sep. 2018, doi: 10.1148/radiol.2018172131.

[36] C. Badve *et al.*, “MR fingerprinting of adult brain tumors: Initial experience,” in *American Journal of Neuroradiology*, American Society of Neuroradiology, Mar. 2017, pp. 492–499. doi: 10.3174/ajnr.A5035.

[37] P. De Blank *et al.*, “Magnetic Resonance Fingerprinting to Characterize Childhood and Young Adult Brain Tumors,” *Pediatr Neurosurg*, vol. 54, no. 5, pp. 310–318, Oct. 2019, doi: 10.1159/000501696.

[38] T. Y. Su *et al.*, “Evaluating whole-brain tissue-property changes in MRI-negative pharmacoresistant focal epilepsies using MR fingerprinting,” *Epilepsia*, vol. 64, no. 2, pp. 430–442, Feb. 2023, doi: 10.1111/epi.17488.

[39] O. Jaubert *et al.*, “T1, T2, and Fat Fraction Cardiac MR Fingerprinting: Preliminary Clinical Evaluation,” *Journal of Magnetic Resonance Imaging*, vol. 53, no. 4, pp. 1253–1265, Apr. 2021, doi: 10.1002/jmri.27415.

[40] O. Jaubert *et al.*, “Multi-parametric liver tissue characterization using MR fingerprinting: Simultaneous T1, T2, T2*, and fat fraction mapping,” *Magn Reson Med*, vol. 84, no. 5, pp. 2625–2635, 2020, doi: 10.1002/mrm.28311.

[41] Y. Chen *et al.*, “MR fingerprinting for rapid quantitative abdominal imaging,” *Radiology*, vol. 279, no. 1, pp. 278–286, Apr. 2016, doi: 10.1148/radiol.2016152037.

[42] P. De Blank *et al.*, “Magnetic Resonance Fingerprinting to Characterize Childhood and Young Adult Brain Tumors,” *Pediatr Neurosurg*, vol. 54, no. 5, pp. 310–318, Oct. 2019, doi: 10.1159/000501696.

[43] G. Körzdörfer *et al.*, “Reproducibility and repeatability of MR fingerprinting relaxometry in the human brain,” *Radiology*, vol. 292, no. 2, pp. 429–437, 2019, doi: 10.1148/radiol.2019182360.

[44] J. Zou, J. M. Balter, and Y. Cao, “Estimation of pharmacokinetic parameters from DCE-MRI by extracting long and short time-dependent features using an LSTM network,” *Med Phys*, vol. c, pp. 1–11, 2020, doi: 10.1002/mp.14222.

[45] J. Zou and Y. Cao, “Joint Optimization of k-t Sampling Pattern and Reconstruction of DCE MRI for Pharmacokinetic Parameter Estimation,” *IEEE Trans Med Imaging*, vol. 41, no. 11, pp. 3320–3331, Nov. 2022, doi: 10.1109/TMI.2022.3184261.

[46] J. Zou and Y. Cao, “Sampling Pattern Optimization for Multi-Contrast MRI with a Fully Unrolled Reconstruction Network,” in *Proceedings of AAPM Annual Meeting*, 2022.

[47] Y. Cao, D. Li, Z. Shen, and D. Normolle, “Sensitivity of Quantitative Metrics Derived from DCE MRI and a Pharmacokinetic Model to Image Quality and Acquisition Parameters,” *Acad Radiol*, vol. 17, no. 4, pp. 468–478, Apr. 2010, doi: 10.1016/j.acra.2009.10.021.

[48] K. Murase, “Efficient method for calculating kinetic parameters using T1-weighted dynamic contrast-enhanced magnetic resonance imaging,” *Magn Reson Med*, vol. 51, no. 4, pp. 858–862, Apr. 2004, doi: 10.1002/mrm.20022.

[49] C. Wang, F. F. Yin, and Z. Chang, “An efficient calculation method for pharmacokinetic parameters in brain permeability study using dynamic contrast-enhanced MRI,” *Magn Reson Med*, vol. 75, no. 2, pp. 739–749, Feb. 2016, doi: 10.1002/mrm.25659.

[50] C. Ulas *et al.*, “Direct Estimation of Pharmacokinetic Parameters from DCE-MRI using Deep CNN with Forward Physical Model Loss,” vol. 11070 LNCS, pp. 39–47, Apr. 2018, doi: 10.1007/978-3-030-00928-1_5.

[51] C. Ulas *et al.*, “Convolutional Neural Networks for Direct Inference of Pharmacokinetic Parameters: Application to Stroke Dynamic Contrast-Enhanced MRI,” *Front Neurol*, vol. 9, no. JAN, Jan. 2019, doi: 10.3389/fneur.2018.01147.

[52] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” *Neural Comput*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[53] J. Donahue *et al.*, “Long-Term Recurrent Convolutional Networks for Visual Recognition and Description,” *IEEE Trans Pattern Anal Mach Intell*, vol. 39, no. 4, pp. 677–691, Apr. 2017, doi: 10.1109/TPAMI.2016.2599174.

[54] J. Chen, L. Yang, Y. Zhang, M. Alber, and D. Z. Chen, “Combining fully convolutional and recurrent neural networks for 3D biomedical image segmentation,” in *Advances in Neural Information Processing Systems*, Sep. 2016, pp. 3044–3052.

[55] N. Antropova, B. Huynh, H. Li, and M. L. Giger, “Breast lesion classification based on dynamic contrast-enhanced magnetic resonance images sequences with long short-term

memory networks," *Journal of Medical Imaging*, vol. 6, no. 01, p. 1, Aug. 2018, doi: 10.1117/1.jmi.6.1.011002.

[56] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[57] M. J. Paldino and D. P. Barboriak, "Fundamentals of Quantitative Dynamic Contrast-Enhanced MR Imaging," *Magn Reson Imaging Clin N Am*, vol. 17, no. 2, pp. 277–289, May 2009, doi: 10.1016/j.mric.2009.01.007.

[58] Y. Cao, "WE-D-T-6C-03: Development of Image Software Tools for Radiation Therapy Assessment," *Med Phys*, vol. 32, no. 6Part19, pp. 2136–2136, Jun. 2005, doi: 10.1118/1.1999737.

[59] T. Fritz-Hansen, E. Rostrup, H. B. W. Larsson, L. Søndergaard, P. Ring, and O. Henriksen, "Measurement of the arterial concentration of Gd-DTPA using MRI: A step toward quantitative perfusion imaging," *Magn Reson Med*, vol. 36, no. 2, pp. 225–231, Aug. 1996, doi: 10.1002/mrm.1910360209.

[60] D. Robben and P. Suetens, "Perfusion parameter estimation using neural networks and data augmentation," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Springer Verlag, 2019, pp. 439–446. doi: 10.1007/978-3-030-11723-8_44.

[61] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," in *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, Dec. 2015.

[62] A. Arbelle and T. R. Raviv, "Microscopy cell segmentation via convolutional LSTM networks," in *Proceedings - International Symposium on Biomedical Imaging*, IEEE, Apr. 2019, pp. 1008–1012. doi: 10.1109/ISBI.2019.8759447.

[63] D. Zhang *et al.*, "A multi-level convolutional LSTM model for the segmentation of left ventricle myocardium in infarcted porcine cine MR images," in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, IEEE, Apr. 2018, pp. 470–473. doi: 10.1109/ISBI.2018.8363618.

[64] J. Simeth and Y. Cao, "GAN and dual-input two-compartment model-based training of a neural network for robust quantification of contrast uptake rate in gadoxetic acid-enhanced MRI," *Med Phys*, pp. 1–11, 2020, doi: 10.1002/mp.14055.

[65] J. J. Van Vaals *et al.*, "'Keyhole' method for accelerating imaging of contrast agent uptake," *Journal of Magnetic Resonance Imaging*, vol. 3, no. 4, pp. 671–675, 1993, doi: 10.1002/jmri.1880030419.

[66] R. P. Lim *et al.*, "3D time-resolved MR angiography (MRA) of the carotid arteries with time-resolved imaging with stochastic trajectories: Comparison with 3D contrast-enhanced bolus-chase MRA and 3D time-of-flight MRA," *American Journal of Neuroradiology*, vol. 29, no. 10, pp. 1847–1854, 2008, doi: 10.3174/ajnr.A1252.

[67] L. Feng *et al.*, "Golden-angle radial sparse parallel MRI: combination of compressed sensing, parallel imaging, and golden-angle radial sampling for fast and flexible dynamic volumetric MRI," *Magn Reson Med*, vol. 72, no. 3, pp. 707–717, 2014, doi: 10.1002/mrm.24980.

[68] Y. Zhu, Y. Guo, S. G. Lingala, R. M. Lebel, M. Law, and K. S. Nayak, "GOCART: GOlden-angle CArtesian randomized time-resolved 3D MRI," *Magn Reson Imaging*, vol. 34, no. 7, pp. 940–950, 2016, doi: 10.1016/j.mri.2015.12.030.

[69] T. Song *et al.*, "Optimal k-space sampling for dynamic contrast-enhanced MRI with an application to MR renography," *Magn Reson Med*, vol. 61, no. 5, pp. 1242–1248, 2009, doi: 10.1002/mrm.21901.

[70] J. Choi and H. Kim, "Implementation of time-efficient adaptive sampling function design for improved undersampled MRI reconstruction," *Journal of Magnetic Resonance*, vol. 273, pp. 47–55, 2016, doi: 10.1016/j.jmr.2016.10.006.

[71] B. Gözcü *et al.*, "Learning-Based Compressive MRI," *IEEE Trans Med Imaging*, vol. 37, no. 6, pp. 1394–1406, Jun. 2018, doi: 10.1109/TMI.2018.2832540.

[72] M. V. W. Zibetti, G. T. Herman, and R. R. Regatte, "Fast Data-Driven Learning of MRI Sampling Pattern for Large Scale Problems," *Sci Rep*, no. 0123456789, pp. 1–19, 2020, doi: 10.1038/s41598-021-97995-w.

[73] F. Sherry *et al.*, "Learning the Sampling Pattern for MRI," *IEEE Trans Med Imaging*, vol. 39, no. 12, pp. 4310–4321, 2020, doi: 10.1109/TMI.2020.3017353.

[74] Y. Bliesener, S. G. Lingala, J. P. Haldar, and K. S. Nayak, "Impact of (k,t) sampling on DCE MRI tracer kinetic parameter estimation in digital reference objects," *Magn Reson Med*, vol. 83, no. 5, pp. 1625–1639, 2020, doi: 10.1002/mrm.28024.

[75] Y. Guo, S. G. Lingala, Y. Bliesener, R. M. Lebel, Y. Zhu, and K. S. Nayak, "Joint arterial input function and tracer kinetic parameter estimation from undersampled dynamic contrast-enhanced MRI using a model consistency constraint," *Magn Reson Med*, vol. 79, no. 5, pp. 2804–2815, 2018, doi: 10.1002/mrm.26904.

[76] S. G. Lingala *et al.*, "Tracer kinetic models as temporal constraints during brain tumor DCE-MRI reconstruction," *Med Phys*, vol. 47, no. 1, pp. 37–51, Jan. 2020, doi: 10.1002/mp.13885.

[77] Y. Guo, S. G. Lingala, Y. Zhu, R. M. Lebel, and K. S. Nayak, "Direct estimation of tracer-kinetic parameter maps from highly undersampled brain dynamic contrast enhanced MRI," *Magn Reson Med*, vol. 78, no. 4, pp. 1566–1578, Oct. 2017, doi: 10.1002/mrm.26540.

[78] C. Y. Lin and J. A. Fessler, "Efficient Dynamic Parallel MRI Reconstruction for the Low-Rank Plus Sparse Model," *IEEE Trans Comput Imaging*, vol. 5, no. 1, pp. 17–26, 2018, doi: 10.1109/tci.2018.2882089.

[79] C. Qin, J. Schlemper, J. Caballero, A. Price, J. V. Hajnal, and D. Rueckert, "Convolutional Recurrent Neural Networks for Dynamic MR Image Reconstruction," *IEEE Trans Med Imaging*, vol. 38, no. 1, pp. 280–290, Dec. 2017, doi: 10.1109/TMI.2018.2863670.

[80] C. Qin *et al.*, "k-t NEXT: Dynamic MR Image Reconstruction Exploiting Spatio-Temporal Correlations," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11765 LNCS, pp. 505–513, 2019, doi: 10.1007/978-3-030-32245-8_56.

[81] J. Zhang *et al.*, "Extending LOUPE for K-space Under-sampling Pattern Optimization in Multi-coil MRI," pp. 1–11, 2020, [Online]. Available: http://arxiv.org/abs/2007.14450

[82] T. Weiss, O. Senouf, S. Vedula, O. Michailovich, M. Zibulevsky, and A. Bronstein, "PILOT: Physics-Informed Learned Optimized Trajectories for Accelerated MRI," pp. 1–23, 2019, [Online]. Available: http://arxiv.org/abs/1909.05773

[83] S. G. Lingala *et al.*, "Tracer Kinetic Models as Temporal Constraints during DCE-MRI reconstruction," 2017.

[84] C. D. Bahadir, A. V. Dalca, and M. R. Sabuncu, "Learning-Based Optimization of the Under-Sampling Pattern in MRI," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11492 LNCS, pp. 780–792, 2019, doi: 10.1007/978-3-030-20351-1_61.

[85] C. Qin, J. Schlemper, J. Caballero, A. N. Price, J. V. Hajnal, and D. Rueckert, "Convolutional Recurrent Neural Networks for Dynamic MR Image Reconstruction," *IEEE Trans Med Imaging*, vol. 38, no. 1, pp. 280–290, Jan. 2019, doi: 10.1109/TMI.2018.2863670.

[86] S. G. Lingala, Y. Hu, E. Dibella, and M. Jacob, "Accelerated dynamic MRI exploiting sparsity and low-rank structure: K-t SLR," *IEEE Trans Med Imaging*, vol. 30, no. 5, pp. 1042–1054, May 2011, doi: 10.1109/TMI.2010.2100850.

[87] J. Schlemper *et al.*, "Data consistency networks for (calibration-less) accelerated parallel MR image reconstruction," vol. 1, pp. 2–7, 2019.

[88] X. Tao, H. Gao, X. Shen, J. Wang, and J. Jia, “Scale-Recurrent Network for Deep Image Deblurring,” in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, Jun. 2018, pp. 8174–8182. doi: 10.1109/CVPR.2018.00853.

[89] M. M. Kim *et al.*, “Developing a Pipeline for Multiparametric MRI-Guided Radiation Therapy: Initial Results from a Phase II Clinical Trial in Newly Diagnosed Glioblastoma,” *Tomography*, vol. 5, no. 1, pp. 118–126, 2019, doi: 10.18383/j.tom.2018.00035.

[90] M. Uecker *et al.*, “ESPIRiT - An eigenvalue approach to autocalibrating parallel MRI: Where SENSE meets GRAPPA,” *Magn Reson Med*, vol. 71, no. 3, pp. 990–1001, 2014, doi: 10.1002/mrm.24751.

[91] T. Zhang, J. M. Pauly, S. S. Vasanawala, and M. Lustig, “Coil compression for accelerated imaging with Cartesian sampling,” *Magn Reson Med*, vol. 69, no. 2, pp. 571–582, 2013, doi: 10.1002/mrm.24267.

[92] R. Bridson, “Fast poisson disk sampling in arbitrary dimensions,” *ACM SIGGRAPH 2007 Sketches, SIGGRAPH’07*, p. 2006, 2007, doi: 10.1145/1278780.1278807.

[93] Y. Bliesener, J. Acharya, and K. S. Nayak, “Efficient DCE-MRI Parameter and Uncertainty Estimation Using a Neural Network,” *IEEE Trans Med Imaging*, vol. 39, no. 5, pp. 1712–1723, 2020, doi: 10.1109/TMI.2019.2953901.

[94] E. Levine, B. Daniel, S. Vasanawala, B. Hargreaves, and M. Saranathan, “3D Cartesian MRI with compressed sensing and variable view sharing using complementary poisson-disc sampling,” *Magn Reson Med*, vol. 77, no. 5, pp. 1774–1785, 2017, doi: 10.1002/mrm.26254.

[95] Y. Le, R. Kroeker, H. D. Kipfer, and C. Lin, “Development and evaluation of TWIST Dixon for dynamic contrast-enhanced (DCE) MRI with improved acquisition efficiency and fat suppression,” *Journal of Magnetic Resonance Imaging*, vol. 36, no. 2, pp. 483–491, 2012, doi: 10.1002/jmri.23663.

[96] J. Zou, J. M. Balter, and Y. Cao, “Estimation of pharmacokinetic parameters from DCE-MRI by extracting long and short time-dependent features using an LSTM network,” *Med Phys*, vol. c, pp. 1–11, 2020, doi: 10.1002/mp.14222.

[97] R. E. Port, M. V. Knopp, and G. Brix, “Dynamic contrast-enhanced MRI using Gd-DTPA: Interindividual variability of the arterial input function and consequences for the assessment of kinetics in tumors,” *Magn Reson Med*, vol. 45, no. 6, pp. 1030–1038, 2001, doi: 10.1002/mrm.1137.

[98] L. Georgiou, D. J. Wilson, N. Sharma, T. J. Perren, and D. L. Buckley, “A functional form for a representative individual arterial input function measured from a population

using high temporal resolution DCE MRI," *Magn Reson Med*, vol. 81, no. 3, pp. 1955–1963, Mar. 2019, doi: 10.1002/mrm.27524.

[99] J. S. Park, E. Lim, S. H. Choi, C. H. Sohn, J. Lee, and J. Park, "Model-Based High-Definition Dynamic Contrast Enhanced MRI for Concurrent Estimation of Perfusion and Microvascular Permeability," *Med Image Anal*, vol. 59, p. 101566, 2020, doi: 10.1016/j.media.2019.101566.

[100] E. Ashton, D. Raunig, C. Ng, F. Kelcz, T. McShane, and J. Evelhoch, "Scan-rescan variability in perfusion assessment of tumors in MRI using both model and data-derived arterial input functions," *Journal of Magnetic Resonance Imaging*, vol. 28, no. 3, pp. 791–796, Sep. 2008, doi: 10.1002/jmri.21472.

[101] X. Liu, J. Wang, H. Sun, S. S. Chandra, S. Crozier, and F. Liu, "On the regularization of feature fusion and mapping for fast MR multi-contrast imaging via iterative networks," *Magn Reson Imaging*, vol. 77, pp. 159–168, Apr. 2021, doi: 10.1016/j.mri.2020.12.019.

[102] W. J. Do, S. Seo, Y. Han, J. C. Ye, S. H. Choi, and S. H. Park, "Reconstruction of multicontrast MR images through deep learning," *Med Phys*, vol. 47, no. 3, pp. 983–997, Mar. 2020, doi: 10.1002/mp.14006.

[103] X. Liu, J. Wang, S. Lin, S. Crozier, and F. Liu, "Optimizing multicontrast MRI reconstruction with shareable feature aggregation and selection," *NMR Biomed*, vol. 34, no. 8, Aug. 2021, doi: 10.1002/nbm.4540.

[104] D. Polak *et al.*, "Joint multi-contrast variational network reconstruction (jVN) with application to rapid 2D and 3D imaging," *Magn Reson Med*, vol. 84, no. 3, pp. 1456–1469, 2020, doi: 10.1002/mrm.28219.

[105] E. Shimron, J. I. Tamir, K. Wang, and M. Lustig, "Implicit data crimes: Machine learning bias arising from misuse of public data," *Proceedings of the National Academy of Sciences*, vol. 119, no. 13, Mar. 2022, doi: 10.1073/pnas.2117203119.

[106] H. Zheng, H. Yong, and L. Zhang, "Deep Convolutional Dictionary Learning for Image Denoising," in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2021, pp. 630–641. doi: 10.1109/CVPR46437.2021.00069.

[107] X. Cao, H. Ye, C. Liao, Q. Li, H. He, and J. Zhong, "Fast 3D brain MR fingerprinting based on multi-axis spiral projection trajectory," *Magn Reson Med*, vol. 82, no. 1, pp. 289–301, Jul. 2019, doi: 10.1002/mrm.27726.

[108] X. Cao *et al.*, "Robust sliding-window reconstruction for Accelerating the acquisition of MR fingerprinting," *Magn Reson Med*, vol. 78, no. 4, pp. 1579–1588, 2017, doi: 10.1002/mrm.26521.

[109] “Deep Learning Initialized Compressed Sensing (Deli-CS) in Volumetric Spatio-Temporal Subspace Reconstruction”, doi: 10.1101/2023.03.28.534431.

[110] D. Ma *et al.*, “Fast 3D magnetic resonance fingerprinting for a whole-brain coverage,” *Magn Reson Med*, vol. 79, no. 4, pp. 2190–2197, Apr. 2018, doi: 10.1002/mrm.26886.

[111] F. Sherry *et al.*, “Learning the Sampling Pattern for MRI,” *IEEE Trans Med Imaging*, vol. 39, no. 12, pp. 4310–4321, 2020, doi: 10.1109/TMI.2020.3017353.

[112] G. Wang, J.-F. Nielsen, J. A. Fessler, and D. C. Noll, “Stochastic Optimization of 3D Non-Cartesian Sampling Trajectory (SNOPY),” Sep. 2022, [Online]. Available: http://arxiv.org/abs/2209.11030

[113] C. Lazarus *et al.*, “SPARKLING: variable-density k-space filling curves for accelerated T2*-weighted MRI,” *Magn Reson Med*, vol. 81, no. 6, pp. 3643–3661, 2019, doi: 10.1002/mrm.27678.

[114] J. Zhang *et al.*, “LARO: Learned acquisition and reconstruction optimization to accelerate quantitative susceptibility mapping,” *Neuroimage*, vol. 268, p. 119886, Mar. 2023, doi: 10.1016/j.neuroimage.2023.119886.

[115] J. Zou, Y. Liu, J. Hamilton, Y. Jiang, N. Seiberlich, and Y. Cao, “A New Framework for 3D MR Fingerprinting with Efficient Subspace Reconstruction and Posterior Distribution Estimation,” in *Proceedings of the International Society of Magnetic Resonance in Medicine*, 2023.

[116] Y. Zhang, H. She, and Y. P. Du, “Dynamic MRI of the abdomen using parallel non-Cartesian convolutional recurrent neural networks,” *Magn Reson Med*, vol. 2020, no. 0999, pp. 1–10, 2021, doi: 10.1002/mrm.28774.

[117] Z. Miller, A. Pirasteh, and K. M. Johnson, “Memory efficient model based deep learning reconstructions for high spatial resolution 3D non-cartesian acquisitions,” *Phys Med Biol*, Apr. 2023, doi: 10.1088/1361-6560/acc003.

[118] Z. Ramzi, G. R. Chaithya, J. L. Starck, and P. Ciuciu, “NC-PDNet: a Density-Compensated Unrolled Network for 2D and 3D non-Cartesian MRI Reconstruction,” *IEEE Trans Med Imaging*, 2022, doi: 10.1109/TMI.2022.3144619.

[119] Z. Chen, Y. Chen, Y. Xie, D. Li, and A. G. Christodoulou, “Data-Consistent Non-Cartesian Deep Subspace Learning for Efficient Dynamic MR Image Reconstruction,” May 2022, doi: 10.1109/ISBI52829.2022.9761497.

[120] C. Zhang *et al.*, “Distributed Memory-Efficient Physics-Guided Deep Learning Reconstruction for Large-Scale 3d Non-Cartesian MRI,” in *2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI)*, IEEE, Mar. 2022, pp. 1–5. doi: 10.1109/ISBI52829.2022.9761485.

[121] M. Kellman *et al.*, “Memory-Efficient Learning for Large-Scale Computational Imaging,” *IEEE Trans Comput Imaging*, vol. 6, pp. 1403–1414, 2020, doi: 10.1109/TCI.2020.3025735.

[122] A. Pramanik and M. Jacob, “Accelerated parallel MRI using memory efficient and robust monotone operator learning (MOL),” Apr. 2023, [Online]. Available: http://arxiv.org/abs/2304.01351

[123] Y. Guo, S. G. Lingala, Y. Zhu, R. M. Lebel, and K. S. Nayak, “Direct estimation of tracer-kinetic parameter maps from highly undersampled brain dynamic contrast enhanced MRI,” *Magn Reson Med*, vol. 78, no. 4, pp. 1566–1578, Oct. 2017, doi: 10.1002/mrm.26540.

[124] B. Zhao, K. Setsompop, H. Ye, S. F. Cauley, and L. L. Wald, “Maximum Likelihood Reconstruction for Magnetic Resonance Fingerprinting,” *IEEE Trans Med Imaging*, vol. 35, no. 8, pp. 1812–1823, Aug. 2016, doi: 10.1109/TMI.2016.2531640.

[125] Y. Chen, Z. Fang, S. C. Hung, W. T. Chang, D. Shen, and W. Lin, “High-resolution 3D MR Fingerprinting using parallel imaging and deep learning,” *Neuroimage*, vol. 206, no. August 2019, p. 116329, 2020, doi: 10.1016/j.neuroimage.2019.116329.

[126] J. I. Hamilton, “A Self-Supervised Deep Learning Reconstruction for Shortening the Breathhold and Acquisition Window in Cardiac Magnetic Resonance Fingerprinting,” *Front Cardiovasc Med*, vol. 9, Jun. 2022, doi: 10.3389/fcvm.2022.928546.

[127] 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,” *Magn Reson Med*, vol. 74, no. 6, pp. 1621–1631, 2015, doi: 10.1002/mrm.25559.

[128] D. H. Kim, E. Adalsteinsson, and D. M. Spielman, “Simple analytic variable density spiral design,” *Magn Reson Med*, vol. 50, no. 1, pp. 214–219, Jul. 2003, doi: 10.1002/mrm.10493.

[129] J. G. Pipe and P. Menon, “Sampling Density Compensation in MRI: Rationale and an Iterative Numerical Solution,” 1999.

[130] Y. Wu and K. He, “Group Normalization,” Mar. 2018, [Online]. Available: http://arxiv.org/abs/1803.08494

[131] P. Micikevicius *et al.*, “Mixed Precision Training,” Oct. 2017, Accessed: Jun. 14, 2023. [Online]. Available: http://arxiv.org/abs/1710.03740

[132] T. Chen, B. Xu, C. Zhang, and C. Guestrin, “Training Deep Nets with Sublinear Memory Cost,” Apr. 2016, [Online]. Available: http://arxiv.org/abs/1604.06174

[133] D. F. McGivney *et al.*, “SVD compression for magnetic resonance fingerprinting in the time domain,” *IEEE Trans Med Imaging*, vol. 33, no. 12, pp. 2311–2322, Dec. 2014, doi: 10.1109/TMI.2014.2337321.

[134] J. I. Tamir *et al.*, “T2 shuffling: Sharp, multicontrast, volumetric fast spin-echo imaging,” *Magn Reson Med*, vol. 77, no. 1, pp. 180–195, 2017, doi: 10.1002/mrm.26102.

[135] B. A. Landman *et al.*, “Multi-parametric neuroimaging reproducibility: A 3-T resource study,” *Neuroimage*, vol. 54, no. 4, pp. 2854–2866, Feb. 2011, doi: 10.1016/j.neuroimage.2010.11.047.

[136] 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, doi: 10.1002/jmri.24619.

[137] F. Ong and M. Lustig, “SigPy: a python package for high performance iterative reconstruction.,” in *Proceedings of the ISMRM 27th Annual Meeting*, 2019.

[138] L. Ying and J. Sheng, “Joint image reconstruction and sensitivity estimation in SENSE (JSENSE),” *Magn Reson Med*, vol. 57, no. 6, pp. 1196–1202, 2007, doi: 10.1002/mrm.21245.

[139] G. Wang, N. Shah, K. Zhu, D. C. Noll, and J. A. Fessler, “MIRTorch: A PyTorch-powered Differentiable Toolbox for Fast Image Reconstruction and Scan Protocol Optimization,” in *Proc. Intl. Soc. Magn. Reson. Med. (ISMRM)*, 2022, p. 4982.

[140] M. J. Muckley, R. Stern, T. Murrell, and F. Knoll, “TorchKbNufft: A High-Level, Hardware-Agnostic Non-Uniform Fast Fourier Transform,” in *ISMRM Workshop on Data Sampling & Image Reconstruction*, 2020.

[141] S. Bai, J. Z. Kolter, and V. Koltun, “Deep Equilibrium Models,” Sep. 2019, [Online]. Available: http://arxiv.org/abs/1909.01377

[142] D. Gilton, G. Ongie, and R. Willett, “Deep Equilibrium Architectures for Inverse Problems in Imaging,” *IEEE Trans Comput Imaging*, vol. 7, pp. 1123–1133, 2021, doi: 10.1109/TCI.2021.3118944.

[143] C. A. Baron, N. Dwork, J. M. Pauly, and D. G. Nishimura, “Rapid compressed sensing reconstruction of 3D non-Cartesian MRI,” *Magn Reson Med*, vol. 79, no. 5, pp. 2685–2692, May 2018, doi: 10.1002/mrm.26928.

[144] N. Seiberlich, F. A. Breuer, M. Blaimer, K. Barkauskas, P. M. Jakob, and M. A. Griswold, “Non-Cartesian data reconstruction using GRAPPA operator gridding (GROG),” *Magn Reson Med*, vol. 58, no. 6, pp. 1257–1265, 2007, doi: 10.1002/mrm.21435.

[145] Y.-H. Shih, G. Wright, J. Andén, J. Blaschke, and A. H. Barnett, “cuFINUFFT: a load-balanced GPU library for general-purpose nonuniform FFTs,” *2021 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, 2021, doi: 10.1109/IPDPSW52791.2021.00105.

[146] P. K. Lee, L. E. Watkins, T. I. Anderson, G. Buonincontri, and B. A. Hargreaves, “Flexible and efficient optimization of quantitative sequences using automatic differentiation of Bloch simulations,” *Magn Reson Med*, vol. 82, no. 4, pp. 1438–1451, 2019, doi: 10.1002/mrm.27832.

[147] J. Z. Bojorquez, S. Bricq, C. Acquitter, F. Brunotte, P. M. Walker, and A. Lalande, “What are normal relaxation times of tissues at 3 T?,” *Magnetic Resonance Imaging*, vol. 35. Elsevier Inc., pp. 69–80, Jan. 01, 2017. doi: 10.1016/j.mri.2016.08.021.

[148] P. M. Johnson *et al.*, “Evaluation of the Robustness of Learned MR Image Reconstruction to Systematic Deviations Between Training and Test Data for the Models from the fastMRI Challenge,” *LNCS*, vol. 12964, pp. 25–34, 2021, doi: 10.1007/978-3-030-88552-6_3.