High Resolution MR Reconstruction

Ramani Tyagi
Stanford University
ramanit@stanford.edu

Abstract

Magnetic resonance imaging (MRI) is an imaging technology that provides detailed, non-invasive visualization of brain structures, giving insight into its structure and function. 3D brain MRI enables diagnosis and monitoring of various neurological conditions and diseases. However, MRI faces limitations in its lengthy acquisition times. Long scan times lead to patient discomfort, increased costs, more motion artifacts degrading image quality, and reduced patient throughput. In this work, I present a new approach for MR super-resolution utilizing a residual dense network as an encoder, then applying a positional encoding to that feature map representation, and passing the encoding through a deep multi-layer perceptron decoder. Through the experiment results, I show that neural representations show promising results for the MRI reconstruction and super-resolution tasks.

1. Introduction

Magnetic Resonance Imaging (MRI) has become an indispensable medical imaging technology for visualizing body structures to aid in diagnosis, monitoring and clinical decision making. Due to the enhanced contrast it provides, it has the powerful ability to distinguish adjacent soft tissue from one another and is particularly effective for imaging soft tissues, unlike other medical imaging modalities. It can also provide multiplanar images, such as direct, sagittal, coronal and oblique images which are not obtainable with other methods such as CT or radiography [4].
Despite all its advantages in medical imaging, it also poses limitations. It can be discomforting to patients to spend a long time in a small, closed space, and patients may experience anxiety and claustrophobia [4, 6]. The long acquisition times lead to reduced clinic efficiency and patient throughput. The long acquisition times also result in more motion artifacts from the patient's movement, which can obscure important details and make a scan more difficult to interpret. For certain organs, such as lung and heart that have internal motion, it is difficult to capture a high resolution MRI scan.

MRI is still one of the primary tools to provide imaging biomarkers for many clinical scenarios, including cancers, neurological disorders such as Alzheimer's [2], physical trauma and injury. This highlights the development of solutions over the past few decades aimed at accelerating MRI acquisition, including reconstruction and super-resolution from under-sampled, low-quality MRI scans. These solutions are paramount in providing a more positive patient experience, higher clinical efficiency, and excellence in scan quality.

2. Related Work

There are two key types of solutions to the MR reconstruction challenge. The classical approaches include parallel imaging and compressed sensing, which have been widely adopted clinically for faster MRI acquisition. Recently, deep learning based solutions have gained popularity due to their excellent performance for the task, surpassing even classical methods. Deep learning approaches still face limitations as they need large, diverse, well-annotated datasets, and the computational resources required make them difficult to be widespread.

2.1. Classical Methods

The classical methods are described in Tab. 1. in Combining parralel imaging methods with compressed sensing results in fast, high-quality MRI reconstrucion.

2.2. Deep Learning Methods

There are many deep learning models for MRI reconstruction and super-resolution, of which some popular ones are summarized in Tab. 2. Evaluation of deep learning methods involves the peak signal noise ratio (PSNR) and structural similarity (SSIM). The difficulty in comparison of these models lies in the lack of one centralized benchmark, due to training and testing on various MRI modalities and datasets.

Table 1. Classical image reconstruction methods for MRI.

Technique	Description
GRAPPA (Parallel Imaging)	GRAPPA reconstructs missing k-space data by applying a linear combination of acquired data from other receiver coils. Calibration for GRAPPA is more extensive as it requires additional autocalibration signal (ACS) lines. [3]
SENSE (Parallel Imaging)	SENSE combines aliased images from each coil using coil sensitivity information. The coil sensitivity profiles are either obtained through a separate calibration scan or estimated using the acquired data. The method is sensitive to inaccuracies in coil sensitivity information. [7]
Compressed Sensing (CS)	CS exploits image sparsity to reconstruct high quality images from undersampled k-space data . CS is dependent on selecting appropriate sparsifying transforms and regularization. [5]

3. Approach

In this project, I propose a novel approach for high-resolution MRI reconstruction that leverages the power of neural representations. This method combines the strengths of deep learning and implicit neural representations to capture spatial relationships and high-frequency details, enabling the generation of high-quality reconstructions from undersampled and low-resolution MRI data.

3.1. Problem Definition

Let us consider a low-resolution image L and its high resolution counterpart R_i. There is a downsampling operation D we can apply that $L = D(R)$. We construct the low resolution image from the high resolution ground truth image using this operation. Then, the objective is given the low resolution image L and the coordinates at some spatial position in the high resolution image to predict the pixel intensities R' for those coordinates in R.

4. Experiments

4.1. Dataset

I used the HCP-1200 dataset [11], which contains brain MRI acquisitions of young adults. Data preprocessing was done as in the ArSSR paper [13].
A scaling factor s is defined as $s = \texttt{random}(2, 4)$. From the MR images, patches of size $10s \times 10s$ were extracted randomly from the image. The ground truth, R was the extracted patch as is. Then the downsampled low resolution image L was created by downsampling R by a factor of s. An example of L and R are shown in Fig. 1.
The dataset contains a 1000 such pairs, of which 800 are randomly sampled for training, 100 for validation, and 100 for testing.

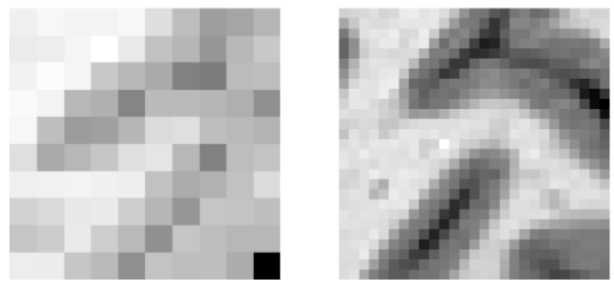

Figure 1. An example of a low-resolution (left) and a high-resolution (right) patch pair.

4.2. Model

I started by using the ArSSR [13] model as a baseline for my model experimentation. The architecture of ArSSR [13] is shown in Fig. 2.

I used the Residual Dense Network for super resolution [14] as my encoder network. I tried experimenting with other popular convolutional neural networks in computer vision, such as U-Net, ResNet, and VGGNet but they did not yield a good performance in training itself, as the losses did not decrease substantially. I decided to use RDN as it is an architecture that is proficient at the super resolution task, and the local and global features extracted from the encoder will be a lot more relevant.

Then the output features from the RDN Encoder Block are trilinear interpolated to form another feature map, which is passed through a positional encoding. The positional encoding process applies a series of sine and cosine functions to enrich the feature vectors and incorporate spatial relationships in the output from the RDN. After this, the encoded vectors are passed through a multi-layer perceptron decoder block with 8 hidden layers, each consisting of a linear transformation followed by a ReLU activation function. This MLP outputs the final predicted intensities

Table 2. Deep learning Image reconstruction methods for MRI.

Model	Modality & Dataset	Novelty
Automap [15]	ImageNet (general images) and MGH-USC HCP (Brain MRI)	Uses manifold learning to develop a fully connected neural network to learn a direct mapping from k-space to image space without utilizing traditional Fourier transform-based reconstruction methods.
Deep Cascade of Convolutional Neural Networks [8]	Cardiac MRI	Proposes a deep cascade of CNNs for MR reconstruction.
CoIL [10]	AAPM (Abdominal CT)	CoIL trains an MLP to encode the measurement field by mapping the coordinates of the measurements to their responses, instead of mapping measurements to the desired image like other methods. A key advantage of CoIL is that it is a self-supervised method, with no training data required.
STINR-MR [9]	XCAT (Simulated Cardiac and Thoracic MRI), UTSW (Liver Motion MRI), 4D THRIVE (Abdomen MRI)	Uses spatial and temporal implicit neural representations with learning-based hash encoding for MRI registration and reconstruction.
Dilated Convolutional Encoder-Decoder Network [1]	Kirby21, ALVIN (adult), BraTS, MSSEG (all Brain MRI)	Uses dilated convolutions and deconvolutions to recover high frequency information of MRI. Also adopts use of residual learning and local skip connections. observation.
Enhanced generative adversarial network for 3D brain MRI super-resolution [12]	HCP (Brain MRI)	Created a SOTA 3D memory efficient residual dense generator and proposed a fully convolutional pyramid pooling discriminator to recover brain image detail.
CuNeRF	Medical Segmentation Decathlon (MSD) Brain Tumour MRI	CuNeRF is a a zero-shot framework that learns the continuous volumetric representation from LR volumes using cube based sampling and rendering.
TransMRSR	IXI MRI (multiple modalities)	Proposes a two-stage network using convolutional blocks to extract local information and transformer blocks to capture long-range dependencies.
ArSSR	HCP (Brain MRI)	Defines the image reconstruction as a continuous implicit voxel function of the low-resolution image.

R'.

4.3. Evaluation Metrics

I evaluated the performance of the model the following quantitative metrics commonly used in image reconstruction and super-resolution tasks:

- Peak Signal-to-Noise Ratio (PSNR): Is the ratio of the maximum possible power of a signal to the power of the corrupting noise. PSNR measures human perception of reconstruction quality.
- Structural Similarity Index (SSIM): Assesses the perceived quality of the reconstructed images by comparing their structural information.
- Mean Squared Error (MSE): Measures the average pixel-wise squared difference between

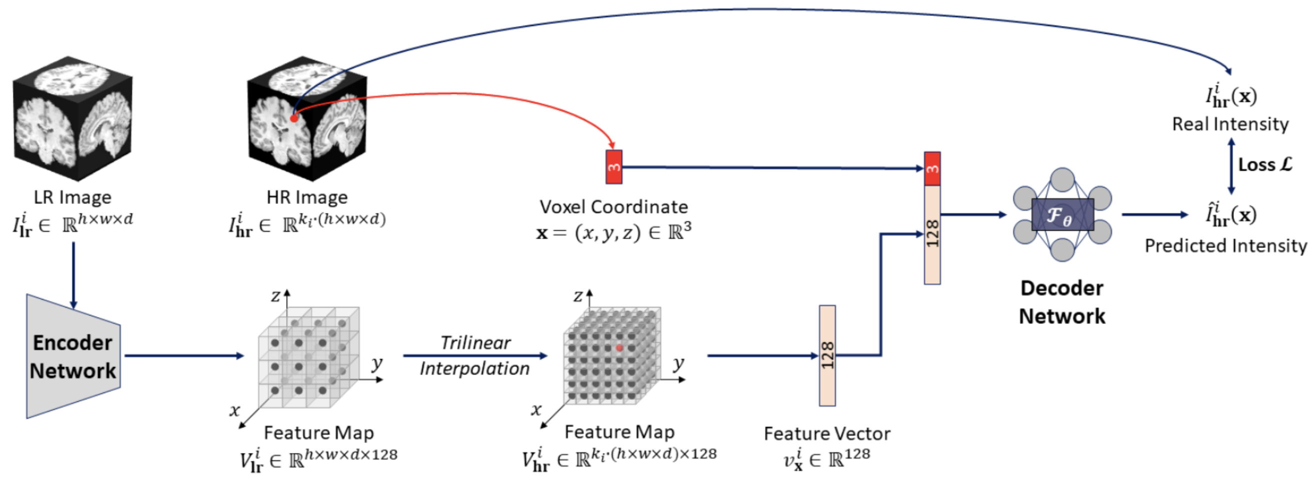

Figure 2. The architecture of ArSSR. [13]

Metric	Value
PSNR	23.38
SSIM	0.7274
MSE	0.001042

Table 3. Results

the reconstructed and ground truth images.

4.4. Quantitative Results

The results of the model evaluated on the test set are shown in Tab. 3. While the PSNR and SSIM results are decent, there is a room for improvement to achieve state-of-the-art performance. Future directions to help improve performance can include designing and comparing different positional encoding functions, as well as trying out different configurations of the MLP and performing hyperparameter optimization.

4.5. Qualitative Results

The predicted outputs of the model are shown in Fig. 3. In terms of a qualitative analysis, overall the shape and structures are reconstructed fairly accurately.
The first example shows a good case of reconstruction, where the model is able to reconstruct a predicted image with similar contrast and structural details as the ground truth.
In the second and third examples, the model predictions capture the overall shapes and intensity variations present in the ground truth images reasonably well. However, some finer details and textures are overly smoothed out in the reconstructions.
The fourth and last example shown illustrates a more challenging case, with the ground truth containing very fine, high-frequency details. While the predicted reconstruction sill captures the overall shape, it struggles with the texture and details, instead generating a blurrier approximation.
Overall, the model is effective for reconstructing a high-resolution image given the low-resolution input. It has limitations in the ability to reconstruct texture details and patterns in the tissue, especially for regions with very high frequencies in the ground truth data. These results indicate that the model has potential for enhancing MRI scans, but further improvements are needed in refining details.

A deeper qualitative analysis with more clinical feedback would be helpful in determining what characteristics of the reconstruction are more important in practice. Determining how the loss of texture and blurred edges might impact real world interpretations from radiologists would be helpful in creating a more clinically relevant evaluation framework and guiding future improvements to the model. By incorporating this clinical perspective, we could refine our evaluation metrics and loss functions to better align with the practical requirements of medical imaging. This could involve developing custom metrics that assign higher weights to clinically important features or designing loss functions that penalize clinically significant errors more heavily.

5. Conclusion

Through this project, I tackled a prevalent challenge in medical imaging using computer vision. I created a model combining a Residual Dense Network encoder, with trilinear interpolation and a positional encoding block, and a multi-layer perceptron to achieve MR super resolution. My approach highlights the potential of positional encodings and neural representations for this task. The quantitative evaluation using PSNR, SSIM, and MSE metrics showed

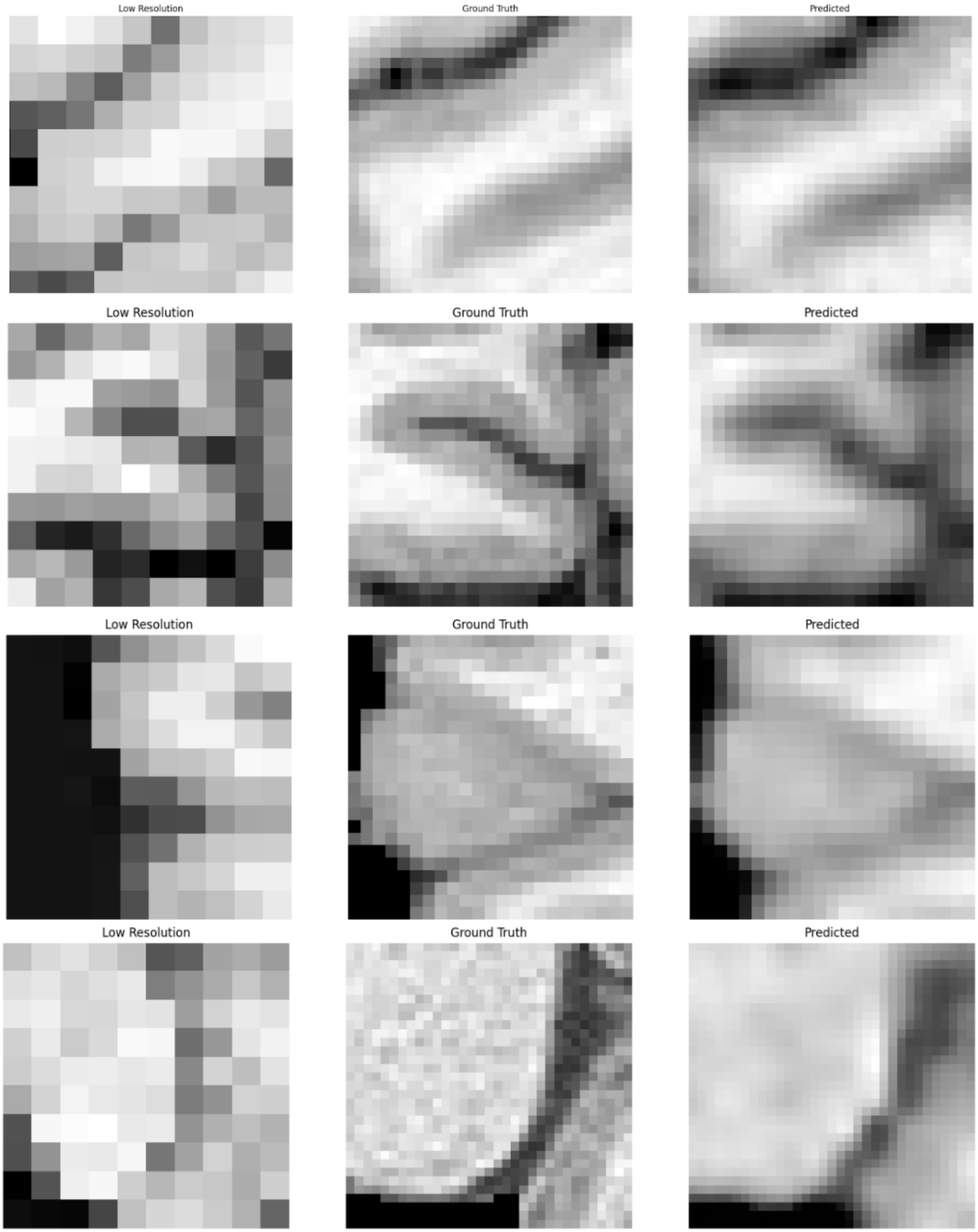

Figure 3. Qualitative image reconstruction results. Each row represents a test example. Leftmost images are low-resolution input images, middle images are ground truth, and rightmost images are the predicted high-resolution output image from the model.

promising results in terms of reconstruction quality. The qualitative analysis of the reconstructed images demonstrates the model's ability to recover overall structural details and intensity variations. There are still limitations to this model, including training limitations, non-generalizability, and lack of finer high-quality details in the reconstructed image.

I learned a lot throughout the course of this project. The biggest challenge and learning for me was working with large-scale real world data. I changed datasets multiple times throughout the course of this project, as sometimes I hit roadblocks with dataset size, or lack of information for correctly preprocessing the data. Despite there being multiple papers that used a certain dataset, there were rarely clear instructions defined on how to process the raw data into a state that could be fed into a neural network. I encountered multiple file formats and processing methods depending on the acquisition and modality. It is not as easy to try out new ideas working with such data. It was challenging to quickly iterate, and my steps had to be much more carefully planned than the other smaller scale projects I was used to. Unlike smaller, well-processed datasets, experimenting with different models, hyperparameters, or architectures on real-world data demanded substantial computational resources and time. Sometimes certain configurations could cause errors which would take a long time to debug, and I had to start from scratch again. I learned to be a lot more strategic in my approaches and to also be more organized – like documenting what worked and what didn't, etc.

With using a VM on GCP, I encountered unexpected challenges with the NVIDIA driver not working at certain times, or the boot disk being out of space and having to clone it and increase the size. I was a lot more methodical in code I wrote training the model, taking steps like making sure to release any CUDA memory I had whenever I could.

I know that I have grown as both a researcher and an engineer, with valuable lessons learned throughout this course.

6. Future Work

This work has potential to explore further directions to improve performance. One such direction is leveraging pre-trained super-resolution networks on generalized image data in the encoder and applying transfer learning to improve model performance and reduce training time. Exploring more neural representations and designing new encoding functions could also lead to better capture of spatial relationships and fine details in the reconstructed images. Another promising direction is to integrate additional anatomical prior knowledge such as segmentation masks or structural information to derive more accurate and clinically relevant outputs. Another extension would be to transfer the model into $3D$ space, working with multiple MRI slices at a time and resulting in $3D$ reconstructions. Generalizing a model to various MRI datasets and configurations such as T1, T2, diffusion modalities or varying coils and channels is a desirable feature. Improving the interpretability of the model and focusing on developing a clinically relevant framework is a natural next step once the model is state-of-the-art.

7. Acknowledgements

Data were provided [in part] by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

References

[1] Jinglong Du, Lulu Wang, Yulu Liu, Zexun Zhou, Zhongshi He, and Yuanyuan Jia. Brain mri super-resolution using 3d dilated convolutional encoder–decoder network. *IEEE Access*, 8:18938–18950, 2020. 3

[2] Giovanni B Frisoni, Nick C Fox, Clifford R Jack Jr, Philip Scheltens, and Paul M Thompson. The clinical use of structural mri in alzheimer disease. *Nature Reviews Neurology*, 6(2):67–77, 2010. 1

[3] Mark A. Griswold, Peter M. Jakob, Robin M. Heidemann, Mathias Nittka, Vladimir Jellus, Jianmin Wang, Berthold Kiefer, and Axel Haase. Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magnetic Resonance in Medicine*, 47(6):1202–1210, 2002. 2

[4] Girish Katti, Syeda Arshiya Ara, and Ayesha Shireen. Magnetic resonance imaging (mri)–a review. *International journal of dental clinics*, 3(1):65–70, 2011. 1

[5] Michael Lustig, David Donoho, and John M. Pauly. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine*, 58(6):1182–1195, 2007. 2

[6] Zachary Munn, Alan Pearson, Zoe Jordan, Fred Murphy, Diana Pilkington, and Amanda Anderson. Patient anxiety and satisfaction in a magnetic resonance imaging department: initial results from an action research study. *Journal of Medical Imaging and Radiation Sciences*, 46(1):23–29, 2015. 1

[7] Klaas P. Pruessmann, Markus Weiger, Markus B. Scheidegger, and Peter Boesiger. SENSE: Sensitivity encoding for fast MRI. *Magnetic Resonance in Medicine*, 42(5):952–962, 1999. 2

[8] Jo Schlemper, Jose Caballero, Joseph V Hajnal, Anthony Price, and Daniel Rueckert. A deep cascade of convolutional neural networks for mr image reconstruction. In *Information Processing in Medical Imaging: 25th International Conference, IPMI 2017, Boone, NC, USA, June 25-30, 2017, Proceedings 25*, pages 647–658. Springer, 2017. 3

[9] Hua-Chieh Shao, Tielige Mengke, Jie Deng, and You Zhang. 3D cine-magnetic resonance imaging using spatial and temporal implicit neural representation learning (STINR-MR). 2023. 3

[10] Yu Sun, Jiaming Liu, Mingyang Xie, Brendt Wohlberg, and Ulugbek S Kamilov. Coil: Coordinate-based internal learning for imaging inverse problems. *arXiv preprint arXiv:2102.05181*, 2021. 3

[11] David C Van Essen, Kamil Ugurbil, Edward Auerbach, Deanna Barch, Timothy EJ Behrens, Richard Bucholz, Acer Chang, Liyong Chen, Maurizio Corbetta, Sandra W Curtiss, et al. The human connectome project: a data acquisition perspective. *Neuroimage*, 62(4):2222–2231, 2012. 2

[12] Jiancong Wang, Yuhua Chen, Yifan Wu, Jianbo Shi, and James Gee. Enhanced generative adversarial network for 3d brain mri super-resolution. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 3627–3636, 2020. 3

[13] Qing Wu, Yuwei Li, Yawen Sun, Yan Zhou, Hongjiang Wei, Jingyi Yu, and Yuyao Zhang. An arbitrary scale super-resolution approach for 3D MR images via implicit neural representation. *IEEE Journal of Biomedical and Health Informatics*, 27(2):1004–1015, 2022. 2, 4

[14] Yulun Zhang, Yapeng Tian, Yu Kong, Bineng Zhong, and Yun Fu. Residual dense network for image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2472–2481, 2018. 2

[15] Bo Zhu, Jeremiah Z Liu, Stephen F Cauley, Bruce R Rosen, and Matthew S Rosen. Image reconstruction by domain-transform manifold learning. *Nature*, 555(7697):487–492, 2018. 3

8. Supplementary Material

The source code can be found here.