

Uncertainty-Aware High-Fidelity Anatomical MRI Synthesis using Deep Convolutional Network with Monte Carlo Dropout

Zichen Xu^{1,a,*}

¹ *Harrow International School Shenzhen Qianhai, Shenzhen, China*

a. denny.xu@harrowshenzhen.cn

**corresponding author*

Abstract: Multi-modality high-resolution MRI is beneficial for studying the brain structure and function in research and clinical settings. However, its acquisition is time-consuming, which reduces its feasibility for wider adoption especially for certain populations who cannot tolerate long scans. In this study, we propose a convolution neural network to obtain high-resolution T1-weighted MRI from lower-resolution T2-weighted input that can be acquired within a shorter scan time. By leveraging Monte Carlo dropout, our model not only produces high-fidelity anatomical T1-weighted image with higher accuracy compared to baseline model, but also generates uncertainty estimation similar to the actual error map. Our method is validated on the Human Connectome Project, and the experiments indicate our method has the potential to improve the robustness and reliability of deep learning image synthesis and accurately accelerate multi-modality MRI to benefit research and clinical practice.

Keywords: Deep Learning, Multi-contrast MRI, Convolutional Neural Network Image Synthesis, Monte Carlo Dropout.

1. Introduction

Magnetic resonance imaging (MRI) is a non-ionizing medical imaging technology that can be used to non-invasively probe brain structure and functional connectivity. Multi-contrast MRI can bring complementary information regarding brain structure; for example, T1-weighted images can provide superior grey-white contrast which would be beneficial for brain region segmentations and cortical surface reconstruction [1], and T2-weighted MRI is useful for clinical practice by suggesting water-related abnormality such as multiple sclerosis [2].

However, acquiring high-quality multi-contrast MRI is time-consuming, which reduces its feasibility in clinical practice. Separate acquisitions of T1, T2-weighted images of ~1 mm isotropic resolution take ~16 minutes, which is impractical especially for subjects intolerant to long scans such as children, the elderly, or other patients who cannot stay still. Especially, the time collection cost of T1 is typically longer than T2 [3], which is mainly reflected in these three aspects: Relaxation Time difference, whether Repetition Time (TR) is required, and data acquisition strategy. T1-weighted imaging relies on longitudinal relaxation time (T1), requires a longer TR (repetition time) to ensure that the signal is mainly contributed by T1 relaxation time, and generally requires more average sampling. In many situations, we have low-resolution T1 but high-resolution T2.

Existing SR(Super-Resolution) methods can be classified into deep-learning-based and non-deep learning based. Although current fast imaging acquisition (e.g., blipped-CAIPI [4], Wave-CAIPI [5]) and reconstruction (e.g., compressed sensing [6]) methods can accelerate the scan, these accelerated acquisitions might suffer from amplified noises and/or artifacts which might confound downstream image analysis and diagnosis. Deep learning-based becomes more popular in recent years because its ability to extract fine-grained MRI features. CNNs have demonstrated superior performance in transferring medical images between different modalities. For example, Emami et al. leveraged CNNs to generate CT scans from MRI [7]. Efforts have also been made to obtain MRI from PET [8]. The efficacy of using CNNs to transfer MRI between different contrasts has also been shown. Studies have successfully generated T2-weighted MRI from T1-weighted MRI and T1-weighted MRI from T2-weighted MRI [9-11].

Moreover, due to the “black-box” nature of deep learning-based methods, the interpretability of decent image synthesis performance in these studies remains limited. The uncertainty of the deep learning models may risk producing hallucinations when generating images, which could confound downstream diagnosis and analysis [12]. There have been growing interests in identifying and avoiding such hallucinations. Uncertainty quantifications represent a promising avenue to reveal the reliability and improve the robustness of deep learning-based methods. For example, Tanno et al. [13] modelled the uncertainty in terms of intrinsic and parameter uncertainty. Intrinsic uncertainty reflects the inherent ambiguity linked to the ill-posed nature of the image translation problem, quantified through the variance of the target conditional distribution estimated from a separate network. Parameter uncertainty accounts for ambiguity in selecting model parameters, reduced using variational dropouts. This approach captures various settings of network parameters for the given training data, providing a more robust result and an uncertainty estimation represented by the standard deviation. The efficacy of dropout layers for estimating uncertainty and improving robustness has been demonstrated in various applications, such as MRI reconstruction [15] and quantitative MRI [15].

In this study, we aim to leverage deep learning to obtain high-quality multi-contrast MRI without extending the scan time. Specifically, we propose an uncertainty-aware convolutional neural network (CNN) which synthesises high-quality anatomical MRI with uncertainty estimation. The uncertainty is quantified using Monte Carlo dropout [16] which introduces variances in different inferences that reflect uncertainty. The results from multiple inferences can be averaged to reduce the variance and improve the performance. Our model is validated on large-scale in-vivo data and shows promises accelerating the 15-minute scan of T1-weighted and T2-weighted images to less than 1 minute, with uncertainty estimation supporting its reliability and robustness for wider adoption to benefit clinical and neuroscientific research.

The contributions of this study are as follows:

1. We propose a deep convolutional neural network architecture for a multi-contrast MRI model.
2. We design a Monte Carlo dropout-based approach to enable the proposed super-resolution (SR) model to be uncertainty-aware.
3. We conduct extensive experiments on an MRI dataset from the Human Connectome Project to empirically demonstrate the effectiveness of our proposed method.

2. Related Work

Image super-resolution is an important task in various fields, including medical imaging and natural image processing. In recent years, deep learning has emerged as a powerful technique for tackling super-resolution problems.

For medical image super-resolution, one of the pioneering works is the application of the SRCNN model to MRI super-resolution by Dong et al. [17]. The authors demonstrated that the shallow SRCNN network can effectively enhance the resolution of MRI scans. Building upon this, Chen et al. proposed a 3D-SRCNN approach for CT image super-resolution, leveraging the 3D convolutional structure to better capture the volumetric information [18].

In the natural image domain, the SRCNN model proposed by Dong et al. [19] was one of the first deep learning-based techniques for single image super-resolution. This work demonstrated the effectiveness of deep convolutional neural networks in learning an end-to-end mapping for upscaling low-resolution images. Extending this, Kim et al. introduced the VDSR model, which utilizes a very deep (20-layer) convolutional network to achieve state-of-the-art performance on natural image super-resolution benchmarks [20]. Furthermore, Shi et al. proposed the ESPCN model, which employs an efficient sub-pixel convolution layer to enable real-time single image and video super-resolution [21].

These works illustrate the significant progress made in deep learning-based image super-resolution, encompassing both medical and natural image domains. The advancements in network architectures, training strategies, and computational efficiency have led to substantial improvements in super-resolution quality and applicability. As the field continues to evolve, further research in this area is expected to bring even more powerful and practical solutions for various image enhancement tasks.

Moreover, Generative adversarial networks (GANs) have been a popular choice for the contrast transfer tasks [7-9,11]. The adversarial loss in GAN has been utilized to synthesize images with more realistic features and visual quality. However, it is recently demonstrated that the use of the adversarial loss does not necessarily lead to better quantitative performance for image synthesis [22]. For medical imaging tasks which demand image accuracy over visual quality, the use of GANs requires careful considerations.

The use of deep learning-based methods has become increasingly prevalent in this domain. However, due to the inherent "black box" nature of such techniques, the interpretability of their performance remains a challenge. To address this limitation, this study introduces the application of Monte Carlo dropout as a means to enhance the interpretability of the developed model.

3. Methods

Given a LR T2, we first feed it into a U-Net and then use monte carlo dropout in the inference time.

3.1. U-Net Architecture

We employ a U-Net [23] as the baseline model to synthesise T1-weighted MRI from T2-weighted input (Fig. 1). The U-Net architecture, a widely used deep learning framework in medical image analysis, is characterized by its multi-level design, which incorporates both multi-scale encoding and decoding capabilities. This hierarchical structure enables the network to capture both global and local features of the input data, facilitating comprehensive image analysis. Specifically, the lower levels of the U-Net are dedicated to encoding the global information of the brain image, allowing the network to understand the overarching context of the data. On the other hand, the upper levels focus on encoding finer details and local features, enabling the network to capture intricate patterns within the image.

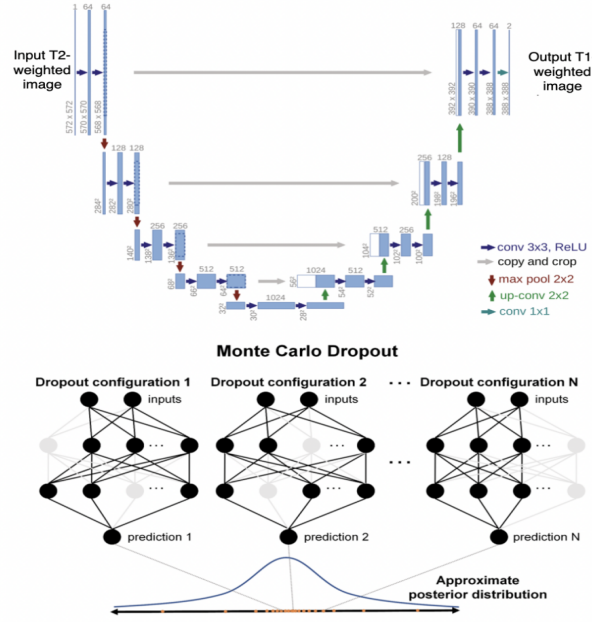


Figure 1: Network architecture. The U-Net [23] with Monte Carlo Dropout [15] entitled “DU-Net” is illustrated.

In our implementation, we adhered closely to the original design principles of the U-Net architecture. At the initial level, the U-Net employs 48 convolution kernels, a configuration that is commonly used to balance computational efficiency and feature representation capacity. Subsequently, the number of convolution kernels is doubled after each max-pooling layer during the encoding process, allowing the network to progressively extract more abstract features from the input data. Conversely, during the decoding phase, the number of convolution kernels is halved after each upsampling layer, ensuring that the network can effectively reconstruct the input data while preserving important features.

One of the key advantages of utilizing 3D convolutions within the U-Net architecture is their enhanced effectiveness in leveraging information from an additional spatial dimension in volumetric MRI data. This spatial context is crucial for accurately capturing the complex structures and relationships present in three-dimensional medical images. Moreover, 3D convolutions have been shown to generate images with reduced boundary artifacts compared to their 2D counterparts, leading to improved overall image quality.

However, it's important to note that common graphics processing units (GPUs) often have limited memory capacity, which can pose challenges when working with large 3D volumes at high spatial resolutions. To address this limitation, we employ a strategy where the inputs to the 3D U-Net are divided into smaller blocks, each consisting of $64 \times 64 \times 64$ voxels. This approach allows us to efficiently process volumetric MRI data within the memory constraints of standard GPUs.

3.2. Loss Function

In this study, we use the mean absolute error (MAE) as a loss function to optimize the model. It calculates the absolute differences between the predicted values and the actual values, and then averages these differences. The formula for MAE is:

$$\mathcal{L}_{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (1)$$

Among them: \mathcal{L}_{MAE} is the value of the MAE loss function, n is the number of samples,

y_i is the true value for the i sample, \hat{y}_i is the predicted value for the i sample.

MAE provides a straightforward interpretation of the error magnitude, as it treats all individual differences equally by considering their absolute values.

3.3. Evaluation Method

We use the mean square error (MSE) as an evaluation function to measure the predictive performance of the model. The mean square error mainly takes the mean of the square variance between the predicted value and the actual value to calculate the loss value of the forecast model, as shown in the equation (2) below.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2)$$

Where: n is the number of samples; y_i is the true value for sample i ; \hat{y}_i is the predicted value for sample i .

A smaller MSE indicates that the predicted values are closer to the actual values, implying better model performance.

3.4. Monte Carlo Dropout

To enhance our network's capacity for uncertainty quantification, we have integrated Monte Carlo Dropout layers, a technique introduced by Avci et al. [15], into our architecture. These dropout layers are strategically inserted after each convolutional layer at every level of the U-Net framework. Traditionally, dropout layers are employed solely during the training phase to mitigate overfitting and enhance model robustness, as outlined by Srivastava et al. [24]. However, in our implementation, we extend the use of dropout layers to both training and inference stages. This enables our network to generate results with diverse configurations for different inference scenarios, as illustrated in Figure 1.

The utilization of dropout layers during inference introduces variability in the network's predictions, effectively simulating results from slightly different network architectures. This variability arises from the activation or deactivation of different nodes during random dropout repetitions. Consequently, each inference yields slightly different outcomes, reflecting the inherent uncertainty in the model's predictions. By performing multiple inferences and subsequently averaging the results, we can mitigate errors stemming from network variance, thereby obtaining more reliable predictions. Additionally, by computing the standard deviation across multiple inferences, we can quantify the uncertainty associated with the network's estimations, providing valuable insights into the reliability of the model's predictions under varying conditions. This comprehensive approach to uncertainty quantification equips our network with the capability to not only produce accurate predictions but also assess the confidence level associated with each prediction, enhancing the overall reliability and interpretability of the model.

4. Experiments

4.1. Dataset

We used MRI data from the Human Connectome Project (<https://www.humanconnectome.org/>) in this study. The T1-weighted images were acquired on 3T MRI scanners using a 3D magnetization-prepared gradient echo (MPRAGE) sequence at 0.7 mm isotropic resolution. The T2-weighted images were acquired within the same session using a 3D turbo spin-echo (TSE) sequence with slab selective, variable excitation pulse at 0.7 mm isotropic resolution. The scan time for T1- and T2-

weighted images is 7 and 8 minutes respectively. The T1- and T2-weighted images are pre-processed and co-registered.

4.2. Implementation Details

The network illustrated in Fig. 1 was implemented using Keras library (keras.io) with a TensorFlow backend(tensorflow.org) using a NVIDIA GeForce RTX 4090 with 24GB memory.

Network was trained using 40 randomly selected healthy subjects with 20% of the data used for validation. The DU-Net and U-Net were both trained with T2-weighted image as input and T1-weighted image as output and optimised with an Adam optimiser with default parameters for 10 epochs using mean absolute error as the loss function (~ 2 minutes per epoch). To validate our model's potential in further accelerating the MRI acquisition, we retrospectively downsampled the input T2-weighted image to 1.4 mm isotropic resolution and used the low-resolution T2-weighted image as the input of the network (denoted as “undersampled”). The networks with the lowest validation error were saved and applied. The dropout layers in the DU-Net were activated during both training and inference (dropout rate=0.2).

Another 20 subjects unseen during training were used for the evaluation of the network. During the inference, DU-Net was applied 20 times with dropout layers activated, which led to 20 output images. The 20 output images were averaged to generate the final output of DU-Net with reduced variance. The standard deviation of the 20 images was calculated to reflect the model uncertainty. The mean squared error (MSE) of network output images with the native T1-weighted images was calculated to evaluate image similarity and model performance.

5. Results

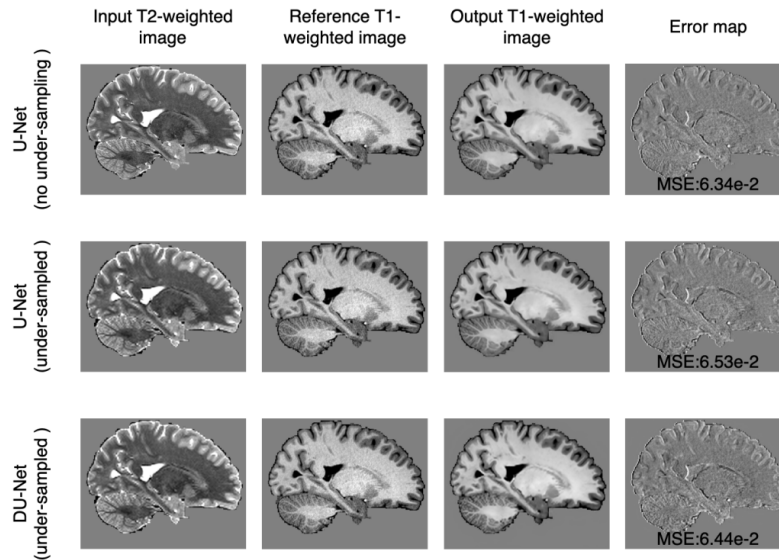


Figure 2: Image results. The input T2-weighted images, reference T1-weighted images, network output T1-weighted images, and the difference between reference and network output from different models are demonstrated. “Under-sampled” indicates the input was retrospectively under-sampled to generate a low-resolution image which takes shorter acquisition time. The mean squared error (MSE) between reference and network output is listed to quantify the image similarity.

Figure 2 demonstrates the image results of different models. Both U-Net and DU-Net produce high-fidelity T1-weighted images from T2-weighted input. Under-sampling T2-weighted input slightly

compromises the network performance, due to the information loss in lower-resolution input. However, a lower-resolution input (1.4 mm isotropic resolution in our study) requires shorter acquisition time (~ 1 min), which is helpful especially for clinical practice. DU-Net produces images with lower MSE compared to U-Net, because the averaging of multiple inferences reduces the variance and improves the accuracy and robustness. The mean MSE for 20 evaluation subjects is 0.0585, 0.0605, and 0.0600 for U-Net (no under-sampling), U-Net (under-sampled), and DU-Net (under-sampled) respectively.

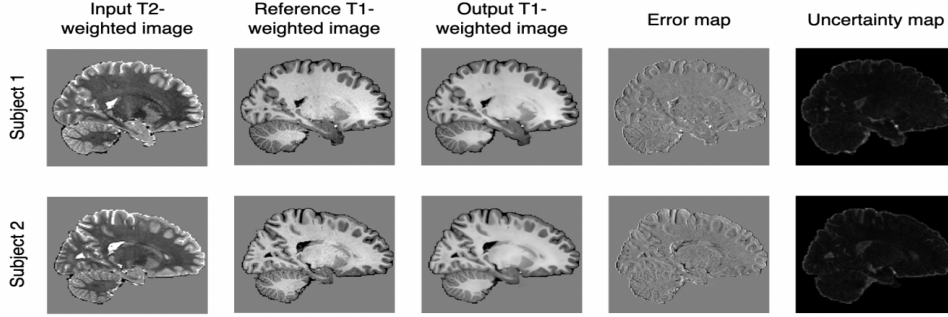


Figure 3: Image results and uncertainty estimation from multiple subjects. The under-sampled low-resolution T2-weighted image input, reference T1-weighted image, DU-Net output T1-weighted image, the difference between reference and network output, and the uncertainty estimation from 2 evaluation subjects are demonstrated.

Figure 3 shows the image results and uncertainty estimation from DU-Net with under-sampled low-resolution T2-weighted images as input from 2 evaluation subjects. The model exhibits decent performance on different subjects, generating T1-weighted images highly similar to the native T1-weighted images. Notably, the estimated uncertainty map shows similar patterns to the error map, indicating our method successfully captures the model uncertainty.

6. Discussion

Our proposed approach introduces a sophisticated deep learning model, DU-Net, aimed at facilitating image synthesis and uncertainty estimation for high-resolution multi-modality MRI. Leveraging the power of deep learning, DU-Net demonstrates remarkable capabilities in generating high-fidelity T1-weighted images from under-sampled, low-resolution T2-weighted images, closely resembling the reference native T1-weighted images. Notably, DU-Net achieves this with significantly lower Mean Squared Error (MSE) compared to the baseline U-Net, underscoring its superior performance in image synthesis tasks. Additionally, our model produces uncertainty maps that exhibit similar patterns to the corresponding error maps, providing valuable insights into the reliability and confidence associated with each synthesized image. Such capabilities hold immense potential to revolutionize the field of MRI imaging by significantly enhancing imaging speed without compromising diagnostic accuracy, thereby benefiting both clinical practice and scientific research endeavors.

Central to the design of our model is the incorporation of Monte-Carlo dropout, a technique commonly used to mitigate network overfitting and enhance training robustness. In our implementation, dropout layers are activated during both training and inference stages, enabling the network to generate diverse outputs for different inference scenarios. These variations in inference outcomes reflect the inherent variance within the model, which can be effectively reduced by averaging multiple inferences to improve model accuracy. Furthermore, quantifying the deviation among these inferences allows for the estimation of uncertainty, providing valuable information

regarding the reliability of synthesized images in different regions. In practical scenarios where reference images are unavailable, and error maps cannot be computed, uncertainty maps serve as invaluable tools for assessing the trustworthiness of synthesized images, thereby enhancing clinical decision-making and diagnostic confidence.

Our method has potential to accelerate the 15-min imaging to less than 1 minute. By retrospective undersampling, we demonstrate the model produces images with high quality even with lower-resolution inputs (1.4 mm, Fig. 2). Acquiring T2-weighted images at 1.4 mm resolution requires substantially shorter scan time using fast imaging techniques such as echo planar imaging (EPI) [25]. For reference, in the Human Connectome Project, the acquisition of T2-weighted images using EPI at 1.25 mm isotropic resolution only requires 8 seconds. Even if multiple volumes are required for correcting geometric distortions and/or reducing the noise level, the acquisition time will still be significantly shortened.

This study underscores the transformative potential of deep learning in advancing the field of multi-modality high-resolution MRI, offering faster and more reliable imaging solutions. Future endeavors may focus on conducting systematic analyses of network-generated images, such as segmentation and lesion quantification, to further elucidate their clinical utility and pave the way for their broader adoption in medical practice and research settings. The generalization of the proposed model also warrants further investigation. Theoretically, the improvement in robustness brought by the Monte-Carlo dropout should also benefit the network generalizability, further promoting its use in research and clinical practice.

While GANs have emerged as a popular choice for contrast transfer tasks in medical imaging, our decision to employ a plain U-Net in lieu of GANs is informed by recent findings indicating that GANs may not necessarily improve quantitative performance. Moreover, integrating GANs into our model would introduce additional complexities in training and generalization, potentially hindering its efficiency and effectiveness. However, the prospect of combining adversarial loss in GANs with Monte-Carlo dropout presents an intriguing avenue for future research, warranting further exploration and investigation into its potential synergies and benefits.

7. Conclusion

In conclusion, our study presents DU-Net, a deep learning model for high-resolution multi-modality MRI image synthesis and uncertainty estimation. DU-Net excels in generating high-fidelity T1-weighted images from under-sampled T2-weighted inputs, outperforming baseline methods with lower MSE. Additionally, DU-Net produces uncertainty maps mirroring error maps, enhancing confidence in image reliability. This advancement promises accelerated imaging without sacrificing diagnostic accuracy, benefiting clinical practice and research. While GANs are popular, our use of U-Net over GANs is informed by their limited quantitative improvement and added complexity. Monte-Carlo dropout integration enhances model robustness and enables diverse inference outputs for uncertainty estimation. Our method reduces a 15-minute procedure to less than 1 minute, underscoring its potential clinical impact. Further research may explore systematic analysis of network-generated images for broader adoption in medical practice and research.

References

- [1] Fischl, B. (2012). *FreeSurfer*. *Neuroimage*, 62(2), 774-781.
- [2] Filippi, M., Paty, D. W., Kappos, L., Barkhof, F., Compston, D. A. S., Thompson, A. J., ... & Miller, D. H. (1995). *Correlations between changes in disability and T2-weighted brain MRI activity in multiple sclerosis: a follow-up study*. *Neurology*, 45(2), 255-260.
- [3] Brown, R. W., Cheng, Y. C. N., Haacke, E. M., Thompson, M. R., & Venkatesan, R. (2014). *Magnetic resonance imaging: physical principles and sequence design*. John Wiley & Sons.

- [4] Setsompop, K., Gagoski, B. A., Polimeni, J. R., Witzel, T., Wedeen, V. J., & Wald, L. L. (2012). Blipped-controlled aliasing in parallel imaging for simultaneous multislice echo planar imaging with reduced g-factor penalty. *Magnetic resonance in medicine*, 67(5), 1210-1224.
- [5] Bilgic, B., Gagoski, B. A., Cauley, S. F., Fan, A. P., Polimeni, J. R., Grant, P. E., ... & Setsompop, K. (2015). Wave-CAIPI for highly accelerated 3D imaging. *Magnetic resonance in medicine*, 73(6), 2152-2162.
- [6] Lustig, M., Donoho, D. L., Santos, J. M., & Pauly, J. M. (2008). Compressed sensing MRI. *IEEE signal processing magazine*, 25(2), 72-82.
- [7] Emami, H., Dong, M., Nejad-Davarani, S. P., & Glide-Hurst, C. K. (2018). Generating synthetic CTs from magnetic resonance images using generative adversarial networks. *Medical physics*, 45(8), 3627-3636.
- [8] Choi, H., & Lee, D. S. (2018). Generation of structural MR images from amyloid PET: application to MR-less quantification. *Journal of Nuclear Medicine*, 59(7), 1111-1117.
- [9] Dai, X., Lei, Y., Fu, Y., Curran, W. J., Liu, T., Mao, H., & Yang, X. (2020). Multimodal MRI synthesis using unified generative adversarial networks. *Medical physics*, 47(12), 6343-6354.
- [10] Iglesias, J. E., Billot, B., Balbastre, Y., Tabari, A., Conklin, J., González, R. G., ... & Alzheimer's Disease Neuroimaging Initiative. (2021). Joint super-resolution and synthesis of 1 mm isotropic MP-RAGE volumes from clinical MRI exams with scans of different orientation, resolution and contrast. *Neuroimage*, 237, 118206.
- [11] Yurt, M., Dar, S. U., Erdem, A., Erdem, E., Oguz, K. K., & Çukur, T. (2021). mustGAN: multi-stream generative adversarial networks for MR image synthesis. *Medical image analysis*, 70, 101944.
- [12] Cohen, J. P., Luck, M., & Honari, S. (2018). Distribution matching losses can hallucinate features in medical image translation. In *Medical Image Computing and Computer Assisted Intervention—MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part I* (pp. 529-536). Springer International Publishing.
- [13] Tanno, R., Worrall, D. E., Kaden, E., Ghosh, A., Grussu, F., Bizzi, A., ... & Alexander, D. C. (2021). Uncertainty modelling in deep learning for safer neuroimage enhancement: Demonstration in diffusion MRI. *NeuroImage*, 225, 117366.
- [14] Schlemper, J., Castro, D. C., Bai, W., Qin, C., Oktay, O., Duan, J., ... & Rueckert, D. (2018). Bayesian deep learning for accelerated MR image reconstruction. In *Machine Learning for Medical Image Reconstruction: First International Workshop, MLMIR 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Proceedings 1* (pp. 64-71). Springer International Publishing.
- [15] Avci, M. Y., Li, Z., Fan, Q., Huang, S., Bilgic, B., & Tian, Q. (2021). Quantifying the uncertainty of neural networks using Monte Carlo dropout for deep learning based quantitative MRI. *arXiv preprint arXiv:2112.01587*.
- [16] Gal, Y., & Ghahramani, Z. (2016, June). Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In *international conference on machine learning* (pp. 1050-1059). PMLR.
- [17] Dong, C., Loy, C. C., & Tang, X. (2016). Accelerating the super-resolution convolutional neural network. In *Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II 14* (pp. 391-407). Springer International Publishing.
- [18] Chen, Y., Xie, Y., Zhou, Z., Shi, F., Christodoulou, A. G., & Li, D. (2018). Brain MRI super-resolution using 3D deep densely connected neural networks. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)* (pp. 739-742). IEEE.
- [19] Dong, C., Loy, C. C., He, K., & Tang, X. (2014). Learning a deep convolutional network for image super-resolution. In *Computer Vision—ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part IV 13* (pp. 184-199). Springer International Publishing.
- [20] Kim, J., Lee, J. K., & Lee, K. M. (2016). Accurate image super-resolution using very deep convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1646-1654).
- [21] Shi, W., Caballero, J., Huszár, F., Totz, J., Aitken, A. P., Bishop, R., ... & Wang, Z. (2016). Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1874-1883).
- [22] Li, Z., Fan, Q., Bilgic, B., Wang, G., Wu, W., Polimeni, J. R., ... & Tian, Q. (2023). Diffusion MRI data analysis assisted by deep learning synthesized anatomical images (DeepAnat). *Medical Image Analysis*, 86, 102744.
- [23] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18* (pp. 234-241). Springer International Publishing.
- [24] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1), 1929-1958.
- [25] Stehling, M. K., Turner, R., & Mansfield, P. (1991). Echo-planar imaging: magnetic resonance imaging in a fraction of a second. *Science*, 254(5028), 43-50.

- [26] Schilling, K. G., Blaber, J., Hansen, C., Cai, L., Rogers, B., Anderson, A. W., ... & Landman, B. A. (2020). Distortion correction of diffusion weighted MRI without reverse phase-encoding scans or field-maps. *PLoS One*, 15(7), e0236418.