

Accelerating Brain DTI and GYN MRI Studies Using Neural Network

by

Yuhao Yan

Medical Physics Graduate Program
Duke University

Date: _____

Approved:

Fang-Fang Yin, Co-advisor

Zheng Chang, Co-advisor

Junzo Chino

Oana Craciunescu

Chunhao Wang

Thesis submitted in partial fulfillment of
the requirements for the degree of
Master of Science in the Graduate Program in
Medical Physics in the Graduate School
of Duke University

2021

ABSTRACT

Accelerating Brain DTI and GYN MRI Studies Using Neural Network

by

Yuhao Yan

Medical Physics Graduate Program
Duke University

Date: _____

Approved:

Fang-Fang Yin, Co-advisor

Zheng Chang, Co-advisor

Junzo Chino

Oana Craciunescu

Chunhao Wang

An abstract of a thesis submitted in partial fulfillment of the requirements for the degree of Master of Science Graduate Program in Medical Physics in the Graduate School of Duke University

2021

Copyright by
Yuhao Yan
2021

Abstract

There always exists a demand to accelerate the time-consuming MRI acquisition process. Many methods have been proposed to achieve this goal, including deep learning method which appears to be a robust tool compared to conventional methods. While many works have been done to evaluate the performance of neural networks on standard anatomical MR images, few attentions have been paid to accelerating other less conventional MR image acquisitions.

This work aims to evaluate the feasibility of neural networks on accelerating Brain DTI and Gynecological Brachytherapy MRI. Three neural networks including U-net, Cascade-net and PD-net were evaluated. Brain DTI data was acquired from public database RIDER NEURO MRI while cervix gynecological MRI data was acquired from Duke University Hospital clinic data. A 25% Cartesian undersampling strategy was applied to all the training and test data. Diffusion weighted images and quantitative functional maps in Brain DTI, T1-spgr and T2 images in GYN studies were reconstructed. The performance of the neural networks was evaluated by quantitatively calculating the similarity between the reconstructed images and the reference images, using the metric Total Relative Error (TRE).

Results showed that with the architectures and parameters set in this work, all three neural networks could accelerate Brain DTI and GYN T2 MR imaging. Generally,

PD-net slightly outperformed Cascade-net, and they both outperformed U-net with respect to image reconstruction performance. While this was also true for reconstruction of quantitative functional diffusion weighted maps and GYN T1-spgr images, the overall performance of the three neural networks on these two tasks needed further improvement. To be concluded, PD-net is very promising on accelerating T2-weighted-based MR imaging.

Future work can focus on adjusting the parameters and architectures of the neural networks to improve the performance on accelerating GYN T1-spgr MR imaging and adopting more robust undersampling strategy such as radial undersampling strategy to further improve the overall acceleration performance.

Contents

Abstract	iv
List of Tables	ix
List of Figures	x
Acknowledgements	xii
1. Introduction.....	1
1.1 MRI acceleration strategies	2
1.1.1 Efficient manipulation of spin dynamics	2
1.1.2 Sparse sampling of k-space.....	3
1.2 Fast MRI algorithms.....	4
1.3 Deep learning basics.....	5
1.3.1 Neural network layers.....	7
1.3.1.1 Convolution layer.....	7
1.3.1.2 Residual block.....	8
1.3.1.3 Pooling layer	9
1.3.1.4 Transpose convolution layer.....	9
1.3.1.5 Data consistency layer	10
1.3.2 Activation function	11
1.4 Deep learning on MR image reconstruction.....	12
1.5 Diffusion Tensor Imaging (DTI).....	13
1.5.1 Diffusion weighted imaging (DWI).....	13

1.5.2	Diffusion tensor	16
1.5.3	Quantitative functional map.....	18
1.5.4	Motivation to accelerate DTI	19
1.6	Gynecologic brachytherapy MRI study (GYN).....	19
1.7	Project objective	20
2.	Method.....	21
2.1	Raw data	22
2.1.1	DTI dataset: RIDER NEURO MRI.....	22
2.1.2	GYN dataset: Duke GYN MRI study.....	22
2.2	Data processing.....	23
2.2.1	k-space data acquisition	23
2.2.2	Undersampled Image	23
2.2.3	File format conversion and partitioning	25
2.2.4	Dataset partitioning	25
2.3	Neural network architecture.....	26
2.3.1	U-net.....	27
2.3.2	Cascade-net	28
2.3.3	PD-net	29
2.4	Training.....	30
2.5	Reconstruction	31
2.6	Evaluation.....	33
3.	Results.....	34

3.1	DTI reconstruction.....	34
3.1.1	Qualitative evaluation	34
3.1.2	Quantitative evaluation.....	41
3.2	GYN reconstruction.....	43
3.2.1	Qualitative evaluation	43
3.2.1.1	T1-spgr reconstruction.....	43
3.2.1.2	T2 reconstruction	47
3.2.2	Quantitative evaluation.....	53
4.	Discussion.....	54
5.	Conclusion.....	56
	Appendix.....	57
	Bibliography	58

List of Tables

Table 1: Dataset partitioning	26
Table 2: TRE evaluation of DTI reconstruction.....	42
Table 3: TRE evaluation of GYN reconstruction.....	53

List of Figures

Figure 1: An example of aliasing effect when image is undersampled using Cartesian strategy. (a): An example of Cartesian undersampling strategy. The white line represents data to be sampled. (b): The reference image, which is reconstructed from the fully sampled k-space data. (c): Image reconstructed from the undersampled k-space data using the undersampling strategy in (a). Aliasing artifact is obvious.....	4
Figure 2: An easy demonstration of neural network	7
Figure 3: Residual block.....	9
Figure 4: Data consistency layer	10
Figure 5: Sketch of Stejskal-Tanner diffusion weighted pulse sequence	14
Figure 6: An example of diffusion weighted images. (a): Image without diffusion gradient. Noting that it is T2-weighted. (b): Image with diffusion gradient, b-value set at 1000 s/mm ²	15
Figure 7: Illustration of diffusion ellipsoid.....	17
Figure 8: Examples of quantitative functional maps. (a): ADC map. (b): FA map. (c): Colored FA maps.	18
Figure 9: General workflow.....	21
Figure 10: Undersampling workflow with 25% Cartesian undersampling strategy. (a): An example of RIDER data. (b): An example of GYN data.	24
Figure 11: Architecture of U-net	27
Figure 12: Architecture of Cascade-net.....	28
Figure 13: Architecture of PD-net. \mathbf{f}_0 is initial guess of image space data, \mathbf{h}_0 is initial guess of k-space data, and \mathbf{g} is the true k-space data.	29
Figure 14: Qualitative evaluation for diffusion weighted images, b=0 s/mm ²	36
Figure 15: Qualitative evaluation for diffusion weighted images, b=1000 s/mm ² , gradient #5	37

Figure 16: Qualitative evaluation for ADC maps.....	38
Figure 17: Qualitative evaluation for FA maps.....	39
Figure 18: Qualitative evaluation for colored FA maps	40
Figure 19: Example DTI slices around eye level, from left to right: slices # 13-16, inferior to superior. Moderate distortions are visible.	41
Figure 20: Qualitative evaluation of GYN T1 reconstruction	44
Figure 21: Sample slices for GYN T1-spgr Cascade-net reconstruction. Noting that the applicators circled in reference images are not visible in reconstructed images.	45
Figure 22: Sample slices for GYN T1-spgr PD-net reconstruction. Noting that the applicators circled in reference images are not visible in reconstructed images.	46
Figure 23: Qualitative evaluation of GYN T2 reconstruction	49
Figure 24: Sample slices for GYN T2 Cascade-net reconstruction. Noting that the applicators circled in reference images are clearly visible in reconstructed images.	50
Figure 25: Sample slices for GYN T2 PD-net reconstruction. Noting that the applicators circled in reference images are clearly visible in reconstructed images.....	51
Figure 26 (a): HRCTV contour based on a pre-treatment T2 exam, denoted in red. (b): Sample slices for GYN T2 PD-net reconstruction relevant to HRCTV level. The reconstructed images are still blur to some extent.	52

Acknowledgements

I would like to express my sincere gratitude towards my advisor, Dr. Zheng Chang, for his careful education, patient guidance, warm encouragement when I encountered problems, and trust on my independency to solve the problems. I wish you all well. I would also like to thank my co-advisor, Dr. Fang-Fang Yin, for his instructions and advice on my later research and thesis. Also, many thanks to my MS thesis committee members, Dr. Junzo Chino, Dr. Oana Craciunescu, and Dr. Chunhao Wang, for their guidance and support through my thesis.

Thanks to my seniors, Xinyi Li and Yushi Chang, and my classmates, Yijie Yuan and Ruilin Li, et al., for their accompany and encouragement, especially throughout this unprecedented pandemic. Also, thank you all for your help with my academic courses and research. I have leant tons of knowledge from you.

I also would like to thank all the administrators and faculties at the Duke Medical Physics Graduate Program for providing such excellent education resources and research environment. I believe this would be my valuable education experience that would benefit me lifelong.

Last but not least, thanks to all my family and friends for their supports. Special thanks to Lei Huang for her mental support and spiritual guidance.

1. Introduction

Magnetic Resonance Imaging (MRI) is a non-invasive and non-ionizing imaging modality which is preferred for providing excellent soft-tissue contrast for anatomic and functional information, compared to other modalities like x-ray and CT. [1-2] It has been widely applied to imaging sites such as brain, breast, prostate and cervix, where soft-tissue contrast is appreciated.

Despite its advantages, MRI tends to cost a longer time to acquire all the data needed to reconstruct images compared to other modalities. [1-3] The imaging speed of MRI is essentially restricted by the MR physics and physiological constraints. [3] MR signals are acquired by repeatedly manipulating the spin dynamics of atoms with unpaired nuclei (typically Hydrogen atoms), where ideally excited spins are expected to recover to the equilibrium before the next manipulation. This recovery normally takes a long time. Besides, the repeated manipulation is realized by adopting rapidly changing gradients, which may lead to physiological problems like periphery nerve stimulation (PNS). [4]

Long acquisition time would lead to many problems other than low efficiency. Motion artifacts resulted from the patient's irregular respiration pattern and unconscious or involuntary move would be introduced to the images. In addition, long acquisition time makes it difficult to capture moving objects, e.g., the heart. [1] Furthermore, constantly being restrained within a closed space would lower the patient's comfort.

Finally, being placed in a rapidly changing magnetic field would increase the risk for the patient to suffer from physiological problems.

To be concluded, there exists an urgency to accelerate the MR imaging process, in other words, shorten the acquisition time and reduce the amount of data to be acquired while maintain comparable image quality.

1.1 MRI acceleration strategies

MRI acceleration strategies can be categorized as follows.

1.1.1 Efficient manipulation of spin dynamics

The efficient manipulation of spin dynamics is mainly achieved by developing fast pulse sequences. The acquisition time of MRI depends on the number and the length of repetition time (TR), which corresponds to sampling along phase encoding direction. Very conventionally, only one line in the k-space domain for each slice is filled during one TR, and the bulk magnetization is expected to fully recover to the equilibrium before the next acquisition. This would lead to a great number of long TRs, of which there is a great spatial and temporal redundancy. This redundancy can be exploited in several ways. One example is to fill more than one line in the k-space domain for each slice during one TR. Many fast pulse sequences fall under this category, e.g., Fast Spin Echo (FSE) [5] and Echo Planar Imaging (EPI) [6]. This category of methods is not the focus of this project and would not be discussed in detail.

1.1.2 Sparse sampling of k-space

Another way to accelerate the MR imaging process is to utilize the spatial redundancy and sparsely sample the k-space. This would significantly reduce the data to be collected, in other words, sampling along the phase encoding direction, which essentially shorten the acquisition time.

A common undersampling strategy is Cartesian undersampling strategy, which was utilized in this project. In this strategy, k-space is fully sampled along frequency encoding direction while sparsely sampled along phase encoding direction, in consideration that the acquisition time is governed by the number and the length of TRs, while the sampling along frequency encoding direction is typically very fast.

However, the Nyquist-Shannon criterion would be unavoidably violated during this undersampling process, which would lead to image artifacts such as aliasing. [1,3] The Nyquist-Shannon sampling theorem is originally described as “If a function $f(t)$ contains no frequencies higher than W cps, it is completely determined by giving its ordinates at a series of points spaced $1/2 W$ seconds apart.” [7] For spatial sampling in discrete system, which is the case of MR imaging, it can be described as that the spatial sampling rate should be larger than $1/\Delta X$, where ΔX is the pixel size along the sampling direction. Apparently, this would be violated if the sampling becomes sparse, which would lead to aliasing artifacts. **Figure 1** shows an example of aliased image adopting Cartesian undersampling strategy.

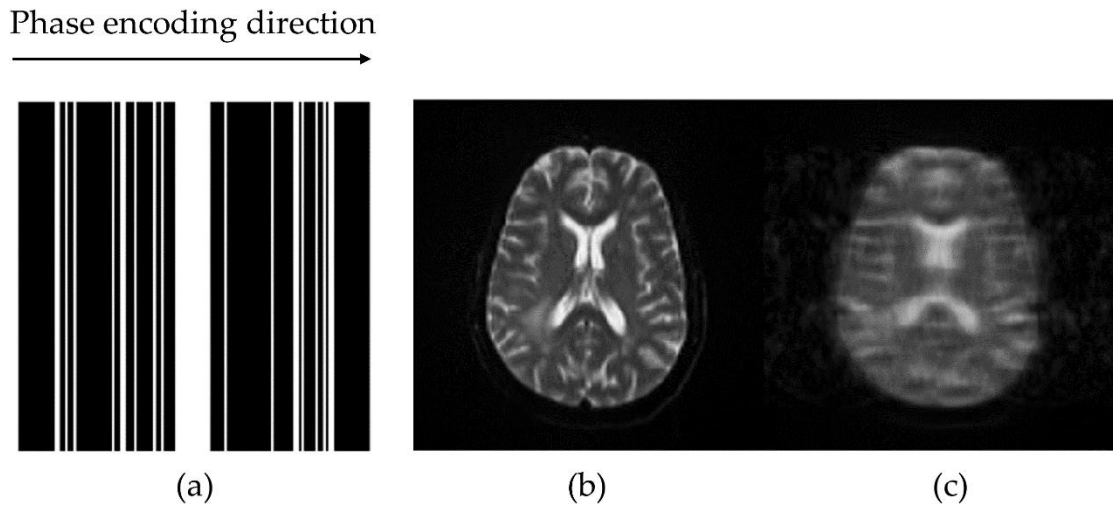


Figure 1: An example of aliasing effect when image is undersampled using Cartesian strategy. (a): An example of Cartesian undersampling strategy. The white line represents data to be sampled. (b): The reference image, which is reconstructed from the fully sampled k-space data. (c): Image reconstructed from the undersampled k-space data using the undersampling strategy in (a). Aliasing artifact is obvious.

1.2 Fast MRI algorithms

Many methods have been proposed to undersample the k-space data while reconstructing images with comparable image quality and free from aliasing artifact, for instance, Parallel Imaging (PI) [1] and Compressed Sensing (CS). [2-3]

There are several PI methods out there, such as Sensitive Encoding (SENSE) and Generalized Autocalibrating Partially Parallel Acquisition (GRAPPA). The common principle for PI is that the undersampled k-space data is acquired by an array of receiver coils simultaneously rather than a single receiver coil, each locates at different direction relative to the imaging object so that each is more sensitive to the partial volume nearest

to itself. Then an unaliased image is reconstructed by a specific algorithm, combining the undersampled data from all the receiver coils and corresponding sensitivities. [1]

The common principle for CS, on the other hand, is to iteratively reconstruct images from undersampled k-space data with flexible penalty. It can also be comprehend as finding the desired solution to a constrained optimization problem. [2-3] Filtration and data enforcement are often applied. Prior knowledge about the unique properties of the images is often appreciated to better facilitate the reconstruction. [2]

While PI and CS both show promising applications to MR image reconstruction, they both face their own limits. The problem of PI is that the noise would sabotage the image quality at high acceleration factor, while in CS an optimal solution to the constraint optimization problem might not always be achievable. [8]

Recently, deep learning has become a robust tool in the field of computer vision. Many deep neural networks have been applied to image reconstruction and have achieved considerable success.

1.3 Deep learning basics

Deep learning is mainly achieved by training and adopting the deep neural networks. Deep neural network, by its name, is a mimic of human neurological system. It is called deep because it contains multiple layers. In each layer, there exist a numerous of neurons, which can be abstractly assumed as a function that generates one output from several inputs. Typically, the deeper layer connects with the shallower layer in a way that

each neuron in the deeper layer takes all the output of the neurons in the previous layer with bias, then generate an output through a specific function (usually convolution), followed by being applied with an activation function. **Figure 2** is an easy version of neural network for demonstration purpose.

Due to the multiple layers and enormous neurons, deep neural networks can realize the learning of a giant, complicated function. Underlying abstract features of the data can be better extracted. These factors account for the high performance of deep neural networks. [9]

Several important layers and activation functions involved in this project are introduced below.

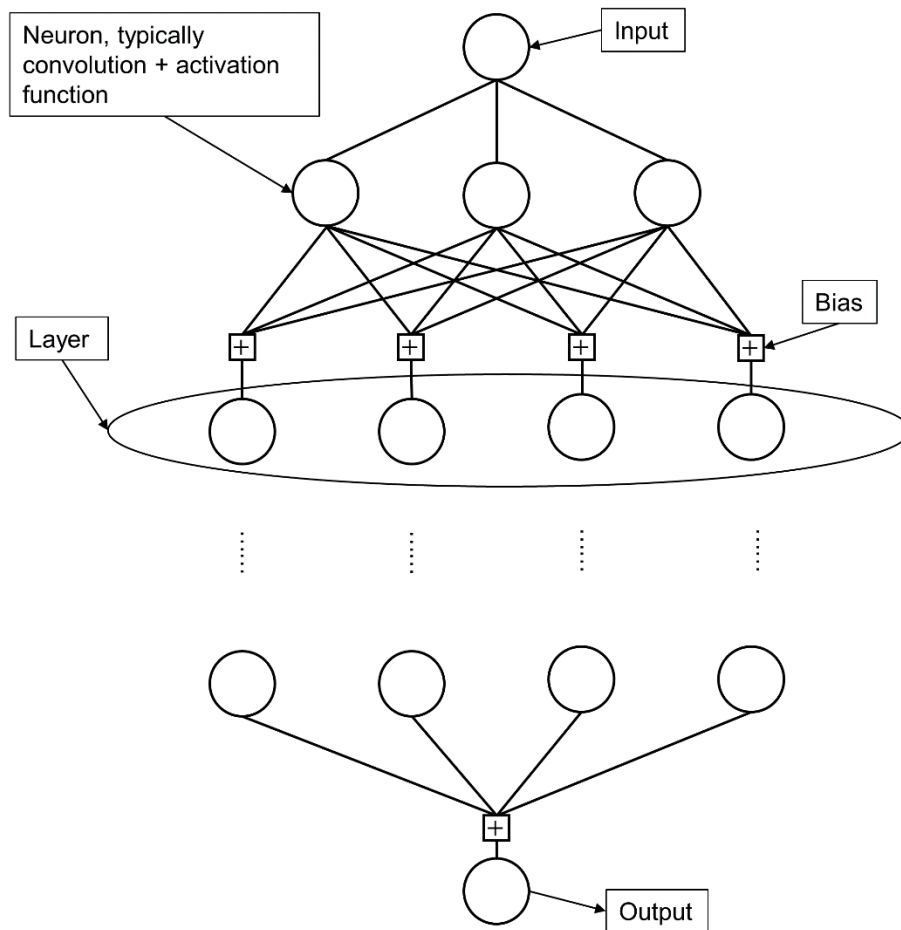


Figure 2: An easy demonstration of neural network

1.3.1 Neural network layers

1.3.1.1 Convolution layer

Convolution layer is the most important and signatred layer in deep learning. It is widely used in almost every deep learning architecture for different kinds of tasks, such as image classification, image segmentation and image reconstruction. It mainly serves to extract local features, for instance, edge information. [10]

Mathematically, convolution is the integral of one function and the 'flip' of another function. In discrete domain, it can be expressed as element-wise multiplication between an array and the flip of another array. However, in deep learning implement, there is no flip to the convolved function, which is called kernel in this case. The mathematic processing of convolution layer is pure element-wise multiplication. The convolution of the image and kernel can thus be expressed as:

$$y(i, j) = \sum_m \sum_n x(i + m, j + n) \times K(m, n) \quad (1)$$

Where x is the image pixel value before convolution, y is the image pixel value after convolution, and K is the convolution kernel.

Another thing needed to be mentioned is that this convolution would reduce the original image size. It might not be an issue in other tasks such as image classification, but since the task here is image reconstruction, image size is preferred unchanged. Thus, a zero padding is applied to every intermediate image to be convolved in this project. This is probably safe due to the fact that the periphery of MR images commonly contains no signals, in other words, pixel values are 0 at the boundary of MR images.

1.3.1.2 Residual block

Residual block is firstly proposed to fix the unexpected degradation problem aroused from the increase of layer depth. The architecture is shown as **Figure 3**. [11] By

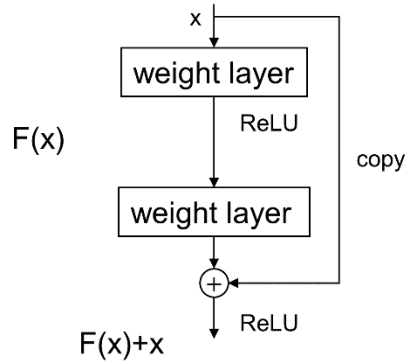


Figure 3: Residual block

copying the input into the output, the desired output turns to $F(x) + x$, which makes previous mapping $F(x)$ becomes residual mapping describing the changing of x indeed. It is believed that such residual mapping with a reference of input x is easier to optimize compared to the original unreferenced mapping. [11]

1.3.1.3 Pooling layer

Pooling layer is often used to reduce the image size combined with many types of mathematical processing. [12] A kernel size and a stride describing the distance between two consecutive positions of pooling kernels should be assigned. Max pooling is a common pooling layer, which outputs the maximum value among a selected grid in the input image.

1.3.1.4 Transpose convolution layer

Transpose convolution is also referred as up-sampling or deconvolution. It is realized by purely switching the passing order of the convolution, [12] in other words, switching the mapping direction. It is usually used to enlarge the image size.

1.3.1.5 Data consistency layer

Data consistency layer is a unique layer specifically designed for MR image reconstruction. The idea is similar to data enforcement in CS. Data consistency layer introduce the original k-space data, which can be viewed as true information, into the neural network to improve the accuracy and prevent overfitting. The layer structure is shown in **Figure 4**. [13]

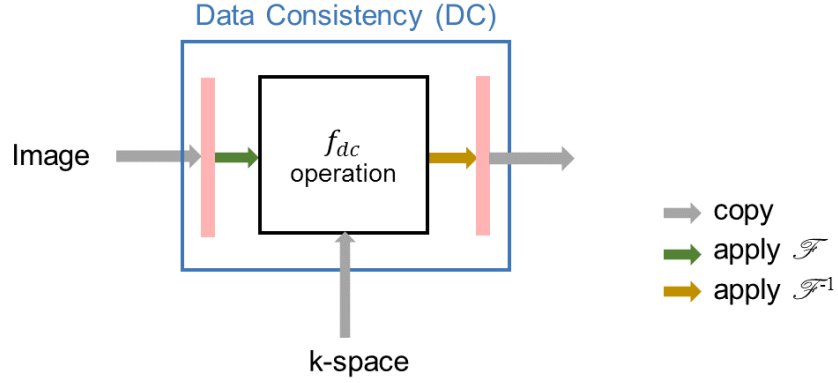


Figure 4: Data consistency layer

The image input is firstly Fourier transformed into k-space data, and then undergo the f_{dc} operation. f_{dc} operation is the main step to realize data consistency, which can be essentially expressed as [13]

$$S_{rec}(j) = \begin{cases} S_{cnn}(j), & \text{if } j \notin \Omega \\ \frac{S_{cnn}(j) + \lambda S_0(j)}{1 + \lambda}, & \text{if } j \in \Omega \end{cases} \quad (2)$$

Where S_{rec} is the reconstructed k-space data after operation, S_{cnn} is the k-space data transformed from the image predicted by the previous convolutional neural network

(CNN), S_0 is the original k-space data, Ω is the original sampled space, and $\lambda = q/\sigma$, q is a hyperparameter and σ^2 is the power of additive white gaussian noise (AWGN).

If $j \in \Omega$, which means the k-space data has originally been sampled, then the output S_{rec} can be expressed as a linear combination of the original k-space data S_0 and the k-space data predicted by the previous CNN. For the limit $\lambda \rightarrow \infty$, the output would be the same as the original k-space data. If $j \notin \Omega$, the output would be the same as the k-space data predicted by the previous CNN. [13]

1.3.2 Activation function

The purpose of applying activation function after each layer is to introduce non-linearity to the neural networks. While many activation functions have been proposed and put into practice, Rectified Linear Unit (ReLU) is the most popular and successful one. [14]

ReLU activation function can be expressed as

$$ReLU(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (3)$$

ReLU activation function turns all the negative input to zero while leave the positive input untouched. An easy and intuitive interpretation is that in image related task, the input would be the pixel value, which normally would not be negative. It can be observed that ReLU partially preserves the characteristic of linear models, which makes the models relatively easier to optimize. [14]

1.4 Deep learning on MR image reconstruction

Many efforts have been devoted to adopting deep learning methods in the medical image reconstruction. Alternating Direction Method of Multipliers (ADMM-net) is proposed to reconstruct MR images from CS-based sparsely sampled k-space data, based on the ADMM algorithm. [15] Deep Cascade of Convolutional Neural Network (Cascade-net) is proposed to reconstruct MR images from undersampled k-space data by introducing data sharing layer and data consistency layer into the conventional CNN. [13] KIKI-net is proposed to process data cross k-space and image space domain iteratively and serve to reconstruct MR images from undersampled k-space data. [16] Primal-Dual-net (PD-net) is originally proposed to reconstruct CT images from sparse raw data based on PD algorithm, [17] then applied to reconstruct MR images from undersampled k-space data. [18] U-net, named after its shape, is originally proposed for image segmentation task [19] and then applied to reconstructing MR images from image input reconstructed from undersampled k-space data. [20]

Despite many works have been done with MR reconstruction, most of them only focus on standard anatomical MR images, e.g., conventional T1 and T2 images. Many other types of MR-based images obtain different features, on which the feasibility of deep learning neural networks remains unevaluated. This project is particularly interested in two types of MR-based images, which are diffusion tensor imaging (DTI) and Gynecologic brachytherapy (GYN) MR images.

1.5 Diffusion Tensor Imaging (DTI)

Diffusion imaging, described by its name, embeds diffusion information in MR images, which is often altered by pathology especially in Brain site. Thus, diffusion imaging can greatly help with the diagnosis of certain diseases. In fact, it is reported to show great sensitivity to acute ischemic stroke and other cerebral and neurological disease. [21] DTI, as a category of diffusion imaging, reveals even more diffusion information than conventional diffusion weighted imaging (DWI) by introducing diffusion tensor and quantitative functional maps. DTI is reported to show great information about white matter connectivity and improve people's knowledge of neurological pathology. [21]

To understand DTI, the underlying physics mechanism of diffusion weighted imaging is introduced.

1.5.1 Diffusion weighted imaging (DWI)

Diffusion is defined as the random motion of molecular on a microscopic level. It is well known as 'Brownian motion', which is governed by heat. The diffusion weighted pulse sequence is originally proposed by Stejskal-Tanner [22] and illustrated as **Figure 5**. [23]

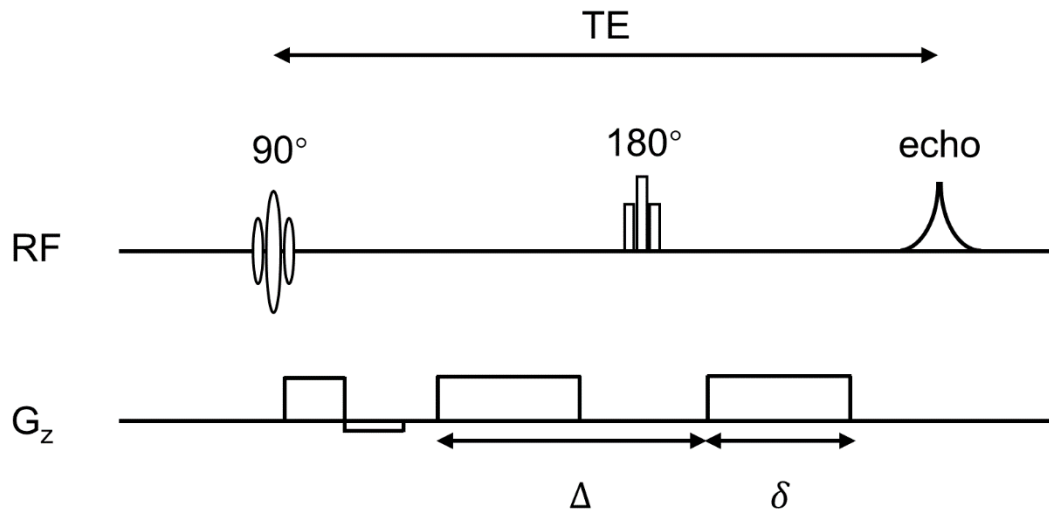


Figure 5: Sketch of Stejskal-Tanner diffusion weighted pulse sequence

This pulse sequence adds two diffusion gradients based on a T2-weighted spin echo pulse sequence. The two diffusion gradients are of the same amplitude, duration and direction. With the presence of the 180° flip RF pulse, the diffusion gradients actually work on the spins towards two opposite directions. Noting that in current implement, EPI is often used instead of a T2-weighted spin echo pulse sequence.

Accordingly, the acquisition of diffusion information is illustrated as follows. Assume within a volume, there is no net water movement, the volume of water would experience two diffusion gradients with the same amplitude and duration towards the opposite directions since the water volume stays at the same position, which would cancel out with each other. Thus, the water volume would not feedback any signal caused by the diffusion gradients, which means the system would not detect any diffusion information.

On the opposite, assume there is net water movement along the gradient direction, the water would experience the first diffusion gradient at one location and the second diffusion gradient at a different location along the gradient direction. This means the two gradients acting on the net water volume would have different amplitudes, thus cannot cancel out with each other. This means spins would not rephase completely, hence a signal characterizing the diffusion information would be recorded. Since the gradient is along one single direction, diffusion can be detected along only one direction at a time. An example of diffusion weighted images is shown in **Figure 6**.

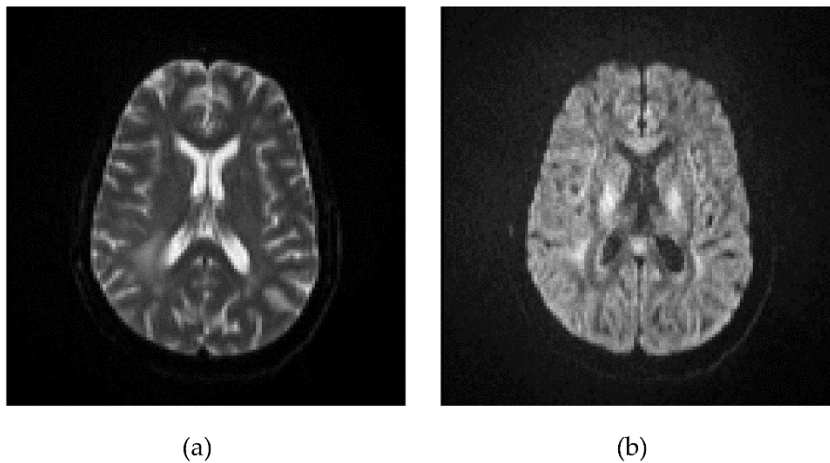


Figure 6: An example of diffusion weighted images. (a): Image without diffusion gradient. Noting that it is T2-weighted. (b): Image with diffusion gradient, b-value set at 1000 s/mm².

The diffusion weighted signal can be quantitatively expressed as [21]

$$S_i = S_0 \times e^{-b \times D} \quad (4)$$

Where S_i is the signal with the presence of diffusion gradients, S_0 is the signal without the presence of diffusion gradients, b is diffusion sensitivity factor, and D is the diffusion coefficient. b can be expressed as [21, 23]

$$b = \gamma^2 G^2 \delta^2 \left(\Delta - \frac{\delta}{3} \right) \quad (5)$$

Where γ is the gyromagnetic ratio, G is the amplitude of the diffusion gradient, δ is duration of the diffusion gradient, and Δ is the time interval between the starts of the two diffusion gradients.

In reality, the net water movement is not completely contributed by diffusion. Other factors like pressure gradients, thermal gradients or ionic interactions would also contribute to the net water movement, which cannot be differentiated from diffusion. Thus, the diffusion coefficient detected is called apparent diffusion coefficient (ADC), expressed as [21]

$$S_i = S_0 \times e^{-b \times ADC} \quad (6)$$

Traditionally b-value is set as 1000 s/mm², and typically it can be set within a range of 600 to 1500 s/mm² depending on interested pathologies. [23]

1.5.2 Diffusion tensor

As stated before, diffusion coefficient D can be measured along only one direction at a time. In fact, diffusion coefficient can be expressed as a 2-D tensor denoted as [23]

$$Diffusion\ Tensor = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{bmatrix} \quad (7)$$

The diagonal elements describe the diffusion along the three orthogonal directions, while the off-diagonal elements describe the correlation of diffusion between two directions. Diffusion tensor is a conjugate symmetric tensor, which means $D_{xy} = D_{yx}$, $D_{xz} = D_{zx}$, $D_{yz} = D_{zy}$. Thus, there are indeed only six independent elements in diffusion tensor. To acquire the diffusion tensor, at least six images with diffusion gradients along unparallel directions and one image without diffusion gradient are required.

A vivid way to show the physical meaning of diffusion tensor is to show the diffusion ellipsoid as **Figure 7**. [23] The length of the three axes of the diffusion ellipsoid are the eigenvalues of the diffusion tensor, while the directions of the three axes are corresponding eigenvector. For isotropic diffusion, diffusion along each direction is equivalent, thus the diffusion ellipsoid turns to a sphere, which shows no direction preference. For anisotropic diffusion, the diffusion ellipsoid shows a direction preference along the direction of eigenvector with the largest eigenvalue.

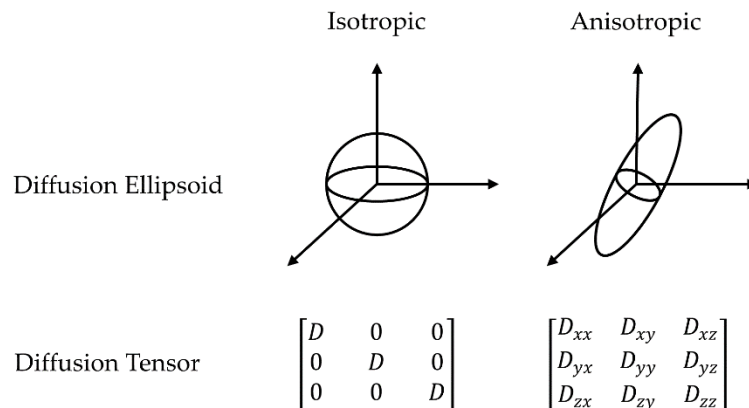


Figure 7: Illustration of diffusion ellipsoid

1.5.3 Quantitative functional map

Denoting the three eigenvalues of diffusion matrix as λ_1 , λ_2 and λ_3 , some quantitative functional maps can be calculated. Typically, ADC value can be calculated as

$$ADC = D_{av} = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3} \quad (8)$$

Another important diffusion tensor parameter, fractional anisotropy (FA), which reflects the degree of directionality of intravoxel diffusivity, can be calculated as [23]

$$FA = \frac{\sqrt{3} \sqrt{(\lambda_1 - D_{av})^2 + (\lambda_2 - D_{av})^2 + (\lambda_3 - D_{av})^2}}{\sqrt{2} \sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}} \quad (9)$$

Based on the FA map, a colored FA map can be constructed. Three colors are assigned to three orthogonal directions, typically green represents the anterior-posterior direction, red represents the left-right direction and blue represents the superior-inferior direction. An example of above quantitative functional maps is shown in **Figure 8**.

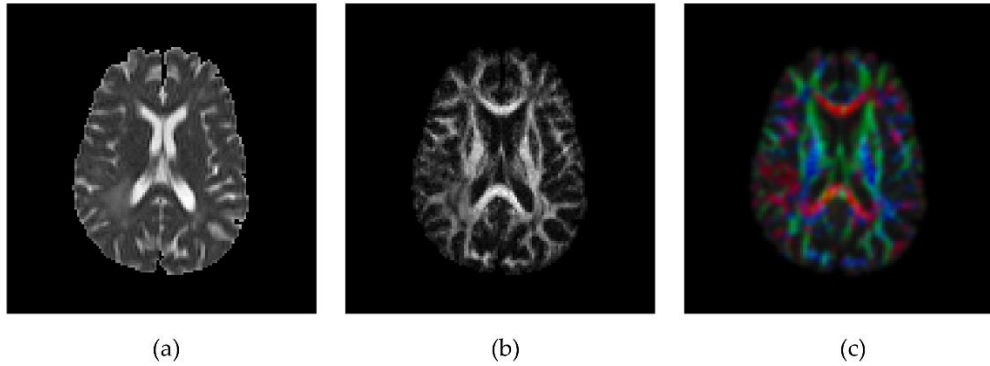


Figure 8: Examples of quantitative functional maps. (a): ADC map. (b): FA map. (c): Colored FA maps.

1.5.4 Motivation to accelerate DTI

DTI has been widely applied in clinic. As stated before, to solve the diffusion tensor, at least 7 sets of images should be acquired. Despite it would cost around 5-10 s to acquire one set of DTI images, depending on the data matrix size, number of slices and whether or not the gating technique is adopted, it is reported that at least 30 sets of DTI images along uniformly distributed directions should be acquired to accurately estimate DTI metrics, which leads to a long acquisition time. [24-25] The long scanning time would lead to a series of problems stated before, e.g., motion artifact, patient discomfort and low efficiency. There exists a demand to accelerate DTI scanning process, which would fix the previously stated problems and make room for improving intrinsic resolution and SNR.

1.6 Gynecologic brachytherapy MRI study (GYN)

Gynecologic Brachytherapy is used to treat cancer at cervix, uterine body and vagina. Brachytherapy applicators such as tandems and ovoids are inserted through vagina and placed near the tumor. These applicators can change the shape and structure of human tissue to some extent and introduce new features into medical images.

There is a continuous effort to adopt MRI to GYN treatment planning. [26] Since new features are introduced, it would be good to re-evaluate the neural network accelerating methodology on GYN MR images.

1.7 Project objective

This work aims to evaluate the feasibility of neural networks on accelerating DTI and Gynecologic MRI study (shortened as GYN). Recently, Ramzi et al. benchmarked four neural networks including KIKI-net, U-net, Cascade-net and PD-net on NYU fast MRI database. [27] The methodology in this work was inspired by Ramzi's. Three neural networks including U-net, Cascade-net and PD-net were evaluated. A Cartesian undersampling strategy was applied to all the training and test data. In DTI, individual diffusion weighted images and quantitative functional maps including ADC maps, FA maps and colored FA maps were reconstructed. In GYN, axial T1 and T2 MR images with applicator were reconstructed. The performance of each neural network was evaluated by qualitatively and quantitatively evaluating the similarity between the reference images and reconstructed images using metric called total relative error (TRE).

2. Method

The general workflow of this work is illustrated as **Figure 9**. The raw data was processed to generate training data, validation data, test data and reference data. Training data and validation data was then fed into the neural networks to train the models. After the models were trained, the test data was then fed into the trained models to reconstruct images. Finally, the reconstruct data was evaluated compared with the reference to assess the performance of corresponding neural networks. Details are explained in the following sections.

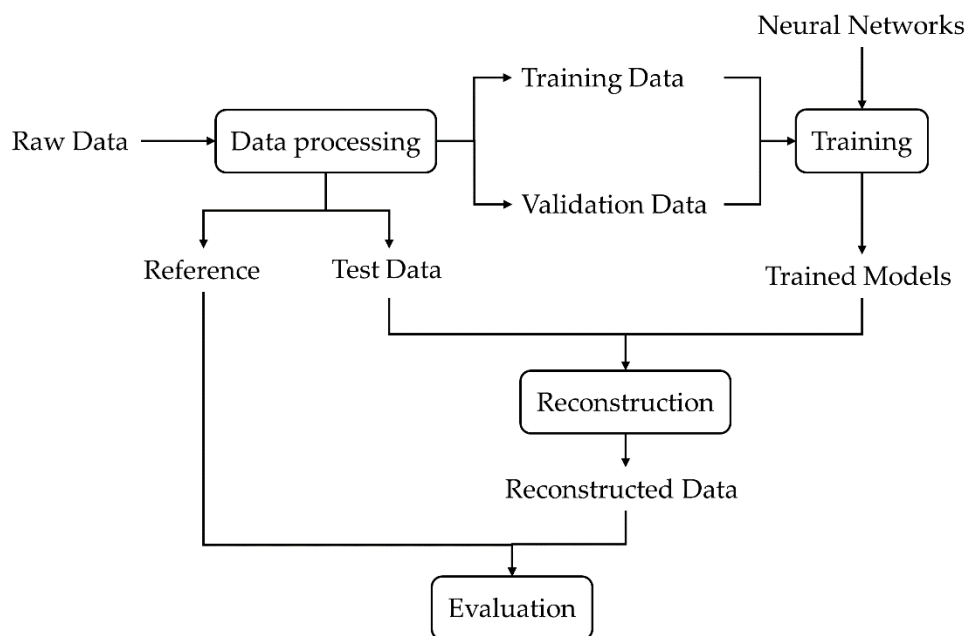


Figure 9: General workflow

2.1 Raw data

2.1.1 DTI dataset: RIDER NEURO MRI

DTI data was acquired from a public database RIDER NEURO MRI. It contains image data of 19 patients with recurrent glioblastoma who were acquired multiple imaging sets. [28] 16 of them were each acquired of Brain DTI scans twice. Each DTI set obtains 36 axial slices, except that in one exam only 32 axial slices were acquired in each DTI set. Each set was acquired with an image matrix size of 128×128 , one zero gradient, and 12 diffusion gradients with b-value of 1000 s/mm^2 . Thus, there are a total of 416 sets, 14924 slices. Gradient directions were not published, hence directly acquired from the dataset author and listed in appendix. Only image space data is available, no k-space data available.

2.1.2 GYN dataset: Duke GYN MRI study

GYN data was acquired from Duke Hospital clinic data under IRB protocol #Pro00101712. A select of 32 patients were each acquired of cervical T1-spgr and T2 images for several exams. There are a total of 148 image data sets for T1-spgr and T2 each. Each T1-spgr set obtains 192 axial slices with an image matrix size of 320×320 , and each T2 set obtains 96 axial slices with an image matrix size of 320×320 . Only image space data is available.

2.2 Data processing

2.2.1 k-space data acquisition

k-space data is the underlying essential information of an MR image. During MRI acquisition, k-space data is initially acquired and then inverse Fourier transformed to image space data with orthonormalization. Since no k-space data is available originally in our database, which is a common situation with MRI data, we need to transform the image space data back to the k-space. To achieve this, a 2-D Fourier transform with orthonormalization was applied to the original image space data.

2.2.2 Undersampled Image

Cartesian undersampling strategy was applied to all the training and test data in this work. Theoretically, k-space is fully sampled along frequency encoding direction while sparsely sampled along phase encoding direction. Since the images involved in this work were all square, phase encoding direction was arbitrarily designated.

Mathematically, Cartesian undersampling was realized by element-wisely multiplying a 0-1 binary mask with the original k-space data. The undersampled k-space data was then inverse Fourier transformed back to image space data, which is called zero-filled reconstruction, to serve as a baseline of image quality.

The workflow of undersampling is illustrated in **Figure 10**. 25% Cartesian undersampling strategy (acceleration factor = 4) was applied to all the data. For RIDER data, the central 8% of the k-space was sampled, and the rest 17% was sampled in the

periphery of k-space following uniform distribution. For GYN data, the central 12.5% of the k-space was sampled, and the rest 12.5% was sampled in the periphery of k-space following uniform distribution. The masks were fixed for RIDER data and GYN data respectively.

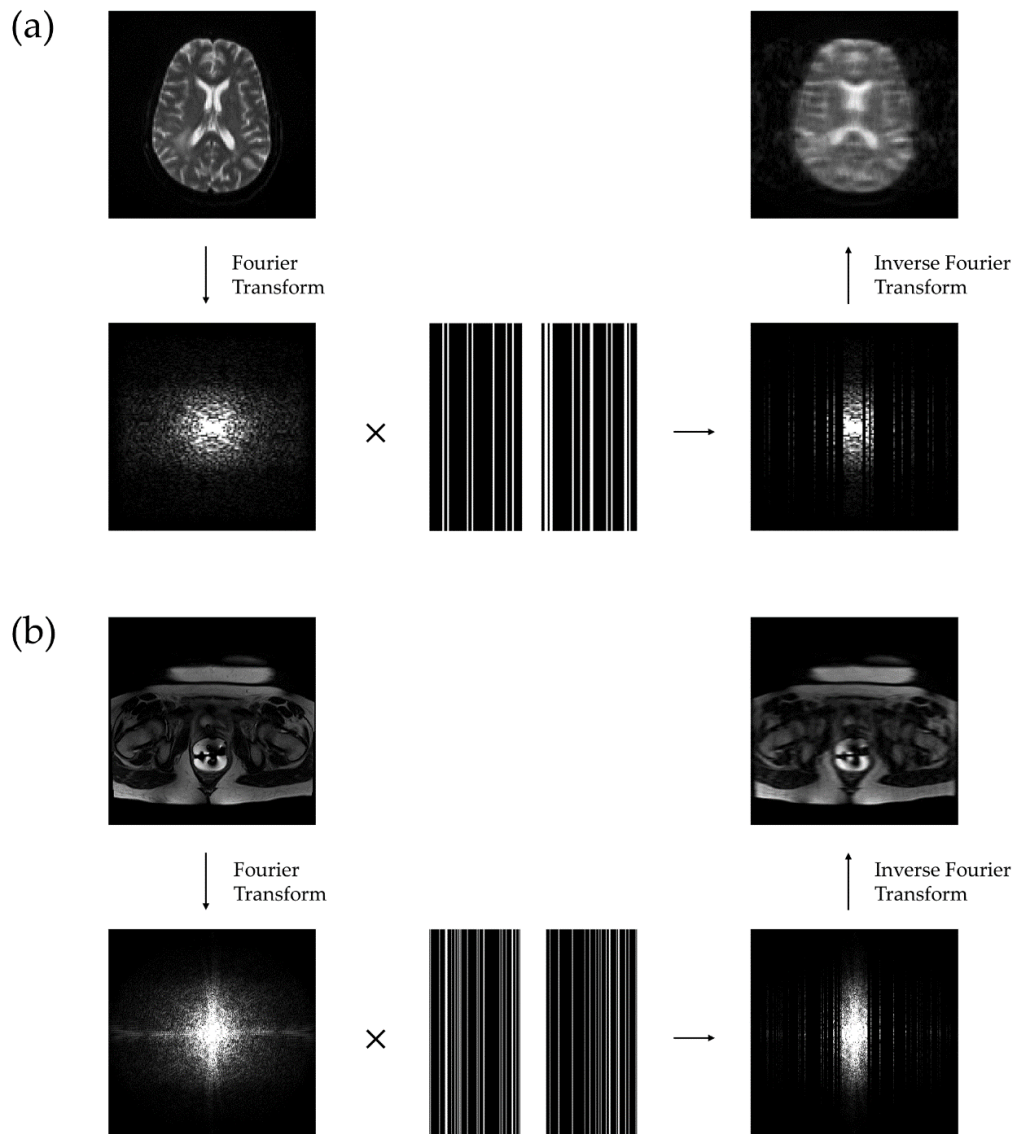


Figure 10: Undersampling workflow with 25% Cartesian undersampling strategy. (a): An example of RIDER data. (b): An example of GYN data.

2.2.3 File format conversion and partitioning

The data was all originally in Dicom file format. It was transformed to h5 file format. Each h5 file contained several axial images in superior-inferior order. For RIDER data, each h5 file contained a complete DTI set consisted of 32 or 36 slices of 128×128 diffusion weighted images. For GYN data, it was ideal to compact all images from one exam into one h5 file, e.g., compacting 192 320×320 images into one h5 file. However, in actual implement, each h5 file would be in a size of 225 Mb, which slowed the training process significantly. Thus, for GYN data, each file was reduced to containing only 32 slices. For T1-spgr exams, each exam was divided into 6 h5 files, each containing 32 320×320 images. For T2-spgr exams, each exam was divided into 3 h5 files, each containing 32 320×320 images. Thus, there are a total of 416 h5 files for RIDER, 888 h5 files for GYN T1-spgr and 444 h5 files for GYN T2.

2.2.4 Dataset partitioning

For RIDER data, each h5 file contained a complete DTI set consisted of 32 or 36 slices of 128×128 diffusion weighted images. Two exams from one random patient, id #1863082306, were excluded from the training process, and one of them was designated as test data, which would be 13 files in total. For the rest of the data, 327 random files were selected as training dataset, while the other 63 files were selected as validation dataset.

For GYN T1-spgr data, each file contained only 32 slices. As a result, each exam was divided into 6 h5 files, each containing 32 slices. 3 exams from one random patient,

id enclosed, were designated as test data and excluded from the training process, which are a total of 18 files. For the rest of the data, 725 random h5 files were selected as training dataset, while 145 random h5 files were selected as validation dataset.

For GYN T2 data, each file contained only 32 slices. As a result, each exam was divided into 3 h5 files, each containing 32 slices. 3 exams from one random patient, id enclosed, were designated as test data and excluded from the training process, which would be 13 files in total. For the rest of the data, 363 random h5 files were selected as training dataset, while 72 random h5 files were selected as validation dataset.

Number of files in each dataset and number of slices in each file are summarized in **Table 1**.

Table 1: Dataset partitioning

Dataset	MRI sequence	Image size	Slices per file	# Files in training dataset	# Files in validation dataset	# Files in test dataset
RIDER Brain	DTI	128×128	36 or 32	327	63	13
GYN Cervix	T1-spgr	320×320	32	725	145	18
	T2	320×320	32	363	72	9

2.3 Neural network architecture

The architecture of the neural networks was inspired by Ramzi et al. [27] Parameters of each network were selected the same for all the training.

2.3.1 U-net

U-net is named after its shape. It was firstly proposed for image segmentation task [19] and then applied to image reconstruction task [20].

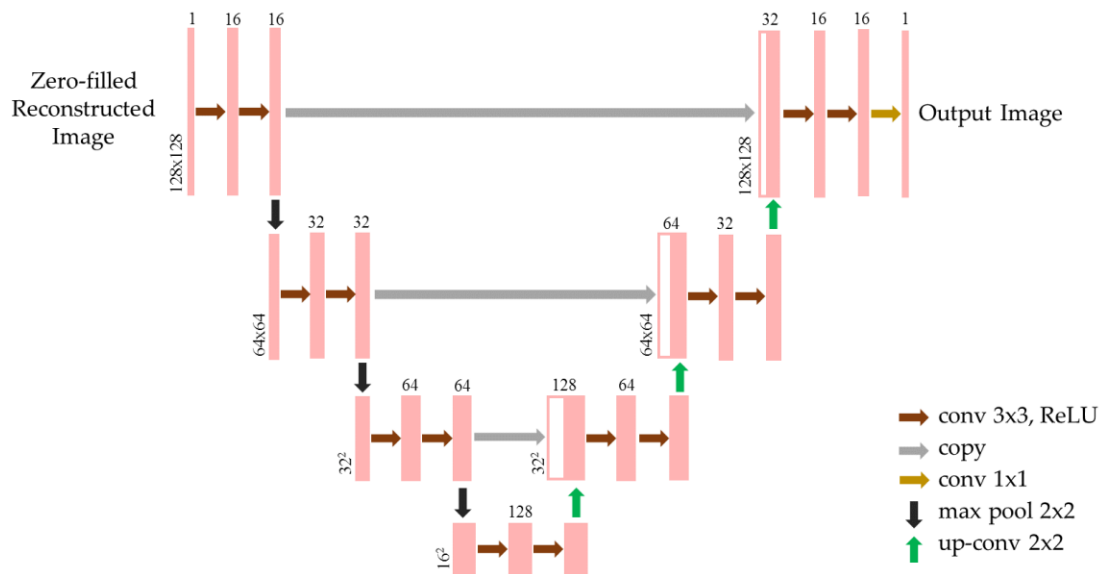


Figure 11: Architecture of U-net

The architecture of U-net is shown in **Figure 11**. [19-20, 27] The zero-filled reconstructed image is input into the neural network. U-net is consisted of several levels. Each level is consisted of several convolutional layers. Zero padding is applied so that the image size remains the same after convolution. On the left side, the level goes down by applying a max pooling layer with a size of 2-by-2 and a stride of 2, which reduces the image size into half. This is to concentrate local features. The number of filters doubles as level goes down. On the right side, the level goes up by applying a transpose convolution layer (also called up-convolution layer) with a size of 2-by-2 and a stride of 2, which would

expand the image size by twice. This is to recover the image information from concentrated bottom levels. In each level, the filters from the left side are copied to the right side to introduce more original information for the purpose of improving accuracy. Finally, a reconstructed image is generated by applying a merge layer. The number of filters in each level is [16, 32, 64, 128] from top to bottom.

2.3.2 Cascade-net

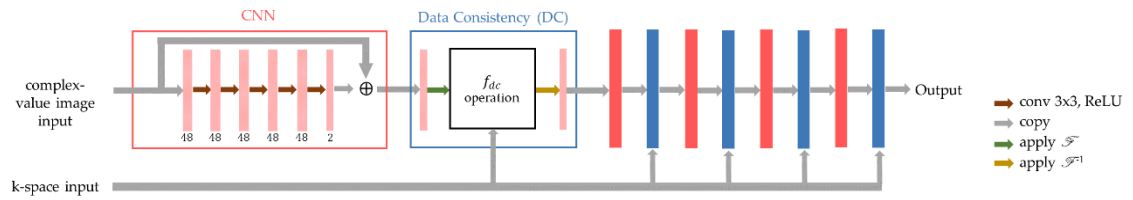


Figure 12: Architecture of Cascade-net

The Deep Cascade Convolutional Neural Network, shortened as Cascade-net, is initially proposed by [13]. The architecture is illustrated in **Figure 12**. [13, 27] Cascade-net deals with both image space data and k-space data. Complex image data and complex k-space data are both input into the network. Since the input would be complex number while Tensorflow does not accept complex number, the complex input is divided into two channels. Cascade-net consists of 5 iterations of 2 blocks, which are CNN block, denoted in red and representing processing in image space, and Data Consistency (DC) block, denoted in blue and representing processing in k-space. The CNN block is consisted of 5 convolutional layers with 48 filters, 1 convolutional layer with 2 filters, and 1 residual

layer. The DC is described in section 1.3.1.5 to enforce the k-space data. The output is the magnitude of the complex reconstructed image.

2.3.3 PD-net

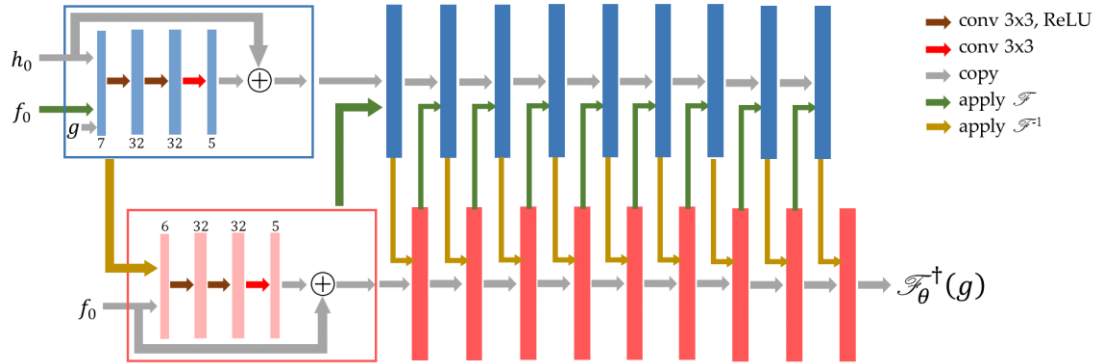


Figure 13: Architecture of PD-net. f_0 is initial guess of image space data, h_0 is initial guess of k-space data, and g is the true k-space data.

Primal-Dual net (PD-net) is firstly proposed to reconstruct CT image [17] and then applied to MR image reconstruction [18]. The architecture is illustrated in **Figure 13**. [17-18, 27] PD-net deals with both image space data and k-space data. Real image data and complex k-space data are both input into the network. In the diagram, f_0 represents initial guess of image space data, which is zero tensor, h_0 represents initial guess of k-space data, which is also zero tensor, and g represents the true k-space data. PD-net consists of 10 iterations of 2 blocks, which are primal block, denoted in blue and representing processing in k-space, and dual block, denoted in red and representing processing in image space. The primal block is consisted of 4 convolutional layers with filters number 7→32→32→5 and 1 residual layer. The reason it has 7 filters in the first convolutional layer is that the

five of the filters are copies of h , which allows redundancy, one is the Fourier transform of f , and one is the true k-space data g . The dual block has the same architecture as the primal block, except that the number of the filters in the first convolutional layer is 6 instead of 7 due to lack of true data input. The output is the magnitude of the inverse Fourier transform of reconstructed k-space data.

2.4 Training

For all the training, the parameters for each neural network were selected the same. For all the neural networks, loss function was defined as mean absolute error and can be expressed as

$$loss = \frac{\sum_m^M \sum_n^N |x_{ref}(m, n) - x_{rec}(m, n)|}{M \times N} \quad (10)$$

where x_{ref} is the reference image, x_{rec} is the reconstructed image, and M and N are two dimensions of image matrix respectively. The models were then optimized using Rectified Adam method. [29]

All the models were trained using python programming, Tensorflow framework and an Intel Xeon W-3223 CPU card. For Cascade-net and PD-net, only random one of the eight innermost slices of each file was used for training in each epoch out of efficiency consideration. For all the trainings, the training loss approximately stopped dropping after 30 epochs. Hereby we chose 50 training epochs for all of the trainings in all of the tasks for stability and consistency consideration.

2.5 Reconstruction

After training, the test data was fed into the trained neural networks, and corresponding images were reconstructed. The reconstruction by neural networks was realized using python programming, Tensorflow framework and an Intel Xeon W-3223 CPU card.

For DTI evaluation, quantitative functional maps such as ADC maps, FA maps and colored FA maps were reconstructed and evaluated. This reconstruction is realized using MATLAB R2019a programming and an Intel Xeon W-3223 CPU card. The method to reconstruct these maps from a combination of diffusion weighted images is illustrated below.

To reconstruct the quantitative functional maps, the diffusion tensor should be figured out. Based on the equation in section 1.5.1

$$S_i = S_0 \times e^{-b \times ADC} \quad (6)$$

It can be furtherly expressed as

$$\ln\left(\frac{S_0}{S_i}\right) = b \begin{bmatrix} g_{xi} & g_{yi} & g_{zi} \end{bmatrix} \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{bmatrix} \begin{bmatrix} g_{xi} \\ g_{yi} \\ g_{zi} \end{bmatrix} \quad (11)$$

where $[g_{xi} \ g_{yi} \ g_{zi}]$ is the normalized diffusion gradient direction vectors. Furtherly,

$$\frac{\ln(S_0/S_i)}{b} = g_{xi}^2 D_{xx} + g_{yi}^2 D_{yy} + g_{zi}^2 D_{zz} + 2g_{xi}g_{yi}D_{xy} + 2g_{xi}g_{zi}D_{xz} + 2g_{yi}g_{zi}D_{yz} \quad (12)$$

Combining all the diffusion gradients together, we have [30]

$$\begin{bmatrix} \frac{\ln(S_0/S_1)}{b} \\ \frac{\ln(S_0/S_2)}{b} \\ \vdots \\ \frac{\ln(S_0/S_M)}{b} \end{bmatrix} = \begin{bmatrix} g_{x1}^2 & g_{y1}^2 & g_{z1}^2 & 2g_{x1}g_{y1} & 2g_{x1}g_{z1} & 2g_{y1}g_{z1} \\ g_{x2}^2 & g_{y2}^2 & g_{z2}^2 & 2g_{x2}g_{y2} & 2g_{x2}g_{z2} & 2g_{y2}g_{z2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ g_{xM}^2 & g_{yM}^2 & g_{zM}^2 & 2g_{xM}g_{yM} & 2g_{xM}g_{zM} & 2g_{yM}g_{zM} \end{bmatrix} \begin{bmatrix} D_{xx} \\ D_{yy} \\ D_{zz} \\ D_{xy} \\ D_{xz} \\ D_{yz} \end{bmatrix} \quad (13)$$

This can be solved by applying a pseudo inverse of the M-by-6 gradient matrix on the right to left-multiply the M-by-1 matrix on the left. After solving for the diffusion tensor, the eigenvalues and eigenvectors can be calculated. These two steps can be easily done by MATLAB. Then, ADC maps and FA maps can be calculated using the equations in section 1.5.3

$$ADC = D_{av} = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3} \quad (8)$$

$$FA = \frac{\sqrt{\frac{3}{2} \sqrt{(\lambda_1 - D_{av})^2 + (\lambda_2 - D_{av})^2 + (\lambda_3 - D_{av})^2}}}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}} \quad (9)$$

Based on the FA map, a colored FA map can be constructed. Three colors are assigned to three orthogonal directions, typically green represents the anterior-posterior direction, red represents the left-right direction and blue represents the superior-inferior direction. The color pixel values are designated as

$$\begin{cases} red = FA \times \cos(\alpha) \\ green = FA \times \cos(\beta) \\ blue = FA \times \cos(\gamma) \end{cases} \quad (14)$$

where α, β, γ are the angles between the principle eigenvector with the biggest eigen value and the three orthogonal Cartesian axes. Instead of combing the three colors together, in this work a single color was designated to each pixel, which was the color that obtains the

largest value among the three colors. For instance, for a pixel with value $red > green > blue$, it is designated as *red* only.

2.6 Evaluation

The performance of neural networks was evaluated by quantitatively calculating the similarity between the reconstructed images and corresponding reference images. A metric called total relative error (TRE) was used. TRE can be expressed as

$$TRE = \frac{\sqrt{\sum_{x,y} [M_0(x,y) - M_g(x,y)]^2}}{\sum_{x,y} M_g(x,y)} \quad (3)$$

where $M_0(x,y)$ is the reconstructed image and $M_g(x,y)$ is the reference image reconstructed from the fully sampled k-space data, both in magnitude.

Qualitative evaluation was also adopted for reference.

3. Results

3.1 DTI reconstruction

3.1.1 Qualitative evaluation

A designated slice, slice #21 of patient id #1863082306, exam ax tensor 53188, was reconstructed for qualitative evaluation. The diffusion weighted images, ADC maps, FA maps and colored FA maps reconstructed using the three neural networks are shown in **Figure 14-18**, compared to the reference images and zero-filled reconstructed images. The corresponding differences between the reconstructed images and the reference images are also shown.

For diffusion weighted images where $b=0$ s/mm², it can be told that the difference of zero-filled reconstructed image is very significant, while the difference of U-net reconstructed image decreases greatly. The differences of Cascade-net and PD-net furtherly decrease and cannot be told which is better. By visual check, the quality of images reconstructed from three neural networks is comparable to the reference and in an order of [*Cascade – net, PD – net* > *U – net*].

For diffusion weighted images where $b=1000$ s/mm², it can be told that the difference of zero-filled reconstructed image is very significant, while the differences of images reconstructed from 3 neural networks decrease greatly and cannot be told which is the best. By visual check, the quality of images reconstructed from three neural networks is comparable to the reference.

For ADC maps, it can be told that the difference of zero-filled reconstructed image is very significant, while the difference of U-net reconstructed image decreases greatly. The differences of Cascade-net and PD-net furtherly decrease and cannot be told which is better. By visual check, the quality of images reconstructed from three neural networks is comparable to the reference and in an order of [*Cascade – net, PD – net > U – net*]. Noting that there are some white dots in the periphery of the brain in difference images. By taking a close look, it can be found that the regions of brain are different between the reference image and reconstructed images, and the white points are generated because there is signal on one image while no signal on the other. This is due to the different performance of auto-segmentation algorithm on images. It automatically defines region of interest (ROI), which is brain tissue, based on the diffusion weighted images with $b=0$ s/mm² and calculation will only be done in ROI. Since the diffusion weighted images are different from the reference image to the reconstructed images, the ROIs can be different.

For FA maps, it can be told that the difference of zero-filled reconstructed image is very significant, while the differences of images reconstructed from 3 neural networks decreases. By visual check, the quality of images reconstructed from 3 neural networks is comparable to the reference, and PD-net slightly outperformed the other two networks.

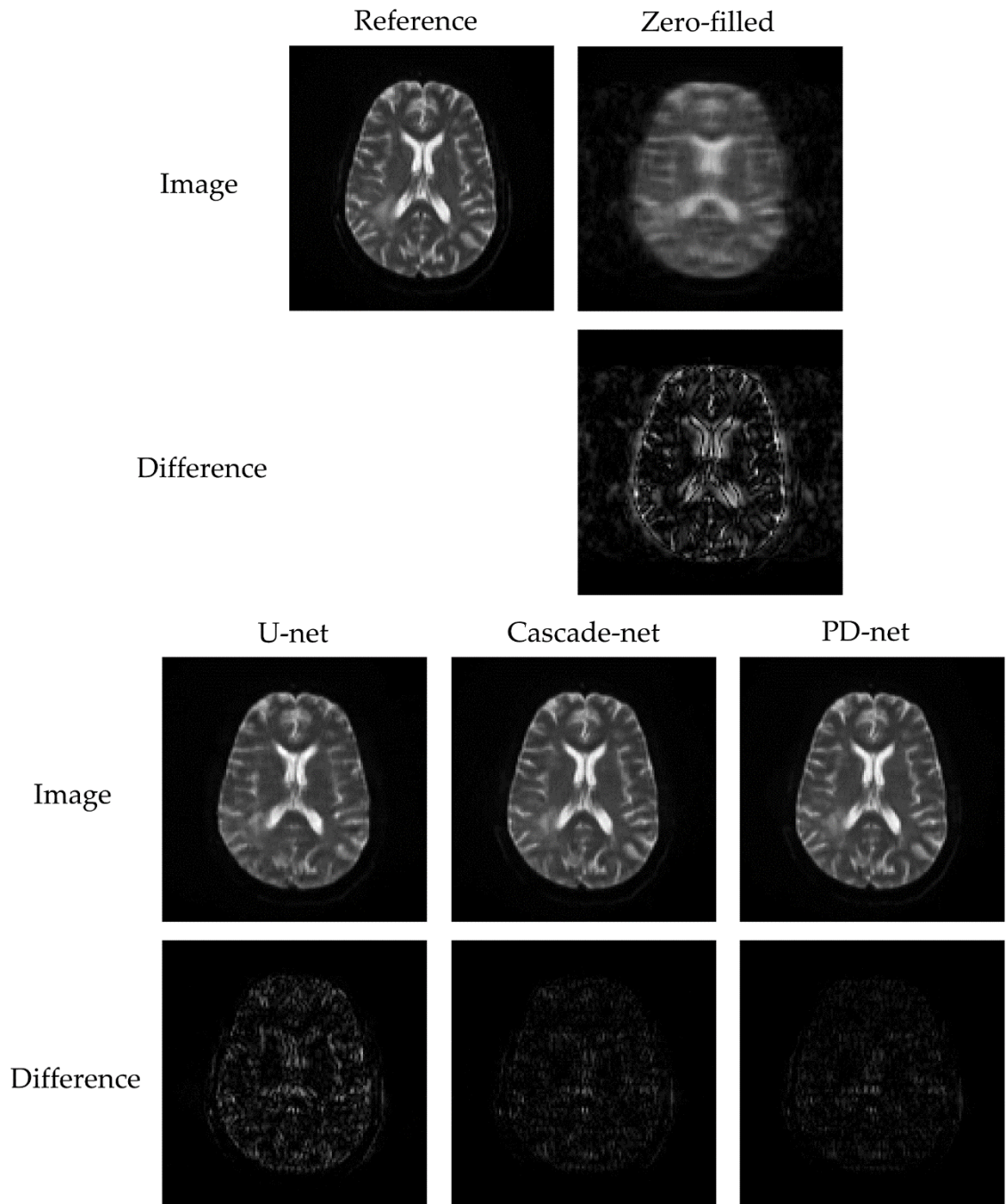


Figure 14: Qualitative evaluation for diffusion weighted images, $b=0 \text{ s/mm}^2$

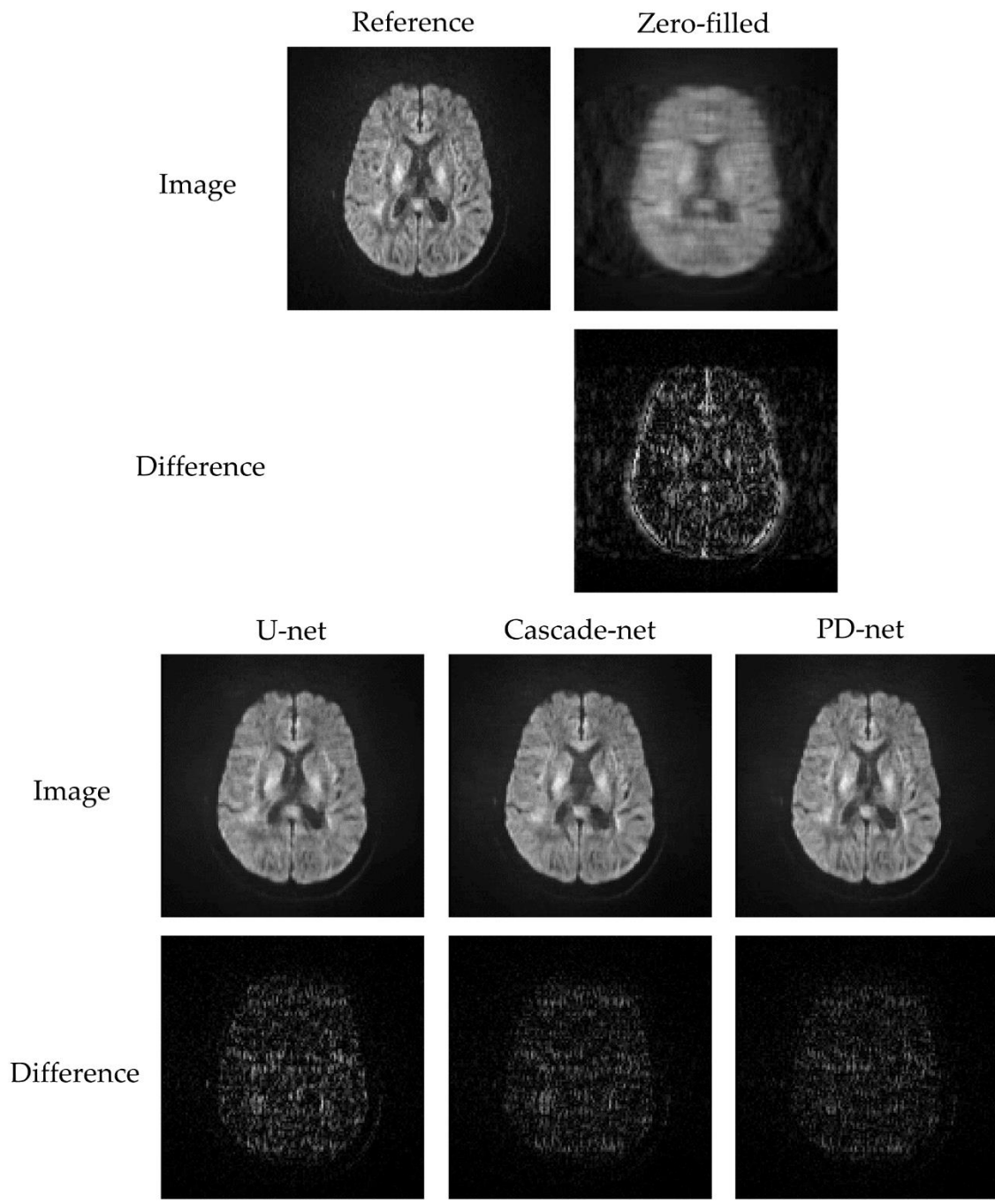


Figure 15: Qualitative evaluation for diffusion weighted images, $b=1000 \text{ s/mm}^2$, gradient #5

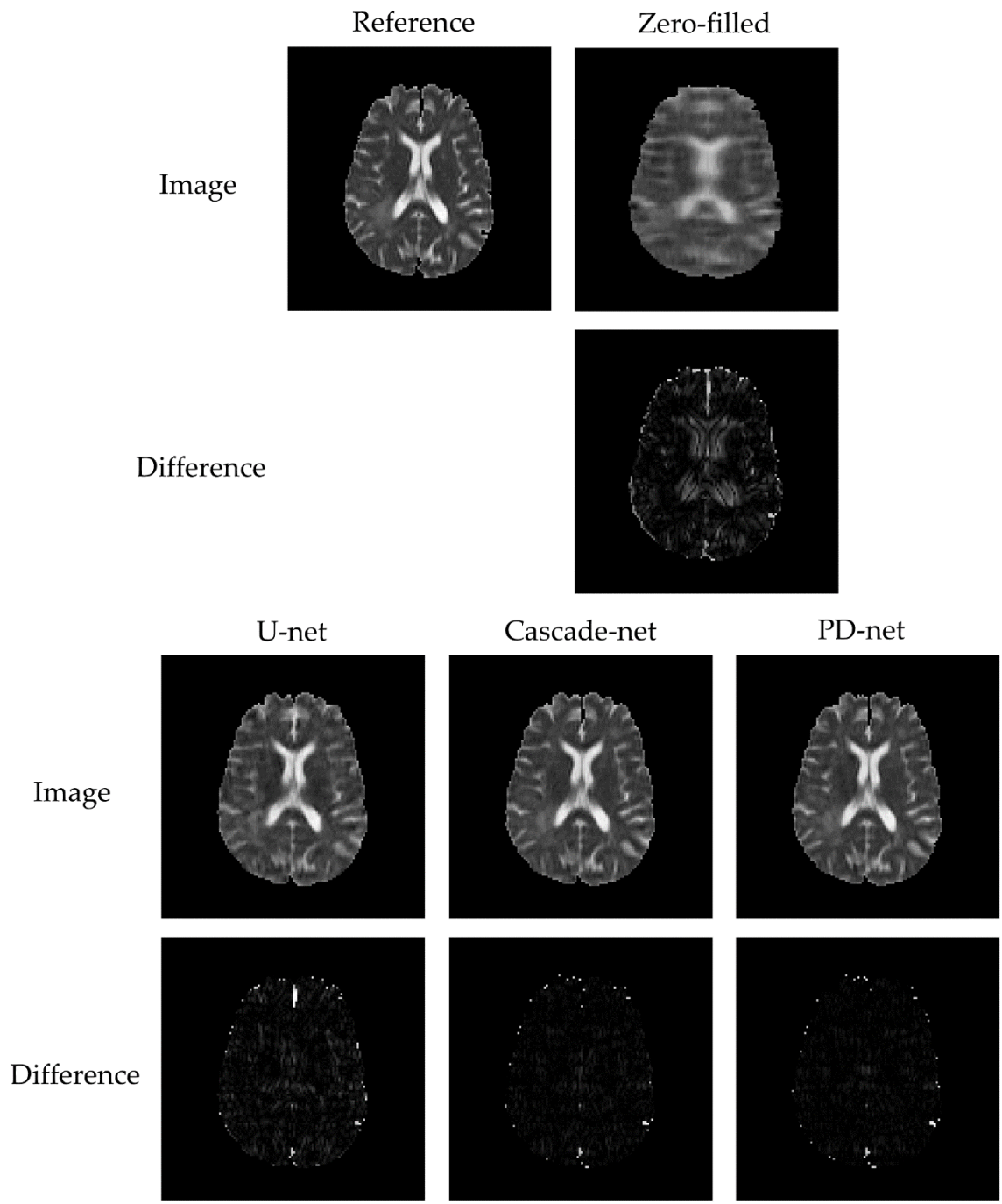


Figure 16: Qualitative evaluation for ADC maps

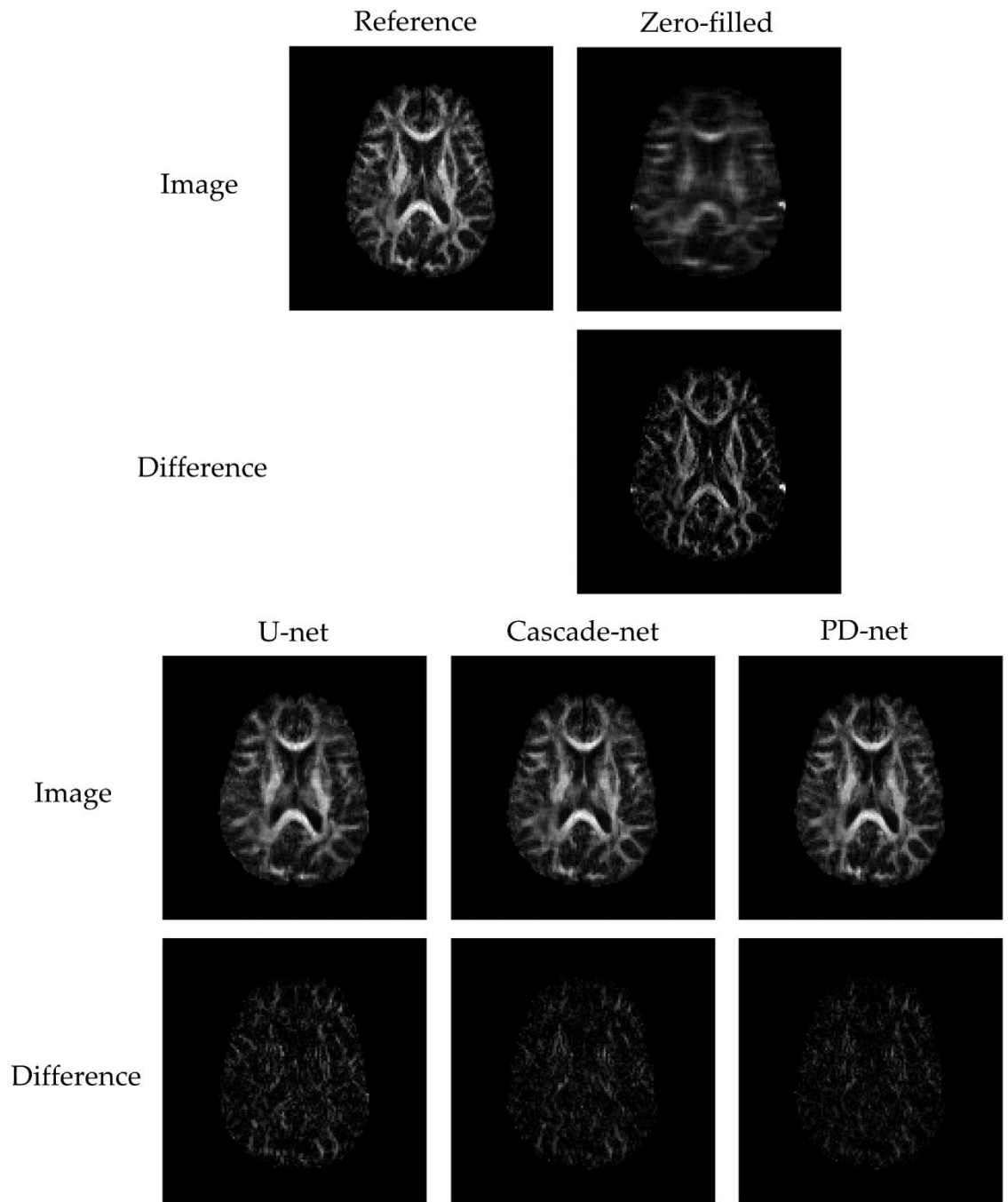


Figure 17: Qualitative evaluation for FA maps

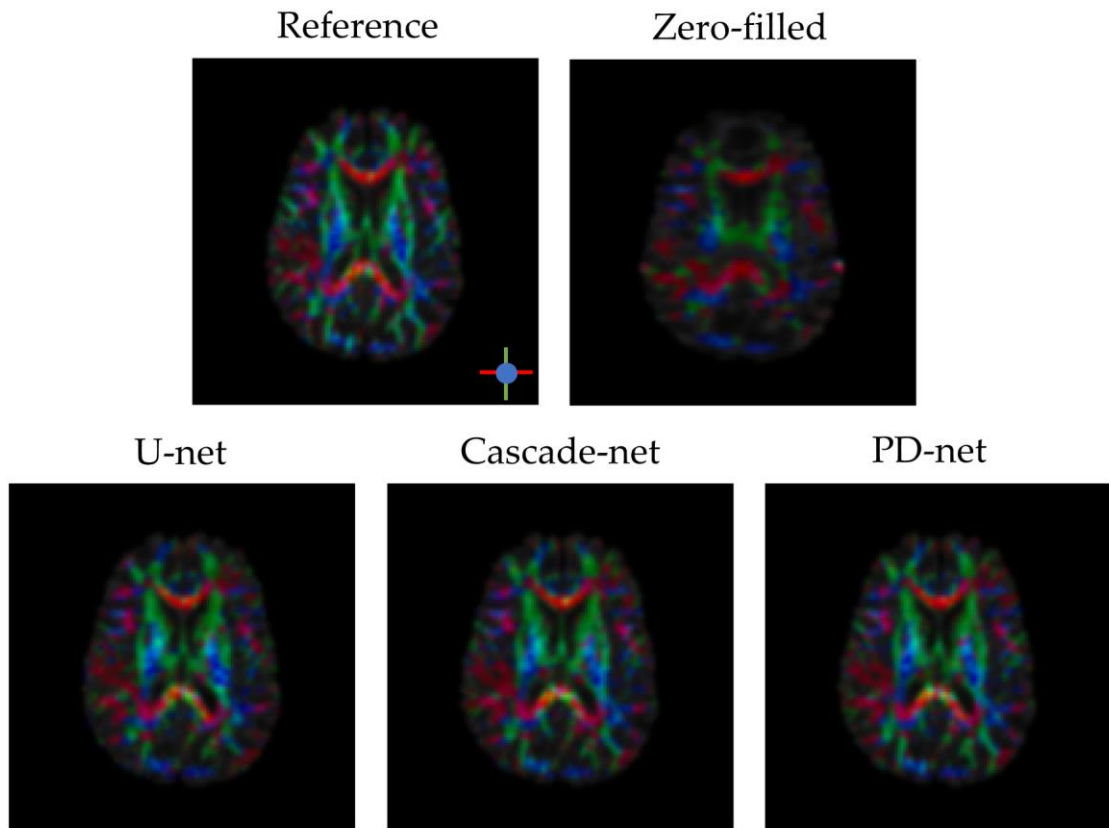


Figure 18: Qualitative evaluation for colored FA maps

For colored FA maps, it can be told that the zero-filled reconstructed image is of bad quality, while the quality of images reconstructed from 3 neural networks improves a lot, but there are still some obvious differences from the reference image. Since we did not come up with a reliable method to evaluate the similarity of colored images, colored FA maps are only generated for reference purpose and are not quantitatively evaluated.

Another common concern in clinic regarding DTI is the image distortion. There are mainly two sources of image distortions in DTI. One is the Eddy-current induced distortion, which is caused by the strong, fast-switching diffusion gradients and would

only affect images to which the diffusion gradients are applied. The other one, which is more of a concern, is EPI distortion. Due to the lack of refocusing RF pulses, phase error is easier to accumulate in EPI, thus EPI often requires a highly homogeneous magnetic field. However, the inhomogeneities of local magnetic field are often inevitable, especially at the magnetic interfaces inside human heads, e.g., between soft tissues and air. Since all DTI images are acquired using EPI sequence, EPI distortions would affect both images with and without diffusion gradients. Both types of the above distortions are more or less correctable. [6, 31] In our RIDER dataset, the distortions do exist but do not compromise the image quality significantly. This is partially because the image matrix is coarse (only 128×128). The report of the dataset does not record any distortion correction processing. Since the distortions do not severely damage image quality and are not the focus of our project, these would not be furtherly discussed.

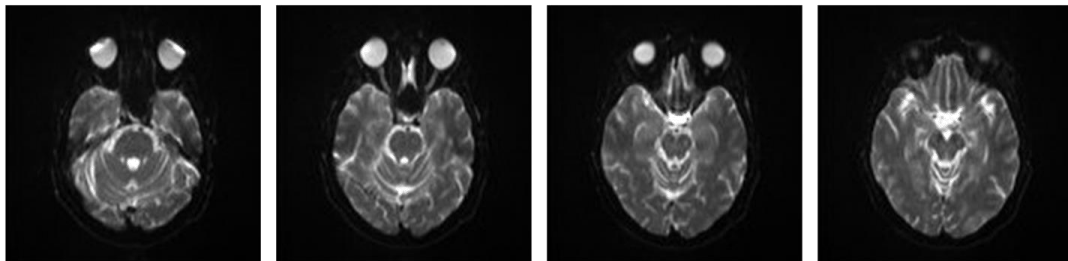


Figure 19: Example DTI slices around eye level, from left to right: slices # 13-16, inferior to superior. Moderate distortions are visible.

3.1.2 Quantitative evaluation

The TRE of diffusion weighted images, ADC maps and FA maps are tabulated in **Table 3** in a format of (mean \pm standard deviation). Noting that only part of the slices, #7-

31, were included in TRE analysis. This is because in the very beginning and ending slices, there is not much tissue, hence the ADC maps and FA maps would not give much information, thus excluded from the evaluation.

Table 2: TRE evaluation of DTI reconstruction

Evaluation Metric Reconstruction Method	TRE (% , mean+std.)			
	Zero-filled	U-net	Cascade-net	PD-net
Diffusion weighted images	0.187±0.031	0.106±0.014	0.084±0.008	0.081±0.009
ADC maps	0.448±0.123	0.305±0.070	0.254±0.072	0.238±0.070
FA maps	0.511±0.046	0.295±0.038	0.239±0.030	0.231±0.038

Comparing across reconstruction methods, PD-net showed a slightly better performance than Cascade-net, and they both showed better performance than U-net. Comparing across image types, it can be found that the reconstruction was decent for diffusion weighted images but not ideal for quantitative functional maps. This is because the quantitative functional maps were reconstructed from the diffusion weighted images. During this process, the error in diffusion weighted images would naturally add up. This suggests that to get decent quantitative functional maps, it would require diffusion weighted images to be of extremely high quality.

3.2 GYN reconstruction

3.2.1 Qualitative evaluation

3.2.1.1 T1-spgr reconstruction

A designated slice, slice #137 of a random patient (id and exam enclosed), was reconstructed for qualitative evaluation. The T1 images reconstructed using neural networks are shown in **Figure 20**, compared to the reference images and zero-filled reconstructed images. The corresponding differences between the reconstructed images and the reference images are also shown.

It can be told that compared to the zero-filled reconstructed images, the differences of the three images reconstructed using neural networks do decrease but are still outstanding. In fact, if taken a close look at, it can be found that even for Cascade-net and PD-net reconstructed images, which show a better quality than U-net and zero-filled reconstructed images, many structures were not reconstructed, and brachytherapy digitizers for image registration which are particularly concerned in T1 images are very vague, which is not desired. Multiple slices of Cascade-net and PD-net reconstructed images compared with reference are shown in **Figure 21-22** to demonstrate this finding. This suggests that the three neural networks with current parameter settings might not be a good tool for GYN T1-spgr acceleration. Similarity between reconstructed images and reference were furtherly quantitatively evaluated using TRE, which could not reflect the similarity of local features like applicators.

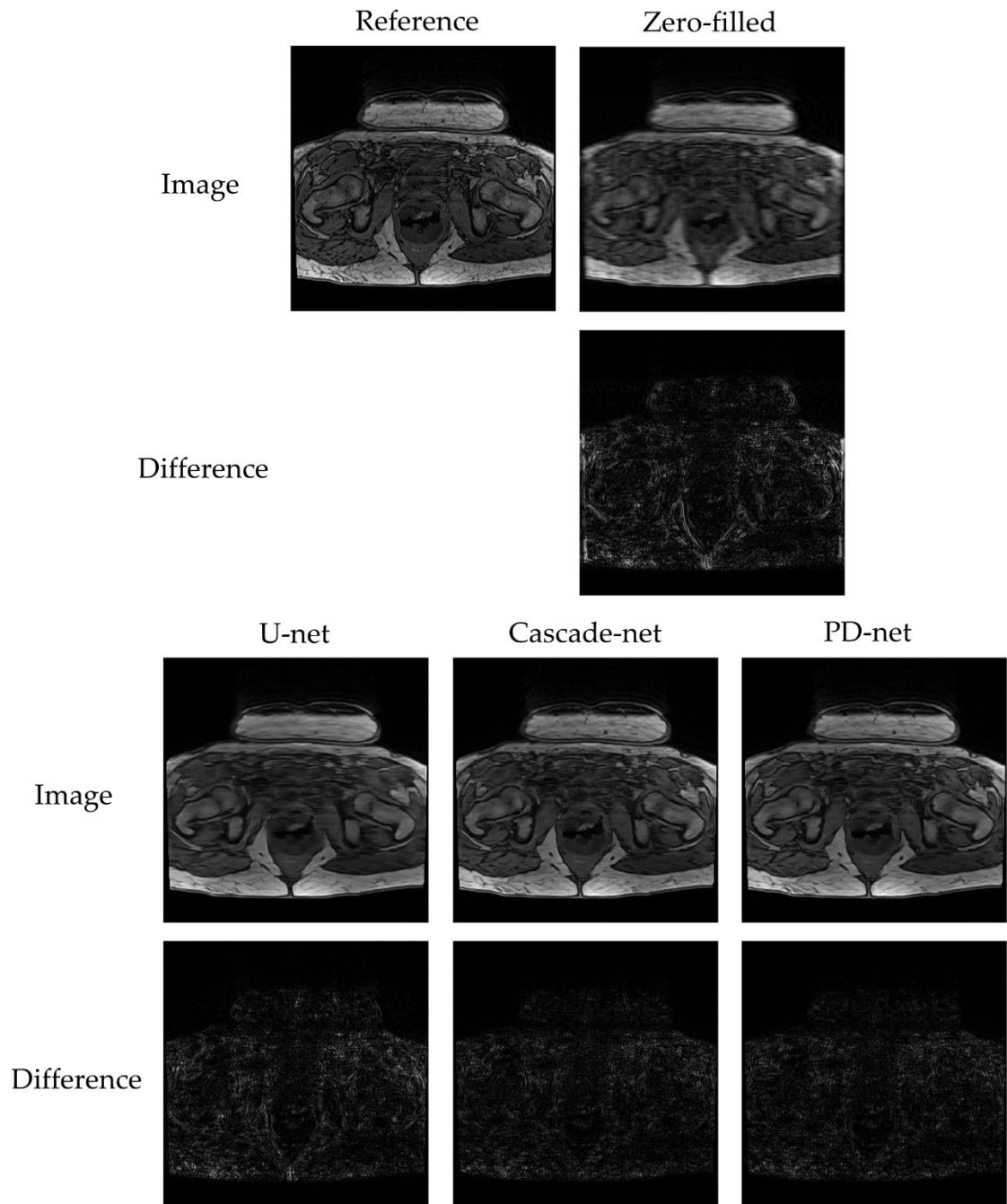


Figure 20: Qualitative evaluation of GYN T1 reconstruction

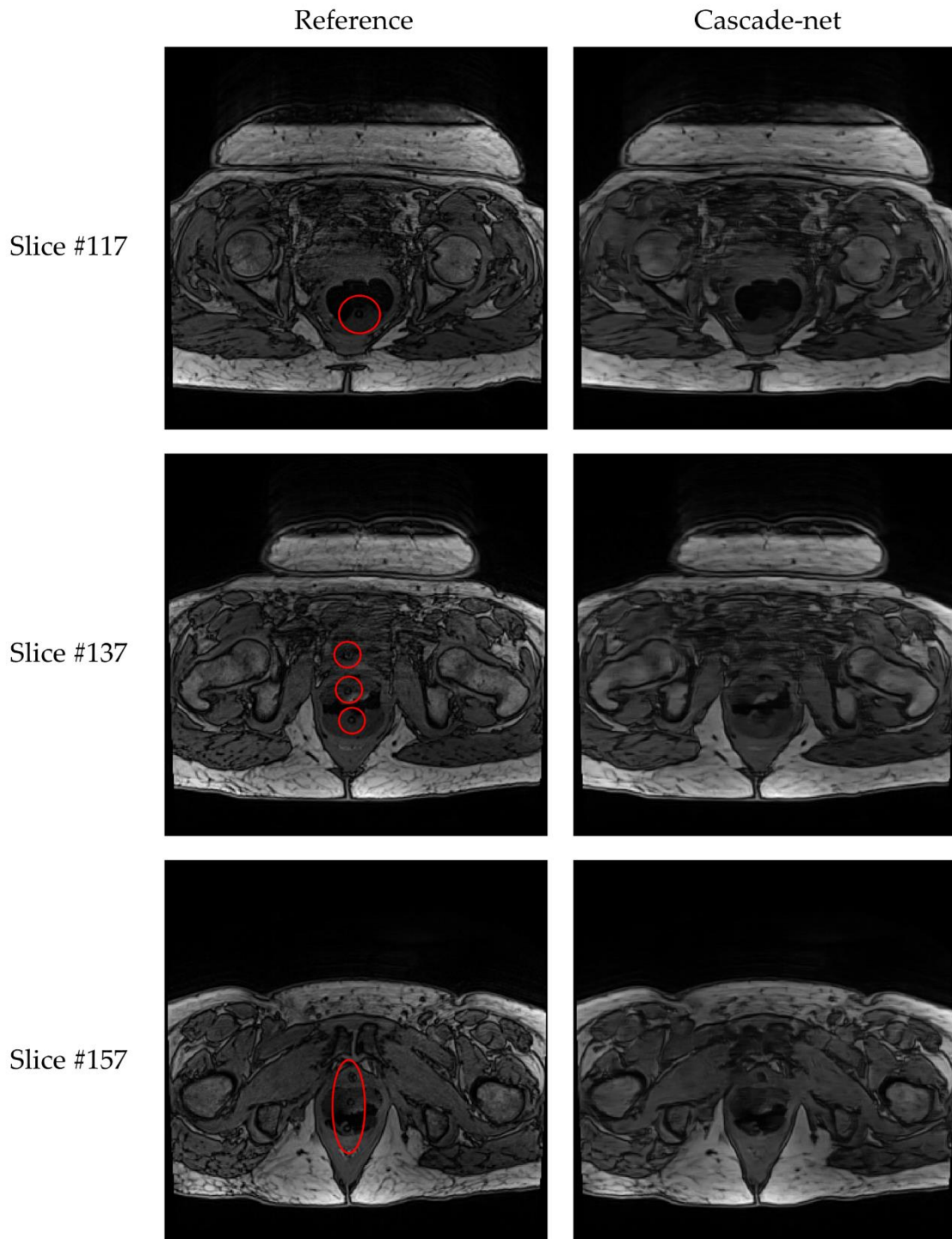


Figure 21: Sample slices for GYN T1-spgr Cascade-net reconstruction. Noting that the applicators circled in reference images are not visible in reconstructed images.

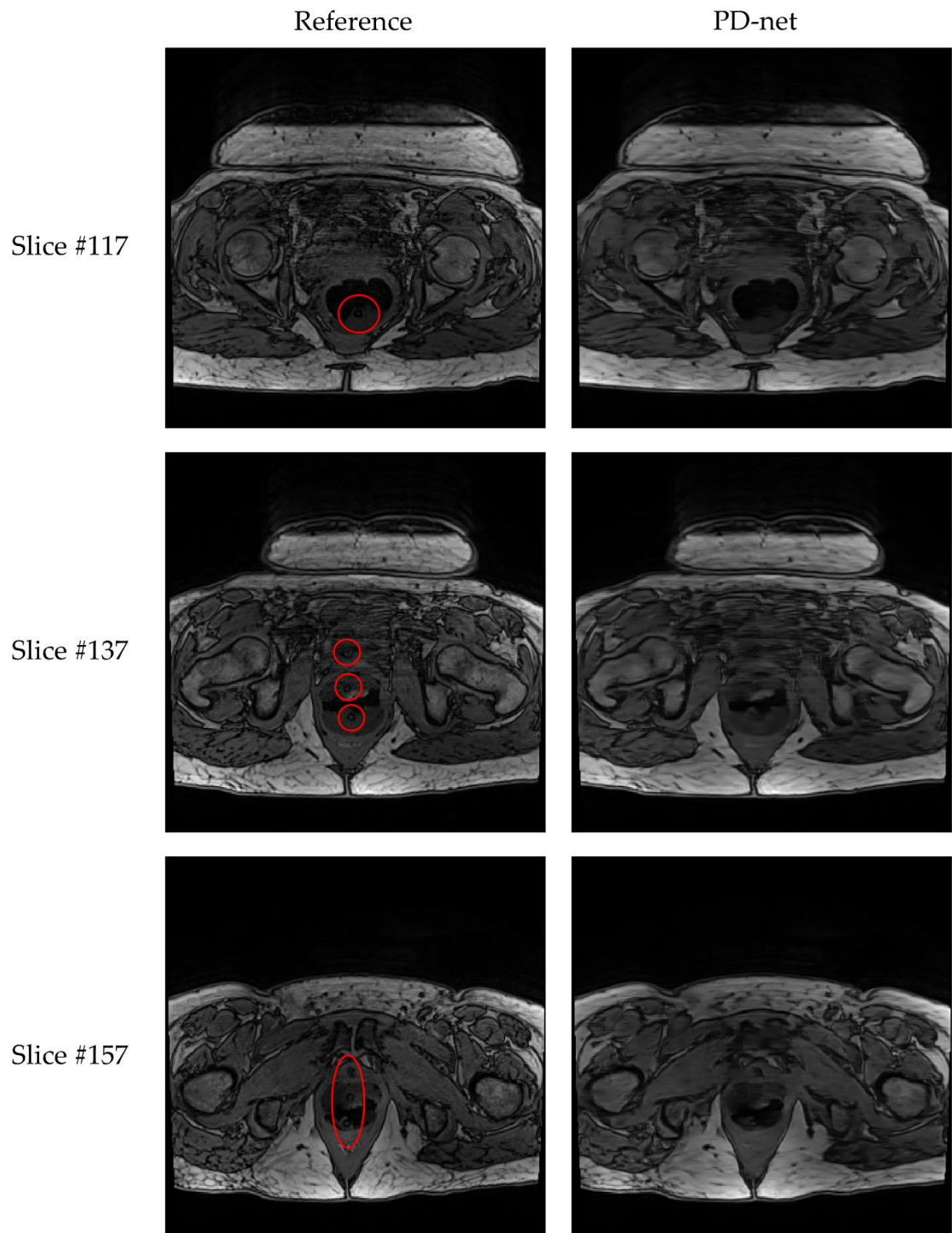


Figure 22: Sample slices for GYN T1-sgr PD-net reconstruction. Noting that the applicators circled in reference images are not visible in reconstructed images.

3.2.1.2 T2 reconstruction

A designated slice, slice #63 of a random patient (id and exam enclosed), was reconstructed for qualitative evaluation. The T2 images reconstructed using neural networks are shown in **Figure 23**, compared to the reference images and zero-filled reconstructed images. The corresponding differences between the reconstructed images and the reference images are also shown.

It can be told that compared to the zero-filled reconstructed images, the differences of the three images reconstructed using neural networks decrease significantly. Cascade-net and PD-net reconstructed images show better image quality than U-net reconstructed one. The applicators are also clearly visible in Cascade-net and PD-net reconstructed images, which is different from the situation in T1-spgr reconstruction. Multiple slices of Cascade-net and PD-net reconstructed images compared with reference are shown in **Figure 24-25** to compare with **Figure 21-22** for demonstration purpose.

In T2 images, we are particularly interested in pathologies, in our case, the high-risk clinic target volume (HRCTV). **Figure 26 (a)** shows the HRCTV contour based on a pre-treatment T2-w exam without applicators, directly screenshotted from Eclipse. **Figure 26 (b)** shows a few PD reconstructed slices from the exam at the second treatment fraction near the HRCTV level compared with reference for demonstration purpose. It can be told that the tissue in reconstructed images is still blur and fine structures are not perfectly reconstructed. The boundary between tissues is relatively clear. Since brachytherapy

applicators would change the shape of tissues, the contour in the pre-treatment T2 exam could not be directly interpreted on the images with applicators. Since the images with applicators are not contoured, and pathologies are not discernible due to a lack of medical background, these images are just shown to demonstrate general reconstruction performance and would not be quantitatively evaluated specific to reconstruction of pathologies.

Similarity between reconstructed images and reference were furtherly quantitatively evaluated using TRE.

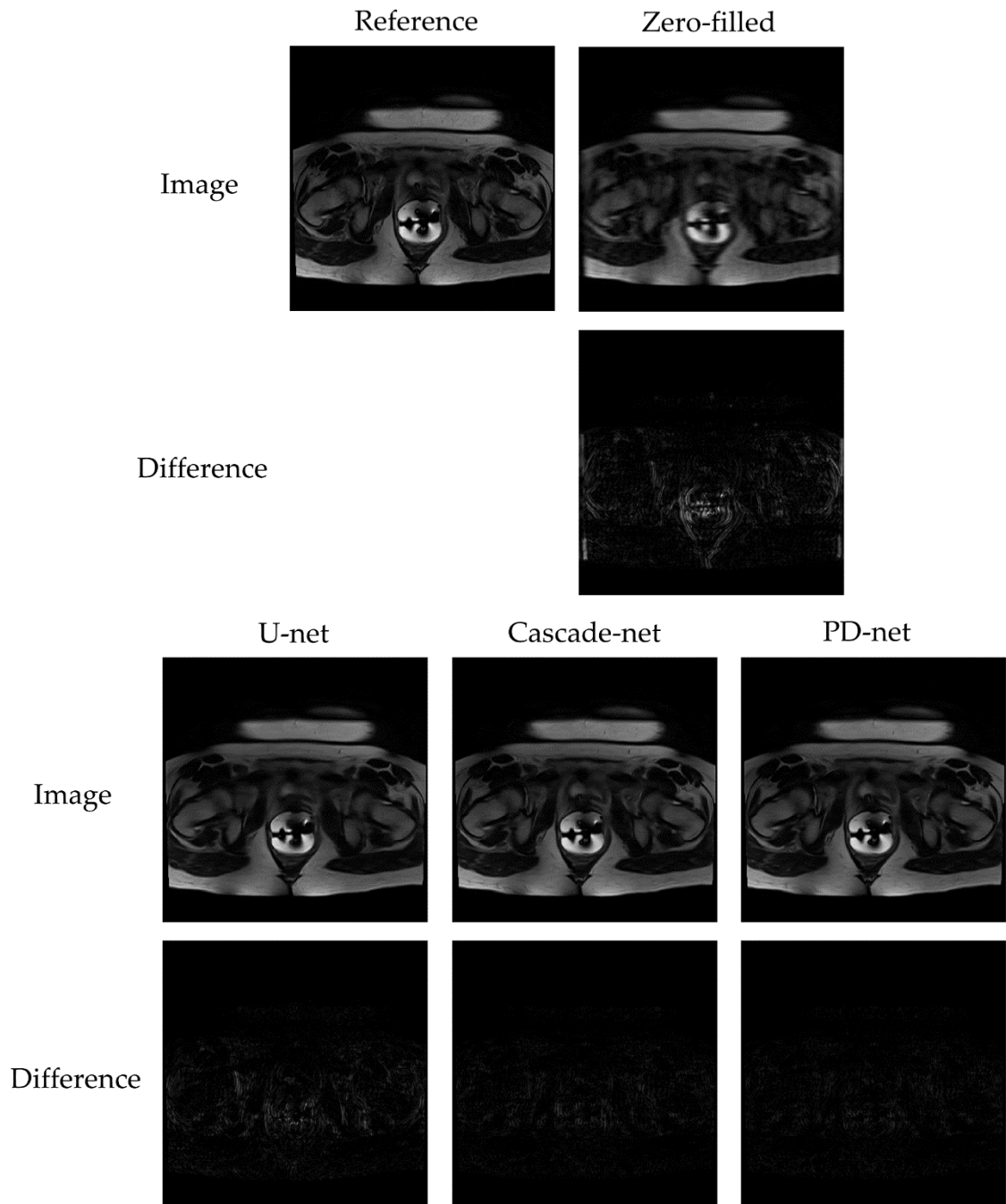


Figure 23: Qualitative evaluation of GYN T2 reconstruction

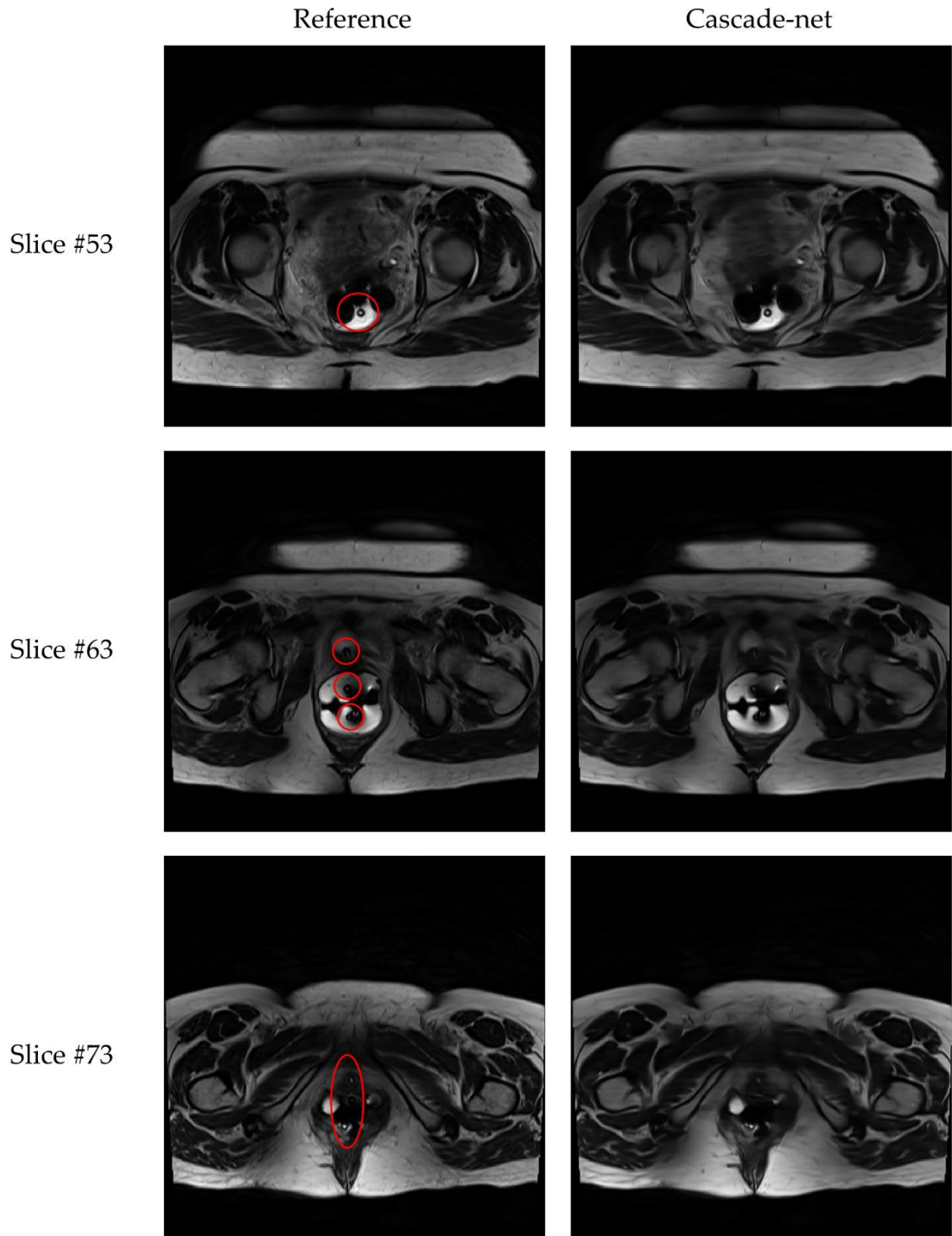


Figure 24: Sample slices for GYN T2 Cascade-net reconstruction. Noting that the applicators circled in reference images are clearly visible in reconstructed images.

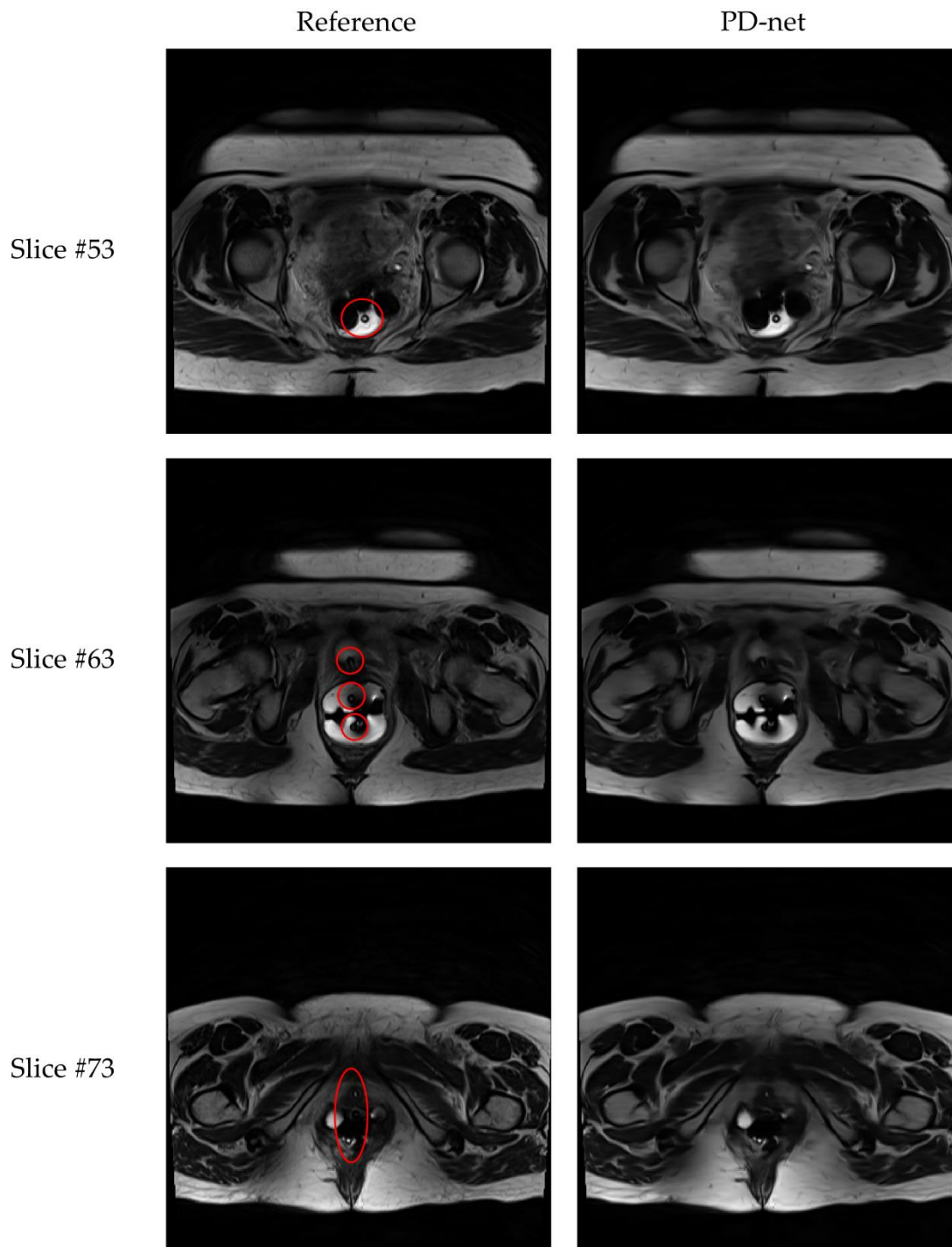


Figure 25: Sample slices for GYN T2 PD-net reconstruction. Noting that the applicators circled in reference images are clearly visible in reconstructed images.

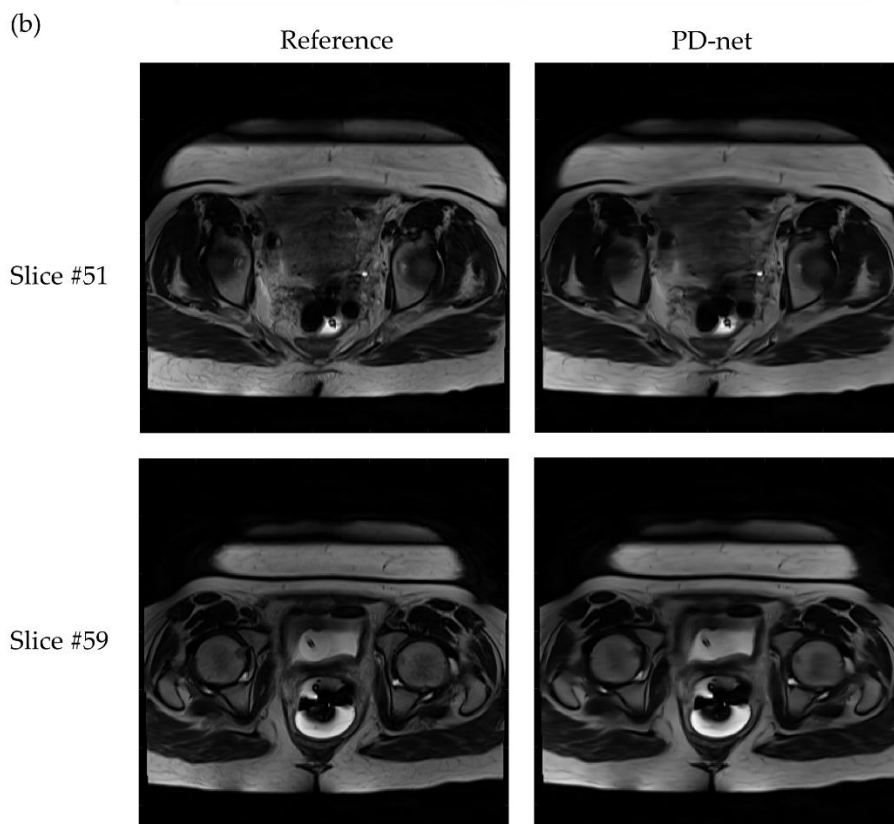
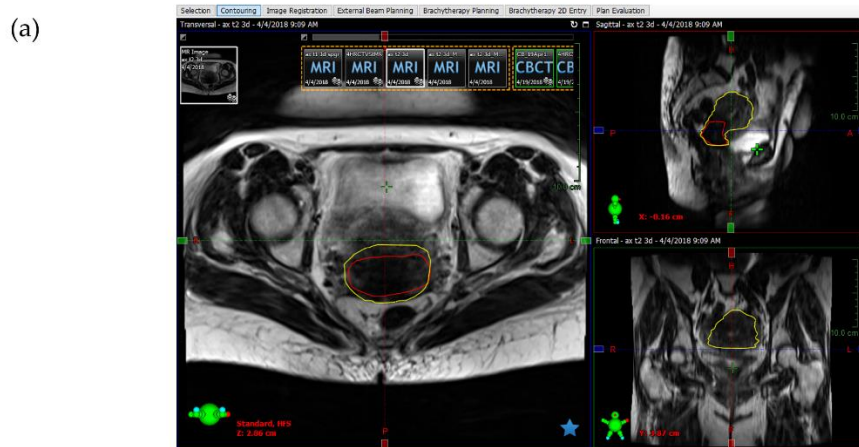


Figure 26 (a): HRCTV contour based on a pre-treatment T2 exam, denoted in red. (b): Sample slices for GYN T2 PD-net reconstruction relevant to HRCTV level. The reconstructed images are still blur to some extent.

3.2.2 Quantitative evaluation

The TRE of T1 and T2 images are tabulated in **Table 4** in a format of (mean± standard deviation). All of the slices from 3 exams of one patient were included into analysis.

Table 3: TRE evaluation of GYN reconstruction

Evaluation Metric	TRE (% , mean+std.)			
Reconstruction Method	Zero-filled	U-net	Cascade-net	PD-net
T1-spgr	0.215±0.021	0.176±0.015	0.146±0.017	0.145±0.016
T2	0.152±0.027	0.092±0.013	0.066±0.010	0.062±0.009

Comparing across reconstruction methods, PD-net generally worked slightly better than Cascade-net, and they both showed better performance than U-net. Comparing across image types, T2 reconstruction was significantly better than T1-spgr reconstruction, which is consistent with the qualitative evaluation. Standard deviations are all relatively small which indicates stable performance. These suggest that the three neural networks work well on GYN T2 reconstruction, while for GYN T1-spgr reconstruction task, the neural networks need further tuning.

4. Discussion

The three neural networks evaluated in this work have demonstrated excellent performance on standard MR anatomical image reconstruction, in other words, conventional T1 and T2 image reconstruction. [8, 13, 17-20] In this work, the networks are furtherly evaluated for their performance on other types of MRI including Brain DTI and Gynecological brachytherapy MR images. The structures and parameters were largely inspired by a previous benchmarking work which evaluated the performance of the neural networks on Brain and Knee T1 and T2 images. [27]

The performance of the neural networks was qualitatively and quantitatively evaluated respectively. Generally speaking, PD-net slightly outperformed Cascade-net, and they both outperformed U-net in all the evaluation. The performance can be furtherly discussed in two parts.

For Brain DTI reconstruction, the reconstruction of individual diffusion weighted image set was of good quality. However, the reconstruction of quantitative functional maps was not very ideal, since the quantitative functional maps were reconstructed based on 13 individual diffusion weighted image set, where misalignment and error can easily add up. This indicate that to achieve reconstruction of high quality, the error of diffusion weighted image reconstruction should be as low as possible, which propose a great requirement for neural network performance.

For cervix GYN reconstruction, the reconstruction of T2 images was of very great quality, while the reconstruction of T1 images was not ideal. Applicators, especially digitizers, which were perfectly shown in reference images, could be clearly reconstructed in T2 images but failed to be clearly reconstructed in T1 images. Combining the fact that Brain DTI set was indeed T2-weighted and reconstruction on Brain DTI set achieved great results on diffusion weighted image reconstruction, there is a chance that current settings of architecture and parameters made the neural networks tend to reconstruct T2-weighted images better than T1-weighted images.

Several works can be done to further improve the performance of these two tasks. Firstly, neural networks can be furtherly fine-tuned to achieve a better performance on Brain diffusion weighted image reconstruction and cervix GYN T1-spgr image reconstruction. Secondly, better evaluation tools can be developed, such as tools that can evaluate the similarity between colored images and tools that can automatically focus on local features such as brachytherapy applicators and pathologies. Thirdly, the undersampling strategy can be furtherly improved. Radial sampling has shown better performance compared to Cartesian undersampling strategy with the same acceleration factor. [32-33] Radial sampling would grasp more information in the central k-space which would improve the inherent image quality.

5. Conclusion

Three neural networks including U-net, Cascade-net and PD-net showed great performance on accelerating Brain DTI and GYN T2 MR images adopting a Cartesian undersampling strategy. PD-net slightly outperformed Cascade-net, and they both showed better performance than U-net. This is still true regarding GYN T1-spgr MR image reconstruction, but the overall performance of T1 image reconstruction needs further improvement. Overall, PD-net with current architecture and parameters is promising on T2-weighted-base MR image reconstruction.

Appendix

The diffusion gradient directions for RIDER NEURO MRI DTI scan were not published. They were acquired by directly contacting the original author and are showed as follows.

Diffusion gradient #	x	y	z
1	0	0	0
2	0.894	0	0.447
3	0	0.447	0.894
4	0.447	0.894	0
5	0.894	0.447	0
6	0	0.894	0.447
7	0.447	0	0.894
8	0.894	0	-0.447
9	0	-0.447	0.894
10	-0.447	0.894	0
11	0.894	-0.447	0
12	0	0.894	-0.447
13	-0.447	0	0.894

Bibliography

- [1] Deshmane, Anagha, et al. "Parallel MR Imaging." *Journal of Magnetic Resonance Imaging*, vol. 36, no. 1, 2012, pp. 55–72., doi:10.1002/jmri.23639.
- [2] Geethanath, Sairam, et al. "Compressed Sensing MRI: A Review." *Critical Reviews in Biomedical Engineering*, vol. 41, no. 3, 2013, pp. 183–204., doi:10.1615/critrevbiomedeng.2014008058.
- [3] Lustig, Michael, et al. "Sparse MRI: The Application of Compressed Sensing for Rapid MR Imaging." *Magnetic Resonance in Medicine*, vol. 58, no. 6, 2007, pp. 1182–1195., doi:10.1002/mrm.21391.
- [4] Glover, P M. "Interaction of MRI Field Gradients with the Human Body." *Physics in Medicine and Biology*, vol. 54, no. 21, 2009, doi:10.1088/0031-9155/54/21/r01.
- [5] Haase, A, et al. "FLASH Imaging. Rapid NMR Imaging Using Low Flip-Angle Pulses." *Journal of Magnetic Resonance (1969)*, vol. 67, no. 2, 1986, pp. 258–266., doi:10.1016/0022-2364(86)90433-6.
- [6] Poustchi-Amin, Mehdi, et al. "Principles and Applications of Echo-Planar Imaging: A Review for the General Radiologist." *RadioGraphics*, vol. 21, no. 3, 2001, pp. 767–779., doi:10.1148/radiographics.21.3.g01ma23767.
- [7] Shannon, C.e. "Communication In The Presence Of Noise." *Proceedings of the IEEE*, vol. 86, no. 2, 1998, pp. 447–457., doi:10.1109/jproc.1998.659497.
- [8] Zhang, HM., Dong, B. "A Review on Deep Learning in Medical Image Reconstruction." *J. Oper. Res. Soc. China*, 8, 2020, pp. 311–340., doi:10.1007/s40305-019-00287-4
- [9] Lecun, Yann, et al. "Deep Learning." *Nature*, vol. 521, no. 7553, 2015, pp. 436–444., doi:10.1038/nature14539.
- [10] Wu, J. "Introduction to Convolutional Neural Networks." (2017).
- [11] He, Kaiming, et al. "Deep Residual Learning for Image Recognition." *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, doi:10.1109/cvpr.2016.90.
- [12] Dumoulin, Vincent, and Francesco Visin. "A guide to convolution arithmetic for deep learning." 2016. Available: <http://arxiv.org/abs/1603.07285>.

- [13] Schlemper, Jo, et al. "A Deep Cascade of Convolutional Neural Networks for Dynamic MR Image Reconstruction." *IEEE Transactions on Medical Imaging*, vol. 37, no. 2, 2018, pp. 491–503., doi:10.1109/tmi.2017.2760978.
- [14] Nwankpa Chigozie, et al. "Activation Functions: Comparison of Trends in Practice and Research for Deep Learning." 2018. Available: <https://arxiv.org/abs/1811.03378>
- [15] Yang, Yan, et al. "ADMM-CSNet: A Deep Learning Approach for Image Compressive Sensing." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 3, 2020, pp. 521–538., doi:10.1109/tpami.2018.2883941.
- [16] Eo, Taejoon, et al. "KIKI -Net: Cross-Domain Convolutional Neural Networks for Reconstructing Undersampled Magnetic Resonance Images." *Magnetic Resonance in Medicine*, vol. 80, no. 5, 2018, pp. 2188–2201., doi:10.1002/mrm.27201.
- [17] Adler, Jonas, and Ozan Oktem. "Learned Primal-Dual Reconstruction." *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, 2018, pp. 1322–1332., doi:10.1109/tmi.2018.2799231.
- [18] Cheng, Jing, et al. "Model Learning: Primal Dual Networks for Fast MR Imaging." *Lecture Notes in Computer Science Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, 2019, pp. 21–29., doi:10.1007/978-3-030-32248-9_3.
- [19] Ronneberger, Olaf, et al. "U-Net: Convolutional Networks for Biomedical Image Segmentation." *Lecture Notes in Computer Science Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, 2015, pp. 234–241., doi:10.1007/978-3-319-24574-4_28.
- [20] Hyun, Chang Min, et al. "Deep Learning for Undersampled MRI Reconstruction." *Physics in Medicine & Biology*, vol. 63, no. 13, 2018, p. 135007., doi:10.1088/1361-6560/aac71a.
- [21] Schaefer, P W et al. "Diffusion-weighted MR imaging of the brain." *Radiology* vol. 217,2 (2000): 331-45. doi:10.1148/radiology.217.2.r00nv24331
- [22] Stejskal, E. O., and J. E. Tanner. "Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient." *The Journal of Chemical Physics*, vol. 42, no. 1, 1965, pp. 288–292., doi:10.1063/1.1695690.

- [23] Mukherjee, P., et al. "Diffusion Tensor MR Imaging and Fiber Tractography: Theoretic Underpinnings." *American Journal of Neuroradiology*, vol. 29, no. 4, 2008, pp. 632–641., doi:10.3174/ajnr.a1051.
- [24] Tian, Qiyuan, et al. "DeepDTI: High-Fidelity Six-Direction Diffusion Tensor Imaging Using Deep Learning." *NeuroImage*, vol. 219, 2020, p. 117017., doi:10.1016/j.neuroimage.2020.117017.
- [25] Jones, Derek K., et al. "White Matter Integrity, Fiber Count, and Other Fallacies: The Do's and Don'ts of Diffusion MRI." *NeuroImage*, vol. 73, 2013, pp. 239–254., doi:10.1016/j.neuroimage.2012.06.081.
- [26] Report 89, *Journal of the International Commission on Radiation Units and Measurements*, Volume 13, Issue 1-2, April 2013, Page NP, <https://doi.org/10.1093/jicru/ndw042>
- [27] Ramzi, Zaccharie, et al. "Benchmarking MRI Reconstruction Neural Networks on Large Public Datasets." *Applied Sciences*, vol. 10, no. 5, 2020, p. 1816., doi:10.3390/app10051816.
- [28] Barboriak, Daniel. (2015). Data From RIDER_NEURO_MRI. The Cancer Imaging Archive. <http://doi.org/10.7937/K9/TCIA.2015.VOSN3HN1>
- [29] Zhang R. Michael, et al. "Lookahead Optimizer: k steps forward,1 step back." 2019. Available: <https://arxiv.org/abs/1907.08610>
- [30] Kingsley, Peter B. "Introduction to Diffusion Tensor Imaging Mathematics: Part III. Tensor Calculation, Noise, Simulations, and Optimization." *Concepts in Magnetic Resonance Part A*, vol. 28A, no. 2, 2006, pp. 155–179., doi:10.1002/cmr.a.20050.
- [31] Tsougos, Ioannis. "Artifacts and Pitfalls in Diffusion MRI." *Advanced MR Neuroimaging*, 2017, pp. 29–54., doi:10.1201/9781351216548-2.
- [32] Feng, Li, et al. "Golden-Angle Radial Sparse Parallel MRI: Combination of Compressed Sensing, Parallel Imaging, and Golden-Angle Radial Sampling for Fast and Flexible Dynamic Volumetric MRI." *Magnetic Resonance in Medicine*, vol. 72, no. 3, 2013, pp. 707–717., doi:10.1002/mrm.24980.

- [33] Freitas, Andreia C., et al. "Comparison of Cartesian and Non-Cartesian Real-Time MRI Sequences at 1.5T to Assess Velar Motion and Velopharyngeal Closure during Speech." *Plos One*, vol. 11, no. 4, 2016, doi:10.1371/journal.pone.0153322.