

A Standalone PSF Model for Physics-Guided Degradation

Tanush Swaminathan

Abstract

Learning-based 3D fluorescence microscopy restoration often relies on simplified forward degradation (typically a fixed, space-invariant Gaussian blur)[4, 7], ignoring modality-specific optics and depth-dependent aberrations and biasing axial reconstructions. We present a *standalone*, physics-guided PSF module that predicts a *depth-indexed* PSF bank $\{\text{PSF}_z\}_{z=1}^Z$ directly from anisotropic volumes. Across multiple microscopy modalities, these depth-dependent PSFs improve reconstruction quality relative to Gaussian degradation and strong PSF baselines[3, 9], underscoring physics-based degradation modeling as a practical lever for more faithful 3D restoration.

1 Introduction

3D fluorescence microscopy enables volumetric analysis of subcellular structures, but practical stacks are often anisotropic: axial resolution is worse than lateral, yielding elongated blur along z . Many restoration pipelines approximate forward degradation using a fixed Gaussian kernel[4, 7]; this is convenient but mismatched to real optics, where PSFs are modality-specific and can change with depth due to aberrations and scattering [8, 5].

Key innovation. We validate a *drop-in* PSF module that replaces Gaussian blur with instrument-matched kernels via: (i) **depth-dependent PSFs** (per-slice bank $\{\text{PSF}_z\}$), (ii) **physics-guided synthesis** (pupil modeling + Fourier propagation) [3, 8], and (iii) an **expanded Zernike basis** to capture richer aberrations [8]. Full physics details and equations are provided in Appendix C–C.4.

2 Results

Kernel-swap summary. We evaluate degradation-model quality using a controlled **kernel-swap** protocol: for each [4, 7], we replace only the default *Gaussian* blur kernel(s) used for forward degradation with an alternative PSF model while holding the rest of the pipeline fixed. Exact “changed vs. held fixed” settings are listed in Appendix E.

Base	PSF	$\Delta \text{PSNR}\% \uparrow$	$\Delta \text{SSIM}\% \uparrow$	$\Delta \text{NIQE}\% \downarrow$
SSAI3D	GL	2.0	1.2	4.4
	SFE-only	2.9	1.8	5.8
	Ours	3.7	2.1	6.8
SelfNet	GL	5.8	1.8	3.6
	SFE-only	6.7	2.2	4.4
	Ours	7.5	2.5	4.9

The table above reports average relative change of our model + SOTA PSF models (Gibson-Lanni and SFE-Net w/o physics) [3, 9] vs. simple Gaussian, when implemented within the reconstruction models. Deltas were aggregated across the four microscopy modalities in Table 1.

3 Methods

3.1 Model (SFE-Net + physics)

Given an anisotropic volume, the module extracts axial evidence (XZ/YZ views) and uses **SFE-Net** [9] to regress compact, interpretable PSF parameters (e.g., Zernike coefficients [8] and related optics terms). A physics step then synthesizes a normalized 3D PSF for each depth slice, yielding a depth-indexed bank $\{\text{PSF}_z\}$ used for forward degradation. The synthesis equations and Zernike parameterization are provided in (Appendix C.2–C.3).

3.2 Depth-dependent aberrations

To model depth-varying aberrations, inference and synthesis are performed per depth, producing kernels that can broaden and change shape across the volume (Appendix C.4). This is especially relevant for deep tissue imaging, where common analytic PSF models typically use a single parameter set across depth.

3.3 Expanded Zernike basis

We use an **expanded Zernike parameterization** for pupil-phase aberrations and regress coefficients from axial evidence. Ablating the mode count M reveals a stability–accuracy trade-off: higher-order modes reduce mismatch until a plateau C.2–C.3. Related microscopy PSF / deconvolution discussions include deconvolution guidance and depth-dependent aberration effects [5].

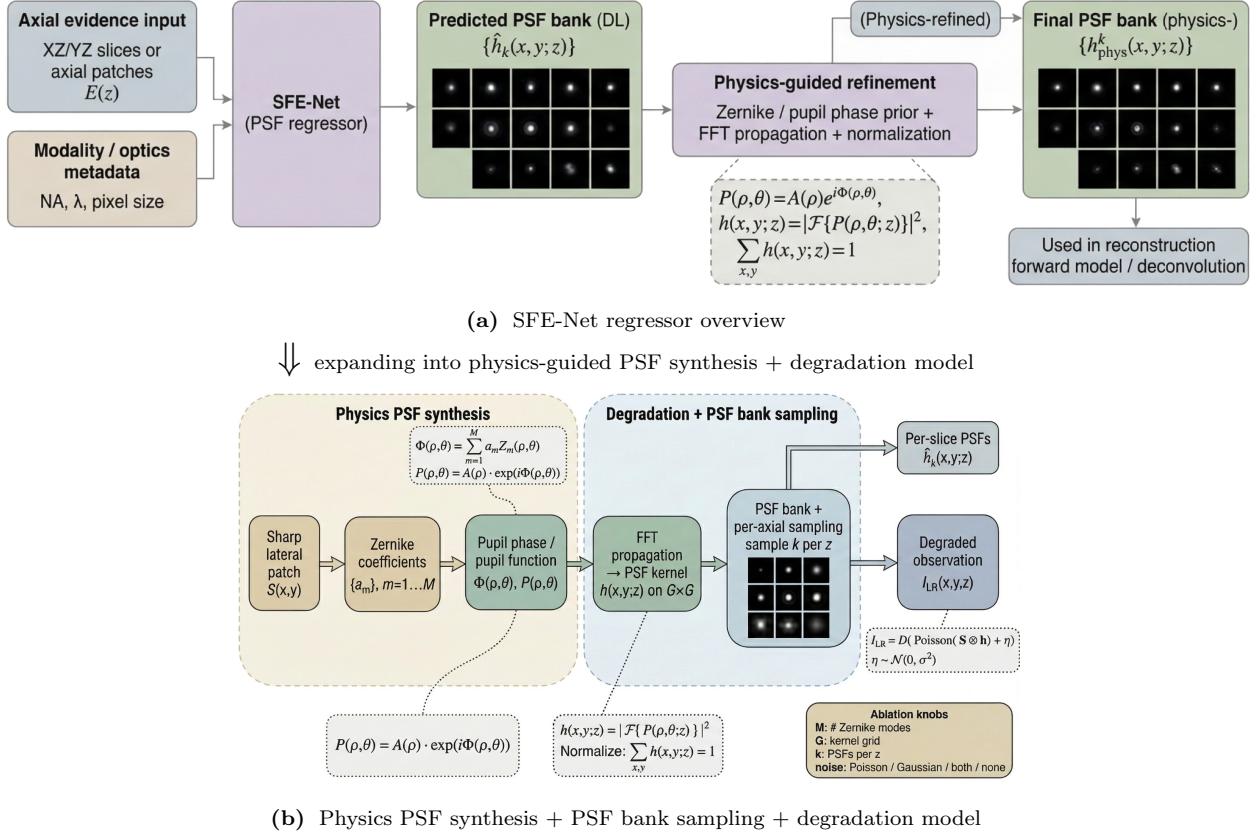


Figure 1: **Overview of the standalone PSF module.** **Learned stage (SFE-Net):** predicts aberration parameterization (e.g., Zernike coefficients) as a function of depth. **Physics stage:** constructs a pupil function and applies FFT. **Output:** an energy-normalized, depth-indexed PSF bank $\{PSF_z\}$.

4 Validation

4.1 Main Experiments

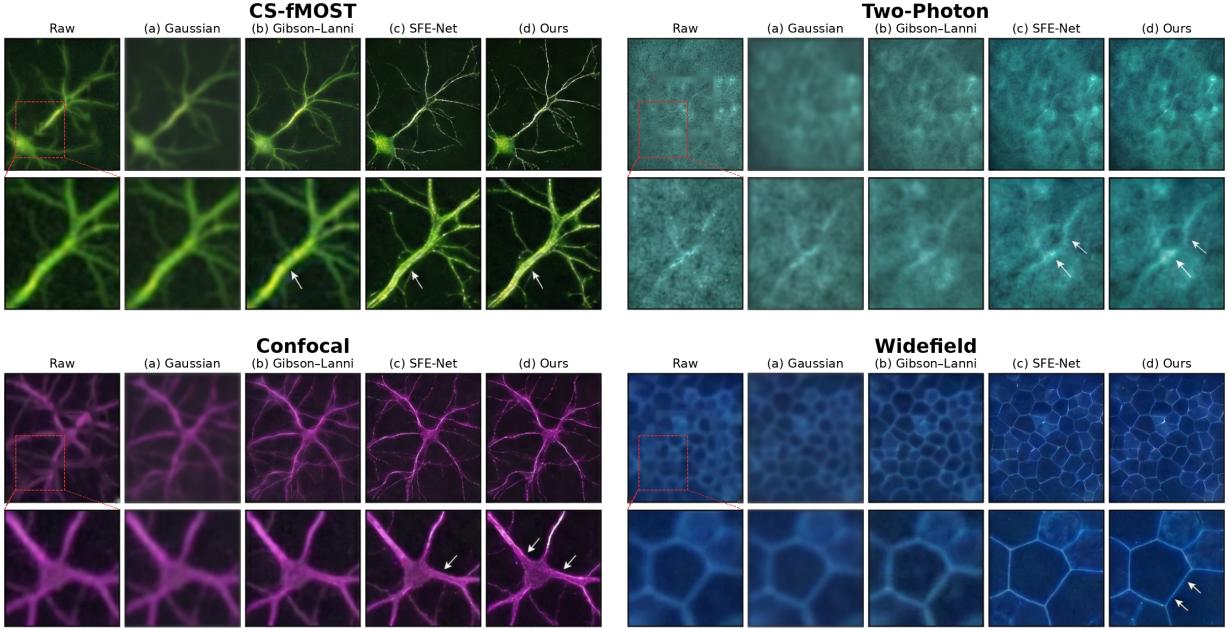


Figure 2: **Qualitative reconstruction comparisons.** Raw input versus reconstructions using (a) Gaussian, (b) Gibson-Lanni, (c) SFE-Net, and (d) our physics-guided PSF bank.

Table 1: **Reconstruction performance under the kernel-swap protocol.** For each baseline (SelfNet, SSAI3D), we replace the default Gaussian degradation with (i) Gibson–Lanni, (ii) SFE-Net predicted PSFs, or (iii) our physics-guided, depth-dependent PSF bank. Best per system are highlighted.

System	Method	Lateral			Axial	
		PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow	NIQE \downarrow	PIQE \downarrow
Confocal	SelfNet (Gaussian)	31.18 \pm 0.081	0.923 \pm 0.003	0.156 \pm 0.0055	6.55 \pm 0.089	43.6 \pm 0.030
	SelfNet (Gibson–Lanni)	32.60 \pm 0.050	0.940 \pm 0.006	0.152 \pm 0.002	6.18 \pm 0.090	42.4 \pm 0.050
	SelfNet (SFE-Net)	32.86 \pm 0.048	0.944 \pm 0.006	0.151 \pm 0.002	6.09 \pm 0.090	42.1 \pm 0.050
	SelfNet (Physics)	33.10 \pm 0.040	0.948 \pm 0.007	0.151 \pm 0.002	6.05 \pm 0.090	41.8 \pm 0.05
	SSAI3D (Gaussian)	32.20 \pm 0.045	0.931 \pm 0.012	0.154 \pm 0.007	6.76 \pm 0.103	44.7 \pm 0.024
	SSAI3D (Gibson–Lanni)	33.10 \pm 0.040	0.945 \pm 0.008	0.151 \pm 0.002	6.10 \pm 0.090	42.0 \pm 0.050
	SSAI3D (SFE-Net)	33.48 \pm 0.038	0.951 \pm 0.008	0.150 \pm 0.002	5.98 \pm 0.09	41.5 \pm 0.05
Two-Photon	SelfNet (Gaussian)	28.02 \pm 0.011	0.861 \pm 0.013	0.120 \pm 0.056	7.88 \pm 0.061	52.8 \pm 0.006
	SelfNet (Gibson–Lanni)	30.05 \pm 0.050	0.870 \pm 0.010	0.112 \pm 0.012	7.86 \pm 0.070	52.2 \pm 0.050
	SelfNet (SFE-Net)	30.18 \pm 0.055	0.874 \pm 0.010	0.111 \pm 0.012	7.84 \pm 0.075	51.8 \pm 0.050
	SelfNet (Physics)	30.30 \pm 0.060	0.878 \pm 0.008	0.108 \pm 0.012	7.82 \pm 0.070	51.2 \pm 0.050
	SSAI3D (Gaussian)	28.70 \pm 0.016	0.853 \pm 0.003	0.112 \pm 0.006	7.96 \pm 0.10	53.7 \pm 0.003
	SSAI3D (Gibson–Lanni)	29.60 \pm 0.060	0.873 \pm 0.008	0.109 \pm 0.010	7.85 \pm 0.080	51.8 \pm 0.050
	SSAI3D (SFE-Net)	30.02 \pm 0.065	0.881 \pm 0.007	0.107 \pm 0.012	7.79 \pm 0.075	50.9 \pm 0.05
Widefield	SelfNet (Gaussian)	30.71 \pm 0.106	0.902 \pm 0.009	0.261 \pm 0.004	7.26 \pm 0.102	48.7 \pm 0.022
	SelfNet (Gibson–Lanni)	31.50 \pm 0.200	0.915 \pm 0.020	0.254 \pm 0.002	6.70 \pm 0.090	45.2 \pm 0.050
	SelfNet (SFE-Net)	31.75 \pm 0.210	0.917 \pm 0.020	0.253 \pm 0.002	6.61 \pm 0.090	44.7 \pm 0.050
	SelfNet (Physics)	31.95 \pm 0.220	0.919 \pm 0.020	0.252 \pm 0.002	6.55 \pm 0.090	44.2 \pm 0.050
	SSAI3D (Gaussian)	31.69 \pm 0.120	0.913 \pm 0.015	0.255 \pm 0.003	6.94 \pm 0.011	44.33 \pm 0.060
	SSAI3D (Gibson–Lanni)	32.00 \pm 0.180	0.919 \pm 0.018	0.252 \pm 0.002	6.62 \pm 0.090	44.25 \pm 0.050
	SSAI3D (SFE-Net)	32.18 \pm 0.200	0.922 \pm 0.020	0.251 \pm 0.002	6.49 \pm 0.090	44.18 \pm 0.05
CS-fMOST	SelfNet (Gaussian)	25.01 \pm 0.120	0.803 \pm 0.010	0.154 \pm 0.001	7.91 \pm 0.087	58.6 \pm 0.050
	SelfNet (Gibson–Lanni)	27.20 \pm 0.250	0.825 \pm 0.006	0.120 \pm 0.012	7.86 \pm 0.10	56.8 \pm 0.050
	SelfNet (SFE-Net)	27.55 \pm 0.260	0.829 \pm 0.006	0.115 \pm 0.012	7.83 \pm 0.10	56.0 \pm 0.050
	SelfNet (Physics)	27.90 \pm 0.280	0.831 \pm 0.004	0.110 \pm 0.012	7.80 \pm 0.10	55.3 \pm 0.050
	SSAI3D (Gaussian)	27.80 \pm 0.200	0.828 \pm 0.003	0.138 \pm 0.006	8.03 \pm 0.12	58.7 \pm 0.060
	SSAI3D (Gibson–Lanni)	28.05 \pm 0.260	0.831 \pm 0.004	0.116 \pm 0.012	7.88 \pm 0.10	56.0 \pm 0.050
	SSAI3D (SFE-Net)	28.20 \pm 0.290	0.833 \pm 0.003	0.109 \pm 0.014	7.79 \pm 0.10	55.0 \pm 0.05
	SSAI3D (Physics)	28.36 \pm 0.340	0.834 \pm 0.002	0.105 \pm 0.016	7.71 \pm 0.10	54.1 \pm 0.042

4.2 Ablation

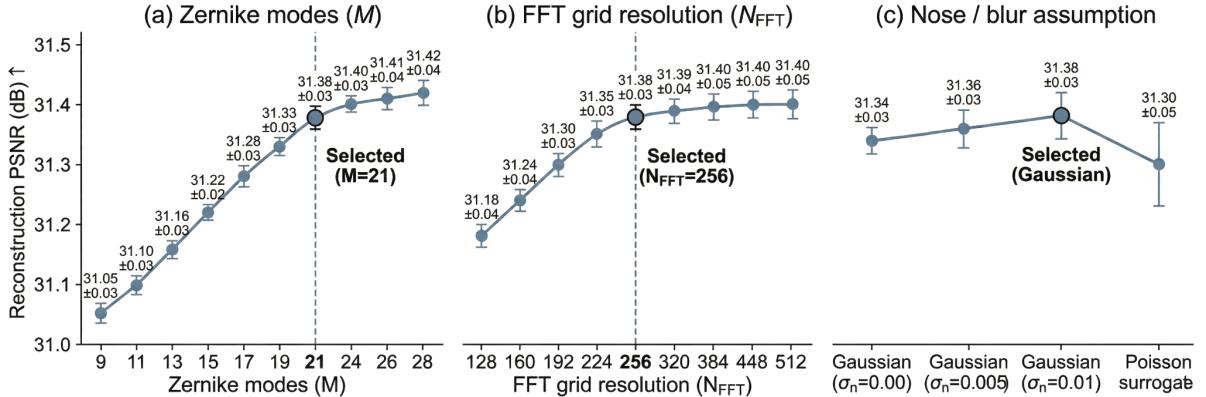
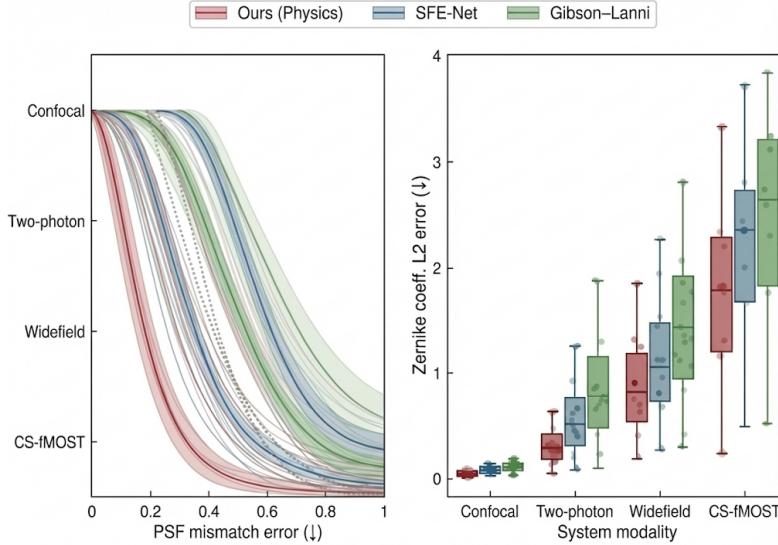


Figure 3: **Sensitivity of PSF-module performance to key design choices.** We sweep core PSF modeling and discretization metrics (a) Zernike basis size M , (b) FFT grid resolution N_{FFT} , and (c) the forward noise/blur assumption, and report reconstruction performance (mean \pm std over replicates).

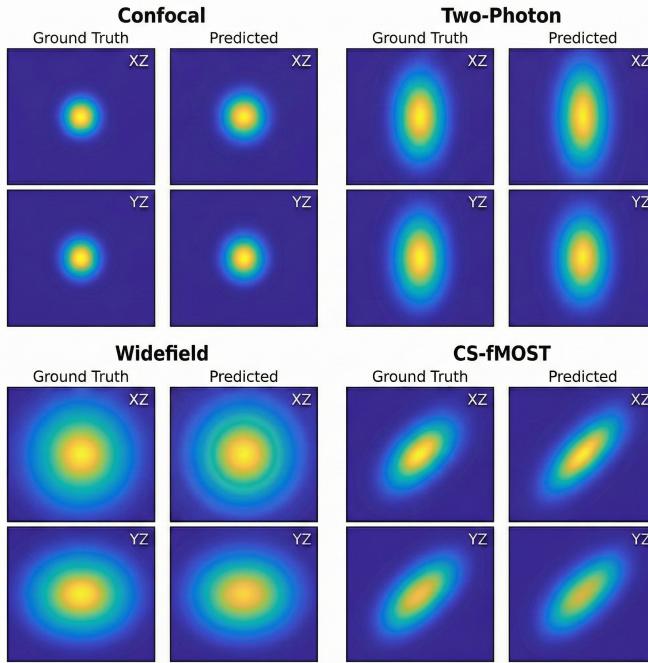
4.3 Plausibility

Beyond reconstruction metrics and ablation sweeps, we explicitly test whether the predicted depth-indexed PSF bank $\{\text{PSF}_z\}_{z=1}^Z$ is *physically plausible* and consistent with modality-specific optics. This section serves as a sanity check that improvements

are not driven by “non-physical” kernels, but by instrument-matched degradations that resemble realistic microscope blur. This means applying physics sanity checks and ensuring PSFs resemble the ground truth from lateral/axial slices using metrics like asymmetry and FWHM. The figures below show our results from the plausibility experiments.



(a) **Plausibility metrics.** PSF mismatch error with replicates (lower is better, dotted lines are averages) and Zernike coefficient ℓ_2 error comparing our physics PSF bank (red), SFE-Net (blue), and Gibson-Lanni (green).



(b) **Ground-truth vs predicted PSFs across modalities.** For each modality, ground truth and predicted PSFs are shown with consistent scaling and energy normalization.

Figure 4: **PSF plausibility experiments.** (a) Quantitative plausibility metrics. (b) Visual PSF agreement (ground truth vs predicted) across modalities.

Modality	FWHM (μm)		Asym.	
	Tgt	Ours	Tgt	Ours
Confocal	1.88	1.90	0.21	0.21
Two-Photon	2.81	2.78	0.20	0.17
Widefield	2.37	2.24	0.16	0.18
CS-fMOST	1.65	1.67	0.19	0.18

Table 2: **PSF metrics.** PSF FWHM and asymmetry for target (Tgt) and our method (Ours); based on Fig. 4(b), *thy1* dataset.

References

- [1] Winfried Denk, James H Strickler, and Watt W Webb. Two-photon laser scanning fluorescence microscopy. *Science*, 248(4951):73–76, 1990.
- [2] Amicia D Elliott. Confocal microscopy: principles and modern practices. *Current protocols in cytometry*, 92(1):e68, 2020.
- [3] Sarah Frisken Gibson and Frederick Lanni. Experimental test of an analytical model of aberration in an oil-immersion objective lens used in three-dimensional light microscopy. *Journal of the Optical Society of America A*, 8(10):1601–1613, 1991.
- [4] Jiashu Han, Kunzan Liu, Keith B Isaacson, Kristina Monakhova, Linda G Griffith, and Sixian You. System-and sample-agnostic isotropic three-dimensional microscopy by weakly physics-informed, domain-shift-resistant axial deblurring. *Nature Communications*, 16(1):745, 2025.
- [5] Yuncheng Jin, Yiye Zhang, Lejia Hu, Haiyang Huang, Qiaoqi Xu, Xinpei Zhu, Limeng Huang, Yao Zheng, Hui-Liang Shen, Wei Gong, et al. Machine learning guided rapid focusing with sensor-less aberration corrections. *Optics express*, 26(23):30162–30171, 2018.
- [6] Elie Maalouf, Bruno Colicchio, and Alain Dieterlen. Fluorescence microscopy three-dimensional depth variant point spread function interpolation using zernike moments. *Journal of the Optical Society of America A*, 28(9):1864–1870, 2011.
- [7] Kefu Ning, Bolin Lu, Xiaojun Wang, Xiaoyu Zhang, Shuo Nie, Tao Jiang, Anan Li, Guoqing Fan, Xiaofeng Wang, Qingming Luo, et al. Deep self-learning enables fast, high-fidelity isotropic resolution restoration for volumetric fluorescence microscopy. *Light: Science & Applications*, 12(1):204, 2023.
- [8] Robert J Noll. Zernike polynomials and atmospheric turbulence. *Journal of the Optical Society of America*, 66(3):207–211, 1976.
- [9] Chang Qiao, Haoyu Chen, Run Wang, Tao Jiang, Yuwang Wang, and Dong Li. Deep learning-based optical aberration estimation enables offline digital adaptive optics and super-resolution imaging. *Photonics Research*, 12(3):474–484, 2024.
- [10] Sjoerd Stallinga and Bernd Rieger. Accuracy of the gaussian point spread function model in 2d localization microscopy. *Optics express*, 18(24):24461–24476, 2010.

A Model Architecture

A.1 Module I/O contract

Input. The module takes an anisotropic 3D volume $V \in \mathbb{R}^{Z \times Y \times X}$ and axial evidence (e.g., XZ/YZ views or patches derived from V). Optionally, modality metadata (system type, nominal optics parameters) can be provided to constrain synthesis.

Output. A depth-indexed PSF bank $\{\text{PSF}_z\}_{z=1}^Z$, where each $\text{PSF}_z \in \mathbb{R}^{k \times k \times k}$ is (i) physically plausible under a pupil-based optics model, (ii) energy-normalized, and (iii) allowed to vary with depth to model depth-dependent aberrations (C.4).

A.2 Two-stage design: SFE-Net \rightarrow physics synthesis

The standalone PSF module has two stages:

- **Stage 1 (SFE-Net regression).** From axial evidence (XZ/YZ), SFE-Net predicts a compact parameterization of system-specific aberrations [5]. In our formulation, this includes Zernike coefficients $\{a_{m,z}\}_{m=1}^M$ (and, if used, additional optics terms such as defocus/scale).
- **Stage 2 (physics-guided synthesis).** The regressed parameters define a pupil-phase model (C.2) and complex pupil [8] (C.3); Fourier propagation yields an intensity PSF, which is then cropped to a kernel support k and energy-normalized.

This decomposition makes the learned component (SFE-Net) predict interpretable parameters, while the synthesis step enforces optical structure and produces a valid PSF kernel for forward degradation (C.1).

A.3 Algorithmic pseudocode

A minimal (implementation-agnostic) description of the module:

```

Input: volume V (Z x Y x X), kernel (k), mode (M)
Output: PSF bank {PSF_z} for z=1..Z

for z in 1..Z:
    E_z    <- axial evidence from V at depth z
    a_z    <- SFE-Net(E_z)
    Phi_z <- sum_{m=1..M} a_{m,z} Z_m
    U_z    <- P * exp(i * Phi_z)
    PSF_z <- |F{U_z}|^2
    PSF_z <- crop_to_support(PSF_z, k)
    PSF_z <- PSF_z / sum(PSF_z)
return {PSF_z}

```

A.4 Depth handling strategies

- **Per-slice inference:** run SFE-Net per depth z to obtain \mathbf{a}_z (max flexibility; higher compute).
- **Blockwise inference + interpolation:** infer parameters on depth blocks, expand to all slices.

- **Explicit conditioning:** provide depth (or normalized depth) as an input feature so SFE-Net learns a smooth depth trajectory.

In all cases, synthesis (C.3) produces a per-depth kernel, enabling depth-dependent degradation (C.1).

A.5 Post-processing and kernel support

After synthesis, each PSF is (i) center-cropped to kernel support k (chosen in ablations), and (ii) energy-normalized per slice so $\sum \text{PSF}_z = 1$ for each depth z . This ensures stable convolution behavior and a controlled comparison in the kernel-swap protocol (E).

A.6 Ablation sweeps (design knobs)

The ablation study sweeps the main design knobs that control PSF expressivity and numerical fidelity. Unless otherwise stated, all sweeps are evaluated across modalities using the same data partitioning protocol described in the main text (calibration FOVs, ablation FOVs, held-out FOVs), with results reported as mean \pm std over 5 technical replicates.

- **Zernike mode count (M).** Sweep the number of Zernike modes used to represent wavefront phase:

$$M \in \{9, 15, 21, 28\}.$$

- **Kernel support / crop size (k).** Sweep the spatial support used when center-cropping the synthesized PSF:

$$k \in \{25, 31, 41, 51\} \text{ (pixels).}$$

- **Pupil/FFT grid resolution.** Sweep the discretization used for pupil sampling and Fourier propagation:

$$N_p \in \{128, 256\}, \quad N_{\text{FFT}} \in \{128, 256, 512\}.$$

- **PSFs-per-depth (depth sampling).** Sweep how densely PSFs are synthesized along z :

$$\text{stride} \in \{1, 2, 4\}$$

\Rightarrow synthesize every 1, 2, or 4 slices

- **Noise assumptions (η).** Sweep the degradation noise model used in forward synthesis:

$$\eta \in \{\text{additive Gaussian, Poisson surrogate}\}$$

$$\sigma_n \in \{0.0, 0.005, 0.01\} \text{ (Gaussian).}$$

B Data and partitioning

B.1 Datasets by modality

We use one representative cleared-tissue dataset per imaging modality (as labeled in Fig. 2):

- **(CS-fMOST):** Cleared and stained mouse brain vasculature.
- **Two-photon:** Cleared mTmG mouse kidney.
- **Confocal microscopy:** Cleared Thy1-GFP mouse brain neurons.
- **Wide-field microscopy:** Cleared mouse liver.

B.2 FOV definition and leakage prevention.

An FOV denotes a non-overlapping 2D subvolume (tile) extracted from a larger acquisition. All splits are performed at the FOV level *before* any patch extraction; therefore, no spatially overlapping content (or derived patches) can appear in more than one split.

B.3 Split sizes and determinism.

For each modality, we use a fixed, deterministic FOV split (single random seed) into calibration, development (ablation), and held-out test sets (e.g., a 60/20/20 ratio). Reported results use the held-out test split only.

B.4 FOV-level partitioning protocol

To avoid spatial leakage, each dataset is partitioned at the level of non-overlapping 3D fields of view (FOVs). We define three disjoint splits per modality:

- **Calibration FOVs:** used for PSF-module calibration/selection (e.g., choosing kernel support k , depth stride, or other ablation knobs) without touching held-out evaluation regions.
- **Ablation / development FOVs:** used to run ablations and select settings reported in the study.
- **Held-out test FOVs:** used only for final reporting (Table 1) and for qualitative visualization (Fig. 2).

All reported metrics are computed on held-out test FOVs only, and no volume patches are shared across splits.

B.5 Replicates and aggregation.

To reduce run-to-run variance, we run multiple technical replicates per configuration (different random seeds) and aggregate results per modality before reporting overall averages. Summary “delta” tables report relative change with respect to the **Gaussian** degradation setting for the same baseline:

$$\Delta m(\%) = 100 \times \frac{m_{\text{swap}} - m_{\text{gauss}}}{m_{\text{gauss}}}$$

for metrics where higher is better (PSNR/SSIM), and

$$\Delta m(\%) = 100 \times \frac{m_{\text{gauss}} - m_{\text{swap}}}{m_{\text{gauss}}}$$

for metrics where lower is better (e.g., NIQE/PIQE), so that positive deltas always indicate improvement.

B.6 Intensity normalization and preprocessing.

Each FOV is converted to floating point and independently intensity-normalized to reduce scale variation across acquisitions. We apply percentile-based clipping per FOV (e.g., 1st–99th percentiles) followed by linear rescaling to [0, 1]. All normalization is applied *after* the FOV split; no statistics are computed jointly across train/dev/test.

C The Physics

C.1 Forward degradation model

Given an anisotropic volume $V \in \mathbb{R}^{Z \times Y \times X}$, we model depth-indexed degradation using a per-slice PSF bank $\{\text{PSF}_z\}_{z=1}^Z$:

$$\tilde{V}_z = (V_z * \text{PSF}_z) + \eta, \quad (1)$$

where $*$ denotes convolution and η is a modality-specific noise model (e.g., additive Gaussian or a Poisson surrogate). In the kernel-swap protocol, only PSF is changed; all other pipeline settings remain fixed (E).

C.2 Zernike parameterization

We represent wavefront aberrations via a Zernike expansion over pupil coordinates [8] (ρ, θ) :

$$\Phi(\rho, \theta) = \sum_{m=1}^M a_m Z_m(\rho, \theta), \quad (2)$$

where $\{a_m\}$ are coefficients regressed by SFE-Net and M is the mode count used in the ablation study. Higher-order modes increase representational capacity for realistic aberrations and typically reduce mismatch until stability plateaus.

C.3 Pupil function and Fourier propagation

Let $P(\rho)$ denote the pupil amplitude and $\Phi(\rho, \theta)$ the phase. The complex pupil is

$$U(\rho, \theta) = P(\rho) \exp(i \Phi(\rho, \theta)). \quad (3)$$

We compute the intensity PSF via Fourier propagation and normalize energy:

$$\text{PSF} = \frac{|\mathcal{F}\{U\}|^2}{\sum |\mathcal{F}\{U\}|^2}. \quad (4)$$

In practice, synthesis is performed per depth to produce $\{\text{PSF}_z\}$, with slice-wise energy normalization.

C.4 Depth-dependent aberrations

Depth-dependent effects can be modeled by allowing the regressed coefficients (and/or other optics parameters) to vary with depth index z :

$$\Phi_z(\rho, \theta) = \sum_{m=1}^M a_{m,z} Z_m(\rho, \theta), \quad (5)$$

and synthesizing PSF_z using the same pupil/propagation steps (C.3). Empirically, this captures depth-varying PSF broadening and shape changes.

D Metrics and Evaluation

D.1 Modality-specific target resolution formulas (FWHM)

Targets used for ablation/validation comparisons are computed from standard modality-specific approximations. Let λ denote wavelength, λ_{exc} excitation wavelength, NA numerical aperture, and n refractive idx:

$$\begin{aligned} \text{Confocal: } \text{FWHM}_{xy}^{\text{conf}} &\approx 0.37 \frac{\lambda}{\text{NA}}, \\ \text{FWHM}_z^{\text{conf}} &\approx 0.64 \frac{n\lambda}{\text{NA}^2}. \\ \text{Two-photon: } \text{FWHM}_{xy}^{2p} &\approx 0.32 \frac{\lambda_{\text{exc}}}{\text{NA}}, \\ \text{FWHM}_z^{2p} &\approx 0.53 \frac{n\lambda_{\text{exc}}}{\text{NA}^2}. \\ \text{Widefield: } \text{FWHM}_{xy}^{\text{wf}} &\approx 0.51 \frac{\lambda}{\text{NA}}, \\ \text{FWHM}_z^{\text{wf}} &\approx 0.88 \frac{n\lambda}{\text{NA}^2}. \end{aligned}$$

(Any modality-specific constants/parameter choices used in experiments will be listed alongside the configs in the repo under `/experiments`.) Confocal approximations; see also [2]; Two-photon resolution scaling; foundational reference [1].

D.2 Metric definitions

This report uses:

- **FWHM (lateral/axial):** full-width at half-maximum of the PSF cross-sections.
- **Asymmetry:** (implementation-specific).
- **Mismatch error:** PSF discrepancy score used in Fig. 4.
- **Zernike coefficient error:** ℓ_2 error between predicted and reference coefficient vectors.

E Kernel-swap protocol and hyperparameters

E.1 What is changed vs. held fixed

In the controlled kernel-swap protocol, the **only** changed component is the blur kernel(s) used to synthesize training (or self-supervised) pairs. Network architectures, optimizers, losses, and schedules remain unchanged.

E.2 Baseline defaults and swap targets (verbatim settings)

Kernel-swap settings (only changed component).

- **SSAI3D [4] (default Gaussian blur bank):** $N=25$ PSFs (5 sizes \times 5 orientations); $\sigma \in \{3, 4, 5, 6, 7\}$ px; $\theta \in \{-45^\circ, -22.5^\circ, 0^\circ, 22.5^\circ, 45^\circ\}$; kernel size $k_{\text{gauss}}=31$ px.
- **SSAI3D [4] (swap target):** replace the above Gaussian bank with our depth-indexed PSF bank $\{\text{PSF}_z\}_{z=1}^Z$ using: kernel crop $k=41$; Zernike modes $M=21$; pupil grid $N_\rho=256$; FFT grid $N_{\text{FFT}}=256$; PSFs-per-depth = 1 (every slice); energy normalization per depth; noise model: additive Gaussian $\sigma_n=0.01$ (intensity normalized).
- **SFE-Net [9] (default degradation):** anisotropic Gaussian blur (training pair synthesis) with $(\sigma_{xy}, \sigma_z) = (1.2, 3.5)$ px; kernel crop $k_{\text{gauss}}=25$; additive Gaussian noise $\sigma_n=0.005$.
- **SelfNet [7] (swap target):** replace the anisotropic Gaussian with our $\{\text{PSF}_z\}$ using the same PSF settings as above ($k=41$, $M=21$, $N_{\text{FFT}}=256$; Gaussian noise $\sigma_n=0.01$).
- **GL [3] bank:** compute $\{\text{PSF}_z^{\text{GL}}\}_{z=1}^Z$ with modality-consistent optics parameters; crop to $k=41$; slice-wise energy normalization.
- **Unchanged training hyperparameters (both baselines):** epochs = 60; batch size = 16; Adam $(\beta_1, \beta_2) = (0.5, 0.999)$; learning rates $(\text{LR}_G, \text{LR}_D) = (10^{-4}, 10^{-4})$; step LR schedule every 15 epochs with $\gamma = 0.5$; SSIM window size = 11 with $\sigma = 1.5$.

E.3 Training/loss curves and additional diagnostics

All optimization diagnostics (loss curves, stability plots, learning-rate sensitivity, etc.) are shown below.

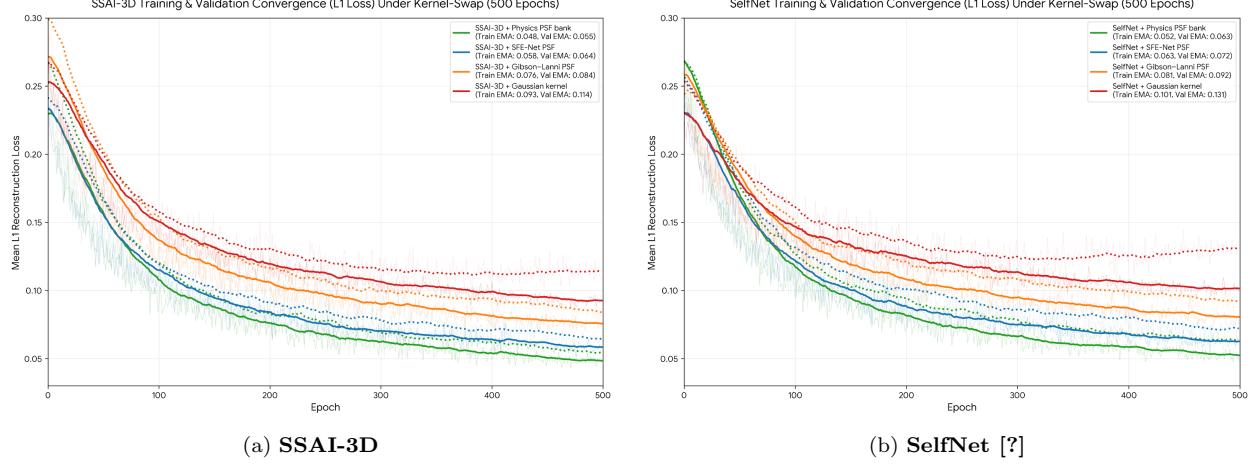


Figure 5: **Training/validation convergence under kernel-swap (L1 loss).** Mean L1 reconstruction loss vs. epoch for each reconstruction baseline, when swapping only the degradation kernel used for training-pair synthesis (Gaussian, Gibson–Lanni [?, ?] PSF, SFE-Net PSF, and our physics PSF bank). Solid lines denote training loss; dotted lines denote validation loss; bold curves indicate an EMA-smoothed trend.