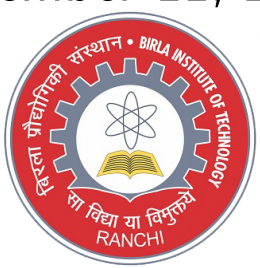
Enhancing Fluorescence Microscopy Resolution Beyond the Diffraction Limit via Cascaded Up- and Down-Sampling Networks

Presented By:
Hazique Aetesam

Department of CSE, Birla Institute of Technology Mesra, Patna Campus

10^{th} **International Conference on Computer Vision and Image Processing (CVIP-2025), IIT Ropar**

December 11, 2025



Outline

Introduction

A Brief Description

- A typical microscopy is suited to the visualization of structures at the cellular and tissue level.
- Fluorescence microscopy is a more recent development to observe the biological samples beyond the cellular level.
- It helps in the visualization of sub-cellular structures in the nanoscale range.
- Fluorescence is the ability of a substance to emit light of a particular wavelength when it is subjected to a light source.
- These substances, also known as fluorophores are genetically modified proteins which emit light on excitation.

General Working Principle

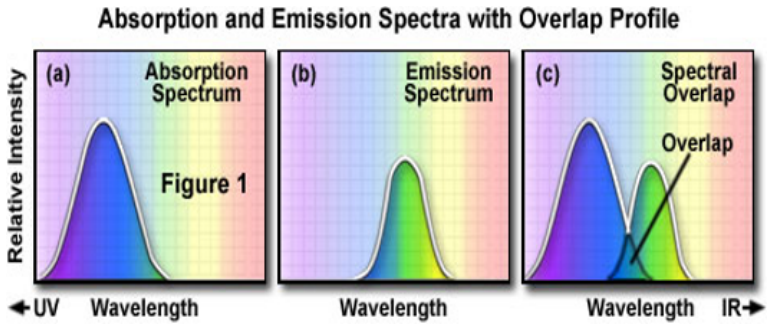


Figure 1: Absorption and Emission Spectra[1]

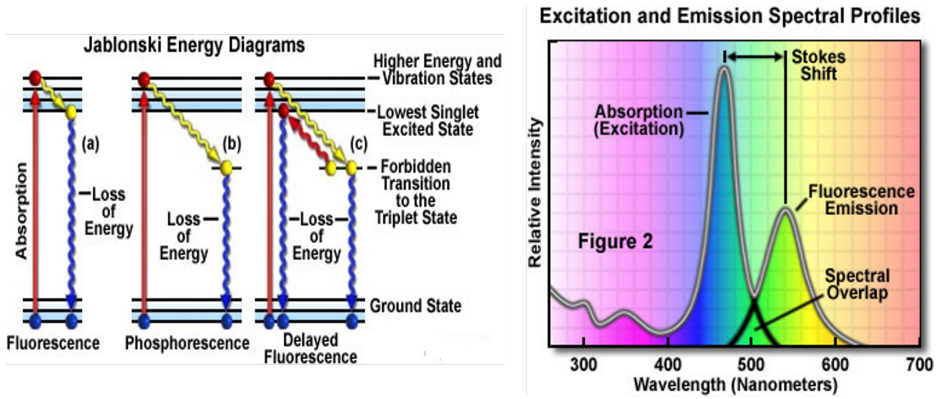


Figure 2: Jablonski Diagram (left) and Stokes Shift (right)[2]

[1] **Photo courtesy:** https://evidentscientific.com/en/microscope-resource/knowledge-hub/fluoroexcitation

[2] **Photo courtesy:** https://evidentscientific.com/en/microscope-resource/tutorials/jablonski

Microscopy Types

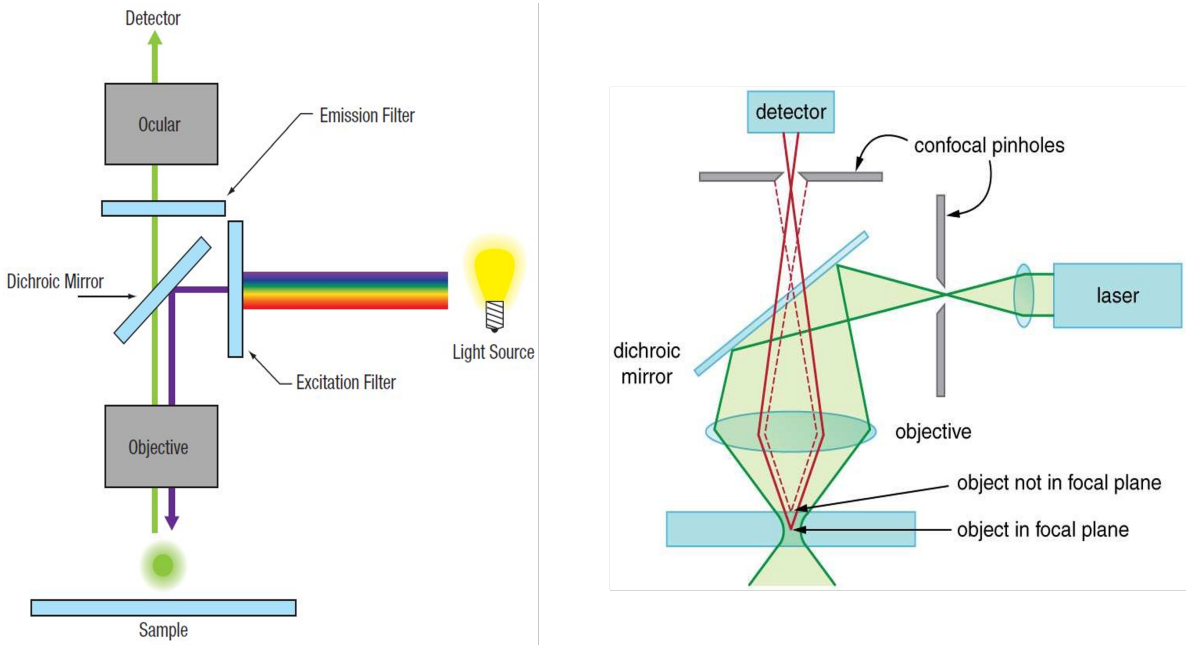
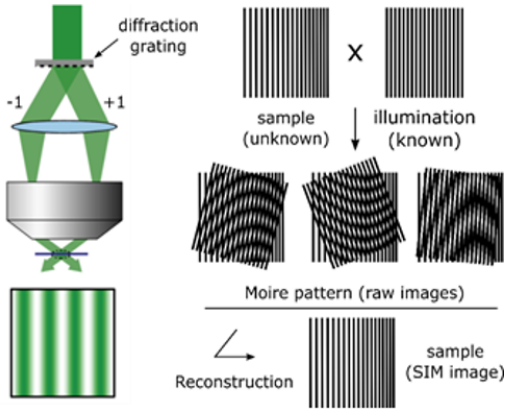


Figure 3: Widefield (left) and confocal microscopy (right)[3].

[1] **Photo courtesy:** https://www.ptglab.com/news/blog/if-imaging-widefield-versus-confocal-microscopy/

Structured Illumination Microscopy



- Super-resolution structured illumination microscopy (SR-SIM) is a powerful imaging technique that surpasses the diffraction limit of conventional light microscopy.
- SR-SIM uses patterned illumination (structured light) to excite the sample. These patterns are typically sinusoidal and can be shifted and rotated to create multiple images.
- The structured light interacts with the sample, producing moir´e fringes, which are lower frequency interference patterns that contain high-frequency information about the sample.

Limitations

- **Blur**, modelled as **point spread function (PSF)** introduces limit to the lateral and axial resolution achievable.
- Depending upon the phenomenon generating noise, noise in fluorescence microscopy can be of two different types: **read noise** and **photon noise**.
- Artefacts in the form of **photobleaching** and **phototoxicity**.

Limitations of SR Microscopy

- Require sophisticated equipment and computational algorithms for image acquisition and reconstruction.
- The need for many activation cycles can result in long acquisition time; hence limiting temporal resolution.

Related Works

Table 1: Conventional SR methods under fluorescence microscopy

Article Title	**Key method/ innovation**
DAOSTORM: an algorithm for high-density super-resolution microscopy	A multi-PSF model enabling resolution at densities of 10 molecules μm^2
Faster storm using compressed sensing	A compressed sensing technique to achieve temporal resolutions of 3 seconds per image in single-molecule localization microscopy (SMLM)
Parameter-free image resolution estimation based on decorrelation analysis	A spatial resolution estimation method using correlation analysis,
Efficient super-resolution volumetric imaging by radial fluctuation bayesian analysis light-sheet microscopy.	Super-resolution radial fluctuation (SRRF) proposed to optimize Bayesian analysis of blinking and bleaching events to achieve precise optical sectioning.

Limitations

- Demands precise a priori knowledge of acquisition parameters, including point spread function (PSF) characteristics and photon-sample interaction dynamics.
- Inaccuracies in parameter estimation degrade reconstruction quality.

Related Works *(cont...)*

Table 2: Learning-based SR methods under fluorescence microscopy.

Article Title	**Key method/ innovation**
Content-aware image restoration: pushing the limits of fluorescence microscopy (2018)	A U-Net-based framework applicable to noise reduction, isotropic resolution enhancement, and super-resolution
3D residual channel attention networks denoise and sharpen fluorescence microscopy image volumes (2021)	Achieved 2.5× lateral resolution enhancement in time-lapse volumes using residual channel attention network (RCAN) with STED ground truths.
Single-frame deep-learning super- resolution microscopy for intracellular dynamics imaging (2023)	Attained \$30˜nm/10\$ ms spatiotemporal resolution using sub-pixel edge priors and multi-component optimization, eliminating multi-frame requirements.
Zero-shot learning enablesinstant denoising and super-resolution in optical fluorescence microscopy (2024)	Unsupervised deconvolution networks (ZS-DeconvNet) that enhance the resolution by more than 1.5-fold

Limitations

- Dependence on large, high-quality training data.
- Simulation-to-real domain gap.
- Interpretability and parameter transparency.

Motivation

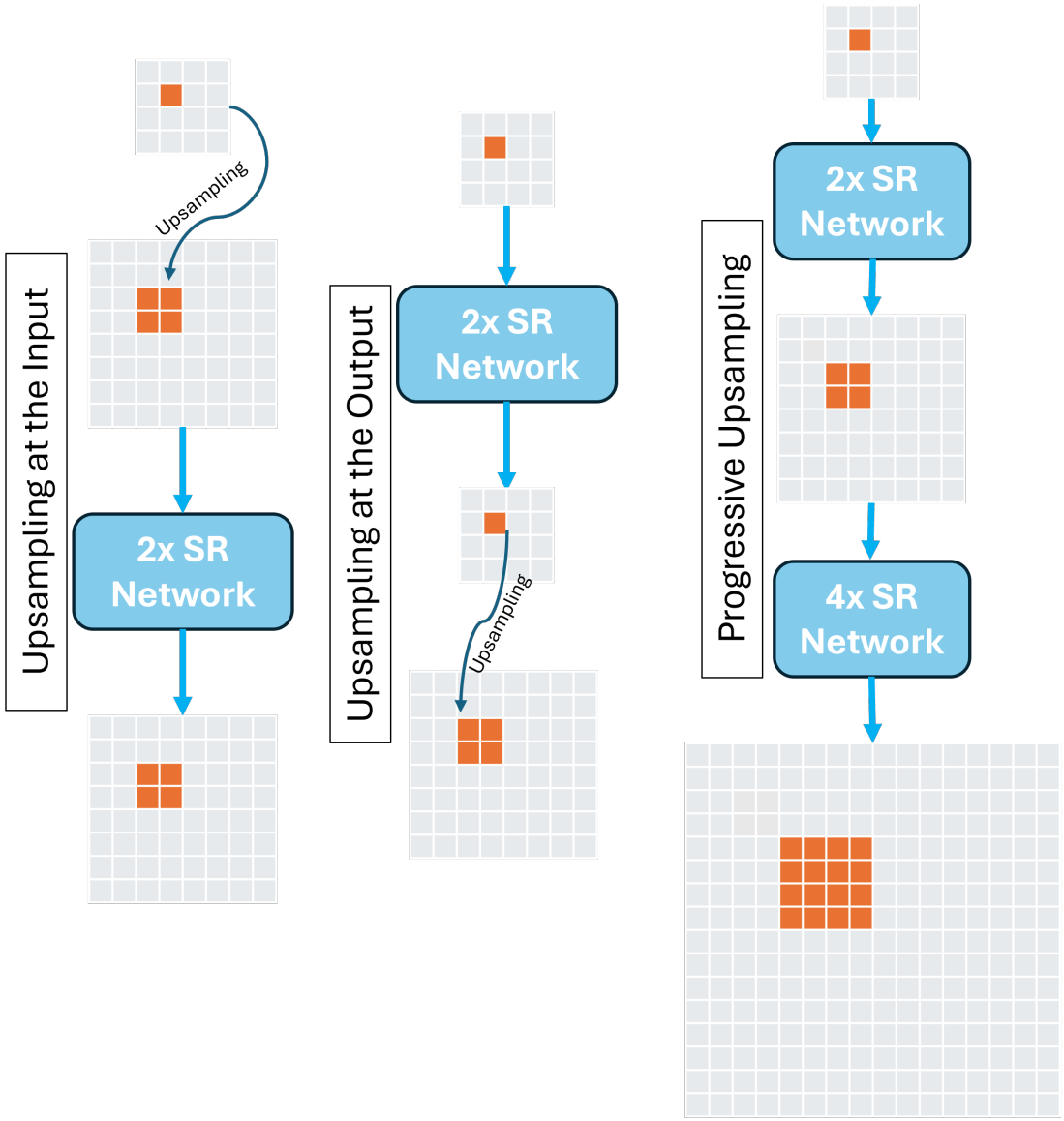


Figure 5: Different Upsampling strategies in a typical SR network.

Aims and Objectives

- Our model iteratively refines reconstruction errors through successive upsampling and downsampling layers, enhancing feature refinement.
- Multi-scale external attention integration from different depths selectively amplify salient features while suppressing noise, improving contextual awareness during reconstruction.
- The model super-resolves blurred, noisy widefield images to achieve spatial resolution comparable to structured illumination microscopy (SIM) ground truth.
- The proposed model is trained and tested on a fluorescence microscopy dataset spanning four structurally complex biological targets: clathrin-coated pits (CCP), endoplasmic reticulum (ER), microtubules (MT) and F-actin fibers.

The Proposed Methodology

Up-Projection Unit

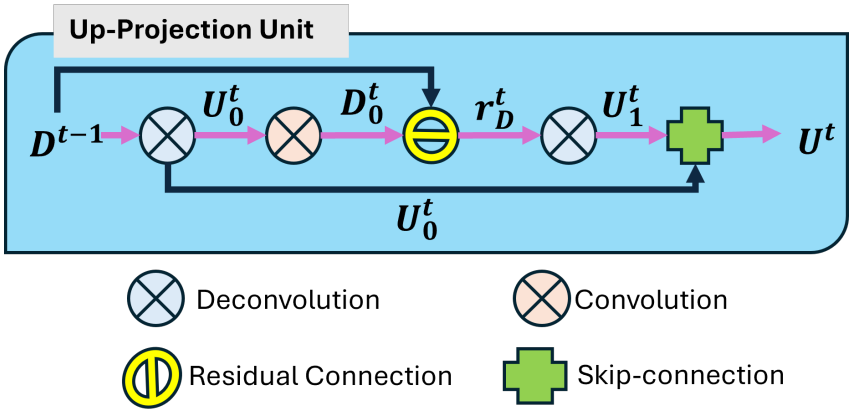


Figure 6: Up-Projection Units in the main SR network.

- **Intermediate Scale Up:** $U_0^t = (D^{t-1} \otimes_{\uparrow} u_0^t) \uparrow_s$
- **Intermediate Scale Down:** $D_0^t = (U_0^t \otimes_{\downarrow} d_0^t) \downarrow_s$
- **Residual:** $r_t^D = D_0^t - D^{t-1}$
- **Residual Scale Up:** $U_1^t = (r_t^D \otimes_{\uparrow} u_1^t) \uparrow_s$
- **Up-Projection Output:** $U^t = U_0^t + U_1^t$

Down-Projection Unit

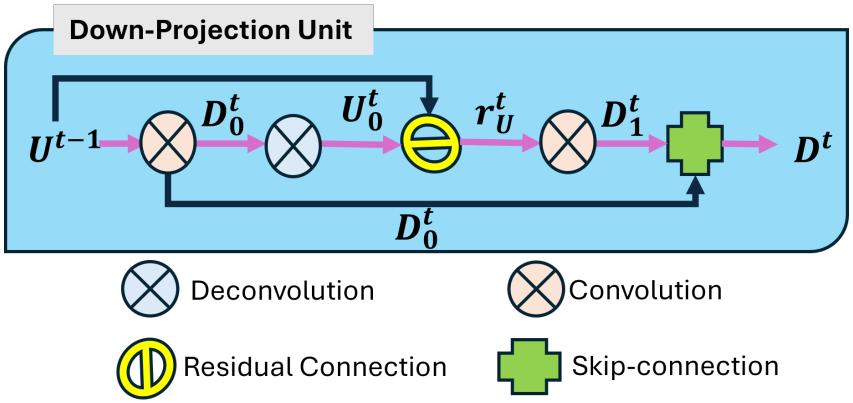


Figure 7: Down-Projection Units in the main SR network.

- **Intermediate Scale Down:** $D_0^t = (U^{t-1} \otimes_{\downarrow} d_0^t) \downarrow_s$
- **Intermediate Scale Up:** $U_0^t = (D_0^t \otimes_{\uparrow} u_0^t) \uparrow_s$
- **Residual:** $r_t^U = U_0^t - U^{t-1}$
- **Residual Scale Down:** $D_1^t = (r_t^U \otimes_{\downarrow} d_1^t) \downarrow_s$
- **Down-Projection Output:** $D^t = D_0^t + D_1^t$

External Attention Multiple Layer Perceptron

External Attention

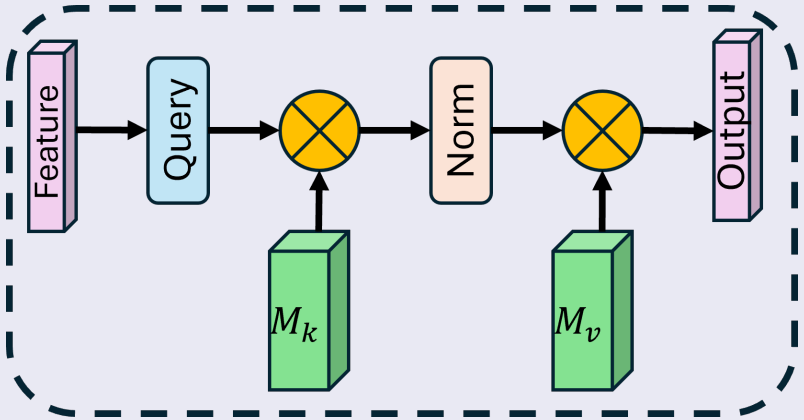


Figure 8: External Attention Module

Self Attention

$$A = (\alpha)_{i,j} = \text{softmax}(QK^T) \tag{1a}$$

$$F_{\text{out}} = AV \tag{1b}$$

where $A \in \mathbb{R}^{N \times N}$ is the attention map and $F_{\text{out}} \in \mathbb{R}^{R \times d}$ is the output feature space.

External Attention

$$A = \text{Norm}(FM_k^T) \tag{2a}$$

$$F_{\text{out}} = AM_v \tag{2b}$$

where $\{M_k,\ M_v\} \in \mathbb{R}^{s \times d}$ and $\text{Norm}(\cdot)$ is instance normalization.

Network Architecture

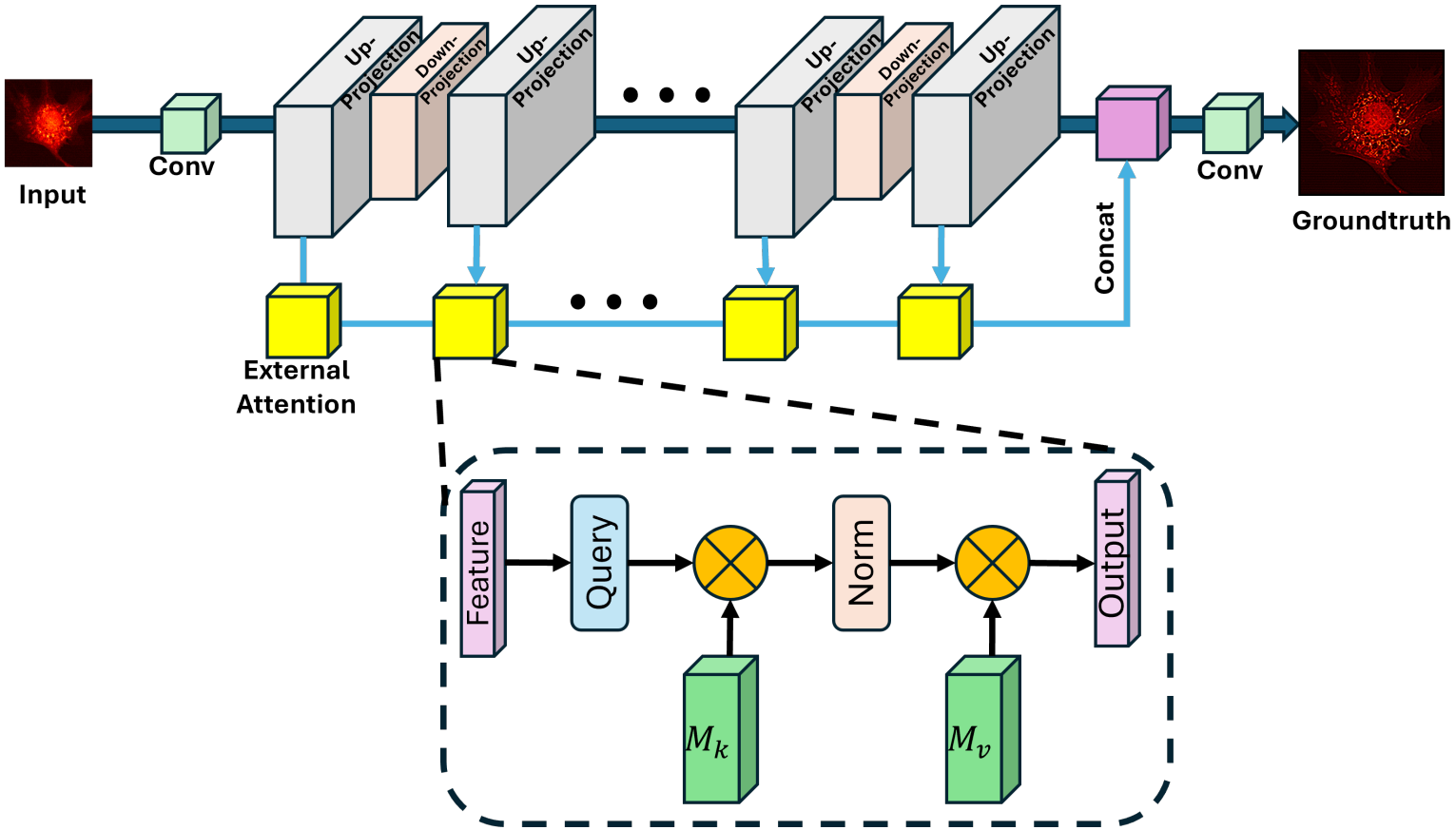


Figure 9: The complete model architecture with up-projection, down-projection and pluggable external attention blocks.

Experimental Results

Hyperparamters and Training Options

- Convolution and deconvolution achieved using standard and transposed 2D convolution, respectively.
- Kernel size: 3×3, stride=up/downsampling factor, number of filters $= 64$.
- The linear layers M_k and M_v within the attention blocks are initialized with random values.
- Number of epochs $= 100$ and ADAM as the optimizer.
- Criterion Function: Frequency distribution Loss [4]

[1] Ni, Zhangkai, et al. "Misalignment-robust frequency distribution loss for image transformation." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024.*

Dataset Description

- All experiments utilized the publicly available BioSR dataset [1], comprising well-registered LR-HR image pairs.
- Image pairs acquired through multi-modality structured illumination microscopy (SIM) in two configurations: grazing incidence (GI) and total internal reflectance fluorescence (TIRF).
- The dataset features nine distinct signal-to-noise ratio (SNR) levels across four biological specimens of increasing structural complexity.
- Four types of samples are captured: clathrin-coated pits (CCPs), endoplasmic reticulum (ER), microtubules (MTs), and linear/non-linear F-actin fibers.
- For each specimen type, 50 fields of view (FOVs) were captured at escalating excitation intensities to simulate SNR variation.

[1] **Link to the dataset:** https://figshare.com/articles/dataset/BioSR/

Dataset Description *(cont...)*

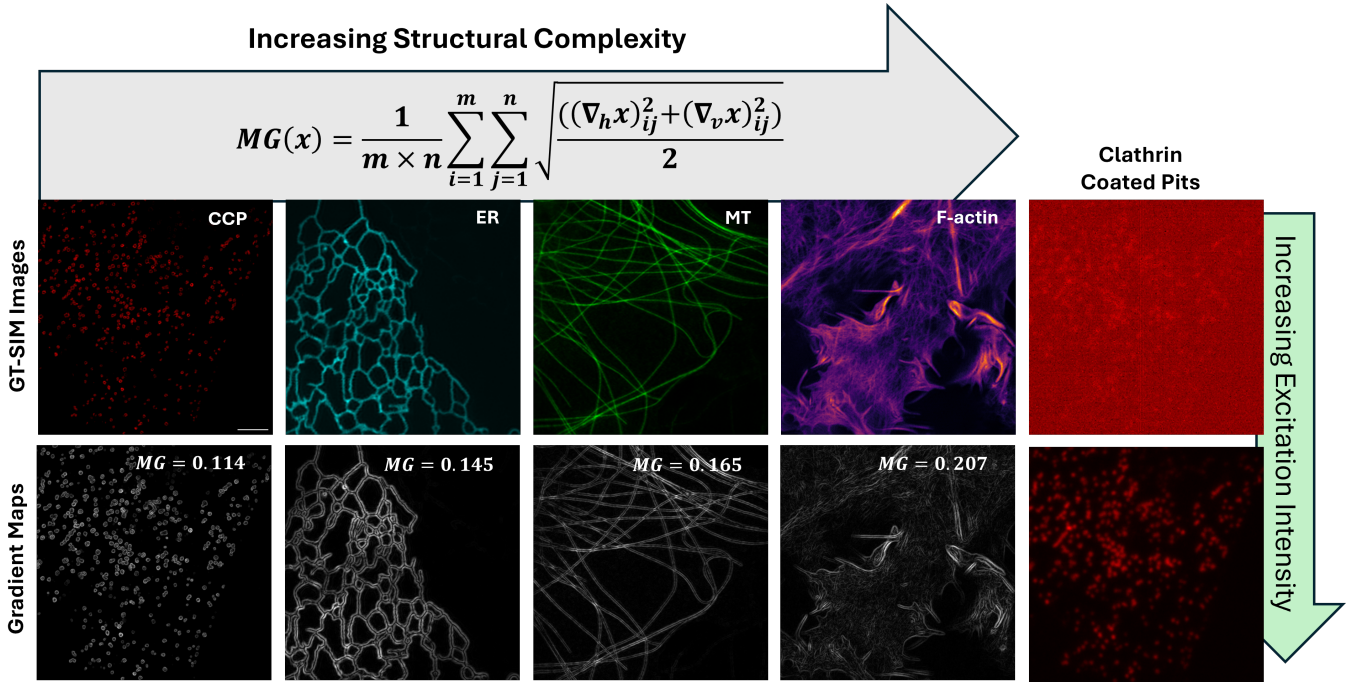


Figure 10: **Row 1:** Different samples obtained from structured illumination miscroscopy (SIM) (acting as groundtruth data) in increasing levels of structural complexity: clathrin-coated pits (CCP), endoplastic reticulum (ER), microtubules (MT) and F-actin fibers. **Row 2:** The corresponding gradient maps where structural complexity is measured in terms of grayscale mean gradient (MG) where $\nabla_h x$ and $\nabla_v x$ are the gradients across horizontal and vertical axes respectively. The representative widefield CCP images for two different excitation intensities serving as input samples during model training (last column).

Baseline Models

Baseline Models

- Content-aware image restoration (CARE)[1]
- Three-dimensional residual channel attention networks for denoising and sharpening (RCAN)[2]
- Deep neural network-based super-resolution evaluation and development (DFCAN)[3]
- Single-frame deep-learning super-resolution microscopy for intracellular dynamics imaging (SFSRM)[4]

---

[1] Weigert, Martin, et al. "Content-aware image restoration: pushing the limits of fluorescence microscopy." *Nature methods* 15.12 (2018): 1090-1097.

[2] Chen, Jiji, et al. "Three-dimensional residual channel attention networks denoise and sharpen fluorescence microscopy image volumes." *Nature methods* 18.6 (2021): 678-687.

[3] Qiao, Chang, et al. "Evaluation and development of deep neural networks for image super-resolution in optical microscopy." *Nature methods* 18.2 (2021): 194-202.

[4] Chen, Rong, et al. "Single-frame deep-learning super-resolution microscopy for intracellular dynamics imaging." *Nature Communications* 14.1 (2023): 2854.

Assessment Metrics and Uncertainty Quantification

- **Superresolution Quantitative Image Rating and Reporting of Error Locations (SQUIRREL) metric**[1]:
  - Resolution scaling function (RSF) →Resolution scaled error map (RSM)
  - Two summary metrics are then derived: the resolution-scaled error (RSE), which is sensitive to contrast and brightness variations, and the resolution-scaled Pearson's correlation coefficient (RSP), ranging from $[-1, 1]$, enabling robust cross-modality quality assessment.
- **The rolling Fourier-Ring Correlation (rFRC)**[2] to achieve spatially heterogeneous reconstruction quality. By analyzing local frequency-domain correlations, rFRC quantifies reconstruction uncertainty at the super-resolution scale.

[1] Culley, Siân, et al. "Quantitative mapping and minimization of super-resolution optical imaging artifacts." *Nature methods* 15.4 (2018): 263-266.

[2] Zhao, Weisong, et al. "Quantitatively mapping local quality of super-resolution microscopy by rolling Fourier ring correlation." *Light: Science & Applications* 12.1 (2023): 298.

Visual Results

Clathrin Coated Pits (CCPs)

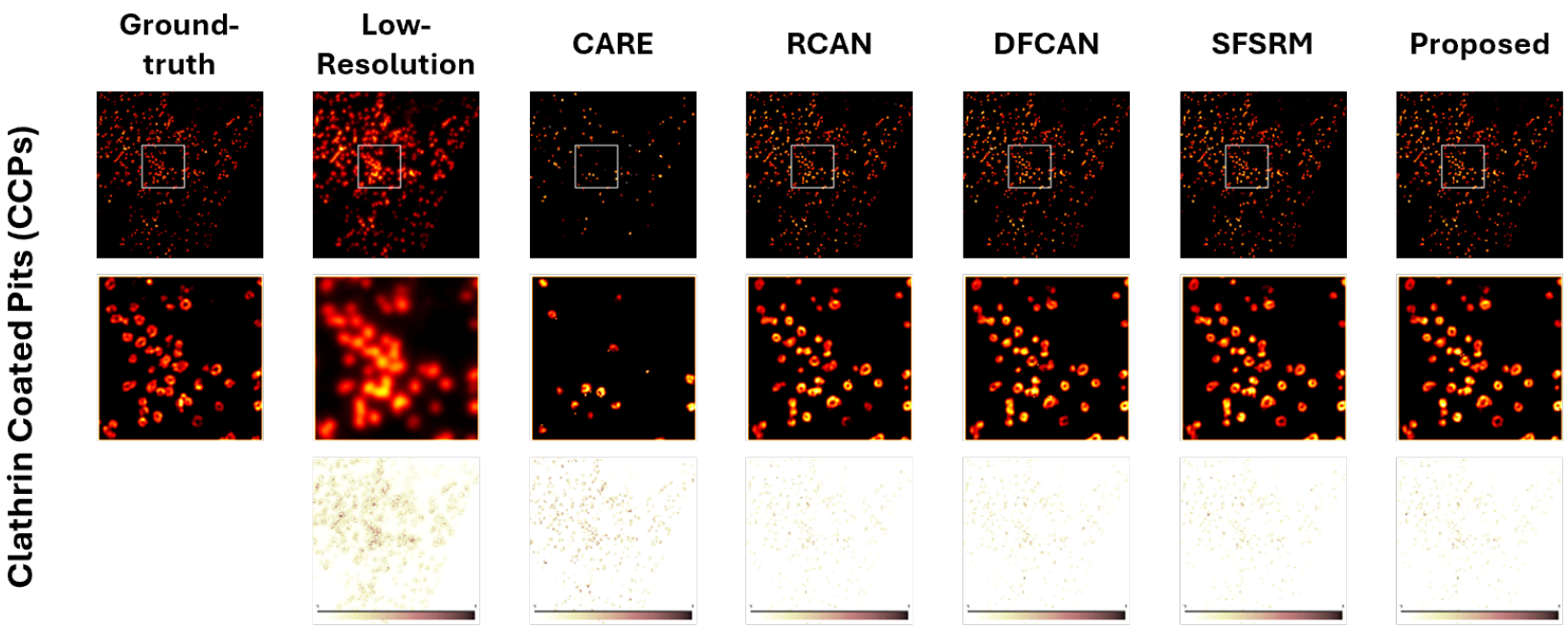


Figure 11: **Row 2:** Zoomed Regions of row 1 and **Row 3:** Mean absolute error between the reconstructed and the groundtruth images.

Endoplasmic Reticulum (ER)

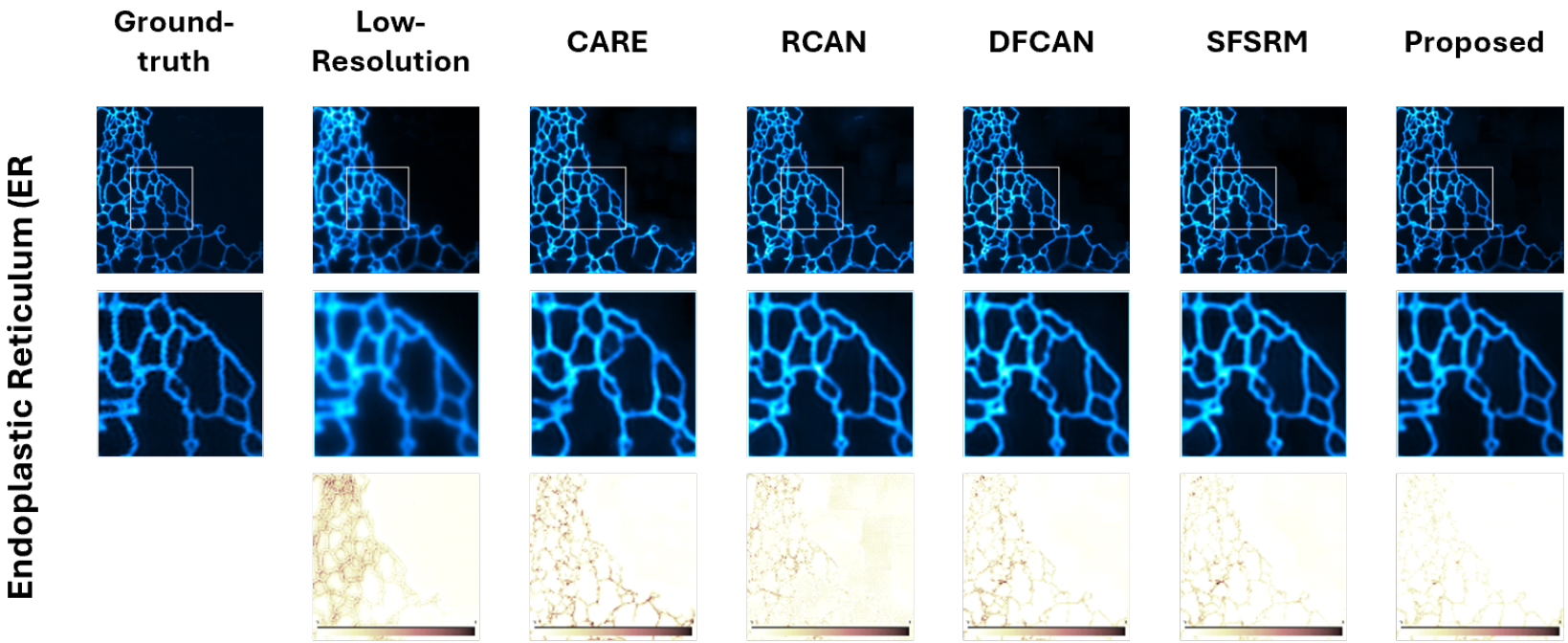


Figure 12: **Row 2:** Zoomed Regions of row 1 and **Row 3:** Mean absolute error between the reconstructed and the groundtruth images.

Microtublues (MT)

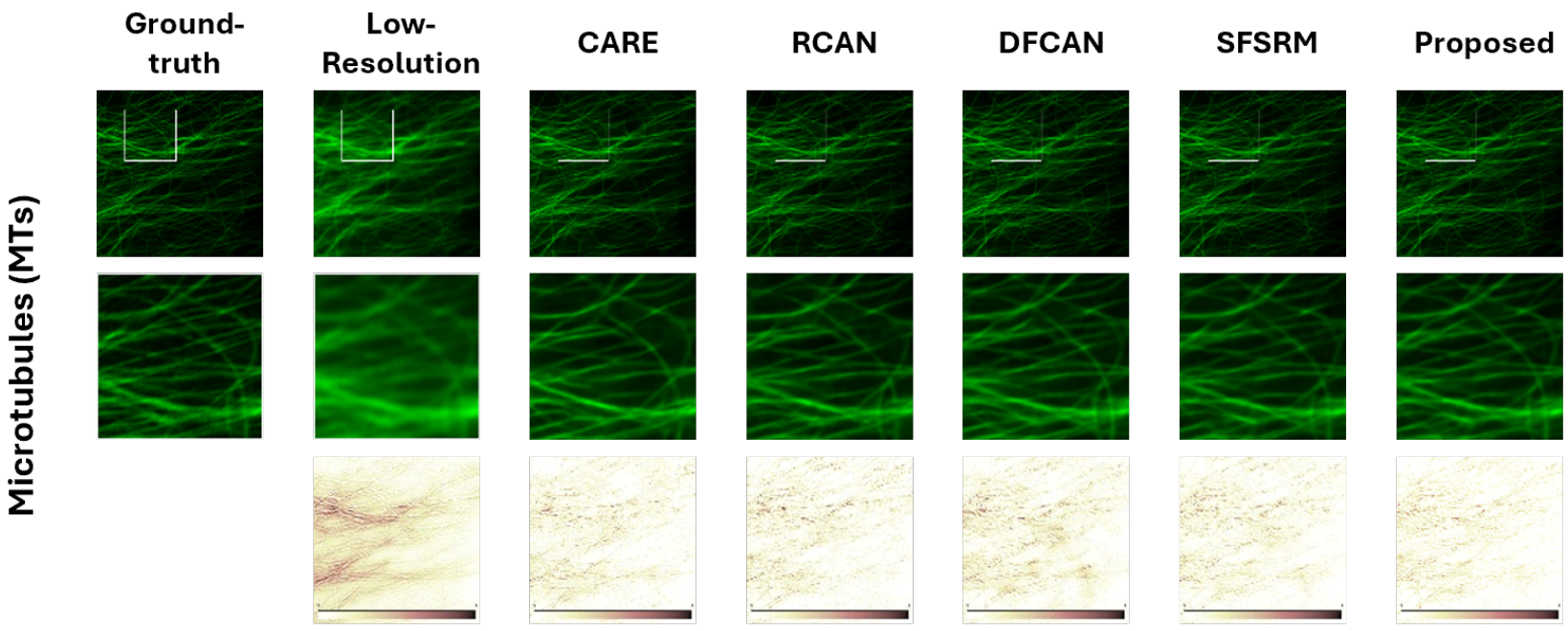


Figure 13: **Row 2:** Zoomed Regions of row 1 and **Row 3:** Mean absolute error between the reconstructed and the groundtruth images.

F-actin

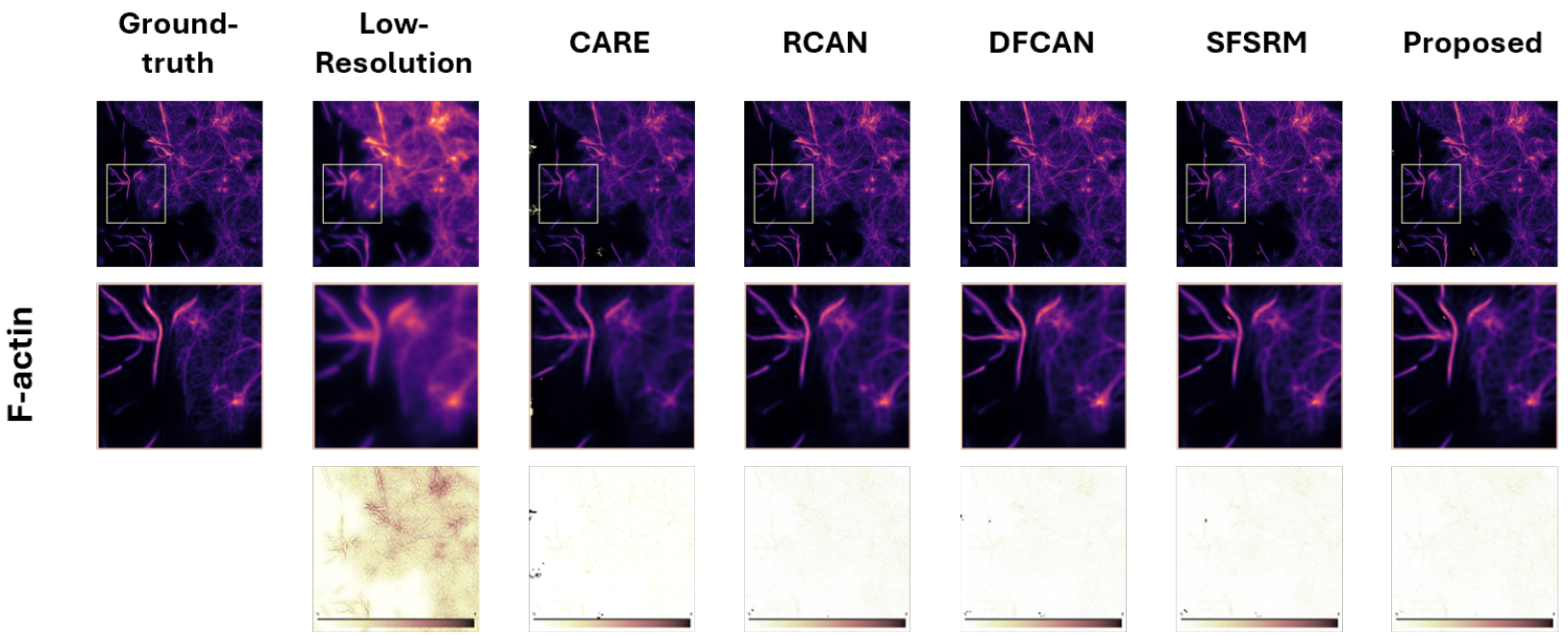


Figure 14: **Row 2:** Zoomed Regions of row 1 and **Row 3:** Mean absolute error between the reconstructed and the groundtruth images.

Quantitative Assessment

Quantitative Assessment

Table 3: Metric Values over Different Sample Types

		Noisy	CARE	RCAN	DFCAN	SFSRM	Proposed
CCPs	MAE	0.054	0.023	0.012	0.013	0.012	0.012
	PSNR	21.88	21.59	26.30	26.17	26.47	26.49
	NRMSE	2.805	2.951	1.707	1.731	1.669	1.674
	MS-SSIM	0.285	0.769	0.907	0.911	0.916	0.918
ER	MAE	0.174	0.043	0.032	0.030	0.028	0.018
	PSNR	14.67	21.92	24.03	25.26	26.58	32.47
	NRMSE	1.586	0.703	0.548	0.489	0.420	0.207
	MS-SSIM	0.356	0.758	0.868	0.827	0.822	0.889
MTs	MAE	0.103	0.040	0.032	0.035	0.031	0.030
	PSNR	18.13	25.14	26.81	26.35	27.51	27.56
	NRMSE	0.829	0.369	0.305	0.322	0.281	0.279
	MS-SSIM	0.621	0.720	0.768	0.786	0.808	0.840
F-actin	MAE	0.103	0.031	0.024	0.025	0.022	0.019
	PSNR	17.78	24.16	28.64	28.57	28.77	28.97
	NRMSE	0.930	0.324	0.256	0.261	0.240	0.226
	MS-SSIM	0.639	0.708	0.810	0.805	0.831	0.851

Quantitative Analysis and Uncertainty Quantification

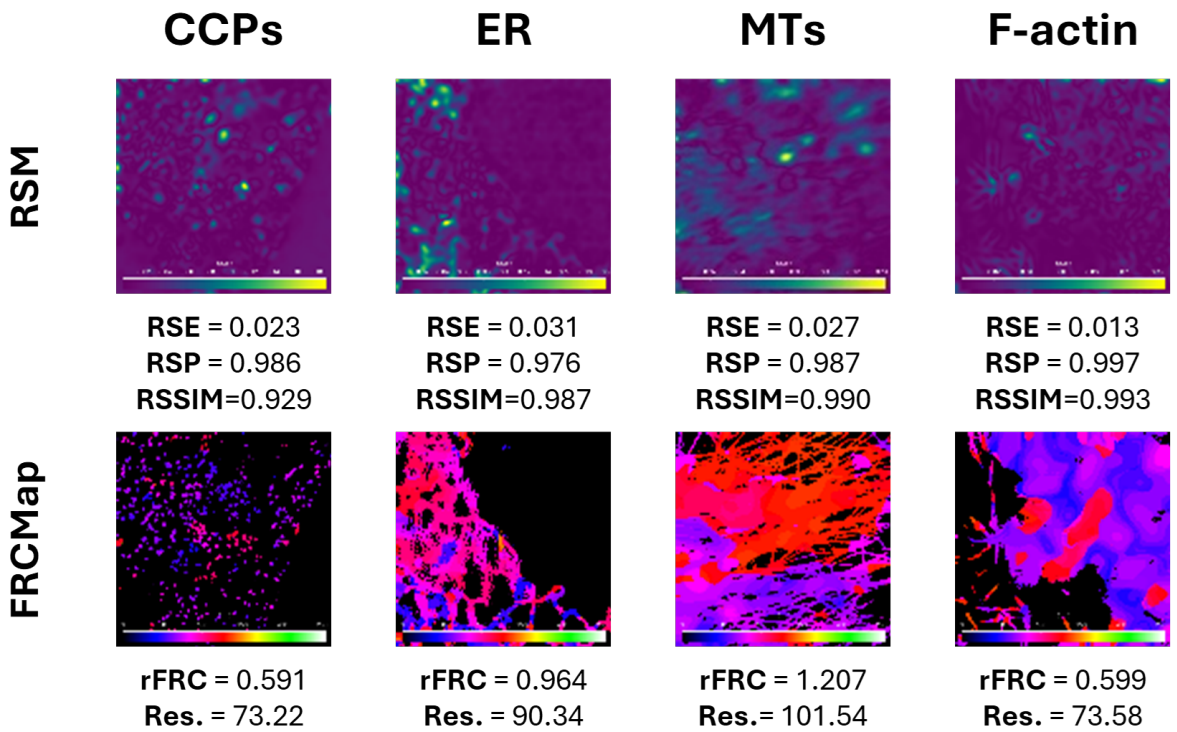


Figure 15: Resolution Scaled Error Maps (**RSM**) and rolling Fourier Ring Correlation Maps (**rFRC Maps**) for different specimen types under the proposed methodology. Also provided are the corresponding metric values: resolution scaled error (**RSE**), Pearson correlation coefficient (**RSP**) and structural similarity index (**RSSIM**). Uncertainty quantification metrics: **rFRC value** (higher value indicates higher confidence in the reconstructed structures) and mean resolution (**Res.**) obtained for the given input LR image. **Scalebar:** [0, 1] for RSM and [0, 256] for rFRCMap from left to right.

Conclusion and Future Works

Conclusion

- A deep learning framework is introduced for super-resolution in fluorescence microscopy.
- It leverages real widefield images as inputs and supervise the training with ground truth data acquired from SIM microscopy.
- Successive up- and down-sampling helps preserve contextual dependencies between and LR and HR images.

Future Scope

- Decoupling noise removal and super-resolution task to avoid noise amplification.
- Incorporating uncertainty prediction as part of the training process.
- Design training objective matching the data fidelity in a model-driven optimization technique.

References I