

# DL-ZTE: Towards Deep Learning-based Methods for Dead Time Gap Recovery in Zero TE MRI

*Ana Cismaru*



Electrical Engineering and Computer Sciences  
University of California, Berkeley

Technical Report No. UCB/EECS-2024-98

<http://www2.eecs.berkeley.edu/Pubs/TechRpts/2024/EECS-2024-98.html>

May 13, 2024

Copyright © 2024, by the author(s).  
All rights reserved.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission.

---

**DL-ZTE: Towards Deep Learning-based Methods for Dead Time Gap  
Recovery in Zero TE MRI**

by Ana Cismaru

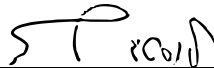
---

**Research Project**

Submitted to the Department of Electrical Engineering and Computer Sciences,  
University of California at Berkeley, in partial satisfaction of the requirements for the  
degree of **Master of Science, Plan II.**

Approval for the Report and Comprehensive Examination:

**Committee:**



---

Professor Michael Lustig  
Research Advisor

05/11/2024

---

(Date)

\* \* \* \* \*



---

Professor Chunlei Liu  
Second Reader

05/12/2024

---

(Date)

DL-ZTE: Towards Deep Learning-based Methods for Dead Time Gap Recovery in Zero TE  
MRI

by

Ana Cismaru

A thesis submitted in partial satisfaction of the

requirements for the degree of

Master of Science

in

Electrical Engineering Computer Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Michael Lustig,  
Professor Chunlei Liu

Spring 2024

DL-ZTE: Towards Deep Learning-based Methods for Dead Time Gap Recovery in Zero TE  
MRI

Copyright 2024  
by  
Ana Cismaru

Abstract

DL-ZTE: Towards Deep Learning-based Methods for Dead Time Gap Recovery in Zero TE MRI

by

Ana Cismaru

Master of Science in Electrical Engineering Computer Science

University of California, Berkeley

Professor Michael Lustig,

Zero Echo Time (ZTE) MRI is a rapid and near-silent 3D radial MRI sequence that is based on constant gradients incremented by small amounts between TRs. Because the readout gradients are active during the RF excitation, the initial readout points for each radial spoke are missed, resulting in a spherical Dead Time Gap (DTG) at the center of  $k$ -space. This gap results in incorrect image contrast if left uncorrected. Traditional approaches to correcting the inaccurate contrast from the DTG often require additional scan time or complex hardware adjustments. This limits ZTE's practical applicability, especially in settings such as dynamic contrast-enhanced where image contrast changes rapidly. This work presents multiple deep-learning frameworks designed to reconstruct large DTG solely from multi-coil  $k$ -space data, without the need for explicit coil sensitivities or additional acquisitions. A large ZTE dataset is simulated by applying the ZTE forward model to a Cartesian fully sampled MRI dataset. Various learning approaches were explored to assess their abilities to recover the missing center of  $k$ -space, including UNet, Interlacer, and Diffusion models. The proposed methods demonstrated strong 2D reconstruction results on simulated datasets with the Diffusion model performing best followed closely by the UNet approach. Due to the significant memory and computational demands of 3D Diffusion models, a 3D UNet model was chosen for reconstructing in-vivo data. This 3D UNet model demonstrated strong performance in reconstructing simulated 3D ZTE data. However, application to in-vivo data revealed challenges such as the model hallucinating auras that correlate with the field of view of the training data. This work contributes to the ongoing development of efficient and accurate reconstruction techniques for ZTE MRI, aiming to extend this fast and silent sequence to applications such as DCE, where image contrast is dynamically changing.

# Contents

Contents	i
List of Figures	iii
List of Tables	v
<b>1 Introduction</b>	<b>1</b>
1.1 MRI Background . . . . .	1
1.2 Zero Echo Time MRI . . . . .	2
1.3 Related Work . . . . .	3
1.4 Our Approach . . . . .	4
<b>2 Methods</b>	<b>5</b>
2.1 The ZTE Forward Model . . . . .	5
2.2 Datasets . . . . .	5
2.3 Simulated Training Data Generation . . . . .	6
2.4 In-Vivo Acquisition . . . . .	7
<b>3 Learning Frameworks</b>	<b>8</b>
3.1 UNet Architecture . . . . .	8
3.2 Interlacer Architecture . . . . .	9
3.3 Diffusion Architecture . . . . .	10
<b>4 2D Results</b>	<b>13</b>
4.1 2D UNet and Interlacer Training Details . . . . .	13
4.2 2D Diffusion Training Details . . . . .	13
4.3 2D Simulated Data UNet Experiments . . . . .	14
4.4 2D Comparison of Learning Frameworks . . . . .	16
<b>5 3D Results</b>	<b>18</b>
5.1 3D Training Details . . . . .	18
5.2 3D Simulated UNet Experiments . . . . .	19
5.3 In-vivo Experiments . . . . .	19

<b>6 Conclusion</b>	<b>23</b>
6.1 Discussion . . . . .	23
6.2 Conclusion . . . . .	24
<b>Bibliography</b>	<b>25</b>



# List of Figures

1.1	(Left) ZTE sequence characterized by constant gradients stepped by small amounts and short RF pulses. These contribute to the appearance of a dead time gap where the initial samples of $k$ -space are not acquired. (Right) The ZTE $k$ -space sampling trajectory is simplified in a 2D representation. The initial missed samples along each radial spoke correspond to the center of $k$ -space where the low-frequency information of the image resides. . . . .	2
1.2	A coil combined T2 image with simulated dead time gaps of varying sizes. The dead time gap is simulated by removing $N$ samples from the center of $k$ -space in each coil. As the size of the dead time gap increases, the contrast of the final coil combined image deviates from the ground truth. . . . .	3
2.1	The ZTE forward model. A 3D volume $x$ is multiplied with the coil sensitivity maps to generate coil images. A 3D NUFFT is applied to each coil image to generate radial $k$ -space. The initial $N$ samples of the $k$ -space data, which are affected by the dead time gap, are typically discarded. . . . .	6
3.1	The UNet model pipeline. The ZTE forward model $A$ is applied to a fully sampled coil-combined image to generate low-resolution coil $k$ -space containing a dead time gap. An adjoint NUFFT $I_\theta$ is then applied to yield gridded low-resolution coil images. Each coil image is individually input into the UNet alongside any additional information (in this case the RSS of all the coils from that image). The model outputs the reconstructed low-resolution coil. $H$ combines the saved high-frequency information of the coil with the low-resolution reconstructed coil. The high-resolution reconstructed coils can be combined via RSS or by estimating sensitivity maps and performing optimal coil combining[1]. . . . .	9
3.2	(reproduced from Singh et al.[22]) The Interlacer architecture depicting one layer of the model. The model takes in a coil image $u_0$ from which its $k$ -space $v_0$ is computed via an inverse Fourier transform. Each network then calculates a weighted combination of data from both the image and frequency domains after which it applies additional convolutions. . . . .	10
4.1	An example 2D input in its original resolution for clarity. The sensitivity map was calculated from the ground truth coil image shown on the right. . . . .	14

4.2	Reconstruction results on 3 UNet models trained with different inputs. Instead of presenting the raw multi-coil data, the images in this figure show the RSS of the multi-coil data. The model trained on coil images alone has the highest reconstruction error while the model trained on coil images with the RSS of all the coil inputs performs best. . . . .	15
4.3	Reconstruction results across the 3 learning frameworks presented. Instead of presenting the raw multi-coil data, the images in this figure show the RSS of the multi-coil data. Only the best UNet model trained on coil + RSS data is shown as the other experiments are compared in Figure 5. The Interlacer model qualitatively shows the worst reconstruction error and the Diffusion model shows the best reconstruction error. . . . .	16
5.1	3D Reconstruction of Simulated ZTE Knee Data with a 5-Sample Dead Time Gap using a UNet trained on 3D knee data. The reconstruction was individually performed on each coil image using a 3D UNet model, which received as input the individual coils and the RSS of all coils. Displayed from top to bottom are the axial, coronal, and sagittal views of the RSS from the reconstructed coil images, demonstrating accurate reconstruction across all orientations. The rightmost column illustrates the reconstruction error relative to the ground truth, amplified by a factor of 5 for clearer visualization. For this experiment, the reconstruction error is minimal. . . . .	20
5.2	3D Reconstruction of Simulated ZTE Knee Data with a 5-Sample Dead Time Gap using a UNet trained on 3D noise. The reconstruction was individually performed on each coil image using a 3D UNet model, which received as input the individual coils and the RSS of all coils. Displayed from top to bottom are the axial, coronal, and sagittal views of the RSS from the reconstructed coil images, demonstrating consistent reconstruction across all orientations. The rightmost column illustrates the reconstruction error relative to the ground truth, amplified by a factor of 5 for clearer visualization. For this experiment, errors are visible in both the higher and lower frequencies of the image, as evidenced by the visible anatomical details of the knee in the error plot. . . . .	21
5.3	In-vivo ZTE Reconstruction using the 3D UNet model trained on synthetic data. A head scan was acquired using a 12-channel head coil on a GE-3T MR750W (with $\pm 31.25\text{kHz}$ rBW, 2.3ms TR, flip angle 2 degrees, 24 cm FOV). The ground truth was reconstructed using additional acquisitions and WASPI. The region highlighted by the red box shows a failure point of the model where the reconstructed image is significantly brighter than the ground truth. . . . .	22

# List of Tables

- 4.1 Performance comparison of various UNet inputs on validation set. NRMSE and PSNR were computed on 100 validation samples. The Coil + RSS UNet yielded the lowest NRMSE and highest PSNR which indicates the best performance. . . . . 15
- 4.2 Performance comparison of the three learning frameworks on the validation set. The table shows the mean and standard deviation of the NRMSE and PSNR computed on 100 validation samples. The Diffusion model had the best NRMSE and the UNet model had the best PSNR . . . . . 17

## Acknowledgments

I would like to thank my advisor Professor Michael Lustig for his valuable insights and for exposing me to the world of MRI research. My heartfelt thanks also go to Alfredo de Goyeneche and Shreya Ramachandran for their exceptional mentorship and guidance throughout my research journey. It was such a pleasure to learn from and work with you all. I would like to express my appreciation to Professor Chunlei Liu for his thorough review and insightful feedback on my thesis, to Nalini Singh for her mentorship on the Interlacer architecture, and to the rest MikGroup for being such a friendly and supportive research group. I would also like to thank my family for being a great support system, introducing me to the world of Computer Science, and encouraging me to pursue grad school. Finally, I would like to thank all the other 5th Year M.S. students who made this year so amazing. To Robert and Fred, thank you for being my lab besties, to Derek and Ava, thank you for being a critical part of the dartay squad and the many fun nights at Bonita House. Lastly, to Devan and Tarun, thank you for being like my brothers and for all the adventures, projects, late nights, and mems from this year.

# Chapter 1

## Introduction

### 1.1 MRI Background

MRI offers unparalleled imaging of soft tissues without harmful ionizing radiation. The underlying technology, Nuclear Magnetic Resonance (NMR), harnesses magnetic fields to excite the hydrogen spins of atomic nuclei within the body. MRI images are generated by measuring the signal emitted by these spins. To acquire images, gradient fields are applied across different spatial dimensions resulting in samples in the spatial frequency domain.

MRI data are acquired in the spatial frequency domain, or  $k$ -space, by measuring the object's magnetization  $M(k)$ . The relationship between the measured  $k$ -space data and the final image is expressed by the Fourier transform with the equation:

$$M(k) = \int_{-\infty}^{\infty} m(r)e^{-2\pi ikr} dk$$

In most clinical scans,  $k$ -space is collected row-by-row in a Cartesian trajectory. Cartesian trajectories are slow and inefficient, which has led to the exploration of non-Cartesian trajectories, such as Golden-Angle Radial[7], which traverse  $k$ -space in a significantly shorter time. Even with more efficient trajectories, not all  $k$ -space is sampled, which introduces the need for techniques to accurately reconstruct an image from undersampled data.

One technique used to improve the acquisition of  $k$ -space and the reconstruction of the final MRI image is parallel imaging with multi-coil arrays[5][19]. In parallel imaging, multiple receive coils are placed around the anatomy of interest with each coil being more sensitive to the tissue nearest to the coil. This configuration not only accelerates the MRI scan time by allowing simultaneous data acquisition from multiple coils but also enriches the collected data with additional spatial information. This extra information is crucial during the reconstruction process, as it helps to accurately recreate the image from the undersampled  $k$ -space data. Consequently, parallel imaging with multi-coil arrays significantly improves both the efficiency and quality of MRI scans.

## 1.2 Zero Echo Time MRI

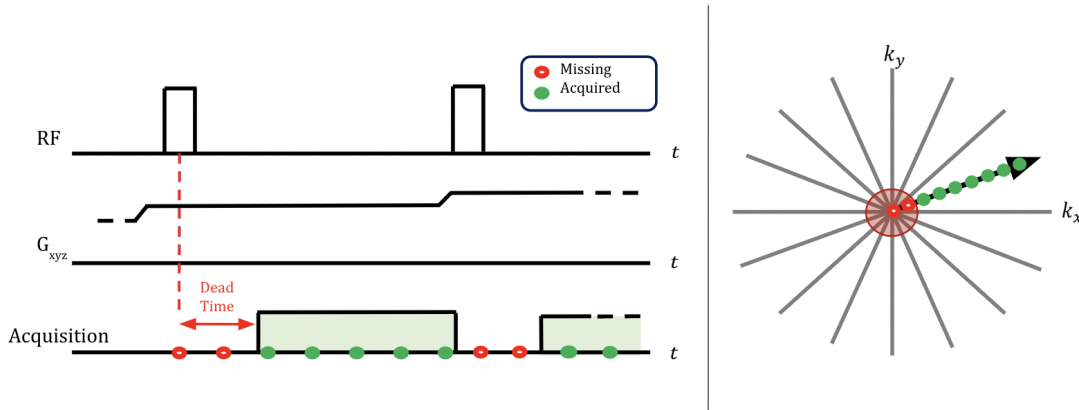


Figure 1.1: (Left) ZTE sequence characterized by constant gradients stepped by small amounts and short RF pulses. These contribute to the appearance of a dead time gap where the initial samples of  $k$ -space are not acquired. (Right) The ZTE  $k$ -space sampling trajectory is simplified in a 2D representation. The initial missed samples along each radial spoke correspond to the center of  $k$ -space where the low-frequency information of the image resides.

Traditional MRI techniques are limited by long acquisition times and significant acoustic noise, which stem from the inherent high sampling requirements and the Lorentz force interactions exerted on the hardware during gradient switching. These limitations can compromise patient comfort, increase the susceptibility to motion artifacts, and hinder the effective imaging of tissues characterized by short T2 relaxation times, where the signal decays rapidly before it can be adequately captured.

Zero Echo Time (ZTE) is a fast and silent 3D radial sequence that aims to address these limitations. It is characterized by the use of a low gradient slew rate and an ultra-short radiofrequency (RF) pulse. The low slew rate minimizes the rapid changes in gradient magnetic fields and reduces the Lorentz force interactions that lead to vibration and noise. This thereby can reduce acoustic noise generated during the scan from 30-40dB to 1-3 dB[16]. The ultra-short RF pulse allows the readout gradients to be active during excitation, a technique that facilitates the immediate commencement of data acquisition post-excitation, thereby speeding up the overall imaging process.

A key consideration of ZTE imaging is that the readout gradients are on during the RF excitation. Since MRI coils cannot transmit RF pulses and receive the resulting MR signal

at the same time, the first samples at the center of  $k$ -space, where most of the energy is concentrated, are missed. The absence of these central  $k$ -space samples is called the dead time gap (DTG), the size of which increases with the duration of the RF pulse. Without the key low-frequency components of the image, the resulting image contrast would be incorrect.

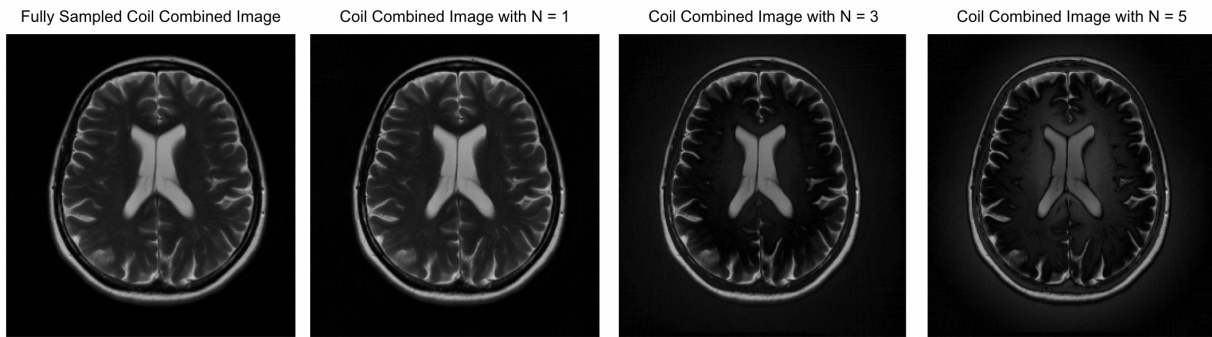


Figure 1.2: A coil combined T2 image with simulated dead time gaps of varying sizes. The dead time gap is simulated by removing  $N$  samples from the center of  $k$ -space in each coil. As the size of the dead time gap increases, the contrast of the final coil combined image deviates from the ground truth.

### 1.3 Related Work

Most commonly, the DTG is infilled using methods like WASPI[27] or PETRA[10], which leverage additional acquisitions to read the center of  $k$ -space. These present limitations when the contrast of the image changes over time, such as in Dynamic Contrast-Enhanced imaging (DCE). In DCE imaging, the contrast of the image changes over time as a contrast agent is administered and circulates through the body. The changing contrast can pose challenges for these infilling methods because the additional acquisitions may not sync precisely with the peak contrast phases, potentially leading to inaccuracies in the final images. This temporal mismatch can affect the diagnostic utility of the MRI, especially in assessing vascular health and tumor responses to treatments. Therefore, while methods like WASPI and PETRA are effective in static conditions, their application in dynamic imaging requires careful consideration and potential adaptation to accommodate the rapidly changing contrast levels.

Other reconstruction techniques, such as PPE-ZTE[8], utilize a  $k$ -space trajectory characterized by opposing spokes. This configuration allows the method to leverage the geometric relationship between these spokes to algebraically reconstruct the missing center of  $k$ -space, providing a unique approach to addressing data gaps. While this trajectory constraint optimizes the use of available data, the technique may face challenges when dealing with larger

dead time gaps, where the extrapolation can become more complex.

Recently, parallel imaging-based approaches have demonstrated effectiveness in recovering data from large dead time gaps without any intermediate dead time infilling[20] [26]. These methods take advantage of the fact that the convolution of  $k$ -space with coil sensitivity kernels can distribute the missing low spatial frequency information to higher frequency areas. This technique, however, hinges on the accurate estimation of sensitivity maps. Typically, auto-calibrating parallel imaging methods, such as ESPIRiT[24], depend on a fully-sampled central region of  $k$ -space to accurately estimate these maps. This requirement poses a challenge in ZTE imaging, where the central  $k$ -space is often undersampled or missing, complicating the generation of reliable sensitivity maps. Consequently, while these parallel imaging strategies are promising, their effectiveness in the ZTE context is limited by the feasibility of obtaining precise sensitivity maps without the complete central  $k$ -space data.

Additionally, the aforementioned parallel imaging techniques employ classical optimization frameworks that are often constrained by large memory requirements. This limitation becomes particularly significant in high-resolution imaging scenarios or when processing large datasets.

## 1.4 Our Approach

We propose several deep learning approaches to reconstruct the dead time gap leveraging multi-coil data. First, we contribute a method to simulate a large ZTE dataset that is still compatible with operations that rely on Cartesian data. Additionally, we perform a survey into different deep learning architectures to reconstruct ZTE data without any information aside from the multi-coil undersampled data. The survey explores the effectiveness of UNet-based models, a joint frequency and image domain model, and diffusion models. This work aims to address the limitations of current Zero Echo Time reconstruction techniques by focusing on larger dead time gaps, eliminating the need for additional acquisitions or detailed knowledge of sensitivity maps, and improving memory and compute efficiency through the application of deep learning methods.



# Chapter 2

## Methods

### 2.1 The ZTE Forward Model

MRI reconstruction relies on a forward model to describe the process by which the observable  $k$ -space data is generated from the spatial distribution of nuclear spins within a subject. This model encapsulates the physics of MRI signal generation, including the effects of magnetic fields and coil sensitivities, and is instrumental in forming the basis for image reconstruction. In practical terms, the forward model translates the object represented in the spatial domain, into the data acquired in the frequency domain ( $k$ -space), using a series of transformations such as coil sensitivity multiplication, Fourier transforms, and modeling of system imperfections like the dead time gap.

We mathematically formulate the ZTE reconstruction problem as an inverse problem to solve for  $x$  in  $A(x) = y$ , where  $A$  is our forward model,  $y$  represents the acquired  $k$ -space with the dead time gap, and  $x$  is the object to reconstruct. The optimization goal is to minimize  $\|A(x) - y\|_2^2$  with the forward model defined as:

$$A = D * F_\theta * S$$

where  $S$  is a multiplication with coil sensitivity maps,  $F$  is the 3D Non-Uniform Fast Fourier Transform (NUFFT) onto the radial trajectory, and  $D$  models the dead time gap by removing the center of  $k$ -space.

### 2.2 Datasets

Preliminary 2D experiments use Cartesian MRI data from NYU FastMRI[28] brain and knee datasets. The dataset consists of Cartesian fully sampled brain MRIs obtained on 3 and 1.5 Tesla magnets and includes axial T1 weighted, T2 weighted, and FLAIR multi-coil images.

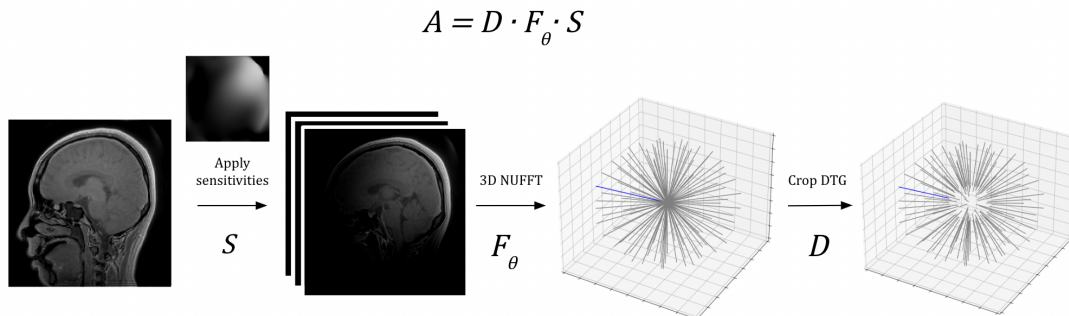


Figure 2.1: The ZTE forward model. A 3D volume  $x$  is multiplied with the coil sensitivity maps to generate coil images. A 3D NUFFT is applied to each coil image to generate radial  $k$ -space. The initial  $N$  samples of the  $k$ -space data, which are affected by the dead time gap, are typically discarded.

The 3D experiments leverage the mridata.org Fully sampled 3D FSE Knees dataset[6]. The images were acquired on 3T magnets using a Cartesian trajectory. The dataset consists of 19 volumes with 8 coils and FOV of 160 mm x 160 mm x 153.6 mm.

To address the constraints imposed by the limited size of the 3D dataset, which complicates the training of deep learning models, we developed a 3D synthetic dataset comprising of 300 noise-like multi-coil images. This approach draws inspiration from similar methodologies, such as those utilized in Resonet[9] and Synthmorph[12], which have shown that training on synthetic data can significantly enhance a model’s ability to generalize. The data generation technique follows the protocol outlined in Resonet. The synthetic images were designed to have similar statistical properties as real MR images. The random images are produced by applying a Fast Fourier Transform to exponentially weighted random complex data. To create the final image, multiple random images are combined to generated edges with various sharpness levels. To create multi-coil data, random sensitivity maps are generated from weighted random SPIRiT[17] kernels and then applied to the final image.

## 2.3 Simulated Training Data Generation

The availability of large 3D radial MRI datasets, let alone ZTE datasets, is limited. This scarcity poses significant challenges for training deep learning models that require extensive and diverse data. To address this limitation, ZTE data was simulated by applying the forward model  $A$  to the original Cartesian datasets. In our case, the radial trajectory  $F$  is the same trajectory used to collect in-vivo data. The DTG is simulated by removing the initial

samples from the radial data generated from the NUFFT.

Another complication of processing ZTE images arises from the radial trajectory used to acquire data. This polar coordinate system introduces complexities in data representation that are not inherently compatible with conventional deep learning techniques, especially convolutional neural networks (CNNs). CNNs, which excel in extracting patterns and features from data structured in Cartesian coordinates, struggle with the polar representation due to their reliance on convolutions that assume a uniform, grid-like spatial organization.

To mitigate this issue, the simulated ZTE data is gridded back to Cartesian coordinates using a density-compensated adjoint NUFFT. The same gridding process is applied during inference on the in-vivo ZTE data. Although the adjoint NUFFT introduces minor gridding artifacts, these are considered negligible in the broader context of the analysis, as all data – simulated and in-vivo – undergoes this procedure. In addition to simply gridding the data, the adjoint NUFFT is used to reduce the resolution of the image to accommodate memory and compute limitations during training. In this instance, the resolution is reduced by a factor of 4. The high-frequency  $k$ -space information is saved prior to performing the adjoint NUFFT so that the full high-resolution image can be reconstructed from the low-resolution model output.

To complete the preprocessing, the data was normalized to the 99th percentile, and the magnitude and phase information were separated across 2 channels. For the 3D data, the volumes were masked prior to normalization with a circular mask to create a finite structure and field of view (FOV).

## 2.4 In-Vivo Acquisition

To evaluate our approach on in-vivo data, we acquired a head scan of a healthy volunteer using a 12-channel head coil on a GE-3T MR750W (with  $\pm 31.25\text{kHz}$  rBW, 2.3ms TR, flip angle 2 degrees, 24 cm FOV). We discarded the first one or two samples along each acquired spoke to avoid effects from coil ringing and ripple from the digital filter applied by the scanner receiver unit. We retrospectively removed 3 additional readout samples per spoke for a final dead time gap of 5 samples to demonstrate the network’s capabilities on a larger dead time gap.

# Chapter 3

## Learning Frameworks

We investigated three distinct deep learning frameworks suitable for ZTE MRI reconstruction: UNet[21][4], Interlacer[22], and Diffusion[23][11]. Each framework offers unique approaches to handling the complex data structures inherent in ZTE imaging. The UNet model, known for its effectiveness in medical image segmentation, is adapted here for image reconstruction by leveraging its powerful encoder-decoder architecture. The Interlacer model introduces a novel approach performing joint learning over the image and frequency domains. Lastly, the Diffusion model, typically used for generative tasks, is explored for its potential to iteratively refine MRI reconstructions via denoising. This chapter provides the foundational understanding necessary to understand the strengths and potential limitations of each model in the context of ZTE reconstruction.

### 3.1 UNet Architecture

UNet models have been shown to be highly effective at extracting features from medical imaging data. In the following experiments, a 4-layer UNet is used to reconstruct the missing center of  $k$ -space in simulated ZTE images. This pipeline can be used with both a 2D UNet or a 3D UNet depending on input dimensions.

Since it is difficult to estimate sensitivity maps from coil data containing a DTG, it is difficult to generate a coil-combined image that includes phase directly from ZTE coil data. To navigate this issue, the model reconstructs the missing center of  $k$ -space for each coil individually by stacking the coil dimension with the batch dimension. The input into the network is thus structured as  $(B * N, 2 + C, H, W, D)$  where  $B$  represents the batch size,  $N$  denotes the number of coils,  $C$  signifies the channel size of additional inputs, and  $D$  is the optional depth dimension for the 3D model. The model supports the inclusion of additional data in the channel dimension, such as the root sum of squares (RSS) of the coils, which provides more contextual information, including the FOV, to enhance the final reconstruction.

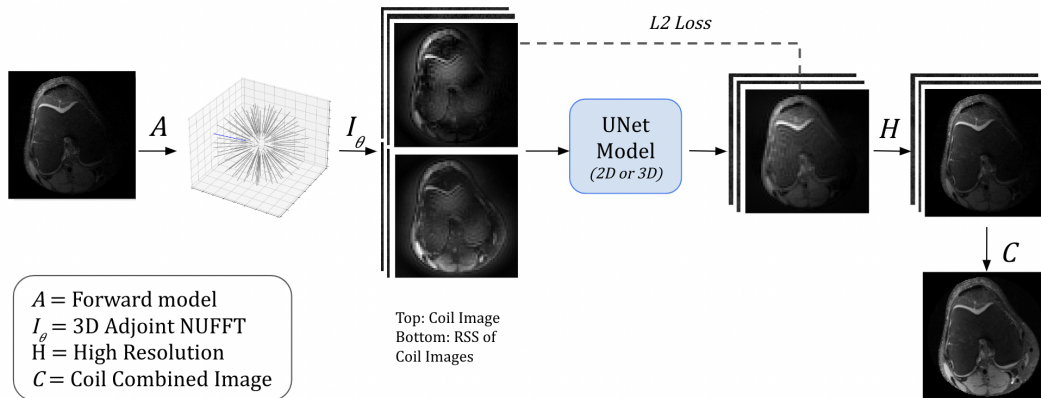


Figure 3.1: The UNet model pipeline. The ZTE forward model  $A$  is applied to a fully sampled coil-combined image to generate low-resolution coil  $k$ -space containing a dead time gap. An adjoint NUFFT  $I_\theta$  is then applied to yield gridded low-resolution coil images. Each coil image is individually input into the UNet alongside any additional information (in this case the RSS of all the coils from that image). The model outputs the reconstructed low-resolution coil.  $H$  combines the saved high-frequency information of the coil with the low-resolution reconstructed coil. The high-resolution reconstructed coils can be combined via RSS or by estimating sensitivity maps and performing optimal coil combining[1].

The model reconstructs the missing center of  $k$ -space of each coil individually; the output reconstructed coil is of dimensions  $(B * N, 2, H, W, D)$ . The loss is computed between the reconstructed coil and the original coil input. This reconstructed data can then be utilized to estimate sensitivity maps, enabling the computation of a complex-data coil-combined image that retains both amplitude and phase information.

## 3.2 Interlacer Architecture

Another architecture that proved more effective at reconstructing undersampled MRI data compared to techniques that treat frequency and image domain separately is the Interlacer model proposed by Singh et al[22]. This model consists of two networks that convolve separately over the image and frequency domain. The information from each network is shared with the other at each layer in an interleaving fashion.

This model learns features in both  $k$ -space and the image domain solely based on a coil image input. Additionally, since the model performs convolutions on both  $k$ -space and the

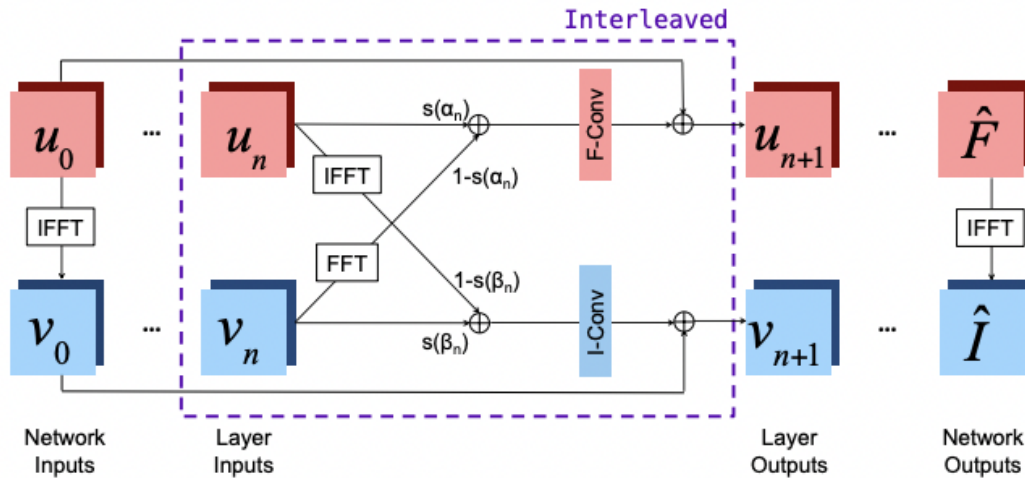


Figure 3.2: (reproduced from Singh et al.[22]) The Interlaced architecture depicting one layer of the model. The model takes in a coil image  $u_0$  from which its  $k$ -space  $v_0$  is computed via an inverse Fourier transform. Each network then calculates a weighted combination of data from both the image and frequency domains after which it applies additional convolutions.

image, it can learn features of both domains at different resolutions. The final output can be specified to either be the output from the image network or the  $k$ -space network.

This work leveraged a 6-layer Interlaced model with 9x9 convolution kernels and intermediary layers consisting of 64 channels. The model was configured to output the reconstruction in the image domain and an L2 loss was computed on the reconstructed coil image and the input coil image.

### 3.3 Diffusion Architecture

More recently, numerous works[15][2] have explored diffusion generative models for MRI reconstruction. Diffusion models can be thought as 2 different processes:

1. The forward process where noise is incrementally added to a sample from the data distribution until only noise remains.
2. The reverse process which methodically removes noise to reconstruct the original image. This reverse process is facilitated by a neural network that has learned the specifics of how noise was added during the forward process.

The intuition of diffusion models is that training a model to add noise to an image inherently teaches it how the data transitions from a structured, meaningful state to a randomized

noise state. In understanding this transition, the model implicitly captures the underlying data distribution. The knowledge acquired during the forward process equips the model to restore the original image from a noise-corrupted state effectively. This approach leverages the learned data distribution as a prior, enabling the reconstruction of high-quality images even from significantly corrupted initial data.

The forward process can be thought of as a stochastic differential equation (SDE) where noise is progressively added to an image at each step:

$$d\mathbf{x} = \frac{s'(t)}{s(t)} \mathbf{x} dt + s(t) \sqrt{2\sigma'(t)\sigma(t)} d\omega$$

In this equation,  $\frac{s'(t)}{s(t)}$  is the drift coefficient, which determines how the data  $x$  changes deterministically over time. The term  $s(t)\sqrt{2\sigma'(t)\sigma(t)}$  represents the diffusion coefficient, which scales the random fluctuations added to the data via the Wiener process  $d\omega$ , and  $d\omega$  itself is the Brownian motion process that introduces these stochastic variations.

The reverse process is defined by the ordinary differential equation (ODE):

$$d\mathbf{x} = \left( \frac{s'(t)}{s(t)} \mathbf{x} - s^2(t)\sigma'(t)\sigma(t)\nabla_{\mathbf{x}} \log p(\tilde{\mathbf{x}}; \sigma(t)) \right) dt$$

where  $\tilde{\mathbf{x}} = \frac{\mathbf{x}}{s(t)}$  represents the image data being recovered. It is important to note that  $s(t)$  and  $\sigma(t)$  are parameters analytically defined in the forward process. Thus, the only portion of the reverse process that needs to be learned is the score  $\nabla_{\mathbf{x}} \log p(\mathbf{x}; \sigma(t))$  at each step. This score function is typically learned by training a neural network  $D_{\theta}(\mathbf{x}; \sigma(t_i))$ [25], which is adept at capturing the complex dependencies involved in reversing the diffusion process.

With an estimated score function, a denoised data point can be sampled from the posterior distribution  $p(\mathbf{x}|\mathbf{y})$ . One of the key challenges in this approach is that the likelihood  $p(\mathbf{y}|\mathbf{x}, \sigma(0))$  is only analytically known at the initial time  $t = 0$ . To address this, the method leverages diffusion posterior sampling techniques. These techniques assume that the intermediate steps approximate the likelihood to be  $p(\hat{\mathbf{x}}; \sigma(0))$ , where  $\hat{\mathbf{x}}$  is an estimate of the denoised image at time  $t = 0$  and is given by Tweedie's formula:

$$E[\mathbf{x}(0)|\mathbf{x}(t)] = \frac{1}{s(t)} \left( \mathbf{x}(t) + s^2(t)\sigma^2(t)\nabla_{\mathbf{x}} \log p(\mathbf{x}(t); \sigma(t)) \right)$$

In addition to denoising the image, the diffusion model is constrained to ensure fidelity to the original input through data consistency. When applying the forward model  $A$  to the output of the model, the result should correspond closely with the original input data.

With this, the full inference algorithm is presented in Algorithm 1.

**Algorithm 1** Diffusion Posterior Sampling (DPS)

---

```

1: procedure DPS( $D_\theta(\mathbf{x}; \sigma(t)), \sigma(t), s(t), t_{i \in \{0, \dots, N\}}, \mathbf{y}$ )
2:    $x(t_0) \sim N(0, \sigma^2(t_0) s^2(t_0) I)$  ▷ Initialize the noise
3:   for  $i \leftarrow 0$  to  $N - 1$  do
4:      $\tilde{\mathbf{x}} \leftarrow \frac{\mathbf{x}(t_i)}{s(t_i)}$ 
5:      $\nu_i \leftarrow \frac{\sigma'(t_i)}{\sigma(t_i)} + \frac{s'(t_i)}{s(t_i)}$ 
6:      $\zeta_i \leftarrow \frac{\sigma'(t_i) s(t_i)}{\sigma(t_i)}$ 
7:      $\mathbf{d}_P \leftarrow \nu_i \mathbf{x}(t_i) - \zeta_i D_\theta(\tilde{\mathbf{x}}; \sigma(t_i))$ 
8:      $\hat{\mathbf{x}} \leftarrow \frac{1}{s(t_i)} \left( \mathbf{x}(t_i) + s(t_i)^2 \sigma(t_i)^2 \frac{D_\theta(\tilde{\mathbf{x}}; \sigma(t_i)) - \mathbf{x}(t_i)}{\sigma(t_i)^2} \right)$ 
9:      $\mathbf{d}_L \leftarrow \nabla_x \|A\hat{\mathbf{x}}(\mathbf{x}(t_i)) - \mathbf{y}\|_2^2$ 
10:     $\mathbf{x}(t_{i+1}) \leftarrow \mathbf{x}(t_i) + (t_{i+1} - t_i) \mathbf{d}_P + \gamma_t \mathbf{d}_L$ 
11:  end for
12: end procedure

```

---

Diffusion models appear to be highly suitable for ZTE reconstruction, building on their proven effectiveness in traditional MRI reconstruction tasks. The ZTE forward model essentially involves a specific type of undersampling. Given this context, a diffusion model that is adeptly trained on a diverse array of MRI images is likely to excel in reconstructing ZTE data. By learning from the broader MRI data distributions, these models can adapt their reconstruction capabilities to effectively address and correct the undersampling in ZTE images.



# Chapter 4

## 2D Results

This chapter presents preliminary results of applying deep learning frameworks to simulated 2D ZTE MRI reconstruction, leveraging the greater availability of 2D data to optimize and test these models. However, it is important to note a significant limitation: there is currently no established in-vivo use case for 2D ZTE imaging. This gap underscores the experimental nature of our findings and suggests a need for caution in directly extrapolating these results to clinical scenarios.

### 4.1 2D UNet and Interlacer Training Details

For the following 2D data experiments, the UNet and Interlacer models were trained on 64x64 images with a batch size of 8. Training used the synthetic ZTE data with a DTG of 5 generated from the FastMRI brain dataset (see Figure 4.1). The dataset was pre-generated to optimize memory usage and speed during training. The models were trained for 200 epochs with a learning rate of 1e-3 and using the Adam optimizer[14]. All training used Pytorch[18] on an NVIDIA Titan Xp GPU.

### 4.2 2D Diffusion Training Details

The Diffusion model was trained on the 5000 middle slices from the FastMRI brain dataset. The images were center cropped to be 256x256 pixels. The model adopted a UNet style architecture as described in Karras et al.[13], configured with 2 input/output channels to effectively handle both the real and imaginary components of the MRI images. We used a batch size of 8 for training the diffusion model and training was done across 2 NVIDIA GeForce RTX 3090 lasting 2 days. During inference, ZTE data with a DTG of 5 was simulated from the FastMRI brain dataset by applying the forward model described in Chapter 2. Inference was performed on one NVIDIA Titan Xp GPU.

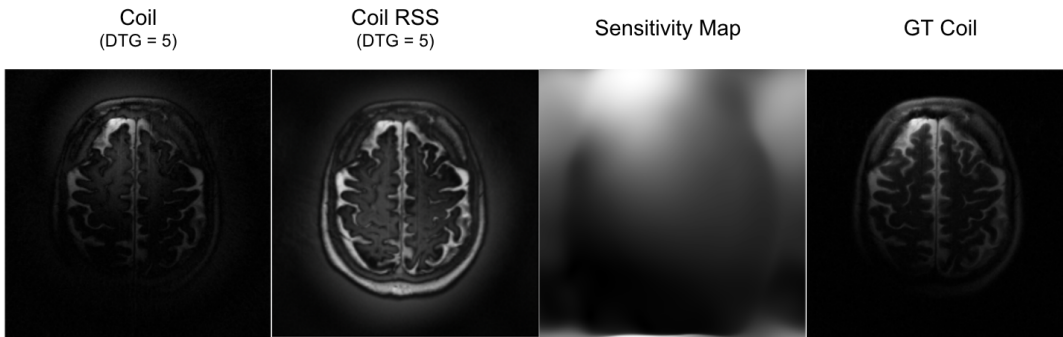


Figure 4.1: An example 2D input in its original resolution for clarity. The sensitivity map was calculated from the ground truth coil image shown on the right.

### 4.3 2D Simulated Data UNet Experiments

An initial set of experiments focused on identifying what additional input data to the UNet is necessary to yield the most accurate reconstruction results. We explored three input options: single coils with no additional information, single coils with their sensitivity maps — which were calculated from fully sampled ground truth data — and single coils paired with RSS of all the coils. The incorporation of additional input data serves to provide the model with extra contextual information about the FOV and structure of the final image, enhancing the overall reconstruction accuracy.

From the qualitative results presented in Figure 4.2, it is evident that the UNet trained with both coil data and additional RSS input performed best. Models trained solely on coil data, or on coil data combined with sensitivity maps, exhibited errors across both low and high-frequency ranges. Notably, the errors in the first two models were more pronounced and varied, affecting the depiction of anatomical details and resulting in higher overall error rates. In contrast, the model trained with coil data and RSS input demonstrated more uniform errors that were predominantly confined to the lower frequencies of the image. This suggests a more stable and reliable reconstruction.

To further confirm these qualitative results, the Normalized Root Mean Squared Error (NRMSE) and Peak Signal to Noise Ratio (PSNR) of the RSS of the model outputs were measured across 100 samples. Table 4.1 shows the mean and standard deviation of the NRMSE and PSNR across the 3 experiments.

Reconstruction using coil data only is a more difficult task since the coil sensitivity maps weigh the structure of the anatomy non-uniformly. Consequently, the full structure of the

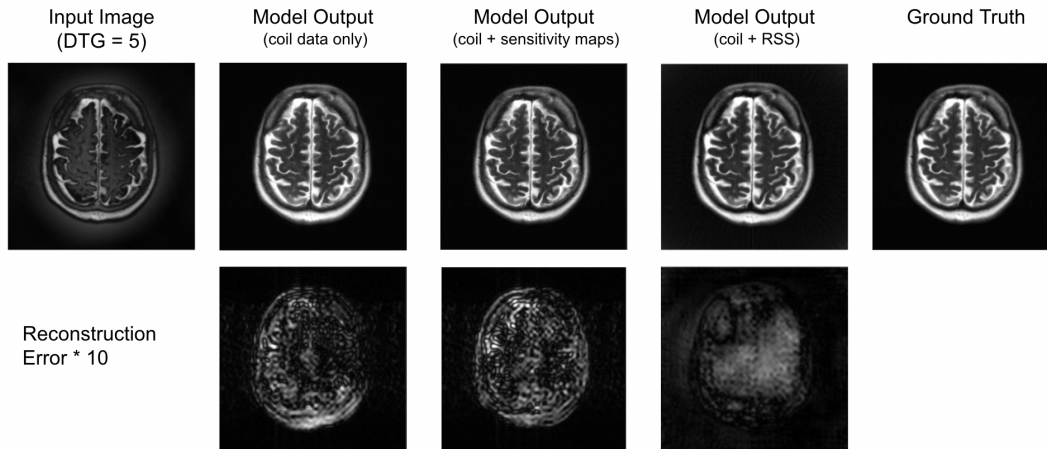


Figure 4.2: Reconstruction results on 3 UNet models trained with different inputs. Instead of presenting the raw multi-coil data, the images in this figure show the RSS of the multi-coil data. The model trained on coil images alone has the highest reconstruction error while the model trained on coil images with the RSS of all the coil inputs performs best.

UNet Input	NRMSE ( $\downarrow$ )	PSNR ( $\uparrow$ )
Coil Only	$0.0168 \pm 0.0084$	$36.02 \pm 3.15$
Coil + Sensitivities	$0.0169 \pm 0.0091$	$36.74 \pm 3.31$
Coil + RSS	<b><math>0.0087 \pm 0.0040</math></b>	<b><math>42.29 \pm 3.11</math></b>

Table 4.1: Performance comparison of various UNet inputs on validation set. NRMSE and PSNR were computed on 100 validation samples. The Coil + RSS UNet yielded the lowest NRMSE and highest PSNR which indicates the best performance.

coil input image often appears unclear, especially in the presence of larger dead time gaps as shown in Figure 4.1. Additionally, the fact that coils are input individually complicates the reconstruction task, as each coil’s data must be integrated accurately to form a cohesive image. More unexpectedly, the model trained on coil data and respective sensitivity maps also exhibited higher errors. Initially, the inclusion of the coil sensitivity map in the training input was intended to provide additional information about how the coil images are weighted, potentially enhancing the model’s ability to accurately reconstruct the anatomy. However, if the image affected by the DTG lacks substantial information about parts of the anatomy distant from the respective receiver coils, then merely adding information on how the image is weighted does not substantially aid the reconstruction process. This suggests that the effectiveness of sensitivity maps may be limited in scenarios where the underlying

coil data are inherently sparse or incomplete, which aligns with the observed increase in reconstruction errors.

Incorporating the RSS of the coils into the model input proved beneficial, as it provided critical information that enhanced the model’s performance. This result is beneficial in practice as the RSS model does not rely on additional information which is harder to acquire such as sensitivity maps. The inclusion of RSS data gives the model a clear understanding of the FOV of the image, which is crucial for accurate image reconstruction. Additionally, the RSS allows the model to better infer the final contrast from the coil image. Even though some low-frequency information may be missing due to the DTG, the sensitivity maps distribute this information across the other coils. By integrating the RSS, which includes information from all coils, the model can leverage these distributed low frequencies, leading to a more accurate and complete reconstruction of the image.

## 4.4 2D Comparison of Learning Frameworks

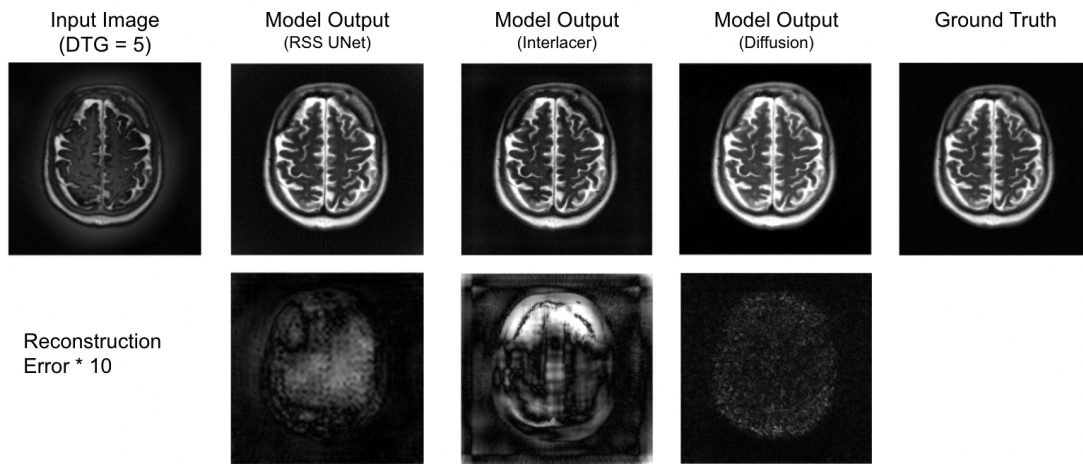


Figure 4.3: Reconstruction results across the 3 learning frameworks presented. Instead of presenting the raw multi-coil data, the images in this figure show the RSS of the multi-coil data. Only the best UNet model trained on coil + RSS data is shown as the other experiments are compared in Figure 5. The Interlacer model qualitatively shows the worst reconstruction error and the Diffusion model shows the best reconstruction error.

With the best UNet result established, we performed a comparison across the three learning frameworks discussed in Chapter 3. The key NRMSE and PSNR metrics were again computed across 100 data points and their means and standard deviations are presented in Table 4.2. The Diffusion and UNet frameworks exhibited very similar results quantitatively,

Method	NRMSE ( $\downarrow$ )	PSNR ( $\uparrow$ )
UNet (Coil + RSS)	$0.0087 \pm 0.0040$	<b><math>42.29 \pm 3.11</math></b>
Interlacer	$0.0225 \pm 0.0114$	$34.23 \pm 3.45$
Diffusion	<b><math>0.0075 \pm 0.0025</math></b>	$42.13 \pm 2.33$

Table 4.2: Performance comparison of the three learning frameworks on the validation set. The table shows the mean and standard deviation of the NRMSE and PSNR computed on 100 validation samples. The Diffusion model had the best NRMSE and the UNet model had the best PSNR

but the qualitative analysis in Figure 4.3 shows that the error from the Diffusion model resembled noise and is less biased than the error from the UNet model.

The Interlacer approach, which performs reconstruction based on only coil data as input, presents high errors in both the high and low frequencies of the output. Qualitatively and quantitatively, the results from the Interlacer are the least favorable, performing even worse than the UNet trained solely with coil data. Unlike the other methods, the Interlacer outputs present errors both within the anatomy and in the surrounding areas. This is due to the method’s use of convolutions in both the image and frequency domains. Consequently, this biases the quantitative metrics and suggests that a method to compute error only in the anatomy region might provide a more accurate comparison of the different frameworks.

The Diffusion approach demonstrated the best results, which is impressive given that the input is not in distribution with the training data, since the model was trained to denoise fully sampled data. Despite this mismatch, the observed error is minimal and predominantly noise-like, likely a residual effect from the denoising diffusion process. Overall, the errors are uniformly minimal across the image contrast, indicating that the model has efficiently learned to reconstruct images from coil data with a DTG. This achievement underscores the model’s robustness and adaptability in handling complex imaging scenarios beyond its initial training conditions. The limitation of this method is that it relies on knowledge of the sensitivity maps for data consistency. This introduces the need for a calibration scan to compute sensitivity maps from the auto-calibrating signal.

# Chapter 5

## 3D Results

This chapter focuses on 3D ZTE reconstruction, which represents the real use case in clinical practice. A significant challenge encountered in this chapter is the limited availability of 3D data. For instance, the 3D dataset explored in this study comprised only 20 volumes—a stark contrast to the 4470 scans available in the 2D brain dataset. Additionally, while it might be tempting to consider using a stack of 2D slices to augment 3D data, this approach is flawed for ZTE imaging. The continuity of the signal cannot be guaranteed across slices due to variations in slice thickness, which can affect the homogeneity of the reconstructed volume. Moreover, slicing through 3D radial data does not translate effectively into 2D planes, as the non-Cartesian nature of radial sampling complicates any straightforward slice-based reconstruction, disrupting the integrity and accuracy of the resulting images.

In the 2D experiments, the Diffusion model yielded the best results, followed closely by the UNet with RSS. Considering the complexities involved in training a 3D diffusion model, such as data and memory availability, a 3D UNet with RSS was selected for initial trials due to its comparatively straightforward and faster training process. This approach facilitated the efficient creation of a reliable performance baseline, setting the stage for further investigation into the more sophisticated 3D diffusion models.

### 5.1 3D Training Details

For the following 3D data experiments, a UNet model was trained on 64x64x64 images with a batch size of 4. The model received as input a coil image along with the RSS of all coils. One experiment used the synthetic ZTE data with a DTG of 5 generated from the Stanford Fullysampled 3D FSE Knees dataset. The other used random noise-like images generated by the method described in De Goyeneche et al.[9] The dataset was pre-generated to optimize memory usage and speed during training. The models were trained for 200 epochs with a learning rate of 1e-3 and using the Adam optimizer. All training used Pytorch on an NVIDIA Titan Xp GPU.

## 5.2 3D Simulated UNet Experiments

When comparing the results of the 3D UNets trained on knee data and noise, as presented in Figure 5.1 and 5.2, it is evident that the model trained on knee data exhibits superior reconstruction capabilities within the targeted area. The model trained on noise demonstrates a compromised ability to reconstruct the knee, and shows errors in both the high and low frequencies of the image. Errors are more pronounced in anatomical regions characterized by complex structures and significant variations in contrast. This highlights the advantage of task-specific training in deep learning models for medical image reconstruction.

An additional observation from the experiments is the model's limitation due to the use of a circular mask, which inherently restricts the full view of the knee's structure. Despite this, a notable advantage is that the reconstruction process did not introduce artifacts or alterations to the regions outside the masked area. This characteristic suggests that the model's learning is well-contained within the region of interest and preserves the integrity of surrounding tissues, ensuring that only the area within the mask is subjected to the reconstruction process.

## 5.3 In-vivo Experiments

To assess the usability of this model in a clinical setting, inference was performed on in-vivo data, as depicted in Figure 5.3, using the model trained on synthetic noise data. The model reconstructed images with enhanced brightness, offering a clearer visualization of brain structures such as the temporal lobes, respective glands, thalamus, hypothalamus, and ventricles. However, the brightness of the region outside the anatomy presented an aura-like artifact; this phenomenon was also present in the input data, suggesting that the model has not fully corrected this issue. Moreover, the model displayed incorrect contrast in the sinus, pharynx, and other orifices, which appear darkened yet not with the same fidelity as shown in the ground truth images. These areas of concern are highlighted in the figure, where the anatomy's contrast differs notably from the expected results, indicating a need for further model refinement to enhance clinical applicability.

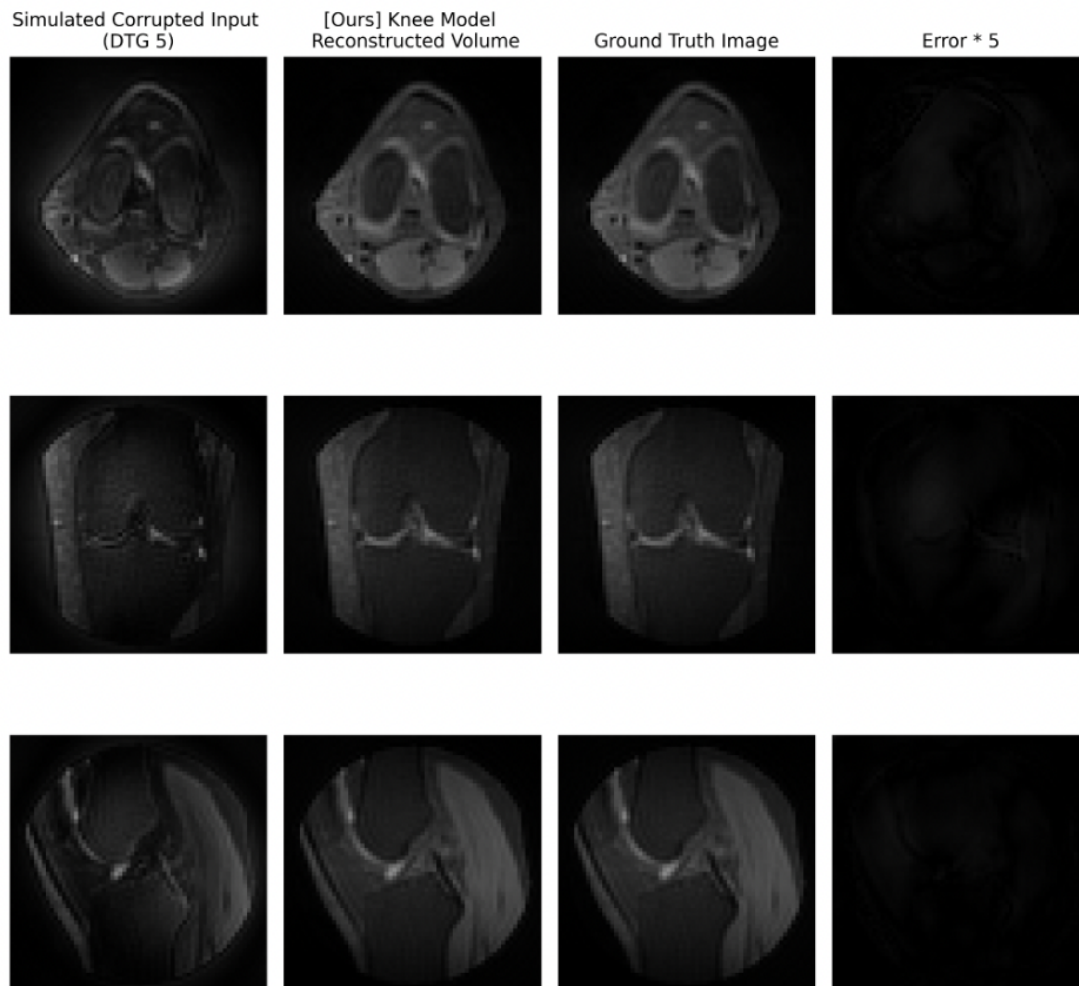


Figure 5.1: 3D Reconstruction of Simulated ZTE Knee Data with a 5-Sample Dead Time Gap using a UNet trained on 3D knee data. The reconstruction was individually performed on each coil image using a 3D UNet model, which received as input the individual coils and the RSS of all coils. Displayed from top to bottom are the axial, coronal, and sagittal views of the RSS from the reconstructed coil images, demonstrating accurate reconstruction across all orientations. The rightmost column illustrates the reconstruction error relative to the ground truth, amplified by a factor of 5 for clearer visualization. For this experiment, the reconstruction error is minimal.



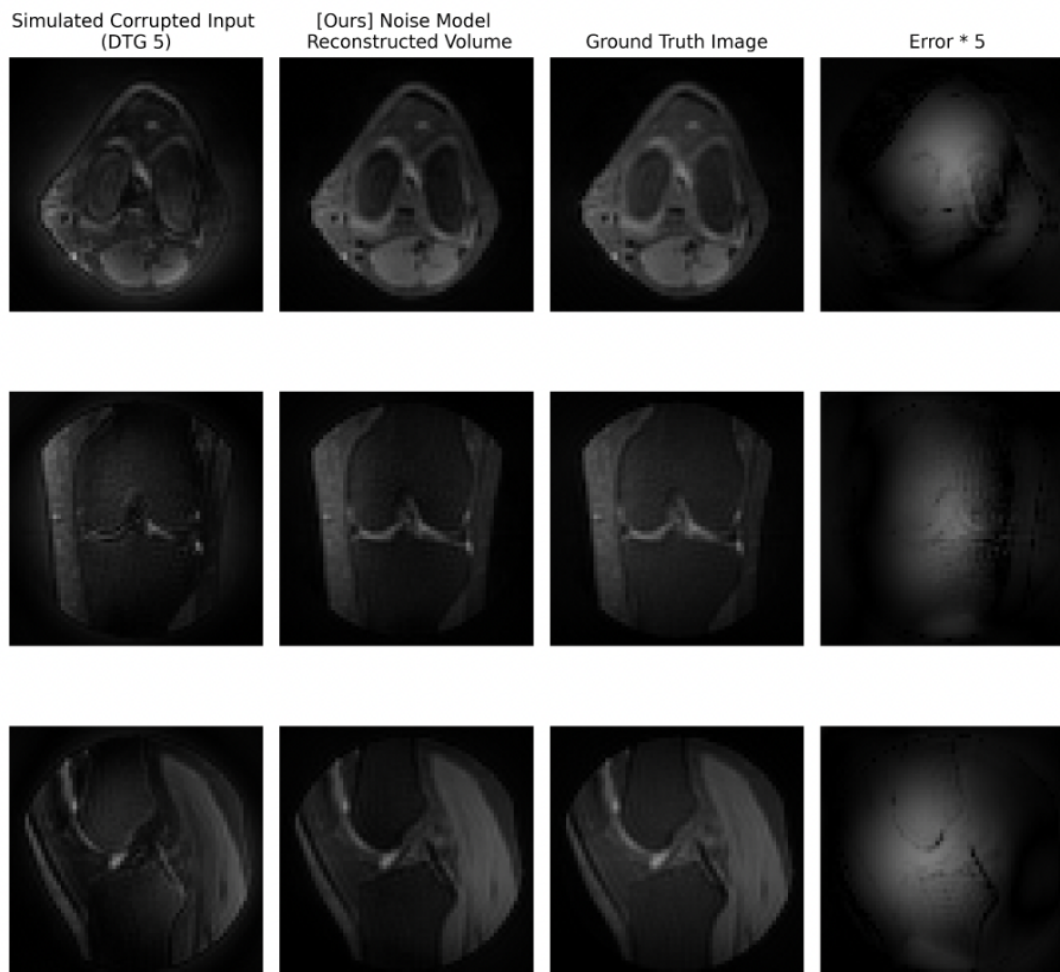


Figure 5.2: 3D Reconstruction of Simulated ZTE Knee Data with a 5-Sample Dead Time Gap using a UNet trained on 3D noise. The reconstruction was individually performed on each coil image using a 3D UNet model, which received as input the individual coils and the RSS of all coils. Displayed from top to bottom are the axial, coronal, and sagittal views of the RSS from the reconstructed coil images, demonstrating consistent reconstruction across all orientations. The rightmost column illustrates the reconstruction error relative to the ground truth, amplified by a factor of 5 for clearer visualization. For this experiment, errors are visible in both the higher and lower frequencies of the image, as evidenced by the visible anatomical details of the knee in the error plot.

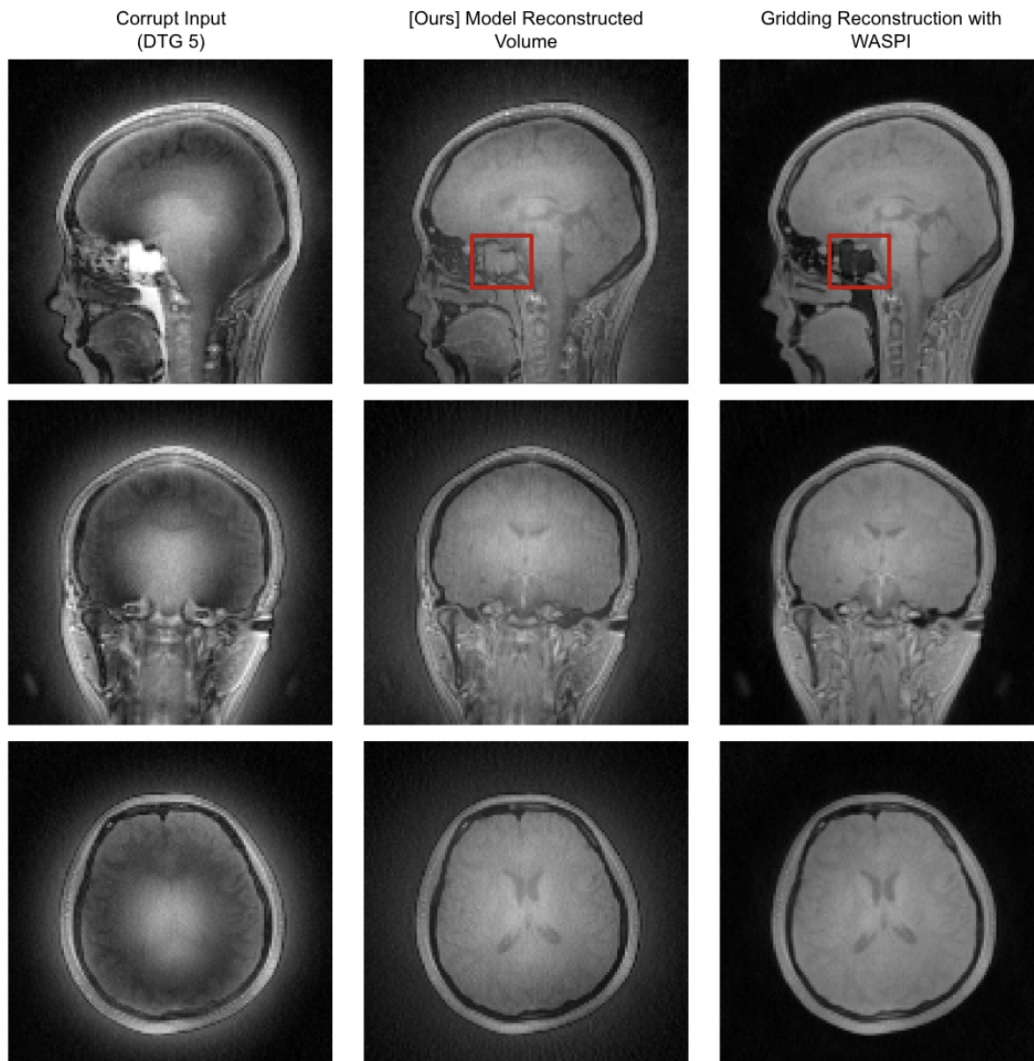


Figure 5.3: In-vivo ZTE Reconstruction using the 3D UNet model trained on synthetic data. A head scan was acquired using a 12-channel head coil on a GE-3T MR750W (with  $\pm 31.25\text{kHz}$  rBW, 2.3ms TR, flip angle 2 degrees, 24 cm FOV). The ground truth was reconstructed using additional acquisitions and WASPI. The region highlighted by the red box shows a failure point of the model where the reconstructed image is significantly brighter than the ground truth.

# Chapter 6

## Conclusion

### 6.1 Discussion

In exploring the 3D UNet methods for reconstructing ZTE data, the primary challenge faced was the model’s tendency to ”brighten” the entire field of view. This artifact was due to the divergence between the distribution of in-vivo data and the simulated data used for training, highlighting a critical limitation in the model’s ability to generalize to new, unseen data types. A first future study should explore the 3D UNet model’s effectiveness on in-vivo data that matches the training data exactly, meaning the same anatomy, FOV, resolution, number of coils, and ideally as close of a matching contrast as possible to the training set.

To enhance model robustness further, a next step involves implementing 3D diffusion techniques, building upon the success of 2D Diffusion models, which have demonstrated less bias in reconstruction results. One promising approach, as described by Chung et al. [3], reconstructs 3D volumes by processing each slice individually using a 2D Diffusion model. This strategy is necessitated by the high memory and computational demands of fully 3D Diffusion models. To ensure consistency across slices, the method incorporates a model-based prior in the z-direction during testing. This augmentation aims to align all 2D-generated slices accurately, ensuring they stack correctly to form a coherent 3D volume.

Another important analysis involves reconstructing outlier data, such as a ZTE scan of a brain with a tumor. While existing methods may accurately reconstruct in-vivo scans of healthy anatomy, they must also prove robust when dealing with such outliers, which could significantly affect contrast.

Lastly, leveraging the premise of convolutional spread can aid the model in recovering the low-frequency energy that gets spread to higher frequencies in the coil data, thus reducing reconstruction errors. If not limited by memory requirements, a more efficient strategy to process multi-coil data would be to input all coil data together rather than separating it

across the batch dimension. This unified approach could enhance the model's ability to learn from the complex spatial relationships inherent in multi-coil MRI data. Additionally, training the model to apply the inverse of the sensitivity maps to the coil  $k$ -space, in order to recover the original  $k$ -space data, presents a promising direction. However, this approach is limited by the need for precise sensitivity maps, which typically require additional calibration scans.

## 6.2 Conclusion

This thesis investigated the application of deep learning techniques for reconstructing ZTE MRI data, which is crucial for enhancing the practical applicability of ZTE in clinical settings. By not relying on additional acquisitions or extensive knowledge of sensitivity maps, our approach simplifies the ZTE reconstruction process, thus reducing scan times and enhancing patient comfort without sacrificing image quality.

Our experiments with simulated data have demonstrated effective results, suggesting that the proposed models can robustly estimate the missing center of  $k$ -space, thereby recovering high-quality images. This capability is instrumental in advancing ZTE MRI technology, making it a more viable option for rapid, dynamic contrast-enhanced imaging used in various medical scenarios.

In conclusion, this work not only enhances our understanding of the capabilities and limitations of deep learning in medical imaging reconstruction but also takes a pivotal step towards making advanced MRI techniques more accessible in everyday clinical practice. Future research will build on these foundations, further integrating deep learning to meet clinical demands, thus realizing the potential of ZTE MRI as a faster, quieter, and more patient-friendly imaging modality.

# Bibliography

- [1] M Bydder, DJ Larkman, and JV Hajnal. “Combination of signals from array coils using image-based estimation of coil sensitivity profiles”. In: *Magnetic Resonance in Medicine* 47.3 (2002), pp. 539–548. DOI: 10.1002/mrm.10092.
- [2] Hyungjin Chung and Jong Chul Ye. “Score-based diffusion models for accelerated MRI”. In: *Medical Image Analysis* 80 (2022), p. 102479.
- [3] Hyungjin Chung et al. “Solving 3D Inverse Problems Using Pre-Trained 2D Diffusion Models”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2023, pp. 22542–22551.
- [4] Özgün Çiçek et al. “3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation”. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*. Springer International Publishing, 2016, pp. 424–432. DOI: 10.1007/978-3-319-46723-8\_49.
- [5] Anagha Deshmane et al. “Parallel MR imaging”. In: *Journal of Magnetic Resonance Imaging* 36.1 (2012), pp. 55–72. DOI: <https://doi.org/10.1002/jmri.23639>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/jmri.23639>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/jmri.23639>.
- [6] K. Epperson et al. “Creation Of Fully Sampled MR Data Repository For Compressed Sensing Of The Knee”. In: *Proceedings of the 22nd Annual Meeting for Section for Magnetic Resonance Technologists*. Salt Lake City, Utah, USA, 2013.
- [7] L Feng. “Golden-Angle Radial MRI: Basics, Advances, and Applications”. In: *Journal of Magnetic Resonance Imaging* 56.1 (2022), pp. 45–62. DOI: 10.1002/jmri.28187.
- [8] R. Froidevaux, M. Weiger, and K. P. Pruessmann. “Algebraic reconstruction of missing data in zero echo time MRI with pulse profile encoding (PPE-ZTE)”. In: *Proceedings of the 30th Annual Meeting of ISMRM*. 2020.
- [9] Alfredo De Goyeneche et al. “ResoNet: a Physics-Informed DL Framework for Off-Resonance Correction in MRI Trained with Noise”. In: *Thirty-seventh Conference on Neural Information Processing Systems*. 2023. URL: <https://openreview.net/forum?id=Ia4dmqst0Z>.

- [10] David M. Grodzki et al. “Ultrashort echo time imaging using pointwise encoding time reduction with radial acquisition (PETRA)”. In: *Magnetic Resonance in Medicine* 67.2 (2012), pp. 510–518. DOI: 10.1002/mrm.23017.
- [11] Jonathan Ho, Ajay Jain, and Pieter Abbeel. “Denoising diffusion probabilistic models”. In: *Advances in Neural Information Processing Systems*. Vol. 33. 2020, pp. 6840–6851.
- [12] Malte Hoffmann et al. “SynthMorph: Learning Contrast-Invariant Registration without Acquired Images”. In: *IEEE Transactions on Medical Imaging* 41.3 (2021), pp. 543–558. DOI: 10.1109/TMI.2021.3116677.
- [13] Tero Karras et al. “Elucidating the design space of diffusion-based generative models”. In: *Advances in Neural Information Processing Systems*. Vol. 35. 2022, pp. 26565–26577.
- [14] Diederik P Kingma and Jimmy Ba. “Adam: A Method for Stochastic Optimization”. In: *arXiv preprint arXiv:1412.6980* (2014).
- [15] Brett Levac et al. “Accelerated motion correction with deep generative diffusion models”. In: *arXiv preprint arXiv:2211.00199* (2022).
- [16] Emma Ljungberg et al. “Silent zero TE MR neuroimaging: Current state-of-the-art and future directions”. In: *Progress in Nuclear Magnetic Resonance Spectroscopy* 123 (Mar. 26, 2021), pp. 73–93. DOI: 10.1016/j.pnmrs.2021.03.002. URL: <https://doi.org/10.1016/j.pnmrs.2021.03.002>.
- [17] Michael Lustig and John M. Pauly. “SPIRiT: Iterative Self-Consistent Parallel Imaging Reconstruction from Arbitrary k-Space”. In: *Magnetic Resonance in Medicine* 64.2 (2010), pp. 457–471. DOI: 10.1002/mrm.22428.
- [18] Adam Paszke et al. *PyTorch: An Imperative Style, High-Performance Deep Learning Library*. NeurIPS. Available: <https://pytorch.org/>. 2019.
- [19] Klaas P. Pruessmann et al. “SENSE: Sensitivity Encoding for Fast MRI”. In: *Magnetic Resonance in Medicine* 42.5 (Nov. 1999), pp. 952–962.
- [20] Shreya Ramachandran et al. “T1-weighted ZTE MRI with phase-modulated RF and a parallel imaging, inverse problem-based reconstruction”. In: *Proceedings of the 31st Annual Meeting of ISMRM*. 2022.
- [21] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. “U-Net: Convolutional Networks for Biomedical Image Segmentation”. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Ed. by Nassir Navab et al. Vol. 9351. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2015, pp. 234–241. DOI: 10.1007/978-3-319-24574-4\_28.
- [22] Nalini Singh et al. “Joint Frequency and Image Space Learning for MRI Reconstruction and Analysis”. In: *Machine Learning for Biomedical Imaging* (2022).

- [23] Jascha Sohl-Dickstein et al. “Deep unsupervised learning using nonequilibrium thermodynamics”. In: *International conference on machine learning*. PMLR. 2015, pp. 2256–2265.
- [24] Martin Uecker et al. “ESPIRiT—an eigenvalue approach to autocalibrating parallel MRI: where SENSE meets GRAPPA”. In: *Magnetic Resonance in Medicine* 71.3 (2014), pp. 990–1001. DOI: 10.1002/mrm.24751.
- [25] Pascal Vincent. “A connection between score matching and denoising autoencoders”. In: *Neural Computation* 23.7 (2011), pp. 1661–1674. DOI: 10.1162/NECO\_a\_00142.
- [26] Tobias C. Wood, Emil Ljungberg, and Mark Chiew. “ZTE Infilling From Auto-calibration Neighbourhood Elements”. In: *Journal Name* (2023). Neuroimaging, King’s College London, London, United Kingdom; Wellcome Centre for Integrative Neuroimaging, FMRIB, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, United Kingdom.
- [27] Yaotang Wu et al. “Water- and fat-suppressed proton projection MRI (WASPI) of rat femur bone”. In: *Magnetic Resonance in Medicine* 57.3 (2007), pp. 554–567. DOI: 10.1002/mrm.21174.
- [28] Jure Zbontar et al. *fastMRI: An Open Dataset and Benchmarks for Accelerated MRI*. <https://fastmri.org/>. 2018.