

MRI Scan Time Reduction Using Complex Valued AI in Image Space

*Thesis to be submitted in the partial fulfillment of the
requirements for the degree*

BACHELOR OF TECHNOLOGY

IN

AGRICULTURAL AND FOOD ENGINEERING

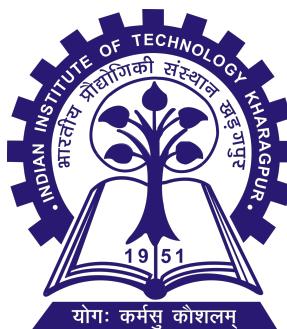
by

Md Tauheed Ansari

(21AG3AI06)

Under the guidance of

Prof. Mahesh Mohan M R



Department of Artificial Intelligence
Indian Institute of Technology, Kharagpur
29th April, 2025



Department of Artificial Intelligence
Indian Institute of Technology,
Kharagpur, India - 721302

CERTIFICATE

This is to certify that I have examined the thesis entitled MRI Scan Time Reduction Using Complex Valued AI in Image Space, submitted by **Md Tauheed Ansari (21AG3AI06)** an undergraduate student of the **Department of Agricultural and Food Engineering** in partial fulfillment for the award of degree of Bachelor of Technology. I hereby accord my approval of it as a study carried out and presented in a manner required for its acceptance in partial fulfillment for the undergraduate Degree for which it has been submitted. The thesis has fulfilled all the requirements as per the regulations of the Institute and has reached the standard needed for submission.

Guide's Signature

Date _____

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my supervisor, Prof. Mahesh Mohan M R, who introduced me to the exciting field of complex valued AI. His profound expertise and insightful guidance have been instrumental in the progress and success of my research.

I am also indebted to the faculty and staff, whose assistance have been invaluable throughout my academic journey. Their willingness to share knowledge and offer advice has greatly enriched my learning experience.

Special thanks are due to my peers and colleagues who have provided me with support and collaboration. Their perspectives and feedback have been a source of encouragement and growth.

Finally, and most importantly, my heartfelt thanks go to my family. Their unwavering support, endless patience, and belief in my abilities have been my strength and perseverance. This accomplishment is as much theirs as it is mine.

ABSTRACT

Magnetic Resonance Imaging (MRI) reconstruction remains a computationally intensive problem, particularly under accelerated acquisition protocols. Traditional techniques, such as SENSE (Pruessmann et al., 1999) and ESPIRiT (Uecker et al., 2014), solve inverse problems using physical models that exploit multi-coil sensitivity to resolve aliasing in undersampled data. However, with increasing demand for high-resolution, fast, and artifact-free imaging, there’s a pressing need to bridge model-driven and data-driven approaches. This work lays the foundation for applying Complex-Valued AI methods to MRI reconstruction—algorithms that operate directly on complex-valued inputs, preserving crucial phase information often discarded or poorly approximated by real-valued models.

To support this transition, a modular reconstruction framework was built, integrating both classical and modern tools. Core contributions include a custom PyTorch dataloader for multi-coil k-space datasets, preprocessing pipelines for real-time undersampling using fixed and learnable masks, and an interface for consistent image-domain reconstruction via RSS and inverse Fourier transforms. The Berkeley Advanced Reconstruction Toolbox (BART) was also employed to implement baseline methods like SENSE, ESPIRiT, and Total Variation regularization, providing solid benchmarks ahead of deep learning integration.

The system remains fully compatible with components of the fastMRI pipeline (Zbontar et al., 2018) and supports compressed sensing methods (Lustig et al., 2007), enabling rapid prototyping of deep learning architectures like DeepSense, GAN-SENSE, Swin Transformer, and ReconFormer. By establishing a robust mathematical and experimental foundation, this project paves the way for future research into complex-valued deep networks that remain faithful to the underlying physics of MRI.

Contents

1	Complex-Valued Neural Networks	1
1.1	Introduction	2
1.2	Motivation for Complex-Valued Signals and Their Applications	2
1.3	Advantages of using complex valued NNs	6
2	Magnetic Resonance Imaging and Problem Motivation	10
2.1	Introduction to MRI and its Importance in Medical Imaging	10
2.2	Justification: Clinical and Operational Motivations for Accelerated MRI	12
2.3	The MRI Signal Acquisition and Reconstruction Workflow	13
2.4	K-Space Domain and Undersampling	16
2.5	Relevance of Complex-Valued AI in MRI Reconstruction	17
2.6	Project Objectives and End Goals	18
3	Background and Literature Review	20
3.1	MRI Sampling Process Recall	20
3.2	K-Space Sampling Theory and Masking Techniques	21
3.3	Inverse Fourier Transform Theory for MRI Reconstruction	23
3.4	Undersampling and Aliasing in MRI Reconstruction	25
3.5	Review of MRI Reconstruction Methods and Advances	27
3.6	Chapter Conclusion	35
4	Methodology	36
4.1	Dataset (fastMRI)	36
4.2	BART: Berkeley Advanced Reconstruction Toolbox	39
4.3	Overview of the Reconstruction Pipeline	42
4.4	Custom Dataloader for Deep Learning-Based MRI Reconstruction . .	49
4.5	Chapter Conclusion	52

5 Results and Discussions	53
5.1 BART Implementation Results	53
5.2 Knee Scan Reconstruction Comparisons	56
5.3 Brain Scan Reconstruction Comparisons	57
5.4 Impact of Acceleration on ESPIRiT Reconstructions	58
5.5 Network Training	58
5.6 Discussion	59
6 Conclusion	62
Bibliography	66

Chapter 1

Complex-Valued Neural Networks

Complex-valued neural networks (CVNNs) have emerged as an essential paradigm in the landscape of computational intelligence and machine learning (1). Unlike traditional neural networks, which primarily operate on real-valued data, CVNNs extend this concept to the complex domain, harnessing the potential of complex numbers in representing and processing information. This transition is not merely a straightforward extension of well-understood real-valued neural networks but a fundamental shift in how neural networks understand and interact with data.

Complex numbers, characterized by their real and imaginary components, provide a richer framework for representing data. In the realm of CVNNs, this richness translates into an enhanced ability to capture intricate patterns and relationships in data that would otherwise be elusive or overly simplified in a real-valued context. The implications of this are profound, opening up new avenues for solving complex problems across various scientific and technological domains.

At the heart of CVNNs lies the recognition that many natural and technological phenomena are inherently complex-valued. From electromagnetic signals to quantum mechanical systems, the complex domain offers a more natural and expressive language for describing these phenomena. By leveraging this language, CVNNs offer a more nuanced and effective approach to tasks such as pattern recognition, signal processing, and complex system modeling.

This report delves into the world of complex-valued neural networks, exploring their foundational concepts, motivations, and diverse applications. It aims to shed light on the unique characteristics of CVNNs.

1.1 Introduction

CVNNs mark a significant departure from traditional neural network architectures. They are not merely an adaptation of existing networks to handle complex numbers but represent a fundamentally different approach to neural computation. This section provides a comprehensive overview of the architecture, operation, and unique aspects of CVNNs.

In CVNNs, both the neurons and the synaptic weights are complex-valued, allowing the network to process and learn from data represented as complex numbers. This change necessitates rethinking various aspects of neural network design, from activation functions to learning algorithms. For instance, the complex nature of the data requires the use of activation functions that can operate effectively in the complex domain, such as complex versions of the rectified linear unit (ReLU) or hyperbolic tangent functions (1).

The learning algorithms in CVNNs also undergo significant modifications to accommodate the complex nature of the weights and inputs. Gradient descent, a cornerstone of neural network training, is adapted to handle complex derivatives, often employing techniques based on Wirtinger calculus.

The implications of these architectural and algorithmic changes are far-reaching. CVNNs exhibit distinct behaviors and learning dynamics compared to their real-valued counterparts, often resulting in different patterns of convergence and model performance. These differences underscore the importance of developing specialized techniques and methodologies for designing, training, and deploying CVNNs.

As we explore further, the report will highlight the theoretical underpinnings of optimizing CVNNs, its practical implementations, and the unique challenges and opportunities they present.

1.2 Motivation for Complex-Valued Signals and Their Applications

The motivation for exploring and employing complex-valued neural networks (CVNNs) stems from the inherent nature of complex-valued signals in various scientific and technological fields. These signals, characterized by their amplitude and phase, provide

a more comprehensive understanding of the phenomena they represent. This section delves into the specific cases where complex-valued signals are crucial and the applications that benefit from the advanced processing capabilities of CVNNs.

1.2.1 Synthetic Aperture Radar for Agriculture

Synthetic Aperture Radar (SAR) technology, widely used in remote sensing, significantly benefits from complex-valued signal processing. SAR (Synthetic Aperture Radar) is a radar imaging technique that uses the movement of the radar antenna to synthesize a large aperture. This allows SAR to produce high-resolution images, regardless of weather conditions or light levels (see Fig. 1.1).

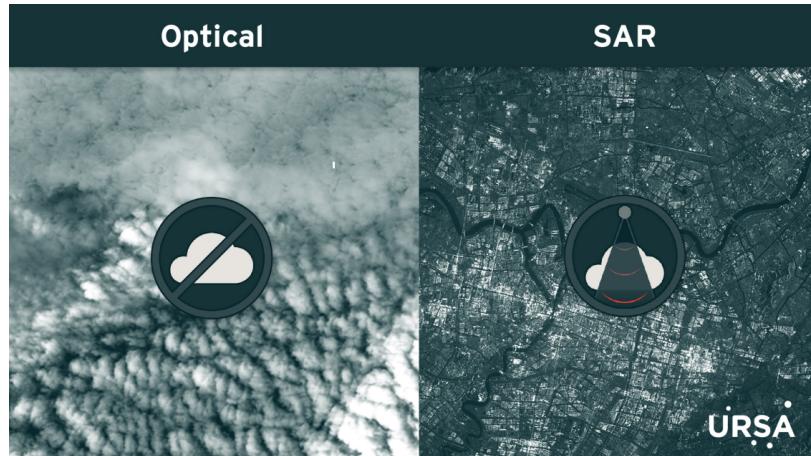


Figure 1.1: Optical vs SAR Imaging for the case of cloudy situation. SAR is able to capture the information whereas optical imaging struggles (Source: NASA).

In agriculture, SAR is instrumental in monitoring and managing agricultural activities. The complex-valued signals in SAR contain vital information about the land surface, including crop growth, soil moisture, and overall agricultural health.

The phase information in SAR images, often disregarded in traditional analysis, can provide unique insights into surface structures and vegetation patterns. CVNNs, with their ability to process and interpret these complex-valued signals, can extract meaningful information that aids in precision farming, crop yield estimation, and sustainable land management practices (2) .

1.2.2 Magnetic Resonance Imaging for Medicine

Magnetic Resonance Imaging (MRI) is another domain where complex-valued signals play a pivotal role. MRI technology relies on complex-valued data to generate detailed images of internal body structures. The challenge in MRI is the reconstruction of high-quality images from the raw data, which is inherently complex-valued.

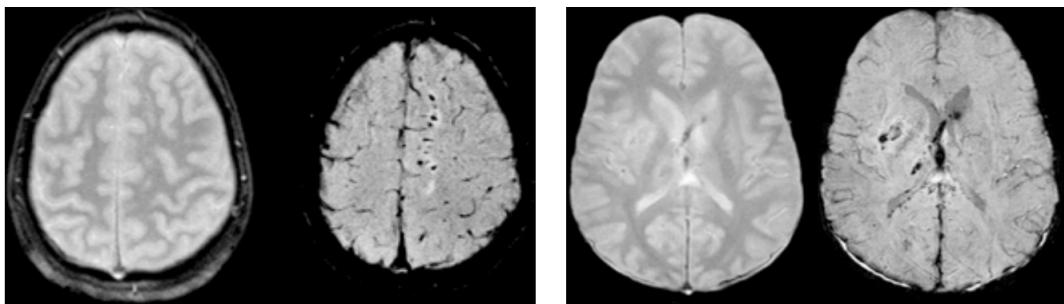


Figure 1.2: Condition: Axonal Injury vs Hemorrhage. For each condition, the left image shows typical MRI (without phase information) and right shows SW-MRI (which incorporates both magnitude and phase information). Note the condition evident in SW-MRI expressed by black patches.

Typical MRIs often discard the phase information. Using Susceptibility Weighted MR phase information is also taken into account. Phase information in such images is influenced by the magnetic properties of the tissues being imaged. This means that different tissues or substances within the body, such as veins, hemorrhage areas, or calcifications, will exhibit distinct signals, sometimes appearing as anti-phase signals compared to their surroundings (see Fig. 1.2). Anti-phase signals occur when the phase of the signal is inverted, indicating variations in susceptibility and often correlating with pathological changes. The MRI images shown compare conditions of axonal injury and hemorrhage, highlighting how phase information can help differentiate between different types of brain injuries and conditions, providing valuable diagnostic information as shown in Fig. 1.2.

CVNNs may offer a sophisticated approach to MRI image reconstruction. By processing the complex-valued MRI data, CVNNs can enhance the clarity and accuracy of the images, leading to better diagnosis and treatment planning in healthcare. The ability of CVNNs to preserve and utilize the phase information in MRI data is particularly beneficial, as it can lead to finer details and contrasts in the reconstructed images.

1.2.3 Image Transforms: Fourier Transform

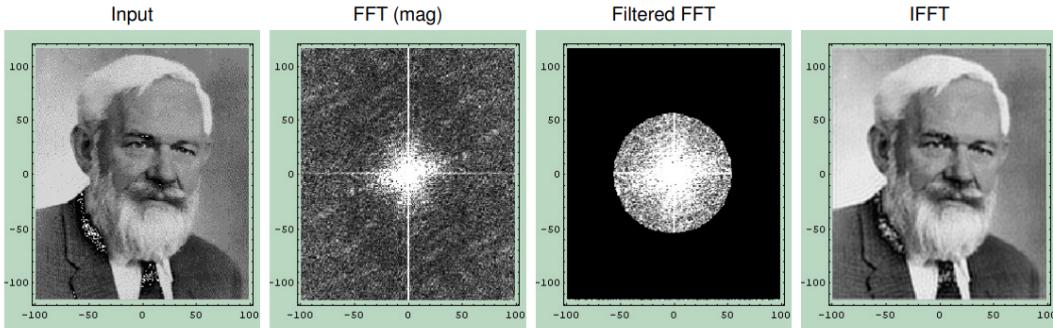


Figure 1.3: Process of Image Restoration using Fourier transform. As natural images have predominantly low frequency (central part of Fourier spectrum) and noise have predominantly high frequency, a low pass filtering in frequency domain removes substantial noise.

Figure 1.3 shows an application where solving in Fourier transformed domain is easier as compared to that in raw or image domain. Starting with the original grayscale image of a noisy portrait image, the process moves to a Fast Fourier Transform (FFT), which translates the image into its frequency domain, revealing the distribution of frequencies as varying intensities in the FFT magnitude representation. A circular low-pass filter is then applied to the FFT to isolate and remove high-frequency noise, which is typically spread outside the central region in the frequency domain. Finally, the Inverse FFT (IFFT) reconstitutes the filtered image back to its spatial domain, resulting in a restored image that ideally has less noise and clearer features than the original.

In applications such as audio processing, telecommunications, and signal filtering, the Fourier Transform's complex-valued output is essential. CVNNs, with their capacity to handle complex-valued data, can effectively process these spectra. This ability is crucial for tasks like noise reduction, signal enhancement, and feature extraction in various signal processing applications.

1.3 Advantages of using complex valued NNs

1.3.1 Orthogonal decision boundaries:

Remark 1: (a) A single complex-valued neuron with n -inputs is equivalent to two real-valued neurons with $2n$ -inputs which have a restriction on a set of weight parameters. (3)

(b) The decision boundary of a single complex-valued neuron consists of two hyper-surfaces which intersect orthogonally.(1)

Proof for the case of a single input single neuron (Remark 1(a)).

Proof. Consider a single neuron with a complex weight $W = w_1 + iw_2$ and a complex input $a + ib$. The output is a complex number $c + id$. The matrix form of the transformation is represented as:

$$\begin{pmatrix} w_1 & -w_2 \\ w_2 & w_1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} c \\ d \end{pmatrix}$$

This can be interpreted as the multiplication of a complex number W with a complex input $(a + ib)$, resulting in a complex output $(c + id)$.

Let M be the matrix representing the transformation:

$$M = \begin{pmatrix} w_1 & -w_2 \\ w_2 & w_1 \end{pmatrix}$$

The magnitude of the complex weight $|W|$ can be represented as $\sqrt{w_1^2 + w_2^2}$. Therefore, we can normalize the matrix M to get:

$$M = |W| \begin{pmatrix} \frac{w_1}{\sqrt{w_1^2 + w_2^2}} & \frac{-w_2}{\sqrt{w_1^2 + w_2^2}} \\ \frac{w_2}{\sqrt{w_1^2 + w_2^2}} & \frac{w_1}{\sqrt{w_1^2 + w_2^2}} \end{pmatrix}$$

This normalized matrix can be further expressed in terms of the magnitude $|W|$ and the arguments (or phase) of W , denoted by $\arg(W)$:

$$M = |W| \begin{pmatrix} \cos(\arg(W)) & -\sin(\arg(W)) \\ \sin(\arg(W)) & \cos(\arg(W)) \end{pmatrix}$$

Finally, we illustrate this with the transformation of a vector $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ in the complex plane, which is rotated by the matrix M to a new position while preserving its

magnitude. This shows that the decision boundary created by a single neuron with complex weights is orthogonal.

The complex number $c + id$, is:

$$c = aw_1 - bw_2,$$

$$d = aw_2 + bw_1.$$

Let activation function (Step function) ϕ applied to the output can be defined as:

$$\phi(x) = \begin{cases} 1 & \text{if } x > c_1, \\ 0 & \text{if } x \leq c_1, \end{cases}$$

where c_1 is a threshold value.

The decision boundary in the input space can then be determined by setting the real part of the neuron's output equal to the threshold c_1 :

$$aw_1 - bw_2 = c_1.$$

Solving for b in terms of a gives the equation of the decision boundary:

$$b = \frac{aw_1 - c_1}{w_2}.$$

Given $w_1, w_2 \neq 0$, the slope of the decision boundary is $-\frac{w_1}{w_2}$, which is orthogonal to the direction of the weight vector in the complex plane. \square

Proof for the Case of a n-input single neuron (Remark 1(b))

Proof. Consider a single complex-valued neuron with n complex inputs $z_j = x_j + iy_j$ for $j = 1, \dots, n$, where i is the imaginary unit, and each z_j corresponds to a real part x_j and an imaginary part y_j . The neuron's weight vector is given by complex numbers $w_j = w_j^R + iw_j^I$, where w_j^R and w_j^I are the real and imaginary parts of the weight associated with input z_j . The output of the neuron before activation is a complex number represented by the dot product of the inputs and weights.

The real part of the output is given by:

$$\operatorname{Re} \left(\sum_{j=1}^n w_j z_j \right) = \sum_{j=1}^n (w_j^R x_j - w_j^I y_j),$$

and the imaginary part of the output is:

$$\text{Im} \left(\sum_{j=1}^n w_j z_j \right) = \sum_{j=1}^n (w_j^R y_j + w_j^I x_j).$$

The decision boundary is then determined by the set of points for which the real part of the output equals a threshold c_1 , and the imaginary part equals another threshold c_2 . These conditions can be expressed as two linear equations representing two hypersurfaces in the $2n$ -dimensional real space of the inputs $(x_1, \dots, x_n, y_1, \dots, y_n)$:

$$\sum_{j=1}^n (w_j^R x_j - w_j^I y_j) = c_1, \quad \text{and} \quad \sum_{j=1}^n (w_j^R y_j + w_j^I x_j) = c_2.$$

To show that these hypersurfaces intersect orthogonally, consider the gradients of the two functions defining the decision boundaries. The gradients are normal to the hypersurfaces and are given by:

$$\nabla_1 = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_p} \\ \frac{\partial f}{\partial y_1} \\ \vdots \\ \frac{\partial f}{\partial y_n} \end{pmatrix} = \begin{pmatrix} w_1^R \\ \vdots \\ w_n^R \\ -w_1^I \\ \vdots \\ -w_n^I \end{pmatrix}, \quad \nabla_2 = \begin{pmatrix} w_1^I \\ \vdots \\ w_n^I \\ w_1^R \\ \vdots \\ w_n^R \end{pmatrix}.$$

The orthogonality of the hypersurfaces can be verified by showing that the dot product of the gradients is zero:

$$\langle \nabla_1, \nabla_2 \rangle = \sum_{j=1}^n w_j^R w_j^I - w_j^I w_j^R = 0.$$

This dot product is zero, implying that the gradients, and thus the decision boundaries, are orthogonal. Therefore, a single complex-valued neuron with n inputs creates decision boundaries that consist of two orthogonal hypersurfaces in the real \mathbb{R}^{2n} .

1.3.2 Complex Activation Functions

Real-valued non-linear activations are incapable of maintaining the magnitude and phase information of complex-valued inputs. This inability highlights the need for complex-valued activations in neural networks, particularly to preserve the intricate relationships embedded in complex data's magnitude and phase(1). Such complex activations are essential for accurately handling and interpreting the characteristics of complex-valued inputs.

1.3.3 Robustness to Noise

Complex-valued neural networks (CV-CNNs) are robust to noise and data distortions, a key advantage in processing real-world data. Their ability to handle complex representations allows them to discern and filter out irrelevant fluctuations in the data more effectively(2) . This feature is particularly valuable in applications where signal integrity is crucial, such as in medical imaging or telecommunications. By maintaining the integrity of the underlying data structure, CV-CNNs can deliver more accurate and reliable outputs, even in challenging noisy environments.

Chapter 2

Magnetic Resonance Imaging and Problem Motivation

2.1 Introduction to MRI and its Importance in Medical Imaging

Magnetic Resonance Imaging (MRI) is one of the most powerful and non-invasive imaging modalities used in modern medicine. It provides detailed anatomical and functional information about soft tissues, without exposing patients to ionizing radiation. Unlike X-rays or CT scans that are based on density differences, MRI works by detecting the nuclear magnetic resonance (NMR) signals emitted by hydrogen nuclei in the body when placed in a strong magnetic field and excited by radiofrequency pulses.

The unique advantage of MRI lies in its ability to generate multi-contrast images, enabling the visualization of diverse tissue types with high spatial resolution. It is particularly valuable in neuroimaging, musculoskeletal imaging, cardiovascular studies, and oncology. Applications include brain tumor diagnosis, ligament injury detection, cardiac perfusion imaging, and even early identification of neurodegenerative diseases.

2.1.1 Basic Working Principle of MRI

MRI scanners function by aligning hydrogen protons in the body using a strong magnetic field (typically 1.5 to 3 Tesla in clinical machines). A radiofrequency (RF) pulse is then applied to disturb this alignment. When the RF pulse is turned off, the protons return to their equilibrium state, emitting energy signals in the process.



Figure 2.1: Overview of the MRI Acquisition and Reconstruction Process: Left – MRI scanning procedure involving patient and radiologist; Right – Resulting MRI image used for medical diagnosis

These emitted signals are detected by receiver coils and encoded in the frequency domain—known as *k-space*—using gradient magnetic fields.

An inverse Fourier transform is then applied to convert the k-space data into spatial images. The raw data collected is inherently *complex-valued*, containing both magnitude and phase components, which are essential for accurate image reconstruction.

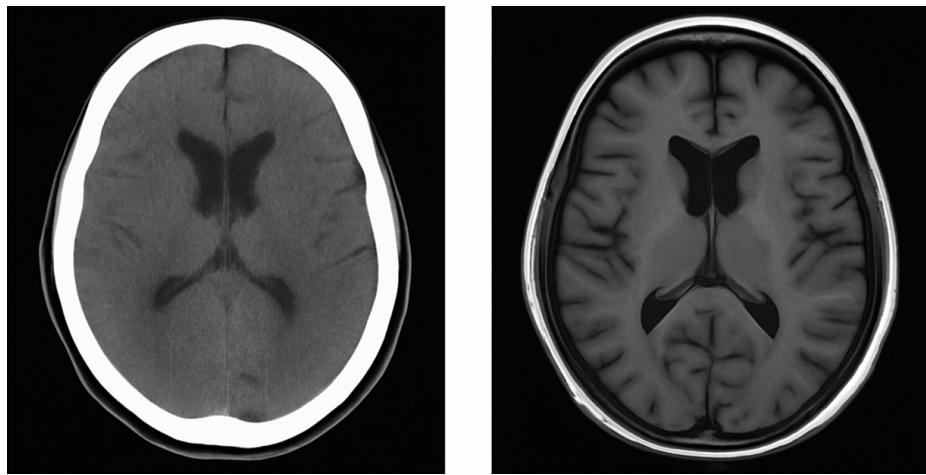


Figure 2.2: Brain Imaging Modality Comparison: Left – CT Scan highlighting dense structures like bone; Right – MRI Scan offering superior soft tissue contrast.

2.1.2 Key Advantages of MRI

- **Non-ionizing Radiation:** Safe for repeated scanning, particularly useful in pediatric and neurological cases.
- **High Soft-Tissue Contrast:** Superior to CT and ultrasound in many soft-tissue applications.
- **Multiplanar Imaging:** MRI enables data acquisition in axial, coronal, sagittal, and oblique planes without repositioning the patient.
- **Functional Imaging Capabilities:** Functional MRI (fMRI), diffusion-weighted imaging (DWI), and spectroscopy expand its diagnostic utility.

2.2 Justification: Clinical and Operational Motivations for Accelerated MRI

While Magnetic Resonance Imaging (MRI) is renowned for its diagnostic precision and safety, it remains one of the slowest imaging modalities in clinical practice. Standard MRI scans can take anywhere from several minutes to over an hour, depending on the anatomical region and protocol. This leads to several critical limitations in real-world settings, both for patients and healthcare systems.

2.2.1 Patient Discomfort and Compliance Issues

Long scan durations cause significant discomfort, especially for patients in pain, trauma cases, pediatric patients, and individuals with neurological disorders or claustrophobia. Maintaining a still posture for extended periods is challenging, often leading to patient motion, which degrades image quality.

- **Motion artifacts** introduced by patient movement are a common issue in long scans, causing blurring, ghosting, or loss of critical diagnostic information.
- **Reduced compliance** can require rescheduling or repeating scans, further delaying diagnosis and treatment.

2.2.2 Operational Constraints and Healthcare Burden

MRI scan rooms are expensive to operate and limited in availability. Long acquisition times directly translate to lower patient throughput, making it difficult to meet clinical demand in high-volume hospitals or under-resourced regions.

- **Lower throughput** increases wait times and reduces the number of patients that can be scanned per day.
- **Increased costs per scan** make MRI less feasible for routine follow-ups or screening in many healthcare systems.
- **Limited accessibility** further exacerbates diagnostic inequality, especially in rural or economically challenged regions.

2.2.3 Need for Technological Intervention

These challenges have spurred a global research effort toward developing methods that can significantly reduce MRI scan times without compromising image quality. A major strategy involves acquiring fewer samples in the frequency domain (k -space) and reconstructing the full-resolution image using advanced computational models.

This project is motivated by the need to develop efficient, robust, and generalizable MRI reconstruction frameworks that leverage the power of **complex-valued deep learning**. By learning from both the magnitude and phase components of MR signals, these models can deliver high-fidelity reconstructions from undersampled data, paving the way for faster, more accessible, and patient-friendly MRI workflows.

2.3 The MRI Signal Acquisition and Reconstruction Workflow

Magnetic Resonance Imaging (MRI) is grounded in the principles of nuclear magnetic resonance (NMR) and involves a sophisticated pipeline to convert physical phenomena into high-resolution diagnostic images. This section outlines the step-by-step workflow of MRI acquisition, from tissue excitation to image reconstruction.

2.3.1 Magnetization and Alignment

The MRI process begins with the patient being placed inside a strong, uniform magnetic field B_0 . This field causes the hydrogen protons in the body — primarily found in water and fat — to align with or against the field direction.

- Most protons align *parallel* (low-energy state) or *antiparallel* (high-energy state) to B_0 .
- This distribution creates a net magnetization vector \mathbf{M} along the direction of B_0 , typically the longitudinal (z) axis.

2.3.2 Excitation Using Radiofrequency (RF) Pulses

To generate a detectable MRI signal, a radiofrequency (RF) pulse is applied perpendicular to the main magnetic field.

- The RF pulse perturbs the equilibrium state by tipping the net magnetization vector away from the z-axis into the transverse (x-y) plane.
- This process excites the protons to a higher energy state, initiating precession about the main magnetic field.
- The angle to which the vector tips is called the *flip angle*, and is controlled by the duration and amplitude of the RF pulse.

2.3.3 Signal Detection and Relaxation

After the RF pulse is switched off, the protons begin to return to their equilibrium state in a process known as *relaxation*.

- Two main relaxation processes occur: $T1$ (longitudinal) and $T2$ (transverse).
- As protons relax, they emit electromagnetic signals at the Larmor frequency, which are picked up by receiver coils placed around the body.
- These complex-valued signals encode the spatial distribution of the proton spins and form the raw data of MRI.

2.3.4 K-Space Sampling and Frequency Encoding

The detected signals are mapped into a domain called *k-space*, which represents the spatial frequency components of the image.

- Gradients in magnetic fields are applied to spatially encode the position of each signal in the x, y (and sometimes z) directions.
- K-space is a matrix where each point captures information about spatial frequencies — not the image itself.
- The central region of k-space corresponds to low-frequency information (contrast), while edges encode high-frequency details (edges, sharpness).

2.3.5 Image Reconstruction via Fourier Transform

To convert k-space data into a meaningful spatial image, reconstruction algorithms are used.

- The standard method is the 2D or 3D Inverse Fast Fourier Transform (IFFT), which transforms frequency components into spatial intensity maps.
- The output is a grayscale image representing the tissue characteristics (T1 or T2 weighting), based on the pulse sequence used.
- Advanced methods, including compressed sensing and deep learning, can accelerate and enhance reconstruction, especially when data is undersampled.

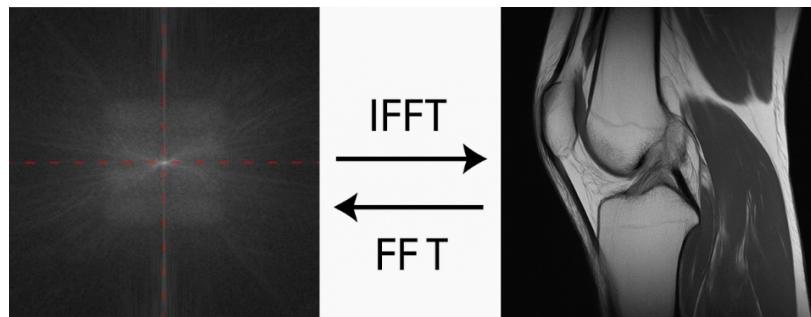


Figure 2.3: Forward and Inverse Fourier Transform in MRI: From k-space to Image Domain and Back

This workflow sets the foundation for understanding the challenges of accelerated MRI, where acquiring full k-space data is often infeasible due to scan time limitations. In such cases, data-driven models — such as those based on Complex-Valued AI — offer promising solutions for reconstructing high-fidelity images from undersampled data.

2.4 K-Space Domain and Undersampling

Following the acquisition of MRI signals, the raw data is stored in a domain known as *k-space*. Understanding this domain is crucial for appreciating the challenges of accelerated MRI and the motivation behind modern reconstruction techniques.

2.4.1 Understanding K-Space

K-space is a frequency-encoded domain that does not visually resemble the final MRI image but contains all the spatial frequency components needed to reconstruct it. Each point in k-space corresponds to a specific combination of phase and frequency information acquired via gradient manipulation during scanning.

- The center of k-space captures low-frequency components — which encode image contrast and gross structures.
- The periphery holds high-frequency information — which contributes to fine details and edge sharpness.
- A complete, densely sampled k-space grid yields high-resolution images via inverse Fourier transformation.

2.4.2 The Challenge of Full K-Space Acquisition

Acquiring every point in k-space can be time-consuming, especially in high-resolution or multi-slice MRI. Longer acquisition times not only increase patient discomfort but also introduce artifacts due to patient motion.

Undersampling is thus introduced to reduce scan time, where only a subset of k-space is acquired — often following structured or random sampling patterns.

- Undersampling accelerates imaging but results in aliasing artifacts and loss of structural fidelity in the reconstructed image.

- Traditional reconstruction techniques like parallel imaging or compressed sensing attempt to restore image quality using prior assumptions.

The challenge, therefore, is to recover high-fidelity images from undersampled k-space data. This motivates the use of more expressive and adaptive reconstruction models, particularly those grounded in deep learning and complex-valued representations, which are discussed next.

2.5 Relevance of Complex-Valued AI in MRI Reconstruction

Magnetic Resonance Imaging inherently produces complex-valued data. Both the raw k-space measurements and intermediate signal representations contain real and imaginary components that together capture essential magnitude and phase information.

2.5.1 Limitations of Real-Valued Models

Most classical deep learning architectures are designed for real-valued inputs and parameters. When applied to complex MRI data:

- They often process only the magnitude, discarding valuable phase information.
- Alternatively, they separate real and imaginary parts, treating them independently—losing crucial interdependencies.

These simplifications can lead to suboptimal reconstruction performance, especially under aggressive undersampling where subtle signal correlations are vital.

2.5.2 Why Complex-Valued AI is Well-Suited

Complex-valued AI (CVAI) models—built using complex-valued convolutions, activations, and weights—retain and exploit both the amplitude and phase characteristics of the input signal. For MRI reconstruction, this leads to several advantages:

- **Faithful representation of signal structure:** MRI physics is inherently complex-valued; modeling it natively reduces domain mismatch.

- **Improved expressiveness:** CV models can encode rotational symmetries and phase-coupled features naturally.
- **Robustness to undersampling:** Richer signal modeling aids in recovering information lost due to partial k-space acquisition.

2.5.3 Integrating CVAI into the Reconstruction Pipeline

Given the preprocessing tools and modular pipeline developed in this study—including data loaders, masking functions, and a structured reconstruction framework—our work lays the groundwork for integrating state-of-the-art CVAI methods. In particular, architectures that perform image-space reconstruction using complex-valued networks (e.g., complex UNet or ReconFormer-inspired transformer backbones) are now easier to prototype and evaluate.

This alignment between the mathematical nature of MRI data and the modeling capability of CVAI forms a key motivation for the project and serves as the foundation for the proposed deep learning extensions in subsequent stages.

2.6 Project Objectives and End Goals

This project is driven by the clinical and computational need to accelerate MRI acquisition while preserving image fidelity. With rising demands on diagnostic throughput and patient comfort, the integration of advanced reconstruction strategies—particularly those leveraging complex-valued deep learning—offers a promising direction.

The core objectives of this study are as follows:

1. **Accelerate MRI Scan Acquisition:** Reduce scan time through effective undersampling strategies in k-space while ensuring diagnostic usability of the reconstructed images.
2. **Leverage Complex-Valued AI Architectures:** Develop and integrate deep learning models capable of processing complex-valued MRI data directly, utilizing both magnitude and phase information during reconstruction.

3. **Preserve Image Quality Under Undersampling:** Maintain or improve image quality despite significant undersampling, as quantified by standard metrics such as Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM).
4. **Enable Real-Time Inference:** Build a scalable inference pipeline that supports near real-time MRI reconstruction, making the approach practical for clinical deployment.
5. **Compare Across Methodologies:** Evaluate the proposed CVAI-based approach against baseline models (e.g., zero-filled reconstruction, real-valued UNet) to demonstrate relative performance improvements.
6. **Bridge Research and Deployment:** Create an open-source-compatible pipeline with modularity for training, inference, and benchmarking, laying the groundwork for future clinical collaborations or integration into platforms such as fastMRI.

Together, these objectives seek not only to advance the performance of MRI reconstruction but also to make complex-valued AI more accessible and applicable to real-world medical imaging problems.

Chapter 3

Background and Literature Review

3.1 MRI Sampling Process Recall

Magnetic Resonance Imaging (MRI) is a non-invasive imaging modality that relies on the response of hydrogen protons under a strong magnetic field and controlled radiofrequency (RF) excitations. While the physical process of signal generation occurs in the spatial domain, the raw measurements are acquired in the frequency domain, more precisely referred to as *k-space*.

The acquisition of k-space data is not performed in a linear raster fashion. Instead, it involves sequential excitation and spatial encoding using magnetic field gradients. Two types of gradients are primarily responsible for k-space traversal: **frequency encoding gradients** and **phase encoding gradients**, which operate along orthogonal directions.

- **Frequency Encoding:** Applied along the readout direction (typically the x-axis), this gradient allows the scanner to resolve spatial locations by varying the precession frequency of proton spins.
- **Phase Encoding:** Applied orthogonally (typically the y-axis), this gradient modifies the phase of the proton spins, encoding spatial information across successive repetitions.

Each excitation pulse acquires a single line (or trajectory) in k-space, defined by the specific combination of phase and frequency encoding gradients used during that acquisition. The complete set of these encoded signals forms the k-space matrix, which

is then transformed into the final image using the Inverse Fast Fourier Transform (IFFT).

In the subsequent sections, we delve deeper into the encoding mechanisms (Section 3.2), k-space sampling strategies and undersampling, and reconstruction techniques that form the basis of modern MRI acceleration research.

3.2 K-Space Sampling Theory and Masking Techniques

MRI data acquisition in k-space is inherently time-consuming, as each line of k-space typically requires a separate RF excitation and readout. The desire to reduce scan time—especially in clinical settings—has led to the development of undersampling strategies that acquire only a subset of the full k-space, with the goal of reconstructing the full image from incomplete data.

3.2.1 Full vs Undersampled K-Space

In conventional MRI, k-space is sampled densely and uniformly across both phase and frequency encoding directions. This ensures adherence to the Nyquist-Shannon sampling criterion, allowing for alias-free image reconstruction via the inverse Fourier transform. However, full sampling comes at the cost of longer acquisition times.

Undersampling reduces the number of acquired phase-encoded lines, which shortens scan duration but violates the Nyquist criterion, leading to artifacts such as aliasing and noise amplification in the reconstructed image.

3.2.2 Random and Variable-Density Sampling Masks

To effectively undersample k-space, specialized binary masks are employed to indicate which frequency components are acquired. These masks fall into different categories:

- **Uniform Undersampling:** Skips lines in a regular pattern (e.g., every 2nd or 4th line), but tends to produce coherent aliasing artifacts.
- **Random Undersampling:** Randomly samples k-space lines, leading to incoherent aliasing that is more amenable to advanced reconstruction algorithms like compressed sensing and deep learning.

- **Variable-Density Sampling:** Concentrates sampling density in the center of k-space (low spatial frequencies), which contains most of the image's energy, and sparsifies sampling in the outer regions (high frequencies). This approach balances scan efficiency and reconstruction fidelity.

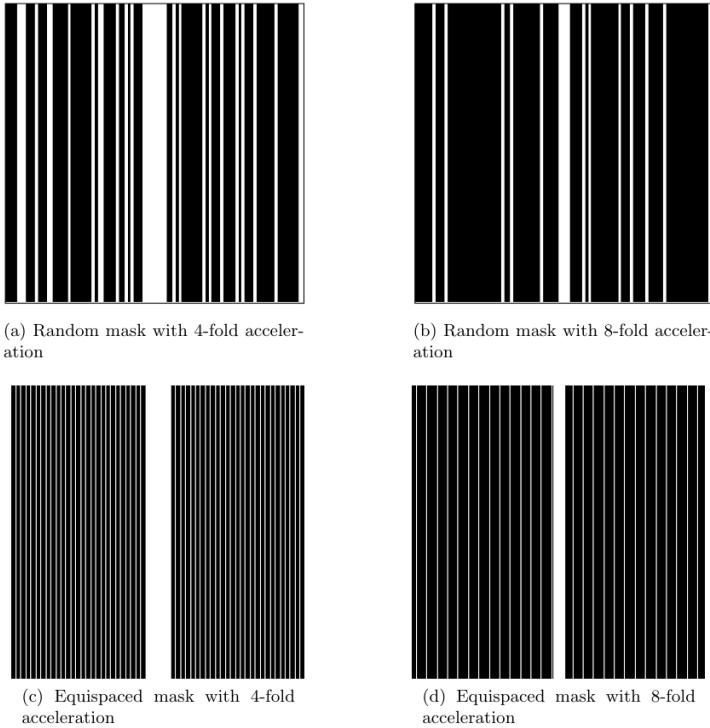


Figure 3.1: Examples of undersampled k-space trajectories(4)

3.2.3 Acceleration Factor and Center Fraction

Two key parameters define an undersampling mask:

- **Acceleration Factor (AF):** Defined as the ratio of the number of k-space lines in a fully sampled acquisition to that in an undersampled one. For example, an AF of 4 implies that only 25% of the original k-space lines are acquired.
- **Center Fraction (CF):** Represents the proportion of low-frequency (central) lines that are always sampled. A typical CF might range from 0.04 to 0.1, depending on image content and reconstruction strategy.

These parameters control the trade-off between acquisition speed and image quality. In the context of this study, we employ fixed-ratio Cartesian sampling masks with random variable-density designs—similar to those proposed in the *fastMRI* benchmark dataset—to simulate accelerated acquisitions.

3.2.4 Implications for Reconstruction

While undersampling allows for faster scans, it necessitates sophisticated reconstruction algorithms to recover the missing data while minimizing artifacts. The upcoming sections discuss the role of Fourier theory (Section 3.4), limitations of direct inverse transforms under undersampling (Section 3.5), and how various classical and learning-based strategies have evolved to overcome these challenges (Sections 3.6 and 3.7).

3.3 Inverse Fourier Transform Theory for MRI Reconstruction

The central principle that bridges k-space and the image domain in MRI is the Fourier Transform. MRI data acquisition inherently occurs in the spatial frequency domain (k-space), and the transformation to the spatial domain (i.e., the actual image) is carried out via the inverse Fourier transform (IFFT).

3.3.1 Forward Process: From Object to K-Space

When an object is placed in a magnetic field and excited using RF pulses, the resulting magnetization signals are spatially encoded using gradient fields. The data collected by receiver coils corresponds to samples in the k-space.

Mathematically, the signal $S(k_x, k_y)$ acquired during MRI can be modeled as the two-dimensional Fourier transform of the object $\rho(x, y)$:

$$S(k_x, k_y) = \int \int \rho(x, y) e^{-i2\pi(k_x x + k_y y)} dx dy$$

This process effectively encodes spatial information in terms of frequency components, building up the k-space representation line by line.

3.3.2 Inverse Process: Reconstructing Image via IFFT

Once the k-space is acquired, image reconstruction is performed by applying the inverse two-dimensional Fourier transform:

$$\rho(x, y) = \int \int S(k_x, k_y) e^{i2\pi(k_x x + k_y y)} dk_x dk_y$$

In practice, since we work with discrete samples, a discrete 2D inverse Fourier transform (IFFT) is used to reconstruct the image from sampled k-space data. This transformation reassembles spatial structures like edges, textures, and tissue contrasts from the encoded frequency information.

3.3.3 Implications in the Context of Undersampling

When k-space is fully sampled, IFFT alone suffices to yield high-fidelity reconstructions. However, in the case of undersampled k-space, direct IFFT leads to images with significant artifacts, particularly aliasing due to violation of the Nyquist criterion.

This limitation necessitates more advanced reconstruction techniques that can fill in the missing information or correct artifacts introduced by undersampling. These approaches form the focus of subsequent sections.

3.3.4 Summary and Transition

To summarize:

- MRI data is acquired in the frequency domain.
- Reconstruction into image domain is achieved through 2D IFFT.
- Full-sampled k-space leads to artifact-free reconstruction.
- Undersampling breaks this pipeline, necessitating smarter reconstruction models.

The next section (3.5) explores the aliasing artifacts introduced during naive undersampling, and motivates the need for reconstruction strategies in both k-space and image-space.

3.4 Undersampling and Aliasing in MRI Reconstruction

3.4.1 Need for Accelerated MRI via Undersampling

One of the main goals in modern MRI is to reduce scan times without compromising image quality. This is typically attempted by **undersampling the k-space**, i.e., acquiring fewer frequency-encoded lines than required by the Nyquist criterion. Fewer samples imply faster scans, reduced patient discomfort, and increased throughput.

3.4.2 Aliasing Artifacts due to Naive Undersampling

While undersampling accelerates acquisition, it introduces a major downside: **aliasing artifacts** in the reconstructed images. Since the k-space is underdetermined, directly applying an inverse Fourier transform (IFFT) on undersampled data leads to overlapping signal components in the spatial domain. This is visually observed as ghosting, blurring, or distortion—significantly hampering diagnostic quality.

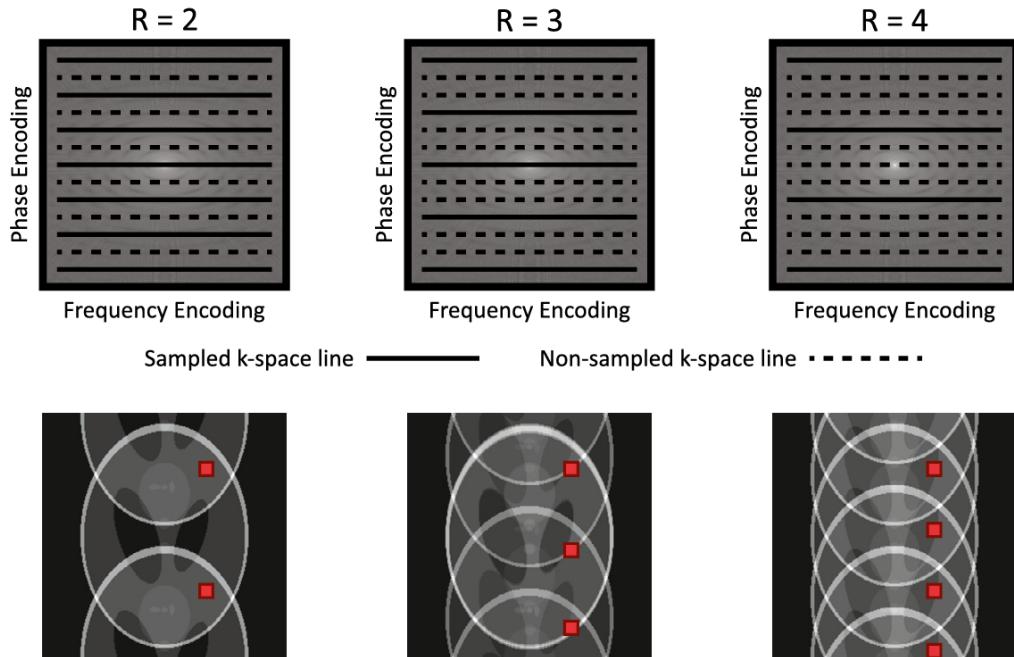


Figure 3.2: Sampling and aliasing patterns for various acceleration factors. (top) Uniform Cartesian sampling schemes for various acceleration factors. (bottom) Aliasing patterns resulting from these undersampling schemes;

3.4.3 Reconstruction Strategies: Two Broad Categories

To mitigate the effects of undersampling, advanced reconstruction algorithms are employed that aim to recover or infer the missing k-space data or learn mappings directly in image space. These solutions are broadly categorized into:

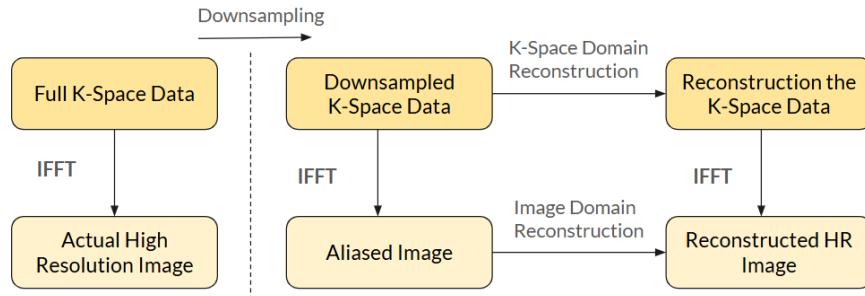


Figure 3.3: Reconstruction Strategies: Two Broad Categories

(a) K-Space-Based Reconstruction Methods

These methods operate directly in the frequency domain (k-space), attempting to interpolate or estimate the missing k-space lines using spatial correlations and calibration data.

Examples include:

- SMASH (Simultaneous Acquisition of Spatial Harmonics)
- AUTO-SMASH
- GRAPPA (Generalized Autocalibrating Partially Parallel Acquisitions)
- GRAPPANet (deep-learning variant of GRAPPA)
- RAKI (Robust Artificial-neural-networks for K-space Interpolation)
- SPARK (Self-supervised Prediction and Reconstruction in K-space)

(b) Image-Space-Based Reconstruction Methods

These techniques operate after transforming the acquired k-space data to the image domain using IFFT, followed by learning or modeling the recovery of alias-free images

using prior knowledge, sparsity, or deep neural networks. This category includes our primary area of focus.

Examples include:

- SENSE (Sensitivity Encoding)
- ESPIRiT (Extended SENSE with eigenvalue maps)
- Compressed Sensing (CS-MRI)
- Deep SENSE
- SENSE with GANs (Generative Adversarial Networks)
- CVNN-based (Complex-Valued Neural Network) reconstructions

3.4.4 Transition to Image-Space-Based Approaches

In this study, we focus on the second family—**image-space-based methods**, with special emphasis on *complex-valued neural networks* that preserve and exploit the intrinsic structure of MRI data for improved reconstruction quality. These models aim to learn optimal inverse mappings from corrupted (aliased) images to their clean counterparts using supervised training paradigms.

3.5 Review of MRI Reconstruction Methods and Advances

3.5.1 Parallel Imaging

In MRI, k-space data are acquired using receiver coils as described by Parallel Imaging. When multiple coils are used in an array, they not only improve signal-to-noise ratio (SNR) but also enable spatial localization of signals due to their localized sensitivity patterns. This spatial encoding property forms the basis of *parallel imaging*, a technique used to reduce scan time by undersampling k-space.

Each coil has a unique complex-valued *coil sensitivity map*, $C_j(r)$, which depends on coil shape, placement, and even patient-specific loading effects. The MRI signal from the j -th coil is given by:

$$S_j(k) = \int_V C_j(r) M_{xy}(r) e^{-2\pi i k \cdot r} dr = \mathcal{F}[C_j(r) M_{xy}(r)] \quad (3.1)$$

This can also be represented in discretized matrix form as:

$$s = BFCx = Ex \quad (3.2)$$

where B is the sampling operator, F the Fourier transform, C the coil sensitivity matrix, and x the spatial image. The encoding matrix E combines all these transformations and is used in many reconstruction techniques.

Undersampling is often applied in the phase encoding direction, characterized by the acceleration factor R . An $R = 2$ acceleration implies skipping every alternate k-space line, reducing acquisition time but introducing aliasing. Theoretical reconstruction limits depend on the number of coils and their spatial encoding diversity.

To reconstruct the full image, data from individual coils are combined using techniques like Roemer's SNR-optimal recombination or simpler root-sum-of-squares (RSS) methods. More advanced algorithms like SENSE (Sensitivity Encoding) and GRAPPA (Generalized Autocalibrating Partial Parallel Acquisition) solve systems of equations using coil sensitivity maps or calibration data.

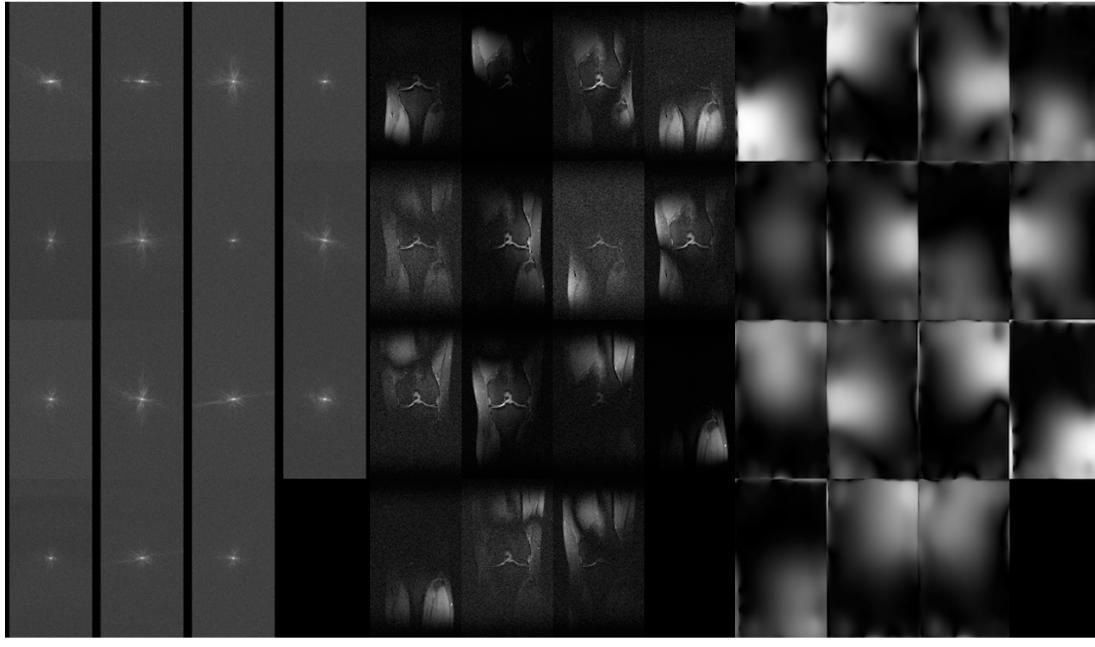


Figure 3.4: Parallel Imaging: Multi-coil MRI representation

Parallel imaging can suffer from reduced SNR, computed as:

$$\text{SNR}_{\text{PI}} = \frac{\text{SNR}_0}{g\sqrt{R}} \quad (3.3)$$

where g is the geometry factor reflecting coil arrangement and aliasing patterns. Techniques like ESPIRiT and Deep SENSE build on these fundamentals with improved robustness and reconstruction fidelity.

3.5.2 SENSE (Sensitivity Encoding)

Concept: SENSE- Sensitivity Encoding (5) is a parallel imaging technique that uses spatial sensitivity variations from multiple receiver coils to reconstruct MR images from undersampled k-space data.

Core Idea: It enables faster scans by reducing the number of k-space lines that need to be acquired, serving as a complement to Fourier encoding.

Reconstruction: SENSE provides a generalized image reconstruction technique applicable to any coil arrangement and k-space sampling trajectory.

Advantage: It allows accelerated imaging without the need to increase k-space traversal speed, thus maintaining image quality.

Theory Overview

SENSE Basics: SENSE combines conventional gradient encoding with coil sensitivity encoding to reconstruct images from undersampled data, minimizing scan times.

Reduction Factor (R): The reduction factor R defines how much the k-space sampling is decreased. Although undersampling introduces aliasing, SENSE exploits coil sensitivity profiles to resolve these artifacts.

Cartesian Sampling and Aliasing

- **Aliased Image Formation:** Each coil produces an aliased image by applying the Discrete Fourier Transform (DFT) to the undersampled data.
- **Unfolding Aliased Pixels:** Overlapping signal contributions from different spatial locations are separated using:
 - Sensitivity matrix \mathbf{S} : Describes spatial sensitivity profiles of receiver coils.

- Unfolding matrix \mathbf{U} : Derived from \mathbf{S} , used to mathematically unmix the aliased pixels.
- Noise covariance matrix \mathbf{C} : Accounts for noise levels and correlations among coils to improve Signal-to-Noise Ratio (SNR).

Steps in SENSE Reconstruction:

1. Generate coil sensitivity maps.
2. Acquire undersampled k-space data.
3. Reconstruct aliased images for each coil.
4. Unfold the aliased images via matrix inversion techniques.

Sensitivity Map Estimation

Accurate sensitivity maps are vital for successful SENSE reconstruction. These maps reflect each coil's spatial sensitivity and are obtained as follows:

1. **Acquisition:** Full Field-of-View (FOV) images are acquired using individual coils or a body coil reference.
2. **Initial Estimation:** Sensitivity maps are calculated by dividing each coil's image by the root-sum-of-squares (RSS) of all coils:

$$S_i(x, y) = \frac{I_i(x, y)}{\sqrt{\sum_{j=1}^N |I_j(x, y)|^2}}$$

3. **Polynomial Fitting:** A 2D polynomial is locally fit to refine sensitivity values, suppressing noise while preserving spatial accuracy.
4. **Weighting Function:** The weighting function incorporates both spatial proximity and pixel reliability (e.g., based on noise or tissue presence), commonly modeled using a Gaussian:

$$w(x, y) = \exp\left(-\frac{(x - x_0)^2 + (y - y_0)^2}{2\sigma^2}\right)$$

5. Edge Handling: In regions lacking reliable data, extrapolation is used to extend sensitivity values, ensuring continuity and avoiding reconstruction artifacts.

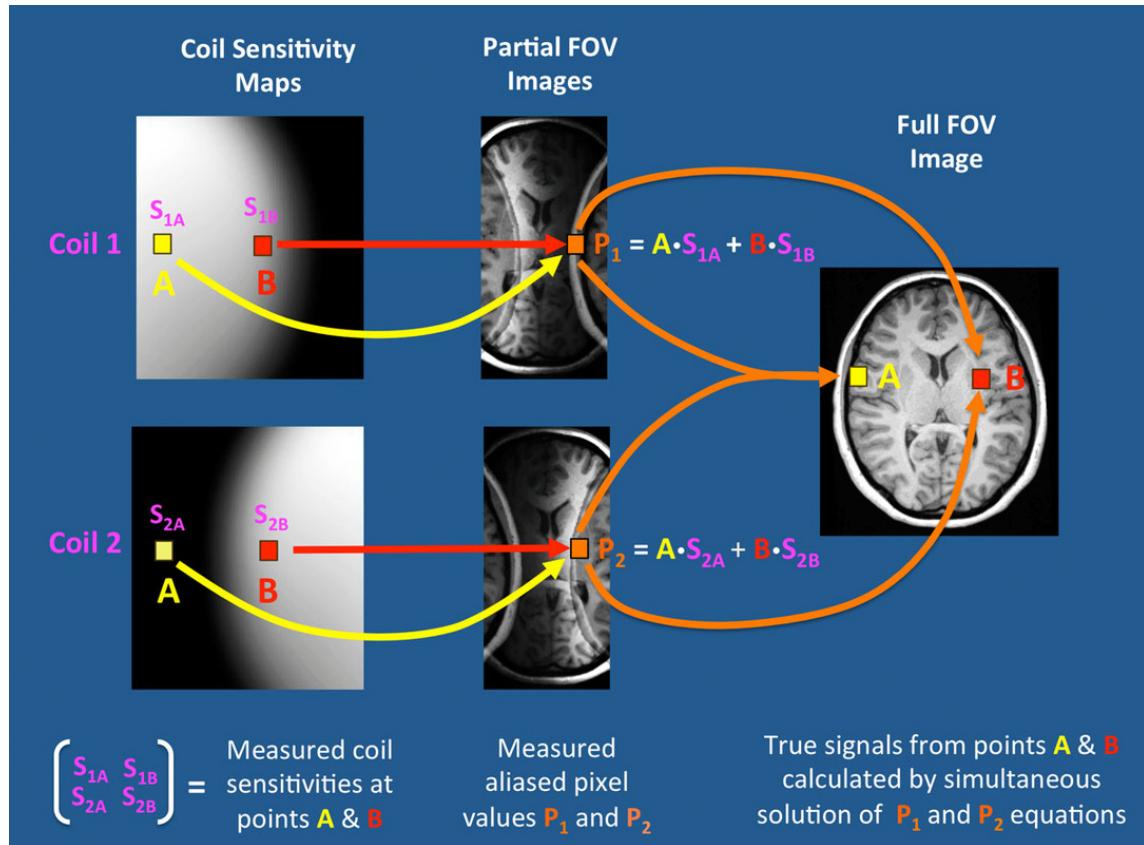


Figure 3.5: SENSE: Unfolding of aliased pixels(simplified)

This sensitivity map refinement enhances the accuracy and robustness of the SENSE reconstruction process, improving image quality and preserving diagnostic detail.

3.5.3 ESPIRiT — An Eigenvalue Approach to Autocalibrating Parallel MRI

Concept: ESPIRiT(6) bridges the SENSE and GRAPPA frameworks by offering an eigenvalue-based method for robust, auto-calibrated coil sensitivity map estimation.

Core Idea: Improved and accurate estimation of coil sensitivity maps from autocalibration data, followed by SENSE-like reconstruction using these high-quality maps.

Advantage: Combines the precise unfolding of SENSE with the robustness and flexibility of GRAPPA, especially in dealing with inconsistencies or imperfect data.

Theory Overview

ESPIRiT Fundamentals: ESPIRiT uses autocalibration signals (ACS) to construct a calibration matrix, where the null space of this matrix encodes k-space correlations between coils. Sensitivity maps are extracted as the principal eigenvectors of a reconstruction operator derived from this null space.

Calibration Matrix and Eigenvalue Decomposition: From the ACS region, a calibration matrix is constructed. A reconstruction operator is then computed, whose eigenvectors correspond to the spatial coil sensitivities. Typically, eigenvalues close to 1 signify the true signal space, while other eigenvalues indicate inconsistencies or noise within the data.

Improved Sensitivity Map Estimation: Unlike traditional methods, ESPIRiT does not require a separate reference scan. Instead, it generates spatially smooth sensitivity maps directly from the acquired data. The use of multiple eigenvectors allows ESPIRiT to model complex sensitivity structures more accurately and account for system imperfections.

Reconstruction with ESPIRiT: Once the sensitivity maps are obtained, image reconstruction proceeds using a SENSE-like approach. The reconstruction problem is solved through a generalized matrix inversion that can incorporate multiple maps when necessary. This methodology yields high-fidelity images and exhibits robustness to noise, motion artifacts, and calibration errors.

Key Benefits: ESPIRiT eliminates the need for external sensitivity calibration scans, offering a more efficient approach. It is also robust against inconsistencies and noise, making it highly reliable in real-world scenarios. By combining the theoretical rigor of SENSE with the empirical strength of GRAPPA in dealing with imperfect data, ESPIRiT provides a powerful and adaptable solution for MRI reconstruction.

In summary, ESPIRiT enhances parallel MRI by improving sensitivity map estimation through an eigenvalue-based formulation and enabling robust, SENSE-style reconstruction from undersampled k-space data.

3.5.4 Compressed Sensing MRI (C-SENSE)

Concept: C-SENSE(Iuga et al.) combines Compressed Sensing (CS) and SENSE with deep learning techniques to enhance MR image reconstruction.

Core Idea: The approach integrates sparsity constraints and data consistency by employing iterative algorithms, which are further enhanced by neural networks.

Advantage: This method enables high-quality reconstruction from heavily undersampled, pseudorandomly acquired k-space data, making it more efficient than traditional methods.

Key Principles: C-SENSE incorporates several key principles. **Pseudorandom Undersampling** replaces the uniform undersampling used in SENSE with incoherent pseudorandom sampling, transforming aliasing into noise-like artifacts rather than structured image distortions. **Sparsity in Transform Domain** exploits the inherent sparsity of MRI signals in domains such as wavelets, enabling accurate reconstruction even from fewer samples. Lastly, **Iterative Reconstruction** combines regularization techniques (such as sparsity priors) with data fidelity through iterative optimization, which is further enhanced by convolutional neural networks (CNNs) to improve the quality of the reconstructed images.

Outcome: C-SENSE achieves faster scan times and improved image quality by suppressing noise-like artifacts during reconstruction. CNNs guide the optimization by learning from data priors, effectively balancing artifact removal and detail preservation.

3.5.5 Deep SENSE

Concept: Deep SENSE(7) employs deep neural networks to either estimate coil sensitivity maps or directly reconstruct MR images from undersampled k-space data.

Core Idea: The approach learns a nonlinear mapping from undersampled data to high-quality images, addressing the limitations of SENSE in handling noise and aliasing.

Advantage: It offers higher reconstruction fidelity and reduced noise/artifacts compared to traditional methods.

Highlights: This technique is trained on brain imaging datasets, ensuring robustness to typical brain anatomy. It either enhances coil map estimation or replaces the traditional unfolding step with learned end-to-end mappings. Moreover, the method

generalizes well across varying undersampling patterns and noise levels, making it highly adaptable to different imaging scenarios.

Outcome: DeepSENSE outperforms classical SENSE in reconstruction quality and stability, especially under aggressive undersampling regimes.

3.5.6 SENSE with Generative Adversarial Networks (GANs)

Concept: SENSE with Generative Adversarial Networks (GANs)(8) utilizes Generative Adversarial Networks (GANs) to post-process SENSE-reconstructed images, effectively suppressing g-factor artifacts.

Core Idea: The GANs are trained to learn and distinguish aliasing and noise patterns that arise from high reduction factors in SENSE, allowing for more accurate image reconstruction.

Advantage: This approach enables higher acceleration in MRI scans while improving image quality, mitigating the traditional limitations of SENSE, particularly with regard to noise and aliasing.

Details: High reduction factors in SENSE often lead to g-factor artifacts, which are structured noise resulting from the ill-conditioned nature of matrix inversion in the reconstruction process. To address this, a GAN is trained to remove these artifacts, refining the output of SENSE through an adversarial framework. The generator component of the GAN works to improve image realism, while the discriminator enforces perceptual similarity to the ground truth, ensuring high-quality, artifact-free images.

Outcome: Enhances image fidelity without requiring changes to the acquisition or initial reconstruction pipeline.

3.5.7 CVNN-based Reconstruction (Complex-Valued Neural Networks)

Concept: This approach uses complex-valued neural networks (CVNNs)(9) to fully exploit the native complex structure of MRI data for reconstruction.

Core Idea: CVNNs process both the real and imaginary components of MRI data jointly in learned representations, preserving phase and magnitude relationships to ensure accurate reconstructions.

Advantage: By more accurately modeling k-space data and coil sensitivities, CVNNs lead to improved reconstructions, particularly under undersampled conditions, where traditional methods might struggle.

Highlights: Traditional convolutional neural networks (CNNs) treat the real and imaginary parts of MRI data separately or stack them, which leads to the loss of important interactions between these components. In contrast, CVNNs use complex weights and activations to maintain the structure of MRI signals, allowing for a more effective processing of the data. This method is suitable for both end-to-end reconstructions and sensitivity map estimation, making it a versatile approach for MRI image reconstruction.

Outcome: Demonstrates superior performance in preserving fine anatomical details and reducing artifacts, especially in phase-sensitive applications.

3.6 Chapter Conclusion

In this chapter, we laid the mathematical and conceptual foundation for several classical and modern MRI reconstruction techniques. Starting from fundamental approaches like Fourier-based reconstruction, we explored advanced parallel imaging methods such as SENSE and ESPIRiT, delving into their underlying formulations, assumptions, and reconstruction pipelines. The mathematical baselines established here provide the theoretical groundwork necessary for interpreting and evaluating reconstruction quality.

We also introduced recent innovations that integrate compressed sensing and deep learning with classical frameworks. Methods such as C-SENSE, Deep SENSE, GAN-based SENSE, and CVNN-based reconstructions represent the forefront of MRI research, aiming to achieve higher acceleration with minimal compromise in image quality. While we have outlined their theoretical underpinnings and design motivations, the implementation and comparative performance evaluation of these data-driven techniques remain part of our upcoming work.

Overall, this chapter sets the stage for transitioning from model-based methods to data-driven approaches, bridging theory with practical experimentation in subsequent sections.

Chapter 4

Methodology

4.1 Dataset (fastMRI)

The fastMRI dataset aims to serve as a foundational step toward advancing machine learning-based reconstruction of accelerated MR images. It comprises a large-scale release of raw MRI data, including 8,344 volumes (167,375 slices), and over 20,000 processed DICOM images, representing more than 1.57 million slices from clinical knee and brain examinations.

Before delving into the dataset specifics, a brief overview of MRI acquisition and reconstruction is provided to bridge the understanding between the MRI and machine learning communities.

4.1.1 Data Types

The dataset includes four main data types:

- **Raw multi-coil k-space data:** Complex-valued, unprocessed data collected using multiple coils.
- **Emulated single-coil k-space data:** Derived from multi-coil data to approximate single-coil acquisitions.
- **Ground-truth images:** Reconstructed from fully sampled multi-coil data using the root-sum-of-squares (RSS) method.
- **DICOM images:** Provided for broader scanner diversity, acquired clinically with unknown reconstruction parameters.

4.1.2 Reconstruction Tasks

Two reconstruction tasks are defined:

1. **Single-coil reconstruction:** From undersampled emulated single-coil k-space data.
2. **Multi-coil reconstruction:** From undersampled multi-coil data.

Each task includes official training and validation splits with fully sampled data, as well as test and challenge splits that are undersampled. Ground-truth images are not provided for test and challenge sets.

4.1.3 Knee k-space Data

This subset includes 1,594 multi-coil scans acquired using Siemens 3T and 1.5T MRI systems with a 15-channel knee coil. Two pulse sequences are included:

- Coronal proton-density (PD) weighting without fat suppression (796 scans)
- Coronal proton-density fat-suppressed (PDFS) weighting (798 scans)

All scans follow clinical acquisition protocols with matrix size 320×320 , 0.5mm in-plane resolution, and 3mm slice thickness. Timing parameters (TR and TE) vary across scanners.

4.1.4 Brain k-space Data

This subset includes 6,970 fully sampled brain MRI scans across 11 scanners from multiple sites, including both 1.5T and 3T systems. Scans include:

- T1-weighted
- T2-weighted
- FLAIR
- T1 POST (with contrast)

All data is axial 2D and de-identified via cropping and inspection.

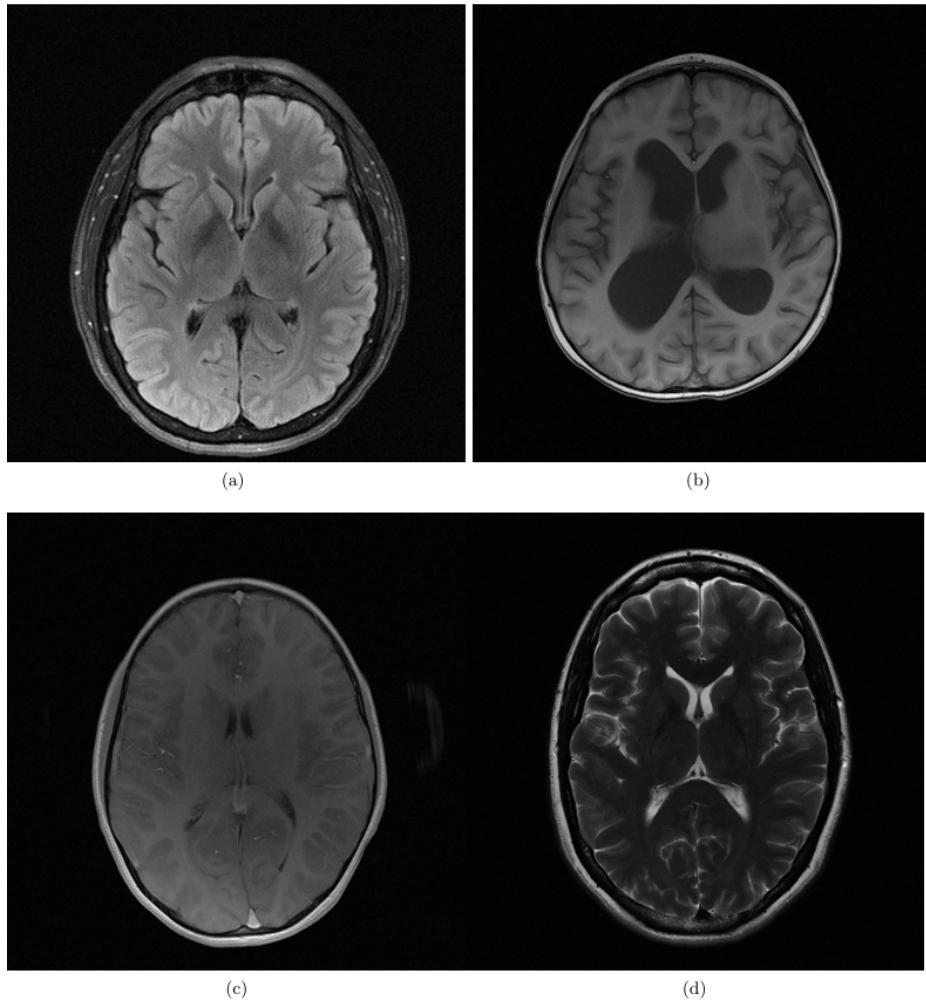


Figure 4.1: Axial brain MRI images with different contrasts: (a) FLAIR, (b) T1 weighted (c) T1 weighted with contrast agent (T1 POST), and (d) T2 weighted.

4.1.5 Knee Emulated Single-Coil Data

Single-coil data is simulated from multi-coil acquisitions using emulated single-coil (ESC) methodology, based on a least-squares fit to the RSS ground truth.

4.1.6 Knee DICOM Data

The dataset also contains 10,000 clinical knee DICOM scans including five pulse sequences per exam:

1. Coronal PD without fat suppression
2. Coronal PD with fat suppression
3. Sagittal PD without fat suppression
4. Sagittal T2 with fat suppression
5. Axial T2 with fat suppression

These are not directly mappable to raw k-space data due to vendor-specific reconstruction steps.

4.1.7 Brain DICOM Data

10,000 additional brain DICOM studies are provided, covering a variety of T1, T2, and FLAIR sequences. Each study is from a unique individual, representing a wide range of pathologies. Identifying features are excluded via cropping and manual validation.

4.1.8 Ground Truth

Ground truth for multi-coil data is computed using the RSS method:

$$m_i = \mathcal{F}^{-1}(y_i)$$

$$m_{\text{rss}} = \sqrt{\sum_{i=1}^{n_c} |m_i|^2}$$

where y_i is the k-space data from the i th coil and n_c is the number of coils.

All ground truth images are cropped to 320×320 to remove oversampling. ESC and RSS reconstructions are both provided for the single-coil dataset.

4.2 BART: Berkeley Advanced Reconstruction Toolbox

The Berkeley Advanced Reconstruction Toolbox (BART) is a free and open-source framework designed for Computational Magnetic Resonance Imaging (MRI). It is intended for research use only and not for diagnostic purposes. BART provides a comprehensive suite of commands for various MRI reconstruction tasks and has

strong support for Fast Fourier Transform (FFT) and Non-Uniform FFT (NUFFT) operations, parallel imaging, and compressed sensing.

This software is highly recognized and regarded in the MRI research community for its versatility and performance. It is widely used to implement and evaluate state-of-the-art MRI reconstruction methods, making it a critical tool for advancing research in this field.

We utilized BART to implement baseline methods and several of the most regarded and important MRI reconstruction techniques, comparing their results to evaluate the effectiveness of various algorithms.

4.2.1 Key Features

- Nonlinear inversion and compressed sensing reconstruction.
- Deep learning integration with TensorFlow and Reconet.
- Synthetic data generation (phantoms, fake k-space).
- Image evaluation tools (NMRSE, ROI statistics, MIP views).
- Wavelet transforms, TGV, and ICTV regularization.
- Fast, lightweight, and suitable for rapid MRI research prototyping.
- Support for parallel imaging algorithms such as SENSE and POCSENSE.
- Scriptable via command-line interface for easy automation.
- Provides tools for efficient data manipulation, transformation, and filtering.

4.2.2 Available Commands

BART includes a wide range of commands for MRI reconstruction and related tasks. Some of the most commonly used commands are:

- **avg** - Average images across multiple slices.
- **bench** - Benchmarking tool for measuring performance.
- **bin** - Convert data to a binary format.

- `bitmask` - Apply bitmask operations for data manipulation.
- `cabs` - Compute the absolute value of complex numbers.
- `calc` - Perform calculations on k-space data.
- `fft` - Compute Fast Fourier Transform (FFT).
- `nufft` - Perform Non-Uniform FFT (NUFFT) operations.
- `reconet` - Deep learning-based MRI reconstruction.
- `pocsense` - Parallel imaging using compressed sensing (POCSENSE).
- `phantom` - Generate synthetic MRI phantoms.
- `noise` - Add noise to images or k-space data.
- `normalize` - Normalize image intensities.
- `wavelet` - Apply wavelet transforms for image denoising.
- `tgv` - Apply Total Generalized Variation (TGV) regularization.
- `mip` - Create Maximum Intensity Projection (MIP) images.
- `morphop` - Morphological operations for image processing.
- `reshape` - Reshape data matrices.

4.2.3 Usage

BART provides a powerful and flexible environment for MRI research prototyping. It supports a wide range of data processing, image reconstruction, and evaluation tasks, making it suitable for both academic and industrial research. The toolbox was designed for ease of use with a command-line interface, making it easy to automate processes and integrate into larger research workflows.

4.3 Overview of the Reconstruction Pipeline

The goal of the reconstruction pipeline is to process multi-coil MRI data and generate high-quality images from undersampled k-space data. This process is essential in accelerating MRI scans by using advanced reconstruction techniques. In this work, two types of MRI datasets are used: one in the `.cfl` format, which is native to BART (Berkeley Advanced Reconstruction Toolbox), and the other in the `.h5` format, which is commonly used in the fastMRI dataset. The pipeline is designed to work seamlessly with both formats, allowing for flexible and efficient image reconstruction.

The primary tools used in this pipeline include the BART toolbox for image reconstruction and sensitivity map estimation, as well as the fastMRI library for handling `.h5` formatted datasets. The core steps in the pipeline include preprocessing of k-space data, applying undersampling masks, and performing image reconstruction using techniques such as SENSE and ESPIRiT. Additionally, various comparison metrics, including visual inspection and difference maps, are used to evaluate the quality of the reconstructions.

The workflow for `.cfl` files and `.h5` files shares common steps, such as coil sensitivity map estimation and the application of undersampling masks. However, the data format-specific preprocessing steps differ, particularly when handling fastMRI’s `.h5` files, which require additional steps for converting the data into a format compatible with BART. Despite these differences, the fundamental reconstruction approach remains the same across both datasets, ensuring consistent and reliable results.

4.3.1 Preprocessing and Input Handling

Preprocessing is a critical step in the reconstruction pipeline, ensuring that input k-space data—regardless of its original format—is compatible with downstream reconstruction tools. This section outlines the input handling procedures for two types of MRI datasets: native BART `.cfl` format and the fastMRI `.h5` format. The goal is to structure the raw multi-coil k-space data into a unified form for subsequent reconstruction steps using BART.

4.3.1.1 Native BART `.cfl` Input (Knee MRI)

For datasets already stored in the `.cfl` format, the BART utility `readcfl` is used to load the raw multi-coil k-space data. These files typically follow a standard four-

dimensional layout, where dimensions represent spatial frequency components and coil channels. After reading the data, the dimensionality is inspected to verify the number of coils and the k-space resolution.

This native format is directly compatible with BART’s reconstruction tools, requiring no additional preprocessing. It facilitates rapid experimentation with various reconstruction algorithms such as SENSE and ESPIRiT, and supports direct visualization and transformation workflows.

4.3.1.2 fastMRI .h5 Input (Brain and Knee MRI)

The fastMRI dataset stores raw k-space volumes in the .h5 format. These volumes are typically four-dimensional, containing multiple slices and coil channels. The file is first accessed using the `h5py` library, from which the k-space volume is extracted.

To prepare the data for use with BART, an individual slice is selected from the 4D volume. Since BART expects a different memory layout, the extracted slice undergoes reformatting using NumPy: a transpose operation is performed to reorder the axes into `[height, width, coils]`, followed by an `expand_dims` operation to add an extra singleton dimension, yielding a shape compatible with BART’s conventions.

Finally, the reformatted slice is written to disk using `writecf1`, making it usable for BART-based reconstruction pipelines. This conversion process ensures interoperability between fastMRI datasets and BART utilities.

4.3.2 Sensitivity Map Estimation

Accurate coil sensitivity estimation is a prerequisite for parallel imaging reconstructions such as SENSE and ESPIRiT. Sensitivity maps characterize the spatial response of each coil in the array and enable the decoupling of overlapping spatial encoding. This section describes the procedure used to estimate sensitivity maps using BART’s `ecalib` command.

4.3.2.1 ESPIRiT Calibration

ESPIRiT-based calibration is performed using BART’s `ecalib` tool, which operates on multi-coil k-space data to produce spatial sensitivity maps. The calibration process focuses on a designated low-frequency region of k-space, which contains the most

coherent energy across coils. In this work, a calibration width of 24 central phase-encoding lines was typically used, specified via the `-r24` flag.

The `-m1` flag configures the estimation to output a single sensitivity map per coil, as opposed to multiple eigenvector-based maps. This simplification was suitable for standard SENSE reconstructions and provides a stable estimate of coil profiles without requiring eigenvalue thresholding.

The output of this calibration step is a complex-valued multi-coil image, where each channel represents the spatial sensitivity of the corresponding coil. These maps are subsequently used for both image-domain and k-space-domain parallel imaging reconstructions.

4.3.3 Masking Strategies for Undersampling

To simulate accelerated MRI acquisitions and assess the performance of reconstruction techniques under data scarcity, k-space undersampling was introduced using Cartesian masks. This section outlines the two primary strategies employed for creating undersampled data: (i) the variable-density Cartesian masks inspired by the fastMRI framework, and (ii) custom-designed centered masks tailored for controlled experimentation.

4.3.3.1 Variable-Density Cartesian Mask (fastMRI Style)

The fastMRI dataset provides a standardized approach for generating realistic undersampling patterns through variable-density masks. These were implemented using the `RandomMaskFunc` utility from the `fastmri` library. The mask is defined by two key parameters:

- **center_fraction**: Fraction of low-frequency k-space lines that are always retained.
- **acceleration**: The overall subsampling factor applied to the full k-space.

By applying `RandomMaskFunc` to fully-sampled k-space slices, a binary mask is produced and multiplied with the input to simulate acquisition with missing lines. This allows for flexible experimentation under varying sampling densities. Visualizations of the generated masks were used to verify sampling density profiles and validate consistency across slices.

4.3.3.2 Custom Centered Cartesian Mask

In parallel, a manually constructed centered mask was developed to provide intuitive control over sampling geometry. This method creates a 1D binary mask along the frequency-encoding direction, retaining a contiguous central portion of k-space while discarding the periphery.

The resulting 1D mask is then replicated along the orthogonal phase-encoding dimension to form a 2D mask. This method facilitates direct control over the sampling pattern and enhances interpretability when comparing reconstruction outcomes. Like the fastMRI masks, these were visualized to ensure spatial coherence and verify effective acceleration.

4.3.4 Image Reconstruction

Image reconstruction from multi-coil k-space involves converting frequency-domain data into spatial-domain images while effectively handling coil sensitivities. In this work, reconstructions were carried out using the BART toolbox, with both conventional and advanced parallel imaging techniques. This section details the procedures followed for reference, SENSE, and ESPIRiT-based reconstructions.

4.3.4.1 Reference (Fully Sampled) Reconstruction

As a baseline, fully sampled k-space data was transformed into images using an inverse Fourier transform followed by coil combination. Specifically, the following steps were executed:

1. **Inverse FFT:** A 2D inverse Fourier transform was performed using BART's `fft` command along spatial dimensions to obtain complex coil images.
2. **Coil Combination:** The resulting coil images were combined using `fmac`, which performs a weighted sum based on estimated coil sensitivity maps. This produces a reference image used for evaluation of undersampled reconstructions.

4.3.4.2 SENSE-based Reconstruction

SENSE (Sensitivity Encoding) leverages the known coil sensitivity profiles to solve the inverse problem of recovering an image from undersampled k-space. The reconstruction was performed using BART's `pics` module:

- **Input:** Undersampled k-space and corresponding single-set sensitivity maps (from `ecalib -m1`).
- **Method:** Solving a regularized least-squares problem where the encoding model incorporates both the undersampling pattern and the coil profiles.
- **Command:** `bart pics -R T:0:0:0.01 kspace sensmaps recon`

The primary trade-off observed in SENSE reconstruction was between acceleration and image quality. While increased undersampling accelerates acquisition, it leads to amplified noise and residual aliasing, particularly in high-frequency regions with low coil sensitivity diversity.

4.3.4.3 ESPIRiT-based Reconstruction

ESPIRiT (Enhanced Sensitivity encoding with Iterative Reconstruction) extends the SENSE approach by incorporating multiple eigenvector-based sensitivity maps to improve robustness. The reconstruction followed a similar pipeline as SENSE, with the key differences being:

- **Multiple Maps:** Sensitivity estimation via `ecalib` without the `-m1` flag yields multiple maps accounting for uncertainties.
- **Improved Conditioning:** ESPIRiT naturally handles inconsistencies due to noise or model mismatch, often producing cleaner images in highly undersampled regimes.
- **Execution:** `bart pics` uses all available maps for data consistency and coil combination.

ESPIRiT proved particularly advantageous in regions with low SNR or overlapping sensitivity, leading to improved reconstruction fidelity and reduced artifact propagation compared to single-map SENSE.

4.3.5 Comparative Analysis of Reconstructions

To evaluate the effectiveness of different reconstruction strategies, both visual and quantitative comparisons were performed between the reference (fully sampled), SENSE, and ESPIRiT-based images. This comparative analysis highlights the strengths and limitations of each technique in the context of parallel imaging.

4.3.5.1 Visual Comparison

Reconstructed images were displayed side-by-side to enable direct qualitative assessment. The key reconstructions compared included:

- **Reference Image:** Obtained from fully sampled k-space using coil combination.
- **SENSE Reconstruction:** Based on single-map sensitivity encoding.
- **ESPIRiT Reconstruction:** Incorporating multiple eigenvector maps for robustness.

The visual inspection focused on assessing the following aspects:

1. **Aliasing Artifacts:** SENSE images exhibited residual aliasing at higher acceleration factors.
2. **Blurring:** Some SENSE reconstructions showed mild spatial blurring, especially near edges.
3. **Sharpness and Fidelity:** ESPIRiT reconstructions preserved finer details and were closer to the reference visually.

These observations provided intuitive insight into reconstruction quality before resorting to numerical metrics.

4.3.5.2 Quantitative and Difference Maps

To quantify reconstruction errors and localize discrepancies, difference maps were generated:

- **Reference-Based Differences:** Subtraction of SENSE and ESPIRiT reconstructions from the fully sampled reference.
- **Inter-method Differences:** Direct comparison between SENSE and ESPIRiT to highlight method-specific artifacts.

The magnitude of the complex-valued difference was calculated and visualized using heatmap-style plots. These maps emphasized regions with higher deviation, such as tissue boundaries and areas with low SNR.

Such comparative metrics serve as a diagnostic tool for analyzing reconstruction performance and guiding future improvements in masking, regularization, or sensitivity estimation.

4.3.5.3 Generalization Across Modalities

The reconstruction pipeline was designed to be modality-agnostic, ensuring consistent applicability across both knee and brain MRI datasets. By maintaining a modular structure for each processing stage—ranging from k-space formatting and masking to sensitivity estimation and reconstruction—the framework seamlessly handled variations in anatomy and scan protocols.

A few key considerations ensured this generalization:

- **Consistent Data Formatting:** Conversion of both native .cfl and .h5 files into a BART-compatible format was critical. Proper dimensionality alignment (e.g., transpose and expansion of axes) ensured that subsequent tools operated as intended across datasets.
- **Masking Strategy Adaptability:** Both fastMRI-style variable-density masks and manually designed centered masks were applied successfully to different datasets, highlighting the pipeline’s flexibility in undersampling simulation.
- **Reconstruction Robustness:** The ESPIRiT-based reconstruction, leveraging multiple eigenvector maps, showed superior performance across modalities—particularly in high-noise brain images. This robustness underscores ESPIRiT’s capability to generalize better than traditional SENSE under varying imaging conditions.

This generalization capacity makes the proposed reconstruction pipeline suitable for broader deployment in clinical and research workflows involving multicoil MRI data.

In summary, the reconstruction pipeline presented in this chapter integrates key stages of data preprocessing, sensitivity estimation, undersampling, and image reconstruction using both SENSE and ESPIRiT frameworks. Through consistent application

across multiple datasets and anatomies, the pipeline demonstrates not only technical rigor but also adaptability and robustness. These foundational steps lay the groundwork for advanced evaluations and potential clinical translation, which will be further explored in the subsequent sections.

4.4 Custom Dataloader for Deep Learning-Based MRI Reconstruction

As we transition from traditional and compressed sensing methods to deep learning-based approaches for MRI reconstruction, the need for an efficient and modular data pipeline becomes paramount. Deep learning models require large amounts of training data, processed in a consistent and memory-efficient manner. To address this, we developed a custom PyTorch dataset class named `Custom_FastMRIDataset`. This dataloader is designed to flexibly handle coil-wise multi-slice MRI data, perform on-the-fly preprocessing, and generate input-target pairs for a variety of reconstruction paradigms, such as k-space domain, image-space domain, and root-sum-of-squares (RSS) image domain.

Dataset Class Design

The `Custom_FastMRIDataset` class inherits from `torch.utils.data.Dataset`, and it is initialized with the following parameters:

- `data_path`: Path to a `.npy` file containing the MRI dataset.
- `mask_func`: Mask generator, defaulting to `EquiSpacedMaskFunc` with center fraction 0.08 and acceleration factor 2.
- `transform`: Optional transformation to apply to data.
- `K, I, rss_combine`: Boolean flags to control output data (k-space, image-space, or RSS images).

Efficient Data Loading

To handle large datasets efficiently, we use `numpy.load(..., mmap_mode='r')`, enabling memory-mapped access to the MRI slices without fully loading the data into

memory. This method ensures scalability when dealing with large MRI datasets. The number of slices is derived from the shape of the loaded array.

Masking and Transformations

The `apply_mask` function applies a subsampling mask to the k-space data. For consistency, a "full mask" is generated by applying the mask to a tensor of ones. Additionally, an inverse FFT converts both masked and unmasked k-space data into image-space.

RSS Combination

The root-sum-of-squares (RSS) technique is applied to both unmasked and masked image-space data to generate `rss_combined` and `masked_rss_combined` respectively. This step is essential for multi-coil image reconstruction, ensuring that the information from different coils is combined into a single image.

Flexible Output Control

The dataset class offers flexible output options based on the configuration flags:

- If `K=True`, only k-space data and the mask are returned.
- If `I=True`, image-space data and the mask are returned.
- If `rss_combine=True`, only RSS-combined images are returned.
- If no flags are set, all data representations (k-space, masked k-space, image-space, RSS) are returned.

4.4.1 Key Features of the Custom Dataset Loader

- **Memory Efficiency:** Memory-mapped dataset loading via `numpy.load(..., mmap_mode='r')`, ensuring large datasets are handled efficiently.
- **Masking Strategy:** Subsampling masks (default: EquiSpaced with center fraction 0.08 and acceleration 2) are applied using `apply_mask()`. A "full mask" is also created for reference.

- **Multi-Representation Support:** Depending on the flags, the loader can return:
 - k-space and masked k-space data
 - Image-space data (IFFT reconstructions)
 - RSS-combined images from multi-coil data
- **Flexible Output Options:** Users can control what is returned, e.g., masked RSS images for training a U-Net model.

4.4.2 DataLoader Integration and Visualization

To enable mini-batch training, we integrated a custom dataset into PyTorch’s `DataLoader`. The `DataLoader` is used to load data in batches, which is essential for efficient training, especially when working with large datasets. The batch is fetched using the `next()` function, which retrieves the next batch from the `train_loader`. In this case, we specify a batch size of 4 and enable shuffling of the data.

Once the batch is fetched, the true and masked RSS (Root Sum of Squares) images are extracted from the batch. These images are then visualized using `matplotlib`’s `imshow` function, with the images displayed in grayscale. This process allows you to visualize both the ground truth (true RSS) and the masked RSS images from the batch.

Integration with PyTorch DataLoader

The dataset is wrapped using PyTorch’s `DataLoader` to facilitate easy batching and shuffling. The `batch_size` is set to 4, meaning that each mini-batch will contain 4 samples. The `shuffle` parameter is set to `True`, which ensures that the data is randomly shuffled before each epoch to improve training performance and prevent the model from overfitting to any specific order of data. Additionally, the `num_workers` is set to 0, indicating that the data loading will be done in the main process without using additional worker processes. This setup enables randomized mini-batch training with minimal overhead, allowing the model to be trained efficiently with a manageable memory footprint.

Data Validation and Visualization

The ground truth and masked RSS images are visualized using `matplotlib.pyplot.imshow()` in grayscale. To ensure correct formatting, the shapes and types of the data are verified. The shape of the data is expected to be `[batch_size, height, width]`, for example, `[4, 640, 320]`, where the batch size is 4, and the height and width correspond to the dimensions of each image. Additionally, the data type is verified to be `torch.float32`, ensuring that the images are represented with the correct precision for further processing and training.

This custom dataloader offers a robust, scalable, and flexible backend for deep learning-based MRI reconstruction. Its efficient data loading, support for multiple data representations, and integration with PyTorch’s `DataLoader` make it a powerful tool for developing and experimenting with various MRI reconstruction models.

4.5 Chapter Conclusion

This chapter presented the complete methodology adopted for deep learning-based MRI reconstruction. We began by describing the dataset used — the `fastMRI` dataset — which provides high-quality, coil-wise raw k-space data suitable for supervised learning. We then introduced BART (Berkeley Advanced Reconstruction Toolbox), an established tool for traditional and compressed sensing-based MRI reconstruction, highlighting its utility in generating reference outputs and benchmarking.

Following that, we provided an end-to-end overview of the proposed deep learning pipeline, emphasizing the modularity and extensibility required for modern reconstruction tasks. The pipeline design integrates data preprocessing, k-space masking, coil combination strategies, and the generation of ground truth and masked inputs.

A key contribution detailed in this chapter was the design and implementation of a custom PyTorch-based dataloader. This component ensures flexible data retrieval in different reconstruction paradigms (k-space, image-space, and RSS) while maintaining memory efficiency through NumPy memory mapping. With runtime masking and output control via user-specified flags, this dataloader forms the backbone of the training and validation pipeline for neural networks.

Altogether, this methodology lays the technical groundwork for the experiments and results discussed in the next chapter, ensuring reproducibility, scalability, and adaptability for future developments in MRI reconstruction.

Chapter 5

Results and Discussions

In this chapter, we present and analyze the outcomes of classical and deep learning-based MRI reconstruction methods on both knee and brain datasets. The experiments were performed using the BART toolbox, ESPIRiT, SENSE algorithms, and our custom deep learning dataloader. Visualizations highlight the reconstruction quality under various acceleration factors and provide comparative insights using difference maps.

5.1 BART Implementation Results

We begin by showcasing the outputs from BART reconstructions using both CFL-format data and fastMRI-converted data.

Reference vs Downsampled Scan (AF = 2): Knee

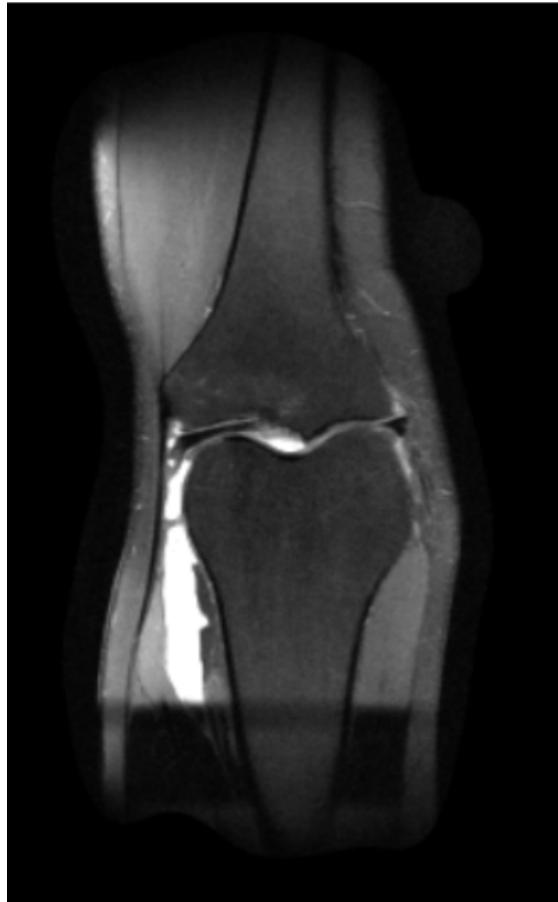


Figure 5.1: *
(a) Knee Reference Scan



Figure 5.2: *
(b) Knee Downsampled Scan (AF = 2)

Figure 5.3: Comparison between fully sampled and downsampled (AF = 2) knee MRI scans.

Reference vs Downsampled Scan (AF = 4): Brain

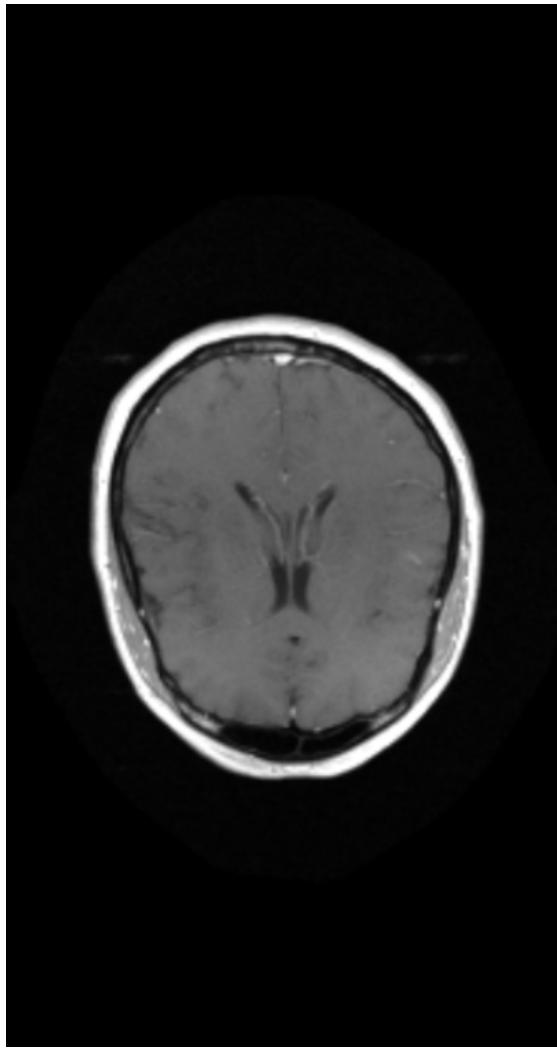


Figure 5.4: *
(a) Brain Reference Scan

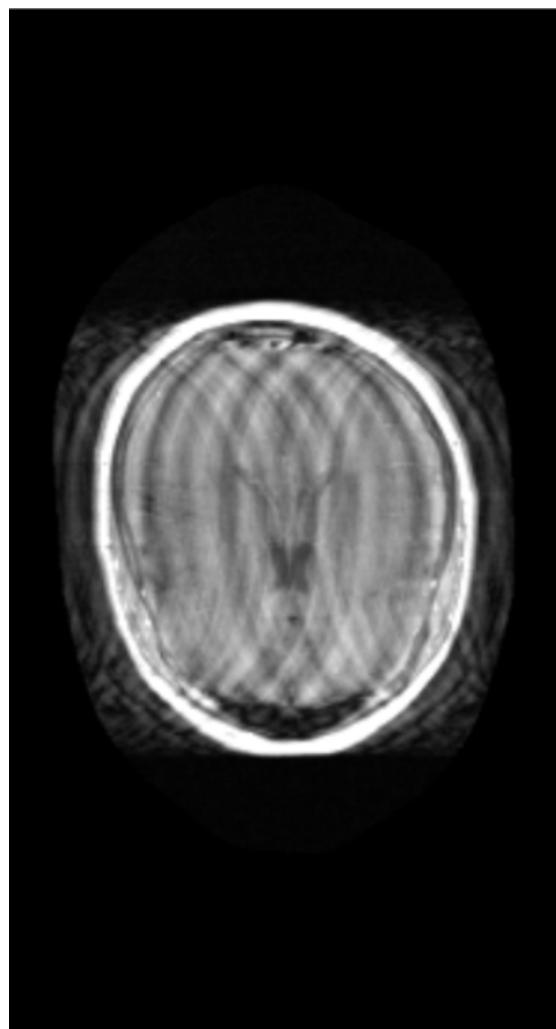


Figure 5.5: *
(b) Brain Downsampled Scan (AF = 4)

Figure 5.6: Comparison between fully sampled and downsampled (AF = 4) brain MRI scans.

5.2 Knee Scan Reconstruction Comparisons

Reconstruction Outputs (Side-by-Side)

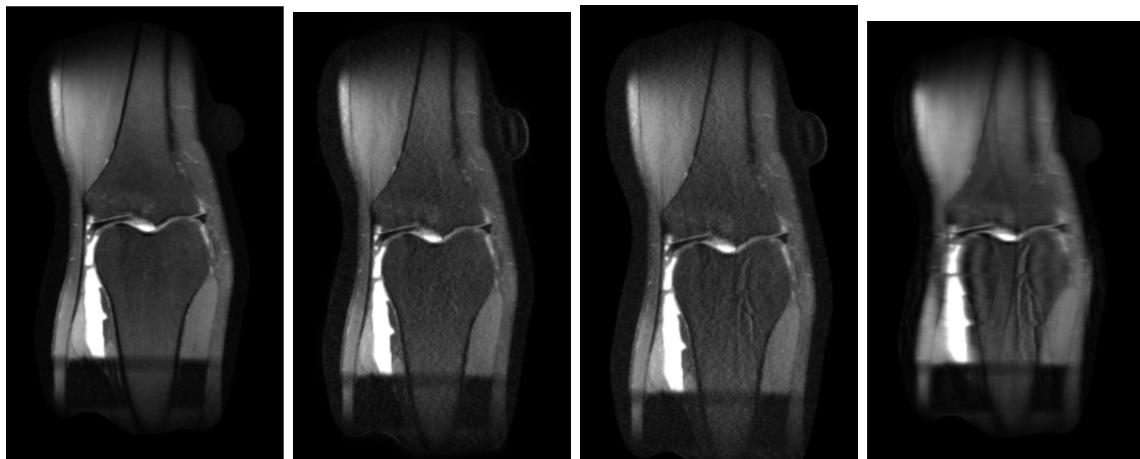


Figure 5.7: Knee Scan Reconstruction Comparison: Reference, ESPIRiT, SENSE, and Downsampled (w/o correction).

Difference Map Comparisons

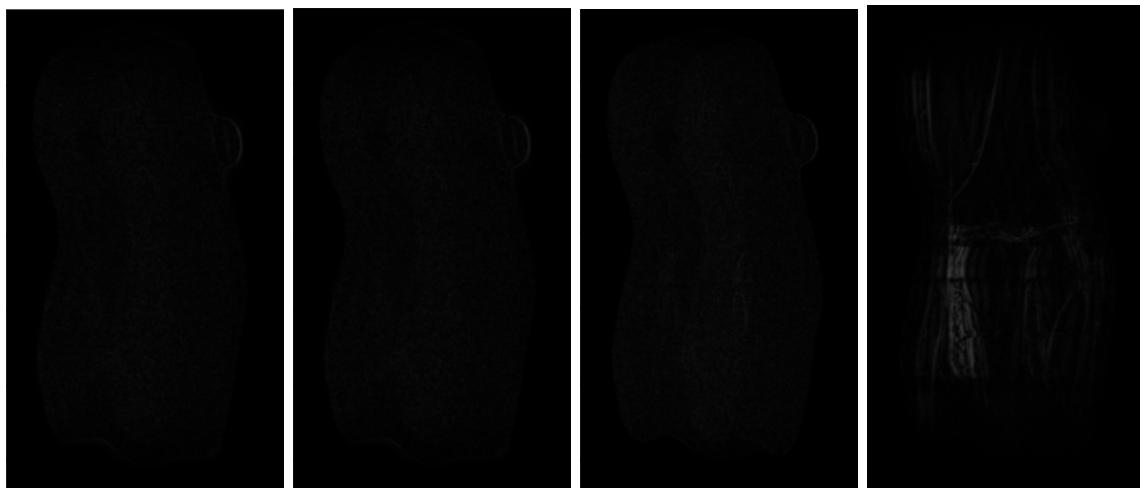


Figure 5.8: Difference Maps: Reference vs ESPIRiT, SENSE, and Downsampled reconstructions.

5.3 Brain Scan Reconstruction Comparisons

Reconstruction Outputs (Side-by-Side)

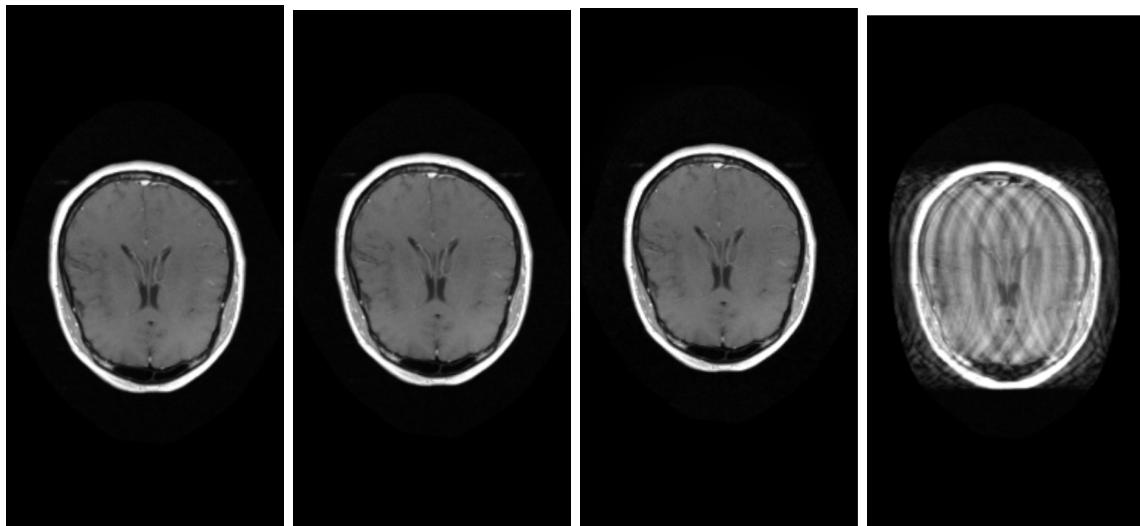


Figure 5.9: Brain Scan Reconstruction Comparison: Reference, ESPIRiT, SENSE, and Downsampled (w/o correction).

Difference Map Comparisons

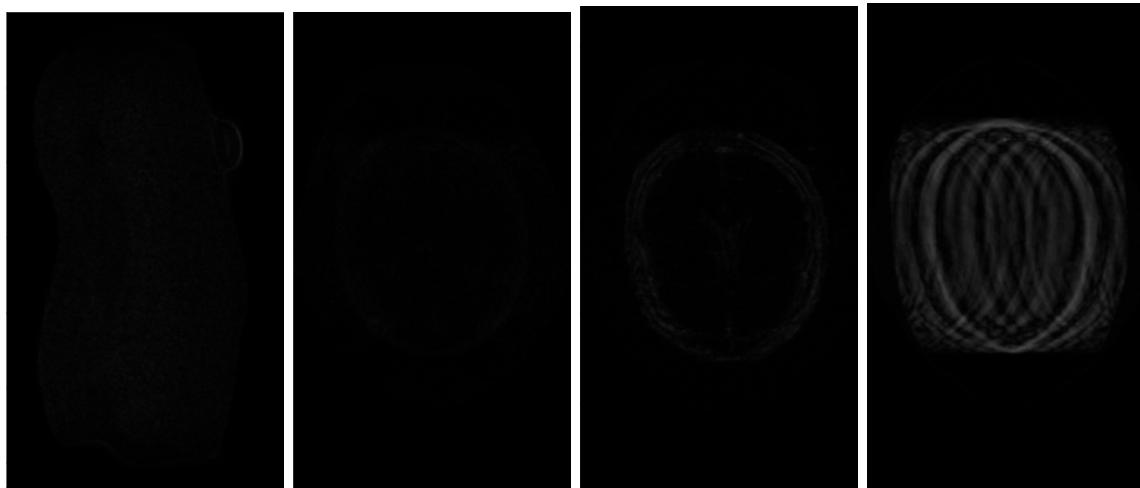


Figure 5.10: Difference Maps: Reference vs ESPIRiT, SENSE, and Downsampled reconstructions.

5.4 Impact of Acceleration on ESPIRiT Reconstructions

As acceleration factor increases, even robust classical methods like ESPIRiT begin to show degradation in output quality, leading to the necessity of deep learning-based approaches.

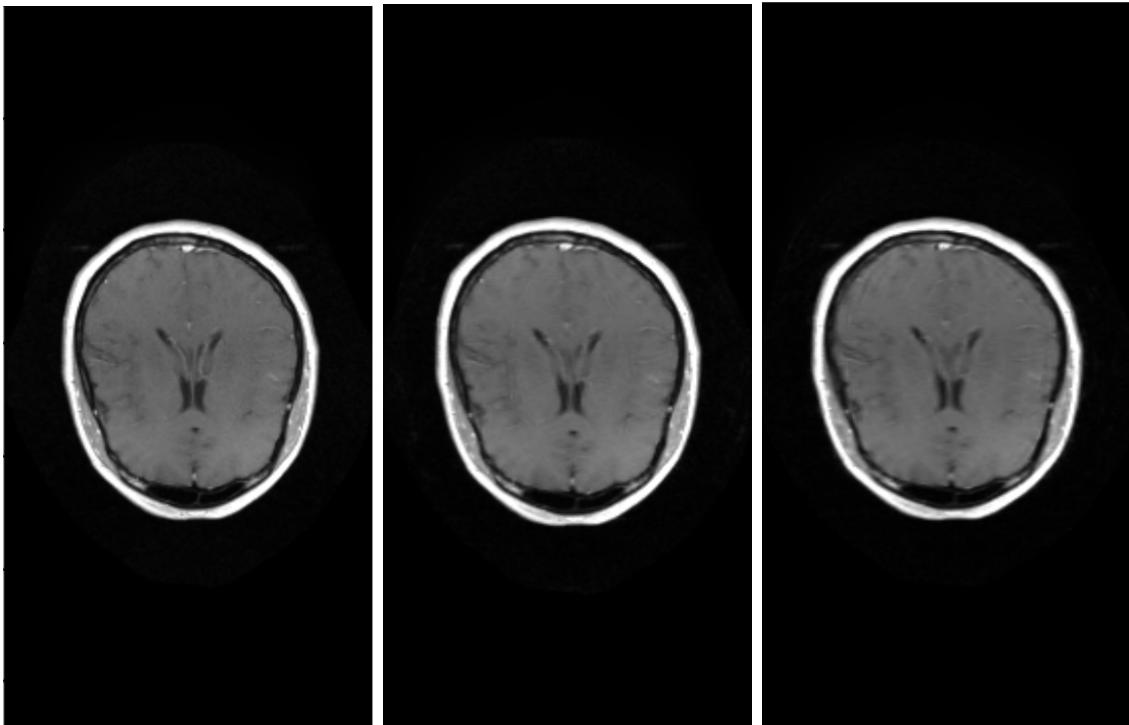


Figure 5.11: ESPIRiT Reconstructions at Increasing Acceleration Factors: Left to Right — AF=2, AF=3, AF=4.

5.5 Network Training

The network is trained using a complex-valued U-Net designed for $2\times$ upscaling. The key components and flow of the network are as follows:

1. **Encoder Path (Downsampling):** The encoder progressively reduces the spatial resolution of the input while increasing the feature depth. The complex convolutions are applied using `Complex_DoubleConv` to extract low-level features, maintaining both phase and magnitude. The `Complex_Down_DoubleConv2d` layer combines

complex convolution and max pooling to halve the spatial dimensions at each stage:

$$50 \times 50 \rightarrow 25 \times 25 \rightarrow 12 \times 12 \rightarrow 6 \times 6 \rightarrow 3 \times 3$$

Additionally, feature channels increase at each stage:

$$16 \rightarrow 32 \rightarrow 64 \rightarrow 128 \rightarrow 128 \quad (\text{or } 256 \text{ if bilinear interpolation is not used})$$

2. ****Bottleneck:**** Positioned at the deepest part of the network, the bottleneck consists of a `Complex_DoubleConv` layer, capturing the most abstract features at the lowest spatial resolution. This serves as the bridge between the encoder and decoder.

3. ****Decoder Path (Upsampling with Skip Connections):**** The decoder mirrors the encoder structure, progressively upsampling the feature maps back to the original resolution. The `Complex_Up` layer performs upsampling (either via bilinear interpolation or transposed convolution) followed by complex convolutions. Skip connections are used to concatenate the corresponding encoder features, which helps preserve spatial details and improve reconstruction accuracy. The feature dimensions decrease as the network ascends:

$$128 \rightarrow 64 \rightarrow 32 \rightarrow 16$$

4. ****Output Layer:**** The final output is produced using the `Complex_OutConv` layer, a complex 1×1 convolution that reduces the feature maps to the desired output channels, maintaining both the real and imaginary parts of the output. This layer is suitable for generating either a reconstructed complex-valued image or a processed k-space map.

5.6 Discussion

The classical reconstruction methods, ESPIRiT and SENSE, perform reliably at moderate acceleration factors, preserving structural integrity with minimal artifacts. However, as the acceleration factor increases (e.g., AF = 4), their limitations become evident through the emergence of aliasing artifacts and noticeable degradation in image quality. This is particularly critical in diagnostic imaging, where fine anatomical details are essential.

These results highlight the potential of deep learning-based approaches, which can exploit learned priors from large datasets to better handle aggressive undersampling.

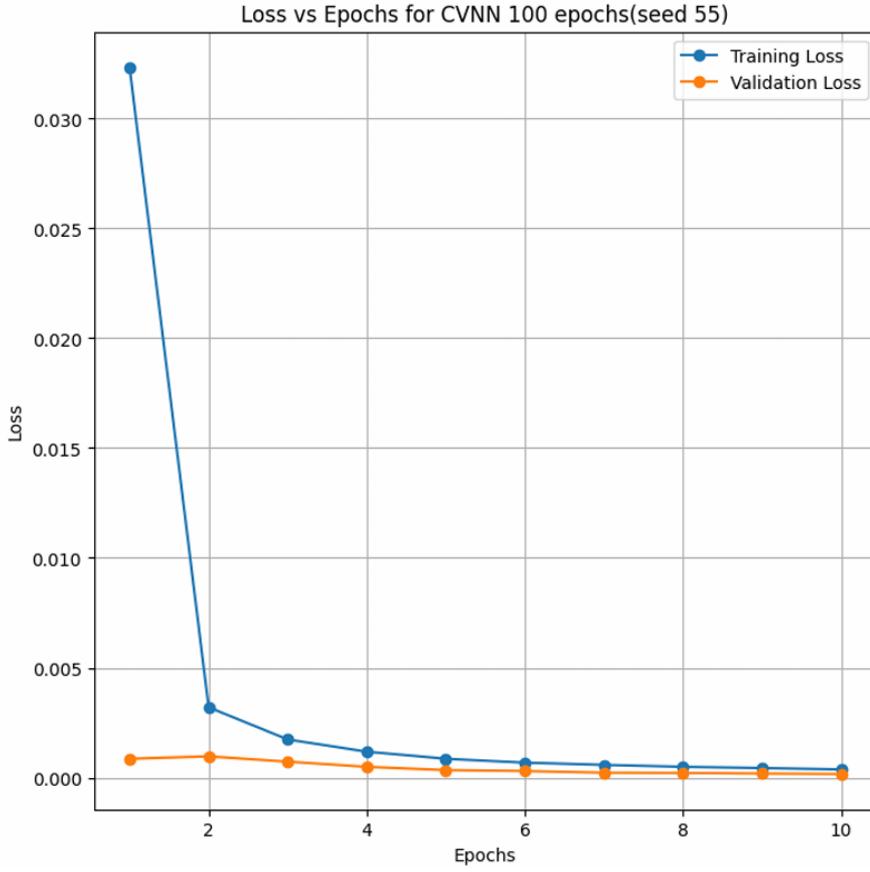


Figure 5.12: Training and validation loss curves versus the number of epochs. The loss curves show rapid convergence, with the network reaching near-convergence within 10 epochs, even when run for up to 100 epochs.

By learning complex, non-linear mappings from aliased inputs to fully sampled outputs, such methods show promise in restoring high-fidelity reconstructions even in highly undersampled regimes.

Our findings underscore the necessity of a flexible and modular reconstruction pipeline—capable of integrating classical methods as well as data-driven alternatives. The implemented framework and dataloader facilitate consistent benchmarking across datasets and anatomical regions (e.g., knee, brain), laying the groundwork for future experimentation and advancement in deep learning-based MRI reconstruction.

As illustrated in Figure 5.12, the training and validation loss curves for the Complex-Valued Neural Network (CVNN) over 10 epochs demonstrate rapid convergence. The training loss shows a steep decline within the first few epochs, sta-

bilizing thereafter. The validation loss remains consistently low and closely follows the training loss, indicating good generalization and minimal overfitting. This highlights the capability of CVNNs to efficiently learn stable mappings for reconstructing high-quality images from undersampled k-space data.

Chapter 6

Conclusion

This thesis presented a complete and modular MRI reconstruction pipeline integrating classical parallel imaging methods with efficient data preprocessing and loading mechanisms, aimed at exploring and extending into deep learning-based approaches. The major takeaways and contributions of this work are as follows:

- **Pipeline Development:** An end-to-end MRI reconstruction workflow was developed using the BART toolbox. Classical reconstruction techniques such as SENSE, ESPIRiT, and PICS were implemented and evaluated on both knee and brain scans using the fastMRI dataset.
- **Foundational Understanding:** The project solidified key theoretical foundations in MRI reconstruction, including:
 - Fourier and inverse Fourier transforms in MRI signal processing
 - Coil sensitivity maps and their impact on image formation
 - Parallel imaging acceleration using multiple receiver coils
- **Custom Masking and Preprocessing:** Realistic undersampling was simulated using equispaced masks configurable by acceleration factor and center fraction. These were paired with clean and consistent preprocessing steps like inverse FFT, k-space manipulation, and RSS image generation.
- **Efficient Dataloading:** A custom PyTorch-compatible dataloader was designed to support multiple data formats (k-space, image-space, RSS). It was memory-efficient and highly configurable, supporting downstream training of deep learning models.

- **Network Training and Generalization:** Training experiments with a Complex-Valued Neural Network (CVNN) on undersampled MRI data demonstrated rapid convergence, as indicated by a steep initial drop in training loss. The validation loss remained low and closely tracked the training curve, suggesting strong generalization and minimal overfitting. This underscores the potential of complex-valued networks in learning stable and efficient mappings for image reconstruction, especially in high undersampling regimes.

SENSE vs ESPIRiT: A Comparative Perspective

Visual comparisons in Chapter 5 clearly demonstrated that ESPIRiT outperformed SENSE, especially at higher acceleration factors. This aligns with theoretical expectations:

- **SENSE** operates in the image domain and directly solves a linear system using coil sensitivity maps. While intuitive and fast, it is highly sensitive to inaccuracies in sensitivity estimation and suffers from noise amplification in regions where coil profiles overlap significantly. It also requires explicit calibration data and suffers from g-factor related SNR penalties.
- **ESPIRiT**, on the other hand, is a more robust autocalibrating method. It combines the benefits of both SENSE and GRAPPA by estimating sensitivity maps using eigenvalue decomposition of a calibration matrix. Unlike SENSE, it can handle inconsistencies in coil data more gracefully and supports multi-map reconstructions for improved error tolerance.
- Experimental results showed that SENSE reconstructions often had residual aliasing and structural artifacts, especially at higher undersampling ($AF = 4$), while ESPIRiT maintained superior structural fidelity and contrast.

Need for Deep Learning-Based Solutions

While ESPIRiT remains a strong classical baseline, its performance also deteriorates with increasing undersampling. To push beyond this limitation, the thesis reviewed several modern deep learning methods that leverage data priors, non-linear mappings, and learning-based regularization strategies:

- **C-SENSE:** Combines compressed sensing with SENSE and integrates deep convolutional priors to enhance reconstruction.
- **DeepSENSE:** Learns coil sensitivity maps or performs end-to-end reconstruction using neural networks.
- **GAN-based SENSE:** Applies generative adversarial networks to suppress g-factor artifacts and hallucinate finer structural details.
- **CVNNs:** Use complex-valued neural networks to exploit the inherent complex structure of MRI data, showing promising generalization and stability during training, as demonstrated in our experiments.

These methods have shown promise in literature and commercial applications but require robust pipelines for preprocessing, batching, and evaluation—exactly what was built in this work.

Future Works

While this thesis focused primarily on classical reconstruction techniques and the foundational development of an adaptable MRI reconstruction pipeline, several promising directions remain open for exploration:

- **Compressed Sensing MRI:** Incorporate compressed sensing techniques into the current classical framework to exploit sparsity priors and improve image recovery, especially under high acceleration factors.
- **Deep Learning-based Reconstruction:** Implement deep learning-based models such as DeepSENSE and GAN-augmented SENSE architectures to learn complex mappings from under-sampled k-space data to fully reconstructed image space.
- **Advanced Neural Architectures:** Investigate the use of modern architectures like self-attention CNNs and high-frequency-aware diffusion models to better preserve structural details in reconstructed images.
- **Transformer-based Designs:** Explore densely connected and transformer-inspired networks such as DenseNet, Swin Transformer, and ReconFormer for building efficient, scalable, and high-fidelity MRI reconstruction systems.

- **Comprehensive Benchmarking:** Establish a thorough evaluation framework to benchmark these deep learning models against classical methods (SENSE, ESPIRiT, PICS) on multiple datasets, anatomical regions, and acceleration factors.

These directions aim to bridge the gap between physics-based mathematical models and data-driven deep learning paradigms, ultimately improving MRI reconstruction accuracy and robustness in real-world clinical scenarios.

In conclusion, this work lays a strong foundation for MRI reconstruction research, bridging theory with implementation, and classical methods with deep learning. It opens the path for future experiments to advance reconstruction quality under increasingly challenging acquisition constraints.

Bibliography

- [1] A. Hirose, “Complex-valued neural networks: The merits and their origins,” in *2009 International joint conference on neural networks*, pp. 1237–1244, IEEE, 2009.
- [2] B. N. Dash and N. Khare, “Deep complex neural network applications in remote sensing: an introductory review,” *Radar Sensor Technology XXV*, vol. 11742, pp. 34–44, 2021.
- [3] A. D. Dola, *Complex Valued Convolutional Neural Networks for Image Restoration Thesis to be submitted in the partial fulfillment of the requirements for the degree MASTER OF SCIENCE*. PhD thesis, Indian Institute of Technology, Kharagpur, 2023.
- [4] J. Zbontar, F. Knoll, A. Sriram, T. Murrell, Z. Huang, M. J. Muckley, A. Defazio, R. Stern, P. Johnson, M. Bruno, *et al.*, “fastmri: An open dataset and benchmarks for accelerated mri,” *arXiv preprint arXiv:1811.08839*, 2018.
- [5] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, “Sense: sensitivity encoding for fast mri,” *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 42, no. 5, pp. 952–962, 1999.
- [6] M. Uecker, P. Lai, M. J. Murphy, P. Virtue, M. Elad, J. M. Pauly, S. S. Vasanawala, and M. Lustig, “Espirit—an eigenvalue approach to autocalibrating parallel mri: where sense meets grappa,” *Magnetic resonance in medicine*, vol. 71, no. 3, pp. 990–1001, 2014.
- [7] X. Peng, B. P. Sutton, F. Lam, and Z.-P. Liang, “Deepsense: Learning coil sensitivity functions for sense reconstruction using deep learning,” *Magnetic resonance in medicine*, vol. 87, no. 4, pp. 1894–1902, 2022.

- [8] J. Lv, G. Li, X. Tong, W. Chen, J. Huang, C. Wang, and G. Yang, “Transfer learning enhanced generative adversarial networks for multi-channel mri reconstruction,” *Computers in Biology and Medicine*, vol. 134, p. 104504, 2021.
- [9] C. Lee, H. Hasegawa, and S. Gao, “Complex-valued neural networks: A comprehensive survey,” *IEEE/CAA Journal of Automatica Sinica*, vol. 9, no. 8, pp. 1406–1426, 2022.