Super-resolution of biomedical volumes with 2D supervision

Cheng Jiang[1*]   Alexander Gedeon[1*]   Yiwei Lyu[1]   Eric Landgraf[1]   Yufeng Zhang[1]

Xinhai Hou[1]   Akhil Kondepudi[1]   Asadur Chowdury[1]   Honglak Lee[1]   Todd Hollon[1]

[1]University of Michigan   [*]Equal Contribution

{chengjia,tocho}@umich.edu   https://mlins.org/msdsr/

Abstract

Volumetric biomedical microscopy has the potential to increase the diagnostic information extracted from clinical tissue specimens and improve the diagnostic accuracy of both human pathologists and computational pathology models. Unfortunately, barriers to integrating 3-dimensional (3D) volumetric microscopy into clinical medicine include long imaging times, poor depth/z-axis resolution, and an insufficient amount of high-quality volumetric data. Leveraging the abundance of high-resolution 2D microscopy data, we introduce masked slice diffusion for super-resolution (MSDSR), which exploits the inherent equivalence in the data-generating distribution across all spatial dimensions of biological specimens. This intrinsic characteristic allows for super-resolution models trained on high-resolution images from one plane (e.g., XY) to effectively generalize to others (XZ, YZ), overcoming the traditional dependency on orientation. We focus on the application of MSDSR to stimulated Raman histology (SRH), an optical imaging modality for biological specimen analysis and intraoperative diagnosis, characterized by its rapid acquisition of high-resolution 2D images but slow and costly optical z-sectioning. To evaluate MSDSR's efficacy, we introduce a new performance metric, SliceFID, and demonstrate MSDSR's superior performance over baseline models through extensive evaluations. Our findings reveal that MSDSR not only significantly enhances the quality and resolution of 3D volumetric data, but also addresses major obstacles hindering the broader application of 3D volumetric microscopy in clinical diagnostics and biomedical research.

1. Introduction

Biomedical microscopy is an essential imaging method and diagnostic modality in clinical medicine and biomedical research. Digital pathology and whole-slide images (WSIs) are now ubiquitous in computational pathology, leading to an increased role of computer vision and ma-
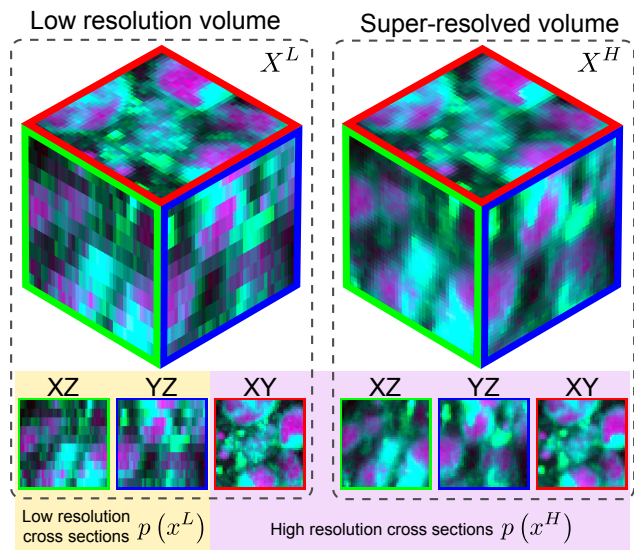

Figure 1. **Super-resolution of biomedical volumes with 2D supervision.** Volumetric microscopy images have a data distribution agnostic to tissue orientation and spatial dimension. Here, we present a method for leveraging this intrinsic characteristic by super-resolving low-resolution volumes using a conditional diffusion model trained on 2D high-resolution images.

chine learning-based approaches for analyzing microscopy data. Recent research has shown that 3-dimensional (3D) volumetric microscopy can improve the diagnostic yield of surgical specimens and diagnostic accuracy of both human pathologists and computational pathology models [14, 53]. Diagnostic histoarchitectural and cytologic structures are three-dimensional, such as chromatin structure, microvilli, and perivascular rosette formation [2]. The major barriers to integrating 3D volumetric microscopy into clinical medicine and biomedical research are (1) long imaging times, (2) poor depth (z-plane) resolution, and (3) insufficient high-quality 3D volumetric data. Importantly, high-quality, high-resolution, open-source 2D microscopy data is abundantly available, such as images from The Cancer

Genome Atlas (TCGA), The Digital Brain Tumor Atlas (DBTA) [42], and OpenSRH [22].

Here, we explore the open computer vision question of how to use high-resolution 2D microscopy data alone to improve the resolution of 3D volumetric microscopy, especially in the low-resolution z-plane or depth axis. We introduce *masked slice diffusion for super-resolution (MSDSR)*, which leverages the observation that all three spatial dimensions share the same underlying data-generating distribution for biological specimens. For example, biological specimens sampled at the time of surgery for cancer diagnosis lack a spatial orientation and, therefore, microscopy images obtained in any imaging plane are valid images for cancer diagnoses. The lack of orientation (e.g. up-down, left-right, front-back) allows super-resolution models trained in any given 2D plane, such as XY, to generalize to any other, such as XZ or YZ.

We evaluate MSDSR using a label-free optical imaging modality that is used for biological specimen analysis and intraoperative diagnosis, called stimulated Raman histology (SRH) [12]. SRH is ideally suited for volumetric super-resolution because high-resolution 2D images are readily obtained, but z-sections through the depth of the specimen are slow and costly to obtain. Our major contributions are:

1. We introduce *MSDSR*, a conditional diffusion-based, 3D volumetric super-resolution method that only requires 2D supervision.
2. We introduce a new volumetric, unpaired, perceptual quality metric, *SliceFID*.
3. MSDSR outperforms both interpolation and UNet baselines on image quality metrics, including SliceFID.

2. Related Work

2.1. Denoising diffusion models

Generating high-fidelity, high-resolution images is a challenging task in computer vision. Generative models have recently gained popularity and media attention for natural image generation given a prompt or condition [41, 45]. In particular, Denoising Diffusion Probabilistic Models (DDPMs) [18] have shown state-of-the-art results on image synthesis [8], using a UNet architecture to iteratively transform random noise into the learned data distribution. However, these models suffer from heavy computational requirements compared to earlier methods such as variational autoencoders (VAEs) [26] and generative adversarial networks (GANs) [15] due to the iterative sampling process. Denoising diffusion implicit models (DDIMs) [49] accelerate the sampling process of DDPM using non-Markovian processes. Latent diffusion models (LDMs) [43] further improve image quality and reduce computational requirements by performing the diffusion process on a latent space with lower dimensionality.

2.2. Image super-resolution

Super-resolution is the process of increasing the pixel resolution of an image. There exist several non-parametric methods for image super-resolution, such as nearest neighbor, linear, bilinear, and bicubic interpolation [48].

Regression-based methods such as SRCNN [10], LIIF [4], EDSR [31] and SwinIR [30] directly learn mappings from low-resolution to high-resolution images with a pixel-wise loss. Generation-based methods are trained to generate a new image based on the input low-resolution image, such as SRGAN [28]; CycleGAN [58], an image-to-image translation method, can be used to convert low-resolution images to high-resolution with unpaired data [25].

Image super-resolution with diffusion models is typically achieved with conditional diffusion, where the original image is used as a condition during the reverse diffusion process [44]. Stable Diffusion [43] enabled efficient high-resolution image super-resolution by combining latent diffusion and conditional diffusion. Recent work combines conditional diffusion with GANs for better quality and faster inference speed [56]. Aside from conditional diffusion, there are also diffusion-based super-resolution methods that either incorporate the input image in the denoising objective, such as DDRM [24], or directly perform iterative reverse diffusion on the low-resolution image, such as IDM [13], and ResShift [57].

2.3. Super-resolution for biomedical imaging

In radiological imaging, acquiring low-resolution images has the advantage of decreasing imaging time, radiation exposure, and motion artifacts. Super-resolution can then be used to interpolate information corrupted or lost during image acquisition. The majority of previous work on biomedical super-resolution has focused on radiological images, such as computed tomography (CT) or magnetic resonance imaging (MRI). Classical interpolation and reconstruction methods for medical images based on image processing methods and modeling the acquisition process have been studied [16]. More recently, deep learning-based methods have gained prominence, including UNet [9, 37], autoencoder [46, 47], and GAN frameworks [1, 17, 32, 34, 35, 59].

Progress has been made in enhancing 3D medical images, such as within radiology. Due to high computational complexity, many existing works combine 2D methods with some modifications for 3D consistency. Spatially aware interpolation network (SAINT) [40] utilized a 2D convolutional network for slice interpolation and a residual fusion network to ensure 3D consistency. Sood et al. [50] used a GAN to generate novel fields-of-view using neighboring slices as conditioning, resulting in through-plane super-resolution and a more detailed 3D volume. Kudo et al. [27] applied a conditional GAN to enhance the generative diver-
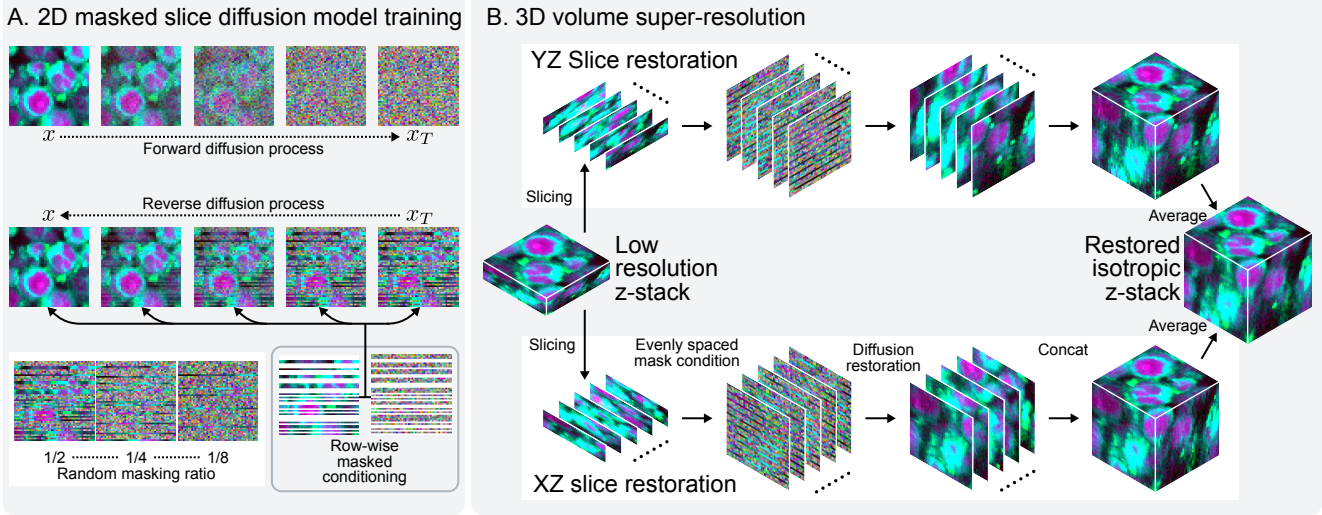
Figure 2. **MSDSR overview. A.** MSDSR is trained with a diffusion network conditioned on row-wise masks of the ground-truth high-resolution image. During the reverse diffusion process, a random masking ratio from 1/2 to 1/8 introduces these rows at random locations to give contextual structure when de-noising. The model then learns to interpolate the noised data in between the mask to produce a high-fidelity 2D image. **B.** During 3D inference, the low-resolution z-stack volume is sliced in both the XZ and YZ dimensions, producing low-resolution 2D images. The rows of these images are then treated as an evenly spaced mask interlacing random noise when individually fed into the model. These mixtures are then up-scaled by the model to produce high-resolution volumes and are then averaged together to generate a restored isotropic z-stack.

sity by incorporating the image information. Xia et al. [55] combined optical flow interpolation with a GAN to generate an auxiliary image as supervision to guide image synthesis. Finally, ArSSR [54] allowed arbitrary scale super-resolution of 3D MRIs via implicit neural representation.

Most recently, diffusion models have demonstrated remarkable effectiveness, and there have been a few studies to use diffusion models for 3D medical image super-resolution [3, 5–7, 29, 39, 52]. [7, 29, 39] are particularly relevant to our work, as they all attempted to use 2D diffusion models for 3D super-resolution. [7, 29] performed super-resolution on 3D MRI/CT by performing 2D super-resolution on perpendicular slices, but still require ground truth 3D high-resolution images to supervise model training. DiffuseIR [39] trained a diffusion model for super-resolution using 2D slices, but did not perform 3D reconstruction of the entire volume. Furthermore, the previous works focused on single-channel MRI and electron microscopy data, whereas SRH generates multi-channel images.

2.4. Deep Learning applications in SRH

There have been many existing works on applying deep learning methods to SRH images, most of them focusing on classification tasks such as brain tumor subtype classification [20, 38], molecular classification of diffuse glioma types [19], or whole-slide classification [21, 23]. Some prior works applied generative methods to denoising 2D SRH images [33, 36]. Lyu et al. [33] applied diffusion-based image restoration to 3D SRH z-stacks, but only to denoise each XY slice independently. To the best of our knowledge, this work is the first that attempts to super-resolve entire 3D SRH z-stack volumes.

3. Methods

The key motivation behind MSDSR is that different spatial dimensions of biomedical microscopy are equivalent and all views of 3D structures are orientation invariant. Thus, different slices of an isotropic 3D microscopy volume are 2D slices from the same underlying data-generating distribution. MSDSR leverages this observation and models the distribution using 2D images. The overall model architecture consists of the masked slice diffusion model and volume super-resolution inference, as described in figure 2.

Given a low-resolution volume $X^L \in \mathbb{R}^{n \times n \times \ell}$, where $\ell < n$, we want to predict an isotropic high-resolution volume $X^H \in \mathbb{R}^{n \times n \times n}$. Without high-resolution volume supervision, it is challenging to model $p(X^H)$ directly. We approximate the conditional probability of high-resolution volumes by modeling slices independently:

$$p\left(X^H \middle| X^L\right) \approx \prod_{i=1}^{n} p\left(X_{[i,:,:]}^H \middle| X_{[i,:,:]}^L\right); \tag{1}$$

$$p\left(X^H \middle| X^L\right) \approx \prod_{j=1}^{n} p\left(X_{[:,j,:]}^H \middle| X_{[:,j,:]}^L\right), \tag{2}$$

where $X^L_{[i,:,:]}, X^L_{[:,j,:]} \in \mathbb{R}^{\ell \times n}$ are YZ and XZ cross sections of the low-resolution volume X^L, respectively; and $X^H_{[i,:,:]}, X^H_{[:,j,:]} \in \mathbb{R}^{n \times n}$ are high-resolution YZ, XZ slices of X^H, respectively. Here, $p(X^H_{[i,:,:]}|X^L_{[i,:,:]})$ and $p(X^H_{[:,j,:]}|X^L_{[:,j,:]})$ are the conditional probability of the high-resolution YZ and XZ slices, given low-resolution YZ and XZ images with less rows in the Z dimension. Since XY, XZ, and YZ are from the same underlying distribution, it is equivalent to maximize the likelihood $p(x^H|x^L)$, where $x^H \in \mathbb{R}^{n \times n}$ is a high-resolution 2D image, and $x^L \in \mathbb{R}^{\ell \times n}$ is a lower resolution observation that can be simulated by downsampling or masking at training time. Since high-resolution 3D data is challenging and expensive to acquire, training with 2D images allows us to learn a better model by leveraging more data.

3.1. Masked slice diffusion

We train a DDPM to generate a high-resolution 2D image x^H, conditioned on the paired low-resolution image x^L. During training, we simulate the paired low-resolution image by removing rows of the high-resolution image. We obtain high-resolution images from the XY plane, and the trained model still applies well to XZ and YZ super-resolution due to the dimensional equivalence of SRH microscopy. Following the key results in [18], forward diffusion is a fixed process that gradually adds Gaussian noise to the image following a noise schedule β, for a total of T steps. At each step t,

$$x^H_t \sim \mathcal{N}\left(\sqrt{\bar{\alpha}_t}x^H, (1-\bar{\alpha}_t)\,\boldsymbol{I}\right), \tag{3}$$

where $\alpha_t = 1 - \beta_t$, and $\bar{\alpha}_t = \prod_{s=1}^t \alpha_s$.

During the reverse diffusion process, we condition x^H_t by interlacing it with the simulated low-resolution image. We sample $\ell \sim \text{Uniform}([\ell_{\min}, \ell_{\max}])$ number of rows to include as the condition, where $\ell_{\min}, \ell_{\max}$ are hyperparameters such that $0 < \ell_{\min} < \ell_{\max} < n$. $S \sim \text{Uniform}([1,n], \ell)$ is a set of random ℓ indices drawn without replacement for each row to be interlaced into x^H_t. We create a row-wise binary mask b to combine the partially denoised image at timestep t and the low-resolution image condition:

$$b = [\mathbf{1}_S(1), \ldots, \mathbf{1}_S(n)]^\top \tag{4}$$

$$c(x^H_t, x^H, b) = b * x^H + (1-b) * x^H_t, \tag{5}$$

where $\mathbf{1}(\cdot)$ is the indicator function, and $*$ denotes element-wise multiplication with broadcasting.

The masked slice diffusion model ϵ_θ in the reverse diffusion process is optimized using the variational lower bound on the negative log-likelihood with the objective function (i.e. the simplified objective from [18]):

$$\mathcal{L} = \mathbb{E}_{x^H, b, \epsilon \sim \mathcal{N}(\mathbf{0},\mathbf{I}), t}\left[\left\|b * \epsilon - \epsilon_\theta\left(c(x^H_t, x^H, b), t\right)\right\|\right]. \tag{6}$$

3.2. Volume super-resolution inference

To generate high-resolution volumes, we use our masked slice diffusion model to infer high-quality YZ and XZ slices along the X and Y axes:

$$\hat{X}^H_{YZ} = \text{Concat}\left[f\left(X^L_{[1,:,:]}\right), \ldots, f\left(X^L_{[n,:,:]}\right)\right]; \tag{7}$$

$$\hat{X}^H_{XZ} = \text{Concat}\left[f\left(X^L_{[:,1,:]}\right), \ldots, f\left(X^L_{[:,n,:]}\right)\right], \tag{8}$$

where f is the full reverse diffusion restoration process of our masked slice diffusion model, including sampling x_T, denoising and interlacing the observed low-resolution image at each time step. During the restoration process, each of the m slices along an axis is super-resolved independently. This independence does not reflect the physical structure we are trying to render, as neighboring slices are correlated. As a result, concatenation artifacts may form along the slices orthogonal to the inference axes. To eliminate the independence of inferences between planes, we use averaging to combine the volumes super-resolved in both directions:

$$\hat{X}^H = \frac{\hat{X}^H_{YZ} + \hat{X}^H_{XZ}}{2}. \tag{9}$$

As shown in section 5, this straightforward method achieves good empirical results in reducing inconsistencies and concatenation artifacts on the super-resolved volumes. A pseudocode of the MSDSR inference process is in algorithm 1.

Algorithm 1 MSDSR volume inference in PyTorch style.

```
def superresolve_along_axis(x):
    # x: transposed low res image of shape [n 3 l n]

    high_res_ims = []
    mask = arange(0, n, n // l) # assume l is a factor of n

    for i in range(n): # for each low res slice

        # interlace random noise with the observation
        x_T = randn_like(x[i])
        x_T[:, mask, :] = x[i]

        # full reverse diffusion restoration process
        high_res_ims.append(msdsr.restore(x_T))

    return stack(high_res_ims)

def superresolve_volume(xl):
    # xl: low res image of shape [3 n n l] (CHWZ)

    # transpose and superresolve xl in XZ slices
    xl_xz = rearrange(xl, "c h w z -> h c z w")
    xh_xz = superresolve_along_axis(xl_xz)
    xh_xz = rearrange(xh_xz, "h c z w -> c h w z")

    # transpose and superresolve xl in YZ slices
    xl_yz = rearrange(xl, "c h w z -> w c z h")
    xh_yz = superresolve_along_axis(xl_yz)
    xh_yz = rearrange(xh_yz, "w c z h -> c h w z")

    # return high resolution volume
    return (xh_yz + xh_xz) / 2
```

---

4. Experimentation

We evaluated MSDSR on a z-stacked volumetric stimulated Raman histology (SRH) dataset and compared the results to interpolation and UNet baselines.

4.1. Data description

Our z-stacked SRH dataset was collected using tumor specimens from patients who underwent brain tumor biopsy or resection at the University of Michigan. This study was approved by the Institutional Review Board (HUM00083059), and informed consent was obtained from each patient before imaging. The z-stacked SRH imaging of fresh surgical specimens follows the imaging protocol described in [22]. Each slide has a 0.5×0.5 mm^2 field of view and a 1000×1000 pixel resolution. The slides are imaged at an initial depth of 20 μm, and the laser focus is adjusted for each subsequent slice for virtual sectioning. Each z-stacked volume has a z-resolution of 1 μm for 20 z-sections. These z-stacks are not isotropic due to the physical limitations of optical sectioning, imaging time, and cost. Our data ingest and image processing pipeline also follows [22], with each whole-slide image volume being patched into $256 \times 256 \times 20$ pixels^3 tiles. Every patch is also subsequently denoised using RSCD [33] before model development and validation.

Our z-stacked dataset consists of 1129 whole-slide images from a total of 300 patients, spanning a wide range of brain tumor types, including glioma, meningioma, metastases, pituitary adenoma, schwannoma, and other less common central nervous system tumors. The dataset is split into training and validation sets, with 241 patients for model training and 59 patients for validation. Additionally, a key advantage of MSDSR is that it does not require training data to be volumetric. As a result, we also utilized a larger 2D SRH dataset consisting of SRH images from 1021 patients for masked slice training.

4.2. Implementation details

MSDSR architecture and training. Our MSDSR model is implemented using DDPM with 274M parameters. Based on the depth field of view of our SRH images, we crop our input images to 48×48. Our DDPM model utilizes a cosine noise scheduler with $T = 1000$ steps. The model is supervised with an L1 loss and optimized using AdamW, with a base learning rate of 10^{-7} and a cosine learn rate scheduler with a 10% warmup. The model was trained until convergence with an effective batch size of 256.

At training time, we condition the reverse diffusion process by randomly masking rows of high-resolution images, as described in section 3.1. The number of rows to be used as the condition is randomly drawn from a uniform distribution, with $\ell_{min} = 5$ and $\ell_{max} = 20$. These parameters were selected based on our z-stacked SRH dataset: $\ell_{max} = 20$ matches the resolution of existing data (at half resolution relative to XY slices); and $\ell_{min} = 5$ matches data collected with $4\times$ speed up, resulting in a z-resolution that is 1/8 of X- and Y-resolution.

MS-UNet baseline. We applied our masked slice training approach to a UNet architecture (MS-UNet). We use the same strategy to train the model using high-resolution images interlaced with random noise rows as input. An L1 loss between the prediction and the ground truth high-resolution 2D image was used to supervise model training. The UNet model was trained with a base learning rate of 10^{-3}, with other hyperparameters kept the same as MSDSR training.

End-to-end (E2E) UNet baseline. In addition to masked slice training, we also trained an end-to-end UNet to interpolate two different slides. The end-to-end UNet has the same architecture as the MS-UNet, except it takes a six-channel input consisting of two RGB slices X_i and X_{i+2}. The model outputs the slice X_{i+1} between the two input slices and is supervised with an L1 loss function. All hyperparameters are the same as MS-UNet training.

4.3. Evaluation protocol

Paired 2D evaluation. We evaluate the model by comparing super-resolved images with their high-resolution ground-truths. We uniformly mask 24, 36, and 42 rows from high-resolution 2D images for $2\times$, $4\times$, and $8\times$ super-resolution tasks, respectively. We report FID and SSIM as quantitative metrics.

Unpaired 3D evaluation. We evaluate 3D super-resolution using the z-stacked SRH images as described in section 4.1. It is challenging to apply FID, an unpaired metric, to volumetric SRH data because it requires a learned embedding space to measure the similarity of the data distribution. Prior work, such as 3D-FID [11] and FVD [51], use specialized volumetric feature extractors to compute embeddings, which is not feasible for our z-stacked SRH dataset. Thus, we propose an evaluation metric, *SliceFID*, to measure the quality of restored z-stack microscopy data.

SliceFID is motivated by the domain knowledge that XY, YZ, and XZ slices of an isotropic 3D volume are different views of the same underlying biological structures and follow the same data distribution. Therefore, we can utilize the FID to compute the distance between a set of high-quality two-dimensional images and each of the XY, YZ, and XZ slices of the generated image. The FID score in each axis informs the perceptual quality of images along each axis, and

	2× super-resolution		4× super-resolution		8× super-resolution	
	FID	SSIM	FID	SSIM	FID	SSIM
NN	90.0	0.714	244.8	0.426	450.4	0.228
Bilinear	31.6	0.727	134.0	0.442	314.6	0.241
MS-UNet	24.9	**0.825**	66.8	**0.628**	163.4	**0.419**
MSDSR (Ours)	**21.0**	0.678	**21.5**	0.486	**22.3**	0.284

Table 1. **Paired 2D evaluation metrics**. We present the FID and SSIM scores of our models and baselines on 2D paired data with a scaling factor ranging from 2× to 8×. These comparisons come from inference on high-resolution 2D images. While MS-UNet achieves a higher SSIM, images super-resolved by MS-UNet are perceptually blurry. NN, nearest neighbor, bilinear, bilinear interpolation.

SliceFID is defined as the average of these per-axis metrics:

$$\text{SliceFID}(x^H, \hat{X}) = \frac{1}{3}\left[\text{FID}(x^H, \text{slice}_{XY}(\hat{X})) + \right.$$
$$\left. \text{FID}(x^H, \text{slice}_{XZ}(\hat{X})) + \text{FID}(x^H, \text{slice}_{YZ}(\hat{X}))\right], \quad (10)$$

where $x^H \in \mathbb{R}^{k \times h \times w}$ is a set of k high-quality ground truth 2D images, $\hat{X} \in \mathbb{R}^{m \times h \times w \times z}$ is a set of m super-resolved volumes, and $\text{slice}_{XY}(\hat{X}) \in \mathbb{R}^{mz \times h \times w}$, $\text{slice}_{XZ}(\hat{X}) \in \mathbb{R}^{mh \times w \times z}$ and $\text{slice}_{YZ}(\hat{X}) \in \mathbb{R}^{mw \times h \times z}$ are the generated volumes in XY, XZ, and YZ slices, respectively.

5. Results

5.1. MSDSR paired 2D evaluation

In this section, we evaluate MSDSR on a paired 2D super-resolution task and compare it to interpolation and UNet baselines. Table 1 summarizes the quantitative metrics, and a panel of super-resolved examples with various numbers of conditioning rows is shown in figure 3.

Quantitatively, MSDSR outperforms all baseline methods on FID across all super-resolution tasks. MS-UNet achieves the best SSIM metric across all tasks but produces perceptually blurry images. MSDSR outperforms nearest neighbor and bilinear interpolation baselines in SSIM with a larger super-resolution scaling factor (4× and 8×).

Visually, MSDSR generates high-fidelity images similar to the paired ground truth, where NN and bilinear interpolation generate images with significant blurring and artifacts. MS-UNet produces overly smooth images, with missing details in cellular structures and background objects. MSDSR consistently recovers relevant cellular features (e.g., shape, chromatin, cytoplasm) with increasingly lower-resolution input images while maintaining realistic details, making it the only robust super-resolution method benchmarked.

5.2. MSDSR 3D Inference Evaluation

To evaluate the quality of generated z-stack volumes, we use SliceFID to assess the volumetric image quality along each of the XY, YZ, and XZ planes, as well as their average for a holistic evaluation. Table 2 shows SliceFID
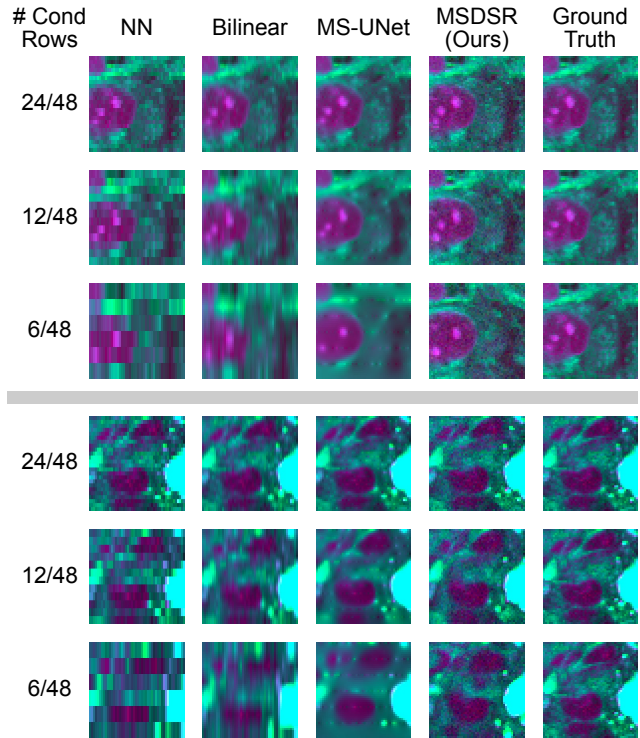

Figure 3. **Paired 2D evaluation.** We compare the images generated by MSDSR and other baselines to the paired ground truth image. # cond rows, number of conditioning rows, NN, nearest neighbor, bilinear, bilinear interpolation.

and its components, and figure 4 shows a sample super-resolved volume for each model, across three different super-resolution scaling factors.

Overall, MSDSR achieves the best SliceFID score across all super-resolution tasks. Non-parametric interpolation methods (i.e., NN and bilinear) have a low FID score along the XY plane because they naively interpolate information between the XY slices in the z-direction, leaving the input data intact. As a result, YZ and XZ images generated by these methods are unrealistic and contain jittering and stretching artifacts due to the sparsity of the input data on the plane. Both UNet-based models achieved better perfor-

	2× super-resolution FID				4x super-resolution FID				8× super-resolution FID			
	XY ⊗	YZ ⬦	XZ ⬦	SliceFID	XY ⊗	YZ ⬦	XZ ⬦	SliceFID	XY ⊗	YZ ⬦	XZ ⬦	SliceFID
NN	24.7	197.9	166.2	129.6	**24.8**	357.7	354.6	245.7	**25.1**	530.8	447.2	334.4
Bilinear	30.6	62.0	60.6	51.1	29.3	179.1	162.7	123.7	27.3	355.0	357.3	246.5
E2E UNet	43.9	73.0	63.4	60.1	58.7	179.8	167.9	135.4	70.9	311.9	298.7	227.2
MS-UNet	38.8	46.2	55.0	46.6	58.0	98.9	95.8	84.2	103.9	224.4	189.6	172.7
MSDSR (Ours)	**16.2**	**28.9**	**31.4**	**25.5**	25.5	**37.1**	**35.8**	**32.8**	107.4	**61.9**	**56.7**	**75.3**

Table 2. **3D super-resolution metrics.** We present the SliceFID score and its components for MSDSR and MS-UNet along with our baseline methods. NN, nearest neighbor, bilinear, bilinear interpolation, E2E UNet, end-to-end UNet.
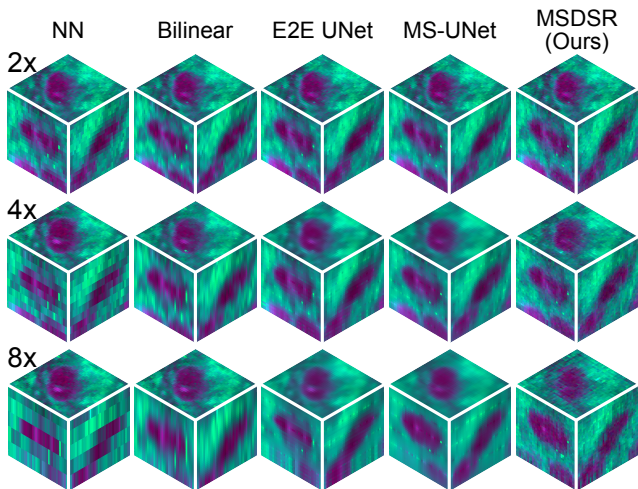

Figure 4. **3D super-resolution results.** We compare 3D volumetric super-resolution inference across three different input scalings. NN, nearest neighbor, bilinear, bilinear interpolation, E2E UNet, end-to-end UNet.
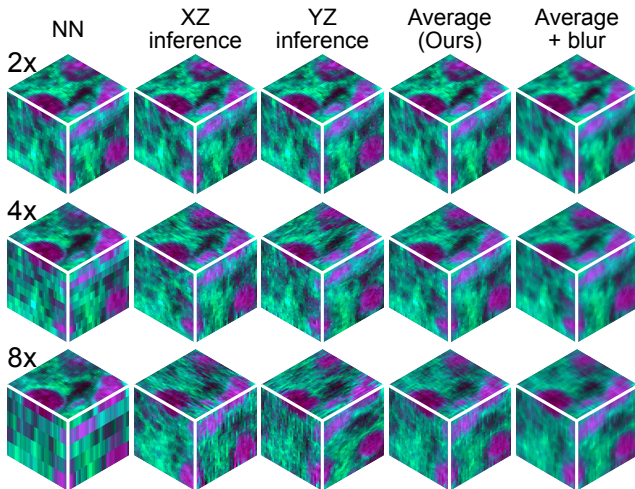

Figure 5. **Ablation study on inference direction and Gaussian blur.** We compare 3D volumetric super-resolution inference using ablated models across three different input scalings. NN, nearest neighbor, average, averaging XZ and YZ inference.

mance compared to the non-parametric methods, but they both generate overly smooth images, lacking detail in the nuclei and cytoplasm of the cells. The naive E2E UNet had worse results because it only models images in the XY plane and is only trained using a fixed 2-micron interval. When inferencing with a lower resolution condition, the E2E model relies on a recursive strategy to predict intermediate slices and amplifies any mistakes made during the process. While MS-UNet performs better, it still generates smoothed images that are not realistic. In comparison, MSDSR performs the best and generates high-fidelity volumes, with cellular and background textures remaining available. As condition image resolution decreases, MS-DSR still maintains consistent and reasonable images, albeit at a mildly worse quality.

Ablation studies. We investigate the effect of averaging inferences from both XZ and YZ directions, as well as using Gaussian blur as a post-processing step. Quantitative

SliceFID metrics are reported in tables 3 and 4, respectively. Examples of super-resolved volumes are shown in figure 5. Inferencing along a single direction results in noticeable stitching artifacts for 4× and 8× super-resolved images, especially in the planes orthogonal to the inference plane, both visually and quantitatively. This results from aggregating independent inferences that are plausible solutions during individual slice inference, given a lower resolution condition. Averaging XZ and YZ alleviates the artifact, but does not completely remove it. Applying Gaussian blur is another way to reduce the stitching artifact, especially in 8× super-resolution task, but it also reduces the overall sharpness of the prediction.

6. Conclusion

Our study presents a novel approach, masked slice diffusion for super-resolution (MSDSR), to enhance the resolution of 3D volumetric biomedical images utilizing 2D supervision. We demonstrated that MSDSR can leverage the

	2× super-resolution FID				4× super-resolution FID				8× super-resolution FID			
	XY	YZ	XZ	SliceFID	XY	YZ	XZ	SliceFID	XY	YZ	XZ	SliceFID
MSDSR-XZ	14.5	28.1	**27.9**	23.5	23.5	38.8	**28.1**	30.1	88.5	103.9	**25.6**	72.7
MSDSR-YZ	15.1	**27.1**	28.0	23.4	26.6	**28.6**	36.3	30.5	102.8	**27.3**	100.7	76.9

Table 3. **3D ablation FID metrics.** We present the SliceFID score and its components for MSDSR ablated for inference in only a single dimension. MSDSR-XZ represents performing inference solely in the XZ dimension, with the analogous MSDSR-YZ. We can observe a significant performance drop on the planes orthogonal to the inference plane, most likely due to a stitching artifact.

SR scale	Method	XY (Δ)	YZ (Δ)	XZ (Δ)	SliceFID (Δ)
	MSDSR-XZ + blur	23.5 (+9.0)	34.3 (+6.2)	37.9 (+10.0)	31.9 (+8.4)
2×	MSDSR-YZ + blur	24.0 (+9.0)	38.5 (+11.4)	37.3 (+9.4)	33.3 (+9.9)
	MSDSR + blur	30.7 (+14.5)	44.2 (+15.3)	45.4 (+14.0)	40.1 (+14.6)
	MSDSR-XZ + blur	26.3 (+2.8)	39.1 (+0.3)	34.7 (+6.7)	33.4 (+3.3)
4×	MSDSR-YZ + blur	28.5 (+1.9)	36.1 (+7.5)	41.0 (+4.7)	35.2 (+4.7)
	MSDSR + blur	32.7 (+7.2)	43.5 (+6.3)	45.0 (+9.2)	40.4 (+7.6)
	MSDSR-XZ + blur	72.5 (-16.0)	88.5 (-15.3)	30.2 (+4.6)	63.7 (-8.9)
8×	MSDSR-YZ + blur	84.7 (-18.1)	30.9 (+3.5)	79.7 (-21.0)	65.1 (-11.9)
	MSDSR + blur	101.2 (-6.2)	62.1 (+0.2)	59.5 (+2.8)	74.3 (-1.1)

Table 4. **3D Gaussian blur ablation FID metrics.** $3 \times 3 \times 3$ Gaussian blur post-processing degrades the model performance for 2× and 4× super-resolution, but offers a boost for the 8× task, especially in the planes orthogonal to the inference plane. Differences to FID scores before Gaussian blurring are reported in parentheses (Δ). SR scale, super-resolution scale.

inherent similarity in data-generating distributions across spatial dimensions of biological specimens, enabling effective generalization from 2D to 3D. Our proposed method significantly surpasses traditional interpolation and UNet baselines across various image quality metrics, notably through our newly introduced SliceFID metric, emphasizing MSDSR's efficacy in generating high-quality, realistic volumetric reconstructions from low-resolution inputs.

Limitations. While MSDSR has shown promising results, it is not without limitations. The primary challenge lies in the method's current reliance on synthesizing slices independently, which can lead to inconsistencies in 3D volumetric reconstructions. This approach, while effective for enhancing individual slices, does not fully exploit the spatial correlations inherent in 3D structures, potentially affecting the consistency of the reconstructed volumes. Furthermore, the computational demands of our method, particularly for high-resolution volumetric data, pose challenges for real-time clinical applications.

Broader impact. The broader impact of MSDSR extends beyond the technical achievements in biomedical imaging. By significantly improving the resolution and quality of 3D volumetric microscopy, our work has the potential to advance diagnostic accuracy, enhance the understanding of complex biological structures, and facilitate the development of novel therapeutic strategies. Furthermore, by reducing the dependency on high-resolution 3D data, MSDSR can democratize access to advanced imaging technologies, particularly in resource-constrained settings, ultimately contributing to the global efforts to bridge the healthcare divide.

Acknowledgements and Competing Interests

References

[1] Waqar Ahmad, Hazrat Ali, Zubair Shah, and Shoaib Azmat. A new generative adversarial network for medical images super resolution. *Scientific Reports*, 12(1):9533, 2022. 2

[2] Stavros J Baloyannis and Ioannis S Baloyannis. The fine structure of ependymomas. *CNS oncology*, 3(1): 49–59, 2014. 1

[3] Trevor J Chan and Chamith S Rajapakse. A super-resolution diffusion model for recovering bone microstructure from ct images. *Radiology: Artificial Intelligence*, 5(6):e220251, 2023. 3

[4] Yinbo Chen, Sifei Liu, and Xiaolong Wang. Learning continuous image representation with local implicit image function. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 8628–8638, 2021. 2

[5] Hyungjin Chung and Jong Chul Ye. Score-based diffusion models for accelerated mri. *Medical image analysis*, 80:102479, 2022. 3

[6] Hyungjin Chung, Eun Sun Lee, and Jong Chul Ye. Mr image denoising and super-resolution using regularized reverse diffusion. *IEEE Transactions on Medical Imaging*, 42(4):922–934, 2022.

[7] Hyungjin Chung, Dohoon Ryu, Michael T McCann, Marc L Klasky, and Jong Chul Ye. Solving 3d inverse problems using pre-trained 2d diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 22542–22551, 2023. 3

[8] Prafulla Dhariwal and Alexander Nichol. Diffusion models beat gans on image synthesis. *Advances in neural information processing systems*, 34:8780–8794, 2021. 2

[9] Pak Lun Kevin Ding, Zhiqiang Li, Yuxiang Zhou, and Baoxin Li. Deep residual dense u-net for resolution enhancement in accelerated mri acquisition. In *Medical Imaging 2019: Image Processing*, pages 110–117. SPIE, 2019. 2

[10] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence*, 38(2):295–307, 2015. 2

[11] Zolnamar Dorjsembe, Hsing-Kuo Pao, Sodtavilan Odonchimed, and Furen Xiao. Conditional diffusion models for semantic 3d brain mri synthesis. *IEEE Journal of Biomedical and Health Informatics*, 2024. 5

[12] Christian W Freudiger, Wei Min, Brian G Saar, Sijia Lu, Gary R Holtom, Chengwei He, Jason C Tsai, Jing X Kang, and X Sunney Xie. Label-free biomedical imaging with high sensitivity by stimulated raman scattering microscopy. *Science*, 322(5909):1857–1861, 2008. 2

[13] Sicheng Gao, Xuhui Liu, Bohan Zeng, Sheng Xu, Yanjing Li, Xiaoyan Luo, Jianzhuang Liu, Xiantong Zhen, and Baochang Zhang. Implicit diffusion models for continuous super-resolution. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10021–10030, 2023. 2

[14] Adam K Glaser, Nicholas P Reder, Ye Chen, Erin F McCarty, Chengbo Yin, Linpeng Wei, Yu Wang, Lawrence D True, and Jonathan TC Liu. Light-sheet microscopy for slide-free non-destructive pathology of large clinical specimens. *Nature biomedical engineering*, 1(7):0084, 2017. 1

[15] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks. *Communications of the ACM*, 63(11):139–144, 2020. 2

[16] Hayit Greenspan. Super-resolution in medical imaging. *The computer journal*, 52(1):43–63, 2009. 2

[17] Yuchong Gu, Zitao Zeng, Haibin Chen, Jun Wei, Yaqin Zhang, Binghui Chen, Yingqin Li, Yujuan Qin, Qing Xie, Zhuoren Jiang, et al. Medsrgan: medical images super-resolution using generative adversarial networks. *Multimedia Tools and Applications*, 79: 21815–21840, 2020. 2

[18] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020. 2, 4

[19] Todd Hollon, Cheng Jiang, Asadur Chowdury, Mustafa Nasir-Moin, Akhil Kondepudi, Alexander Aabedi, Arjun Adapa, Wajd Al-Holou, Jason Heth, Oren Sagher, et al. Artificial-intelligence-based molecular classification of diffuse gliomas using rapid, label-free optical imaging. *Nature Medicine*, 29 (4):828–832, 2023. 3

[20] Todd C Hollon, Balaji Pandian, Arjun R Adapa, Esteban Urias, Akshay V Save, Siri Sahib S Khalsa, Daniel G Eichberg, Randy S D'Amico, Zia U Farooq, Spencer Lewis, Petros D Petridis, Tamara Marie, Ashish H Shah, Hugh J L Garton, Cormac O Maher, Jason A Heth, Erin L McKean, Stephen E Sullivan, Shawn L Hervey-Jumper, Parag G Patil, B Gregory Thompson, Oren Sagher, Guy M McKhann, 2nd, Ricardo J Komotar, Michael E Ivan, Matija Snuderl, Marc L Otten, Timothy D Johnson, Michael B Sisti, Jeffrey N Bruce, Karin M Muraszko, Jay Trautman, Christian W Freudiger, Peter Canoll, Honglak Lee, Sandra Camelo-Piragua, and Daniel A Orringer. Near real-time intraoperative brain tumor diagnosis using

stimulated raman histology and deep neural networks. *Nat. Med.*, 2020. 3

[21] Xinhai Hou, Cheng Jiang, Akhil Kondepudi, Yiwei Lyu, Asadur Zaman Chowdury, Honglak Lee, and Todd C Hollon. A self-supervised framework for learning whole slide representations. *arXiv preprint arXiv:2402.06188*, 2024. 3

[22] Cheng Jiang, Asadur Chowdury, Xinhai Hou, Akhil Kondepudi, Christian Freudiger, Kyle Conway, Sandra Camelo-Piragua, Daniel Orringer, Honglak Lee, and Todd Hollon. Opensrh: optimizing brain tumor surgery using intraoperative stimulated raman histology. *Advances in neural information processing systems*, 35:28502–28516, 2022. 2, 5

[23] Cheng Jiang, Xinhai Hou, Akhil Kondepudi, Asadur Chowdury, Christian W Freudiger, Daniel A Orringer, Honglak Lee, and Todd C Hollon. Hierarchical discriminative learning improves visual representations of biomedical microscopy. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 19798–19808, 2023. 3

[24] Bahjat Kawar, Michael Elad, Stefano Ermon, and Jiaming Song. Denoising diffusion restoration models. In *Advances in Neural Information Processing Systems*, 2022. 2

[25] Gwantae Kim, Jaihyun Park, Kanghyu Lee, Junyeop Lee, Jeongki Min, Bokyeung Lee, David K Han, and Hanseok Ko. Unsupervised real-world super resolution with cycle generative adversarial network and domain discriminator. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 456–457, 2020. 2

[26] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013. 2

[27] Akira Kudo, Yoshiro Kitamura, Yuanzhong Li, Satoshi Iizuka, and Edgar Simo-Serra. Virtual thin slice: 3d conditional gan-based super-resolution for ct slice interval. In *Machine Learning for Medical Image Reconstruction: Second International Workshop, MLMIR 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 17, 2019, Proceedings 2*, pages 91–100. Springer, 2019. 2

[28] Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4681–4690, 2017. 2

[29] Suhyeon Lee, Hyungjin Chung, Minyoung Park,

Jonghyuk Park, Wi-Sun Ryu, and Jong Chul Ye. Improving 3d imaging with pre-trained perpendicular 2d diffusion models. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 10710–10720, 2023. 3

[30] Jingyun Liang, Jiezhang Cao, Guolei Sun, Kai Zhang, Luc Van Gool, and Radu Timofte. Swinir: Image restoration using swin transformer. *arXiv preprint arXiv:2108.10257*, 2021. 2

[31] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee. Enhanced deep residual networks for single image super-resolution. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2017. 2

[32] Qing Lyu, Hongming Shan, and Ge Wang. Mri super-resolution with ensemble learning and complementary priors. *IEEE Transactions on Computational Imaging*, 6:615–624, 2020. 2

[33] Yiwei Lyu, Sung Jik Cha, Cheng Jiang, Asadur Chowdury, Xinhai Hou, Edward Harake, Akhil Kondepudi, Christian Freudiger, Honglak Lee, and Todd C Hollon. Step-calibrated diffusion for biomedical optical image restoration. *arXiv preprint arXiv:2403.13680*, 2024. 3, 5

[34] Dwarikanath Mahapatra and Zongyuan Ge. Mr image super resolution by combining feature disentanglement cnns and vision transformers. In *International Conference on Medical Imaging with Deep Learning*, pages 858–878. PMLR, 2022. 2

[35] Dwarikanath Mahapatra, Behzad Bozorgtabar, Sajini Hewavitharanage, and Rahil Garnavi. Image super resolution using generative adversarial networks and local saliency maps for retinal image analysis. In *Medical Image Computing and Computer Assisted Intervention- MICCAI 2017: 20th International Conference, Quebec City, QC, Canada, September 11-13, 2017, Proceedings, Part III 20*, pages 382–390. Springer, 2017. 2

[36] Bryce Manifold, Elena Thomas, Andrew T Francis, Andrew H Hill, and Dan Fu. Denoising of stimulated raman scattering microscopy images via deep learning. *Biomed. Opt. Express*, 10(8):3860–3874, 2019. 3

[37] Evan M Masutani, Naeim Bahrami, and Albert Hsiao. Deep learning single-frame and multiframe super-resolution for cardiac mri. *Radiology*, 295(3):552–561, 2020. 2

[38] Daniel A Orringer, Balaji Pandian, Yashar S Niknafs, Todd C Hollon, Julianne Boyle, Spencer Lewis, Mia Garrard, Shawn L Hervey-Jumper, Hugh J L Garton, Cormac O Maher, Jason A Heth, Oren Sagher, D Andrew Wilkinson, Matija Snuderl, Sriram Venneti, Shakti H Ramkissoon, Kathryn A McFadden, Amanda

Fisher-Hubbard, Andrew P Lieberman, Timothy D Johnson, X Sunney Xie, Jay K Trautman, Christian W Freudiger, and Sandra Camelo-Piragua. Rapid intraoperative histology of unprocessed surgical specimens via fibre-laser-based stimulated raman scattering microscopy. *Nat Biomed Eng*, 1, 2017. 3

[39] Mingjie Pan, Yulu Gan, Fangxu Zhou, Jiaming Liu, Ying Zhang, Aimin Wang, Shanghang Zhang, and Dawei Li. Diffuseir: Diffusion models for isotropic reconstruction of 3d microscopic images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 323–332. Springer, 2023. 3

[40] Cheng Peng, Wei-An Lin, Haofu Liao, Rama Chellappa, and S Kevin Zhou. Saint: spatially aware interpolation network for medical slice synthesis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 7750–7759, 2020. 2

[41] Aditya Ramesh, Prafulla Dhariwal, Alex Nichol, Casey Chu, and Mark Chen. Hierarchical text-conditional image generation with clip latents. *arXiv preprint arXiv:2204.06125*, 1(2):3, 2022. 2

[42] Thomas Roetzer-Pejrimovsky, Anna-Christina Moser, Baran Atli, Clemens Christian Vogel, Petra A Mercea, Romana Prihoda, Ellen Gelpi, Christine Haberler, Romana Höftberger, Johannes A Hainfellner, et al. The digital brain tumour atlas, an open histopathology resource. *Scientific Data*, 9(1):55, 2022. 2

[43] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10684–10695, 2022. 2

[44] Chitwan Saharia, Jonathan Ho, William Chan, Tim Salimans, David J Fleet, and Mohammad Norouzi. Image super-resolution via iterative refinement. *arXiv:2104.07636*, 2021. 2

[45] Chitwan Saharia, William Chan, Saurabh Saxena, Lala Li, Jay Whang, Emily L Denton, Kamyar Ghasemipour, Raphael Gontijo Lopes, Burcu Karagol Ayan, Tim Salimans, et al. Photorealistic text-to-image diffusion models with deep language understanding. *Advances in Neural Information Processing Systems*, 35:36479–36494, 2022. 2

[46] Jörg Sander, Bob D de Vos, and Ivana Išgum. Unsupervised super-resolution: creating high-resolution medical images from low-resolution anisotropic examples. In *Medical Imaging 2021: Image Processing*, pages 82–88. SPIE, 2021. 2

[47] Jörg Sander, Bob D de Vos, and Ivana Išgum. Autoencoding low-resolution mri for semantically smooth interpolation of anisotropic mri. *Medical Image Analysis*, 78:102393, 2022. 2

[48] Wan-Chi Siu and Kwok-Wai Hung. Review of image interpolation and super-resolution. In *Proceedings of The 2012 Asia Pacific Signal and Information Processing Association Annual Summit and Conference*, pages 1–10, 2012. 2

[49] Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. In *International Conference on Learning Representations*, 2020. 2

[50] Rewa R Sood, Wei Shao, Christian Kunder, Nikola C Teslovich, Jeffrey B Wang, Simon JC Soerensen, Nikhil Madhuripan, Anugayathri Jawahar, James D Brooks, Pejman Ghanouni, et al. 3d registration of pre-surgical prostate mri and histopathology images via super-resolution volume reconstruction. *Medical image analysis*, 69:101957, 2021. 2

[51] Thomas Unterthiner, Sjoerd van Steenkiste, Karol Kurach, Raphaël Marinier, Marcin Michalski, and Sylvain Gelly. Fvd: A new metric for video generation. 2019. 5

[52] Jueqi Wang, Jacob Levman, Walter Hugo Lopez Pinaya, Petru-Daniel Tudosiu, M Jorge Cardoso, and Razvan Marinescu. Inversesr: 3d brain mri super-resolution using a latent diffusion model. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 438–447. Springer, 2023. 3

[53] Martin Weigert, Uwe Schmidt, Tobias Boothe, Andreas Müller, Alexandr Dibrov, Akanksha Jain, Benjamin Wilhelm, Deborah Schmidt, Coleman Broaddus, Siân Culley, et al. Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nature methods*, 15(12):1090–1097, 2018. 1

[54] 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. 3

[55] Yan Xia, Nishant Ravikumar, John P Greenwood, Stefan Neubauer, Steffen E Petersen, and Alejandro F Frangi. Super-resolution of cardiac mr cine imaging using conditional gans and unsupervised transfer learning. *Medical Image Analysis*, 71:102037, 2021. 3

[56] Heng Xiao, Xin Wang, Jun Wang, Jing-Ye Cai, Jian-Hua Deng, Jing-Ke Yan, and Yi-Dong Tang. Single image super-resolution with denoising diffusion gans. *Scientific Reports*, 14(1):4272, 2024. 2

[57] Zongsheng Yue, Jianyi Wang, and Chen Change Loy. Resshift: Efficient diffusion model for image super-

resolution by residual shifting. *Advances in Neural Information Processing Systems*, 36, 2024. 2

[58] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pages 2223–2232, 2017. 2

[59] Yongpei Zhu, Zicong Zhou, Guojun Liao, and Kehong Yuan. Csrgan: Medical image super-resolution using a generative adversarial network. In *2020 IEEE 17th International Symposium on Biomedical Imaging Workshops (ISBI Workshops)*, pages 1–4, 2020. 2