Project Report : CS 7643

Sonakshi Gupta
sgupta786@gatech.edu

Anastasiia Alokhina
aalokhina3@gatech.edu

Sandaru Fernando
sfernando8@gatech.edu

Abstract

Accelerating Magnetic Resonance Imaging (MRI) by taking fewer measurements has the potential to reduce medical costs and speed up the current MRI procedure. Recently, a lot of research has been done on solving the task of MRI reconstruction from undersampled data using Neural Networks including models such as U-Net and VarNet [8]. We focus on the task of reconstruction of the ground truth data from the single-coil images and using U-Net with added attention blocks. We also change the loss function from L1 to SSIM (Structure Similarity) and compare the results.

1. Introduction/Background/Motivation

MRI scans are an important diagnostic tool in medical imaging, but their long scan times and associated high costs can pose challenges. Deep neural networks offer the potential to significantly reduce the time required for MRI scans, thereby increasing availability and reducing costs. U-Net, a popular architecture for image segmentation and reconstruction, has been shown to be effective for reconstructing fully-sampled MRI images from under-sampled k-space data. By utilizing U-Net for fast MRI reconstruction, medical professionals could potentially improve patient outcomes by providing faster and more accurate diagnoses, while also making this technology more accessible to a wider range of patients.

MRI scans create images of a patients anatomy by placing a patient in a magnetic field, and measuring the frequency of resulting radio signals in a receiver coil. In order to construct images from the collected signals, it is necessary to acquire many samples to apply an inverse Fourier transform to the samples. If not enough samples are acquired, there will be artifacts and a decrease in the quality of the image. However, the cost of an MRI scan is related to the time it takes to acquire the necessary amount of samples. By creating networks that can take undersampled data and construct accurate images, the time needed to acquire an MRI scan will decrease, resulting in a substantial cost saving to the patient.

There are a couple different types of data collection, networks, and other techniques used for reconstructing MRI images. First, state of the art models utilize multiple-coil scans, where multiple receiver coils are used in an MRI machine to simultaneously collect more data. U-Net is a common baseline model for reconstructing MRI images, but there are other networks like Variational Networks[2][6] (VarNet) which are designed to approximate equations used in the construction of MRI images. In addition to different architectures, there is research into reducing or correcting artifacts in the output images, such as adversarial training to reduce the appearance of banding[1]. Additionally, most image reconstruction networks are only capable of using a fixed set of measurements, but there is research into creating additional networks [5][10] that can be trained alongside the image reconstruction networks to determine at inference what samples to take. This technique is called active acquisition, and it allows the number of samples to be reduced, as the MRI scan will only take samples that best increase the image quality.

For our project, we utilized the dataset provided by NYU Langone, which contains data for knee, brain, and prostate. Specifically, we focused on the knee dataset to develop our model. The dataset [8] contains raw multi-coil k-space data, emulated single-coil k-space data, ground-truth images and DICOM images. Further details can be found in [3].

2. Approach

We started by utilizing the pre-trained U-Net model from the fastMRI collection [8] [9] to perform the image reconstruction task. We decided to experiment with different loss functions and changed the original absolute error loss (L1) to structural similarity index (SSIM) to enhance the image quality of the reconstructed images[1].

In contrast to L1 loss, SSIM does not assume pixel-wise independence and accounts for the structural and luminance similarities between images. This is a particularly important criteria for medical imaging applications where accurate interpretation and diagnosis are essential. More details are provided in Section 5.

[1]Code available at: https://github.gatech.edu/sfernando8/DL-Fastmri-UNet-Project

Additionally, we modified the U-Net architecture to include attention mechanisms, which help the model to focus on the most relevant regions of the input image during training, enhancing the accuracy and quality of the output. We trained this modified U-Net model from the pretrained model, using SSIM as the loss function. We made a decision to not train it from scratch because autors in [8] claimed that training from scratch on 32 GPUs took a week which we could not afford to replicate.

Initially, we anticipated encountering some setbacks during the model training process due to the large size of the data, which amounted to approximately 500 GB. To mitigate this issue, we decided to proceed with the smaller single-coil knee dataset, which was only around 100 GB in size. This helped to accelerate the training process by reducing the data load.

Despite our efforts to streamline the training process, we encountered numerous unanticipated errors when setting up the virtual machine on Google Cloud. The setup process was time-consuming, primarily due to the complexity of the infrastructure involved and the constraints on memory. These challenges made it difficult to configure the virtual machine optimally, causing significant delays in the setup process. However, we were able to resolve this issue after some troubleshooting and were able to successfully train the model using the single-coil knee dataset.

3. Experiments and Results

In our experiments, we tested how changing the loss from L1 to SSIM and adding the attention blocks change the performance. We tested three different setups.

- Using the pre-trained U-Net that was originally trained with L1 loss and training it with SSMI loss.
- Adding to the pre-trained U-Net attention blocks, freezing all the original U-Net weights and training it with L1 loss.
- Adding to the pre-trained U-Net attention blocks, freezing all the original U-Net weights and training it with SSIM loss.

Our decision behind adding attention blocks was motivated by the paper [4] that used the mentioned approach for MRI tasks. Our decision to train with SSIM loss was motivated by the fact that it may better capture the structure of the image that it very important for the MRI reconstuction.

Our results are concluded in Table 2. Results of the MRI reconstruction for different models are presented in Table 3

Based on the loss and reconstruction images, our models didn't outperform the baseline model. It is clear that the image quality of our reconstructions looks worse. We have several assumption that may explain our results. First of all, we didn't have enough time to retrain the model completely from scratch over the entire training dataset after changing its architecture or hyperparameters as it would take over a week as reported in [8]. Consequently, we chose to use a subset of the training data to shorten the time to train the models. Second of all, for SSIM loss more hyperparameter tuning is required such as trying more options for window size.

4. Attention Block

For our model, we implemented attention mechanics. We implemented the approach of attention mechanism for U-Net presented in [4] adding attention blocks between the transpose convolution blocks and the convolution blocks. See Figure 1 for the U-Net architecture with attention gates and Figure 2 for the Attention Gate itself.

5. Structural Similarity Loss

We used SSIM loss as a loss function for our two of our training setups. This loss is based on Structural Similarity Index that is an approach to image comparison introduced in [7]. The loss function combines Structural Similarity Index for different patches of the image. We assumed that this loss function would be beneficial because the structural similarity of MRI reconstructions would be more important than just L1 difference. We used the implementation of SSIM from `torch.geometry`.

5.1. Structural Similarity Index

Compares images using the following factors.

- **Luminance** of the image x is defined like $\mu_x = \frac{1}{N}\sum_{i=1}^{N} x_i$. The luminance comparison function $l(\mathbf{x}, \mathbf{y})$ is then a function of μ_x and μ_y:

$$l(\mathbf{x}, \mathbf{y}) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}$$

  where $C_1 = (K_1 L)^2$ where L is the dynamic range of the pixel values (255 for 8-bit grayscale images), and $K_1 \ll 1$ is a small constant.

- **Contrast** of the image x measured by taking the standard deviation of all the pixel values.

$$\sigma_x = \left(\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu_x)^2\right)^{\frac{1}{2}}$$

  The contrast comparison function is the following:

$$c(\mathbf{x}, \mathbf{y}) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}, \quad \text{where} \quad C_2 = (K_2 L)^2$$

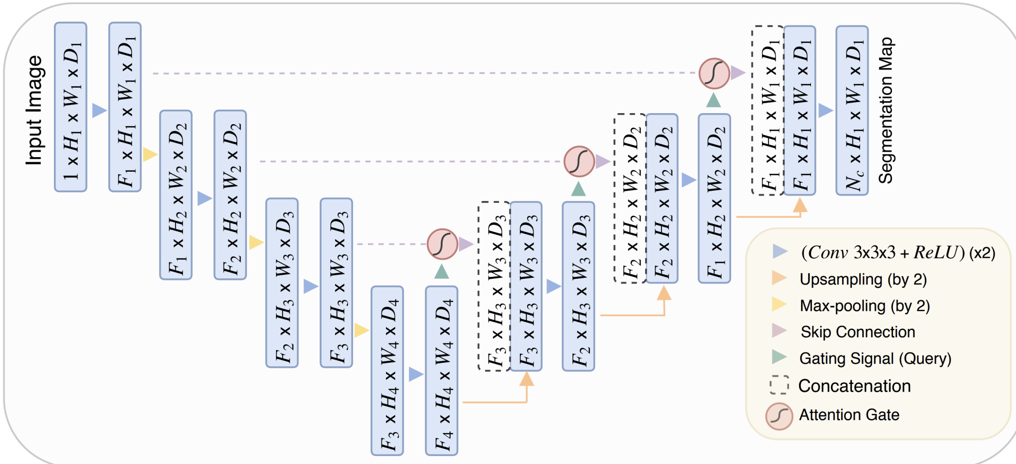

Figure 1. Original architecture of U-Net with attention presented in [4]

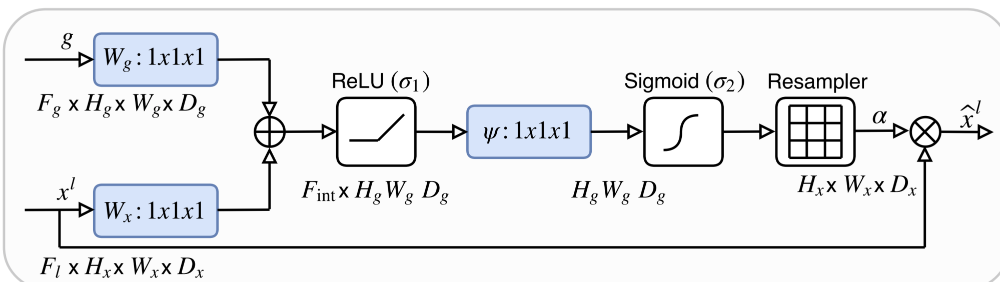

Figure 2. Attention gate

Experiment	Type of training Loss	Val L1 loss	Val SSIM
Baseline model (pre-trained)	L1 loss	0.775424	0.499933
Baseline model	SSIM	0.775421	0.499923
Baseline model (frozen weights) + attention block	L1 loss	0.567888	0.408449
Baseline model (frozen weights) + attention block	SSIM	0.5331835	0.346999

Table 1. Results of experiments.

Experiment	Learning Rate	LR Gamma	LR Step Size	Weight Decay	Dropout	Epochs
Unet + SSIM	0.001	0.1	40	0	0	5
Frozen Unet + attention block + L1 loss	0.0214	0.2215	30	0.0001	0.0909	5
Frozen Unet + attention block + SSIM	0.0016	0.4835	50	0.0023	0.0404	5

Table 2. Hyperparameter Tuning Results

- **Structure** comparison function: It is defined by the function $s(x, y)$ is shown below. σ denotes the standard deviation of a given image, x and y are the two images being compared.

$$s(\mathbf{x}, \mathbf{y}) = \frac{\sigma_{xy} + C_3}{\sigma_x \sigma_y + C_3}$$

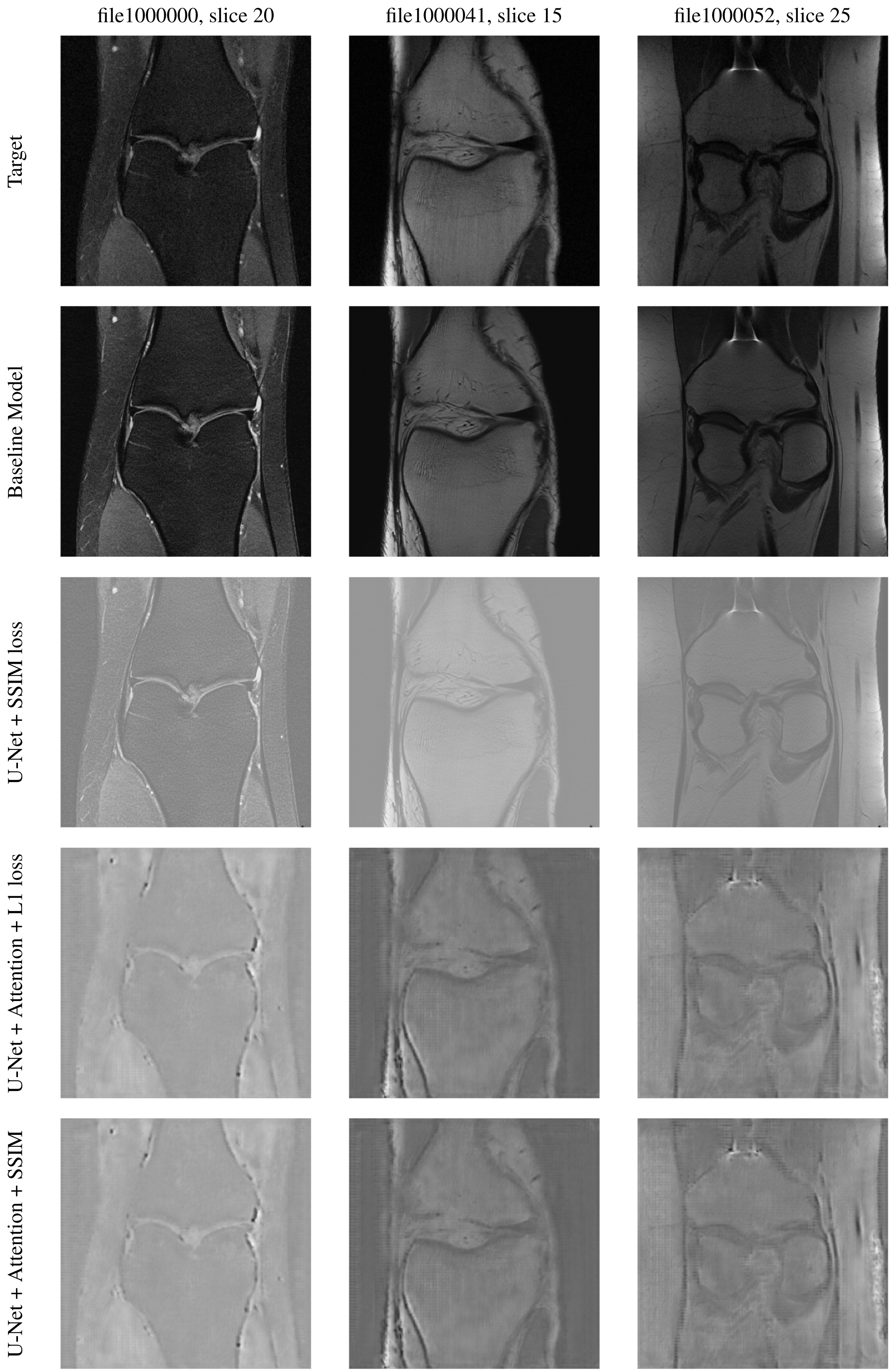

Table 3. Examples of reconstructions.

where $\sigma(\mathrm{xy})$ is defined as,

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \mu_x)(y_i - \mu_y)$$

The total SSIM score is given by,

$$\mathrm{SSIM}(\mathbf{x}, \mathbf{y}) = [l(\mathbf{x}, \mathbf{y})]^{\alpha} \cdot [c(\mathbf{x}, \mathbf{y})]^{\beta} \cdot [s(\mathbf{x}, \mathbf{y})]^{\gamma}$$

6. Work Division

See Figure 4

References

[1] Aaron Defazio, Tullie Murrell, and Michael P. Recht. Mri banding removal via adversarial training, 2020. 1

[2] Reza Khodayi-Mehr and Michael M. Zavlanos. Varnet: Variational neural networks for the solution of partial differential equations, 2019. 1

[3] Florian Knoll, Jure Zbontar, Aditya Sriram, Matthew J Muckley, Miguel Bruno, Anthony Defazio, Michelle Parente, Krzysztof J Geras, Joseph Katsnelson, Hersh Chandarana, Zhongnan Zhang, Michal Drozdzalv, Adriana Romero, Michael Rabbat, Pascal Vincent, Ryan Pinkerton, David Wang, Nargiza Yakubova, Evan Owens, C Lawrence Zitnick, Michael P Recht, Daniel K Sodickson, and Yvonne W Lui. fastmri: A publicly available raw k-space and dicom dataset of knee images for accelerated mr image reconstruction using machine learning. *Radiology: Artificial Intelligence*, 2(1):e190007, 2020. 1

[4] Ozan Oktay, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Mattias Heinrich, Kazunari Misawa, Kensaku Mori, Steven McDonagh, Nils Y Hammerla, Bernhard Kainz, Ben Glocker, and Daniel Rueckert. Attention u-net: Learning where to look for the pancreas, 2018. 2, 3

[5] Luis Pineda, Sumana Basu, Adriana Romero, Roberto Calandra, and Michal Drozdzal. Active MR k-space sampling with reinforcement learning. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, pages 23–33. Springer International Publishing, 2020. 1

[6] Anuroop Sriram, Jure Zbontar, Tullie Murrell, Aaron Defazio, C. Lawrence Zitnick, Nafissa Yakubova, Florian Knoll, and Patricia Johnson. End-to-end variational networks for accelerated mri reconstruction. In Anne L. Martel, Purang Abolmaesumi, Danail Stoyanov, Diana Mateus, Maria A. Zuluaga, S. Kevin Zhou, Daniel Racoceanu, and Leo Joskowicz, editors, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, pages 64–73, Cham, 2020. Springer International Publishing. 1

[7] Zhou Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004. 2

[8] Jure Zbontar, Florian Knoll, Anuroop Sriram, Matthew J. Muckley, Mary Bruno, Aaron Defazio, Marc Parente, Krzysztof J. Geras, Joe Katsnelson, Hersh Chandarana, Zizhao Zhang, Michal Drozdzal, Adriana Romero, Michael G. Rabbat, Pascal Vincent, James Pinkerton, Duo Wang, Nafissa Yakubova, Erich Owens, C. Lawrence Zitnick, Michael P. Recht, Daniel K. Sodickson, and Yvonne W. Lui. fastmri: An open dataset and benchmarks for accelerated MRI. *CoRR*, abs/1811.08839, 2018. 1, 2

[9] Jure Zbontar, Florian Knoll, Anuroop Sriram, Tullie Murrell, Zhengnan Huang, Matthew J. Muckley, Aaron Defazio, Ruben Stern, Patricia Johnson, Mary Bruno, Marc Parente, Krzysztof J. Geras, Joe Katsnelson, Hersh Chandarana,

Student Name	Contributed Aspects	Details
Anastasiia Alokhina	Data Creation, Infrastructure Set Up and Implementation	Loaded the dataset for this paper. Helped to set up the cloud platform. Adapted training to use SSIM loss. Debugged the training script.
Sandaru Fernando	Attention Implementation, Infrastructure, Analysis	Helped set up, administrate, and debug virtual machines. Created implementation of Attention Unet, training script, and hyperparameter tuner.
Sonakshi Gupta	Infrastructure Set Up, Testing, Analysis	Helped set up the virtual machine. Assisted in hyperparameter search. Set up the testing script

Table 4. Contributions of team members.

Zizhao Zhang, Michal Drozdzal, Adriana Romero, Michael Rabbat, Pascal Vincent, Nafissa Yakubova, James Pinkerton, Duo Wang, Erich Owens, C. Lawrence Zitnick, Michael P. Recht, Daniel K. Sodickson, and Yvonne W. Lui. fastMRI: An open dataset and benchmarks for accelerated MRI, 2018. 1

[10] Zizhao Zhang, Adriana Romero, Matthew J. Muckley, Pascal Vincent, Lin Yang, and Michal Drozdzal. Reducing uncertainty in undersampled mri reconstruction with active acquisition, 2019. 1