

# **Optimized Implementation of Multi-layer Convolutional Sparse Coding Framework For High Dimensional Data**



By

**Abdul Wahid Tareen**

Reg.No. 4-FET/PhDEE/F-16

**A dissertation submitted to I.I.U. in partial fulfillment  
of the requirements for the degree of**

**DOCTOR OF PHILOSOPHY**

**Department of Electrical Engineering**

**Faculty of Engineering and Technology**

**INTERNATIONAL ISLAMIC UNIVERSITY**

**ISLAMABAD**

**2021**

PhD

6213

ABO

Accession No TH25469

✓

Electrical engineering  
Power electronics  
Multi-layer

Copyright © 2021 by Abdul Wahid Tareen

All rights reserved. No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without the permission from the author.

DEDICATED TO

The memory of Professor Ijaz Mansoor Qureshi

# CERTIFICATE OF APPROVAL

**Title of Thesis:** Optimized Implementation of Multi-layer Convolutional Sparse Coding Framework for High Dimensional Data

**Name of Student:** ABDUL WAHID TAREEN

**Registration No:** 4-FET/PHDEE/F16

Accepted by the Department of Electrical Engineering, Faculty of Engineering and Technology, International Islamic University (IIU), Islamabad, in partial fulfillment of the requirements for the Doctor of Philosophy degree in Electronic Engineering.

**Viva voce committee:**

**Dr. Jawad Ali Shah (Supervisor)**

Assistant Professor, DEE, FET, IIU Islamabad.

**Dr. Adnan Umar Khan (Co-Supervisor)**

Assistant Professor DEE, FET, IIU Islamabad.

**Dr. Waseem Khan (Internal)**

Assistant Professor DEE, FET, IIU Islamabad.

**Dr. M.Usman (External-I)**

Ex-Director NESCOM , Islamabad

**Dr. Hammad Umar ( External-II)**

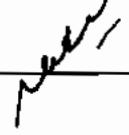
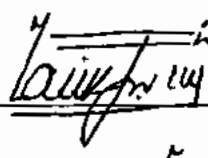
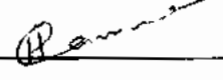
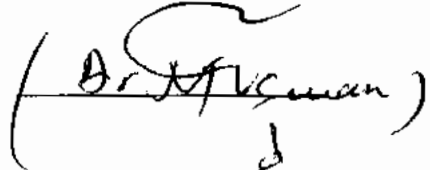
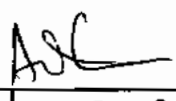
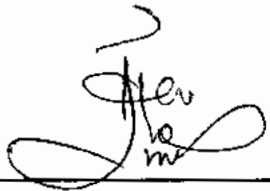
Tenured Associate Professor, COMSATS, Islamabad.

**Dr. Suheel Abdullah Malik (Chairman, DEE)**

Assistant Professor DEE, FET, IIU Islamabad.

**Prof. Dr. Nadeem Ahmad Sheikh (Dean, FET)**

Professor DME, FET, IIU Islamabad.



## ABSTRACT

Deep learning has been employed for different image processing problems in recent years. Although there is impressive performance of these architectures, their inner working largely remains heuristic. The celebrated sparse coding theory works on the premise of first representing data with a over-complete matrix called dictionary and the corresponding sparse maps employing pursuit algorithms for optimal solutions. This theory which has been extensively researched in last decade has been extended recently to theoretically model deep neural networks and extract insights about their inner working. The convolutional sparse coding (CSC) and the multi-layer convolutional sparse coding (ML-CSC) has been presented as theoretical equivalent to forward pass of convolutional neural networks (CNNs). The ML-CSC has been analyzed for convergence analysis and feature extraction of publicly available datasets for performance comparison with legacy CNN based classification frameworks. Although, this research area has grabbed significant attention in recent years by image processing community, the frameworks have not been applied in application scenarios like biomedical imaging.

In this research work, we propose ML-CSC based frameworks for inverse problem of compressive sensing MRI (CS-MRI) and biomedical image classification problems. An optimized ML-CSC framework employing multi-layer basis pursuit algorithms is implemented for CS-MRI restoration tasks on different datasets and shown its improved performance in terms of fast convergence, improved PSNR/SSIM and fast restoration times as compared to the state of the art. The performance of layered basis pursuit employing global pursuit and local pursuit is given to demonstrate the effectiveness of the proposed restoration frameworks. Furthermore, comparative analysis is given for CS-MRI restoration task of knee MR images by training and testing state of the art restoration frameworks alongside our proposed framework. Utilizing self-acquired knee MR dataset and different brain MR datasets, the restoration performance of the proposed framework is demonstrated. In further work, a biomedical image classification framework is proposed and implemented for

classification of knee anterior cruciate ligament (ACL) tear pathology on self-acquired and labeled knee MR dataset. With the help of extensive empirical results, the classification efficiency of proposed framework is compared with the state-of-the-art deep learning-based architectures with improved results.

## LIST OF PUBLICATIONS AND SUBMISSIONS

- [1] **A. Wahid**, A. U. Khan, S. Khan, J. Shah, et al., "A multilayered convolutional sparse coding framework for modeling of pooling operation of convolution Neural Networks," in 2019 IEEE International Conference on Smart Instrumentation, Measurement and Application (ICSIMA), IEEE, 2019, pp.1–5.
- [2] S. Ikram, S. Zubair, J. A. Shah, I. M. Qureshi, **A. Wahid**, and A. U. Khan, "Enhancing MR image reconstruction using block dictionary learning," IEEE Access, vol. 7, pp. 158434–158444, 2019.
- [3] **A. Wahid**, J. A. Shah, A. U. Khan, M. Ahmed and H. Razali, "Multi-layer basis pursuit for compressed sensing MR image reconstruction," in IEEE Access, vol. 8, pp. 186222-186232, 2020, doi: 10.1109/ACCESS.2020.3028877.
- [4] **A. Wahid**, J. A. Shah, A. U. Khan, M. Ullah and M. Z. Ayob, "Multi-layered basis pursuit algorithms for classification of MR images of knee ACL tear," in IEEE Access, vol. 8, pp. 205424-205435, 2020, doi: 10.1109/ACCESS.2020.3037745.
- [5] **A. Wahid**, K. Kadir, A. W. U. Ullah, and S. Saadain, "Multi-layer convolutional sparse coding framework for restoration of under-sampled MR images," International Instrumentation and Measurement Technology Conference (I2MTC-2021).
- [6] **A. Wahid**, J. A. Shah, A. U. Khan and S. Ikram "A performance comparative analysis of MRI compressive sensing image restoration frameworks based on convolutional sparse coding and deep learning techniques." Journal of Computerized Medical Imaging and Graphics.(Submitted)



## ACKNOWLEDGEMENTS

*I would like to thank my advisor Dr. Jawad Ali Shah, who is supportive, generous, and is always there to provide help and guidance to his students. He continually and convincingly conveyed a spirit of adventure in regard to open minded research. Without his guidance and persistent help, this dissertation would not have been possible. I would like to thank my co-advisor Dr. Adnan Umar Khan for extending every possible help in facilitating my research work and advising me at every critical point of stay at IIU Islamabad. I would also like to thank Prof. Mukhtar Ullah for valuable suggestions and guidance. I would always be grateful to Prof. Ijaz Mansoor Qureshi (Late) for his encouragement and support at hardest time in research and taking care of me during my stay at IIUI. I would like to express my appreciation to my friendly and inspiring lab mates. I extend my acknowledgments and gratitude to my teachers Dr. Aqdas Naveed Malik, Dr. Abdul Jalil and Dr. Nouman for their assistance during my studies.*

*Lastly, I would like extend my gratitude to Higher Education Commission of Pakistan (HEC) and British Malaysian Institute (BMI), University of Kuala Lumpur, Malaysia for their financial support for completion of this research work.*

**Abdul Wahid Tareen**

May, 2021

## TABLE OF CONTENTS

<b>Abstract</b> . . . . .	v
<b>Publications</b> . . . . .	vii
<b>Acknowledgments</b> . . . . .	viii
<b>List of Figures</b> . . . . .	xiv
<b>List of Tables</b> . . . . .	xvi
<b>Chapter 1: Introduction</b> . . . . .	17
1.1 Contributions . . . . .	19
1.2 Thesis organization . . . . .	21
<b>Chapter 2: Literature Review</b> . . . . .	23
2.1 Convolutional Neural Networks, Sparse Coding and Multi-layer Convolutional Sparse Coding Theory . . . . .	23
2.1.1 Convolutional Neural Networks . . . . .	24
2.1.2 Sparse Coding . . . . .	25
2.1.3 Convolutional Sparse Coding and Multi-layer Convolutional Sparse Coding . . . . .	27
2.2 Magnetic Resonance Imaging Compressive Sensing Image Restoration Problem-Image Restoration Techniques . . . . .	30

2.2.1	Optimization Based Reconstruction . . . . .	33
2.2.2	Learning Based Reconstruction with Deep Neural Networks . . . . .	33
2.2.3	Multi-layered Thresholding Algorithms with Unfoldings . . . . .	34
2.2.4	Convolutional Sparse Coding and Inverse Problems . . . . .	36
2.3	Feature Learning and Classification Problem . . . . .	41
2.3.1	Deep Neural Networks for Feature Learning and Classification . . . . .	43
<b>Chapter 3: Implementation of Multi-Layer Convolutional Sparse Coding Based Framework for Compressed Sensing MR Image Reconstruction</b>		<b>45</b>
3.1	MRI Image Acquisition and CS Reconstruction Model . . . . .	46
3.1.1	Solvers for CS MRI Reconstruction Problem . . . . .	47
3.1.2	ML-CSC as a Theoretical Foundation For Deep Learning . . . . .	48
3.1.3	Multi Layered Basis Pursuit . . . . .	49
3.1.4	Iterative Shrinkage Algorithms . . . . .	50
3.1.5	Layered Basis Pursuit . . . . .	50
3.2	Proposed ML-CSC Based Framework for CS-MRI . . . . .	51
3.2.1	Model . . . . .	51
3.2.2	Architecture . . . . .	55
3.2.3	Training Loss for CS-MRI Framework . . . . .	56
3.2.4	Datasets . . . . .	57
3.3	Empirical Results . . . . .	57
3.3.1	Clinical Application . . . . .	58
3.3.2	Discussion . . . . .	60

<b>Chapter 4: Implementation of Multi-layer Convolutional Sparse Coding based Framework for Biomedical Image Classification Problem . . . . .</b>	<b>63</b>
4.1 Anterior Cruciate Ligament Tear - Background . . . . .	65
4.2 Deep Neural Networks for Classification . . . . .	66
4.3 Convolutional Sparse Coding Model-The Multilayered Basis Pursuit . . . . .	68
4.3.1 ML-CSC for Classification . . . . .	70
4.4 Algorithms . . . . .	70
4.4.1 Layered Basis Pursuit . . . . .	70
4.4.2 Iterative Thresholding Algorithms (ISTA) . . . . .	71
4.4.3 Multi-layer ISTA and FISTA . . . . .	72
4.4.4 Unfolded Iterative Algorithms as Neural Networks . . . . .	75
4.5 Experiments and Results on Knee MR Dataset . . . . .	76
4.5.1 Multi-layered Iterative Thresholding Algorithms with Unfoldings . . . . .	78
4.5.2 Layered Basis Pursuit . . . . .	78
4.5.3 An Adaptive Learning Framework . . . . .	78
4.5.4 Data Augmentation . . . . .	84
4.5.5 Class Imbalance . . . . .	85
4.5.6 Deep Learning Architecture and the Challenge of Overfitting . . . . .	85
4.5.7 Discussion . . . . .	86
<b>Chapter 5: Multi-layer Convolutional Sparse Coding Framework for Restoration of Under-sampled MR Images . . . . .</b>	<b>88</b>
5.1 CS MR Image Reconstruction as Inverse Problem . . . . .	89
5.1.1 Dictionary Learning-Based Approaches . . . . .	90

5.1.2	Low-Rank-Based Approaches . . . . .	90
5.1.3	Deep Learning and Multi-layer Thresholding Algorithms for MR Image Reconstruction . . . . .	90
5.2	ML-CSC Based Framework for CS-MRI Image Reconstruction Problem . .	92
5.2.1	Iterative Shrinkage Algorithms for Basis Pursuit . . . . .	92
5.2.2	Model . . . . .	93
5.2.3	Datasets and Parameter Settings . . . . .	95
5.2.4	Training Loss . . . . .	96
5.3	Empirical Results . . . . .	96
5.4	Discussion . . . . .	98
 <b>Chapter 6: Comparative Analysis of ML-CSC Based CS-MRI Framework With State of the Art . . . . .</b>		 <b>99</b>
6.1	Convolutional Sparse Coding and Inverse Problems . . . . .	100
6.2	State of the Art CS-MRI Restoration Framework and ML-CSC Based CS-MRI Model . . . . .	101
6.2.1	Ista-Net:The Layered Basis Pursuit (ML-BP) . . . . .	102
6.2.2	Multi-layer Convolutional Sparse Coding Framework for CS-MRI Image Restoration . . . . .	102
6.2.3	Multi Layered Basis Pursuit . . . . .	103
6.2.4	Iterative Shrinkage Algorithms . . . . .	104
6.2.5	Model . . . . .	105
6.2.6	Architecture . . . . .	108
6.2.7	Training Loss for CS-MRI Framework . . . . .	109
6.2.8	Dataset . . . . .	110

6.3 Empirical Results . . . . .	110
6.4 Summary . . . . .	111
<b>Chapter 7: Conclusion and Future Work . . . . .</b>	<b>112</b>
7.1 Summary of the Contributions . . . . .	112
7.2 Future Directions . . . . .	113
<b>Appendix A: Experimental Equipment . . . . .</b>	<b>115</b>
<b>References . . . . .</b>	<b>116</b>

## LIST OF FIGURES

2.1	Forward pass of a Convolutional Neural Network . . . . .	25
2.2	Sparse coding image representation model . . . . .	26
2.3	A multi-layer convolutional sparse coding model . . . . .	28
3.1	Framework for CS-MRI. . . . .	54
3.2	Train set losses for brain MR (left) and knee MR (right) datasets. . . . .	56
3.3	Knee MR ground truth images (top) and reconstructed images (bottom), for ACL tear identification with 5-fold acceleration factor . . . . .	58
3.4	Examples of undersampling mask . . . . .	59
3.5	PSNR/SSIM of brain MR[68] test set . . . . .	59
3.6	PSNR/SSIM of Knee MR test set . . . . .	60
4.1	Knee anatomy with normal and torn ACL tears (Coronal view) . . . . .	65
4.2	Multilayered basis pursuit based framework for MR image classification. . .	71
4.3	A two layer ISTA model as illustrated in [1] . . . . .	75
4.4	Labeled Knee MR images from Knee MR dataset . . . . .	76
4.5	Confusion matrices of CNN and All-Free framework . . . . .	79
4.6	Confusion matrices of ML-ISTA with unfoldings . . . . .	80
4.7	Confusion matrices of ML-FISTA with unfoldings . . . . .	81

4.8	Confusion matrices of ML-BP with unfoldings . . . . .	82
4.9	Training accuracy for unfolding = 1 (left), and unfolding = 2 (right) . . . . .	83
4.10	Training loss for unfolding = 1 (left), and unfolding = 2 (right) . . . . .	83
4.11	Validation loss for unfolding = 1 (left), and unfolding = 2 (right) . . . . .	84
5.1	Training framework for CS-MRI image restoration with ML-BP . . . . .	95
5.2	Example images from test set . . . . .	95
5.3	Image-wise PSNR and SSIM for 20% CS ratio (5-fold acceleration factor) for brain MR images . . . . .	97
6.1	Number of publication on Scopus searched with "CSC and ML-CSC" search term. . . . .	100



## LIST OF TABLES

3.1	Average PSNR/SSIM and restoration times of reconstructed brain and Knee MR images . . . . .	60
3.2	Algorithms along with the PSNR and reconstruction times of brain MR images . . . . .	61
4.1	Precision, Recall, Accuracy and F-1 scores of iterative thresholding algorithms with unfoldings . . . . .	77
4.2	Class-wise accuracies and test loss . . . . .	77
5.1	Average PSNR/SSIM and restoration times of reconstructed brain MR images . . . . .	96
6.1	Average PSNR/SSIM and restoration times of reconstructed Knee MR images . . . . .	110

## CHAPTER 1

### INTRODUCTION

Deep neural networks have been the focus of research for solving an array of problems during the last decade. The enhancements in the computational power of computers further enabled these frameworks to be used for different image processing tasks. With the successful applicability of these networks, interest grew in the theoretical foundation of these architectures which largely remain heuristic in their implementations. Thus, the resurgence of the application of time-tested exiting tools of sparse coding theory has been applied to better analyze and model these architectures. The celebrated sparse coding theory has been presented as one key candidate for the purpose. With its successful implementation in signal and image processing research, the theory has been further extended to model deep neural networks i.e. convolutional neural networks (CNNs), and analyze their performance for better understating of their inner working and possible further improvements. The convolutional sparse coding (CSC) and the multi-layer convolutional sparse coding (ML-CSC) which is a special case of sparse coding theory have been presented as theoretical frameworks which can be used to model CNNs theoretically. In the seminal works, the CSC and ML-CSC has been implemented for general datasets with legacy CNNs as the baseline for performance comparison [1].

Although the preliminary work has shown the impressive performance of ML-CSC based frameworks, open problems need further investigations for effects of batch normalization, regularization techniques, application of pooling operations, and implementation of ML-CSC on applied problems of image processing.

Existing deep learning-based architectures suffer from the following limitations.

#### **Deep neural networks for inverse problems**

- The deep learning architectures are mainly trained with heuristic techniques that require theoretical analysis to improve feature learning for an accurate solution to the inverse problems, especially in the case of biomedical images where the reconstructions should be of reasonable quality for the clinicians to accurately interpret for diagnosis.
- The frameworks require extensive training before being implemented in application scenarios on custom datasets.
- A generalized architecture that can be adapted to other datasets and protocols is required to pave the way for better integration of Artificial intelligence (AI)-assisted tools in the biomedical imaging pipeline.
- Once the model parameters are learned, the test/ restore framework is required to be implemented with minimum effort in clinical settings for wider applicability.

#### **Deep Neural networks for classification of biomedical images**

- The frameworks are mostly applied on general imaging datasets for classification which requires a large number of labeled images. In the case of biomedical images, the availability of large datasets labeled by specialist radiologists is a challenging issue.
- Due to the limited availability of datasets, the problem of imbalanced classification becomes even more challenging which is crucial for decision making for post-diagnosis treatment of patients. Furthermore, in the case of classification of biomedical images for anterior cruciate ligament tears problem (ACL), recent works mostly address the binary classification problem of presence or absence of ACL tears, wherein

the classification task of partial tear presents an additional challenge for classification algorithms.

- The deep learning (DL) architectures are mainly trained with heuristic techniques which require theoretical analysis to improve feature learning for accurate classification, especially in the case of biomedical images where the error margin should be as small as possible.
- Lastly medical imaging plays a crucial role in diagnosis, treatment planning, treatment delivery, and follow-ups. To increase efficiency, there is an urgent need to integrate DL-based automation tools into all stages of the medical imaging pipeline ranging from image acquisition and reconstruction to analysis and interpretation.

## **1.1 Contributions**

This research work aims at developing algorithms for compressive sensing magnetic resonance imaging (MRI) image reconstruction for their corresponding undersampled measurement and biomedical image classification frameworks based on ML-CSC theory employing multi-layer basis pursuit algorithms. In the proposed framework the insights are taken from the celebrated sparse coding theory and its recently proposed version of the multi-layered sparse coding to design a CS-MRI framework and a classification framework optimally designed and tuned for grayscale labeled MR images obtained at Hospital Kuala Lumpur (HKL). Specifically, this work makes the following contributions,

- The multi-layer sparse coding theory which has been recently used to explain the theoretical foundations of deep learning is extended to inverse problem namely CS-MRI to demonstrate its effectiveness. Furthermore, unlike the traditional deep learning frameworks based on CNNs, the multi-layered iterative thresholding algorithm (ML-ISTA) based framework benefits from network unfoldings to increase network depth without incurring the cost of additional parameters and computational complexity.

This in turn improves reconstruction performance demonstrated through experiments on multiple datasets.

- A generalized framework is designed, trained, and images restored on four benchmark datasets.
- The framework is tested to demonstrate the fast restoration of MR images on GPU and CPU machines with reasonable restoration time and PSNR/SSIM.
- The framework has been optimized successfully on the labeled knee MR dataset and an average test set accuracy of 92% has been achieved without adding regularization techniques.
- The unrolling ML-CSC framework with ML-ISTA, its fast version FISTA along with multilayered Basis Pursuit (ML-BP) is demonstrated to perform better than the CNN-based classification framework taken as baseline, without increasing network depth.
- The solution to challenging partial ACL tear classification problem, where the classifiers generally do not give sound accuracies, is optimized with data augmentation techniques and accuracy of more than 85% on this specific class is achieved by proposing for multilayer iterative thresholding algorithm (ML-ISTA) framework, which outperformed traditional CNN with the same number of parameters.
- Data augmentation technique especially suited for training a DL framework on biomedical images is applied, which shows optimal performance as compared to other transforms used in image classification algorithms.
- Lastly, a classification framework based on a recurrent architecture with the same depth as the generative models described above is trained and analyzed for comparison. The accuracies of all models are compared with CNN, demonstrating the viability and effectiveness of the MR image classification framework.

Taking insights from recent research works in theoretical foundations of deep learning, we have proposed a CS-MRI restoration framework based on multilayered convolutional sparse coding (ML-CSC) with Iterative Shrinkage Thresholding Algorithms based global basis pursuit to learn a fast mapping between the training set of MR images and their corresponding undersampled measurements in  $k$ -space. With the global pursuit of ML-CSC and faster convergence, the framework is first trained on a GPU machine in a short time and the learned model is successfully used for CS-MRI image restoration with different CS-ratios using random masks. The test set reconstruction is also carried out on the CPU with a reasonable restoration time without employing GPUs after the trained parameters are available.

## **1.2 Thesis organization**

The rest of the dissertation is organized as follows.

### **Chapter 2: Literature Review**

The chapter covers the background for the other chapters of the dissertation. A brief overview of sparse coding theory, theoretical foundations of deep neural networks, and multi-layer convolutional sparse coding is given. The feature learning of convolutional neural networks and the ML-CSC-based model is introduced with a focus on implementation algorithms for inverse problems and classification problems in the biomedical imaging domain.

### **Chapter 3: Multi-Layer Basis Pursuit for Compressed Sensing MR Image Reconstruction**

This chapter introduces a generic framework for MR image reconstruction from under-sampled measurements which is applied for different CS-MRI problems. Experiments on different MR datasets acquired in different imaging planes are performed to demonstrate the merits of the proposed restoration framework as compared to the state-of-the-art. For multi-layer basis pursuit, the global pursuit algorithm employing an iterative thresholding

algorithm is utilized for CS-MRI image reconstruction.

#### **Chapter 4: Implementation of multi-layer convolutional sparse coding framework for classification problem**

A classification framework based on ML-CSC, employing iterative thresholding algorithms for basis pursuit is proposed and implemented for classification of custom knee dataset acquired at Hospital Kuala Lumpur (HKL). A series of detailed experiments are conducted to show the merits of the proposed classification framework as compared to baseline convolutional neural network-based architecture.

#### **Chapter 5: Multi-layer Convolutional Sparse coding Framework for Restoration of Under-sampled MR Images**

Further extending the work of chapter 3, a CS-MRI framework utilizing iterative thresholding algorithms for layer-wise pursuit is proposed. Extensive experiments on two benchmark datasets of brain MR images are shown to demonstrate the merits of the proposed framework.

#### **Chapter 6: Comparative analysis of ML-CSC based CS-MRI framework with state of the art**

The frameworks proposed in Chapter 3 and Chapter 5 are compared with the state-of-the-art custom knee MR dataset acquired for the research work at Hospital Kuala Lumpur (HKL). The analysis was done by training the proposed CS-MRI restoration and state-of-the-art MRI restoration frameworks on the same dataset and presenting the restoration results of both in tabular form demonstrating the better restoration efficiency of our proposed framework.

#### **Chapter 7: Summary and Future Direction**

The chapter summarizes the contributions presented in this dissertation and recommends future directions about how to employ deep sparse coding to further improve the entire MRI imaging pipeline and workflow and discuss future work.

## **CHAPTER 2**

### **LITERATURE REVIEW**

Deep neural networks have been the focus of extensive research works in recent years due to the availability of supporting hardware and large-scale datasets. Due to the heuristic structure of deep learning algorithms, the theoretical aspect of working of these frameworks has been researched to better model the inner working and develop insights for improvement. Sparse coding theory [2], which has been the focus of image processing research during the last decade has been presented as the theoretical foundation of deep neural networks. In the following, Convolutional Neural Networks (CNNs), the sparse coding theory, and the convolutional sparse coding the inverse problem of MRI image reconstruction and classification of MRI images are briefly reviewed.

#### **2.1 Convolutional Neural Networks, Sparse Coding and Multi-layer Convolutional Sparse Coding Theory**

In the following, first a review of CNNs, is given along with sparse coding theory. The connection between the two is given through a brief review of convolutional sparse coding and multi-layer convolutional sparse coding. Lastly the chapter-wise organization of dissertation is given.

Convolutional neural networks [3, 4] are utilized for data that exhibit grid-like topology. For one-dimensional signals, the neural networks can be modeled as 1D-CNN [5], whereas image data can be modeled as 2D-CNN [6]. For video data, the kernel or feature dimension can be further extended to model video data[7].



### 2.1.1 Convolutional Neural Networks

The ability of machine learning algorithms that can learn and improve based on the experience of the complexity of the problem at hand and their adaptation to that specific problem gives tremendous opportunities in an array of applications. The CNNs are a fundamental part of representation learning and are briefly explained below. The forward pass is the fundamental part of the CNNs (Figure 2.1), where an input signal  $X \in \mathbb{R}^N$  is convolved with set of  $m_1$  learned filters of chosen size  $n_1$  giving output as feature maps or kernels. In matrix vector form, this can be written as  $W_1^T X \in \mathbb{R}^{N \times N m_1}$ , where  $W_1 \in \mathbb{R}^{N m_1}$ , is a convolutional matrix (transposed) with learned filters as columns with all their shifts. After convolution, a bias term  $b_1 \in \mathbb{R}^{N m_1}$  is added to resulting vector and a nonlinear operation (here Rectified Linear Unit-ReLU) is applied. For a two layers forward pass of CNN, the operation is given by Equation (2.1),

$$f(X, [W_i]_{i=1}^2, [b_i]_{i=1}^2) = Z_2 = ReLU(W_2^T ReLU(W_1^T X + b_1) + b_2) \quad (2.1)$$

The output of first stage/layer is then treated as input to another stage with convolutional matrix  $W_2^T \in \mathbb{R}^{N m_1 \times N m_2}$  and bias term  $b_2 \in \mathbb{R}^{N m_2}$ . The operation is extended up to desired number of layers and feature maps are then used for problems like classification and inverse problems. For the problem of classification, the output of last layer is fed to train a classifier which tries to predict the label  $h(X)$  associated with given image  $X$ . For given dataset of images  $(X_j)_j$ , the task of CNN including filters  $[W_i]_{i=1}^K$  and biases  $[b_i]_{i=1}^K$ , parameter of the classifiers  $U$  can be written as, Equation (2.2)

$$\min_{[W_i]_{i=1}^2, [b_i]_{i=1}^2, U} \sum_j l(f(h(X_j), U, f(X_j), [W]_{i=1}^2, [b]_{i=1}^2)) \quad (2.2)$$

The task of the optimization algorithm is to minimize the mean of the loss function  $l$ . Recent research works have utilized CNNs, especially in computer vision. These frameworks

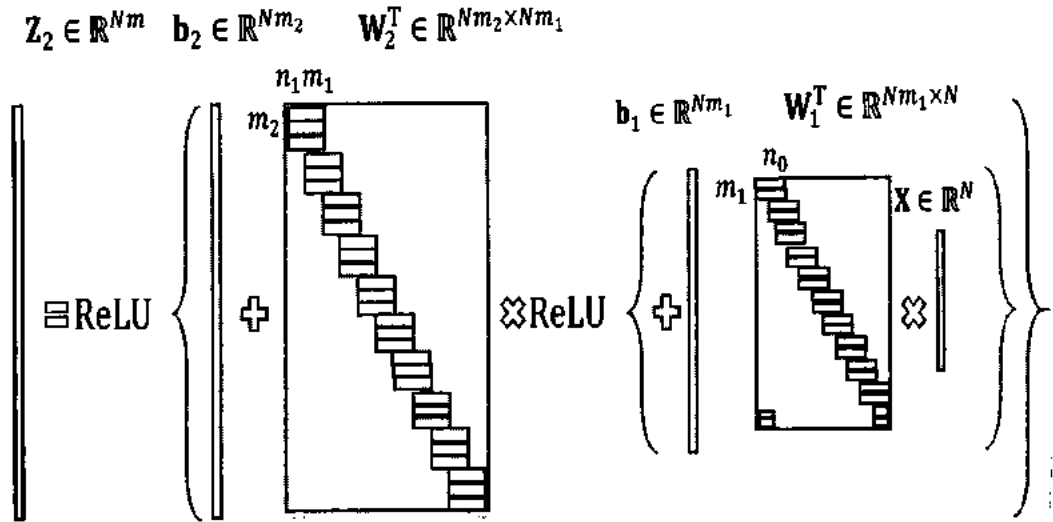


Figure 2.1: Forward pass of a Convolutional Neural Network

have been used in deep learning research problems ranging from image processing tasks of segmentation, classification, general object detection to inverse problems [8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20]. The successful application of CNNs in range of image processing applications emphasized the need for their broader theoretical understanding to improve the performance. Sparse coding theory [2] has been successfully utilized for image processing tasks of denoising, de-blurring, image recognition, image segmentation and image classification during the last decade [21, 22, 23, 24, 25, 26, 27, 28, 29, 30].

### 2.1.2 Sparse Coding

Sparse representations for image processing work on the premise that images can be represented as a sparse linear combination of elements from a redundant dictionary. For an image of size  $X$ , the patch extractor operator  $R_i$  extracts a patch of a specific size, which can be represented as,

$$x_i = R_i X$$

The sparse coding theory further represent the  $i$ th patch of the underlying image as:

$$R_i X = D \gamma_i$$

Where  $D$  is the dictionary of atoms  $\gamma_i$  is the  $i$ th sparse feature map representing the  $i$ th patch extracted from the image. The sparse coding image representation modal is given in Figure 2.2.

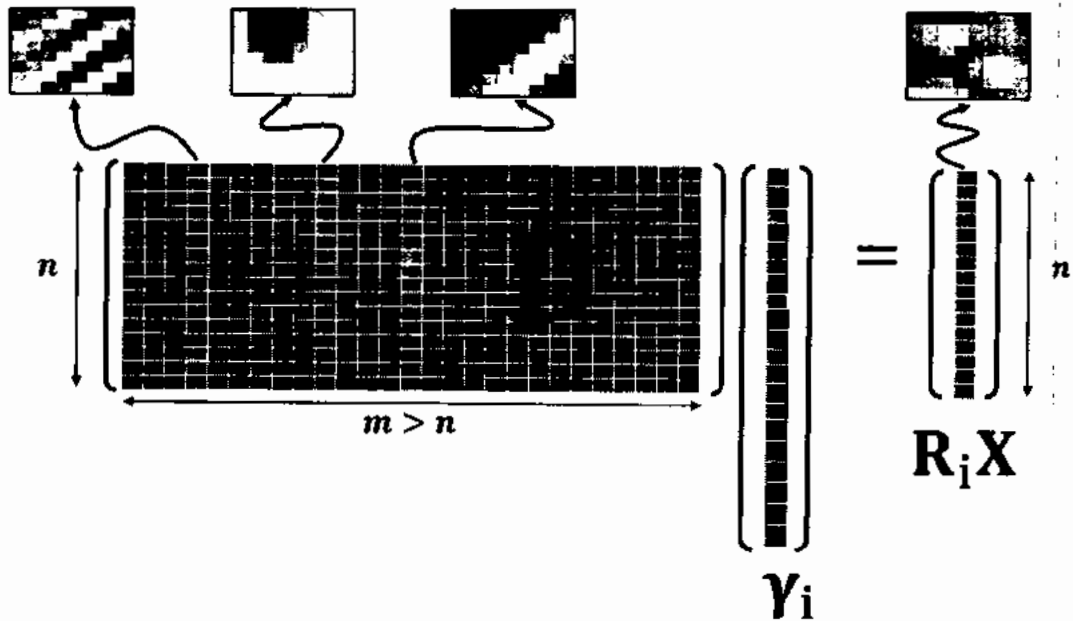


Figure 2.2: Sparse coding image representation model

The greedy basis algorithm pursuit can be written with  $\ell_0$  norm as:

$$P_0 : \min_{\gamma_i} \|\gamma_i\|_0 \quad s.t. \quad R_i X = D \gamma_i, \quad (2.3)$$

The optimization of Equation (2.3) is NP-hard therefore greedy methods such as Orthog-

onal Matching Pursuit (OMP), thresholding or convex relaxations such as basis pursuit are employed for solutions. Over the years the sparse coding-based model has been successfully used for image denoising problems [31] and other inverse problems in image processing [32, 33, 34]. Inverse problems in image processing like image de-blurring, super-resolution, and image denoising have been successfully tackled with a sparse representation model. Under the model assumption, natural signals can be modeled as a sparse linear combination of a few columns taken from a matrix defined as a dictionary. Sparse coding or pursuit is the process of recovering the sparse decomposition of a given signal over a dictionary (typically over complete). This inverse problem is formulated usually as an optimization objective with minimizing a  $\ell_0$  pseudonorm or the  $\ell_1$  norm as its convex relation resulting in signal reconstruction of acceptable quality. For deployment of an effective sparse representation model, the discovery of a dictionary consistent with the underlying data is required resulting in the dictionary learning problem. This problem gets complicated with high dimensional signals (i.e. complete images) due to requirement of more memory and high computations. Consequently, the global pursuits fail to keep track of local varying behaviors resulting in poor performance of the dictionary learning frameworks. Thus, to circumvent this issue, the large images are first split into small overlapping patches, dictionaries are learnt, sparse coded, reconstructed before averaging the patches back for complete image (global signal) reconstruction. Although this strategy is widely implemented, the limitation of introducing artifacts at the patch averaging stage and not keeping track of dependencies between adjacent patches limits its applicability.

### **2.1.3 Convolutional Sparse Coding and Multi-layer Convolutional Sparse Coding**

The celebrated sparse coding theory has been the focus of research works to theoretically explain deep neural networks. Convolutional sparse coding (CSC) framework which is a special case of sparse coding theory [35, 36], has shown to exhibit superior performance to traditional patch-based image processing frameworks. This framework (Figure 2.3) pro-

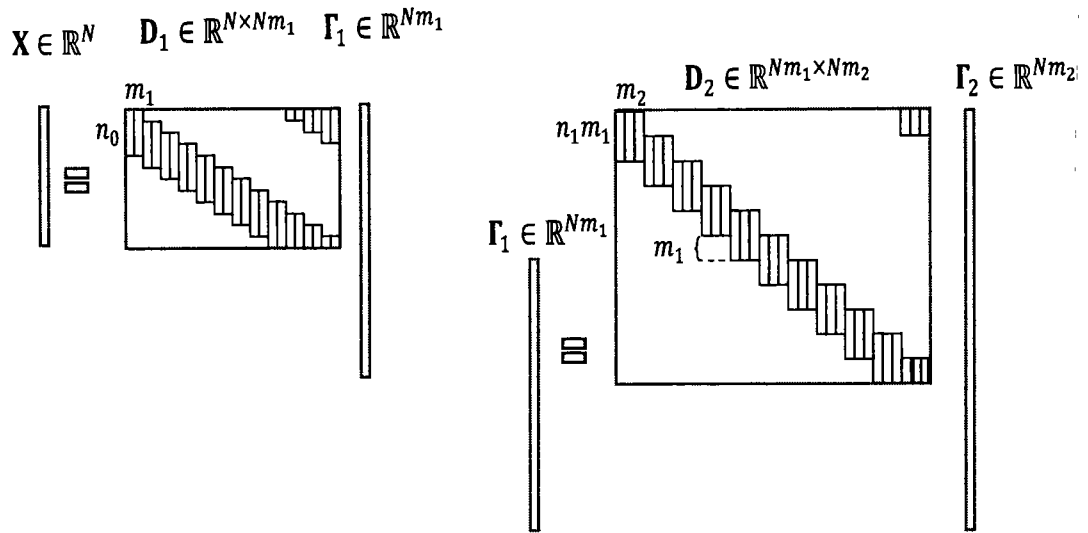


Figure 2.3: A multi-layer convolutional sparse coding model

vides a global, yet tractable, model that operates on the whole images instead of image patches, resulting in addressing several limitations of the patch-based sparse models while achieving superior performance in various applications. The global model also makes the process translation-invariant w.r.t underlying structures of input images and addresses the local-global problem in image modeling present in traditional patch-based processing. In CNNs the images are convolved with filters and dimensions of resulting feature maps are decreased through max or mean pooling. This process is repeated layered wise until high-level feature maps are extracted from input images and then used for classification or other tasks. In convolutional sparse coding [37, 35], first the convolutional dictionaries are learned from images, and then pursuit algorithms are used to learn sparse feature maps of the given images. These feature maps are then used for image classification and the solution of inverse problems. The CSC framework has been presented in recent research works to establish the connection between convolutional neural networks and sparse coding theory. The forward pass of CNNs and CSC was shown completely equivalent in [38], wherein it was elaborated that CNNs are a special case of sparse modeling with special

structures imposed on dictionaries of sparse models, and stability of solution bounds on the success of pursuit algorithms. In multilayered CSC (ML-CSC), which is a layered extension of CSC, the convolutional dictionaries are trained and feature maps obtained on input images, and pursuit algorithms are applied in layers to obtain high-level features. In each successive layer, the ML-CSC takes the feature maps of the preceding layer as inputs and repeats the process of dictionary learning and sparse coding to successive layers. Although earlier work of [38, 39] did not incorporate effects of pooling operation on ML-CSC model, the work in [40] theoretically modeled pooling operation in ML-CSC framework and emphasized the benefits of the addition of pooling layer to CSC model. Work in [40] applied pooling operation on one layer only and took feature maps of successive layers in original feature dimensions. Rey-Otero *et al.* in [41] proposed two convex alternatives by merging global norms with local penalties and constraints to existing CSC formulations resulting in more efficient and converging algorithms. The work in [41] examined the proposed model on tasks of natural image inpainting and cartoon texture separation.

Although the CSC and MLC have been investigated for theoretical insights in the context of deep learning frameworks in recent years, the implementations of the models thus proposed largely remain limited to general datasets and natural images. Thus in our research work, the theoretical insights of ML-CSC were applied to challenges of biomedical image processing tasks, such as the restoration of undersampled MRI pictures and the classification of knee MRI images for anterior cruciate ligament (ACL) rupture in our work. Specifically, we have proposed a compressive sensing MRI (CS-MRI) image restoration framework and a classification framework for biomedical images incorporating the multi-layer convolutional sparse coding and computationally efficient first-order pursuit algorithms. The CS-MRI framework has been trained on multiple brain datasets and a knee MR dataset and its performance has been compared with state-of-the-art CS-MRI frameworks. The classification framework has been trained on the knee MR dataset for the classification of ACL tear types of normal, complete, and partial tears. Extensive experiments on multiple datasets

695841

show the effectiveness of the proposed frameworks and their possible integration into MRI imaging pipelines for improving the efficiency of diagnosis.

## **2.2 Magnetic Resonance Imaging Compressive Sensing Image Restoration Problem+ Image Restoration Techniques**

MRI is a non-invasive imaging modality, which is frequently used in academic and clinical settings for different applications. The imaging modality with the help of flexible signal encodings is being used to visualize soft-tissue contrast showing anatomical, metabolic, and functional information. Unlike other radiology modalities like X-rays, CT scans, and positron emission tomography (PET), MRI is radiation-free which makes it favorable to reduce risk factors. Although these benefits have resulted in the wide adoption of MRI in clinical applications, some challenges need to be addressed to make MRI more accessible and efficient. One of the challenges in using MRI is its longer scan time as compared to the radiology-based techniques, as the technique requires more measurements to reconstruct high-quality images for better diagnosis. To address this issue, a great deal of effort has gone into designing optimization-based algorithms, where the reconstruction function is a regularized minimization problem. Similarly, techniques like parallel imaging (PI) and compressive sensing (CS) have also been proposed to reduce the acquisition time through undersampling. In PI, the objective of undersampling is achieved with the placement of more receiver coils in the MRI scanner and processing the resulting signals for reconstructions. In contrast, the CS technique uses existing hardware for signal acquisitions and employs sparse coding techniques for reconstructions. The advantage of CS is its ready applicability to the existing acquisition systems, unlike PI, which requires hardware modifications for parallel acquisitions.

Reconstructing signals from their measurements is an inverse imaging problem where the aim is to recover an estimate of original image of specific size  $x$  from the measurements  $y$ .

In its basic form, the problem of reconstruction is formulated as:

$$R_{obj}\{y\} = \arg \min_x f(H(x), y), \quad (2.4)$$

where the objective of the optimization is to develop a system model  $H$  with minimum possible errors with the aim to recover image  $x$  of specific size from its undersampled measurements  $y$ . The challenge for recovering a signal from its under-sampled observation is to design and implement Equation (2.1) for a specific application. For sparse measurements of signals that exhibit sparsity in some domains, the Equation (2.1) can be further added with a regularized term that enforces the sparsity constraint on required solutions. Therefore, any new application necessitates the development of new system models and corresponding regularizers resulting in much effort going into developing general-purpose minimization algorithms and their regularizers in recent years. The reconstruction algorithms for CS-MRI are tuned for a non-linear optimization problem, which uses time-consuming iterative techniques. These techniques use handcrafted penalty terms as regularizers, which are difficult to tune. Furthermore, due to the non-quantitative nature of MR images in most cases, MR images can vary between protocols and scanners resulting in difficulty for clinicians in quantitative analysis of MRI scans. CS theory works on the premise that if a signal is sparse in some transform domain, it can be reconstructed with high probability after taking fewer measurements as compared to the Nyquist sampling theorem. With the capability of undersampling during acquisition and compression at the same time, CS theory has been applied in different applications like video compressive sensing [42], cognitive radio communication [43], ECG signal compression [44] and MRI [45].

An alternative to optimization-based approaches for recovering an undersampled signal is learning-based approaches. Given a set of ground truth images as train set and their corresponding undersampled measurements  $\{x_n, y_n\}_{n=1}^N$ , the objective of the learning-based



approach is to solve the reconstruction algorithm of the form:

$$R_{learn} = \arg \min_{D_\theta, \theta \in \Theta} \sum_{n=1}^N f(\{x_n, y_n\}), D_\theta + g(\theta) \quad (2.5)$$

where  $R_{learn}$  is the parametric reconstruction model,  $f(\cdot)$  is a measure of the error to be minimized,  $\Theta$  is the set of all learnable parameters,  $g(\Theta)$  is the regularizer added to avoid overfitting of the learning model. Once the learning/training step is complete, the model can be readily used to reconstruct new images from their corresponding undersampled measurements. Fueled by the learning and approximation capability of deep neural networks, learning-based techniques have been proposed for the restoration of CS signals in compressed form to the original signal domain [46, 47]. For MRI acquisition, where the signal is acquired in  $k$ -space, the challenge for successful CS-MRI reconstruction is to train a neural network for nonlinear mapping with fully sampled  $k$ -space data [48]. One of the advantages of learning-based algorithms is their ability to overcome many limitations of legacy optimization-based approaches like handcrafted features for forward models and the design of problem-specific cost functions and optimizers. In other words, the strength of the learning-based approaches lies in their generalization capability for an array of problem types, unlike the optimization-based approaches, which are generally problem-specific. The availability of datasets on large scale for MRI research in recent years has enabled the utilization of learning-based models with the design of algorithms for efficient solutions of a myriad of inverse problems including compressive sensing. Generally, the CS reconstruction works have been grouped into three categories with the emergence of deep learning-based architectures. In the following section, a brief review is given for these categories with an emphasis on research related to our work.

### 2.2.1 Optimization Based Reconstruction

Optimization-based or objective function-based reconstruction approaches use sparse regularization to design transform sparsity for CS reconstruction. These methods employ the transform sparsity in transform domain [49, 50], and dictionary learning based subspace [51, 52, 53] for CS reconstruction. Methods based on these techniques have the advantage of fast optimization at the cost of introducing staircase artifacts in images constructed from compressive measurements [54]. The issue has been further addressed through the grouping of similar patches to better keep the sparsity structure in non-local patch-based reconstruction methods [55]. The dictionary learning-based methods have better reconstruction quality although are dependent upon reference images to improve the CS reconstruction performance and are difficult to generalize. For methods that rely on traditional transforms like wavelets, or DCT, the convergence is fast, but results are less accurate. The dictionary learning-based methods produce better quality reconstruction results at the cost of slow convergence. Spatiotemporal sparse Bayesian learning techniques have also been employed for simultaneous recovery of multi-channel physiological signals [56]. Employing sparse Bayesian learning, a high-resolution Electrical Impedance Tomography (EIT) image reconstruction method exploiting the structure sparsity of the underlying signals for estimation of images was proposed in [57]. Liu *et al.* in [58] proposed a Bayesian learning-based approach to reconstruct EIT signals utilizing Bayesian spatiotemporal priors, resulting in enhanced reconstruction performance and reduced computational complexity as compared to state of the art.

### 2.2.2 Learning Based Reconstruction with Deep Neural Networks

Deep learning algorithms have been successfully applied to the solutions of inverse problems in image processing in recent years [59, 60]. In [46], A. Musavi *et al.* used stacked autoencoders to reconstruct natural images from ImageNet dataset. Iliadis *et al.* in [47] presented a deep learning-based framework for recovering compressive sensing video.

These CS recovery frameworks were usually constrained by a fixed measurement matrix and image size. To resolve this issue, Jin *et al.* in [60] proposed a convolutional neural network(CNN) based algorithm to learn mappings from preliminary CS reconstruction to high-quality reconstruction. More recently the generative adversarial networks (GANs) are proposed for CS-MRI reconstruction [61, 62, 63]. These works use mean square error in the pixel domain as a loss function to achieve a better quality of CS reconstruction. G.Luo *et al.* in [64] proposed a deep learning framework based on the Bayesian approach by using the likelihood of priors as training loss and the objective function to improve the reconstruction quality of MR images. Wolterink *et al.* in [61] proposed a framework based on training a CNN along adversarial CNN to better estimate routine dose CT images from low dose CT images. The DAGAN framework in [62] coupled adversarial loss with the content loss to better preserve texture and edges in reconstructed images. For improved performance and fast reconstruction time, Mardani *et al.* in [63] compared a GAN-based framework with dictionary learning, wavelet-based, and deep learning frameworks incorporating pixel-wise mapping. Adler *et al.* in [65] presented a deep learning framework based on block-based CS, which jointly optimized sensing matrix and non-linear reconstruction operator giving improved performance compared to state of the art. In the absence of fully sampled datasets, Yaman *et al.* in [66] proposed a self-supervised learning framework for MR image reconstruction without using fully sampled data. By dividing the sub-sampled points for each MR scan into two sets, one for data consistency term and the other for training loss, MR images reconstruction was achieved with comparable quality to fully supervised learning-based approaches.

### **2.2.3 Multi-layered Thresholding Algorithms with Unfoldings**

The discriminative learning methods based on the concept of unfolding an iterative sparse coding algorithm initially proposed in [67], use insights from model-based and deep learning-based approaches. In these algorithms, one iteration is equivalent to traditional CNN-based

learning frameworks. As the iterations are increased, the frameworks learn the required mapping in the training phase with the same number of parameters as the traditional CNNs, with flexibility of increasing network depth without incurring the cost of increased parameters as encountered in traditional CNNs. Zhan *et al.* [68] proposed a multilayered basis pursuit algorithm based on unfolded neural architecture for CS recovery of natural images used by Kulkarni *et al.* in [69] along with its generalization to MR images. This work employed an iterative thresholding algorithm (ISTA) [70], to recover natural images from their undersampled measurements. Inspired from primal-dual hybrid gradient methods, Adler *et al.* [71] proposed a learned primal-dual algorithm replacing proximal operators with a CNN. Gupta [72] *et al.* presented a relaxed version of projected gradient descent, wherein measurement consistency of reconstructed images with their corresponding measurements was enforced to achieve convergence under certain conditions, with improved results for sparse-view computed-tomography (CT) reconstruction. Hammernik *et al.* [73] embedded a variational model in an unrolled gradient descent scheme to reconstruct knee MR images with an undersampling factor of 4. The work demonstrated improved computational performance for undersampled reconstruction on a single graphic card. Metzler *et al.* [74] developed an unrolled approximate message passing (D-AMP) algorithm, defining it as the learned-AMP, where a deep CNN replaced the denoising operator of each iteration of AMP [75]. Further extending their preliminary work in [76], Yang *et al.* [77] proposed a CS recovery framework based on an unrolled alternating direction method of multipliers (ADMM) algorithm for CS recovery of MR and natural images. The proposed model where parameters of CS recovery and ADMM are learned discriminately achieved favorable results on complex-valued MR images and real-valued natural images. In the parallel imaging (PI) domain, Sriram *et al.* [78] integrated traditional PI methods with a deep neural network for acceleration factors of 4-Fold and 8-Fold (CS-ratio of 25% and 12.5%, respectively), achieving comparable results with the state of the art for CS-MRI.

#### 2.2.4 Convolutional Sparse Coding and Inverse Problems

The success of deep learning architectures for the solution of inverse problems in image processing motivated the research direction of designing deep learning architectures incorporating domain-specific knowledge. Convolutional sparse coding theory has been successfully applied to image super-resolution, reconstruction of hyperspectral images, denoising of seismic data, and reconstruction of biomedical images like MRI and CT scans. Marivani *et al.* in [79] proposed a multimodal deep learning architecture for the solution of the CSC problem by employing an unfolding proximal method. Using RGB images as side information, the proposed method was applied for the super resolution problem of near-infrared images. For single image super-resolution (SISR), Ahmed *et al.* in [80] proposed a convolutional sparse coding-based model in the wavelet domain. The work employed a wavelet integrated convolutional sparse coding approach to better capture contextual information of images. Simon *et al.* in [81] suggested a Bayesian connection between the CSC model and the patch-based frameworks for denoising the problem of natural images. Using strided convolutions, the authors proposed a feed-forward network that followed an MMSE approximation to the CSC model. Serrano *et al.* in [82] employed convolutional sparse coding to reconstruct videos from single frame which is coded with temporal information. The full video sequence was then reconstructed from the single coded image and a trained dictionary. The work imposed an additional constraint in the temporal dimension, enforcing sparsity of the first order overtime for the proposed CSC-based high-speed video acquisition problem. Extending the sparse coding theory to hyper-spectral domain, Arun *et al.* in [83] implemented sparse coding in convolutional encoder-decoder framework. The proposed models employing shrinkage thresholding algorithm and structured sparse regularization utilized 3D convolutions for modeling spatial and spectral features of hyper-spectral data. The authors showed with extensive experiments the applicability of 2D and 3D filters along with sparse regularization for improved learning capability on tasks of super-resolution and dimensionality reduction on hyper-spectral datasets. Al-Madani.M *et*

*al.* in [84] implemented a convolutional sparse coding (CSC) framework for denoising task of seismic data with Local Block Coordinate Descent (LoBCoD) algorithm. The proposed framework presented better inversion capability of the CSC model by exploiting its treatment of input images globally, unlike the legacy patch base approaches resulting in better inversion of noisy seismic data. Furthermore, work in [84] compared the performance of the CSC model with dictionary learning-based algorithm K-SVD for benchmarks of peak signal-to-noise ratio (PSNR) and the relative L2-norm of the error (RLNE) showing better performance for the former algorithm. Addressing the ill-posed image super-resolution problem, Marivani *et al.* in [85] proposed a multi-modal deep learning framework that incorporates sparse priors and relies on a deep unfolding operator similar to the CSC model with side information. The framework was used for super-resolution of near-infrared and multi-spectral images with RGB as side information resulting in better performance as compared to the state-of-the-art. Carrera *et al.* in [86] compared the convolutional sparse coding model which seeks sparsity over the entire dictionary via a global optimization with traditional techniques requiring separate sparsity w.r.t. each translate of the orthonormal basis and solving multiple partial optimizations. Through experiments, it was shown that the CSC model performed better on sparse signals and the performance of the model in terms of the variance of the global solution increased rapidly with the original signal becoming less sparse. Xiong *et al.* in [87] proposed a two-stage model utilizing convolutional sparse coding for image restoration problems of de-blurring, denoising, and MRI compressive sensing. The authors used Augmented Lagrangian and alternating direction method of multipliers for an optimal solution to the non-linear optimization problem showing the effectiveness of CSC-priors for the tasks of inverse problems as compared to the state of the art techniques. To extract transient events from the EEG data, Prokopiou *et al.* in [88] employed convolutional sparse coding (CSC) analysis along with simultaneous EEG-fMRI data. The authors used the events in voxel-wise fMRI analysis and also employed FIR models to obtain HRF estimates using convolutional sparse coding events showing the

CSC model's ability to detect reliable events in EEG data. Nguyen-Duc *et al.* in [89] proposed an alternating method that concurrently builds a multi-scale 3D CSC dictionary as the MRI reconstruction from undersampled measurements proceed using a variant of the ADMM algorithm. The authors demonstrated the reconstruction performance as compared to dictionary learning-based methods for undersampled MRI reconstruction.

To recover highly undersampled dynamic fMRI data, Nguyen-Duc *et al.* in [90] employed a multi-scale 3D CSC framework and a spectral decomposition technique. Using a shared 3D convolution-based dictionary built progressively during the reconstruction process, the technique used unsupervised learning to recover high-frequency components while the low-frequency components of MRI were restored using total variation-based energy minimization method using temporal coherence of dynamic MRI data. To better approximate the sparse input data, elastic-net regularization was used. The performance of the proposed method was demonstrated for different sampling ratios on cardiac and brain MR images as compared to the state of the art. Xiong *et al.* in [91] employed convolutional sparse coding in gradient domain to address the limitations of patch-based CS-MRI techniques. In the proposed method, which operated on whole gradient image for capturing correlation between local neighborhoods and exploiting gradient image global correlation, the authors demonstrated the better reconstruction of MR images in terms of better edges and sharp features. Existing methods of image synthesis problems are limited by the requirement of large training sets, are tailored for specific applications, or need patch-based processing. These limitations were addressed by Huang *et al.* in [92] with the proposed framework namely Dual convolutional filter learning (DOTE) approach. The framework's effectiveness was demonstrated on tasks of image super-resolution and cross-modality synthesis. To address the challenge of requirement of large amount of data for fast and accurate reconstruction of magnetic resonance (MR) images from under-sampled data, Guo *et al.* in [93] proposed a federated learning based approach utilizing data from multiple institutions. Authors in [93] used the learnt intermediate latent features among distinct source sites and

aligned them with the distribution of the latent features at the target site in a cross-site modelling for MR image reconstruction for generalizability of the proposed framework.

Quan *et al.* in [94] used 3D convolutional sparse coding to extract and adapt temporal information from the MRI data using compact shift-invariant 3D filters. The authors demonstrated the effectiveness of CSC-based algorithms as compared to the dictionary learning-based state-of-the-art. Quan *et al.* in [95] proposed a CSC based method to reconstruct undersampled dynamic contrast-enhanced (DCE) MRI data. By using an extension to the ADMM algorithm, authors introduced energy formation based on the learning over time-varying DCE-MRI images on graphical processing units comparing the performance with dictionary learning-based state-of-the-art methods.

To simultaneously solve the problems of super-resolution and cross-modality image synthesis, Huang *et al.* in [96] proposed weakly-supervised joint convolutional sparse coding based on semi-supervised learning with few registered multi-modal image pairs. The authors used an additional maximum mean discrepancy term for minimizing dissimilarity between feature distributions of paired images which were enhanced in the joint learning process with unpaired data. A sparse model namely Bidirectional Convolutional Sparse Coding (BiCSC) proposed by Zhang *et al.* in [97] tackled the semi-supervised problem of learning latent features from ultrasound and CT images. These features are then used to formulate a relationship between the two image modalities. The resulting images are then used for multi-needle detection in three-dimensional (3D) ultrasound (US) with reduced noise and artifacts. Finally for true needle position, clustering algorithms were used and model a needle per region of interest(ROI), the random sample consensus algorithm (RANSAC) was used. The framework was used on the prostate image dataset giving its efficacy for accurate needle detection for ultrasound-guided high-dose-rate prostate brachytherapy.

Convolutional sparse coding was used by Bao *et al.* in [98] to reconstruct undersampled CT image reconstruction. The proposed framework exploits the CSC capability of working on whole images instead of patches thus avoiding the artifacts caused by patch aggrega-



tion in traditional dictionary learning-based reconstruction methods. Yan *et al.* in [99] addressed the limitations of Sliding window-based dynamic brain connectivity frameworks by proposing a framework employing that automatically learned a low rank and sparse brain functional connectivity patterns from raw fMRI data. The model with the help of learned sparse coded convolutional filters was able to measure different length dynamic brain functional connectivity patterns in an equal space. The authors demonstrate the applicability of the proposed methods in real clinical applications for brain disorder diagnosis. The limitations of methods using small overlapping patches in sparse coding techniques were recently addressed through convolutional sparse coding using shift-invariant dictionaries and  $\ell_0, \ell_1$  norms. These norms resulted in an imbalanced sparsity across various regions in the given image. This issue was addressed with employing  $\ell_0, \ell_\infty$  norms which operated locally while thinking locally as an alternate to  $\ell_0, \ell_1$  norms. Plaut and Giryes *et al.* in [100] proposed and demonstrated a greedy algorithm for specially tailored for  $\ell_0, \ell_\infty$  norm for image inpainting and salt-and-pepper noise removal. Sun *et al.* [101] proposed an adversarial defensive method to counter adversarial attacks employing the convolutional sparse coding technique. Authors used convolutional dictionaries and projected both clean and adversarially attacked input images to quasi-natural image space and employed a Sparse Transformation Layer (STL) between the first layer of the network the input. The proposed image projection technique onto quasi-natural image space was demonstrated for image detail preservation and adversarial perturbations removal as compared to state-of-the-art transformation-based adversarial defenses.

Liu *et al.* in [102] developed online dictionary learning(ODL) based on first and second-order methods. In contrast to batch-based learning methods, which sample batches of training signals, the ODL techniques process training samples in a streaming fashion. Compared to the state-of-the-art convolutional dictionary learning algorithms based on batch-processing, authors proposed CDL based on first order ODL requiring lower learning time and memory. Furthermore, a second-order ODL method was proposed proving

the convergence of earlier research work[103]. Quan *et al.* in [104] employed hierarchical multi-scale 3D convolutional sparse coding to generate up to 75 dimensions feature vectors for voxel classification problem. The generated features were then fed to a random forest-based classifier for classification. The proposed method was compared against the conventional intensity-based voxel classification methods on real-world volume and synthetic datasets. Chun *et al.* in [105] proposed Block Proximal Gradient method using a Majorizer (BPG-M) for Convolutional dictionary learning (CDL or sparsifying CDL). The method was employed for learning filter for image denoising problem and was compared with the augmented Lagrangian (AL) method or the variant alternating direction method of multipliers (ADMM) based methods. With experiments, it was shown to converge quickly, with fewer memory requirements for large datasets and requiring no parameters tuning as compared to ADMM based methods. Liu *et al.* in [106] presented a synthesis and analysis s deconvolutional network (SADN) enforcing sparse convolutions between filters and their underlying images. The work employed Markov Random fields(MRF) and sparse coding to represent general images with generative and discriminative abilities. The proposed model was shown to be effective on the problem of image compressive sensing.

### **2.3 Feature Learning and Classification Problem**

One of the most common sports injuries in young adults is an ACL tear. A study which spanned over 21 years, discovered an incidence of 68.8 per 100,000 person-years in the general population [107]. The diagnosis requires surgical intervention such as reconstruction or enhanced primary repair to avoid further damage and degeneration of injury into osteoarthritis and subsequent chronic instability [108, 109, 110]. The frequent occurrence of the ACL tear in the sports community and general public requires an accurate diagnosis of complete and partial ACL tears. This is also important for therapeutic decision-making and avoidance of further damage. In addition to examinations by experienced sports medicine specialists for exams like pivot shift tests, magnetic resonance (MR) imaging is routinely

used to complement and confirm clinical diagnosis and assess the status of associated injuries. MR imaging plays a crucial part in diagnosing, treatment planning, treatment delivery, and follow-ups. Consensus is building among researchers for the stronger need of using automated tools to reduce costs, increase efficiency and provide higher diagnostic and prognostic accuracy for clinical decision making. Although for an experienced musculoskeletal (MSK)-trained radiologist, MR imaging is specific and accurate in diagnosing ACL tears [111], the diagnosis becomes challenging for non-MSK radiologists and clinicians without access to sub-specialty radiology. People who do not have access to specialists for diagnosing ACL tear injuries remain at risk of further deterioration of injuries without timely and proper diagnosis.

Deep learning (DL) has emerged as a powerful tool for image processing tasks in recent years complemented with the development of graphical processing hardware. The subsequent DL algorithms developed so far have been successfully used in tasks like object detection [112], MR image reconstruction [15, 113] and classification of biomedical images [114, 115]. The primary advantage of learning representations through deep neural networks is their ability to learn semantically meaningful patterns and features in underlying data without explicit human intervention. These models once trained successfully on training datasets can be effectively used for the solution of a range of problems like image recognition and image classification on (unseen) test data. With the tremendous success of DL architectures i.e. the Convolutional Neural Networks (CNNs) until recent years, their working has largely remained heuristic, and a deeper understanding is required to model their working and improve performance. Sparse coding theory [2] developed over the last decade has been applied to a range of problems in image processing [27, 116]. The theory is based on constructing models that represent signals as linear combinations of a few columns, called atoms from a given redundant matrix termed as a dictionary. This theory applied successfully in problems of image processing tasks over the last decade has been recently extended to explain the theoretical foundation of DL. The convolutional sparse

coding (CSC) and its multilayered version ML-CSC have been introduced to explain the theoretical foundation of DL and its association with the sparse coding theory. Specifically, the CNNs are interpreted as approximations of multi-layer basis pursuit problem [35, 38]. The ability of DL architectures to learn from the wider availability of datasets demonstrate their capacity to become part of biomedical imaging workflow and help revolutionize the healthcare industry, especially for communities with limited access to specialized facilities.

### 2.3.1 Deep Neural Networks for Feature Learning and Classification

With improvements in hardware for processing a large number of images, it became feasible to train large neural networks for classification tasks on datasets of different sizes. The seminal work of [117] significantly improved the state-of-the-art classification of general images using graphical processing units. The work achieved an error rate of more than 15% on the ImageNet dataset. This error rate has been improved significantly since then on general datasets available for research purposes. In addition to general datasets, Rajpurkar *et al.* in [118] implemented a 121 layer deep neural network for radiologist-level pneumonia detection on chest X-rays. The algorithm was trained on a large dataset of Chestx-rays14 [119]. Hannun *et al.* in [120] developed a deep neural network (DNN) to classify 12 rhythm classes using 91,232 single-lead ECGs from 53,549 patients. Liu *et al.* in [121] used classification performance of DL networks as compared to clinical reports for binary classification (tear or no tear presence), and concluded that there is no significant difference between the two. The study by Chang *et al.* in [122], demonstrated the feasibility of a high-performing CNN tool to detect complete ACL injury with over 96% test accuracy for binary classification problems. The study, which excluded cases with partial tear and mucoid pathologies demonstrated the feasibility of high performing CNN tool, with customized CNN architecture and dynamic patched-based sampling with five-sliced 3-D input. The results of the study in [123] suggested the usefulness of pre-operative MRI-detected lateral meniscal extrusion (LME) for estimating lateral meniscus

posterior root tear (LMPRT) in injured knees with an ACL tear. Although there is a significant improvement in application to classifications and inverse problems in the context of DL architectures, the theoretical foundations of DL largely remain heuristic. One such very useful heuristic technique which is widely applied in DL architectures as regularization to avoid overfitting the model on test data is the dropout. This technique randomly discards activations to improve classification accuracy on tests sets and avoid over-fitting by the learning model. These regularization techniques have been improved recently with the proposal of stochastic techniques to further reduce overfitting by DL networks [124].

The recent research addressing the limitations of DL architectures has focused on a theoretical explanation of the working of deep learning frameworks. In [35, 38], authors elaborated the significance of theoretical understanding of deep learning and proved the connection between widely used CNN architectures and celebrated sparse coding theory. The sparse coding theory which has been successfully used in an inverse problem in imaging and classification tasks was shown in [35] to be tightly connected to CNN. The work established a connection between CNN and sparse coding theory and further gave insights to the multilayered version of sparse coding. Further work by [125] pointed out the suboptimal performance of the model presented in [35]. The work in [125] analyzed the proposed multilayered basis pursuit in the context of a combination of synthesis and analysis. Further extending the work on multilayered basis pursuit and its application to explain CNNs and performance on applied problems of classification, Sulam *et al.* in [1] introduced a multilayered basis pursuit framework wherein an  $l_1$  norm penalty was proposed on intermediate representations of the multilayered framework. [1] showed that iterative thresholding algorithms can be used for multilayer basis pursuit and demonstrated the framework effectiveness on classification tasks of general datasets of MNIST, SVHN, and CIFAR-10 with improved performance of thresholding algorithms as compared to the CNNs.

## **CHAPTER 3**

### **IMPLEMENTATION OF MULTI-LAYER CONVOLUTIONAL SPARSE CODING BASED FRAMEWORK FOR COMPRESSED SENSING MR IMAGE RECONSTRUCTION**

Compressive sensing (CS) is a widely used technique in biomedical signal acquisition and reconstruction. The technique is especially useful for reducing acquisition time for MRI signal acquisitions and reconstruction, where effects of patient fatigue and Claustrophobia need mitigation. In addition to improving patient experience, faster MRI scans are important for time-sensitive imaging, such as functional or cardiac MRI, where target movement is unavoidable. Inspired from recent research works of ML-CSC theory utilized for the modeling of deep neural networks, this work proposes a multi-layer basis pursuit framework that combines the benefit from objective-based CS reconstructions and deep learning-based reconstruction by employing iterative thresholding algorithms for successfully training a CS-MRI restoration framework on GPU and reconstruct test images using parameters of the trained model. Extensive experiments show the effectiveness of the proposed framework on four MRI datasets in terms of faster convergence, improved PSNR/SSIM, and better restoration efficiency as compared to the state of the art frameworks with different CS ratios.

The convolutional neural networks have been interpreted with the help of a recently proposed multi-layer convolutional sparse coding framework which comprises a cascade of convolutional sparse layers. This model interprets the forward pass of CNN as pursuit algorithms that aim to estimate the feature maps or nested sparse representation from underlying given data. Although this model has proved pivotal in explaining the theoretical foundations of deep learning frameworks, Its efficient implementations of real-world problems, such as undersampled MR-image reconstruction, are yet to be investigated.

In this chapter a CS-MRI image reconstruction framework based on the ML-CSC model is proposed and implemented multiple MR datasets showing its effectiveness with the help of extensive empirical evidence.

### 3.1 MRI Image Acquisition and CS Reconstruction Model

In biomedical signal capture and reconstruction, compressed sensing (CS) is a widely utilized approach. The technique is particularly beneficial for lowering acquisition time for magnetic resonance imaging (MRI) signal acquisition and reconstruction when patient weariness and Claustrophobia must be considered. In time-sensitive imaging, such as functional or cardiac MRI, where target movement is unavoidable, faster MRI scans are critical. This work proposes a multi-layer basis pursuit framework that combines the benefits of objective-based CS reconstructions and deep learning-based reconstruction by employing iterative thresholding algorithms for successfully training a CS-MRI restoration framework on GPU and reconstructing test images using parameters of the trained model, which is inspired by recent research works on multi-layer convolutional sparse coding (ML-CSC) theory to model deep neural networks. Extensive experiments on four MRI datasets show that the proposed framework is more effective than state-of-the-art frameworks with different CS ratios in terms of faster convergence, improved Peak signal-to-noise ratio(PSNR)/ structural similarity index measure(SSIM), and better restoration efficiency. The MR-CS image acquisition problem can be modelled as:

$$\hat{s} = \arg \min_s \frac{1}{2} \|\Phi\Psi s - y\|_2^2 + \lambda \|s\|_1, \quad (3.1)$$

where  $\Phi$  is measurement matrix,  $\Psi$  are sparsifying basis,  $s$  are sparse coefficients i.e.  $x = \Psi s$ , and  $\lambda \geq 0$ . Typical sparsifying basis  $\Psi$  consists of wavelet, DCT or any other learned dictionary. Substituting sparsifying transform and its sparse coefficient vectors with  $x$ , i.e.

$x = \Psi s$ , and orthonormal sparsifying transform as  $\Psi \cdot \Psi^+ = I$ , in Equation (3.1), we have:

$$\hat{x} = \arg \min_x \frac{1}{2} \|\Phi x - y\|_2^2 + \lambda \|\Psi^+ x\|_1, \quad (3.2)$$

The compressive sensing incorporates the compression into acquisition with measurement matrix  $\Phi \in \mathbb{C}^{M \times N}$ , ( $M \ll N$ ), to infer original signal  $x \in \mathbb{C}^N$  from its measurements  $y = \Phi x \in \mathbb{C}^M$ . The compression ratio is defined as  $M/N$ .

Equation (3.2) is further generalized by replacing the regularization term with learned convolutional filters.

$$\hat{x} = \arg \min_x \frac{1}{2} \|\Phi x - y\|_2^2 + \sum_{l=1}^L \lambda_l g(D_l x), \quad (3.3)$$

The transform matrix in Equation (3.3) is denoted by  $D_l$ , which can be a gradient transform, discrete wavelet transform (DWT), or Discrete Cosine Transform (DCT).  $L$  is the number of layers of the transformation matrix. Here,  $g(\cdot)$  is a  $l_q$  sparse regularization function where  $q \in [0, 1]$ .

### 3.1.1 Solvers for CS MRI Reconstruction Problem

One way to solve Equation (3.3) is the popular variable splitting and auxiliary variables method of ADMM as proposed in [76, 77]. The objective function can be written as:

$$\begin{aligned} \min_x \quad & \frac{1}{2} \|\Phi x - y\|_2^2 + \sum_{l=1}^L \lambda_l g(z_l), \\ \text{s.t.} \quad & z_l = D_l x \quad \forall l \in 1, 2, \dots, L, \end{aligned} \quad (3.4)$$

where  $z_l$  is defined as an auxiliary variable. The minimization problem for a two layer multi-layer basis pursuit problem with constraints can be written as:

$$\min_x \frac{1}{2} \|y - D_1 D_2 x\|_2^2 + \lambda_1 \|x\|_1 + \lambda_2 \|x\|_1, \text{ s.t. } x_1 = D_2 x_2, \quad (3.5)$$



To solve Equation (3.5), an augmented Lagrangian (in normalized form) is constructed to minimize the loss given by:

$$\min_{x_1, x_2, u} \frac{1}{2} \|y - D_1 D_2 x_2\|_2^2 + \lambda_1 \|x_1\|_1 + \lambda_2 \|x_2\|_1 + \frac{\rho}{2} \|x_1 - D_2 x_2 + u\|_2^2, \quad (3.6)$$

After merging the  $l_2$  norm terms, the pursuit of innermost representations is carried out with any basis pursuit formulations like the iterative shrinkage thresholding algorithm (ISTA). This algorithm then updates intermediate representations followed by updating of dual variable  $u$ . As the minimization function is convex, convergence is guaranteed [126].

### 3.1.2 ML-CSC as a Theoretical Foundation For Deep Learning

The heuristic techniques applied to problems in deep learning frameworks have been recently investigated for theoretical explanations of deep learning with the help of celebrated sparse coding theory. Sparse coding theory [2] works on the premise of first learning filters (weights/dictionaries) from given data and then finding the sparse representation maps from those dictionaries for representations of the underlying structure of the data. Once the underlying structure is successfully modeled, the problems of reconstruction on images from noisy measurements, retrieving/reconstructing a signal in compressive sensing domain, and classification of test sets on already trained dictionaries and sparse maps can be done successfully with the help of different algorithms developed over the years and applied successes fully in different domains. The sparse coding theory has been further extended to theoretically explain widely used CNNs namely convolutional sparse coding. The resulting CSC model, where a special circulant and convolutional structure is imposed on dictionaries (which are otherwise traditionally unstructured in sparse coding theory) is defined as a forward pass of CNN[35]. Further work in a multilayered version of CSC has been shown in [1] for convergence analysis and multi-layer basis pursuit for classification performance comparison with CNNs on three public datasets.

### 3.1.3 Multi Layered Basis Pursuit

A signal  $y$  admitting a sparse representation in terms of a dictionary  $D$ , can be represented as  $y = Dx$ , with  $x$  having a sparse structure (solution with fewer non-zero entries). After employing an  $l_1$  norm penalty and considering the problem as NP-hard, the basis pursuit problem is given as:

$$\min_x \|x\|_1 \text{ s.t. } \|y - Dx\|_2^2, \quad (3.7)$$

Papayan *et al.* in [35] extended the basis pursuit problem to multi-layer settings, where a signal  $y$  expressed as  $y = D_1x_1$ , for sparse representations  $x_1$  and possibly convolutional dictionary  $D_1$ . The sparse representations can be further expressed as  $x_1 = D_2x_2$  for another dictionary  $D_2$  with sparse representations  $x_2$ . This framework can be extended to  $L$  number of layers. Under this framework, for an observed signal  $y$ , the deep coding problem in multi-layer settings can be expressed as:

$$\min_{x_i} \|y - D_1x_1\|_2^2 \text{ s.t. } [x_{i-1} = D_ix_i, \|x_i\|_0 \leq s_i]_{i=1}^L, \quad (3.8)$$

This generative multi-layer basis pursuit framework has provided a platform for conveniently analyzing deep neural networks. Alternatively, the work in [125] developed a global pursuit algorithm based on projection interpretation. By imposing an analysis prior on deepest representation, Equation (3.7) can be written as,

$$\begin{aligned} \min_{x_i} \|y - D_{(1,L)}x_L\|_2^2, \\ \text{s.t. } \|x_L\|_0 \leq s_L, [\|D_{(i,L)}x_L\|_0 \leq s_{i-1}]_{i=1}^L, \end{aligned} \quad (3.9)$$

The greedy pursuit algorithm presented in [125] does not scale for high dimensional settings. Sulam *et al.* [1] presented a multi-layer basis pursuit algorithm that could leverage the symbiotic relationship of analysis and synthesis-priors on sparse representations of different layers according to the depth of the multi-layer basis pursuit framework. Specifically,

a convex relation was proposed for (3.9), resulting in multi-layer basis pursuit:

$$\min_x \frac{1}{2} \|y - D_1 D_2 x\|_2^2 + \lambda_1 \|D_2 x\|_1 + \lambda_2 \|x\|_1, \quad (3.10)$$

The model imposed mixture of analysis and synthesis priors, with  $\lambda_1 = 0$  and  $\lambda_2 > 0$  recovering traditional pursuit with factorized global dictionary  $D$ .

### 3.1.4 Iterative Shrinkage Algorithms

Iterative shrinkage algorithms (ISTA) are first-order methods that require matrix-vector multiplications and entry-wise operations. This gives a clear advantage over interior point methods and other solvers, which depend upon second-order information making them computationally expensive in high dimensional settings. ISTA originally proposed in [70] gives convergence in order of  $\mathcal{O}(1/k)$  in functional value and its fast version FISTA [127] has improved convergence in the order of  $\mathcal{O}(1/k^2)$ . The ISTA algorithm is given as follows:

---

#### Algorithm 1 ISTA

---

```

Init  $x^0 \in f(x)$ 
1: for any  $k=0,1,2..$  do
2:    $x^{k+1} = \text{prox}_{\frac{1}{L_k}} g(x - \frac{1}{L_k} \nabla f(x))$ 
3: end for

```

---

### 3.1.5 Layered Basis Pursuit

The layered basis pursuit in context of sparse models and deep learning has been proposed addressing pursuit problems of the form:

$$\hat{x}_i \leftarrow \arg \min_{x_i} \|\hat{x}_{i-1} - D x_i\|_2^2 + \lambda_i \|x_i\|_1, \quad (3.11)$$

[68] and [128] attempt to unfold neural networks with iterative thresholding and minimizing ML-BP. As a result, each representation estimate is required to explain the immediate

layer only and a signal based on generation in global multilayer sparse model settings is not possible.

### 3.2 Proposed ML-CSC Based Framework for CS-MRI

We propose a multi-layer basis pursuit algorithm based on an iterative thresholding algorithm. The algorithm considers the merits of objective-based approaches and network-based approaches to achieve CS-MRI reconstruction. Each ISTA update is fed to an ML-CSC network employing multi-layer ISTA for basis pursuit. The deep network learns the mappings from previous ISTA updates with a fixed number of unfoldings which corresponds to the iterations of ISTA.

#### 3.2.1 Model

Given a decomposed loss function of the form;

$$F(x) = f(x) + g(x),$$

where  $f(x)$  is convex and smooth and  $g(x)$  is smooth,  $L$  is Lipchitz constant, the ISTA algorithm as proximal gradient method finds the minimizer of  $F = f + g$  by iterating the updates given by proximal operator  $g(\cdot)$ :

$$x^{k+1} = \text{prox}_{\frac{1}{L}} g(x^k - \frac{1}{L} \nabla f(x)), \quad (3.12)$$

The traditional ISTA formulation cannot be applied to (3.10) due to presence of no separable composite term in  $g(\cdot)$ . In order to tackle this issue and solve (3.10), gradient mapping<sup>1</sup> approach [129] is analyzed.

---

<sup>1</sup>For a detailed description of gradient operators, see [129], chapter 10.

Given a function of the form  $F = f + g$ , the gradient mapping operator is given by:

$$G_L^{f,g}(x) = L(x - \text{prox}_{\frac{1}{L}}(g(x - \frac{1}{L}\nabla f(x))))), \quad (3.13)$$

The ISTA update for (3.13) can be written as:

$$x^{k+1} = x^k - \frac{1}{L}G_L^{f,g}(x),$$

This update can be considered as gradient mapping step. Since,

- $G_L^{f,g}(x) = \nabla F(x) = \nabla f(x)$  as  $g(x) \equiv 0$ ,
- $G_L^{f,g}(x) = 0$  iff  $x$  is minimizer of  $F(x)$ ,

Essentially  $\Phi^T(\Phi x^{k-1} - y)$  is gradient of the data fidelity term in Equation (3.10). The objective function of ML-BP, can be generally expressed as minimization problem of the form,

$$\min F(x) = f(D_i x) + g_1(D_i x) + g_2(x),$$

for the  $i$ th layer of ML-BP, the proximal gradient mapping method to minimize above objective function takes the update of the form:

$$x_i^{k+1} = \text{prox}_{t g_i}(x_i^k - t \cdot D_i^T G_{1/c}^{f(\cdot), g_{i-1}(D_i \cdot)}(x_i^k)), \quad (3.14)$$

for  $t > 0$  and  $c > 0$ , which are learnable parameters of ML-CSC framework. The proximal of  $g_{i-1}(D_i \cdot)$  involves a composite term, an approximation for representations  $x_i$  is proposed such that  $x_{i-1} = D_i x_i$ . Therefore, for  $i$ th layer of the framework the update of (3.14) can be modified as,

$$x_i^{k+1} = \text{prox}_{t g_i}(x_i^k - t \cdot D_i^T G_{1/c}^{f(\cdot), g_{i-1}}(x_{i-1}^k)), \quad (3.15)$$

with approximation of composite term, the calculation of proximal mapping of  $g$  term becomes soft thresholding that is  $\text{prox}_{g(x)} = \tau_\lambda(x)$ . For a multi-layer model, the ISTA

update step can be written as:

$$x_i^{k+1} \leftarrow \mathcal{T}_{t,\lambda_i} \left[ x_i^k - \frac{t}{c} D_i^T (x_{i-1}^k - \mathcal{T}_{t,\lambda_i} (x_{i-1}^k - c D_{i-1}^T (D_i x_{i-1}^k - y))) \right], \quad (3.16)$$

The algorithm employing multi-layer ISTA takes images in  $k$ -space along with their CS measurements, learns the mappings from CS measurements with ground-truth images, and saves the mappings in the form of model parameters of ML-CSC. The testing module reconstructs the test images with the help of trained ML-ISTA learned parameters and gives PSNR/SSIM of recovered images. The framework initializes the dictionaries with Xavier initialization [130] for all layers. The multilayered basis pursuit algorithm employing an iterative thresholding algorithm for multi-layer pursuit is given in Algorithm-2.

---

**Algorithm 2** ML-ISTA for CS-MRI

---

**Require:**  $\{y_i, x_i\}_{i=1}^N$ , dictionaries  $D_i$  and  $\lambda_i$

**Ensure:** Model containing parameters of neural network for CS-MRI image restoration.

$\{y_i, x_i\}_{i=1}^N$  are images in  $k$ -space and their corresponding CS measurements,  $N$  is size of trainset

1: Init Set  $x_0^k = y$ ,  $\forall k$  and  $x_L^1 = 0$

2: **for**  $k = 1 : K$  **do**

$\hat{x}_i \leftarrow D_{(i,L)} x_L^k \quad \forall_i [0, L - 1]$

3: **for**  $i=1:L$  **do**

$x_i^{k+1} \leftarrow \mathcal{T}_{c_i, \lambda_i} (\hat{x}_i - c_i D_i^T (D_i \hat{x}_i - x_{i-1}^{k+1}))$  ▷ Find set of representations

4: **end for**

5: **end for** ▷ Save model with Epoch number

---

The training module (Figure-3.1) takes pairs of images, their corresponding CS measurements in  $k$ -space along with undersampling masks and trains the model parameters. The learning Algorithm-2 uses ML-ISTA unfoldings to increase depth of learning framework without incurring cost of additional parameters. The testing framework takes the test

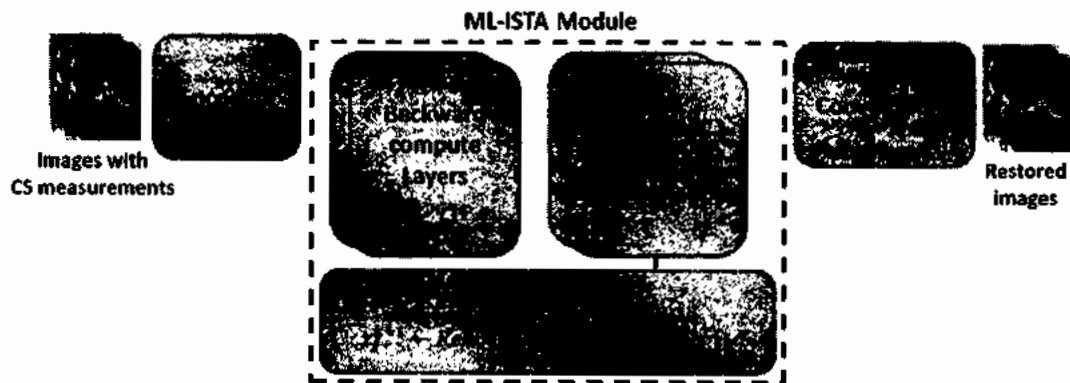


Figure 3.1: Framework for CS-MRI.

set images, undersampling masks, and the parameters of the saved model to reconstruct undersampled test images and comparing the metrics of PSNR and SSIM with ground truth images along with restoration time.

The pseudo-code for training and testing the CS-MRI framework is given below.

---

**ML-ISTA CS-MRI PSEODOCODE**

---

- **Training ML-ISTA with MRI Train Data:**

**Require:** Input: MR images, their CS measurements, random masks,  $D_i$  and  $\lambda_i$

**Ensure:** Model containing parameters of ML-ISTA for CS-MRI reconstruction.

- Run encoding to find representations for given layers
- for** Unfoldings= 1,2..., **do**
- Find representations using ML-ISTA with global pursuits
  - Train and save model with EPOCH number and minimum loss

**end for**

---

- **Reconstruct Test set with trained model.**

**Require:** Trained model, MR test images, undersampling masks

**Ensure:** Reconstructed images with PSNR , SSIM and reconstruction CPU/GPU time

**for** i=1:N **do**

- Run ML-ISTA with learned dictionaries from trained model

- Reconstruct MR images
- Calculate PSNR and SSIM of testset ▷ N is the size of test set;

**end for**

---

### 3.2.2 Architecture

The proposed framework takes advantage of iterative thresholding-based pursuit algorithms along with the learning capability of the ML-CSC model and maps each learning update to an ML-ISTA-based unfolding. The sparsifying transform is replaced with ML-ISTA learning, which is denoted by  $\mathcal{D}(\cdot)$ . In the multilayered basis pursuit algorithm,  $\mathcal{D}(\cdot)$  consists of six convolutional layers with ISTA-based pursuit (Equation (3.16)), as given in Algorithm-2. After replacing sparsifying transform with learnable parameter  $\mathcal{D}(\cdot)$ , Equation (3.2) becomes:

$$\hat{x} = \arg \min_x \frac{1}{2} \|\Phi x - y\|_2^2 + \lambda \|\mathcal{D}(x)\|_1, \quad (3.17)$$

The ISTA update for CS reconstruction is given by;

$$\mathcal{I}^k = x^{k-1} - \rho \nabla \left( \frac{1}{2} \|\Phi x - y\|_2^2 \right), x^k = \arg \min_x \frac{1}{2} \|\Phi x - \mathcal{I}^k\|_2^2 + \lambda \|\mathcal{D}(x)\|_1, \quad (3.18)$$

The ISTA update consists of the gradient of data fidelity term used for intermediate reconstruction of images.

#### ISTA-Module

ISTA module generates a representation of intermediate results with learnable step size  $\rho$ .

#### Proximal Mapping

The  $x^k$  takes ISTA update as input and computes  $x$  according to Equation (3.17).



### 3.2.3 Training Loss for CS-MRI Framework

Given pairs of MR images with their corresponding CS measurements, the framework produces intermediate reconstructions denoted by  $x_i^{rec}$ . The loss function is designed to seek reduced discrepancy between the input images and the intermediate reconstruction satisfying symmetry constraint of  $\tilde{\mathcal{D}}^k \mathcal{D}^k = I, \forall k = 1, 2, \dots$ , in check. The end to end loss is thus defined as:

$$\begin{aligned} \mathcal{L}_{discrepancy} &= \frac{1}{IN} \sum_{i=1}^N \|x_i^{rec} - x_i\|_2^2 \\ \mathcal{L}_{constraint} &= \frac{1}{IN} \sum_{i=1}^N \sum_{k=1}^U \|\hat{\mathcal{D}}^k(\mathcal{D}^k(x_i)) - x_i\|_2^2 \\ \mathcal{L}_{total} &= \mathcal{L}_{discrepancy} + \lambda \mathcal{L}_{constraint} \end{aligned} \quad (3.19)$$

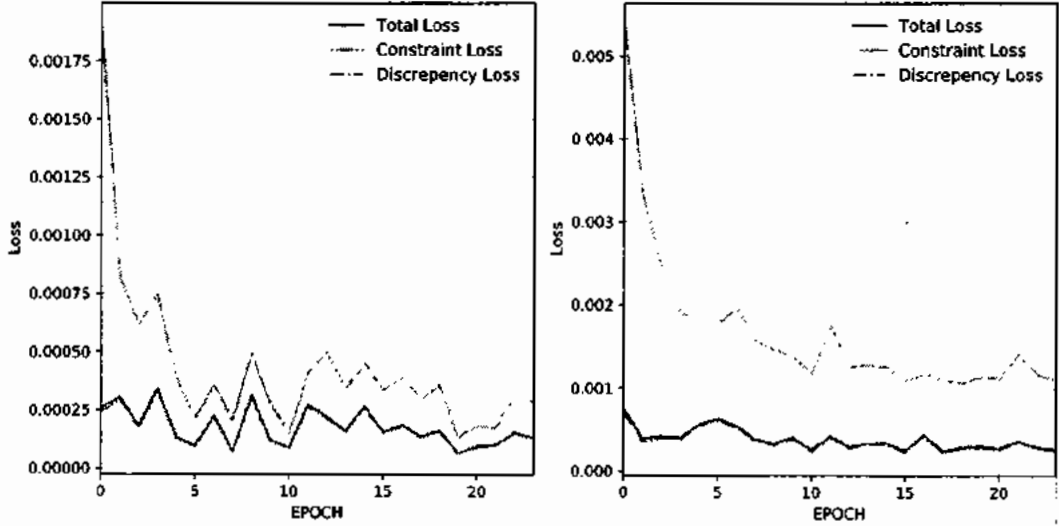


Figure 3.2: Train set losses for brain MR (left) and knee MR (right) datasets.

where image size  $I$ , number of images  $N$ , ML-ISTA unfoldings  $U$ , learning rate  $\lambda$  are parameters used in above equations. A learning rate of  $1e-4$  with Adam optimizer [131] and learning rate decay factor of 0.2 is used in Pytorch implementation. The discrepancy,

constraint and total losses along with the number of EPOCHs are given in Figure 3.2 for brain and knee MR datasets.

### 3.2.4 Datasets

The first dataset used for training the CS-MRI framework comprises 622 knee MR images of adult patients (Male and Female) aged between 18 to 40 years, with the coronal view and Proton density (PD) fat saturation, collected from a 1.5T imaging unit (Siemens MAGNETOM Symphony) at Hospital Kuala Lumpur (HKL). The patients underwent knee MRI scans for anterior cruciate ligament (ACL) tear diagnosis and a certified radiologist labeled the images for ACL tear types of complete, partial, and normal knee categories. The second dataset has been taken from [68], with 850 brain images acquired in axial and sagittal planes. A test set size of 5% is used to validate the performance of the testing framework in both data sets for analyzing PSNR and SSIM on reconstructed images. The third dataset comprising brain images is of the normal aging coronal plane with 123 slices, matrix size ( $256 \times 256$ ), and is publicly available on the AANLIB database of Harvard medical school at <http://www.med.harvard.edu/AANLIB/home.html> [132] used by Murad *et al.* in [133]. We have also used a single image consisting of a single-slice (axial T2-weighted reference brain image) dataset of size ( $256 \times 256$ ), Vivo MR scans from American Radiology Services as used by Prasad *et al.* in [51]. The framework was trained on a workstation with core i9-9900 and RTX-2080Ti GPU. PSNR along with CS-ratios and reconstruction time is given in Table-3.2 for ADMM-Net [134], ISTA-Net [68] and our proposed framework.

### 3.3 Empirical Results

Figure-3.3 gives ground truth images along with the restored images emphasizing the ACL and three tear types. The three types of masks which are used for undersampling in  $k$  space for CS-MRI reconstruction algorithms [77] are given in Figure 3.4. In our proposed framework, the random mask has been used for undersampling in  $k$  space, and results were

obtained for different CS ratios. Results for CS-MRI reconstruction of four different benchmark test sets, CS ratios along with PSNR/SSIM, and restoration times are given in Table 3.1. PSNR/SSIM for the brain MR dataset of [68] is given in Table 3.2 for comparison. The image-wise PSNR/SSIM along with the average PSNR/SSIM plots of reconstructed images for both datasets are given in Figure-3.5 for Brain MR images and Figure-3.6 for Knee MR images.

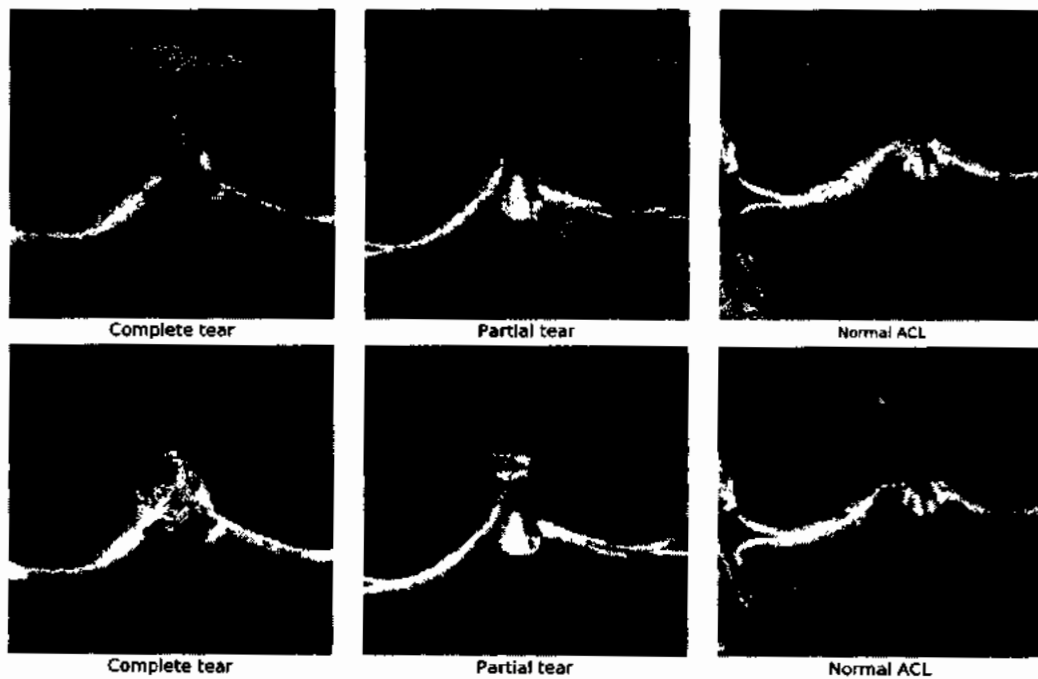


Figure 3.3: Knee MR ground truth images (top) and reconstructed images (bottom), for ACL tear identification with 5-fold acceleration factor

### 3.3.1 Clinical Application

MR imaging plays a crucial role in the medical imaging pipeline comprising diagnosis, treatment planning, treatment delivery, and follow-ups. ACL tear is one of the common sports injuries, diagnosed with MRI scans by clinicians to compliment and support physical examinations like the Lachman, pivot shift, and anterior drawer tests of ACL tear patients for assessment of the knee injury.

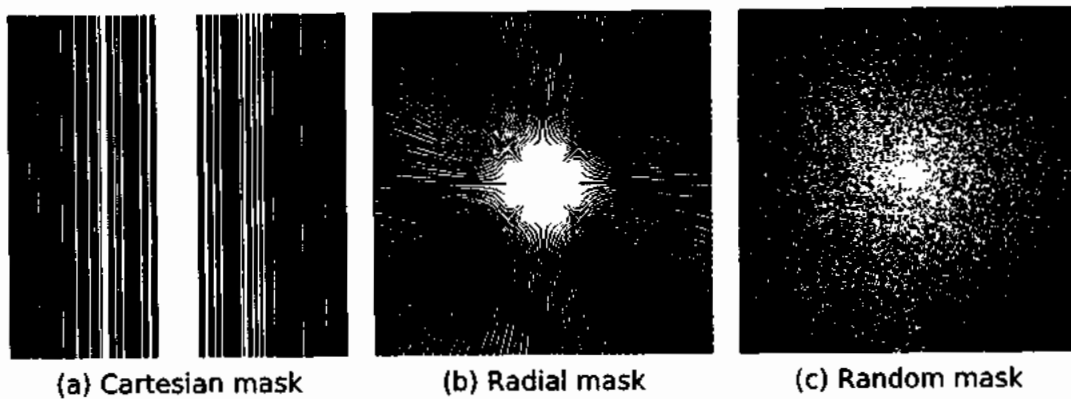


Figure 3.4: Examples of undersampling mask

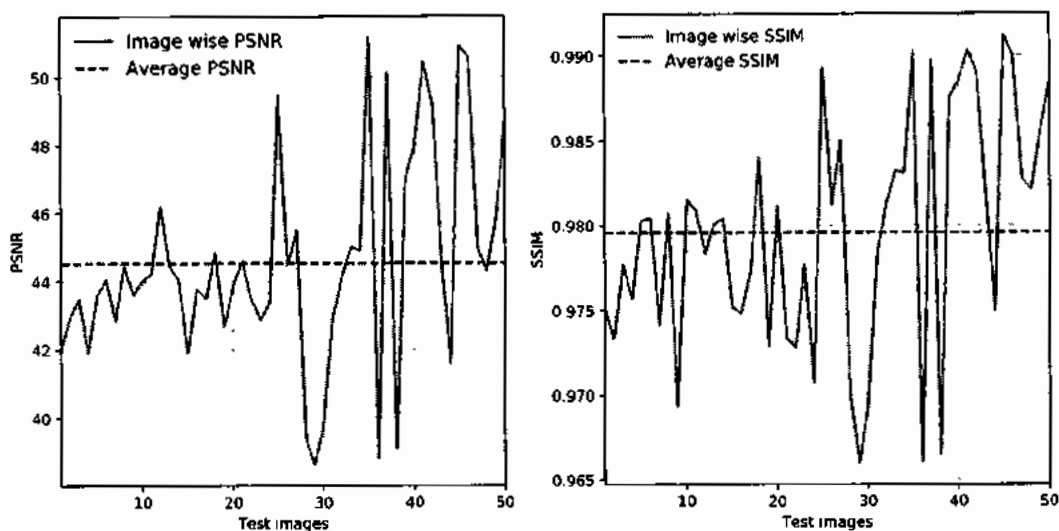


Figure 3.5: PSNR/SSIM of brain MR[68] test set

For CS-MRI practical applications, the reconstructed images from undersampled acquisitions must be of high quality for radiologists to successfully diagnose different types of ACL tears, the status of the meniscus, and pathologies of patients. The proposed framework successfully restores undersampled images and gives a good PSNR/SSIM for the 5-Fold acceleration factor (20% CS-ratio), as shown in Table-3.1, and Figure-3.3 for both datasets.

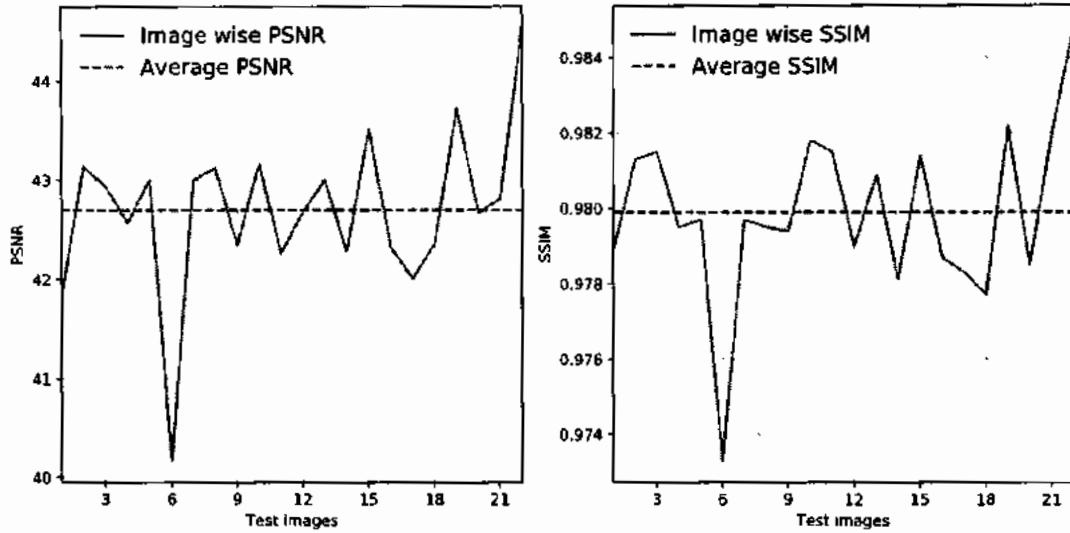


Figure 3.6: PSNR/SSIM of Knee MR test set

Table 3.1: Average PSNR/SSIM and restoration times of reconstructed brain and Knee MR images

Dataset	CS Ratio				GPU Time
	20%	30%	40%	50%	
	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	
Brain[51]	38.97/0.9641	42.30/0.9799	45.16/0.9876	49.06/0.9924	0.0216s
Brain[68]	39.25/0.9551	41.50/0.9689	43.66/0.9774	45.96/0.9855	0.0688s
Brain[132]	31.68/0.8782	35.21/0.9234	38.35/0.9554	41.60/0.9724	0.0409s
Knee	36.93/0.9262	39.68/0.9569	42.04/0.9738	43.44/0.9824	0.0977s

### 3.3.2 Discussion

The proposed framework has been trained on a brain MR dataset of 800 images and tested on three benchmark test sets with encouraging and comparable results demonstrating its generalizability. The zoomed part of regions of interest for ACL tear pathology showed reasonable reconstruction quality of images in terms of PSNR/SSIM as given in Table 1. The results of knee datasets exhibit better performance as compared to the brain dataset as shown in Figure 5 and Figure 6. One reason for these improved results can be the acquisition of the knee MR dataset, as all the train and test sets contain images from the

Table 3.2: Algorithms along with the PSNR and reconstruction times of brain MR images

Algorithms	CS Ratio				GPU/CPU Time
	20%	30%	40%	50%	
	PSNR	PSNR	PSNR	PSNR	
ADMM-Net	37.17	39.84	41.56	43.00	0.9535s/-.-
ISTA-Net <sup>+</sup>	38.73	40.89	42.52	44.09	0.1437s/-.-
<b>ML-CSC</b>	<b>39.25</b>	<b>41.50</b>	<b>43.66</b>	<b>45.96</b>	<b>0.0688s/2.8s</b>

coronal plane. On the other hand, the brain dataset comprises images from axial and sagittal planes. The restoration results showed better performance in terms of average PSNR/SSIM with state-of-the-art, comparative performance on a single image, and dataset of 123 MR images of the brain. The framework also demonstrates the relative benefits of training frameworks on different custom datasets acquired in specific planes (e.g. axial or coronal) and restoring images in relevant planes with better quality. As the test framework only uses saved parameters from the training framework, it can be readily integrated into clinical settings for improving the efficiency of the MR imaging pipeline.

### Summary

In this chapter, a CS-MRI restoration framework based on multi-layer convolutional sparse coding, employing iterative thresholding algorithms for basis pursuits to learn parameters of nonlinear mappings from undersampled MR images acquired in  $k$ -space was proposed. Empirical results showed that the proposed MRI-CS framework learned the desired mapping from CS measurements effectively for brain and knee MR images after training on masks with different CS ratios. Furthermore, at the cost of a smaller increase in learnable parameters of a deep neural network, the CS reconstruction results are improved with reasonable reconstruction time as compared to the state of the art. The images can also be successfully restored on a CPU-based machine with parameters learned from models trained on GPU in a reasonable amount of time. The generalized structure and application to four different datasets also signify the applicability of the framework in practical appli-

cations in the MR imaging pipeline for increasing the efficiency of clinical diagnosis.

## **CHAPTER 4**

### **IMPLEMENTATION OF MULTI-LAYER CONVOLUTIONAL SPARSE CODING BASED FRAMEWORK FOR BIOMEDICAL IMAGE CLASSIFICATION**

#### **PROBLEM**

Deep learning architectures have been extensively used in recent years for the classification of biomedical images to assist clinicians in the diagnosis and treatment management of patients with different health conditions. These architectures have demonstrated expert-level diagnosis, and in some cases, surpassed human experts in diagnosing health conditions. The automation tools based on deep learning frameworks have the potential to transform all stages of the medical imaging pipeline from image acquisition to interpretation and analysis. One of the most common areas where these techniques are applied is knee MR image classification for different types of Anterior Cruciate Ligament (ACL) tears. If properly and timely managed, the diagnosis and treatment of ACL tear can avoid further degradation of patients' knee joints and can help slow the process of subsequent occurrence of knee arthritis. In this work, we have implemented a novel classification framework based on multilayered basis pursuit algorithms inspired by recent research work in the area of the theoretical foundation of deep learning with the help of celebrated sparse coding theory. We implement an optimal multilayered Convolutional Sparse Coding (ML-CSC) framework for the classification of a labeled dataset of knee MR images with the coronal view and compare the results with a traditional convolutional neural network (CNN) based classifiers. Empirical results demonstrate the effectiveness of the ML-CSC framework and show that the framework can successfully learn distinct features on small a dataset and achieves a good efficiency of more than 92% without employing regularization techniques and extensive training on large datasets. In addition to 95% average accuracy on the presence and absence of ACL tears, the framework also performs well on the imbalanced and



challenging classification of partial ACL tear with 85% accuracy.

Deep learning techniques have been used for classification problems of images with improving accuracies on unseen test sets. In biomedical imaging, the potential of deep learning techniques to assist clinicians in diagnosis and treatment management has been explored in research works. One of the areas where deep neural networks are applied for classification is labeled MRI images of knee Anterior Cruciate ligament (ACL) tear. In this chapter, a classification framework based on a multilayer convolutional sparse coding framework is proposed. With the help of extensive empirical experiments on knee ACL tear data set, the effectiveness of the proposed model is shown as compared to the state-of-the-art traditional CNN-based classifiers.

In this work, we have implemented an optimal framework for multilayered basis pursuit algorithms and demonstrated through experiments its applicability to the classification of biomedical images. The novel architecture, which is optimized for the classification of biomedical images, trained on an original dataset of knee MR images, achieves a good average test accuracy of more than 92% and class-wise test accuracy of 95%, outperforming traditional CNN without adding regularization parameters and computational complexity. In the following, a brief overview of the clinical background of ACL and its tears, a brief overview of CNNs, the multilayered sparse coding model, and image classification in the context of CNN and ML-CSC are given. Furthermore, overviews of iterative thresholding algorithms for single layer basis pursuit and its extended version in multilayer settings along with multilayered basis pursuit are given. Lastly the ML-CSC model, its implementations for classification of biomedical image dataset, and experimental results of image classification of the knee ACL tear are given.

#### 4.1 Anterior Cruciate Ligament Tear - Background

An anterior cruciate ligament (ACL) is one of the key ligaments that help stabilize the knee joint. These ligaments connect the thighbone (femur) to the shinbone (tibia) (Figure 1). Injuries of ACL are most often a result of low-velocity, noncontact, deceleration injuries, and contact injuries with a rotational component. A complete tear is characterized by rupture of the ligaments and partial tear by stretching of the ligaments becoming loose and damaged. The MR images with a partial tear, normal knee, and complete ACL tears are given from dataset used in this work (Figure 4.1)

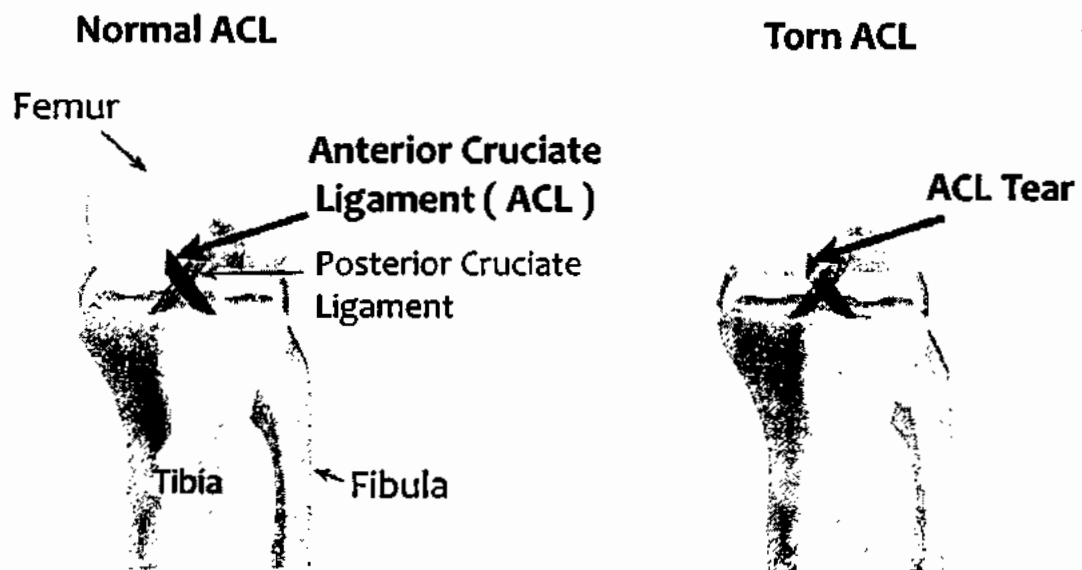


Figure 4.1: Knee anatomy with normal and torn ACL tears (Coronal view)

The diagnosis process involves an emphasis on the history and physical examination of affected patients. The Lachman, pivot shift, and anterior drawer tests are three types of physical examinations performed on ACL tear patients for assessment of the injury. Out of these three tests, the anterior drawer test has the highest sensitivity of 94% [135]. MRI examination coupled with physical examination helps clinicians in identifying ACL tear

types in addition to identifying bone bruising, which is present in most of the patients with an ACL tear. Once the ACL tear is diagnosed, the treatment plan is devised by clinicians for rehabilitation or surgical intervention according to patient condition and medical profile. Studies have been reported that, in some cases, an average radiologist must interpret one image every 3-4 seconds in an 8-hour workday to meet the workload demand. Under these conditions, the errors are inevitable for radiology tasks where visual perception and decision making are involved [136]. An integrated AI system in the imaging pipeline, which enables the trained radiologist to receive pre-screened images would enable better decision making especially in heavy workloads in addition to helping in the diagnosis of ACL tear of knee injury patients in regions where trained MSK - radiologists are difficult to access. Besides, the ability of machines to scan large amounts of data enables them to generalize the classification algorithms for better decision-making.

#### **4.2 Deep Neural Networks for Classification**

With improvements in hardware for processing of a large number of images, it became feasible to train large neural networks for classification tasks on datasets of different sizes. The seminal work of [117] significantly improved the state of the art classification of general images using graphical processing units. The work achieved an error rate of more than 15% on the ImageNet dataset. This error rate has been improved significantly since then on general datasets available for research purposes. In addition to general datasets, Rajpurkar *et al.* in [118] implemented a 121 layer deep neural network for radiologist-level pneumonia detection on chest X-rays. The algorithm was trained on a large dataset of Chestx-rays14 [119]. Rajpurkar *et al.* in [137] developed a 33 layer CNN for detecting a wide range of heart arrhythmias from single-lead ECG records. Liu *et al.* in [121] used classification performance of DL networks as compared to clinical reports for binary classification (tear or nor tear presence), and concluded that there is no significant difference between the two. The study by Chang *et al.* in [122], demonstrated the feasibility of a high-performing

CNN tool to detect complete ACL injury with over 96% test accuracy for binary classification problems. The study, which excluded cases with partial tear and mucoid pathologies demonstrated the feasibility of high performing CNN tool, with customized CNN architecture and dynamic patched-based sampling with five-sliced 3-D input. The results of the study in [123] suggested the usefulness of preoperative MRI-detected lateral meniscal extrusion (LME) for estimating lateral meniscus posterior root tear (LMPRT) in injured knees with an ACL tear. Although there is a significant improvement in application to classifications and inverse problems in the context of DL architectures, the theoretical foundations of DL largely remain heuristic. One such very useful heuristic technique which is widely applied in DL architectures as regularization to avoid overfitting the model on test data is a dropout. This technique randomly discards activations to improve classification accuracy on tests sets and avoid over-fitting by the learning model. These regularization techniques have been improved recently with the proposal of stochastic techniques to further reduce overfitting by DL networks [124].

The recent research addressing the limitations of DL architectures has focused on a theoretical explanation of the working of deep learning frameworks. In [35, 38], authors elaborated the significance of theoretical understanding of deep learning and proved the connection between widely used CNN architectures and celebrated sparse coding theory. The sparse coding theory which has been successfully used in an inverse problem in imaging and classification tasks was shown in [35] to be tightly connected to CNN. The work established a connection between CNN and sparse coding theory and further gave insights to the multilayered version of sparse coding. Further work by [125] pointed out the suboptimal performance of the model presented in [35]. The work in [125] analyzed the proposed multilayered basis pursuit in the context of a combination of synthesis and analysis. Further extending the work on multilayered basis pursuit and its application to explain CNNs and performance on applied problems of classification, Sulam *et al.* in [1] introduced a multilayered basis pursuit framework wherein an  $l_1$  norm penalty was proposed on inter-

mediate representations of the multilayered framework. [1] showed that iterative thresholding algorithms can be used for multilayer basis pursuit and demonstrated the framework effectiveness on classification tasks of general datasets of MNIST, SVHN, and CIFAR-10 with improved performance of thresholding algorithms as compared to the CNNs.

In this work, we have implemented an optimal framework for multilayered basis pursuit algorithms and demonstrate through experiments its applicability to the classification of biomedical images. The novel architecture, which is optimized for classification of biomedical images, trained on an original dataset of knee MR images, achieves a good average test accuracy of more than 92% and class-wise test accuracy of 95%, outperforming traditional CNN without adding regularization parameters and computational complexity.

### 4.3 Convolutional Sparse Coding Model-The Multilayered Basis Pursuit

Sparse coding theory works on the premise of first learning filters (weights/dictionaries) from given data and then finding their sparse representation from those dictionaries for the representation of given images. Once the underlying structure is successfully modeled through sparse coding theory, the problems of reconstruction on images from noisy measurements, retrieving/reconstructing a signal in compressive sensing domain, and classification of test sets on already training dictionaries and sparse maps can be done successfully with the help of different algorithms developed over the years and applied successes fully in different domains. In sparse coding theory, formally, a given signal  $y$  admits a sparse representation in terms of a dictionary  $D$ , if  $y = Dx$ , and  $x$  is sparse. Given dictionary  $D$ , the celebrated basis pursuit problem with  $l_1$  norm penalty is formulated as,

$$\min_x \|x\|_1 \text{ s.t. } \|y - Dx\|_2^2 \quad (4.1)$$

This modeling theory was extended in [35] to multilayer settings, providing a connection between sparse coding theory and state-of-the-art DL architectures. The traditional sparse

coding model assumes the dictionaries without any structure. Whereas in CSC, which is a special form of sparse coding [2], a special structure on learned dictionaries is imposed with filters banded together and concatenated in circulant form. In the multilayered version of CSC, which is an extended version of CSC, the sparse feature maps thus obtained from one layer are then treated as input to the second layer, and dictionary learning and sparse coding steps are repeated for subsequent layers. The CSC model represents a signal of interest as multiplication of dictionaries  $D$  and sparse vectors  $x$ . The deep learning problem in the context of sparse coding theory, which is shown as a theoretical explanation of CNNs [35], can be formulated as follows. For a global signal  $X$ , convolutional dictionaries  $D$  and sparse vectors  $x$ , and  $k$  number of layers, the deep pursuit problem is defined as [125]:

$$\begin{aligned}
 \text{Find } [x_i]_{i=1}^k \quad \text{s.t.} \quad & \|y - D_1 x_1\|_2^2 < \epsilon \\
 & x_{i-1} = D_i x_i \quad \forall \quad 2 \leq i \leq k \\
 & \|x_i\|_0 \leq s_i \quad \forall \quad 1 \leq i \leq k
 \end{aligned} \tag{4.2}$$

A convex relaxation proposed in [1] for deep pursuit problem result in multilayered basis pursuit. For a two-layer model, the problem can be formulated as,

$$\min_x \frac{1}{2} \|y - D_1 D_2 x\|_2^2 + \lambda_1 \|D_2 x\|_1 + \lambda_2 \|x\|_1 \tag{4.3}$$

In case  $\lambda_1, \lambda_2 = 0$  and  $\lambda_1 > 0$ , the above formulation is equivalent to traditional basis pursuit with global dictionary. With  $\lambda_1, \lambda_2 > 0$ , analysis priors are imposed on set of sought after representations  $x$  with regularized solutions as a result.

### 4.3.1 ML-CSC for Classification

Given sparse vectors  $\Gamma^*$  and dictionaries  $D$ , the classifications problem can be formulated in deep sparse coding context as,

$$\min_{(D,U)} \sum_j l(f(h(X_j), U, \Gamma^*(X_j), D)) \quad (4.4)$$

Where sparse representations  $\Gamma$  are fed to the classifier after dictionary learning, multilayer basis pursuit, and training of the classifier.

## 4.4 Algorithms

DL architectures and algorithms traditionally deal with high dimensional settings where second-order methods result in prohibitive computational complexity and slow convergence rates. The proximal gradient descent which uses first-order approximations for updating its optimization steps is, therefore, a suitable choice for multilayer basis pursuit due to its dependence on sparse prior terms instead of the convex term [129]. This algorithm only needs to calculate the sub-gradients of convex term and proximal mapping associated with update depending on sparse prior. The convergence analysis is done in terms of the number of iterations of the algorithm.

### 4.4.1 Layered Basis Pursuit

The layered basis pursuit (L-BP) given in [35], addresses sequence of pursuit of the form:

$$\hat{x}_i \leftarrow \arg \min_{x_i} \|\hat{x}_{i-1} - Dx_i\|_2^2 + \lambda_i \|x_i\|_1 \quad (4.5)$$

where  $x_0 = y$  and  $i=1$  to  $k$ .

These algorithms [68],[128], which present heuristic approximation do not minimize Equation (4.3) and each layer is required to explain next layer only so cannot be used to generate

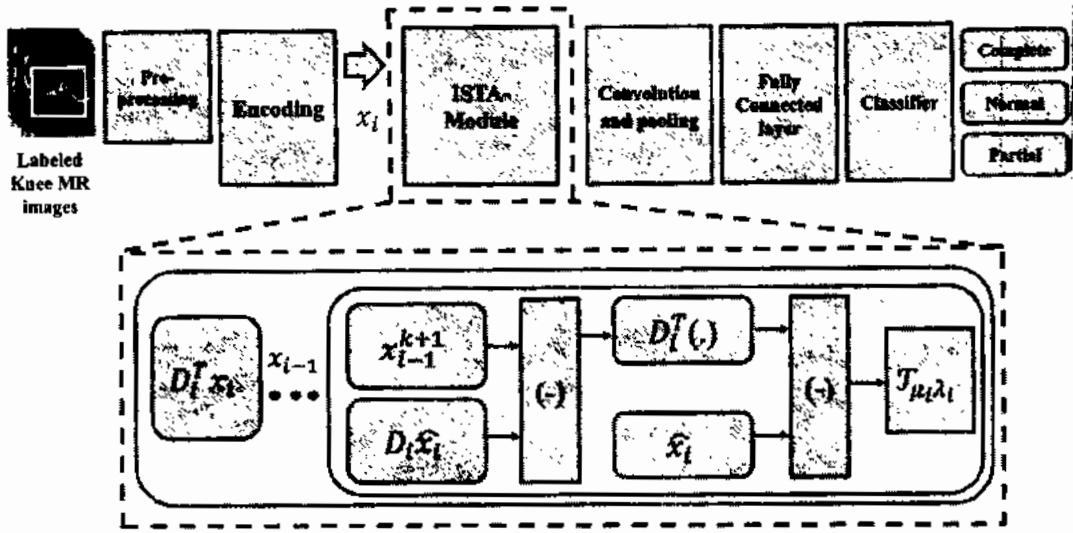


Figure 4.2: Multilayered basis pursuit based framework for MR image classification.

signal according to multilayer sparse model. Algorithm-3 for layered basis pursuit which

---

#### Algorithm 3 ML-BP

---

Input signal  $y$ , dictionaries  $D_i$   
 Init Set  $\hat{x}_0 = y$ ,  
 1: **for**  $i = 1 : k$  **do**  
 2:  $\hat{x}_i = \begin{cases} \mathcal{H}(D_i^T \hat{x}_{i-1}) \\ P_1(D_i, \hat{x}_{i-1}, \lambda_i) \end{cases}$   
 3: **return**  $[\hat{x}_i]_{i=1}^k$  ▷ Set of representations  
 4: **end for**

---

seeks sparse maps  $x$ , subject to constraints given in of Equation(4.2) for  $P_1$  term (which is basis pursuit) and thresholding operator  $\mathcal{H}$  at each layer of the neural network.

#### 4.4.2 Iterative Thresholding Algorithms (ISTA)

ISTA is a first-order method for optimizing functions comprising composite terms originally proposed in [70]. A faster version of this algorithm FISTA was proposed in [127] introduced momentum term, giving improved convergence rates. These algorithms require matrix-vector multiplications, therefore are appealing due to their low complexity. The ISTA provides convergence in function value in the order of  $\mathcal{O}(1/k)$  and its fast version



FISTA provides better convergence rate in the order of  $\mathcal{O}(1/k^2)$ . The proximal gradient method ISTA works by iterating the updates given by the proximal operator. As  $g(\cdot)$  in

---

**Algorithm 4** ISTA

---

Init  $x^0 \in f(x)$   
1: **for** any  $k=0,1,2$ . **do**  
2:      $x^{k+1} = \text{prox}_{\frac{1}{L}} g(x - \frac{1}{L_k} \nabla f(x))$   
3: **end for**

---

Equation(4.3) is sum of  $l_1$  composite terms so, application of ISTA algorithm is not feasible. Another feasible alternative the generalized LASSO [138] can also be computationally expensive due to the requirement of inversions of linear operators during optimization. The iterative algorithm employing re-weighted  $l_2$  norm approaches proposed in [139], for compressive sensing also requires iterative matrix inversions and is thus computationally expensive.

#### 4.4.3 Multi-layer ISTA and FISTA

For a composite model comprising a smooth and convex term  $f(x)$  and convex and not necessarily smooth term  $g(x)$ , the objective function is given by

$$F(x) = f(x) + g(x)$$

The gradient mapping is the operator given by:

$$G_L^{f,g}(x) = L(x - \text{prox}_{\frac{1}{L}} g(x - \frac{1}{L_k} \nabla f(x))) \quad (4.6)$$

Where  $L$  is Lipschitz constant.

The ISTA update step for Equation(4.6) is given by,

$$x^{k+1} = x^k + \frac{1}{L} G_L^{f,g}(x) \quad (4.7)$$

The optimization problem of multilayer basis pursuit is given by

$$\min_x F(x) = f(D_2x) + g_1(D_2x) + g_2(x)$$

The sparse representations for second layer are

$$F(x_2) = f(D_2x_2) + g_1(D_2x_2) + g_2(x_2)$$

The update for the gradient mapping method is given by:

$$x_2^{k+1} = \text{prox}_{tg_2}(x_2^k - t.G_{1/c}^{f(\cdot),g_1(D_2\cdot)}x_2^k) \quad (4.8)$$

Here  $c$  and  $t$  are constants with specific bounds for convergence of the subject algorithms. As  $g_1(D_2\cdot)$  is composite term in Equation (4.8), in order to avoid calculating its proximal mapping, the term  $D_2x_2$  is approximated with  $x_1$  that is  $x_1 = D_2x_2$ . The approximation results in calculation of proximal mapping of  $x_1$  in in Equation (4.8). The update for Equation (4.8) becomes:

$$x_2^{k+1} = \text{prox}_{tg_2}(x_2^k - t.G_{1/c}^{f(\cdot),g_1(D_2\cdot)}x_2^k) \quad (4.9)$$

Consequently, the proximal mapping of composite term after approximation becomes soft thresholding of  $x_1$  which is equal to  $x_2 = \mathcal{T}_{t\lambda_1}x_1$ . The update step for ML-ISTA after above approximation becomes:

$$x_2^{k+1} \leftarrow \mathcal{T}_{t\lambda_2}[x_2^k - \frac{t}{c}D_2^T(x_1^k - cD_1^T(D_2x_1^k - y))] \quad (4.10)$$

Algorithms for ISTA and FISTA in multilayer settings are described in Algorithm 5 and Algorithm 6 respectively.

---

**Algorithm 5 ML-ISTA**

---

Input signal  $y$ , dictionaries  $D_i$  and  $\lambda_i$ ;

Init Set  $x_0^k = y, \forall k$  and  $x_L^1 = 0$

1: **for**  $k=1:K$  **do**

2:  $\hat{x}_i \leftarrow D_{(i,L)} x_L^k \quad \forall_i [0, L-1]$

3: **for**  $i = 1 : L$  **do**

4:  $x_i^{k+1} \leftarrow \mathcal{T}_{c_i \lambda_i}(\hat{x}_i - c_i D_i^T (D_i \hat{x}_i - x_{i-1}^{k+1}))$

▷ Set of representations

5: **end for**

6: **end for**

---

---

**Algorithm 6 ML-FISTA**

---

Input signal  $y$ , dictionaries  $D_i$  and  $\lambda_i$ ;

Init: Set  $x_0^k = y, \forall k$  and  $z = 0$ ;

1: **for**  $k=1:K$  **do**

2:  $\hat{x}_i \leftarrow D_{(i,L)} z \quad \forall_i [0, L-1]$

3: **for**  $i = 1 : L$  **do**

4:  $x_i^{k+1} \leftarrow \mathcal{T}_{\mu_i \lambda_i}(\hat{x}_i - \mu_i D_i^T (D_i \hat{x}_i - x_{i-1}^{k+1}))$

5:  $t_{k+1} \leftarrow \frac{1 + \sqrt{1 + 4t_k^2}}{2}$

6:  $z \leftarrow x_L^{k+1} + \frac{t_k - 1}{t_{k+1}} (x_L^{k+1} - x_L^k)$

▷ Set of representations

7: **end for**

8: **end for**

---

The FISTA algorithms incorporate the momentum term which improves the convergence rate. The framework for classification with iterative thresholding algorithms is given in Figure 4.3 and pseudocode is presented in Appendix-A. The ISTA module described in Figure 4.3 computes representations according to Algorithm 4. First, the encoded feature maps are backward computed for the three layers framework, and the iterations and unfoldings progress according to the ML-ISTA algorithm in Figure 4.3. The (-) sign given in

Figure 4.3 depicts subtraction of resulting representations after convolution and transposed convolution operations are carried out with dictionaries  $D$ . The number of unfoldings inside the ISTA module enables the shallow network to increase depth without having any impact on the number of parameters.

#### 4.4.4 Unfolded Iterative Algorithms as Neural Networks

Unfolded iterative algorithms are successfully used in recent research works [67, 140, 141, 142], for solving sparse recovery problems. To speed up the computational cost associated with approximation algorithms, the work in [140] showed a combination of optimization and neural networks to produce deterministic functions to successfully approximate parsimonious/sparse models resulting in a significant reduction in computational time for applications requiring real-time performance such as image modeling, robust face modeling, audio sources separation and robust speaker recognition. The work in [141], proposed a partial weight coupling structure to learned iterative thresholding algorithms (LISTA) and support selection for improved convergence rate with experimental demonstrations. A two-layer ISTA network is given in Figure 4.3. The classification framework implements a multilayer ISTA and FISTA framework with two unfoldings.

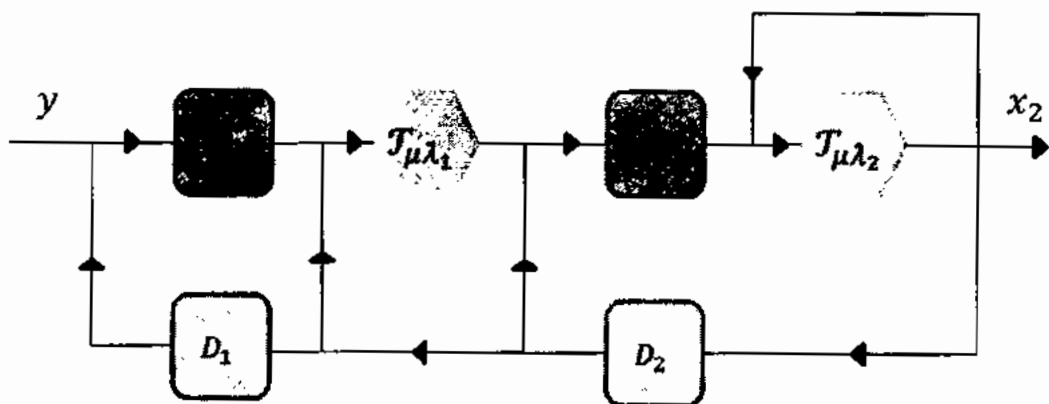


Figure 4.3: A two layer ISTA model as illustrated in [1]

#### 4.5 Experiments and Results on Knee MR Dataset

We use a dataset of 623 MR images comprising 205 (complete tear), 205 (normal), and 213 (partial tear) images with the coronal view. Data collected in the study include adult patients, aged between 18 to 40 years (Male and Female), with Proton density (PD)-weighted images and fat saturation. The images were labeled by a certified MSK-radiologist at Hospital Kuala Lumpur (HKL). An 80-20 split is applied for train and test. This work does

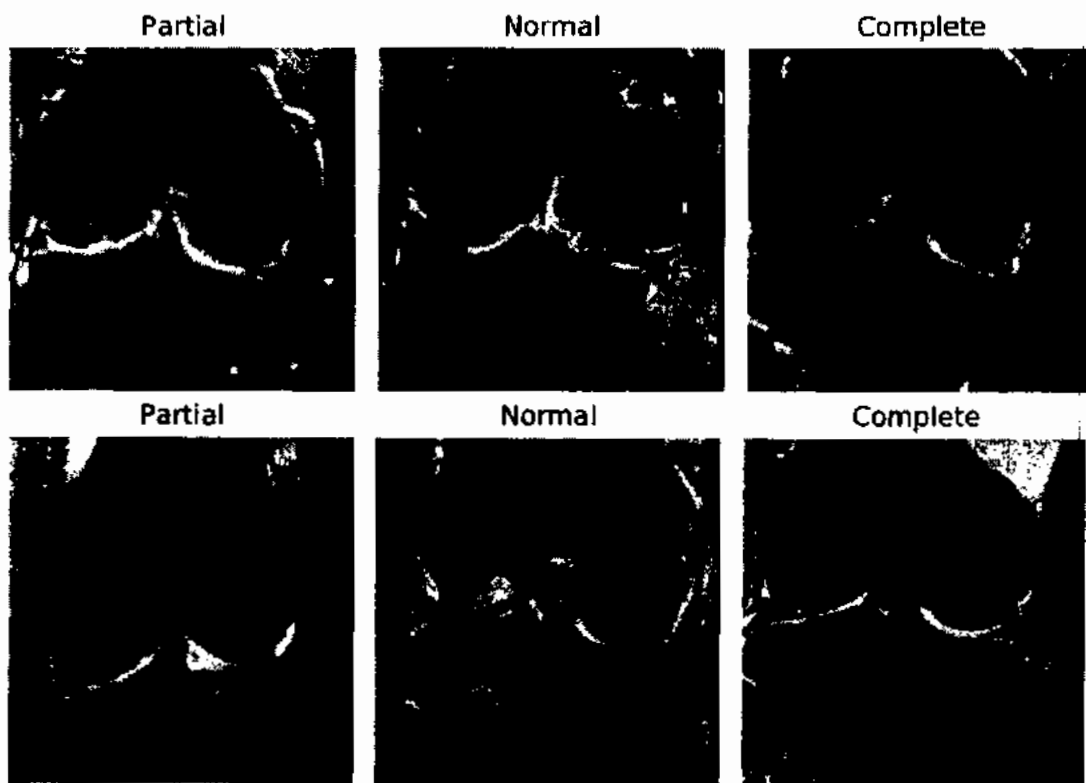


Figure 4.4: Labeled Knee MR images from Knee MR dataset

not employ regularization techniques of dropout and batch normalization to provide a clear experimental setup and demonstrate its effectiveness on a framework with application to biomedical image classification. All algorithms use three convolutional layers, with a filter size of 5 in each layer, and the number of feature maps of 16,32 and 32 size for layers one, two, and three respectively. These parameters have empirically experimented for optimal

performance on this dataset. Similarly, all algorithms use a learning rate of 0.001 and batch size of 3. The optimizer parameters of weight decay (an  $l_2$  weight regularization) and learning rate scheduler values have also been kept the same for all algorithms. All models have been trained with stochastic gradient descent. Table-4.1 gives precision, recall, average accuracies, and F-1 scores for baseline CNN, and All-Free learning framework and proposed ML-ISTA, ML-FISTA, and ML-BP with network unfoldings.

Table 4.1: Precision, Recall, Accuracy and F-1 scores of iterative thresholding algorithms with unfoldings

Algorithms	Precision	Recall	Accuracy	F1-Score
CNN	87.79	90.30	88.19	88.34
All Free	88.77	89.50	89.89	89.58
ML-FISTA-unfolding= 1	82.29	82.18	82.00	82.14
ML-FISTA-unfolding = 2	90.45	90.55	89.43	90.45
ML-ISTA-unfolding = 1	87.50	88.10	88.00	87.30
ML-ISTA-unfolding = 2	91.00	91.29	91.96	91.03
<b>ML-BP-unfolding = 1</b>	<b>92.42</b>	<b>92.68</b>	<b>92.20</b>	<b>92.42</b>
ML-BP-unfolding = 2	90.01	90.01	90.90	89.58

Table 4.2: Class-wise accuracies and test loss

Algorithms	Partial ACL tear	Complete ACL tear	Average accuracy	Test Loss
CNN	71	95	88	0.1291
All Free	71	97	89	0.0814
ML-ISTA	85	93	92	0.0785
ML-FISTA	73	98	89	0.1845
<b>Layered BP</b>	<b>82</b>	<b>100</b>	<b>92</b>	<b>0.0814</b>

The classification metrics of the framework with the highest average accuracy have been highlighted to emphasize the effectiveness and better accuracy of proposed frameworks as compared to baseline. Class-wise accuracies, average accuracies, and test losses of CNN and proposed algorithms are given in Table-4.2, with emphasis on the framework with

better classification accuracies on complete ACL tear and partial ACL tear.

#### **4.5.1 Multi-layered Iterative Thresholding Algorithms with Unfoldings**

The results of the classifier based on features extracted by multilayered iterative thresholding algorithms are given in Figure 4.6 and Figure 4.7. The classifier performance is given for two unfoldings (1 and 2) and a further increase in unfolding value results in divergence of algorithms. The training accuracy of implemented frameworks is given in Figure 4.9 for unfoldings 1 and 2. Train losses and validation losses for the number of unfoldings are depicted in Figure 4.10 and Figure 4.11 respectively. The empirical results show improvement in learning performance of the proposed classification framework, a sharp decrease in loss curves, and better classification accuracy as unfoldings are increased. The ML-ISTA framework with two unfoldings outperforms CNN and ML-FISTA with reasonable margins as shown in Table-4.1.

#### **4.5.2 Layered Basis Pursuit**

The layered basis pursuit algorithm is incorporated with the same architecture and hyperparameters of baseline CNN and results are added for comparison. The algorithm as proposed in [35], is implemented with each iterative shrinkage iteration unrolling at each layer. Whereas in the case of ML-ISTA and ML-FISTA, unrolling of iterations is done for the entire multi-layer basis pursuit problem. The experimental framework uses two iterations and results are given for comparison with CNN in Figure 4.8. The ML-BP framework with single unfolding has comparable results with the ML-ISTA framework and outperforms baseline with reasonable margins in terms of average accuracy.

#### **4.5.3 An Adaptive Learning Framework**

In addition to the three generative models described above, an all-free learning framework consisting of three layers with the same number of feature maps as CNN is also imple-

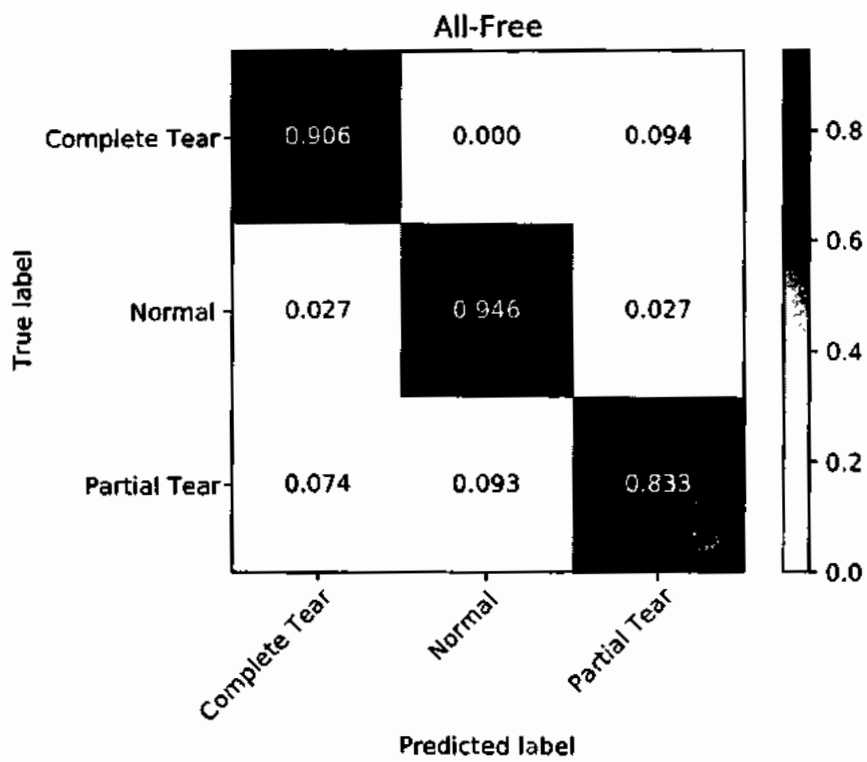
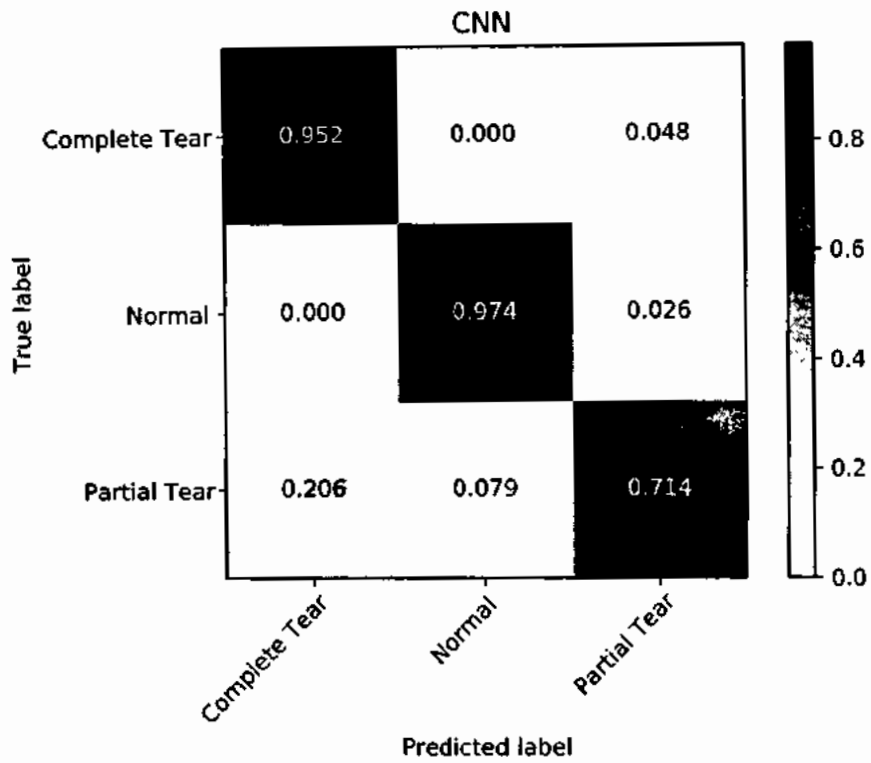


Figure 4.5: Confusion matrices of CNN and All-Free framework



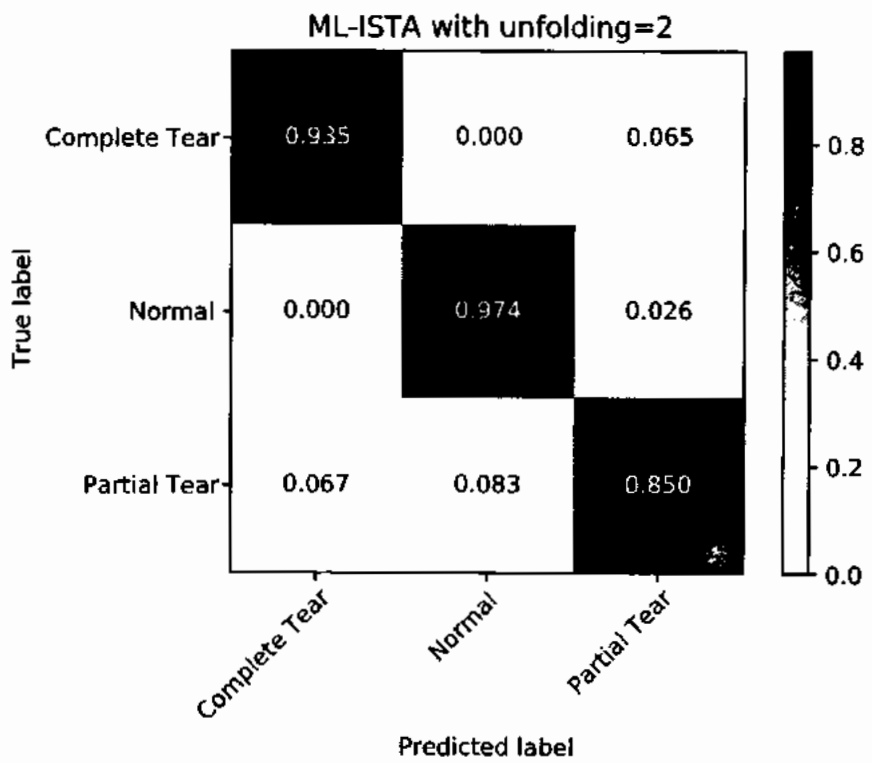
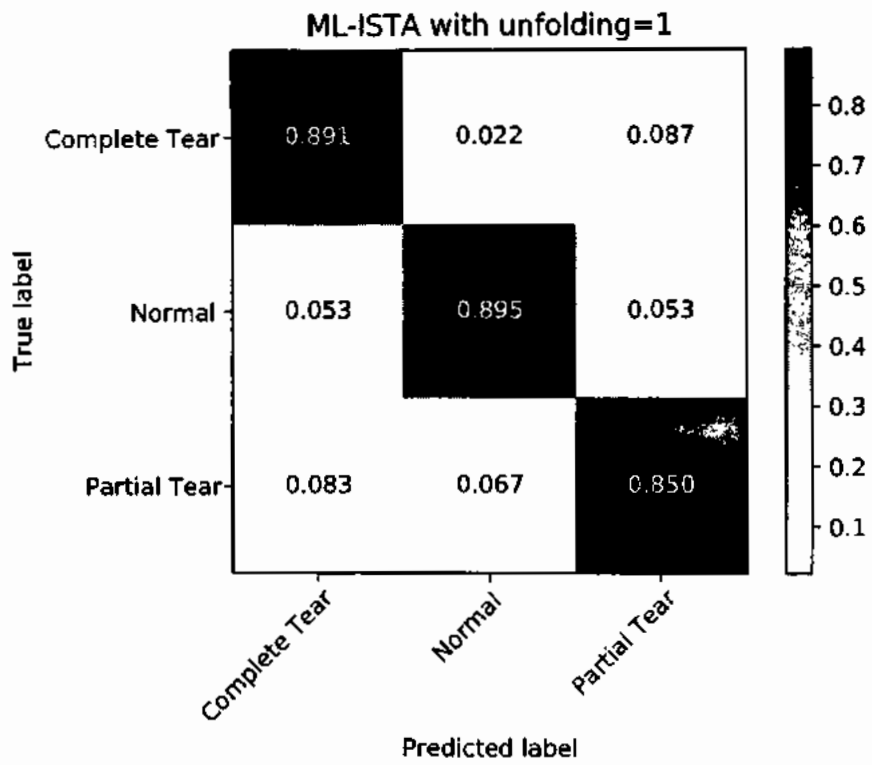


Figure 4.6: Confusion matrices of ML-ISTA with unfoldings

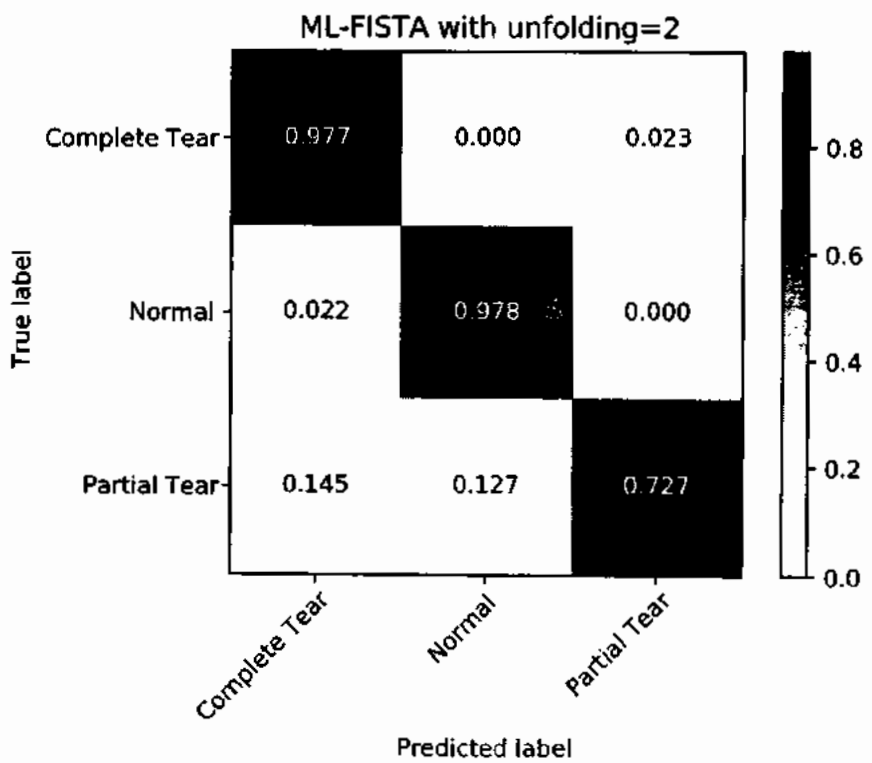
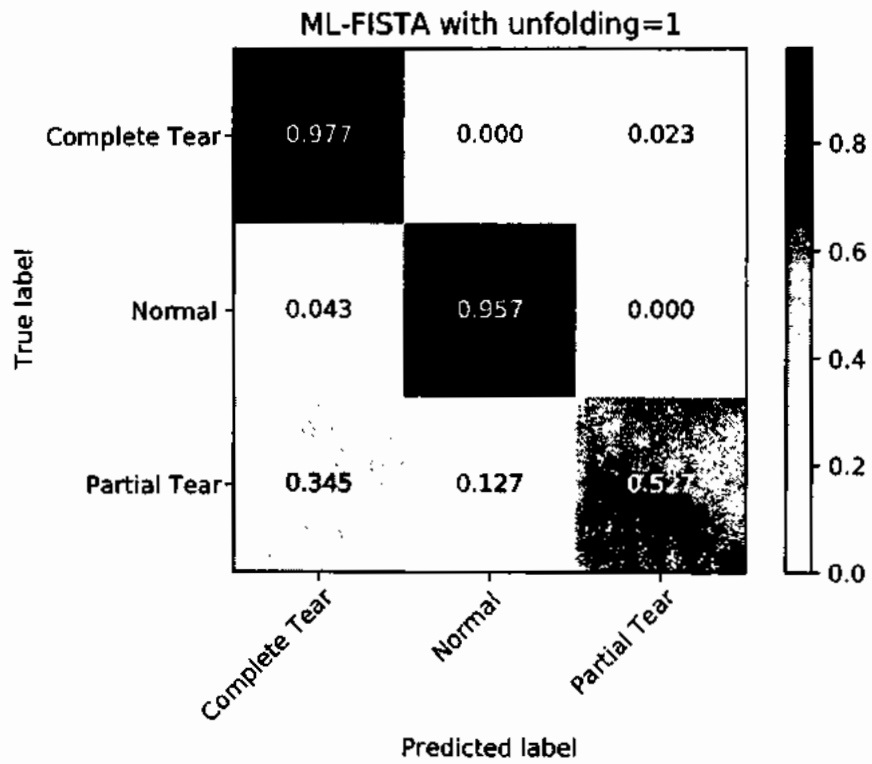


Figure 4.7: Confusion matrices of ML-FISTA with unfoldings

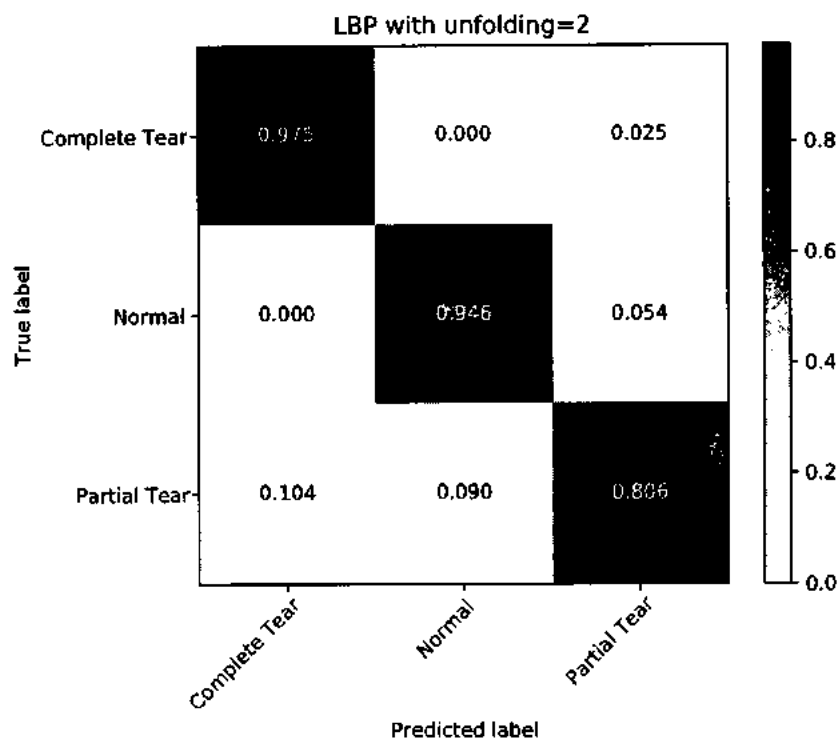
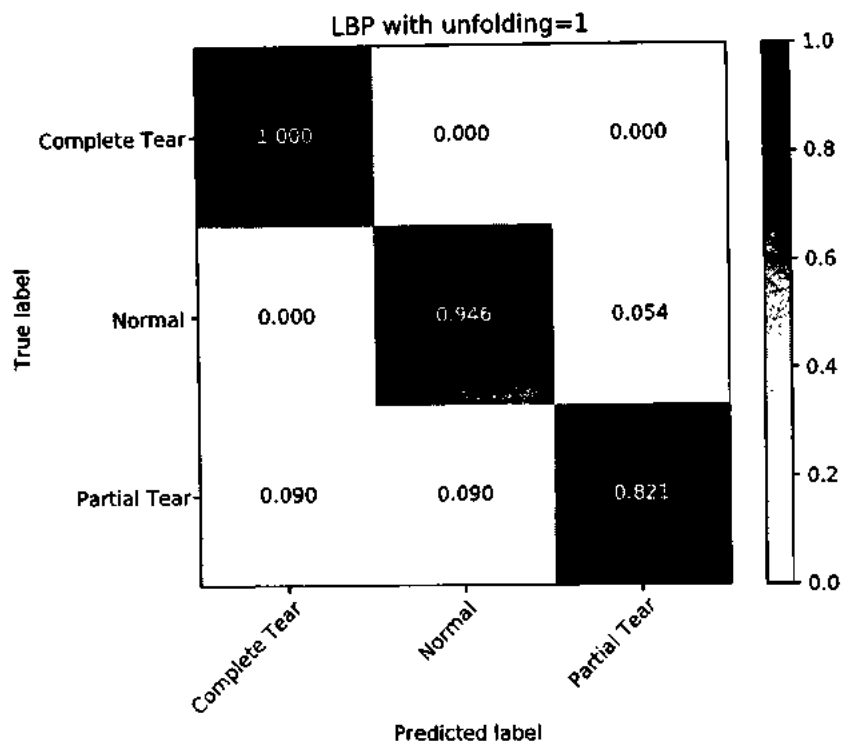


Figure 4.8: Confusion matrices of ML-BP with unfoldings

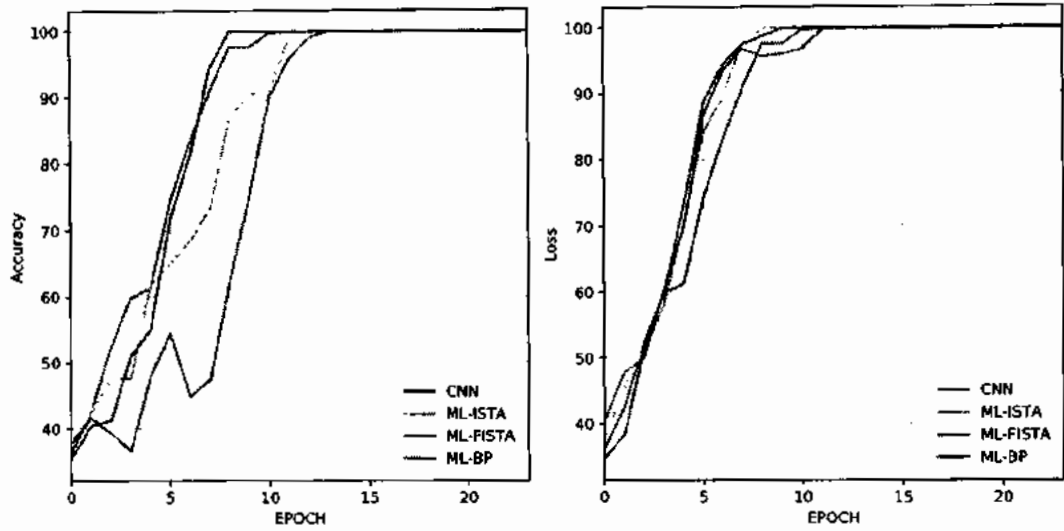


Figure 4.9: Training accuracy for unfolding = 1 (left), and unfolding = 2 (right)

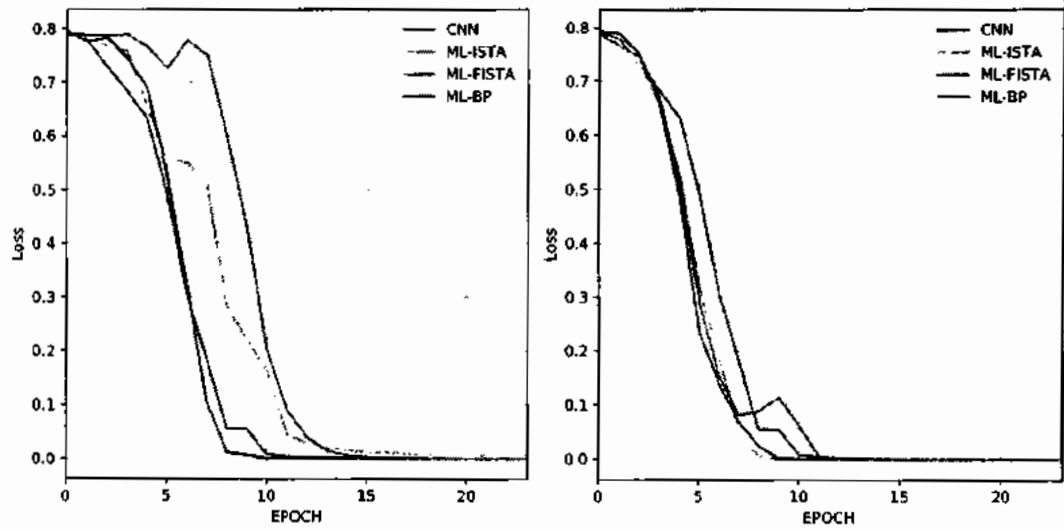


Figure 4.10: Training loss for unfolding = 1 (left), and unfolding = 2 (right)

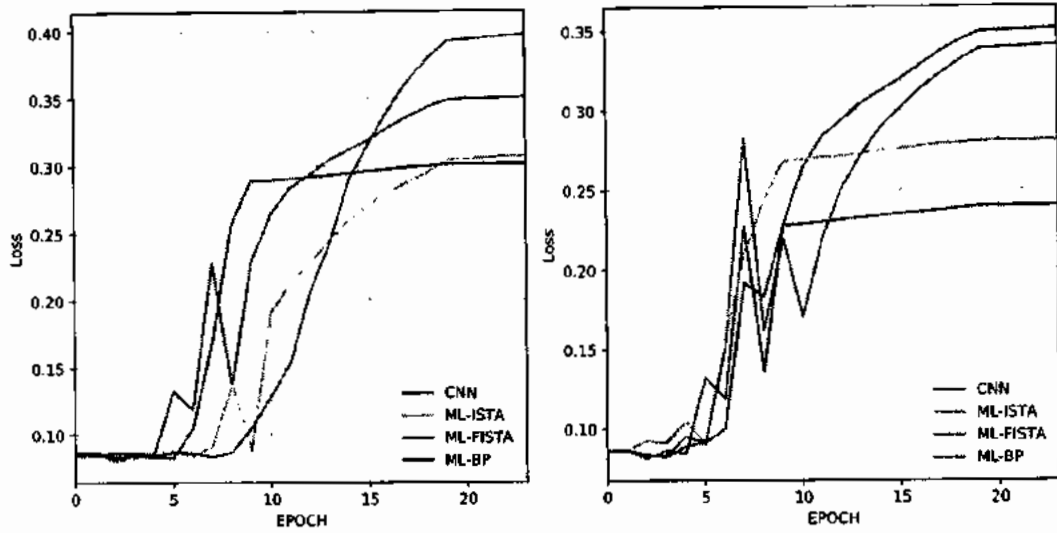


Figure 4.11: Validation loss for unfolding = 1 (left), and unfolding = 2 (right)

mented. In this framework, the dictionaries and corresponding representations are adaptively learned for the subject dataset. The all-free model is trained on the same number of layers and relevant parameters for a recurrent architecture. Framework for ML-ISTA, ML-FISTA, and layered BP has the same number of parameters as of CNN, whereas all free framework has  $\mathcal{O}(LK)$  parameters.  $L$  is the number of layers (here 3 layers are used) and  $K$  is the number of unfoldings. The results for classification accuracies for baseline CNN and all-free recurrent architecture are given in Figure 4.5, Table-4.1, and Table-4.2. The number of iterations of the framework slightly improves the accuracies on given classes as observed in ISTA and FISTA architectures.

#### 4.5.4 Data Augmentation

MR data used in the architecture is first center cropped for an image size of 320x320 and then normalized with the mean and standard deviation of the dataset. Of many transforms available in Pytorch, the center crop proves very effective in this architecture, and the learning curve for the training dataset follows a steady pattern of decrease with an increase in the number of epochs.

#### **4.5.5 Class Imbalance**

When there is an imbalance in classes, DL frameworks give poor classification accuracy in certain classes. For the knee ACL tear classification problem, the classifiers give poor accuracy for the challenging problem of partial ACL tear classification as compared to the other two classes. To circumvent this issue, the oversampling of the partial tear category is done during the training phase. This technique significantly increases test accuracy for partial ACL tear class when equally applied for CNN as well as the proposed framework, resulting in improved results of the proposed framework with two unfoldings in comparison to other classification algorithms. The accuracies and effectiveness of biomedical classification algorithms can be further improved by incorporating the learning framework with different datasets comprising different age profiles (young and aged population) and gender profiles for a more specific and accurate diagnosis.

#### **4.5.6 Deep Learning Architecture and the Challenge of Overfitting**

The DL architecture was chosen for this problem of classification, the size of filters and number of feature maps were observed to be optimal for this dataset of knee MR images. Deeper architectures for CNN, with a higher number of feature maps, resulted in the underfitting of the learning model. The effect of the performance of the classifier is given to demonstrate the effectiveness of ISTA, FISTA, and layered basis pursuit with unfoldings, which uses the iterations of the algorithm to extend the depth of the network without increasing the number of parameters. With improved results in terms of accuracy of classifiers as shown in confusion matrices are given in Figures 4.6, 4.7, and 4.8, the unfolding increases the classification performance of the implemented framework, especially on an imbalanced class of partial ACL tear. Unlike the CNN, where the increase in depth of neural network results in underfitting of learning model, this framework successfully can be implemented with two unfolding for ISTA (Figure-4.6), FISTA (Figure 7), and L-BP (Figure 4.8).

#### 4.5.7 Discussion

The challenging problem of identification of partial ACL tear, which is characterized by stretching and weakening of the knee ligaments is diagnosed by clinicians with clinical tests along with MR imaging and arthroscopic examinations. In our work, the MR images with coronal PD were used for training and testing the framework, as the coronal imaging plane is mainly used by radiologists to trace ACL fibers from origin to insertion. In the proposed framework, the partial tear class is successfully identified with 85% accuracy by ML-ISTA and 82% accuracy by ML-BP. The cases of complete ACL tear class which are characterized by the rupture of the knee ligaments are identified with 98% accuracy by ML-FISTA followed by ML-BP framework which has an accuracy of 97%. Overall, the ML-BP algorithm results in the highest average classification accuracy on all classes as shown in Table 4.1 and Table 4.2.

Generally, the presence of notch origin tears makes the diagnosis of a complete ACL tear difficult for radiologists to detect in clinical settings. Another possible reason for misclassification of the complete ACL tear class is the mild focal intrasubstance degeneration rather than a complete tear.

As MRI-based pathology is localized to small regions of interest, the image crop operation applied in our work significantly improves the learning network training accuracy. This insight can be used to further improve the framework with the incorporation of training images comprising of sagittal and axial planes, which are part of standard knee imaging protocol used in clinical applications. The evaluation and interpretation of three-dimensional (3D) data is another unique feature associated with cross-sectional imaging. For musculoskeletal injuries, the combination of 3D contextual information of ligaments in the imaging pipeline is especially useful for diagnosing ACL tears. Furthermore, the performance and generalizability of the framework may be improved with the incorporation of different magnetic field strengths, scanning protocols, and vendors of MRI scanners.

## Summary

In this chapter, a multilayered convolutional sparse coding (ML-CSC) framework employing iterative thresholding pursuit algorithms was implemented and demonstrated its effectiveness in terms of classification accuracy in comparison to traditional CNN-based frameworks. Algorithms of gradient mapping schemes like iterative thresholding algorithm (ISTA), fast iterative thresholding algorithm (FISTA) along multilayered basis pursuit were implemented for feature extraction and training of the classifier. The framework was applied to a labeled dataset of knee MR images for classification and accuracies were given for different types of ACL tears. In absence of larger labeled datasets, this work demonstrated the effectiveness of the classification framework's learning capability with the same number of features as the baseline CNN, and without adding regularization hyperparameters and computational complexity to the neural network architecture. The framework also demonstrated the effectiveness of unfolding on neural networks' performance, improving classification accuracies on imbalanced classification problems of a partial ACL tear.



## CHAPTER 5

### MULTI-LAYER CONVOLUTIONAL SPARSE CODING FRAMEWORK FOR RESTORATION OF UNDER-SAMPLED MR IMAGES

Magnetic resonance imaging plays an important role in the diagnosis of different pathologies associated with human anatomy. The need to acquire images with higher temporal and spatial resolution induce longer scan times resulting in patient fatigue and claustrophobia. In addition to long scan times, the induced motion artifacts further necessitate the reduction in scan time for better image quality in case the process is repeated. To circumvent the longer scan time, parallel imaging and compressive sensing techniques have been proposed enabling 2- to 3-fold scan time accelerations. The emergence of deep learning-based techniques that rely on a large number of fully sampled MR images to learn image priors and key parameters. In this chapter, we propose a multi-layer convolutional sparse coding framework (ML-CSC) utilizing layered basis pursuit for CS-MRI reconstruction and demonstrate its effectiveness with different acceleration factors. The generic architecture is shown to provide successful reconstruction from undersampled images which can be used for clinical interpretations. Magnetic resonance imaging (MRI) is a noninvasive imaging modality widely used in clinical applications to assist clinicians in diagnosis and treatment management of different health conditions [136]. One of the challenges of MRI image reconstruction is the longer scan time needed for better quality image reconstruction. This results in claustrophobia in some patients and becomes a challenging issue in cardiac MR imaging due to the presence of motion artifacts. To address the issue, techniques like parallel imaging (PI) and compressive sensing have been proposed. In PI, the scan time is decreased by a reduction in increment steps and utilizing the multiple receiver coil sensitivity for MRI image reconstruction. The compressive sensing (CS) technique works on the premise of acquiring  $k$ -space data being randomly undersampled has a sparse

representation in some basis or dictionary which are pre-defined. The non-linear image reconstruction is done to enforce sparsity and consistency in acquired MR data. The The Food and Drug Administration (FDA) in 2017, cleared the use of CS technology opening the door for broader use of the technique in clinical settings [143, 144].

### 5.1 CS MR Image Reconstruction as Inverse Problem

The MR image reconstruction can be given as linear model of equations where the task of reconstruction is to recover an image from its undersampled observations. The MR-CS image acquisition problem can be modeled as:

$$\hat{s} = \arg \min_s \frac{1}{2} \|\Phi \Psi s - y\|_2^2 + \lambda \|s\|_1, \quad (5.1)$$

where  $\Phi$  is measurement matrix,  $\Psi$  are sparsifying basis,  $s$  are sparse coefficients i.e.  $x = \Psi s$ , and  $\lambda \geq 0$ . Typical sparsifying basis  $\Psi$  consists of wavelet, DCT or any other learned dictionary. Substituting sparsifying transform and its sparse coefficient vectors with  $x$ , i.e.  $x = \Psi s$ , and orthonormal sparsifying transform as  $\Psi \cdot \Psi^+ = I$ , in Equation (5.2), we have:

$$\hat{x} = \arg \min_x \frac{1}{2} \|\Phi x - y\|_2^2 + \lambda \|\Psi^+ x\|_1, \quad (5.2)$$

The compressive sensing incorporates the compression into acquisition with measurement matrix  $\Phi \in \mathbb{C}^{M \times N}$ , ( $M \ll N$ ), to infer original signal  $x \in \mathbb{C}^N$  from its measurements  $y = \Phi x \in \mathbb{C}^M$ . The compression ratio (CS-ratio) is defined as  $M/N$ .

Equation (5.2) is further generalized by replacing the regularization term with learned convolutional filters.

$$\hat{x} = \arg \min_x \frac{1}{2} \|\Phi x - y\|_2^2 + \sum_{l=1}^L \lambda_l g(D_l x), \quad (5.3)$$

The transform matrix in Equation (5.3) is denoted by  $D_l$ , which can be a gradient transform, discrete wavelet transform (DWT), or Discrete Cosine Transform (DCT). Here,  $g(\cdot)$  is a  $l_q$

sparse regularization function where  $q \in [0, 1]$ . The CS reconstruction problem can be categorized into different groups with the emergence of deep learning-based techniques. In the following sections, a brief overview is given for different approaches applied for the CS-MRI problem in recent research works.

### **5.1.1 Dictionary Learning-Based Approaches**

In traditional CS reconstruction techniques, the transform or fixed dictionary is known as a priori, whereas the dictionary learning techniques adaptively learn the dictionary from the underlying MR images [53, 52]. The traditional techniques relying on fixed transform are fast but give less accurate results, while in dictionary learning-based methods, the convergence guarantees are difficult to achieve due to nonconvex optimization.

### **5.1.2 Low-Rank-Based Approaches**

This technique utilizes the inherent high degree of correlations of MR images, representing them with the union of low-dimensional sub-spaces [145]. In [146], patch-based reconstruction for undersampled 2D cine MR images has been used.

### **5.1.3 Deep Learning and Multi-layer Thresholding Algorithms for MR Image Reconstruction**

One of the main drawbacks of traditional CS reconstruction methods is dependent on the choice of the sparsifying transform, leading to over smoothness of images when high acceleration factors are used. Other drawbacks include a non-generalization of algorithms as they depend upon specific images. For recovering an undersampled signal, an alternative approach is learning-based techniques. For the learning-based approach, given a set of ground truth images as train set and their corresponding undersampled measurements  $\{x_n, y_n\}_{n=1}^N$ , the objective of the learning-based approach is to solve the reconstruction al-

gorithm of the form:

$$R_{learn} = \arg \min_{D, \Theta} \sum_{n=1}^N f(\{x_n, y_n\}), D_{\Theta} + g(\Theta) \quad (5.4)$$

where  $f(\cdot)$  is a measure of the error to be minimized,  $\Theta$  is the set of all learnable parameters,  $g(\Theta)$  is the regularizer added to avoid overfitting of the learning model. The model can be used successfully for CS signal reconstruction after the training phase is complete. In the CS-MRI image reconstruction problem, the task is to reconstruct images from the undersampled  $k$  space measurements. This is done with help of training a neural network for non-linear mapping between fully sampled  $k$ -space data and under-sampled  $k$ -space data [113, 147, 148]. The successful implementations of deep neural networks for inverse problems in image processing, the theoretical understanding of these networks have been recently explored to improve their performance [38, 35, 1]. Taking insights from the celebrated spars coding theory, the multilayered convolutional sparse coding frameworks have been proposed to theoretically model deep neural networks. This model has been used in [149] to reconstruct undersampled MR images with improved performance as compared to traditional CNN-based frameworks. In this work, we implement the multi-layer basis pursuit with layer-wise pursuit i.e ML-BP, and demonstrate with empirical experiments its viability in the restoration task of CS-MRI. The proposed generic framework once trained on GPU enabled machine to successfully restore undersampled MR images with desirable quality for clinical interpretation.

The contribution of this work is the implementation of layered wised basis pursuit instead of global pursuit for finding sparse representations of MR images and learning the non linear mapping form fully sampled dataset to under sampled datasets.

## 5.2 ML-CSC Based Framework for CS-MRI Image Reconstruction Problem

In ML-CSC-based framework employing a layered basis pursuit, a signal  $y$  admitting a sparse representation in terms of a dictionary  $D$ , can be represented as  $y = Dx$ , with  $x$  having a sparse structure (solution with fewer non-zero entries). After employing an  $l_1$  norm penalty and considering the problem as NP-hard, the basis pursuit problem is given as:

$$\min_x \|x\|_1 \text{ s.t. } \|y - Dx\|_2^2, \quad (5.5)$$

Each representation estimate is required to explain the immediate layer only and a signal is not generated based on the global multilayer sparse model. The ML-BP algorithm is implemented for six layers comprising dictionaries of the same size. The pursuit is done on the individual layer for feature extraction for underlying MR datasets.

### 5.2.1 Iterative Shrinkage Algorithms for Basis Pursuit

Iterative shrinkage algorithms (ISTA) require matrix-vector multiplications and entry-wise operations. This gives a clear advantage over computationally expensive second-order methods like the interior point method [68, 129]. We propose a multi-layer basis pursuit algorithm based on an iterative thresholding algorithm. The algorithm takes into account the merits of objective-based approaches and network-based approaches to achieve CS-MRI reconstruction. Each ISTA update is fed to an ML-CSC network employing multi-layer ISTA for basis pursuit. The deep network learns the mappings from previous ISTA updates with a fixed number of unfoldings which corresponds to the iterations of ISTA.

---

**Algorithm 7** Iterative Shrinkage Thresholding Algorithm

---

Init  $x^0 \in f(x)$   
1: **for** any  $k=0,1,2..$  **do**  
2:      $x^{k+1} = \text{prox}_{\frac{1}{L_k}} g(x - \frac{1}{L_k} \nabla f(x))$   
3: **end for**

---

### 5.2.2 Model

Given a decomposed loss function of the form;

$$F(x) = f(x) + g(x),$$

where  $f(x)$  is convex and smooth and  $g(x)$  is smooth,  $L$  is Lipchitz constant, the ISTA algorithm as proximal gradient method finds the minimizer of  $F = f + g$  by iterating the updates given by proximal operator  $g(\cdot)$ :

$$x^{k+1} = \text{prox}_{\frac{1}{L}}(g(x^k - \frac{1}{L_k} \nabla f(x))), \quad (5.6)$$

The traditional ISTA formulation cannot be applied to (5.5) due to the presence of non-separable composite term in  $g(\cdot)$ . To tackle this issue and solve (5.5), gradient mapping approach is analyzed.

Given a function of the form  $F = f + g$ , the gradient mapping operator is given by:

$$G_L^{f,g}(x) = L(x - \text{prox}_{\frac{1}{L}}(g(x - \frac{1}{L_k} \nabla f(x)))), \quad (5.7)$$

The ISTA update for (5.7) can be written as:

$$x^{k+1} = x^k - \frac{1}{L} G_L^{f,g}(x),$$

This update can be considered as gradient mapping step. Since,

- $G_L^{f,g}(x) = \nabla F(x) = \nabla f(x)$  as  $g(x) \equiv 0$ ,
- $G_L^{f,g}(x) = 0$  iff  $x$  is minimizer of  $F(x)$ ,

Essentially  $\Phi^T(\Phi x^{k-1} - y)$  is gradient of the data fidelity term in Equation (5.5). The objective function of ML-BP, can be generally expressed as minimization problem of the

form,

$$\min F(x) = f(D_i x) + g_1(D_i x) + g_2(x),$$

for  $i$ th layer of ML-BP, the proximal gradient mapping method to minimize above objective function takes the update of the form:

$$x_i^{k+1} = \text{prox}_{t g_i}(x_i^k - t \cdot D_i^T G_{1/c}^{f(\cdot), g_{i-1}(D_i \cdot)}(x_i^k)), \quad (5.8)$$

for  $t > 0$  and  $c > 0$ , which are learnable parameters of ML-CSC framework. The proximal of  $g_{i-1}(D_i \cdot)$  involves a composite term, an approximation for representations  $x_i$  is proposed such that  $x_{i-1} = D_i x_i$ . Therefore, for  $i$ th layer of the framework the update of (5.8) can be modified as,

$$x_i^{k+1} = \text{prox}_{t g_i}(x_i^k - t \cdot D_i^T G_{1/c}^{f(\cdot), g_{i-1}} x_{i-1}^k), \quad (5.9)$$

with approximation of composite term, the calculation of proximal mapping of  $g$  term becomes soft thresholding that is  $\text{prox}_{g(x)} = \tau_\lambda(x)$ . For a multi-layer model, the ISTA update step can be written as:

$$x_i^{k+1} \leftarrow \mathcal{T}_t \lambda_i [x_i^k - \frac{t}{c} D_i^T (x_{i-1}^k - \mathcal{T}_t \lambda_i \mu (x_{i-1}^k - D_{i-1}^T (D_i x_{i-1}^k - y)))]), \quad (5.10)$$

The ML-BP algorithms employing multi-layer ISTA trains with images in  $k$ -space along with their CS measurements learns the mappings from CS measurements with ground-truth images and saves the mappings in the form of model parameters of the ML-CSC framework. Once the parameters are saved, the testing module reconstructs the test images with the help of learned parameters. The framework (Figure 5.1) initializes the dictionaries with Xavier initialization for all layers.

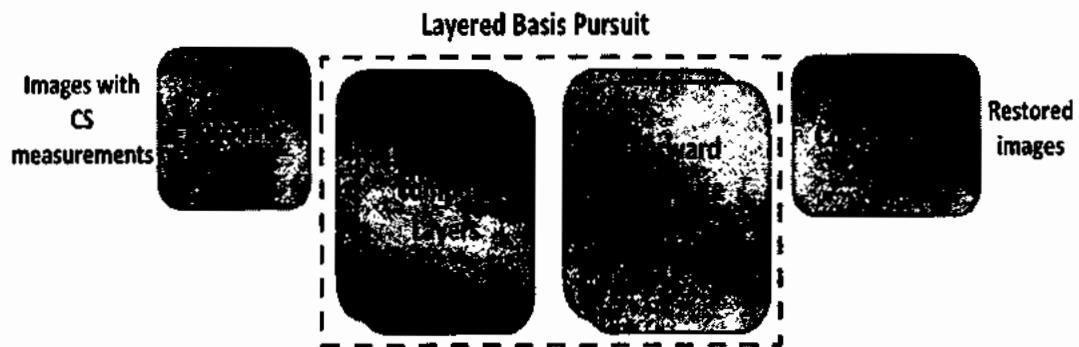


Figure 5.1: Training framework for CS-MRI image restoration with ML-BP

### 5.2.3 Datasets and Parameter Settings

The first dataset has been taken from [68], with 850 brain images acquired in axial and sagittal planes (Figure 5.2). The second benchmark comprises a single image consisting of a single-slice (axial T2-weighted reference brain image) dataset of size (256x256), Vivo MR scans from American Radiology Services as used by Prasad *et al.* in [53]. All algo-

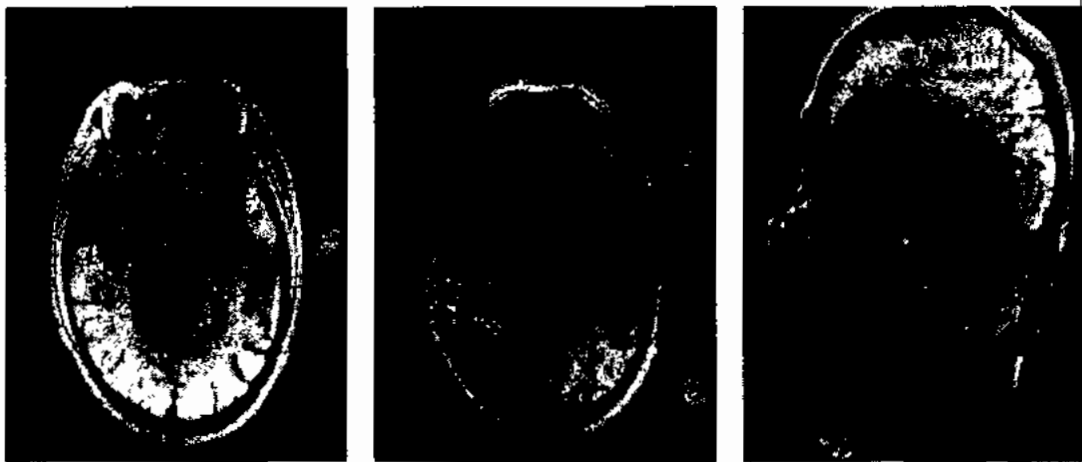


Figure 5.2: Example images from test set

rithms use three convolutional layers, with a filter size of 5 in each layer, and the number of feature maps of 16,32 and 32 size for layers one, two, and three respectively. These parameters have empirically experimented for optimal performance on these datasets.



### 5.2.4 Training Loss

The framework produces intermediate reconstructions denoted by  $x_i^{rec}$  when given pairs of MR images with their corresponding undersampled measurements. The loss function seeks to reduced discrepancy between the input images and the intermediate reconstruction satisfying symmetry constraint of  $\tilde{\mathcal{D}}^k \mathcal{D}^k = I, \forall k = 1, 2, \dots$ , in check. The end to end loss is thus defined as:

$$\begin{aligned} \mathcal{L}_{discrepancy} &= \frac{1}{IN} \sum_{i=1}^N \|x_i^{rec} - x_i\|_2^2 \\ \mathcal{L}_{constraint} &= \frac{1}{IN} \sum_{i=1}^N \sum_{k=1}^U \|\tilde{\mathcal{D}}^k(\mathcal{D}^k(x_i)) - x_i\|_2^2 \\ \mathcal{L}_{total} &= \mathcal{L}_{discrepancy} + \lambda \mathcal{L}_{constraint} \end{aligned} \quad (5.11)$$

where image size  $I$ , number of images  $N$ , ML-ISTA unfoldings  $U$ , learning rate  $\lambda$  are parameters used in above equations.

A learning rate of 1e-4 with Adam optimizer and learning rate decay factor of 0.2 is used in Pytorch implementation.

### 5.3 Empirical Results

The results of CS-MRI restoration framework based on multi-layer basis pursuit are given in Table-5.1. Image-wise PSNR and SSIM for 20% CS ratio (5-fold acceleration factor) for brain MR images is given in Figure 5.3. First the IstaNet [68] is trained on the brain

Table 5.1: Average PSNR/SSIM and restoration times of reconstructed brain MR images

Dataset	CS Ratio				GPU Time
	20%	30%	40%	50%	
	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	
Brain [53]	38.90/0.9598	42.69/0.9825	46.24/0.9914	49.39/0.9945	0.103596s
Brain [68]	39.25/0.9551	41.52/0.9693	44.03/0.9798	46.08/0.9862	0.07268s

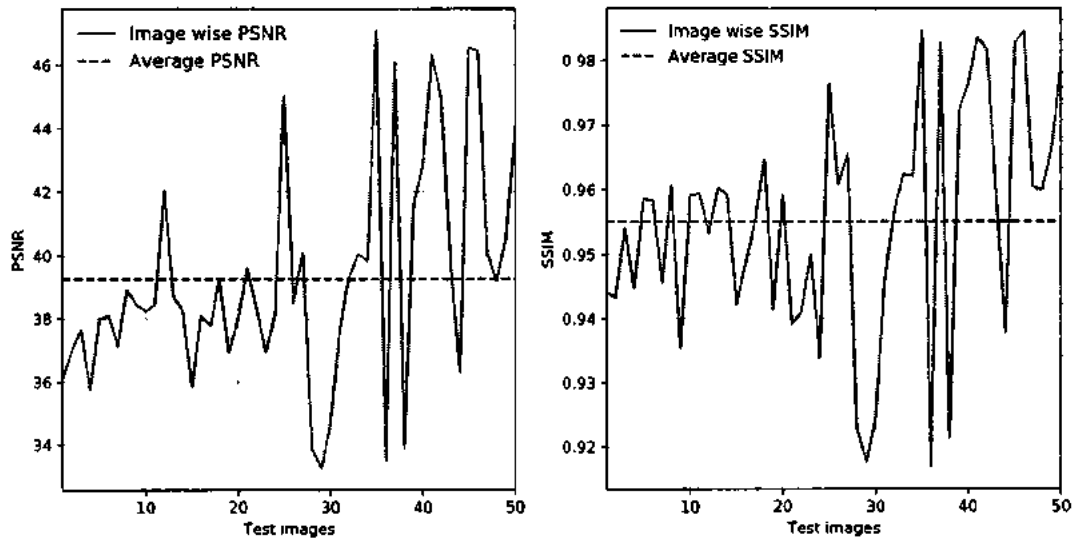


Figure 5.3: Image-wise PSNR and SSIM for 20% CS ratio (5-fold acceleration factor) for brain MR images

dataset and parameters are saved. Then the the proposed CS-MRI framework is trained on the same dataset with same number of EPOCHs. The results of restoration are then taken for testset images for the relevant test codes of IstaNet and our proposed frameworks. The results are given for four CS-ratios with PSNR and SSIM for the both datasets in Table 5.1. The pseudocode for CS-MRI training and testing frameworks based on multi-layer basis pursuit algorithm is given below.

---

#### ML-ISTA CS-MRI PSEODOCODE

---

- **Training ML-ISTA with MRI Train Data:.**

**Require:** Input: MR images, their CS measurements, random masks,  $D_i$  and  $\lambda_i$

**Ensure:** Model containing parameters of ML-ISTA for CS-MRI reconstruction.

- Run encoding to find representations for given layers
- Find representations using ML-ISTA with global pursuits
- Train and save model with EPOCH number and minimum loss

---

- **Reconstruct Test set with trained model.**

**Require:** Trained model, MR test images, undersampling masks

**Ensure:** Reconstructed images with PSNR , SSIM and reconstruction CPU/GPU time

**for**  $i = 1:N$  **do**

- Run ML-BP with learned dictionaries from trained model

- Reconstruct MR images

- Calculate PSNR and SSIM of testset

▷  $N$  is testset size

**end for**

---

## 5.4 Discussion

In this work, we have implemented a generic CS-MRI restoration framework and shown its restoration efficiency on multiple datasets. The trained parameters of the proposed framework for specific MR imaging anatomy can be used for restoration same anatomy in a clinical setting successfully. The work can be further extended for utilizing different planes and datasets for its wider usage and application in clinical settings.

## Summary

The deep learning-based architectures have paved the way for generic acceleration frameworks, where the algorithms once trained on fully sampled datasets can be used to restore undersampled MR images with acceptable quality as compared to fully sampled MR images. These techniques have the advantage of learning the distinct non-linear mapping between fully sampled data with undersampled data without explicit fine-tuning or hand-crafted features for different imaging modalities and protocols. This work implemented the multi-layer basis pursuit algorithm with layer-wise pursuit. The algorithm trained on a GPU-enabled machine is shown to converge quickly and the restoration framework successfully restores images in a small-time on CPU based machine.

## **CHAPTER 6**

### **COMPARATIVE ANALYSIS OF ML-CSC BASED CS-MRI FRAMEWORK WITH STATE OF THE ART**

In image processing research, deep neural networks have been employed extensively for inverse problems. Following that, there was a need to conceptually model deep nets using well-known sparse coding theory, leading to the development of convolutional sparse coding (CSC) theory. The CSC theory, which is a subset of sparse coding theory, is based on the idea of learning filters/dictionaries and sparse feature maps to represent underlying data (natural or biological images). Unlike their sparse coding counterparts, dictionaries have a unique structure, and pursuit algorithms work on a global scale rather than in patches.

This global pursuit results in mitigating the effects of the patch aggregation process during traditional regularization-based techniques. Further extending the CSC model, the learning features are again processed through the CSC model representing them with another layer of filters/dictionaries and their corresponding sparse maps. This process is continued until the last layers of the model making the multi-layer convolutional sparse coding model. The pursuit algorithms can be employed layer-wise of a global pursuit algorithm can be utilized using iterative thresholding algorithms.

In our previous works, we have implemented an ML-CSC model on classification tasks of biomedical images and restoration tasks of CS-MRI images. This work will further extend the proposed model and compare its performance on the custom dataset acquired at Hospital Kuala Lumpur (HKL). The restoration efficiency will be compared with the art and extensive empirical results will be presented to show the effectiveness of the proposed model.

## 6.1 Convolutional Sparse Coding and Inverse Problems

Deep neural networks have been used extensively for inverse problems in image processing research. Subsequently, the interest has been shown to theoretically model deep nets with celebrated sparse coding theory resulting in the emergence of convolutional sparse coding theory. Convolutional sparse coding [150, 151], has been the focus of recent research interest and has been presented as theoretical foundations of deep neural networks. Figure 6.1 shows the number of publications from the Scopus database for the theme of convolutional sparse coding and multi-layer convolutional sparse coding.

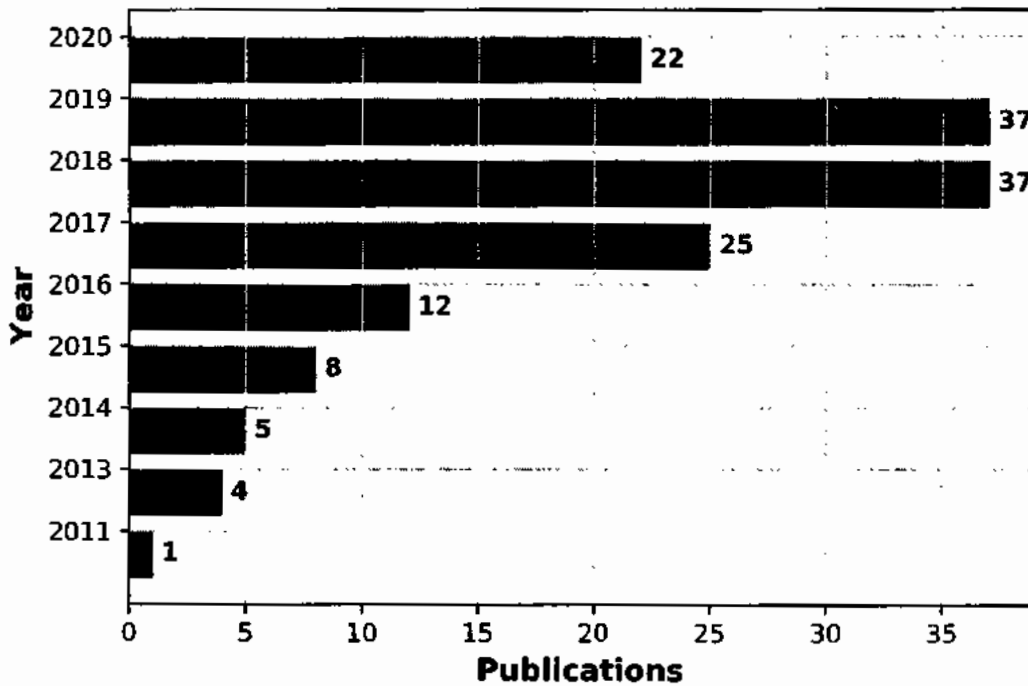


Figure 6.1: Number of publication on Scopus searched with "CSC and ML-CSC" search term.

In our previous works, we implemented ML-CSC models on classification and CS-MRI restoration tasks of biomedical images. This work will further extend the proposed model and compare its performance with the state-of-the-art custom dataset acquired at HKL.

The restoration efficiency will be compared on benchmarks of peak signal to noise ratio (PSNR), structure similarity index (SSIM), and MR image restoration times, presenting extensive empirical results to show the effectiveness of the proposed model.

## 6.2 State of the Art CS-MRI Restoration Framework and ML-CSC Based CS-MRI Model

For reconstruction of a signal acquired with undersampled measurements, the purpose of reconstruction is to infer original signal from its randomized measurements. For MRI signals which are acquired in  $k$ -space, the problem can be formulated mathematically as :

$$\hat{s} = \arg \min_s \frac{1}{2} \|\Phi \Psi s - y\|_2^2 + \lambda \|s\|_1, \quad (6.1)$$

where  $\Phi$  is measurement matrix,  $\Psi$  are sparsifying basis,  $s$  are sparse coefficients i.e.  $x = \Psi s$ , and  $\lambda \geq 0$ . Typical sparsifying basis  $\Psi$  consists of wavelet, DCT or any other learned dictionary. Substituting sparsifying transform and its sparse coefficient vectors with  $x$ , i.e.  $x = \Psi s$ , and orthonormal sparsifying transform as  $\Psi \cdot \Psi^+ = I$ , in Equation (6.2), we have:

$$\hat{x} = \arg \min_x \frac{1}{2} \|\Phi x - y\|_2^2 + \lambda \|\Psi^+ x\|_1, \quad (6.2)$$

The compressive sensing incorporates the compression into acquisition with measurement matrix  $\Phi \in \mathbb{C}^{M \times N}$ , ( $M \ll N$ ), to infer original signal  $x \in \mathbb{C}^N$  from its measurements  $y = \Phi x \in \mathbb{C}^M$ . The compression ratio is defined as  $M/N$ .

Equation (6.2) is further generalized by replacing the regularization term with learned convolutional filters.

$$\hat{x} = \arg \min_x \frac{1}{2} \|\Phi x - y\|_2^2 + \sum_{l=1}^L \lambda_l g(D_l x), \quad (6.3)$$

The transform matrix in Equation (6.3) is denoted by  $D_l$ , which can be a gradient transform, discrete wavelet transform (DWT), or Discrete Cosine Transform (DCT). Here,  $g(\cdot)$  is a  $l_q$

sparse regularization function where  $q \in [0, 1]$ .

### 6.2.1 Ista-Net: The Layered Basis Pursuit (ML-BP)

The layered basis pursuit in context of sparse models and deep learning has been proposed for addressing pursuit problems of the form:

$$\hat{x}_i \leftarrow \arg \min_{x_i} \|\hat{x}_{i-1} - Dx_i\|_2^2 + \lambda_i \|x_i\|_1, \quad (6.4)$$

[68] proposed Ista-Net, an algorithm for compressive sensing (CS) reconstruction of natural images using insights from traditional optimization-based and network-based methods. Employing the iterative shrinkage thresholding algorithm for optimizing a general  $\ell_1$  norm CS reconstruction model, a structured deep network was proposed for the reconstruction of undersampled images in the pixel domain. Instead of using hand-crafted features, authors utilized a customized CNN for learning the parameters of the non-linear mapping operator and end-to-end learning of transforms, shrinkage thresholds, and step sizes. The proposed framework was employed for CS reconstruction of 91 images dataset used by Kulkarni *et al.* in [69] and CS-MRI reconstruction of brain images used by Yang *et al.* in [76]. [68] and [128] attempt to unfold neural networks with iterative thresholding and minimizing ML-BP. As a result, each representation estimate is required to explain the immediate layer only and a signal based on generation in global multilayer sparse model settings is not possible.

### 6.2.2 Multi-layer Convolutional Sparse Coding Framework for CS-MRI Image Restoration

The heuristic techniques applied to problems in deep learning frameworks have been recently investigated for theoretical explanations of deep learning with the help of celebrated sparse coding theory. Sparse coding theory [2] works on the premise of first learning filters (weights/dictionaries) from given data and then finding the sparse representation maps

from those dictionaries for representations of the underlying structure of the data. Once the underlying structure is successfully modeled, the problems of reconstruction on images from noisy measurements, retrieving/reconstructing a signal in compressive sensing domain, and classification of test sets on already trained dictionaries and sparse maps can be done successfully with the help of different algorithms developed over the years and applied successfully in different domains. The sparse coding theory has been further extended to theoretically explain widely used CNNs namely convolutional sparse coding. The resulting CSC model, where a special circulant and convolutional structure is imposed on dictionaries (which are otherwise traditionally unstructured in sparse coding theory) is defined as a forward pass of CNN. Further work in a multilayered version of CSC has been shown in [1] for convergence analysis and multi-layer basis pursuit for classification performance comparison with CNNs on three public datasets.

### 6.2.3 Multi Layered Basis Pursuit

A signal  $y$  admitting a sparse representation in terms of a dictionary  $D$ , can be represented as  $y = Dx$ , with  $x$  having a sparse structure (solution with fewer non-zero entries). After employing an  $l_1$  norm penalty and considering the problem as NP-hard, the basis pursuit problem is given as:

$$\min_x \|x\|_1 \text{ s.t. } \|y - Dx\|_2^2, \quad (6.5)$$

V.Papayan *et al.* in [35] extended the basis pursuit problem to multi-layer settings, where a signal  $y$  expressed as  $y = D_1x_1$ , for sparse representations  $x_1$  and possibly convolutional dictionary  $D_1$ . The sparse representations can be further expressed as  $x_1 = D_2x_2$  for another dictionary  $D_2$  with sparse representations  $x_2$ . This framework can be extended to  $L$  number of layers. Under this framework, for an observed signal  $y$ , the deep coding problem in multi-layer settings can be expressed as:

$$\min_{x_i} \|y - D_1x_1\|_2^2 \text{ s.t. } [x_{i-1} = D_ix_i, \|x_i\|_0 \leq s_i]_{i=1}^L, \quad (6.6)$$



This generative multi-layer basis pursuit framework has provided a platform for conveniently analyzing deep neural networks. Alternatively, the work in [125] developed a global pursuit algorithm based on projection interpretation. By imposing an analysis prior on deepest representation, Equation (6.5) can be written as,

$$\begin{aligned} \min_{x_i} & \|y - D_{(1,L)}x_{(L)}\|_2^2, \\ \text{s.t.} & \|x_L\|_0 \leq s_L, [\|D_{(i,L)}x_L\|_0 \leq s_{i-1}]_{i=1}^L, \end{aligned} \quad (6.7)$$

The greedy pursuit algorithm presented in [125] does not scale for high dimensional settings. J.Sulam *et al.* [1] presented a multi-layer basis pursuit algorithm that could leverage the symbiotic relationship of analysis and synthesis-priors on sparse representations of different layers according to the depth of the multi-layer basis pursuit framework. Specifically, a convex relation was proposed for (6.7), resulting in multi-layer basis pursuit:

$$\min_x \frac{1}{2} \|y - D_1 D_2 x\|_2^2 + \lambda_1 \|D_2 x\|_1 + \lambda_2 \|x\|_1, \quad (6.8)$$

The model imposed mixture of analysis and synthesis priors, with  $\lambda_1 = 0$  and  $\lambda_2 > 0$  recovering traditional pursuit with factorized global dictionary  $D$ .

#### 6.2.4 Iterative Shrinkage Algorithms

Iterative shrinkage algorithms (ISTA) are first-order methods that require matrix-vector multiplications and entry-wise operations. This gives a clear advantage over interior point methods and other solvers, which depend upon second-order information making them computationally expensive in high dimensional settings.

ISTA originally proposed in [70] gives convergence in order of  $\mathcal{O}(1/k)$  in functional value and its fast version FISTA [127] has improved convergence in the order of  $\mathcal{O}(1/k^2)$ . The ISTA algorithm is given as Algorithm-8. In the following, the proposed model for ML-CSC employing multi-layer basis pursuit-based learning framework of [149] is used

---

**Algorithm 8** ISTA

---

Init  $x^0 \in f(x)$   
**for any**  $k=0,1,2..$  **do**  
     $x^{k+1} = \text{prox}_{\frac{1}{L_k}} g(x - \frac{1}{L_k} \nabla f(x))$   
**end for**

---

for comparative analysis with state of the art CS-MRI restoration framework [68] trained on for brain and knee MR datasets. Both frameworks are tested on different datasets and reconstruction efficiency in terms of peak signal to noise ratios(PSNR), structure similarity index (SSIM), and average restoration time is given for different undersampling ratios. The ML-CSC-based algorithm incorporates the merits of network-based approaches and objective-based approaches for CS-MRI reconstruction. The ML-CSC framework takes each ISTA update and employs multi-layer ISTA for basis pursuit. With a fixed number of unfoldings, the framework learns the mappings from previous ISTA updates which correspond to the iterations of ISTA.

### 6.2.5 Model

Given a decomposed loss function of the form:

$$F(x) = f(x) + g(x),$$

where  $f(x)$  is convex and smooth and  $g(x)$  is smooth,  $L$  is Lipchitz constant, the ISTA algorithm as proximal gradient method finds the minimizer of  $F = f + g$  by iterating the updates given by proximal operator  $g(\cdot)$ :

$$x^{k+1} = \text{prox}_{\frac{1}{L_k}} g(x^k - \frac{1}{L_k} \nabla f(x)), \quad (6.9)$$

The traditional ISTA formulation cannot be applied to (6.8) due to presence of no separable composite term in  $g(\cdot)$ . In order to tackle this issue and solve (6.8), gradient mapping<sup>1</sup>

---

<sup>1</sup>For a detailed description of gradient operators, see [129], chapter 10.

approach [129] is analyzed.

Given a function of the form  $F = f + g$ , the gradient mapping operator is given by:

$$G_L^{f,g}(x) = L(x - \text{prox}_{\frac{1}{L}}g(x - \frac{1}{L}\nabla f(x))), \quad (6.10)$$

The ISTA update for (6.10) can be written as:

$$x^{k+1} = x^k - \frac{1}{L}G_L^{f,g}(x),$$

This update can be considered as gradient mapping step. Since,

- $G_L^{f,g}(x) = \nabla F(x) = \nabla f(x)$  as  $g(x) \equiv 0$ ,
- $G_L^{f,g}(x) = 0$  iff  $x$  is minimizer of  $F(x)$ ,

Essentially  $\Phi^T(\Phi x^{k-1} - y)$  is gradient of the data fidelity term in Equation (6.8). The objective function of ML-BP, can be generally expressed as minimization problem of the form,

$$\min F(x) = f(D_i x) + g_1(D_i x) + g_2(x),$$

for  $i$ th layer of ML-BP, the proximal gradient mapping method to minimize above objective function takes the update of the form:

$$x_i^{k+1} = \text{prox}_{t g_i}(x_i^k - t.D_i^T G_{1/c}^{f(\cdot), g_{i-1}(D_i \cdot)}(x_i^k)), \quad (6.11)$$

for  $t > 0$  and  $c > 0$ , which are learnable parameters of ML-CSC framework. The proximal of  $g_{i-1}(D_i \cdot)$  involves a composite term, an approximation for representations  $x_i$  is proposed such that  $x_{i-1} = D_i x_i$ . Therefore, for  $i$ th layer of the framework the update of (6.11) can be modified as,

$$x_i^{k+1} = \text{prox}_{t g_i}(x_i^k - t.D_i^T G_{1/c}^{f(\cdot), g_{i-1}} x_{i-1}^k), \quad (6.12)$$

with approximation of composite term, the calculation of proximal mapping of  $g$  term becomes soft thresholding that is  $\text{prox}_{g(x)} = \tau_\lambda(x)$ . For a multi-layer model, the ISTA update step can be written as:

$$x_i^{k+1} \leftarrow \mathcal{T}_t \lambda_i \left[ x_i^k - \frac{t}{c} D_i^T (x_{i-1}^k - \mathcal{T}_t \lambda_i \mu (x_{i-1}^k - D_{i-1}^T (D_i x_{i-1}^k - y))) \right], \quad (6.13)$$

The algorithm employing multi-layer ISTA takes images in  $k$ -space along with their CS measurements, learns the mappings from CS measurements with ground-truth images, and saves the mappings in the form of model parameters of ML-CSC. The testing module reconstructs the test images with the help of trained ML-ISTA learned parameters and gives PSNR/SSIM of recovered images. The framework initializes the dictionaries with Xavier initialization [130] for all layers. The multilayered basis pursuit algorithm employing iterative thresholding algorithm for multi-layer pursuit is given in Algorithm-9.

---

**Algorithm 9** ML-ISTA for CS-MRI

---

**Require:**  $\{y_i, x_i\}_{i=1}^N$ , dictionaries  $D_i$  and  $\lambda_i$

**Ensure:** Model containing parameters of neural network for CS-MRI image restoration.

$\{y_i, x_i\}_{i=1}^N$  are images in  $k$  space and their corresponding CS measurements,  $N$  is size of trainset

1: Init Set  $x_0^k = y$ ,  $\forall k$  and  $x_L^1 = 0$

2: **for**  $k=1:K$  **do**

$\hat{x}_i \leftarrow D_{(i,L)} x_L^k \quad \forall_i [0, L-1]$

3: **for**  $i=1:L$  **do**

$x_i^{k+1} \leftarrow \mathcal{T}_{\mu_i, \lambda_i} (\hat{x}_i - \mu_i D_i^T (D_i \hat{x}_i - x_{i-1}^{k+1}))$  ▷ Find set of representations

4: **end for**

5: **end for** ▷ Save model with Epoch number

---

The training module of the proposed CS-MRI restoration framework takes pairs of images, their corresponding CS measurements in  $k$ -space along with undersampling masks

and trains the model parameters. The learning Algorithm-9, uses ML-ISTA unfoldings to increase depth of learning framework without incurring cost of additional parameters. The testing framework takes the test set images, undersampling masks, and the parameters of the saved model to reconstruct undersampled test images and comparing the metrics of PSNR and SSIM with ground truth images along with restoration time. The pseudo-code for training and testing the CS-MRI framework is given below.

---

**ML-ISTA CS-MRI PSEODOCODE**

---

• **Training ML-ISTA with MRI Train Data:**

**Require:** Input: MR images, their CS measurements, random masks,  $D_i$  and  $\lambda_i$

**Ensure:** Model containing parameters of ML-ISTA for CS-MRI reconstruction.

- Run encoding to find representations for given layers

**for** Unfoldings= 1,2..., **do**

- Find representations using ML-ISTA with global pursuits
- Train and save model with EPOCH number and minimum loss

**end for**

---

• **Reconstruct Test set with trained model.**

**Require:** Trained model, MR test images, undersampling masks

**Ensure:** Reconstructed images with PSNR , SSIM and reconstruction CPU/GPU time

**for** i=1:N **do**

- Run ML-ISTA with learned dictionaries from trained model
- Reconstruct MR images
- Calculate PSNR and SSIM of testset

▷ N is the size of test set;

**end for**

---

### 6.2.6 Architecture

The proposed framework takes advantage of iterative thresholding-based pursuit algorithms along with the learning capability of the ML-CSC model and maps each learning update to

an ML-ISTA-based unfolding. The sparsifying transform is replaced with ML-ISTA learning, which is denoted by  $\mathcal{D}(\cdot)$ . In the multilayered basis pursuit algorithm,  $\mathcal{D}(\cdot)$  consists of six convolutional layers with ISTA-based pursuit (Equation (6.13)), as given in Algorithm-9. After replacing sparsifying transform with learnable parameter  $\mathcal{D}(\cdot)$ , Equation (6.4) becomes:

$$\hat{x} = \arg \min_x \frac{1}{2} \|\Phi x - y\|_2^2 + \lambda \|\mathcal{D}(x)\|_1, \quad (6.14)$$

The ISTA update for CS reconstruction is given by;

$$\begin{aligned} \mathcal{I}^k &= x^{k-1} - \rho \nabla \left( \frac{1}{2} \|\Phi x - y\|_2^2 \right), \\ x^k &= \arg \min_x \frac{1}{2} \|\Phi x - \mathcal{I}^k\|_2^2 + \lambda \|\mathcal{D}(x)\|_1, \end{aligned} \quad (6.15)$$

The ISTA update consists of the gradient of data fidelity term used for intermediate reconstruction of images.

### ISTA-Module

ISTA module generates a representation of intermediate results with learnable step size  $\rho$ .

### Proximal Mapping

The  $x^k$  takes ISTA update as input and computes  $x$  according to Equation (6.14).

## 6.2.7 Training Loss for CS-MRI Framework

Given pairs of MR images with their corresponding CS measurements, the framework produces intermediate reconstructions denoted by  $x_i^{rec}$ . The loss function is designed to seek reduced discrepancy between the input images and the intermediate reconstruction satisfying symmetry constraint of  $\tilde{\mathcal{D}}^k \mathcal{D}^k = I, \forall k = 1, 2, \dots$ , in check. The end to end loss is thus defined as:

$$\mathcal{L}_{discrepancy} = \frac{1}{IN} \sum_{i=1}^N \|x_i^{rec} - x_i\|_2^2$$

$$\mathcal{L}_{constraint} = \frac{1}{IN} \sum_{i=1}^N \sum_{k=1}^U \|\tilde{D}^k(D^k(x_i)) - x_i\|_2^2$$

$$\mathcal{L}_{total} = \mathcal{L}_{discrepancy} + \lambda \mathcal{L}_{constraint} \quad (6.16)$$

where  $I$  is the the image size ,  $N$  are the number of images,  $U$  are ML-ISTA unfoldings ,  $\lambda$  is the learning rate in above equations.

The framework employed a learning rate of 1e-4 with Adam optimizer [131] and learning rate decay factor of 0.2 for Pytorch implementation.

Table 6.1: Average PSNR/SSIM and restoration times of reconstructed Knee MR images

Framework	CS Ratio				GPU Time
	20%	30%	40%	50%	
	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	
ISTA-Net[68]	34.45/0.8621	36.87/0.9132	39.16/0.9474	41.74/0.9707	0.04289s
<b>Ours</b>	<b>36.93/0.9262</b>	<b>39.68/0.9569</b>	<b>42.04/0.9738</b>	<b>43.44/0.9824</b>	<b>0.0977s</b>

### 6.2.8 Dataset

For training of the CS-MRI framework, a dataset comprising 622 knee MR images of adult patients (Male and Female) aged between 18 to 40 years, with the coronal view and Proton density (PD) fat saturation, collected from a 1.5T imaging unit (Siemens MAGNETOM Symphony). The MR images were acquired to diagnose ACL tear at Hospital Kuala Lumpur (HKL) and a certified radiologist labeled the images for ACL tear types of complete, partial, and normal knee categories.

## 6.3 Empirical Results

The proposed framework employed a random mask for undersampling of MR images in  $k$  space for CS-MRI reconstruction algorithms and results were obtained for different CS-ratios. Results for CS-MRI reconstruction of knee MR dataset, CS ratios, PSNR/SSIM,

and restoration times are given in Table 6.1. The proposed framework has been trained on a knee dataset of 622 images and restoration performance is shown on the test set with encouraging and comparable results as compared to the state-of-the-art framework which is also trained and tested with the same dataset and other parameters.

#### **6.4 Summary**

In this chapter, a CS-MRI restoration framework based on multi-layer convolutional sparse coding, employing iterative thresholding algorithms for basis pursuits to learn parameters of nonlinear mappings from undersampled MR images acquired in  $k$ -space proposed in our earlier work was compared with state of the art. Empirical results on custom acquired knee MR datasets showed that the proposed CS-MRI framework successfully trained for the desired mapping from CS measurements effectively for knee MR images after training on masks with different CS ratios with improved convergence and reconstruction results at the cost of a smaller increase in learnable parameters of a deep neural network. The successful application of proposed methods shows that it can be used for different anatomies due to its generalizability and can be easily integrated into the MRI imaging pipeline to improve acquisitions.



## CHAPTER 7

### CONCLUSION AND FUTURE WORK

#### 7.1 Summary of the Contributions

In this research work, we proposed and implemented multi-layer convolutional sparse coding frameworks for the inverse problem of compressive sensing MRI restoration and classification of biomedical images. Our contributions include,

- A compressive sensing MRI framework for restoration of undersampled MR images was presented employing multi-layer convolutional sparse coding (ML-CSC) theory and compared it with state-of-the-art for multiple benchmark datasets. With help of extensive empirical results, the effectiveness of the proposed frameworks has been demonstrated for fast convergence and improved peak signal to noise ratios and structure similarity index measurements benchmarks on multiple datasets.
- Furthermore, a biomedical image classification framework is proposed employing ML-CSC and pursuit algorithms compared with state-of-the-art deep learning-based architectures. Extensive experiments and results were given to show the effectiveness of the proposed frameworks employing different pursuit algorithms and comparing them with state of the art self-labeled datasets acquired at Hospital Kuala Lumpur.
- Finally, the CS-MRI restoration framework proposed in chapter-3 was analyzed and compared on knee MR dataset with state of the art keeping all parameters the same and showing the better restoration quality of our proposed framework on benchmarks of PSNR/SSIM and restoration times.

## 7.2 Future Directions

In future work, the research work can be extended with

- Implementation of the CS-MRI restoration framework for parallel coil MRI using datasets acquired in specific planes as training and testsets to further improve the models.
- Framework used to restore images directly taken in undersampled form.
- In addition, the research work can be extended to design more generalized classifiers using transfer learning, which can adapt to different datasets without requiring training from scratch and use them for improvement in the performance of classifiers of biomedical images.
- In summary, the models can be further improved with training on diverse datasets and their potential integration into imaging pipelines in coordination with clinicians for improving the efficiency of biomedical imaging.

## **Appendices**

## **APPENDIX A**

### **EXPERIMENTAL EQUIPMENT**

The experimental setup for the research work consisted HP Omen workstation with i9-9990 and NVIDIAI 2080 Ti GPU.

## REFERENCES

- [1] J. Sulam, A. Aberdam, A. Beck, and M. Elad, "On multi-layer basis pursuit, efficient algorithms and convolutional neural networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019.
- [2] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*, 1st. Springer Publishing Company, Incorporated, 2010, ISBN: 144197010X.
- [3] Y. LeCun *et al.*, "Generalization and network design strategies," *Connectionism in Perspective*, vol. 19, pp. 143–155, 1989.
- [4] Y. LeCun, P. Haffner, L. Bottou, and Y. Bengio, "Object recognition with gradient-based learning," in *Shape, contour and grouping in computer vision*, Springer, 1999, pp. 319–345.
- [5] A. Shoeibi *et al.*, "Automatic diagnosis of schizophrenia in EEG signals using CNN-LSTM models," *Frontiers in Neuroinformatics*, vol. 15, 2021.
- [6] D. Perdios, M. Vonlanthen, F. Martinez, M. Arditì, and J.-P. Thiran, "CNN-Based ultrasound image reconstruction for ultrafast displacement tracking," *IEEE Transactions on Medical Imaging*, vol. 40, no. 3, pp. 1078–1089, 2021.
- [7] B. Barros, P. Lacerda, C. Albuquerque, and A. Conci, "Pulmonary covid-19: Learning spatiotemporal features combining CNN and LSTM networks for lung ultrasound video classification," *Sensors*, vol. 21, no. 16, p. 5486, 2021.
- [8] F.-P. An and Z.-W. Liu, "Medical image segmentation algorithm based on feedback mechanism cnn," *Contrast media & molecular imaging*, vol. 2019, 2019.
- [9] F. Renard, S. Guedria, N. De Palma, and N. Vuillerme, "Variability and reproducibility in deep learning for medical image segmentation," *Scientific Reports*, vol. 10, no. 1, pp. 1–16, 2020.
- [10] M. H. Hesamian, W. Jia, X. He, and P. Kennedy, "Deep learning techniques for medical image segmentation: Achievements and challenges," *Journal of Digital Imaging*, vol. 32, no. 4, pp. 582–596, 2019.
- [11] A. Zhao, G. Balakrishnan, F. Durand, J. V. Guttag, and A. V. Dalca, "Data augmentation using learned transformations for one-shot medical image segmentation," in

*Proceedings of the IEEE conference on computer vision and pattern recognition*, 2019, pp. 8543–8553.

- [12] S. S. Yadav and S. M. Jadhav, “Deep convolutional neural network based medical image classification for disease diagnosis,” *Journal of Big Data*, vol. 6, no. 1, p. 113, 2019.
- [13] A. Kumar, J. Kim, D. Lyndon, M. Fulham, and D. Feng, “An ensemble of fine-tuned convolutional neural networks for medical image classification,” *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 1, pp. 31–40, 2016.
- [14] L. Faes *et al.*, “Automated deep learning design for medical image classification by health-care professionals with no coding experience: A feasibility study,” *The Lancet Digital Health*, vol. 1, no. 5, e232–e242, 2019.
- [15] M. T. McCann, K. H. Jin, and M. Unser, “Convolutional neural networks for inverse problems in imaging: A review,” *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 85–95, 2017.
- [16] A. Lucas, M. Iliadis, R. Molina, and A. K. Katsaggelos, “Using deep neural networks for inverse problems in imaging: Beyond analytical methods,” *IEEE Signal Processing Magazine*, vol. 35, no. 1, pp. 20–36, 2018.
- [17] Y. Bai, W. Chen, J. Chen, and W. Guo, “Deep learning methods for solving linear inverse problems: Research directions and paradigms,” *Signal Processing*, p. 107 729, 2020.
- [18] I. Häggström, C. R. Schmittlein, G. Campanella, and T. J. Fuchs, “Deeppet: A deep encoder–decoder network for directly solving the pet image reconstruction inverse problem,” *Medical Image Analysis*, vol. 54, pp. 253–262, 2019.
- [19] C. M. Hyun, H. P. Kim, S. M. Lee, S. Lee, and J. K. Seo, “Deep learning for undersampled mri reconstruction,” *Physics in Medicine & Biology*, vol. 63, no. 13, p. 135 007, 2018.
- [20] T. M. Quan, T. Nguyen-Duc, and W.-K. Jeong, “Compressed sensing mri reconstruction using a generative adversarial network with a cyclic loss,” *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1488–1497, 2018.
- [21] Q. Liu, C. Zhang, Q. Guo, H. Xu, and Y. Zhou, “Adaptive sparse coding on pca dictionary for image denoising,” *The Visual Computer*, vol. 32, no. 4, pp. 535–549, 2016.

- [22] Y. Liu, S. Canu, P. Honeine, and S. Ruan, "Mixed integer programming for sparse coding: Application to image denoising," *IEEE Transactions on Computational Imaging*, vol. 5, no. 3, pp. 354–365, 2019.
- [23] L. Pfister and Y. Bresler, "Automatic parameter tuning for image denoising with learned sparsifying transforms," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2017, pp. 6040–6044.
- [24] Y. Lei *et al.*, "A denoising algorithm for ct image using low-rank sparse coding," in *Medical Imaging 2018: Image Processing*, International Society for Optics and Photonics, vol. 10574, 2018, 105741P.
- [25] J. Tong, Y. Zhao, P. Zhang, L. Chen, and L. Jiang, "Mri brain tumor segmentation based on texture features and kernel sparse coding," *Biomedical Signal Processing and Control*, vol. 47, pp. 387–392, 2019.
- [26] Y. Li, M. Tofighi, J. Geng, V. Monga, and Y. C. Eldar, "An algorithm unrolling approach to deep blind image deblurring," *arXiv preprint*, 2019.
- [27] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 31, no. 2, pp. 210–227, 2008.
- [28] Y.-F. Yu, D.-Q. Dai, C.-X. Ren, and K.-K. Huang, "Discriminative multi-scale sparse coding for single-sample face recognition with occlusion," *Pattern Recognition*, vol. 66, pp. 302–312, 2017.
- [29] H. Zhang, J. Yang, J. Xie, J. Qian, and B. Zhang, "Weighted sparse coding regularized nonconvex matrix regression for robust face recognition," *Information Sciences*, vol. 394, pp. 1–17, 2017.
- [30] R. Zhang, J. Shen, F. Wei, X. Li, and A. K. Sangaiah, "Medical image classification based on multi-scale non-negative sparse coding," *Artificial Intelligence in Medicine*, vol. 83, pp. 44–51, 2017.
- [31] M. Elad and M. Aharon, "Image denoising via learned dictionaries and sparse representation," in *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, IEEE, vol. 1, 2006, pp. 895–900.
- [32] J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," *IEEE Transactions on Image Processing*, vol. 17, no. 1, pp. 53–69, 2008.
- [33] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Non-local sparse models for image restoration," in *2009 IEEE 12th International Conference on Computer Vision*, 2009, pp. 2272–2279.

- [34] V. Pappyan and M. Elad, "Multi-scale patch-based image restoration," *IEEE Transactions on Image Processing*, vol. 25, no. 1, pp. 249–261, 2016.
- [35] V. Pappyan, Y. Romano, and M. Elad, "Convolutional neural networks analyzed via convolutional sparse coding," *The Journal of Machine Learning Research*, vol. 18, no. 1, pp. 2887–2938, 2017.
- [36] V. Pappyan, J. Sulam, and M. Elad, "Working locally thinking globally: Theoretical guarantees for convolutional sparse coding," *IEEE Transactions on Signal Processing*, vol. 65, no. 21, pp. 5687–5701, 2017.
- [37] F. Heide, W. Heidrich, and G. Wetzstein, "Fast and flexible convolutional sparse coding," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 5135–5143.
- [38] V. Pappyan, Y. Romano, J. Sulam, and M. Elad, "Theoretical foundations of deep learning via sparse representations: A multilayer sparse model and its connection to convolutional neural networks," *IEEE Signal Processing Magazine*, vol. 35, no. 4, pp. 72–89, 2018.
- [39] J. Sulam, V. Pappyan, Y. Romano, and M. Elad, "Multilayer convolutional sparse modeling: Pursuit and dictionary learning," *IEEE Transactions on Signal Processing*, vol. 66, no. 15, pp. 4090–4104, 2018.
- [40] M. Kabkab, "The case for spatial pooling in deep convolutional sparse coding," in *2017 51st Asilomar Conference on Signals, Systems, and Computers*, IEEE, 2017, pp. 367–371.
- [41] I. Rey-Otero, J. Sulam, and M. Elad, "Variations on the convolutional sparse coding model," *IEEE Transactions on Signal Processing*, vol. 68, pp. 519–528, 2020.
- [42] R. G. Baraniuk, T. Goldstein, A. C. Sankaranarayanan, C. Studer, A. Veeraraghavan, and M. B. Wakin, "Compressive video sensing: Algorithms, architectures, and applications," *IEEE Signal Processing Magazine*, vol. 34, no. 1, pp. 52–66, 2017.
- [43] X. X. Zhu and R. Bamler, "Demonstration of super-resolution for tomographic SAR imaging in urban environment," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 8, pp. 3150–3157, 2012.
- [44] L. F. Polanía and R. I. Plaza, "Compressed sensing ECG using restricted Boltzmann machines," *Biomedical Signal Processing and Control*, vol. 45, pp. 237–245, 2018.
- [45] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine: An Of-*



*ficial Journal of the International Society for Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1182–1195, 2007.

- [46] A. Mousavi, A. B. Patel, and R. G. Baraniuk, “A deep learning approach to structured signal recovery,” in *2015 53rd annual allerton conference on communication, control, and computing (Allerton)*, IEEE, 2015, pp. 1336–1343.
- [47] M. Iliadis, L. Spinoulas, and A. K. Katsaggelos, “Deep fully-connected networks for video compressive sensing,” *Digital Signal Processing*, vol. 72, pp. 9–18, 2018.
- [48] S. Wang *et al.*, “Accelerating magnetic resonance imaging via deep learning,” in *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, IEEE, 2016, pp. 514–517.
- [49] J. Yang, Y. Zhang, and W. Yin, “A fast alternating direction method for TVL1-L2 signal reconstruction from partial Fourier data,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 288–297, 2010.
- [50] X. Qu *et al.*, “Undersampled MRI reconstruction with patch-based directional wavelets,” *Magnetic Resonance Imaging*, vol. 30, no. 7, pp. 964–977, 2012.
- [51] 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, 2010.
- [52] Y. Wang and L. Ying, “Compressed sensing dynamic cardiac cine MRI using learned spatio temporal dictionary,” *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 4, pp. 1109–1120, 2014.
- [53] S. Ikram, S. Zubair, J. A. Shah, I. M. Qureshi, A. Wahid, and A. U. Khan, “Enhancing MR image reconstruction using block dictionary learning,” *IEEE Access*, vol. 7, pp. 158 434–158 444, 2019.
- [54] S.-M. Gho, Y. Nam, S.-Y. Zho, E. Y. Kim, and D.-H. Kim, “Three dimension double inversion recovery gray matter imaging using compressed sensing,” *Magnetic Resonance Imaging*, vol. 28, no. 10, pp. 1395–1402, 2010.
- [55] E. M. Eksioğlu, “Decoupled algorithm for MRI reconstruction using nonlocal block matching model: BM3D-MRI,” *Journal of Mathematical Imaging and Vision*, vol. 56, no. 3, pp. 430–440, 2016.
- [56] Z. Zhang, T.-P. Jung, S. Makeig, Z. Pi, and B. D. Rao, “Spatiotemporal sparse Bayesian learning with applications to compressed sensing of multichannel physiological signals,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 6, pp. 1186–1197, 2014.

- [57] S. Liu, J. Jia, Y. D. Zhang, and Y. Yang, "Image reconstruction in electrical impedance tomography based on structure-aware sparse Bayesian learning," *IEEE Transactions on Medical Imaging*, vol. 37, no. 9, pp. 2090–2102, 2018.
- [58] S. Liu, R. Cao, Y. Huang, T. Ouypornkochagorn, and J. Jia, "Time sequence learning for electrical impedance tomography using Bayesian spatiotemporal priors," *IEEE Transactions on Instrumentation and Measurement*, 2020.
- [59] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3142–3155, 2017.
- [60] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, "Deep convolutional neural network for inverse problems in imaging," *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4509–4522, 2017.
- [61] J. M. Wolterink, T. Leiner, M. A. Viergever, and I. Išgum, "Generative adversarial networks for noise reduction in low-dose CT," *IEEE Transactions on Medical Imaging*, vol. 36, no. 12, pp. 2536–2545, 2017.
- [62] G. Yang *et al.*, "DAGAN: Deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1310–1321, 2017.
- [63] M. Mardani *et al.*, "Deep generative adversarial neural networks for compressive sensing MRI," *IEEE Transactions on Medical Imaging*, vol. 38, no. 1, pp. 167–179, 2018.
- [64] G. Luo, N. Zhao, W. Jiang, E. S. Hui, and P. Cao, "MRI reconstruction using deep Bayesian estimation," *Magnetic Resonance in Medicine*, 2020.
- [65] A. Adler, D. Boubilil, and M. Zibulevsky, "Block-based compressed sensing of images via deep learning," in *2017 IEEE 19th International Workshop on Multimedia Signal Processing (MMSP)*, IEEE, 2017, pp. 1–6.
- [66] B. Yaman, S. A. H. Hosseini, S. Moeller, J. Ellermann, K. Uğurbil, and M. Akçakaya, "Self-supervised physics-based deep learning MRI reconstruction without fully-sampled data," in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, IEEE, 2020, pp. 921–925.
- [67] K. Gregor and Y. LeCun, "Learning fast approximations of sparse coding," in *Proceedings of the 27th International Conference on International Conference on Machine Learning*, 2010, pp. 399–406.

- [68] J. Zhang and B. Ghanem, "ISTA-Net: Interpretable optimization-inspired deep network for image compressive sensing," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 1828–1837.
- [69] K. Kulkarni, S. Lohit, P. Turaga, R. Kerviche, and A. Ashok, "Reconnet: Non-iterative reconstruction of images from compressively sensed measurements," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 449–458.
- [70] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, vol. 57, no. 11, pp. 1413–1457, 2004.
- [71] J. Adler and O. Öktem, "Learned primal-dual reconstruction," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1322–1332, 2018.
- [72] H. Gupta, K. H. Jin, H. Q. Nguyen, M. T. McCann, and M. Unser, "CNN-based projected gradient descent for consistent CT image reconstruction," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1440–1453, 2018.
- [73] K. Hammernik *et al.*, "Learning a variational network for reconstruction of accelerated MRI data," *Magnetic Resonance in Medicine*, vol. 79, no. 6, pp. 3055–3071, 2018.
- [74] C. Metzler, A. Mousavi, and R. Baraniuk, "Learned D-AMP: Principled neural network based compressive image recovery," in *Advances in Neural Information Processing Systems*, 2017, pp. 1772–1783.
- [75] C. A. Metzler, A. Maleki, and R. G. Baraniuk, "From denoising to compressed sensing," *IEEE Transactions on Information Theory*, vol. 62, no. 9, pp. 5117–5144, 2016.
- [76] Y. Yang, J. Sun, H. Li, and Z. Xu, "Deep ADMM-Net for compressive sensing MRI," in *Proceedings of the 30th international conference on neural information processing systems*, 2016, pp. 10–18.
- [77] Y. Yang, J. Sun, H. Li, and Z. Xu, "ADMM-CSNet: A deep learning approach for image compressive sensing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 3, pp. 521–538, 2020.
- [78] A. Sriram, J. Zbontar, T. Murrell, C. L. Zitnick, A. Defazio, and D. K. Sodickson, "GrappaNet: Combining parallel imaging with deep learning for multi-coil MRI reconstruction," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 14 315–14 322.

- [79] I. Marivani, E. Tsiligianni, B. Cornelis, and N. Deligiannis, "Learned multimodal convolutional sparse coding for guided image super-resolution," in *2019 IEEE International Conference on Image Processing (ICIP)*, 2019, pp. 2891–2895.
- [80] A. Ahmed, S. Kun, R. A. Memon, J. Ahmed, and G. Tefera, "Convolutional sparse coding using wavelets for single image super-resolution," *IEEE Access*, vol. 7, pp. 121 350–121 359, 2019.
- [81] D. Simon and M. Elad, "Rethinking the CSC model for natural images," *arXiv preprint arXiv:1909.05742*, 2019.
- [82] A. Serrano, E. Garces, B. Masia, and D. Gutierrez, "Convolutional sparse coding for capturing high-speed video content," in *Computer Graphics Forum*, Wiley Online Library, vol. 36, 2017, pp. 380–389.
- [83] P. V. Arun, K. B., and A. Porwal, "Spatial-spectral feature based approach towards convolutional sparse coding of hyperspectral images," *Computer Vision and Image Understanding*, vol. 188, p. 102 797, 2019.
- [84] M. Al-Madani, U. b. Waheed, and M. Masood, "Fast and accurate dictionary learning for seismic data denoising using convolutional sparse coding," in *SEG Technical Program Expanded Abstracts 2019*, Society of Exploration Geophysicists, 2019, pp. 4645–4649.
- [85] I. Marivani, E. Tsiligianni, B. Cornelis, and N. Deligiannis, "Multimodal deep unfolding for guided image super-resolution," *IEEE Transactions on Image Processing*, vol. 29, pp. 8443–8456, 2020.
- [86] D. Carrera, G. Boracchi, A. Foi, and B. Wohlberg, "Sparse overcomplete denoising: Aggregation versus global optimization," *IEEE Signal Processing Letters*, vol. 24, no. 10, pp. 1468–1472, 2017.
- [87] J. Xiong, Q. Liu, Y. Wang, and X. Xu, "A two-stage convolutional sparse prior model for image restoration," *Journal of Visual Communication and Image Representation*, vol. 48, pp. 268–280, 2017.
- [88] P. C. Prokopiou and G. D. Mitsis, "Modeling of the BOLD signal using event-related simultaneous EEG-fMRI and convolutional sparse coding analysis," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019, pp. 181–184.
- [89] T. Nguyen-Duc and W. Jeong, "Compressed sensing dynamic MRI reconstruction using multi-scale 3D convolutional sparse coding with elastic net regularization," in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, 2018, pp. 332–335.

- [90] T. Nguyen-Duc, T. M. Quan, and W.-K. Jeong, "Frequency-splitting dynamic MRI reconstruction using multi-scale 3D convolutional sparse coding and automatic parameter selection," *Medical Image Analysis*, vol. 53, pp. 179–196, 2019.
- [91] J. Xiong, H. Lu, M. Zhang, and Q. Liu, "Convolutional sparse coding in gradient domain for MRI reconstruction," *Acta Automatica Sinica*, vol. 43, no. 10, pp. 1841–1849, 2017.
- [92] Y. Huang, L. Shao, and A. F. Frangi, "DOTE: Dual convolutional filter learning for super-resolution and cross-modality synthesis in MRI," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2017, pp. 89–98.
- [93] P. Guo, P. Wang, J. Zhou, S. Jiang, and V. M. Patel, "Multi-institutional collaborations for improving deep learning-based magnetic resonance image reconstruction using federated learning," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 2423–2432.
- [94] T. M. Quan and W.-K. Jeong, "Compressed sensing dynamic MRI reconstruction using GPU-accelerated 3D convolutional sparse coding," in *International conference on medical image computing and computer-assisted intervention*, Springer, 2016, pp. 484–492.
- [95] T. M. Quan and W. Jeong, "Compressed sensing reconstruction of dynamic contrast enhanced MRI using GPU-accelerated convolutional sparse coding," in *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, 2016, pp. 518–521.
- [96] Y. Huang, L. Shao, and A. F. Frangi, "Simultaneous super-resolution and cross-modality synthesis of 3D medical images using weakly-supervised joint convolutional sparse coding," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 6070–6079.
- [97] Y. Zhang *et al.*, "Weakly supervised multi-needle detection in 3D ultrasound images with bidirectional convolutional sparse coding," in *Medical Imaging 2020: Ultrasonic Imaging and Tomography*, International Society for Optics and Photonics, vol. 11319, 2020, p. 1131914.
- [98] P. Bao *et al.*, "Convolutional sparse coding for compressed sensing CT reconstruction," *IEEE Transactions on Medical Imaging*, vol. 38, no. 11, pp. 2607–2619, 2019.
- [99] J. Yan and Y. Zhu, "Convolutional sparse coded dynamic brain functional connectivity," *Neural Processing Letters*, vol. 52, no. 3, pp. 1881–1892, 2020.

- [100] E. Plaut and R. Giryes, "A greedy approach to  $\ell_{0,\infty}$  based convolutional sparse coding," *SIAM Journal on Imaging Sciences*, vol. 12, no. 1, pp. 186–210, 2019.
- [101] B. Sun, N.-h. Tsai, F. Liu, R. Yu, and H. Su, "Adversarial defense by stratified convolutional sparse coding," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 11 447–11 456.
- [102] J. Liu, C. Garcia-Cardona, B. Wohlberg, and W. Yin, "First-and second-order methods for online convolutional dictionary learning," *SIAM Journal on Imaging Sciences*, vol. 11, no. 2, pp. 1589–1628, 2018.
- [103] J. Liu, C. Garcia-Cardona, B. Wohlberg, and W. Yin, "Online convolutional dictionary learning," in *2017 IEEE International Conference on Image Processing (ICIP)*, 2017, pp. 1707–1711.
- [104] T. M. Quan, J. Choi, H. Jeong, and W.-K. Jeong, "An intelligent system approach for probabilistic volume rendering using hierarchical 3D convolutional sparse coding," *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, no. 1, pp. 964–973, 2017.
- [105] I. Y. Chun and J. A. Fessler, "Convolutional dictionary learning: Acceleration and convergence," *IEEE Transactions on Image Processing*, vol. 27, no. 4, pp. 1697–1712, 2017.
- [106] Q. Liu and H. Leung, "Synthesis-analysis deconvolutional network for compressed sensing," in *2017 IEEE International Conference on Image Processing (ICIP)*, 2017, pp. 1940–1944.
- [107] T. L. Sanders *et al.*, "Incidence of anterior cruciate ligament tears and reconstruction: A 21-year population-based study," *The American Journal of Sports Medicine*, vol. 44, no. 6, pp. 1502–1507, 2016.
- [108] G. S. Perrone, B. L. Proffen, A. M. Kiapour, J. T. Sieker, B. C. Fleming, and M. M. Murray, "Bench-to-bedside: Bridge-enhanced anterior cruciate ligament repair," *Journal of Orthopaedic Research*, vol. 35, no. 12, pp. 2606–2612, 2017.
- [109] H. Louboutin *et al.*, "Osteoarthritis in patients with anterior cruciate ligament rupture: A review of risk factors," *The Knee*, vol. 16, no. 4, pp. 239–244, 2009.
- [110] K Satku, V. Kumar, and S. Ngoi, "Anterior cruciate ligament injuries. to counsel or to operate?" *The Journal of Bone and Joint Surgery. British volume*, vol. 68, no. 3, pp. 458–461, 1986.

- [111] J. H. Mink, T. Levy, and J. Crues 3rd, "Tears of the anterior cruciate ligament and menisci of the knee: MR imaging evaluation.," *Radiology*, vol. 167, no. 3, pp. 769–774, 1988.
- [112] Z.-Q. Zhao, P. Zheng, S.-t. Xu, and X. Wu, "Object detection with deep learning: A review," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 11, pp. 3212–3232, 2019.
- [113] 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 transactions on Medical Imaging*, vol. 37, no. 2, pp. 491–503, 2017.
- [114] J. B. Carvalho, J.-M. Moreira, M. A. Figueiredo, and N. Papanikolaou, "Automatic detection and segmentation of lung lesions using deep residual CNNs," in *2019 IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE)*, IEEE, 2019, pp. 977–983.
- [115] Y. Zhou *et al.*, "Collaborative learning of semi-supervised segmentation and classification for medical images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 2079–2088.
- [116] W. Luo, J. Li, J. Yang, W. Xu, and J. Zhang, "Convolutional sparse autoencoders for image classification," *IEEE transactions on neural networks and learning systems*, vol. 29, no. 7, pp. 3289–3294, 2017.
- [117] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.
- [118] P. Rajpurkar *et al.*, "CheXNet: Radiologist-level pneumonia detection on chest X-Rays with deep learning," *arXiv preprint arXiv:1711.05225*, 2017.
- [119] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "Chestx-ray8: Hospital-scale chest X-Ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2097–2106.
- [120] A. Y. Hannun *et al.*, "Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network," *Nature Medicine*, vol. 25, no. 1, p. 65, 2019.
- [121] F. Liu *et al.*, "Fully automated diagnosis of anterior cruciate ligament tears on knee MR images by using deep learning," *Radiology: Artificial Intelligence*, vol. 1, no. 3, p. 180091, 2019.

- [122] P. D. Chang, T. T. Wong, and M. J. Rasiej, "Deep learning for detection of complete anterior cruciate ligament tear," *Journal of Digital Imaging*, vol. 32, no. 6, pp. 980–986, 2019.
- [123] Y. Kamatsuki *et al.*, "Complete tear of the lateral meniscus posterior root is associated with meniscal extrusion in anterior cruciate ligament deficient knees," *Journal of Orthopaedic Research*, vol. 36, no. 7, pp. 1894–1900, 2018.
- [124] N. Khan, J. Shah, and I. Stavness, "Bridgeout: Stochastic bridge regularization for deep neural networks," *IEEE Access*, vol. 6, pp. 42 961–42 970, 2018.
- [125] A. Aberdam, J. Sulam, and M. Elad, "Multi-layer sparse coding: The holistic way," *SIAM Journal on Mathematics of Data Science*, vol. 1, no. 1, pp. 46–77, 2019.
- [126] 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.
- [127] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.
- [128] C. Murdock, M. Chang, and S. Lucey, "Deep component analysis via alternating direction neural networks," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 820–836.
- [129] A. Beck, *First-Order Methods in Optimization*. Philadelphia, PA: volume 25 of *MOS-SIAM Series on Optimization*. Society for Industrial and Applied Mathematics, 2017.
- [130] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, 2010, pp. 249–256.
- [131] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [132] D Summers, "Harvard whole brain atlas: [Www. med. harvard. edu/aanlib/home.html](http://www.med.harvard.edu/aanlib/home.html)," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 74, no. 3, pp. 288–288, 2003.
- [133] M. Murad, M. Bilal, A. Jalil, A. Ali, K. Mehmood, and B. Khan, "Efficient reconstruction technique for multi-slice CS-MRI using novel interpolation and 2D sampling scheme," *IEEE Access*, vol. 8, pp. 117 452–117 466, 2020.



- [134] J. Sun, H. Li, Z. Xu, *et al.*, “Deep ADMM-Net for compressive sensing MRI,” in *Advances in neural information processing systems*, 2016, pp. 10–18.
- [135] H. Makhmalbaf, A. Moradi, S. Ganji, and F. Omidi-Kashani, “Accuracy of lachman and anterior drawer tests for anterior cruciate ligament injuries,” *Archives of Bone and Joint Surgery*, vol. 1, no. 2, p. 94, 2013.
- [136] A. Hosny, C. Parmar, J. Quackenbush, L. H. Schwartz, and H. J. Aerts, “Artificial intelligence in radiology,” *Nature Reviews Cancer*, vol. 18, no. 8, pp. 500–510, 2018.
- [137] P. Rajpurkar, A. Y. Hannun, M. Haghpahani, C. Bourn, and A. Y. Ng, “Cardiologist-level arrhythmia detection with convolutional neural networks,” *arXiv e-prints*, arXiv-1707, 2017.
- [138] R. J. Tibshirani, J. Taylor, *et al.*, “The solution path of the generalized LASSO,” *The Annals of Statistics*, vol. 39, no. 3, pp. 1335–1371, 2011.
- [139] R. Chartrand and W. Yin, “Iteratively reweighted algorithms for compressive sensing,” in *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, IEEE, 2008, pp. 3869–3872.
- [140] P. Sprechmann, A. M. Bronstein, and G. Sapiro, “Learning efficient sparse and low rank models,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 9, pp. 1821–1833, 2015.
- [141] X. Chen, J. Liu, Z. Wang, and W. Yin, “Theoretical linear convergence of unfolded ISTA and its practical weights and thresholds,” in *Advances in Neural Information Processing Systems*, 2018, pp. 9061–9071.
- [142] P. Ablin, T. Moreau, M. Massias, and A. Gramfort, “Learning step sizes for unfolded sparse coding,” in *Advances in Neural Information Processing Systems*, 2019, pp. 13100–13110.
- [143] A. Bustin, N. Fuin, R. M. Botnar, and C. Prieto, “From compressed-sensing to artificial intelligence-based cardiac MRI reconstruction,” *Frontiers in Cardiovascular Medicine*, vol. 7, p. 17, 2020.
- [144] *FDA clears compressed sensing MRI acceleration technology from Siemens healthineers*, <https://www.siemens-healthineers.com/en-us/news/fda-clears-mri-technology-02-21-2017.html>, 2017.
- [145] A Bustin, T. Correia, I Rashid, G Cruz, R Neji, and R Botnar, “Highly accelerated 3D whole-heart isotropic submillimeter CMRA with non-rigid motion correction,” in *Proceedings of the 27th Annual Meeting of ISMRM*, 2019, pp. 11–16.

- [146] Y. Q. Mohsin, S. G. Lingala, E. DiBella, and M. Jacob, "Accelerated dynamic MRI using patch regularization for implicit motion compensation," *Magnetic Resonance in Medicine*, vol. 77, no. 3, pp. 1238–1248, 2017.
- [147] F. Knoll *et al.*, "Deep-learning methods for parallel magnetic resonance imaging reconstruction: A survey of the current approaches, trends, and issues," *IEEE Signal Processing Magazine*, vol. 37, no. 1, pp. 128–140, 2020.
- [148] A. S. Lundervold and A. Lundervold, "An overview of deep learning in medical imaging focusing on MRI," *Zeitschrift für Medizinische Physik*, vol. 29, no. 2, pp. 102–127, 2019.
- [149] A. Wahid, J. A. Shah, A. U. Khan, M. Ahmed, and H. Razali, "Multi-Layer Basis pursuit for compressed sensing MR image reconstruction," *IEEE Access*, vol. 8, pp. 186 222–186 232, 2020.
- [150] M. S. Lewicki and T. J. Sejnowski, "Coding time-varying signals using sparse shift-invariant representations," *Advances in Neural Information Processing Systems*, pp. 730–736, 1999.
- [151] B. Wohlberg, "Efficient algorithms for convolutional sparse representations," *IEEE Transactions on Image Processing*, vol. 25, no. 1, pp. 301–315, 2016.

