Title

Deep Learning-Based Image Enhancement with Uncertainty Estimation for Rapid T2-Weighted Knee MRI Using a 3D DESS Sequence

Author

Deb, Ashmita

Publication Date

2023

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
Los Angeles

Deep Learning-Based Image Enhancement with Uncertainty Estimation for Rapid T_2-Weighted Knee MRI Using a 3D DESS Sequence

A thesis submitted in partial satisfaction of the
requirements for the degree Master of Science
in Bioengineering

by

Ashmita Deb

2023

ABSTRACT OF THE THESIS

Deep Learning-Based Image Enhancement with Uncertainty Estimation for Rapid T_2-Weighted Knee MRI Using a 3D DESS Sequence

by

Ashmita Deb

Master of Science in Bioengineering

University of California, Los Angeles, 2023

Professor Holden H. Wu, Chair

Magnetic resonance imaging (MRI) is a powerful imaging technique for visualizing soft tissues and characterizing tissue properties. It can be used for the assessment of meniscal, ligamentous, and cartilaginous lesions in the knee. When the extracellular matrix of articular cartilage is compromised, water moves more freely within the cartilage, which leads to prolonged T_2 relaxation times. Consequently, quantitative T_2 maps have been used to understand the pathophysiology of osteoarthritis or follow-up monitoring of knee cartilage after surgery. Standard T_2 quantification techniques usually use spin echo-based sequences which require long acquisition times. Three-dimensional (3D) dual echo steady state (DESS) MRI has shown its potential for time-efficient T_2 mapping in the knee. DESS acquires images with distinct contrasts ("FID" and "Echo").

This feature of DESS allows for both morphological T_2-weighted (T2w) imaging and quantitative T_2 mapping in a single scan. Our research group at UCLA has developed a protocol to acquire 3D DESS scans with isotropic high-resolution. However, the scan time still needs to be reduced to facilitate eventual translation.

In order to reduce acquisition time in T_2 mapping sequences, previous works developed methods to reconstruct images using undersampled data. Previous proposals such as compressed sensing (CS) reconstruction have been used to reconstruct images from undersampled data while mitigating undersampling artifacts. However, CS reconstruction methods are time-consuming and the results are sometimes overregularized. Deep learning (DL) based image enhancement or reconstruction methods provide a solution to the shortcomings of CS reconstruction by learning the mapping between undersampled input data and high-quality output images from large reference datasets, and providing rapid inference times. Since the isotropic high-resolution 3D DESS acquisition designed by our group at UCLA is relatively recent and has a limited number of datasets, strategies such as transfer learning, i.e., pretraining on a larger available reference dataset, may be needed to obtain a DL network that can produce high-quality output images from our undersampled high-resolution 3D DESS MRI data.

Another main challenge of DL-based image enhancement or reconstruction methods is that it is difficult to understand why/how they work or to predict its performance. In medical image enhancement, image fidelity must be prioritized. Obstruction or elimination of image details may confound diagnostic decisions. To overcome this problem, recently there are some works on developing DL networks with uncertainty estimation to predict error in enhancement tasks.

Thus, the aim of this work is to shorten the scan time of 3D DESS MRI by reconstructing high quality images from undersampled data using DL networks that incorporate uncertainty estimation and transfer learning.

The thesis of Ashmita Deb is approved.

Kyunghyun Sung

Dan Ruan

Holden H. Wu, Committee Chair

University of California, Los Angeles

2023

Dedicated to my parents:

Biswa Mohan Deb &

Diya Deb

Table of Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

ACKNOWLEDGEMENTS

I would like to start by expressing my sincere gratitude to my advisor, Dr. Holden Wu. His leadership and advice has helped me grow in leaps and bounds over the two years of pursuing my Master's degree, not just as a professional in the field, but as an individual as well. He has provided me with endless support, patience and guidance, that I am forever grateful for. His immense knowledge and far-sightedness lead me to an amazing research field. His optimism and enthusiasm for research motivates me to be the best version of myself, every day. I am extremely thankful to have Dr. Wu as my advisor.

I would like to thank the rest of my committee members, Dr. Kyunghyun Sung and Dr. Dan Ruan. They helped me expand my horizon and gave me the opportunity to explore different aspects of my research topic. Their advice allowed me to refine my project and improve the technical developments.

I would also like to thank all the members of the Magnetic Resonance Research Labs, who have supported me and provided valuable insights for my research. I would like to give special thanks to ShuFu Shih. As a senior member of Dr. Wu's lab, he has helped me improve my technical understanding of various concepts in MRI research.

Lastly, I would like to thank my father, Dr. Biswa Mohan Deb, whose love and wisdom remains with me, even today, and my mother, Mrs. Diya Deb, for being my rock and my biggest cheerleader.

Chapter 1: Introduction

1.1 Significance and Motivation

Osteoarthritis (OA) is a degenerative joint disease that primarily affects the cartilage, which is the protective tissue that covers the ends of bones within a joint. OA commonly affects weight-bearing joints such as the knees, hips, and spine, as well as the hands and feet. The cartilage gradually wears away, causing the bones to rub against each other. This can result in pain, stiffness, swelling, and reduced range of motion in the affected joint. OA is one of the major causes of disability and loss of independence among middle aged and elderly individuals[1-5].

Treatment for OA focuses on managing symptoms, reducing pain, and improving joint function. This can include a combination of non-pharmacological approaches such as exercise, physical therapy, weight management, and assistive devices, as well as medications for pain relief and inflammation. In severe cases, joint replacement surgery may be recommended[6-8]. Early diagnosis and appropriate management of OA can help individuals maintain an active and fulfilling lifestyle while minimizing pain and disability associated with the condition.

Surowiec et al[9] discusses the current methods for assessing cartilage health and the need for non-invasive imaging tools to track early stages of OA. Clinical evaluation of pain symptoms and subjective outcome scoring are commonly used, but more robust methods are required. Arthroscopy is the gold standard for diagnosing cartilage degeneration, but it is invasive. Radiographs are limited to detecting moderate to severe cartilage loss[10-11].

Magnetic Resonance Imaging (MRI) has shown a strong potential to provide

reliable and accurate information about soft tissues in the human body[12]. MRI T_2 relaxation reflects water molecule interactions at the cellular level and has the potential to detect pathological tissue alterations. Conventional MRI methods acquire T_2-weighted (T2w) images for anatomical assessment, but it is less sensitive to early OA-related biochemical changes. Quantitative MRI (qMRI) techniques, such as T_2 mapping, have shown sensitivity to biochemical changes in cartilage[13] and are helpful in assessing cartilage degeneration and tissue repair. Conventional T_2 mapping methods, such as spin echo (SE)-based sequences, acquire images at multiple echo times (TE) which are fitted to an exponential model to generate T_2 maps. However, SE-based acquisitions are time-consuming and it is difficult to achieve high spatial resolution and large 3D coverage. Using a three-dimensional (3D) dual echo steady state (DESS)[14] MRI technique for rapid T_2 mapping addresses these drawbacks in the context of knee imaging[15-28]. However, the scan time for isotropic high-resolution 3D knee imaging using DESS MRI still needs to be reduced.

While MRI knee presents opportunities for characterization of OA, the technique has its own set of limitations; for example, long scan acquisition times. Undersampling data to reduce the scan time would reduce motion artifacts and increase patient comfort. However, this would result in poor image quality and aliasing artifacts in the image. Advanced methods like compressed sensing (CS)[29-31] and deep learning (DL)[32] based image enhancement or reconstruction can help mitigate the artifacts of accelerated scans. In this work, we studied DL-based image enhancement of undersampled knee MRI.

However, DL networks are complex and it is often difficult to interpret why/how the network is performing a certain way. In the context of medical image enhancement/reconstruction, this can potentially cause severe problems as elimination of

minute details within the image may be misleading and impede diagnostic decisions. To combat this drawback of DL networks, uncertainty estimation has been explored in recent years. Hence, we tried to quantify the model uncertainty of the DL networks used in this work to estimate the error in image enhancement.

Another drawback of DL networks is the need for large training datasets to perform well. As mentioned previously, knee imaging requires high spatial resolution and large, homogeneous 3D coverage. To address this, our research group at UCLA has designed a 3D DESS sequence with isotropic high-resolution[34]. Since this DESS acquisition method was recently designed, only a limited number of scans have been acquired. Hence, we used the transfer learning approach to perform DL-based enhancement on the undersampled 3D isotropic DESS data by pretraining on a large, publicly available reference dataset and then fitting the model to our dataset[35-37].

1.2 Specific Aims

The main objective of this work was to achieve accelerated 3D DESS MRI of the knee by training DL network models to enhance undersampled 3D DESS MRI and characterizing the DL model uncertainty to estimate error in the enhanced images. This objective was achieved via the following specific aims:

Specific Aim 1: Accelerate scan time for 3D DESS knee MRI by training DL networks to produce enhanced images from undersampled data and characterize the error in enhanced images through uncertainty quantification.

Approach: DL networks based on the U-Net and the Shifted windows transformer for Image Restoration (SwinIR) architectures were trained and tested on the publicly

available Stanford Knee Dataset[38] with retrospective undersampling. The Structural Similarity index (SSIM) and Peak Signal-to-Noise Ratio (PSNR) of the enhanced images with respect to the reference images were used to quantify the quality of the enhanced images. To estimate the epistemic uncertainty (i.e., model uncertainty), Monte Carlo Dropout was applied to the network and the Pearson R correlation scores were reported between the enhancement error and the uncertainty map values in the knee cartilage to assess the potential of characterizing errors during image enhancement.

Specific Aim 2: Accelerate scan time for a limited set of 3D isotropic high-resolution DESS knee MRI by using transfer learning-based deep learning models to produce enhanced images from undersampled data.

Approach: DL network models were pre-trained with a processed version of the Stanford Knee DESS dataset, and then fine-tuned for a limited set of 3D isotropic high-resolution DESS knee MRI acquired at UCLA. Retrospective undersampling was performed to generate input images to the DL networks. The SSIM and PSNR of the enhanced images with respect to the reference images were calculated to assess quality of the enhanced 3D isotropic high-resolution DESS images.

Chapter 2: Background

2.1 Dual Echo Steady State MRI

The authors of Bruder et al[14] proposed using a three-dimensional (3D) dual echo steady state (DESS) MRI technique for rapid T_2 mapping. The DESS sequence acquires two steady-state free precession echoes within one repetition time (TR), generating images with different contrasts: free-induction decay (FID) and Echo. This sequence provides the advantage of obtaining 3D morphological and functional information from the same dataset in a relatively short imaging time[15-18]. A detailed modelling of the DESS FID and Echo signals is required to obtain a curve to map Echo/FID signal intensity ratios to T_2 values in each pixel. The ratio of the Echo signal to the FID signal intensity in each pixel is thus fitted to the curve to obtain the T_2 map. The sequence diagram for the 3D DESS sequence is shown in **Figure 2-1**.

DESS MRI offers multiple advantages over spin echo-based methods for T_2 mapping. The 3D DESS sequence benefits from a high SNR efficiency compared to conventional 2D multi-slice spin-echo sequences, which leads to improved image quality and enhanced visualization of structures within the knee joint. The high SNR enables better differentiation of tissues, making it suitable for evaluating cartilage, ligaments, tendons, and other structures with fine detail[19-25].

DESS sequences typically have shorter acquisition times compared to multi-echo spin echo-based T_2 mapping sequences. This is because DESS utilizes a steady-state magnetization and allows for rapid imaging. Spin echo-based methods require the acquisition of multiple echo times, sometimes necessitating multiple acquisitions, for T_2 mapping. Thus, due to reduced scan times, DESS imaging has less motion artifacts[26-28].

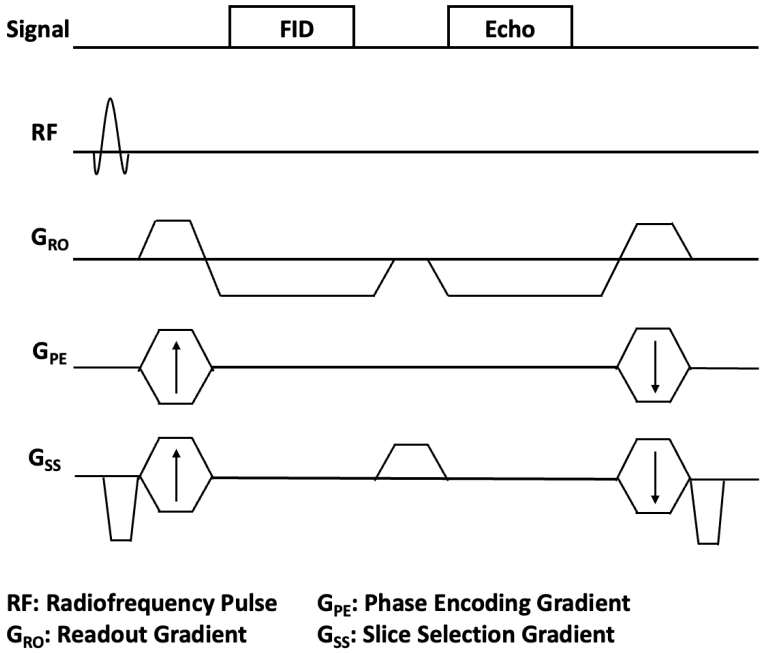

Figure 2-1: 3D DESS sequence diagram. Two contrasts, FID and Echo, are acquired within one TR.

In MRI, a shorter acquisition time would reduce the likelihood of motion artifacts, and would also improve patient comfort and feasibility for clinical translation. This can be achieved by undersampling the data. However, data undersampling introduces aliasing artifacts in the image. These artifacts can degrade image quality and affect the diagnostic accuracy of the exam.

In order to maintain image quality as it would be when the data is fully sampled, advanced reconstruction techniques can be applied to reconstruct images from undersampled data to shorten the scan acquisition time. MR images often exhibit sparsity in an appropriate transform domain. The transform sparsity of MR images and the inherently compressible nature of MR acquisition make compressed sensing (CS) a suitable method for MRI reconstruction[29-30].

2.2 Compressed Sensing Reconstruction for MRI

Due to the compressible nature of the MR image, the effective number of degrees of freedom is much lower than the nominal number of samples. The CS framework formulates the problem of signal recovery as an optimization problem. The goal is to find the sparsest solution that satisfies the given measurements or constraints. The optimization problem is typically solved using algorithms such as iterative thresholding, basis pursuit, or convex relaxation methods[31]. As a result, accurate signal reconstruction can be achieved with a nonlinear procedure using relatively few measurements. MRI reconstruction as a CS optimization problem is illustrated using **Equation 2-1**.

Equation 2-1

$$x = argmin(\left|\left|F(x) - y\right|\right|_2 + \alpha * TV(x) + \beta * W(x)) ,$$

Where, x is the reconstructed image, $F(.)$ is the Fourier Transform, y is the measured signal, $\left|\left|.\right|\right|_2$ is the l_2 norm, $TV(.)$ is the Total Variation (TV), $W(.)$ is the Wavelet Transform and α and β are the TV penalty weight and l_1 penalty weight respectively.

CS reconstruction usually prefers undersampled data with incoherent aliasing energy, which can be facilitated by using a pseudo-random undersampling pattern to acquire k-space data points. Various k-space sampling schemes can be employed, providing flexibility in the trade-off between scan time reduction and image quality. CS can also be combined with parallel imaging techniques to further accelerate the acquisition and improve image reconstruction quality[39-40]. This combination allows for even greater scan time reduction while maintaining image fidelity.

Although CS may improve reconstructed image quality, it poses a set of limitations in the context of medical image reconstruction[41-45]. First, it is computationally expensive.

The computational requirements for solving the optimization problem can be high, especially for high resolution datasets or datasets with a large volumetric coverage, leading to increased processing time. Second, CS algorithms often require tuning regularization parameters to achieve optimal results. Selecting appropriate values for these parameters require expertise or extensive parameter exploration.

DL-based image enhancement (or reconstruction) has certain advantages over CS reconstruction techniques, such as requiring short inference times and not relying on assumptions about sparse representation of input data. DL networks learn data characteristics from a large reference dataset as opposed to CS algorithms where sparsity is assumed in a transform domain and applied separately to each case.

2.3 DL-Based Image Enhancement

In recent years, there have been many works employing DL-based image enhancement for improving quality from undersampled (accelerated) MRI acquisitions[32,46]. The inference time of DL models is generally faster than the iterative optimization required by CS. DL models can learn complex patterns and relationships from the training data, allowing them to capture the inherent structure and characteristics of MR images more effectively. These models, with their ability to learn non-linear mappings, can effectively capture and model these non-linear relationships, producing high-quality images, similar to CS[47-50]. DL-based methods can learn to suppress noise and mitigate artifacts by exploiting the complex patterns and structures in the training data. They can learn without relying on explicit assumptions about the data or the imaging process, and effectively produce high-quality images even from highly undersampled or noisy measurements[51-58].

In recent years, Convolutional Neural Networks (CNNs)[59-60] have revolutionized many computer vision tasks due to their ability to automatically learn and extract meaningful features from the images. CNNs are a type of deep learning model specifically designed for processing grid-like data, such as images or time series data.

CNNs are composed of multiple layers, including convolutional layers, pooling layers, and fully connected layers. The core operation in a CNN is the convolution operation, which includes a mathematical operation of applying a set of filters or kernels to the images to extract features. Each filter slides across the input, performing element-wise multiplications and summations to produce a feature map. These filters learn to capture different patterns and features present in the input data. CNNs typically consist of multiple convolutional layers.

One of the key advantages of CNNs is their ability to automatically learn hierarchical representations of data. The initial layers capture low-level features such as edges and textures, while deeper layers capture more complex and abstract features. This hierarchical representation learning makes CNNs highly effective in capturing and understanding the visual patterns present in images.

CNNs are commonly used for image enhancement tasks in MRI. These models learn to map the undersampled input data to a high-quality output image. For example, the U-Net architecture[61] is well suited for medical image enhancement tasks due to its skip connections enabling the network to capture fine-grained details and high-level contextual information. U-Net has shown robust performance even when training data is limited. The architecture can be adapted and customized for various MRI image enhancement/reconstruction tasks[62-65].

Vision Transformers (ViTs)[66] are another class of deep learning models that apply the transformer architecture, originally developed for natural language processing (NLP), to computer vision tasks. Unlike CNNs, which have been the dominant approach in computer vision for a long time, ViTs operate on the entire input image as a sequence of patches, rather than relying on localized convolutions. The core idea behind ViTs is to decompose the input image into a grid of fixed-size patches and linearly transform them into a sequence of embeddings. These embeddings are then processed by a transformer architecture, which consists of multiple layers of self-attention and feed-forward neural networks. One of the advantages of ViTs is their ability to model long-range dependencies in the image, which can be challenging for CNNs.

The Shifted windows transformer for Image Restoration (SwinIR) network[67], based on vision transformers, has several advantages over CNN structures, while potentially providing image enhancement performance comparable to CNNs. SwinIR architecture is capable of capturing long-range dependencies and modeling complex relationships in the image data, which can be advantageous for MRI image reconstruction tasks that involve capturing global context and dependencies. SwinIR adopts a hierarchical structure by dividing the input image into patches and processing them hierarchically. This enables the model to capture both local and global information, which can be beneficial for preserving fine details and spatial relationships in the reconstructed images. SwinIR incorporates self-attention mechanisms, enabling the model to capture global relationships and dependencies between patches, which is important for understanding the context and semantics of the image. This is particularly advantageous for image enhancement or reconstruction tasks in MRI, where capturing context information and long-range dependencies can be crucial[68,69].

In spite of the benefits of DL-based image enhancement over other advanced reconstruction techniques as mentioned previously, DL methods lack interpretability. DL models are sometimes hard to explain, making it challenging to understand the reasoning behind their decisions. This can raise ethical and legal concerns in the medical domain, where the interpretability of decisions and transparency in the decision-making process are critical[70-71]. Ensuring the robustness and fidelity of deep learning-based reconstruction algorithms is very important, but conventional DL-based methods typically do not provide explicit uncertainty measures or confidence intervals for the predictions[72]. Uncertainty quantification is a potential solution as it aims to provide additional information about the reliability and confidence of the reconstructed images as well[73].

2.4 Uncertainty Quantification

The inherent complexity and nonlinearity of deep learning models make it challenging to assess the uncertainty of the reconstructed images solely based on the network architecture. To address this, uncertainty quantification methods have been developed to capture and quantify the uncertainty in deep learning-based image enhancement. One common approach is to incorporate Bayesian inference into the deep learning framework[33,74-76]. Bayesian deep learning treats the model weights as random variables and estimates their posterior distribution given the observed data. This allows for the estimation of uncertainty by sampling multiple plausible weight configurations from the posterior distribution. Monte Carlo Dropout[77] is another popular technique used for uncertainty quantification in deep learning.

Uncertainty quantification methods can potentially provide valuable information in MRI enhancement/reconstruction from undersampled data. The estimated uncertainty can

be used for various purposes, such as identifying regions of high uncertainty that may require further examination, guiding downstream analysis or decision-making, and improving the overall reliability and trustworthiness of the reconstructed images[78].

Chapter 3: Methodology

3.1 DL-Based Image Enhancement for DESS MRI

The U-Net and the SwinIR networks were used to perform image enhancement on undersampled 3D knee DESS MRI. Monte Carlo Dropout (MCDO) was used to estimate the epistemic uncertainty for both networks.

3.1.1 Defining Uncertainty

The "epistemic" or "model" uncertainty was explored in this study to estimate the error in DL-based image enhancement of undersampled MR images. The MCDO method was applied to obtain the uncertainty estimates during inference. The MCDO method utilizes a concept called dropout, a technique commonly used for model regularization. When dropout is applied, a fraction of the neurons is randomly selected and "dropped out". This is implemented by temporarily ignoring the outputs from those neurons of the given layer in the neural network. The fraction of neurons dropped out can be selected as hyperparameter. In model regularization, dropout is turned on during training, to prevent the model from overfitting the training data.

In the context of uncertainty estimation, dropout is turned off while training the network and all neurons are used. While running model inferencing, dropout is turned on. This leads to ignoring the outputs from a stochastically selected group of neurons while obtaining an output image. When this process is run again, a different group of neurons is stochastically selected to be turned off and another image is obtained at the output. This image is slightly differently from the first output image as they were not obtained using the exact same networks. When this process is run multiple times, the variance of the output images within each pixel is an estimate of how uncertain the model is during inferencing.

The pixel-wise variance calculation is given in **Equation 3-1**.

Equation 3-1

$$Uncertainty = Var[X] = \frac{1}{T}\sum_{t=1}^{T}(X_t - E[X])^2 \ ,$$

Where X_t is the t[th] network output and $E[.]$ is the expectation operator.

In this study, dropout was only used for uncertainty estimation and not as a regularization technique and was thus turned off during training. During inferencing, 50% dropout was used, which means that 50% of the neurons within particular layers were excluded. The inferencing was run 20 times to generate 20 instances of inference results. As dropout selects neurons at random, the 20 inferences were made by 20 slightly different networks. The mean of these 20 instances was taken as the final output of the network (i.e., the enhanced image). The variance of the 20 instances was used as the epistemic uncertainty map for the output enhanced image.

3.1.2 2D U-Net Architecture with Monte Carlo Dropout

The U-Net architecture[61] was originally introduced by Olaf Ronneberger et al. in 2015. It has a U-shaped design, with a contracting path followed by an expanding path (**Figure 3-1(A)**).

Contracting Path (Encoder): The contracting path consists of several convolutional layers which helps capture the context and high-level features of the input image. It is followed by two main operations: a downsampling operation such as max pooling to reduce the spatial dimensions and a second operation to increase the number of feature channels. Each convolutional layer is typically followed by a rectified linear unit (ReLU) activation function, which introduces non-linearity and allows the network to learn nonlinear relationships.

Expanding Path (Decoder): The expanding path is responsible for integrating information from the extracted feature maps into the resultant image that has the same dimensions as the input image. It consists of upsampling operations such as transposed convolutions or bilinear interpolation to increase the spatial dimensions. Each upsampling operation is followed by a concatenation step that combines the feature maps from the corresponding contracting path layer. This skip connection allows the network to preserve the fine-grained details learned during the contracting path. The concatenated feature maps are then processed by a series of convolutional layers to refine the final image.

The skip connections in the U-Net architecture facilitates the integration of both local and global context information. By combining features from different scales, the U-Net can capture both fine-grained details and high-level semantic information, making it particularly effective for image enhancement and reconstruction tasks.

Training and Inferencing: The U-Net is implemented using a training phase and an inference phase. The input to the network is each of the 2D axial slices individually passed from the 3D undersampled DESS knee image, with FID and Echo magnitude images stacked as 2 channels. During the training phase, the network learns to optimize its parameters by minimizing the mean squared error (MSE) loss, which measures the discrepancies between the enhanced image and the reference image using **Equation 3-2**. The Adam optimizer[79] was used with a learning rate of 0.05. In the inference phase, the trained U-Net takes a previously unseen set of 2D undersampled images as input and produces a set of enhanced 2D images as output. **Figure 3-1(B)** shows the inferencing process for image enhancement along with generating uncertainty maps.

Equation 3-2

$$MSE\ Loss = Var[X] = \frac{1}{n}\sum_i (y_i - \hat{y}_i)^2,$$

Where y_i is the i[th] reference value, $\hat{y}_i$ is the predicted value and n is the number of observation.

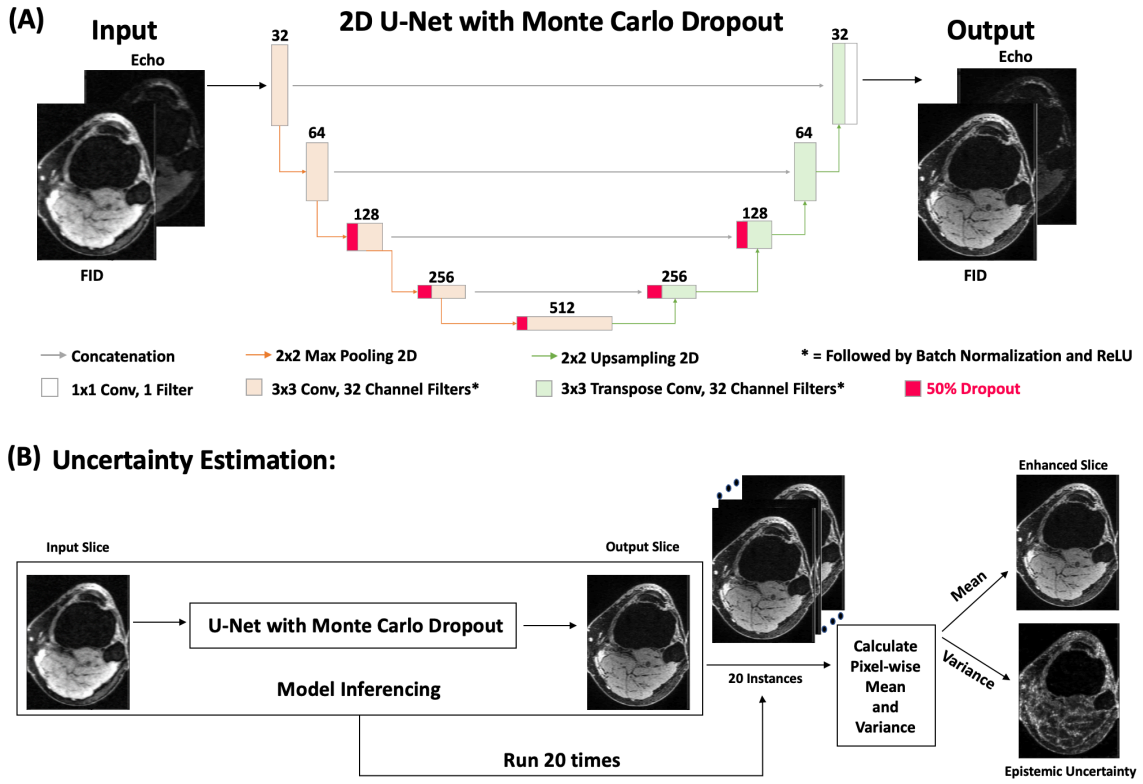

Figure 3-1: (A) 2D U-Net architecture implemented with Monte Carlo Dropout, with FID and Echo images stacked as two channels for input and output to the network. (B) The process of estimating the epistemic uncertainty by running inference 20 times on a single input instance

3.1.3 2D SwinIR Architecture with Monte Carlo Dropout

The SwinIR architecture, introduced by Huang et al.[67], is a deep learning model designed for image reconstruction tasks, including MR image reconstruction and enhancement (**Figure 3-2(A)**). It is based on the Swin Transformer, which is a variant of the transformer architecture that was developed for natural language processing tasks. The

SwinIR architecture incorporates several key components and techniques to achieve effective image enhancement: (a) Patch-based hierarchical structure: SwinIR divides the input image into smaller patches and processes them hierarchically. This allows the model to capture both local and global information. The patches are processed through multiple stages or layers, each consisting of a set of Swin blocks. (b) Swin block: The Swin block is the basic building block of the SwinIR architecture. It consists of two main components: a shifted window-based self-attention mechanism and a feed-forward neural network. The self-attention mechanism helps the model capture long-range dependencies and spatial relationships within the patches, while the feed-forward network is responsible for feature extraction and non-linear mapping. (c) Shifted window-based self-attention: The self-attention mechanism in the Swin block is designed to handle large images efficiently. It employs a shifted window-based approach, where the attention is computed locally within a window and then shifted across different positions. This approach reduces the computational complexity compared to the standard self-attention mechanism, making it more feasible for processing large images. (d) Positional encoding: SwinIR uses positional encoding to provide the model with spatial information about the input image. This helps the model understand the relative positions of different patches and capture spatial dependencies effectively. (e) Skip connections: SwinIR incorporates skip connections between different stages or layers of the architecture. These connections allow the model to access low-level features and fine details from earlier stages, which can be beneficial for preserving important information during the image enhancement process.

Overall, the SwinIR architecture combines the advantages of the Swin Transformer, such as capturing long-range dependencies and hierarchical processing, with specific

adaptations for image enhancement tasks. This allows it to achieve high-quality image enhancements and reconstructions in MRI.

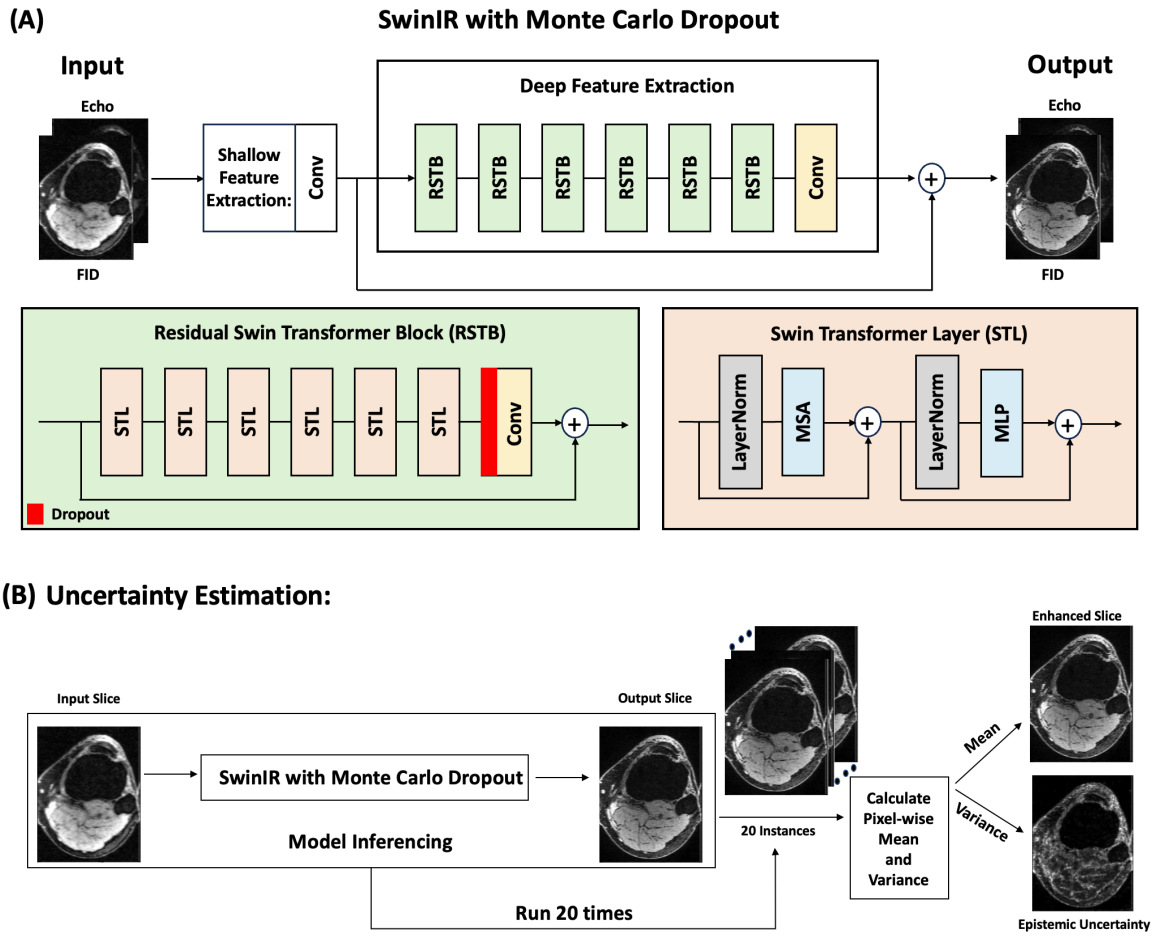

Figure 3-2: (A) 2D SwinIR architecture implemented with Monte Carlo Dropout, with FID and Echo images stacked as two channels for input and output to the network. (B) The process of estimating the epistemic uncertainty by running inference 20 times on a single input instance

Training and Inferencing: The working principle of SwinIR is a supervised learning task. The model learns to map the undersampled images to enhanced, clean images through the hierarchical processing and self-attention mechanisms. The input to the network is each of the 2D axial slices individually passed from the 3D undersampled DESS knee image, with FID and Echo magnitude images stacked as 2 channels. Adam optimizer

was used to train the network. The network was optimized using the MSE loss function (**Equation 3-2**). The process of estimating the epistemic uncertainty using the 2D SwinIR network with MCDO is shown in **Figure 3-2(B)**.

3.2 Dataset

We evaluated the performance of U-Net and SwinIR image enhancement with uncertainty estimation in two different 3D DESS MRI datasets.

3.2.1 3D DESS Knee Dataset (Public)

The Stanford Knee MRI (SKM) dataset[38] was used for training and testing DL networks for image enhancement and uncertainty quantification in this study. This dataset was also used for pretraining DL networks in a transfer learning approach. Out of 155 available subjects, 95 were used for this study: 65 subjects for training, 10 for validation, and 20 for hold-out testing.

The 3D DESS MRI datasets in the SKM were acquired at 3 T in the sagittal orientation and scan time was 9 minutes 25 seconds with 2-fold parallel imaging. After parallel imaging reconstruction (to obtain all k-space data samples), each k-space dataset was retrospectively undersampled by factors of 4, 6 and 8 using variable density undersampling masks provided within this dataset (see **Figure 3-3**). The undersampling masks were applied in the k_y-k_z dimensions (corresponding to axial slices), achieving equivalent scan acquisition times of 5 minutes, 3 minutes 16 seconds and 2 minutes 26 seconds respectively. The reference images used as the ground truth for all models using this dataset were reconstructed using the SENSE[81] method for parallel imaging. **Table 3-1** shows the DESS sequence parameters for this dataset.

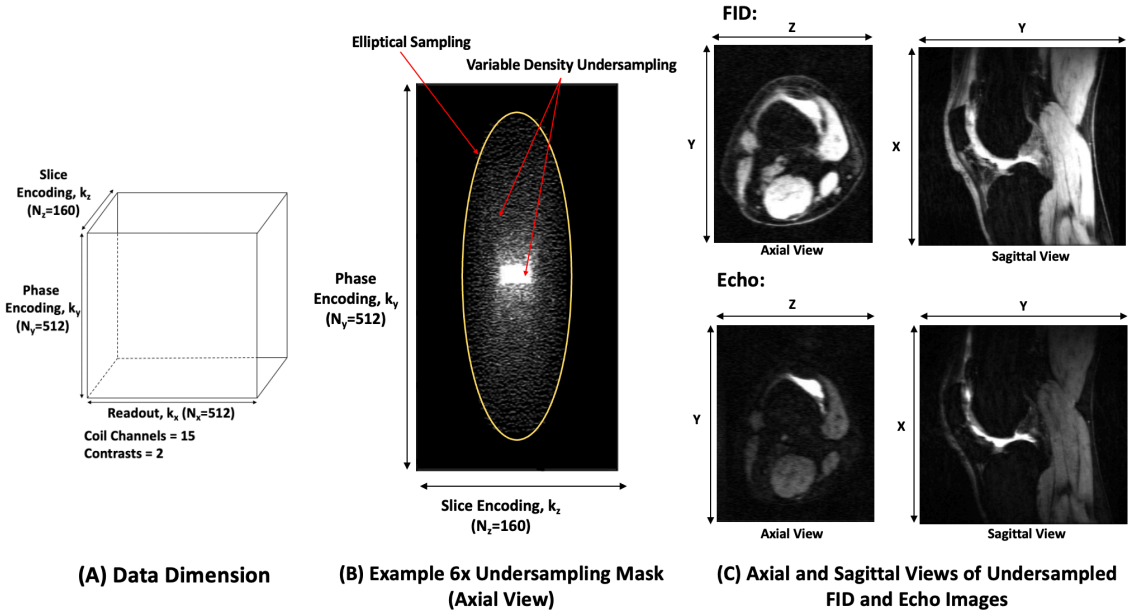

Figure 3-3: Stanford Knee Dataset[38] details showing (A) the raw k-space data dimensions; (B) an example undersampling mask showing 6-fold variable density undersampling over an elliptical data sampling trajectory; (C) A single axial and sagittal slice showing the results of applying the undersampling mask to the entire volume.

qDESS Sequence Parameters for Stanford Knee Dataset[38]	
Matrix (RO x PE)	416 x 512
Resolution (mm^2)	0.38 x 0.31
Echo Time, TE – Echo 1(ms)	5.7
Number of Echoes	2
Repetition Time, TR (ms)	17.9
Flip Angle (°)	20
Parallel Imaging	2 x 1
Scan Time	9 min, 25 sec

Table 3-1: Sequence parameters used to acquire the 3D knee DESS scans in the SKM dataset.

3.2.2 3D Isotropic High-Resolution DESS Knee Dataset (UCLA)

Our research group at UCLA has designed a 3D DESS sequence with isotropic high-resolution (0.66x0.66x0.66 mm^3)[34]. Development and evaluation of this sequence is still in progress and thus very limited knee DESS MRI data is available currently (approximately 13 subjects). Thus, a transfer learning approach was adopted to apply DL-based image enhancement to the UCLA dataset.

The original 3D isotropic high-resolution DESS MRI datasets were acquired at 3 T in the sagittal orientation and scan time was 7 minutes 48 seconds with 2-fold parallel imaging. **Table 3-2** shows the DESS scan parameters. Each dataset, after parallel imaging reconstruction, was retrospectively undersampled in k-space by a factor of 4, with an accelerated scan time of 4 minutes 18 seconds. The undersampling was done using a variable density undersampling mask that is different from the undersampling mask used for the SKM data (see **Figure 3-4**). The undersampling mask was applied to each 2D axial slice within the whole volume. The undersampling mask was applied on the reference data and Fourier transformed to obtain the undersampled images. The images used as reference for the image enhancement tasks were reconstructed using GRAPPA[82] (GeneRalized Autocalibrating Partial Parallel Acquisition). For the image enhancement task, the data was split into 9, 1 and 3 subjects for training, validation and testing, respectively.

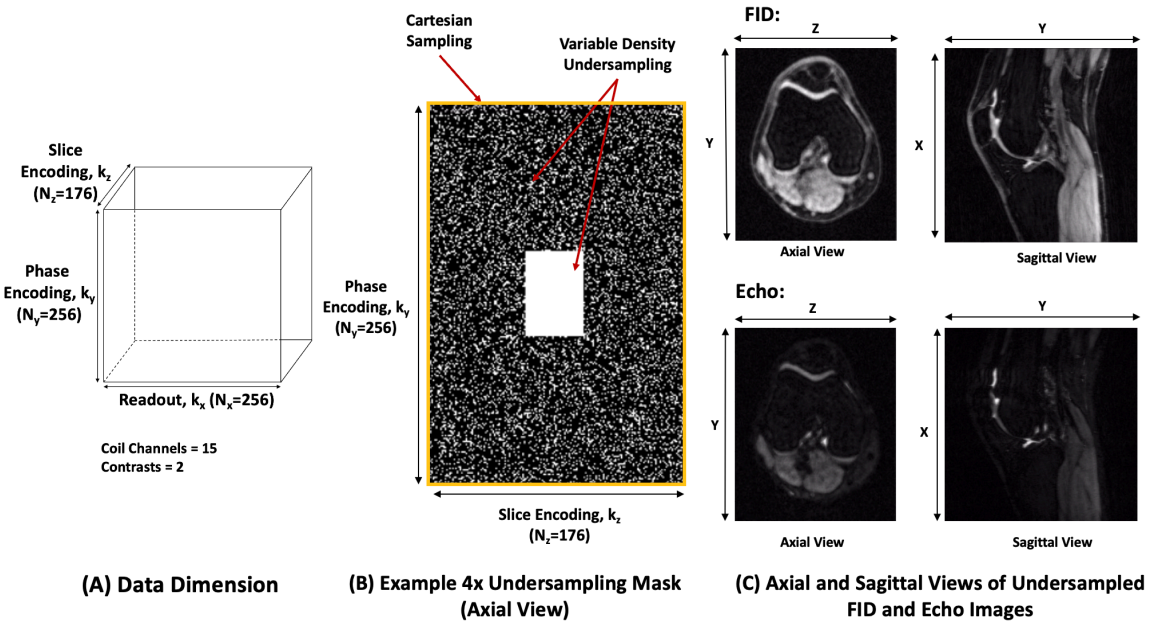

Figure 3-4: 3D isotropic DESS data details showing (A) the raw k-space data dimensions; (B) an example undersampling mask showing 4-fold variable density undersampling; (C) A single axial and sagittal slice showing the results of applying the undersampling mask to the entire volume.

qDESS Sequence Parameters for UCLA Dataset[34]	
Matrix (RO x PE)	256 x 256
Resolution (mm^2)	0.66 x 0.66
Slice thickness (mm)	0.66
Echo Time, TE – Echo 1(ms)	5.2
Number of Echoes	2
Repetition Time, TR (ms)	19
Flip Angle (°)	25
Parallel Imaging	2 x 1
Scan Time	7 min, 48 sec

Table 3-2: Sequence parameters used to acquire the 3D isotropic high-resolution knee DESS scans at the UCLA.

3.3 Transfer Learning

To enhance undersampled 3D isotropic high-resolution DESS MRI from the UCLA dataset, the U-Net and SwinIR transformer networks were trained using the weights from pretrained models. The pretraining was done using the SKM Dataset, with a training and validation data split of 65 and 10 subjects.

The SKM datasets had different matrix sizes and acquired resolution compared to the UCLA datasets (**Figure 3-5**). To account for these differences, the k-space data from SKM datasets were resized from 512 phase encoding lines to 256 phase encoding lines by extracting the central 256 phase encoding lines. The 160 sagittal slices were zero-padded in k-space to obtain 176 sagittal slices. Since enhancement was done on each 2D axial slice, the number of axial slices in the SKM dataset (512) did not need to be matched to the UCLA data dimension (256). The resized k-space dataset was undersampled and then transformed to the image domain and passed to the DL network with the FID and Echo contrasts stacked as two separate channels.

After pre-training using SKM data, the models were then fine-tuned for the UCLA datasets. While training the model on the UCLA data, 9 subjects were used for training, 1 for validation and 3 for testing. The data was passed to the network as 2D axial slices with FID and Echo stacked as 2 channels.

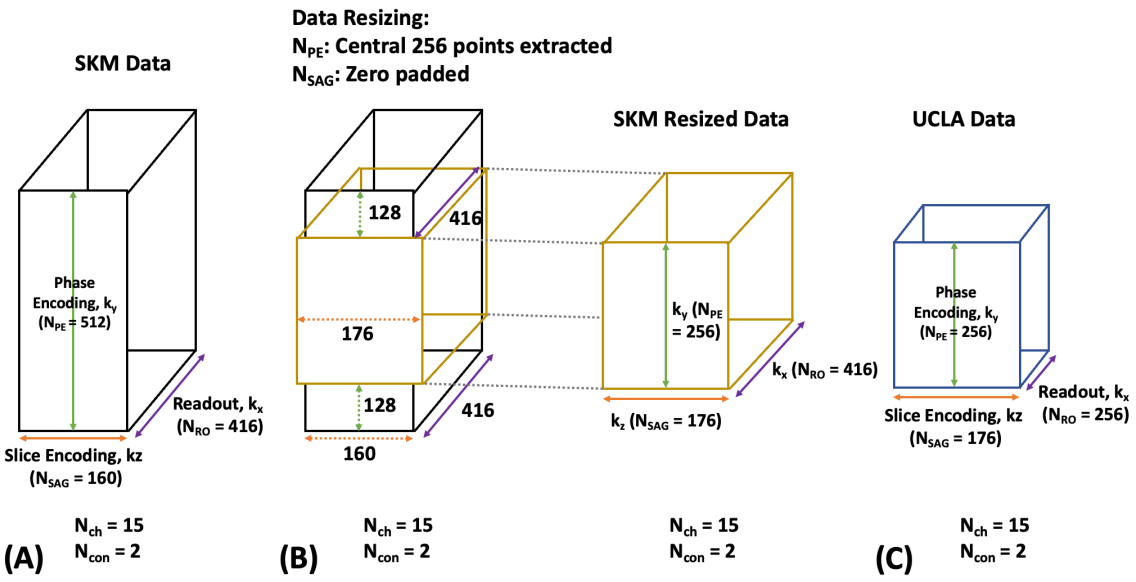

Figure 3-5: Stanford Knee dataset dimensions in (A) are transformed in (B) to make it similar to the UCLA DESS data dimensions in (C) to pretrain networks for transfer learning.

3.4 Analysis of Results

The peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) index were calculated to quantify enhanced image quality with respect to the reference GRAPPA reconstructed image. T_2 maps were analyzed using a Bland-Altman plot between the T_2 values within the cartilage of the reference image and the T_2 values within the cartilage of the enhanced image. The average T_2 value for the entire cartilage volume was measured for each subject, giving 20 data points for the SKM data.

The uncertainty maps were first visually compared with the difference images. Difference images were calculated as the absolute difference between the enhanced images and corresponding reference images. For the SKM data, the cartilage segmentation masks provided in the SKM dataset was used to extract the cartilage region in each subject and the average difference and uncertainty was calculated within the cartilage. This "region of interest" (ROI) approach allowed us to obtain a relationship between uncertainty values

and the enhancement error in knee cartilage. The Pearson R correlation score was calculated between the uncertainty and error values to assess the strength of the correlation.

For the UCLA dataset, it was difficult to manually perform cartilage segmentations over 176 sagittal slices of each subject. Therefore, for each subject, only one slice with maximum cartilage was segmented using a cartilage segmentation mask for that slice obtained using the OsiriX^{TM} based open-source medical image viewer, Horos (v3.3.6). The T_2 values within the cartilage (ROI approach) for the enhanced and reference images were displayed as boxplots for each of the 3 test subjects individually to assess the accuracy of T_2 maps.

Ablation studies were performed to measure the effectiveness of transfer learning. Undersampled images from the 3 test subjects in the UCLA dataset were passed through: (1) the model trained with transfer learning, (2) the model trained without transfer learning (model trained directly on UCLA data without pretrained weights), and (3) the model trained only on the Stanford dataset. The visual image quality was compared along with the PSNR and SSIM values of the enhanced images with respect to the reference images. The boxplots for the T_2 values within the cartilage was also shown for these models to assess the performance of transfer learning.

Chapter 4: Results and Discussion

4.1 Deep Learning Image Enhancement with Uncertainty Estimation

The total training time was 36 hours for the U-Net and 8 hours for the SwinIR, for each undersampling factor, using an NVIDIA A6000 GPU with 48 GB memory. The inference time was roughly 8 ms/slice for both networks. **Figure 4-1** and **Figure 4-2** show the U-Net and SwinIR results, respectively, for image enhancement of 4, 6 and 8-fold undersampled DESS knee MRI images. These results are from models trained and tested on the SKM dataset only. For each network, the orange boxes in the reference image, undersampled image and enhanced image show a zoomed-in patch of the sagittal slice with fine detail. We can clearly see that details obscured in the undersampled images have been recovered in the enhanced images, when compared with the corresponding reference images. **Table 4-1** shows the PSNR and SSIM values averaged over 20 subjects for each model under consideration. The SwinIR network seemed to perform slightly better than the U-Net for all 3 acceleration factors. Upon closer inspection of the enhanced image from each model in **Figures 4-1 and 4-2**, 4-fold and 6-fold undersampling results for the U-Net and SwinIR are comparable. The statistical difference between the PSNR and SSIM values can be analyzed in the future to estimate whether there is really a drop in performance from 4-fold to 6-fold undersampled image enhancement.

The drop in performance for the 8-fold undersampling results was expected since there is always a tradeoff between image quality and acceleration factor. Consequently, we observed a noticeable drop in the SSIM and PSNR values of the images enhanced by models trained on 8-fold undersampled images in **Table 4-1**.

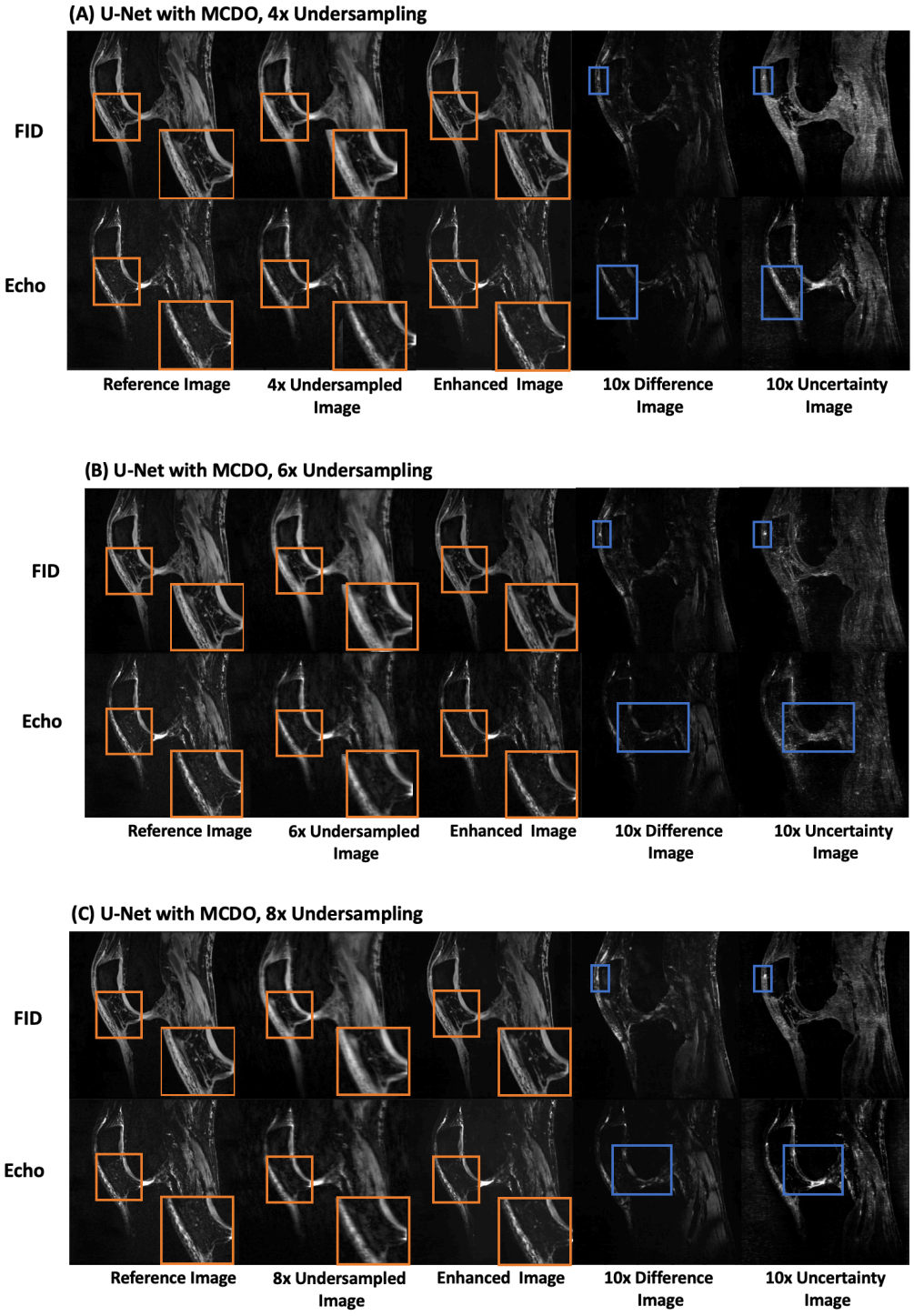

Figure 4-1: The image enhancement results using the U-Net with Monte Carlo Dropout architecture are shown for undersampling factors of (A) 4x, (B) 6x and (C) 8x. The corresponding difference image and uncertainty map are also shown.

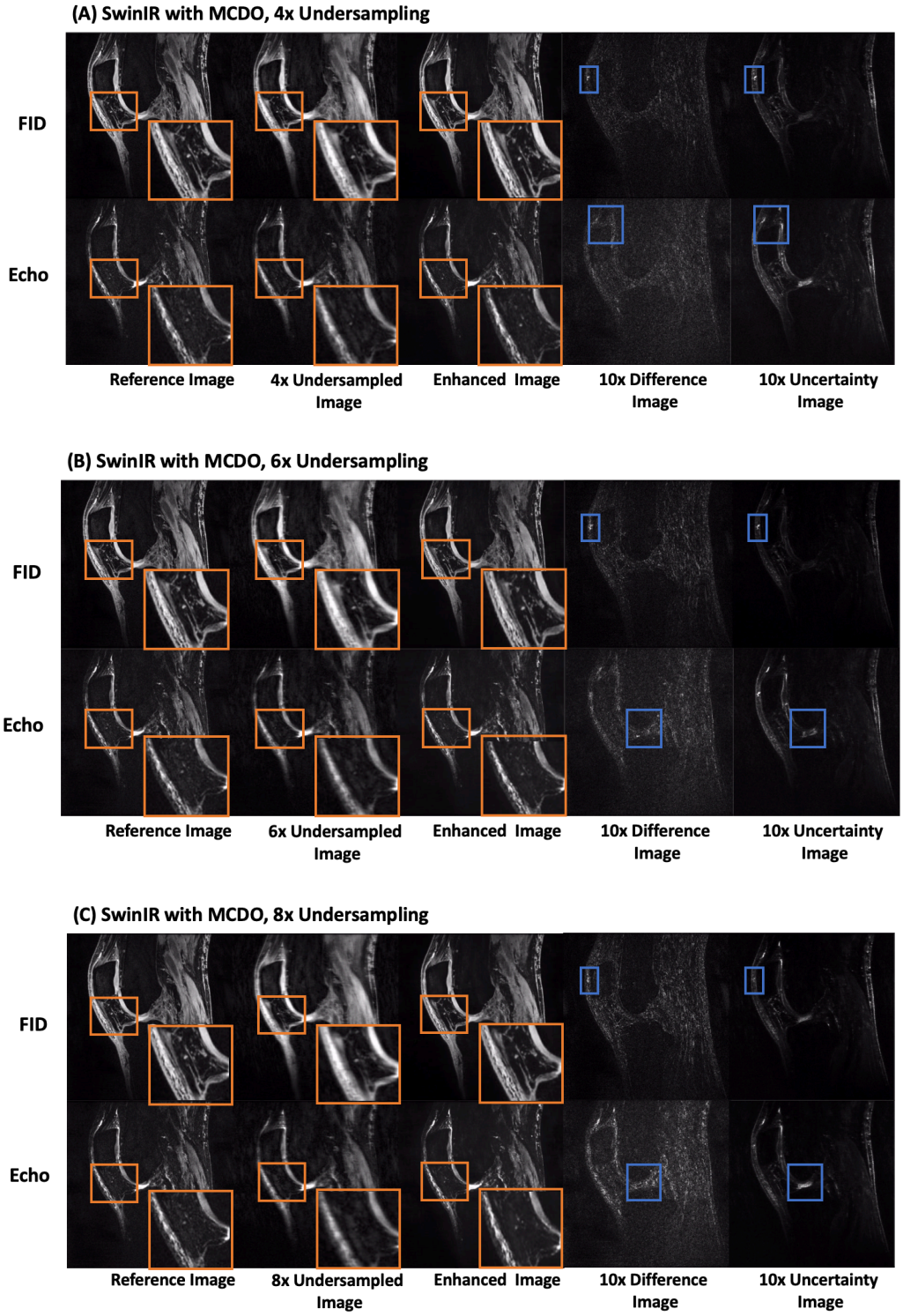

Figure 4-2: The image enhancement results using the SwinIR with Monte Carlo Dropout architecture are shown for undersampling factors of (A) 4x, (B) 6x and (C) 8x. The corresponding difference image and uncertainty map are also shown.

The difference images in **Figure 4-1** and **Figure 4-2** show the absolute difference between the enhanced and the reference images, dubbed as "difference image". These differences were small and were scaled 10 times for displaying purposes. **Figure 4-1** and **Figure 4-2** also show the uncertainty image. The uncertainty values were also scaled by 10 before displaying.

The difference image helps us visualize the error in enhancement for all the models under consideration. The uncertainty image helps us visualize the model (i.e., epistemic) uncertainty in different pixels of the enhanced image. It is observed that the areas with bright patches in the difference image correspond to bright patches in the same areas in the uncertainty image. Examples of these patches are marked by blue boxes in **Figure 4-1** and **Figure 4-2**. The elevated brightness in difference images is an indication of relatively higher enhancement error in the difference images and high uncertainty in the uncertainty image. Therefore, visually, pixels with higher enhancement error seem to have higher uncertainty compared to the rest of the image.

Model		DESS FID		DESS Echo	
		SSIM	PSNR	SSIM	PSNR
4x Acceleration	U-Net	0.913 ± 0.033	37.999 ± 2.394	0.922 ± 0.034	41.682 ± 2.291
	SwinIR	0.916 ± 0.031	38.671 ± 2.413	0.925 ± 0.032	42.043 ± 2.374
6x Acceleration	U-Net	0.888± 0.041	36.646 ± 2.196	0.900 ± 0.045	41.432 ± 2.280
	SwinIR	0.896 ± 0.039	37.292 ± 2.407	0.909 ± 0.039	41.154 ± 2.394
8x Acceleration	U-Net	0.879 ± 0.046	36.005 ± 2.537	0.895 ± 0.046	40.394 ± 2.382
	SwinIR	0.878 ± 0.042	36.336 ± 2.446	0.895 ± 0.041	40.413 ± 2.432

Table 4-1: The SSIM and PSNR values of the enhanced images with respect to the reference images for all the models

To analyze the relationship between the difference image (enhancement error) and the uncertainty maps in more depth, the cartilage region was segmented within each

volume and the average difference and uncertainty within the cartilage was compared over 20 test subjects, as explained previously. These 20 data points were plotted in a scatter plot for each model to visualize the trend (**Figure 4-3** and **Figure 4-4)**. The Pearson R correlation score was calculated between the error and uncertainty values and all models showed a strong and significant correlation between enhancement error and epistemic uncertainty for DESS FID and Echo contrasts.

The T_2 maps of the reference image and enhanced image were analyzed in the cartilage. **Figure 4-5** shows an example of the T_2 map around the region of the cartilage along with the difference between the T_2 values in the reference image and enhanced image. These difference images are indicative of the error in T_2 maps for each model. Using the ROI approach with the average T_2 value within the cartilage taken for 20 subjects, the Bland-Altman plots from all the models are shown in **Figure 4-6**. We observe that the U-Net and SwinIR had similar results for 4-fold undersampling. For higher undersampling factors, the mean difference of T_2 values between the enhanced and reference image was around the same for the U-Net. However, for SwinIR, we observed that the bias (mean difference) increased as the undersampling factor increased. At higher undersampling factors, there are expected to be higher undersampling artifacts. This might lead to more errors in estimating the T_2 values. Thus, there may be a tradeoff between acceleration factor and T_2 values.

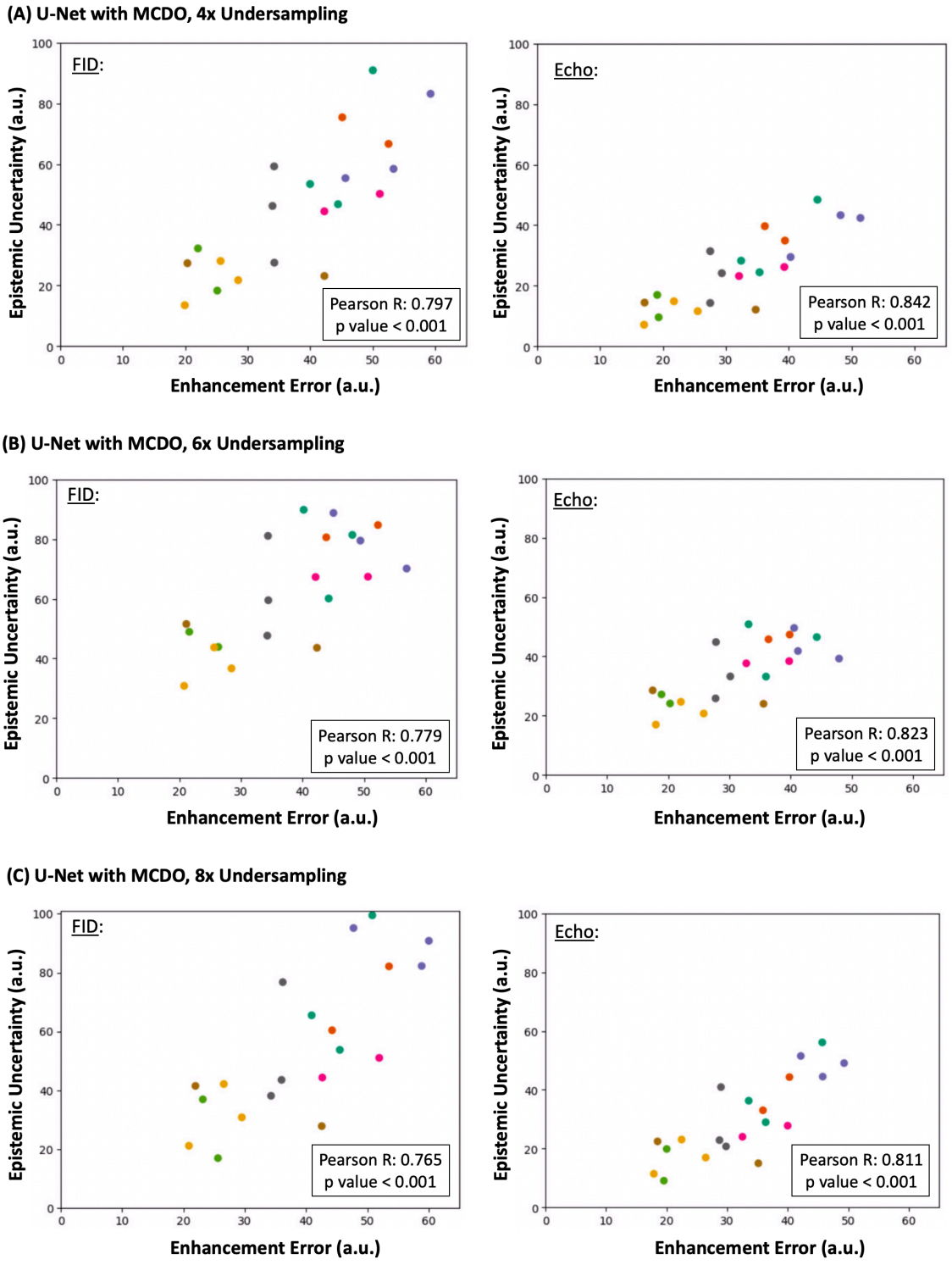

Figure 4-3: Scatter plots showing the correlation between enhancement error and epistemic uncertainty over 20 subjects for U-Net with (A) 4x, (B) 6x, (C) 8x acceleration.

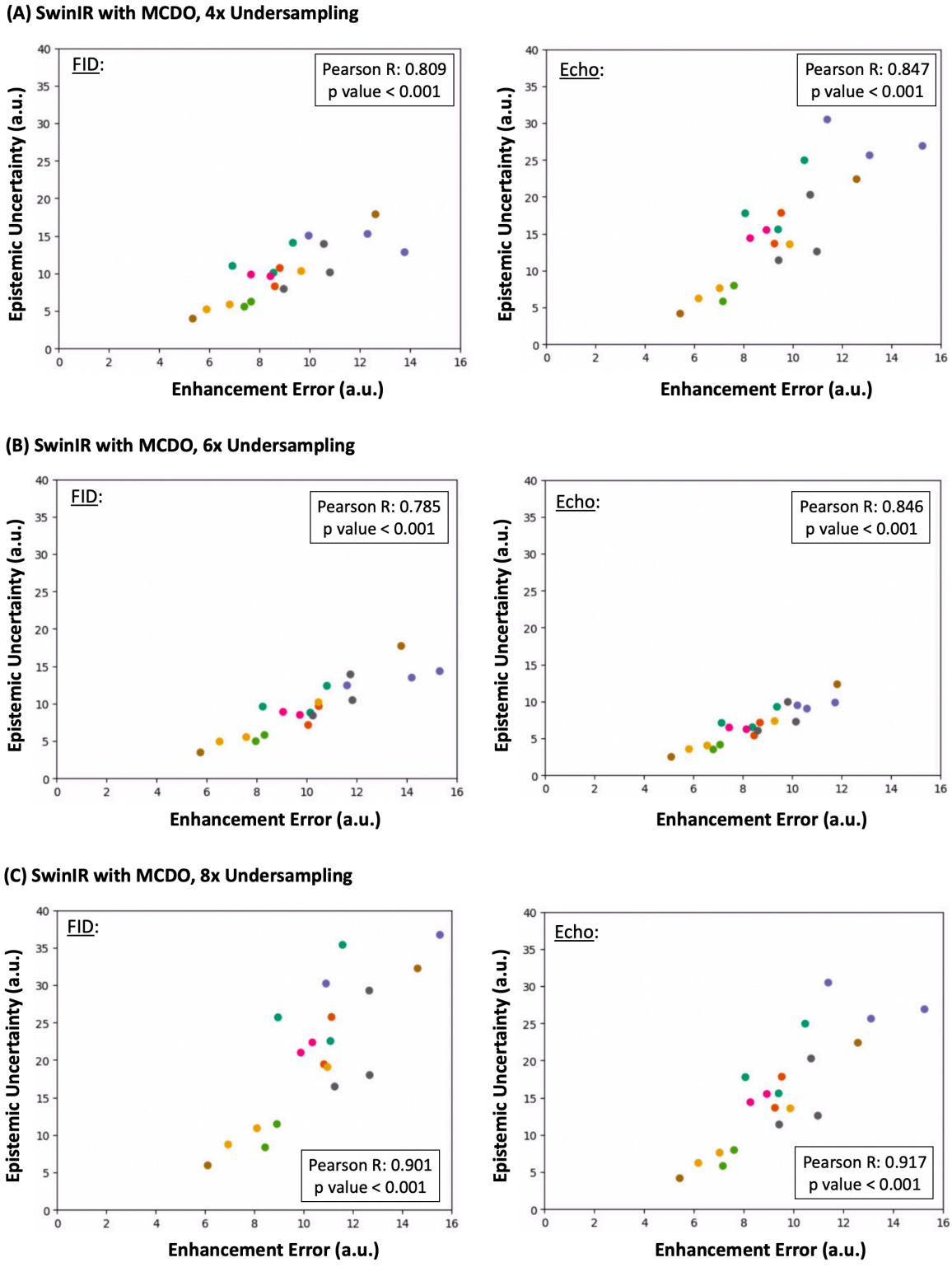

Figure 4-4: Scatter plots showing the correlation between enhancement error and epistemic uncertainty over 20 subjects for SwinIR with (A) 4x, (B) 6x, (C) 8x acceleration.

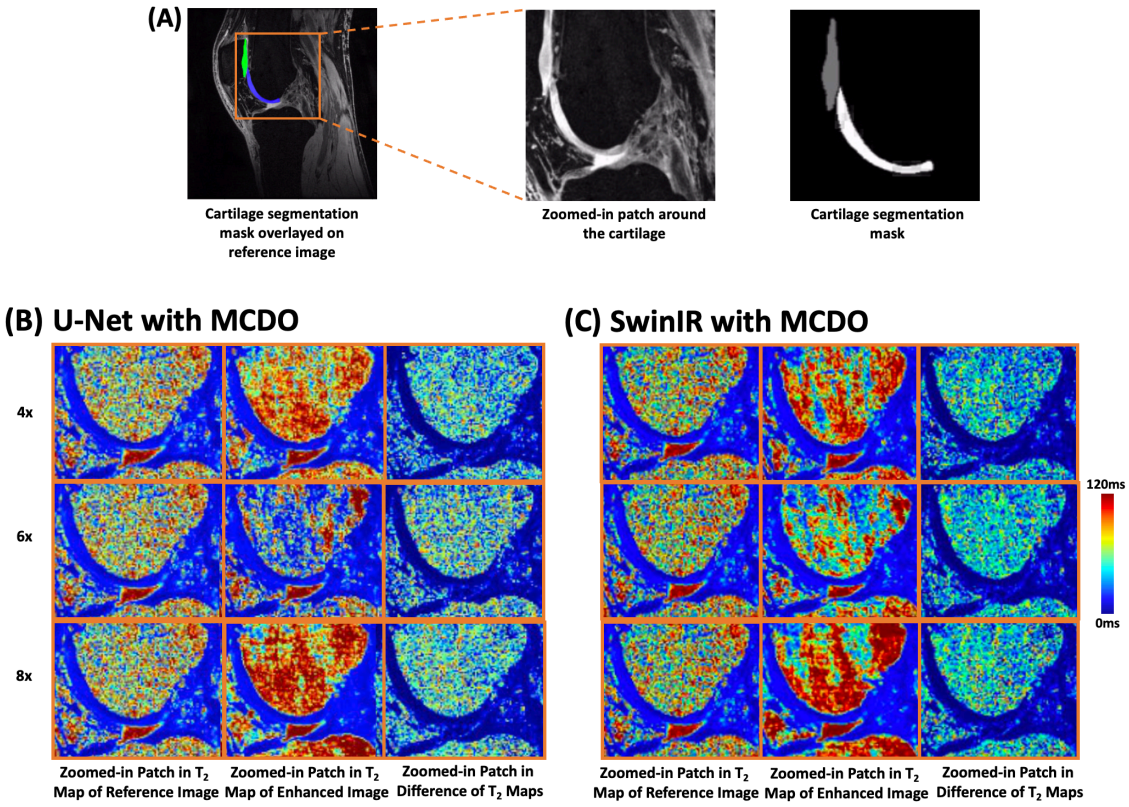

Figure 4-5: T_2 maps of the reference images, enhanced images from all models and the difference between the T_2 values of the reference and enhanced image in a single sagittal slice from the SKM dataset.

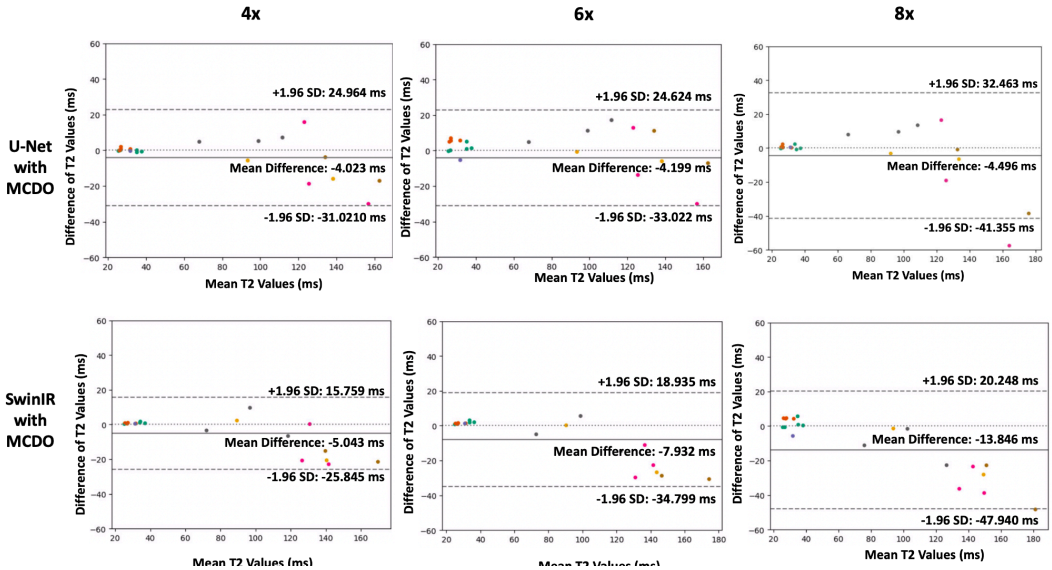

Figure 4-6: Bland-Altman plots comparing the T_2 values in the reference image and enhanced image from all models in the cartilage of 20 subjects (testing set) from the SKM dataset.

4.2 Transfer Learning

The results of the transfer learning approach using the U-Net are shown in **Figure 4-7** and results using the SwinIR are shown in **Figure 4-8**. The green boxes show regions within the FID and Echo images where the differences in the level of detail in each image is noticeable. Compared to the reference images, the images enhanced using the transfer learned model shows the most amount of detail recovered. The images enhanced by the model directly trained on the UCLA dataset without transfer learning capture similar level of details compared to the transfer learned model in the marked green patches. The images enhanced using the model trained only on the SKM dataset clearly failed to capture the minute details. These observations are justified by the SSIM and PSNR values in **Table 4--2**. The U-Net seemed to perform better than the SwinIR.

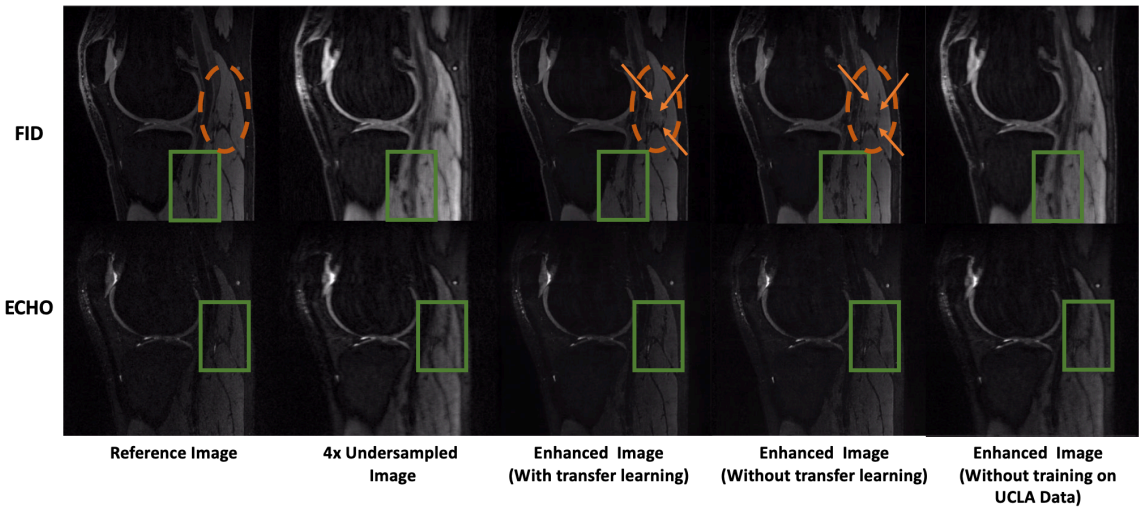

Figure 4-7: FID and Echo enhanced images from 4x undersampled data using the transfer learning model with U-Net architecture. The results were compared with the outputs from the model trained on UCLA data without transfer learning and the model trained only on the SKM data.

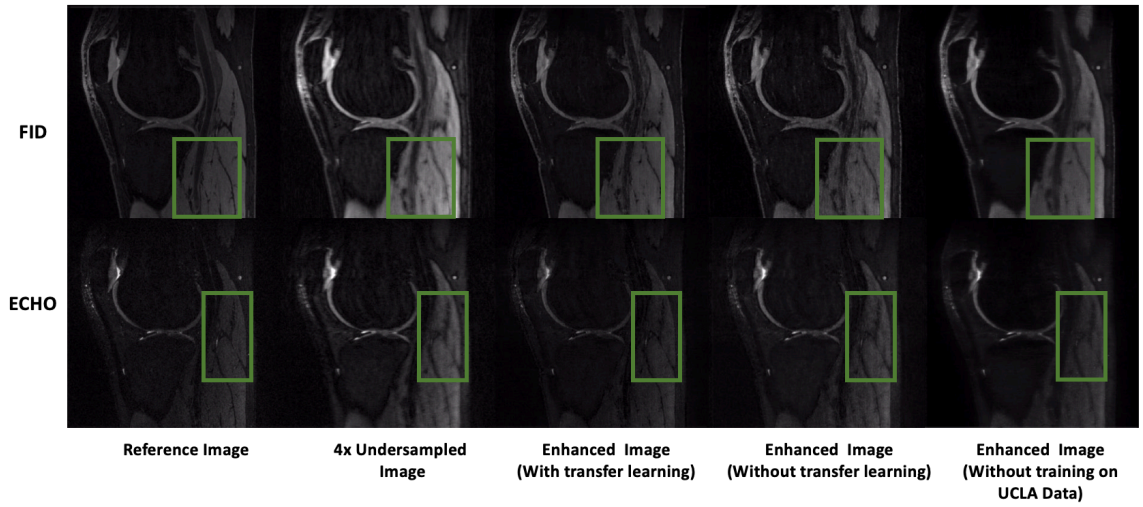

Figure 4-8: FID and Echo enhanced images from 4x undersampled data using the transfer learning model with U-Net architecture. The results were compared with the outputs from the model trained on UCLA data without transfer learning and the model trained only on the SKM data.

Model		DESS FID		DESS Echo	
		SSIM	PSNR	SSIM	PSNR
U-Net	With Transfer Learning	0.902 ± 0.101	35.55 ± 1.16	0.884 ± 0.010	37.70 ± 0.70
	Without Transfer Learning	0.883 ± 0.008	34.58 ± 1.00	0.873 ± 0.009	37.16 ± 0.67
	Trained on SKM Data	0.687 ± 0.026	22.62 ± 0.95	0.824 ± 0.014	31.24 ± 1.60
SwinIR	With Transfer Learning	0.877 ± 0.004	34.50 ± 0.92	0.848 ± 0.037	35.68 ± 1.84
	Without Transfer Learning	0.857 ± 0.009	35.56 ± 0.43	0.861 ± 0.021	36.41 ± 1.21
	Trained on SKM Data	0.748 ± 0.027	22.67 ± 1.13	0.769 ± 0.03	30.45 ± 1.24

Table 4-2: Image Quality of the Transfer Learning Results and Ablation Studies

The networks were trained using 2D axial slices. Comparing the regions circled in orange in the reference image, the image enhanced with transfer learning and the image enhanced without transfer learning, we can see that there are light horizontal lines present (pointed out by orange arrows in **Figure 4-7**) when transfer learning is not used. This means that although there is no inconsistency within a single enhanced axial slice, there

is some scaling inconsistency across different slices. Although transfer learning has helped overcome this issue for the U-Net, other methods might be more suitable to eradicate such errors. One such method might be to use 3D networks instead of 2D networks. 3D networks would allow the model to capture information across all slices and that may eliminate the inconsistency we observe over different slices. This might have also shown up in the model trained without transfer learning due to the limited training dataset size. Data augmentation provides a solution to having limited training datasets and can be explored in the future. Thus, going forward, incorporating data augmentation and the use of a 3D model, the performance of the model trained without transfer learning might be at par with the performance of the model trained with transfer learning.

Figure 4-9 shows the T_2 maps from the results of the transfer learning approach using the U-Net and the SwinIR networks. The images show zoomed-in patches around the cartilage as the we want to observe T_2 values in the cartilage mainly. For the Bland-Altman analysis, an ROI approach over the 3 test subjects was done, similar to the previous section. However, we did not have segmentation masks for the entire volume and thus analyzed the cartilage in the central slice using manual cartilage segmentation masks. **Figure 4-9(D)** shows the boxplots for the T_2 values in the cartilage of each subject for the models trained with and without transfer learning, compared to the T_2 cartilage values in the reference image of the respective subject. We observed that the average T_2 value in the cartilage dropped slightly compared to the reference T_2 when the images were enhanced using the transfer learning approach. The T_2 values dropped further without transfer learning for the U-Net and was much higher than the reference T_2 for

the SwinIR. There is a significant drop in T_2 values for the images enhanced using the model trained on the SKM dataset. This shows that the T_2 values in the cartilage are closest to the reference when the model is trained with transfer learning, compared to the models trained without transfer learning.

Regardless, the transfer learning results can be further improved. Data augmentation is a popular technique used to increase the number of datasets, and help generalize the model. The uncertainty maps estimated using MCDO in the previous section can be incorporated in the model training process during the validation step to improve results in the future.

Overall, for transfer learning, the U-Net performed better than the SwinIR. Comparing the image quality of the enhanced images from the two networks, the U-Net showed more promise. While training the models, it was observed that the SwinIR network is extremely sensitive to slight changes in hyperparameters. This made it extremely hard to train the model and find the optimal hyperparameters. Small changes in the learning rate and number of epochs specifically, widely changed the enhancement results. The enhanced images showed artifacts arising from scaling inconsistencies in the axial slices for a set of hyperparameters and showed blurring and elimination of details when the hyperparameters were tweaked a little. Going forward, a more extensive search of the hyperparameter space might improve the enhancement results using the SwinIR.

Lastly, the input data used to train and test all models was in the image domain. In the future, unrolled networks[82] can be considered, where the k-space data is directly used as the input to the models for training and testing. Unrolled networks may improve the performance compared to the results in this work as they incorporate imaging physics

details by iteratively optimizing k-space inconsistencies within each pixel.

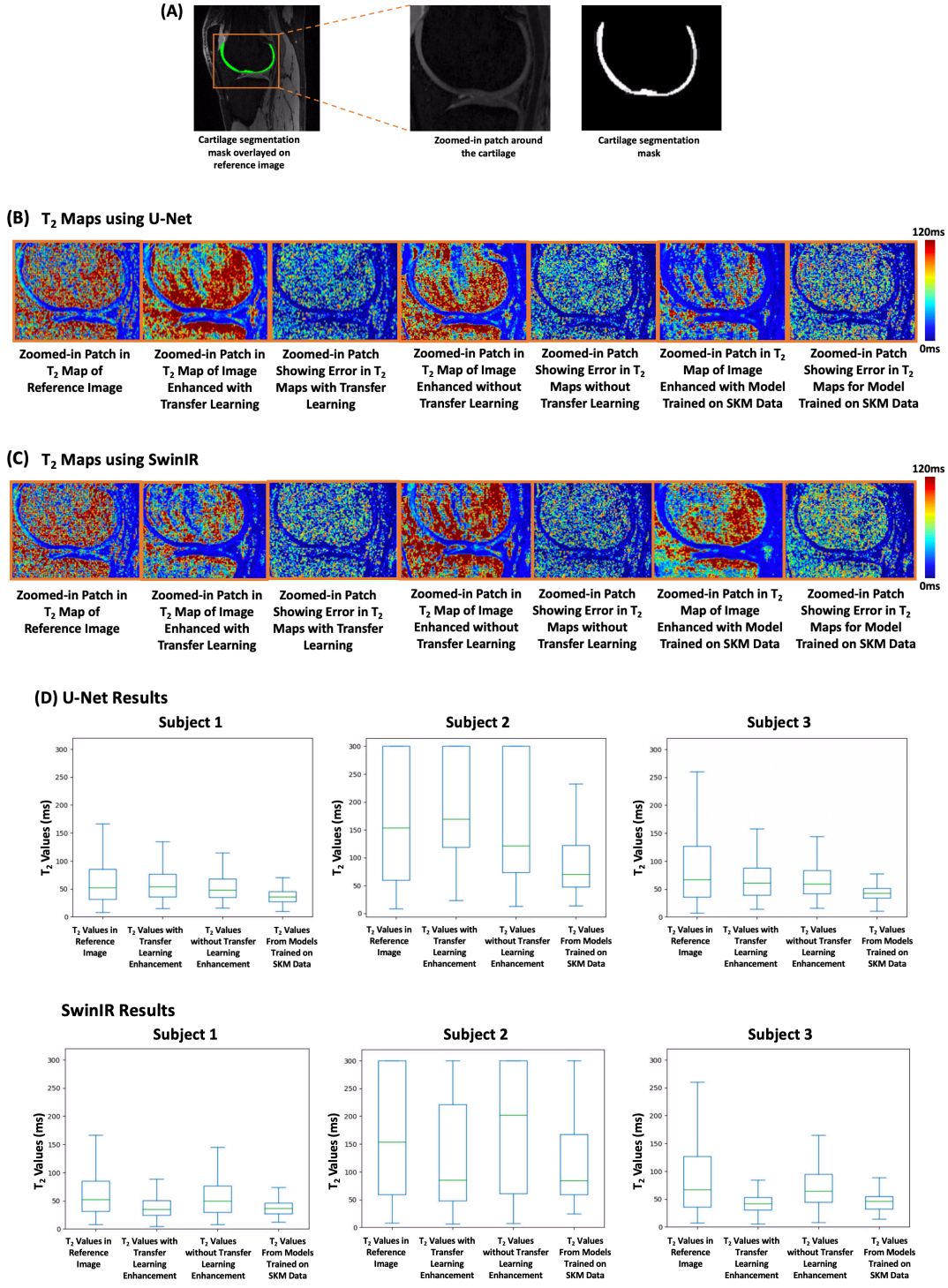

Figure 4-9: (A) Region around cartilage zoomed in and the cartilage segmentation mask. (B) and (C) are the T_2 results from the U-Net and SwinIR respectively. (D) Boxplots of T_2 values in the cartilage of each subject of U-Net and SwinIR networks (with transfer learning and ablation studies)

Chapter 5: Summary and Future Directions

5.1 Summary of Main Finding

This study investigated DL-based image enhancement with epistemic uncertainty measurements using MCDO. The enhanced images for 4-fold, 6-fold and 8-fold retrospectively accelerated scans showed acceptable image quality (SSIM above 0.875 for all models and PSNR above 35 for all models). These metrics, paired with the difference images, show that the undersampled images were successfully enhanced and recovered a certain amount of detail accurately. The Bland-Altman plots showed low bias (mean difference) for T_2 values within the cartilage of the enhanced images compared to reference images.

The transfer learning approach also showed acceptable image quality for 4-fold retrospectively accelerated scans for isotropic high-resolution 3D DESS imaging using U-Net and SwinIR architectures, with the U-Net model performing better than the SwinIR model, for image enhancement and T_2 mapping.

The epistemic, or model, uncertainty, was found to show significant correlation with the image enhancement error in the cartilage ROIs using the U-Net with MCDO network and the SwinIR with MCDO network, for three acceleration factors. This shows that the epistemic uncertainty quantified in this work has the potential to provide model confidence in the cartilage of DL-based undersampled knee DESS enhancement.

Thus, DL-based image enhancement is a promising solution to overcoming the challenges of accelerating isotropic high-resolution 3D DESS knee MRI scans.

5.2 Future Work

It should be noted that while these results on retrospectively undersampled data are promising, translating this work to prospectively undersampled data may have very different results and should be investigated separately.

While the DL-based enhancement of undersampled knee DESS MRI using the U-Net and SwinIR architectures have promising results as seen in **Figure 4-1**, **Figure 4-2** and **Table 3**, implementing DL networks comes with its own set of limitations, some of which have not been fully addressed in this work and need to be investigated in the future. The U-Net and SwinIR networks may sometimes overfit the training data even though a validation set is used to observe the validation loss and choose the best model. Data augmentation techniques can help reduce the overfitting problem. Another solution would be to use ensembling methods, where the information from multiple networks is combined to achieve the best possible results. Going forward, other network architectures can be explored and compared to the results obtained using U-Net and SwinIR.

Although DL methods have advantages over CS reconstruction in terms of inference time, supervised DL methods require a large amount of data for training. In this case, the convergence process becomes very computationally intensive and memory expensive. Model compression and optimization techniques like network pruning, quantization, or knowledge distillation can be applied to reduce the model's memory requirements and improve its efficiency. These techniques would make it more practical to construct and train DL-models.

All epistemic uncertainty estimates in this work were obtained using Monte Carlo Dropout (MCDO) method. Going forward, this can be extended to other, more complex

methods of calculating the epistemic uncertainty such as using generative adversarial networks (GANs) and variational autoencoders[70-71], as MCDO has multiple limitations. MCDO comes with a high computational overhead. Applying MCDO involves multiple forward passes through the network with dropout activated. This process can be computationally expensive, especially for large networks or high-resolution MRI. The increased computational overhead can limit the network inferencing time as well. MCDO introduces additional hyperparameters that need to be tuned, such as the dropout rate and the number of Monte Carlo samples. Selecting appropriate values for these hyperparameters can impact the quality and reliability of uncertainty estimation. Finding the optimal hyperparameter settings may require extensive experimentation and validation.

This work only explored epistemic uncertainty or model uncertainty. Other types of uncertainty, such as "aleatoric" uncertainty or "data" uncertainty should be incorporated as well to provide a complete picture of the uncertainty in the deep learning-based image enhancement results. Further, radiologists' assessment can be performed for the enhanced images as well as the uncertainty maps. Radiologists' scoring may be incorporated in the future to assess how helpful it would be to analyze the uncertainty maps to estimate the model's confidence in enhancement and whether the maps can be reliable in a clinical setting.

The uncertainty maps might also be able to help with the image enhancement task in more complex models, leading to developing semi-supervised or unsupervised models for MRI image enhancement from undersampled data for protocols with limited datasets. This remains to be explored in the future.

It can be noted that majority of the analysis for the results presented in this study were reported for the cartilage region as the T_2 values within the cartilage is monitored to estimate the progression of OA in the knee. This analysis can be extended to other tissues and body parts as well.

5.3 Conclusion

The 3D isotropic high-resolution knee DESS MRI scans can be accelerated by using undersampled acquisitions and incorporating deep learning-based image enhancement to improve the image quality. Transfer learning is a promising solution to enhancement tasks for limited datasets.

The epistemic uncertainty maps show a strong correlation with the error in enhancement. Thus, these maps have the potential to convey the level of model confidence in the enhanced images.

References

[1] Chien, A., Weaver, J. S., Kinne, E., & Omar, I. (2020). Magnetic resonance imaging of the knee. Polish Journal of Radiology, 85(1), e509–e531.

[2] Hunter, D. J., Zhang, W., Conaghan, P. G., Hirko, K., Menashe, L., Reichmann, W. M., & Losina, E. (2011). Responsiveness and reliability of MRI in knee osteoarthritis: A meta-analysis of published evidence. Osteoarthritis and Cartilage, 19(5), 589–605.

[3] Link, T. M., Steinbach, L. S., Ghosh, S., Ries, M., Lu, Y., Lane, N., & Majumdar, S. (2003). Osteoarthritis: MR imaging findings in different stages of disease and correlation with clinical findings. Radiology, 226(2), 373–381.

[4] Nacey, N. C., Geeslin, M. G., Miller, G. W., & Pierce, J. L. (2017). Magnetic resonance imaging of the knee: An overview and update of conventional and state of the art imaging. Journal of Magnetic Resonance Imaging, 45(5), 1257–1275.

[5] Peterfy, C. G., Schneider, E., & Nevitt, M. (2008). The osteoarthritis initiative: report on the design rationale for the magnetic resonance imaging protocol for the knee. In Osteoarthritis and Cartilage (Vol. 16, Issue 12, pp. 1433–1441).

[6] Roemer, F. W., Crema, M. D., Trattnig, S., & Guermazi, A. (2011). Advances in imaging of osteoarthritis and cartilage. In Radiology (Vol. 260, Issue 2, pp. 332–354).

[7] Winters, K., & Tregonning, R. (2005). Reliability of magnetic resonance imaging of the traumatic knee as determined by arthroscopy. The New Zealand Medical Journal, 118(1209), U1301.

[8] Guermazi, A., Roemer, F. W., & Hayashi, D. (2011). Imaging of osteoarthritis. Current Opinion in Rheumatology, 23(5), 484–491.

[9] Surowiec, R. K., Lucas, E. P., & Ho, C. P. (2014). Quantitative MRI in the evaluation

of articular cartilage health: reproducibility and variability with a focus on T_2 mapping. Knee Surgery, Sports Traumatology, Arthroscopy, 22(6), 1385–1395.

[10] David-Vaudey, E., Ghosh, S., Ries, M., & Majumdar, S. (2004). T_2 relaxation time measurements in osteoarthritis. Magnetic Resonance Imaging, 22(5), 673–682.

[11] Mosher, T. J., Liu, Y., Yang, Q. X., Yao, J., Smith, R., Dardzinski, B. J., & Smith, M. B. (2004). Age dependency of cartilage magnetic resonance imaging T_2 relaxation times in asymptomatic women. Arthritis & Rheumatism, 50(9), 2820–2828.

[12] Li, X., Cheng, J., Lin, K., Saadat, E., Bolbos, R. I., Jobke, B., Ries, M. D., Horvai, A., Link, T. M., & Majumdar, S. (2011). Quantitative MRI using $T_1\rho$ and T_2 in human osteoarthritic cartilage specimens: correlation with biochemical measurements and histology. Magnetic Resonance Imaging, 29(3), 324–334.

[13] Eckstein, F., Hudelmaier, M., Wirth, W., Kiefer, B., Jackson, R., Yu, J., Eaton, C. B., & Schneider, E. (2006). Double echo steady state magnetic resonance imaging of knee articular cartilage at 3 Tesla: A pilot study for the Osteoarthritis Initiative. Annals of the Rheumatic Diseases, 65(4), 433–441.

[14] Bruder, H., Fischer, H., Graumann, R., & Deimling, M. (1988). A new steady-state imaging sequence for simultaneous acquisition of two MR images with clearly different contrasts. Magnetic Resonance in Medicine, 7(1), 35–42.

[15] Dregely, I., Margolis, D. A. J., Sung, K., Zhou, Z., Rangwala, N., Raman, S. S., & Wu, H. H. (2016). Rapid quantitative T_2 mapping of the prostate using three-dimensional dual echo steady state MRI at 3T. Magnetic Resonance in Medicine, 76(6), 1720–1729.

[16] Welsch, G. H., Scheffler, K., Mamisch, T. C., Hughes, T., Millington, S., Deimling, M., & Trattnig, S. (2009). Rapid estimation of cartilage T_2 based on double echo at steady

state (DESS) with 3 Tesla. Magnetic Resonance in Medicine, 62(2), 544–549.

[17] Sveinsson, B., Chaudhari, A. S., Gold, G. E., & Hargreaves, B. A. (2017). A simple analytic method for estimating T_2 in the knee from DESS. Magnetic Resonance Imaging, 38, 63–70.

[18] Moriya, S., Miki, Y., Kanagaki, M., Matsuno, Y., & Miyati, T. (2014a). 90°-Flip-angle three-dimensional double-echo steady-state (3D-DESS) magnetic resonance imaging of the knee: Isovoxel cartilage imaging at 3 T. European Journal of Radiology, 83(8), 1429–1432.

[19] Ruehm, S., Zanetti, M., Romero, J., & Hodler, J. (1998). MRI of patellar articular cartilage: Evaluation of an optimized gradient-echo sequence (3D-DESS). Journal of Magnetic Resonance Imaging, 8(6), 1246–1251.

[20] Scheffler, K. (1999). A pictorial description of steady-states in rapid magnetic resonance imaging. Concepts in Magnetic Resonance, 11(5), 291–304.

[21] Shapiro, L. M., McWalter, E. J., Son, M.-S., Levenston, M., Hargreaves, B. A., & Gold, G. E. (2014). Mechanisms of osteoarthritis in the knee: MR imaging appearance. Journal of Magnetic Resonance Imaging, 39(6), 1346–1356.

[22] Van Dyck, P., Vanhevel, F., Vanhoenacker, F. M., Wouters, K., Grodzki, D. M., Gielen, J. L., & Parizel, P. M. (2015). Morphological MR imaging of the articular cartilage of the knee at 3 T—comparison of standard and novel 3D sequences. Insights into Imaging, 6(3), 285–293.

[23] Verstraete, K. L., Almqvist, F., Verdonk, P., Vanderschueren, G., Huysse, W., Verdonk, R., & Verbrugge, G. (2004). Magnetic resonance imaging of cartilage and cartilage repair. Clinical Radiology, 59(8), 674–689.

[24] Braun, H. J., & Gold, G. E. (2012). Diagnosis of osteoarthritis: Imaging. Bone, 51(2), 278–288.

[25] Hargreaves, B. A., Bangerter, N. K., Shimakawa, A., Vasanawala, S. S., Brittain, J. H., & Nishimura, D. G. (2006). Dual-acquisition phase-sensitive fat–water separation using balanced steady-state free precession. Magnetic Resonance Imaging, 24(2), 113–122.

[26] Imhof, H., Nöbauer-Huhmann, I.-M., Krestan, C., Gahleitner, A., Sulzbacher, I., Marlovits, S., & Trattnig, S. (2002). MRI of the cartilage. European Radiology, 12(11), 2781–2793.

[27] Keenan, K. E., Biller, J. R., Delfino, J. G., Boss, M. A., Does, M. D., Evelhoch, J. L., Griswold, M. A., Gunter, J. L., Hinks, R. S., Hoffman, S. W., Kim, G., Lattanzi, R., Li, X., Marinelli, L., Metzger, G. J., Mukherjee, P., Nordstrom, R. J., Peskin, A. P., Perez, E., … Sullivan, D. C. (2019). Recommendations towards standards for quantitative MRI (qMRI) and outstanding needs. Journal of Magnetic Resonance Imaging, 49(7), e26–e39.

[28] Kornaat, P. R., Reeder, S. B., Koo, S., Brittain, J. H., Yu, H., Andriacchi, T. P., & Gold, G. E. (2005). MR imaging of articular cartilage at 1.5T and 3.0T: Comparison of SPGR and SSFP sequences. Osteoarthritis and Cartilage, 13(4), 338–344.

[29] Donoho, D. L. (2006). Compressed sensing. IEEE Transactions on Information Theory, 52(4), 1289–1306.

[30] Lustig, M., Donoho, D. L., Santos, J. M., & Pauly, J. M. (2008). Compressed Sensing MRI. IEEE Signal Processing Magazine, 25(2), 72–82.

[31] Lustig, M., Donoho, D., & Pauly, J. M. (2007). Sparse MRI: The application of compressed sensing for rapid MR imaging. Magnetic Resonance in Medicine, 58(6), 1182–

1195.

[32] Hyun, C. M., Kim, H. P., Lee, S. M., Lee, S., & Seo, J. K. (2018). Deep learning for undersampled MRI reconstruction. Physics in Medicine and Biology, 63(13).

[33] Schlemper, J., Castro, D. C., Bai, W., Qin, C., Oktay, O., Duan, J., Price, A. N., Hajnal, J., & Rueckert, D. (2018). Bayesian Deep Learning for Accelerated MR Image Reconstruction (pp. 64–71).

[34] Shih, S., Zhang Z., Deb A., Zhong X., Ryan T.W., Wu H.H. (2023). Rapid Isotropic 3D T_2 Mapping of the Knee using Dual-Echo Steady-State MRI with Compressed Sensing Reconstruction. International Society for Magnetic Resonance in Medicine.

[35] Zhang, Y. N., & An, M. Q. (2017). Deep Learning- and Transfer Learning-Based Super Resolution Reconstruction from Single Medical Image. Journal of Healthcare Engineering, 2017.

[36] Wang, S., Su, Z., Ying, L., Peng, X., Zhu, S., Liang, F., Feng, D., & Liang, D. (2016). Accelerating magnetic resonance imaging via deep learning. Proceedings - International Symposium on Biomedical Imaging, 2016-June, 514–517.

[37] Arshad, M., Qureshi, M., Inam, O., & Omer, H. (2021). Transfer learning in deep neural network based under-sampled MR image reconstruction. Magnetic Resonance Imaging, 76, 96–107.

[38] Desai, A. D., Schmidt, A. M., Rubin, E. B., Sandino, C. M., Black, M. S., Mazzoli, V., Stevens, K. J., Boutin, R., Ré, C., Gold, G. E., Hargreaves, B. A., & Chaudhari, A. S. (n.d.). SKM-TEA: A Dataset for Accelerated MRI Reconstruction with Dense Image Labels for Quantitative Clinical Evaluation.

[39] Otazo, R., Kim, D., Axel, L., & Sodickson, D. K. (2010). Combination of compressed

sensing and parallel imaging for highly accelerated first-pass cardiac perfusion MRI. Magnetic Resonance in Medicine, 64(3), 767–776.

[40] Knoll, F., Hammernik, K., Zhang, C., Moeller, S., Pock, T., Sodickson, D. K., & Akcakaya, M. (2020). Deep-Learning Methods for Parallel Magnetic Resonance Imaging Reconstruction: A Survey of the Current Approaches, Trends, and Issues. IEEE Signal Processing Magazine, 37(1), 128–140.

[41] Candes, E. J., Romberg, J., & Tao, T. (2006). Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE Transactions on Information Theory, 52(2), 489–509.

[42] Islam, R., Islam, M. S., & Uddin, M. S. (2022). Compressed Sensing in Parallel MRI: A Review. International Journal of Image and Graphics, 22(04).

[43] Qu, X., Hou, Y., Lam, F., Guo, D., Zhong, J., & Chen, Z. (2014). Magnetic resonance image reconstruction from undersampled measurements using a patch-based nonlocal operator. Medical Image Analysis, 18(6), 843–856.

[44] Ravishankar, S., & Bresler, Y. (2011). MR Image Reconstruction From Highly Undersampled k-Space Data by Dictionary Learning. IEEE Transactions on Medical Imaging, 30(5), 1028–1041.

[45] Ravishankar, S., & Bresler, Y. (2013). Sparsifying transform learning for Compressed Sensing MRI. 2013 IEEE 10th International Symposium on Biomedical Imaging, 17–20.

[46] Pal, A., & Rathi, Y. (2021). A review and experimental evaluation of deep learning methods for MRI reconstruction.

[47] Pezzotti, N., Yousefi, S., Elmahdy, M. S., van Gemert, J. H. F., Schuelke, C., Doneva, M., Nielsen, T., Kastryulin, S., Lelieveldt, B. P. F., van Osch, M. J. P., Weerdt, E. D. E.,

& Staring, M. (2020). An adaptive intelligence algorithm for undersampled knee MRI reconstruction. IEEE Access, 8, 204825–204838.

[48] Qiu, D., Zhang, S., Liu, Y., Zhu, J., & Zheng, L. (2020). Super-resolution reconstruction of knee magnetic resonance imaging based on deep learning. Computer Methods and Programs in Biomedicine, 187.

[49] Zeng, G., Guo, Y., Zhan, J., Wang, Z., Lai, Z., Du, X., Qu, X., & Guo, D. (2021). A review on deep learning MRI reconstruction without fully sampled k-space. In BMC Medical Imaging (Vol. 21, Issue 1). BioMed Central Ltd.

[50] Chaudhari, A. S., Fang, Z., Kogan, F., Wood, J., Stevens, K. J., Gibbons, E. K., Lee, J. H., Gold, G. E., & Hargreaves, B. A. (2018). Super-resolution musculoskeletal MRI using deep learning. Magnetic Resonance in Medicine, 80(5), 2139–2154.

[51] Hammernik, K., Klatzer, T., Kobler, E., Recht, M. P., Sodickson, D. K., Pock, T., & Knoll, F. (2018). Learning a variational network for reconstruction of accelerated MRI data. Magnetic Resonance in Medicine, 79(6), 3055–3071.

[52] Mardani, M., Giannakis, G. B., & Ugurbil, K. (2016). Tracking Tensor Subspaces with Informative Random Sampling for Real-Time MR Imaging.

[53] Hossain, Md. B., Kwon, K.-C., Imtiaz, S. M., Nam, O.-S., Jeon, S.-H., & Kim, N. (2022). De-Aliasing and Accelerated Sparse Magnetic Resonance Image Reconstruction Using Fully Dense CNN with Attention Gates. Bioengineering, 10(1), 22.

[54] Schlemper, J., Caballero, J., Hajnal, J. V., Price, A., & Rueckert, D. (2017). A Deep Cascade of Convolutional Neural Networks for MR Image Reconstruction (pp. 647–658).

[55] Jin, K. H., McCann, M. T., Froustey, E., & Unser, M. (2017). Deep Convolutional Neural Network for Inverse Problems in Imaging. IEEE Transactions on Image Processing,

26(9), 4509–4522.

[56] Chen, Y., Shi, F., Christodoulou, A. G., Xie, Y., Zhou, Z., & Li, D. (2018). Efficient and Accurate MRI Super-Resolution Using a Generative Adversarial Network and 3D Multi-level Densely Connected Network (pp. 91–99).

[57] Yang, G., Yu, S., Dong, H., Slabaugh, G., Dragotti, P. L., Ye, X., Liu, F., Arridge, S., Keegan, J., Guo, Y., & Firmin, D. (2018). DAGAN: Deep De-Aliasing Generative Adversarial Networks for Fast Compressed Sensing MRI Reconstruction. IEEE Transactions on Medical Imaging, 37(6), 1310–1321.

[58] Kokkotis, C., Moustakidis, S., Papageorgiou, E., Giakas, G., & Tsaopoulos, D. E. (2020). Machine learning in knee osteoarthritis: A review. Osteoarthritis and Cartilage Open, 2(3), 100069.

[59] Cun, L., Henderson, J., Le Cun, Y., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (n.d.). Handwritten Digit Recognition with a Back-Propagation Network.

[60] Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278–2324.

[61] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation (pp. 234–241).

[62] Dixon, A. K., Keene, G. S., Hollingworth, W., Lomas, D. J., & Villar, R. N. (1996). Magnetic Resonance Imaging of the Knee: Assessment of Effectiveness. In Clinical Radiology (Vol. 51).

[63] Ghodrati, V., Shao, J., Bydder, M., Zhou, Z., Yin, W., Nguyen, K.-L., Yang, Y., & Hu, P. (2019). MR image reconstruction using deep learning: evaluation of network

structure and loss functions. Quantitative Imaging in Medicine and Surgery, 9(9), 1516–1527.

[64] Cole, E., Cheng, J., Pauly, J., & Vasanawala, S. (2021). Analysis of deep complex-valued convolutional neural networks for MRI reconstruction and phase-focused applications. Magnetic Resonance in Medicine, 86(2), 1093–1109.

[65] Mackenzie, R., Dixon, A. K., Keene, G. S., Hollingworth, W., Lomas, D. J., & Villar, R. N. (1996). Magnetic resonance imaging of the knee: Assessment of effectiveness. Clinical Radiology, 51(4), 245–250.

[66] Vaswani, A., Brain, G., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (n.d.). Attention Is All You Need.

[67] Liang, J., Cao, J., Sun, G., Zhang, K., Van Gool, L., & Timofte, R. (n.d.). SwinIR: Image Restoration Using Swin Transformer.

[68] Huang, J., Fang, Y., Wu, Y., Wu, H., Gao, Z., Li, Y., Ser, J. Del, Xia, J., & Yang, G. (2022). Swin transformer for fast MRI. Neurocomputing, 493, 281–304.

[69] Huang, J., Wu, Y., Wu, H., & Yang, G. (2022). Fast MRI Reconstruction: How Powerful Transformers Are? 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2066–2070.

[70] Abdar, M., Pourpanah, F., Hussain, S., Rezazadegan, D., Liu, L., Ghavamzadeh, M., Fieguth, P., Cao, X., Khosravi, A., Acharya, U. R., Makarenkov, V., & Nahavandi, S. (2021). A review of uncertainty quantification in deep learning: Techniques, applications and challenges. Information Fusion, 76, 243–297.

[71] Loftus, T. J., Shickel, B., Ruppert, M. M., Balch, J. A., Ozrazgat-Baslanti, T., Tighe, P. J., Efron, P. A., Hogan, W. R., Rashidi, P., Upchurch, G. R., & Bihorac, A. (2022).

Uncertainty-aware deep learning in healthcare: A scoping review. PLOS Digital Health, 1(8), e0000085.

[72] Lakshminarayanan, B., Pritzel, A., & Deepmind, C. B. (n.d.). Simple and Scalable Predictive Uncertainty Estimation using Deep Ensembles.

[73] Sensoy, M., Kaplan, L., & Kandemir, M. (n.d.). Evidential Deep Learning to Quantify Classification Uncertainty.

[74] Gal, Y., & Uk, Z. A. (2016). Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning Zoubin Ghahramani.

[75] Kendall, A., & Gal, Y. (n.d.). What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?

[76] Maddox, W. J., Garipov, T., Izmailov, P., Vetrov, D., & Wilson, A. G. (n.d.). A Simple Baseline for Bayesian Uncertainty in Deep Learning.

[77] Avci, M. Y., Li, Z., Fan, Q., Huang, S., Bilgic, B., & Tian, Q. (n.d.). Quantifying the uncertainty of neural networks using Monte Carlo dropout for deep learning based quantitative MRI.

[78] Ovadia, Y., Research, G., Fertig, E., Ren, J., Nado Google Research, Z., Nowozin Google Research, S., Dillon Google Research, J. V, Lakshminarayanan, B., & Snoek, J. (n.d.). Can You Trust Your Model's Uncertainty? Evaluating Predictive Uncertainty Under Dataset Shift.

[79] Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

[80] Pruessmann, K. P., Weiger, M., Scheidegger, M. B., & Boesiger, P. (1999). SENSE: sensitivity encoding for fast MRI. Magnetic Resonance in Medicine: An Official Journal

of the International Society for Magnetic Resonance in Medicine, 42(5), 952-962.

[81] Griswold, M. A., Jakob, P. M., Heidemann, R. M., Nittka, M., Jellus, V., Wang, J., … & Haase, A. (2002). Generalized autocalibrating partially parallel acquisitions (GRAPPA). Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine, 47(6), 1202-1210.

[82] Liang, D., Cheng, J., Ke, Z., & Ying, L. (2019). Deep MRI reconstruction: unrolled optimization algorithms meet neural networks. arXiv preprint arXiv:1907.11711