

Multiscale Super-Resolution for Harmonized Diffusion Magnetic Resonance Imaging

Pratham Soni

`psoni@mit.edu`

under the direction of

Dr. Suheyla Karayumak

Psychiatry Neuroimaging Laboratory

Brigham & Women's Hospital, Harvard Medical School

Research Science Institute

July 27, 2019

Abstract

Though high resolution medical imaging provides detailed anatomical information, it comes with the cost of higher scan times and patient discomfort. The single image super-resolution (SISR) task aims to generate a high-resolution image from a low-resolution one. In this paper, we explore a novel deep learning method, Flexible Medical Image Super-Resolution (FMISR), for the SISR task as it pertains to Rotation Invariant Simple Harmonic (RISH) features derived from diffusion MRI data. In testing workloads with variable scale, FMISR delivers superior performance when compared to other super-resolution methods using standard metrics for measuring the quality of image restoration.

Summary

Super-Resolution is a technique for artificially enhancing the quality of images. In this paper, we propose a novel deep learning model for super-resolution that applies to medical imaging. This model, Flexible Medical Image Super-Resolution (FMISR), is constructed in order to restore images at a variety of degradation levels. Through testing, we show FMISR outperforms other super-resolution techniques using standard metrics for measuring the quality of image restoration.

1 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, 2, 3, 4].

Deep learning methods have become prominent solutions for the super-resolution problem (Section 2.2). 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, Section 2.1) feature maps generated from Diffusion MRI data.

Diffusion MRI, or dMRI for short, is a unique method for mapping white matter composition (tractography, Figure 1) in brain tissue. In making a dMRI, the relative strength of water movement is recorded. Water diffusion is restricted through white matter, and as a result, the signal strength decays faster in white matter tissue. The raw dMRI data gives signal values in the direction of a multitude of directions and are represented by scalar values on a sphere located at each voxel (3D pixel). From that data, RISH features can be calculated that serve to standardize measurements across varying imaging machines and input parameters.

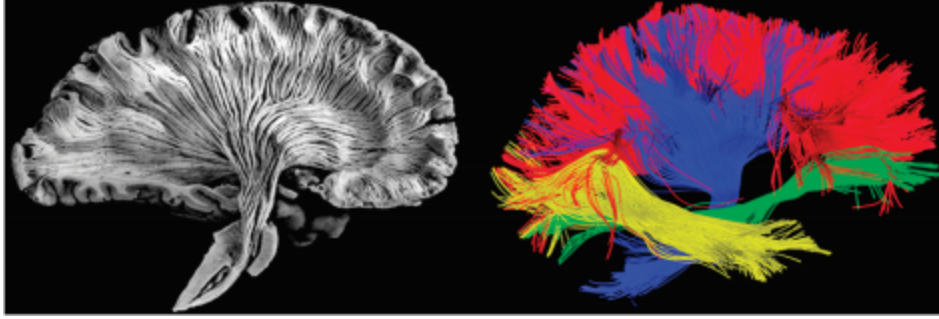


Figure 1: Tractography of white matter fibers [5]

In light of 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 (Section 6).

2 Background

2.1 Rotation Invariant Spherical Harmonics

There is an inherent flaw in dMRIs in that it is difficult to standardize calculations across different imaging machines and input parameters. This is where the technique of harmonization is extremely effective. Harmonization uses Spherical Harmonics (SHs, Figure 2) in order to decompose the raw dMRI signal into standardized units. Spherical Harmonics are a set of functions on the sphere that act like sine and cosine on the circle. These can be used to decompose spherical data into simpler components in the same vein as Fourier Transformations on two dimensional functions. Diffusion signals are both symmetric and real, and as such, can be represented by even order harmonics.

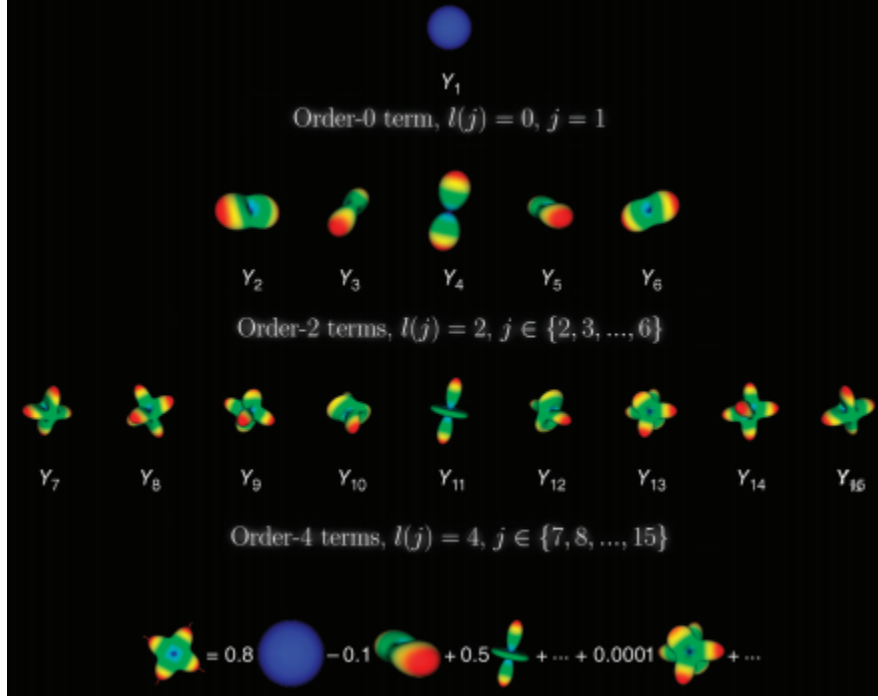


Figure 2: Illustration of various order Spherical Harmonics [5]

For a given signal \mathbf{S} we may represent it in SH space with the approximation $\mathbf{S} \approx \sum_l \sum_m C_{lm} Y_{lm}$ [6], where C_{lm} are the coefficients for SH basis functions Y_{lm} of order l and degree m . From there, Rotation Invariant Simple Harmonic (RISH) features can be calculated at each voxel using the following formula [7]:

$$\mathcal{F} = [\|C_0\|^2, \|C_2\|^2, \|C_4\|^2, \|C_6\|^2, \dots] \quad \text{where } \|C_l\|^2 = \sum_{m=1}^{2l+1} (C_{lm})^2.$$

These RISH features can be used to scale dMRI data without changing the diffusion directions of fibers [7]. Furthermore, it has been shown that low order RISH features capture much of the energy of the diffusion signal [8]. In addition, higher order features consist of mostly noise. As such, we utilize RISH features of order 0, 2, and 4 in this study. An example of a RISH feature map is shown in 3.

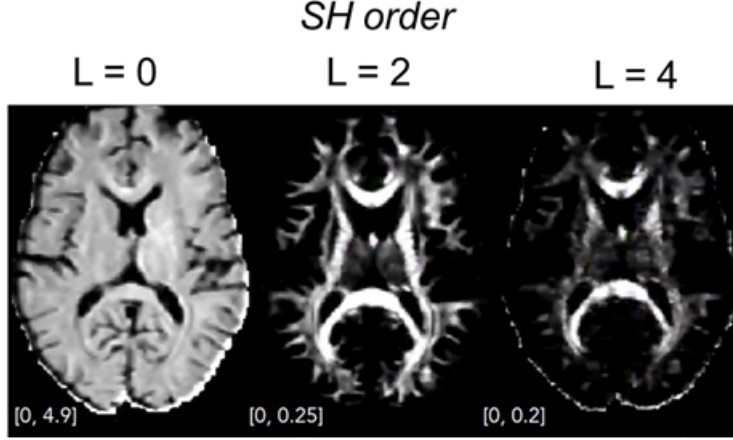


Figure 3: 2D slices of RISH feature map, adapted from [6]

2.2 Super-Resolution Problem & Deep Learning

Single image super-resolution (SISR) is a foundational area of study in the field of computer vision considered to be “a notoriously challenging ill-posed problem” [9]. The goal of SISR is to artificially generate a high resolution image from a low resolution one. We define the high-resolution ground truth as \mathbf{Y} and the low-resolution as \mathbf{X} , we have the relationship $\mathbf{X} = f(\mathbf{Y})$, where f is an interpolation method that degrades the quality of the original image. Our goal is develop a model M so that the generated image $\hat{\mathbf{X}} = M(\mathbf{X})$ approximates the original \mathbf{Y} . We can represent this relationship as $\mathbf{Y} = f^{-1}(\mathbf{X}) = \hat{\mathbf{X}} + \mathbf{r}$, where \mathbf{r} represents the residual between the generated image and the ground truth.

The SISR problem is challenging due to the multiplicity of the solution space. For a given low-resolution image there are many possibilities for the inverse functions, f^{-1} , that may be used to restore the high-resolution image. Traditionally, the issue of multiplicity is handled by constraining the solution space through the use of prior knowledge. The classic representative model is sparse coding [10, 11, 12, 13], which uses example-based algorithms. These

algorithms take pairs of low- and high- resolution patches and attempt to learn a mapping to translate the low-resolution patches into high resolution ones. Example pair techniques are able to be created for both general images as well as for a specific type of image, say medical images, depending on the training set of examples provided. However, the existence of a large enough dataset of pairs of correspondent low- and high-resolution images for example-based super-resolution methods is rare, and this limits the possibilities and applications of super-resolution.

Recently, deep learning techniques like Convolutional Neural Networks (CNNs [14]) have come to prominence as solutions to the SISR problem. The pathway for learning-based SISR consists of three fundamental steps [15]:

1. Extract image features from \mathbf{X}
2. Map the feature vector to a feature space
3. Reconstruct \mathbf{Y} from the feature space.

It has been shown that CNNs can handle all of these operations in a unified pipeline [14], thus enabling greater capacity for optimization. As a result, the model is able to learn the mapping between images through these steps by training in order to decrease \mathbf{r} . Furthermore, deep learning techniques use the technique of self-exemplars [11], where the input is generated from the ground truth, in order to build larger libraries for model training where independent high-resolution and low-resolution do not exist. With increased data, more complex functions can be modeled, and performance is improved.

Deep learning based super-resolution models, however, struggle with flexibility and generalizability. In general, models are trained for particular upscaling ratios (eg. $2\times$, $3\times$, $4\times$) as learning multiple ratios requires a larger number of features, which increases computational load and the size of models. This limitation necessitates that the model must be retrained

for each ratio. The ability to restore at continuous small non-integer scales is critical to real-world application to medical pipelines.

Recently, however, models like Very Deep Super-Resolution (VDSR) [16] (Figure 4) have shown the ability to perform in multiscale applications due to optimization of model structure that allows a larger number of total parameters. Increased parameters enable the model to better learn structural information for varied degradation levels.

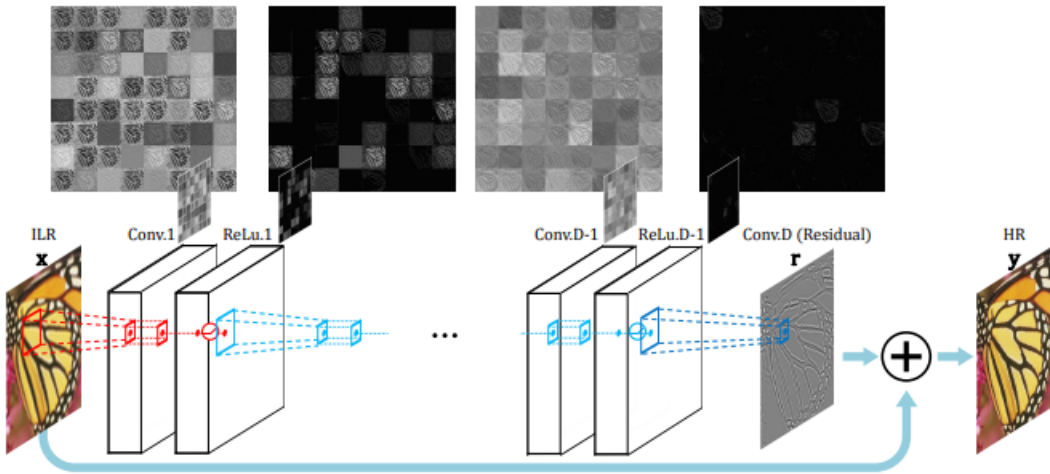


Figure 4: VDSR block diagram [16]

Applications of SISR in the medical field are also hindered by the fact that many medical images, like MRIs, are 3D, while most SISR models focus on applications to 2D images. SISR approaches for medical images in the past have utilized the idea of splitting the 3D volume into 2D slices and then recombining the 2D outputs into a final output. However, models like 3D Densely Connected Super-Resolution Networks (DCSRN, Section 3.2), have achieved high accuracy in 3D workloads by taking advantage of whole 3D volume inputs.

3 Related Work

3.1 Enhanced Deep Super-Resolution (EDSR)

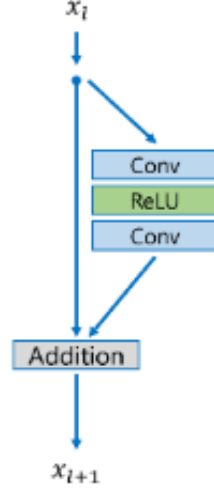


Figure 5: Structure of the EDSR Residual Block [17]

The central feature of EDSR is the Residual Block (Figure 5), which consists of two convolutional layers sandwiching a Rectified Linear Unit (ReLU) [18] activation layer, which can be quantified by the function $\max(0, x)$. Each Residual Block consists of such a system, which enables the construction of deeper models as each Residual Block is not very computationally demanding. However, increased number of parameters can lead to instability in training. In order to combat this, a scaling factor of 0.1 is used. If we represent the first and second convolutional layers as f and g respectively, then we have the relationship $x_{i+1} = x_i + 0.1 \cdot g(\max(0, f(x_i)))$. The role of the scaling factor is to minimize the impact of poor individual training batches in order to hinder potential cascade away from convergence.

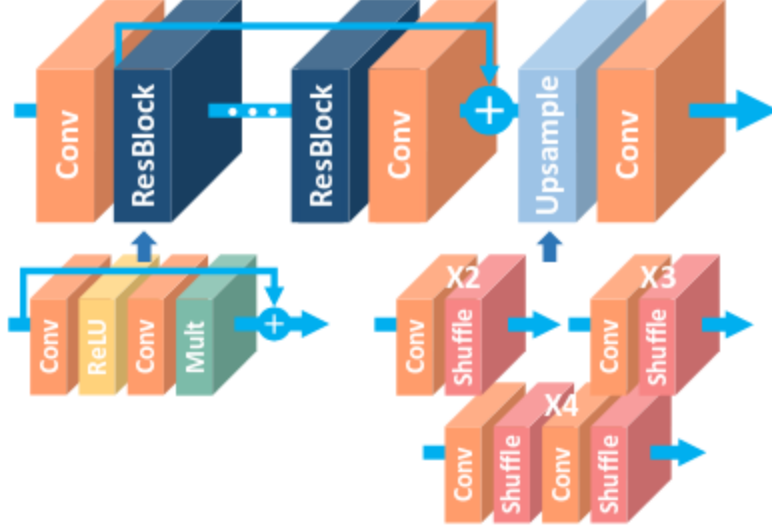


Figure 6: EDSR model showing layer architecture [17]

The EDSR model, Figure 6, is an extension of the standard Residual Network that has been trimmed for efficiency. By removing features like Batch Normalization layers from the SRResNet [19] 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.

This pruning allows for the implementation a larger number of total parameters. The higher number of parameters 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.

3.2 Densely Connected Super-Resolution Networks (DCSRN)

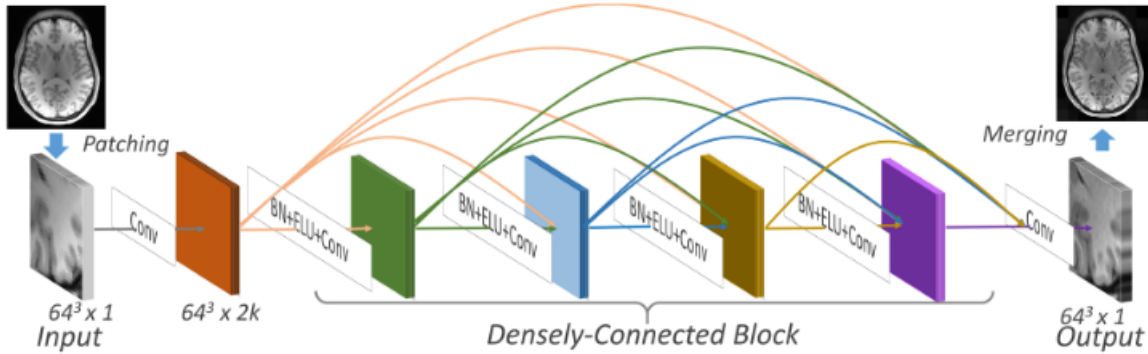


Figure 7: DCSRN model showing skip connections [15]

The DCSRN (Figure 7) 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.

4 Flexible Medical Image Super-Resolution (FMISR)

4.1 Proposed Model Architecture

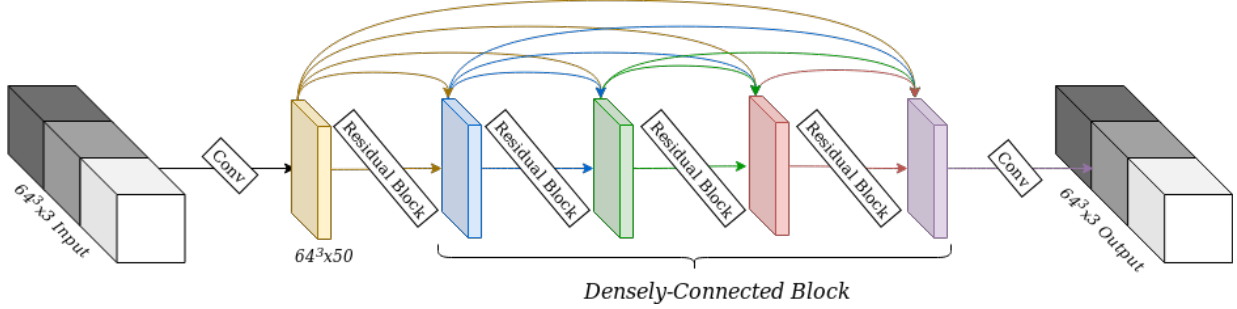


Figure 8: FMISR model block diagram with Residual Blocks and skip connections

The structure of FMISR (Figure 8) combines the best features from EDSR and DCSRN in order to achieve scalability with harmonized 3D diffusion MRI inputs. As seen in Figure 8, the structure of FMISR 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 $3 \times 3 \times 3$ 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 $3 \times 3 \times 3$ and filter number 50 sandwiching a ReLU activation layer. The final convolutional layer gives the SISR output.

4.2 Preprocessing

As an initial preprocessing step, input images are first normalized to 99th percentile values to account for noise in the image, and grayscale values are clipped to the range $[0, 1]$. Low-resolution generation is performed by downscaling and upscaling the patches with the use of bicubic interpolation in order to degrade image quality. This method differs from the process used for DCSRN. In DCSRN, a Fast Fourier Transform (FFT) is first applied, from which the outer parts are zeroed along two axes, and finally, inverse FFT is applied. Though this process mimics low-resolution MRI capturing, it does not allow for variable scale degradation. With the use of bicubic interpolation in both directions, we are able to specifically select the levels of degradation used for training and testing.

4.3 Training

From each original high-resolution training image, 100 three channel patches of dimension $64 \times 64 \times 64$ are selected randomly from the larger volume. Random selection represents the translational step. Each channel contains one of the SH features, and the model takes in all features in simultaneously. As a result, the model is forced to optimize for all order SH features together.

To learn the end-to-end mapping between the low-resolution and high-resolution images we must estimate our set of parameters Θ . To do this, we wish to minimize our loss function J between the reconstructed images using Θ , $F(\mathbf{X}; \Theta)$, and the corresponding ground-truth high-resolution image \mathbf{Y} .

5 Experiments

5.1 Evaluation Metric

As a measurement of loss we use MSE (Mean-Squared Error). For a set of low-resolution inputs $\{\mathbf{X}_i\}$ and corresponding high-resolution images $\{\mathbf{Y}_i\}$ we define MSE for a given Θ as

$$J(\Theta) = \frac{1}{n} \sum_{i=1}^n (F(\mathbf{X}_i; \Theta) - \mathbf{Y}_i)^2$$

where n is the number of images in the batch. Though we are not limited in our choice of loss function, we choose MSE as our loss function because a low MSE predicts a high Peak Signal to Noise Ratio (PSNR), a widely-used metric for quantitatively evaluating image restoration quality and is indicative of perceived quality as well:

$$\text{PSNR} = 20 \cdot \log \left(\frac{\text{MAX_SIGNAL}}{\sqrt{\text{MSE}}} \right)$$

where MAX_SIGNAL represents the highest possible intensity value in the image, in this case 1.0. MSE also serves as a favorable indicator of Structural Similarity Index (SSIM) [20], robust measure of image restoration focusing on the reconstruction of structural details [20]:

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

where μ_x , μ_y are the average of x , and y , respectively. Similarly, σ_x^2 , σ_y^2 are the variance of x , and y , respectively, and σ_{xy} is the covariance of x and y .

The weights are optimized using Adam optimizer [21] initialized with a learning rate, η , of 0.0001, exponential decay rates with β_1 to be .9 and β_2 to be .999, and ϵ of 10^{-8} in order to minimize MSE. The Adam optimizer utilizes adaptive moment estimates and individualized learning rates for each parameter. This provides better performance and efficiency than traditional stochastic gradient descent optimization. The weights at training iteration t are updated in the form

$$\Theta_t = \Theta_{t-1} - \eta \cdot \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}}$$

where \hat{m}_t, \hat{v}_t are the bias corrected first and second order moments, respectively. Each is derived from the stochastic gradient of the previous weights. The weights are initialized using a truncated normal distribution centered around 0.

5.2 Dataset

The image data for this study consists of 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. Each channel has dimensions $145 \times 174 \times 145$ and voxel size 1.25mm isotropic.

5.3 Training Details

The model is trained from scratch with 143,000 patches for 100 epochs (72 hours), where each epoch consists of 350 training iterations each, after which little improvement in loss is observed. At each epoch, a checkpoint is taken with 2 randomly selected patches used for validation. The patches are fed into the model in batches of size 2, and padding is used to prevent loss of edge details. Thus, the model output is the same size as the model input.

The model checkpoint with the lowest validation loss during training is saved, and the weights from this checkpoint are used on the test set to measure performance.

TensorFlow [22] is used for model implementation using Python. CUDA and CuDNN libraries are utilized for GPU-accelerated computation on an array of four Nvidia GTX 1080Ti GPUs.

5.4 Continuous versus Discrete Interpolation Ratios

The goal of this study is to develop a novel model that can be utilized for image restoration on a variety of degradation levels. As such, we explicitly train our model using inputs that have been degraded to differing amounts. However, we are presented with a choice between inputs at discrete common interpolation levels and a continuous spectrum of interpolation levels. We explore both options, and the results are enumerated in Section 6. Performance of the model is compared to an existing deep learning method, DCSRN, as well as baseline interpolation techniques, nearest neighbor (NN), and bicubic interpolation.

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.

In terms of the specific scales used, the resolution of medical images is represented by the voxel size. A smaller voxel size implies a higher resolution image. For FMISR, we utilize interpolation ratios of 1.2, 1.6, and 2.0, which represent quality degradation from 1.25mm isotropic to 1.5mm, 2.0mm, and 2.5mm, respectively. In each training step, one of these three values is selected for each image. For FMISR+, we randomly select the interpolation ration from the range $(1.0, 2.0]$.

6 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

We compare testing values for whole-image interpolation using the metrics of PSNR and SSIM. These PSNR and SSIM results are summarized in Tables 1 and 2. We can see that both FMISR and FMISR+ models outperform other image restoration methods, both deep learning based, as well as traditional, in scale-variable testing environments. Furthermore, FMISR+ outperforms or matches FMISR in all testing scenarios.

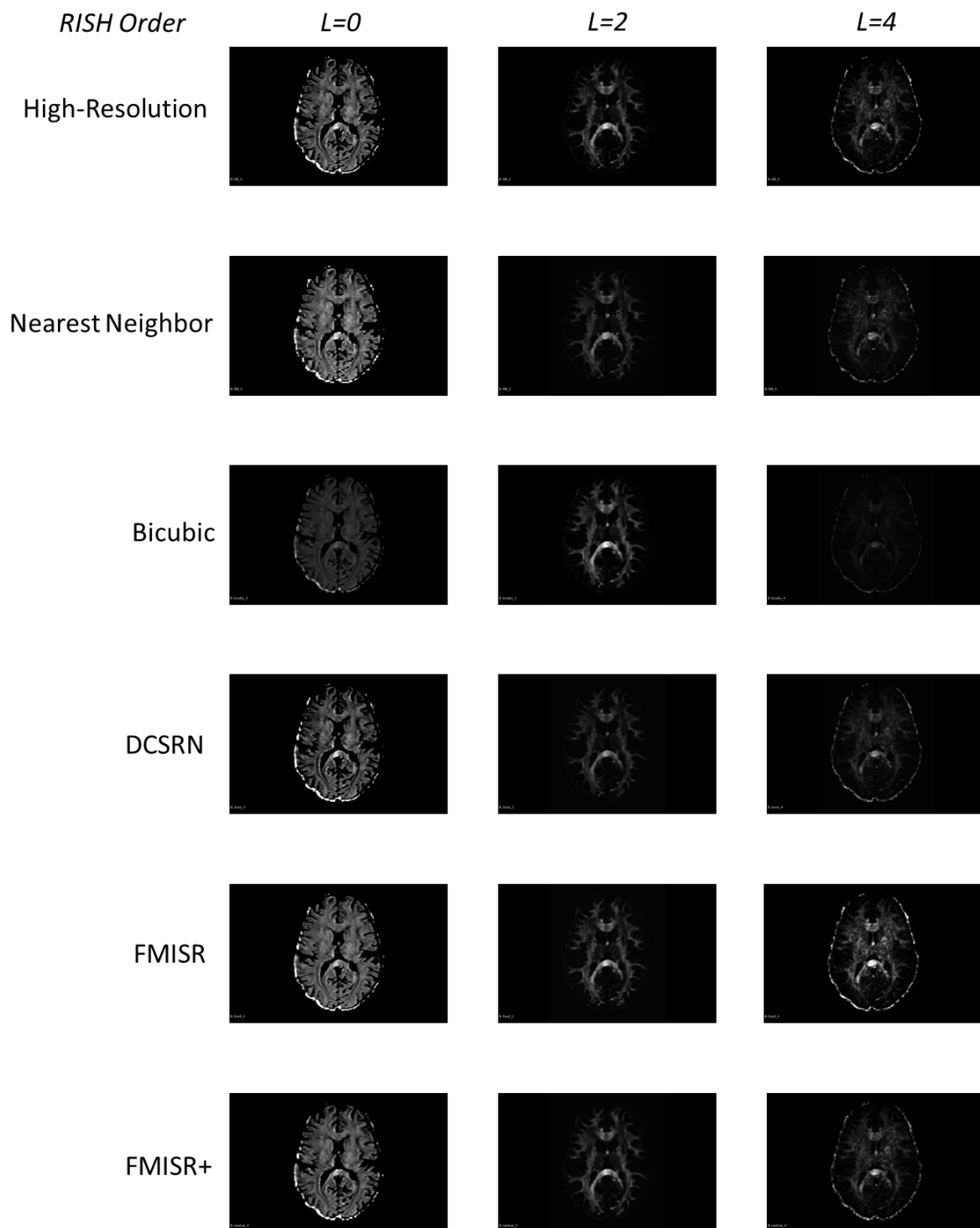


Figure 9: Upsampled outputs compared to high-resolution ground truth at each RISH order

7 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.

8 Future Work

Recent innovations in the field of deep learning have lead to breakthroughs in SISR of 2D images. Techniques such as FALSIR [23] and SRGAN [19] have demonstrated significant performance improvements. In particular, SRGAN utilizes a Generative Adversial Network, which can potentially help to recreate complex structures found in medical images. We plan to apply 3D-based optimizations, such as those found in FMISR, to these models in order to integrate them for 3D workloads.

In addition, there is significant opportunity to further improve the performance of FMISR with the use of better tuned hyperparameters. Hyperparameters can strongly effect the convergence and performance of models during training. However, it is difficult to find the optimal hyperparameters that will deliver the best testing results as small changes can induce large variation. As such, precise tuning can be extremely beneficial in terms of increased image reconstruction performance.

9 Acknowledgments

This work would be impossible without the mentorship of Dr. Suheyila Karayumak of the Psychiatric Neuroimaging Laboratory at Brigham and Women’s Hospital, Harvard Medical School. The resources of the laboratory were indispensable in the completion of this project. I would like to thank Annelise Silva, the Research Assistant I worked with for ensuring timely execution. I would also like to thank Dr. Jenny Y. Sendova for her tutelage and proofreading that has dramatically improved the quality of this paper. Finally, I would like to thank the Research Science Institute, the Center for Excellence in Education, and the Massachusetts Institute of Technology for their gracious sponsorship. I thank Cynthia Chen, Devrath Iyer, Cailyn Van Zyl, and Andrew Yang, my RSI peers, for their advice and encouragement. In addition, I thank my sponsors; Citadel; Admiral and Mrs. Bobby R. Inman, USN (Ret.); Drs. Clay and Janice Semenkovich; and Diane Tang, for financially providing me the opportunity to be at RSI. Finally, I would like to thank my parents and my teachers for supporting me every step of the way.

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] H. Greenspan. Super-resolution in medical imaging. *The Computer Journal*, 52(1):43–63, 2008.
- [2] P. Milanfar. *Super-resolution imaging*. CRC press, 2010.
- [3] K. J. Murphy, J. A. Brunberg, and R. H. Cohan. Adverse reactions to gadolinium contrast media: a review of 36 cases. *AJR. American journal of roentgenology*, 167(4):847–849, 1996.
- [4] A. J. Einstein. Medical imaging: the radiation issue. *Nature Reviews Cardiology*, 6(6):436, 2009.
- [5] M. Descoteaux. High angular resolution diffusion imaging (HARDI). *Wiley Encyclopedia of Electrical and Electronics Engineering*, pages 1–25, 1999.
- [6] S. C. Karayumak, S. Bouix, L. Ning, A. James, T. Crow, M. Shenton, M. Kubicki, and Y. Rathi. Retrospective harmonization of multi-site diffusion mri data acquired with different acquisition parameters. *NeuroImage*, 184:180–200, 2019.
- [7] H. Mirzaalian, A. de Pierrefeu, P. Savadjiev, O. Pasternak, S. Bouix, M. Kubicki, C.-F. Westin, M. E. Shenton, and Y. Rathi. Harmonizing diffusion mri data across multiple sites and scanners. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 12–19. Springer, 2015.
- [8] M. Descoteaux, E. Angelino, S. Fitzgibbons, and R. Deriche. Regularized, fast, and robust analytical q-ball imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 58(3):497–510, 2007.
- [9] W. Yang, X. Zhang, Y. Tian, W. Wang, J.-H. Xue, and Q. Liao. Deep learning for single image super-resolution: A brief review. *IEEE Transactions on Multimedia*, 2019.
- [10] H. Chang, D.-Y. Yeung, and Y. Xiong. Super-resolution through neighbor embedding. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages I–I. IEEE, 2004.
- [11] J.-B. Huang, A. Singh, and N. Ahuja. Single image super-resolution from transformed self-exemplars. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2015.
- [12] R. Zeyde, M. Elad, and M. Protter. On single image scale-up using sparse-representations. In *International conference on curves and surfaces*, pages 711–730. Springer, 2010.

- [13] J. Yang, K. Yu, Y. Gong, and T. Huang. Linear spatial pyramid matching using sparse coding for image classification. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 1794–1801. IEEE, 2009.
- [14] C. Dong, C. C. Loy, K. He, and X. Tang. Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence*, 38(2):295–307, 2015.
- [15] Y. Chen, Y. Xie, Z. Zhou, F. Shi, A. G. Christodoulou, and D. Li. Brain mri super resolution using 3d deep densely connected neural networks. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pages 739–742. IEEE, 2018.
- [16] J. Kim, J. Kwon Lee, and K. Mu Lee. Accurate image super-resolution using very deep convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1646–1654, 2016.
- [17] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee. Enhanced deep residual networks for single image super-resolution. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, volume 1, page 3, 2017.
- [18] V. Nair and G. E. Hinton. Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th international conference on machine learning (ICML-10)*, pages 807–814, 2010.
- [19] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. *arXiv preprint*, 2016.
- [20] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004.
- [21] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [22] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org.
- [23] X. Chu, B. Zhang, H. Ma, R. Xu, J. Li, and Q. Li. Fast, accurate and lightweight super-resolution with neural architecture search. *arXiv preprint arXiv:1901.07261*, 2019.