DEEP LEARNING-BASED RECONSTRUCTION METHODS FOR NEAR-FIELD MIMO RADAR IMAGING

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

İRFAN MANISALI

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
ELECTRICAL AND ELECTRONICS ENGINEERING

JUNE 2022

Approval of the thesis:

**DEEP LEARNING-BASED RECONSTRUCTION METHODS FOR
NEAR-FIELD MIMO RADAR IMAGING**

submitted by **İRFAN MANISALI** in partial fulfillment of the requirements for the
degree of **Master of Science in Electrical and Electronics Engineering Department, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar
Dean, Graduate School of **Natural and Applied Sciences**   _____

Prof. Dr. Ilkay Ulusoy
Head of Department, **Electrical and Electronics Engineering**   _____

Assoc. Prof. Dr. Sevinç Figen Öktem
Supervisor, **Electrical and Electronics Engineering, METU**   _____

Examining Committee Members:

Prof. Dr. A. Aydın Alatan
Electrical and Electronics Engineering, METU   _____

Assoc. Prof. Dr. Sevinç Figen Öktem
Electrical and Electronics Engineering, METU   _____

Prof. Dr. Sencer Koç
Electrical and Electronics Engineering, METU   _____

Assoc. Prof. Dr. Yeşim Serinağaoğlu Doğrusöz
Electrical and Electronics Engineering, METU   _____

Assist. Prof. Dr. Aykut Koç
Electrical and Electronics Engineering, Bilkent University   _____

Date:29.06.2022

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Surname: İrfan Manisalı

Signature   :

ABSTRACT

DEEP LEARNING-BASED RECONSTRUCTION METHODS FOR NEAR-FIELD MIMO RADAR IMAGING

Manisalı, İrfan

M.S., Department of Electrical and Electronics Engineering

Supervisor: Assoc. Prof. Dr. Sevinç Figen Öktem

June 2022, 58 pages

Near-field multiple-input multiple-output (MIMO) radar imaging systems are of interest in diverse fields such as medicine, through-wall imaging, airport security, and surveillance. These computational imaging systems reconstruct the three-dimensional scene reflectivity distribution from the radar data. Hence their imaging performance largely depends on the image reconstruction method. The analytical reconstruction methods suffer from either low image quality or high computational cost. In fact, sparsity-based methods offer better image quality than the traditional direct inversion methods, but their high computational cost is undesirable in real-time applications. In this thesis, we develop two novel deep learning-based reconstruction methods for near-field MIMO radar imaging. The main goal is to achieve high image quality with low computational cost. The first approach has a two-staged structure that consists of an adjoint operation followed by a deep neural network. The adjoint stage exploits the observation model and back project the measurements to the reconstruction space. The second stage employs a deep neural network which is trained to convert the backprojected measurements to a suitable reflectivity image. For comparison, a second approach is also developed which replaces the adjoint stage with a fully con-

nected neural network. In this two-staged structure, the reconstruction is performed directly from the radar measurements using neural networks which are trained end-to-end to learn the direct mapping between the measurements and unknown reflectivity magnitude. For each case, a 3D U-Net is used at the second stage to jointly exploit range and cross-range correlations. We demonstrate the performance of the developed methods using a synthetically generated dataset and compare with the commonly used analytical methods. The developed two-staged method with adjoint provides the best reconstruction quality while enabling fast reconstruction.

ÖZ

YAKIN ALAN MİKRODALGA GÖRÜNTÜLEMEDE DERİN ÖĞRENMEYE DAYALI İMGE GERİÇATIM TEKNİKLERİ

Manisalı, İrfan

Yüksek Lisans, Elektrik ve Elektronik Mühendisliği Bölümü

Tez Yöneticisi: Doç. Dr. Sevinç Figen Öktem

Haziran 2022 , 58 sayfa

Yakın alan mikrodalga görüntüleme sistemleri tıp, duvar arkası görüntüleme, gizli silah tespiti ve gözetleme gibi çeşitli alanlarda kullanılmaktadır. Bu sistemler radardan elde edilen ölçümlerden üç boyutlu bir sahnenin yansıtıcılık dağılımını geri kazanır. Bu görüntüleme sistemlerinin başarımı, kullanılan geriçatım yöntemine büyük ölçüde bağlıdır. Literatürde var olan analitik yöntemlerin dezavantajları yüksek hesaplama yüküne sahip olmaları veya düşük geriçatım başarımı göstermeleridir. Seyreklik düzenlileştirmesine dayalı analitik yöntemlerin direkt evirme yöntemlerinden daha iyi geriçatım başarımı sağlamalarına rağmen, yüksek hesaplama yükleri gerçek zamanlı görüntülemede kullanılmalarını zorlaştırmaktadır. Bu tezde, yakın alan mikrodalga görüntüleme sistemleri için derin öğrenmeye dayalı iki yeni geriçatım yöntemi geliştirilmektedir. Bu tezin ana amacı yüksek kaliteli geriçatıma düşük hesaplama yükü ile ulaşmaktır. Geliştirilen ilk yöntem eklenik işlemini takiben derin sinir ağı içeren iki aşamadan oluşmaktadır. Eklenik işlemini gerçekleştiren ilk aşamada ölçüm modelinden faydalanılarak ölçümlerin görüntü uzayına geri izdüşümü yapılır. İkinci aşamada ise derin bir sinir ağı kullanılarak elde edilen ara sonucun sahnenin gerçek görüntü-

süne dönüştürülmesi amaçlanır. Karşılaştırma amacıyla geliştirilen ikinci yöntemde, geriçatılmak istenen görüntünün direkt olarak ölçümlerden derin sinir ağları kullanılarak elde edilmesi amaçlanır. Bu amaçla ilk geliştirilen yöntemin ilk aşamasında kullanılan eklenik işleci yerine tam bağlantılı sinir ağı kullanılır. Geliştirilen iki yaklaşımda da ikinci aşamada üç boyutlu U-Net yapısı kullanılır, böylece geriçatılan görüntünün her üç yönündeki ilinti bilgisinden de faydalanılır. Geliştirilen geriçatım yöntemlerinin başarımları sentetik veri seti üretilerek gösterilmekte ve analitik yöntemlerle karşılaştırılmaktadır. Eklenik işlecini içeren geliştirilen ilk yöntem, en iyi geriçatım kalitesine hızlı bir hesaplama ile ulaşmaktadır.

Anahtar Kelimeler: hesaplamalı görüntüleme, imgenin geriçatılması, ters problemler, derin öğrenme, evrişimli sinir ağları, çok-girdili çok-çıktılı radar görüntüleme, yakın alan mikrodalga görüntüleme

To my family

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

1D	1 Dimensional
2D	2 Dimensional
3D	3 Dimensional
MIMO	Multiple Input Multiple Output
SAR	Synthetic Aperture Radar
CNN	Convolutional Neural Network
DNN	Deep Neural Network
FFT	Fast Fourier Transform
CPU	Central Processing Unit
GPU	Graphics Processing Unit
HQS	Half Quadratic Splitting
ADMM	Alternating Direction Method of Multipliers
PSNR	Peak Signal-to-Noise Ratio
SSIM	Structural Similarity Index Metric
SNR	SNR Signal-to-Noise Ratio
MSE	Mean Squared Error
TV	Total Variation

CHAPTER 1

INTRODUCTION

Near-field radar imaging systems are of interest in diverse fields such as medicine, through-wall imaging, airport security, concealed weapon detection, and surveillance [1–4]. Earlier near-field radar imaging systems have operated in monostatic mode, i.e., with collocated transmitter and receiver antennas [5–7]. In such systems, a large number of transceiver antennas are required to maintain high image quality with fine range and cross-range resolutions. This results in high system cost, hardware complexity, and long acquisition time [8, 9].

Recently, sparse multiple-input multiple-output (MIMO) arrays with spatially distributed transmit and receive antennas (i.e. multistatic array) have gained more attention since they can offer high resolution with reduced cost, hardware complexity, and acquisition time. The number of antennas in MIMO systems can be significantly reduced compared to the monostatic case while maintaining high image quality [10, 11].

Near-field radar imaging systems, operating in monostatic or multistatic mode, are computational imaging systems that reconstruct the three-dimensional (3D) reflectivity distribution of the scene from the raw radar data. As a result, the imaging performance of such radar systems largely depends on the underlying image reconstruction method. Various analytical reconstruction methods are used to reconstruct the unknown scene reflectivity distribution from the measurements. These methods can be grouped into two categories: traditional direct inversion methods and regularized iterative reconstruction methods.

Traditional direct inversion methods obtain a direct solution for the equation system that models the observations without exploiting any prior information. These meth-

ods generally involve back projecting the measurements to the object plane using the adjoint operator and applying a filter-like operation [12–17]. Back-projection, filtered back-projection, and Kirchhoff migration methods [7, 14, 15, 18–24] can be given as examples of this class of methods for MIMO radar imaging. These traditional methods generally have low computational complexity but as a drawback, they can not offer state-of-the-art reconstruction performance. The reconstruction quality degrades for ill-posed cased with the presence of measurement noise and limited data (as acquired with sparse MIMO arrays).

Regularized iterative reconstruction methods on the other hand incorporate additional prior information (such as sparsity) into the reconstruction process. Motivated by compressed sensing theory [25–27], sparsity-based reconstruction methods are the most commonly used iterative inversion methods and they have been widely studied in various imaging problems including radar imaging [28–35], both for monostatic or far-field imaging settings [30, 34, 36–46], as well as for multistatic and near-field settings [29,33,47–52]. Although these methods provide better reconstruction quality than the traditional direct inversion methods, they suffer from high computational cost and large memory usage, and also require parameter tuning to achieve good reconstruction under different observation scenarios. These aspects are undesirable in real-time applications.

Recently, reconstruction techniques that exploit deep learning have emerged as an alternative to the analytical methods. These methods are shown to simultaneously achieve high reconstruction quality and low computational cost for various imaging problems [53–56]. The existing deep learning-based approaches in the literature can be grouped into three main classes: 1) learning-based direct inversion, 2) plug-and-play regularization, 3) learned iterative reconstruction based on unrolling.

Learning-based direct inversion methods are aimed to perform the reconstruction directly from the measurements using a deep neural network. Hence the neural network is trained to learn the direct mapping from the observations to the desired image solely using training data. However, these methods can not provide successful results whenever the observation model is complex, the unknown image does not look alike observations, or there is not much training data available. For this reason, commonly

an efficient analytical inversion method is first performed and provided to the network as a warm start. Subsequently a deep neural network is employed to improve this intermediate reconstruction. This type of approach is applied to various linear inverse problems in imaging such as noise removal [57], deconvolution [58], super-resolution [59,60], and tomography [12]. It has also been applied to nonlinear inverse problems such as phase-retrieval [61–64].

The key idea in plug-and-play regularization and unrolling-based deep-learning methods is to replace the hand-crafted analytical priors in model-based reconstruction methods with data-driven deep priors. In plug-and-play methods, a deep prior is first learned from training data and then utilized for the regularization of a model-based inversion method. This type of approach is applied to various linear inverse problems such as deconvolution, super-resolution, and image inpainting problems [65–67]. It has also been applied to nonlinear inverse problems such as phase-retrieval [68]. In unrolling-based learned iterative reconstruction methods, an iterative reconstruction method with deep prior is unrolled into an end-to-end trainable network [69]. These methods can also be utilized to learn the optimum regularization parameters [70] or the proximal operator [71]. Thus, unlike plug-and-play and direct inversion methods, unrolling-based approaches use the forward model during training, which comes with the cost of increased training time. This type of approach has also been applied to various inverse problems in imaging [69–71].

In the context of near-field radar imaging, deep learning-based reconstruction methods have not been studied much in the literature. Most of the proposed methods are for far-field settings in SAR/ISAR or MIMO radar imaging [72–80]. In the near-field radar imaging context, there are some works for deep learning-based approaches, but most of them apply to the monostatic setting such as [81, 82]. For near-field and MIMO radar imaging systems, there are fewer works [83] and no comprehensive study. In particular, the work in [83] develops a learning-based direct inversion method to reconstruct point scatterers from near-field MIMO radar data. In this approach, the magnitude and phase of the backprojection images are processed separately using 2D-convolutional layer blocks. Because the used network performs 2D processing only (using two-dimensional kernels that work on the cross-range dimensions), the correlation along the range direction can not be exploited. Moreover, the

training and testing are performed only for simple scenes that consist of point scatterers. These will significantly limit the performance of the method in real practical scenarios that involve complex extended targets (rather than simple point scatterers).

1.1 Proposed Methods

In this thesis, we developed two novel deep learning-based methods to reconstruct the 3D scene reflectivity from the near-field observations of a MIMO imaging radar. We also compared the performance of the developed methods with the commonly used analytical methods. The main goal is to achieve high image quality with low computational cost so that the developed method can be used in real-time applications. Learning-based direct inversion methods can provide such capabilities. For this reason, the developed approaches are based on learned direct reconstruction.

The first approach has a two-staged structure that consists of an adjoint operation followed by a 3D U-Net architecture. The adjoint stage exploits the observation model of the system and back project the measurements to the reconstruction space. The second stage employs a deep neural network which is trained to convert the backprojected measurements to a suitable magnitude-only reflectivity image.

For comparison, a second approach is also developed which replaces the earlier adjoint stage with a fully connected neural network. In this two-staged structure, the reconstruction is performed directly from the radar measurements using only neural networks, and the observation model is not used. These neural networks are trained end-to-end to learn the direct mapping between the measurements and the magnitude of the unknown reflectivity image. In both approaches, a 3D U-Net architecture is used at the second stage to jointly exploit the correlations along both range and cross-range directions.

Numerical simulations are performed for a microwave imaging setting. For both training and testing, a large synthetic dataset is randomly generated to obtain 3D scenes by taking into account the complex-valued and random phase nature of scene reflectivities. Using this dataset, we illustrate the performance of the developed approaches under different imaging scenarios and compare their performance with the commonly

used analytical reconstruction methods.

1.2 Contributions of the Thesis

The main contributions are as follows:

- Development of two novel 3D image reconstruction methods for near-field MIMO radar imaging based on learning-based direct inversion and using 3D convolutional layers

- Generation of synthetic 3D scenes that involves extended targets with random phase to obtain large data for training the neural networks

- Performance comparison with the commonly used analytical methods (i.e. back-projection and sparsity-based reconstruction)

- Comprehensive experiments on synthetic 3D scenes with quantitative and qualitative analysis by considering different compression and noise levels in the observations as well as different network architectures

Compared with the previous works the developed approaches reduce memory usage and computation time, while offering good reconstruction quality. In fact, compared to back-projection and sparsity-based methods, the developed two-staged approach with adjoint operation achieves the best reconstruction quality on the synthetically generated dataset both visually and qualitatively, while also enabling fast reconstruction. All adjustable parameters of the developed approaches are learned end-to-end, which avoids the need for parameter tuning.

1.3 The Outline of the Thesis

The thesis is organized as follows. Chapter 2 describes the working principle of a near-field MIMO radar imaging system by introducing the observation model and the image reconstruction problem. The related work on image reconstruction is presented in Chapter 3, which includes the review of analytical methods for near-field

MIMO radar imaging and deep-learning based reconstruction methods for general inverse problems in imaging. The developed deep learning-based methods for near-field MIMO radar imaging are presented in Chapter 4. Numerical simulation results are presented in Chapter 5, which also contains the details of the simulation setting, synthetic dataset generation and training procedure. Finally, we conclude the thesis and discuss future research directions in Chapter 6.

CHAPTER 2

NEAR-FIELD MIMO RADAR IMAGING

In this chapter, we introduce the observation model for near-field MIMO radar imaging in Section 2.1 and the underlying inverse problem for image reconstruction in Section 2.2. The thesis focuses on the solution of this inverse problem by exploiting deep learning.

2.1 Observation Model

A sample observation geometry for near-field MIMO radar imaging is illustrated in Fig. 2.1. The transmit and receive antennas are spatially distributed on a planar MIMO array located at $z = 0$. Each transmit antenna, located at $(x_t, y_t, 0)$, illuminates a scene that lies in the near-field of the array. Using Born approximation for the scattered field, the signal captured by the receive antenna at $(x_r, y_r, 0)$ due to a single scatterer at (x, y, z) with reflectivity $s(x, y, z)$ can be expressed in the time-domain as follows [19]:

$$r\left(x_t, y_t, x_r, y_r, t\right) = \frac{1}{4\pi d_t d_r}\, s(x, y, z)\, p\left(t - \frac{d_t}{c} - \frac{d_r}{c}\right) \tag{2.1}$$

where

$$d_t = \sqrt{\left(x_t - x\right)^2 + \left(y_t - y\right)^2 + z^2} \tag{2.2}$$

$$d_r = \sqrt{\left(x_r - x\right)^2 + \left(y_r - y\right)^2 + z^2} \tag{2.3}$$

Here $r\left(x_t, y_t, x_r, y_r, t\right)$ denotes the time-domain measurement obtained using the transmitter at $(x_t, y_t, 0)$ and the receiver at $(x_r, y_r, 0)$, where d_t and d_r respectively

denote the distances from the corresponding transmit and receive antenna elements to the scatterer at (x, y, z), $p(t)$ is the transmitted pulse, and c denotes the speed of the light. By applying Fourier transform, the received signal due to a single scatterer
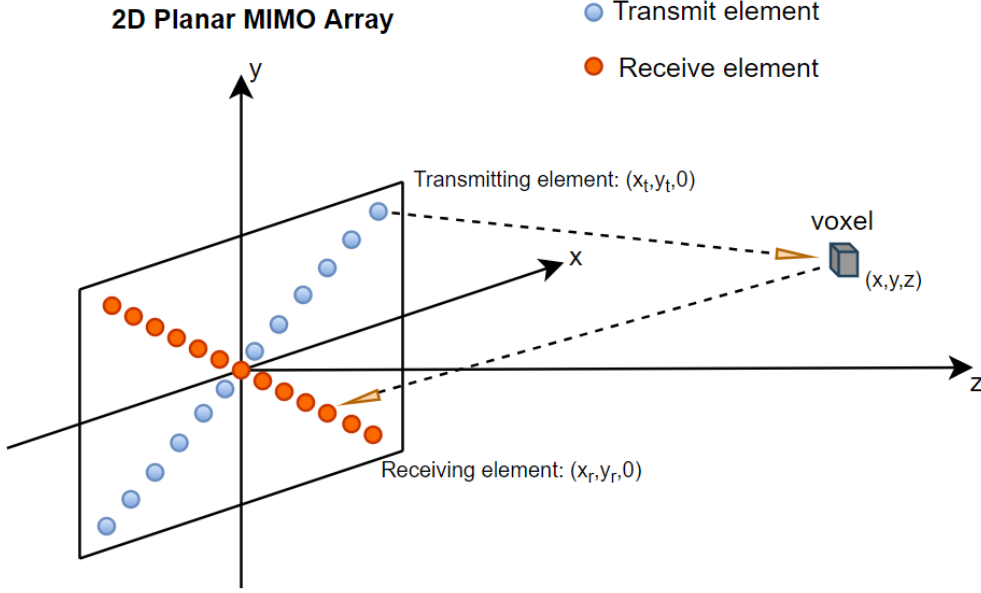
Figure 2.1: Near-field MIMO Radar Imaging System.

can be expressed in the temporal frequency domain [19] as:

$$r\left(x_t, z_t, x_r, z_r, k\right) = \frac{1}{4\pi d_t d_r} s(x, y, z) p(k) \, e^{-jkd_t} e^{-jkd_r} \tag{2.4}$$

Here $p(k)$ denotes the Fourier transform of the transmitted signal, where $k = 2\pi f/c$ denotes the frequency-wavenumber and f denotes the temporal frequency. Then the total received signal $\tilde{r}\left(x_t, z_t, x_r, z_r, k\right)$ due to an extended target is given by

$$\tilde{r}\left(x_t, z_t, x_r, z_r, k\right) = \iiint \frac{1}{4\pi d_t d_r} s(x, y, z) p(k) \, e^{-jkd_t} e^{-jkd_r} \, dx \, dy \, dz \tag{2.5}$$

where $s(x, y, z)$ represents the complex-valued three-dimensional reflectivity distribution of the scene.

Because the measurements will be acquired digitally and image reconstruction will be performed on a computer, a discrete forward model is needed. The continuous forward model in Eq. 2.1 or Eq. 2.5 is converted to a discrete model by replacing the

8

three-dimensional continuous reflectivity function with its discrete representation in terms of voxels. By using lexicographic ordering, the voxel values of the discretized reflectivity function are put into the vector $s \in C^N$. The noisy measurement vector is also denoted by $y \in C^M$, which contains the discrete set of noisy measurements obtained using different transmitter-receiver pairs and frequency steps. Then the linear relation between the complex-valued image vector s and the measurement vector y can be expressed as follows:

$$y = As + w \tag{2.6}$$

where $A \in C^{M \times N}$ is the observation matrix and w is the noise vector. In general, the observation matrix A is rectangular. The total number of rows in A is the length, M, of the measurement vector, which is equal to the multiplication of the number of transmit and receive antennas, and used frequency steps. The number of columns in A is equal to the number of voxels, N.

Using the frequency-domain model given in Eq. 2.5, the (m, n)th element of the observation matrix, representing the contribution of the nth voxel to the mth measurement, can be expressed [84] as:

$$A_{m,n} = \frac{p\left(k_m\right) e^{-jk_m d_{t_m}^{(n)}} e^{-jk_m d_{r_m}^{(n)}}}{4\pi d_{t_m}^{(n)} d_{r_m}^{(n)}} \tag{2.7}$$

Here the measurement index m indicates the locations of the transmitting and receiving antennas, as well as the frequency, k_m, used in this measurement. Moreover, $d_{t_m}^{(n)}$ and $d_{r_m}^{(n)}$ respectively represent the distances from the center of the nth voxel to the transmitter and receiver used in the mth measurement. Note that, for the sampled reflectivity distribution, the voxel size is chosen based on the desired down-range and cross-range resolutions of the MIMO imaging system.

2.2   Image Reconstruction Problem

As in any computational imaging system, we need to solve an image reconstruction problem after mathematically relating the unknown reflectivity distribution of

the scene to the radar measurements as in Eq. 2.6. This requires solving a linear inverse problem. In the inverse problem faced here, the goal is to estimate the unknown 3D reflectivity image, s, from the noisy radar measurements, y. There are different type of approaches for this purpose.

An *analytical reconstruction* approach can be used to solve the corresponding inverse problem. *Traditional direct inversion methods* form the first class of the existing analytical methods. The back-projection method can be given as an example of this class. These traditional methods have generally low computational complexity but as a drawback, they can not offer state-of-the-art reconstruction performance. Also, they are not robust to noise as they directly aim to solve the equation that models the observations (such as Eq. 2.5 or 2.6) without considering the effect of noise and exploiting prior information. *Regularized iterative reconstruction methods* on the other hand incorporates additional prior knowledge (such as sparsity prior) into the reconstruction process. Total variation-based reconstruction can be given as an example of this class of methods. This second class of the analytical methods provides better reconstruction quality but with higher computational cost.

Recently, *reconstruction techniques that exploits deep learning* have emerged as an alternative to analytical methods. These methods are shown to simultaneously achieve high reconstruction quality and low computational cost for various inverse problems in imaging. The existing deep learning-based approaches in the literature can be grouped into three main classes: 1) learning-based direct inversion, 2) plug-and-play regularization, and 3) learned iterative reconstruction based on unrolling. However, in the context of near-field MIMO radar imaging, deep learning-based reconstruction methods have not been studied much in the literature. This thesis aims to develop deep learning-based fast reconstruction methods for near-field MIMO radar imaging and investigate the respective merits and drawbacks of these developed methods in comparison to the existing analytical methods in the literature.

CHAPTER 3

RELATED WORK

The theory and the implementation of the related reconstruction methods are explained in detail in this chapter. First analytical reconstruction methods for near-field MIMO radar imaging are reviewed in Section 3.1. Then in Section 3.2 deep learning-based reconstruction methods for general inverse problems in imaging are examined. The goal is to discuss the merits and drawbacks of the existing methods in the literature and the motivation of the developed methods in this thesis.

3.1 Analytical Reconstruction Methods for Near-field MIMO Radar Imaging

Analytical reconstruction methods for near-field MIMO radar imaging can be classified into two groups: 1) traditional direct inversion methods, and 2) regularized iterative reconstruction methods.

3.1.1 Traditional Direct Inversion Methods

Traditional direct inversion methods are the first class of analytical methods. These methods obtain a direct solution for the equation system that models the observations and do not exploit any prior information. This generally involves back projecting the measurements to the object plane using the adjoint of the scattering operator and applying a filter-like operation [12–17].

There is plenty of work for classical methods in the context of near-field radar imaging. These methods are implemented either in the time domain or in the frequency domain. For monostatic radar systems with collocated transmitter and receiver an-

tennas, commonly used methods are backprojection (also known as diffraction stack migration or delay-and-sum algorithm), Kirchoff migration, range migration methods, and their variants [2, 5, 20]. These direct inversion methods have also been extended for MIMO (multistatic) arrays with spatially distributed transmit and receive antennas [3, 15, 18, 19, 21–23]. The multistatic imaging configuration makes the FFT-based imaging more challenging and requires a multidimensional interpolation process [7, 19]. These traditional methods generally have low computational complexity but as a drawback, they can not offer state-of-the-art reconstruction performance. The reconstruction quality and resolution degrade in the presence of noise or when we have limited data (as acquired with sparse MIMO arrays).

In our experiments, the backprojection (BP) algorithm is chosen from these traditional methods for performance comparison. Considering the discrete forward model in Eq. (2.6) with constant $p(k)$, the three-dimensional reflectivity distribution of the scene can be reconstructed using the BP algorithm as follows [15, 85]:

$$\hat{s}_n = \sum_m y_m \, e^{jk_m d_{t_m}^{(n)}} \, e^{jk_m d_{r_m}^{(n)}} \tag{3.1}$$

Here $\hat{s}$ is the reconstructed complex-valued vector for the sampled 3D reflectivity image and y is the available measurement vector. That is, $\hat{s}_n$ represents nth voxel of the unknown reflectivity field of the scene and y_m is the mth measurement. The measurement index m indicates the locations of the transmitting and receiving antennas, as well as the frequency, k_m, used in this measurement. Moreover, $d_{t_m}^{(n)}$ and $d_{r_m}^{(n)}$ respectively represent the distances from the center of the nth voxel to the transmitter and receiver used in the mth measurement.

3.1.2 Regularized Iterative Reconstruction

Different than direct inversion methods, regularized iterative reconstruction methods incorporate additional prior information (such as sparsity) into the reconstruction process to eliminate uniqueness and noise amplification issues arising from limited data and measurement noise. Motivated by compressed sensing theory [25–27], sparsity-based reconstruction is the most commonly used regularization approach and

has been widely studied in various imaging problems [86–91], including radar imaging [28–35], both for monostatic or far-field imaging settings [30, 34, 36–46], as well as for multistatic and near-field settings [29, 33, 47–52].

In the inverse problem for MIMO radar imaging, the goal is to reconstruct the unknown reflectivity field of the scene, s, from the radar measurements, y. Because scene reflectivity has correlation along both range and cross-range directions, incorporating a sparsity prior can improve the reconstruction quality. To enforce sparsity [87, 92, 93], the inverse problem is formulated as the following regularized least-squares (LS) problem:

$$\min_{s} \|y - As\|_2 + \alpha^2 \|\Phi s\|_1 \tag{3.2}$$

Here Φ represents the sparsifying transform for the reflectivity field of the scene and α denotes the regularization parameter. It is well-known that finite-difference operation (i.e. discrete gradient operator) provides a good sparsifying transform for extended targets. This special case is known as total-variation (TV) regularization.

Various sparsity-based reconstruction algorithms have been developed to solve the optimization problem in Eq. 3.2 for near-field MIMO radar imaging. These algorithms are generally adopted from sparsity-based reconstruction algorithms developed for two-dimensional image restoration problems and mainly differ from each other in their convergence rate. For example, Cheng et al. [29] adopted the split augmented lagrangian shrinkage algorithm (SALSA) [88] and Oktem [32] adopted the half-quadratic regularization approach [94, 95]. There have been also some efforts to reduce the computational cost and memory usage of such sparsity-based reconstruction algorithms by exploiting the special structure of the forward problem [33, 52]. Although sparsity-based methods provide better reconstruction quality than the traditional direct inversion methods, they suffer from higher computational cost and memory usage due to their iterative nature and requiring computation of the forward scattering operator and its adjoint at every iteration. They also require parameter tuning to achieve good reconstruction under different observation scenarios. These aspects are undesirable in real-time applications.

In our experiments, for performance comparison, we use a sparsity-based algorithm

[32] which is a fixed-point iterative approach [94] based on half-quadratic regularization [95]. As sparsifying transform, total variation is chosen since it provides good reconstruction quality for extended targets. The implementation of this sparsity-based algorithm is described in detail below.

First a smooth approximation is applied to the regularization term in Eq. 3.2 to make l_1 norm differentiable, which results in the following minimization problem [32]:

$$\min_s \|y - As\|_2 + \alpha^2 \sum_{i=1}^{3N} \sqrt{|[\Phi s]_i|^2 + \beta} \tag{3.3}$$

This problem is solved with a fixed-point algorithm by alternately calculating the matrix $W(s)$ using Eq. (3.4) and updating the solution s using Eq. (3.5) as follows:

$$W(s_l) = \text{diag}\left(\frac{1/2}{\sqrt{|[\Phi s_l]_i|^2 + \beta}}\right) \tag{3.4}$$

$$s_{l+1} = \left(A^H A + \alpha^2 \Phi^H W(s_l) \Phi\right)^{-1} A^H y \tag{3.5}$$

Here l represents the iteration count. For optimum parameter selection, we choose α parameter that optimizes the average reconstruction performance of the test dataset. Based on this, α is selected as 25 in all experiments.

3.2 Deep Learning-based Reconstruction Methods for Inverse Problems in Imaging

In the literature, deep learning-based methods are grouped into three main classes: 1) learning-based direct inversion, 2) plug-and-play regularization, and 3) learned iterative reconstruction based on unrolling.

3.2.1 Learning-based Direct Inversion Methods

Learning-based direct inversion methods aim to perform the reconstruction directly from the measurements using deep neural networks [12, 57]. The observation model may not be used in this type of approach. That is, the neural network can be trained to learn the direct mapping from the measurement space to the reconstruction space solely using training data. For example, Kulkarni et al. [96] developed a non-iterative fast algorithm to reconstruct images from compressively sensed random measurements. In the end, the developed approach achieved the best reconstruction performance with lower computational complexity compared the to existing iterative CS reconstruction algorithms. However, these type of methods can not provide successful results whenever the observation model is complex, the unknown image does not look alike observations, or there is not much training data available.

For this reason, commonly an efficient analytical inversion method is first performed and provided to the network as a warm start. Subsequently a deep neural network is employed to improve this intermediate reconstruction. This type of physics-based learning approaches are applied to various linear inverse problems in imaging such as noise removal [57], deconvolution [58], super-resolution [59, 60], and tomography [12]. It has also been applied to nonlinear inverse problems such as phase-retrieval [61–64]. These show the reconstruction capability of this type of approach since they provided the state-of-the-art reconstruction performance in various inverse problems. Another advantage of learning-based direct inversion methods is their low computational complexity due to their feed-forward (non-iterative) nature. Their disadvantage is that the performance highly depends on the training data since the learning mostly relies on this.

3.2.2 Plug-and-Play Regularization

In plug-and-play methods, the key idea is to replace the hand-crafted analytical priors in model-based regularized reconstruction methods with data-driven deep priors. A deep prior is first learned from training data and then utilized for the regularization of a model-based iterative method. For this purpose, commonly variable splitting is ap-

plied to the cost function involving the regularization term (as in Eq. 3.2). In this way, the optimization problem is converted to simpler sub-problems. The sub-problem that involves the regularization term can be viewed as a noise removal (denoising) problem. This problem can be solved using an appropriate denoiser. This idea forms the basis of the plug and play regularization [97, 98]. Since deep neural networks have been shown to achieve the state-of-the-art performance in denoising problems in recent years, deep neural network-based denoisers have become the first choice for plug-and-play regularization. This type of approach has been applied to various linear inverse problems in imaging such as deconvolution, super-resolution, and image inpainting problems [65–67]. It has also been applied to nonlinear inverse problems such as phase-retrieval problems [68].

For example, Zhang et al. [65] developed a deep prior-based iterative approach for two-dimensional linear inverse problems such as denoising, deconvolution, and super-resolution. In this approach, the linear inverse problem is formulated as follows:

$$x = \arg\min_x \frac{1}{2}\|y - Ax\|_2^2 + \lambda\Phi(x) \tag{3.6}$$

Here the solution minimizes a cost function composed of a data fidelity term and a regularization term with the trade-off parameter λ. The data fidelity term ensures that the solution is consistent with the forward model, while the regularization term enforces that the solution matches the prior information like sparsity in some transform domain. The fidelity term and regularization term can be decoupled into separate sub-problems with the aid of variable splitting methods. Half quadratic splitting (HQS) method is used by [65] to obtain the following sub-problems:

$$x^{l+1} = \arg\min_x \|y - Ax\|_2^2 + v\left\|x - z^l\right\|_2^2 \tag{3.7}$$

$$z^{l+1} = \arg\min_z \frac{v}{2}\left\|z - x^{l+1}\right\|_2^2 + \lambda\Phi(z) \tag{3.8}$$

The inverse problem can then be solved by alternately solving these sub-problems. The first subproblem is a regularized least-squares problem which has a direct solution and can be expressed in closed form. Based on Bayesian estimation, the second sub-problem corresponds to a denoising problem. A gaussian denoiser can be used to

solve this problem. An alternative way of solving this denoising problem is to train a network for the denoising task and then use this deep denoiser. In [65], ResNet architecture is used for solving the denoising sub-problem. The reconstruction is performed by alternating between the denoising and data fidelity steps as shown in Fig. 3.1. Although plug-and-play methods often perform better than learning-based direct-
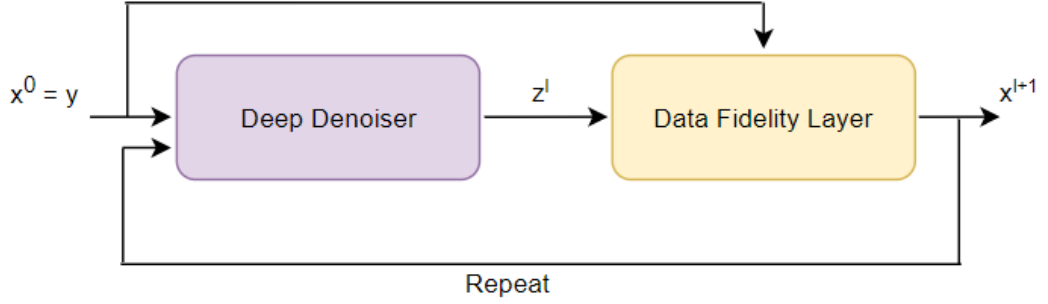
Figure 3.1: The plug-and-play regularization method.

rect inversion methods, these approaches may require high memory usage and high computational complexity due to their iterative nature and requiring computation of the forward scattering operator and its adjoint at every iteration.

3.2.3   Learned Iterative Reconstruction Based on Unrolling

In unrolling-based learned iterative reconstruction methods, an iterative reconstruction method with deep prior is unrolled into an end-to-end trainable network [69]. For example, after fixing the iteration number L, the algorithm described by Eq. (3.7) and (3.8) can be unrolled into an end-to-end trainable network as shown in Fig. 3.2. This type of approach is applied to various imaging problems [69–71]. These methods are also utilized to learn the optimum regularization parameters [70] or the proximal operator [71] from the training data.   Similar to plug-and-play methods, unrolled reconstruction methods may have high computational complexity if they involve computation with large sensing matrices to evaluate the forward scattering operator and its adjoint. Unlike plug-and-play and direct inversion methods, unrolling-based ap-
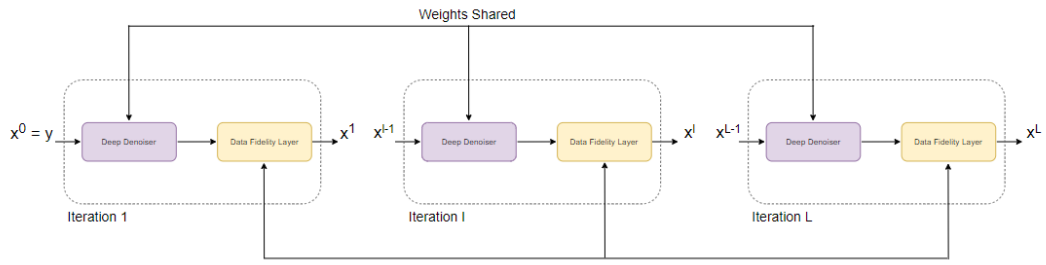
Figure 3.2: End-to-end learned reconstruction with unrolling.

proaches use the forward model during training, which comes with the cost of increased training time and complexity. But this generally provides the advantage of better performance.

CHAPTER 4

DEVELOPED DEEP LEARNING-BASED RECONSTRUCTION METHODS

The developed deep learning-based methods for near-field MIMO radar imaging are presented in this chapter. The main goal is to achieve high image quality with low computational cost so that the developed method can be used in real-time applications. Both approaches are based on learned direct reconstruction. The first approach is a physics-based learning approach. It has a two-staged structure that consists of an adjoint operation stage followed by a 3D U-Net architecture for refinement. The second approach replaces the adjoint operation stage with a fully connected neural network for comparison. Hence this approach aims to perform the reconstruction directly from the radar measurements using only deep neural networks and does not use the physics-based model. The details of these two DNN-based approaches are presented.

4.1 First Approach: DNN-based Two-Stage (Deep2S) Reconstruction

The first approach has two consecutive stages as shown in Fig. 4.1 where the first stage performs an analytical computation whereas the second stage is a neural network. In the first stage, the adjoint operation is used to back project the measurements to the reconstruction space using the observation model of the system. This intermediate result is worse compared to a sparsity-based reconstruction like total-variation but the adjoint operation has the benefit of fast computation due to its non-iterative nature. Subsequently, a deep neural network is employed to improve this intermediate result. This second stage employs a DNN which is trained to convert the back-projected measurements to a suitable magnitude-only reflectivity image. Although

First Stage: Analytical Computation

Complex-valued Measurements → Adjoint → Magnitude $|\cdot|$ → 3D Intermediate Reconstruction

Second Stage: Neural-network based Computation

3D U-Net → The Final Reconstruction
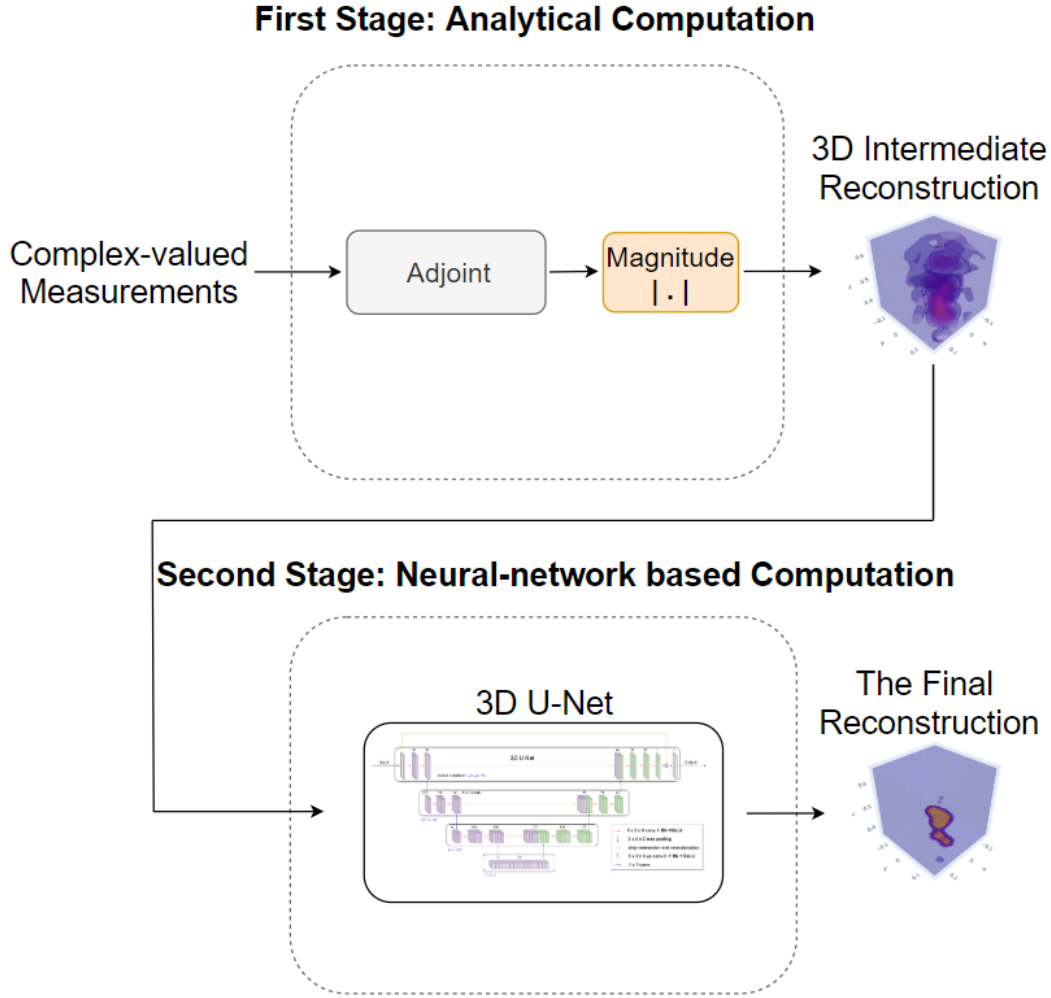
Figure 4.1: Architecture of DNN-based Two-Stage Reconstruction Approach.

scene reflectivities are complex-valued, in most applications they have random phase nature, and as a result it is generally sufficient to reconstruct the magnitude of the reflectivity distribution. Because of this, the DNN is trained to improve the reflectivity magnitudes obtained from the first stage. As DNN, a modified 3D U-net architecture is used to jointly exploit the correlations along both range and cross-range directions.

For training, the simulated radar measurements are first passed through the adjoint operation stage. The magnitude of the 3D reflectivity image obtained with the adjoint operation is then input to the second stage. The DNN in the second stage is trained using these reflectivity magnitudes together with the corresponding ground-truth reflectivity magnitude of the scene, which form our training input and output respectively. For the training procedure, a synthetically generated dataset is used.

The generation of this synthetic dataset is explained in detail in Section 5.1.2.

After training, we can process our radar measurements in the test dataset with the proposed method to reconstruct the reflectivity magnitude of the unknown scene. In what follows, we provide the details of each stage in our approach.

4.1.1 First Stage: Adjoint Operation

Since it is generally a difficult task for a network to learn the direct mapping from the measurement space to the reconstruction space, we first apply the adjoint operator to the measurements to provide the network in the second stage a warm start. The adjoint operation encapsulates the physical model of the near-field MIMO imaging system and involves an analytical fast computation. This stage simplifies the learning process of the 3D U-Net architecture since it back projects the radar measurements to the object plane. In this way, the network can be trained to improve the reflectivity magnitude obtained from the first stage by directly working in the object plane.

In our experiments, the adjoint operation is applied to the radar measurements by using the hermitian of the system matrix. The hermitian matrix applied on the discrete radar measurements gives us a 3D intermediate result, $\hat{s}$, in the reconstruction space as follows:

$$\hat{s} = A^H y \tag{4.1}$$

4.1.2 Second Stage: 3D U-Net Architecture

For the second stage, we develop a 3D U-Net architecture based on [99] which was originally designed for segmentation. This 3D U-Net architecture can capture the correlation information along both range and cross-range directions of a 3D extended target unlike the 2D U-Net architecture proposed in [99] which can only exploit the correlation in a two-dimensional space. To achieve this, three-dimensional convolution kernels, max pooling, and up-sampling operations are used in the network architecture instead of their two-dimensional counterparts. The 3D U-Net architecture and its details are illustrated in Fig. 4.2. This architecture has an encoding and decoding
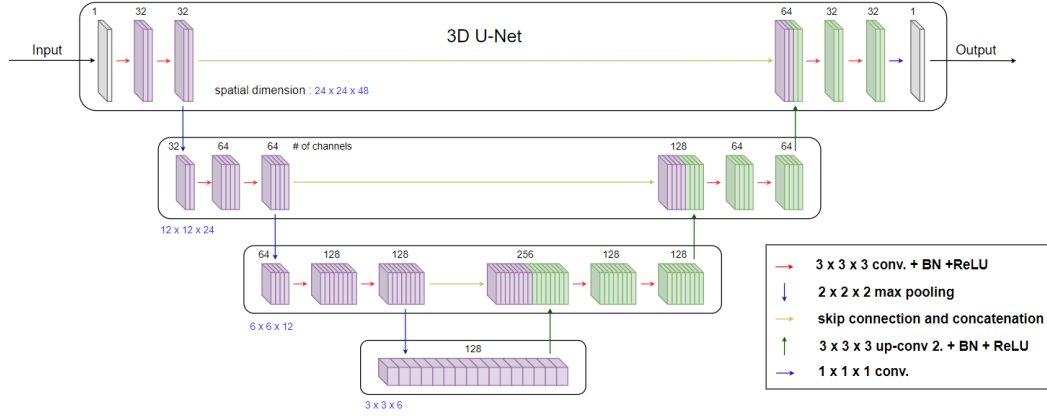
Figure 4.2: Architecture of the 3D U-Net.

path. The encoding path contains the repeated application of $3 \times 3 \times 3$ convolutions, batch normalization (BN), rectified linear unit (RELU), followed by a $2 \times 2 \times 2$ max pooling with strides of two. The decoding path consists of $3 \times 3 \times 3$ upconvolution with strides of two in each dimension, which is followed by a RELU. In the decoding path, there are also concatenations with the cropped feature maps from the encoding path. The size of the input and output of the network is $25 \times 25 \times 49$ voxels in x, y, and z directions.

The 3D U-Net architecture has three properties that will be suited for our imaging problem. Firstly, due to the decoding path, the effective receptive field of the network increases. In our problem, our main purpose is the refinement of the input image. Having a large receptive field over the input image can improve the quality of the output image [100]. We can use encoder-decoder architecture to take the advantage of large receptive fields. Secondly, the 3D U-Net employs multichannel filters. By this way it can better extract the feature maps of its input. This increases the dimension of the latent representation of our input images, which increases the expressive power of the network [101]. Thirdly, the 3D U-Net architecture can capture the correlation along both range and cross-range directions of the three-dimensional target, unlike the existing approaches in the literature. This 3D U-Net architecture is used in the second stages of the developed DNN-based algorithms.

In the end, a feed-forward approach is obtained while incorporating the physics-based knowledge of the MIMO imaging system through the usage of the adjoint of the

system matrix. The approach has low computational complexity as desired.

4.1.3 Computational Complexity

For an image size of $N \times M \times L$ and measurement size of $T \times R \times F$ where T is the number of transmitter antennas, R is the number of receiving antennas, and F is the number of frequency steps, the computational cost of the adjoint operation is $O(N\,M\,LT\,RF)$. The main operations in the 3D U-Net architecture are 3D convolutions, batch normalization, application of the RELU function, max pooling, and upconvolution. The computational cost of the network is dominated by the 3D convolution operations, which are performed in the 3D spatial domain. When the kernel size is $K \times K \times K$, number of filter per layer is S with Q layers, the computational cost of the network is $O(N\,M\,L\,K^3\,S^2\,Q)$ [102]. Overall, the proposed algorithm also has $O(N\,M\,L\,K^3\,S^2\,Q)$ complexity. The computational cost of the algorithm is dominated by the 3D U-Net.

4.2 Second Approach: DNN-based Direct Inversion (DeepDI)

Direct inversion from the measurement domain to the reconstruction domain is also possible by using a proper network. For comparison, a second approach is developed which replaces the earlier adjoint stage with a fully connected neural network. In this two-staged structure, the reconstruction is performed directly from the radar measurements using only neural networks, and the observation model is not used. These neural networks are trained end-to-end to learn the direct mapping between the measurements and the magnitude of the unknown reflectivity image. This algorithm is inspired by the CNN-based approach proposed in [96] for the CS recovery problem. Hence the developed algorithm contains two consecutive stages: a fully connected layer followed by a 3D U-Net architecture. The only difference from the first reconstruction approach (Deep2S) is that we use neural network-based computation in the first stage of the reconstruction approach. The overall method is illustrated in Fig. 4.3.

First Stage: Neural-network based Computation

Concatenation of Real
and Imaginary Part

Complex-valued
Measurements

Re

Im

Fully Connected
Layer

Second Stage: Neural-network based Computation

3D U-Net

The Final
Reconstruction
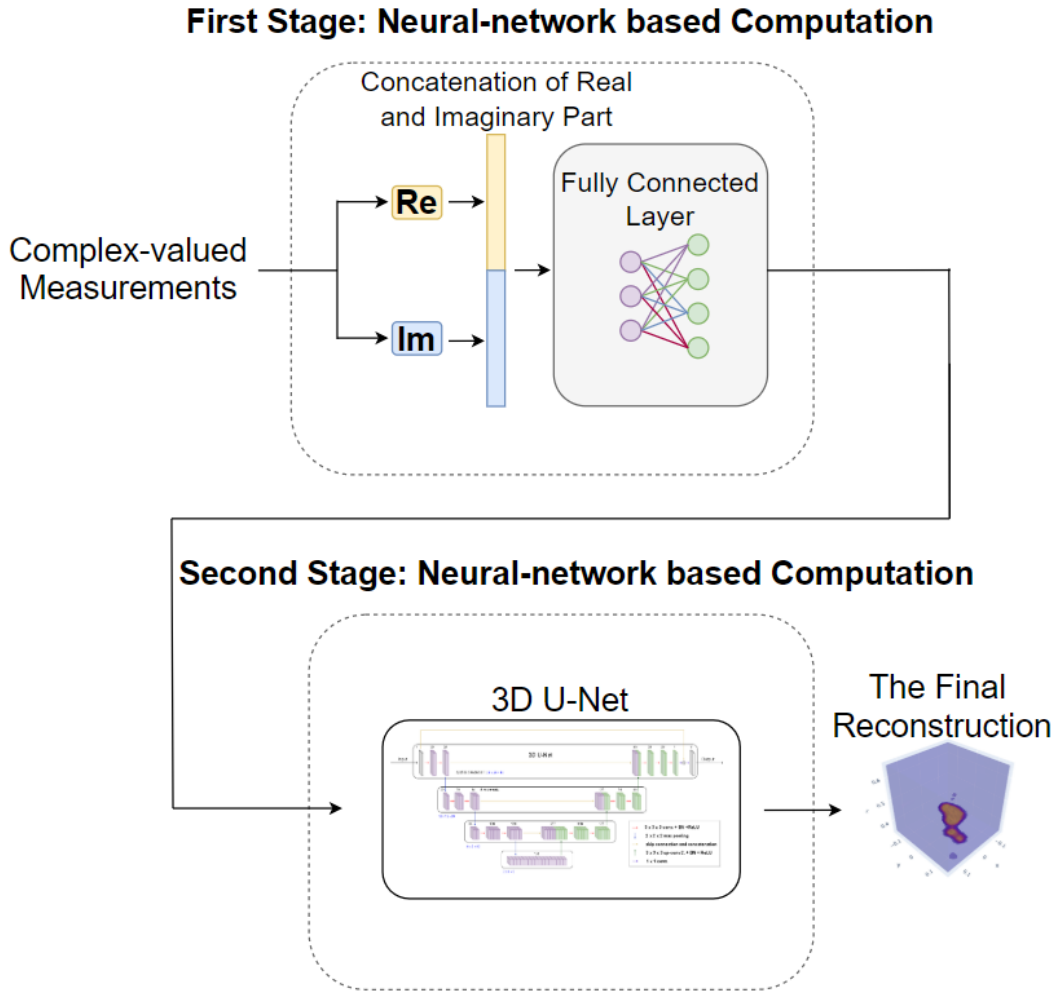
Figure 4.3: Architecture of DNN-based Direct Inversion Approach.

4.2.1 First Stage: Fully Connected Layer

The first stage contains a fully connected layer that takes the radar measurements as input and outputs a 3D intermediate reconstruction for the magnitude of the unknown reflectivity image. Since the measurements are complex-valued, the input of the fully connected layer is provided as the concatenation of the real and imaginary parts of the measurements. This stage learns a mapping from the complex-valued measurements to the 3D reconstruction space (which represents the real-valued reflectivity magnitude). Then, the 3D intermediate reconstruction is fed into the second stage.

4.2.2   Second Stage: 3D U-Net Architecture

In the second stage, a deep neural network is employed to improve this 3D intermediate reconstruction. Same 3D U-Net architecture in Fig. 4.2 is used for this purpose. Instead of using a single end-to-end fully connected layer structure, we introduce a second stage which contains the 3D U-Net architecture. The main reason is to increase the reconstruction performance using a CNN structure and decrease the training time of the whole network.

The two DNNs are trained end-to-end. For training, the simulated radar measurements and the corresponding ground-truth reflectivity images are used as the input and output respectively. For training, a synthetically generated dataset is used as explained in Section 5.1.2.

For testing, we can pass our radar measurements into the trained network to directly obtain the reflectivity magnitude of the unknown scene as reconstruction. This approach has similar computational complexity as the first one, but since it does not exploit physics-based knowledge, it can yield poor results.

CHAPTER 5

NUMERICAL RESULTS

Simulation results are presented in this chapter. Firstly, we describe the simulation setting, synthetic dataset generation, and training procedure of the neural networks in Section 5.1. Then we present the performance of the developed methods in comparison with the commonly used analytical methods in Sections 5.2 and 5.3. We also investigate the effect of SNR and different network architectures on the performance in Sections 5.4 and 5.5.

5.1 Preparation for Experiments

In this section, we describe the details of the simulation setting, synthetic scene generation, and the training procedure.

5.1.1 Simulation Setting

The sketch of the used experimental setting is shown in Fig. 2.1. As a sparse MIMO array topology, we consider a Mills Cross array since it is commonly used due to its imaging performance [19, 32], but other alternatives could also be considered [103]. The width of the planar array is 0.3 m, which includes 12 uniformly spaced transmit antennas and 13 uniformly spaced receive antennas along its diagonals in a cross configuration. The target center is located approximately 0.5 m away from the 2D MIMO array. The frequency ranges from 4 to 16 GHz with uniformly sampled steps. In the numerical simulation, the number of frequency steps is selected as 7, 15, and 31 respectively to investigate the performance of the methods at different compression

levels.

For this imaging setting, the theoretical resolution [19] is 1.25 cm in the cross-range directions, x and y, and 2.5 cm in the down-range direction, z. Our goal is to infer the reflectivity image within a cube of size $0.3m \times 0.3m \times 0.3m$, where the voxel size is chosen as $1.25cm \times 1.25cm \times 0.625cm$ in x, y, and z directions, which is half of the expected theoretical resolution in each direction. The cube that we want to infer the reflectivity image contains $25 \times 25 \times 49$ voxels in x, y, and z directions respectively.
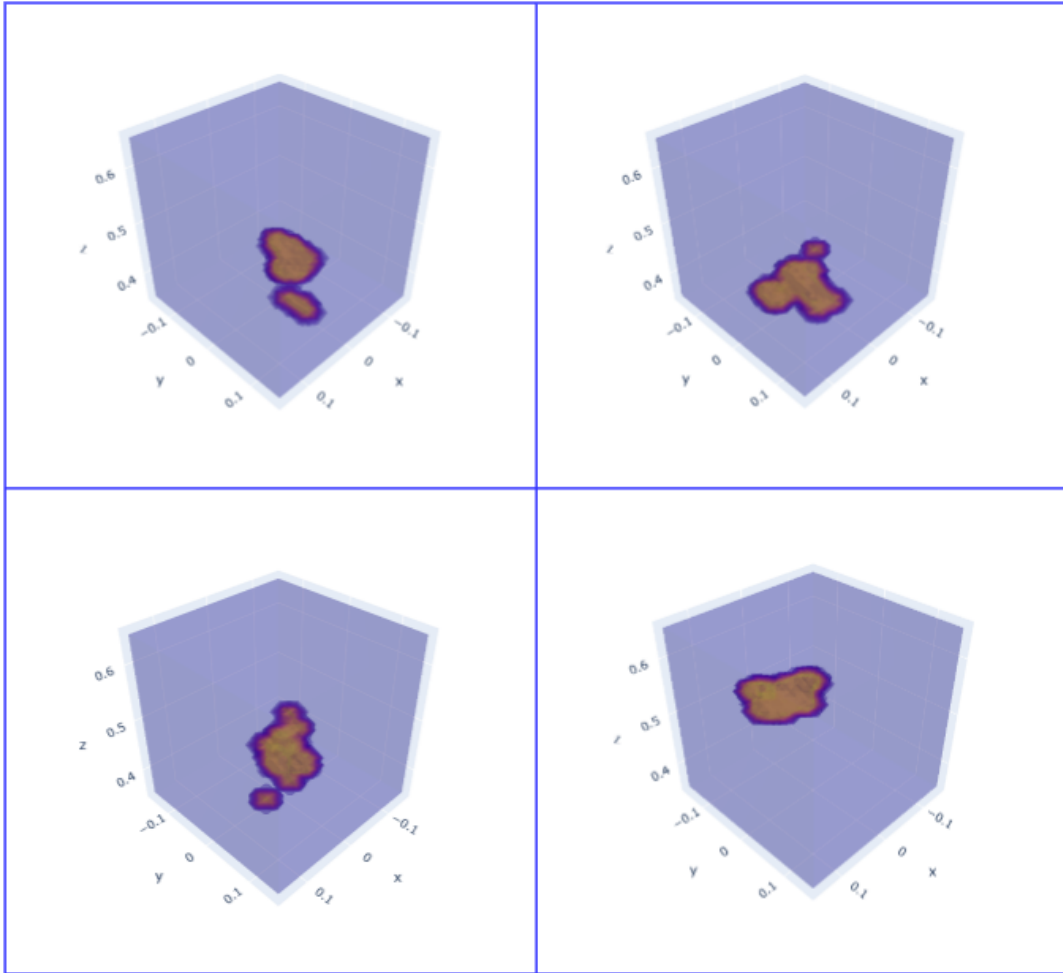
5.1.2 Synthetic Dataset Generation

Figure 5.1: Samples of Synthetically Generated Dataset (The units of x, y, and z-axis are meters).

A real-world dataset is the most appropriate way of training a network. However, a
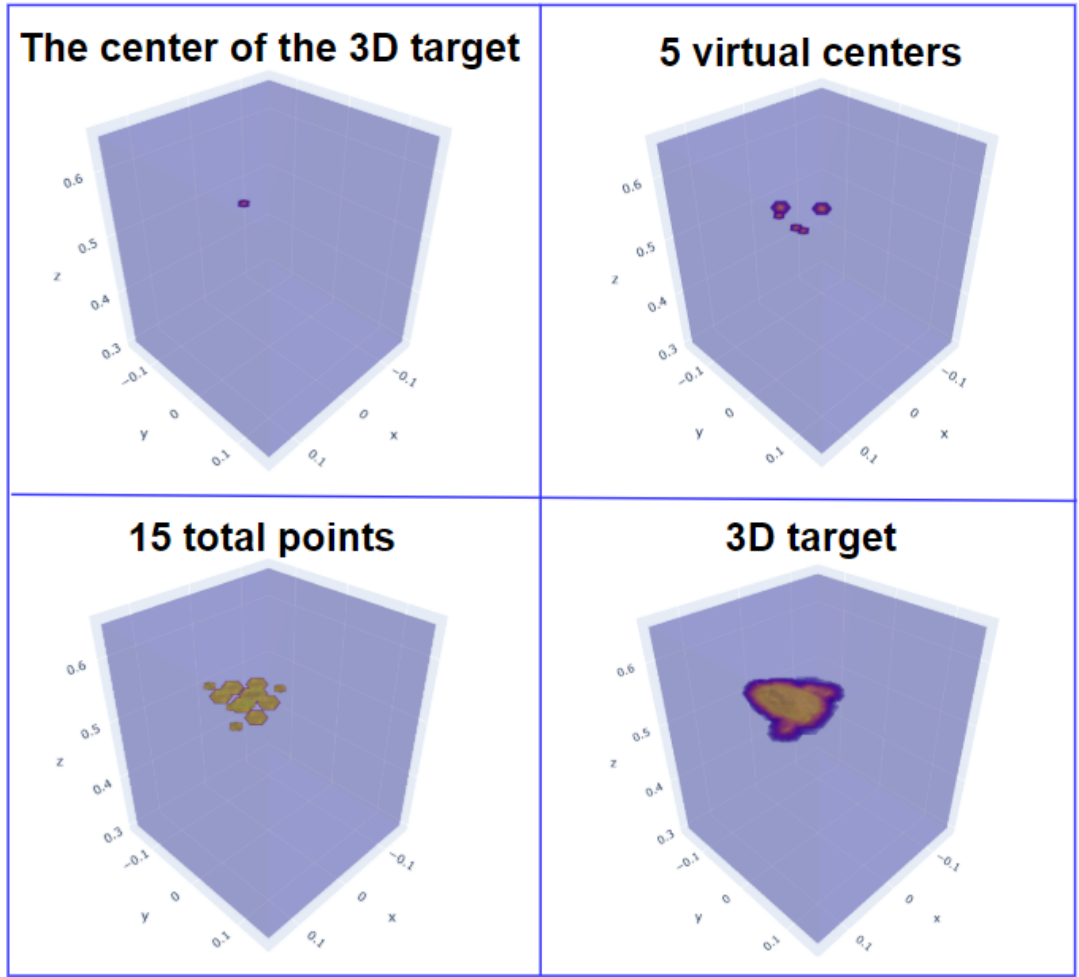
Figure 5.2: Synthetic Dataset Generation Procedure (The units of x, y, and z-axis are meters).

large dataset is often not available in many radar imaging applications. Therefore, the neural networks are trained and tested using a synthetically generated dataset. To enable generalizability to different target scenarios, 3D target images are randomly generated within the cube we want to infer as illustrated in Fig. 5.2. Firstly, the center of the object is randomly chosen from a uniform distribution in the range -0.05 to $0.05m$ for the x and y-axis and, 0.41 to $0.59m$ for the z-axis. Then, around the center of the object 5 virtual centers are chosen according to a gaussian distribution with zero mean and standard deviation of 2. Moreover, for every virtual center, 3 points are generated according to a gaussian distribution with zero mean and standard deviation of 1.5. The variance parameter of the gaussian distributions determines the size of the distributed target. For a large variance, generated distributed targets take up more vol-

ume within the cube compared to the selection of a small variance. For one synthetic scene, we totally have 15 points chosen randomly within the cube. To obtain volumetric objects, these 15 points are passed through a three-dimensional gaussian filter with a standard deviation of 1.3. This is then passed through the sigmoid function which performs the amplitude normalization of the generated three-dimensional targets to a maximum value of 1. In this way, the amplitude values of the generated targets are in the range of zero to one. With this approach, we obtain different three-dimensional objects which spread within the cube from a randomly chosen center.

The training, validation, and test datasets contain 800, 100, and 100 images which were randomly generated in this way.

5.1.3 Training Procedure

The implementation and the training of the networks are done using Jupyterlab 3.1.7 environment and default Adam Optimizer. The training of the networks used in the experiments took about 8 hours for both approaches using NVDIA GeForce RTX 3060. The learning rate is chosen as 10^{-3}, batch size equals 16, loss function is chosen as mean square error (MSE) (l_2 loss), maximum number of epochs are set to 100, and the early stopping criterion is selected as 15 consecutive epochs with no drop in validation loss. Fig. 5.3 shows the learning curve of the 3D U-Net model trained for the Deep2S approach. Fig. 5.4 shows the learning curve of the DeepDI approach.

5.2 Performance Comparison of Different Approaches

In the experiments, we use 3D peak-signal-to-noise (PSNR) and the structural similarity index (SSIM) as quantitative metrics. A higher SSIM and PSNR value corresponds to a better reconstruction or higher similarity between the reconstructed and the ground truth image. SSIM between the three-dimensional images is calculated using [104]. PSNR between the three-dimensional images of size $N \times M \times L$ (ground-truth image s and the reconstructed image $\hat{s}$) is calculated using the following
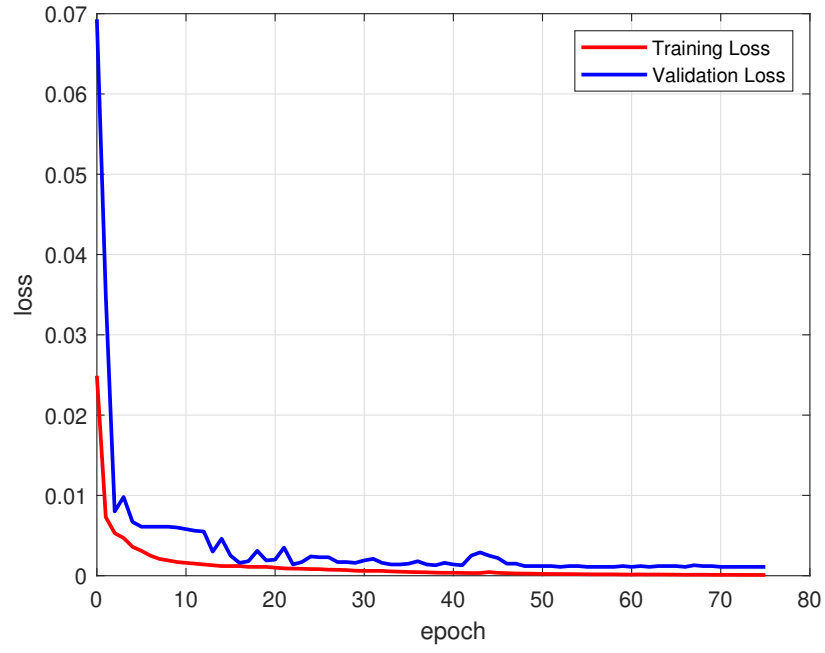
Figure 5.3: The training and validation loss versus epoch for the 3D U-Net model in the Deep2S Approach (when the number of frequency steps is 15 and SNR is 30 dB).
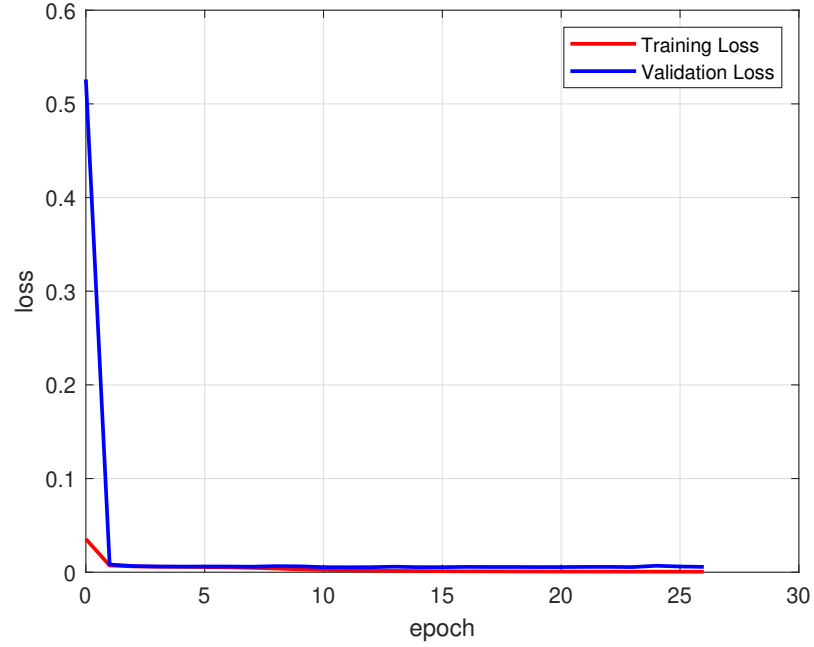
Figure 5.4: The training and validation loss versus epoch for the 3D U-Net model in the DeepDI Approach (when the number of frequency steps is 15 and SNR is 30 dB).
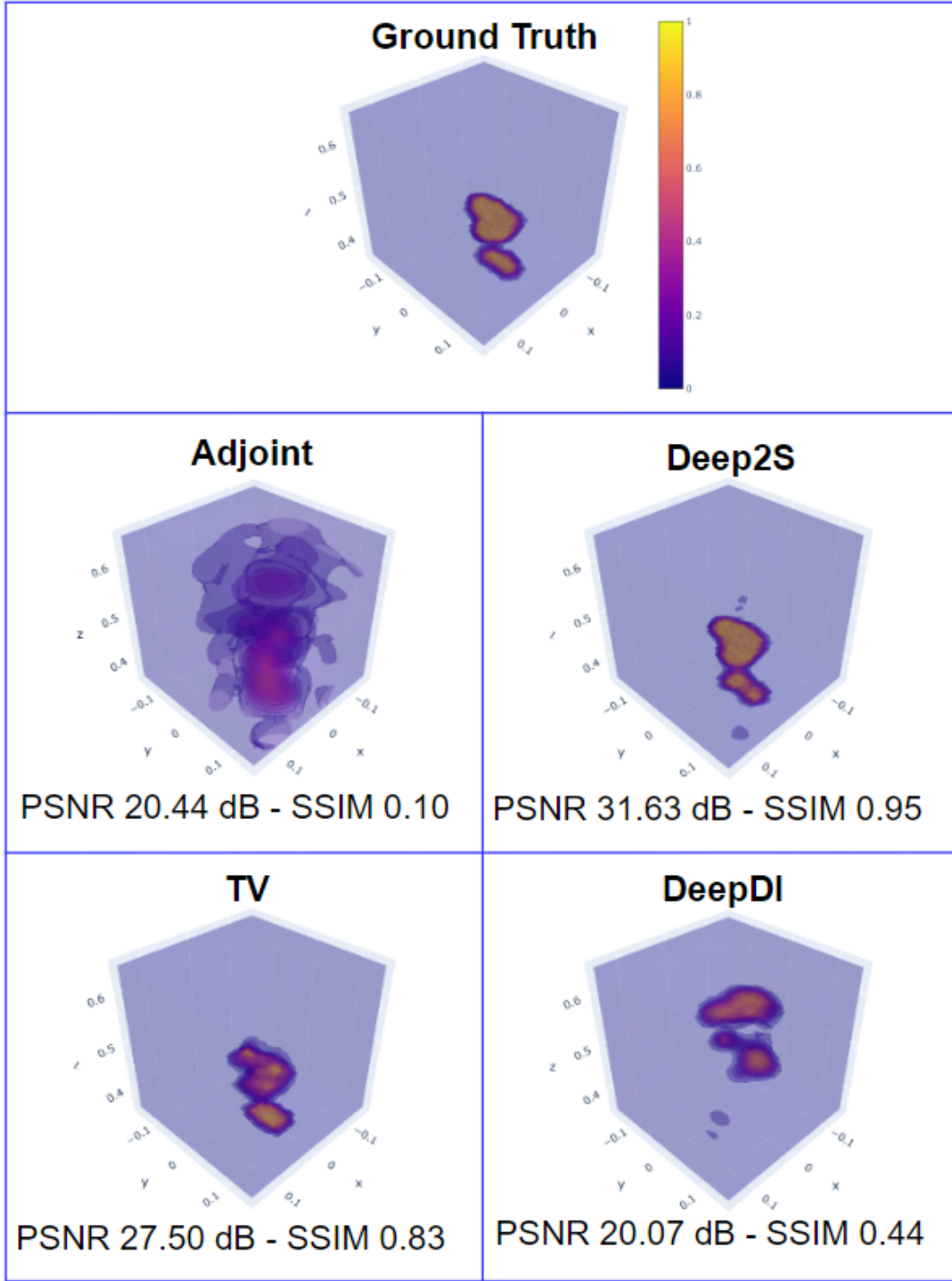
Figure 5.5: Reconstructions of the different algorithms for the first test image of the synthetically generated dataset at 30 dB SNR (Number of Frequency Steps: 15), (The units of x, y, and z-axis are meters).

formula [105]:

$$PSNR = 10 \log_{10} \left(\frac{s_{max}^2}{MSE} \right) \tag{5.1}$$
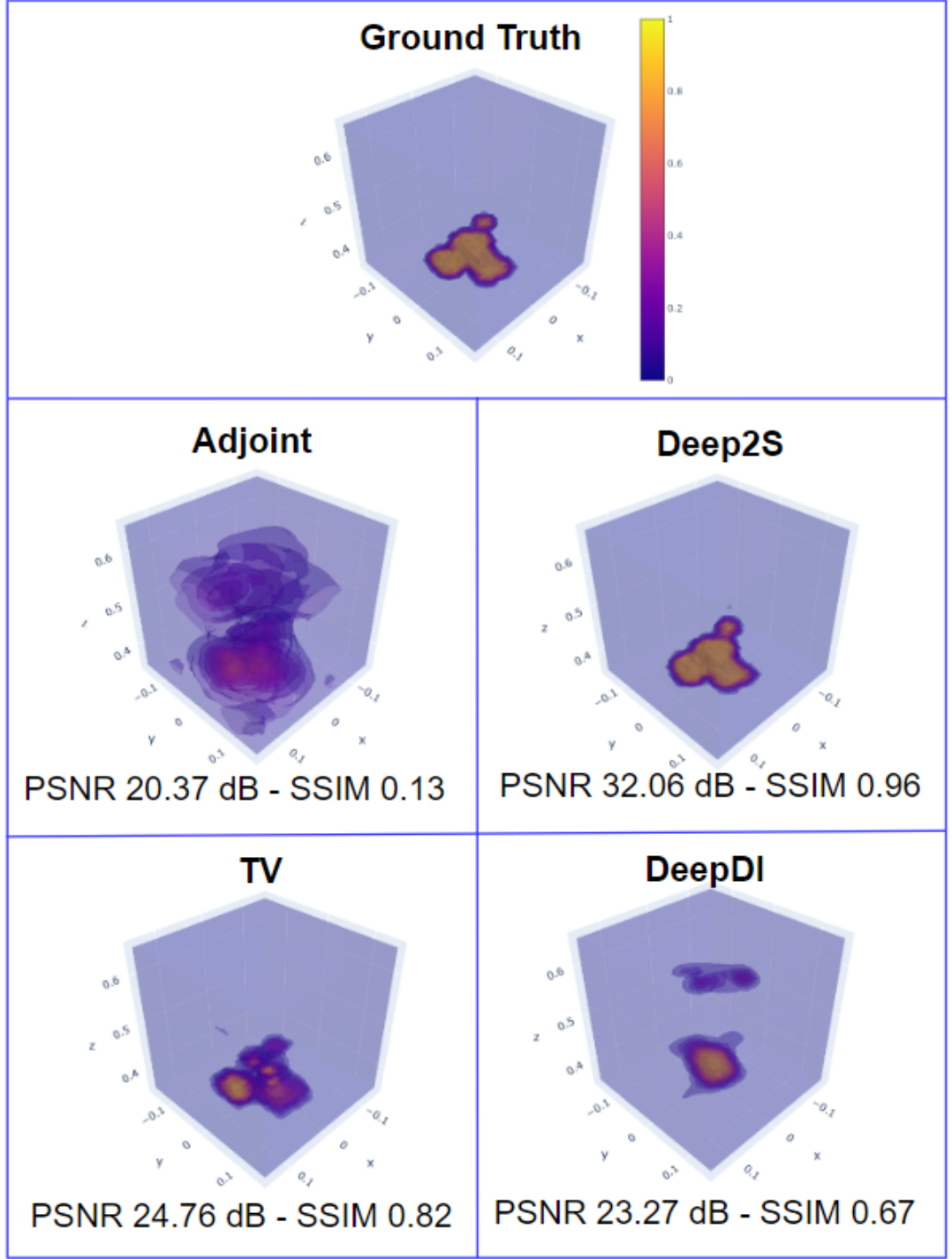
Figure 5.6: Reconstructions of the different algorithms for the second test image of the synthetically generated dataset at 30 dB SNR (Number of Frequency Steps: 15).

where s_{max} is the maximum value of the image voxels (which is 1 in our experiments). MSE is the mean square error, which is calculated by the following formula [106]:

$$MSE = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \sum_{l=0}^{L-1} \frac{(s(n,m,l) - \hat{s}(n,m,l))^2}{N \cdot M \cdot L} \qquad (5.2)$$

For performance comparison, the results are obtained using the adjoint operation, DeepDI, and Deep2S approaches, as well as TV method as we discuss its details in Chapter 3.

Figure 5.7: Reconstructions of the different algorithms for the third test image of the synthetically generated dataset at 30 dB SNR (Number of Frequency Steps: 15).
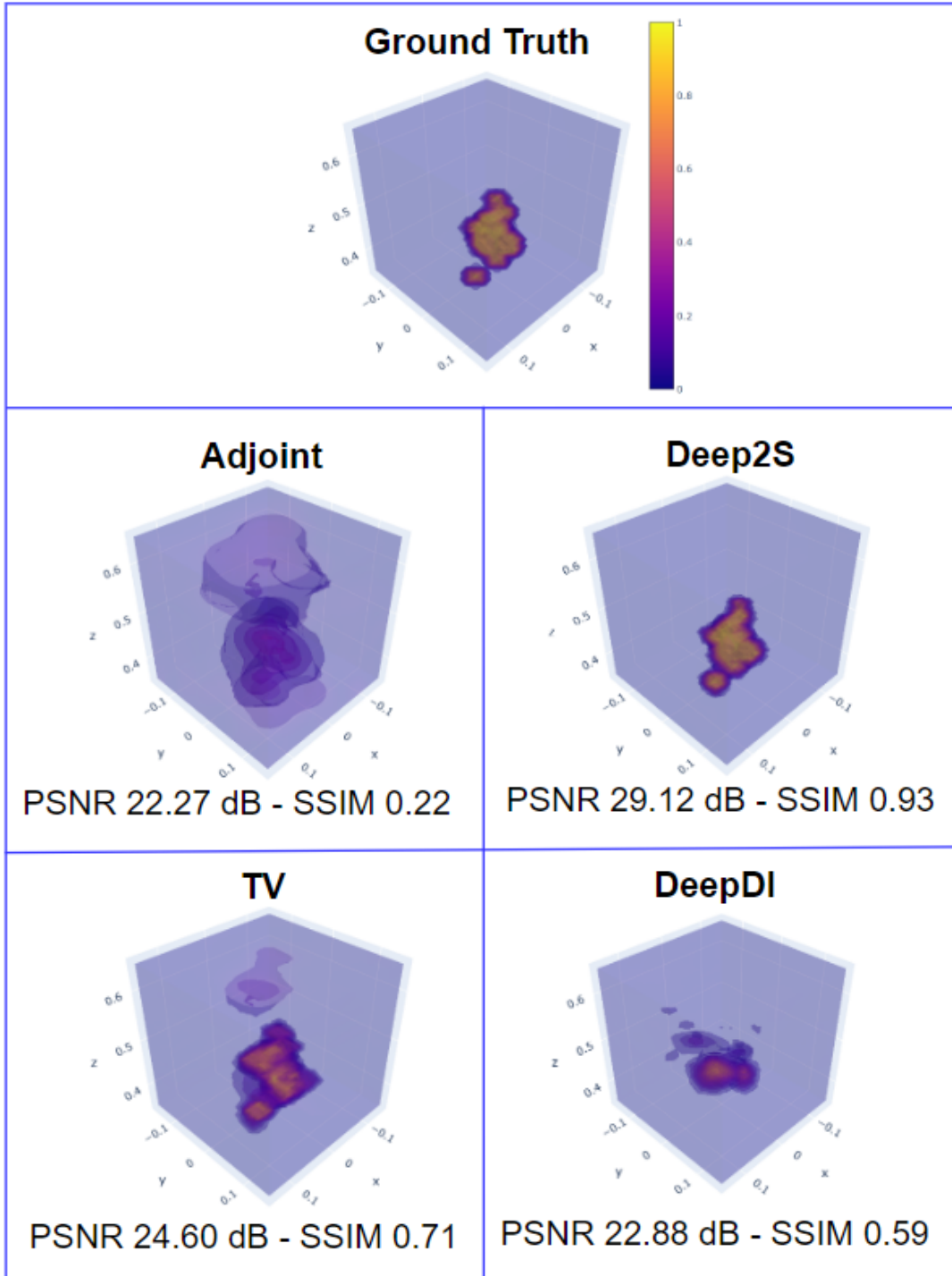
Table 5.1: Average PSNR and SSIM Values for Different Number of Frequency Steps at 30 dB SNR. Best results are shown in bold.

Number of Frequency Steps	Method	PSNR (dB)	SSIM
31	Adjoint	22.95	0.35
	DeepDI	23.25	0.62
	TV	26.25	0.83
	Deep2S	**30.35**	**0.94**
15	Adjoint	21.77	0.19
	DeepDI	23.18	0.61
	TV	25.47	0.74
	Deep2S	**30.39**	**0.94**
7	Adjoint	20.33	0.12
	DeepDI	22.58	0.65
	TV	22.71	0.44
	Deep2S	**29.13**	**0.89**

In Table 5.1, the average reconstruction performance of the algorithms for 100 test images is given for the different frequency steps at 30 dB SNR (where SNR is defined as $10 \log_{10} \left(\frac{\sigma_s^2}{\sigma_w^2} \right)$). The frequency step numbers are increased twice (7, 15, and 31 respectively). Hence the compression ratio is respectively %4, %8, and %16. In all cases, the Deep2S approach outperforms the other approaches in terms of PSNR and SSIM metrics. When the number of frequency steps is decreased, the reconstruction performance of all approaches decreases accordingly because we have fewer measurements to perform the reconstruction. Our problem becomes more and more ill-posed. In the worst case when the number of frequency steps is seven, the Deep2S approach surpasses the other approaches. Deep2S approach gives a PSNR of 29.13 dB and an SSIM value of 0.89 at the average of 100 test images. When the number of frequency steps increases Deep2S approach passes 30 dB PSNR and 0.90 SSIM at the average which gives the best reconstruction performance. Also, increasing the number of frequency steps from fifteen to thirty-one does not make a significant change in the reconstruction performance. As a result, fifteen frequency steps are enough for obtaining the best reconstruction performance for the Deep2S approach.

Figures 5.5, 5.6 and 5.7 show the results of reconstructed images for three samples in the test dataset. In these figures, both the TV algorithm and the Deep2S approach give the best reconstruction performance. These two methods give nearly artifacts-free re-

Table 5.2: Average PSNR and SSIM Values of Adjoint Operation and BP Algorithm at 30 dB SNR (Number of Frequency Steps: 15). Best results are shown in bold.

	BP	Adjoint
PSNR	20.34	**21.77**
SSIM	0.12	**0.19**

construction, while the adjoint operation and DeepDI approach show significant volume artifacts. The adjoint operation and DeepDI approach give poor reconstructions which are not comparable with the TV and the Deep2S approach. In fact, the DeepDI approach could not predict the initial position of the targets which gives worse reconstruction results. Both the TV and the Deep2S approach reduce volume artifacts significantly. Moreover, the Deep2S approach gives a better reconstruction both visually and qualitatively compared to the state-of-the-art TV reconstruction algorithm which have low computational complexity.

Testing two DNN-based algorithms in conditions different from the training dataset examined in the experiment as shown in Figure 5.8. For this purpose, we introduce a 3D target image that is not contained in the synthetically generated dataset. The 3D target image is an ellipsoid that is centered on the cube that we want to infer its reflectivity. The introduced target differs from the training dataset due to the following features: The introduced target image takes up more volume than the training dataset images within the cube. The training dataset contains targets that are shapeless and randomly scattered in the volume. However, the introduced ellipsoid is a bulk volume that covers a large volume within the cube.

Figure 5.8 shows the visual result of reconstructed images for different algorithms at 30 dB SNR (Number of Frequency Steps: 15). The best reconstruction is obtained by the Deep2S approach both visually and qualitatively. Deep2S approach has a robust performance when we introduce 3D target images that differ from the synthetically generated dataset. As a result, the developed approach can be used in different imaging scenarios, although the network is trained on a specific configuration. The adjoint operation, DeepDI approach, and TV reconstruction show significant volume artifacts. Moreover, the TV reconstruction algorithm failed to fill the volume inside of the ellipsoid.
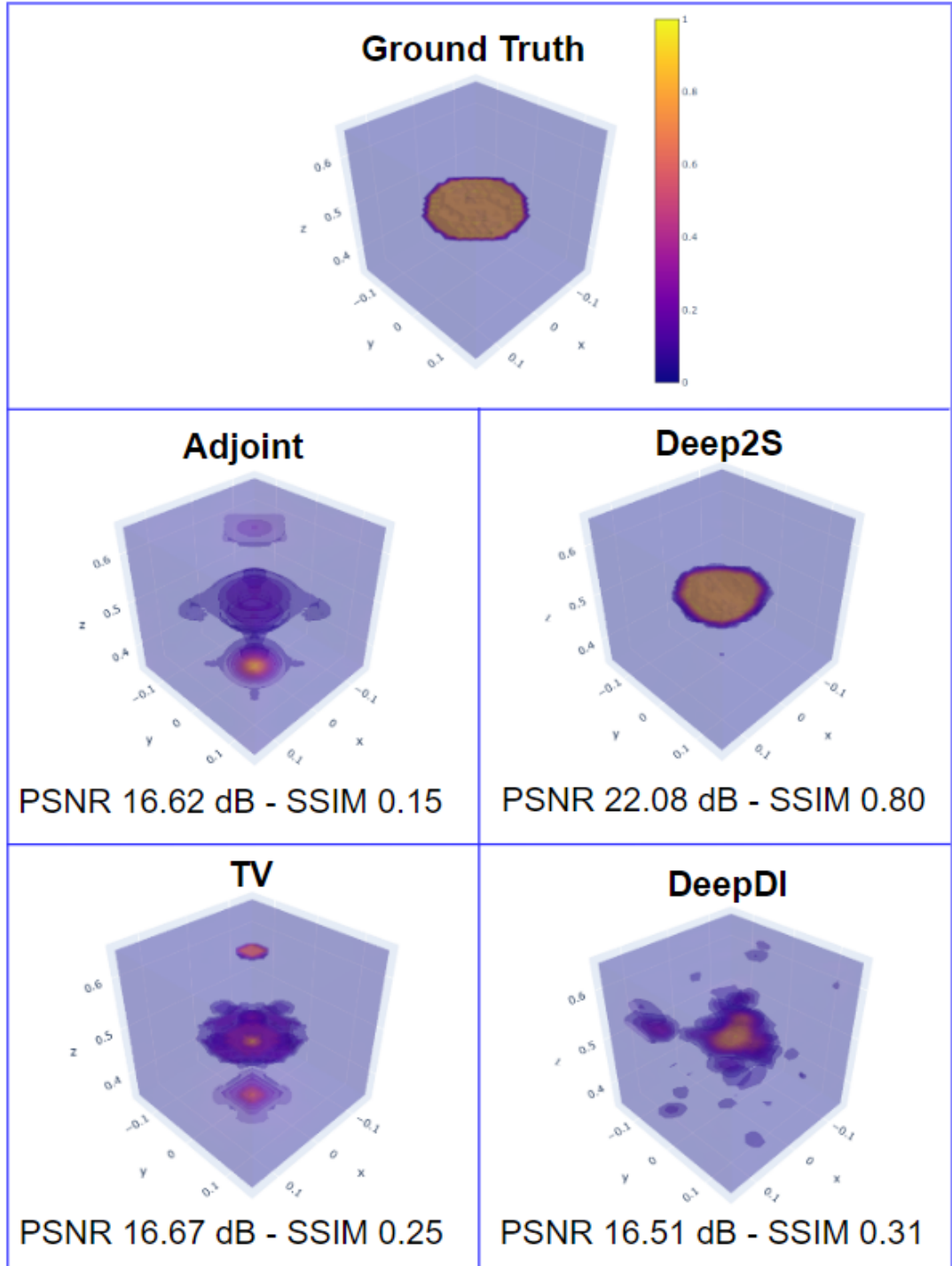
Figure 5.8: Reconstructions of the different algorithms for the ellipsoid test image at 30 dB SNR (Number of Frequency Steps: 15).

In Table 5.2, the average reconstruction performance of the adjoint operation and the classical BP algorithm for 100 test images is given for fifteen frequency steps at 30 dB SNR. The reconstruction performance of the adjoint operation outperforms the clas-

Table 5.3: Average Runtimes for 100 Test Images at 30 dB SNR (Number of Frequency Steps: 15).

Adjoint	DeepDI	TV	Deep2S
32ms	111ms	165.04s	77ms

sical BP algorithm in terms of PSNR and SSIM metrics. As a result, we use adjoint operation as the first stage of the Deep2S approach. The performance comparison also carried out with adjoint operation instead of the classical BP algorithm.

Table 5.3 shows the comparison of average runtimes across different methods for 100 test images. Note that the Deep2S approach and DeepDI approach require less than a second to reconstruct the reflectivity field of a $25 \times 25 \times 49$ scene on the CPU. While the TV reconstruction requires a runtime on the order of minutes. The Deep2S approach and DeepDI approach are three to four orders of magnitude faster than the TV algorithm. Hence the Deep2S approach not only surpasses the other approaches in terms of PSNR and SSIM metrics but also is computationally more efficient than the others except for the adjoint operation which gives poor performance. The Deep2S approach overperforms the other approaches in terms of PSNR and SSIM which have low computational complexity. Moreover, it also gives a better reconstruction visually.

5.3 Performance Comparison with Random Phase

To work in a more realistic scenario that takes into account the complex-valued and random phase nature of scene reflectivities, we add random phase to our synthetically generated dataset. Random phase is added for every voxel from a uniform distribution in the range of $[-\pi, \pi]$. From now on, we will only examine this realistic imaging scenario where we add random phase to our synthetically generated dataset. The training procedure of the networks used in these experiments are done using transfer learning of the existing trained networks used in Section 5.2 where we don't consider random phase. Transfer learning is done by unfreezing the weights of the trained networks used in Section 5.2, and re-trained the networks on the new random phase added data within five epochs.

Table 5.4: Average PSNR and SSIM Values for Different Number of Frequency Steps at 30 dB SNR. Best results are shown in bold.

Number of Frequency Steps	Method	PSNR (dB)	SSIM
31	Adjoint	24.23	0.69
	DeepDI	23.90	0.83
	Deep2S	**28.65**	**0.94**
15	Adjoint	23.69	0.49
	DeepDI	23.37	0.81
	Deep2S	**29.22**	**0.93**
7	Adjoint	22.56	0.26
	DeepDI	22.70	0.73
	Deep2S	**28.36**	**0.93**

Table 5.5: Average PSNR and SSIM Values of Adjoint Operation and BP Algorithm at 30 dB SNR (Number of Frequency Steps: 15). Best results are shown in bold.

	BP	Adjoint
PSNR	23.26	**23.69**
SSIM	0.39	**0.49**

In Table 5.4, the average reconstruction performance of the algorithms for 100 test images is given for the different frequency steps at 30 dB SNR. In all cases, the Deep2S approach outperforms the other approaches in terms of PSNR and SSIM metrics. The reconstruction performance of all approaches decreases when the number of frequency steps is decreased. We have observed that Table 5.4 and Table 5.1 share similar results in terms of reconstruction performance of different algorithms. The general trend is maintained when are dealing with a more realistic imaging scenario.

Figure 5.9 show the reconstructions of the different algorithms for the first test image of the synthetically generated dataset at 30 dB SNR. Deep2S approach gives the best reconstruction performance which gives nearly artifacts-free reconstruction, while the adjoint operation and DeepDI approach show significant volume artifacts.

Lastly, in Table 5.5, the average reconstruction performance of the adjoint operation and the classical BP algorithm for 100 test images is given for fifteen frequency steps at 30 dB SNR. The reconstruction performance of the adjoint operation outperforms the classical BP algorithm in terms of PSNR and SSIM metrics. As a result, we use
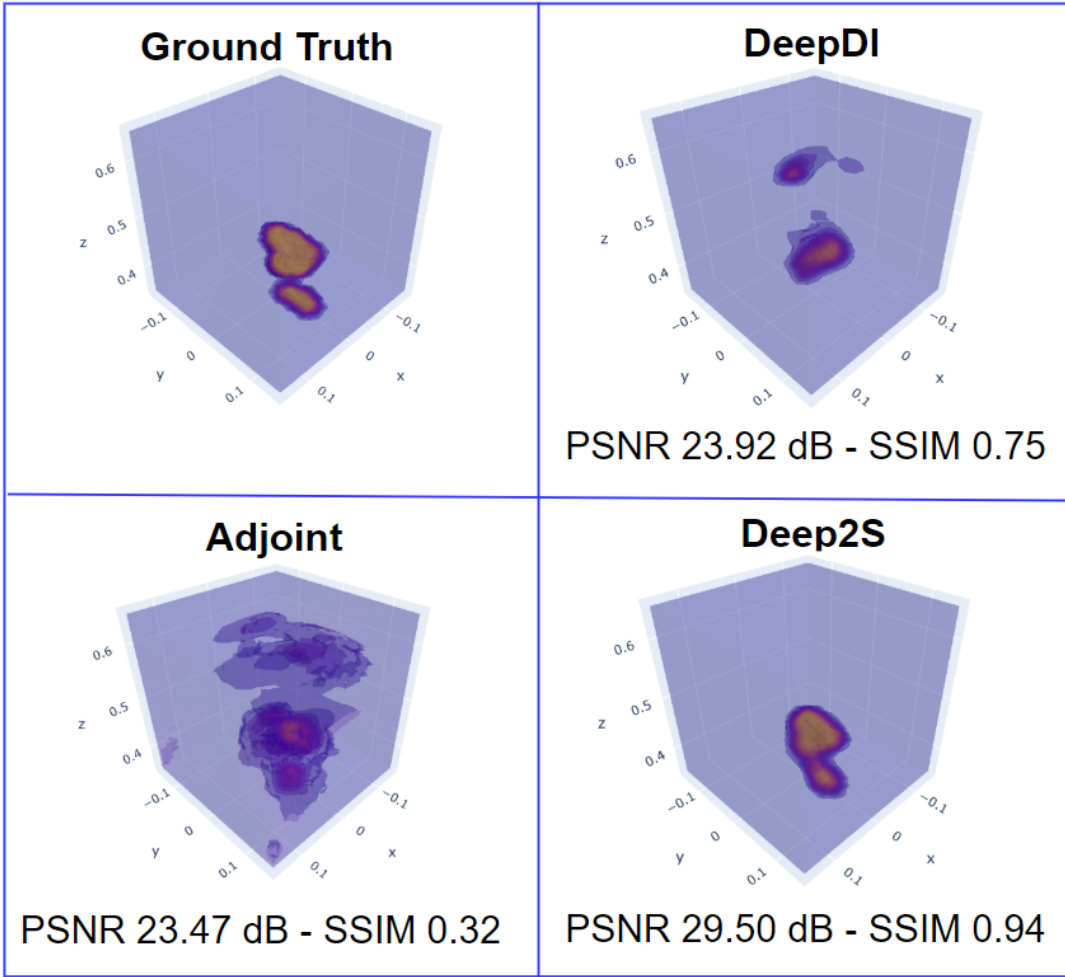
Figure 5.9: Reconstructions of the different algorithms for the first test image of the synthetically generated dataset at 30 dB SNR (Number of Frequency Steps: 15), (The units of x, y, and z-axis are meters).

adjoint operation as the first stage of the Deep2S approach. The performance comparison also carried out with adjoint operation instead of the classical BP algorithm. From now on, we will only examine the reconstruction performance of the Deep2S approach in different experimental settings.

5.4   SNR Analysis

Deep2S approach appears to be robust to different noise levels. Figures 5.10 and 5.11 show the performance analysis of the Deep2S approach at different SNR settings. In this experiment, we trained the network on data with a fixed number of frequency

steps which is fifteen but varied the noise level. We obtained five different networks which are trained for different SNR values -10, 0, 10, 20, and 30 dB.

To train our 3D U-Net Architecture the radar measurements of the generated dataset corrupted with additive white Gaussian noise (AWGN) with the specified SNR values. In the end, the average reconstruction performance of the Deep2S approach and adjoint operation in different SNR settings is given in terms of PSNR and SSIM metrics. Figure 5.10 shows the degradation of the PSNR metric along with a decrease in SNR value for a fixed number of frequency steps. Figure 5.11 shows the degradation of the SSIM metric along with a decrease in SNR value. The results show that even the SNR value in our experimental setting is 0 dB, the reconstruction performance of the Deep2S approach gives a PSNR of 28 dB and an SSIM value of 0.90 at the average of 100 test images. The Deep2S approach gives an acceptable reconstruction when the power of the measurements is equal to the power of corrupting noise. Also, the reconstruction performance does not corrupt significantly until the SNR of the radar measurements drops 10 dB. Until that point, the Deep2S approach gives the best reconstruction performance in terms of PSNR and SSIM. From this analysis, we conclude that training and testing for a high noise level configuration do not degrade the reconstruction performance significantly. Figure 5.12 illustrates the reconstructed images for the adjoint operation and the Deep2S approach visually for the specified SNR values -10, 10, and 30 dB. We observe that the network trained at -10 dB SNR gives an over smoothed reconstruction whereas the network trained at 30 dB preserves fine details and structures. As a result, in low SNR values, the Deep2S approach gives acceptable reconstruction performance where the target initial position, target shape, and size are correctly captured by the 3D U-Net Architecture. At high SNR values, the Deep2S approach gives a high-quality reconstruction where the reconstruction captures all fine details and structures of the ground-truth image.

5.5 Performance Comparison with Different Network Architectures

There is various type of Deep CNN architectures that are used for regression problems. To understand the effect of choosing different architectures on the reconstruction performance. The Residual Net [107] architecture is also trained to regress the

Figure 5.10: Average PSNR values of adjoint operation and, Deep2S approach for 100 test images of the synthetically generated dataset versus measurement SNR.

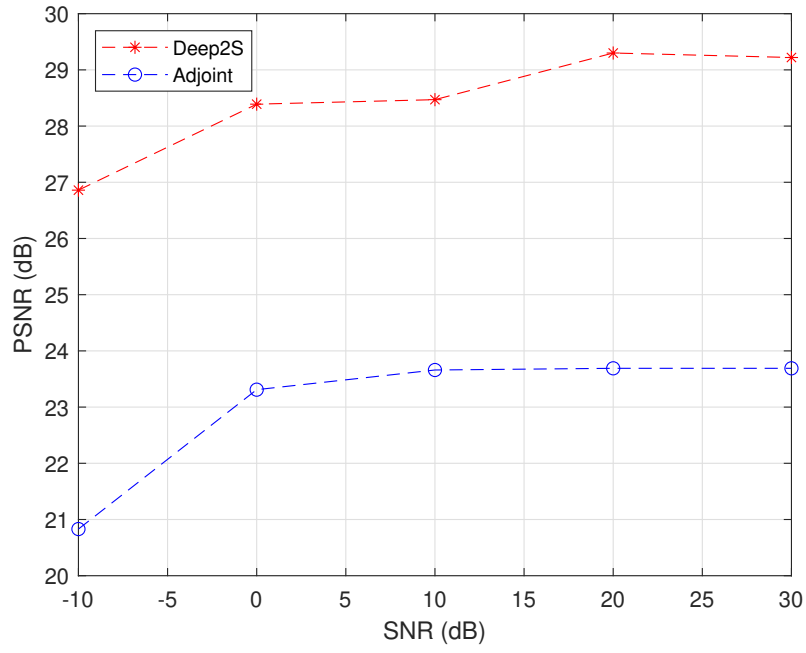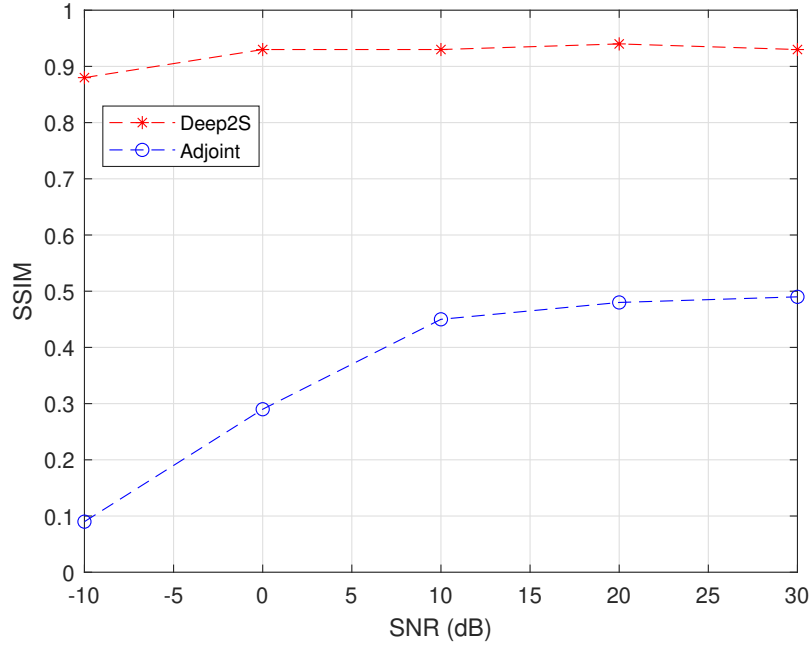Figure 5.11: Average SSIM values of adjoint operation and, Deep2S approach for 100 test images of the synthetically generated dataset versus measurement SNR.
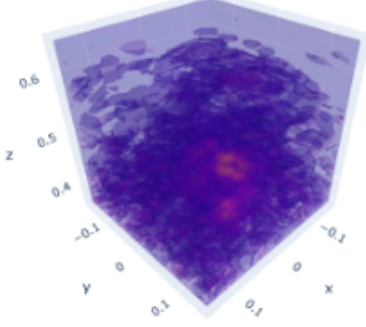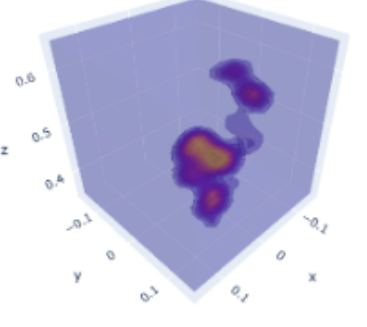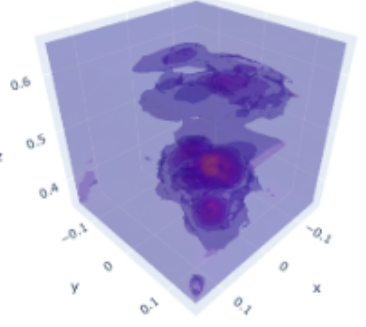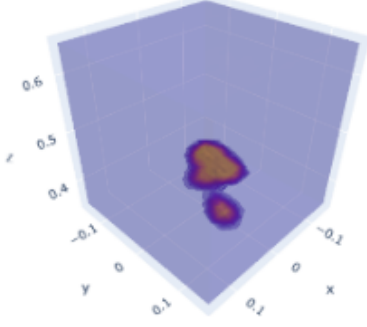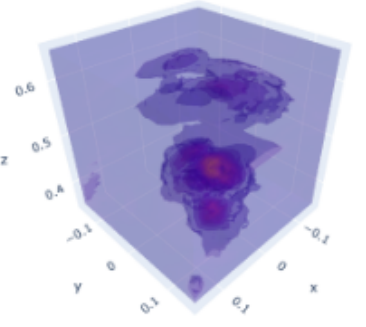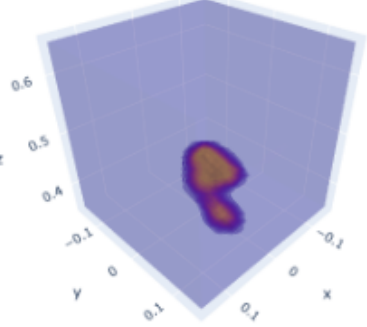
Figure 5.12: Reconstructions of the adjoint operation and, Deep2S approach for the first test image of the synthetically generated dataset (Number of Frequency Steps: 15).

adjoint operation result to a suitable ground truth image. Briefly, we use residual net architecture instead of 3D U-Net architecture in the second stage of the Deep2S

Table 5.6: Average Reconstruction Performances for 100 Test Images for ResNet at 30 dB SNR.

	ResNet	U-Net
PSNR	26.08	**29.22**
SSIM	0.66	**0.93**

approach. The implementation of ResNet architecture is based on [107]. In the experiments, a modified version of the original ResNet architecture is used. 3D convolutions are used instead of 2D convolutions. The network layer size is selected to be 10 in order to achieve a similar training time with the 3D U-Net architecture which is seven to eight hours.

In Table 5.6, the average reconstruction performance of the Residual Net architecture and the 3D U-Net architecture for 100 test images are given for fifteen frequency steps at 30 dB SNR. We can conclude that the 3D U-Net architecture reconstruction performance surpasses the Residual Net architecture in terms of PSNR and SSIM. This comparison demonstrates the power of the 3D U-Net architecture which is used in the second stage of the developed approaches. Figure 5.13 illustrates the reconstructed images for Residual Net architecture and the developed 3D U-Net architecture visually for the first test image. By looking at Figure 5.13, Residual Net architecture gives a poor reconstruction which has some severe volume artifact. The reconstruction obtained with Residual Net is not comparable with the 3D U-Net Architecture. The proposed 3D U-Net Architecture provides the best reconstruction performance compared to Residual Net architecture.

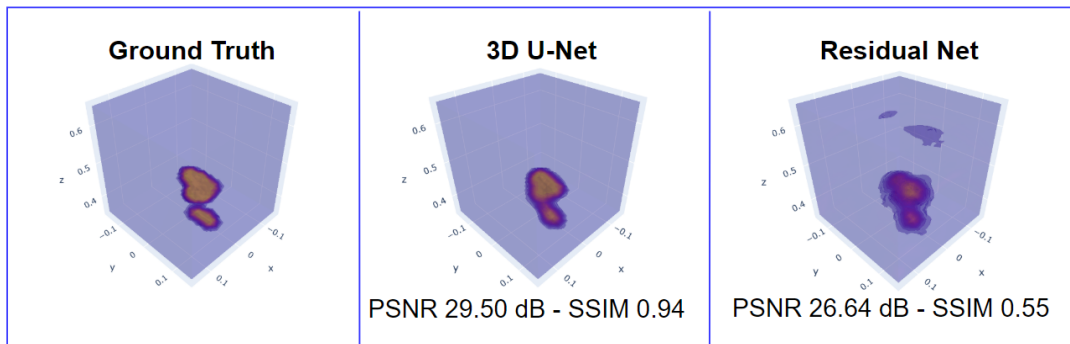Figure 5.13: Reconstructions of the Deep2S Approach for the first test image of the synthetically generated dataset at 30 dB SNR (Number of Frequency Steps: 15).

CHAPTER 6

CONCLUSIONS

In this thesis, we developed two deep learning-based approaches for near-field multiple-input multiple-output (MIMO) radar imaging based on learned direct reconstruction. The main motivation of the developed approaches was to reconstruct the 3D complex-valued reflectivity magnitude with high image quality and low computational cost. For this reason, we focused on learned direct reconstruction due to their feed-forward (non-iterative) nature. Plug-and-play and unrolled methods are not studied in this thesis because these approaches may require higher memory usage and computational complexity due to their iterative nature and computation of the forward scattering operator and its adjoint at every iteration. As future work, these type of methods can also be studied for near-field MIMO radar imaging and compared with the developed methods in this thesis.

In the first developed approach, we exploit the observation model of the physical system by using the adjoint operation stage. The adjoint operation has the benefit of fast computation due to its non-iterative nature and also exploits the physics-based knowledge. This result is also refined with a DNN in the second stage, which works as a deep denoiser. Here, it must be noted that in real practical scenarios we are interested in imaging 3D complex extended targets which generally have correlations along both range and cross-range directions. Therefore, the structure of the DNN architecture is critical for obtaining a good reconstruction quality. Unlike the existing approaches in the literature, in this thesis, we use a modified 3D U-net architecture that involves 3D operations to exploit the correlations in the 3D object space.

We demonstrate the performance of the developed method using a synthetically generated dataset and compare its performance with the commonly used analytical meth-

ods. It has been observed that the developed method provides the best reconstruction quality while enabling fast reconstruction. Hence due to its low computational complexity, it can be used in real-time applications. All adjustable parameters of the developed approach are also learned end-to-end, which avoids performance deterioration caused by optimum parameter selection.

For comparison, a second approach is also developed which replaces the adjoint stage with a fully connected neural network. As expected, this pure neural network approach was outperformed by the first developed method that involves adjoint. This result is consistent with the existing literature which emphasizes the significance of incorporating physics-based knowledge into deep-learning based reconstruction. However, exploring the performance with different types of 3D network architectures (such as those mimicking transforms) may improve the reconstruction quality and is a topic for future study.

Another contribution of the thesis is the generation of synthetic 3D scenes that involves extended targets with random phase in order to obtain large data for training. A real-world dataset is the most appropriate way of training a network. However, a large dataset is often not available in many radar imaging applications including near-field MIMO radar imaging. Our approach generates different three-dimensional distributed objects that spread within the reconstruction cube from a randomly chosen center. Another topic for future study is to modify this synthetic dataset generation method to obtain a more realistic dataset.

As a last note, numerical simulations are performed in a specific physical setup where we consider Mill Cross array topology and fix the number of transmitting and receiving antennas. As future work, the whole system optimization can be considered jointly with image reconstruction task. It would be interesting to optimize he number of antennas and their positions together with the DNN architecture in an end-to-end fashion.

REFERENCES

[1] Y. Zhou, J. Yu, Z. Xu, Y. Wang, and Q. Cao, "Fast Modeling Methods for Estimating Imaging Performance of Whole Body Screening," *IEEE Antennas and Wireless Propagation Letters*, vol. 18, no. 1, pp. 39–43, 2018.

[2] S. S. Ahmed, A. Schiessl, F. Gumbmann, M. Tiebout, S. Methfessel, and L.-P. Schmidt, "Advanced Microwave Imaging," *IEEE Microwave Magazine*, vol. 13, no. 6, pp. 26–43, 2012.

[3] X. Zhuge and A. G. Yarovoy, "A Sparse Aperture MIMO-SAR-Based UWB Imaging System for Concealed Weapon Detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 1, pp. 509–518, 2010.

[4] M. Klemm, J. A. Leendertz, D. Gibbins, I. J. Craddock, A. Preece, and R. Benjamin, "Microwave Radar-Based Differential Breast Cancer Imaging: Imaging in Homogeneous Breast Phantoms and Low Contrast Scenarios," *IEEE Transactions on Antennas and Propagation*, vol. 58, no. 7, pp. 2337–2344, 2010.

[5] D. M. Sheen, D. L. McMakin, and T. E. Hall, "Three-dimensional millimeter-wave imaging for concealed weapon detection," *IEEE Transactions on Microwave Theory and Techniques*, vol. 49, no. 9, pp. 1581–1592, 2001.

[6] M. Soumekh, "Bistatic synthetic aperture radar inversion with application in dynamic object imaging," in *Acoustics, Speech, and Signal Processing, IEEE International Conference on*, pp. 2577–2578, IEEE Computer Society, 1991.

[7] Y. Alvarez, Y. Rodriguez-Vaqueiro, B. Gonzalez-Valdes, S. Mantzavinos, C. M. Rappaport, F. Las-Heras, and J. Á. Martínez-Lorenzo, "Fourier-Based Imaging for Multistatic Radar Systems," *IEEE Transactions on Microwave Theory and Techniques*, vol. 62, no. 8, pp. 1798–1810, 2014.

[8] M. E. Yanik and M. Torlak, "Near-Field MIMO-SAR Millimeter-Wave Imag-

ing With Sparsely Sampled Aperture Data," *IEEE Access*, vol. 7, pp. 31801–31819, 2019.

[9] E. Anadol, I. Seker, S. Camlica, T. O. Topbas, S. Koc, L. Alatan, F. Oktem, and O. A. Civi, "UWB 3D Near-Field Imaging with a Sparse MIMO Antenna Array for Concealed Weapon Detection," in *Radar Sensor Technology XXII*, vol. 10633, p. 106331D, International Society for Optics and Photonics, 2018.

[10] X. Zhuge and A. Yarovoy, "Near-field ultra-wideband imaging with two-dimensional sparse MIMO array," in *Proceedings of the Fourth European Conference on Antennas and Propagation*, pp. 1–4, IEEE, 2010.

[11] S. S. Ahmed, A. Schiessl, and L.-P. Schmidt, "A novel fully electronic active real-time imager based on a planar multistatic sparse array," *IEEE Transactions on Microwave Theory and Techniques*, vol. 59, no. 12, pp. 3567–3576, 2011.

[12] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, "Deep convolutional neural network for inverse problems in imaging," *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4509–4522, 2017.

[13] S. Basu and Y. Bresler, "O (n/sup 2/log/sub 2/n) filtered backprojection reconstruction algorithm for tomography," *IEEE Transactions on Image Processing*, vol. 9, no. 10, pp. 1760–1773, 2000.

[14] D. L. Marks, O. Yurduseven, and D. R. Smith, "Fourier accelerated multistatic imaging: A fast reconstruction algorithm for multiple-input-multiple-output radar imaging," *IEEE Access*, vol. 5, pp. 1796–1809, 2017.

[15] Y. Liu, X. Xu, and G. Xu, "MIMO radar calibration and imagery for near-field scattering diagnosis," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 1, pp. 442–452, 2017.

[16] C. J. Nolan and M. Cheney, "Microlocal analysis of synthetic aperture radar imaging," *Journal of Fourier Analysis and Applications*, vol. 10, no. 2, pp. 133–148, 2004.

[17] L.-I. Sorsa, S. Pursiainen, and C. Eyraud, "Analysis of full microwave propagation and backpropagation for a complex asteroid analogue via single-point quasi-monostatic data," *Astronomy & Astrophysics*, vol. 645, p. A73, 2021.

[18] X. Zhuge, A. G. Yarovoy, T. Savelyev, and L. Ligthart, "Modified Kirchhoff Migration for UWB MIMO Array-Based Radar Imaging," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 6, pp. 2692–2703, 2010.

[19] X. Zhuge and A. G. Yarovoy, "Three-Dimensional Near-Field MIMO Array Imaging Using Range Migration Techniques," *IEEE Transactions on Image Processing*, vol. 21, no. 6, pp. 3026–3033, 2012.

[20] J. M. Lopez-Sanchez and J. Fortuny-Guasch, "3-D radar imaging using range migration techniques," *IEEE Transactions on Antennas and Propagation*, vol. 48, no. 5, pp. 728–737, 2000.

[21] T. Sakamoto, T. Sato, P. J. Aubry, and A. G. Yarovoy, "Ultra-Wideband Radar Imaging Using a Hybrid of Kirchhoff Migration and Stolt F-K Migration With an Inverse Boundary Scattering Transform," *IEEE Transactions on Antennas and Propagation*, vol. 63, no. 8, pp. 3502–3512, 2015.

[22] K. Tan, S. Wu, X. Liu, and G. Fang, "Omega-K algorithm for near-field 3-D image reconstruction based on planar SIMO/MIMO array," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 4, pp. 2381–2394, 2018.

[23] K. Tan, S. Wu, X. Liu, and G. Fang, "A Modified Omega-K Algorithm for Near-Field MIMO Array-Based 3-D Reconstruction," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 10, pp. 1555–1559, 2018.

[24] Y. Álvarez, Y. Rodriguez-Vaqueiro, B. Gonzalez-Valdes, F. Las-Heras, and A. García-Pino, "Fourier-Based Imaging for Subsampled Multistatic Arrays," *IEEE Transactions on Antennas and Propagation*, vol. 64, no. 6, pp. 2557–2562, 2016.

[25] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[26] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[27] E. J. Candès and M. B. Wakin, "An Introduction To Compressive Sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, 2008.

[28] L. Qiao, Y. Wang, Z. Shen, Z. Zhao, and Z. Chen, "Compressive sensing for direct millimeter-wave holographic imaging," *Applied Optics*, vol. 54, no. 11, pp. 3280–3289, 2015.

[29] Q. Cheng, A. Alomainy, and Y. Hao, "Near-Field Millimeter-Wave Phased Array Imaging With Compressive Sensing," *IEEE Access*, vol. 5, pp. 18975–18986, 2017.

[30] L. C. Potter, E. Ertin, J. T. Parker, and M. Cetin, "Sparsity and Compressed Sensing in Radar Imaging," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1006–1020, 2010.

[31] I. Stojanovic, M. Çetin, and W. C. Karl, "Compressed Sensing of Monostatic and Multistatic SAR," *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no. 6, pp. 1444–1448, 2013.

[32] F. S. Oktem, "Sparsity-based three-dimensional image reconstruction for near-field MIMO radar imaging," *Turkish Journal of Electrical Engineering & Computer Sciences*, vol. 27, no. 5, pp. 3282–3295, 2019.

[33] S. Li, G. Zhao, H. Li, B. Ren, W. Hu, Y. Liu, W. Yu, and H. Sun, "Near-Field Radar Imaging via Compressive Sensing," *IEEE Transactions on Antennas and Propagation*, vol. 63, no. 2, pp. 828–833, 2014.

[34] C. Ma, T. S. Yeo, Y. Zhao, and J. Feng, "MIMO Radar 3D Imaging Based on Combined Amplitude and Total Variation Cost Function With Sequential Order One Negative Exponential Form," *IEEE Transactions on Image Processing*, vol. 23, no. 5, pp. 2168–2183, 2014.

[35] J. H. Ender, "On compressive sensing applied to radar," *Signal Processing*, vol. 90, no. 5, pp. 1402–1414, 2010.

[36] S.-J. Wei, X.-L. Zhang, J. Shi, and K.-F. Liao, "Sparse array microwave 3-D imaging: Compressed sensing recovery and experimental study," *Progress In Electromagnetics Research*, vol. 135, pp. 161–181, 2013.

[37] E. Yiğit, "Compressed Sensing for Millimeter-wave Ground Based SAR/ISAR Imaging," *Journal of Infrared, Millimeter, and Terahertz Waves*, vol. 35, no. 11, pp. 932–948, 2014.

[38] G. Zhao, S. Li, B. Ren, Q. Qiu, and H. Sun, "Cylindrical Three-Dimensional Millimeter-Wave Imaging via Compressive Sensing," *International Journal of Antennas and Propagation*, vol. 2015, 2015.

[39] S. Samadi, M. Çetin, and M. A. Masnadi-Shirazi, "Sparse Representation-Based Synthetic Aperture Radar Imaging," *IET Radar, Sonar & Navigation*, vol. 5, no. 2, pp. 182–193, 2011.

[40] Z. Yang and Y. R. Zheng, "A comparative study of compressed sensing approaches for 3-D synthetic aperture radar image reconstruction," *Digital Signal Processing*, vol. 32, pp. 24–33, 2014.

[41] J. Guo, J. Zhang, K. Yang, B. Zhang, W. Hong, and Y. Wu, "Information Capacity and Sampling Ratios for Compressed Sensing-Based SAR Imaging," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 4, pp. 900–904, 2014.

[42] L. Guo and A. Abbosh, "Microwave Stepped Frequency Head Imaging Using Compressive Sensing With Limited Number of Frequency Steps," *IEEE Antennas and Wireless Propagation Letters*, vol. 14, pp. 1133–1136, 2015.

[43] X. Hu, N. Tong, B. Song, S. Ding, and X. Zhao, "Joint sparsity-driven three-dimensional imaging method for multiple-input multiple-output radar with sparse antenna array," *IET Radar, Sonar & Navigation*, vol. 11, no. 5, pp. 709–720, 2017.

[44] C. Ma, T. S. Yeo, and B. P. Ng, "Multiple input multiple output radar imaging based on multidimensional linear equations and sparse signal recovery," *IET Radar, Sonar & Navigation*, vol. 12, no. 1, pp. 3–10, 2018.

[45] Z. Jiao, C. Ding, X. Liang, L. Chen, and F. Zhang, "Sparse Bayesian Learning Based Three-Dimensional Imaging Algorithm for Off-Grid Air Targets in MIMO Radar Array," *Remote Sensing*, vol. 10, no. 3, p. 369, 2018.

[46] P. Huang, X. Li, and H. Wang, "Tensor-based Match Pursuit Algorithm for MIMO Radar Imaging," *Radioengineering*, vol. 27, no. 2, pp. 580–586, 2018.

[47] J. Chen, Y. Li, J. Wang, Y. Li, and Y. Zhang, "An accurate imaging algorithm for millimeter wave synthetic aperture imaging radiometer in near-field.," *Progress In Electromagnetics Research*, vol. 141, pp. 517–535, 2013.

[48] A. C. Gurbuz, J. H. McClellan, and W. R. Scott Jr, "Compressive sensing for subsurface imaging using ground penetrating radar," *Signal Processing*, vol. 89, no. 10, pp. 1959–1972, 2009.

[49] W. Zhang and A. Hoorfar, "A generalized approach for SAR and MIMO radar imaging of building interior targets with compressive sensing," *IEEE Antennas and Wireless Propagation Letters*, vol. 14, pp. 1052–1055, 2015.

[50] R. Coifman, V. Rokhlin, and S. Wandzura, "The fast multipole method for the wave equation: A pedestrian prescription," *IEEE Antennas and Propagation magazine*, vol. 35, no. 3, pp. 7–12, 1993.

[51] S. Li, G. Zhao, H. Sun, and M. Amin, "Compressive sensing imaging of 3-D object by a holographic algorithm," *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 12, pp. 7295–7304, 2018.

[52] E. A. Miran, F. S. Oktem, and S. Koc, "Sparse Reconstruction for Near-Field MIMO Radar Imaging Using Fast Multipole Method," *IEEE Access*, vol. 9, pp. 151578–151589, 2021.

[53] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 2, pp. 295–307, 2015.

[54] D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrell, and A. A. Efros, "Context encoders: Feature learning by inpainting," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2536–2544, 2016.

[55] H. Yao, F. Dai, S. Zhang, Y. Zhang, Q. Tian, and C. Xu, "Dr2-net: Deep residual reconstruction network for image compressive sensing," *Neurocomputing*, vol. 359, pp. 483–493, 2019.

[56] M. Hradiš, J. Kotera, P. Zemcık, and F. Šroubek, "Convolutional neural networks for direct text deblurring," in *Proceedings of BMVC*, vol. 10, pp. 6.1–6.13, 2015.

[57] A. Lucas, M. Iliadis, R. Molina, and A. K. Katsaggelos, "Using deep neural networks for inverse problems in imaging: beyond analytical methods," *IEEE Signal Processing Magazine*, vol. 35, no. 1, pp. 20–36, 2018.

[58] C. J. Schuler, M. Hirsch, S. Harmeling, and B. Schölkopf, "Learning to deblur," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 7, pp. 1439–1451, 2015.

[59] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4681–4690, 2017.

[60] E. Nehme, L. E. Weiss, T. Michaeli, and Y. Shechtman, "Deep-STORM: super-resolution single-molecule microscopy by deep learning," *Optica*, vol. 5, no. 4, pp. 458–464, 2018.

[61] Y. Rivenson, Y. Zhang, H. Günaydın, D. Teng, and A. Ozcan, "Phase recovery and holographic image reconstruction using deep learning in neural networks," *Light: Science & Applications*, vol. 7, no. 2, pp. 17141–17141, 2018.

[62] A. Sinha, J. Lee, S. Li, and G. Barbastathis, "Lensless computational imaging through deep learning," *Optica*, vol. 4, no. 9, pp. 1117–1125, 2017.

[63] T. Nguyen, Y. Xue, Y. Li, L. Tian, and G. Nehmetallah, "Deep learning approach for fourier ptychography microscopy," *Optics express*, vol. 26, no. 20, pp. 26470–26484, 2018.

[64] Ç. Işıl, F. S. Oktem, and A. Koç, "Deep iterative reconstruction for phase retrieval," *Applied Optics*, vol. 58, no. 20, pp. 5422–5431, 2019.

[65] K. Zhang, W. Zuo, S. Gu, and L. Zhang, "Learning deep CNN denoiser prior for image restoration," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3929–3938, 2017.

[66] T. Meinhardt, M. Moller, C. Hazirbas, and D. Cremers, "Learning proximal operators: Using denoising networks for regularizing inverse imaging prob-

lems," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1781–1790, 2017.

[67] J. Rick Chang, C.-L. Li, B. Poczos, B. Vijaya Kumar, and A. C. Sankaranarayanan, "One network to solve them all–solving linear inverse problems using deep projection models," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 5888–5897, 2017.

[68] C. Metzler, P. Schniter, A. Veeraraghavan, *et al.*, "prDeep: robust phase retrieval with a flexible deep network," in *International Conference on Machine Learning*, pp. 3501–3510, PMLR, 2018.

[69] H. K. Aggarwal, M. P. Mani, and M. Jacob, "MoDL: Model-based deep learning architecture for inverse problems," *IEEE Transactions on Medical Imaging*, vol. 38, no. 2, pp. 394–405, 2018.

[70] J. Adler and O. Öktem, "Solving ill-posed inverse problems using iterative deep neural networks," *Inverse Problems*, vol. 33, no. 12, p. 124007, 2017.

[71] J. Adler and O. Öktem, "Learned primal-dual reconstruction," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1322–1332, 2018.

[72] M. B. Alver, A. Saleem, and M. Çetin, "Plug-and-play synthetic aperture radar image formation using deep priors," *IEEE Transactions on Computational Imaging*, vol. 7, pp. 43–57, 2020.

[73] C. Hu, L. Wang, Z. Li, and D. Zhu, "Inverse synthetic aperture radar imaging using a fully convolutional neural network," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 7, pp. 1203–1207, 2019.

[74] M. B. Alver, A. Saleem, and M. Çetin, "A novel plug-and-play SAR reconstruction framework using deep priors," in *2019 IEEE Radar Conference (RadarConf)*, pp. 1–6, IEEE, 2019.

[75] L. Peng, X. Qiu, C. Ding, and W. Tie, "Generating 3d point clouds from a single SAR image using 3D reconstruction network," in *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*, pp. 3685–3688, IEEE, 2019.

[76] S. Wang, J. Guo, Y. Zhang, Y. Hu, C. Ding, and Y. Wu, "Single target SAR 3D reconstruction based on deep learning," *Sensors*, vol. 21, no. 3, p. 964, 2021.

[77] J. Chen, L. Peng, X. Qiu, C. DING, and Y. WU, "A 3D building reconstruction method for SAR images based on deep neural network," *Scientia Sinica Informationis*, vol. 49, no. 12, pp. 1606–1625, 2019.

[78] S. Zhou, Y. Li, F. Zhang, L. Chen, and X. Bu, "Automatic reconstruction of 3-D building structures for tomoSAR using neural networks," in *2019 IEEE International Conference on Signal, Information and Data Processing (ICSIDP)*, pp. 1–5, IEEE, 2019.

[79] T. Weiss, N. Peretz, S. Vedula, A. Feuer, and A. Bronstein, "Joint optimization of system design and reconstruction in MIMO radar imaging," in *2021 IEEE 31st International Workshop on Machine Learning for Signal Processing (MLSP)*, pp. 1–6, IEEE, 2021.

[80] J. Gao, B. Deng, Y. Qin, H. Wang, and X. Li, "Enhanced radar imaging using a complex-valued convolutional neural network," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 1, pp. 35–39, 2018.

[81] G. Sun, F. Zhang, B. Gao, Y. Zhou, Y. Xiang, and S. Pan, "Photonics-based 3D radar imaging with CNN-assisted fast and noise-resistant image construction," *Optics Express*, vol. 29, no. 13, pp. 19352–19361, 2021.

[82] M. Wang, S. Wei, J. Liang, X. Zeng, C. Wang, J. Shi, and X. Zhang, "RMIST-Net: Joint range migration and sparse reconstruction network for 3-D mmw imaging," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–17, 2021.

[83] Q. Cheng, A. A. Ihalage, Y. Liu, and Y. Hao, "Compressive sensing radar imaging with convolutional neural networks," *IEEE Access*, vol. 8, pp. 212917–212926, 2020.

[84] M. B. Kocamis and F. S. Oktem, "Optimal design of sparse MIMO arrays for near-field ultrawideband imaging," in *2017 25th European Signal Processing Conference (EUSIPCO)*, pp. 1952–1956, IEEE, 2017.

[85] E. Anadol, I. Seker, S. Camlica, T. O. Topbas, S. S. Koç, L. Alatan, S. F. Öktem, and H. Aydın Çivi, "UWB 3D near-Field imaging with a sparse MIMO antenna array for concealed weapon detection," vol. 10633, 2018.

[86] S. J. Wright, R. D. Nowak, and M. A. Figueiredo, "Sparse reconstruction by separable approximation," *IEEE Transactions on Signal Processing*, vol. 57, no. 7, pp. 2479–2493, 2009.

[87] J. A. Tropp and S. J. Wright, "Computational methods for sparse solution of linear inverse problems," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 948–958, 2010.

[88] M. V. Afonso, J. M. Bioucas-Dias, and M. A. Figueiredo, "Fast image recovery using variable splitting and constrained optimization," *IEEE Transactions on Image Processing*, vol. 19, no. 9, pp. 2345–2356, 2010.

[89] M. K. Ng, P. Weiss, and X. Yuan, "Solving constrained total-variation image restoration and reconstruction problems via alternating direction methods," *SIAM Journal on Scientific Computing*, vol. 32, no. 5, pp. 2710–2736, 2010.

[90] D. Krishnan and R. Fergus, "Fast image deconvolution using hyper-Laplacian priors," *Advances in Neural Information Processing Systems*, vol. 22, pp. 1033–1041, 2009.

[91] L. B. Montefusco, D. Lazzaro, S. Papi, and C. Guerrini, "A fast compressed sensing approach to 3D MR image reconstruction," *IEEE Transactions on Medical Imaging*, vol. 30, no. 5, pp. 1064–1075, 2010.

[92] A. Beck and M. Teboulle, "Gradient-based algorithms with applications to signal recovery," *Convex Optimization in Signal Processing and Communications*, pp. 42–88, 2009.

[93] F. S. Oktem, L. Gao, and F. Kamalabadi, "Computational spectral and ultrafast imaging via convex optimization," in *Handbook of Convex Optimization Methods in Imaging Science*, pp. 105–127, Springer, 2018.

[94] C. R. Vogel and M. E. Oman, "Fast, robust total variation-based reconstruction of noisy, blurred images," *IEEE Transactions on Image Processing*, vol. 7, no. 6, pp. 813–824, 1998.

[95] D. Geman and C. Yang, "Nonlinear image recovery with half-quadratic regularization," *IEEE Transactions on Image Processing*, vol. 4, no. 7, pp. 932–946, 1995.

[96] K. Kulkarni, S. Lohit, P. Turaga, R. Kerviche, and A. Ashok, "Reconnet: Non-iterative reconstruction of images from compressively sensed measurements," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 449–458, 2016.

[97] S. V. Venkatakrishnan, C. A. Bouman, and B. Wohlberg, "Plug-and-play priors for model based reconstruction," in *2013 IEEE Global Conference on Signal and Information Processing*, pp. 945–948, IEEE, 2013.

[98] A. Danielyan, V. Katkovnik, and K. Egiazarian, "Deblurring by augmented langrangian with BM3D frame prior," in *Workshop on Information Theoretic Methods in Science and Engineering*, pp. 16–18.

[99] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *International Conference on Medical Image Computing and Computer-assisted Intervention*, pp. 234–241, Springer, 2015.

[100] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *European Conference on Computer Vision*, pp. 694–711, Springer, 2016.

[101] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 25, p. 84–90, 2012.

[102] J. Cong and B. Xiao, "Minimizing computation in convolutional neural networks," in *International Conference on Artificial Neural Networks*, pp. 281–290, Springer, 2014.

[103] X. Zhuge and A. G. Yarovoy, "Study on two-dimensional sparse mimo uwb arrays for high resolution near-field imaging," *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 9, pp. 4173–4182, 2012.

[104] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[105] K. R. Rao and P. C. Yip, *The Transform and Data Compression Handbook*. CRC Press, 2018.

[106] D. Ravichandran, M. G. Ahamad, and M. A. Dhivakar, "Performance analysis of three-dimensional medical image compression based on discrete wavelet transform," in *2016 22nd International Conference on Virtual System & Multimedia (VSMM)*, pp. 1–8, IEEE, 2016.

[107] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising," *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3142–3155, 2017.