






Exploring the feasibility of zero-shot super-resolution in preclinical imaging

Omar A. M. Gharib¹, Samuel W. Remedios¹, Blake E. Dewey²,
Jerry L. Prince¹, and Aaron Carass¹

¹The Image Analysis and Communications Laboratory,
Johns Hopkins University, Baltimore, MD 21218, USA

²Department of Neurology,
Johns Hopkins School of Medicine, Baltimore, MD 21287, USA

Abstract. Preclinical imaging studies are vital to the research, development, and evaluation of new medical therapies. Images acquired during these studies often have high in-plane resolution but low through-plane resolution, resulting in highly anisotropic volumes that hamper downstream volumetric analysis. Additionally, since there are no image acquisition standards across studies, training data for conventional supervised super-resolution (SR) methods is limited. In this work, we compare two SR methods that do not require additional training data. The first is ECLARE, a self-SR approach that creates its own training data from in-plane patches drawn from the anisotropic volume. The second is Biplanar Denoising diffusion null space model (DDNM) Averaging (BiDA), a proposed method leveraging two independently pre-trained denoising diffusion probabilistic models and the DDNM posterior sampling technique. We evaluate both methods first on rat data at two scale factors ($2.5\times$ and $5\times$) and compare signal recovery and downstream task performance. We further evaluate these methods on a different species (mice) to measure their generalizability. Both methods experimentally resulted in good signal recovery performance, but only the images super-resolved by BiDA were accurately skullstripped downstream. Although both methods performed well on the in-domain rat data, BiDA did not fully generalize to the out-of-domain mouse data.

Keywords: Super-resolution · DDNM · DDPM · MRI · Preclinical

1 Introduction

Magnetic resonance imaging (MRI) is a prevalent modality in preclinical studies [14] owing to its excellent ability to image soft tissue in the head and brain [12, 36]. MRI is flexible due to the variety of available pulse sequences. Often, through-plane resolution is sacrificed due to the signal-to-noise ratio, scan time,

O. A. M. Gharib and S. W. Remedios—Equal contribution.

and the ethics of sedation duration or constraining the animal in the scanner. Thus, while anisotropic image volumes are usually acquired for in-plane reading, they are subsequently unsuitable for volumetric analysis. For the purposes of this paper, such anisotropic images are considered to be low-resolution (LR) along the through-plane axis.

Super-resolution (SR) is the estimation of the high-resolution (HR) image given the LR image by recovering the lost high-frequency information. B-spline [30] interpolation increases the number of voxels in the image, but does not estimate high-frequency information and therefore is not strictly SR. Nonetheless, interpolation approaches are popular due to their simplicity and ease of use. In contrast, learning-based approaches [3, 4, 8, 13, 15, 18, 23–25, 34, 35] aim to solve the SR problem directly. These approaches conventionally learn to estimate the HR image on average over a training dataset with an ℓ_p loss function. Optimizing the ℓ_p norm reduces the ability of such models to estimate high-frequency coefficients, resulting in images that are blurrier than desired. Learning-based SR approaches may also suffer from domain shift (e.g., SR scale factor, pathology, field strength, species imaged, etc.). This domain shift issue makes it difficult to generalize traditional ℓ_p -trained networks to data outside the training set, a problem that is pervasive in preclinical imaging.

Zero-shot SR approaches report strong generalization for SR on unseen data. These include methods based on compressed sensing with generative models [2, 17, 20, 33], internally supervised methods [22–25, 27], and null-space models [31, 32]. The generative model approaches relying on denoising diffusion probabilistic models (DDPMs) [9] have recently achieved the most realistic results. DiffusionMBIR [6] is an approach that leverages 2D generative models to estimate the underlying 3D volume using iterative reconstruction. This is one method to address the well-known “slice inconsistency” problem that occurs when using 2D methods for 3D image volumes [5]. Null-space models, also relying on generative models, have the additional benefit of being cycle-consistent with respect to the degradation operator; that is, only the lost information is estimated and the observed information present in the LR image is untouched. The most prominent of these methods is the denoising diffusion null-space model (DDNM) [32], which forms the basis of our proposed approach.

In this work, we train two independent DDPMs to generate HR 2D slices from orthogonal planes of rat MR images. Then, to mitigate slice inconsistency, we use DDNM inference on through-plane slices to generate two volumes that are averaged to produce a final result. We name our method Biplanar DDNM Averaging (BiDA) and evaluate its performance on withheld rat and mouse data. Specifically, we: 1) curate and preprocess HR 2D slices of rat data and train two DDPMs to approximate these distributions; 2) propose BiDA for through-plane SR; 3) evaluate BiDA on out-of-domain rat data and compare to other state-of-the-art SR methods with respect to signal recovery metrics and a downstream task; and 4) evaluate BiDA on out-of-domain mouse data.

2 Methods

First, we write our model of the degradation process. In 2D multi-slice MRI, the loss of resolution due to slice selection can be modeled as $y = (x * h) \downarrow_r$, where x is the HR through-plane signal, h is the slice selection profile (i.e., the through-plane point-spread function), $*$ is convolution, r is the down-sampling factor, and y is the LR through-plane signal. Since this equation is linear, it is often written as $y = Ax$, where A is the matrix form of this convolution.

Our proposed method, BiDA, is the application of two DDPMs in different orientations to generate two volume estimates of the HR image that are averaged into the final result.

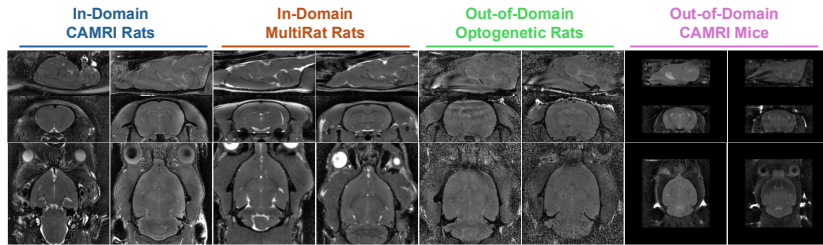


Fig. 1: Representative isotropic rodent images from each of the datasets used in this study. Here, in-domain data refers to data used in training the DDPM, and out-of-domain data is withheld from the DDPM training. At this resolution, the out-of-domain CAMRI mouse dataset requires zero-padding to have the same field-of-view extents as the available rat data.

Data We train and validate with publicly available preclinical MRI datasets. Example images from the datasets are shown in Fig. 1.

CAMRI Rat and Mouse Brain MRI The CAMRI Dataset [10] consists of high-resolution MRI scans of rat brains acquired on a Bruker 9.4T MRI system. We used 63 rat and 16 mouse T2w RARE scans that were acquired at a 200 μ m isotropic resolution.

OpenNeuro MultiRat Resting-State fMRI This dataset [7] includes 65 high-resolution isotropic MRI datasets from 46 institutions. We use a subset of 149 isotropic structural T2w scans acquired at five resolutions: 250 μ m (49 scans), 200 μ m (30), 190 μ m (20), 160 μ m (30), and 110 μ m (20).

OpenNeuro Optogenetic fMRI of STN DBS in Parkinson’s Disease Rats This dataset [16] consists of functional and structural MRI acquisitions from rodent models of Parkinson’s disease. We used a subset of nine isotropic structural T2w MRI scans acquired at 200 μ m isotropic resolution.

Preprocessing and Standardization To ensure consistency and optimize the dataset for training, we applied a series of data selection and preprocessing steps. 1) Only isotropic 3D scans were selected, with anisotropic images being

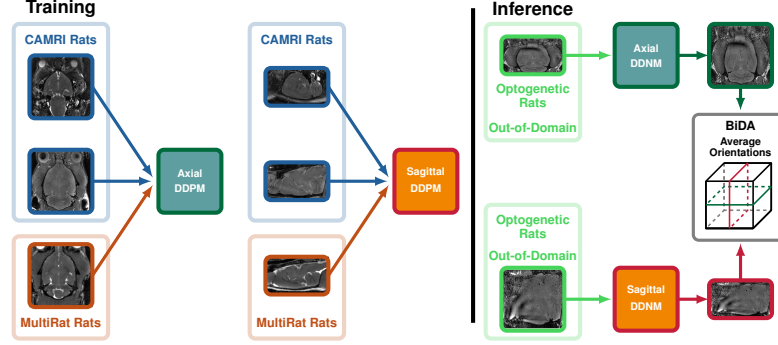


Fig. 2: **BiDA Method** (Left) The training phase generates two independent 2D slice-wise DDPMs. (Right) The inference phase, uses the DDPM models in DDNM frameworks to perform super-resolution on all slices independently. The resulting volumes are then stacked and averaged to produce the final result.

reserved for validation, resulting in 197 isotropic volumes. 2) N4 bias field correction [29] was applied to remove inhomogeneities. 3) To ensure consistency, all images were interpolated to a $200\mu\text{m}$ isotropic digital resolution using cubic B-spline interpolation. 4) Pixel intensities were linearly normalized to a range of $[-1, 1]$ by clipping extreme values based on the 1st and 99th percentiles. 5) Scans were reoriented to the standard RAS format and center-padded to $144 \times 144 \times 64$ voxels. 6) To train the 2D DDPM, 2D slices were extracted from axial and sagittal planes, yielding 12,608 axial and 28,370 sagittal slices.

DDPM Training We first trained two independent DDPMs on the in-domain rat data described above, one for each of the axial and sagittal orientations. We followed the standard procedure for diffusion training with forward Gaussian noise, but rather than $\hat{\epsilon}$ prediction, we used flow prediction [11]. Training data augmentation included small rigid transformations with a random rotation up to 10° and random translation up to 5 pixels in each direction.

DDNM Inference Here, we summarize the DDNM posterior sampling framework [32]. Given an estimate of A , for each time t in the diffusion process, every x_t is split into its range and null-space components,

$$x_{t,\text{range}} = A^\dagger A x, \quad \text{and} \quad x_{t,\text{null}} = (I - A^\dagger A) x. \quad (1)$$

Since $y = Ax$, DDNM replaces the range component to produce an estimate of the clean image,

$$\hat{x}_{0|t} = A^\dagger y + (I - A^\dagger A) x. \quad (2)$$

This is repeated for all $t = T, T-1, \dots, 1$.

At inference time, we use this procedure to independently super-resolve each through-plane slice of the anisotropic volume. Then, all super-resolved slices for each through-plane are stacked into their respective volumes, and then averaged to produce the final result. This process is outlined in Fig. 2, which shows both

the training and inference steps. We approximate A with h as a Gaussian kernel with full-width-at-half-max equal to the slice thickness and by assuming no slice gap in acquisition. During inference, we use the DDIM sampling scheme [28] with 16 timesteps.

3 Experiments and Results

We evaluated our approach using peak signal-to-noise ratio (PSNR) to assess reconstruction quality and a consistency Dice Similarity Coefficient (cDSC) for skull stripping consistency. For the cDSC, we used [10] to skull strip both the HR and super-resolved images, treating the mask generated from the HR image as a silver standard. We compared BiDA with ECLARE [25] (a self-SR method), and cubic interpolation [30]. On an NVIDIA A16, BiDA processes a $144 \times 144 \times 64$ voxel volume in ~ 10 s (< 10 GB vRAM) while ECLARE takes ~ 5 min at similar memory. For significance testing, we used a two-sided Wilcoxon signed rank sum test between ECLARE and BiDA on both resolutions and for the in- and out-of-domain data, with a significance level of 0.001.

We performed these evaluations at two SR scales $2.5\times$ (i.e., going from $500\mu\text{m}$ to $200\mu\text{m}$) and $5\times$ (i.e. going from $1000\mu\text{m}$ to $200\mu\text{m}$). These scales were obtained by simulating LR volumes from the HR in-domain rat isotropic volumes. Simulation was done by convolution with a slice selection profile generated with the Shinnar-Le Roux algorithm [19, 21, 26] of the full-width-half-max equal to the desired slice thickness ($500\mu\text{m}$ and $1000\mu\text{m}$ respectively) followed by subsampling by the desired slice spacing ($2.5\times$ and $5\times$, respectively).

First, we comment on the overall results in Figs. 5 and 7. Qualitatively, the results generated by BiDA appear sharper and have more details than the other methods. This appears to affect the skull stripping algorithm as well, which achieved a more accurate brain mask compared to the ECLARE results. This potentially could be due to the sensitivity of the downstream task to the textures in the image. Quantitatively, PSNR was lower for BiDA compared to ECLARE. This is expected, due to the perception-distortion tradeoff [1].

Next, we separate our analysis into in-domain and out-of-domain. The in-domain evaluation included 21 rat volumes from the same cohorts used to train BiDA, and the out-of-domain evaluations included data not used in training (nine volumes from the Optogenetic dataset) and the CAMRI Mice dataset. The out-of-domain evaluations represented different contrast and different species.

The in-domain rat PSNR results are shown in Fig. 3(a) and the out-of-domain results are shown in Fig. 3(b). Representative results for in-domain rat data are shown in the first and second rows of Fig. 5.

The results of the in-domain cDSC rat data are shown in Fig. 4(a) and the out-of-domain results are shown in Fig. 4(b). Representative out-of-domain rat data are shown in the third and fourth rows of Fig. 5.

The out-of-domain mice data results are shown in Fig. 6, with Fig. 6(a) showing PSNR results and Fig. 6(b) shown the cDSC results. PSNR was calculated on images that had their zero-padding removed to avoid artificial metric inflation.

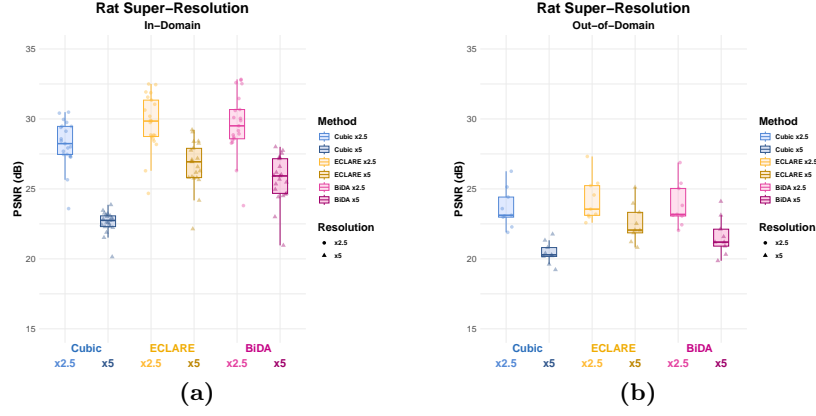


Fig. 3: The PSNR values for cubic interpolation (Cubic), ECLARE, and the proposed method (BiDA) are shown for the in-domain (a) and the out-of-domain (b) rat data at two different SR scales $500\mu\text{m}$ to $200\mu\text{m}$ ($\times 2.5$) and $1000\mu\text{m}$ to $200\mu\text{m}$ ($\times 5$).

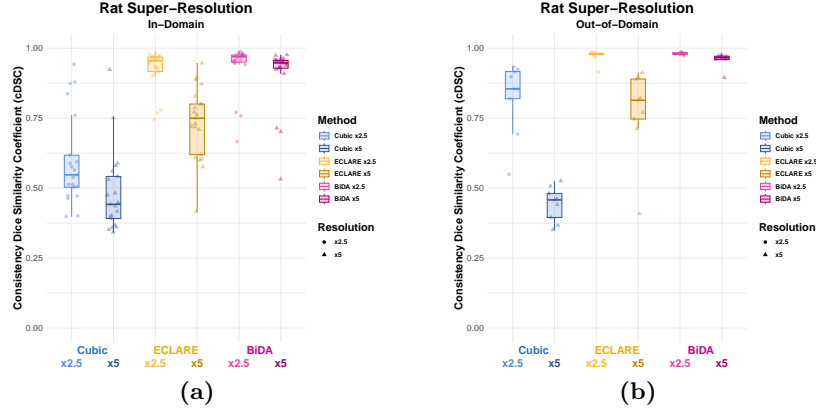


Fig. 4: The cDSC scores are shown for the in-domain (a) and the out-of-domain (b) rat data. The skull stripping masks are generated by [10] and evaluated at two SR scales.

Qualitative results for mouse data are shown in Fig. 7. The segmentation result of cubic interpolation entirely failed on average. This is potentially due to the downstream network’s sensitivity to texture shift at the border of the zero-padding or throughout the image. Nonetheless, both ECLARE and BiDA correct for this and achieve more accurate brain masks. Interestingly, despite the less realistic appearance of the BiDA results on the out-of-domain mice data, the resultant images are more amenable to segmentation than those generated by ECLARE.

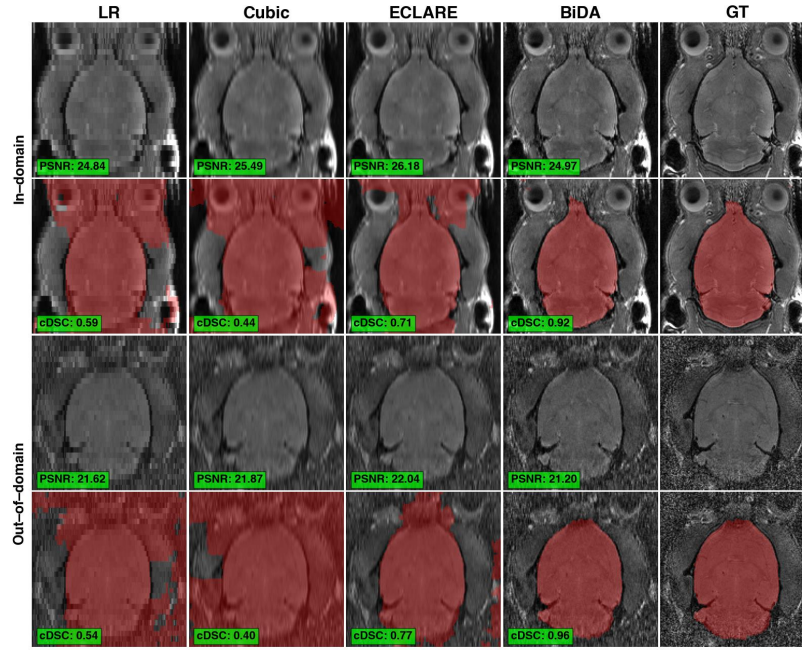


Fig. 5: Qualitative results from a representative in-domain and out-of-domain rat going from $1000\mu\text{m}$ to $200\mu\text{m}$ ($5\times$ scale factor). The red overlay is the skull stripping mask generated by [10].

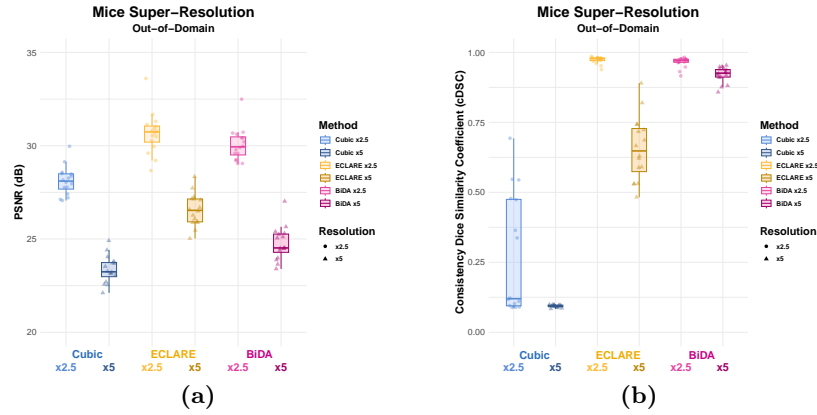


Fig. 6: (a) PSNR and (b) cDSC for the out-of-domain mouse data are shown. The skull stripping masks are generated by [10] and evaluated on two SR scales.

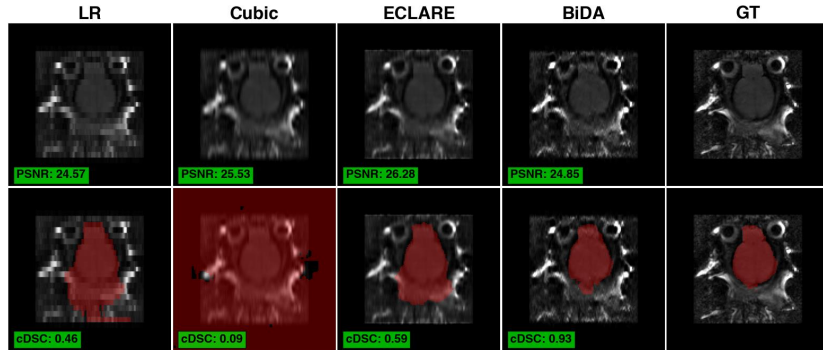


Fig. 7: Qualitative results from a representative out-of-domain mouse going from $1000\mu\text{m}$ to $200\mu\text{m}$ ($5\times$ scale factor). The red overlay is the skull stripping mask generated by [10]. PSNR was calculated after cropping the boundaries to avoid artificial metric inflation from zero-padding.

4 Conclusion

In this work, we explored the feasibility of zero-shot super-resolution (SR) in preclinical imaging using a denoising diffusion null-space model (DDNM). We introduced Biplanar DDNM Averaging (BiDA), a method leveraging two orthogonal 2D diffusion models for volumetric SR, and evaluated its performance across multiple datasets, including both in-domain and out-of-domain contrast and species experiments. Our results demonstrated that BiDA effectively preserves anatomical structure and outperforms traditional interpolation methods in both pixel-wise reconstruction fidelity and the downstream task performance of skull stripping. BiDA generally had worse mean PSNR than the ECLARE [25] method but comparable cDSC results. Notably, BiDA showed strong generalization within the rat dataset but had worse performance when applied to mouse data. This suggests species-specific adaptations may be necessary for broader applicability.

While both BiDA and ECLARE represent viable options for SR in preclinical imaging, some challenges remain. First, the domain shift between species, particularly in anatomical size and resolution, presents difficulties achieving consistent performance. Future work should explore techniques to enhance the generalization of diffusion-based models, including domain adaptation strategies and training on diverse multi-species datasets. Second, although our approach focuses on 2D slice-wise SR, a fully volumetric 3D diffusion model could further improve spatial coherence and reduce artifacts from slice-wise processing. Addressing these challenges will be crucial for advancing zero-shot SR methods in biomedical imaging, ultimately facilitating more robust and generalizable preclinical research.

Acknowledgments. This work is supported in part by the NIH under NIBIB grant R01-EB036013 (PI: J.L. Prince) and by the Congressionally Directed Medical Research Programs Grant W81XWH2010912 (PI: J.L. Prince). This material is partially supported

by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1746891 (S.W. Remedios).

Disclosure of Interests. The authors have no competing interests to declare.

References

1. Blau, Y., Michaeli, T.: The perception-distortion tradeoff. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 6228–6237 (2018)
2. Bora, A., et al.: Compressed sensing using generative models. In: International Conference on Machine Learning. pp. 537–546. PMLR (2017)
3. Chen, Y., et al.: Efficient and accurate MRI super-resolution using a generative adversarial network and 3D multi-level densely connected network. In: 21st International Conference Intervention (MICCAI 2018). pp. 91–99 (2018)
4. Chen, Y., et al.: MRI super-resolution with GAN and 3D multi-level DenseNet: smaller, faster, and better (2020), arXiv preprint arXiv:2003.01217
5. Choo, K., et al.: Slice-Consistent 3D Volumetric Brain CT-to-MRI Translation with 2D Brownian Bridge Diffusion Model. In: 27th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI 2024). pp. 657–667. Springer (2024)
6. Chung, H., et al.: Solving 3D inverse problems using pre-trained 2D diffusion models. In: 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). vol. 8, pp. 10–12 (2023)
7. Grandjean, J., et al.: MultiRat Rest Dataset (2022)
8. Han, Z., Huang, W.: Arbitrary scale super-resolution diffusion model for brain MRI images. *Computers in Biology and Medicine* **170**, 108003 (2024)
9. Ho, J., et al.: Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems* **33**, 6840–6851 (2020)
10. Hsu, L.M., et al.: Automatic skull stripping of rat and mouse brain MRI data using U-Net. *Front. Neurosci.* **14**, 568614 (2020)
11. Hu, J., et al.: Generative prediction of flow field based on the diffusion model. arXiv preprint arXiv:2407.00735 (2024)
12. Immonen, R.J., et al.: Distinct MRI pattern in lesional and perilesional area after traumatic brain injury in rat—11 months follow-up. *Experimental Neurology* **215**(1), 29–40 (2009)
13. Jog, A., et al.: Self super-resolution for magnetic resonance images. In: 19th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI 2016). vol. 9902, pp. 553–560 (2016)
14. Kiessling, F., Pichler, B.J. (eds.): *Small Animal Imaging: Basics and Practical Guide*. Springer, Berlin, Germany (2011)
15. Kim, Y., Son, D.: Noise conditional flow model for learning the super-resolution space. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 424–432 (2021)
16. Li, Y., et al.: Optogenetic functional magnetic resonance imaging reveals therapeutic feedforward circuits of the subthalamic nucleus deep brain stimulation in the hemiparkinsonian rat (2024)
17. Liu, G.H., et al.: I²SB: image-to-image Schrödinger bridge. In: Proceedings of the 40th International Conference on Machine Learning. pp. 22042–22062 (2023)

18. Lugmayr, A., et al.: SRFlow: Learning the super-resolution space with normalizing flow. In: 2020 European Conference on Computer Vision. pp. 715–732 (2020)
19. Martin, J., et al.: SigPy. RF: Comprehensive open-source RF pulse design tools for reproducible research (2020), ISMRM Annual Meeting
20. Menon, S., et al.: PULSE: Self-supervised photo upsampling via latent space exploration of generative models. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 2437–2445 (2020)
21. Pauly, J., et al.: Parameter relations for the Shinnar-Le Roux selective excitation pulse design algorithm (NMR imaging). *IEEE Trans. Med. Imag.* **10**(1), 53–65 (1991)
22. Remedios, S.W., et al.: Deep filter bank regression for super-resolution of anisotropic MR brain images. In: 25th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI 2022). vol. 13436, pp. 613–622 (2022)
23. Remedios, S.W., et al.: Self-supervised super-resolution for anisotropic MR images with and without slice gap. In: Workshop on Simulation and Synthesis in Medical Imaging (SASHIMI) held in conjunction with the. vol. 14288, pp. 118–128 (2023)
24. Remedios, S.W., et al.: Pushing the limits of zero-shot self-supervised super-resolution of anisotropic MR Images. In: Proceedings of SPIE Medical Imaging (SPIE-MI 2024). vol. 12926, pp. 29–35 (2024)
25. Remedios, S.W., et al.: ECLARE: Efficient cross-planar learning for anisotropic resolution enhancement. *arXiv preprint arXiv:2503.11787* (2025)
26. Shinnar, M., et al.: The synthesis of pulse sequences yielding arbitrary magnetization vectors. *Mag. Reson. Med.* **12**(1), 74–80 (1989)
27. Shocher, A., et al.: “Zero-shot” super-resolution using deep internal learning. In: IEEE Conf. on Computer Vision and Pattern Recognition. pp. 3118–3126 (2018)
28. Song, J., et al.: Denoising Diffusion Implicit Models. In: International Conference on Learning Representations (2021)
29. Tustison, N.J., et al.: N4ITK: Improved N3 Bias Correction. *IEEE Trans. Med. Imag.* **29**(6), 1310–1320 (2010)
30. Unser, M.: Splines: A perfect fit for signal and image processing. *IEEE Signal Processing Magazine* **16**(6), 22–38 (1999)
31. Wang, Y., et al.: GAN prior based null-space learning for consistent super-resolution. In: AAAI Conference on Artificial Intelligence. vol. 37(3), pp. 2724–2732 (2023)
32. Wang, Y., et al.: Zero-Shot Image Restoration Using Denoising Diffusion Null-Space Model. In: The 11th International Conference on Learning Representations (2023)
33. Wu, Z., et al.: TS-SR3: Time-strided denoising diffusion probabilistic model for MR super-resolution. In: Machine Learning in Medical Imaging (MLMI 2024). vol. 15241, pp. 248–258 (2024)
34. Wu, Z., et al.: Super-resolution of brain MRI images based on denoising diffusion probabilistic model. *Biomedical Signal Processing and Control* **85**, 104901 (2023)
35. Zhang, H., et al.: StackGAN: Text to photo-realistic image synthesis with stacked generative adversarial networks. In: 2017 IEEE International Conference on Computer Vision (ICCV). pp. 5907–5915 (2017)
36. Zhang, Z.G., et al.: Magnetic resonance imaging and neurosphere therapy of stroke in rat. *Annals of Neurology* **53**(2), 259–263 (2003)