

---

# CellINR: Implicitly Overcoming Photo-induced Artifacts in 4D Live Fluorescence Microscopy

---

Anonymous Author(s)

## Abstract

1        4D live fluorescence microscopy is often compromised by prolonged high intensity  
2        illumination which induces photobleaching and phototoxic effects that generate  
3        photo-induced artifacts and severely impair image continuity and detail recovery.  
4        To address this challenge, we propose the CellINR framework, a case-specific  
5        optimization approach based on implicit neural representation. The method em-  
6        ploys blind convolution and structure amplification strategies to map 3D spatial  
7        coordinates into the high frequency domain, enabling precise modeling and high-  
8        accuracy reconstruction of cellular structures while effectively distinguishing true  
9        signals from artifacts. Experimental results demonstrate that CellINR significantly  
10      outperforms existing techniques in artifact removal and restoration of structural  
11      continuity, and for the first time, a paired 4D live cell imaging dataset is provided  
12      for evaluating reconstruction performance, thereby offering a solid foundation for  
13      subsequent quantitative analyses and biological research. The code and dataset will  
14      be public.

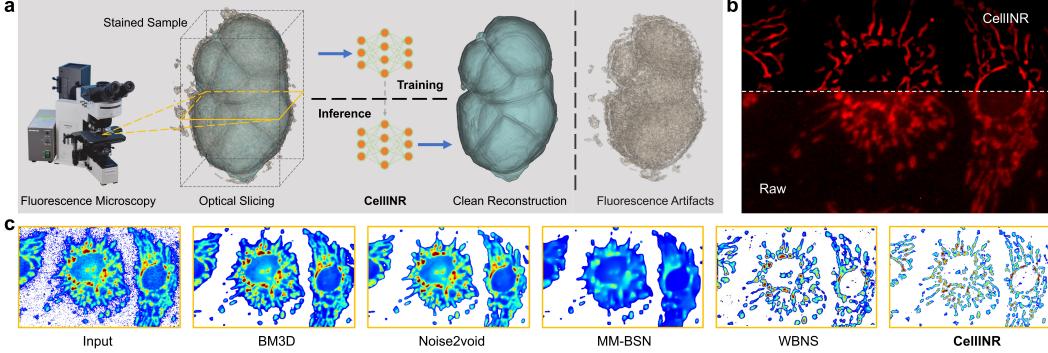
15      

## 1 Introduction

16      4D live fluorescence microscopy imaging exploits specific wavelengths of light to excite fluorescent  
17      markers in biological samples [1], and it acquires 3D image stacks via a computed tomography  
18      (CT)-like scanning approach (see Fig. 1a). Subsequently, by capturing these 3D image stacks at  
19      uniform time intervals, a 4D live imaging sequence is constructed.

20      However, compared with conventional imaging methods, 4D live fluorescence imaging requires  
21      prolonged repetitive exposures, which subject the sample to continuous high-intensity excitation.  
22      This significantly amplifies phototoxicity and photobleaching effects [2, 3, 4]. Continuous and high-  
23      intensity illumination not only causes irreversible photobleaching of fluorescent probes—resulting  
24      in gradual signal decay and a reduction in the image signal-to-noise ratio [5]—but may also induce  
25      phototoxic reactions that damage cellular structures and functions, thereby severely impairing the  
26      physiological state and dynamic behavior of the cells [6, 7]. These additional noise and artifact  
27      issues produced by excessive illumination become particularly pronounced in long-term imaging  
28      and high-frequency dynamic monitoring, constituting a major technical barrier to high-precision 4D  
29      imaging applications. To address these challenges, various experimental strategies have been devised,  
30      such as employing specialized culture media [8, 9], utilizing pulsed excitation [10], and adopting  
31      more advanced controlled-illumination microscopy techniques [3, 11]. Moreover, the 4D imaging  
32      process often fails to obtain pixel-level matched data, thereby hindering the optimization of imaging  
33      results via supervised learning.

34      Furthermore, artifacts exhibit distinct spatiotemporal systematic characteristics, including non-  
35      uniform local signal attenuation, time-dependent drift, and prominent structured noise patterns  
36      in the images [17, 18]. Currently, most mainstream denoising methods primarily target the in-  
37      herent random noise of the imaging system (e.g., low photon-count noise and readout noise  
38      [12, 19, 20, 21, 22, 23, 14, 24, 25]), but they often overlook the pseudo-signals produced by process-  
39      dependent photobleaching and phototoxicity under continuous illumination (see Fig. 1c) [26, 27].



40 Therefore, there is an urgent need to develop novel denoising strategies that both account for the  
 41 intrinsic noise characteristics of the imaging system and adequately accommodate the dynamic artifact  
 42 patterns induced by illumination conditions, thereby enabling precise restoration and quantitative  
 43 analysis of 4D live imaging data.

44 Fluorescence imaging typically employs a tomographic method to acquire 3D image stacks. Despite  
 45 the presence of artifacts, the spatial distribution of signals in 3D space exhibits consistency—including  
 46 both adjacent pixels within different scanning layers and the continuity of signals within the same  
 47 scanning layer. This adjacent information can be used to eliminate the discontinuities caused by  
 48 noise [28, 14]. Therefore, our research focus has shifted from directly removing conventional noise  
 49 to utilizing the overall image structure to clearly reconstruct the target signal from noisy images.

50 Based on these considerations, we employ Implicit Neural Representation (INR) [29] to model  
 51 the entire cell structure, as INR demonstrates exceptional efficacy in capturing position-dependent  
 52 correlations. We optimize the INR on a case-by-case basis using blind convolution [14], which enables  
 53 a faithful representation of the unique three-dimensional features of each cell structure. Furthermore,  
 54 its powerful capability to recover information from incomplete data effectively addresses issues of  
 55 structural discontinuity [30].

56 To further mitigate the impact of artifact on image reconstruction, we introduce a structure am-  
 57plification scheme. This approach not only enhances the core cellular structures but also reduces  
 58 the influence of artifact during the modeling process. These design choices enable INR to more  
 59 effectively decouple the signal from the artifact while maintaining continuity and detail accuracy  
 60 in the reconstructed image (see Figure 1b). Additionally, we have constructed a paired 4D noise  
 61 dataset for evaluating denoising performance, and we have validated the robustness of our denoising  
 62 approach on several unpaired public datasets. Our major contribution can be concluded as follows:

- 63 1. **CellINR**, an INR-based case-specific optimization framework, effectively recovers the core  
   64 cell structures from 4D fluorescence images through self-supervised methods.
- 65 2. We optimize the INR reconstruction process using blind convolution and structure amplifica-  
   66 tion, constructing clear 3D representations of noisy fluorescence images through a complete  
   67 forward propagation process.
- 68 3. We have constructed a paired dataset of 3D images acquired under both low-exposure  
   69 and high-exposure conditions, as well as a synthetic dataset for 4D imaging. In addition,  
   70 we have also collected an unpaired dataset containing background artifacts. Extensive  
   71 experimental results demonstrate that CellINR significantly outperforms state-of-the-art  
   72 methods in eliminating photo-induced artifacts.

73 We optimize each cell’s INR to ensure the representation closely aligns with actual measurements.  
74 Once optimization is complete, we can accurately recover and visualize cell structures at their  
75 corresponding locations for biological analysis.

## 76 2 Related work

77 **Implicit Neural Representations (INR).** Implicit neural representations typically use a multi-layer  
78 perceptron (MLP) to map local coordinates to their corresponding values [29, 31], such as pixel  
79 intensities for images or occupancy values for 3D volumes. An overview of INR fundamentals is  
80 provided below. An INR is a neural network  $F_\theta : \mathbf{x} \mapsto F_\theta(\mathbf{x})$  that continuously maps coordinates  $\mathbf{x}$   
81 to the desired quantity. Given data expressed as a function  $\Phi : \mathbf{x} \mapsto \Phi(\mathbf{x})$ , the INR  $F_\theta$  aims to solve  
82  $F_\theta(\mathbf{x}) - \Phi(\mathbf{x}) = 0$ . The weights  $\theta$  of the INR are typically found through optimization. INRs have  
83 proven effective in parameterizing geometry and learning shape priors, as seen in DeepSDF [32],  
84 Occupancy Networks [33], and IM-Net [34]. Many subsequent works have proposed volumetric ren-  
85 dering techniques for 3D implicit representations, including Neural Radiance Fields (NeRF) [35] and  
86 its variants for faster rendering [36, 37], sparse reconstruction [38, 39], and improved quality [40, 41].  
87 Recent advancements include CuNeRF [42], which achieves zero-shot super-resolution at arbitrary  
88 scales. McGinnis et al.[43] improved multi-contrast MRI imaging using implicit representations.  
89 While Kniesel et al. [44] reconstructed 3D structures from STEM images by modeling mechanical  
90 noise, their method struggles with local accidental excitation fluorescence noise. Though NeRF and  
91 its variants excel in 3D reconstruction, further research is needed to address complex artifacts in  
92 fluorescence microscopy super-resolution.

93 **Fluorescence Image Denoising.** Based on the statistical characteristics of optical imaging noise,  
94 researchers modeled noisy images as the superposition of deterministic signals and independent  
95 Poisson-Gaussian distributions [12, 45]. Traditional denoising methods such as Non-Local Means  
96 (NLM) [46] and wavelet transforms [47] have been widely applied in biological imaging. Dabov  
97 et al. [13] proposed a sparse three-dimensional transform-domain collaborative filtering algorithm,  
98 achieving remarkable results in Gaussian noise suppression. With the rapid development of deep  
99 learning technology, supervised learning-based denoising methods have demonstrated superior  
100 performance in learning complex noise patterns and achieving higher quality image reconstruction [20,  
101 48, 49]. However, their dependence on high-quality annotated data limits their application in  
102 biological imaging. Recently, unsupervised learning algorithms have made breakthrough progress in  
103 image denoising: Noise2void [14] achieve self-supervised denoising using blind convolution; methods  
104 such as Noisier2Noise and Recorrupted-to-recorrupted [50, 51] enable training with unpaired noisy  
105 data by introducing enhanced noise. These algorithms excel in diverse noise environments. However,  
106 for background artifacts removal, traditional techniques[52, 53, 16] remain the most effective.

## 107 Preliminaries

108 Fluorescence microscopy images can be modeled as the sum of the true signal and various noise  
109 components [1]:

$$I(x, y, z) = P \otimes S(x, y, z) + N_{\text{conv}}(x, y, z) + N_{\text{artifacts}}(x, y, z), \quad (1)$$

110 where  $P$  denotes the point spread function,  $I(x, y, z)$  is the observed image, and  $S(x, y, z)$  is the  
111 ideal fluorescence signal.  $N_{\text{conv}}(x, y, z)$  represents conventional noise (e.g., from photon scarcity  
112 and vibrations), typically following a Poisson–Gaussian distribution that is either uniform or signal-  
113 dependent.

114 In contrast,  $N_{\text{artifacts}}(x, y, z)$  represents artifacts induced by photobleaching and phototoxicity,  
115 arising from signal degradation and complex biological interference [13, 46, 28, 54]. Fluorescence  
116 signals are usually localized in specific structures (e.g., membranes and nuclei) and exhibit high-  
117 frequency features, whereas artifacts tend to be uniformly low-frequency[17, 18]. This distinction  
118 forms the basis for designing targeted artifact removal methods to separate structured signals from  
119 noise.

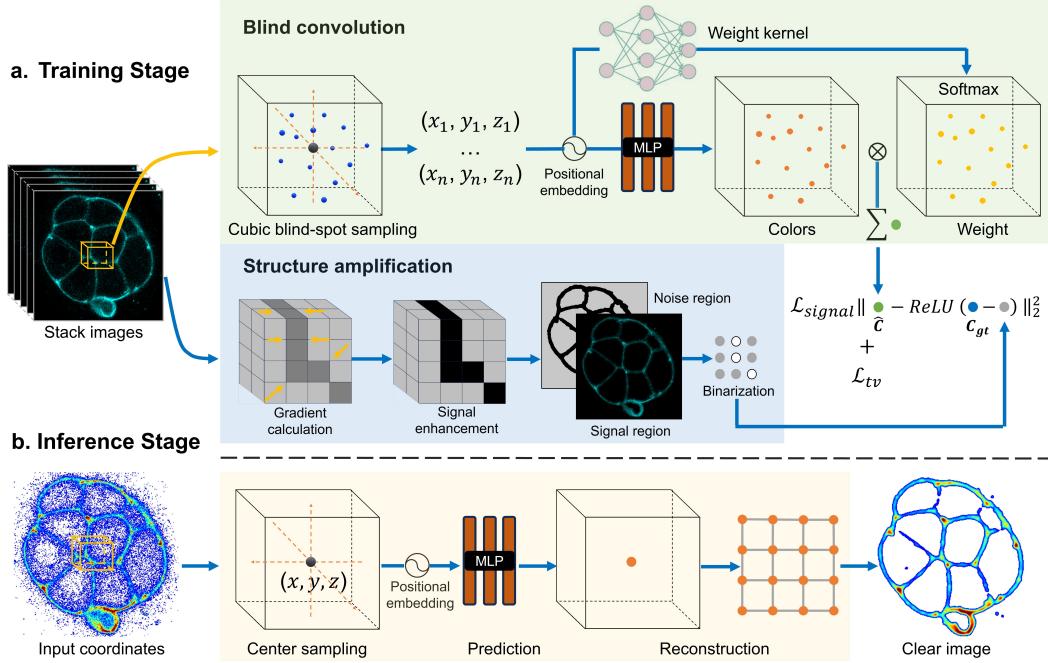


Figure 2: Overview of the CellINR framework. (a) Training Stage: Cubic blind-spot sampling is performed on noisy images, with MLP predicting sample point colors, and blind deconvolution with structure amplification reconstructing center pixel values. (b) Inference Stage: Spatial coordinates of target pixels are sampled to output clear reconstruction results.

### 120 3 Proposed methodology

121 As shown in Fig.2a, CellINR adopts a case-by-case optimization strategy to construct clear 3D  
 122 representations from fluorescence images during INR’s forward propagation. During training, cubic  
 123 sampling randomly selects multiple sampling points around each target pixel and performs con-  
 124 volution summation to eliminate noise interference to ensure signal smoothness and robust noise  
 125 suppression. Meanwhile, structural amplification is applied to fluorescence images to eliminate  
 126 jagged effects at signal boundaries and locally uniform low-brightness distribution of artifacts.  
 127 During inference stage, INR is directly sampled to generate clear continuous 3D representation.

#### 128 3.1 CellINR

129 The backbone of our CellINR uses a Multi-Layer Perceptron (MLP) [55] with parameters  $\varrho$  to map  
 130 spatial coordinates  $\rho$  of fluorescence images to their corresponding implicit expressions. We perform  
 131 uniform sampling in the fluorescence image space to obtain  $N$  discrete coordinate points  $\rho$ , denoted  
 132 as  $N_\rho$ . Subsequently, we map the coarse sampling point set  $N_\rho$  to high-frequency domain space to  
 133 enhance the model’s ability to capture details [35], represented here as:

$$\gamma(\rho) = [\sin(\rho), \cos(\rho), \dots, \sin(2^{\epsilon-1}\rho), \cos(2^{\epsilon-1}\rho)]^T, \quad (2)$$

134 where  $\epsilon$  represents the dimension of positional encoding. Through simple mapping, we obtain:

$$c = F_\varrho(\gamma(\rho)), \quad (3)$$

135 where  $c$  is the color value of the pixel at spatial position  $\rho$  in the implicit representation. However,  
 136 INR constructed through this method alone tends to behave as identity mappings, failing to effectively  
 137 model clear representations of fluorescence images.

#### 138 3.2 Blind convolution

139 Based on the discrete nature of photo-induced artifacts and continuous distribution characteristics  
 140 of signal features, we recognize that noise disrupts continuity between adjacent pixels. Meanwhile,

141 fluorescence signal distribution maintains structural continuity within regions. Inspired by [45, 14],  
 142 we model 3D continuous representations of fluorescence images using INR based on 3D blind-spot  
 143 convolution. This method estimates target point values through the convolution of points around the  
 144 target pixel while ensuring noise information from target pixels is not input to the neural network.  
 145 Traditional 3D blind-spot convolution can be represented as:

$$I_{(x,y,z)} = \sum_{i=-h}^h \sum_{j=-h}^h \sum_{k=-h}^h C_{x-i,y-j,z-k} \cdot W_{x-i,y-j,z-k}, \quad (4)$$

146 where  $C_{x-i,y-j,z-k}$  represents the pixel value of sampling points with coordinate offset  $(i, j, k)$   
 147 within a cubic space centered at point  $(x, y, z)$  with radius  $h$ ,  $W_{x-i,y-j,z-k}$  is the weight convolution  
 148 kernel corresponding to each pixel position, and  $I_{(x,y,z)}$  is the pixel value at position  $(x, y, z)$  obtained  
 149 after convolution operation between  $C$  and  $W$ , used to estimate the pixel value at the target point  
 150 in noise-free conditions. To avoid blind spot invalidation caused by target pixels being input to the  
 151 network during convolution [45], and to maintain INR continuity while preventing mesh artifacts  
 152 and modeling holes during sampling [42], we improved INR's sampling method to approximate the  
 153 convolution kernel.

154 Specifically, we perform random coarse sampling on spatial coordinates in  $S_{cube}$  to obtain  $N$  discrete  
 155 coordinate points, denoted as  $N_c$ . Subsequently, we input the positionally encoded coarse sampling  
 156 points into coarse MLP  $F_\Theta$  to predict volume density  $\sigma = F_\Theta(\gamma(N_c))$ . Based on the predicted  
 157 volume density, we construct a cumulative distribution function (CDF) and apply importance sampling  
 158 strategy [42], sampling  $N_f$  fine sampling points  $N_f = \xi(N_c, \sigma)$ . Here,  $\xi$  is the function performing  
 159 fine sampling based on the volume density  $\sigma(N_c)$  of coarse sampling points. Merging coarse  
 160 sampling point set  $N_c$  with fine sampling point set  $N_f$ , we obtain the final sampling point set  
 161  $N_{(x,y,z)} = N_c \cup N_f$ , where  $N$  represents the blind-spot cube sampling point set centered at  $(x, y, z)$ .  
 162 Then, the merged sampling point set is input to fine MLP  $F_\delta$  to predict corresponding color values:

$$c(\rho) = F_\delta [\gamma (\rho \cup \xi (\rho, F_\Theta [\gamma (\rho)]))], \quad \text{where } \rho \in N. \quad (5)$$

163 Considering that convolution kernel weights are typically spatially varying in complex noisy images,  
 164 we choose to use MLP  $F_\phi$  to model spatially varying kernels. To ensure weight normalization, we  
 165 apply the Softmax function to generated weights, making their sum equal to 1. Specifically, for each  
 166 input view, we predict the weights as:

$$w(\rho) = \frac{\exp(F_\phi(\gamma(\rho)))}{\sum_{\rho \in N} \exp(F_\phi(\gamma(\rho)))}. \quad (6)$$

167 Finally, our blind convolution can be expressed as:

$$I_{\text{clean}} = \sum_{\rho \in N} [c(\rho) \cdot w(\rho)], \quad (7)$$

168 This method effectively models continuous representations of fluorescence images in INR.

### 169 3.3 Structural amplification

170 Another characteristic of photo-induced artifact is its uniform low-frequency distribution in local  
 171 regions[17, 18]. In contrast, the features of the fluorescence signal are confined to specific cellular  
 172 structures and generally exhibit a concentrated high-frequency distribution. To further suppress  
 173 photo-induced artifact, a Hessian matrix enhancement technique is applied to the input image  
 174 [56, 57, 58]. to input images. Specifically, we calculate second-order partial derivatives of images and  
 175 perform eigenvalue decomposition to select pixels conforming to planar structures, thereby enhancing  
 176 fluorescent feature signals. For a 3D fluorescence image  $I^f(x, y, z)$ , we compute the Hessian matrix  
 177 as:

$$H = \begin{bmatrix} \frac{\partial^2 I^f}{\partial x^2} & \frac{\partial^2 I^f}{\partial x \partial y} & \frac{\partial^2 I^f}{\partial x \partial z} \\ \frac{\partial^2 I^f}{\partial x \partial y} & \frac{\partial^2 I^f}{\partial y^2} & \frac{\partial^2 I^f}{\partial y \partial z} \\ \frac{\partial^2 I^f}{\partial x \partial z} & \frac{\partial^2 I^f}{\partial y \partial z} & \frac{\partial^2 I^f}{\partial z^2} \end{bmatrix} = E \Lambda E^T, \quad (8)$$

178 where  $\lambda_1, \lambda_2, \lambda_3$  are the eigenvalues satisfying  $|\lambda_1| < |\lambda_2| < |\lambda_3|$ , and  $e_1, e_2, e_3$  are the correspond-  
 179 ing eigenvectors. By selecting the largest eigenvalue  $\lambda_3$ , we enhance the fluorescent signals:

$$I_{\text{en}}(x, y, z) = \frac{|\lambda_3(x, y, z)|}{\max\{|\lambda_3(x, y, z)| \mid x, y, z \in I\}}. \quad (9)$$

180 The enhanced image  $I_{\text{en}}$  is then binarized, assigning signal regions a value of 255 and noise regions a  
 181 value of 0, resulting in a binary image  $I_{\text{binary}}$ . The binarization process is represented as:

$$I_{\text{binary}}(x, y, z) = \begin{cases} 255, & \text{if } I_{\text{en}}(x, y, z) > \mu, \\ 0, & \text{otherwise,} \end{cases} \quad (10)$$

182 where  $\mu$  is the Otsu thresholding[59] of the largest eigenvalues.

### 183 3.4 Loss function

184 To effectively train the CellINR framework, we design a hybrid three-dimensional loss function  
 185 that optimizes the reconstruction effect in signal regions while maintaining the three-dimensional  
 186 structural consistency of the image. The loss function primarily consists of two components.

187 **Clean reconstruction loss:** For pixel points labeled as signal regions based on the binarized image  
 188  $I_{\text{binary}}$ , we calculate the MSE between the reconstructed image  $I_{\text{clean}}$  and the rectified original noisy  
 189 image  $I_{\text{raw}}$  activated by ReLU:

$$\mathcal{L}_{\text{signal}} = \frac{1}{N} \sum_I (I_{\text{clean}} - \text{ReLU}(I_{\text{raw}} - I_{\text{binary}}))^2, \quad (11)$$

190 where  $N$  represents the total number of pixel points in the signal regions, and  $\text{ReLU}(x) = \max(0, x)$   
 191 ensures that only the signal regions contribute to the error calculation.

192 **Structural consistency loss:** To enhance the structural consistency of the reconstructed image,  
 193 we introduce a Structural Consistency Loss based on the Total Variation (TV) method. Here,  $X$   
 194 represents the pixel matrix of the reconstructed image, while  $i$  and  $j$  denote the horizontal and vertical  
 195 pixel indices, respectively. The TV loss is defined as

$$\mathcal{L}_{\text{TV}}(X) = \sum_{i,j} (|X_{i+1,j} - X_{i,j}| + |X_{i,j+1} - X_{i,j}|). \quad (12)$$

196 This loss function encourages smooth transitions between adjacent pixels, thereby suppressing  
 197 artifacts and noise, and ultimately enhancing the continuity of the image structure.

198 Combining these two components, the final total loss is defined as:

$$\mathcal{L} = \mathcal{L}_{\text{signal}} + \lambda \mathcal{L}_{\text{TV}}, \quad (13)$$

199 where  $\lambda$  is a balancing coefficient, set to 0.15 in our experiments. This loss function enables CellINR  
 200 to maintain detailed signal regions while ensuring three-dimensional structural consistency, achieving  
 201 high-quality image denoising and clean reconstruction.

### 202 3.5 Implementation details

203 We implemented the CellINR framework based on NeRF’s PyTorch implementation [60],  
 204 CuNeRF [42] and Deblur-NeRF [61]. The specific settings are as follows: within a cubic space of  
 205 radius 1 around the blind spot,  $N = 27$  coordinate points are uniformly sampled during experiments,  
 206 inputting coarse predictions of volume density via  $F_\theta$ . In the fine volume, an additional  $N$  coordinate  
 207 points are sampled, totaling  $2N$  points input into the fine  $F_\delta$  and kernel modules to output color  
 208 values and corresponding weights. To ensure that the kernel module outputs only weights, a Softmax  
 209 function is applied to ensure the sum of the  $2N$  weights equals 1. The optimizer used is Adam with a  
 210 weight decay of  $10^{-6}$ , a batch size of 4096, a maximum of 500,000 iterations, and a learning rate  
 211 linearly annealing from  $2 \times 10^{-3}$  to  $2 \times 10^{-5}$ . Training one (256,356,160) volume image on a single  
 212 NVIDIA RTX 4090 GPU required 50,000 iterations, 20 minutes. The MLP structure is identical  
 213 to NeRF, comprising 8 fully connected hidden layers with 256 channels each and ReLU activation  
 214 functions. The kernel module follows a similar structure, with center point coordinates input to the  
 215 seventh layer to assist in weight generation. The specific architecture is detailed in *Supplementary*  
 216 *materials*. During inference stage (Fig.2b), the fine  $F_\delta$  predicts the corresponding pixel values using  
 217 center point coordinates, and kernel weights are set to 1. By varying sampling densities, output  
 218 images at different resolution can be obtained.

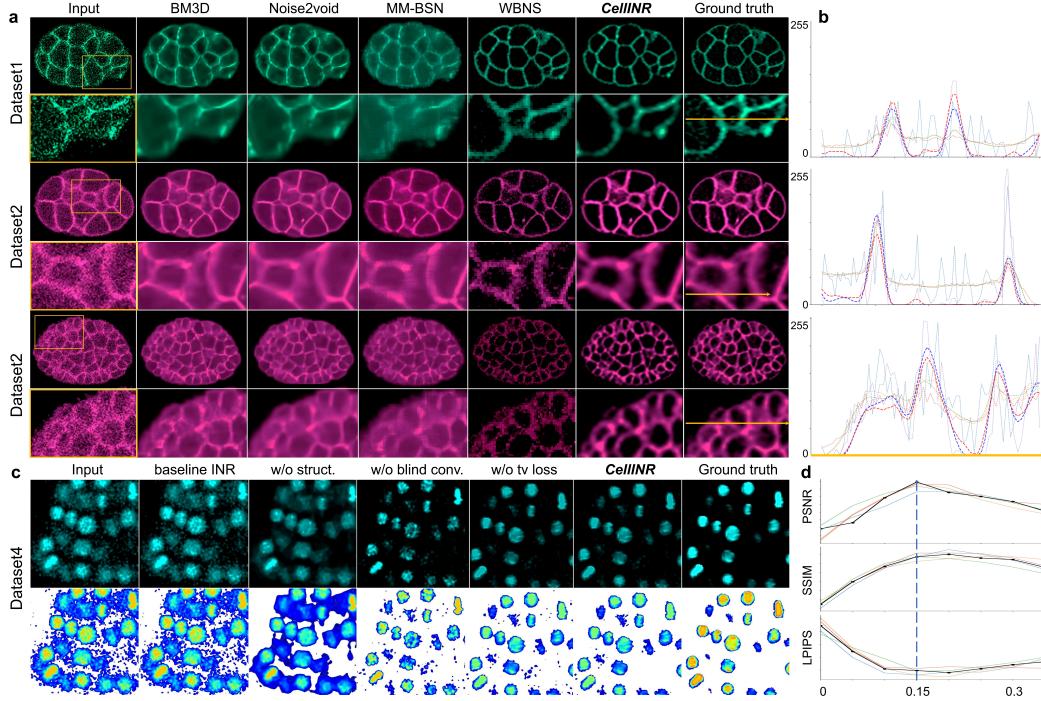


Figure 3: Visual comparison with existing methods. (a) Visualization of the benchmark on Dataset 1 and Dataset 2. (b) Distribution of pixel values along selected line segments for different methods. The curves are color-coded as follows: **Input**, **BM3D**, **Noise2Void**, **MM-BSN**, **WBNS**, **CellINR**, and **Ground truth**. (c) Visualization of ablation experiments on Dataset 3. (d) Sensitivity test of the hyperparameter  $\lambda$  in structural consistency loss.

## 219 4 Experiments

### 220 4.1 Datasets

| Dataset  | Model system            | Excitation | Type                     | Exposure Time | Number | Image Size    | Source              | GT                       |
|----------|-------------------------|------------|--------------------------|---------------|--------|---------------|---------------------|--------------------------|
| dataset1 | C.elegans cell membrane | 561nm      | Static 3D Imaging        | 176s          | 40     |               | paired real dataset |                          |
| dataset2 | C.elegans cell membrane | 561nm      | Time-Resolved 3D Imaging | 176s          | 40     |               |                     |                          |
| dataset3 | C.elegans cell membrane | 561nm      | Time-Resolved 3D Imaging | 752s          | 80     | (256,356,214) | Ours                | paired synthetic dataset |
| dataset4 | C.elegans cell nucleus  | 488nm      | Time-Resolved 3D Imaging | 176s          | 40     |               |                     |                          |
| dataset5 | C.elegans cell membrane | 561nm      | Time-Resolved 3D Imaging | /             | 40     | (256,356,160) | [62]                |                          |
| dataset6 | BAPE                    | /          | Static 2D Imaging        | /             | 20     | (512,512)     | [12]                |                          |
| dataset7 | Zebrafish               | /          | Static 2D Imaging        | /             | 20     | (512,512)     | [12]                | unpaired dataset         |
| dataset8 | Mouse                   | /          | Static 2D Imaging        | /             | 2      | (299,298)     | [63]                |                          |

Table 1: The composition of the qualitative/quantitative datasets used in this study to evaluate the denoising results.

221 In this study, datasets 1, 2, 3, and 4 were acquired from live *C. elegans* embryo imaging using a Leica  
222 SP5II confocal microscope equipped with a  $63\times/1.4$  NA objective lens. The embryos were placed on  
223 a 3–5% (w/v) agarose pad, treated with 0.5% tetramisole, and sealed with vaseline. Datasets 1, 2,  
224 and 3 were labeled with mCherry to mark the cell membrane and excited by a 561 nm laser, whereas  
225 dataset 4 employed green fluorescent protein (GFP) to label the cell nucleus and was excited by a 488  
226 nm laser. All four datasets have a resolution of  $712\times512$  pixels in the  $xy$  plane, with 44, 44, 94, and  
227 44 focal planes acquired along the  $z$ -axis, respectively (the  $xy$  resolution is  $0.09\ \mu\text{m}/\text{pixel}$ , and the  $z$   
228 resolutions are  $0.42$ ,  $0.42$ ,  $0.24$ , and  $0.42\ \mu\text{m}/\text{pixel}$ , respectively). Dataset 1 is a static paired dataset  
229 obtained under low- and high-illumination conditions. By controlling the exposure time of individual  
230 2D slices (10 ms for the low-exposure, low-SNR group and 100 ms for the high-exposure, high-SNR  
231 group), 20 paired groups of low-noise and high-noise images were obtained, yielding a total of  
232 40 image sets. Datasets 2 to 4 are time-resolved 3D imaging datasets acquired with an exposure  
233 time of 100 ms per 2D slice and sampled repeatedly from a single embryo at 90-second intervals.

| Benchmark                 |                              | Dataset1    |             |             | Dataset2    |             |             | Dataset4    |             |             |
|---------------------------|------------------------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
|                           |                              | PSNR↑       | SSIM↑       | LPIPS↓      | PSNR↑       | SSIM↑       | LPIPS↓      | PSNR↑       | SSIM↑       | LPIPS↓      |
| Comparison                | BM3D (Traditional)           | 22.737±3.27 | 0.678±0.086 | 0.170±0.022 | 20.630±3.06 | 0.718±0.027 | 0.111±0.016 | 19.936±2.00 | 0.757±0.067 | 0.158±0.021 |
|                           | Noise2void (Self-supervised) | 22.815±2.28 | 0.618±0.066 | 0.137±0.019 | 21.466±3.15 | 0.759±0.095 | 0.118±0.017 | 18.554±1.86 | 0.735±0.074 | 0.147±0.019 |
|                           | MM-BSN (Self-supervised)     | 21.408±2.14 | 0.599±0.045 | 0.233±0.028 | 20.416±2.04 | 0.659±0.067 | 0.197±0.025 | 18.809±1.88 | 0.665±0.066 | 0.283±0.022 |
|                           | WBNS (Traditional)           | 33.180±1.66 | 0.842±0.042 | 0.141±0.007 | 34.898±1.75 | 0.904±0.045 | 0.037±0.002 | 32.698±1.64 | 0.902±0.045 | 0.033±0.002 |
| Ablation                  | baseline INR                 | 23.804±3.38 | 0.526±0.079 | 0.259±0.021 | 23.302±3.32 | 0.681±0.096 | 0.209±0.010 | 20.869±3.97 | 0.578±0.047 | 0.242±0.011 |
|                           | w/o blindconv.               | 30.040±3.40 | 0.826±0.053 | 0.159±0.013 | 32.210±2.22 | 0.831±0.058 | 0.109±0.015 | 30.640±3.54 | 0.878±0.078 | 0.100±0.012 |
|                           | w/o struct.                  | 27.730±2.37 | 0.751±0.057 | 0.137±0.010 | 22.230±3.22 | 0.619±0.062 | 0.172±0.012 | 21.550±2.76 | 0.723±0.072 | 0.175±0.012 |
|                           | w/o TV loss                  | 33.460±3.35 | 0.868±0.087 | 0.129±0.014 | 37.640±3.76 | 0.913±0.091 | 0.053±0.007 | 33.020±3.30 | 0.900±0.090 | 0.050±0.004 |
| CellINR (Self-supervised) |                              | 35.470±1.77 | 0.881±0.044 | 0.079±0.004 | 38.620±1.93 | 0.925±0.046 | 0.029±0.001 | 33.170±1.66 | 0.919±0.046 | 0.025±0.001 |

Table 2: Quantitative comparison of PSNR, SSIM, and LPIPS across three datasets. CellINR scores are highlighted in orange and WBNS scores in cyan (indicating the best and second best performance, respectively). Mean±std is reported for each metric.

234 Specifically, datasets 2 and 4 each include 40 time points (with cumulative exposure times of 176 s,  
235 respectively), while dataset 3 comprises 80 time points (with a cumulative exposure time of 752 s).  
236 After processing with Leica LAS X software and ImageJ, the images were resampled using trilinear  
237 interpolation to achieve a uniform resolution of 0.18  $\mu\text{m}/\text{pixel}$  in all three dimensions, ultimately  
238 yielding 200 sets of 3D NIFTI data with dimensions (256, 356, 214). Dataset 1 is primarily used to  
239 evaluate denoising performance under low-light conditions and is constructed based on the prominent  
240 photo-induced artifacts observed in the long-exposure 4D dataset 3. To simulate the photo-induced  
241 artifacts encountered in live fluorescence microscopy, Poisson noise (factor = 0.1) and Gaussian noise  
242 (with a standard deviation of 20 added to non-zero pixel regions) were incorporated into datasets 2  
243 and 4, respectively. For quantitative evaluation, the paired datasets consist of datasets 1, 2, and 4  
244 (comprising 20, 40, and 40 pairs, respectively), while datasets 3, 5 [62], 6 [12], 7 [12], and 8 [63]  
245 are used solely for qualitative evaluation; corresponding experimental results are provided in the  
246 supplementary material.

## 247 4.2 Quantitative Evaluation

248 To assess denoising performance, we compared CellINR with mainstream denoising models[13, 14,  
249 15, 16]. These methods were applied to datasets 1, 2, and 4 (visual results are shown in Fig. 3a and  
250 quantitative results in Table 2). All methods were implemented using the open-source codes provided  
251 by the respective authors. A comprehensive visual comparison reveals that CellINR offers significant  
252 advantages in removing photo-induced artifacts. In terms of suppressing jagged edges at microscopic  
253 structures, eliminating pseudo-signals, and overall reconstruction quality, CellINR provides the best  
254 interpretation of cellular structural signals. Specifically, CellINR successfully eliminates artifacts  
255 that are difficult for mainstream denoising methods. Although WBNS can remove some artifacts, its  
256 wavelet decomposition process inevitably disrupts signal continuity; in contrast, CellINR accurately  
257 reconstructs the continuous signal structure within the noise, preserving the integrity of fine cellular  
258 features without introducing aliasing. Figure 3b shows the pixel intensity profiles along a yellow  
259 line for three samples denoised by different methods, where the CellINR result closely matches the  
260 ground truth with minimal artifacts. For denoising results on datasets 3, 5, 6, 7, and 8, please refer to  
261 the supplementary material. For the three paired datasets, we conducted a comprehensive evaluation  
262 using PSNR, SSIM, and LPIPS metrics, repeating experiments on all data within each dataset and  
263 reporting the mean and standard deviation. In the real 3D dataset (dataset 1), CellINR not only  
264 showed approximately a 30% improvement in mean values over other mainstream denoising methods  
265 (approximately 10% improvement over WBNS) but also exhibited higher robustness in fine detail  
266 restoration, with minimal standard deviation in all metrics. In the synthetic 4D datasets (datasets  
267 2 and 4), CellINR also demonstrated stable and excellent quantitative performance, confirming its  
268 suitability for removing photo-induced artifacts in live fluorescence microscopy under unsupervised  
269 conditions.

## 270 4.3 Ablation Experiments

271 To further quantitatively assess the effectiveness of each component in CellINR framework, we  
272 conducted ablation experiments using the same evaluation metrics and datasets as in the quantitative  
273 evaluation. We compared variants in which only the baseline INR is employed, the blind convolution  
274 module is removed, the structural amplification module is removed, or the TV loss is omitted  
275 against the complete CellINR model. Using only the baseline INR resulted in the model behaving

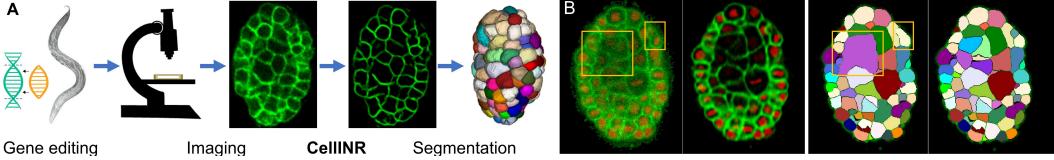


Figure 4: (a) Cell morphology modeling using 4D live fluorescence microscopy during *Caenorhabditis elegans* embryonic development. (b) Segmentation results before and after denoising. The yellow box shows that artifact removal reduces misidentification in downstream analysis.

essentially as an identity mapping, consistent with the discussion in Section 4.1, and it performed significantly worse across all metrics. Removal of the blind convolution module led to a marked decline in the ability to maintain signal continuity and anti-aliasing at edges, underscoring its critical role in constructing a continuous representation. Eliminating the structural amplification module significantly reduced the model’s capacity to remove artifacts—especially low-frequency background pseudo-signals—with experimental results showing an approximate PSNR drop of 10 dB and a SSIM decrease of about 0.2, thereby demonstrating the necessity of incorporating artifact-based structural amplification during training. Omitting the TV loss slightly decreased both PSNR and SSIM, indicating a weakened suppression of high-frequency noise and isolated noise points, which confirms the importance of smoothing for signal optimization. Detailed quantitative results for the ablation experiments are provided in Table 1.

#### 4.4 Downstream Tasks

The results of 4D live fluorescence microscopy are often used for subsequent downstream analyses to obtain time-series 3D shapes, which serve as the basis for physical modeling and studies in quantitative biology and developmental biology [62, 64, 65, 66]. In this study, we employed gene editing to label the cell membrane and nucleus with fluorescent proteins and performed 3D segmentation for quantitative analysis of cellular morphology (see Fig. 4a). The 3D segmentation pipeline follows a standard procedure: first, Mask-RCNN [67] was used for automatic segmentation of the nuclei and to extract their centroids as seeds; subsequently, an avascular enhancement filter [68] (with scale parameters set sequentially to 2.0, 2.25, and 2.5) was applied to amplify the boundary signals in the cell membrane region, followed by a seeded watershed algorithm to accurately extract cell contours, thus obtaining the complete cell shape; finally, manual corrections were performed using Labkit to refine the edge details and ensure the segmentation results closely match the actual cellular morphology. As shown in Fig. 4b, artifact removal significantly reduced blurred regions, and the cell membrane structures became clearer with enhanced boundary signals. This effectively reduced errors such as signal misrecognition and segmentation inaccuracies caused by artifacts in downstream tasks, thereby greatly improving the accuracy of subsequent analyses.

#### Conclusion

This study presents the **CellINR** framework based on implicit neural representations, which systematically integrates blind convolution and structural amplification strategies to effectively suppress photo-induced artifacts and accurately restore the continuity of cellular structures in 4D live fluorescence microscopy. The proposed method has achieved notable success in noise optimization and the recovery of critical structures, thereby providing high-quality data support for quantitative analyses of embryonic development, cell differentiation, and other dynamic biological processes. Rigorous quantitative experiments confirm that CellINR meets the stringent requirements for precise capture of fine structures in biological imaging, thereby establishing a robust theoretical and practical foundation for multi-scale and multi-modal biological information analysis. However, CellINR may misclassify data where artifacts and genuine signals are ambiguous; thus, it is best applied to datasets with clearly distinguishable artifacts. Additionally, for multi-channel composite images, processing without separate single-channel data may conflate fluorescence signals and compromise reconstruction quality. Future work will further incorporate biophysical models to extend its applicability in complex biological systems.

318 **References**

- 319 [1] J. Mertz, *Introduction to Optical Microscopy*, pp. 142–160. Cambridge University Press, 2019. [1](#), [3](#)
- 320 [2] J. Y. Tinevez, J. Dragavon, L. Baba-Aissa, P. Roux, E. Perret, A. Canivet, and S. Shorte, “A quantitative  
321 method for measuring phototoxicity of a live cell imaging microscope,” in *Methods in Enzymology*, vol. 506,  
322 pp. 291–309, Academic Press, 2012. [1](#)
- 323 [3] R. A. Hoebe, C. H. Van Oven, T. W. J. Gadella, *et al.*, “Controlled light-exposure microscopy reduces  
324 photobleaching and phototoxicity in fluorescence live-cell imaging,” *Nature Biotechnology*, vol. 25,  
325 pp. 249–253, 2007. [1](#)
- 326 [4] R. A. Hoebe, H. T. M. Van Der Voort, J. Stap, C. J. F. Van Noorden, and E. M. M. Manders, “Quantitative  
327 determination of the reduction of phototoxicity and photobleaching by controlled light exposure  
328 microscopy,” *Journal of Microscopy*, vol. 231, no. 1, pp. 9–20, 2008. [1](#)
- 329 [5] C. Eggeling, J. Widengren, R. Rigler, and C. A. Seidel, “Photobleaching of fluorescent dyes under  
330 conditions used for single-molecule detection: evidence of two-step photolysis,” *Analytical Chemistry*,  
331 vol. 70, no. 13, pp. 2651–2659, 1998. [1](#)
- 332 [6] A. H. Wright, J. M. DeLong, J. L. Franklin, R. R. Lada, and R. K. Prange, “A new minimum fluorescence  
333 parameter, as generated using pulse frequency modulation, compared with pulse amplitude modulation:  
334  $f_{alpha}$  versus  $f_0$ ,” *Photosynthesis Research*, vol. 97, pp. 205–214, 2008. [1](#)
- 335 [7] R. Dixit and R. Cyr, “Cell damage and reactive oxygen species production induced by fluorescence  
336 microscopy: effect on mitosis and guidelines for non-invasive fluorescence microscopy,” *The Plant  
337 Journal*, vol. 36, no. 2, pp. 280–290, 2003. [1](#)
- 338 [8] A. M. Bogdanov, E. A. Bogdanova, D. M. Chudakov, *et al.*, “Cell culture medium affects gfp photostability:  
339 A solution,” *Nature Methods*, vol. 6, pp. 859–860, 2009. [1](#)
- 340 [9] A. M. Bogdanov, E. I. Kudryavtseva, and K. A. Lukyanov, “Anti-fading media for live cell gfp imaging.”  
341 *PLoS One*, vol. 7, p. e53004, 2012. [1](#)
- 342 [10] T. Nishigaki, C. D. Wood, K. Shiba, *et al.*, “Stroboscopic illumination using light-emitting diodes reduces  
343 phototoxicity in fluorescence cell imaging,” *Biotechniques*, vol. 41, pp. 191–197, 2006. [1](#)
- 344 [11] W. Caarls, B. Rieger, A. H. B. De Vries, *et al.*, “Minimizing light exposure with the programmable array  
345 microscope,” *Journal of Microscopy*, pp. 101–110, 2010. [1](#)
- 346 [12] Y. Zhang, Y. Zhu, E. Nichols, Q. Wang, S. Zhang, C. Smith, and S. Howard, “A poisson-gaussian denoising  
347 dataset with real fluorescence microscopy images,” in *Proceedings of the IEEE/CVF Conference on  
348 Computer Vision and Pattern Recognition*, pp. 11710–11718, 2019. [1](#), [2](#), [3](#), [7](#), [8](#)
- 349 [13] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, “Image denoising by sparse 3-d transform-domain  
350 collaborative filtering,” *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2080–2095, 2007. [2](#), [3](#),  
351 [8](#)
- 352 [14] A. Krull, T. O. Buchholz, and F. Jug, “Noise2void—learning denoising from single noisy images,” in  
353 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 2129–2137,  
354 2019. [1](#), [2](#), [3](#), [5](#), [8](#)
- 355 [15] D. Zhang, F. Zhou, Y. Jiang, and Z. Fu, “Mm-bsn: Self-supervised image denoising for real-world with  
356 multi-mask based on blind-spot network,” in *Proceedings of the IEEE/CVF Conference on Computer  
357 Vision and Pattern Recognition*, pp. 4189–4198, 2023. [2](#), [8](#)
- 358 [16] M. Hüpfel, A. Yu. Kobitski, W. Zhang, and G. U. Nienhaus, “Wavelet-based background and noise  
359 subtraction for fluorescence microscopy images,” *Biomedical Optics Express*, vol. 12, no. 2, pp. 969–980,  
360 2021. [2](#), [3](#), [8](#)
- 361 [17] J. Icha, M. Weber, J. C. Waters, and C. Norden, “Phototoxicity in live fluorescence microscopy, and how to  
362 avoid it,” *BioEssays*, vol. 39, no. 8, p. 1700003, 2017. [1](#), [3](#), [5](#)
- 363 [18] P. P. Laissue, R. A. Alghamdi, P. Tomancak, E. G. Reynaud, and H. Shroff, “Assessing phototoxicity in  
364 live fluorescence imaging,” *Nature Methods*, vol. 14, no. 7, pp. 657–661, 2017. [1](#), [3](#), [5](#)
- 365 [19] E. H. K. Stelzer, “Contrast, resolution, pixelation, dynamic range and signal-to-noise ratio: fundamental  
366 limits to resolution in fluorescence light microscopy,” *Journal of Microscopy*, vol. 189, no. 1, pp. 15–24,  
367 1998. [1](#)

- 368 [20] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, “Beyond a Gaussian denoiser: Residual learning of  
 369 deep CNN for image denoising,” *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3142–3155,  
 370 2017. [1](#), [3](#)
- 371 [21] M. Weigert, U. Schmidt, T. Boothe, *et al.*, “Content-aware image restoration: pushing the limits of  
 372 fluorescence microscopy,” *Nature Methods*, vol. 15, pp. 1090–1097, 2018. [1](#)
- 373 [22] T. Plotz and S. Roth, “Benchmarking denoising algorithms with real photographs,” in *Proceedings of the  
 374 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pp. 99–107, 2017. [1](#)
- 375 [23] J. Lehtinen, J. Munkberg, J. Hasselgren, S. Laine, T. Karras, M. Aittala, and T. Aila, “Noise2noise:  
 376 Learning image restoration without clean data,” in *Proceedings of the 35th International Conference on  
 377 Machine Learning (ICML)*, 2018. [1](#)
- 378 [24] J. Batson and L. Royer, “Noise2self: Blind denoising by self-supervision,” in *International Conference on  
 379 Machine Learning*, pp. 524–533, PMLR, 2019. [1](#)
- 380 [25] J. Ho, A. Jain, and P. Abbeel, “Denoising diffusion probabilistic models,” in *Advances in Neural Information  
 381 Processing Systems*, vol. 33, pp. 6840–6851, 2020. [1](#)
- 382 [26] M. Guo, Y. Wu, C. Hobson, *et al.*, “Deep learning-based aberration compensation improves contrast and  
 383 resolution in fluorescence microscopy,” *Nat Commun*, vol. 16, p. 313, 2025. [1](#)
- 384 [27] G. M. Hagen, J. Bendesky, R. Machado, T. A. Nguyen, T. Kumar, and J. Ventura, “Fluorescence microscopy  
 385 datasets for training deep neural networks,” *GigaScience*, vol. 10, no. 5, p. giab032, 2021. [1](#)
- 386 [28] C. Qiao, Y. Zeng, Q. Meng, X. Chen, H. Chen, T. Jiang, and Q. Dai, “Zero-shot learning enables instant  
 387 denoising and super-resolution in optical fluorescence microscopy,” *Nature Communications*, vol. 15, no. 1,  
 388 p. 4180, 2024. [2](#), [3](#)
- 389 [29] V. Sitzmann, J. Martel, A. Bergman, D. Lindell, and G. Wetzstein, “Implicit neural representations with  
 390 periodic activation functions,” in *Advances in Neural Information Processing Systems*, vol. 33, pp. 7462–  
 391 7473, NeurIPS, 2020. [2](#), [3](#)
- 392 [30] A. Molaei, A. Aminimehr, A. Tavakoli, A. Kazerouni, B. Azad, R. Azad, and D. Merhof, “Implicit neural  
 393 representation in medical imaging: A comparative survey,” in *Proceedings of the IEEE/CVF International  
 394 Conference on Computer Vision*, pp. 2381–2391, 2023. [2](#)
- 395 [31] D. Grattarola and P. Vandergheynst, “Generalised implicit neural representations,” in *Advances in Neural  
 396 Information Processing Systems*, 2022. [3](#)
- 397 [32] J. J. Park, P. Florence, J. Straub, R. Newcombe, and S. Lovegrove, “Deepsdf: Learning continuous signed  
 398 distance functions for shape representation,” in *Proceedings of the IEEE/CVF Conference on Computer  
 399 Vision and Pattern Recognition*, 2019. [3](#)
- 400 [33] L. Mescheder, M. Oechsle, M. Niemeyer, S. Nowozin, and A. Geiger, “Occupancy networks: Learning 3D  
 401 reconstruction in function space,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and  
 402 Pattern Recognition*, 2019. [3](#)
- 403 [34] Z. Chen and H. Zhang, “Learning implicit fields for generative shape modeling,” in *Proceedings of the  
 404 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019. [3](#)
- 405 [35] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, “Nerf: Representing  
 406 scenes as neural radiance fields for view synthesis,” *Communications of the ACM*, vol. 65, no. 1, pp. 99–106,  
 407 2021. [3](#), [4](#)
- 408 [36] T. Müller, A. Evans, C. Schied, and A. Keller, “Instant neural graphics primitives with a multiresolution  
 409 hash encoding,” *ACM Transactions on Graphics (ToG)*, 2022. [3](#)
- 410 [37] A. Chen, Z. Xu, A. Geiger, J. Yu, and H. Su, “TensoRF: Tensorial radiance fields,” in *Proceedings of the  
 411 European Conference on Computer Vision*, 2022. [3](#)
- 412 [38] A. Yu, V. Ye, M. Tancik, and A. Kanazawa, “Pixelnerf: Neural radiance fields from one or few images,” in  
 413 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4578–4587,  
 414 2021. [3](#)
- 415 [39] Q. Wang, Z. Wang, K. Genova, P. P. Srinivasan, H. Zhou, J. T. Barron, and T. Funkhouser, “Ibrnet: Learning  
 416 multi-view image-based rendering,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and  
 417 Pattern Recognition*, pp. 4690–4699, 2021. [3](#)

- 418 [40] J. T. Barron, B. Mildenhall, M. Tancik, P. Hedman, R. Martin-Brualla, and P. P. Srinivasan, “Mip-nerf:  
 419 A multiscale representation for anti-aliasing neural radiance fields,” in *Proceedings of the IEEE/CVF*  
 420 *International Conference on Computer Vision*, pp. 5855–5864, 2021. 3
- 421 [41] J. T. Barron, B. Mildenhall, D. Verbin, P. P. Srinivasan, and P. Hedman, “Zip-NeRF: Anti-aliased grid-based  
 422 neural radiance fields,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*,  
 423 2023. 3
- 424 [42] Z. Chen, L. Yang, J. H. Lai, and X. Xie, “Cunerf: Cube-based neural radiance field for zero-shot medical  
 425 image arbitrary-scale super resolution,” in *Proceedings of the IEEE/CVF International Conference on*  
 426 *Computer Vision*, pp. 21185–21195, 2023. 3, 5, 6
- 427 [43] J. McGinnis, S. Shit, H. B. Li, V. Sideri-Lampretsa, R. Graf, M. Dannecker, and B. Wiestler, “Single-subject  
 428 multi-contrast mri super-resolution via implicit neural representations,” in *International Conference on*  
 429 *Medical Image Computing and Computer-Assisted Intervention*, (Cham), pp. 173–183, Springer Nature  
 430 Switzerland, 2023. 3
- 431 [44] H. Kniesel, T. Ropinski, T. Bergner, K. S. Devan, C. Read, P. Walther, and P. Hermosilla, “Clean implicit  
 432 3d structure from noisy 2d stem images,” in *Proceedings of the IEEE/CVF Conference on Computer Vision*  
 433 and *Pattern Recognition*, pp. 20762–20772, 2022. 3
- 434 [45] G. Zhang, X. Li, Y. Zhang, X. Han, X. Li, J. Yu, and Q. Dai, “Bio-friendly long-term subcellular dynamic  
 435 recording by self-supervised image enhancement microscopy,” *Nature Methods*, vol. 20, no. 12, pp. 1957–  
 436 1970, 2023. 3, 5
- 437 [46] A. Buades, B. Coll, and J. M. Morel, “Non-local means denoising,” *Image Processing On Line*, vol. 1,  
 438 pp. 208–212, 2011. 3
- 439 [47] M. Farge, “Wavelet transforms and their applications to turbulence,” *Annual Review of Fluid Mechanics*,  
 440 vol. 24, no. 1, pp. 395–458, 1992. 3
- 441 [48] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu, “Image super-resolution using very deep residual  
 442 channel attention networks,” in *Proceedings of the European Conference on Computer Vision (ECCV)*,  
 443 pp. 286–301, 2018. 3
- 444 [49] J. Chen, H. Sasaki, H.-C. Lai, et al., “Three-dimensional residual channel attention networks denoise and  
 445 sharpen fluorescence microscopy image volumes,” *Nature Methods*, vol. 18, pp. 678–687, 2021. 3
- 446 [50] J. Lehtinen, “Noise2noise: Learning image restoration without clean data,” *arXiv preprint*  
 447 *arXiv:1803.04189*, 2018. 3
- 448 [51] T. Pang, H. Zheng, Y. Quan, and H. Ji, “Recorrupted-to-recorrupted: Unsupervised deep learning for image  
 449 denoising,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*,  
 450 pp. 2043–2052, 2021. 3
- 451 [52] S. R. Sternberg, “Biomedical image processing,” *Computer*, vol. 16, no. 1, pp. 22–34, 1983. 3
- 452 [53] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *International Journal of Computer*  
 453 *Vision*, vol. 60, no. 2, pp. 91–110, 2004. 3
- 454 [54] W. Zhao, S. Zhao, Z. Han, X. Ding, G. Hu, L. Qu, and H. Li, “Enhanced detection of fluorescence  
 455 fluctuations for high-throughput super-resolution imaging,” *Nature Photonics*, vol. 17, no. 9, pp. 806–813,  
 456 2023. 3
- 457 [55] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the*  
 458 *IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016. 4
- 459 [56] J. Cao, M. K. Wong, Z. Zhao, and H. Yan, “3dmms: robust 3d membrane morphological segmentation of  
 460 *C. elegans* embryo,” *BMC Bioinformatics*, vol. 20, pp. 1–13, 2019. 5
- 461 [57] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, “Multiscale vessel enhancement filtering,”  
 462 in *Medical Image Computing and Computer-Assisted Intervention — MICCAI’98: First International*  
 463 *Conference Cambridge, MA, USA, October 11–13, 1998 Proceedings 1*, pp. 130–137, Springer Berlin  
 464 Heidelberg, 1998. 5
- 465 [58] X. Huang, J. Fan, L. Li, H. Liu, R. Wu, Y. Wu, and L. Chen, “Fast, long-term, super-resolution imaging  
 466 with hessian structured illumination microscopy,” *Nature Biotechnology*, vol. 36, no. 5, pp. 451–459, 2018.  
 467 5

- 468 [59] X. Xu, S. Xu, L. Jin, and E. Song, “Characteristic analysis of otsu threshold and its applications,” *Pattern*  
469 *Recognition Letters*, vol. 32, no. 7, pp. 956–961, 2011. 6
- 470 [60] L. Yen-Chen, “Nerf-pytorch.” <https://github.com/yenchenlin/nerf-pytorch/>, 2020. 6
- 471 [61] L. Ma, X. Li, J. Liao, Q. Zhang, X. Wang, J. Wang, and P. V. Sander, “Deblur-nerf: Neural radiance  
472 fields from blurry images,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern*  
473 *Recognition*, pp. 12861–12870, 2022. 6
- 474 [62] J. Cao, G. Guan, V. W. S. Ho, M. K. Wong, L. Y. Chan, C. Tang, and H. Yan, “Establishment of  
475 a morphological atlas of the *C. elegans* embryo using deep-learning-based 4d segmentation,” *Nature*  
476 *Communications*, vol. 11, no. 1, p. 6254, 2020. 7, 8, 9
- 477 [63] A. Domingo-Muelas, R. M. Skory, A. A. Moverley, G. Ardestani, O. Pomp, C. Rubio, and N. Plachta,  
478 “Human embryo live imaging reveals nuclear dna shedding during blastocyst expansion and biopsy,” *Cell*,  
479 vol. 186, no. 15, pp. 3166–3181, 2023. 7, 8
- 480 [64] G. Guan, Z. Li, Y. Ma, P. Ye, J. Cao, M. K. Wong, and Z. Zhao, “Cell lineage-resolved embryonic mor-  
481 phological map reveals signaling associated with cell fate and size asymmetry,” *Nature Communications*,  
482 vol. 16, no. 1, p. 3700, 2025. 9
- 483 [65] K. McDole, L. Guignard, F. Amat, A. Berger, G. Malandain, L. A. Royer, and P. J. Keller, “In toto imaging  
484 and reconstruction of post-implantation mouse development at the single-cell level,” *Cell*, vol. 175, no. 3,  
485 pp. 859–876, 2018. 9
- 486 [66] A. Wolny, L. Cerrone, A. Vijayan, R. Tofanelli, A. V. Barro, M. Louveaux, and A. Kreshuk, “Accurate and  
487 versatile 3d segmentation of plant tissues at cellular resolution,” *Elife*, vol. 9, p. e57613, 2020. 9
- 488 [67] K. He, G. Gkioxari, P. Dollár, and R. Girshick, “Mask r-cnn,” in *Proceedings of the IEEE International*  
489 *Conference on Computer Vision (ICCV)*, pp. 2961–2969, 2017. 9
- 490 [68] M. Chlebiej, I. Gorczynska, A. Rutkowski, J. Kluczewski, T. Grzona, E. Pijewska, and M. Szkulmowski,  
491 “Quality improvement of oct angiograms with elliptical directional filtering,” *Biomedical Optics Express*,  
492 vol. 10, no. 2, pp. 1013–1031, 2019. 9

493 **NeurIPS Paper Checklist**

494 The checklist is designed to encourage best practices for responsible machine learning research,  
495 addressing issues of reproducibility, transparency, research ethics, and societal impact. Do not remove  
496 the checklist: **The papers not including the checklist will be desk rejected.** The checklist should  
497 follow the references and follow the (optional) supplemental material. The checklist does NOT count  
498 towards the page limit.

499 Please read the checklist guidelines carefully for information on how to answer these questions. For  
500 each question in the checklist:

- 501 • You should answer [Yes] , [No] , or [NA] .  
502 • [NA] means either that the question is Not Applicable for that particular paper or the  
503 relevant information is Not Available.  
504 • Please provide a short (1–2 sentence) justification right after your answer (even for NA).

505 **The checklist answers are an integral part of your paper submission.** They are visible to the  
506 reviewers, area chairs, senior area chairs, and ethics reviewers. You will be asked to also include it  
507 (after eventual revisions) with the final version of your paper, and its final version will be published  
508 with the paper.

509 The reviewers of your paper will be asked to use the checklist as one of the factors in their evaluation.  
510 While "[Yes]" is generally preferable to "[No]", it is perfectly acceptable to answer "[No]" provided a  
511 proper justification is given (e.g., "error bars are not reported because it would be too computationally  
512 expensive" or "we were unable to find the license for the dataset we used"). In general, answering  
513 "[No]" or "[NA]" is not grounds for rejection. While the questions are phrased in a binary way, we  
514 acknowledge that the true answer is often more nuanced, so please just use your best judgment and  
515 write a justification to elaborate. All supporting evidence can appear either in the main paper or the  
516 supplemental material, provided in appendix. If you answer [Yes] to a question, in the justification  
517 please point to the section(s) where related material for the question can be found.

518 **IMPORTANT**, please:

- 519 • **Delete this instruction block, but keep the section heading “NeurIPS paper checklist”,**  
520 • **Keep the checklist subsection headings, questions/answers and guidelines below.**  
521 • **Do not modify the questions and only use the provided macros for your answers.**

522 **1. Claims**

523 Question: Do the main claims made in the abstract and introduction accurately reflect the  
524 paper's contributions and scope?

525 Answer: [Yes]

526 Justification: The main claims accurately reflect our contributions and are fully supported  
527 by the results.

528 Guidelines:

- 529 • The answer NA means that the abstract and introduction do not include the claims  
530 made in the paper.  
531 • The abstract and/or introduction should clearly state the claims made, including the  
532 contributions made in the paper and important assumptions and limitations. A No or  
533 NA answer to this question will not be perceived well by the reviewers.  
534 • The claims made should match theoretical and experimental results, and reflect how  
535 much the results can be expected to generalize to other settings.  
536 • It is fine to include aspirational goals as motivation as long as it is clear that these goals  
537 are not attained by the paper.

538 **2. Limitations**

539 Question: Does the paper discuss the limitations of the work performed by the authors?

540 Answer: [Yes]

541 Justification: Please refer to 4.4 and Supplementary for detailed discussion of limitations.

542 Guidelines:

- 543 • The answer NA means that the paper has no limitation while the answer No means that  
544 the paper has limitations, but those are not discussed in the paper.
- 545 • The authors are encouraged to create a separate "Limitations" section in their paper.
- 546 • The paper should point out any strong assumptions and how robust the results are to  
547 violations of these assumptions (e.g., independence assumptions, noiseless settings,  
548 model well-specification, asymptotic approximations only holding locally). The authors  
549 should reflect on how these assumptions might be violated in practice and what the  
550 implications would be.
- 551 • The authors should reflect on the scope of the claims made, e.g., if the approach was  
552 only tested on a few datasets or with a few runs. In general, empirical results often  
553 depend on implicit assumptions, which should be articulated.
- 554 • The authors should reflect on the factors that influence the performance of the approach.  
555 For example, a facial recognition algorithm may perform poorly when image resolution  
556 is low or images are taken in low lighting. Or a speech-to-text system might not be  
557 used reliably to provide closed captions for online lectures because it fails to handle  
558 technical jargon.
- 559 • The authors should discuss the computational efficiency of the proposed algorithms  
560 and how they scale with dataset size.
- 561 • If applicable, the authors should discuss possible limitations of their approach to  
562 address problems of privacy and fairness.
- 563 • While the authors might fear that complete honesty about limitations might be used by  
564 reviewers as grounds for rejection, a worse outcome might be that reviewers discover  
565 limitations that aren't acknowledged in the paper. The authors should use their best  
566 judgment and recognize that individual actions in favor of transparency play an impor-  
567 tant role in developing norms that preserve the integrity of the community. Reviewers  
568 will be specifically instructed to not penalize honesty concerning limitations.

569 **3. Theory Assumptions and Proofs**

570 Question: For each theoretical result, does the paper provide the full set of assumptions and  
571 a complete (and correct) proof?

572 Answer: [NA]

573 Justification: This paper is primarily empirical and does not contain theoretical results  
574 requiring formal proofs.

575 Guidelines:

- 576 • The answer NA means that the paper does not include theoretical results.
- 577 • All the theorems, formulas, and proofs in the paper should be numbered and cross-  
578 referenced.
- 579 • All assumptions should be clearly stated or referenced in the statement of any theorems.
- 580 • The proofs can either appear in the main paper or the supplemental material, but if  
581 they appear in the supplemental material, the authors are encouraged to provide a short  
582 proof sketch to provide intuition.
- 583 • Inversely, any informal proof provided in the core of the paper should be complemented  
584 by formal proofs provided in appendix or supplemental material.
- 585 • Theorems and Lemmas that the proof relies upon should be properly referenced.

586 **4. Experimental Result Reproducibility**

587 Question: Does the paper fully disclose all the information needed to reproduce the main ex-  
588 perimental results of the paper to the extent that it affects the main claims and/or conclusions  
589 of the paper (regardless of whether the code and data are provided or not)?

590 Answer: [Yes]

591 Justification: Detailed implementation specifics and experimental settings are provided in  
592 Sections 3.5 and 4.1, including key parameters, model configurations, and evaluation metrics.  
593 The complete prompts used in the paper are provided in the supplementary materials.

594 Guidelines:

- 595 • The answer NA means that the paper does not include experiments.
- 596 • If the paper includes experiments, a No answer to this question will not be perceived
- 597 well by the reviewers: Making the paper reproducible is important, regardless of
- 598 whether the code and data are provided or not.
- 599 • If the contribution is a dataset and/or model, the authors should describe the steps taken
- 600 to make their results reproducible or verifiable.
- 601 • Depending on the contribution, reproducibility can be accomplished in various ways.
- 602 For example, if the contribution is a novel architecture, describing the architecture fully
- 603 might suffice, or if the contribution is a specific model and empirical evaluation, it may
- 604 be necessary to either make it possible for others to replicate the model with the same
- 605 dataset, or provide access to the model. In general, releasing code and data is often
- 606 one good way to accomplish this, but reproducibility can also be provided via detailed
- 607 instructions for how to replicate the results, access to a hosted model (e.g., in the case
- 608 of a large language model), releasing of a model checkpoint, or other means that are
- 609 appropriate to the research performed.
- 610 • While NeurIPS does not require releasing code, the conference does require all submissions
- 611 to provide some reasonable avenue for reproducibility, which may depend on the
- 612 nature of the contribution. For example
  - 613 (a) If the contribution is primarily a new algorithm, the paper should make it clear how
  - 614 to reproduce that algorithm.
  - 615 (b) If the contribution is primarily a new model architecture, the paper should describe
  - 616 the architecture clearly and fully.
  - 617 (c) If the contribution is a new model (e.g., a large language model), then there should
  - 618 either be a way to access this model for reproducing the results or a way to reproduce
  - 619 the model (e.g., with an open-source dataset or instructions for how to construct
  - 620 the dataset).
  - 621 (d) We recognize that reproducibility may be tricky in some cases, in which case
  - 622 authors are welcome to describe the particular way they provide for reproducibility.
  - 623 In the case of closed-source models, it may be that access to the model is limited in
  - 624 some way (e.g., to registered users), but it should be possible for other researchers
  - 625 to have some path to reproducing or verifying the results.

626 **5. Open access to data and code**

627 Question: Does the paper provide open access to the data and code, with sufficient instructions

628 to faithfully reproduce the main experimental results, as described in supplemental

629 material?

630 Answer: [Yes]

631 Justification: The link is presented in Supplementary.

632 Guidelines:

- 633 • The answer NA means that paper does not include experiments requiring code.
- 634 • Please see the NeurIPS code and data submission guidelines (<https://nips.cc/public/guides/CodeSubmissionPolicy>) for more details.
- 635 • While we encourage the release of code and data, we understand that this might not be
- 636 possible, so “No” is an acceptable answer. Papers cannot be rejected simply for not
- 637 including code, unless this is central to the contribution (e.g., for a new open-source
- 638 benchmark).
- 639 • The instructions should contain the exact command and environment needed to run to
- 640 reproduce the results. See the NeurIPS code and data submission guidelines (<https://nips.cc/public/guides/CodeSubmissionPolicy>) for more details.
- 641 • The authors should provide instructions on data access and preparation, including how
- 642 to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- 643 • The authors should provide scripts to reproduce all experimental results for the new
- 644 proposed method and baselines. If only a subset of experiments are reproducible, they
- 645 should state which ones are omitted from the script and why.

- 648           • At submission time, to preserve anonymity, the authors should release anonymized  
649            versions (if applicable).  
650           • Providing as much information as possible in supplemental material (appended to the  
651            paper) is recommended, but including URLs to data and code is permitted.

652           **6. Experimental Setting/Details**

653           Question: Does the paper specify all the training and test details (e.g., data splits, hyper-  
654           parameters, how they were chosen, type of optimizer, etc.) necessary to understand the  
655           results?

656           Answer: [Yes]

657           Justification: Experimental details, including model configurations, evaluation metrics,  
658           datasets, and implementation details, are provided in Sections 3.5 and 4.1. Additional details  
659           are included in the supplementary materials.

660           Guidelines:

- 661           • The answer NA means that the paper does not include experiments.  
662           • The experimental setting should be presented in the core of the paper to a level of detail  
663            that is necessary to appreciate the results and make sense of them.  
664           • The full details can be provided either with the code, in appendix, or as supplemental  
665            material.

666           **7. Experiment Statistical Significance**

667           Question: Does the paper report error bars suitably and correctly defined or other appropriate  
668           information about the statistical significance of the experiments?

669           Answer: [Yes]

670           Justification: The STD of each scores are reported in Table.1.

671           Guidelines:

- 672           • The answer NA means that the paper does not include experiments.  
673           • The authors should answer "Yes" if the results are accompanied by error bars, confi-  
674            dence intervals, or statistical significance tests, at least for the experiments that support  
675            the main claims of the paper.  
676           • The factors of variability that the error bars are capturing should be clearly stated (for  
677            example, train/test split, initialization, random drawing of some parameter, or overall  
678            run with given experimental conditions).  
679           • The method for calculating the error bars should be explained (closed form formula,  
680            call to a library function, bootstrap, etc.)  
681           • The assumptions made should be given (e.g., Normally distributed errors).  
682           • It should be clear whether the error bar is the standard deviation or the standard error  
683            of the mean.  
684           • It is OK to report 1-sigma error bars, but one should state it. The authors should  
685            preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis  
686            of Normality of errors is not verified.  
687           • For asymmetric distributions, the authors should be careful not to show in tables or  
688            figures symmetric error bars that would yield results that are out of range (e.g. negative  
689            error rates).  
690           • If error bars are reported in tables or plots, The authors should explain in the text how  
691            they were calculated and reference the corresponding figures or tables in the text.

692           **8. Experiments Compute Resources**

693           Question: For each experiment, does the paper provide sufficient information on the com-  
694           puter resources (type of compute workers, memory, time of execution) needed to reproduce  
695           the experiments?

696           Answer: [Yes]

697           Justification: Experiments compute resources are provided in 3.5.

698           Guidelines:

- 699           • The answer NA means that the paper does not include experiments.  
700           • The paper should indicate the type of compute workers CPU or GPU, internal cluster,  
701            or cloud provider, including relevant memory and storage.  
702           • The paper should provide the amount of compute required for each of the individual  
703            experimental runs as well as estimate the total compute.  
704           • The paper should disclose whether the full research project required more compute  
705            than the experiments reported in the paper (e.g., preliminary or failed experiments that  
706            didn't make it into the paper).

707           **9. Code Of Ethics**

708           Question: Does the research conducted in the paper conform, in every respect, with the  
709           NeurIPS Code of Ethics <https://neurips.cc/public/EthicsGuidelines>?

710           Answer: [Yes]

711           Justification: Our research aims to optimize microscopy imaging, which aligns with the  
712           NeurIPS Code of Ethics regarding data privacy, copyright protection, and responsible AI  
713           development.

714           Guidelines:

- 715           • The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.  
716           • If the authors answer No, they should explain the special circumstances that require a  
717            deviation from the Code of Ethics.  
718           • The authors should make sure to preserve anonymity (e.g., if there is a special consid-  
719            eration due to laws or regulations in their jurisdiction).

720           **10. Broader Impacts**

721           Question: Does the paper discuss both potential positive societal impacts and negative  
722           societal impacts of the work performed?

723           Answer: [Yes]

724           Justification: Yes, our research has optimized the results of live fluorescence microscopy  
725           imaging and has promoted the fields of biological imaging and data analysis.

726           Guidelines:

- 727           • The answer NA means that there is no societal impact of the work performed.  
728           • If the authors answer NA or No, they should explain why their work has no societal  
729            impact or why the paper does not address societal impact.  
730           • Examples of negative societal impacts include potential malicious or unintended uses  
731            (e.g., disinformation, generating fake profiles, surveillance), fairness considerations  
732            (e.g., deployment of technologies that could make decisions that unfairly impact specific  
733            groups), privacy considerations, and security considerations.  
734           • The conference expects that many papers will be foundational research and not tied  
735            to particular applications, let alone deployments. However, if there is a direct path to  
736            any negative applications, the authors should point it out. For example, it is legitimate  
737            to point out that an improvement in the quality of generative models could be used to  
738            generate deepfakes for disinformation. On the other hand, it is not needed to point out  
739            that a generic algorithm for optimizing neural networks could enable people to train  
740            models that generate Deepfakes faster.  
741           • The authors should consider possible harms that could arise when the technology is  
742            being used as intended and functioning correctly, harms that could arise when the  
743            technology is being used as intended but gives incorrect results, and harms following  
744            from (intentional or unintentional) misuse of the technology.  
745           • If there are negative societal impacts, the authors could also discuss possible mitigation  
746            strategies (e.g., gated release of models, providing defenses in addition to attacks,  
747            mechanisms for monitoring misuse, mechanisms to monitor how a system learns from  
748            feedback over time, improving the efficiency and accessibility of ML).

749           **11. Safeguards**

750 Question: Does the paper describe safeguards that have been put in place for responsible  
751 release of data or models that have a high risk for misuse (e.g., pretrained language models,  
752 image generators, or scraped datasets)?

753 Answer: [NA]

754 Justification: The paper poses no such risks.

755 Guidelines:

- 756 • The answer NA means that the paper poses no such risks.
- 757 • Released models that have a high risk for misuse or dual-use should be released with  
758 necessary safeguards to allow for controlled use of the model, for example by requiring  
759 that users adhere to usage guidelines or restrictions to access the model or implementing  
760 safety filters.
- 761 • Datasets that have been scraped from the Internet could pose safety risks. The authors  
762 should describe how they avoided releasing unsafe images.
- 763 • We recognize that providing effective safeguards is challenging, and many papers do  
764 not require this, but we encourage authors to take this into account and make a best  
765 faith effort.

## 766 12. Licenses for existing assets

767 Question: Are the creators or original owners of assets (e.g., code, data, models), used in  
768 the paper, properly credited and are the license and terms of use explicitly mentioned and  
769 properly respected?

770 Answer: [Yes]

771 Justification: All models and datasets used in our experiments are properly cited in the paper,  
772 and we fully comply with their respective licenses and terms of use.

773 Guidelines:

- 774 • The answer NA means that the paper does not use existing assets.
- 775 • The authors should cite the original paper that produced the code package or dataset.
- 776 • The authors should state which version of the asset is used and, if possible, include a  
777 URL.
- 778 • The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- 779 • For scraped data from a particular source (e.g., website), the copyright and terms of  
780 service of that source should be provided.
- 781 • If assets are released, the license, copyright information, and terms of use in the  
782 package should be provided. For popular datasets, [paperswithcode.com/datasets](https://paperswithcode.com/datasets)  
783 has curated licenses for some datasets. Their licensing guide can help determine the  
784 license of a dataset.
- 785 • For existing datasets that are re-packaged, both the original license and the license of  
786 the derived asset (if it has changed) should be provided.
- 787 • If this information is not available online, the authors are encouraged to reach out to  
788 the asset's creators.

## 789 13. New Assets

790 Question: Are new assets introduced in the paper well documented and is the documentation  
791 provided alongside the assets?

792 Answer: [Yes]

793 Justification: Yes, we provide a new 4D imaging dataset in the [4.1](#) section

794 Guidelines:

- 795 • The answer NA means that the paper does not release new assets.
- 796 • Researchers should communicate the details of the dataset/code/model as part of their  
797 submissions via structured templates. This includes details about training, license,  
798 limitations, etc.
- 799 • The paper should discuss whether and how consent was obtained from people whose  
800 asset is used.

- 801           • At submission time, remember to anonymize your assets (if applicable). You can either  
802            create an anonymized URL or include an anonymized zip file.

803          **14. Crowdsourcing and Research with Human Subjects**

804          Question: For crowdsourcing experiments and research with human subjects, does the paper  
805          include the full text of instructions given to participants and screenshots, if applicable, as  
806          well as details about compensation (if any)?

807          Answer: [NA]

808          Justification: The paper does not involve crowdsourcing nor research with human subjects.

809          Guidelines:

- 810           • The answer NA means that the paper does not involve crowdsourcing nor research with  
811            human subjects.  
812           • Including this information in the supplemental material is fine, but if the main contribu-  
813            tion of the paper involves human subjects, then as much detail as possible should be  
814            included in the main paper.  
815           • According to the NeurIPS Code of Ethics, workers involved in data collection, curation,  
816            or other labor should be paid at least the minimum wage in the country of the data  
817            collector.

818          **15. Institutional Review Board (IRB) Approvals or Equivalent for Research with Human  
819            Subjects**

820          Question: Does the paper describe potential risks incurred by study participants, whether  
821          such risks were disclosed to the subjects, and whether Institutional Review Board (IRB)  
822          approvals (or an equivalent approval/review based on the requirements of your country or  
823          institution) were obtained?

824          Answer: [NA]

825          Justification: The paper does not involve crowdsourcing nor research with human subjects.

826          Guidelines:

- 827           • The answer NA means that the paper does not involve crowdsourcing nor research with  
828            human subjects.  
829           • Depending on the country in which research is conducted, IRB approval (or equivalent)  
830            may be required for any human subjects research. If you obtained IRB approval, you  
831            should clearly state this in the paper.  
832           • We recognize that the procedures for this may vary significantly between institutions  
833            and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the  
834            guidelines for their institution.  
835           • For initial submissions, do not include any information that would break anonymity (if  
836            applicable), such as the institution conducting the review.