creative commons

COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

Master's Thesis

Frequency-splitting Dynamic MRI Reconstruction using Multi-scale 3D Convolutional Sparse Coding and Automatic Parameter Selection

Nguyen Duc Van Thanh

Department of Electrical and Computer Engineering
(Computer Science & Engineering)

Graduate School of UNIST

2019

Frequency-splitting Dynamic MRI Reconstruction using Multi-scale 3D Convolutional Sparse Coding and Automatic Parameter Selection

Nguyen Duc Van Thanh

Department of Electrical and Computer Engineering
(Computer Science & Engineering)

Graduate School of UNIST

ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY

Frequency-splitting Dynamic MRI Reconstruction using Multi-scale 3D Convolutional Sparse Coding and Automatic Parameter Selection

A thesis
submitted to the Graduate School of UNIST
in partial fulfillment of the
requirements for the degree of
Master of Science

Nguyen Duc Van Thanh

12/10/2018 of submission
Approved by

Advisor
Won-Ki Jeong

Frequency-splitting Dynamic MRI Reconstruction using Multi-scale 3D Convolutional Sparse Coding and Automatic Parameter Selection

Nguyen Duc Van Thanh

This certifies that the thesis/dissertation of Nguyen Duc Van Thanh is approved.

12/10/2018

signature

Advisor (Chair): Won-Ki Jeong

signature

Jaesik Choi

Committee member: Jaesik Choi #1

signature

Committee member: Seungjoon Yang #2

Abstract

In this thesis, we propose a novel image reconstruction algorithm using multi-scale 3D convolutional sparse coding and a spectral decomposition technique for highly undersampled dynamic Magnetic Resonance Imaging (MRI) data. The proposed method recovers high-frequency information using a shared 3D convolution-based dictionary built progressively during the reconstruction process in an unsupervised manner, while low-frequency information is recovered using a total variation-based energy minimization method that leverages temporal coherence in dynamic MRI. Additionally, the proposed 3D dictionary is built across three different scales to more efficiently adapt to various feature sizes, and elastic net regularization is employed to promote a better approximation to the sparse input data. Furthermore, the computational complexity of each component in our iterative method is analyzed. We also propose an automatic parameter selection technique based on a genetic algorithm to find optimal parameters for our numerical solver which is a variant of the alternating direction method of multipliers (ADMM). We demonstrate the performance of our method by comparing it with state-of-the-art methods on 15 single-coil cardiac, 7 single-coil DCE, and a multi-coil brain MRI datasets at different sampling rates (12.5%, 25% and 50%). The results show that our method significantly outperforms the other state-of-the-art methods in reconstruction quality with a comparable running time and is resilient to noise.

Contents

I Introduction 1

II Background 2

- 2.1 The Compressed Sensing In Dynamic MRI Problem 2
- 2.2 Dictionary Learning Algorithms 3

III Related work 5

IV Proposed Method 7

- 4.1 Reconstruction Process 7
- 4.2 Complexity analysis of proposed reconstruction algorithm 12
- 4.3 Parameter Searching Process 13

V Experiment Results 15

- 5.1 Reconstruction quality evaluation 15
- 5.2 Extension to multi-coil parallel MR 16
- 5.3 Robustness to noise 17
- 5.4 Convergence evaluation 17
- 5.5 Running time evaluation 18

VI Conclusion and Future Work 19

References 19

Acknowledgements 36

List of Figures

1 Overview of the proposed CS-MRI reconstruction process based on multi-scale 3D CSC with spectral decomposition and optimal parameter selection using a genetic algorithm. . . . 2

2 Overview of the CS-MRI problem. . . . 3

3 Overview of patch-based dictionary learning. . . . 4

4 Overview of convolutional sparse coding. . . . 4

5 Overview multi-scale 3D CSC for CS-MRI reconstruction. . . . 6

6 Genetic algorithm for automatically searching optimal parameters. . . . 14

7 MRI datasets and Cartesian undersampling masks used in our experiments 20

8 MSEs, PSNRs, and SSIMs evaluation in three different sampling rates on cardiac MRIs (first column: MSEs; second column: PSNRs; third column: SSIMs). . . . 21

9 MSEs, PSNRs, and SSIMs evaluation in three different sampling rates on tumor DCE MRIs (first column: MSEs; second column: PSNRs; third column: SSIMs). 22

10 Image quality and pixel-wise error comparison of various reconstruction methods on Cardiac 2 dataset (12.5% sampling) 23

11 Image quality and pixel-wise error comparison of various reconstruction methods on Tumor 2 dataset (12.5% sampling) 24

12 Reconstruction quality comparison of Tumor 1 dataset 25

13 PSNRs evaluation on the multi-coil MRI brain dataset 26

14 Image quality and pixel-wise error comparison of multi-coil reconstruction methods on the brain dataset (12.5% sampling) 26

15 Average PSNRs of the images reconstructed from the various noise-corrupted k-space data 27

16 Reconstruction of noise-corrupted Cardiac 3 dataset (at noise level 0.01) 27

17 Reconstruction of noise-corrupted DCE tumor 3 dataset (at noise level 0.01) . . . 28

18 Convergence of proposed method for reconstructing 12.5% sampling Cardiac 1 dataset over 100 epochs. 29

19 Convergence rate comparison of 12.5% sampling Cardiac 2 and Tumor 2 datasets in terms of PSNRs. 30

20 Convergence curves of 12.5% sampling cardiac and tumor data in terms of PSNRs using the optimal parameters found by the genetic algorithm. 31

I Introduction

Dynamic magnetic resonance imaging (MRI), including dynamic contrast-enhanced (DCE) MRI and cardiac MRI, has been widely used to analyze changes in tissue characteristics or the movement of organs over time. Since dynamic MRI's diagnostic performance is highly correlated with its temporal resolution [1], speeding up the acquisition time has been actively studied in the recent decades. More recently, compressed sensing theory [2] has been applied to the MRI reconstruction problem [3] to reduce the acquisition time. The objective of CS-MRI is to achieve close-to-perfect reconstruction from sub-Nyquist sampling [4] of k-space data from the MRI scanner. [5] showed that the computational approach to CS-MRI can also accelerate conventional imaging. Slow numerical computation of CS-MRI reconstruction process can be further accelerated by leveraging parallel computing hardware such as graphics processing units (GPUs) [6].

Early work on CS-MRI exploited the sparsity of signal by applying universal sparsifying transforms such as the Fourier transform, total variation (TV) [7], and Wavelets [8]. For dynamic CS-MRI, spatio-temporal correlations are commonly used (e.g., k-t FOCUSS [9, 10]). Dictionary learning, a technique extracting common feature sets (i.e. atoms) from the training data, has been employed in CS-MRI to replace universal sparsifying transforms [11–15]. Since dictionary learning can generate custom-designed atoms that are better fit to the target image, reconstruction quality improves significantly compared to that of universal transforms. Recently, a filter-based dictionary learning method such as convolutional sparse coding (CSC) [16, 17], has been proposed to overcome the drawbacks of conventional patch-based dictionary learning, which includes generating redundant atoms and longer running time. CSC has been successfully adopted in solving dynamic CS-MRI problems [18, 19].

In this thesis, we introduce a novel dynamic CS-MRI reconstruction method that extends state-of-the-art CSC-based reconstruction methods [18, 19]. First, we employ frequency filtering to separate low- and high-frequency components of the target image, and then use a relatively simple energy minimization process based on a temporal total variation (TV) energy for low-frequency reconstruction while more expensive feature encoding resources (i.e. dictionary of convolutional filters) are dedicated only to recovering high-frequency component of the image (see Fig. 1 for the overview of the reconstruction process). The motivation behind this approach is that, in conventional CS-MRI, low-frequency k-space data is more densely sampled (see Fig. 7c k-space mask) while filter-based dictionary learning can better represent sparse local features in high-frequency data. Second, we employ 3D convolutional filters in various scales and elastic net regularization [20] to further improve the reconstruction quality compared to conventional approaches that only use a single-scale dictionary and l_1 regularization [15, 19]. Third, we propose an automatic optimal parameter-selection method based on a genetic algorithm (GA) for our iterative numerical solver, which is a variant of the alternating direction method of multipliers (ADMM). The proposed GA method belongs to the class of metaheuristics algorithms [21]

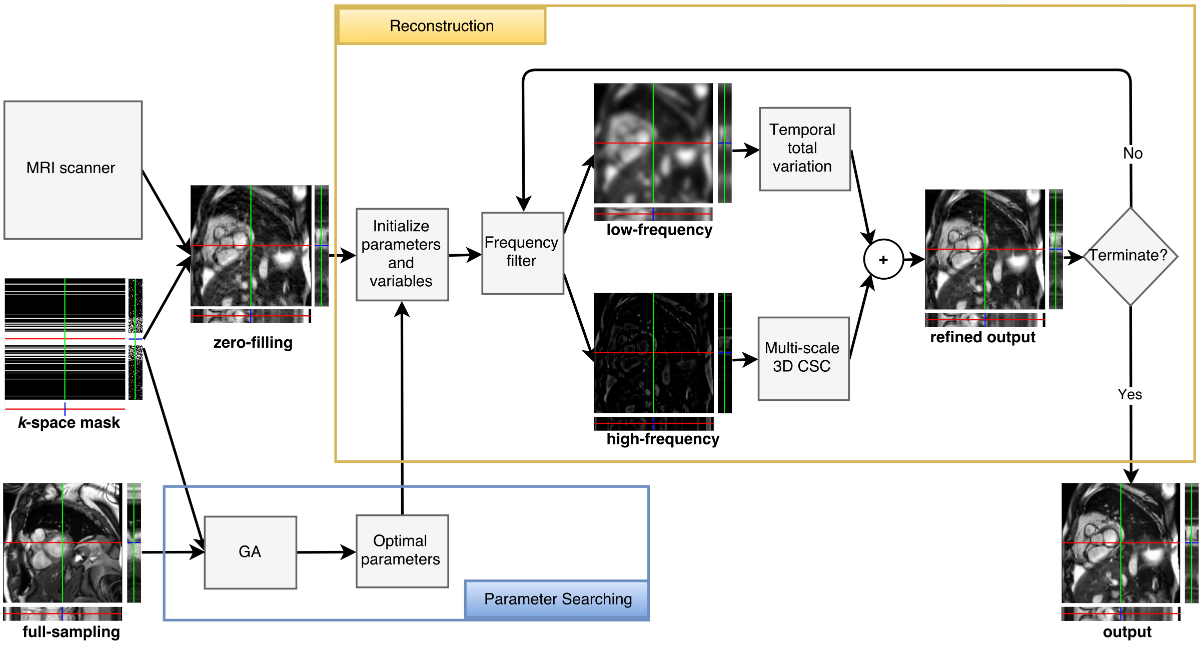

Reconstruction
MRI scanner
k-space mask
zero-filling
full-sampling
Initialize parameters and variables
Frequency filter
low-frequency
high-frequency
Temporal total variation
Multi-scale 3D CSC
+
refined output
Terminate?
No
Yes
output
GA
Optimal parameters
Parameter Searching

Figure 1: Overview of the proposed CS-MRI reconstruction process based on multi-scale 3D CSC with spectral decomposition and optimal parameter selection using a genetic algorithm.

inspired by natural evolution to optimize objective function [22]. Our approach uses a GA only once to automatically find a set of optimal parameters without manual intervention, and then we can reuse them to reconstruct similar datasets. In our experiment on 15 cardiac and 7 dynamic contrast-enhanced (DCE) MRI datasets in three different Cartesian sampling rates (12.5%, 25% and 50%), the proposed reconstruction method produces significantly better image quality, compared to several state-of-the-art methods (e.g., *k*-*t* FOCUSS [9, 10], single-scale 3D CSC [19], blind compressive sensing [13], patch-based dictionary learning [14, 15], and FTVNNR [23]), and runs at an efficient rate with GPU acceleration.

The rest of the thesis is organized as follows. Section II illustrates the background knowledge included compressed sensing problem in dynamic MRI and dictionary learning algorithms. Section III reviews the recent literature related to the proposed method. Details of the proposed reconstruction process are introduced in Section IV. We demonstrate the performance of our reconstruction method and compare it with state-of-the-art methods in Section V. Finally, we summarize our work and propose future research directions in Section VI.

II Background

2.1 The Compressed Sensing In Dynamic MRI Problem

MRI scanner uses strong magnetic fields and radiofrequency pules to generate signals from body in frequency domain (*k*-space) and images are created by taking inverse Fourier transform (F^H), as shown in first row of Fig. 2. However, it takes a long acquisition time because the magnetic fields of system need to align atoms inside organs and then the scanner records what happens as they relax back from excited state to unexcited state. In order to speed up the time of acquiring data, the scanner can only get a few samples despite of causing bad quality images (see second row, Fig. 2). A good reconstruction method is used to overcome this drawback by applying compressed sensing MRI [3] that generates high quality images and fast acquisition time. Compressed sensing MRI [3] is a sampling strategy that will allow perfect reconstruction of discrete signal from a small number of samples. The dynamic MR data is obtained in frequency domain (k-space) as a sequence of 2D images ($s_{f_x} \times s_{f_y}$) acquired at s_{f_t} different time instances which are stacked to become 3D volume s_f. Thus, undersampling mask R is used to speed up acquisition process by sampling $m << s_f$ k-space data. Given a sparsity transformation T, the problem CSMRI can be expressed with 2D Fourier transformation for every instance $\mathcal{F}_2$ as follow:

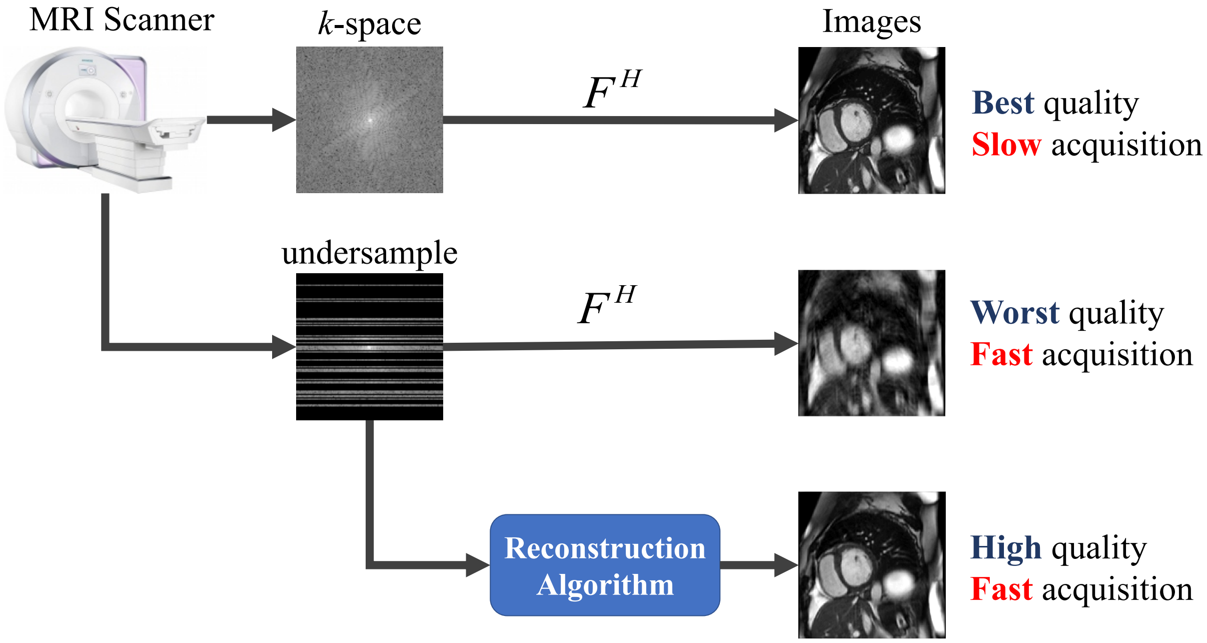


Figure 2: Overview of the CS-MRI problem.

$$\min_{s} \|T(s)\|_p \quad s.t. : \|R\mathcal{F}_2(s) - m\|_2^2 \; < \; \epsilon^2 \tag{1}$$

where $0 \leq p \leq 1$ and ϵ is a small constant representing for noise sampling in k-space. Generally, an important of CSMRI is enforcing sparsity transformation $\|T(s)\|_p$ as much as possible.

2.2 Dictionary Learning Algorithms

Patch-based dictionary learning decompose patches of image into a linear combination of overcomplete basis sets that provides sparse representation of this particular signal. The objective function Eq. 2 shows that the method collects patches and vectorize them as column representation. Output of this method is an approximation between dictionary D and sparse map x, as shown in Fig. 3. However, redundant of atom and slow running time are drawbacks of this method because of using small patches for input.

$$\min_{d,x} \frac{\alpha}{2} \sum_{n}^{N} \left\| s_n - D x_n \right\|_2^2 + \lambda \sum_{n}^{N} \|x_n\|_0 \ s.t.: \|D\|_2^2 \leq 1 \tag{2}$$

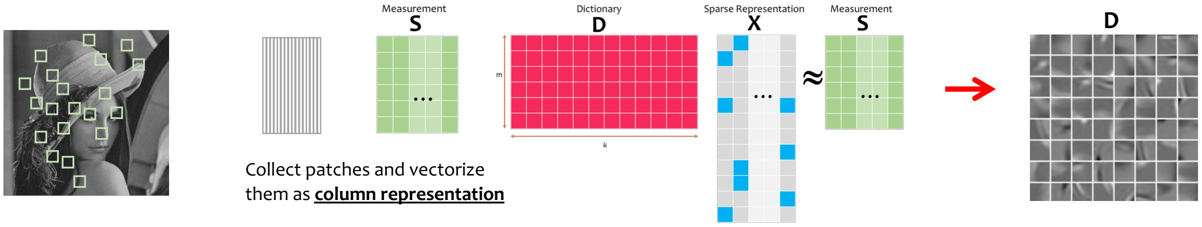


Figure 3: Overview of patch-based dictionary learning.

Convolutional sparse coding (CSC) is learning-based method that learns a shift-invariant dictionary built by convolution filters from data. The brief overview of CSC is finding its best approximation image (s) from the summation of *response map* $\sum_n^N d_n * x_n$, shown in 3 .

$$\min_{d,x} \frac{\alpha}{2} \left\| s - \sum_{n}^{N} d_n * x_n \right\|_2^2 + \lambda \sum_{n}^{N} \|x_n\|_1 \ s.t.: \|d_n\|_2^2 \leq 1 \tag{3}$$

Each *response map* is calculated by convolution between a filter (or atom) d_n and sparse map x_n. The second non-linear norm term enforces sparsity to x_n that helps finding a feasible solution of collection d_n and x_n. The remaining constraint restricts the Frobenius norm of each atom d_n within a unit length. Zeiler et al. [24] proposed a solution by solving series of subproblems between d_n and x_n until convergence; however, performance is significantly affected by complexity convolution operator because the solvers is completely on image domain. The efficient approaches [7, 16, 25] to solve Eq. 3 leverages Fourier Convolution theorem that convolution operator on spatial domain is equivalent to element-wise on frequency domain. For example, Fig. 4 shows that features are extracted by using convolutional sparse coding.

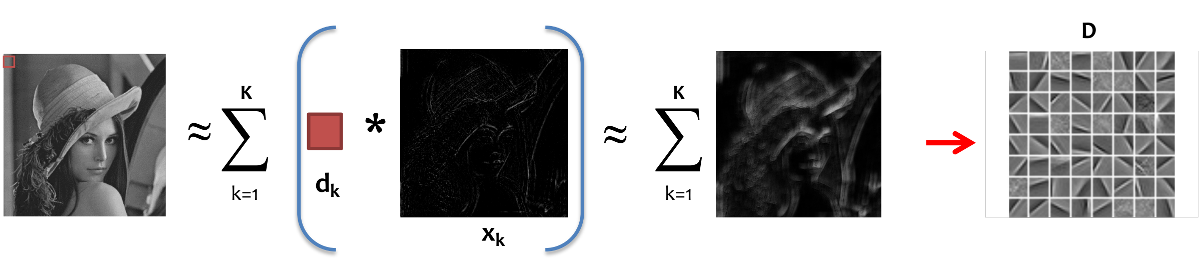


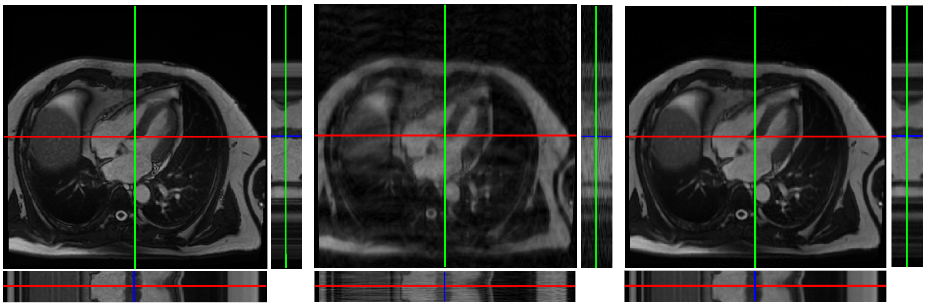
Figure 4: Overview of convolutional sparse coding.

III Related work

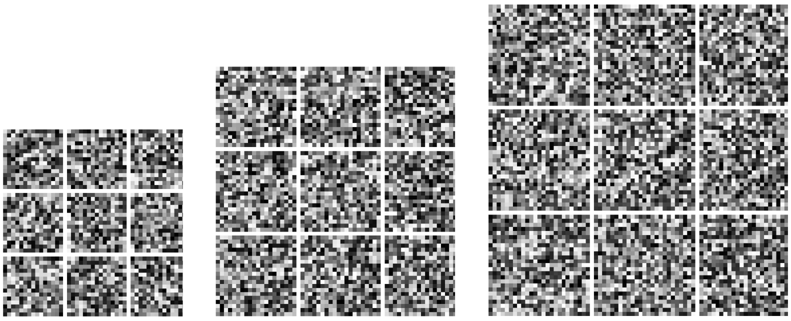
Since its inception in [3]'s seminal work, CS-MRI has been actively studied to accelerate the time-consuming MRI imaging process. Conventional CS-MRI algorithms are mainly based on promoting sparsity in the data by employing l_1 regularization in universal sparsifying transformation models. In such methods, the target image is transformed into a sparse domain by applying universal transforms such as Wavelet and Fourier transforms or total variation operation [26,27]. CS-MRI for dynamic data is also proposed by enforcing spatial and temporal coherence (i.e., k-t FOCUSS [9,10] and [28]). These conventional CS-MRI methods suffer from computational overhead because of solving expensive nonlinear l_1 minimization problems. This leads to the development of efficient numerical algorithms [29,30] that leverage the Alternating Direction Method of Multipliers (ADMMs) to solve this nonlinear problem in CS-MRI [6]. CS-MRI reconstruction methods using nuclear norm and low-rank matrix completion techniques [23,31–35] have also been proposed.

More recently, while the major limitation of universal transform-based methods is just using transformation in general, dictionary learning (DL) [36], an unsupervised learning approach, can be adapted to features from input data by training dictionary. Thus, conventional DL in CS-MRI [11–15] approaches have successfully applied to enhancing MRI reconstruction quality. Caballero et al. [14,15], for instance, demonstrated the efficiency of using patch-based DL for dynamic MRI reconstruction with temporal TV filter for enforcing coherence of time-axis. Efficient convolutional sparse coding (CSC) is then introduced by minimizing an energy objective function using a convolutional operator on the image domain, which leads to element-wise in the frequency domain, derived within ADMMs framework [16,17,25]. [18,19] first employed CSC to solve CS-MRI problems, which significantly improves running time and reconstruction quality by building more compact and expressive shift-invariance convolutional filters. Nevertheless, MR data contains various feature sizes so that multiple dictionary sizes in the multi-scale 3D CSC method can adapt well to the data rather than single-scale atom size methods (e.g., [13,15,19]). The multi-scale 3D CSC in our method builds the dictionary by simultaneously learning shift-invariant multiple sizes of convolutional filters from data (Fig. 5). In this approach, zero-filling (Fig. 5b, zero-filling reconstruction) and randomly initialized filters (e.g., in Fig. 5d, the various filter sizes) are updated iteratively until they converge, as shown in Fig. 5c, 5e, and 5f.

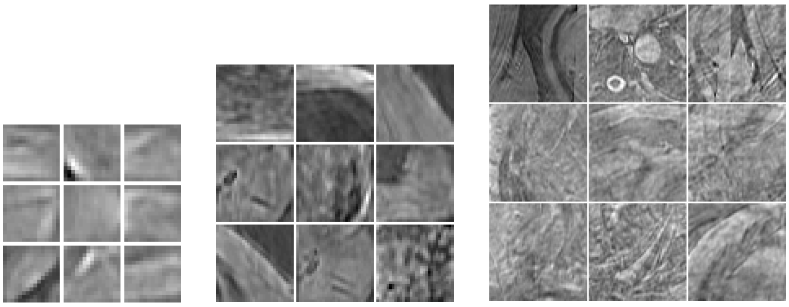
There are a few approaches that leverage different reconstruction strategies based on the frequency range of the image. [37] proposed a two-stage reconstruction method which performs the first-step reconstruction of low-frequency component only from the central region of k-space data, and then combines this result with the remaining k-space data (corresponding to high-frequency component) to conduct the second-step reconstruction from the full k-space data. A filter-based reconstruction approach [38] separates k-space with high- and low-pass filters, reconstructs each frequency component independently, and combines the reconstructed results to generate the final image. Our method shares a similar idea with these previous work by split-



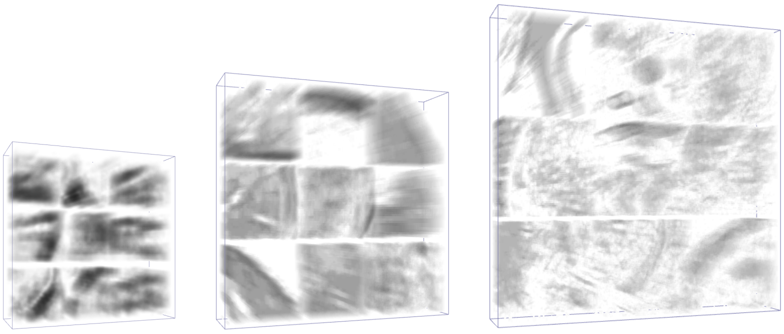
(a) Full-sampling (b) Zero-filling (c) Converged result



(d) Initial filters



(e) 2D view of converged filters



(f) 3D view of converged filters

Figure 5: Overview multi-scale 3D CSC for CS-MRI reconstruction.

ting k-space (i.e. frequency) data, but our method employs advanced reconstruction methods specifically designed to leverage characteristics of each frequency (i.e. multi-scale 3D CSC for

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

sparsity in high-frequency and temporal TV for coherency in low-frequency). Furthermore, our method iteratively updates the result to enforce the measurement consistency (see Algorithm 1 and the constrain term in Eq. 7) while the other methods perform one-time reconstruction for each frequency data and then combining them.

Genetic algorithm (GA) is a population-based method which is inspired by nature's capability to evolve living beings well-adapted to their environment [21,39]. This meta-heuristic algorithm, which is well known in optimization problem with many applications [21,22,39], deals in every iteration with a set of solutions rather than with a single solution (i.e. hill-climbing, simulated annealing and tabu search). The main idea of GA is generating a sequence of populations (i.e. generations) where each individual in a population is a solution to the problem. A new generation is created from the previous generation using natural operations such as mutations and crossovers based on the fitness values from members of the current population. This evolution process stops when the number of generations is greater than the pre-defined number and outputs an individual which has a minimum fitness value. Meta-heuristics in optimization is an immense research field and many different classes of algorithms exist, such as single-state methods, population methods, and hybrid methods [21,39]. To the best of our knowledge, our approach is the first attempt to use GA for automatic parameter selection in the CS-MRI reconstruction problem.

IV Proposed Method

The overview of our proposed CS-MRI reconstruction process is illustrated in Fig. 1. The input to our method is a zero-filling reconstruction that is generated by applying inverse Fourier transform to the undersampled k-space data from the MRI scanner, which suffers from undersampling artifacts. Then, the proposed method consists of two components: *the reconstruction process* that takes a zero-filling reconstruction as an initial guess to improve image quality (i.e. removing undersampling artifact, Fig. 1 yellow box), and *the parameter searching process* using a GA (Fig. 1 blue box).

4.1 Reconstruction Process

In the reconstruction process, the input zero-filling reconstruction image is separated into low- and high-frequency zero-filling images by applying a band-pass frequency filter. The high-frequency image is then reconstructed by using multi-scale 3D CSC with elastic net regularization. This multi-scale CSC method adapts well to features of various sizes, and elastic net regularization can outperform l_1-only regularization without impairing the sparsity of representation [20]. In the meantime, the low-frequency image is reconstructed by minimizing the total variation along the temporal direction (i.e., promoting sparsity in the temporal gradient field), which is based on the observation that the information in dynamic MRI data is sparser in first-order temporal gradients than in spatial gradients [15]. At the end of the reconstruction process, the reconstructed low- and high-frequency images are combined together to generate a refined output (i.e. undersampling artifact is reduced). This update process is repeated (see Algorithm 1) until reaching the stopping criteria in which the number of iterations are greater than a pre-defined number, or primal and dual residuals are less than an absolute value, as discussed in Sec. 3.3 of [30]. More detailed discussions of the proposed method are presented in the following sections.

Algorithm 1 Iterative reconstruction process using spectral decomposition

```
procedure RECONSTRUCTION($m$)                                  ▷ $m$: sparse measurement of $k$-space
    $s_0 \leftarrow \mathcal{F}_2^H(m)$                                    ▷ input: zero-filling reconstruction
    while stopping criteria not met do
        $s_i^l = \mathcal{F}_2^H H \mathcal{F}_2(s_i)$             ▷ $H$: low-pass filter; $\mathcal{F}_2$: 2D Fourier transform
        $s_i^h = s_i - s_i^l$
        $s_{i+1}^l \leftarrow$ find $s_{i+1}^l$ in minimizing temporal TV problem          ▷ Eq. 8
        $s_{i+1}^h \leftarrow$ find $s_{i+1}^h$ in multi-scale 3D CSC problem             ▷ Eq. 12
        $s_i \leftarrow s_{i+1}^l + s_{i+1}^h$
    return $s_i$                                                ▷ $s_i$ with removed artifact
```

Spectral decomposition using a frequency filter

To apply different reconstruction strategies based on the frequency range, we decompose the input image into low- and high-frequency images by applying a two-dimensional Butterworth low-pass filter (BLPF) to every 2D MRI slice in the frequency domain. The transfer function $H(u,v)$ of the BLPF of order n with a cutoff frequency at a specified distance D_0 from the origin is defined as follows:

$$H(u,v) = \frac{1}{1 + [D(u,v)/D_0]^{2n}} \tag{4}$$

where $D(u,v)$ is the distance from a point (u,v) to the center of the frequency domain, and the parameters D_0 and n define how the frequency is cut off. Thus, for the input image s, the low-frequency image s^l can be computed using two-dimensional Fourier transform $\mathcal{F}_2$ and its inverse $\mathcal{F}_2^H$ along the time-axis as follows:

$$s^l = \mathcal{F}_2^H H \mathcal{F}_2(s) \tag{5}$$

and the high-frequency image s^h can be computed as follows:

$$s^h = s - s^l \tag{6}$$

The choice of D_0 and n can be automatically made via our proposed parameter selection method (see Section 4.3).

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

Temporal total variation and multi-scale 3D CSC for reconstruction

After the frequency filter splits the input image s into high (s^h) and low (s^l) frequency images, the reconstruction process reduces undersampling artifact in s by solving the energy minimization problem as follows:

$$\begin{aligned} \min_{d,x,s^h,s^l} \frac{\alpha}{2}\left\| s^h - \sum_n^N \sum_k^K d_{n,k} * x_{n,k} \right\|_2^2 \\ + \lambda_1 \sum_n^N \sum_k^K \left\| x_{n,k} \right\|_1 + \frac{\lambda_2}{2} \sum_n^N \sum_k^K \left\| x_{n,k} \right\|_2^2 + \theta \left\| \nabla_t s^l \right\|_1 \\ s.t. : \|R\mathcal{F}_2(s^h + s^l) - m\|_2^2 \ < \ \epsilon^2, \|d_{n,k}\|_2^2 \leq 1 \end{aligned} \tag{7}$$

where N is the number of filter scales (i.e. levels), K is the number of filters in each scale, $*$ is a convolution operator, $d_{n,k}$ is the k-th filter (or atom) in n^{th} dictionary (i.e. dictionary for level n), and $x_{n,k}$ is its corresponding sparse code (or sparse map) for s^h. Note that the dimension of the filters we used were $15 \times 15 \times 20$, $20 \times 20 \times 25$, and $25 \times 25 \times 30$, and the dimension of the sparse map $x_{n,k}$ was identical to the size of the image. In Eq. 7, the first term measures the difference between s^h and its sparse approximation $s^h - \sum \sum d_{n,k} * x_{n,k}$, weighted by α. The combination of the second and third terms, which are weighted by λ_1 and λ_2 parameters, are called elastic net regularization [20]. The fourth term $\theta \|\nabla_t s^l\|_1$ is the total variation energy that enforces the temporal coherence of the low-frequency image. The rest of this equation is the collection of constraints: the first constraint keeps the consistency between undersampled measurement m and the undersampled reconstructed image using k-space mask R with $\mathcal{F}_2$ operator; the second constraint restricts the Frobenius norm of each atom $d_{n,k}$ within a unit length. In following discussion, we simplify the notations without indices n, k and also replace the result of Fourier transform of a given variable by using subscript f (for instance, d_{f_3} represents simplified notation for $\mathcal{F}_3 d$ in the 3D domain and $s^h_{f_2}$ is the simplified notation for $\mathcal{F}_2 s^h$ in the 2D spatial domain). The problem (7) can be split into two sub-optimization problems as follows, which can be iteratively updated for the global minimum solution.

Temporal TV minimization: we minimize Eq. 7 with respect to s^l, which contains the total variation along the time-axis and the measurement constraint term.

$$\begin{aligned} \min_{s^l} \ \theta \|\nabla_t s^l\|_1 \\ s.t. : \|R\mathcal{F}_2(s^h + s^l) - m\|_2^2 \ < \ \epsilon^2 \end{aligned} \tag{8}$$

Eq. 8 can be written in an unconstrained form with γ parameter, as shown in Eq. 9 below:

$$\min_{s^l} \ \theta \|\nabla_t s^l\|_1 + \frac{\gamma}{2} \|R\mathcal{F}_2(s^h + s^l) - m\|_2^2 \tag{9}$$

This problem can be efficiently solved by an iterative clipping algorithm using the primal-dual method [40].

$$
\begin{aligned}
s^l_{(i+1)} &= \frac{\gamma}{2}\mathcal{F}_2^H m^l - \nabla_t^T z_{(i)} \\
z_{(i+1)} &= clip(z_{(i)} + \frac{1}{\eta}\nabla_t s^l_{(i+1)}, \frac{\theta}{2})
\end{aligned}
\tag{10}
$$

where m^l is equal to $m - R\mathcal{F}_2 s^h$ and i is the iteration number. The clipping function is defined in Eq. 11.

$$
clip(a, b) = \begin{cases} a, & \text{if } |a| \leq b \\ b \cdot sign(a), & \text{if } |a| \geq b \end{cases}
\tag{11}
$$

The index i starts with 0, the initial $z_{(0)} = 0$, and $\eta \geq maxeig(\nabla_t \nabla_t^T)$. In this case, the maximum eigenvalue of $\nabla_t \nabla_t^T$ is less than four regardless to the length of signal; thus, we can set $\eta = 4$ and the maximum number of iterations equals 40 in our experiments. Note that θ and γ are optimized by using a GA (see Section 4.3). For more details on parameter and derivation about the algorithm, refer to [40]

Multi-scale 3D CSC with elastic net regularization: in this problem, we find s^h in the energy minimization function, as follows:

$$
\begin{aligned}
\min_{d,x,s^h} \frac{\alpha}{2}\left\| s^h - \sum_n^N \sum_k^K d_{n,k} * x_{n,k} \right\|_2^2 \\
+ \lambda_1 \sum_n^N \sum_k^K \left\| x_{n,k} \right\|_1 + \frac{\lambda_2}{2} \sum_n^N \sum_k^K \left\| x_{n,k} \right\|_2^2 \\
s.t.: \|R\mathcal{F}_2(s^h + s^l) - m\|_2^2 \ < \ \epsilon^2, \|d_{n,k}\|_2^2 \leq 1
\end{aligned}
\tag{12}
$$

Eq. 12 can be re-written using auxiliary variables y and g for x and d, as follows:

$$
\begin{aligned}
\min_{d,x,g,y,s^h} \frac{\alpha}{2}\left\| s^h - \sum\sum d * x \right\|_2^2 + \lambda_1 \|y\|_1 + \frac{\lambda_2}{2}\|y\|_2^2 \\
s.t.: x - y = 0, \|R\mathcal{F}_2 s^h - m^h\|_2^2 \ < \ \epsilon^2, \\
g = Proj(d), \|g\|_2^2 \leq 1
\end{aligned}
\tag{13}
$$

where m^h is equal to $m - R\mathcal{F}_2 s^l$. The g and d variables are related by a projection operator as a combination of a truncated matrix with the corresponding dictionary size followed by a padding-zero in oder to make the dimension of g the same as that of x, and the variable g should also be zero-padded to make its size similar to g_{f_3} and x_{f_3} so we can leverage Fourier transform to solve this problem. The constrained Eq. 13 can be unconstrained by using dual variable u, h, and further regulates the measurement consistency and the dual differences with γ, ρ and σ, respectively:

$$
\begin{aligned}
\min_{d,x,g,y,s^h} \frac{\alpha}{2}\left\| s^h - \sum\sum d * x \right\|_2^2 + \lambda_1 \|y\|_1 + \frac{\lambda_2}{2}\|y\|_2^2 \\
+ \frac{\gamma}{2}\|R\mathcal{F}_2 s^h - m^h\|_2^2 + \frac{\rho}{2}\|x - y + u\|_2^2 \\
+ \frac{\sigma}{2}\|d - g + h\|_2^2 \ s.t.: g = Proj(d), \|g\|_2^2 \leq 1
\end{aligned}
\tag{14}
$$

We solve Eq. 14 by iteratively finding the minimization solution of subproblems, as shown below:

Solve for x:

$$\min_{x} \frac{\alpha}{2}\left\| \sum\sum d * x - s^h \right\|_2^2 + \frac{\rho}{2}\|x - y + u\|_2^2 \tag{15}$$

We apply the Fourier transform to subproblem (15), it becomes:

$$\min_{x_{f_3}} \frac{\alpha}{2}\left\| \sum\sum d_{f_3} x_{f_3} - s^h_{f_3} \right\|_2^2 + \frac{\rho}{2}\|x_{f_3} - y_{f_3} + u_{f_3}\|_2^2 \tag{16}$$

Next, the minimum solution can be found by taking the derivative of Eq. 16 with respect to variable x_{f_3} and setting it to zero. The solution is shown in Eq. 17.

$$(\alpha D^H_{f_3} D_{f_3} + \rho I) x_{f_3} = D^H_{f_3} s^h_{f_3} + \rho(y_{f_3} - u_{f_3}) \tag{17}$$

where D_{f_3} is the concatenation of all diagonalized matrices $d_{f_3\ n,k}$, as illustrated in Eq. 18 and $D^H_{f_3}$ is the Hermitian transpose of D_{f_3}.

$$D_{f_3} = [diag(d_{f_3\ 1,1}), ..., diag(d_{f_3\ 1,k}), ..., diag(d_{f_3\ n,k})] \tag{18}$$

Solve for y:

$$\min_{y} \lambda_1\|y\|_1 + \frac{\lambda_2}{2}\|y\|_2^2 + \frac{\rho}{2}\|x - y + u\|_2^2 \tag{19}$$

In [16–19], l_1 regularization is only used for CSC; however, our subproblem contains both lasso and ridge regularizations. Fortunately, this subproblem can also be solved by using a shrinkage operation:

$$y = S_{\lambda_1/(\lambda_2+\rho)}\left(\frac{\rho(x+u)}{\lambda_2+\rho}\right) \tag{20}$$

Update for u:

The update rule for u can be defined as a fixed-point iteration with the difference between x and y (u converges when x and y converge each other).

$$u = u + x - y \tag{21}$$

Solve for d:

$$\min_{d} \frac{\alpha}{2}\left\| \sum\sum d * x - s^h \right\|_2^2 + \frac{\sigma}{2}\|d - g + h\|_2^2 \tag{22}$$

We solve this subproblem in the Fourier domain, similar to x:

$$\min_{d_{f_3}} \frac{\alpha}{2}\left\| \sum\sum d_{f_3} x_{f_3} - s^h_{f_3} \right\|_2^2 + \frac{\sigma}{2}\|d_{f_3} - g_{f_3} + h_{f_3}\|_2^2 \tag{23}$$

$$(\alpha X^H_{f_3} X_{f_3} + \sigma I) d_{f_3} = X^H_{f_3} s^h_{f_3} + \sigma(g_{f_3} - h_{f_3}) \tag{24}$$

Note that X_{f_3} stands for the concatenated matrix of all diagonal matrices $x_{f_3\ n,k}$ as shown in Eq. 25 and $X^H_{f_3}$ is the Hermitian transpose of X_{f_3}.

$$X_{f_3} = [diag(x_{f_3\ 1,1}), ..., diag(x_{f_3\ 1,k}), ..., diag(x_{f_3\ n,k})] \tag{25}$$

Solve for g:

$$\min_{g} \frac{\sigma}{2}\|d-g+h\|_2^2 \ s.t.: \ g = Proj(d), \|g\|_2^2 \leq 1 \tag{26}$$

g can be solved by using the inverse Fourier transform of d_{f_3}. This projection should be constrained by suppressing the elements which are outside the filter size $d_{n,k}$, and followed by normalizing its l_2-norm to a unit length.

Update for h: Similar to u, we update h as follows:

$$h = h + d - g \tag{27}$$

Solve for s^h:

$$\min_{s^h} \frac{\alpha}{2}\left\| s^h - \sum\sum d * x \right\|_2^2 + \frac{\gamma}{2}\|R\mathcal{F}_2 s^h - m^h\|_2^2 \tag{28}$$

Subproblem (28) can be transformed and solved in the 2D Fourier domain:

$$\min_{s^h_{f_2}} \frac{\alpha}{2}\left\| s^h_{f_2} - \mathcal{F}_2^H \sum\sum d_{f_3} x_{f_3} \right\|_2^2 + \frac{\gamma}{2}\|R s^h_{f_2} - m^h\|_2^2 \tag{29}$$

Previously, d_{f_3} and x_{f_3} were obtained in the 3D Fourier domain, we must bring it onto the same space by applying an inverse Fourier transform along the time-axis $\mathcal{F}_2^H$. Finally, $s^h_{f_2}$ can be found by solving the following linear system:

$$(\gamma R^H R + \alpha I)s^h_{f_2} = \gamma R^H m^h + \alpha \mathcal{F}_2^H \sum\sum d_{f_3} x_{f_3} \tag{30}$$

Note that we can efficiently solve independent linear systems (17), (24), and (30) via the Sherman-Morrison formula, as shown in [17]. After the iteration process, s^h will be the results of applying a 2D inverse Fourier transform $\mathcal{F}_2^H$ to $s^h_{f_2}$.

4.2 Complexity analysis of proposed reconstruction algorithm

Our iterative method consists of spectral decomposition using low-pass filter and solving subproblems. In order to investigate computational complexity, we will carefully analyze these components using Big $\mathcal{O}$ Notation, as shown in Table 1. Suppose P is the number of pixels in dynamic MR image and M is total convolutional filters ($M = N \times K$). Firstly, the cost of fast Fourier transforms (FFTs) of the MR image is $\mathcal{O}(PlogP)$. In spectral decomposition step, low-pass filter is used to separate spectral of MR data which costs $\mathcal{O}(P)$; however, this step needs FFTs operator so it costs $\mathcal{O}(PlogP)$. Temporal TV minimization (s^l) contains FFTs, calculating temporal TV and clipping operator which cost $\mathcal{O}(PlogP)$, $\mathcal{O}(P)$ and $\mathcal{O}(P)$, respectively. For solving subproblems x, d and s^h (Eq. 17, 24 and 30), these can be efficiently solved by Sherman-Morrison formula with $\mathcal{O}(MP)$ cost proposed by [25] so that the dominant operator is FFTs at all M indices $\mathcal{O}(MPlogP)$. The same computational complexity $\mathcal{O}(MPlogP)$ can be interpreted for solving subproblem g because of domination of FFTs operator. Finding solution for y using shrinkage operation which is applied for every element in y; thus, it cost $\mathcal{O}(MP)$. Finally, the addition and subtraction operations for updating u and h cost $\mathcal{O}(MP)$. In conclusion, the proposed method costs $\mathcal{O}(MPlogP)$.

Component	**Complexity**
Spectral decomposition (Eq. 4)	$\mathcal{O}(PlogP)$
Solve s^l (Eq. 10)	$\mathcal{O}(PlogP)$
Solve x, d and s^h (Eq. 17, 24 and 30)	$\mathcal{O}(MPlogP)$
Solve g	$\mathcal{O}(MPlogP)$
Solve y (Eq. 20)	$\mathcal{O}(MP)$
Update u and h (Eq. 21 and 27)	$\mathcal{O}(MP)$

Table 1: Complexity analysis of proposed method.

4.3 Parameter Searching Process

The proposed reconstruction process contains many user-controllable parameters (e.g., α, γ, λ_1, λ_2, ρ, σ, θ, D_0 and n), and manually adjusting these parameters is a time-consuming and laborious process. Thus, we propose an automatic parameter searching process, which can be thought of as a pre-processing step of the reconstruction process. Our searching approach employ a GA, which is a meta-heuristic algorithm inspired by the process of natural evolution, as shown in Fig. 6. In the proposed parameter searching process, we create a simulation based on the objective function which is defined by a fully-sampled k-space data and an undersampling mask, and a GA finds the best parameters to minimize this objective function.

A detailed description of the parameter selection process is introduced below. First, the full (ground-truth) image is undersampled to generate zero-filling reconstruction using the same sampling mask used in the reconstruction process. Then, GA arbitrarily initializes a population with L members where each individual (i) has a set of nine randomly chosen parameters to be used in the reconstruction process. Next, the zero-filling is reconstructed L times by using our method (Fig. 1 yellow box) that each time uses a set of nine parameters which contents in every individual (i) of the current population to get reconstructed results (z). The fitness values of objective function f are then calculated for each member in the current population by comparing between the full-sampled and reconstructed image (z) as well as sparse code x. After ward, GA selects members, called parents, based on the ranking of their calculated fitness values. Finally, there are three types of children that are generated for the next generation such as elite children with the best fitness values, crossover children created by combining a pair of parents and mutation children created by introducing random changes to a single parent. The algorithm continues by creating a loop of generating the new generation based on bio-operators until reaching the terminate condition (i.e., the current generation is greater than pre-defined number). In general, the goal of GA in optimization problem is conducting many trials and errors to find the best individual which minimizes the fitness function. Therefore, we propose a fitness function (f), as shown in Eq. 31, that uses Peak Signal-To-Noise-Ratio $(PSNR)$ between the CS-reconstruction result, the full reconstruction, and the average value of sparse codes $(\bar{x})$

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

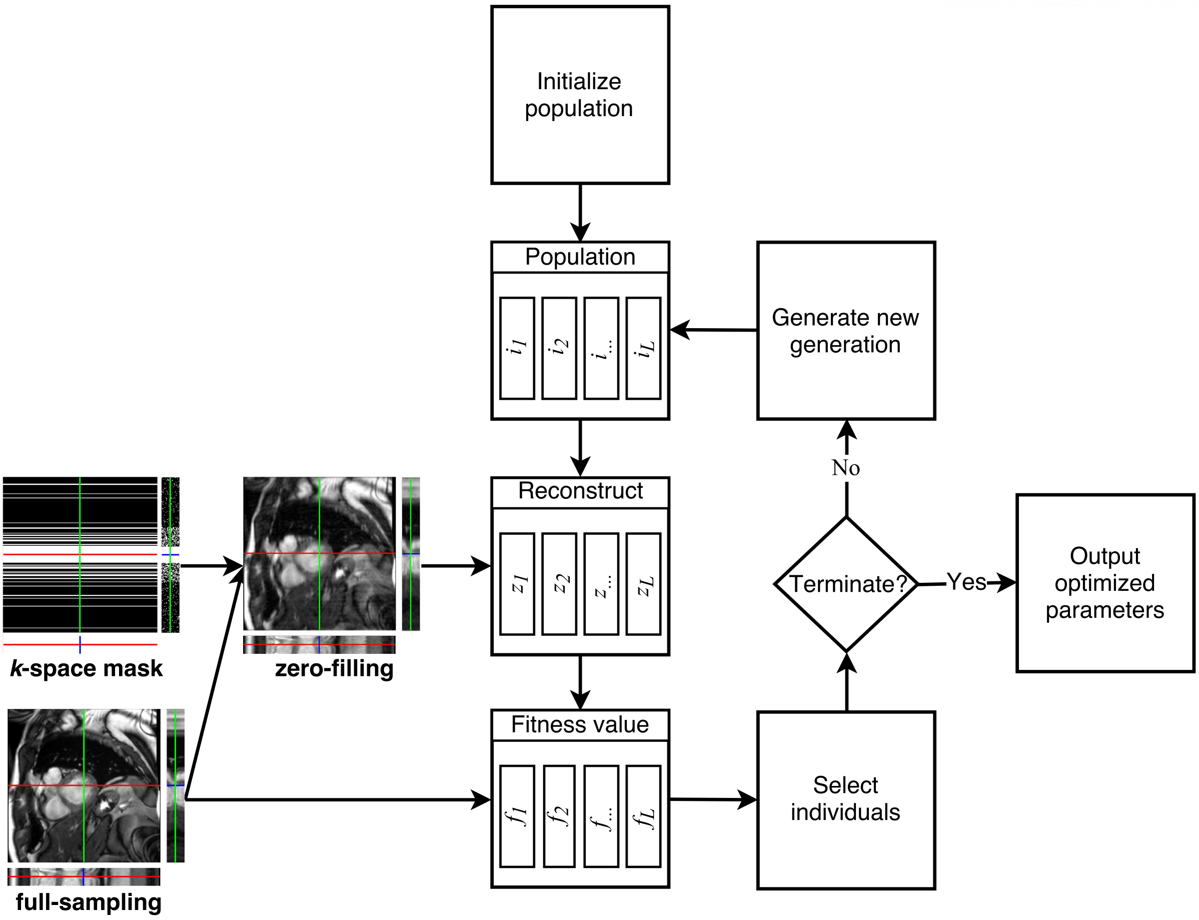

Initialize population
Population
i_1
i_2
$i_{...}$
i_L
Generate new generation
No
k-space mask
zero-filling
Reconstruct
z_1
z_2
$z_{...}$
z_L
Terminate?
Yes
Output optimized parameters
full-sampling
Fitness value
f_1
f_2
$f_{...}$
f_L
Select individuals

Figure 6: Genetic algorithm for automatically searching optimal parameters.

	α	γ	λ_1	λ_2	ρ	σ	θ	D_0	n
Lower	0.001	0.001	0.001	0.001	1	1	0.001	1	1
Upper	5	5	5	5	100	100	1	5	5

Table 2: Lower and upper bounds of parameters used in our experiment.

which are outputs of the reconstruction process (Fig. 1 yellow box) after 100 iterations.

$$f = -PSNR + \tau\bar{x} \tag{31}$$

where τ is the weight for trade-off between 2 values $PSNR$ (bigger is better) and average of x (smaller is better). Moreover, we also set possible lower and upper bounds for each parameter to narrow down the GA's searching space, as illustrated in Table 2. Specially, our experiment shows that a GA can be applied only once to find the optimal parameters that work for similar types of data (i.e., cardiac or DCE MRI data).

V Experiment Results

To assess the performance of our method, we conducted experiments on 15 cardiac MRI datasets from The Data Science Bowl [41] (30 frames of a 256×256 image across the cardiac cycle of a heart for each dataset), 7 tumor DCE MRI datasets (128×128 image with 128 frames of each dataset) and a multi-coil MRI brain dataset [42] (128×128 with 12 frames in which each frame consists of 12 parallel acquisitions) in three different rates of Cartesian k-space undersampling masks (12.5%, 25% and 50% sampling masks), as shown in Fig. 7c. GA is also set to run only one time to search parameters in five generations with $L = 200$ and $\tau = 20$. More specifically, for each kind of dataset, we arbitrarily select one full-sampled data and the sampling mask to optimize parameters, and reuse these found parameters to reconstruct the remaining of MRI datasets. We compare our method with several state-of-the-art methods, including k-t FOCUSS [10], FTVNNR [23], BCS [13], Caballero et al. [15], and 3D CSC [19]. We also compare with the intermediate versions of our method (i.e., incrementally adding new features, such as multi-scale extension of CSC, elastic net regularization, and spectral decomposition, to the baseline version of 3D CSC [19] to assess how each addition of feature affects the overall performance of the method (see Table 3). The proposed prototype system is implemented using MATLAB 2017a with GPU support.

Method	Abbreviation
k-t FOCUSS [10]	k-t FOCUSS
Algorithm using low-rank and total variation regularizations [23]	FTVNNR
Blind compressive sensing [13]	BCS
DLTG [15]	DLTG
3D CSC [19]	3D-CSC
Multi-scale 3D CSC	Multi-scale
Multi-scale 3D CSC with l_1, l_2 regularization	Elastic multi-scale
Elastic multi-scale and temporal TV minimization	Elastic multi-scale tv
Elastic multi-scale on high-frequency, and temporal TV minimization on low-frequency	Our method

Table 3: CS-MRI reconstruction methods compared in our experiment.

5.1 Reconstruction quality evaluation

For a fair comparison, we set up all DLTG [15] and CSC methods with the same number of filters (27 filters). All multi-scale 3D CSC methods used three different filter sizes ($N = 3$) and each size contained nine convolutional filters ($K = 9$). The box-plots below illustrate achievements of Mean Square Errors ($MSEs$), Peak Signal-To-Noise-Ratios ($PSNRs$) and Structural SIMilarities ($SSIMs$) for cardiac and tumor DCE MRIs; on each box, the central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. As can be seen in Fig. 8 and 9, our method gives a higher $PSNR$, $SSIM$ and a lower MSE compared to the other approaches in case of 12.5%, 25%, 50% sampling rates. Fig. 10 and 11 show a qualitative comparison of various reconstruction results. It is shown that our method generates less visual artifact compared to other methods (see the region of interests and pixel-wise error maps), even under an extremely low sampling rate (12.5%). It is also worth noting that our method can reconstruct temporal changes in DCE data more accurately compared to k-t FOCUSS and DLTG at a very low sampling rate (see Fig. 9 12.5%).

The Ak$_{ep}$ value, borrowed from the [43]'s model, is also a commonly used quality metric for dynamic MRI (especially DCE) to judge the consistency of the reconstructed image sequence because this value reflects the degree of MRI signal enhancement and the exchange rate in term of brightness and the contrast. Therefore, Ak$_{ep}$ value characterizes the velocity of MRI signal change in the region of interest (ROI) of the tumor, which is shown to provide the relevant information regarding tumor perfusion and permeability. To assess how the proposed reconstruction method performs, we measure the Ak$_{ep}$ on each reconstructed image (actually, Ak$_{ep}$ is generated per pixel) and generate a least-square fitting curve. By comparing the curve's shape, we can verify the reconstruction method's efficacy. The Fig. 12a shows Ak$_{ep}$ values as a color map of Tumor 1 dataset in Fig. 7b where our result clearly preserves details much better than the others and Fig. 12b and 12c shows the curve fitting profiles at some locations (marked A and B in Fig. 12a). As illustrated, our curve fitting profiles closely approximate the ground truth (full signal). In addition, our method effectively reduces the intensity variation (or noise) of the reconstructed images (see the variation of red circles are much smaller than that of other results) due to enforcing temporal coherence using a TV energy.

Overall, the proposed method achieves better reconstruction quality than other state-of-the-art methods. Shift-invariant convolution filters can represent both spatial and temporal features well, whilst multi-scale 3D CSC with l_1 and l_2 regularization shows its performance with better MSE, $PSNR$ and $SSIM$. Moreover, our frequency-splitting reconstruction approach, using temporal TV minimization for low-frequency and multi-scale 3D CSC with elastic net regularization for high-frequency, can significantly improve the image quality as well as convergence rate, which will be discussed in Section 5.4.

5.2 Extension to multi-coil parallel MR

Although we introduced our method for single coil MRI data in the previous sections, the proposed method can be applied to multi-coil MRI data as well with a small modification. For this, the main objective function (Eq. 7) can be combined with SENSE in the k-t SPARSE-SENSE framework [44] by multiplying the coil sensitivity matrix E after the undersampled

Fourier transform as shown in Eq. 32 below:

$$\begin{aligned}\min_{d,x,s_e^h,s_e^l} \frac{\alpha}{2}\left\| s_e^h - \sum_n^N \sum_k^K d_{n,k} * x_{n,k} \right\|_2^2 \\ + \lambda_1 \sum_n^N \sum_k^K \left\| x_{n,k} \right\|_1 + \frac{\lambda_2}{2} \sum_n^N \sum_k^K \left\| x_{n,k} \right\|_2^2 + \theta \left\| \nabla_t s_e^l \right\|_1 \\ s.t. : \| R\mathcal{F}_2(s_e^h + s_e^l) - m \|_2^2 \ < \ \epsilon^2, \| d_{n,k} \|_2^2 \leq 1\end{aligned} \tag{32}$$

where $s_e^h = Es^h$ and $s_e^l = Es^l$.

We tested the parallel version of our method on a 12-coil MRI brain dataset, and compared with FTVNNR, BCS and 3D-CSC methods for image quality assessment. For all sampling rates we tested (12.5%, 25% and 50%), our method outperformed the other methods (i.e., higher PSNR of reconstructed images, see Fig. 13a and Fig. 13b). Fig. 14 also shows that our method is less prone to pixel-wise errors and generates images closer to the full reconstruction than FTVNNR, BCS and 3D-CSC.

5.3 Robustness to noise

Noise resiliency is another important quality measure for the MR image reconstruction algorithm. To assess the noise resiliency of the proposed method, we synthetically corrupted the undersampled k-space data with Gaussian random noise and compared the reconstruction quality. In this experiment, we used cardiac and DCE MRI datasets at three Cartesian sampling rates, corrupted with Gaussian white noise in various levels (the range of standard deviation of Gaussian is $\{0.01, 0.03, 0.05, 0.07, 0.09, 0.1\}$), as shown in Fig. 7c. We compared the reconstruction results with those of three representative methods (FTVNNR [23], BCS [13] and 3D-CSC [19]) because they represent general transformation (non-learning approach), patch-based dictionary learning, and convolutional sparse coding, respectively. The optimal parameters found in the previous experiments (Sec. 5.1) are used for all methods in this experiment. As shown in Fig. 15, 16 and 17, our method produced the results with much higher PSNR, less noise, and superior image quality compared to the other methods.

5.4 Convergence evaluation

Fig. 18 illustrates the convergence of our method over 100 iterations. The left column contains full low-frequency, full high-frequency, full reconstruction and total time-axis gradient value of full low frequency, whilst right column gives the changing of refined output with its low- and high-frequency as well as the convolutional filters and temporal total variation value. In Fig. 18, the features in high-frequency part are progressively refined along with converging of convolutional filters every twenty epochs. The temporal total variation value in low-frequency also reduces over iterations which enforces the coherence of time-axis. As seen in Fig. 19 and 20, using temporal TV minimization is clearly helpful for the reconstruction process, while our frequency splitting approach can significantly improve the quality with comparable convergence rate. The proposed method converges faster (i.e. require less number of epochs) than 3D-CSC, multi-scale and elastic multi-scale. Note that patch-based dictionary (DLTG [15]) usually converges faster, but the actual running time is much slower because there is no GPU acceleration as in our method (see Table 4). Moreover, Fig. 19 and 20 show adequate convergence curves of the CSC methods with a large $PSNR$ improvement and no divergence. Thus, GA can find optimal parameters for the reconstruction process. In our observation, when we used the searched parameters for all datasets in the same kind of MRIs, the efficiency of convergence rate still remains, as shown in Fig. 20.

5.5 Running time evaluation

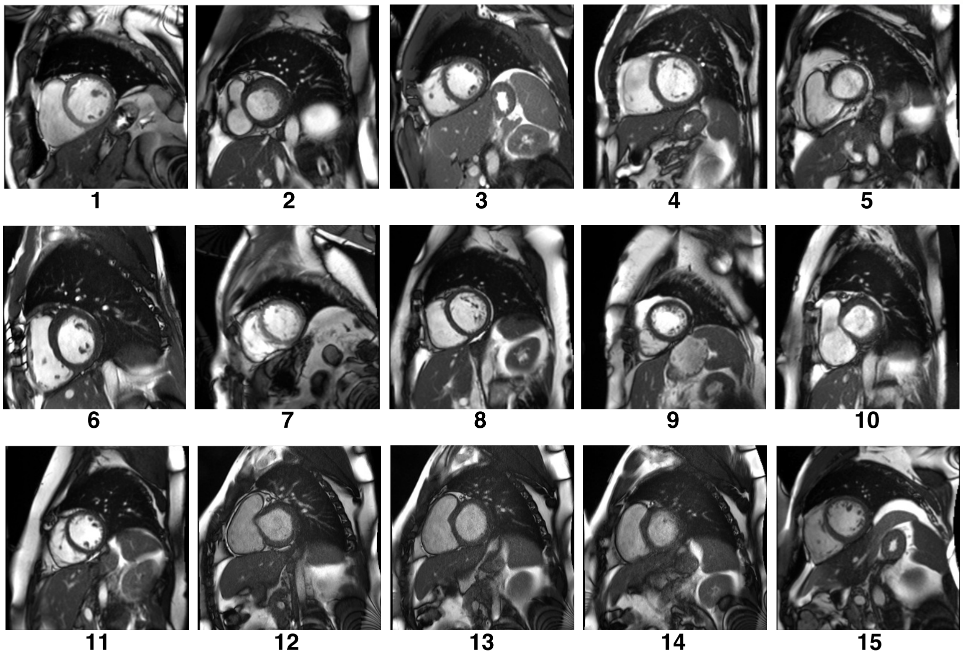
We measured the average reconstruction time of methods on a PC equipped with an Intel i7-7700K CPU and a NVIDIA Titan X GPU. All CSC approaches and DLTG [15] were run for 200 epochs, as shown in Fig. 19 and 20; however, the rate of convergence varied over methods. For example, our method converged in about 130 epochs, while DLTG took about 40 epochs (see Fig. 19) and 3D-CSC took more than 185 epochs (see Fig. 20). As shown in Table 4, the average per-epoch running time of our method on a GPU is 36.9×, 84×, 5×, and 16.5× faster than BCS, DLTG, 3D-CSC (CPU) and the CPU version of our method, respectively. A similar performance gain of our method was observed on DCE MRI data and multi-coil brain MRI date as well. Furthermore, DLTG converged faster (i.e., small number of epochs) than our method, but the actual running time of our method on a GPU to reach the convergence was significantly faster than that of DLTG (speed up of 25.58× for cardiac and 23× for DCE data). However, due to the computational overhead of our method (e.g., frequency filtering and energy optimization in low- and high-frequency), single-scale 3D CSC on a GPU was faster than our method. We also expect a significant performance improvement using NVIDIA CUDA and C/C++ over the current implementation using MATLAB, which is left for the future work.

	Cardiac dataset (second)		DCE dataset (second)		Brain dataset (second)	
	per epoch	total time	per epoch	total time	per epoch	total time
k-t FOCUSS (CPU)	_	177.6	_	256.2	_	_
FTVNNR (CPU)	_	131.1	_	151.5	_	65.1
BCS (CPU)	41.05	410.5	54.01	540.1	29.1	290.1
DLTG (CPU)	66.12	2644.8	76.97	3078.8	_	_
3D-CSC (CPU)	3.94	728.9	5.86	1084.1	2.875	474.38
Our method (CPU)	12.96	1684	18.35	2385.5	6.87	1099.2
3D-CSC (GPU)	0.503	93.06	0.728	134.68	0.361	57.76
Our method (GPU)	0.787	102.31	1.029	133.77	0.403	64.5

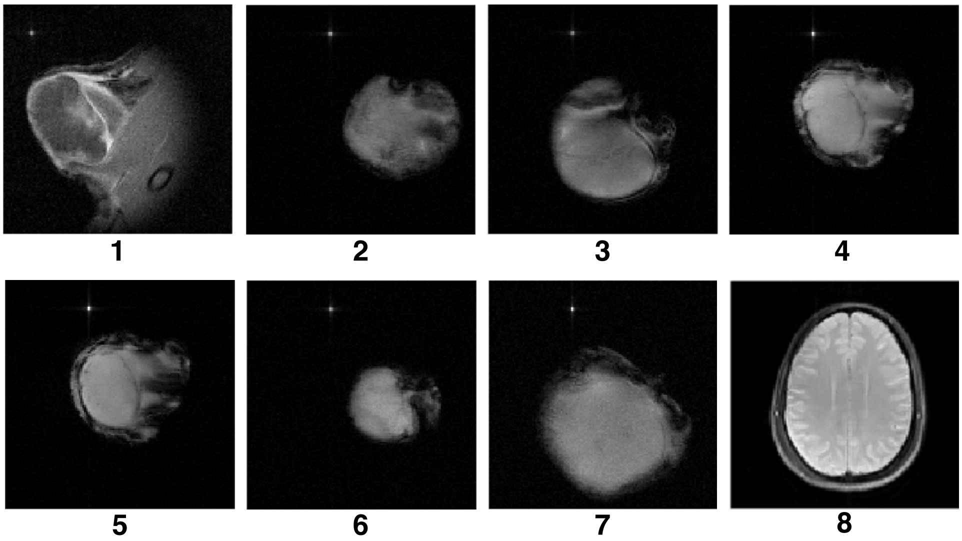
Table 4: Average running time of various reconstruction methods (per epoch: average running time of one epoch; total time: average running time to reach the convergence).

VI Conclusion and Future Work

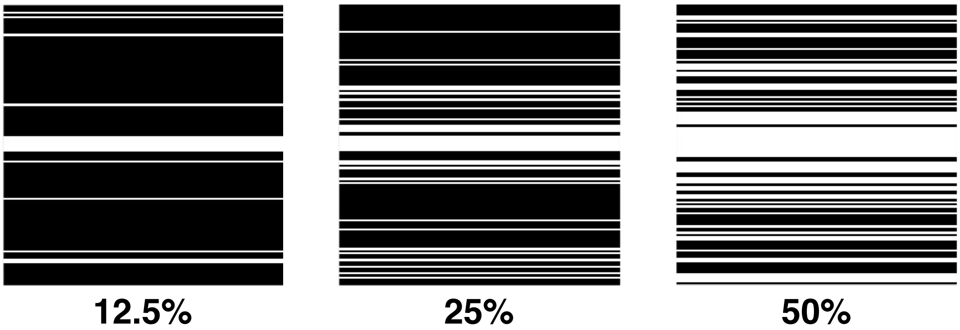
In this thesis, we introduced a novel CS-MRI reconstruction workflow based on an unsupervised learning approach. We discovered that the frequency-dependent reconstruction, i.e., using temporal total variation for low-frequency reconstruction and multi-scale 3D CSC with elastic net regularization for high-frequency reconstruction, played an important role for improving the overall reconstruction quality and the convergence rate. We also analyze computational complexity of our reconstruction algorithm. The proposed automatic parameter searching method using GA significantly reduced user's effort for tuning parameters in the reconstruction process. Furthermore, we showed that the proposed method can be easily extended to parallel MR imaging. The results showed that the proposed method outperformed the state-of-the-art CS-MRI reconstruction methods, such as k-t FOCUSS [9, 10], FTVNNR [23], single-scale 3D CSC [19], blind compressive sensing [13] and patch-based dictionary learning [14, 15], in terms of image quality and noise resiliency. In the future, we plan to improve the speed of proposed method by leveraging the parallel computing technology, such as multi-GPU and cluster systems. Exploring advanced meta-heuristic algorithms to improve the parameter selection process is another interesting future research direction.


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15

(a) Cardiac datasets

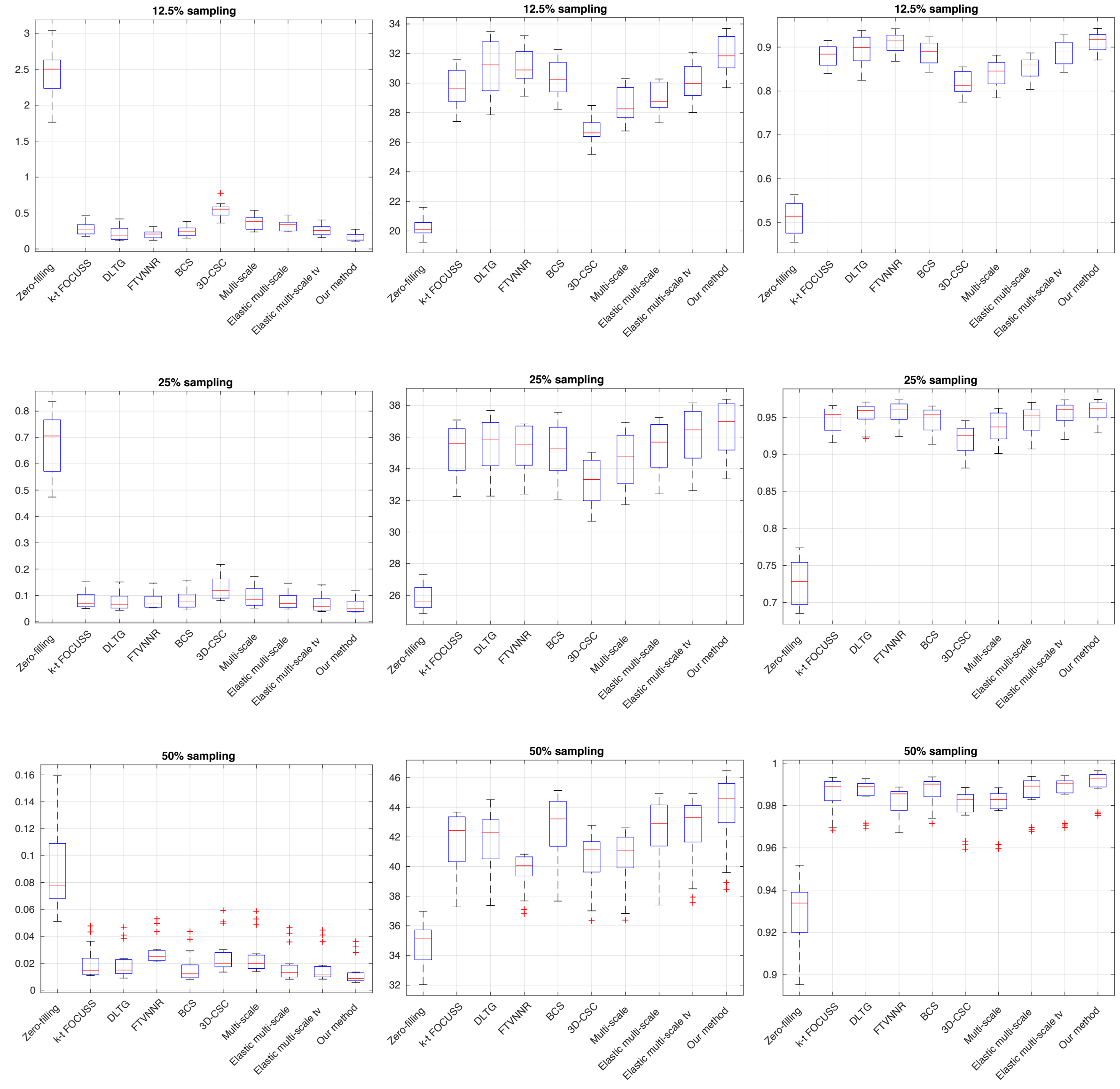


(b) Tumor DCE (1-7) and multi-coil brain (8) datasets



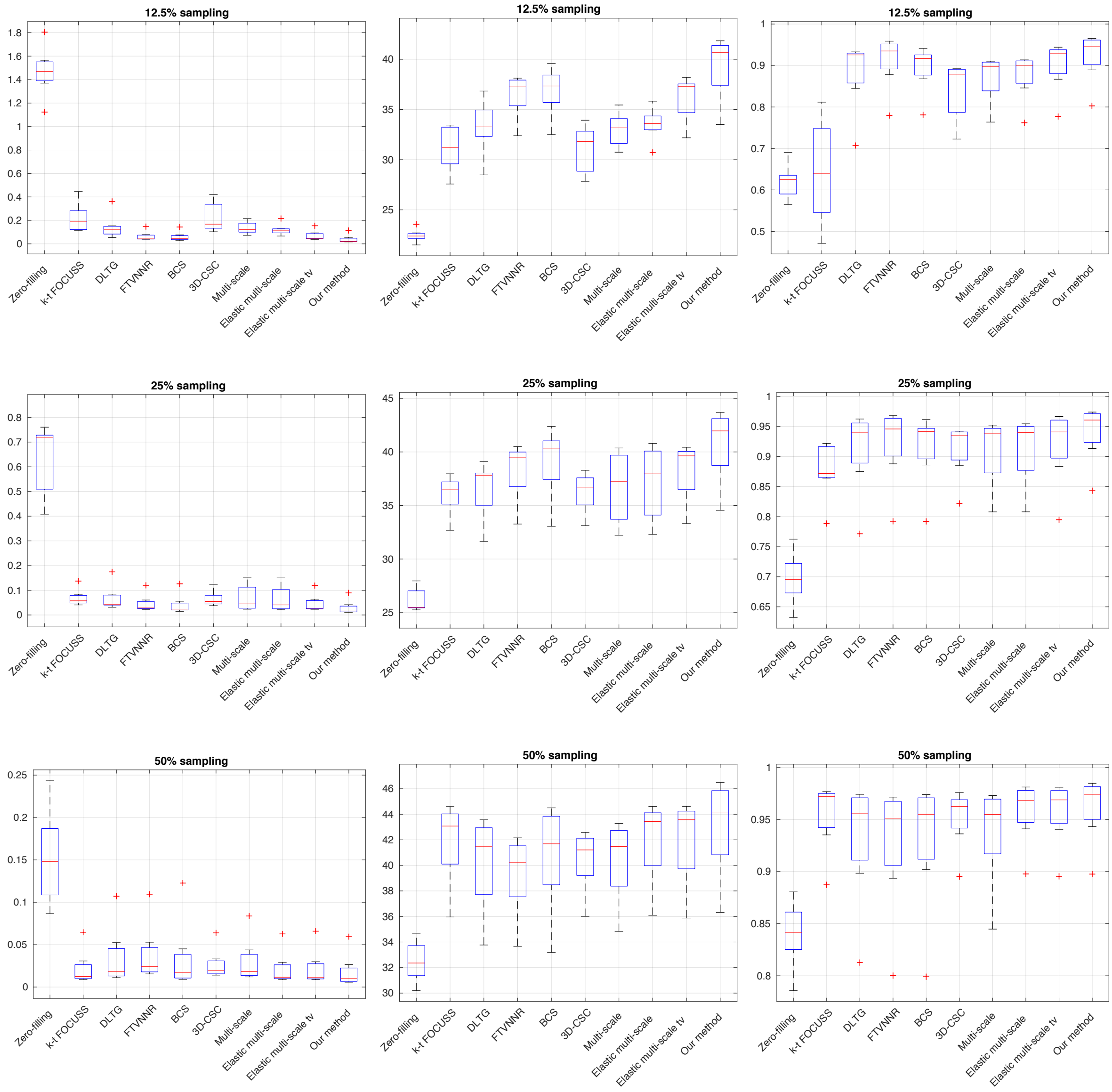
(c) Undersampling k-space masks

Figure 7: MRI datasets and Cartesian undersampling masks used in our experiments



(a) MSEs (b) PSNRs; unit:dB (c) SSIMs

Figure 8: MSEs, PSNRs, and SSIMs evaluation in three different sampling rates on cardiac MRIs (first column: MSEs; second column: PSNRs; third column: SSIMs).



(a) MSEs (b) PSNRs; unit:dB (c) SSIMs

Figure 9: MSEs, PSNRs, and SSIMs evaluation in three different sampling rates on tumor DCE MRIs (first column: MSEs; second column: PSNRs; third column: SSIMs).

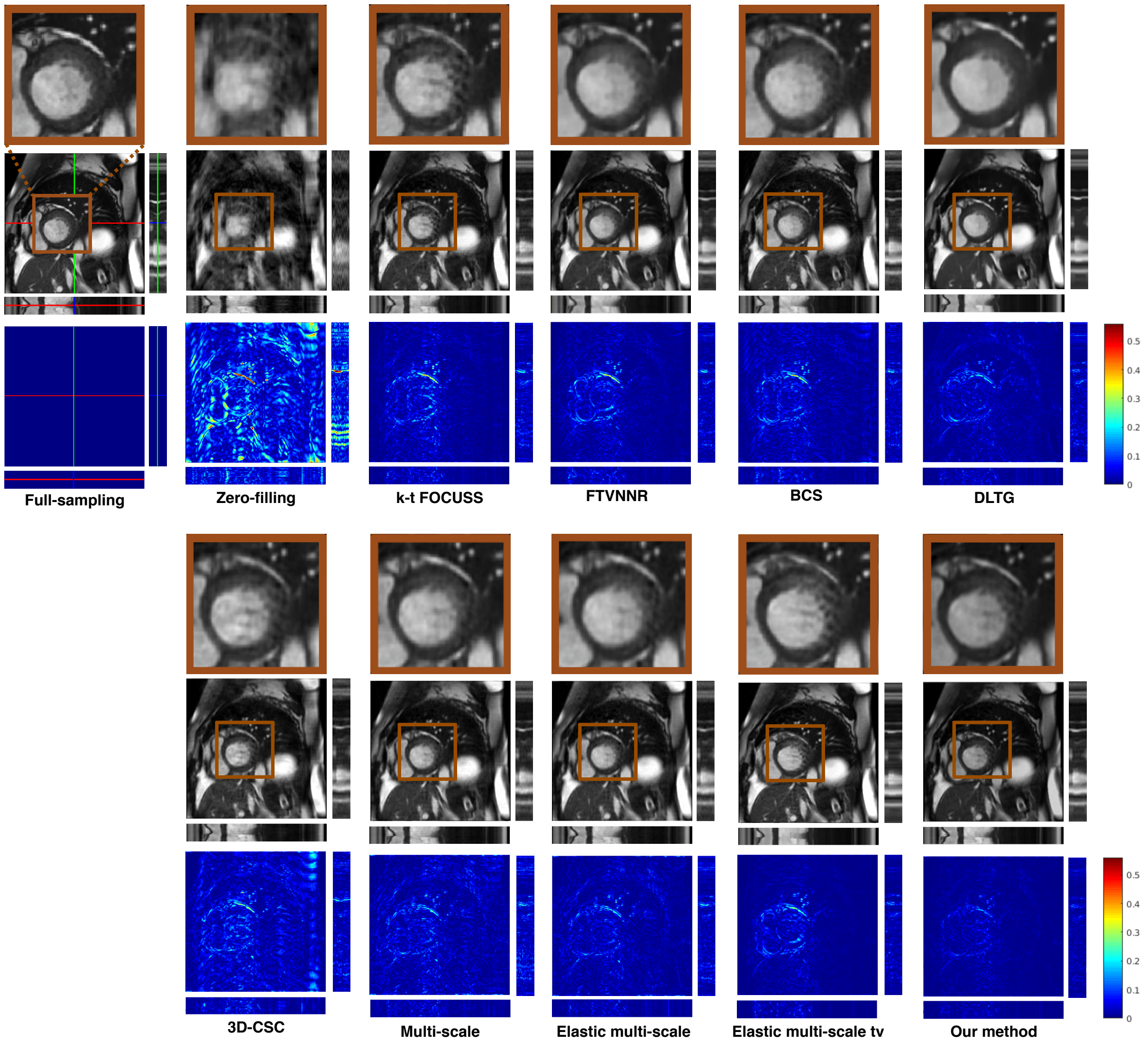

Full-sampling
Zero-filling
k-t FOCUSS
FTVNNR
BCS
DLTG
0.5
0.4
0.3
0.2
0.1
0
3D-CSC
Multi-scale
Elastic multi-scale
Elastic multi-scale tv
Our method
0.5
0.4
0.3
0.2
0.1
0

Figure 10: Image quality and pixel-wise error comparison of various reconstruction methods on Cardiac 2 dataset (12.5% sampling)

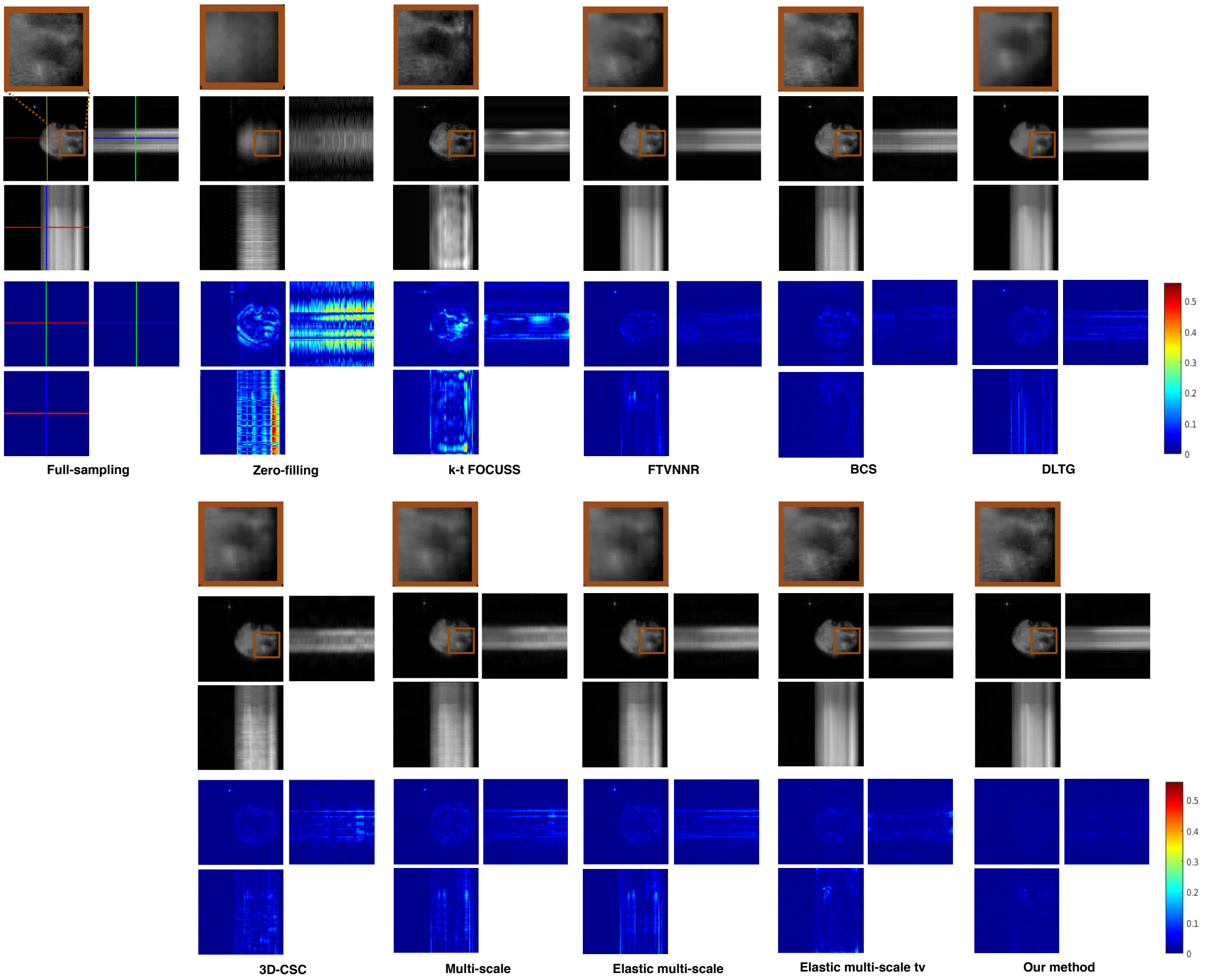
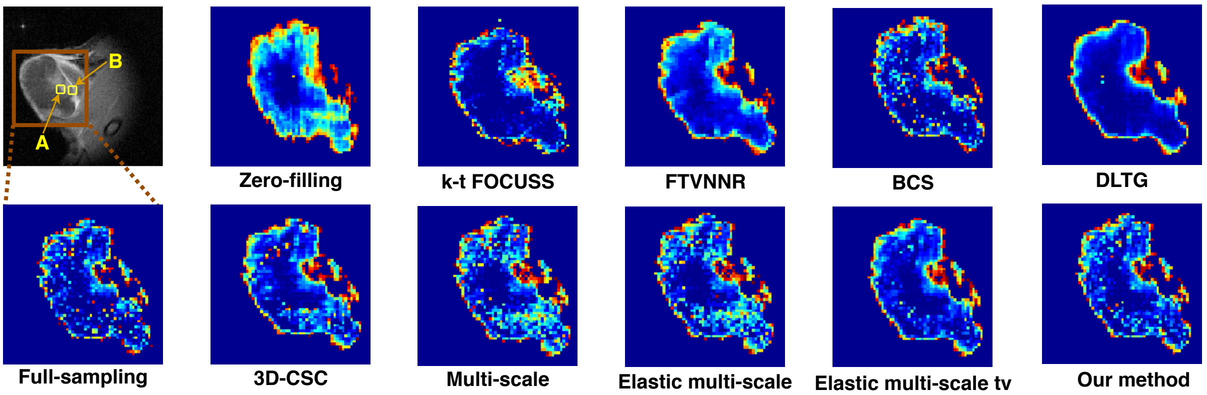
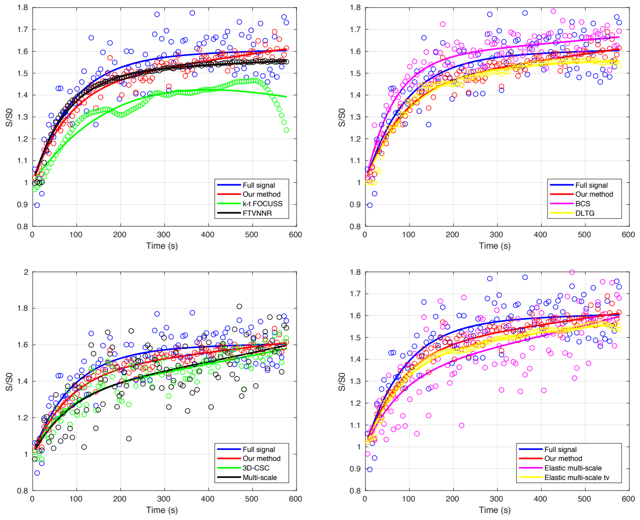

Full-sampling
Zero-filling
k-t FOCUSS
FTVNNR
BCS
DLTG
0.5
0.4
0.3
0.2
0.1
0
3D-CSC
Multi-scale
Elastic multi-scale
Elastic multi-scale tv
Our method
0.5
0.4
0.3
0.2
0.1
0

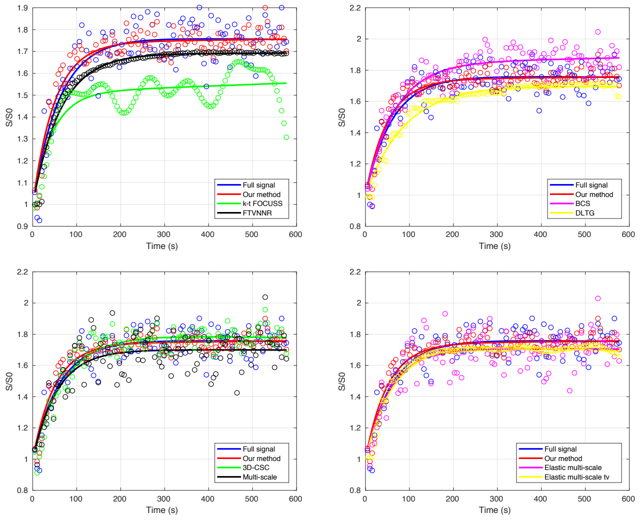
Figure 11: Image quality and pixel-wise error comparison of various reconstruction methods on Tumor 2 dataset (12.5% sampling)


A
B
Zero-filling
k-t FOCUSS
FTVNNR
BCS
DLTG
Full-sampling
3D-CSC
Multi-scale
Elastic multi-scale
Elastic multi-scale tv
Our method

(a) Ak$_{ep}$ maps of 25% sampling reconstruction of Tumor 1 dataset

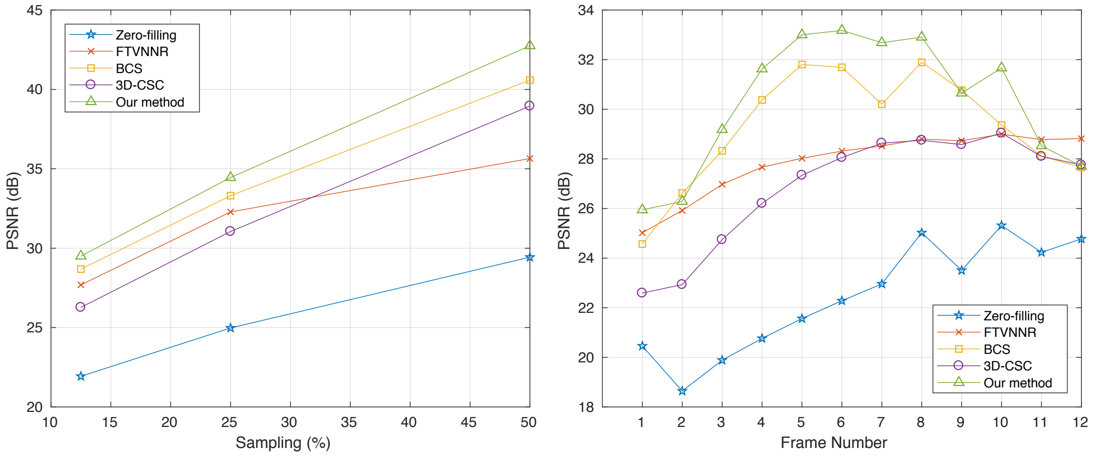


(b) Temporal profile of A



(c) Temporal profile of B

Figure 12: Reconstruction quality comparison of Tumor 1 dataset



(a) Average PSNRs of all frames (b) PSNRs of frames at 12.5% mask

Figure 13: PSNRs evaluation on the multi-coil MRI brain dataset

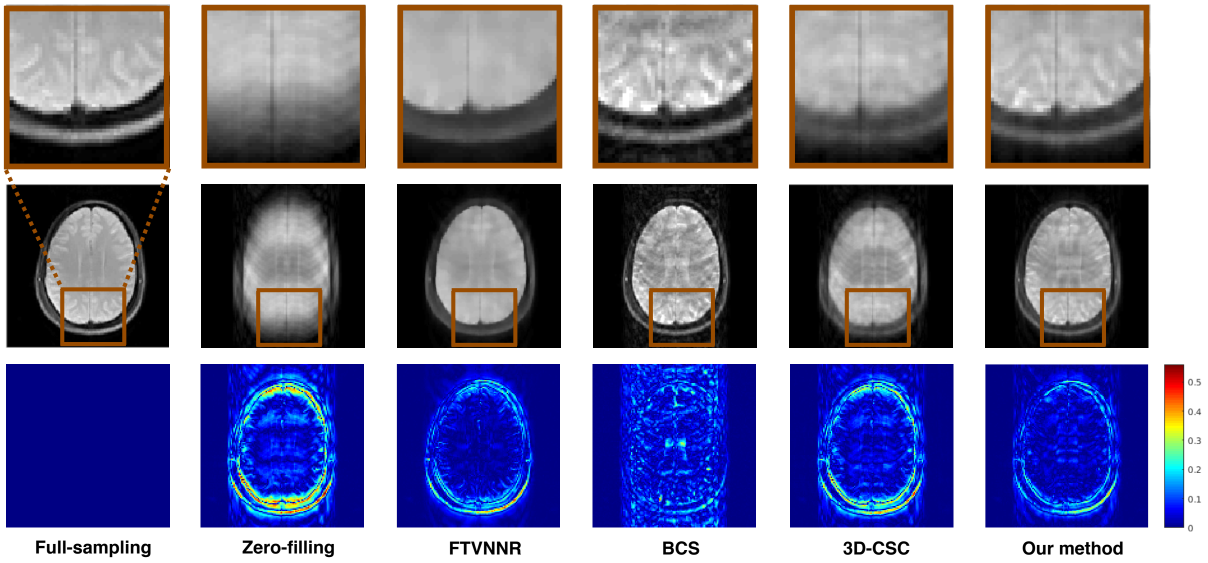
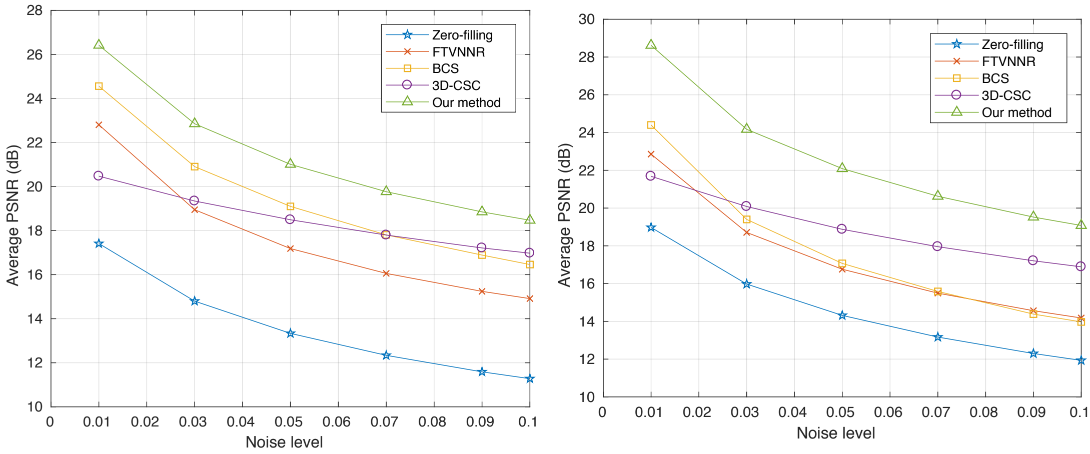

Full-sampling
Zero-filling
FTVNNR
BCS
3D-CSC
Our method
0.5
0.4
0.3
0.2
0.1
0

Figure 14: Image quality and pixel-wise error comparison of multi-coil reconstruction methods on the brain dataset (12.5% sampling)



(a) Cardiac MRI

(b) DCE tumor MRI

Figure 15: Average PSNRs of the images reconstructed from the various noise-corrupted k-space data

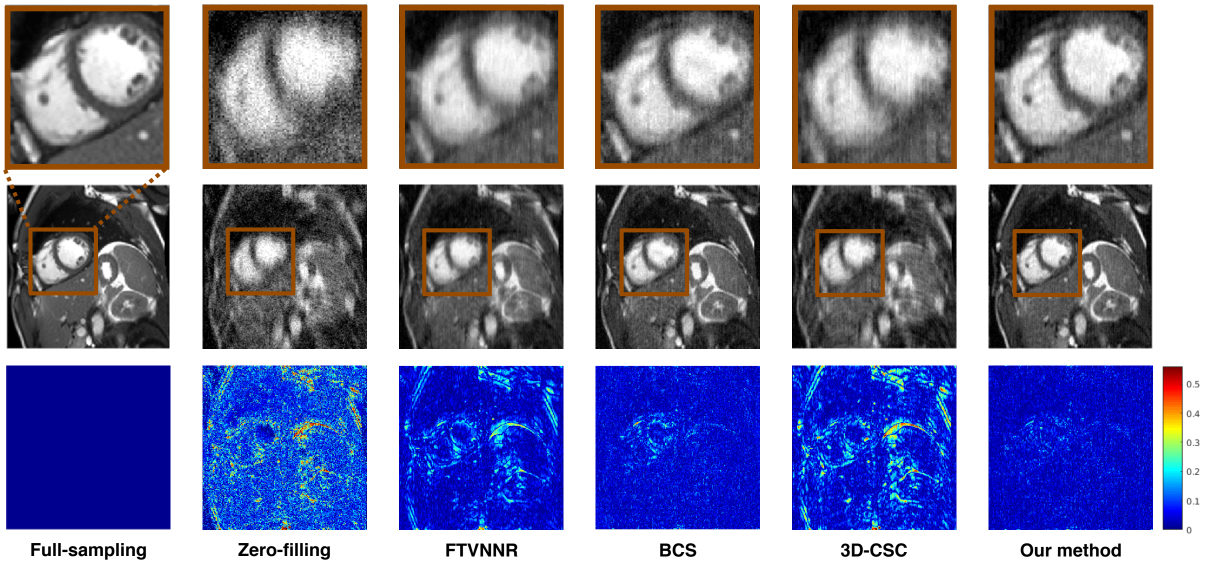

Full-sampling
Zero-filling
FTVNNR
BCS
3D-CSC
Our method
0.5
0.4
0.3
0.2
0.1
0

Figure 16: Reconstruction of noise-corrupted Cardiac 3 dataset (at noise level 0.01)

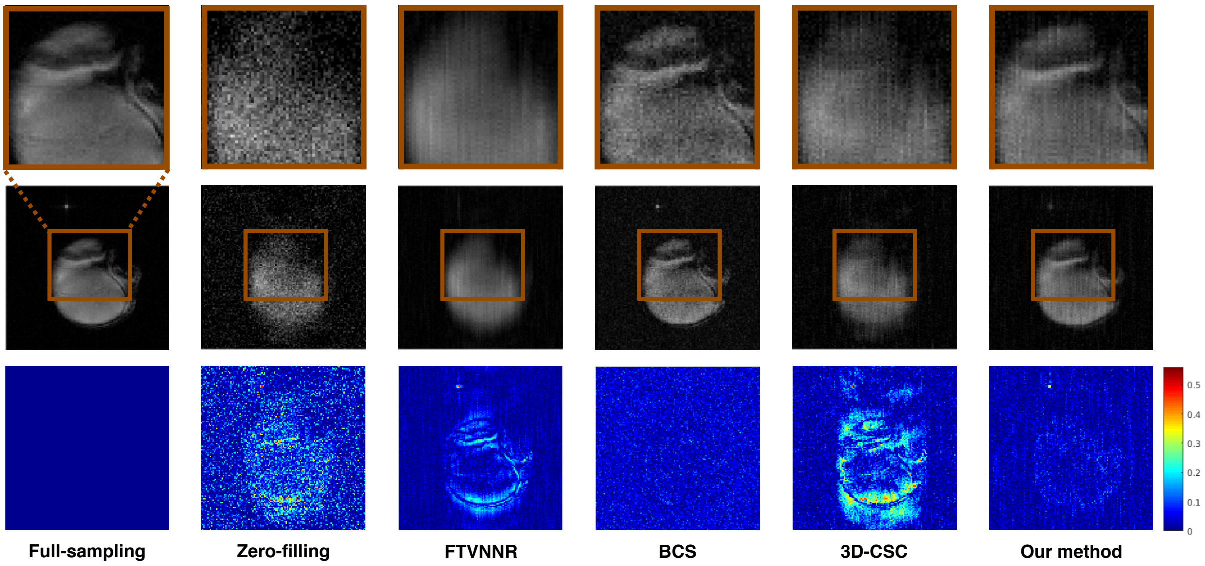

Full-sampling
Zero-filling
FTVNNR
BCS
3D-CSC
Our method
0.5
0.4
0.3
0.2
0.1
0

Figure 17: Reconstruction of noise-corrupted DCE tumor 3 dataset (at noise level 0.01)

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

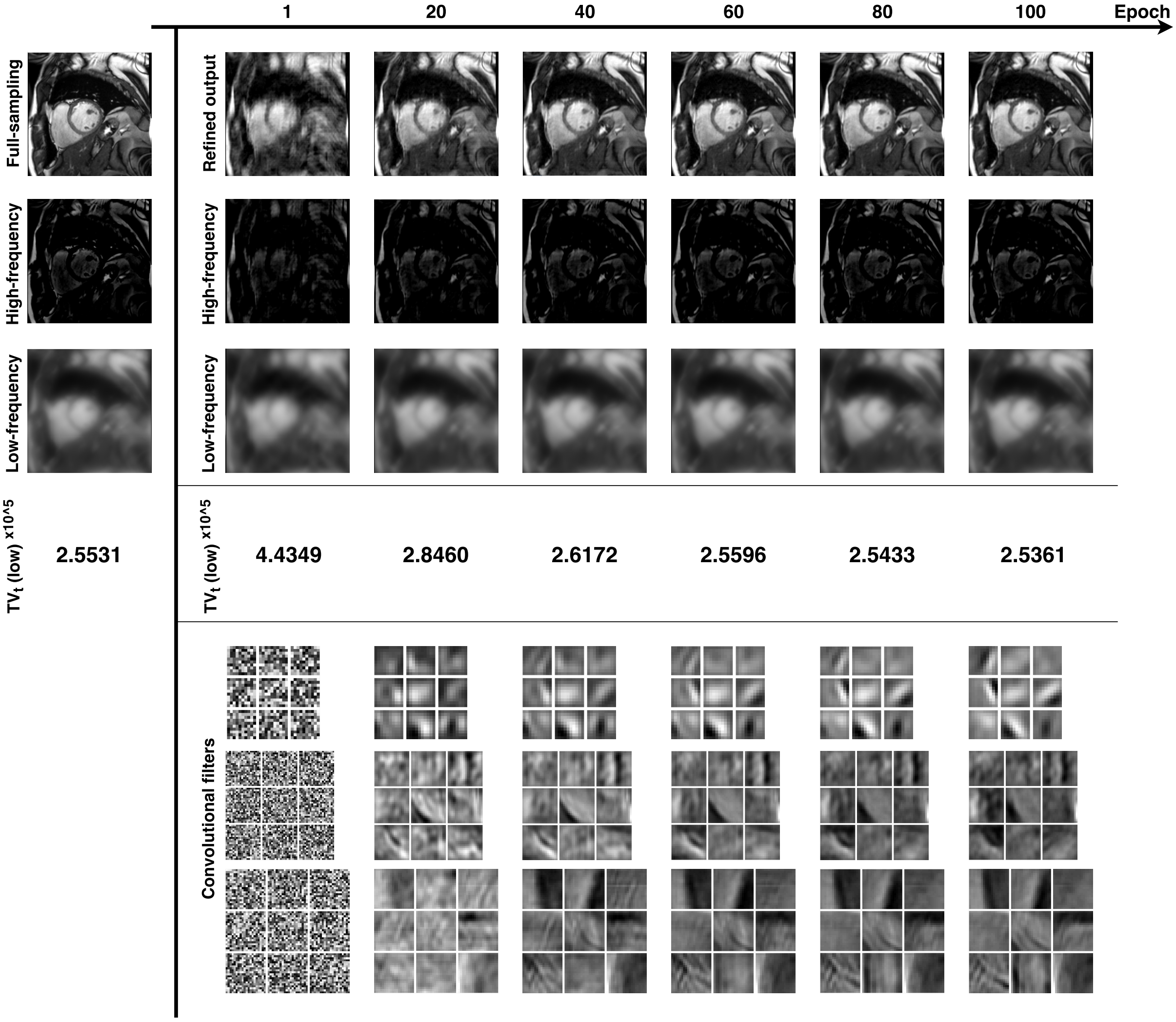
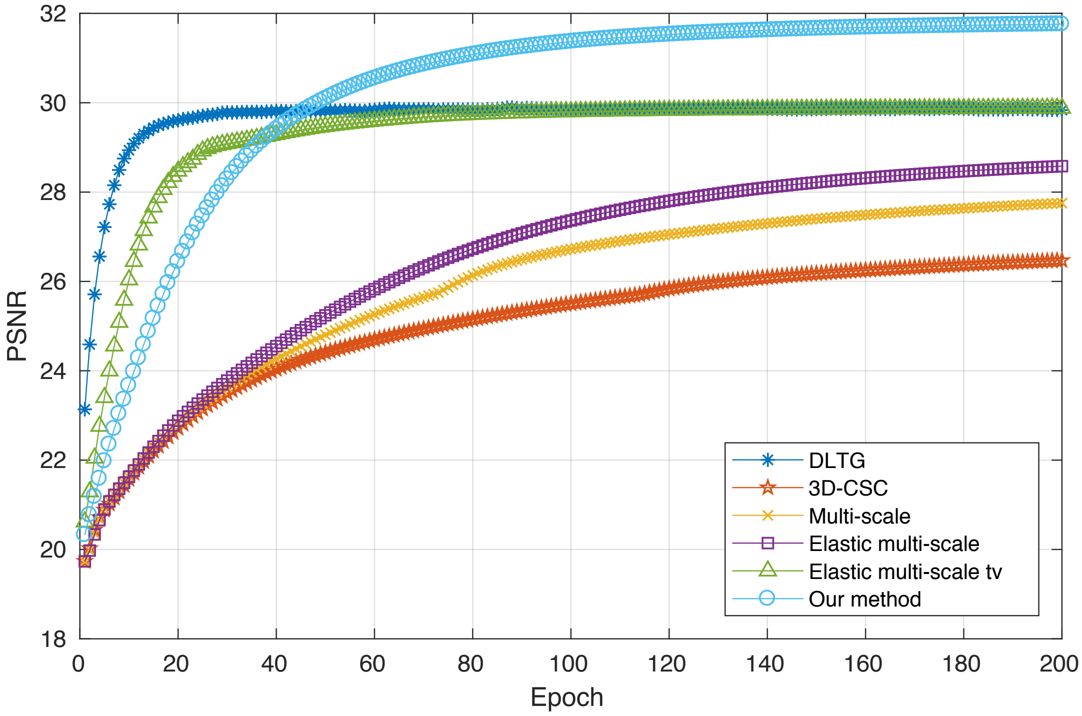
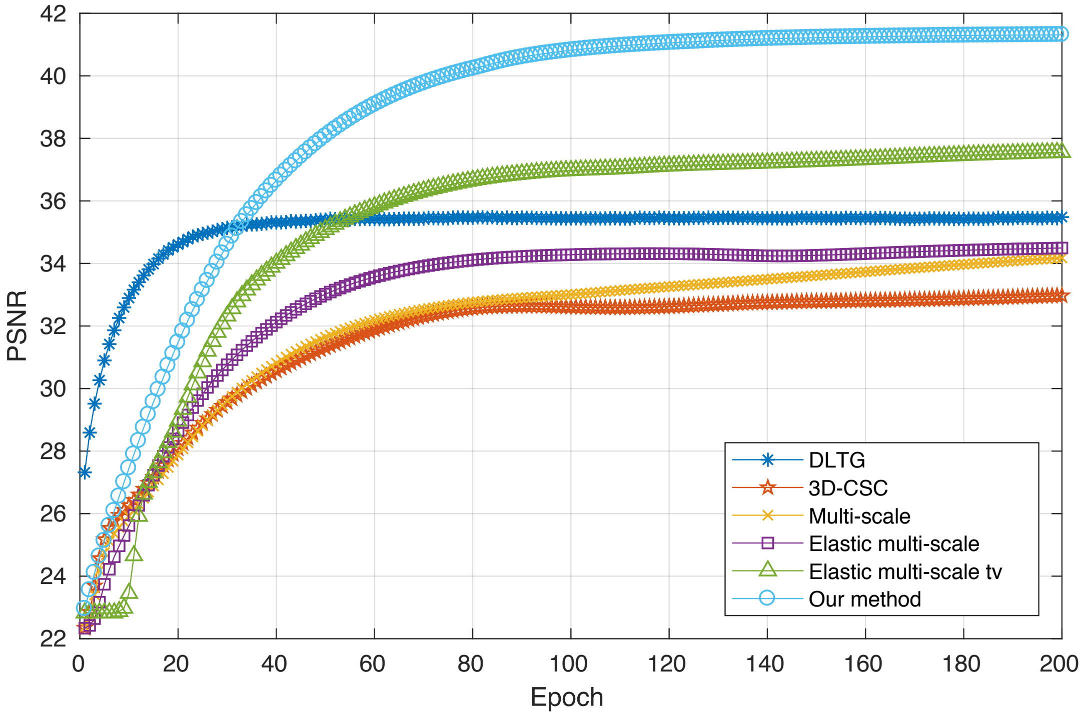

1
20
40
60
80
100
Epoch
Full-sampling
High-frequency
Low-frequency
TVt (low) x10^5
2.5531
Refined output
High-frequency
Low-frequency
TVt (low) x10^5
4.4349
2.8460
2.6172
2.5596
2.5433
2.5361
Convolutional filters

Figure 18: Convergence of proposed method for reconstructing 12.5% sampling Cardiac 1 dataset over 100 epochs.

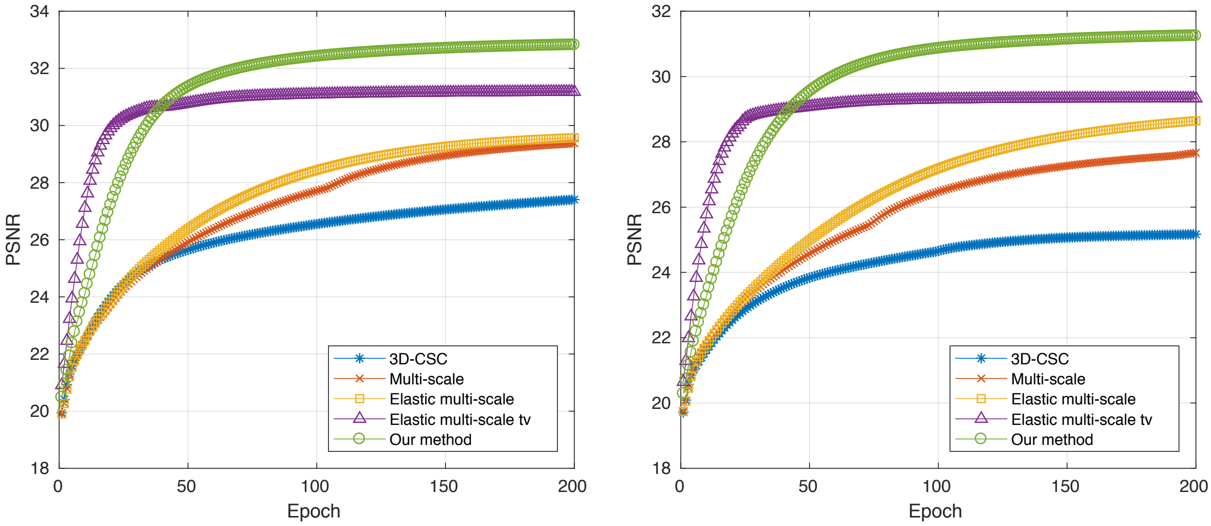


(a) Cardiac 2 dataset

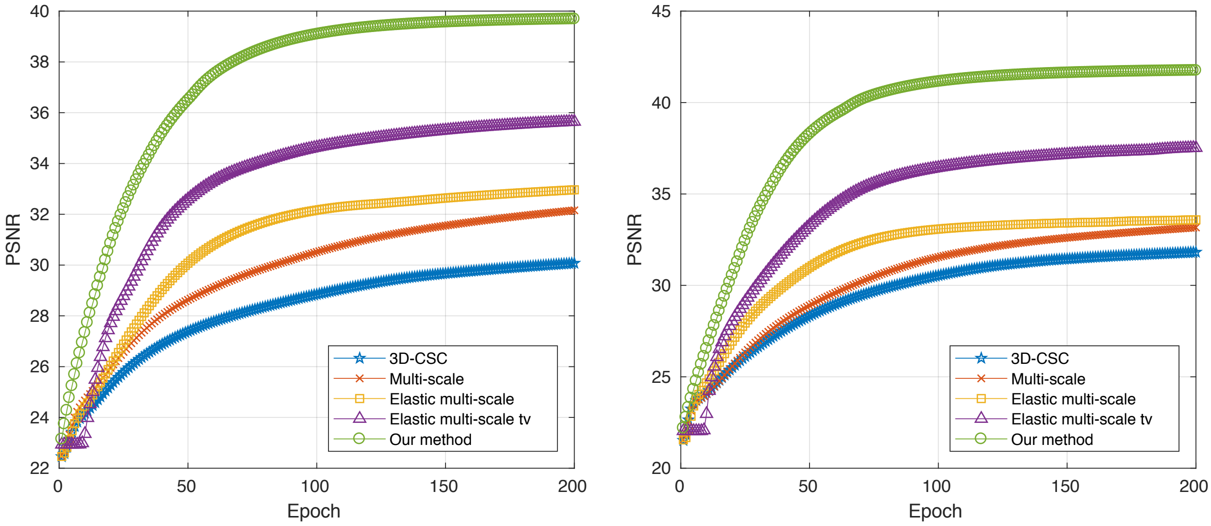


(b) Tumor 2 dataset

Figure 19: Convergence rate comparison of 12.5% sampling Cardiac 2 and Tumor 2 datasets in terms of PSNRs.



(a) Convergence curves of two different Cardiac 3 and 4 datasets



(b) Convergence curves of two different Tumor 3 and 4 datasets

Figure 20: Convergence curves of 12.5% sampling cardiac and tumor data in terms of PSNRs using the optimal parameters found by the genetic algorithm.

References

[1] R. H. El Khouli, K. J. Macura, P. B. Barker, M. R. Habba, M. A. Jacobs, and D. A. Bluemke, "Relationship of temporal resolution to diagnostic performance for dynamic contrast enhanced MRI of the breast," *Journal of Magnetic Resonance Imaging*, vol. 30, no. 5, pp. 999–1004, 2009.

[2] D. L. Donoho, "Compressed sensing," *IEEE Transactions on information theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[3] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, "Compressed sensing MRI," *IEEE signal processing magazine*, vol. 25, no. 2, pp. 72–82, 2008.

[4] C. E. Shannon, "Communication in the presence of noise," *Proceedings of the IRE*, vol. 37, no. 1, pp. 10–21, 1949.

[5] R. Otazo, L. Feng, H. Chandarana, T. Block, L. Axel, and D. K. Sodickson, "Combination of compressed sensing and parallel imaging for highly-accelerated dynamic MRI," in *Proceedings - International Symposium on Biomedical Imaging*, 2012, pp. 980–983.

[6] T. M. Quan, S. Han, H. Cho, and W.-K. Jeong, "Multi-GPU reconstruction of dynamic compressed sensing MRI," in *Proceeding of ISBI 2016.* Springer, 2015, pp. 484–492.

[7] S. Ma, W. Yin, Y. Zhang, and A. Chakraborty, "An efficient algorithm for compressed MR imaging using total variation and wavelets," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on.* IEEE, 2008, pp. 1–8.

[8] I. Daubechies, *Ten lectures on wavelets.* SIAM, 1992.

[9] H. Jung, J. C. Ye, and E. Y. Kim, "Improved k–t blast and k–t sense using focuss," *Physics in medicine and biology*, vol. 52, no. 11, p. 3201, 2007.

[10] H. Jung, K. Sung, K. S. Nayak, E. Y. Kim, and J. C. Ye, "*k-t* FOCUSS: A general compressed sensing framework for high resolution dynamic MRI," *Magnetic resonance in medicine*, vol. 61, no. 1, pp. 103–116, 2009.

[11] S. Ravishankar and Y. Bresler, "MR image reconstruction from highly undersampled k-space data by dictionary learning," *IEEE transactions on medical imaging*, vol. 30, no. 5, pp. 1028–1041, 2011.

[12] S. P. Awate and E. V. DiBella, "Spatiotemporal dictionary learning for undersampled dynamic MRI reconstruction via joint frame-based and dictionary-based sparsity," in *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on.* IEEE, 2012, pp. 318–321.

[13] S. G. Lingala and M. Jacob, "Blind compressive sensing dynamic MRI," *IEEE transactions on medical imaging*, vol. 32, no. 6, pp. 1132–1145, 2013.

[14] J. Caballero, D. Rueckert, and J. V. Hajnal, "Dictionary learning and time sparsity in dynamic MRI," in *Proceeding of MICCAI 2012.* Springer, 2012, pp. 256–263.

[15] J. Caballero, A. N. Price, D. Rueckert, and J. V. Hajnal, "Dictionary learning and time sparsity for dynamic MR data reconstruction," *IEEE transactions on medical imaging*, vol. 33, no. 4, pp. 979–994, 2014.

[16] H. Bristow, A. Eriksson, and S. Lucey, "Fast convolutional sparse coding," in *Proceedings of CVPR*, 2013, pp. 391–398.

[17] B. Wohlberg, "Efficient convolutional sparse coding," in *Proceeding of ICASSP*, 2014, pp. 7173–7177.

[18] T. M. Quan and W.-K. Jeong, "Compressed sensing reconstruction of dynamic contrast enhanced MRI using GPU-accelerated convolutional sparse coding," in *Proceeding of ISBI 2016.* IEEE, 2016, pp. 518–521.

[19] ——, "Compressed Sensing Dynamic MRI Reconstruction Using GPU-accelerated 3D Convolutional Sparse Coding," in *Proceeding of MICCAI*, 2016, pp. 484–492.

[20] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 67, no. 2, pp. 301–320, 2005.

[21] S. Luke, *Essentials of metaheuristics.* Lulu Raleigh, 2009, vol. 113.

[22] H. Mühlenbein, M. Schomisch, and J. Born, "The parallel genetic algorithm as function optimizer," *Parallel computing*, vol. 17, no. 6-7, pp. 619–632, 1991.

[23] J. Yao, Z. Xu, X. Huang, and J. Huang, "An efficient algorithm for dynamic MRI using low-rank and total variation regularizations," *Medical image analysis*, vol. 44, pp. 14–27, 2018.

[24] M. D. Zeiler, D. Krishnan, G. W. Taylor, and R. Fergus, "Deconvolutional networks," 2010.

[25] B. Wohlberg, "Efficient algorithms for convolutional sparse representations," *IEEE Transactions on Image Processing*, vol. 25, no. 1, pp. 301–315, 2016.

[26] M. Lustig, J. M. Santos, D. L. Donoho, and J. M. Pauly, "*k-t* SPARSE: High frame rate dynamic MRI exploiting spatio-temporal sparsity," in *Proceedings of the 13th Annual Meeting of ISMRM, Seattle*, vol. 2420, 2006.

[27] L. B. Montefusco, D. Lazzaro, S. Papi, and C. Guerrini, "A fast compressed sensing approach to 3D MR image reconstruction," *IEEE transactions on medical imaging*, vol. 30, no. 5, pp. 1064–1075, 2011.

[28] C. Chen, Y. Li, L. Axel, and J. Huang, "Real time dynamic mri by exploiting spatial and temporal sparsity," *Magnetic resonance imaging*, vol. 34, no. 4, pp. 473–482, 2016.

[29] T. Goldstein and S. Osher, "The split Bregman method for L1-regularized problems," *SIAM journal on imaging sciences*, vol. 2, no. 2, pp. 323–343, 2009.

[30] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[31] B. Trémoulhéac, N. Dikaios, D. Atkinson, and S. R. Arridge, "Dynamic MR Image Reconstruction–Separation From Undersampled (*k-t*)-Space via Low-Rank Plus Sparse Prior," *IEEE transactions on medical imaging*, vol. 33, no. 8, pp. 1689–1701, 2014.

[32] X. Miao, S. G. Lingala, Y. Guo, T. Jao, and K. S. Nayak, "Accelerated cardiac cine using locally low rank and total variation constraints," in *Proceedings of the 23rd Annual Meeting, Toronto, Canada*, 2015, p. 571.

[33] J. Yao, Z. Xu, X. Huang, and J. Huang, "Accelerated dynamic MRI reconstruction with total variation and nuclear norm regularization," in *Proceeding of MICCAI*. Springer, 2015, pp. 635–642.

[34] S. Ravishankar, B. E. Moore, R. R. Nadakuditi, and J. A. Fessler, "Low-rank and adaptive sparse signal (LASSI) models for highly accelerated dynamic imaging," *IEEE transactions on medical imaging*, vol. 36, no. 5, pp. 1116–1128, 2017.

[35] Z. Zhu, J. Yao, Z. Xu, J. Huang, and B. Zhang, "A simple primal-dual algorithm for nuclear norm and total variation regularization," *Neurocomputing*, vol. 289, pp. 1–12, 2018.

[36] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on signal processing*, vol. 54, no. 11, pp. 4311–4322, 2006.

[37] Y. Yang, F. Liu, W. Xu, and S. Crozier, "Compressed sensing MRI via two-stage reconstruction," *IEEE Transactions on biomedical engineering*, vol. 62, no. 1, pp. 110–118, 2015.

[38] Y.-C. Wu, H. Du, and W. Mei, “Filter-based compressed sensing MRI reconstruction,” *International Journal of Imaging Systems and Technology*, vol. 26, no. 3, pp. 173–178, 2016.

[39] C. Blum and A. Roli, “Metaheuristics in combinatorial optimization: Overview and conceptual comparison,” *ACM Computing Surveys (CSUR)*, vol. 35, no. 3, pp. 268–308, 2003.

[40] A. Chambolle, “Total variation minimization and a class of binary MRF models,” in *EMMCVPR*, vol. 5. Springer, 2005, pp. 136–152.

[41] Kaggle, “Data science bowl cardiac challenge data,” https://www.kaggle.com/c/second-annual-data-science-bowl/data, 2015.

[42] CBIG, “Computational Biomedical Imaging Group,” https://research.engineering.uiowa.edu/cbig/content/codes-blind-compressed-sensing-bcs-dynamic-mri, 2018.

[43] U. Hoffmann, G. Brix, M. V. Knopp, T. Heβ, and W. J. Lorenz, “Pharmacokinetic Mapping of the Breast: A New Method for Dynamic MR Mammography,” *Magnetic Resonance in Medicine*, vol. 33, no. 4, pp. 506–514, 1995.

[44] R. Otazo, D. Kim, L. Axel, and D. K. Sodickson, “Combination of compressed sensing and parallel imaging for highly accelerated first-pass cardiac perfusion MRI,” *Magnetic resonance in medicine*, vol. 64, no. 3, pp. 767–776, 2010.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my advisor Prof. Won-Ki Jeong for the continuous support of my Master study and related research.

Besides my advisor, I would like to thank the rest of my thesis committee: Prof. Jaesik Choi and Prof. Seungjoon Yang, for their insightful comments and also all professors who have taught me in UNIST.

My sincere thanks also goes to my fellow labmates and following institution friends for their supports and for all the memories we have had in the last two years.

Last but not the least, I would like to thank my family: my parents and to my brother and sisters for supporting me spiritually throughout writing this thesis and my life in general.