Deep Learning Applications in Computational MRI: A Thesis in Two Parts

Sukrit Arora
Michael Lustig, Ed.
Stella Yu, Ed.

Electrical Engineering and Computer Sciences
University of California, Berkeley

Technical Report No. UCB/EECS-2021-91
http://www2.eecs.berkeley.edu/Pubs/TechRpts/2021/EECS-2021-91.html

May 14, 2021

Acknowledgement

I would like to thank my Faculty Advisor, Professor Michael (Miki) Lustig, for giving me the opportunity to do research in his lab, and for all his project insights, guidance and support. I would also like to thank my second reader Professor Stella Yu for her guidance on my projects as well.

There are also several people that have helped me with my research work and my thesis that I would like to thank, in no particular order: Volkert Roeloffs, Ekin Karasan, Ke Wang, Shreya Ramachandran, Alfredo De Goyeneche, Jon Tamir, Frank Ong, Zhiyong Zhang, and Mia Mirkovic.

Last but not least, I would like to thank my wonderful grandparents, parents, and little sister for their continual love and emotional support.

Deep Learning Applications in Computational MRI: A Thesis in Two Parts

by

Sukrit Arora

A thesis submitted in partial satisfaction of the

requirements for the degree of

Masters of Science

in

Electrical Engineering and Computer Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Michael Lustig, Chair
Professor Stella Yu

Spring 2021

The thesis of Sukrit Arora, titled Deep Learning Applications in Computational MRI: A Thesis in Two Parts, is approved:

Chair	Michael Lustig	Date	05/12/2021
	Stella Yu	Date	5/12/2021
		Date	

University of California, Berkeley

Deep Learning Applications in Computational MRI: A Thesis in Two Parts

Contents

List of Figures

Acknowledgments

There are several people I would like to thank, for without them my work and this thesis would not exist. First and foremost, I would like to thank my Faculty Advisor, Professor Michael (Miki) Lustig, for giving me the opportunity to do research in his lab, and for all his project insights, guidance and support. I would also like to thank my second reader Professor Stella Yu for her guidance on my projects as well.

There are also several people that have helped me with my research work and my thesis that I would like to thank, in no particular order: Volkert Roeloffs, Ekin Karasan, Ke Wang, Shreya Ramachandran, Alfredo De Goyeneche, Jon Tamir, Frank Ong, Zhiyong Zhang, and Mia Mirkovic.

Last but not least, I would like to thank my wonderful grandparents, parents, and little sister for their continual love and emotional support.

Abstract

Deep Learning Applications in Computational MRI: A Thesis in Two Parts

by

Sukrit Arora

Masters of Science in Electrical Engineering and Computer Science

University of California, Berkeley

Professor Michael Lustig, Chair

Magnetic resonance imaging (MRI) is a powerful medical imaging modality that provides diagnostic information without the use of harmful ionizing radiation. As with any imaging modality, the raw acquired data must undergo a series of processing steps in order to produce an image. During these steps, several problems may arise that impact the quality of the final image. While traditional signal and image processing methods have been employed with great success to address these issues, as the field of deep learning has grown, so too has the research of these methods to address problems in MRI. This thesis, a thesis in two parts, will discuss deep-learning approaches for addressing three common problems in computational MRI: noise, reconstruction, and off-resonance blurring. The first part of the thesis proposes the use of an untrained modified deep decoder network in order to denoise as well as reconstruct MR images. The second part of the thesis investigates the generalizability of a convolutional residual network in its ability to correct for blurring due to off-resonance.

Chapter 1

Introduction

Magnetic resonance imaging (MRI) is a powerful medical imaging modality that provides diagnostic information without the use of harmful ionizing radiation. During an MRI scan, a constant, homogeneous main magnetic field aligns molecular spins in the body. Short radiofrequency pulses then disrupt the alignment of H+ spins in water and generate a signal, due to resonance of these spins at a frequency ω_0 proportional to the main magnetic field strength [18]. This RF signal is received and demodulated at ω_0, and the subsequent baseband signal corresponds to samples of the spatial frequency domain, known as k-space. An image can be reconstructed by applying an inverse Fourier transform to the sampled k-space.

MRI is also an inherently multi-dimensional imaging modality. MRI data is generally volumetric and/or multi-coil: most MRI machines have multiple RF receivers. In multichannel MRI, each of the receiver coils only partially samples the k-space, thereby reducing scan time. The image is reconstructed from all the k-space samples collected from all the coils [32]. In the image domain, multiple coils can be thought of as different illuminations of the same subject.

Problems in MRI

This thesis will discuss deep-learning approaches for addressing three common problems in MRI: noise, reconstruction, and off-resonance blurring.

Noise is commonly attributed to the radio-frequency receiver in the MR imaging system, and can arise from a variety of physical sources [23]. These noise sources are all expected to have a "white" spectrum; that is, over the course of time, the mean-square voltage fluctuations occur with equal amplitude across all frequencies detectable by the system [35]. It has also been shown that a Gaussian model characterizes this systematic noise well [19].

Another complex problem in MRI is reconstruction; namely, reconstruction from incomplete k-space data. As briefly mentioned above, the baseband signal that the MRI machine

receives corresponds to spatial frequency samples. In order to reconstruct an image from these samples, an inverse Fourier transform must be taken. However, this is only possible in the trivial case of a fully sampled k-space domain. Reconstructing MR images from sub-sampled k-space data is appealing since it requires fewer measurements and therefore shorter scan times. Reconstruction is also important in cases where fully sampled data is difficult or impossible to acquire [10].

A third problem encountered in MRI is blurring due to off-resonance. As previously mentioned, our imaging signal model assumes that the nuclear spins resonate at the same frequency of ω_0. However, in reality, not all spins in the body resonate exactly at ω_0, due to either inhomogeneities in the main magnetic field or excitation of H+ in other complex molecules, such as fat [37]. Spins actually resonate at a range of frequencies around ω_0, which is known as off-resonance. Off-resonance causes spatial blurring in the reconstructed image, especially when k-space is sampled with non-Cartesian trajectories [33].

A more detailed description of these problems and this thesis' solution to those problems will be given under the relevant project descriptions.

Conventional Solutions

As electrical engineering and computer science have evolved, so too have the methods used to address these problems in the field of MRI. Historically, these kinds of problems have been solved using non-learning based, established methods from statistical signal processing and image processing, such as different forms of filtering [2], compressed sensing [31], and conjugate phase reconstruction [8]

The Deep Learning Boom, Applied to MRI

However, over the past decade, the development of novel deep learning methods like CNNs, generative models, residual models, etc. [28, 17, 21] has created new solutions to a wide variety of imaging problems, from image compression and denoising to inpainting and superresolution [41, 5, 29], often outperforming methods based on classical image processing.

As these methods gain popularity in the overall imaging field, research on applying these methods to MRI has grown immensely due to their applicability to the particular challenges posed by MRI and medical imaging in general, such as limited data availability and unique data acquisition. My thesis gives two examples of deep learning applications for MRI. Part I discusses a project that uses an unsupervised generative CNN for the tasks of denoising and reconstruction, and Part II covers a project that works on generalizing off-resonance blurring correction using deep residual learning.

Part I

Untrained Modified Deep Decoder for Joint Denoising and Parallel Imaging Reconstruction

Chapter 2

MRI Denoising and Reconstruction

2.1 Abstract

An untrained deep learning model based on a Deep Decoder was used for image denoising and parallel imaging reconstruction. The flexibility of the modified Deep Decoder to output multiple images was exploited to jointly denoise images from adjacent slices and to reconstruct multi-coil data without pre-determined coil sensitivity profiles. Higher PSNR values were achieved compared to the traditional methods of denoising (BM3D) and image reconstruction (Compressed Sensing). This untrained method is particularly attractive in scenarios where access to training data is limited, and provides a possible alternative to conventional sparsity-based image priors.

2.2 Background

The Imaging Forward Model

In order to discuss the tasks of denoising and reconstruction, we must first develop a mathematical model for the MR imaging system. For an MR image $x \in \mathcal{C}$, we can model our measurements of x as $y = Ax + \eta$, where A is a known encoding, or measurement, matrix and η is noise, modeled by a white gaussian process [19]. In the case of denoising, $A = I$ and y is the noisy MR image. In the case of reconstruction, y is a set of noisy k-space measurements, and $A = P_k F$, where F is the Fourier matrix, and P_k is a sampling mask that picks rows of the Fourier matrix. In both cases, we aim to find an x that minimizes $||Ax - y||^2$.

This problem requires regularization for several reasons. In the denoising case, we only have access to a noisy image and want to prevent any solver from overfitting to the noise. In the reconstruction case, because the goal of MRI reconstruction is to reduce the number of measurements needed to produce an image, we have fewer measurements than unknowns, and

the resultant encoding matrix A is underdetermined. So, in order to solve these problems, we need some kind of regularization that serves as a prior. Only then do we have hope of recovering a good approximation of the true x.

Traditional methods have leveraged the idea that MR images are inherently sparse in some transform domain. These methods incorporate this prior by solving

$$\min_x ||Ax - y||_2^2 + \lambda \mathcal{R}(x)$$

where $\mathcal{R}(x)$ is the regularization that incorporates the sparsity prior. Classical choices for this regularization include using the norm of the wavelet coefficients of x [31][15], or the total variation or discrete gradient of x [24].

Deep Learning Regularization

With the rise in the popularity of deep learning, there have also been several methods, such as MoDL [1] and variational networks [20], that suggest using some form of a deep convolutional neural network (CNN) as a prior. Since CNNs are non-linear and more expressive, they may learn prior image statistics that can aid in denoising and reconstruction, potentially generating better results than traditional linear methods.

While these methods perform well, with comparable and usually improved results over established methods, they are dependent on two key things: access to a large training dataset and ground truth data. While this access is readily available for most classes of images due to open source datasets such as ImageNet, Places2, Cifar, etc. [39, 50, 27], the same kind of accessibility usually does not exist for MR images. This is both because ground truth scans require fully sampled data (which is rarely done clinically due to length of scans), and because there is a higher barrier of access to data due to privacy and data sensitivity. For these reasons, computational MRI data is not available on the same scale as other image data. This is what makes untrained or unsupervised methods so appealing for these tasks: they can leverage the complexity of these models without needing large datasets for training.

The first work that explored the idea of using untrained networks for solving these kinds of image inverse problems was Deep Image Prior (DIP) [42]. In this work, they show that a great deal of the image statistics that are essential to solving certain image restoration problems are captured by the structure of generator ConvNets, independent of learning. Heckel later expanded on the idea of an untrained network as a prior, presenting a simpler yet improved network architecture called a Deep Decoder (DD) [22]. Unlike DIP, which uses an overparameterized U-Net, the DD uses an underparameterized decoder network. This enables it to draw strong comparisons to classical image representations such as sparse wavelet representations, which made it appealing to use in an MRI context.

In this work, we modify the DD to create the mDD and investigate its use to perform two different tasks without training: multi-slice denoising and parallel imaging reconstruction.

2.3 Methods

The deep decoder network is a generative decoder model that starts with a fixed, randomly initialized tensor z, and generates an output image through pixel-wise linear combinations of channels, ReLU non-linearities, channel normalizations, and bilinear interpolation upsampling. Specifically, at every layer, the network performs the following operation:

$$T_{l+1} = U_l \texttt{cn}(\texttt{relu}(T_l C_l)) \quad \forall l \in \{0, \ldots, d-1\}$$

where T_l is the tensor at layer l and $C_l \in \mathcal{R}^{k \times k}$ are the weights of the network at layer l. The expression $T_l C_l$ is a linear combination of T_l across the channels, done in a way that is consistent across all pixels.

Finally, the output of the network $G_w(z)$ is formed by $G_w(z) = T_d C_d$, where $C_d \in \mathcal{R}^{k \times k_{out}}$ (Figure 2.1).

The original DD network is modified in three ways: First, non-integer upsampling factors between layers were implemented to enable arbitrary output dimensions; second, the sigmoid in the final layer was removed to support an unconstrained output range; and third, the output channels of the network were identified with either multiple slices (denoising) or multiple channels (image reconstruction).

This modified DD (mDD) is then used to solve inverse problems by minimizing the objective function:

$$f(w) = ||AG_w(z) - y||_2^2$$

with respect to the weights w for an observation y and a given forward model A. As aforementioned, for denoising, we set $A = I$ and identify the network output channels with the set of slices to be jointly denoised. For multi-coil image reconstruction, we set $A = P_k \mathcal{F}$, where P_k is the k-space subsampling mask and $\mathcal{F}$ is the Fourier Transform. Here, we identify the network output channels with the individual coil images. In either case, the network never sees any ground truth data — only the noisy images(s) or the subsampled k-space.

The key innovation of the mDD is the ability to have a variable number of output channels, which are then used to jointly denoise or reconstruct. This enables us to leverage the multidimensionality of MRI data in order to improve the performance of the network. In the case of denoising, we take advantage of the volumetric nature of MRI data and denoise multiple successive slices of the same subject simultaneously: because these slices are adjacent, they have a strong correlation, which the network can leverage to improve denoising performance compared to its results for individual slices. Similarly, in the reconstruction case, the mDD simultaneously learns multiple coil images, taking advantage of common information

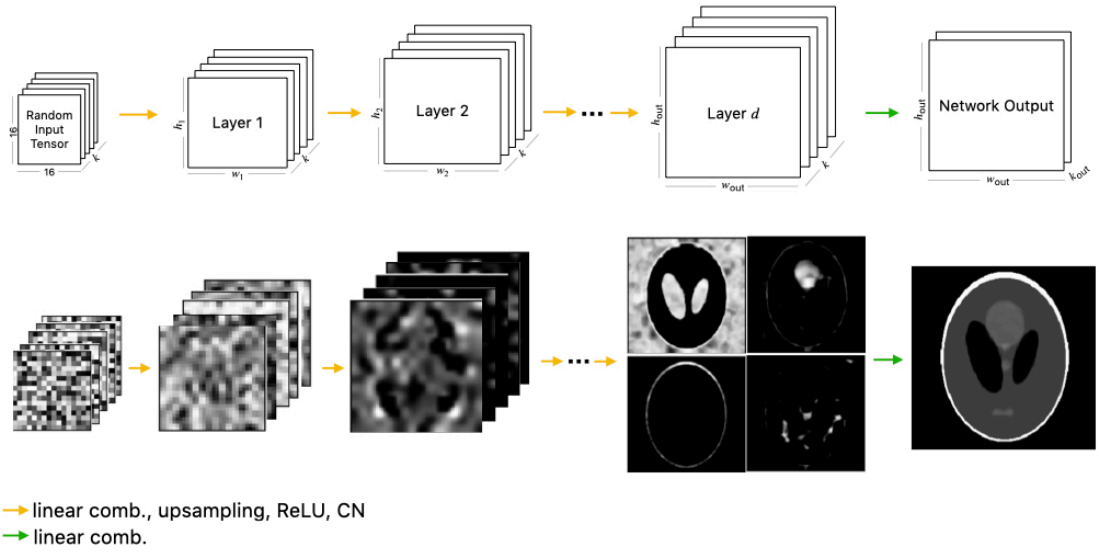

Figure 2.1: (Top) Deep Decoder network architecture visualized. (Bottom) Activation maps for network visualized for a Shepp-Logan phantom with $k = 64$ and $k_{\texttt{out}} = 1$. Selected channels illustrate data flow.

between images. A further advantage to this method is that, unlike some other multi-coil reconstruction methods, there is no need for predetermined coil sensitivity estimates.

2.4 Results

Figure 2.2 shows the effect of jointly denoising 10 adjacent slices of a synthetic data set (BrainWeb) [9]. White Gaussian noise was added to the dataset, and then denoised using BM3D [12], individual mDD (denoising slices separately using the mDD), and joint mDD (denoising slices simultaneously using the mDD). Joint denoising outperforms both BM3D denoising and single-slice mDD denoising, and results in a maximal pSNR improvement of 1.54 dB. BM3D leaves some blocky artifacts, giving a pixelated appearance, while the single-slice mDD instead blurs some of the details. Joint denoising preserves structure better while reducing artifacts.

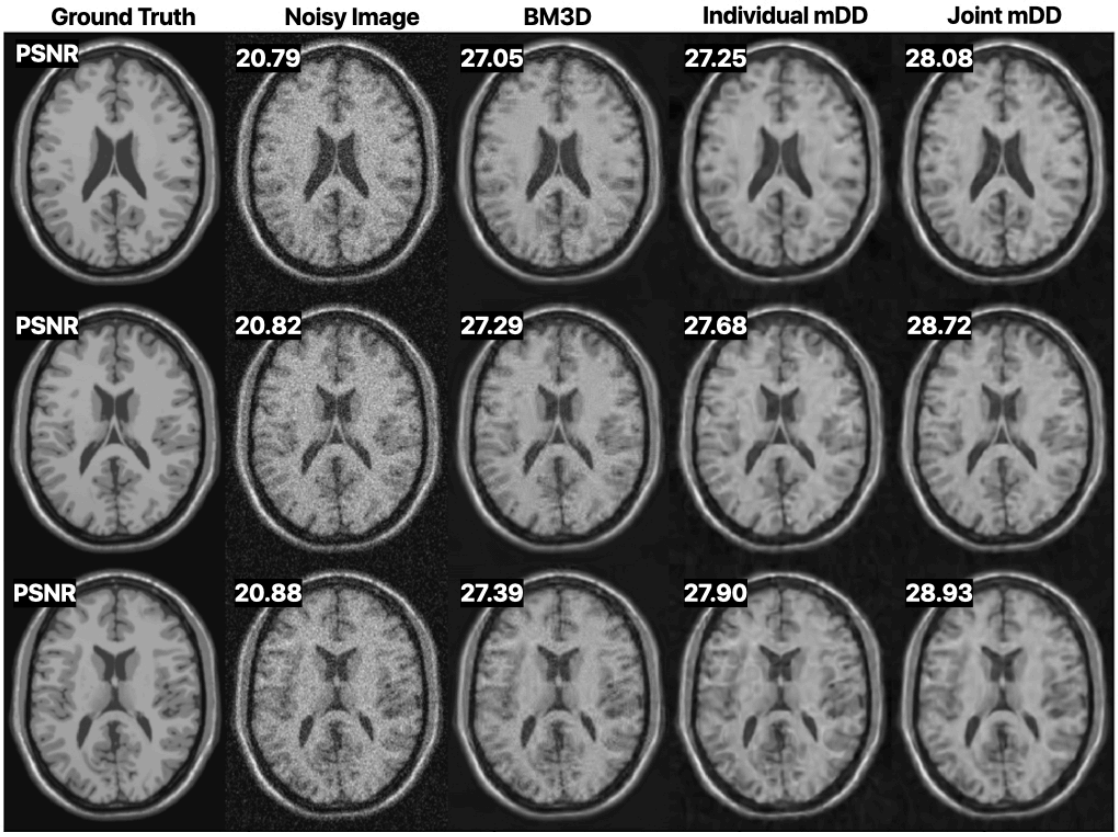

Figure 2.2: Selected 3 out of 10 adjacent slices denoised using BM3D, single-slice denoising ($k = 64$, $k_{\texttt{out}} = 1$) mDD, and joint denoising ($k = 64$, $k_{\texttt{out}} = 10$) mDD where the network denoises all 10 slices simultaneously.

Figure 2.3 shows a 4x and 8x accelerated parallel imaging reconstruction of acquired 15 channel knee data from the FastMRI NYU dataset [46]. The results of the mDD reconstruction are compared to a Zero Filled inverse FFT (ZF) and Parallel Imaging Compressed Sensing (PICS) [31]. Similar to its denoising performance, the mDD performs better than PICS with a maximal pSNR improvement of 1.39 dB in the 4x case and 2.58 dB in the 8x case.

2.5 Discussion

Our results show that joint denoising preserves structure better and reduces the artifact level compared to individually optimizing with the mDD. For multi-coil reconstruction, the mDD generates images with a reduced level of aliasing artifacts. We hypothesize that the network output is biased toward smooth, unaliased images. This is because the network,

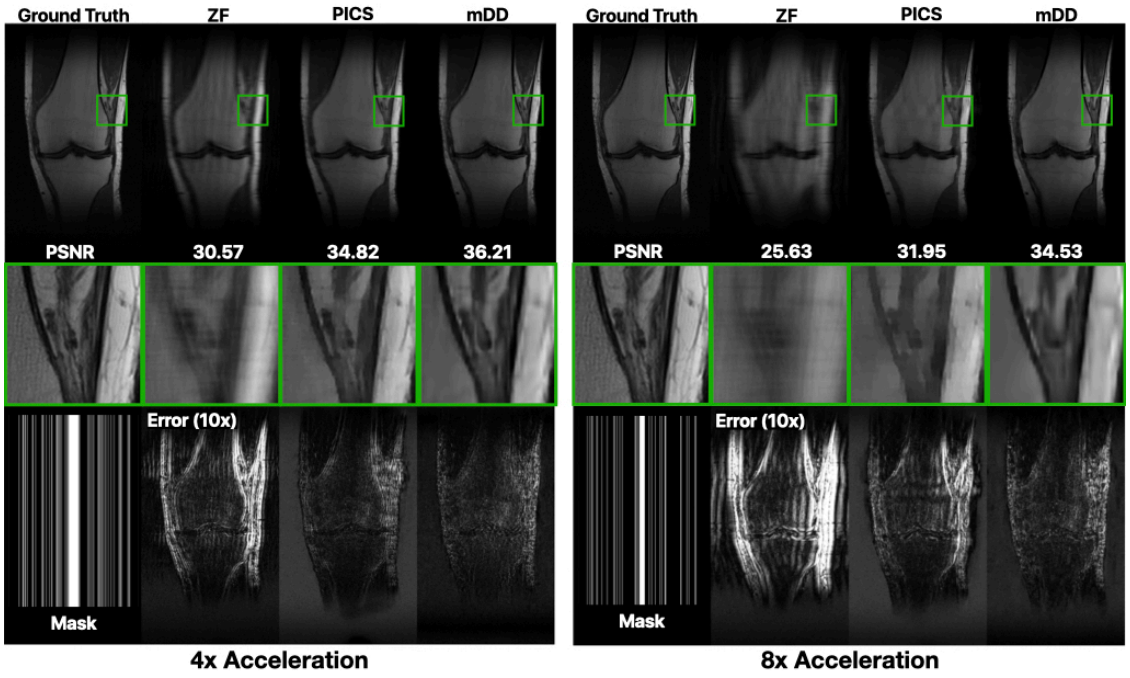

Figure 2.3: Reconstruction of subsampled k-space data with acceleration factors of 4 and 8 respectively. For each acceleration factor, the reconstructions are (from left to right) 1. Zero Filled (ZF), 2. Parallel Imaging Compressed Sensing (PICS), and 3. mDD, with $k = 256$ and $k_{\texttt{out}} = 30$ (complex data). Center row is a zoomed-in region. The bottom row shows the subsampling mask and error maps.

acting as a prior, is unable to express images that are "unnatural," i.e. images that have aliasing or very high frequency content.

While this method is appealing because it does not require large training datasets, which are harder to find in computational MRI than in computer vision, it does have several drawbacks. The first is the inference time. In most deep learning methods, training the network on a large dataset takes a long time; but, after training, inference is very fast. In contrast, because the mDD performs inference via training, it needs to re-train for each instance of a particular task, so there is no post-training speedup as in traditional supervised learning methods. Once the network is trained for a particular reconstruction task, it has only solved the reconstruction problem for that instance of data—given another instance,

the network would have to train from scratch to perform the new reconstruction.

Additionally, while the performance does beat traditional methods, it does not do as well as other supervised methods. Winning results from the FastMRI competition, in which reconstruction was performed on the same dataset, outperformed the mDD by using supervised methods. However, those methods also suffered from over-smoothing, and in some cases introduced hallucinations of anatomy not present in the ground truth data [36]. Although the supervised methods quantitatively outperform the mDD, their potential to be biased by the training data and subsequently hallucinate nonexistent anatomy is further motivation to use an untrained method like the mDD.

Another limitation of the mDD is its limited ability to express high frequency information, which is addressed in the next section.

2.6 MDD Extension and Limitations

While the method is fairly successful in denoising, it seems limited in its ability to reconstruct and express high frequency content, as demonstrated by the loss of fine anatomical details in mDD reconstruction. Further research was done to increase the network's ability to express high frequency content.

A recent work by Tanick et al. investigated the use of "Fourier features" to learn high-frequency functions in low-dimensional problem domains [40]. This work showed that, when mapping from coordinates to a functional representation at those coordinates using a multilayer perceptron (MLP), lifting the coordinates to a Fourier feature space before inputting them into the MLP improved the MLP's ability to express high frequency content. This idea is akin to the positional encoding technique that enabled the success of transformer networks in translation [43].

While it seemed like incorporating these Fourier features into the mDD model could help improve the quality of the reconstruction, it was not immediately clear how to do so: all other works using positional encoding techniques had coordinates as input to the model, while the mDD did not. However, Liu et al. introduced the idea of replacing Conv2D layers with a CoordConv layer, which augments the input to the Conv2D layer by concatenating it with coordinates, showing improved results in a variety of different learning tasks [30]. So, we modified the mDD, replacing Conv2D layers with CoordConv layers. But, instead of appending coordinates to the input of every conv, Fourier lifted coordinates were appended. Specifically, for each coordinate p, the Fourier feature lift can be described as

$$\gamma(p) = \left(\sin\left(\frac{2\pi}{L}kp\right), \cos\left(\frac{2\pi}{L}kp\right)\right)$$

where $k = 0, 1, \cdots, L-1$. Additionally, in order to prevent the network from producing

arbitrary frequencies in regions of k-space for which there is no data (the unsampled regions), a multi-scale loss was introduced. This loss was the same metric as the original, but it was evaluated at different scales of output, i.e. the output of every layer of the generative model had a loss associated with it. The network was then optimized over a sum of these losses, weighted to account for the different different scales of the different resolutions. Figure 2.4 shows the updated network architecture, and Figure 2.5 shows the new network reconstruction results.

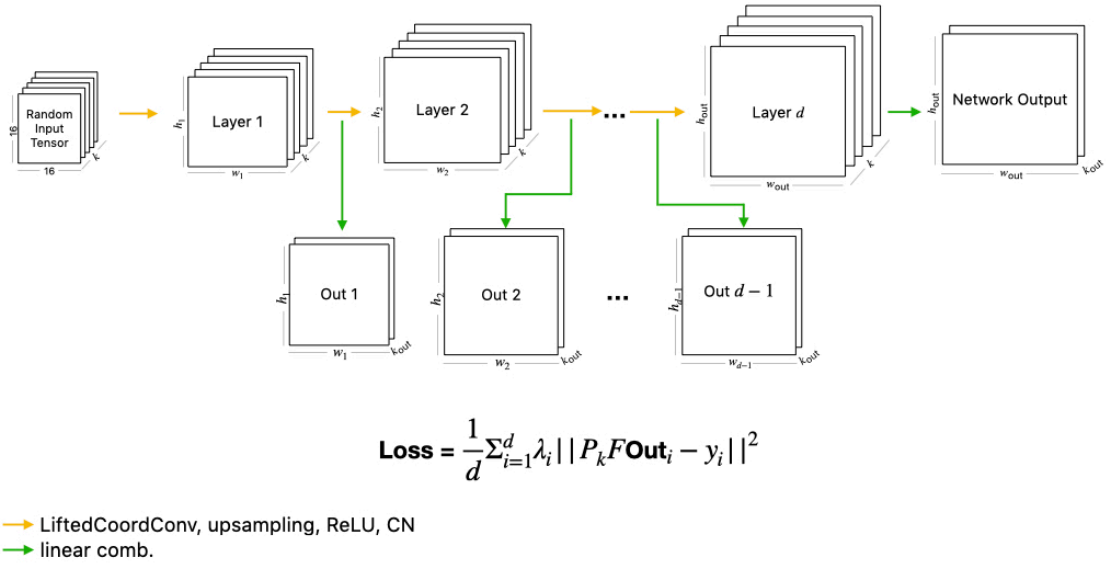

Figure 2.4: The updated high frequency mDD network architecture, with LiftedCoordConv layers instead of Conv2D layers, as well as a multi-scale output and loss.

As we can see, the updated network has an increased ability to express high spatial frequency content. However, increasing the expressivity of the network detracts from the regularization provided by the network structure. And, because in this case the network structure is the only thing regularizing our inverse problem, this method produces a reconstruction that is not "natural." In summary, further modification of the mDD to try to improve its expressivity of high spatial frequencies was successful, but perhaps counter-intuitively results in a less successful reconstruction. Further work could explore adding

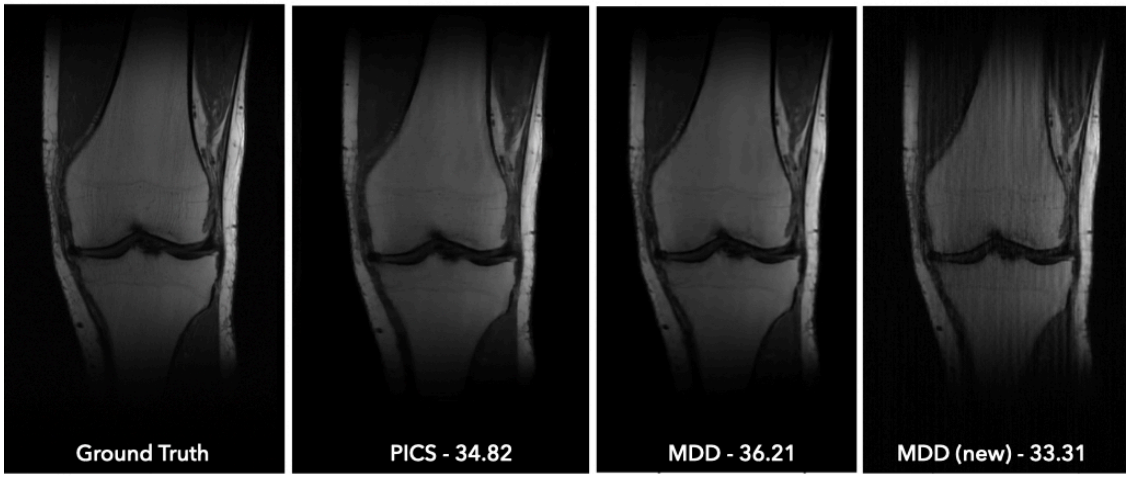

Figure 2.5: 4x Accelerated Reconstruction results with the high frequency mDD. PICS and original mDD reconstruction results and PSNR present for comparison. We see that the network better expresses high frequency textures but performs worse overall in reconstruction.

additional regularization to this new network, but that further complexity might detract from what made this method appealing in the first place.

2.7 Conclusion

The results show that the Modified Deep Decoder architecture allows a concise representation of MR images. The flexibility of this generative image model was successfully leveraged to jointly denoise adjacent slices in 3D MR images and to reconstruct multi-coil data without explicit estimation of coil sensitivities. While the method outperforms traditional methods, it does have some limitations in its ability to express high spatial frequency content when compared to supervised learning methods. This untrained method is particularly attractive in scenarios when access to training data or fully-sampled data is limited, and provides a possible alternative to conventional sparsity-based image priors.

Part II

Generalizing Off-Resonance Blur Correction using Deep Residual Learning

Chapter 3

Off-Resonance De-Blurring

3.1 Abstract

A convolutional residual network was used to correct for blurring due to off-resonance, an artifact common in non-cartesian trajectories. The model was tested in a variety of different experiments in order to evaluate its power and generalizability. Our results indicate that this deep learning method is quite powerful as it shows good results both qualitatively and quantitatively (PSNR, SSIM, NRMSE) across experiments. This method is appealing because robust and generalizable off-resonance correction enables the use of spiral trajectoires, which allows for the rapid imaging necessary for a variety of MRI subfields, such as fMRI, Cardiac Cine MRI, etc.

3.2 Background

In MR imaging, samples from the spatial frequency domain known as k-space are used to reconstruct images. The way in which k-space is sampled is known as a "trajectory," and there is a wealth of research analyzing the benefits and drawbacks of using different types of trajectories [6, 38, 48]. While the most commonly used trajectory is a cartesian sampling pattern, where k-space is sampled row by row, an alternative and commonly researched trajectory is the spiral trajectory, in which k-space is sampled in spirals that start in the center of k-space and spiral outward.

The spiral trajectory is particularly attractive for a few reasons. The first is that each spiral starts by sampling the center of k-space, which is the region of k-space that has the highest signal to noise ratio. Additionally, the acquisition time is shorter on average than a comparable cartesian trajectory [4], and the spiral trajectory is less sensitive to motion artifacts. [45]. This makes spiral trajectories favorable for scans that rely on time series data, such as fMRI or cardiac scanning [16].

However, the MRI signal model relies on the assumption that the resonance of nuclear spins in the magnet is uniform, i.e. that all spins resonate at a frequency ω_0 proportional to the main magnetic field strength [18]. However, in reality, not all spins in the body resonate exactly at ω_0, either due to inhomogeneities in the main magnetic field or due to excitation of H+ in other complex molecules, such as fat [37]. Spins actually resonate at a range of frequencies around ω_0, which is known as off-resonance. Off-resonance causes spatial blurring in the reconstructed image, especially when k-space is sampled with non-Cartesian trajectories, such as spiral trajectories [33] (Figure 3.1).

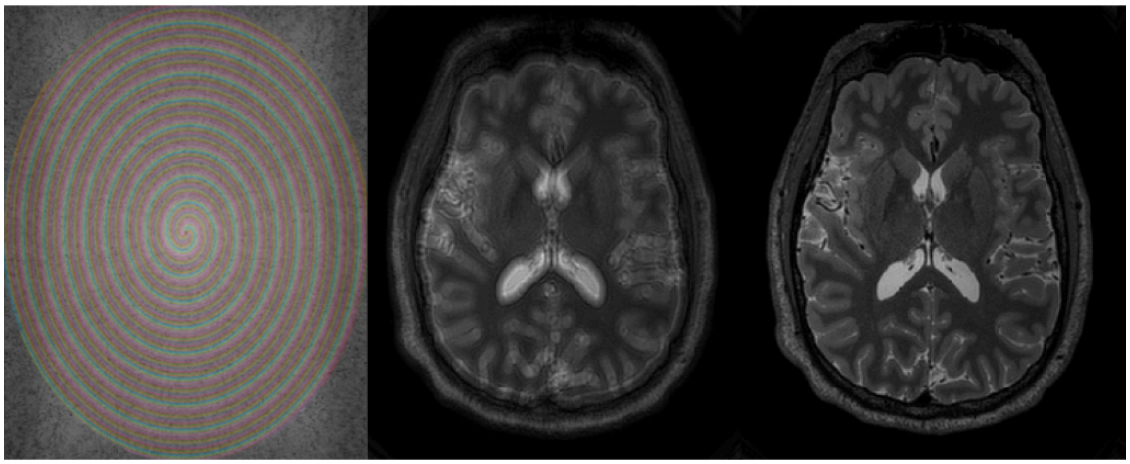

Figure 3.1: Left: A spiral trajectory overlayed on k-space data — Middle: Reconstruction of k-space data with off-resonance blur — Right: Reconstruction of k-space data without off-resonance blur

Several traditional, analytical methods have been proposed for off-resonance correction in the past [37, 33, 34]; however, these methods are not very accurate and can be computationally slow. Recently, Zeng et al. proposed a deep learning based approach to correct for off-resonance blur in 3D Cone trajectories that achieved great success. This work aims to reproduce Zeng's results using 2D spiral trajectories (rather than the 3D cone trajectories used in Zeng et al.) and investigate the generalizability of this deep learning model. This is done by evaluating network performance agnostic to anatomy and image statistics, as well as between different datasets.

3.3 Theory

The typical MR signal model is $s(t) = \int_r M(r)e^{-j2\pi r \cdot k_r(t)}dr$, but the off-resonance signal equation (ignoring relaxation) can be expressed as

$$s(t) = \int_r M(r)e^{-j2\pi r \cdot k_r(t)}e^{-j\omega_r t}dr$$

The difference between these two equations is the introduction of the $e^{-j\omega_r t}$ term in the second signal model, known as the off-resonance term. The off-resonance term multiplies the Fourier imaging term, and therefore has a convolutional effect in the image domain. The degree of off-resonance artifact present in the image depends on the off-resonance frequency ω_r, as well as the length of the readout duration t. This artifact manifests as a phase error in k-space, or, equivalently, blurring in the image domain. Additionally, off-resonance varies spatially, which means that the off-resonance convolutional kernel is different depending on its location in the image. We can therefore model the effect of off-resonance as a nonstationary convolution.

When framed as such, we can think of correcting off-resonance artifacts as solving for a nonstationary deconvolutional (deblurring) operator. This problem naturally lends itself to a deep learning solution. Firstly, the learned weights of the convolutional neural network can be thought of as learned deconvolutional weights of the forward off-resonance blurring model. Secondly, the nonlinearities present in the network can address the nonstationarity of the blurring by adjusting the deconvolutional weights depending on the spatial location. Additionally, because off-resonance can be thought of as a convolution, we know that the original information is present but locally distributed, which again lends itself well to the convolutional kernels typical in CNNs. Finally, because the blurring kernel is only a function of the off-resonance and not the image amplitude, each pixel in the training set can be considered a training example in a fully convolutional neural network [47].

3.4 Methods

CNN Architecture

The CNN architecture used in this work is 2D adaptation of the network used in Zeng et al. The network is a convolutional residual network with 3 residual layers, each consisting of two 5x5 Conv2D layers and ReLU activations. Each of the residual layers has a filter depth of 128, and both the input and outputs of the network are images with 2 channels, corresponding to real and imaginary components. The network is designed to correct for off-resonance, and is trained by mapping blurred images to clean ones (Figure 3.2). The network is designed to be smaller with less capacity in order to promote learning off-resonance correction and prevent overfitting through memorization of anatomy. Additionally, we know

that off-resonance blurring kernels are low rank [34, 37], so a relatively small network should be able to effectively learn the deblurring. The network was trained using the L_1 loss metric because it has been shown to produce better objective and perceptual results [49]. The network was built and trained in PyTorch, optimized with Adam over 8 epochs with a learning rate of 10^{-4} and a batch size of 15 [25].

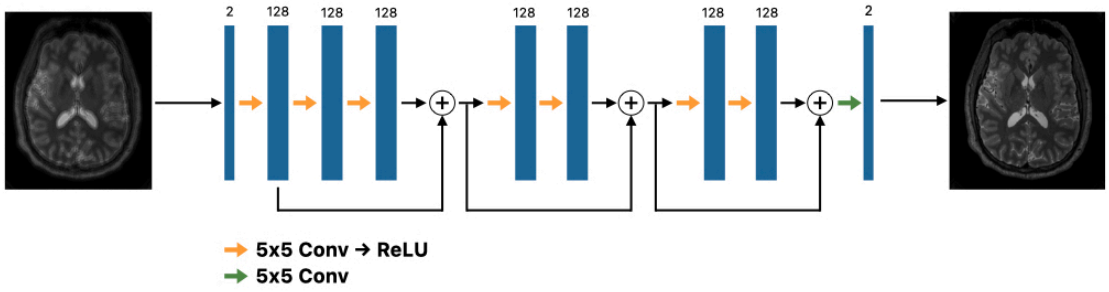

Figure 3.2: Visualization of the network architecture that maps off-resonance blurred images to clean images

Experiments

2D Reproduction

Several experiments were done in order to reproduce and test the generalizability of this method. The first goal was to replicate the work in Zeng et al. in a 2D spiral trajectory setting. The data for this experiment came from the Human Connectome Project [13]. The data consists of 3D T2-SPACE scans of 32 patients with a TR of 3200 ms, a TE of 561 ms, a voxel size of 0.7 mm isotropic, and each subject volume of dimensions $320 \times 320 \times 256$. The central 50 axial slices for each subject were then used to generate the training and test datasets. Each 320×256 slice was a magnitude image.

In order to simulate a spiral trajectory and off-resonance artifacts, each image underwent a density compensated NUFFT [14] with a spiral k-space trajectory; then, off-resonance was simulated by by multiplying the k-space with a complex exponential at a constant frequency (as described in the Theory section); and finally, an inverse density compensated NUFFT was taken to get the resulting image. For each slice, 20 artifact images were created by following the procedure described above for off-resonance frequencies ω_r ranging from -100 Hz to 100 Hz in steps of 10 Hz. The image corresponding to $\omega_r = 0$ was then considered to be the ground truth, because it was simulated with a spiral trajectory but with no introduced off-resonance blur. The goal of the network was then to learn a many-to-one mapping between the 20 artifact images and the one ground truth, for each slice and subject.

The data (32,000 artifact images and 1600 ground truth images) were split in a 30%-70% train-test split across subjects, resulting in 9000 training images and 23,000 test images. Although this split is unconventional, we and Zeng et al. were able to achieve good results with a small training set for several reasons.

First, the network was deliberately chosen to be shallow to reduce the number of training parameters. Second, the network was fully convolutional, such that it could be interpreted that each pixel was a training example for this regression problem. Additionally, each pixel was augmented with 20 simulated off-resonant frequencies. Third, and perhaps most importantly, we have posed off-resonance correction as a nonstationary deconvolution problem; under this problem structure, the network can be considered to learn generalizable kernels as a function of not only the data model, but also the physical model. This paradigm is similar to the one presented in previous works such as AUTOMAP [51], and provides a powerful theoretical framework to operate within. [47].

While the preparation of data described here indicates that the synthetic off-resonance artifacts introduced are not spatially varying, this technique is similar to that of Zeng et al., which demonstrated that this data model is sufficient to learn the deconvolutional deblurring operation.

Noise Data

The next experiment seeks to investigate the generalizability of the method; in order to learn a generalized kernel which is only a function of the physical model (and not the data model), theoretically, the network should not need to see anatomy in order to correct for off-resonance blur.

We do this by replicating the setup of the first experiment, but for pure noise images instead of anatomy; the data was generated in the same way as in the first experiment, but instead of adding an off-resonance artifact to an axial slice of a brain volume, we introduce the off-resonance artifact to a known, fixed noise image. The (`train`, `ground_truth`) pairs thus become (`noise`+`off-res`, `noise`), and the goal of the network—to correct for the introduced off-resonance artifact—is the same. We want to see if the network can still produce a general deblurring operation and correct for off-resonance in anatomy images when trained on images that have the off-resonance artifact, but lack any image anatomy or statistics to learn from.

Additionally, it has been shown that deep convolutional networks can generalize poorly to small image transformations, and that typical data augmentations such as crops and rotations are insufficient in compensating for these transformations [3]. By learning off-resonance correction from white noise images, we hope that the network will not be biased by any specific training dataset, making the network more robust and generalizable under a variety of different dataset sources and conditions.

Inter-Dataset Generalization

Next, we sought to evaluate the model's performance when correcting more realistic, spatially-varying off-resonance blur after training on synthetic spatially-uniform off-resonance artifact images. However, this requires the dataset to include fieldmaps, which describe the off-resonance frequency at each voxel. For this purpose, we used the NIFD dataset [11], which is a series of 2D gradient echo scans for each subject, with a TR of 667 ms, a TE of 7.65 ms, a voxel size of 1.0 mm isotropic, and each subject slice of dimensions 106x106. There were 11 subjects in the dataset, each with 60 slices acquired. Unlike the HCP dataset, the NIFD also included fieldmap data for each slice. The training data was generated as in the first experiment, but with a lower resolution k-space spiral trajectory to account for the difference in resolution.

However, the NIFD dataset did not have a sufficient amount of data to effectively train the network. So, we trained on a modified HCP dataset (HCP data with the new k-space trajectory) and tested on the NIFD dataset to see if it was able to correct for the off-resonance artifact despite the scans being from different data sources.

Fieldmap Data

Once the network was trained on the modified HCP dataset and tested on the NIFD dataset, we could evaluate its performance on more realistic, spatially-varying off-resonance artifact images. Instead of simulating off-resonance at a constant frequency as was done in the previous experiments, we use the fieldmap given in the NIFD dataset to introduce spatially-varying off-resonance. This is performed similarly to the artifact simulation discussed in section 3.4-Experiments-2D Reproduction. We again take a density-compensated NUFFT using a spiral k-space trajectory; but, instead of multiplying the entirety of k-space with a complex exponential with a constant ω_r, we use the fieldmap, which dictates how certain positions are going to accrue phase, to set ω_r. We use that fieldmap data in an Inverse Nonuniform Discrete Fourier Transform to produce an image with spatially varying off-resonance artifacts.

Figure 3.3 demonstrates a practical issue in using fieldmap data: because the fieldmap data is noisy, using the fieldmap without any processing results in additional noise artifacts that the network has not seen before. In order to use the fieldmap without introducing these noisy artifacts, we first denoise the fieldmap. We use a Total Variation denoiser [7], and then use the denoised fieldmap to introduce the spatially-varying off-resonance. We then take the network from the inter-dataset experiment, trained on synthetic-artifact HCP data, and test it on fieldmap-artifact NIFD data.

This experiment evaluates the network's performance in correcting fieldmap artifacts for two training cases simultaneously: (1) training on synthetic-artifact data and (2) training on data from a different source.

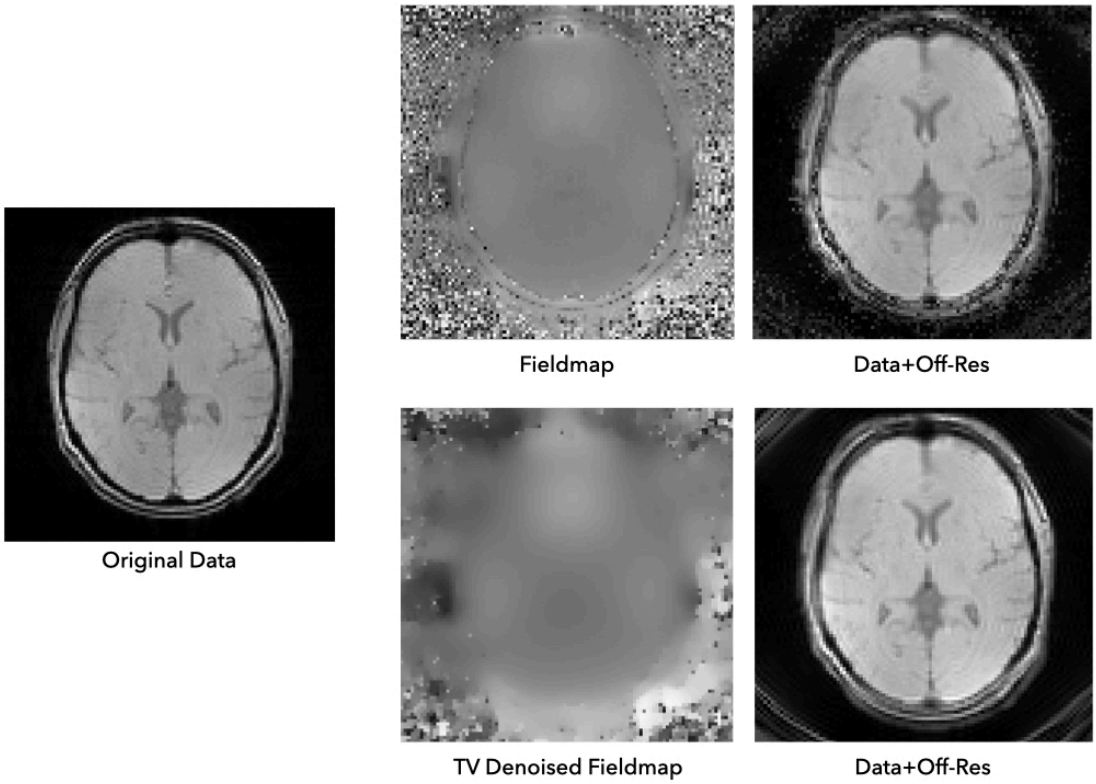

Figure 3.3: Visualization of the fieldmap and denoised fieldmap and their corresponding artifact images. The top row shows how the unprocessed fieldmap results in an artifact image with both blur and noise. The bottom row shows how, after denoising the fieldmap, the artifact image contains a spatially varying blur and not spurious noise artifacts.

In Vivo Data

The method was then evaluated on In Vivo data acquired on a 3T GE scanner, with a TE of 1 ms, a TR of 1 s, and a slice thickness of 5 mm, using the same low-res k-space trajectory that the network trained on in the previous experiment. Ground truth data was not available for comparison, so an alternative no-reference image quality metric (IQA) [26] was used to compare the performance of this method to the baseline, autofocus method.

3.5 Results

All of these experiments were evaluated using three different metrics: Peak Signal to Noise Ratio (PSNR), Structural Similarity (SSIM) [44], and Normalized Mean Squared Error

(NRMSE). All the summary figures include the performance of the autofocus method [37] on the same dataset, which is a traditional method of correcting for off-resonance blur and serves as a baseline.

2D Reproduction

Figure 3.4 and 3.5 show deblurring examples from the 2D reproduction experiment on the HCP dataset. Figure 3.6 shows the summary statistics (mean and standard deviation) for the entire test dataset.

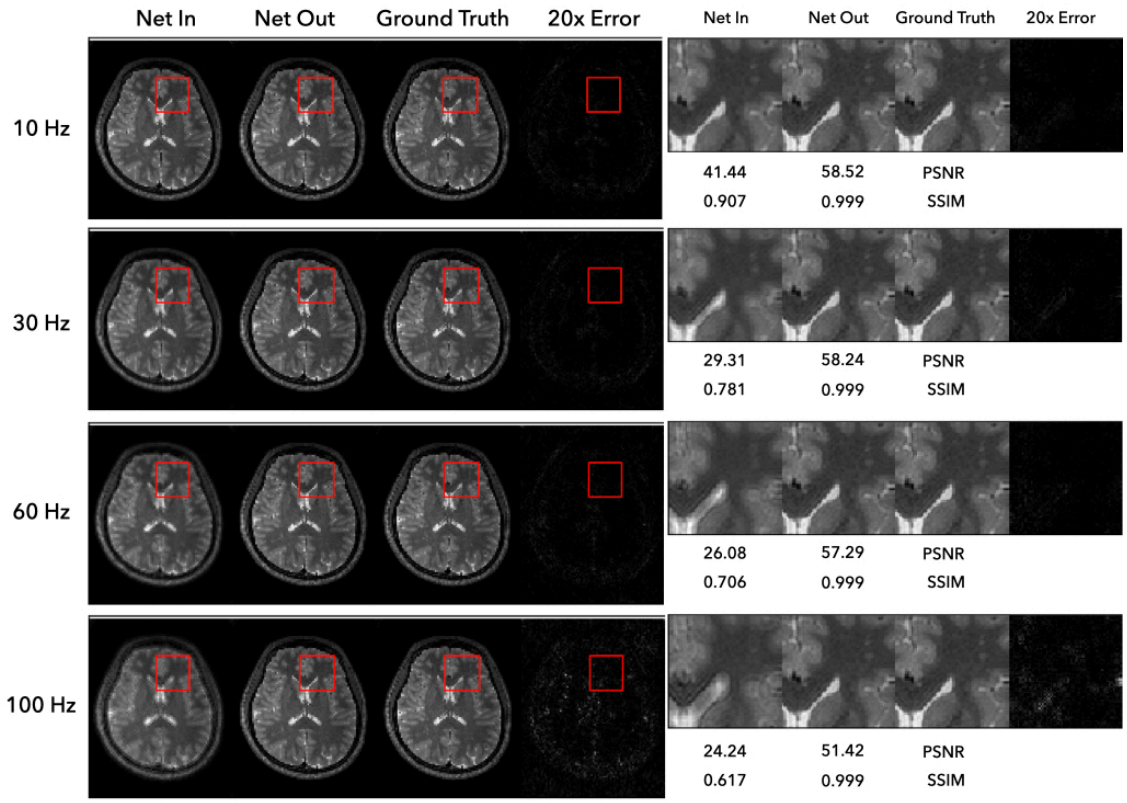

Figure 3.4: Example results of the 2D reproduction experiment. The network corrects for varying degrees of simulated off-resonance (10 Hz, 30 Hz, 60 Hz, 100 Hz). The region in the red box on the left is magnified on the right to show detail. Each artifact image and its corrected image have PSNR and SSIM metrics shown below the magnified region on the right.

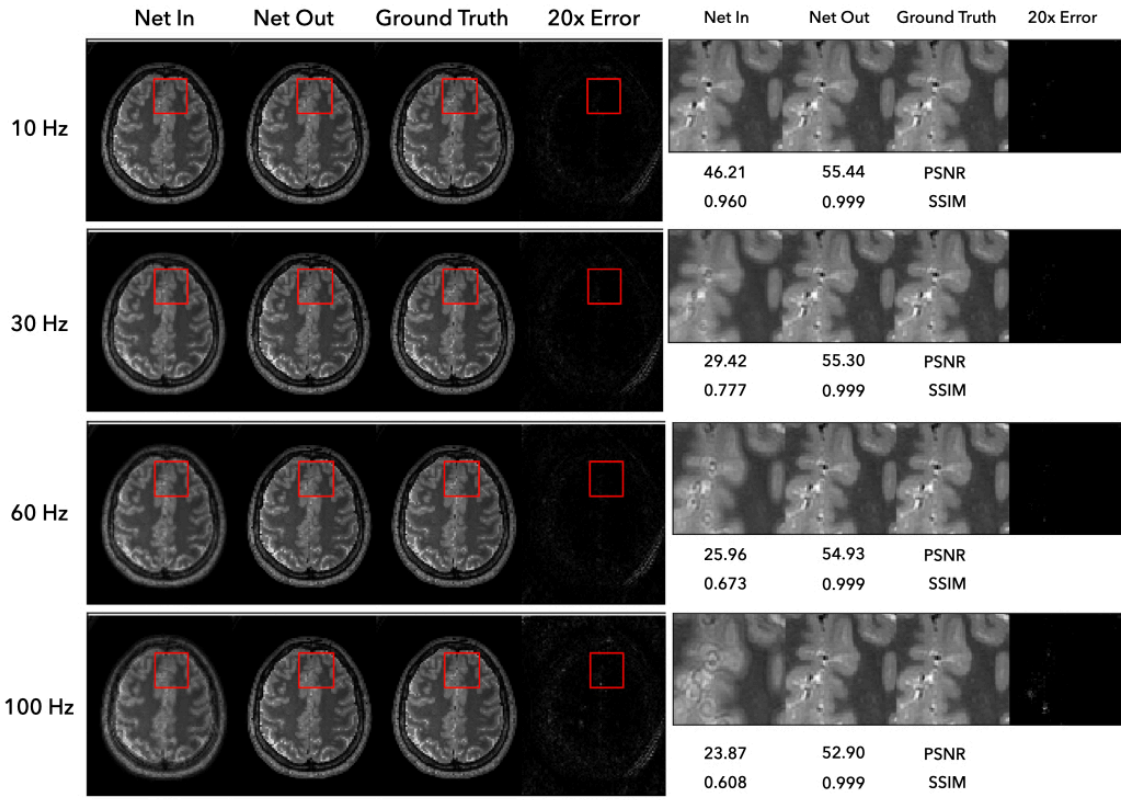

Figure 3.5: Example results of the 2D reproduction experiment. The network corrects for varying degrees of simulated off-resonance (10 Hz, 30 Hz, 60 Hz, 100 Hz). The region in the red box on the left is magnified on the right to show detail. Each artifact image and its corrected image have PSNR and SSIM metrics shown below the magnified region on the right.

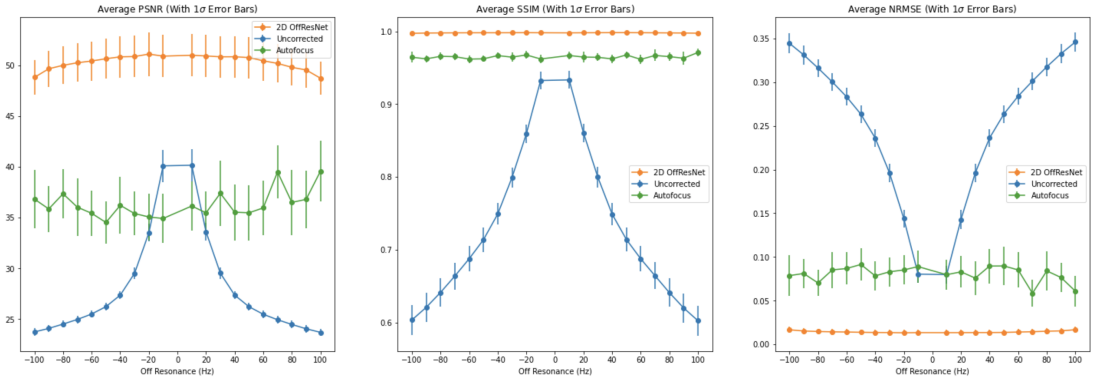

Figure 3.6: Summary statistics of the average PSNR, SSIM, and NRMSE of the artifact correction performance of the 2D reproduction experiment on the test set. The bars show the 1σ standard deviation in performance.

In the summary statistics, we see that across the metrics and the different degrees of off-resonance, the network does a great job of correcting for off resonance blur, beating the autofocus method in terms of SSIM and NRMSE by a slight amount and by a large amount in the PSNR metric. The example images show high fidelity to the anatomy present in the image, and the only errors seen are on the boundary of the anatomy and don't have any structure present. This is fairly impressive considering these results are all on the test dataset, i.e. the network has not seen these images before, only images from other subjects in this study.

Noise Data

Figure 3.7 and 3.8 show deblurring examples of the HCP test dataset when trained on the noise dataset, and Figure 3.9 shows its summary statistics.

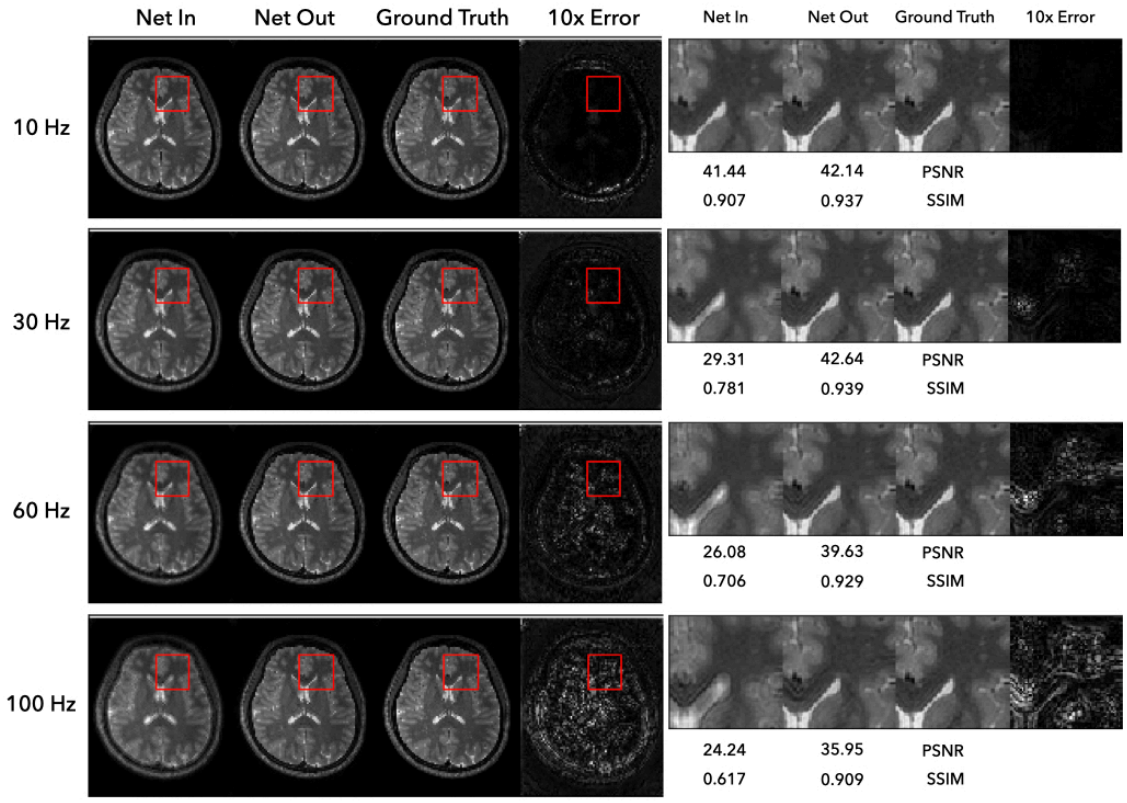

Figure 3.7: Example results of the noise data experiment. The network corrects for varying degrees of simulated off-resonance (10 Hz, 30 Hz, 60 Hz, 100 Hz). The region in the red box on the left is magnified on the right to show detail. Each artifact image and its corrected image have PSNR and SSIM metrics shown below the magnified region on the right.

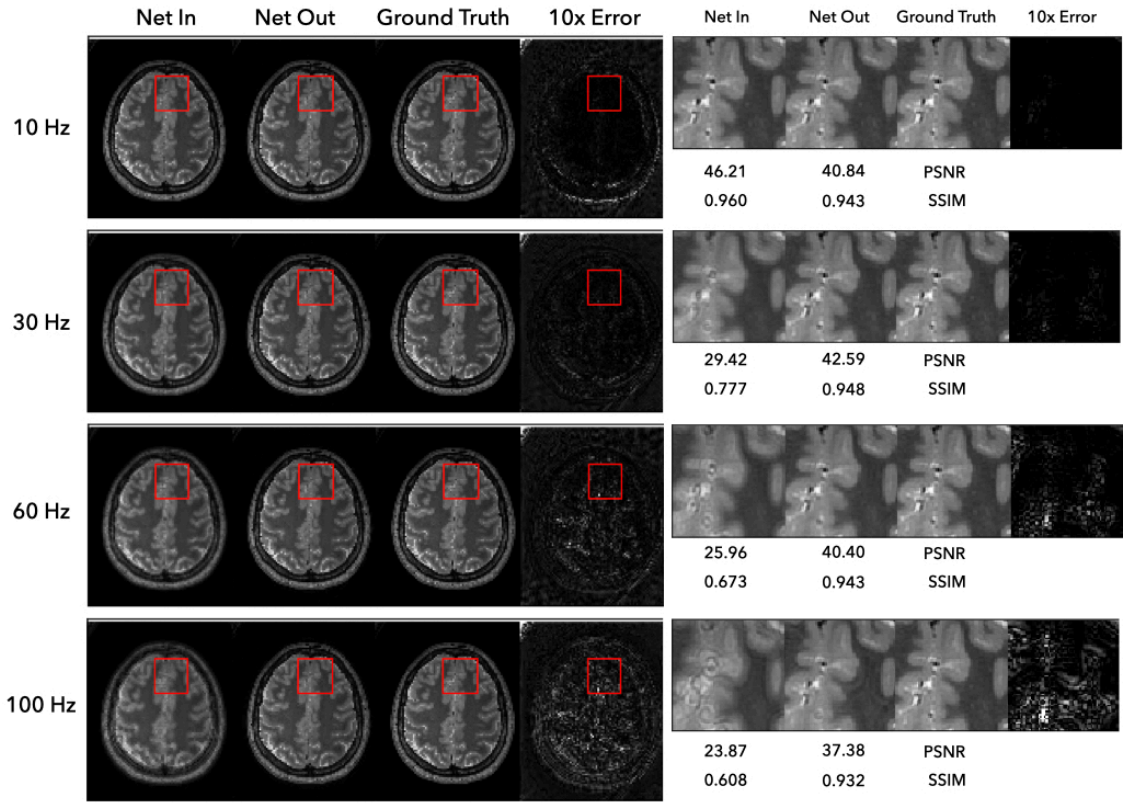

Figure 3.8: Example results of the noise data experiment. The network corrects for varying degrees of simulated off-resonance (10 Hz, 30 Hz, 60 Hz, 100 Hz). The region in the red box on the left is magnified on the right to show detail. Each artifact image and its corrected image have PSNR and SSIM metrics shown below the magnified region on the right.

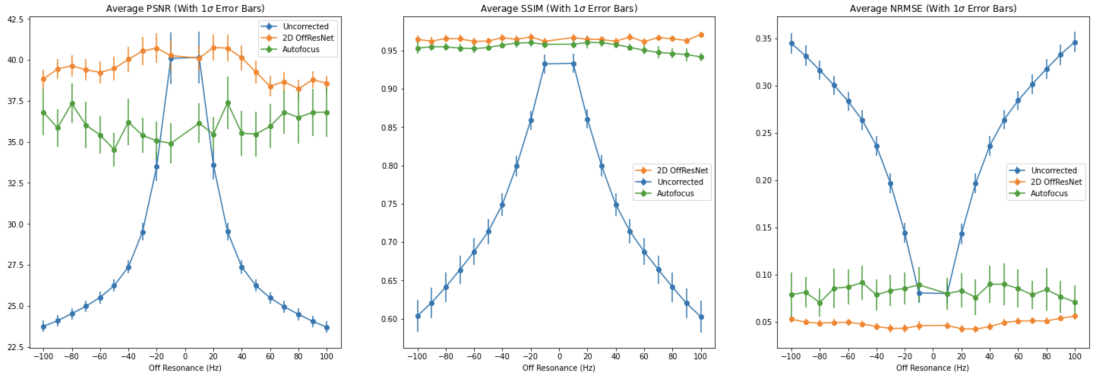

Figure 3.9: Summary statistics of the average PSNR, SSIM, and NRMSE of the artifact correction performance when trained on noise data and tested on the HCP dataset. The bars show the 1σ standard deviation in performance.

Here we see that the model still performs fairly well, albeit not as well as when it is both trained and tested on the HCP dataset. However, despite the fact that the network did not learn any image anatomy or statistics, it is still able to correct for off-resonance in anatomy data fairly well. Again, we that it outperforms autofocus across our metrics, although now with a reduced margin of improvement.

For low levels of off-resonance, the network does not correct well when applied to the HCP dataset. This makes sense, as the network has not been exposed to anatomy before. So, when it sees an image with little off-resonance artifact, the most we can hope for is for it to not corrupt the image, which seems to be the case as the PSNR values essentially overlap for the -10 and 10 Hz cases.

In the image examples, we see that the error is much higher, and there is no some structure to the error when compared to the 2D reproduction, but the network's ability to deblur for high levels of off-resonance without ever learning image anatomy or statistics is impressive.

Inter-Dataset Generalization

Figure 3.10 shows the summary statistics of training on the low-resolution trajectory of the HCP data and testing on the HCP and NIFD data respectively. Figures 3.11 and 3.12 show examples of deblurring the NIFD data when trained on HCP data.

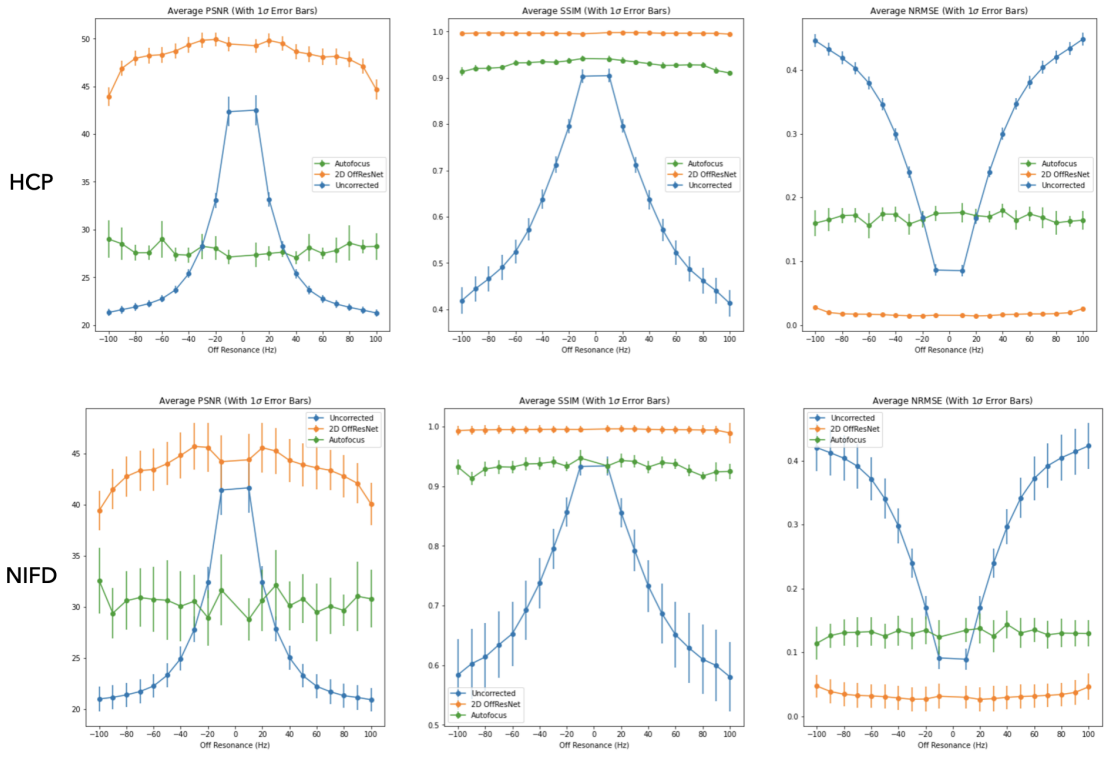

Figure 3.10: Summary statistics of the average PSNR, SSIM, and NRMSE of the artifact correction performance when trained on low resolution HCP data and tested on the low-res test HCP dataset (top) and NIFD dataset (bottom). The bars show the 1σ standard deviation in performance.

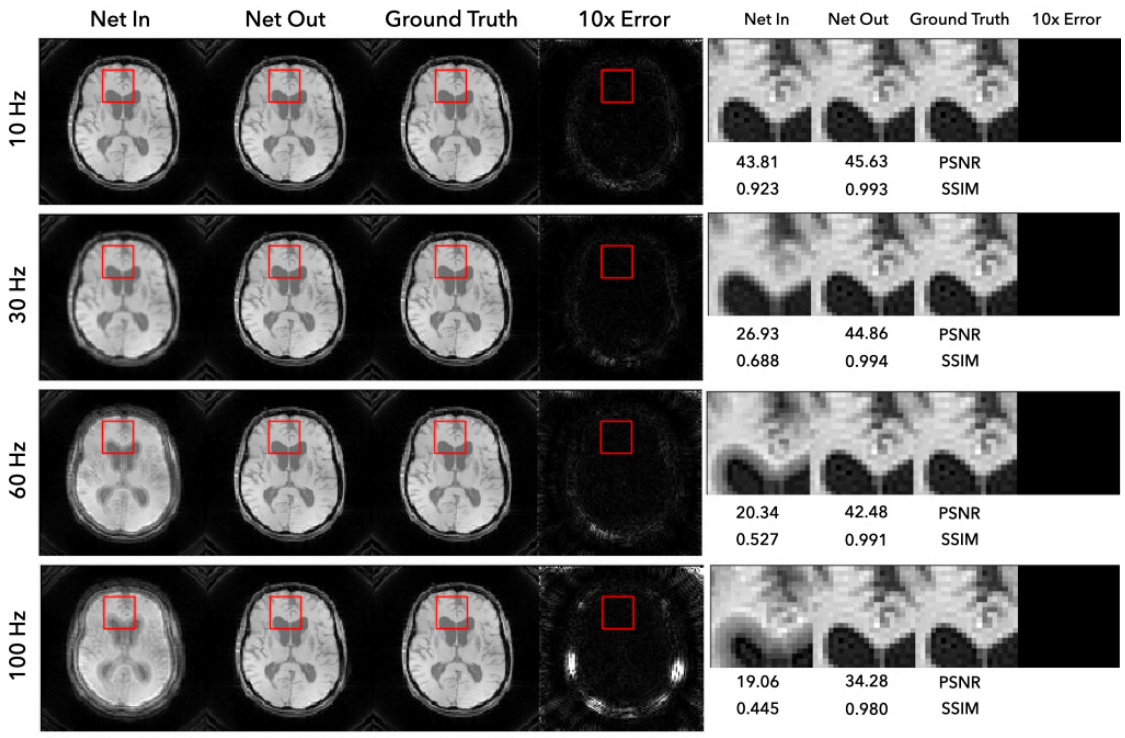

Figure 3.11: Example results of the inter-dataset generalization experiment. The network, trained on low-res HCP data, corrects for varying degrees of simulated off-resonance (10 Hz, 30 Hz, 60 Hz, 100 Hz shown) in the NIFD dataset. The region in the red box on the left is magnified on the right to show detail. Each artifact image and its corrected image have PSNR and SSIM metrics shown below the magnified region on the right.

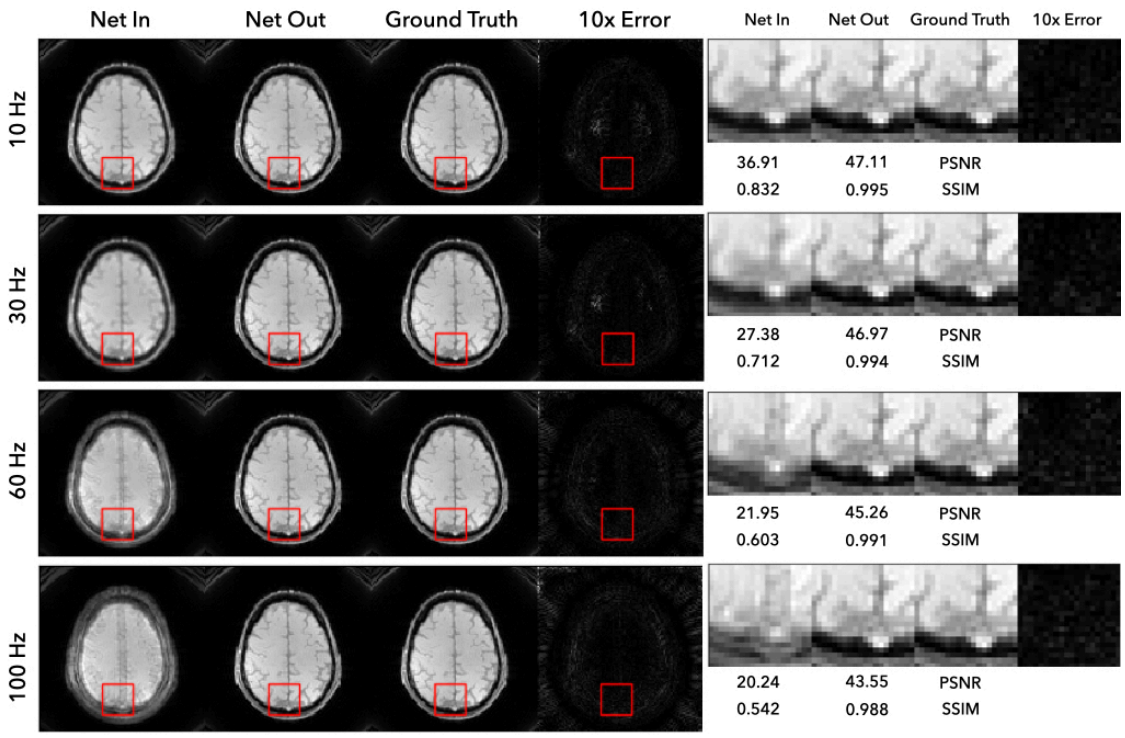

Figure 3.12: Example results of the inter-dataset generalization experiment. The network, trained on low-res HCP data, corrects for varying degrees of simulated off-resonance (10 Hz, 30 Hz, 60 Hz, 100 Hz shown) in the NIFD dataset. The region in the red box on the left is magnified on the right to show detail. Each artifact image and its corrected image have PSNR and SSIM metrics shown below the magnified region on the right.

We again see that the method performs better when trained and tested on the same dataset than when trained on one and tested on the other. However, it still performs quite well in both cases, especially for high levels of off-resonance. And, in both cases, outperforms the autofocus method.

In the deblurring examples, we see results similar to the 2D reproduction experiment in that the error is low overall but highest at the image boundary. Additionally, the error appears unstructured.

Fieldmap Data

Figure 3.13 shows the results of training on the HCP low-resolution data and testing on the NIFD fieldmap-artifact (spatially varying) data. Figure 3.14 shows the summary statistics for the fieldmap artifact correction.

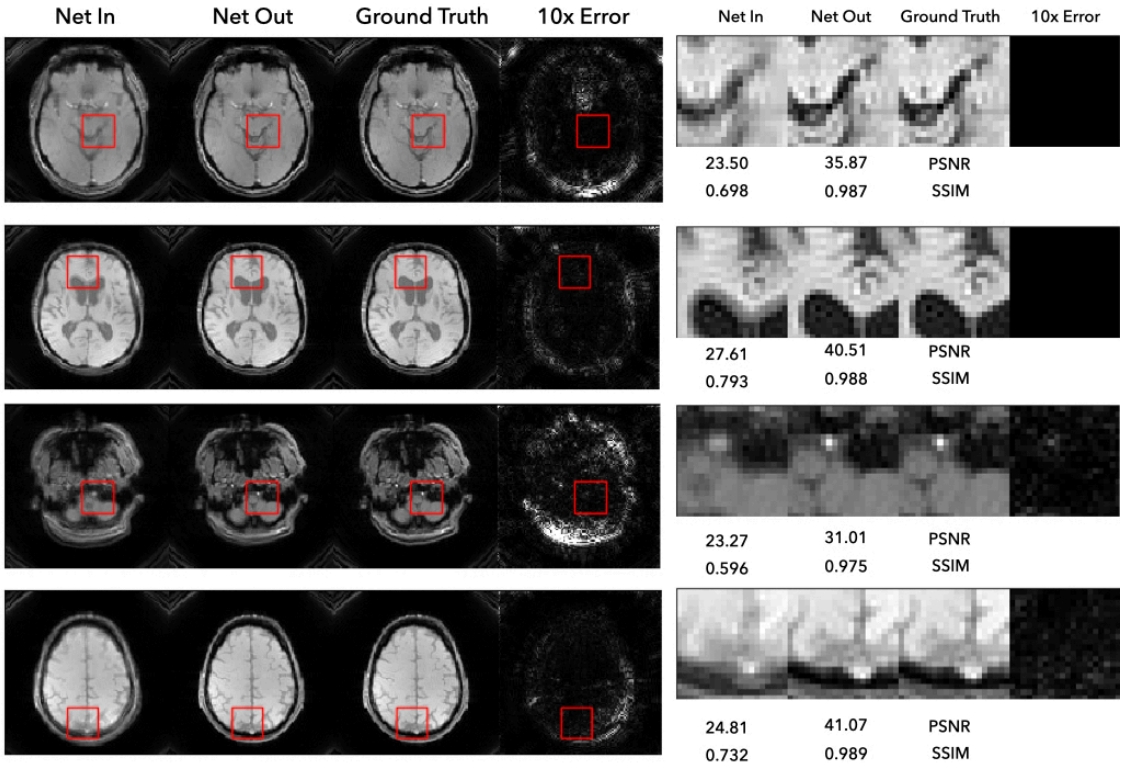

Figure 3.13: Example results of the fieldmap data experiment. The network, trained on low-res HCP data, corrects for spatially varying off-resonance in the NIFD dataset. The region in the red box on the left is magnified on the right to show detail. Each artifact image and its corrected image have PSNR and SSIM metrics shown below the magnified region on the right.

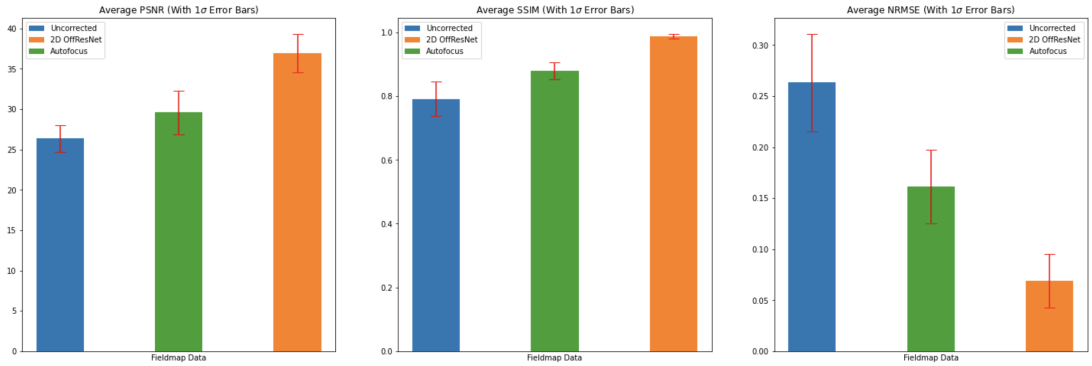

Figure 3.14: Summary statistics of the average PSNR, SSIM, and NRMSE of the artifact correction performance when trained on low-res HCP data and tested on spatially-varying fieldmap off-resonance NIFD data. The bars show the 1σ standard deviation in performance.

We see that again, across all our metrics, the deep learning model corrects the off-resonance artifacts very well, outperforming autofocus. Even in this case, where we train on a different dataset and test on a more realistic off-resonance artifact, the network corrects the images across all metrics. In the examples, we see that the error is high mainly on the boundary of the anatomy. An interesting note is that the third image in Figure 3.13 is of anatomy that the network is unfamiliar with, as only the center slices of the HCP dataset were used to train; this might explain why it performs the worst of the examples shown.

In Vivo Data

Figure 3.15 shows the results of applying the same network in the previous experiment, that was trained on the HCP low-resolution data, to in vivo data.

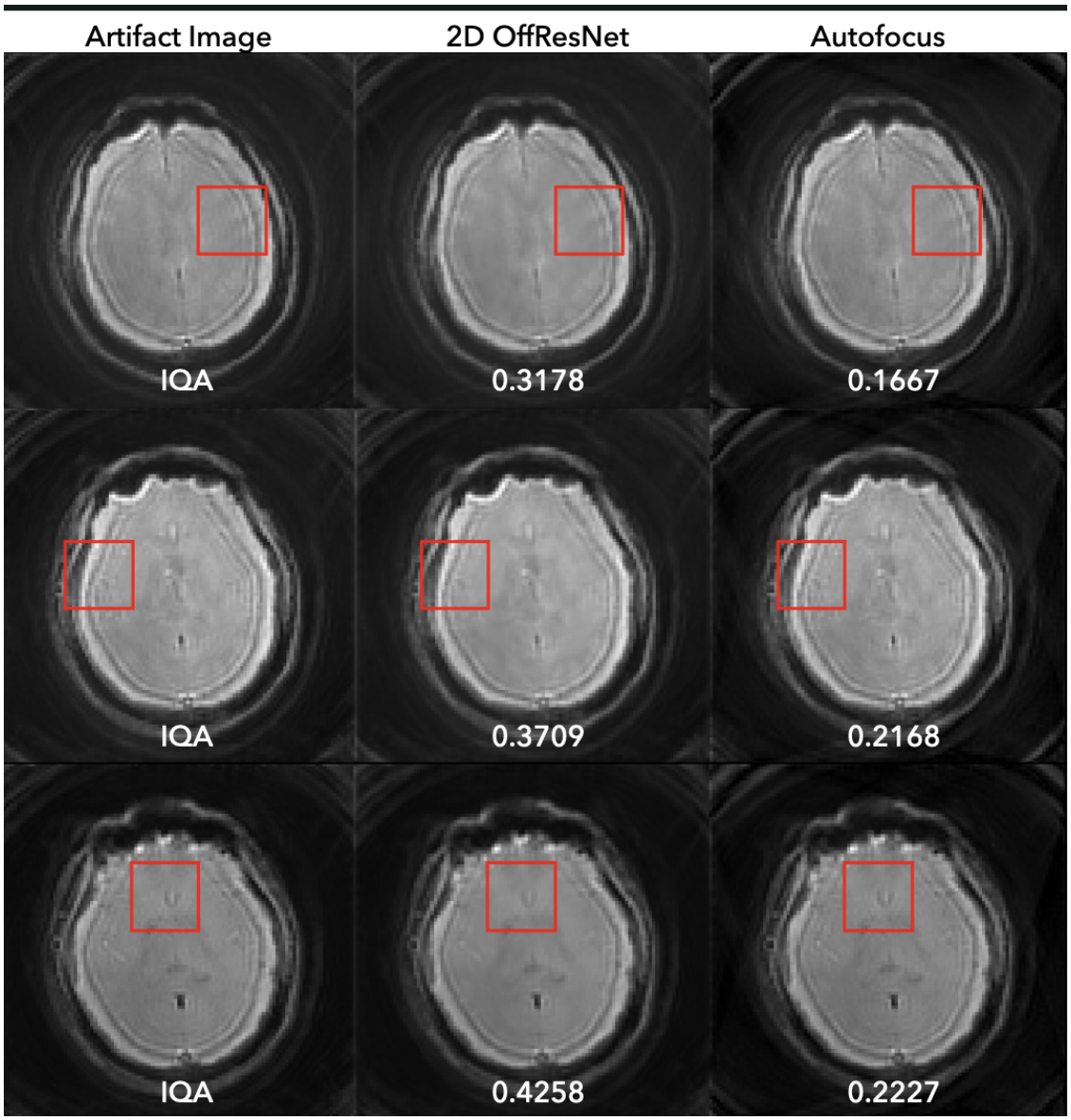

Figure 3.15: Three brain slices of In-Vivo data with off-resonance artifacts shown, alongside their corrected versions using the 2D Off-ResNet and Autofocus methods respectively. Image Quality Assessment (IQA) metrics, normalized between 0 and 1, are also shown.

We see that, overall the network does correct some of the off-resonance and ringing in the images, especially in the regions highlighted with the red boxes. We also see that, according the Image Quality Assessment (IQA) metric scores that have been normalized to be between 0 and 1, that the network outperforms the Autofocus method.

However, while the network outperforms the autofocus method, the off-resonance correction on the in vivo data is not as good overall as the correction on the simulated data in previous experiments. We hypothesize that this is the case for the following reason. The network that trains and works well with high levels of spatially-uniform off-resonance also works well with low levels of spatially-varying off-resonance (fieldmaps). However, the in vivo data seems to have a high level of spatially-varying off-resonance, which the network may not be as well equipped to handle. This issue, however, is probably more indicative of the training data than of the model itself. If the network is trained on spatially-varying off-resonance, then the method would likely work better with in vivo data. However, generating a sufficient amount of spatially-varying off-resonance training data is a time intensive process, and is why spatially-uniform off-resonance was used in the experiments in this thesis.

3.6 Discussion

Having shown that the method performs well in a variety of cases, we can say with some certainty that framing this problem as solving for a nonstationary deconvolutional operator (and using a convolutional residual network to do so) is valid, which is highly convenient since this problem structure lends itself naturally to deep learning solutions. Further deep learning solutions experimenting with different network architectures and loss functions can thus be explored.

Because we are working with lower-resolution 2D data and 2D spiral trajectories rather than 3D cones, this method demands significantly less training time that of Zeng et al: these results all took about 3-4 hours to train, whereas Zeng et al. took 32 days to train.

In certain cases, the method does not work as well for very low levels of off-resonance. This is particularly true for the noise data experiment, and likely occurs because the network has no information about the image statistics or expected anatomy. An additional note is that, for low levels of simulated off-resonance, the autofocus method performs worse than if the image is left uncorrected. However, we see in Figure 3.14 that autofocus always improves the quality of the image when it has spatially-varying off-resonance. This is probably due to the fact that autofocus was designed to correct for spatially-varying off-resonance. So, it can not perform as well on the synthetic off-resonance datasets as the network that trains on that kind of data.

It would be interesting to do further research to see how well the method performs when trained on spatially-varying off-resonance, or on multiple different data sources, and whether

that could help improve the generalizability and robustness of the method. Training on multiple different sources simultaneously could also improve the network's correction when little artifact is present in the image. It could also be interesting to see how the method performs when trained on a traditional computer vision dataset like ImageNet with added off-resonance: would it still do a good job of correcting for off-resonance when then tested using MRI data?

3.7 Conclusion

This work aimed to investigate the generalizability of using a convolutional residual network to correct for blurring due to off-resonance. This was done with a 2D spiral trajectory through several experiments using a variety of data sources. These experiments have shown that the residual model is able to correct off-resonance when trained on noise data with no typical image statistics or anatomy, when trained and tested on different datasets from different sources, and when trained on synthetic, spatially uniform off-resonance and tested on more realistic, spatially varying off-resonance. Establishing a robust, general, and efficient method of correcting for off-resonance from 2D spiral trajectories is important, as it would enable the use of these trajectories in rapid imaging scenarios, such as fMRI and Cardiac Cine MRI, without the concern of blurring artifacts.

Bibliography

[1] Hemant K. Aggarwal, Merry P. Mani, and Mathews Jacob. "MoDL: Model-Based Deep Learning Architecture for Inverse Problems". In: *IEEE Transactions on Medical Imaging* 38.2 (Feb. 2019), pp. 394–405. ISSN: 1558-254X. DOI: 10.1109/tmi.2018.2865356. URL: http://dx.doi.org/10.1109/TMI.2018.2865356.

[2] Hanafy M. Ali. "MRI Medical Image Denoising by Fundamental Filters". In: *High-Resolution Neuroimaging*. Ed. by Ahmet Mesrur Halefoğlu. Rijeka: IntechOpen, 2018. Chap. 7. DOI: 10.5772/intechopen.72427. URL: https://doi.org/10.5772/intechopen.72427.

[3] Aharon Azulay and Yair Weiss. "Why do deep convolutional networks generalize so poorly to small image transformations?" In: *CoRR* abs/1805.12177 (2018). arXiv: 1805.12177. URL: http://arxiv.org/abs/1805.12177.

[4] Matt A Bernstein, Kevin F King, and Xiaohong Joe Zhou. *Handbook of MRI pulse sequences*. Elsevier, 2004.

[5] Harold C. Burger, Christian J. Schuler, and Stefan Harmeling. "Image denoising: Can plain neural networks compete with BM3D?" In: *2012 IEEE Conference on Computer Vision and Pattern Recognition*. 2012, pp. 2392–2399. DOI: 10.1109/CVPR.2012.6247952.

[6] Stephen F. Cauley et al. "Autocalibrated wave-CAIPI reconstruction; Joint optimization of k-space trajectory and parallel imaging reconstruction". In: *Magnetic Resonance in Medicine* 78.3 (2017), pp. 1093–1099. DOI: https://doi.org/10.1002/mrm.26499. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.26499. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.26499.

[7] Antonin Chambolle. "An Algorithm for Total Variation Minimization and Applications". In: *Journal of Mathematical Imaging and Vision* 20.1 (Jan. 2004), pp. 89–97. ISSN: 1573-7683. DOI: 10.1023/B:JMIV.0000011325.36760.1e. URL: https://doi.org/10.1023/B:JMIV.0000011325.36760.1e.

[8] W. Chen, C. T. Sica, and C. H. Meyer. "Fast conjugate phase image reconstruction based on a Chebyshev approximation to correct for B0 field inhomogeneity and concomitant gradients". In: *Magn Reson Med* 60.5 (Nov. 2008), pp. 1104–1111.

[9] Chris A. Cocosco et al. "BrainWeb: Online Interface to a 3D MRI Simulated Brain Database". In: *NeuroImage* 5 (1997), p. 425.

[10] Elizabeth K. Cole et al. *Unsupervised MRI Reconstruction with Generative Adversarial Networks*. 2020. arXiv: 2008.13065 [eess.IV].

[11] Crawford. *NIFD LONI — Data Use Summary*. URL: https://ida.loni.usc.edu/collaboration/access/appLicense.jsp.

[12] Kostadin Dabov et al. "Image denoising with block-matching and 3D filtering". In: *Image Processing: Algorithms and Systems, Neural Networks, and Machine Learning*. Ed. by Nasser M. Nasrabadi et al. Vol. 6064. International Society for Optics and Photonics. SPIE, 2006, pp. 354–365. DOI: 10.1117/12.643267. URL: https://doi.org/10.1117/12.643267.

[13] Jennifer Stine Elam and David Van Essen. "Human Connectome Project". In: *Encyclopedia of Computational Neuroscience*. Ed. by Dieter Jaeger and Ranu Jung. New York, NY: Springer New York, 2013, pp. 1–4. ISBN: 978-1-4614-7320-6. DOI: 10.1007/978-1-4614-7320-6_592-1. URL: https://doi.org/10.1007/978-1-4614-7320-6_592-1.

[14] J. A. Fessler and B. P. Sutton. "Nonuniform fast Fourier transforms using min-max interpolation". In: *IEEE Trans. Sig. Proc.* 51.2 (Feb. 2003), 560–74. DOI: 10.1109/TSP.2002.807005.

[15] M.A.T. Figueiredo and R.D. Nowak. "An EM algorithm for wavelet-based image restoration". In: *IEEE Transactions on Image Processing* 12.8 (2003), pp. 906–916. DOI: 10.1109/TIP.2003.814255.

[16] G. H. Glover. "Spiral imaging in fMRI". In: *Neuroimage* 62.2 (Aug. 2012), pp. 706–712.

[17] Ian J. Goodfellow et al. *Generative Adversarial Networks*. 2014. arXiv: 1406.2661 [stat.ML].

[18] Vijay P.B. Grover et al. "Magnetic Resonance Imaging: Principles and Techniques: Lessons for Clinicians". In: *Journal of Clinical and Experimental Hepatology* 5.3 (2015), pp. 246–255. ISSN: 0973-6883. DOI: https://doi.org/10.1016/j.jceh.2015.08.001. URL: https://www.sciencedirect.com/science/article/pii/S0973688315004156.

[19] H. Gudbjartsson and S. Patz. "The Rician distribution of noisy MRI data". In: *Magn Reson Med* 34.6 (Dec. 1995), pp. 910–914.

[20] Kerstin Hammernik et al. *Learning a Variational Network for Reconstruction of Accelerated MRI Data*. 2017. arXiv: 1704.00447 [cs.CV].

[21] Kaiming He et al. *Deep Residual Learning for Image Recognition*. 2015. arXiv: 1512.03385 [cs.CV].

[22] Reinhard Heckel and Paul Hand. *Deep Decoder: Concise Image Representations from Untrained Non-convolutional Networks*. 2019. arXiv: 1810.03982 [cs.CV].

[23] D.I Hoult and Paul C Lauterbur. "The sensitivity of the zeugmatographic experiment involving human samples". In: *Journal of Magnetic Resonance (1969)* 34.2 (1979), pp. 425–433. ISSN: 0022-2364. DOI: https://doi.org/10.1016/0022-2364(79)90019-2. URL: https://www.sciencedirect.com/science/article/pii/0022236479900192.

[24] Y. Hu et al. "Generalized higher degree total variation (HDTV) regularization". In: *IEEE Trans Image Process* 23.6 (June 2014), pp. 2423–2435.

[25] Diederik P. Kingma and Jimmy Ba. "Adam: A Method for Stochastic Optimization". In: *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*. Ed. by Yoshua Bengio and Yann LeCun. 2015. URL: http://arxiv.org/abs/1412.6980.

[26] Xiangfei Kong et al. "A New Image Quality Metric for Image Auto-denoising". In: *2013 IEEE International Conference on Computer Vision*. 2013, pp. 2888–2895. DOI: 10.1109/ICCV.2013.359.

[27] Alex Krizhevsky, Vinod Nair, and Geoffrey Hinton. "CIFAR-10 (Canadian Institute for Advanced Research)". In: (). URL: http://www.cs.toronto.edu/~kriz/cifar.html.

[28] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. "ImageNet Classification with Deep Convolutional Neural Networks". In: *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*. NIPS'12. Lake Tahoe, Nevada: Curran Associates Inc., 2012, pp. 1097–1105.

[29] Christian Ledig et al. *Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network*. 2017. arXiv: 1609.04802 [cs.CV].

[30] Rosanne Liu et al. *An Intriguing Failing of Convolutional Neural Networks and the CoordConv Solution*. 2018. arXiv: 1807.03247 [cs.CV].

[31] Michael Lustig et al. "Compressed Sensing MRI". In: *IEEE Signal Processing Magazine* 25.2 (2008), pp. 72–82. DOI: 10.1109/MSP.2007.914728.

[32] Angshul Majumdar. "Multi-Coil Parallel MRI Reconstruction". In: *Compressed Sensing for Magnetic Resonance Image Reconstruction*. Cambridge University Press, 2015, pp. 86–119. DOI: 10.1017/CBO9781316217795.005.

[33] LC Man, JM Pauly, and A Macovski. "Improved automatic off-resonance correction without a field map in spiral imaging". In: *Magnetic resonance in medicine* 37.6 (June 1997), pp. 906–913. ISSN: 0740-3194. DOI: 10.1002/mrm.1910370616. URL: https://doi.org/10.1002/mrm.1910370616.

[34] LC Man, JM Pauly, and A Macovski. "Multifrequency interpolation for fast off-resonance correction". In: *Magnetic resonance in medicine* 37.5 (May 1997), pp. 785–792. ISSN: 0740-3194. DOI: 10.1002/mrm.1910370523. URL: https://doi.org/10.1002/mrm.1910370523.

[35] E. R. McVeigh, R. M. Henkelman, and M. J. Bronskill. "Noise and filtration in magnetic resonance imaging". In: *Med Phys* 12.5 (1985), pp. 586–591.

[36] Matthew J. Muckley et al. *State-of-the-Art Machine Learning MRI Reconstruction in 2020: Results of the Second fastMRI Challenge*. 2020. arXiv: 2012.06318 [eess.IV].

[37] Douglas C. Noll et al. "Deblurring for non-2D fourier transform magnetic resonance imaging". In: *Magnetic Resonance in Medicine* 25.2 (1992), pp. 319–333. DOI: https://doi.org/10.1002/mrm.1910250210. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.1910250210. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.1910250210.

[38] Klaas P. Pruessmann et al. "Advances in sensitivity encoding with arbitrary k-space trajectories". In: *Magnetic Resonance in Medicine* 46.4 (2001), pp. 638–651. DOI: https://doi.org/10.1002/mrm.1241. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.1241. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.1241.

[39] Olga Russakovsky et al. "ImageNet Large Scale Visual Recognition Challenge". In: *International Journal of Computer Vision (IJCV)* 115.3 (2015), pp. 211–252. DOI: 10.1007/s11263-015-0816-y.

[40] Matthew Tancik et al. *Fourier Features Let Networks Learn High Frequency Functions in Low Dimensional Domains*. 2020. arXiv: 2006.10739 [cs.CV].

[41] George Toderici et al. *Variable Rate Image Compression with Recurrent Neural Networks*. 2016. arXiv: 1511.06085 [cs.CV].

[42] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. "Deep Image Prior". In: *International Journal of Computer Vision* 128.7 (Mar. 2020), pp. 1867–1888. ISSN: 1573-1405. DOI: 10.1007/s11263-020-01303-4. URL: http://dx.doi.org/10.1007/s11263-020-01303-4.

[43] Ashish Vaswani et al. *Attention Is All You Need*. 2017. arXiv: 1706.03762 [cs.CL].

[44] Zhou Wang et al. "Image quality assessment: from error visibility to structural similarity". In: *IEEE Transactions on Image Processing* 13.4 (2004), pp. 600–612. DOI: 10.1109/TIP.2003.819861.

[45] Y. Yang et al. "A comparison of fast MR scan techniques for cerebral activation studies at 1.5 tesla". In: *Magn Reson Med* 39.1 (Jan. 1998), pp. 61–67.

[46] Jure Zbontar et al. *fastMRI: An Open Dataset and Benchmarks for Accelerated MRI*. 2019. arXiv: 1811.08839 [cs.CV].

[47] David Y. Zeng et al. "Deep residual network for off-resonance artifact correction with application to pediatric body MRA with 3D cones". In: *Magnetic Resonance in Medicine* 82.4 (2019), pp. 1398–1411. DOI: https://doi.org/10.1002/mrm.27825. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.27825. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.27825.

[48] Yantian Zhang et al. "A novel k-space trajectory measurement technique". In: *Magnetic Resonance in Medicine* 39.6 (1998), pp. 999–1004. DOI: https://doi.org/10.1002/mrm.1910390618. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.1910390618. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.1910390618.

[49] Hang Zhao et al. "Loss Functions for Image Restoration With Neural Networks". In: *IEEE Transactions on Computational Imaging* 3.1 (2017), pp. 47–57. DOI: 10.1109/TCI.2016.2644865.

[50] Bolei Zhou et al. "Places: A 10 million Image Database for Scene Recognition". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2017).

[51] Bo Zhu et al. "Image reconstruction by domain transform manifold learning". In: *CoRR* abs/1704.08841 (2017). arXiv: 1704.08841. URL: http://arxiv.org/abs/1704.08841.