Precision-Driven MRI Reconstruction

Bianca Toto, Sofia Muller, Travis Taylor
Georgia Institute of Technology
Atlanta, Georgia, United States of America
{btoto3, sofia.muller, ttaylor99}@gatech.edu

Abstract

In this paper, we build on a model from Facebook Research's fastMRI project. Magnetic Resonance Imaging (MRI) is essential in medical diagnostics, yet long acquisition times lead to high costs and reduce the number of patients that can be diagnosed. The fastMRI project aims to accelerate MRI reconstruction using deep learning algorithms. In particular, they provide a baseline model called a U-net model. We improve on this model in two ways; one, through custom loss functions, and two, with a knowledge distillation model. For the custom loss functions, we added a center-weighted loss function, where diagnosis is generally more likely to occur, and we additionally leveraged fastMRI+ bounding box annotations to focus weights on areas indicating pathology. The knowledge distillation model uses the pre-trained base model as a "teacher" model and trains a "student" model of reduced size, thereby improving accuracy, leading to faster inferences and similar intermediate representations as measured by Centered Kernel Alignment (CKA).

1. Introduction

We aimed to make MRI scans faster and more efficient by improving the way images are reconstructed, focusing primarily on the parts that healthcare professionals need most for diagnosis. MRI scans can take a long time, which is uncomfortable for patients and limits how many people can be scanned in a day. In knee MRI scans, problems often occur near the knee joint, but current methods spend time and effort reconstructing the full image, even the less important areas. Our goal was to develop a method that concentrates on reconstructing these critical areas more accurately while spending less effort on less important regions. By doing this, we hoped to reduce the overall time needed for MRI scans without sacrificing the quality of the image regions that are most important for diagnosing patients.

Currently, MRI images are reconstructed using techniques that treat all parts of the image equally. This means the scanner collects a large amount of data from every region. The reconstruction algorithms then process all data uniformly. While this ensures that the image is of consistent quality, it also means that time and computational resources are spent on areas irrelevant to diagnosis. In knee imaging, for example, health issues are typically localized near the joint. Unfortunately, the current practice does not take advantage of this knowledge. Parallel imaging and compressed sensing techniques speed up MRI scans. However, they still process the entire image uniformly and have limitations in how much they can accelerate scans without degrading image quality. In other words, they may not fully exploit the potential benefits of focusing on diagnostically significant regions. Further, these methods often require complex resources, which makes them less accessible for widespread clinical use globally.

If successful, the approach could have significant benefits. Patients would experience shorter MRI scan times which may reduce discomfort associated with procedures. This is particularly important for individuals who struggle to remain still for extended periods (e.g., children). For healthcare providers, this means high-quality images focused on the critical areas for diagnosis based on their specialty, economic growth, increased patient throughput, and fast clinical decisions. Overall, success would mean contributing to making MRI technology more accessible, cost-effective, and efficient while enhancing patient care.

For our project, we utilized the single-coil knee MRI data from the publicly available fastMRI dataset, provided by NYU Langone Health and Facebook AI Research in 2020 [2]. This dataset is curated explicitly for research purposes and offers high-quality, clinically relevant data.

1.1. Critical aspects of the fastMRI dataset

1.1.1 Type of data

The dataset includes de-identified raw k-space data, which captures the frequency-domain signals collected during MRI scans. Using this raw data allows us to work directly with the fundamental signals of the MRI process, giving us

the opportunity to implement and test advanced reconstruction techniques while maintaining clinical relevance.

1.1.2 Domain and focus

Our project focuses on knee imaging, which is well-suited for studying localized pathologies and anatomical structures near the joint. The knee was chosen because of its clinical significance and the variety of conditions that can be diagnosed through imaging, such as ligament tears and cartilage degeneration.

1.1.3 Diversity

The dataset includes scans from a range of patients and is captured using both 1.5 Tesla and 3 Tesla magnets. This variety in imaging settings and patient characteristics helps make our models more adaptable and reliable for use in real-world situations.

1.1.4 Configuration/Preprocessing

We utilized the single-coil configuration, which combines data from multiple coils into a single simplified view. This configuration reduced computational complexity and was ideal given our resource constraints. Despite this, the data is of high quality and suitable for deep learning (DL) tasks. The knee single-coil dataset included a 72.7 GB training set, a 14.9 GB validation set, and a 1.4 GB test set.

1.1.5 Clinical relevance

For our stretch goals, we incorporated the fastMRI+ dataset, which is an extension of the original dataset that includes clinical pathology annotations. These annotations identify knee abnormalities and help us in training the model to focus on diagnostically significant regions. The fastMRI+ dataset enhances our project by providing additional clinical context, which is crucial for the accurate diagnosis of knee pathologies.

1.1.6 Composition

The knee dataset includes over 1,300 scans with proton density (PD) and PD with fat suppression. Additionally, the dataset supports single-coil and multi-coil reconstruction tasks, enabling comparisons of more straightforward and more advanced methods.

1.2. Prior Work

1.2.1 Compressed Sensing

In 2006, the breakthrough of compressed sensing was applied to MRI imaging to reconstruct images from theoretically incomplete measurements. The Nyquist-Shannon theorem provides a lower bound for the number of samples required in the frequency domain (k-space) to reproduce a signal in the spatial domain. However, this theory assumes no a priori information. In contrast, compressed sensing takes advantage of a priori information to reconstruct images in the spatial domain at a sampling rate that violates the Nyquist-Shannon theorem [5].

1.2.2 U-Net Architecture

When Facebook Research published the fastMRI dataset, they also published baseline models, including a U-Net DL model. A diagram of the U-Net architecture can be seen in Figure 1. U-Net models have been used in MRI segment analysis. The structure of a U-Net model is similar to an encoder-decoder model; the ConvBlocks in the model down-sample to a latent space and then up-sample to reconstruct the image. However, in the U-Net, copies of the output from each downsampling layer are sent to the corresponding upsampling layer.

The Facebook GitHub repository includes code for the baseline models in PyTorch [5]. We used this code as our starting point. Our code is available at: https://github.com/travisdoesmath/fastMRI.

2. Approach

Our primary goal was to develop a deep learning-based MRI reconstruction method that could accelerate imaging workflows by focusing attention on the clinically significant areas of knee scans. Knee pathologies are often localized near the joint, but traditional reconstruction methods treat every pixel in the image equally. By concentrating on reconstruction accuracy in the most critical regions, we aimed to maintain diagnostic quality while potentially reducing overall scan times and computational burden.

2.1. Implementation and Challenges

To achieve our goals, we began by extending the existing fastMRI codebase and adapted it to our needs. There was some difficulty building on top of the fastMRI repository due to deprecated package features and inconsistent package dependencies. However, once we carefully set up a working environment, the team leveraged the flexibility to design custom neural network architectures, hooks, loss functions, and datasets. We employed PyTorch for its flexibility and integrated PyTorch Lightning to streamline routine tasks such as checkpointing, gradient management, and logging.

Using LightningModules, we structured our data pipelines, model logic, and evaluation metrics to ensure modularity and reusability. The Trainer API further simplified the training process by providing support for distributed training and automatic mixed precision (AMP),

which optimized computational efficiency. This combination of PyTorch's customizability and Lightning's automation allowed us to efficiently experiment with various neural network architectures, hooks, loss functions, and data configurations.

By establishing this pipeline, we ensured that our training and experimentation processes were consistent, reproducible, and adaptable to both baseline and advanced approaches. As a baseline, we started with a U-Net model, known for its strong performance in image reconstruction tasks. We then focused on three enhanced models: Center-Weighted, Knowledge-Distillation, and Annotation-Weighted.

The computational resources available to us locally were an AMD-based Linux desktop with an Nvidia RTX 3060 GPU, M2 MacBook Pro laptops, and an Intel-based MacBook Pro laptop. Additionally, we also used the Google Cloud Platform (GCP) to create virtual machines. Using the 3060, training the baseline 32-channel U-Net model on the full single coil knee dataset took 24 minutes per epoch. Training the 256-channel U-Net model on the full single coil knee dataset took around 10 hours per epoch. For comparison, the baseline model was trained to 50 epochs.

We tested several different GPUs available on GCP, including a T4, a P100, and a V100. The T4 and P100 GPUs were not significantly faster than the 3060. The V100 was significantly faster, but prohibitively expensive; training the 256-channel U-Net model for 50 epochs was projected to cost more than $300.

To balance performance and computational efficiency, and to achieve reasonable training times, we used a fixed subsample of 100 training files (out of 975) that were randomly selected. This subset still amounted to roughly 10 GB. We stored the dataset on Huggingface for ease of access. Additionally, we trained to only 10 epochs for evaluating our models. With more time and resources available, we would train to 50 epochs to match the baseline model.

2.2. Center-Weighted Models

To guide the U-Net's attention to the knee joint, we introduced a weighted loss function that increased the penalty for errors in the MRI's central region. This "center-weighting" was designed to ensure that the model, during training, would allocate more representational power to areas of clinical significance. To achieve this, we applied a weight mask that doubled the loss contribution from this region.

The original plan was to train a center-weighted single-coil model with 256 channels over 50 epochs and conduct hyperparameter tuning using a grid search for a fair comparison with the baseline pretrained U-Net from the fastMRI dataset. However, we held the hyperparameters constant due to computational constraints, adopting the default values from the fastMRI implementation. Instead, we experimented with various channel sizes to evaluate how center-weighted reconstructions influence reconstruction quality and denoising capabilities as representational power changes.

2.3. Knowledge Distillation Framework

To handle the growing computational demands of training larger or more complex models, we employed a knowledge distillation framework to transfer representations from a large, pretrained teacher model (256 channels, 496 million parameters) to a smaller student model (32 channels, 7.8 million parameters, with an additional 2.1 million parameters for projection layers). Knowledge distillation enables the student model to learn from the teacher's behavior, achieving similar performance while reducing computational requirements [1]. This approach is particularly advantageous for deployment in resource-constrained environments, where it allows for model compression without significant loss of accuracy, and was a direct response to our realized computational limitations.

Feature maps were extracted from selected layers in both models, including Block 1 of the encoder's downsampling path, the bottleneck, and Block 4 of the decoder's upsampling path. To align the feature dimensions between the models, trainable projection layers were introduced to upsample the student's feature maps to match the size of the corresponding teacher feature maps. During training, hooks were placed in both models to capture the intermediate feature outputs at these layers. The extracted features were normalized for scale consistency and then passed through the projection layers.

The distillation loss was computed by summing the mean squared error (MSE) between the teacher's feature maps and the projected student feature maps at each selected layer. This loss was combined with a weighted reconstruction MSE loss that emphasizes the central regions of the image, which are typically more informative in knee MRI reconstruction tasks. The total loss function is given by:

$$\mathcal{L}_{\text{total}} = \alpha \cdot \mathcal{L}_{\text{reconstruction}} + (1 - \alpha) \cdot \mathcal{L}_{\text{distillation}} \quad (1)$$

where $\alpha = 0.75$, $\mathcal{L}_{\text{reconstruction}}$ is the weighted reconstruction loss, and $\mathcal{L}_{\text{distillation}}$ is the teacher-student distillation loss. The weighted reconstruction loss is defined as:

$$\mathcal{L}_{\text{reconstruction}} = \frac{1}{N} \sum_{i=1}^{N} w_i \cdot (y_i - \hat{y}_i)^2 \quad (2)$$

where y_i and $\hat{y}_i$ represent the target and predicted pixel values, respectively, and w_i is a spatial weight emphasizing the central regions of the image, computed using a central mask. The teacher-student distillation loss is defined as:

$$\mathcal{L}_{\text{distillation}} = \frac{1}{M} \sum_{j=1}^{M} \|\text{Teacher}(\mathbf{F}_j) - \text{Projection}(\mathbf{F}_j)\|_2^2 \tag{3}$$

where $\mathbf{F}_j$ represents the feature maps at the selected layers, and the projection function adjusts the student feature maps to match the teacher's feature map dimensions.

By combining these two loss terms, the student model learns to minimize both the reconstruction error, and the difference between its intermediate representations and the teacher's. This strategy allows the student model to improve its performance while maintaining a smaller model size. The approach balances computational efficiency and reconstruction accuracy as it enables the student to learn from the reconstruction targets and the teacher's richer intermediate representations, thereby reducing model complexity, while improving performance.

2.4. FastMRI+: Clinical Annotation Weighting

The publicly available fastMRI+ data set augments the fastMRI dataset with pathology labels and bounding boxes, providing researchers with focal areas to analyze model performance (Figure A3)[6]. Building on this, we went a step further and leveraged the bounding boxes to create a mask giving greater weight to the areas of the image that indicate a pathology during the training step. Specifically, we adapted our previously implemented center-weighted model and replaced the center weights with pathology-based weights. Whenever pathology bounding boxes were not available, we defaulted to our center weights. We successfully trained, tested, and evaluated a model using a small subsample of five images. However, scaling up to a larger dataset revealed significant time constraints. The current implementation accesses the fastMRI+ data set in every train iteration, leading to prohibitively slow training times. Future research could address this limitation by optimizing the dataset access or applying our approach in settings with greater computational resources. The code for our implementation can be found in the `fastmriplus` branch of our repository.

3. Experiments and Results

To evaluate model success, we measured both quantitative performance metrics and qualitative reconstruction quality. Quantitatively, we used standard metrics for assessing deep learning-based MRI reconstructions, including Mean Squared Error (MSE), Structural Similarity Index (SSIM), Normalized Mean Squared Error (NMSE), and Peak Signal-to-Noise Ratio (PSNR) [5, 4]. These metrics provided insights into the pixel-level fidelity of the reconstructed images compared to ground-truth data and the perceptual quality of the reconstructions.

For experimental validation, we trained the models on a subsampled subset of the fastMRI dataset under various configurations, such as different channel sizes, with center or annotation weighting, and using knowledge distillation. Comparative experiments against the baseline pretrained U-Net model were conducted to understand how the different sizes, training procedures, and loss functions changed feature representations.

The baseline model we used is the pretrained U-Net model with 256 channels trained for 50 epochs. We compare this to our center-weighted models, which used 32, 64, 128, and 256 channels. The results of these models can be seen in Table 1. A chart of the MSE for the different models is shown in Figure 2; we choose to visualize MSE specifically because of its ease of interpretation, and the other metrics show similar behavior. We see that for the center-weighted model, error decreases as the number of channels increases, which is to be expected. Additionally, we also compared the knowledge distillation model with 32 channels. These results can be seen in Table 2. It is worth noting that the knowledge distillation model outperforms the 256 channel center-weighted model, despite only having 32 channels. Note that the doubling of channels in a model increases the model size four-fold, so the knowledge distillation model is a significantly smaller model despite being more accurate. This also leads to a six-fold speedup in inference. This approach was especially appealing given the GPU constraints, costs, and time pressures we faced.

An additional metric, Centered Kernel Alignment, was used to measure the similarity between the teacher and student neural network representations [3].

$$\text{CKA}(K, L) = \frac{\text{HSIC}(K, L)}{\sqrt{\text{HSIC}(K, K) \cdot \text{HSIC}(L, L)}}, \tag{4}$$

Qualitatively, we visually inspected reconstructed images to identify artifacts, assess noise reduction, and evaluate the preservation of clinically relevant details in the center-weighted or annotated regions. This step was critical for understanding the practical implications of the reconstructions and determining whether the models were learning to emphasize diagnostically significant features. Figure 1 shows an example comparison between the 256 channel center-weighted model and the knowledge distillation model after 1 training epoch. Note that the center weighted reconstruction shows significant banding compared to the minimal banding of the knowledge distillation model. Further reconstruction examples and their target images can be found in Figure A1 in the Appendix.

3.1. Knowledge Distillation

The results of the knowledge distillation model indicate strong average performance but highlight areas for improvement due to variability across all metrics (Table 2). While

Metric	256 Channel	128 Channel	64 Channel	32 Channel
Total Model Params	496 million	124 million	31 million	7.8 million
Training Time[a]	12.9hrs, 90hrs	3.9hrs, 28hrs	1.3hrs, 10hrs	0.5hrs, 3.3hrs
MSE	1.54e-10 ± 2.91e-10	1.97e-10 ± 3.73e-10	1.64e-10 ± 3.0e-10	1.99e-10 ± 3.79e-10
NMSE	0.05 ± 0.10	0.07 ± 0.13	0.05 ± 0.10	0.07 ± 0.12
PSNR	30.25 ± 6.63	29.22 ± 6.66	29.98 ± 6.56	29.18 ± 6.42
SSIM	0.87 ± 0.08	0.85 ± 0.08	0.87 ± 0.08	0.85 ± 0.08

Table 1. Results for center weighted knee models for 10 epochs. [a]Training time shows the total training time for the subsampled training set and an estimated time for the full training set, respectively.

Metric	Pretrained fastMRI	Center Knee Reconstruction	Knowledge Distillation
Epochs Trained	50	10	10
Channels Used	256	256	32
Loss Function	L1 Loss	Equation 2	Equation 1
Total Model Params	496 million	496 million	7.8 million
Training Dataset Used	Singlecoil fastMRI	Sampled Singlecoil fastMRI	Sampled Singlecoil fastMRI
Total Train Time	N/A	12.9hrs	7.9hrs
MSE	6.34e-11 ± 1.24e-10	1.54e-10 ± 2.91e-10	1.14e-10 ± 2.05e-10
NMSE	0.02 ± 0.04	0.05 ± 0.10	0.04 ± 0.06
PSNR	34.24 ± 6.02	30.25 ± 6.63	31.52 ± 5.99
SSIM	0.93 ± 0.05	0.87 ± 0.08	0.88 ± 0.07

Table 2. Comparison of model types and metrics.

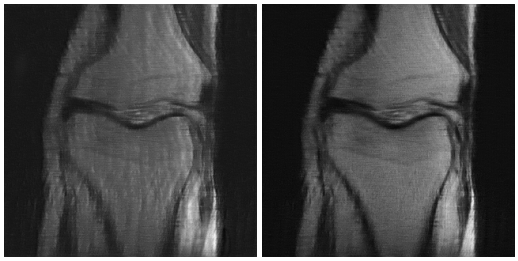

Figure 1. Example reconstructions after one training epoch: 256-channel center-weighted model (left) and 32-channel knowledge distillation model (right).

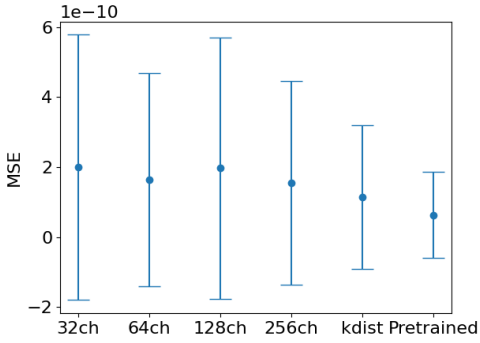

Figure 2. MSE on validation set for all models.

the model generally performs well in pixel-wise accuracy and structural similarity, the observed fluctuations suggest inconsistencies in reconstruction quality. The high standard deviation in MSE and PSNR reflects some variability in pixel-level precision and overall image clarity, possibly caused by dataset imbalance from the sampling process, training instability from the knowledge distillation set up, and/or limited training time/ reduced complexity. The model's NMSE suggests that, while the overall energy distribution in the reconstructed images is generally well-preserved, certain regions, specifically those with less complex structures, may exhibit lower reconstruction accuracy. This could be due to factors such as noise amplification or insufficient information in the undersampled data. Additionally, SSIM results indicate strong structural preservation, but some variability in the reconstruction of anatomical details, especially in regions with less pathological content, may contribute to the observed NMSE values.

The model showed significant improvements in noise reduction and artifact minimization, especially in the central regions with enhanced clarity of key anatomical structures such as lesions, cartilage, and soft tissue. The central areas

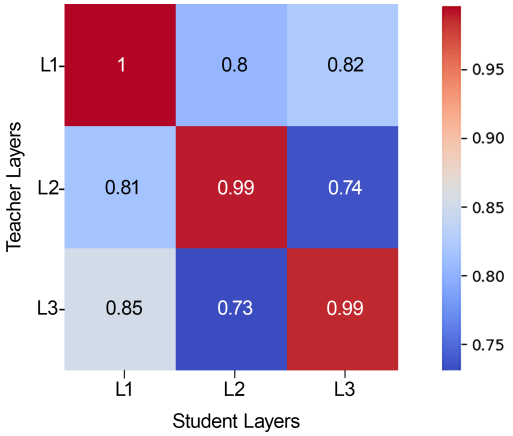

Figure 3. Centered kernel alignment (CKA) matrix for 3 layers (L1: down_sample_layers.0.layers.0, L2: conv.layers.7, and L3: up_conv.3.1) of teacher and student models.

demonstrated sharper detail, reflecting the model's ability to focus on diagnostically relevant regions and learn more complex representations from the teacher model, surpassing models with larger channel sizes (except the pretrained teacher). However, subtle artifacts persisted in the peripheral regions, which could be addressed by refining the loss function to better balance focus between central and peripheral areas or by introducing architectural adjustments like attention mechanisms. Visually, the reconstructions were closer to the ground truth in the central regions where anatomical structures showed improved sharpness, clarity, and less banding (Figure A1). While the peripheral regions exhibited some deviations, they remained clinically acceptable. The enhanced alignment with the ground truth, despite reduced model complexity, suggests that the increased signal from the intermediate representations and the gradients of the weights in the knowledge distillation process helped produce final outputs that more closely resembled the ground truth compared to other model approaches. This is also supported by Figure 3 which highlights how the corresponding teacher and student convolutional layers from the blocks included in the knowledge distillation training had high similarity. Non-corresponding layers were moderately similarity due to resource contraints that prevented including a more diverse, larger sample size in the CKA calculations.

To address these issues, future work could focus on refining the training process, enhancing regularization techniques by increasing batch size or dropout probability, increasing the number of epochs, and training on the full dataset to refine the parameters, improve generalization, and reduce variability. Future efforts could explore ablation or increasing the number of layers included in the knowledge distillation process, potentially experimenting with more specific layer outputs vs. block outputs, or changing the direction of the projections (downsampling the teacher instead of upsampling the student features) which may enhance the student model's ability to capture complex patterns and stabilize training. Additionally, testing alternative loss functions, beyond the current combination of two MSE functions, or different values of alpha, could potentially offer better trade-offs between pixel-wise accuracy, structural preservation, and focus on reconstruction/ learning the intermediate teacher representations.

4. Work Division

A complete summary of team member contributions is provided in Table 3. Please note that all team members contributed meaningfully to both the technical components of the project, including implementation, experimentation, and analysis, and to supporting tasks such as report writing.

References

[1] Geoffrey Hinton, Oriol Vinyals, and Jeff Dean. Distilling the knowledge in a neural network, 2015. 3

[2] Florian Knoll, Jure Zbontar, Anuroop Sriram, Matthew J. Muckley, Mary Bruno, Aaron Defazio, Marc Parente, Krzysztof J. Geras, Joe Katsnelson, Hersh Chandarana, Zizhao Zhang, Michal Drozdzalv, Adriana Romero, Michael 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: 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. PMID: 32076662. 1

[3] Simon Kornblith, Mohammad Norouzi, Honglak Lee, and Geoffrey E. Hinton. Similarity of neural network representations revisited. *CoRR*, abs/1905.00414, 2019. 4

[4] Matt Muckley. fastmri - attentive reconstruction, 2020. 4

[5] 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. 2, 4

[6] Ruiyang Zhao, Burhaneddin Yaman, Yuxin Zhang, Russell Stewart, Austin Dixon, Florian Knoll, Zhengnan Huang, Yvonne W. Lui, Michael S. Hansen, and Matthew P. Lungren. fastmri+: Clinical pathology annotations for knee and brain fully sampled multi-coil mri data, 2021. 4

Student Name	Contributed Aspects	Details
Kevin Eguiluz	Paper/README, Contributions, FastMRI+ Model	FastMRI+ Model, README (covering description, installation and execution, running the scripts), debugging assistance, stretch goals, and paper contributions and editing.
Sofia Muller	FastMRI+ Model, Database Preparation Assistance, Virtual Environment/GPU Research	Directly responsible individual for implementing FastMRI+ Model and providing report summary. Prepared data for upload to hugging face, explored alternative virtual environment of Azure/Inner Eye. Debugging assistance.
Travis Taylor	DevOps, MLOps, Paper Contributions	Managed VMs and codebase, created sampled training set, ran and coordinated training runs, paper contributions and LaTeX typesetting, debugging assistance.
Bianca Toto	Center Weighted Model, Knowledge Distillation, Centered Kernel Alignment (including class and visualization), and Paper Contributions	Created pipeline scripts (data, train, test/ reconstruction, evaluate) for center weighted and knowledge distillation models. Implemented the data modules to work with huggingface, centered weighted loss function and model, knowledge distillation loss functions and model. Center Weighted Model overview and Package Implementation; Knowledge distillation overview and results; Experiment Intro and Tables. Debugging assistance.

Table 3. Contributions of team members.

A. Appendix

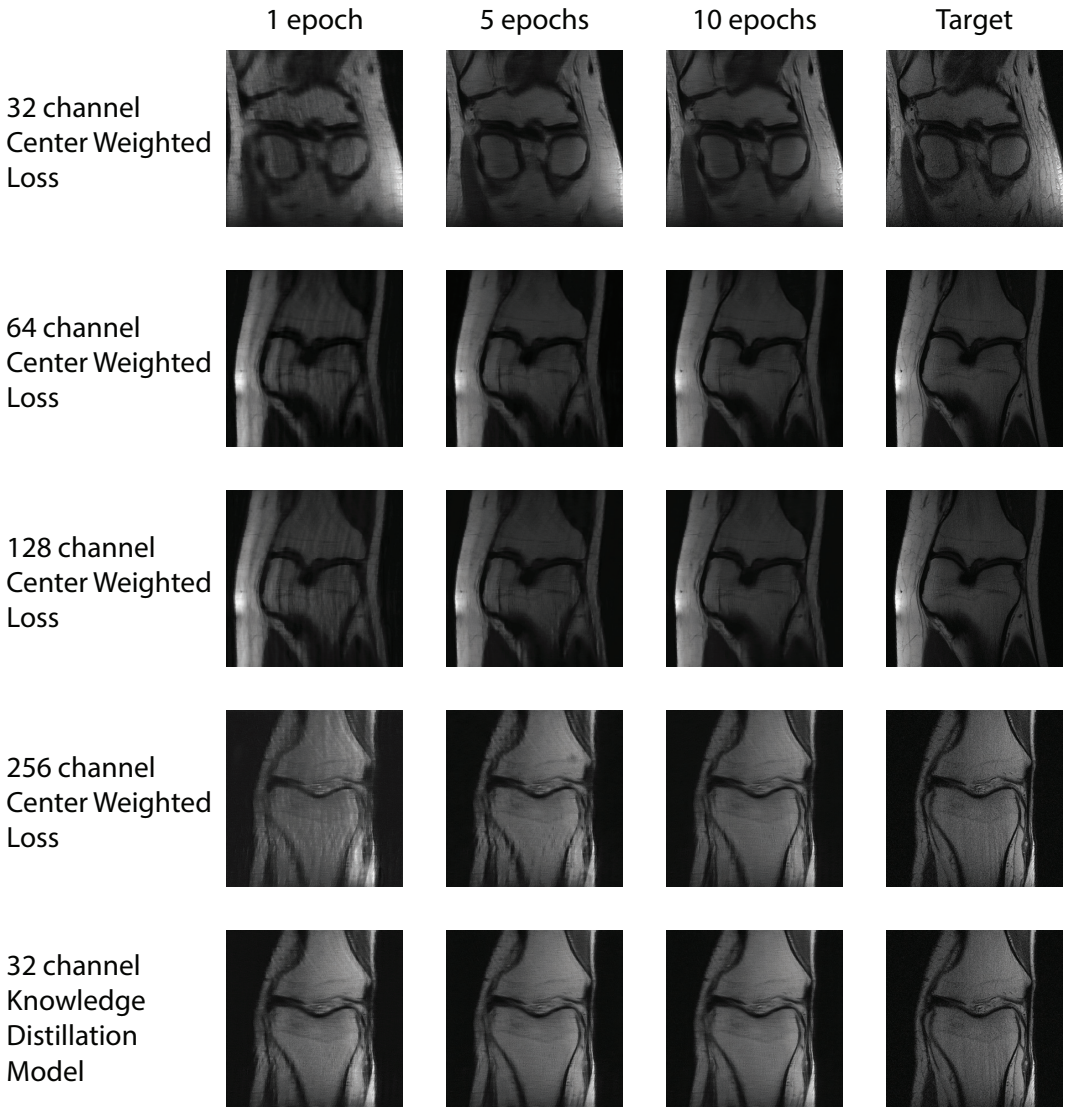

Figure A1. Visualizations across epochs for center weighted models versus ground truth.

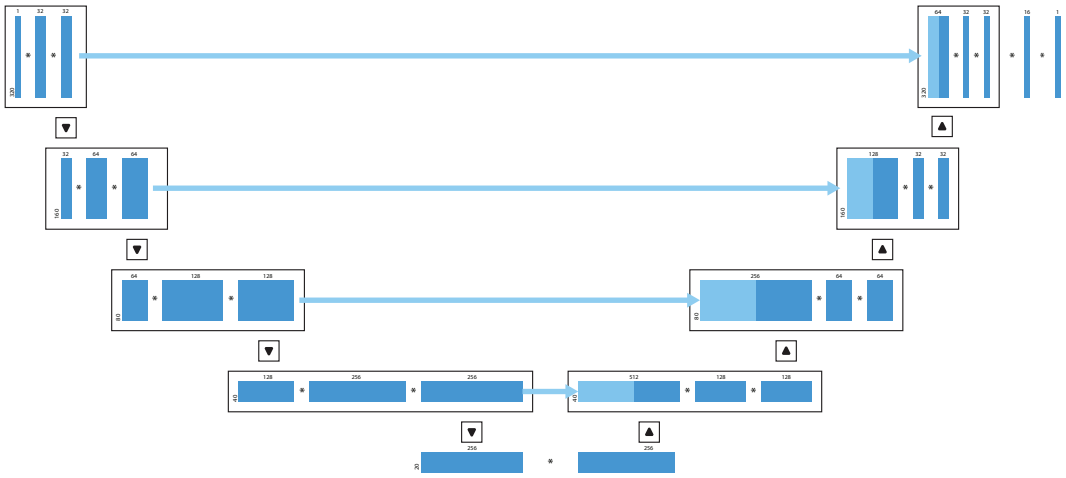

Figure A2. Baseline U-Net architecture; 32 channels.

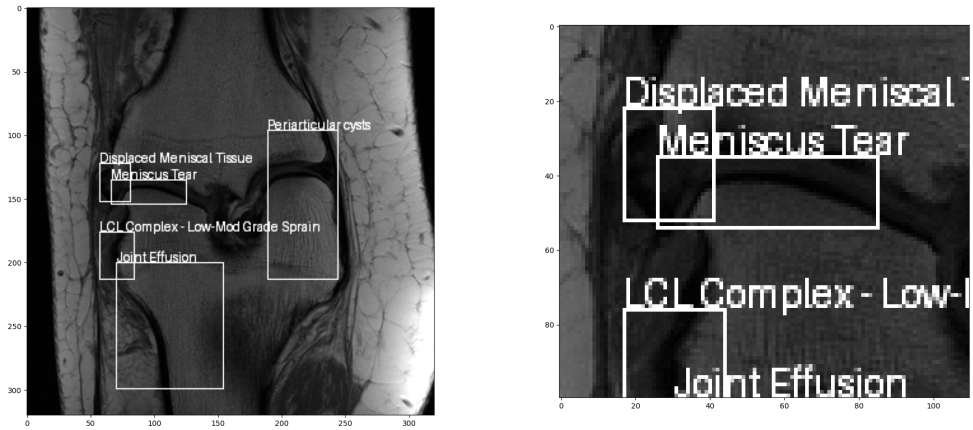

Figure A3. Sample image with multiple bounding boxes from fastMRI+ (left). Zoom into pathology bounding boxes (right).

b