Multiscale Super-Resolution for Harmonized Diffusion Magnetic Resonance Imaging

INTRODUCTION: Though medical imaging like Magnetic Resonance Imaging (MRI) is a powerful tool for medical professionals, there are problems with its usability. The utility of medical images hinges upon their resolution as higher resolution conveys more information to the viewer. On the other hand, high-resolution scans come with the cost of greater time in the machine, which increases patient discomfort and leaves open the possibility for motion artifacts. Super-Resolution, in its ability to artificially increase the quality of medical images, enables greater safety and quality of diagnosis [1].

Deep learning methods have become prominent solutions for the super-resolution problem. These methods, however, have issues when applied to medical images both in terms of flexibility and performance. Rather than working with RGB images, as is often the case for most super-resolution studies, we must directly manage the more sophisticated data structure of Rotation Invariant Spherical Harmonics (RISH) feature maps generated from Diffusion MRI data. Diffusion MRI, or dMRI for short, is a unique method for mapping white matter composition (tractography) in brain tissue. From that data, RISH features can be calculated that serve to standardize measurements across varying imaging machines and input parameters.

Considering these unique challenges, we explore leading SISR models and develop a new superior model, Flexible Medical Image Super-Resolution (FMISR). The proposed model directly learns an end-to-end mapping between low- and high-resolution 3D volumes at variable scales. We experimentally test the model's performance and flexibility with a range of scales and show it outperforms other models, scoring higher on standard metrics for image restoration.

RELATED WORKS: The Convolutional Neural Networks (CNNs) listed below are leading models for SISR.

Enhanced Deep Super Resolution (EDSR): The central feature of EDSR [2] is the Residual Block, which consists of two convolutional layers sandwiching a Rectified Linear Unit (ReLU) activation layer, which enables the construction of deeper models as each Residual Block is not very computationally demanding. In order to combat instability in training, a scaling factor of 0.1 is used to minimize the impact of poor individual training batches in order to hinder potential cascade away from convergence.

The EDSR model is an extension of the standard Residual Network that has been trimmed for efficiency. By removing features like Batch Normalization layers from the SRResNet [3] model, EDSR has greater flexibility for feature learning as normalization decreases learning range. Combined with decreased computational overhead and memory requirements, EDSR is able to achieve superior performance compared to prior models because of a larger number of total parameters, which increases learning capacity. Furthermore, the structure, in terms of width and depth is optimized for greater computational efficiency. In CNN based models, memory load scales at $O(BF)$ while feature size scales at $O(BF^2)$, where B is the model depth (number of layers) and F is the model width (number of feature channels). Thus, very wide models like EDSR are favorable.

Densely Connected Super Resolution Networks (DCSRN): The DCSRN [4] model features skip connects that create bypasses across layers. This feature has several benefits. The first of which is that overall the mean path for calculation is shorter, which decreases computational load and increases rate of learning. Furthermore, the model is lighter overall as many of the weights are shared across paths. This factor leads to a smaller propensity for overfitting as there are fewer parameters.

In addition, DCSRN differs from previous models by using 3D inputs rather than 2D slices, an approach that ignores the fact that continuous structures are present between slices that can serve as features to be learned. The model takes advantage of 3D convolutional filters in order to utilize features that span across multiple slices. This contextual information can significantly add to the feature space available for the model to learn from. All of these factors lead to the DCSRN model's superior speed and performance compared to models not specialized for super-resolution for 3D images.

METHODS & MODELS: We work to combine the best features from EDSR and DCSRN in order to achieve scalability with harmonized 3D diffusion MRI inputs. The structure of the Flexible Medical Image Super Resolution (FMISR) model follows that of DCSRN, but layers in the densely-connected blocks are replaced with the Residual Blocks of the EDSR model adapted for 3D convolution. This enables greater flexibility and performance as computationally heavy batch normalization is removed, therefore reducing regularization.

Each low-resolution input image is passed through an initial convolution layer of filter size 3x3x3 and filter number 50. These outputs are then passed through a densely connected block with four Residual Block units, with each consisting of a convolutional pathway of two convolutional layers with filter size 3x3x3 and filter number 50 sandwiching a ReLU activation layer. The final convolutional layer gives the SISR output.

Training and testing are performed on 173 RISH Feature maps (orders 0, 2, 4), with the dataset split into 143 image volumes for training and 30 image volumes reserved for testing. These maps are calculated from data from the Human Connectome Project. The metrics of PSNR and SSIM [5] are used for evaluation.

$$\mathrm{PSNR} = 20 \cdot \log\left(\frac{\mathrm{MAX\ SIGNAL}}{\sqrt{\mathrm{MSE}}}\right)$$

$$\mathrm{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

We explicitly train our model using inputs that have been degraded to differing amounts for scale flexibility. However, we are presented with a choice between inputs at discrete common interpolation levels and a continuous spectrum of interpolation levels. We differentiate FMISR into two separate training scenarios, FMISR and FMISR+. FMISR represents the model trained on a discrete set of interpolation level, and FMISR+ represents the model trained on randomly selected interpolation levels sourced from a continuous range. We test the models on testing environments of both discrete and continuous scale.

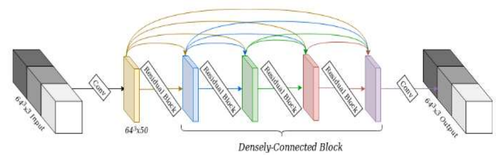
RESULTS:

Scale	NN	Bicubic	DCSRN	FMISR	FMISR+
Discrete	9.87	12.64	31.21	39.96	40.12
Continuous	10.52	14.23	32.16	40.76	40.84

Table 1: Average testing PSNR (dB)

Scale	NN	Bicubic	DCSRN	FMISR	FMISR+
Discrete	0.823	0.853	0.929	0.980	0.984
Continuous	0.834	0.863	0.932	0.986	0.986

Table 2: Average testing SSIM



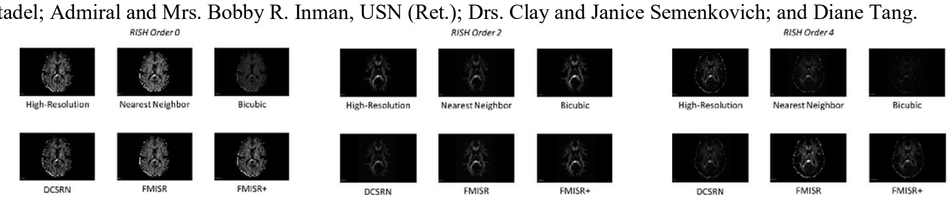
FMISR Block Diagram

DISCUSSION: We have presented a novel deep learning approach for single image super-resolution, FMISR, that takes a low-resolution patch as an input and generates a high-resolution patch as an output. The proposed FMISR model utilizes light weight Residual Block and skip connect architecture in order to maximize parameter space and increase effective receptive field.

Our approach outperforms baseline interpolation methods, nearest neighbor and bicubic interpolation, as well as a leading deep learning model for 3D medical image super-resolution, DCSRN, when evaluated in flexible scale scenarios using the widely popular image restoration metrics of PSNR and SSIM.

REFERENCES: [1] Greenspan, 2008, *The Comp. Journal* 52(1):43-63 [2] Lim, 2017, *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops* 1:3 [3] Ledig, 2016, *arXiv Preprint* [4] Chen, 2018, *ISBI 2018*:739-742 [5] Wang, 2004, *IEEE Transactions on Image Processing* 13(4):600-612

ACKNOWLEDGEMENTS: Dr. Suheyla Karayumak of the Psychiatric Neuroimaging Laboratory at Brigham and Women's Hospital, Harvard Medical School for her mentorship. Annelise Silva, the Research Assistant I worked with, for ensuring timely execution. Dr. Jenny Y. Sendova for her tutelage and proofreading. The Research Science Institute, the Center for Excellence in Education, and the Massacusetts Institute of Technology for their gracious sponsorship. My sponsors: Citadel; Admiral and Mrs. Bobby R. Inman, USN (Ret.); Drs. Clay and Janice Semenkovich; and Diane Tang.



Example testing outputs