Science Advances
AAAS

Supplementary Materials for

Optimal sparsity allows reliable system-aware restoration of fluorescence microscopy images

Biagio Mandracchia *et al.*

Corresponding author: Biagio Mandracchia, biagio.mandracchia@uva.es; Shu Jia, shu.jia@gatech.edu

Sci. Adv. **9**, eadg9245 (2023)
DOI: 10.1126/sciadv.adg9245

The PDF file includes:

Sections S1 to S3
Figs. S1 to S35
Tables S1 to S7
Legends for movies S1 to S9
Legend for supplementary software file
References

Other Supplementary Material for this manuscript includes the following:

Movies S1 to S9
Supplementary software file

Supplementary Figures

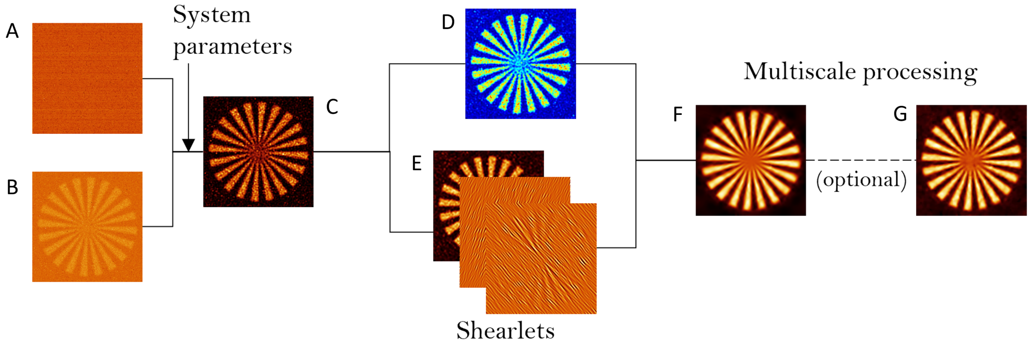

Figure S1: Schematic of the MIRO framework. Calibration images (A) are used to remove the sensor fixed pattern noise from the input image (B). (C) The resulting image is pre-filtered in the Fourier space according to the system parameters. A noise map (D) is evaluated from (C), of which is generated a shearlet decomposition (E). (F) Multiscale pixel-wise domain transform shrinkage is performed to obtain the final image. (G) In case of low SNRs, an optional denoising step can be applied using either unbiased non-local means or a multiscale patch-based Wiener filter.

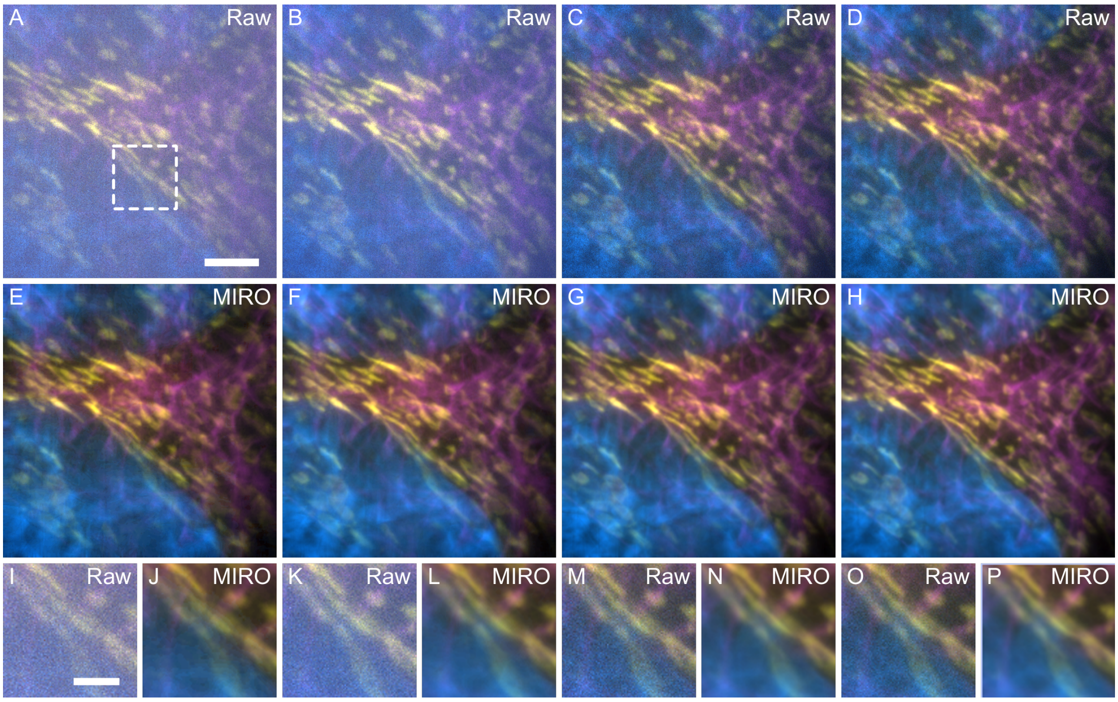

Figure S2: Noise correction of three-color TIRFM images at different SNRs. Raw images of a fixed BPAE cell were obtained at camera exposure times of 10 ms (A), 20 ms (B), 50 ms (C), and 100 ms (D). (e-h) The relative images after MIRO processing. Zoomed-in images corresponding to the dashed area for each SNR level before (I,K,M,O) and after processing (J,L,N,P). Scale bars: 5 µm (A), 2 µm (I).

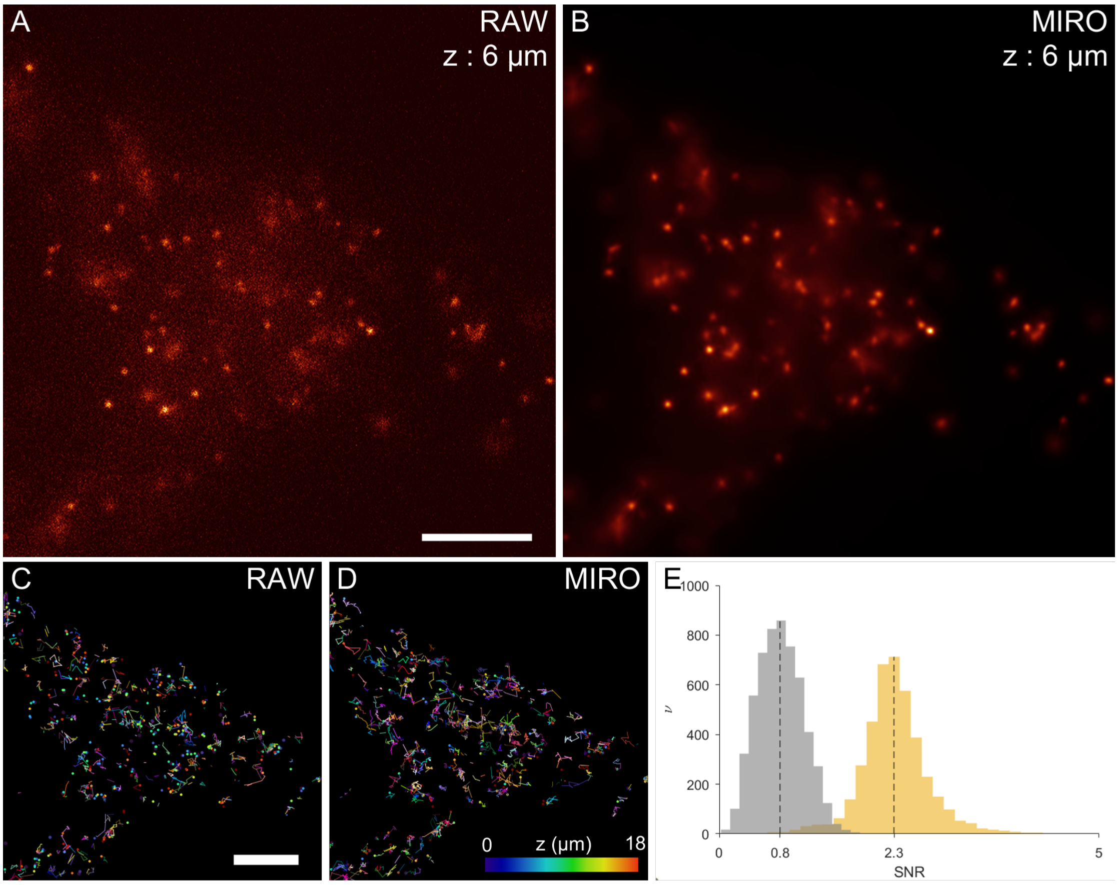

Figure S3: 4D live cell imaging with epifluorescence using an EMCCD camera. Representative image of the 4D data stack of peroxisomes in HEK cell at $t = 0$ and $z = 6\mu m$ (A). The same image after noise correction (B). (C,D) 3D tracking of peroxisomes processed in (A) and (B), respectively. The color of each spot represents its axial position. Tracks in (D) look longer and more connected than (C). (E) At the same time, the histogram of the SNR calculated for each detected spot after noise correction (yellow bars) shows a more than two-fold improvement over the histogram of the raw image (gray bars). Scale bar: 5 µm (A,C).

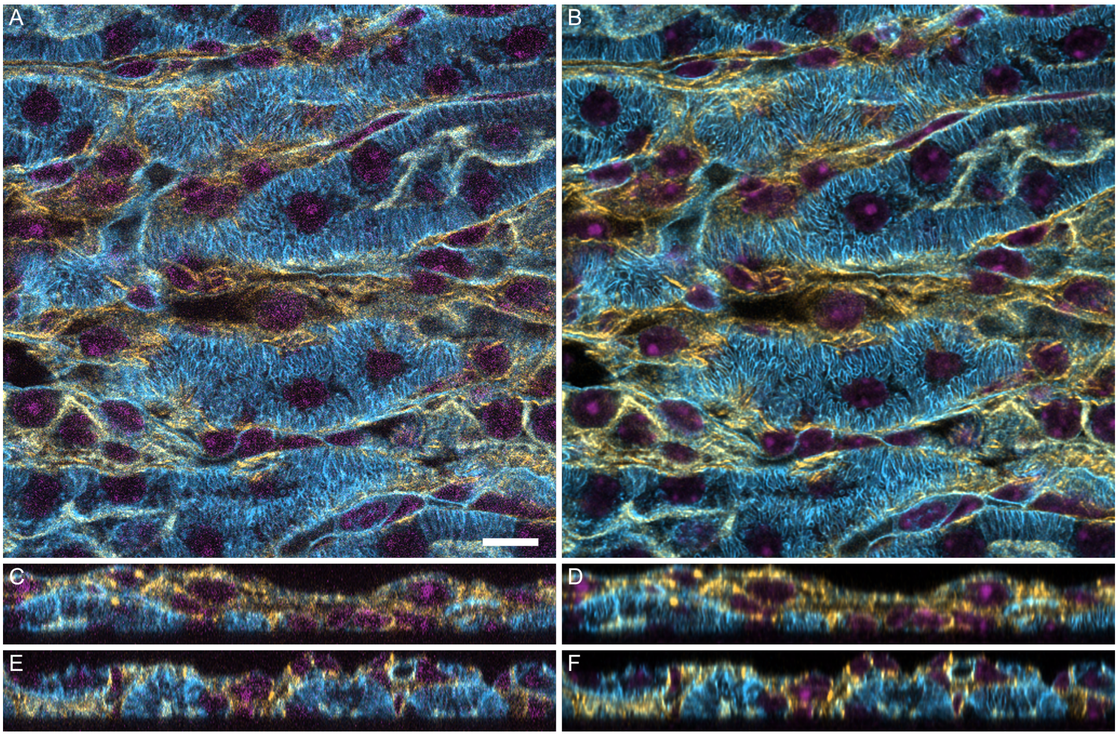

Figure S4: Three-dimensional multi-color tissue imaging with laser scanning confocal microscopy (LSCM). Representative slices of a three-color 3D stack of a mouse kidney section taken along the xy (A), yz (C), and xz (E). (B,D,F) The corresponding slices taken from the MIRO-processed stack. The three color channels correspond to: DAPI (purple), phalloidin (cyan), and wheat germ agglutinin (yellow). Scale bar: 10 µm.

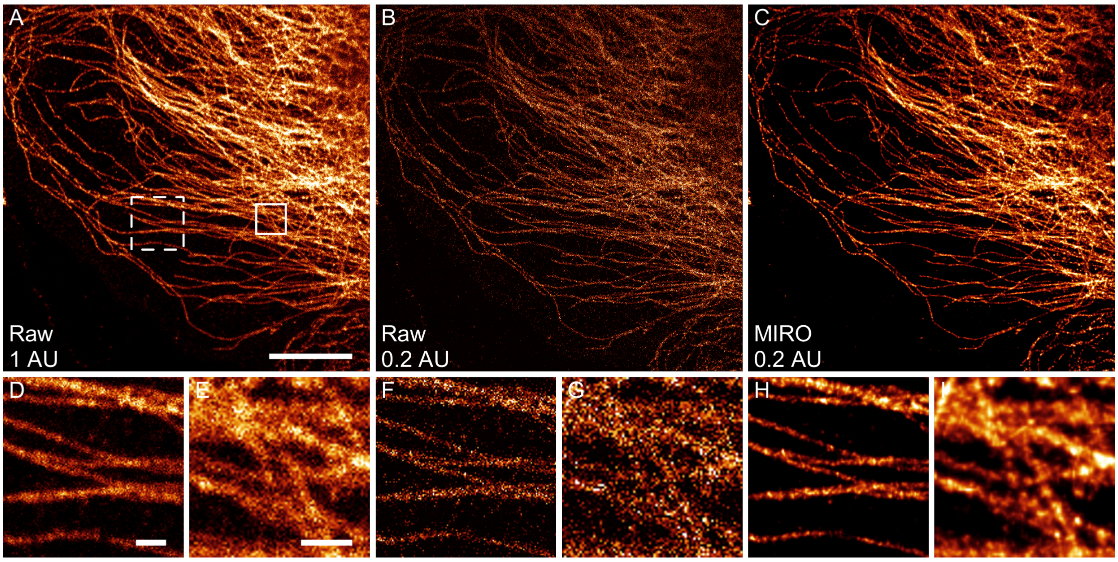

Figure S5: Sub-diffraction-limited confocal microscopy at low illumination power using a GaAsP PMT. Images of fluorescently-labeled microtubules obtained using a LSCM equipped with a GaAsP PMT detector (A, B), where the pinhole size was set to 1 and 0.2 Airy units (AU), respectively. The laser power between images remained unchanged. In image (B), it is clear that the detrimental effect of low SNR cancels the resolution improvement given by the the closed pinhole. After MIRO processing, the image quality is recovered to restore the expected sub-diffraction-limited resolution (C). (D-I) Zoomed-in images relative to the dashed (D, F, H) and solid (E, G, I) boxed areas of (A-C) as marked in (A). Scale bars: 10 µm (A), 1 µm (D,E).

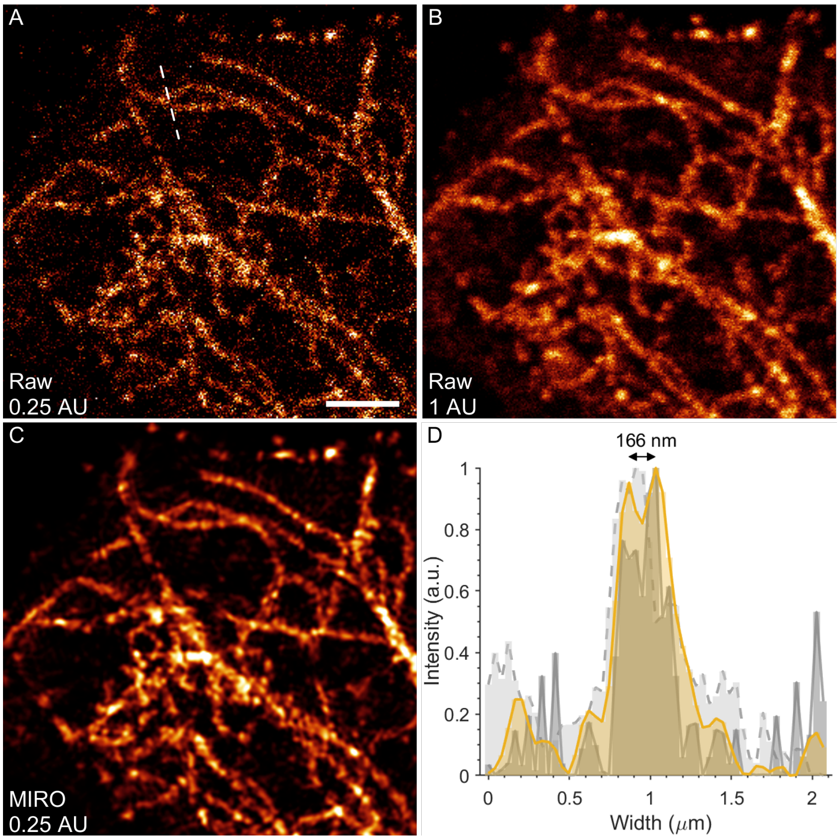

Figure S6: Sub-diffraction-limited LSCM at low illumination power using a standard PMT. Images of fluorescently-labeled microtubules obtained setting the pinhole size to 0.2 (A) and 1 AU (B). The laser power between images remained unchanged. (C) The same image in (A) after MIRO processing. (D) Comparison of the image profile corresponding to the position marked by the dashed line in (A). Because of the low SNR, the profile measured in the raw image with the closed pinhole (dark gray, solid line) does not show a clear resolution improvement compared to the diffraction-limited image acquired with the open pinhole (light gray, dashed line). Instead, this is notably restored in the denoised image (yellow). Scale bar: 2 μm (A).

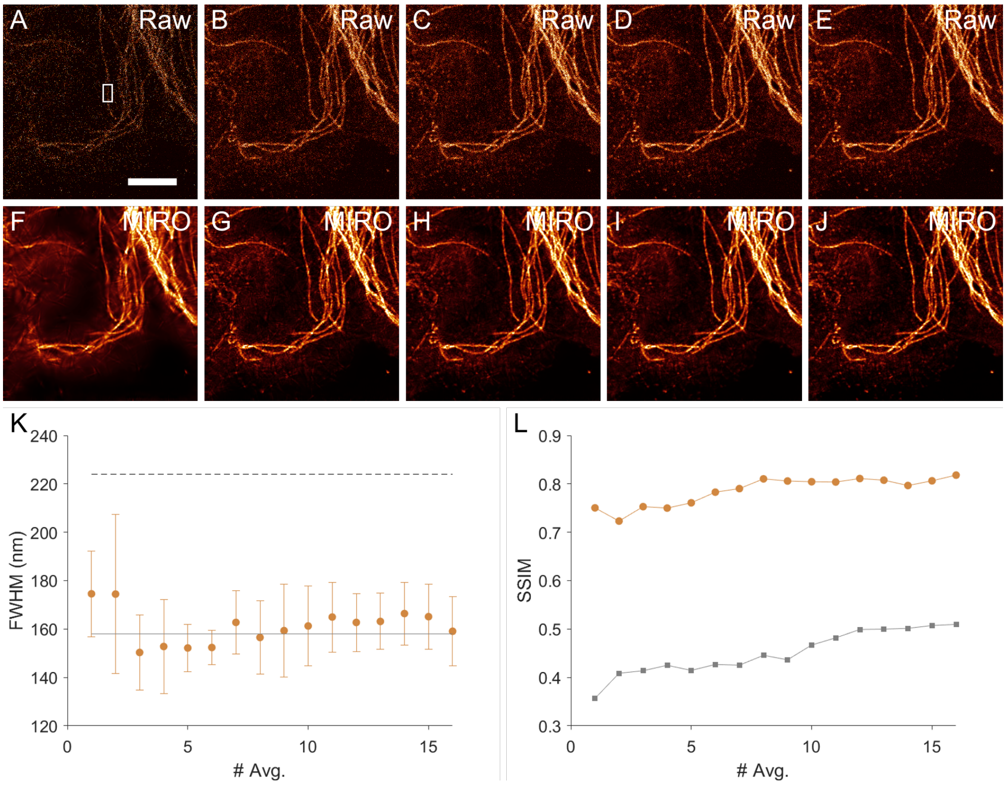

Figure S7: Sub-diffraction-limited resolution recovery of LSCM with MIRO at different SNR conditions. The processed images were the result of the average of an increasing number of captures of the same fluorescently-labeled microtubules. (A-E) depiction of a raw image as well as the result of the average of 4, 8, 12, and 16 frames. All raw images were obtained using LSCM equipped with a standard PMT detector and the pinhole closed at 0.25 AU. (F-J) The same images shown in (A-E) after MIRO processing. (K) Image resolution in the processed images estimated by the average thickness of the microtubule in the boxed area. In all cases, the resolution is consistently lower than the diffraction limit (dashed line) and in agreement with the expected $\sqrt{2}\times$ resolution improvement (solid line). (L) Image quality of the processed images before (gray) and after denoising (yellow). Scale bar: 5 μm.

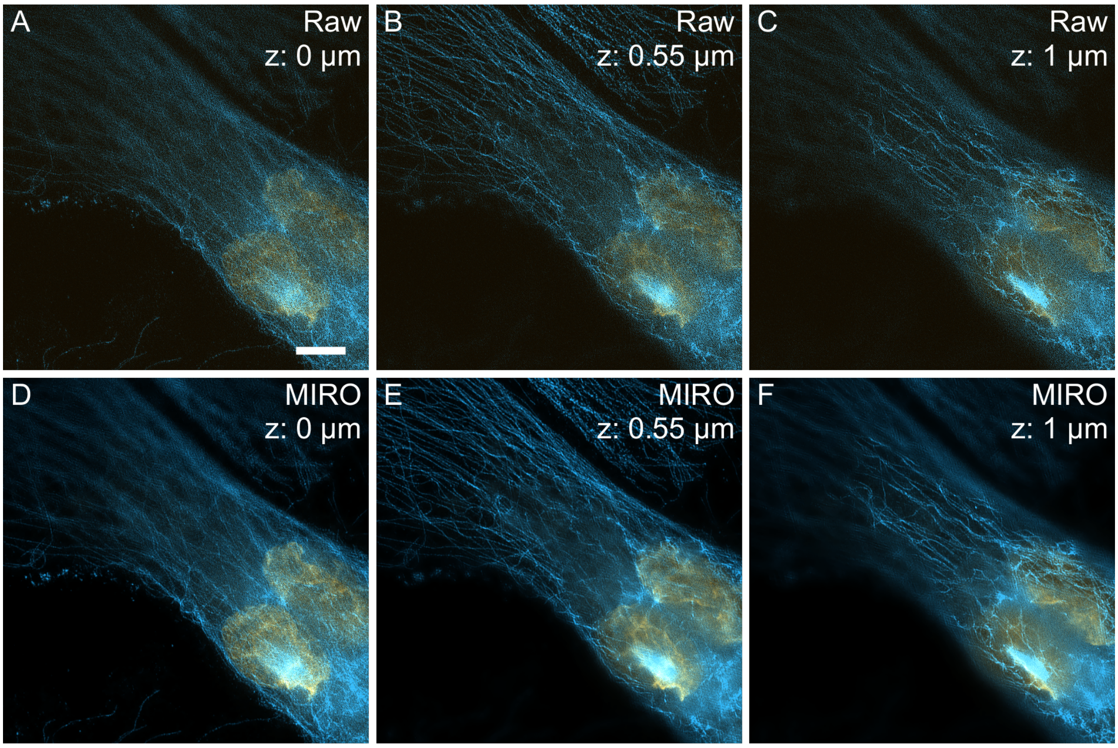

Figure S8: Restoration of 3D SIM reconstructions. (A-C) Representative slices of the SIM reconstruction of fluorescently-labeled microtubules (cyan) and nucleus (yellow) in a BPAE cell. The raw images were recorded at low illumination intensity, which resulted in the presence of noise artifacts in the subsequent reconstruction and were greatly reduced in the MIRO-processed images (D-F). Scale bar: 10 µm.

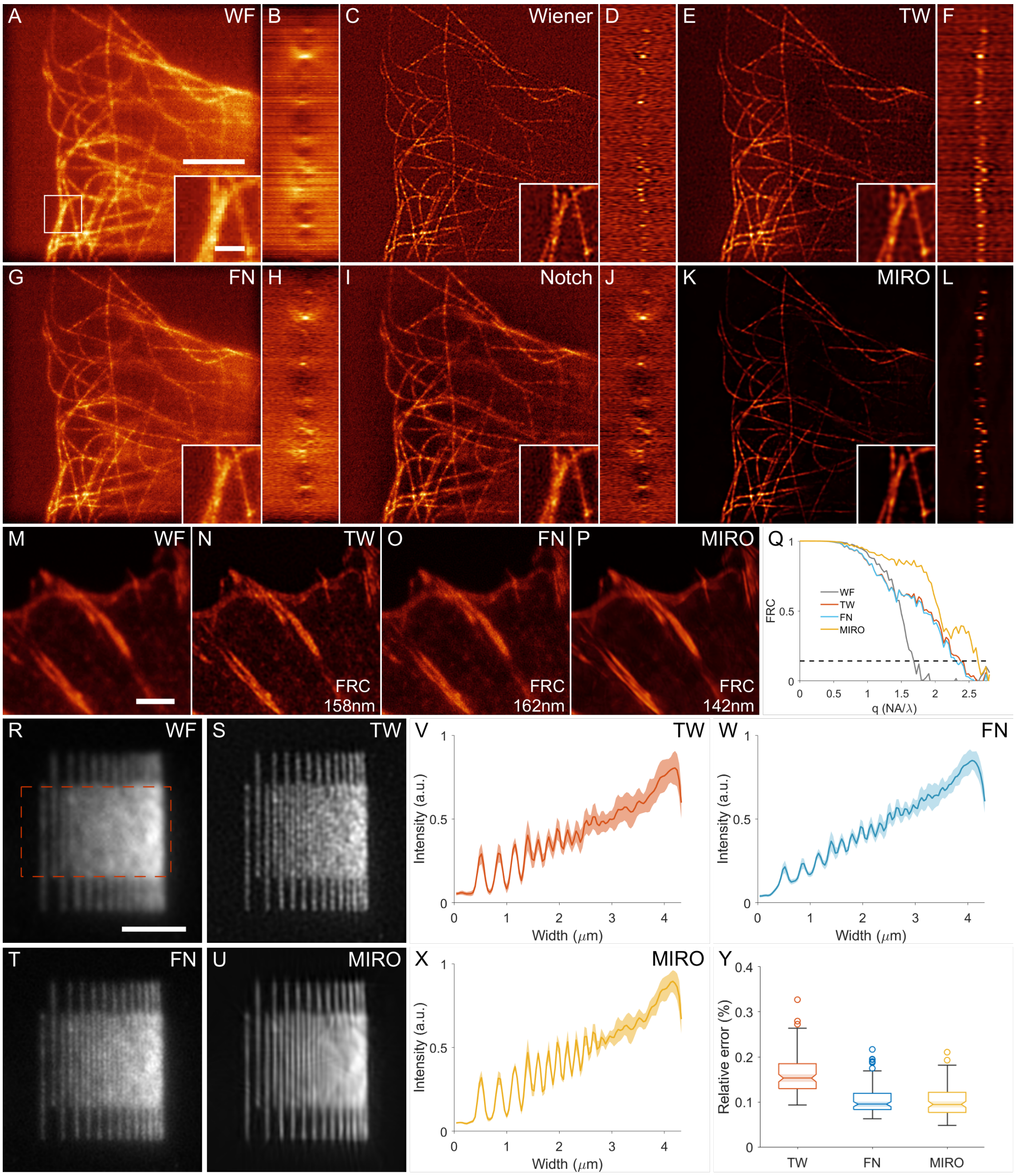

Figure S9: Noise-controlled SIM reconstructions. Cross-sections of 3D-reconstructions (xy and yz) of fluorescently-labeled tubulin filaments: widefield (WF; A,B), SIM reconstruction with regularization parameter fixed to $w = 5 \times 10^{-4}$ (Wiener; C,D), noise-controlled true-Wiener (TW; E,F), flat-noise (FN; G,H), and notch-filtered flat-noise SIM reconstruction (NF; I,J), and MIRO-processed Wiener reconstruction (MIRO; K,L). SIM reconstructions of GFP-zyxin in focal adhesions: WF (M), TW (M), FN (O), and MIRO (P). (Q) Measurement of Fourier ring correlation (FRC) of (M-P) as described in [42]. We see that TW and FN have nearly identical curves with a slightly better performance of the former (158 and 162 nm, respectively). Instead, the MIRO-processed image shows higher correlation in the middle frequency range and higher resolution overall (142 nm). Similarly, from the reconstructions of chirped nanopatterned structures (R-U) we can see that the FN reconstruction shows less noise than TW at the expenses of contrast (V,W). Instead, MIRO shows higher contrast than both methods (X), especially in the middle frequencies as observed also in (Q), with low noise comparable to FN (Y). The plots in (V-X) show the mean and standard deviation of the line pattern along the boxed area. All data are reproduced from the datasets in [42]. Scale bars: 4 μm (A), 0.8 μm (A, inset), 2 μm (M), 1 μm (R).

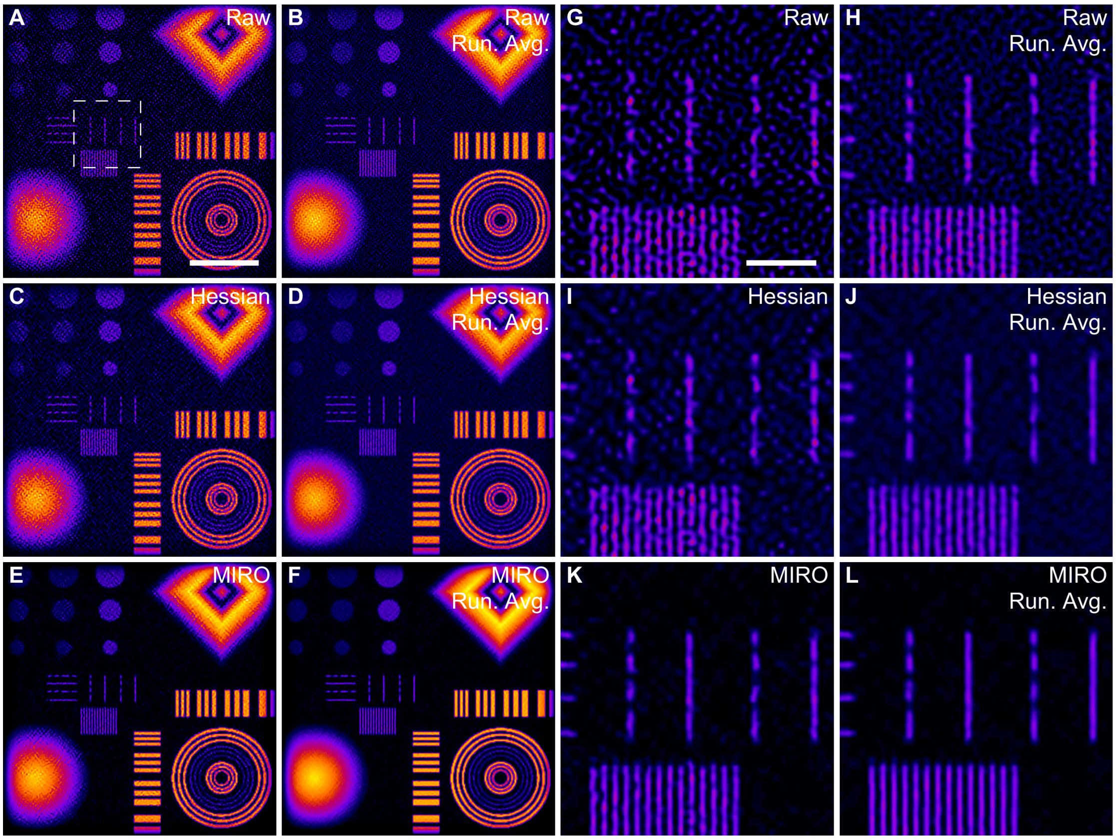

Figure S10: Correction of reconstruction artifacts in simulated SIM images. (A) SIM reconstruction from a representative frame of a simulated SIM dataset from [43]. (B) A frame from the same dataset as (A) reconstructed using running average (i.e. rolling reconstruction). (C,D) The SIM reconstruction in (A,B) after Hessian artifact minimization. This method is based on the computation of Hessian penalty across sequential frames assuming a structural continuity along the x, y, and t axes as a-priori knowledge. Here, rolling reconstruction is usually applied in pre-processing to further mitigate temporal noise fluctuations (D). (E,F) The images in (A,B) after MIRO processing. In this case, each frame is first processed individually (microlocal noise correction). Then, non-local similarities across the sample are evaluated in both space and time in order to apply a Wiener filter to groups of similar patches. The goal of this grouping is to allow the use of a higher-dimensionality transform to maximize the input sparsity during the Wiener filtering. Importantly, patches are grouped only if they have a minimum similarity score and no assumption about temporal continuity or sample dynamics is necessary. (G-L) Zoomed-in images of (A-F) corresponding to the area marked by the dashed box in (A). Scale bars: 4 µm (A), 1 µm (G).

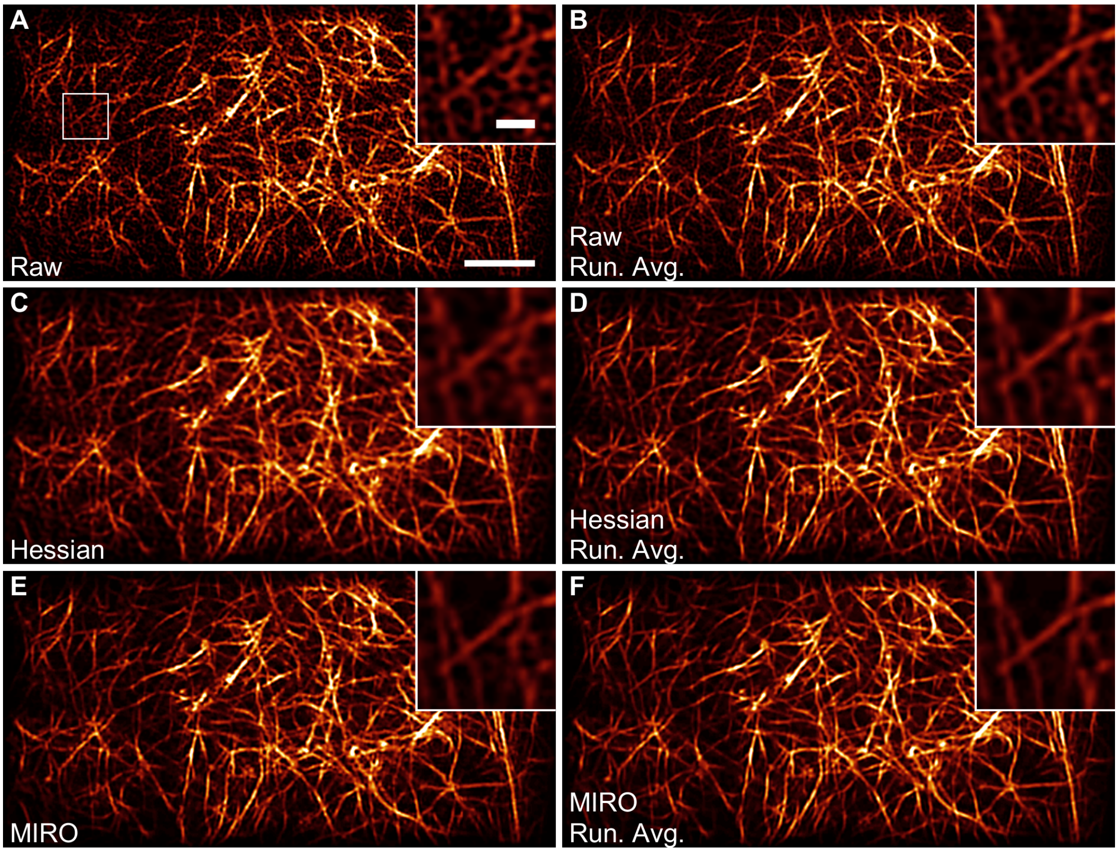

Figure S11: Correction of reconstruction artifacts in experimental SIM images. (A) SIM reconstruction from a dataset of cellular actin structures from [43]. (B) The second frame from the same dataset as (A) reconstructed using running average (i.e. rolling reconstruction). (C,D) The SIM reconstruction in (A,B) after Hessian artifact minimization. From the inset in (D), it is possible to observe the beneficial effect of running reconstruction to resolve features that may be lost under the noise floor, see inset in (C). Nonetheless, with MIRO processing (E,F) the same structure appears visible even without rolling reconstruction (E). Scale bars: 2 μm.

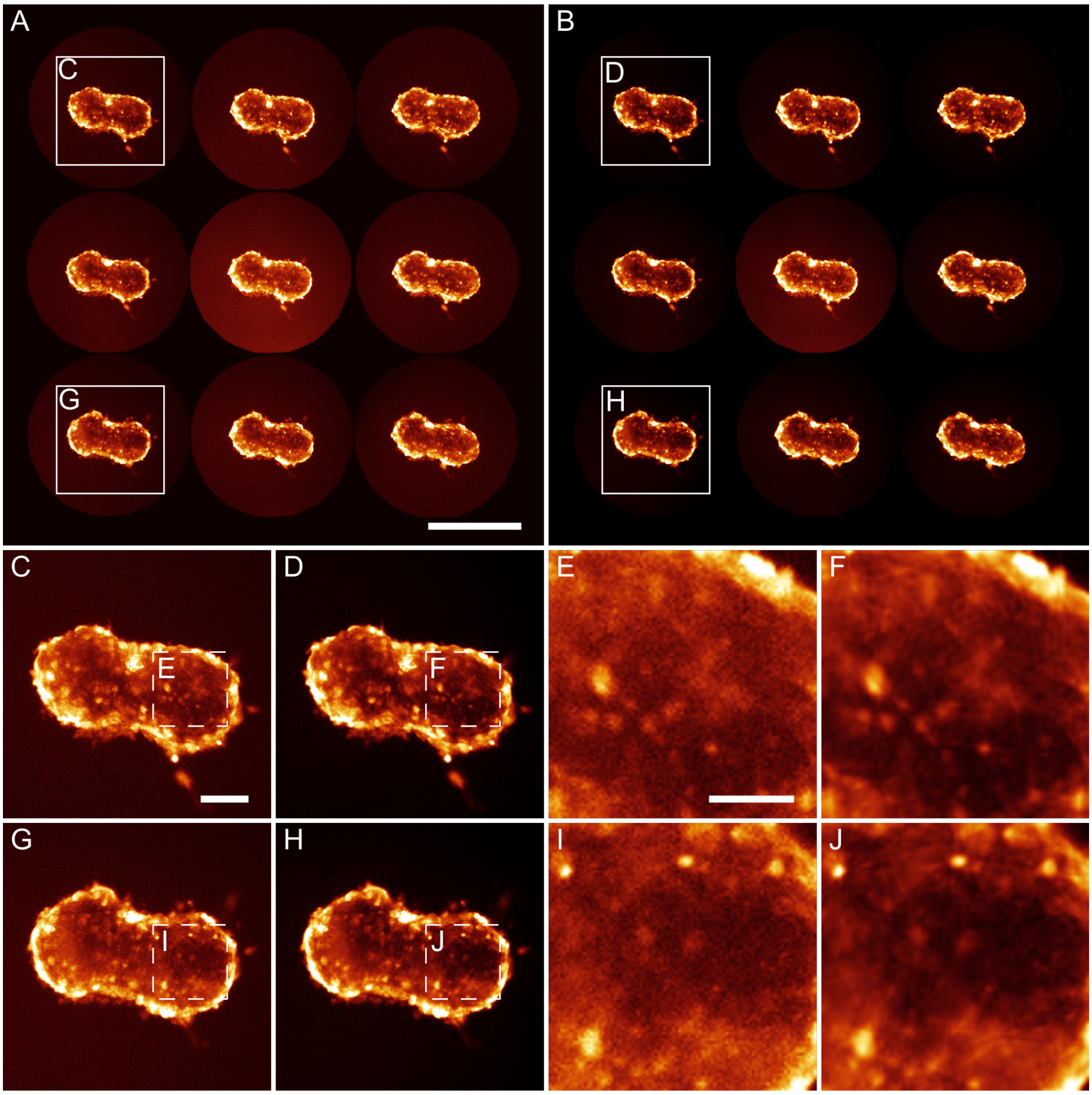

Figure S12: Noise correction of Fourier light-field microscopy (FLFM) images. (A) Simultaneous angular views of a live cardiomyocyte spheroid were obtained at 100 Hz camera frame-rate, resulting in low SNR images. (B) the same image in (A) after MIRO noise correction. (C,D,G,H) Zoomed-in images corresponding to different angular projections marked by the boxed areas in (A) and (B). (E,F,I,J) Zoomed-in images showing further details into the dashed areas in (C,D,G,H), respectively. Scale bars: 500 µm (A), 100 µm (C), 50 µm (E).

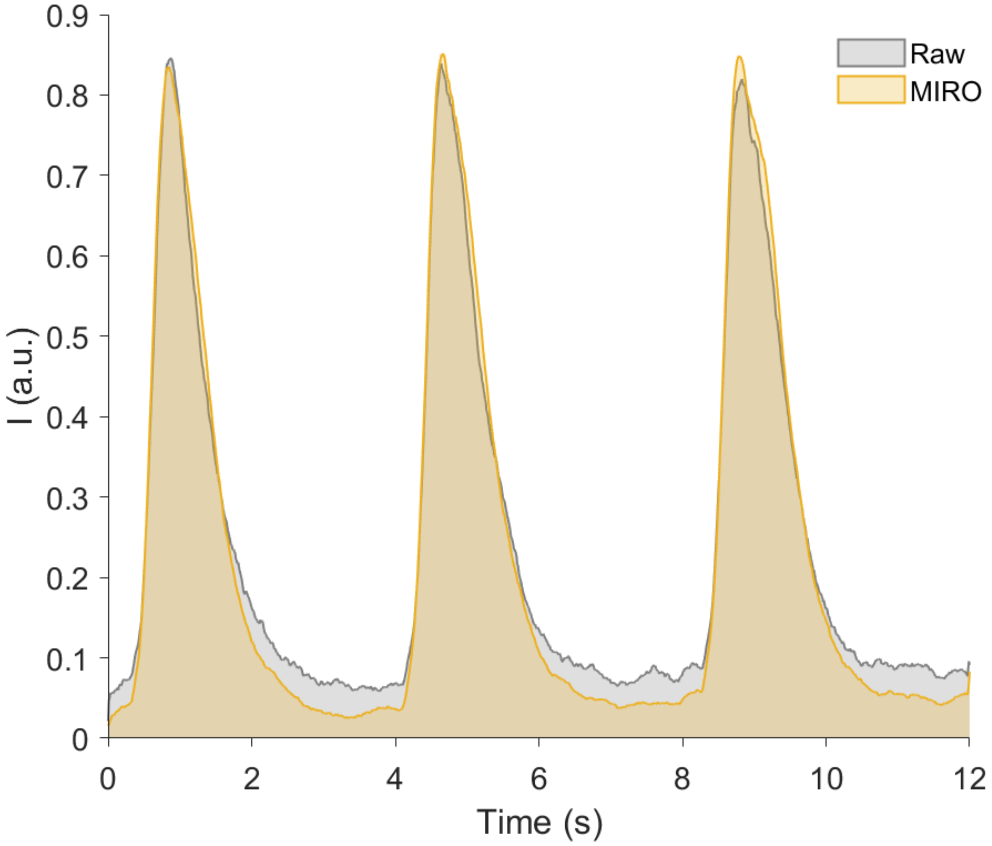

Figure S13: Ca^{2+} transients of the cardiomyocyte indicated by the white arrows in Figs.3M,N of the main text. Comparing the curves measured obtained from the sequence with (yellow) and without (gray) MIRO processing, a deviation in the signal can be noticed especially as soon as the fluorescence starts to drop, resulting in an enhanced peak to baseline contrast for functional characterization. This is due to the decreasing SNR hindering both intensity estimation and tracking of cell location during the spheroid's contractions associated with signaling. Both curves were smoothed using a moving average with a span of 25 time points.

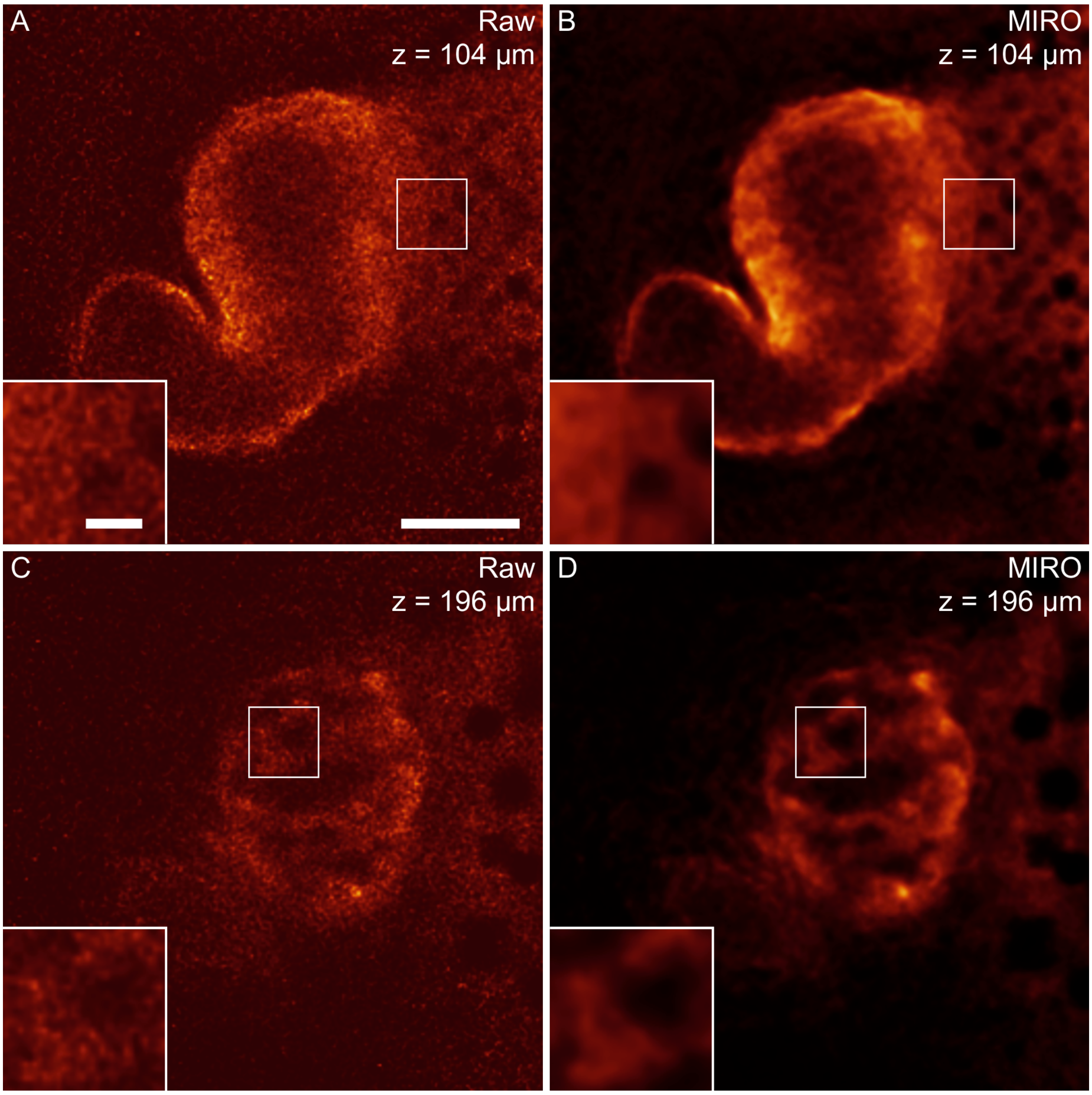

Figure S14: MIRO noise correction improves computational reconstruction of FLFM acquisitions of a tadpole heartbeat. A living tadpole was imaged using a custom FLFM setup for mesoscopic imaging. The laser illumination was kept low while the camera frame-rate was set to 100 Hz to better follow the heart contraction. (A,C) Representative slices (z = 104 µm and z = 196 µm, respectively) of a volumetric stack obtained by numerical reconstruction of a FLFM 2D image. (B,D) The same slices shown in (A) and (C) taken from a volumetric stack obtained by reconstructing the same FLFM image after MIRO denoising. MIRO preprocessing removes spurious signal due to electrical noise while preserving the fluorescence-related sample details. Scale bars: 100 µm (A), 20 µm (A inset).

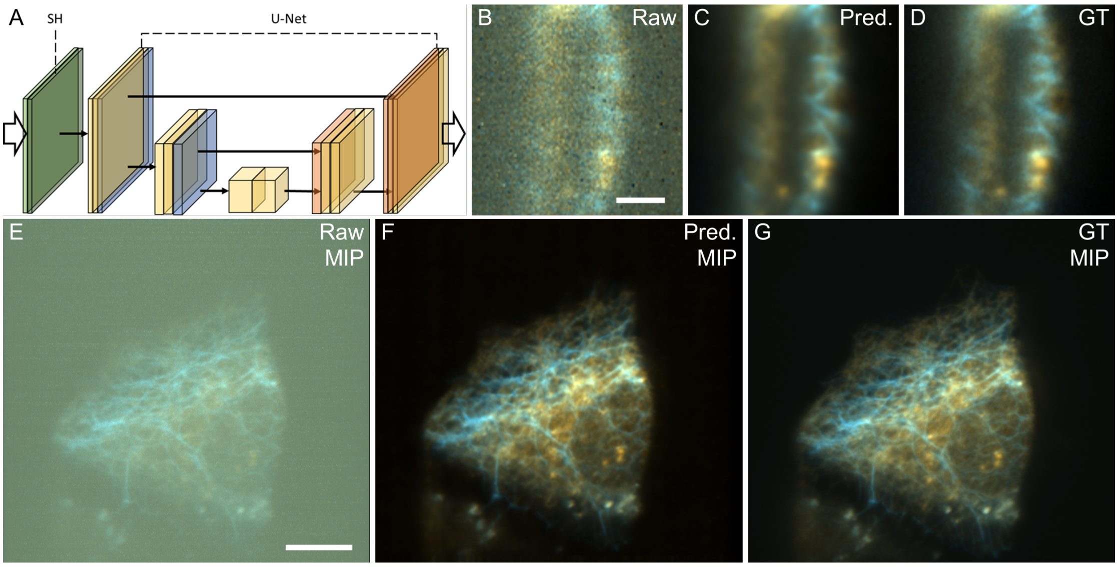

Figure S15: Optimally-sparse image representation facilitates learning-based restoration. (A) Hybrid network architecture. In brief, input images are preprocessed with system-aware parameters and decomposed in the shearlet domain. Than, they are passed through a custom convolutional layer, which performs a learning-based microlocal soft shrinkage (SH). The output of this layer is then fed to a U-Net architecture that filters the sample's relevant features learned across different spa-tial and shearlet scales. (B-D) Example input image (B) shows a zoomed-in slice of volumetric VAPB (orange) and keratins (blue) in fixed epithelial cells imaged by dual-color Lattice Light-Sheet Microscopy. (C, D) show the corresponding predicted output (C) and ground truth (GT) (D). (E-G) Corresponding maximum intensity projection (MIP) images of the whole volumetric dataset as shown in b-d, respectively. Scale bars: 5 µm µm (B), 10 µm (E).

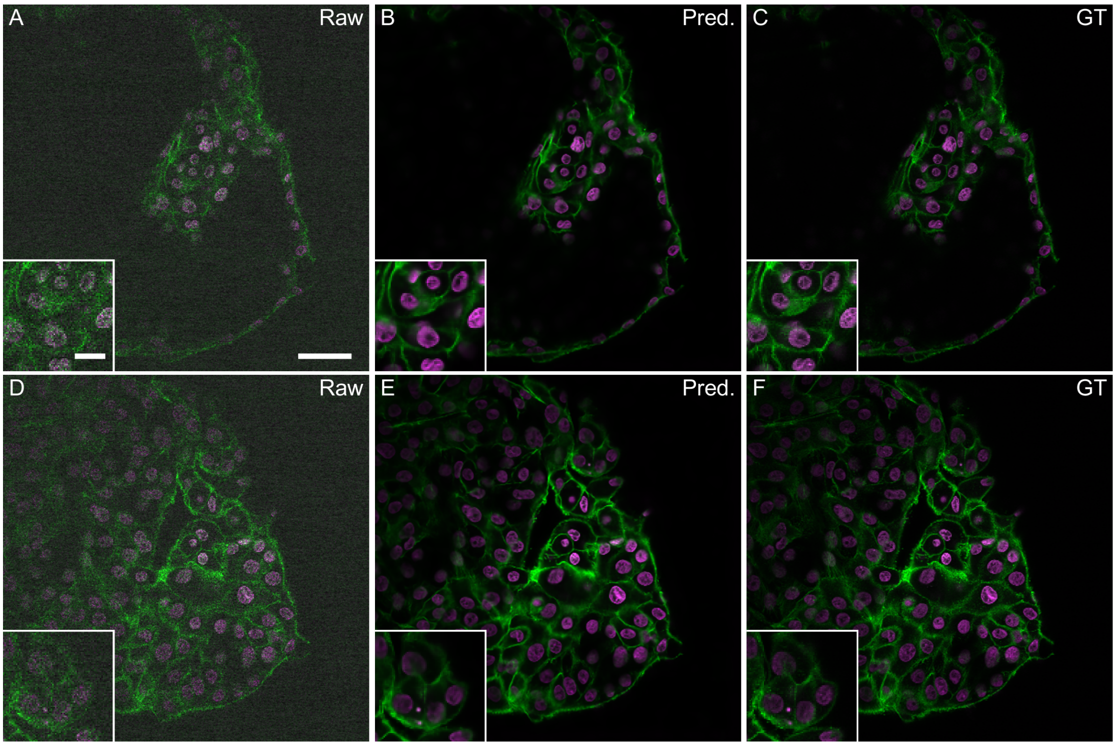

Figure S16: Restoration of MCF10A mammary 3D organoid's LSCM data via MIRO-based hybrid neural network. (A) Representative slice of the low SNR data (z = 12.7 μm) with the relative prediction (B) and ground truth (C) images. (D) Slice of the low SNR data at z = 5.7 μm and the relative prediction (E) and ground truth (F) images. The two channels correspond to F-actin (green) and nuclei (purple). Scale bars: 50 μm (A), 10 μm (A inset).

Supplementary Tables

Exp. time		SSIM		PSNR (dB)	
		Raw	MIRO	Raw	MIRO
10 ms	ch. 1	0.22	0.78	12.65	29.39
	ch. 2	0.50	0.95	20.49	36.39
	ch. 3	0.34	0.87	12.70	29.93
20 ms	ch. 1	0.28	0.86	12.77	31.36
	ch. 2	0.65	0.97	26.07	39.14
	ch. 3	0.39	0.90	15.25	32.14
50 ms	ch. 1	0.33	0.90	19.01	33.68
	ch. 2	0.82	0.98	32.33	40.60
	ch. 3	0.55	0.94	22.97	34.12
100 ms	ch. 1	0.49	0.92	23.15	34.47
	ch. 2	0.90	0.99	35.44	42.99
	ch. 3	0.66	0.95	28.56	35.16

Table S1: Values of SSIM and PSNR before (Raw) and after denoising (MIRO) relative to the datasets shown in Figure S2. Channels 1, 2, and 3 represent the blue (nucleus), yellow (mitochondria), and purple (actin) channels, respectively.

F/F_0	4.33
$\tau_{1/2}$ (s)	0.16
$d_{1/2}$ (s)	0.62
Period (s)	4.09

Table S2: Functional characterization of cardiomyocyte beating and Ca^{2+} transients calculated from MIRO-processed FLFM data shown in Supplementary Video 7. F/F_0: peak/baseline; $\tau_{1/2}$: half rise to peak time; $d_{1/2}$: peak to half decay time.

Heart rate (s^{-1})	1.67 ± 0.03
Diastolic vol. (mm^3)	0.0739
Systolic vol. (mm^3)	0.0413
Ejection fraction (mm^3/min)	2.35

Table S3: Quantification of the embryonic heart function in a embryonic frog *in vivo* observed via FLFM (see Supplementary Video 8). The raw frames of the video used for the measurements were processed with MIRO before image reconstruction.

Dataset	Technique	Detector	Optical parameters			MIRO		Notes
			λ (nm)	px (nm)	NA	Mode	OF	
Beads	Epi	sCMOS	680	65	1.45	soft	-	Fig.1
Siemens Star	Simul.	sCMOS	680	65	1.45	hard	-	Figs.1,S1
BPAE cells	TIRFM	sCMOS	461	65	1.45	soft	-	Fig.S2
BPAE cells	TIRFM	sCMOS	512	65	1.45	soft	-	Fig.S2
BPAE cells	TIRFM	sCMOS	599	65	1.45	soft	-	Fig.S2
Peroxisomes	Epi	EMCCD	513	50	1.46	soft	-	Fig.1,S3
Mitochondria	Epi	sCMOS	515	65	1.45	hard	-	Fig.2
Mouse Kidney	LSCM	PMT	461	120	1.4	soft	-	Figs.2,S4
Mouse Kidney	LSCM	PMT	519	120	1.4	soft	-	Figs.2,S4
Mouse Kidney	LSCM	PMT	600	120	1.4	soft	-	Figs.2,S4
Microtubules	LSCM	GaAsP	513	45	1.46	soft	-	Figs.2,S5
Microtubules	LSCM	PMT	513	40	1.4	soft	-	Figs.S6,S7
BPAE cell	SIM	EMCCD	461	80	1.4	hard	NLM	Figs.2,S8
BPAE cell	SIM	EMCCD	513	80	1.4	hard	NLM	Figs.2,S8
Microtubules	SIM	sCMOS	525	82	1.4	soft	NLM	Fig.S9, [42]
Focal adhesions	SIM	EMCCD	535	80	1.4	soft	-	Fig.S9, [42]
Nanopattern	SIM	EMCCD	535	80	1.4	hard	-	Fig.S9, [42]
Target	Simul.	-	515	32	0.9	hard	WF	Fig.S10, [43]
Actin	SIM	sCMOS	515	32	1.47	soft	WF	Fig.S11, [43]
Cardiomyocytes	FLFM	sCMOS	525	1.8	0.078	hard	NLM	Fig.S12
Tadpole	FLFM	sCMOS	525	1.8	0.078	hard	WF	Fig.S14
Epithelial cells	LLSM	sCMOS	515	65	1.4	MIROnet	-	Fig.S15
MFC10A	LSCM	GaAsP	416	65	1.4	MIROnet	-	Fig.S16
MFC10A	LSCM	GaAsP	515	65	1.4	MIROnet	-	Fig.S16

Table S4: List of all datasets presented along the manuscript with relative reconstruction parameters. OF: optional filter. SF: Supplementary Figure. Simul.: simulation. hard: microlocal hard thresholding. soft: microlocal soft thresholding. MIROnet: MIRO-based hybrid neural network. NLM: unbiased non-local means. WF: Wiener filter.

Supplementary Movies

Movie S1. 4D live-cell imaging of peroxisomes in HEK cell. The video was acquired using a microscope equipped with an EMCCD camera. The axial position of each element has been color-coded.

Movie S2. Three-dimensional tracking of peroxisomes in HEK cell. The video depicts the results of single-particle tracking performed on the dataset shown in Supplementary Video 1. Here, the color code represents the signal-to-noise ratio of the tracked point. The MIRO-processed tracks display a clear improvement in SNR.

Movie S3. Live-cell imaging of mitochondria moving in a human embryonic kidney (HEK) cell. Video sequence before (left) and after (right) MIRO noise correction. Exposure time: 20 ms. Scale bar: 5 µm.

Movie S4. Multi-color three-dimensional stack of a mouse kidney section acquired with laser scanning confocal microscopy (LSCM). The sequence shows the data before (left) and after (right) MIRO noise correction. The three color channels correspond to: DAPI (purple), phalloidin (cyan), and wheat germ agglutinin (yellow). Scale bar: 10 µm.

Movie S5. Three-dimensional renderings of LCSM stacks of fluorescently-labeled microtubules in a BPAE cell. The first sequence shows a 3D rendering of the volumetric scan acquired with the open pinhole (RAW, 1 AU). The second and the third ones show the same area but acquired with the pinhole almost closed and rendered before (RAW, 0.2 AU) and after MIRO processing (MIRO, 0.2 AU).

Movie S6. Artifact minimization in multi-color three-dimensional SIM reconstructions. The sequence shows the reconstructed volumetric dataset before (left) and after (right) MIRO processing. The two channels correspond to the fluorescently-labeled microtubules (cyan) and nucleus (yellow) in a BPAE cell. Scale bar: 10 µm.

Movie S7. Fourier light-field imaging of a live cardiomyocytes spheroid. Maximum intensity projection of a three-dimensional stack computationally reconstructed from 2D light-field images without (left) and with MIRO preprocessing (right). The video was recorded at 100 Hz. Scale bars: 100 µm.

Movie S8. Fourier light-field imaging of the heartbeat of a fluorescently-labeled tadpole. Comparing the two sequences obtained reconstructing raw data (left) and noise-corrected data (right), it is noticeable that MIRO preprocessing improves the robustness of Fourier light-field reconstruction. The video was recorded at 100 Hz. Scale bars: 100 µm.

Movie S9. Restoration of 3D LSCM data of a MCF10A mammary organoid via MIRO-based hybrid neural network. The sequence shows the raw data (left), the predicted results (center) and the ground trouth dataset (right). The two channels correspond to F-actin (green) and nuclei (purple). Scale bar: 50 µm.

Supplementary Software

Supplementary Software. Matlab package for MIRO (Multiscale Image Restoration through Optimally-sparse representation). This is a software for microlocal noise correction designed for fluorescence microscopy. It uses the optimal sparsity of the shearlet domain and a physics-based noise estimation to improve the quality of microscopy images acquired using any of the most common camera sensors. This yields a stable and effective noise correction improving the reliability of image analysis and reconstruction. The software requires a standard computer with enough RAM to support MATLAB 2021b. For minimum performance, this will be a computer with about 4 GB of RAM. For optimal performance, we recomend 16+ GB of RAM, a CPU with 6+ cores and 3.2+ GHz/core. To work properly, the software requires MATLAB 2021b+ with the Image Processing Toolbox installed.

Supplementary Text

CONTENTS

1 MULTISCALE IMAGE RESTORATION

In this section, we summarize the principal noise sources in the most common scientific sensors and delineate the theoretical model underlying our image restoration strategy. Then, we analyze in detail all the elements of the MIRO framework and discuss their role in the restoration process.

1.1 Photon transfer noise sources in digital cameras

The quality of a camera image can be degraded by interactions that produce a signal variance or noise from pixel-to-pixel. These interactions can have several sources so that the total effective noise per pixel is the quadrature sum of different noise components[76]. The photon-detector interaction is related to the presence of *shot noise* and *Fano noise*. However, the latter is usually negligible compared to shot noise and only significant when the average number of electrons generated per pixel is less than 1. Instead, the lack of uniformity in the pixel sensitivity is associated to *fixed pattern noise*. Finally, *read noise* encompasses all other noise sources that are not dependent on signal strength.

SIGNAL SHOT NOISE. Signal shot noise is connected to the statistics describing the way photons spatially arrive on the detector. This is generally a spatially and temporally random phenomenon that obeys the Bose-Einstein statistics. However, for semiconductors at room temperature and wavelengths in the visible spectrum, the signal deviation due to shot noise (σ_{SHOT}) reduces to:

$$\sigma_{SHOT}(p,t) = \sqrt{S(p,t)}\,, \tag{1}$$

where S is the average number of electrons generated in the pixel p at time t as a function of interacting photons. Shot noise is described by the Poisson probability distribution,

$$\mathcal{P}_i(p,t) = \frac{S_i(p,t)}{i!} e^{-S_i(p,t)}\,, \tag{2}$$

where $\mathcal{P}_i$ is the probability that there are i interactions at the pixel p at time t. Importantly, a Poisson distribution can be approximated by a Gaussian distribution when $S > 3$ with an error less than 5%.

DARK SHOT NOISE. Dark current, I_D, is a thermal phenomenon resulting from electrons spontaneously generated within the silicon chip (valence electrons are thermally excited into the conduction band) even when no photons are incident on the camera. The number of dark electrons collected in one frame is termed dark shot noise and it does not depend on the signal level but on the exposure time:

$$\sigma_D(p) = \sqrt{I_D(p) \cdot \tau}\,. \tag{3}$$

For scientific-grade cameras, the contribution of dark current can be neglected in many practical cases, e.g. exposure times below 1 s. However, dark noise is still relevant in situations that require low-cost sensors or long exposure times.

FIXED PATTERN NOISE. Once the photoelectrons are generated, they are collected in the detector's pixels. This process is not perfect so that the collection efficiency changes from pixel to pixel, resulting in a sensitivity non-uniformity. Since this non-uniformity remains constant in time, it forms the same spatial pattern from image to image and is termed fixed pattern noise (FPN). Thus, FPN presents a high temporal correlation that cannot be reduced by frame averaging. Instead, it becomes more significant after some common operations like temporal averaging or maximum intensity projection. This is especially noticeable in CMOS sensors where column amplifiers are responsible for the appearance of the typical vertical stripes. The effect of FPN depends on the electronics of the camera and can be divided in two components: one proportional to the photocurrent, termed photoresponse non-uniformity (PRNU), and one independent from the signal and due to offset voltages due to the amplifiers used, termed dark signal non-uniformity (DSNU). Thus, FPN results from the assumption that signal gain, g, and offset, o, are uniform across the sensor area:

$$g(p) = g\,, \tag{4}$$

$$o(p) = o\,. \tag{5}$$

This generates an error in the signal estimation that is modeled as Gaussian distributed noise whose standard deviation varies linearly with the signal:

$$\sigma_{FPN}(p,t) = \alpha S(p,t) + \beta\,. \tag{6}$$

Here, it is important to notice that the parameters α and β correspond to the PSNU and DSNU components of the sensor respectively, and must not be confused with g and o. Unlike shot noise, which varies with the square root signal, FPN is proportional to the signal. So, it will dominate signal shot noise over most of a sensor's dynamic range for visible application unless the pixel-to-pixel non-uniformity is properly calibrated.

READ NOISE. Read noise represents any noise source that is not a function of the signal. It is generated in producing the electronic signal, which results from the sensor design but also by the design of the camera electronics. For example, CMOS pixels perform charge-to-voltage conversion at each location using additional circuitry that increases readout noise and generates extra fixed-pattern noise sources compared to CCDs. Read noise is Gaussian distributed and independent from the illumination level but for some detectors, e.g. CMOS, it can depend on the pixel position:

$$\sigma_R \equiv \sigma_R(p)\,. \tag{7}$$

1.2 Theoretical model

A camera system is composed by different elements that convert the photon input into the digital signal such as the semiconductor, pixel detector, and the electronics that process the signal[77]. Thus, given a certain number $P(p,t)$ of photons impinging on a pixel p at time t, the final output of that pixel will be:

$$DN(p,t) = K(P(p,t))\,, \tag{8}$$

where DN is the output in digital numbers and K the camera transfer function. For the sake of simplicity, we will consider the input signal in photoelectrons, which are related to the incident photons by the relation:

$$S(p, t) = Q \cdot \eta \cdot P(p, t)\,, \tag{9}$$

where Q and η are the quantum efficiency and the quantum yield gain of the detector, respectively. Thus, equation (8) becomes:

$$DN(p, t) = K'(S(p, t))\,, \tag{10}$$

where

$$K'(\cdot) = K(\cdot) \cdot Q \cdot \eta\,. \tag{11}$$

At this point, we can separate the signal-dependent component of K' from the signal-independent component:

$$DN(p, t) = g(p) \cdot S'(p, t) + o(p)\,, \tag{12}$$

where g is the gain in DN/e^{-} units, o the offset in DN, and S' the total number of electrons generated in the pixel p at time t. The number of electrons S' read at a certain pixel is the result of a statistical process modeled by the relation:

$$S'(p, t) = \mathcal{P}\{S; p, t\} + \mathcal{P}\{I_D\tau; p\} + \mathcal{G}\{0, \sigma_R^2; p\} + \mathcal{G}\{0, \sigma_{FPN}^2; p, t\}\,. \tag{13}$$

If we approximate the Poisson distributions, $\mathcal{P}\{\cdot\}$, of signal and dark shot noise with heteroskedastic Gaussian distributions[78], $\mathcal{G}\{\cdot; p\}$, we obtain:

$$S'(p, t) = \mathcal{G}\{S, S; p, t\} + \mathcal{G}\{I_D\tau, I_D\tau; p\} + \mathcal{G}\{0, \sigma_R^2; p\} + \mathcal{G}\{0, \sigma_{FPN}^2; p, t\}\,. \tag{14}$$

Reminding that the sum of independent Gaussian-distributed random variables is by theory, still a Gaussian-distributed random variable:

$$S'(p, t) = \mathcal{G}\{S + I_D\tau, \sigma_T^2; p, t\}\,, \tag{15}$$

where:

$$\sigma_T^2(p,t) = S(p,t) + I_D(p)\tau + \sigma_R^2(p) + \sigma_{FPN}^2(p,t)\,, \tag{16}$$

As discussed in Sec. 1.1, the contribution of FPN can be canceled by calibrating the camera sensor to determine the gain and offset values across the sensor. This way, we can use the true gain and offset values in each pixel, which implies:

$$\sigma_{FPN}(p,t) = 0\,. \tag{17}$$

Substituting equations (15-17) into equation (12), we obtain:

$$DN(p,t) = g(p) \cdot \mathcal{G}\{S + I_D\tau, S + I_D\tau + \sigma_R^2; p, t\} + o(p)\,, \tag{18}$$

We can see from equation (18) that DN is a biased estimator of S:

$$DN(p,t) \to g(p) \cdot S(p,t) + g(p) \cdot I_D(p)\tau + o(p)\,. \tag{19}$$

This can be fixed by removing the last two terms of equations (19) that do not dependent on the signal and considering the measured intensity in terms of e^-:

$$I(p,t) = \frac{DN(p,t) - o'(p,\tau)}{g(p)}\,, \tag{20}$$

where

$$o'(p,\tau) = o(p) + g(p) \cdot I_D(p)\tau\,. \tag{21}$$

This way, we have that:

$$\begin{cases} \mathbb{E}[I(p,t)] = S(p,t) \\ \mathbb{V}[I(p,t)] = \mathbb{E}[I(p,t)] + \sigma_R^2(p) + \sigma_D^2(p) \end{cases} \tag{22}$$

where $\mathbb{E}[\cdot]$ and $\mathbb{V}[\cdot]$ are the expected value and the variance of the random variable $I(p,t)$, respectively. Equation (22) shows that I is indeed an unbiased estimator of the number of photoelectrons generated in each pixel, S. However, to obtain $\mathbb{E}[I]$ it is necessary to know the noise variance, which, in turn, depends on the knowledge of $\mathbb{E}[I]$. This must be substituted, then, with an approximated estimation $\tilde{\mathbb{E}}[\cdot]$:

$$\mathbb{V}[I(p,t)] = \tilde{\mathbb{E}}[I(p,t)] + \sigma_R^2(p) + \sigma_D^2(p) \tag{23}$$

Starting from the noise model summarized in equation (23), we devised a strategy to improve the estimation accuracy of the signal generated by the incident photons based on three main elements:

- the knowledge of relevant acquisition parameters, e.g. τ;
- an accurate calibration of the camera parameters, i.e. $\sigma_R(p)$, $g(p)$, $o(p)$ and $I_D(p)$;
- a sensible approximation $\tilde{\mathbb{E}}[I(p,t)]$ based on mathematical parameters related to sample morphology (similarity) and image representation (sparsity).

1.2.1 *The heteroskedastic Gaussian approximation*

We modeled camera-related noise as the contribution of different noise components and, throughout the previous section, we derived a series of relations which depend not only on the mean and variance, but also on the particular distribution of the processed samples. In particular, the main contributions come from two different statistics, namely the Gaussian and the Poisson distribution. For the sake of simplicity, we exploit the fact that the Poisson distribution of shot noise is well approximated by a heteroskedastic Gaussian distribution, i.e., with a spatially-varying noise variance. The accuracy of this assumption is increases with the number of photons and already with only 10 photons per pixel it can be seen that there is virtually no difference between the estimation accuracy of a truly Poissonian variable and that of its Gaussian approximation[78]. In general, the minimum required photon flux depends on the considered application and desired accuracy. As a rule of thumb, this can be considered a good approximation already when the photon flux is 3-5 photons per pixel given that the error committed by approximating the variance of Poisson distribution with a Gaussian standard deviation is still small[76, 19]. Furthermore, we tested the robustness of our algorithm at low light levels with both simulations and experimental images seeking for deviations in accuracy due to the heteroskedastic Gaussian approximation. Remarkably, we did not observe any relevant worsening of the MIRO performance down to 1 photon/pixel, neither in terms of mean square error nor structural similarity (see section 2.1).

1.3 Pixel response calibration

As discussed in the previous section, camera noise is related to a few pixel-dependent parameters: $\sigma_R^2(p)$, $g(p)$, $o(p)$ and $I_D(p)$. So, a pixel-wise calibration of these parameters has to be performed to optimize noise correction, especially when using sensors with high pixel-to-pixel non-uniformity as CMOS-based camera[19, 79]. The offset $o(p)$, expressed in DN, can be estimated by recording a series of dark images, DN_0, and calculating the temporal mean for each pixel:

$$o(p) = \mathbb{E}[DN_0(p)]\,. \tag{24}$$

The amplification gain $g(p)$, expressed in DN/e^-, can be obtained from multiple sets of images recorded at different illumination intensities and calculated using the relation[16]:

$$g(p) = \arg\min \sum_{k=1}^{K} \left(\left(\nu^k(p) - v(p) \right) - g(p) \left(D^k(p) - o(p) \right) \right)^2 , \tag{25}$$

where K is the total number of illumination levels acquired, k is the k-th illumination sequence, $D^k(p) \equiv \mathbb{E}[DN_k(p)]$ is the mean DN count obtained from temporal averaging of all frames that are acquired during illumination sequence k at pixel p, $o(p)$ and $\nu(p) \equiv \mathbb{V}[DN_0(p)]$ are the offset and variance values for pixel p, and $\nu^k(p) \equiv \mathbb{V}[DN_k(p)]$ stands for the temporal variance of the DN counts for illumination sequence k in pixel p.

Importantly, some detectors can have multiple amplifiers for different illumination conditions, usually high-gain and low-gain amplifiers for low-light and high-light conditions, respectively. In this case, users must perform multiple gain calibrations within each light range.

Once obtained $o(p)$ and $g(p)$, the read noise variance in e^- units, $\sigma_R^2(p)$ can be calculated from the variance of the dark frames:

$$\sigma_R^2(p) = \frac{\mathbb{V}[DN_0(p)]}{g(p)}\,, \tag{26}$$

where $\mathbb{V}[DN_0(p)]$ is the variance calculated at pixel p from the dark image dataset.

Finally, the dark current $I_D(p)$, expressed in e^-/s can be calculated by multiple sets of dark images recorded at different exposure time using linear regression:

$$DN_0(p,\tau) = g(p) \cdot I_D(p)\tau + o(p)\,. \tag{27}$$

In cooled cameras, $I_D(p)$ can be very low[80], e.g. $\langle I_D \rangle_p = 0.06\ e^-/s$, except for a few "hot" pixels. In these cases we just assumed $I_D(p) = \langle I_D \rangle_p$ and corrected the "hot" pixels by interpolation. Usually, this is not possible with low-cost chips, in which case equation (27) must be used.

1.4 Sparsity-based image restoration

The resolution of an optical system is physically governed by the laws of diffraction, which limit the level of detail that such system can convey. In terms of frequency analysis, this means that the optical transfer function (OTF) will always present a characteristic cutoff frequency, $f_{c,o}$ [81]. At the same time, an analog-to-digital converter, like a digital camera, also presents a cutoff frequency, termed Nyquist frequency f_N, over which any signal is considered under-sampled[82]. Thus a digital imaging system can be considered as a low-pass filter whose effective cutoff frequency is $f_c = \min\{f_{c,o}, f_N\}$. For the sake of simplicity, we consider a system where $f_c = f_{c,o}$ and we will discuss the case of under-sampling separately in section 2.3.

We can say, then, that all the useful signal collected by a digital image is contained within a small region of the Fourier space delimited by f_c. Instead, the components of electric noise present in a digital camera have a constant power spectrum[19, 83], i.e. noise is equally present at all frequencies. The sparsity of the signal in the Fourier space allows to easily reduce noise by removing only the frequencies $f > f_c$, while leaving signal unaltered. Most importantly, the same reasoning can be extended to other transforms that provide a sparser representation of the signal and, in turn, a more efficient image restoration. In the last years, the shearlet transform has been proposed as an extension of the Wavelet transform to address curved singularities[84]. In particular, shearlets provide a directional, multiscale representation of images and an optimally sparse approximations of anisotropic features[25].

The shearlets have been used to solve different image restoration tasks[85, 86, 87, 24], including denoising[26, 24, 27]. However, these approaches assume the presence of only spatially-uniform Gaussian white noise. Here, we propose a novel microlocal domain thresholding approach based on the comprehensive noise model summarized in equation (23).

1.4.1 The shearlet transform

Wavelets provide optimally sparse representations for 1-D functions that are smooth away from point singularities. However, this does not hold true in higher dimensions[88]. Indeed, wavelets, due to their isotropic nature, are not ideally suited for covering anisotropic features such as curve-like singularities in the class of cartoon-like images. This class of images is a subset of $L_2(\mathbb{R}^2)$ functions introduced to provide a simplified model of natural images, which emphasizes anisotropic features and is consistent with many models of the human visual system and fluorescence images[24]. These functions have by now become a standard model for measuring sparse (nonlinear) approximation properties of directional representation systems.

Shearlets have been constructed with the aim of improving on this shortcoming by applying anisotropic scaling to the generating function, i.e. different dimensions can have different scaling factors.

Introducing directional selectivity to a system of representing functions requires the capability to vary the direction. In the theory of shearlets, this is achieved by applying a shearing operator along with the anisotropic scaling operator. Thus, shearlets are very similar to wavelets in the sense that both are constructed from generating functions that can be modified with a certain operator. However, while for wavelets only isotropic scaling is possible, shearlets are subject to both anisotropic scaling and shearing.

A significant problem of the shearlet system is that to capture horizontally aligned anisotropic structures one would have to apply the shearlet matrix a great number of times. To avoid this issue, which would be very difficult to handle in any digital implementation, the so-called cone-adapted continuous shearlet system was introduced[22, 89]. Here, the Fourier-domain is partitioned into four cones (two horizontal, two vertical) and a square-shaped low-pass region (Figure S17). The horizontal and vertical ones are now associated with their own generating functions ψ, $\tilde{\psi}$ while a scaling function ϕ covers the low-pass region.

Let

$$\tilde{A}_a = \begin{pmatrix} a^{1/2} & 0 \\ 0 & a \end{pmatrix} . \tag{28}$$

be another scaling matrix, shearlet generators $\psi, \tilde{\psi} \in L_2(\mathbb{R})$ and scaling function $\phi \in L_2(\mathbb{R})$ be given. The cone-adapted continuous shearlet system $\mathcal{SH}_{cont}(\phi, \psi, \tilde{\psi})$ is given by the union of the following sets:

$$\Phi = \{\phi_t = \phi(\cdot - t) : t \in \mathbb{R}^2\} \tag{29}$$

$$\Psi = \{\psi_{a,s,t} = a^{-\frac{3}{4}} \psi(A_a^{-1} S_s^{-1}(\cdot - t)) : a \in (0.1] \,,\, |s| \leqslant 1 + a^{1/2} \,,\, t \in \mathbb{R}^2\} \tag{30}$$

$$\tilde{\Psi} = \{\tilde{\psi}_{a,s,t} = a^{-\frac{3}{4}} \tilde{\psi}(A_a^{-1} S_s^{-T}(\cdot - t)) : a \in (0.1] \,,\, |s| \leqslant 1 + a^{1/2} \,,\, t \in \mathbb{R}^2\} \tag{31}$$

Interestingly, there exist compactly supported shearlet generators ψ,$\tilde{\psi}$ and a shearlet scaling function ϕ such that the best n-term approximation of functions f in the class of cartoon-like images obeys the relation:

$$\|f - f_n\|^2 = O(n^{-2} (\log n)^3)) \,, \tag{32}$$

which is the optimal decay rate achievable, besides the negligible log factor[25, 89]. This also gives a mathematical justification of the superiority of shearlet systems over wavelet or Fourier bases which cannot guarantee optimal decay rates. Noticeably, shearlets decompose an image in a joint space/orientation representation where points and curves are very different objects. So, even if points and curves may happen to overlap spatially, they are distinct microlocally, i.e., in the orientation direction (Figure S17c,d).

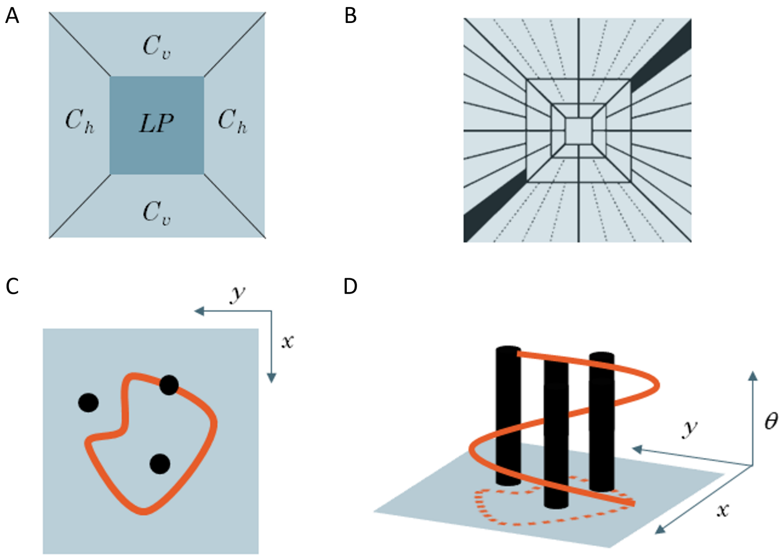

Figure S17: The shearlet transform and its microlocal properties. (A) In the cone-adapted continuous shearlet system the Fourier-domain is partitioned into four cones (two horizontal, C_h, and two vertical, C_v) and a square-shaped low-pass region (LP). (B) Frequency tiling of the cone-adapted shearlet system. This decomposes an image separating spatial features at different scales and orientations. Thus, different features such as lines and points that may appear to overlap spatially (C) happen to be microlocally distinct (D).

1.4.2 *Microlocal shearlet thresholding*

Once the pixel-to-pixel non-uniformity from the camera output has been removed according to equation (20), we want to find $\tilde{\mathbb{E}}[I(p,t)] \approx S(p.t)$. To do this, we propose a shearlet-based domain shrinkage method for microlocal hard thresholding:

$$\tilde{\mathbb{E}}_{\mathcal{S}}^{hard}[I(p,t)] = \mathcal{SH}^{-1}\left\{\check{I}(p,t)\cdot \mathbb{1}_{(|\check{I}(p,t)|>\mathcal{T}(p,t))}\right\}, \tag{33}$$

and soft thresholding:

$$\tilde{\mathbb{E}}_{\mathcal{S}}^{soft}[I(p,t)] = \mathcal{SH}^{-1}\left\{\mathrm{sign}(\check{I}(p,t))\cdot(\check{I}(p,t)-\mathcal{T}(p,t))\cdot\{I;p,t\}\cdot \mathbb{1}_{(|\check{I}(p,t)|>\mathcal{T}(p,t))}\right\}, \tag{34}$$

where the symbol $\mathcal{SH}\{\cdot\}$ represents the discrete shearlet transform operation, $\check{I}(p,t) = \mathcal{SH}\{I;p,t\}$ and:

$$\mathbb{1}_{(|\check{I}(p,t)|>\mathcal{T}(p,t))} = \begin{cases} 1 & \text{if } |\check{I}(p,t)| > \mathcal{T}(p,t) \\ 0 & \text{otherwise} \end{cases}. \tag{35}$$

In both cases, the microlocal thresholding factor $\mathcal{T}(p,t)$, according to equation (23), is:

$$\mathcal{T}(p,t) = h\cdot\sqrt{\tilde{\mathbb{E}}_0[I(p,t)] + \sigma_D^2 + \sigma_R^2}\,, \tag{36}$$

where h is a user-adjustable multiplicative factor that depends on the image SNR and the shearlet system while the estimate $\tilde{\mathbb{E}}_0[I(p,t)]$ is the interpolated median value calculated among the neighboring pixels within half Airy unit. The size of the Airy disk in pixel for the input images are calculated using the user-inserted values of wavelength of the incoming photons, λ, effective numerical aperture of the optical system, NA, and pixel size in the image plane, d_p.

The multiplicative factor h is proportional to the image SNR, which is automatically estimated via phase decorrelation[90]. In brief, this method measures the decorrelation of the phase of the Fourier transform in the input image to estimate its SNR and resolution. Unlike Fourier ring correlation, this method does not need image pairs so it can work also with single images and shows better accuracy. When processing multidimensional data, the estimate is usually generalizable to the whole dataset and by default the MIRO software estimates resolution and SNR only on an individual image. To account for possible SNR changes due, for example, to photobleaching, this behavior can be changed in the options. Moreover, the regularizing parameter h can also be adjusted manually in case of an occasional miscalculation of the SNR. Generally, this manual adjustment requires only a few iterations to reach the optimal value and, also in this case, the parameter can be adjusted simply on one image then used with the whole dataset.

The microlocal thresholding strategy alone already out-performs standard shearlet-thresholding denoising (Figure S18). However, when the SNR is very low, noise can generate patterns that are not completely removed by the thresholding operation. To avoid this issue, we pre-filter the input image $I(p,t)$ in the Fourier space to remove unwanted frequencies:

$$I'(p,t) = I(p,t) \otimes \mathcal{F}\{\Gamma\} + K(p,t) \cdot H(p,t)\,, \tag{37}$$

where:

$$\Gamma(f) = \frac{1}{1 + e^{6\left(\frac{f-f_c^*}{f_c-f_c^*}-1\right)}}\,, \tag{38}$$

$$H:\ \mathcal{F}\{H\} = \begin{cases} \mathcal{F}\{I\} \cdot \Gamma(f) & \text{if } f > f_c^* \\ 0 & \text{otherwise} \end{cases}\,, \tag{39}$$

$$K(p,t) = \begin{cases} k & \text{if } H(p,t) > 0 \\ 0 & \text{otherwise} \end{cases}\,, \tag{40}$$

$\mathcal{F}\{\cdot\}$ is the Fourier transform, f_c is the normalized theoretical cutoff frequency of the system as calculated from the system parameters, and f_c^* is the normalized effective cutoff frequency measured by phase decorrelation[90].

As mentioned at the beginning of this section, MIRO offers the option to operate either soft or hard shearlet thresholding. As shown in Figure S18, we did not observe major differences between the two implementations during the characterization measurements and, as a rule of thumb, we suggest using the soft thresholding method by default and switching to the hard thresholding only in case of low SNR.

The typical MIRO run time depends on the image size and the number of shearlet scales used, so the user can also indicate the maximum number of shearlet scales used in the decomposition and the algorithm will calculate the closer value given the image size. We generally opt for using the maximum available number of shearlet

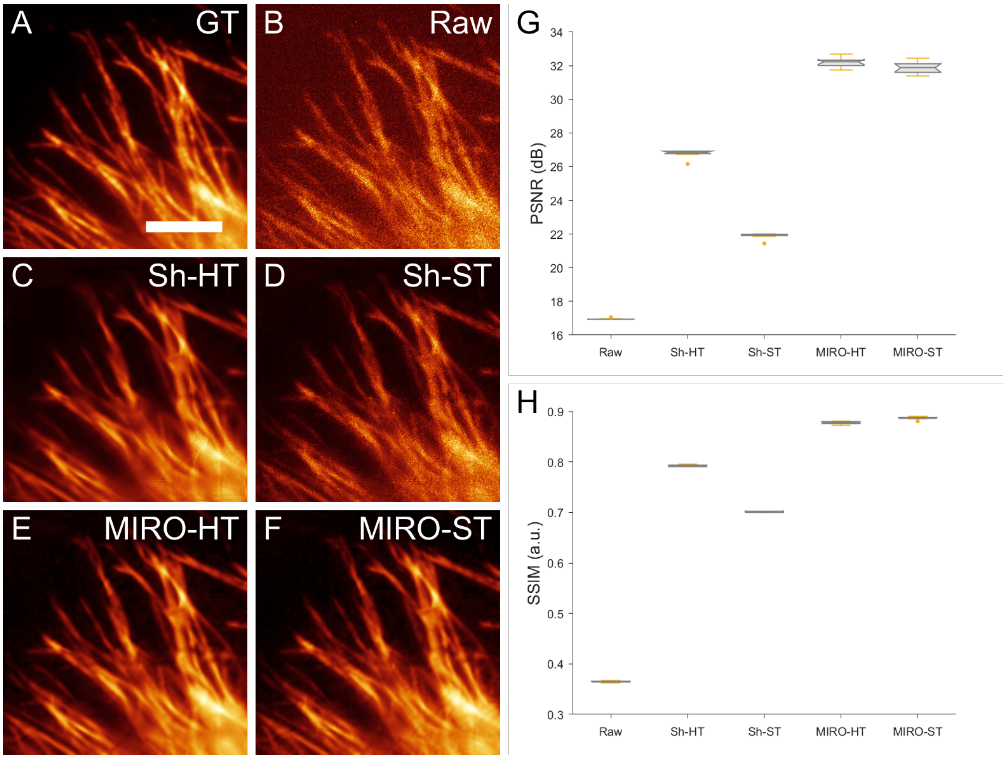

Figure S18: Comparison of shearlet-based thresholding strategies. Image of fluorescently-labeled microtubules in a HeLa cell acquired setting the exposure time to 100 ms (A) and 10 ms (B). Noise-corrected images of (B) using uniform Shearlet Hard-thresholding (C), uniform Shearlet Soft-thresholding (D), microlocal Shearlet Hard-thresholding (E), and microlocal Shearlet Soft-thresholding (F). (G,H) Box plots of the PSNR and SSIM of the images in (B-F) using the image in (A) as ground truth. In both cases, we see that microlocal thresholding outperforms the standard methods with similar performance between hard and soft thresholding versions. Scale bar: 5 µm.

components. Nonetheless, reducing the number of shearlet components considerably improves the processing speed so that using a smaller shearlet scale can be worth testing when processing big datasets.

Finally, it is possible to save runtime by feeding MIRO with a precalculated shearlet system (which can be reused with different datasets bacause depends only on the image size in x and y) so that this is not recalculated every time (Table S5). For an example on how to use this feature, see the demo code provided with the Supplementary Software.

Image size (px)	Shearlet scales	Runtime	Runtime (w/ precalculated shearlet base)
64x64	3	0.07s	0.04s
128x128	4	0.17s	0.10s
256x256	5	1.17s	0.36s

Table S5: MIRO runtime for different image sizes and Shearlet scales. The runtimes are provided both with and without the use of a precalculated shearlet base.

1.5 Additional image processing

As an optional final step, we implemented the possibility to leverage the image non-local similarities to enhance noise correction in case of very low SNR or non-optimal microlocal noise estimation. Real-world objects seldom present a random structure but usually exhibit repeated patterns or a certain degree of non-local similarities. This has been often exploited by image restoration algorithms with notable results also in the field of scientific imaging[19, 91, 92, 93]. In fluorescence microscopy, for example, only one type of structure is imaged at a time so that in principle a good degree of similarity is expected all across the image. Thus, we offer three possible options: unbiased non-local means processing, multiscale collaborative Wiener filtering, and time-based collaborative Wiener filtering.

1.5.1 Unbiased non-local means

The unbiased non-local means (NLM) option provides a fast and effective solution to lingering low intensity noise. NLM denoising of a given image patch is obtained as a weighted average of the surrounding patches, with weights proportional to the their similarity[94]. However, in the classic implementation of NLM, the estimate of the noise-free patch from a set of nearest neighbors is biased by a factor of $2\sigma^2$[95]. This can produce artifacts especially when using a reduced number of neighbors. To avoid this issue we use an unbiased estimate of the patch weights w[96]. Given a noisy reference patch μ_r with N elements, the squared distance from a noisy neighbor γ_k patch is:

$$\delta^2(\mu_r, \gamma_k) = \frac{1}{N} \sum_{i=0}^{N-1} (\mu_r^i - \gamma_k^i)^2 \,, \tag{41}$$

and the noise-free patch is calculated as:

$$\hat{\mu}_r(\mu_r) = \frac{\sum_k w_{\mu_r,\gamma_k} \cdot \gamma_k}{\sum_k w_{\mu_r,\gamma_k}} \tag{42}$$

where

$$w_{\mu_r,\gamma_k} = \frac{1}{h^2} \exp \max[0, \delta^2(\mu_r, \gamma_k) - 2\sigma^2] \,. \tag{43}$$

1.5.2 Multiscale collaborative Wiener filter

A significant advantage of image self-similarity is the possibility to stack small image patches in different groups according to their mutual similarity. This way, a two-dimensional image can be "folded" in a three-dimensional one, with the third dimension being represented by similarity. This higher dimensionality allows to use transforms that provide image representations with higher sparsity, which results in a more efficient image restoration[91].

Here, we calculated the image self-similarity at different scales to obtain a better similarity estimate and also to provide a pixel-wise estimation of the total noise variance according to the noise model in equation (23). Then, a collaborative Wiener filter is used to obtain the noise-free patches that are finally put together to form the restored image (Figure S19). Therefore, the process can be divided into three main parts: multiscale grouping, collaborative filtering and weighted aggregation.

Multiscale grouping. In collaborative filtering, the similarity between different patches is evaluated using Frobenius norm. However, this method becomes less reliable in presence of noise. Here, we use a more robust evaluation by measuring similarity at different scales[97]. We generate for each patch multiple scaled-down versions by Gaussian pyramid reduction and use these scaled-down patches to compute the similarity. Then, the original patches are grouped accordingly in three-dimensional stacks. Specifically, the algorithm scans the input image selecting $k \times k$ patches (P) and searching an $m \times m$ area for other similar patches (Q). Each patch is scaled down two or three times (depending on the patch size) using Gaussian average. These scaled-down patches re used to assess similarity by the normalized square distance in L_2

$$d(P, Q) = \sum_{j=1}^{3} \frac{\|P_j - Q_j\|_2^2}{k_j^2}, \tag{44}$$

where j is the order of the scaled-down patch. This evaluation is then used to stack the original (not scaled-down) patches in a 3D group, $\mathcal{P}[P]$, up to a maximum of 8. This is repeated with different reference patches until all the input images are stacked in 3D groups.

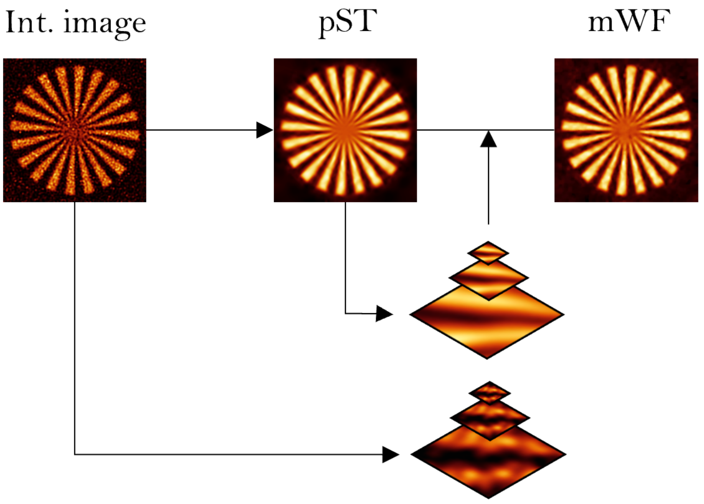

Figure S19: Schematic depiction of the application of multiscale Wiener filter to MIRO. We have implemented a multiscale patch-based Wiener filter to further improve the accuracy of MIRO noise correction in case of images with very low SNR. Firstly, both the intermediate image (Int. image) produced after FPN correction and the image resulting from pixel-wise shearlet thresholding (pST) are divided into patches. For each patch two of more sub-sampled copies are produced according to a Gaussian pyramid representation. Then, the sub-sampled (less noisy) copies are used to evaluate the similarity between patches in order to form groups of similar patches. Finally, these groups are fed into a Wiener filter and the outputs re-aggregated to form the final noise-free image (mWF).

Collaborative Wiener filter. At this point, collaborative filtering employs a 3D transform, τ_{3D}^{h}, to filter the coefficients of each group all at once. Due to the similarity between the patches, the 3D transform results in even sparser representation of the original patches than the 2D transforms whereas the noise still has a constant power spectrum[98]. Collaborative filtering can be expressed by the formula:

$$P^{w}(P) = \tau_{3D}^{w^{-1}}(c_P(\tau_{3D}^{w}(\mathcal{P}[P]))), \tag{45}$$

where the Wiener coefficient is:

$$c_P = \frac{|\tau_{3D}^{w}(\mathcal{P}[P])|^2}{|\tau_{3D}^{w}(\mathcal{P}[P])|^2 + \sigma_N^2}. \tag{46}$$

The noise variance σ_N^2 is the sum the the calibrated component σ_R^2 and the estimated shot-noise variance σ_S^2. The shot-noise variance is determined using a weighted non-local means (Kullback–Leibler divergence) calculated across the patch group:

$$d_{KL}(P, Q) = \frac{1}{2}(P - Q)\ln\frac{P}{Q}. \tag{47}$$

Weighted aggregation. Once the collaborative filtering produces all the patch estimates, the aggregation procedure returns the denoised patches to their original locations. Notably, during the grouping part, one patch can be assigned to more than one group. This implies that pixels can have more than one estimation. Thus, the final estimate is computed as a weighted average of all overlapping pixels where the weights are defined to prioritize homogeneous patches over the ones containing edges and corners:

$$w_P = \begin{cases} N_P^{-1} & N_P \geqslant 1 \\ 1 & \text{otherwise} \end{cases} \tag{48}$$

where N_P is the number of non-zero coefficients in the 3D block after filtering.

1.5.3 Time-based collaborative Wiener filter

When the input images are a video sequence, if the *fast Wiener filter* option is selected, the algorithm performs a time-based collaborative Wiener filter[99]. This is similar to the sparse filtering operation detailed in the previous paragraph with the difference being that during the grouping task the algorithm performs the patch search also in the neighboring frames. This means that the search area for similar patches this time is an $m \times m \times t$ spatio-temporal volume. This way, lingering noise can be further reduced taking advantage of the sample self-similarity, not only in space, but also in time.

To maximize the temporal correlation, during video filtering the search window is shifted according to a motion estimation of the reference patch in time. This is performed using a motion estimator with the sequence down-sampled by a factor of two and computing the motion trajectories using a fast diamond search[100]. Here, the trajectory of a block is defined by the sequence of the most similar blocks in the neighboring frames.

It is important to notice that this video filtering involves only an extension of the search window during the grouping phase and does not involve any time averaging, so that the temporal resolution is preserved.

2 CHARACTERIZATION

Before testing the performance of MIRO with the different sensor types and applications, we characterize the expected performance on both simulations and experimental caliber datasets. In particular, we review the expected accuracy and resolution improvement with the SNR, and the effect of digital sampling on noise removal. Also, we compare our results with the most common denoising algorithms in scientific imaging.

2.1 Accuracy

In fluorescence imaging, the performance of noise correction is intrinsically related to the photon flux of the input image. Indeed, sample details at low photon count start to be comparable to noise-generated random patterns directly affecting the image accuracy. Here, we assess the dependence of the denoising performance on the intensity of the emission light and the corresponding improvement of the image quality. We estimate the effect of MIRO denoising on image accuracy by measuring its mean squared error:

$$\mathrm{MSE}(I, I_{\mathrm{ideal}}) = \frac{1}{N_p} \sum_{p=1}^{N_p} |I(p) - I_{ideal}(p)|^2 \,, \tag{49}$$

the related peak signal-to-noise ratio:

$$\mathrm{PSNR}(I, I_{\mathrm{ideal}}) = 10 \log_{10} \left(\frac{I_{\mathrm{max}}^2}{\mathrm{MSE}(I, I_{\mathrm{ideal}})} \right) , \tag{50}$$

where N_p is the total number of pixels while I and I_{ideal} are the test and ideal image intensity in e^-, respectively. To assess image fidelity respect to image distortion, we used the structural similarity metric (SSIM):

$$\mathrm{SSIM}(X, Y) = \frac{(2\mu_X \mu_Y)\,(2\sigma_{XY})}{\left(\mu_X^2 + \mu_Y^2\right)\left(\sigma_X^2 + \sigma_Y^2\right)} \tag{51}$$

where μ_X, μ_y, σ_X, σ_Y, and σ_{XY} are the local means, standard deviations, and cross-covariance for images X and Y. To avoid errors given by the different dynamic range of low and high SNR images (used as ground truth), all images were normalized first:

$$N(I) = \frac{I - I_{\min}}{I_{\max} - I_{\min}} \,. \tag{52}$$

To begin with, we assessed the performance of MIRO image restoration using simulations which we generated using the ideal image of a Siemens star (Figure S20). We scaled the intensity of the ideal image to simulate different illumination conditions. Then, we added shot, read and fixed-pattern noise (FPN) according to equations (12) and (13). In addition, to show the effect of FPN, we used the mean gain and offset values instead of the pixel-wise calibrated values when converting the raw images in photoelectrons.

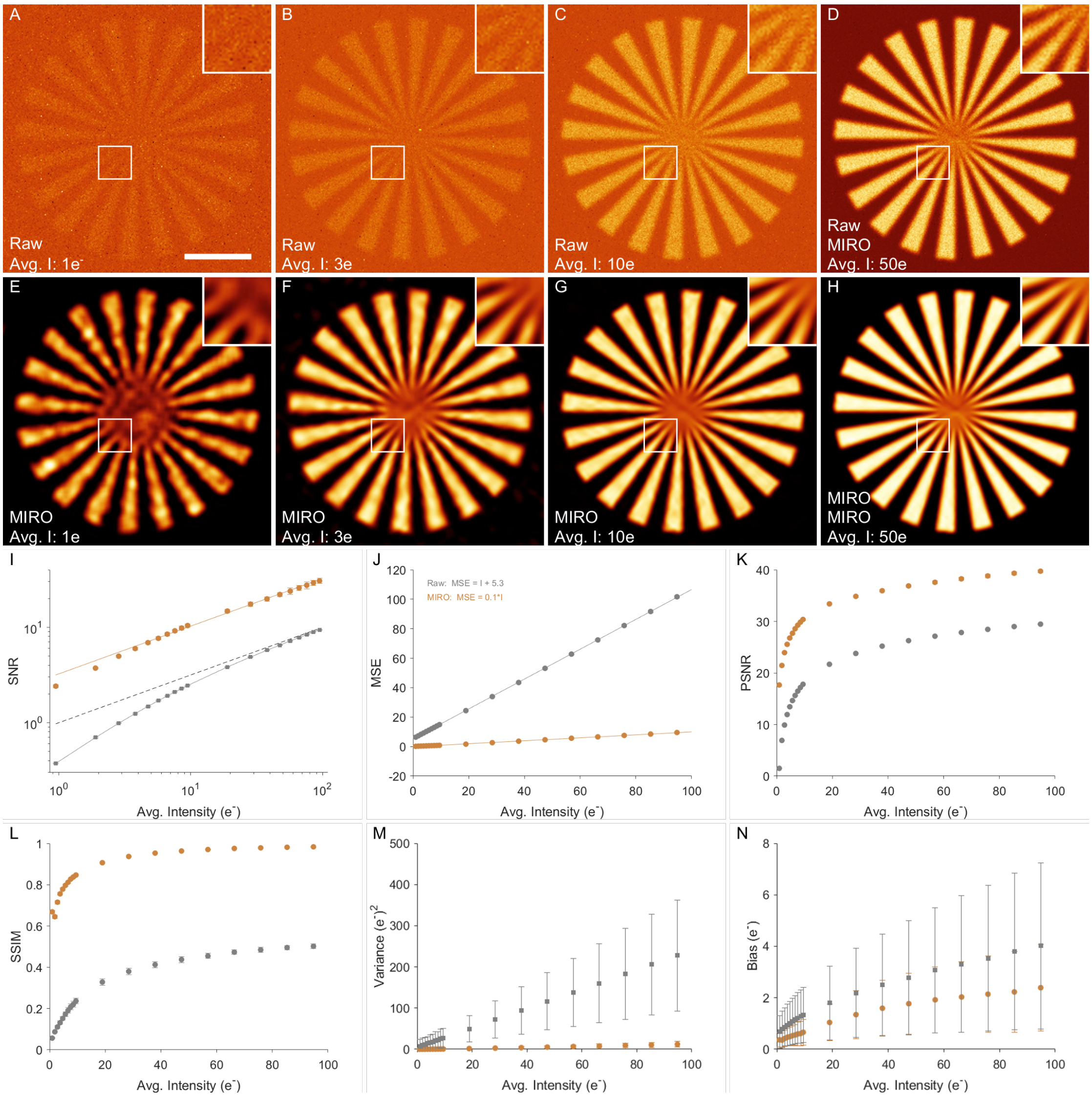

Figure S20: Characterization of MIRO noise correction using simulated data. We generated several simulated images of a Siemens star at different signal-to-noise ratios (SNRs) to estimate the efficiency of MIRO processing. For each SNR condition, we simulated a star generating a certain number of photo-electrons on a camera sensor affected by FPN, signal shot noise, dark shot noise, and readout noise (A-D). (E-H) MIRO-processed images relative to (A-D). (I) Plot of the simulated camera response in terms of SNR with the generated photo-electrons (gray) compared to an ideal camera response (dashed line). After MIRO noise correction, the SNR remains consistently above the ideal-camera scenario (yellow). Measuring the image mean-square error (MSE) at different light conditions, it is possible to estimate the image accuracy and separate the effect of Poisson and Gaussian noise components (J). (K,L) Plots of the peak SNR (PSNR) and structural similarity index measure (SSIM) calculated between ideal and noisy simulations before (gray) and after (yellow) noise correction. (M,N) Plots of the reduction of temporal variance and bias between noisy and denoised simulations calculated in the pixels corresponding to the star. Scale bar: 4 μm.

In all conditions, we observed a super-ideal noise correction with an average improvement of 90% compared to a camera working in the (ideal) Poisson regime. The average quality of the denoised images remains essentially stable also for images with less than 10 photons per pixel, with super-ideal results even when the photon flux is less than 3 photons per pixel. We have obtained similar results with experimental data using standard fluorescent beads imaged in different SNR conditions (Figure S21). The image recovery trend and the average improvement (94%) are in agreement with the simulations, suggesting that MIRO can offer a robust noise correction down to a few photons per pixel in both simulations and experimental data.

However, when the photon budget is very low, it is challenging to generate a representation where the signal is sparser than noisy pixels. Therefore, since a realistic sample is not uniform and the emission photon flux in fluorescence microscopy can vary depending on both the power of the excitation light and the specific fluorophores observed, we further tested MIRO performance in different practical cases of general interest.

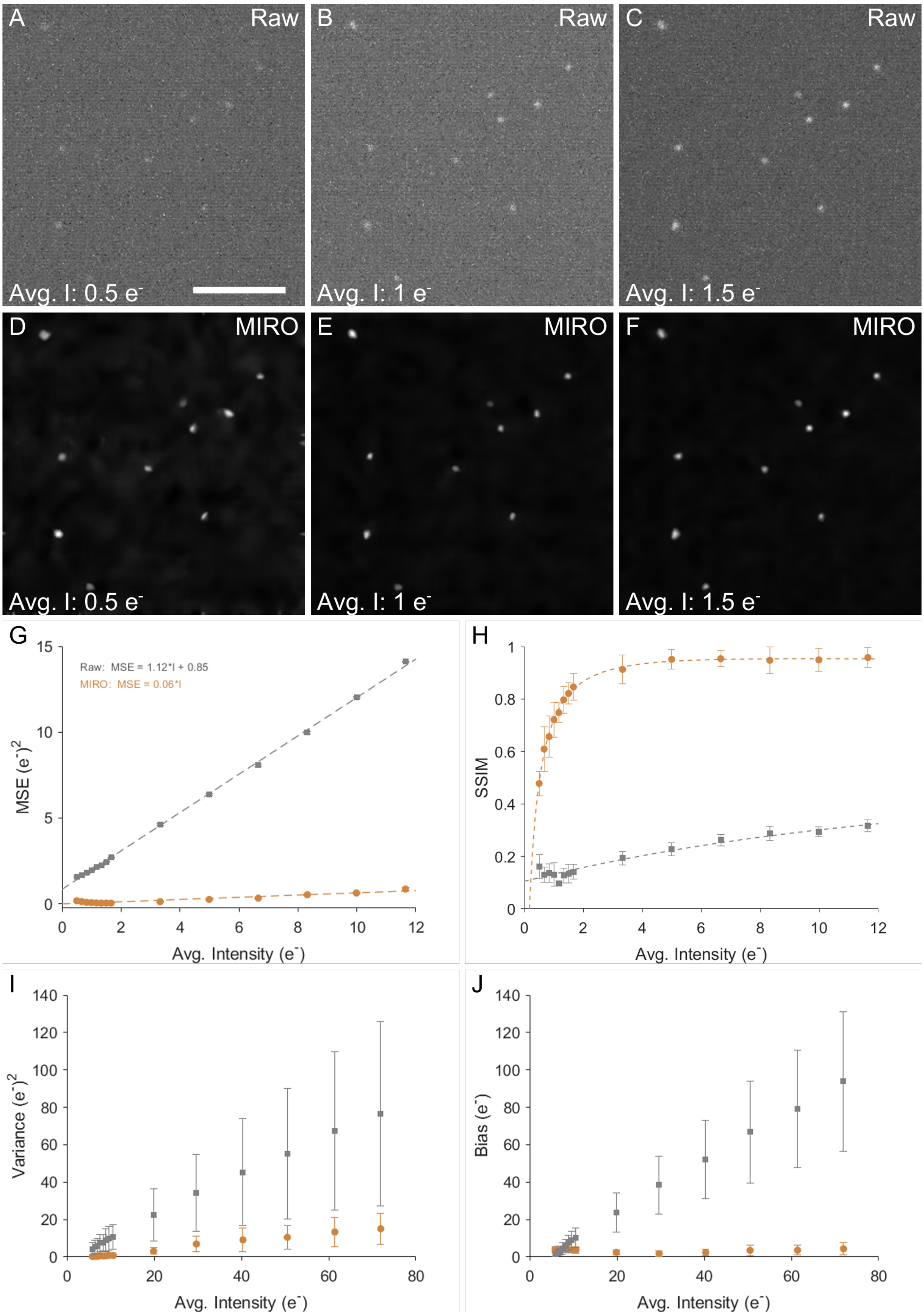

Figure S21: Characterization of MIRO noise correction using experimental data. We recorded several images of fluorescent beads at different illumination conditions (A-C) to validate the efficiency of MIRO processing. (D-F) MIRO-processed images relative to (A-C). (G) Plot of the image MSE compared to a ground-truth image generated by averaging 100 frames acquired at a high SNR. Both the correction of Poisson and Gaussian noise components are in agreement with the simulated data. (H) Plot of the SSIM calculated between the ground truth and the individual frames before (gray) and after (yellow) noise correction. (I,J) Plots showing the reduction of temporal variance and bias between noisy and denoised data calculated in the pixels corresponding to the fluorescent beads. Scale bar: 5 μm.

2.2 Resolution

Noise removal can enhance resolution by improving image accuracy and sharpness. On the flip side, inaccurate image denoising could potentially cancel the fine details of the sample. For instance, low-pass filtering represents a classic example where thorough noise correction can come at the expenses of resolution. We have tested MIRO by measuring the image resolution before and after noise correction in both simulated (Figure S22) and experimental single-molecule data (Figure S23).

To measure resolution in simulated and experimental data, we used Gaussian fitting to evaluate the width of the point-spread function (PSF) and detected no additional blurring after the denoising process. For fair comparison between noisy and corrected data, the intensity profiles were generated by averaging over the vertical dimension of each single molecule. Finally, we further tested the effect of MIRO processing on noisy images recorded from fluorescently-labeled microtubules (Figure S24). The results confirmed that the algorithm does not induce any loss of resolution compared to the original image. On the contrary, we observed the recovery of the system's ideal resolution even in low SNR conditions.

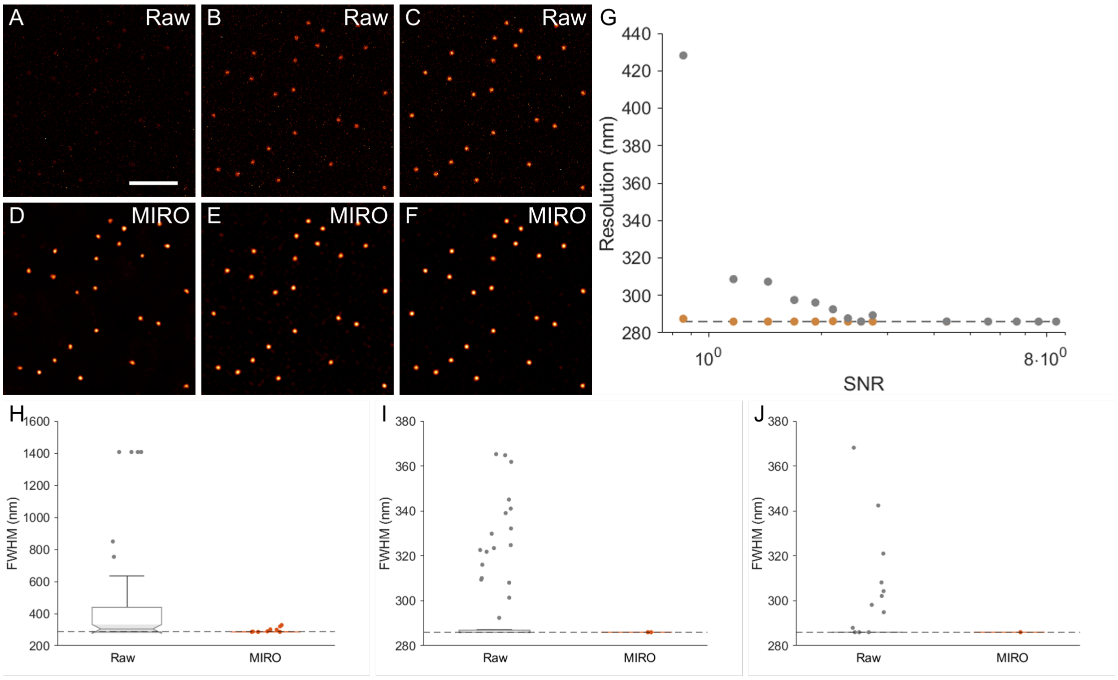

Figure S22: Characterization of the effect of MIRO processing on image resolution using simulated data. We simulated noisy images of fluorescent beads at different SNR (A-C) and processed them with MIRO (D-F). (G) Using two-dimensional Gaussian fitting, we estimated the image resolution as the average full-width half-maximum (FWHM) of the simulated beads before (gray) and after denoising (yellow). We observed an overall agreement with the expected resolution (dashed line). (H-J) The distribution of the measured FWHM of the three representative datasets in (A-C, gray) and (D-F, red). It is noticeable a greatly reduced number of outliers in the processed images. Scale bar: 4 µm.

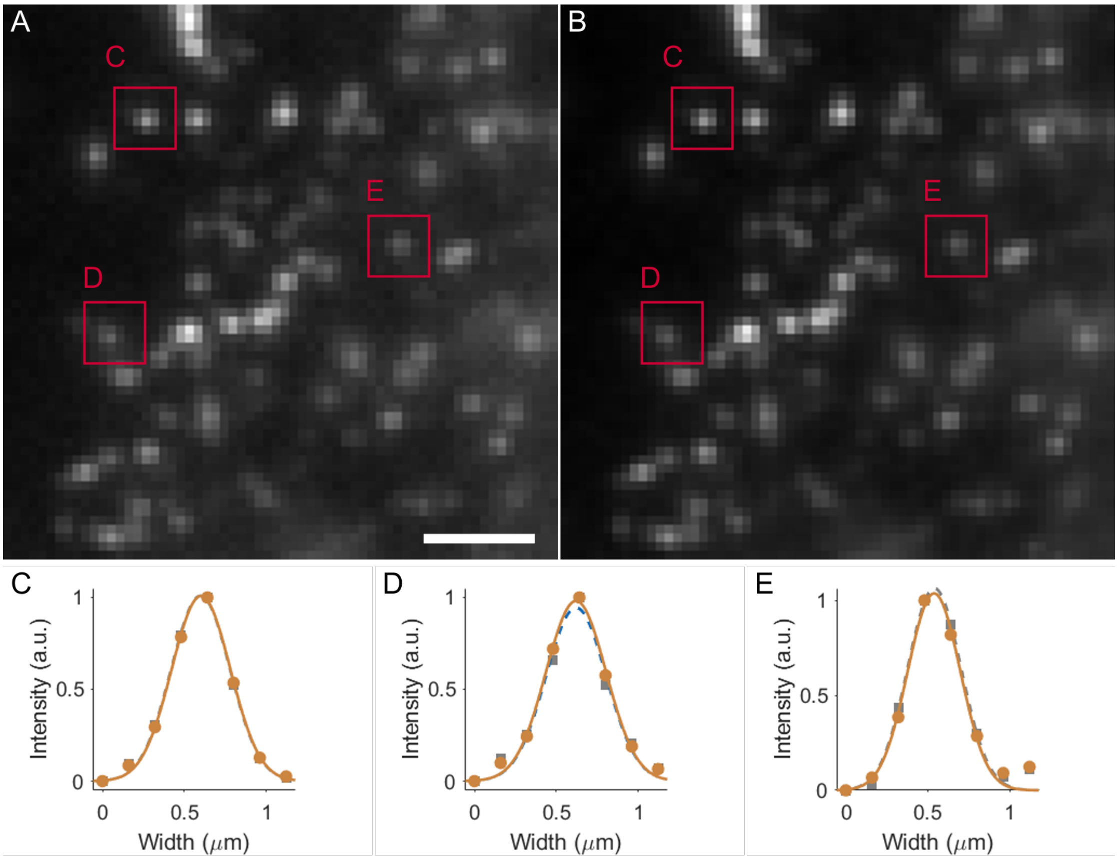

Figure S23: Image resolution after MIRO noise correction on diffraction-limited spots. We imaged fluorescent dyes using a microscope equipped with an EMCCD camera (A) and processed the resulting image with MIRO (B). (C-E) Comparing the profile of three representative diffraction-limited spots, we did not see smoothing related to noise correction. All profiles were calculated by Gaussian fitting of raw data (gray, dashed lines) and MIRO processed images (yellow, solid lines). To minimize the effect of noise in raw data, we calculated each profile as the average along the vertical direction of the boxed areas in (A) and (B). Scale bar: 2 μm.

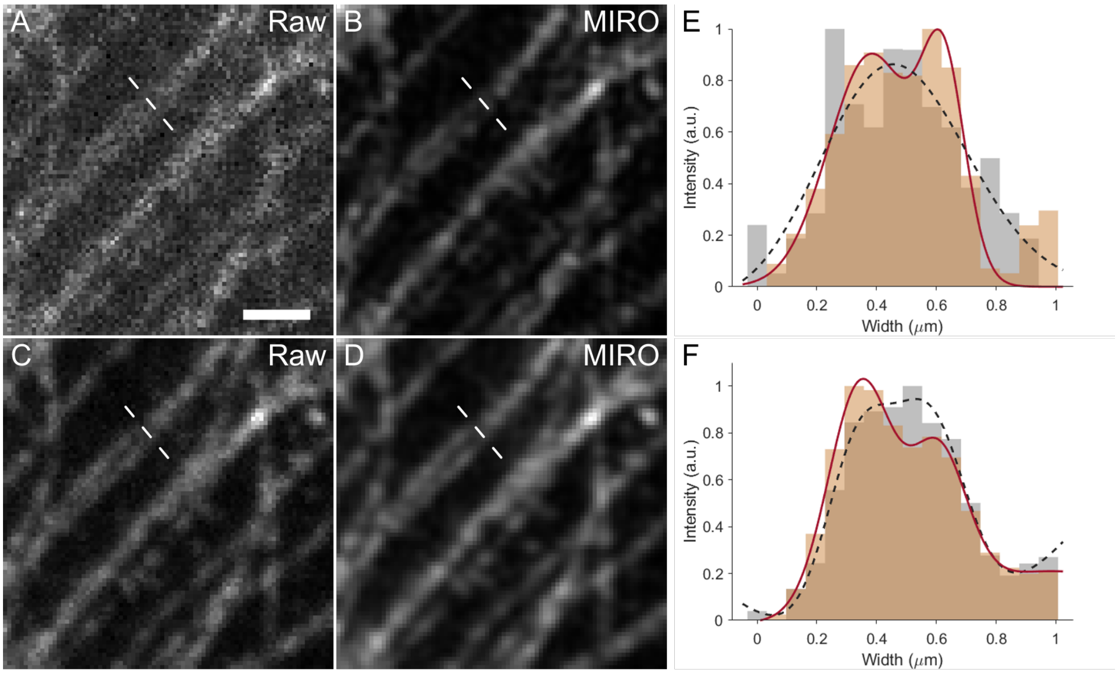

Figure S24: Image resolution after MIRO noise correction on diffraction-limited microtubules imaged at low illumination power. We recorded fluorescently-labeled microtubules in a BPAE cell using a microscope equipped with sCMOS camera. We imaged the sample setting the integration time at 10 ms (A) first and then at 100 ms (C). We processed the resulting images with MIRO (B) and (D), respectively. (E) The cross-sectional profiles corresponding to the dashed lines in (A, gray bars) and (B, red bars). Both data were fitted assuming the presence of two Gaussian profiles. However, only the fit on the MIRO data (red solid line) shows two resolved Gaussians. (F) The cross-sectional profiles corresponding to the dashed line in (C, gray bars) and (D, red bars). Despite the higher SNR of the raw image, the two microtubules are not yet well resolved in the raw data (dashed gray line). Scale bar: 1 μm.

2.3 Sampling

The sampling frequency of the camera pixels should be strictly higher than the Nyquist frequency in order to have a diffraction-limited optical system[101]. Thus, the pixel size calculated in the image plane should be preferably no bigger than half of the diffraction limit. However, this is not necessarily the case for practical imaging systems, so we investigated the performance of our algorithm under various sampling conditions (Figure S25). We define the sampling rate as the ratio between the system resolution, given by the Rayleigh criterion ($0.61\lambda/NA$) and the effective pixel size of the image (physical sensor pixel size divided by the magnification), so that the minimum sampling rate according to the Nyquist criterion should be greater than 2. Low sampling can affect noise correction by reducing the sparsity of the acquired image. Digital under-sampling, i.e., pixel binning, is sometimes used when retaining temporal resolution may be more important than spatial resolution. Usually, pixel binning yields an improvement of the SNR that counterbalance the loss of sparsity. However, this is not true when using a low magnification objective and the spatial information is traded in favor of field-of-view. In this case, the under-sampling does not necessarily involve a gain in SNR because the readout noise is not reduced by pixel binning and low-magnification objectives usually have lower light collection capability. So, we tested MIRO performance with differently sampled images while maintaining a comparable SNR to evaluate the effect of the sampling alone and we did not observe any significant correlation between noise correction and image sampling (Figure S26). In summary, experimental data together with the results obtained by the simulations support the viability of MIRO denoising at the sampling rates normally adopted for most fluorescence microscopy. A sampling rate as close as possible to the Nyquist limit is generally recommended as a good trade-off between SNR and detail preservation. Nonetheless, MIRO noise correction works consistently even on undersampled images (Figure S27). Thus, the upper limit for the range of acceptable sampling rates is mainly related to the photon flux necessary to maintain a feasible SNR of the input data while the lower limit is determined by the cost in terms of loss of details.

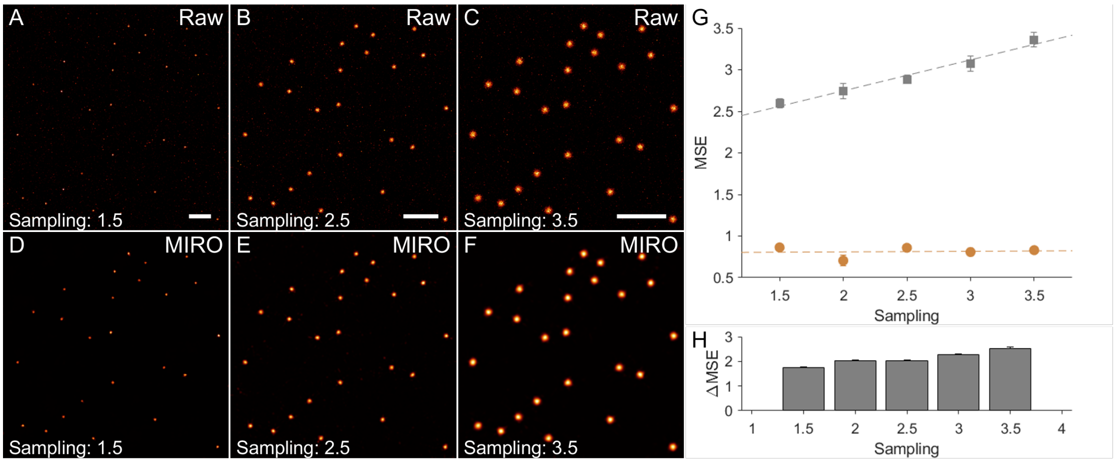

Figure S25: Simulation of the effect of sampling on MIRO processing. We simulated noisy images of fluorescent beads at different sampling rates (A-C) and processed them with MIRO (D-F). (G) Plotting the image MSE for the different sampling rates before (gray) and after denoising (yellow), we did not observe any significant change in the efficiency of MIRO processing. At higher sampling rates, the improvement in image quality is slightly higher (H) due to the fact that at low sampling rates the SNR increases. Scale bars: 4 μm.

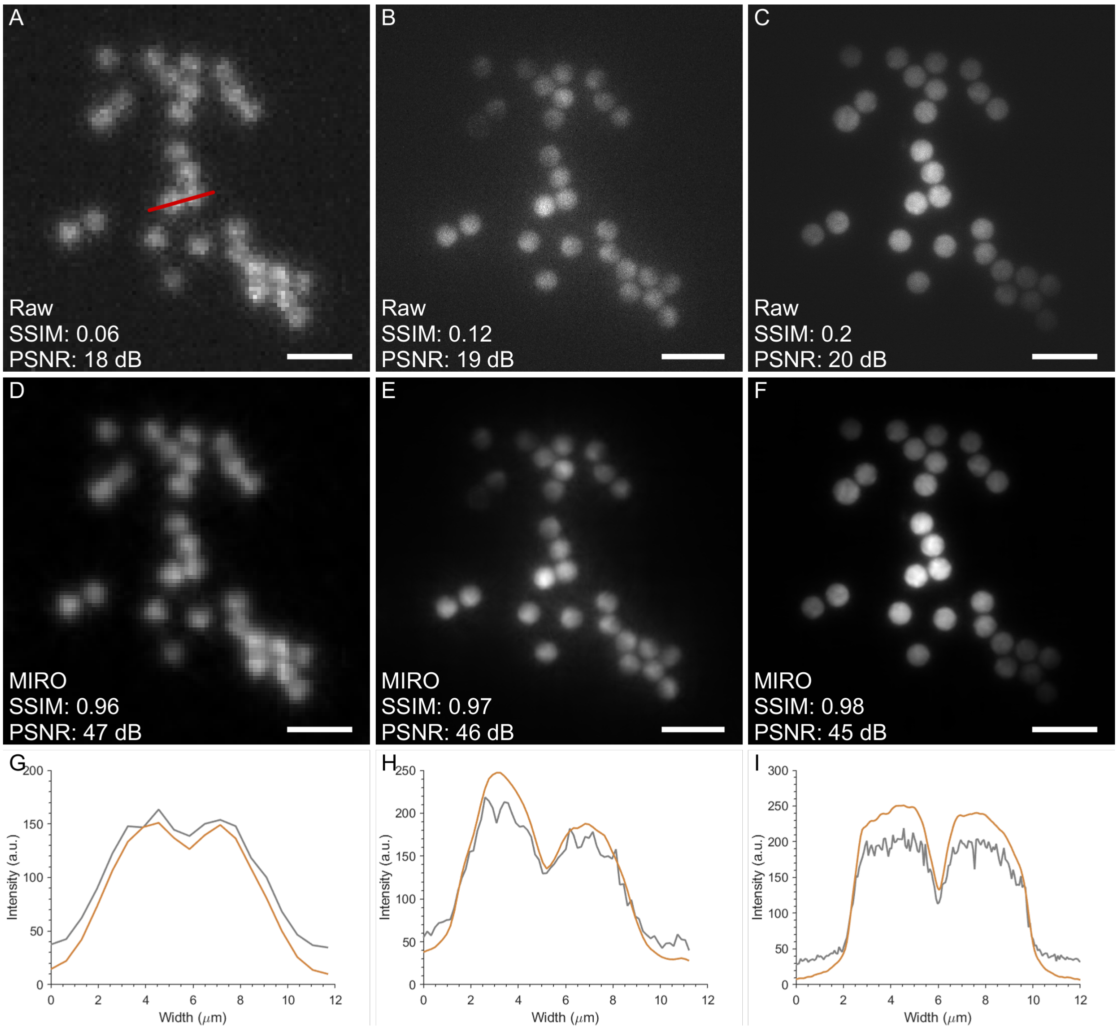

Figure S26: Experimental characterization of the effect of sampling rate on MIRO processing. We imaged fluorescent beads using objectives with different magnification and numerical aperture, namely: 10x/0.3 (A, sampling rate 1.7) 20x/0.5 (B, sampling rate 2.1), and 40x/0.75 (C, sampling rate 2.5). To measure the effect of sampling rate on image restoration in the same SNR conditions, each time we adjusted the illumination to keep a comparable number of impinging photons per pixel. (D-F) MIRO processed images of the frames in (A-C), respectively. Also in this case, we do not observe any significant effect of sampling on noise correction. For each sampling rate, PSNR and SSIM values were measured using as ground truth the average of 100 raw frames. (G-I) cross-sectional profiles corresponding to the line in (A) and measured in (A-C, gray lines) and (D-F, yellow lines). Scale bars: 10 µm.

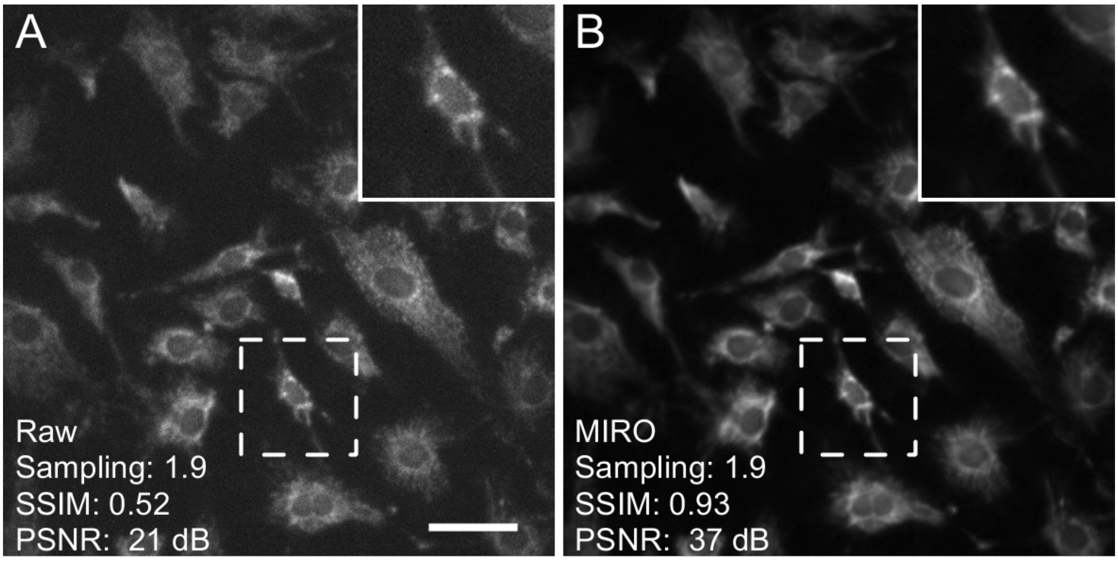

Figure S27: Noise correction of sub-sampled images. As a further test, we imaged fluorescently-labeled mitochondria in BPAE cells at a sub-Nyquist sampling rate (A) and processed it with MIRO (B). The PSNR and SSIM values were measured using as ground truth the average of 100 raw frames. Scale bar: 50 μm.

2.4 Comparison with state-of-the-art

Biomedical images have a significant amount of variability because of the variety of specimens and imaging conditions available. Thus, it is important to establish if the intensity variance seen in measurements is the result of biological variability or measurement error. Therefore, it is necessary to individuate algorithms that can reduce the inaccuracy and imprecision of the acquisition process with minimal assumptions about the sample itself. Here, we compare our denoising approach based on the connections between physical and mathematical properties of image formation with the most commonly-used alternatives.

2.4.1 General-purpose denoising methods

We compared MIRO with the most common general-purpose denoising algorithms used in scientific imaging. We used TetraSpeck$^{\mathrm{TM}}$ microspheres as caliber data (Figure S28) and fluorescently labeled microtubules as biological sample (Figure S29). The Gaussian filter is likely the easiest and most used denoising method in imaging. This is likely due to the fact that as long as the standard deviation of the smoothing kernel is comparable to the PSF radius, it is possible to obtain a good improvement with little or no resolution loss. A very common alternative is the non-local means algorithm[94], which tries to take advantage of the high degree of redundancy of natural images. This is essentially a regularity assumption actually more general than most local filters. It also generalizes a periodicity assumption of the image. Similarly, block matching and 3d filtering (BM3D) extends this concept to entire pixel blocks and combines it with Wiener filtering[91]. While the original algorithm have been extensively adopted in the processing of natural images, it assumes the presence of only spatially-uniform noise and the knowledge of noise variance. Thus, here we compared our algorithm with Poisson BM3D (pBM3D), a generalized version of BM3D implemented to correct mixed Gaussian-Poisson noise and capable to automatically estimate noise variance[102]. It is interesting to notice how the performance of each algorithm vary with noise fingerprint. For instance, non-local filtering methods such as NLM or pBM3D sometimes fail to correct high-frequency noise while Gaussian filtering tends to leave a low-frequency noise mark. MIRO, instead, displayed an overall stable performance at all frequencies and generally better noise correction.

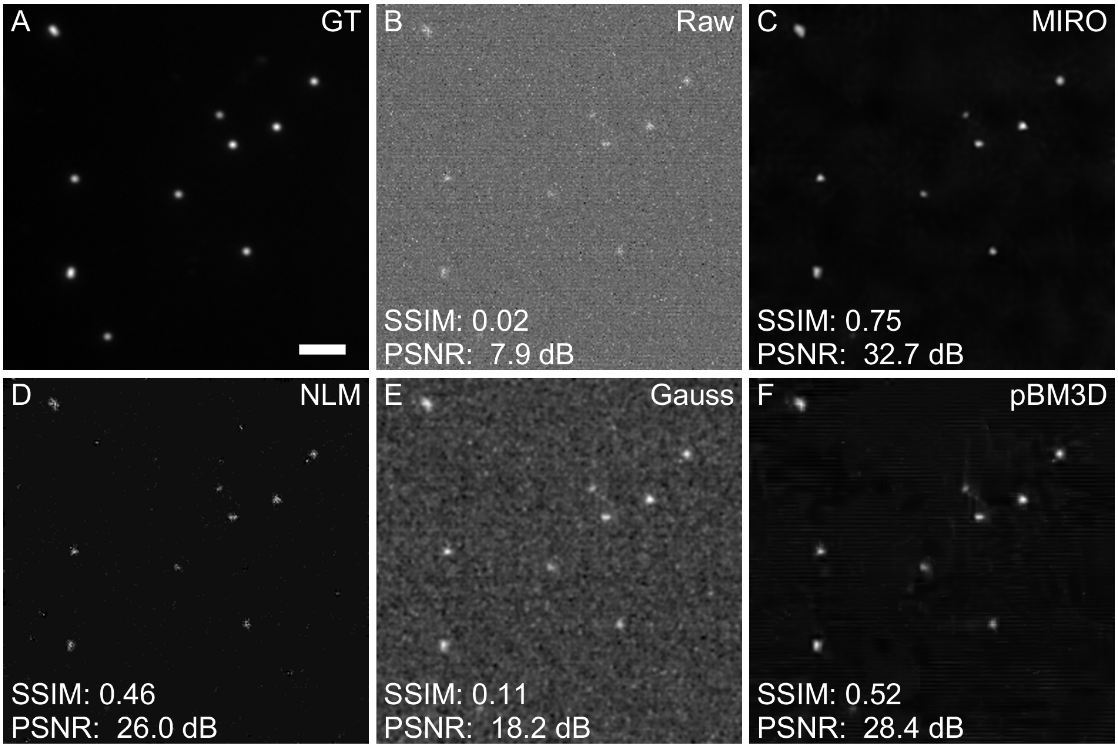

Figure S28: Comparison of noise correction algorithms using a sparse sample. We used an acquisition of sub-diffraction limited beads taken at high (A) and low (B) SNRs to compare the performance of common denoising strategies. We observed that MIRO noise correction (C) delivers the best performance overall both in terms of SSIM and PSNR. Instead, non-local means (NLM) correction fails to remove high-frequency noise (D) while Gaussian filtering (Gauss, $\sigma = 1$ pixel) delivers the worst performance due to the lingering low-frequency noise (E). Poisson-BM3D (pBM3D), while performing better than NLM and Gauss, still features artifacts apparently related to FPN and lingering high-frequency noise in general (F). Scale bar: 2 μm.

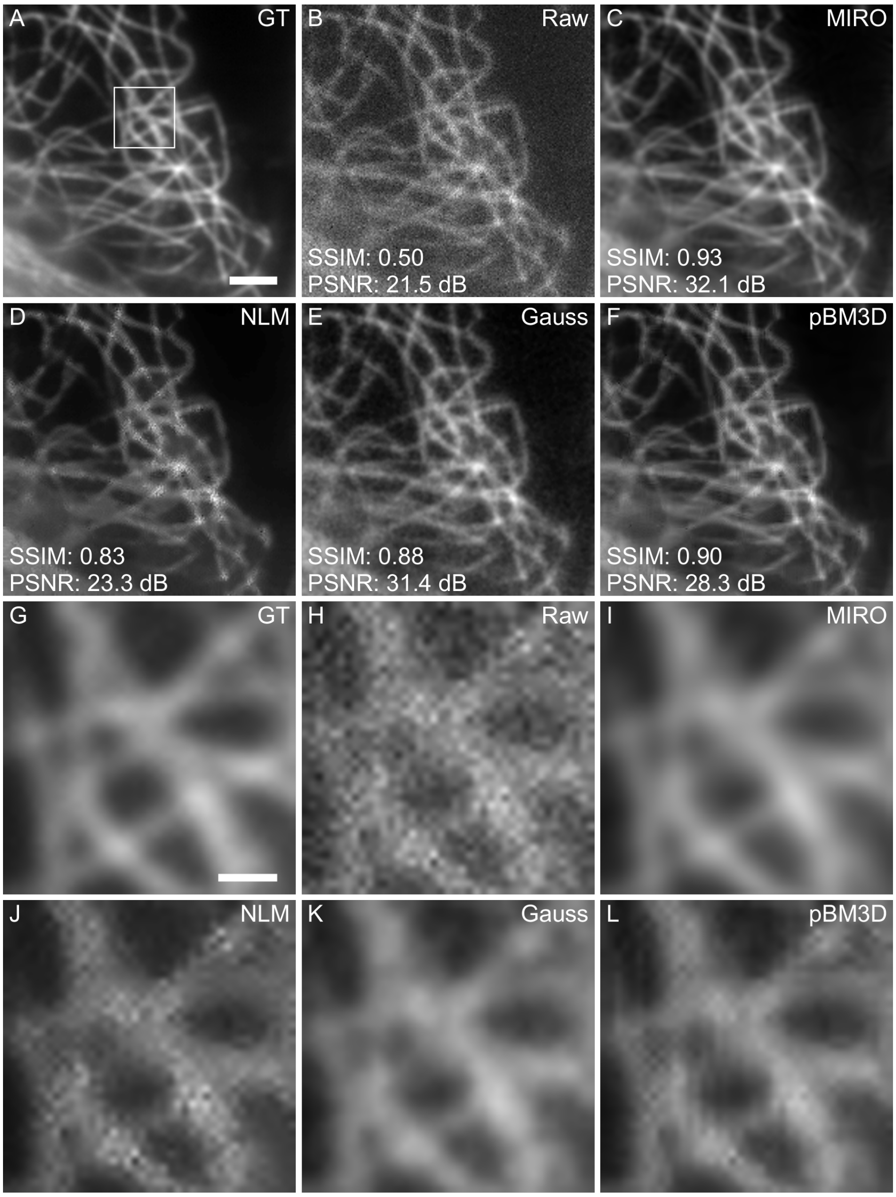

Figure S29: Comparison of noise correction algorithms using a dense biological sample. We used an acquisition of fluorescently-labeled microtubules taken at high (A) and low (B) SNRs to compare the performance of common denoising strategies. MIRO noise correction (C) delivers the best performance overall both in terms of SSIM and PSNR while the other methods' performances are mainly hindered by the presence of their method noise fingerprints. This consists in the presence of lingering high-frequency noise for NLM (D), low-frequency noise for Gauss (E), and structured and high-frequency noise for pBM3D (F). Such issues can be clearly noticed in the zoomed-in images of the boxed area shown in (G-L). Scale bars: 2 μm (A), 500 nm (G).

2.4.2 *sCMOS-based noise correction*

Recently, scientific CMOS camera have gained enormous popularity in optical microscopy. This prompted the development of specific noise correction algorithms for this specific detector architecture[18, 19]. The main common feature of these different algorithms is the combination of physical information about the imaging system within the mathematical processing. A first algorithm proposed by Liu et al.[18] addressed the noise problem in sCMOS cameras using a method composed of two main steps: low-pass filtering followed by the correction of pixel fluctuation via the iterative minimization of a negative log-likelihood function. Both low-pass filtering and iterative minimization are based on the knowledge of pre-recorded calibration data. Later, Mandracchia et al.[19] proposed an alternative non-iterative approach leveraging both spatial and temporal self-similarity of biological structures to recover the underlying clean signal. This marked a noticeable improvement in both accuracy and runtime. Overall, both methods showed a good tolerance for spatially-distributed noise. Nonetheless, they become more unreliable when the imaging conditions depart from the assumption that the camera is working in Gaussian-regime, i.e. noise distribution is mostly uniform.

This limitation results in a less-than-optimal noise correction in many practical cases. We evaluated this using both sCMOS-based simulations and experimental data (Figure S30). Furthermore, given that ACsN can use temporal information in video/timelapse to improve the image restoration, we tested it also using video correction (Figure S31).

In all cases, NCS performed relatively poorly compared to both MIRO and ACsN. Probably, this is due to the fact that the algorithm focuses more on the correction of pixel-level temporal fluctuations than spatial features. ACsN showed superior performances especially in the video mode (which we indicated as ACsN.v). Nonetheless, MIRO performed better than all the other methods. Notably, MIRO performed better than ACsN.v even without the need for any temporal information (Figure S31). This highlights the important improvement given by optimal sparsity and local noise estimation in image restoration even when using predetermined system information, e.g. calibration data, which is even more clear when processing images acquired with detectors such as EMCCDs (Figure S32) or PMTs (Figure S33).

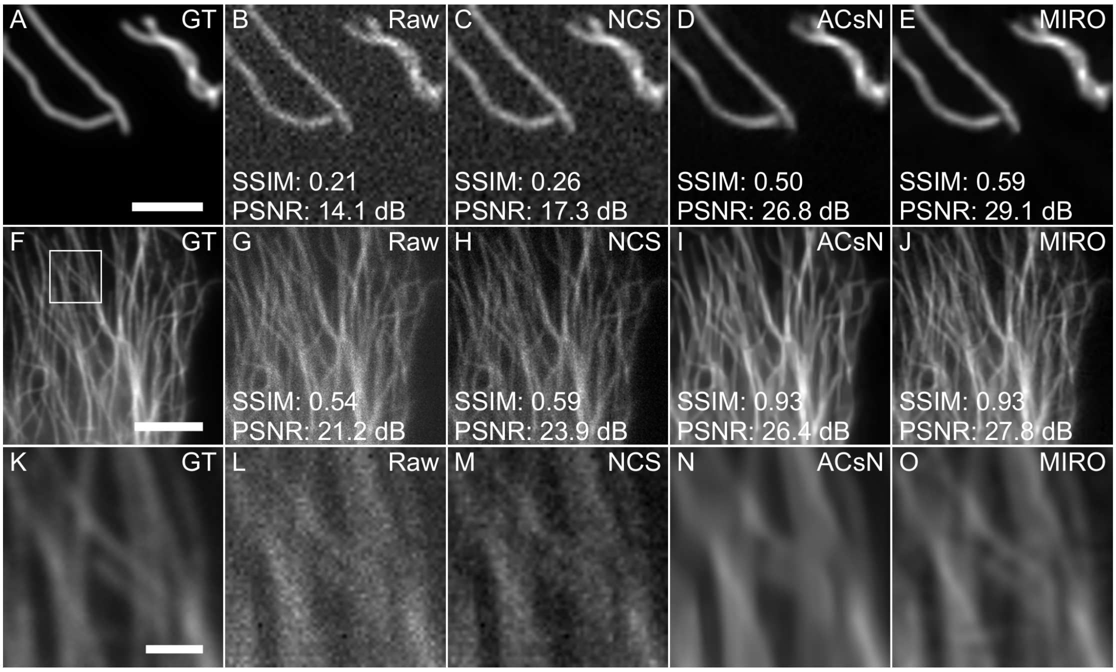

Figure S30: Comparison of sCMOS-specific noise correction methods. For this comparison, we used both simulated (A, B) and experimental images (G, H) of microtubule filaments. We compared the performance of the most recent denoising methods implemented specifically for sCMOS cameras and thoroughly validated for bioimaging, namely: NCS (C,H) and ACsN in single-frame (D, I) mode. We observed that MIRO noise correction (F, J) appears to perform overall better. (K-O) Zoomed regions corresponding to the boxed area in (F) relative to (F-J), respectively. Here, it possible to notice the higher improvement of MIRO in terms of resolution recovery when compared to the other techniques. Scale bars: 2 µm (A), 5 µm (G), 1 µm (K).

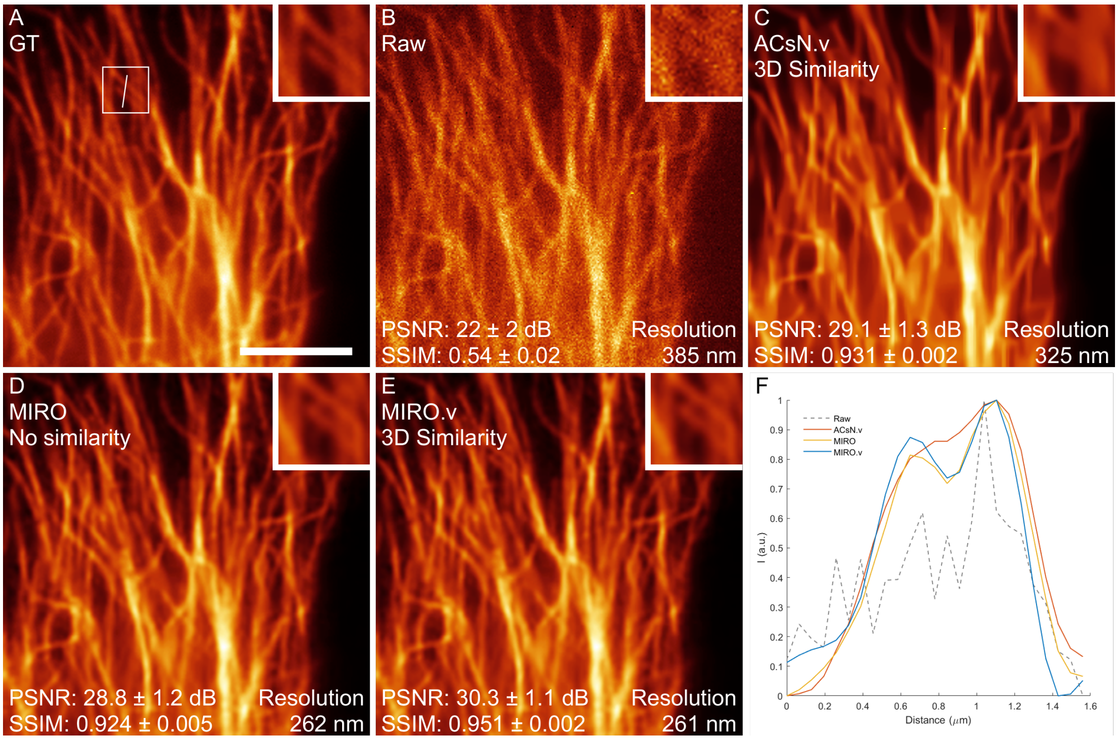

Figure S31: Comparison of sCMOS-specific noise correction methods using 3D similarity. (A,B) fluorescently-labeled microtubules acquired with a sCMOS-equipped widefield microscope. Here, the image in (B) is representative of a temporal stack of 100 frames that has been processed either with ACsN using 3D similarity (C), MIRO using no similarity (D), and MIRO with 3D similarity (E). Clearly, MIRO achieves a better performance than ACsN.v when using 3D similarity. Nonetheless, MIRO denoising with no similarity still achieves comparable SSIM and PSNR values and definitely better results in terms of resolution recovery (F). The effective resolution was calculated using phase decorrelation[90] Scale bar: 5 μm (A).

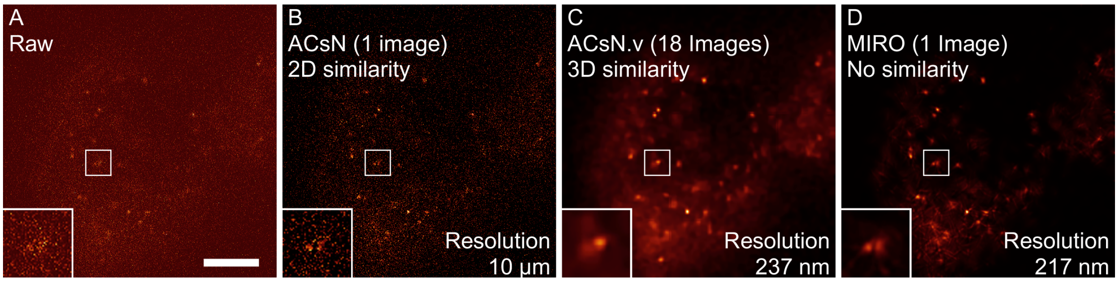

Figure S32: Comparison of ACsN and MIRO noise correction of EMCCD images. (A) Image of fluorescently-labeled peroxisomes in HEK cells acquired using an EMCCD camera. (B,C) ACsN-processed image using 2D (single frame) and 3D (18 frames) similarity. It is clear from the images how ACsN restoration does not estimate noise correctly even when using 3D similarity because its model does not consider the spatial variations of noise. (D) MIRO, on the other hand, estimates noise pixel by pixel delivering optimal results already from single-frame reconstructions. The effective resolution was calculated using phase decorrelation[90]. Scale bar: 5 µm (A).

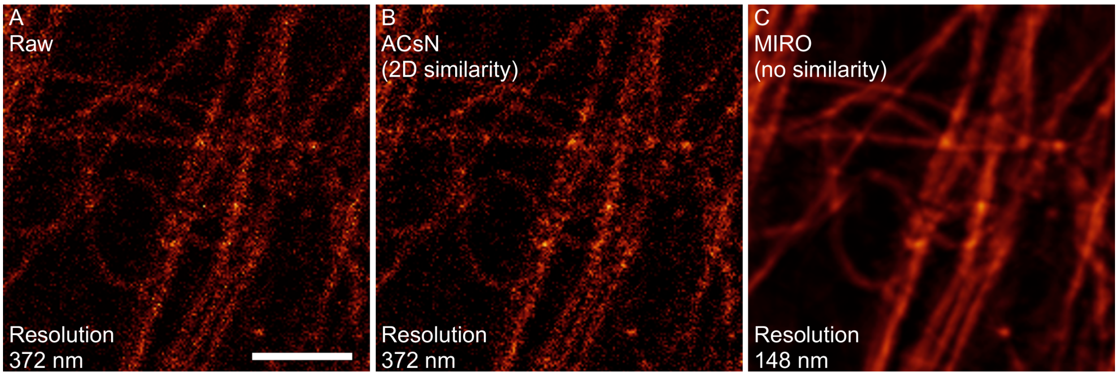

Figure S33: Comparison of ACsN and MIRO noise correction for PMT-based imaging. (A) Microtubules in BPAE cells acquired using a confocal microscope equipped with a PMT and the pinhole set at 0.25 AU. (B,C) noise correction of (A) using ACsN and MIRO, respectively. Image self-similarity used in ACsN does not compensate for the incompatibility of its noise estimation with PMTs. The effective resolution was calculated using phase decorrelation[90]. Scale bar: 2 µm (A).

3 HYBRID IMAGE RESTORATION

3.1 Image restoration using Deep Neural Networks

The aim of image restoration is finding the best possible estimation of the clean image $f(x, y)$ from the degraded observation

$$g(x, y) = f(x, y) + \Delta(x, y)\,, \tag{53}$$

where $\Delta(x, y)$ represents the image degradation. This is generally an ill-posed inverse problem since there can be several potential clean images f for each g.

Residual learning attempts to solve this problem by directly learning the degradation function $\Delta(x, y)$ using deep convolutional networks with skip connections[103]. In presence of a sufficiently large set of pairs (f, g) of noise-free and corrupted versions of the same image, this approach is generally easier to optimize than identity mapping and often allows deeper networks to be used[104].

Convolutional neural networks (CNNs) have gained a significant role in image processing thanks to the impressive results obtained for image classification[105, 106], segmentation[107] and restoration problems[57, 58]. CNNs are built by stacking different layers that interleave convolutions with non-linear activation functions. Usually, the convolutions have small filters (e.g., $3 \times 3 \times c$) that are learned during the training phase. The activation functions, instead, are applied per-pixel and are often fixed. Using this structure, CNNs are able to capture within their inner layers, also called feature layers, relevant structural features that are then used to recover the local image content (Figure S34a).

To increase the efficiency of CNNs, these are often implemented employing an encoder-decoder architecture with skip connections such as U-Net. In this architecture, the network is divided in two parts. In the first part, the encoder, the internal feature layers are compressed by reducing their spatial resolution with pooling layers. These pooling layers perform downsampling by dividing the input into rectangular pooling regions, then computing typically the maximum or the average of each region. Changing the scale enables the network to study features of different sizes permitting to learn more high-level (non-local) features of the image and detect persistent properties across scales. To avoid a possible loss of resolution due to the downsampling, the maps generated by the high-resolution feature layers are retained using skip connections that join them to the subsequent decoder part. Here, upsampling layers are used to concatenate the feature maps generated at the different scales and eventually yield an output with the same resolution as the input (Figure S34b).

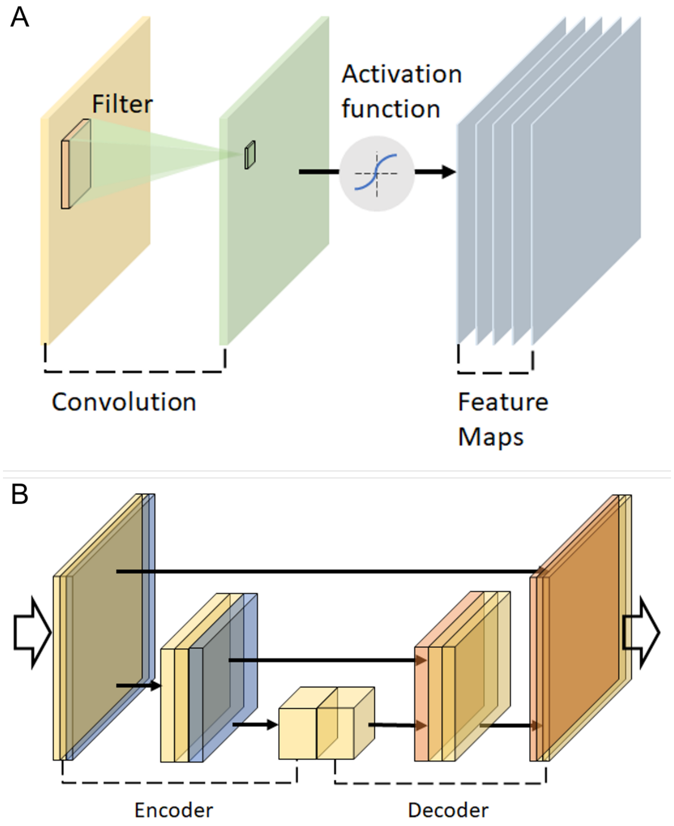

Figure S34: Schematic representation of a U-Net network. (A) A convolutional layer generally performs one or more convolutions on an input image (yellow). Each convolution is performed using a different filter, which is learned during the training phase. The resulting image of each convolution (green) is then processed on a pixel basis by an activation function. All the resulting images, known as feature maps, are finally concatenated and sent to the next layer. (B) A U-Net is formed mostly by convolutional layers (yellow) arranged in an encoder and a decoder part. The encoder part is divided in stages where the input image is progressively shrunk in order to individuate features at different scales. To do so, each stage but the last ends with a pooling layer (blue) that reduces the size of the feature maps of a fixed amount. In the decoder part instead, feature maps are progressively enlarged in order to return output of the same size as the input. For this reason, transposed convolution layers (orange) are placed before each stage but the first. In residual networks, skip connections (black arrows) are used to force the network learn the degradation function instead of the clean image.

3.2 Shearlet decomposition improves learning-based restoration

A CNN can compute complicated functions of the local image content by simply interleaving convolutional layers and activation functions. However, its capability to achieve satisfying results depends on both its architecture and the training dataset. The more complicated the function to compute, the more complicated (deeper) will be the architecture and bigger the training dataset.

A CNN is defined via the parameters θ of all its layers, which are typically chosen by minimizing a loss function $L(\theta)$. The function $L(\theta)$ assigns a cost to each prediction $g_\theta(f^t)$ performed on the couple (f^t, g^t) of the training dataset.

The whole training process is divided in epochs. During each epoch the training data are partitioned into random subsets, also called batches, and computed one by one. Each time a batch is computed, the parameters are updated based on a stochastic gradient descent guided by a predefined step size or learning rate.

CNNs have generally a reduced number of parameters in comparison with networks with fully connected layers. Nonetheless, a deep CNN with many layers will still require the optimization of many parameters, which can greatly increase the complexity of the training process. Indeed, the presence of a great number of parameters makes it more complicated for the gradient descent to reach the global minimum. Usually, the training process is fully data-driven because the CNN learns end-to-end a suitable representation of the data for a specific the task. Thus, a more complex CNN also requires a much bigger dataset to improve the overall performance of the training.

Lately, hybrid methods have been proposed in order to achieve more complicated tasks without increasing the complexity of the network. The idea of hybrid methods is to incorporate *a-priori* information from the task to perform in order to ease the learning task and, at the same time, ensure solutions that fit with the desired model[68]. This can be generally achieved by preprocessing the input images using an appropriate representation system. Importantly, when processing images, the representation system needs to encode not only the position but also the orientation. Thus, the incorporation of a transformation that changes the orientation becomes necessary. The system of shearlets is presumably to date the most widely used system of this type. Shearlets can optimally represent images governed by curvilinear singularities as key features so they can be used as a preprocessing step for CNNs. This type of approach has already been used for various tasks[108, 109, 110], showing a significant advantage over other methods that solve the problem using a pure data-driven approach.

Here, we validate that by just the addition of a shearlet transformation, we can obtain an improvement of the parameters' optimization, reflected by lower values of the validation loss, and a faster convergence. For this test, we have used a probabilistic U-Net architecture that we trained both with and without shearlet preprocessing. Here, the preprocessing consisted simply in a shearlet decomposition of the input images so that the different shearlet components were fed to the network as different channels. As training and test datasets, we used demo data that were already tested to work well with the U-Net[57] and both networks were trained using the same images.

3.3 Integration of Deep Learning within the MIRO framework

Since shearlet decomposition can clearly improve learning-based image restoration, we reasoned that the implementation of hybrid method based on the MIRO framework would be even more effective. To this end, we preprocess the input images by removing FPN and applying a shearlet transformation. This way, we use system-based preprocessing and shearlets to encode sample information in the input dataset for the neural network.

To further integrate the U-Net architecture within the MIRO framework, we have built an alternative architecture that we termed MIROnet. Here, we added an additional convolutional layer paired with an activation function that thresholds each element of the convolution. Overall, we have implemented these layers in such a way as to mimic the pixel-wise soft thresholding step in MIRO algorithm. Indeed, the convolutional layer scans the image with $3 \times 3 \times s$ filters, where s is the number of shearlets components, learning to estimate the appropriate threshold value that will be used then in the pixel-wise activation function. The resulting image is then passed to the U-Net, which in turn replaces the function of the non-local similarity-based denoising step.

As expected, we observed a further improvement of the parameters' optimization ($\sim$ 7%) and a faster convergence ($\sim$ 45%), see Table S6. To compare the U-Net and MIROnet test datasets, we used the SSIM and PSNR metrics. In particular, we have calculated these metrics using both standard normalization (SSIM, PSNR), see (52), and

Network	Min. Loss	Epoch
U-Net	0.0368	22
Shearlet + U-Net	0.0354	17
MIROnet	0.0343	12

Table S6: Training performance of the standard U-Net architecture, U-Net with shearlet preprocessing (Shearlet + U-Net), and the MIRO-modified hybrid U-Net (MIROnet). The shearlet-preprocessed U-Net has the same architecture of the standard U-Net, the difference being that images are fed to the network after shearlet decomposition. The different components given by the shearlet transform are passed to the network as different channels of the same image. Instead, MIROnet features an additional convolutional layer that mimics microlocal thresholding.

percentile normalization (SSIM′, PSNR′), as described in [57]. In both cases, MIROnet outperforms U-Net scores. More importantly, a closer look at the test dataset reveals the correction of recurring artifacts that appear in the reconstruction obtained with U-Net. These artifacts could be related to the presence of structured noise and the fact that the CNN may see this as a feature of the image. In our network, instead, the FPN preprocessing removes this recurring correlated noise facilitating the learning process (Figure S35).

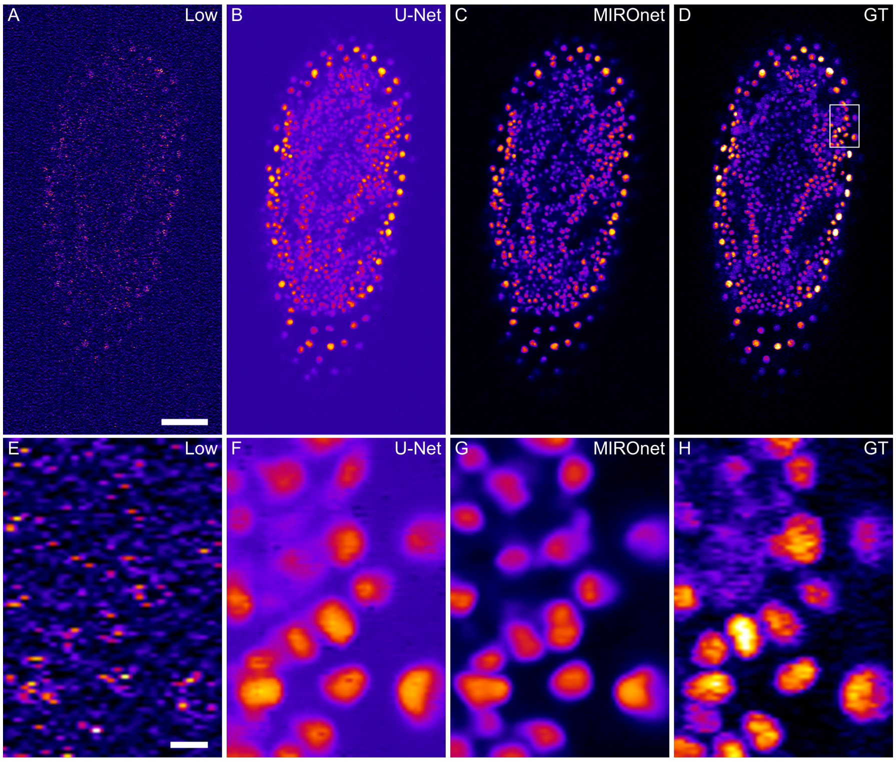

Figure S35: Comparison of MIRO-modified hybrid U-Net (MIROnet) with standard U-Net. Representative slice of the raw stack provided by Weigert et al. [57] and used as input to the networks (A). The relative output of U-Net and MIROnet are shown in (B) and (C), respectively. The ground truth is shown in (D). The image stack restored using MIROnet outperformed U-Net in both SSIM and PSNR scores. (E-H) Zoom-ins corresponding to the boxed area in (D) taken from the slices shown in (A-D), respectively. Scale bars: 50 µm (A), 5 µm (E).

Metric	Raw	U-Net	MIROnet
SSIM	0.23	0.17	0.79
PSNR (dB)	20.9	18.1	27
SSIM′	0.05	0.37	0.84
PSNR′ (dB)	18.2	25.8	26.1

Table S7: Performance comparison of U-Net and MIROnet architectures using the test dataset. SSIM′ and PSNR′ denote the metrics evaluated using percentile normalization.

REFERENCES AND NOTES

1. A. P.-T. Jost, J. C. Waters, Designing a rigorous microscopy experiment: Validating methods and avoiding bias. *J. Cell Biol.* **218**, 1452–1466 (2019).

2. N. Scherf, J. Huisken, The smart and gentle microscope. *Nat. Biotechnol.* **33**, 815–818 (2015).

3. A. R. Hibbs, G. MacDonald, K. Garsha, "Practical confocal microscopy" in *Handbook Of Biological Confocal Microscopy*, J. B. Pawley Ed. (Springer, 2006), chap. 36, pp. 650–671.

4. J. Icha, M. Weber, J. C. Waters, C. Norden, Phototoxicity in live fluorescence microscopy, and how to avoid it. *Bioessays* **39**, 1700003 (2017).

5. P. P. Laissue, R. A. Alghamdi, P. Tomancak, E. G. Reynaud, H. Shroff, Assessing phototoxicity in live fluorescence imaging. *Nat. Methods* **14**, 657–661 (2017).

6. J. C. Waters, Accuracy and precision in quantitative fluorescence microscopy. *J. Cell Biol.* **185**, 1135–1148 (2009).

7. J. R. Janesick, *Photon Transfer* (SPIE, 2007).

8. A. Buades, B. Coll, J.-M. Morel, Image Denoising Methods. A New Nonlocal Principle. *SIAM Review* **52**, 113–147 (2010).

9. K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, "Image denoising with block-matching and 3D filtering" in *Image Processing: Algorithms and Systems, Neural Networks, and Machine Learning*, E. R. Dougherty, J. Y. Astola, K. O. Egiazarian, N. M. Nasrabadi, S. A. Rizvi, Eds. (SPIE, 2006), vol. 6064, p. 606414.

10. R. C. Gonzalez, R. E. Woods, B. R. Masters, Digital image processing (Pearson, 2008).

11. L. Azzari, A. Foi, Variance stabilization for Noisy+estimate combination in iterative poisson denoising. *IEEE Signal Process. Lett.* **23**, 1086–1090 (2016).

12. A. A. Bindilatti, N. D. A. Mascarenhas, A nonlocal poisson denoising algorithm based on stochastic distances. *IEEE Signal Process. Lett.* **20**, 1010–1013 (2013).

13. F. Luisier, T. Blu, M. Unser, Image denoising in mixed Poisson–Gaussian noise. *IEEE Trans. Image Process.* **20**, 696–708 (2011).

14. W. Meiniel, J.-C. Olivo-Marin, E. D. Angelini, Denoising of microscopy images: A review of the state-of-the-art, and a new sparsity-based method. *IEEE Trans. Image Process.* **27**, 3842–3856 (2018).

15. J. M. Murray, "Practical Aspects of Quantitative Confocal Microscopy" in *Digital Microscopy,* G. Sluder, D. E. Wolf, Eds. (Academic Press, 2007), pp. 467–478.

16. F. Huang, T. M. Hartwich, F. E. Rivera-Molina, Y. Lin, W. C. Duim, J. J. Long, P. D. Uchil, J. R. Myers, M. A. Baird, W. Mothes, M. W. Davidson, D. Toomre, J. Bewersdorf, Video-rate nanoscopy using sCMOS camera-specific single-molecule localization algorithms. *Nat. Methods* **10**, 653–658 (2013).

17. P. Coupé, M. Munz, J. V. Manjon, E. S. Ruthazer, D. L. Collins, A CANDLE for a deeper in vivo insight. *Med. Image Anal.* **16**, 849–864 (2012).

18. S. Liu, M. J. Mlodzianoski, Z. Hu, Y. Ren, K. McElmurry, D. M. Suter, F. Huang, sCMOS noise-correction algorithm for microscopy images. *Nat. Methods* **14**, 760–761 (2017).

19. B. Mandracchia, X. Hua, C. Guo, J. Son, T. Urner, S. Jia, Fast and accurate sCMOS noise correction for fluorescence microscopy. *Nat. Commun.* **11**, 94 (2020).

20. J. He, Y. Cai, J. Wu, Q. Dai, Spatial-temporal low-rank prior for low-light volumetric fluorescence imaging. *Opt. Express* **29**, 40721–40733 (2021).

21. G. Kutyniok, W.-Q. Lim, Compactly supported shearlets are optimally sparse. *J. Approx. Theory* **163**, 1564–1589 (2011).

22. W. Q. Lim, The discrete shearlet transform: a new directional transform and compactly supported shearlet frames. *IEEE Trans. Image Process.* **19**, 1166–1180 (2010).

23. G. R. Easley, D. Labate, “Image processing using shearlets” in *Shearlets*, G. Kutyniok, D. Labate, Eds. (Springer, 2012), pp. 283–325.

24. D. Labate, F. Laezza, P. Negi, B. Ozcan, M. Papadakis, Efficient processing of fluorescence images using directional multiscale representations. *Math. Model Nat. Phenom.* **9**, 177–193 (2014).

25. K. Guo, D. Labate, Optimally sparse multidimensional representation using shearlets. *SIAM J. Math. Anal.* **39**, 298–318 (2007).

26. G. R. Easley, D. Labate, F. Colonna, Shearlet-based total variation diffusion for denoising. *IEEE Trans. Image Process.* **18**, 260–268 (2009).

27. H. R. Shahdoosti, O. Khayat, Image denoising using sparse representation classification and non-subsampled shearlet transform. *Signal, Image and Video Process.* **10**, 1081–1087 (2016).

28. M. M. Frigault, J. Lacoste, J. L. Swift, C. M. Brown, Live-cell microscopy - tips and tools. *J. Cell Sci.* **122**, 753–767 (2009).

29. A. Ettinger, T. Wittmann, Fluorescence live cell imaging. *Methods Cell Biol.* **123**, 77–94 (2014).

30. A. E. Y. T. Lefebvre, D. Ma, K. Kessenbrock, D. A. Lawson, M. A. Digman, Automated segmentation and tracking of mitochondria in live-cell time-lapse images. *Nat. Methods* **18**, 1091–1102 (2021).

31. J. A. Conchello, J. W. Lichtman, Optical sectioning microscopy. *Nat. Methods* **2**, 920–931 (2005).

32. J. Pawley, *Handbook of Biological Confocal Microscopy* (Science & Business Media, 2006), vol. 236.

33. J. Jonkman, C. M. Brown, G. D. Wright, K. I. Anderson, A. J. North, Tutorial: Guidance for quantitative confocal microscopy. *Nat. Protoc.* **15**, 1585–1611 (2020).

34. C. J. Sheppard, S. B. Mehta, R. Heintzmann, Superresolution by image scanning microscopy using pixel reassignment. *Opt. Lett.* **38**, 2889–2892 (2013).

35. C. B. Müller, J. Enderlein, Image scanning microscopy. *Phys. Rev. Lett.* **104**, 198101 (2010).

36. Y. Wu, H. Shroff, Faster, sharper, and deeper: Structured illumination microscopy for biological imaging. *Nat. Methods* **15**, 1011–1019 (2018).

37. J. M. Heddleston, J. S. Aaron, S. Khuon, T.-L. Chew, A guide to accurate reporting in digital image acquisition - can anyone replicate your microscopy data? *J. Cell Sci.* **134**, jcs254144 (2021).

38. C. Karras, M. Smedh, R. Förster, H. Deschout, J. Fernandez-Rodriguez, R. Heintzmann, Successful optimization of reconstruction parameters in structured illumination microscopy – A practical guide. *Opt. Commun.* **436**, 69–75 (2019).

39. M. Guo, Y. Li, Y. Su, T. Lambert, D. D. Nogare, M. W. Moyle, L. H. Duncan, R. Ikegami, A. Santella, I. Rey-Suarez, D. Green, A. Beiriger, J. Chen, H. Vishwasrao, S. Ganesan, V. Prince, J. C. Waters, C. M. Annunziata, M. Hafner, W. A. Mohler, A. B. Chitnis, A. Upadhyaya, T. B. Usdin, Z. Bao, D. Colón-Ramos, P. La Riviere, H. Liu, Y. Wu, H. Shroff, Rapid image deconvolution and multiview fusion for optical microscopy. *Nat. Biotechnol.* **38**, 1337–1346 (2020).

40. N. Gustafsson, S. Culley, G. Ashdown, D. M. Owen, P. M. Pereira, R. Henriques, Fast live-cell conventional fluorophore nanoscopy with ImageJ through super-resolution radial fluctuations. *Nat. Commun.* **7**, 12471 (2016).

41. G. Ball, J. Demmerle, R. Kaufmann, I. Davis, I. M. Dobbie, L. Schermelleh, SIMcheck: A toolbox for successful super-resolution structured illumination microscopy. *Sci. Rep.* **5**, 15915 (2015).

42. C. S. Smith, J. A. Slotman, L. Schermelleh, N. Chakrova, S. Hari, Y. Vos, C. W. Hagen, M. Müller, W. van Cappellen, A. B. Houtsmuller, J. P. Hoogenboom, S. Stallinga, Structured illumination microscopy with noise-controlled image reconstructions. *Nat. Methods* **18**, 821–828 (2021).

43. X. Huang, J. Fan, L. Li, H. Liu, R. Wu, Y. Wu, L. Wei, H. Mao, A. Lal, P. Xi, L. Tang, Y. Zhang, Y. Liu, S. Tan, L. Chen, Fast, long-term, super-resolution imaging with Hessian structured illumination microscopy. *Nat. Biotechnol.* **36**, 451–459 (2018).

44. A. Llavador, J. Sola-Pikabea, G. Saavedra, B. Javidi, M. Martínez-Corral, Resolution improvements in integral microscopy with Fourier plane recording. *Opt. Express* **24**, 20792–20798 (2016).

45. L. Cong, Z. Wang, Y. Chai, W. Hang, C. Shang, W. Yang, L. Bai, J. Du, K. Wang, Q. Wen, Rapid whole brain imaging of neural activity in freely behaving larval zebrafish (Danio rerio). *eLife* **6**, e28158 (2017).

46. G. Scrofani, J. Sola-Pikabea, A. Llavador, E. Sanchez-Ortiga, J. C. Barreiro, G. Saavedra, J. Garcia-Sucerquia, M. Martínez-Corral, FIMic: Design for ultimate 3D-integral microscopy of in-vivo biological samples. *Biomed. Opt. Express* **9**, 335–346 (2018).

47. C. Guo, W. Liu, X. Hua, H. Li, S. Jia, Fourier light-field microscopy. *Opt. Express* **27**, 25573–25594 (2019).

48. K. Yanny, N. Antipa, W. Liberti, S. Dehaeck, K. Monakhova, F. L. Liu, K. Shen, R. Ng, L. Waller, Miniscope3D: Optimized single-shot miniature 3D fluorescence microscopy. *Light Sci. Appl.* **9**, 171 (2020).

49. Y. Xue, I. G. Davison, D. A. Boas, L. Tian, Single-shot 3D wide-field fluorescence imaging with a Computational Miniature *Mesoscope Sci. Adv.* **6**, eabb7508 (2020).

50. W. Liu, G.-A. R. Kim, S. Takayama, S. Jia, Fourier light-field imaging of human organoids with a hybrid point-spread function. *Biosens. Bioelectron.* **208**, 114201 (2022).

51. Z. Zhang, L. Bai, L. Cong, P. Yu, T. Zhang, W. Shi, F. Li, J. Du, K. Wang, Imaging volumetric dynamics at high speed in mouse and zebrafish brain with confocal light field microscopy. *Nat. Biotechnol.* **39**, 74–83 (2021).

52. X. Hua, W. Liu, S. Jia, High-resolution Fourier light-field microscopy for volumetric multi-color live-cell imaging. *Optica* **8**, 614–620 (2021).

53. D. C. Nguyen, T. A. Hookway, Q. Wu, R. Jha, M. K. Preininger, X. Chen, C. A. Easley, P. Spearman, S. R. Deshpande, K. Maher, M. B. Wagner, T. C. McDevitt, C. Xu, Microscale generation of cardiospheres promotes robust enrichment of cardiomyocytes derived from human pluripotent stem cells. *Stem Cell Reports* **3**, 260–268 (2014).

54. D. M. Camacho, K. M. Collins, R. K. Powers, J. C. Costello, J. J. Collins, Next-generation machine learning for biological networks. *Cell* **173**, 1581–1592 (2018).

55. E. Moen, D. Bannon, T. Kudo, W. Graf, M. Covert, D. Van Valen, Deep learning for cellular image analysis. *Nat. Methods* **16**, 1233–1246 (2019).

56. B. Midtvedt, S. Helgadottir, A. Argun, J. Pineda, D. Midtvedt, G. Volpe, Quantitative digital microscopy with deep learning. *Appl. Phys. Rev.* **8**, 011310 (2021).

57. M. Weigert, U. Schmidt, T. Boothe, A. Müller, A. Dibrov, A. Jain, B. Wilhelm, D. Schmidt, C. Broaddus, S. Culley, M. Rocha-Martins, F. Segovia-Miranda, C. Norden, R. Henriques, M. Zerial, M. Solimena, J. Rink, P. Tomancak, L. Royer, F. Jug, E. W. Myers, Content-aware image restoration: Pushing the limits of fluorescence microscopy. *Nat. Methods* **15**, 1090–1097 (2018).

58. J. Chen, H. Sasaki, H. Lai, Y. Su, J. Liu, Y. Wu, A. Zhovmer, C. A. Combs, I. Rey-Suarez, H.-Y. Chang, C. C. Huang, X. Li, M. Guo, S. Nizambad, A. Upadhyaya, S.-J. J. Lee, L. A. G. Lucas, H. Shroff, Three-dimensional residual channel attention networks denoise and sharpen fluorescence microscopy image volumes. *Nat. Methods* **18**, 678–687 (2021).

59. J. Lehtinen, J. Munkberg, J. Hasselgren, S. Laine, T. Karras, M. Aittala, T. Aila, Noise2noise: Learning image restoration without clean data. arXiv:1803.04189 [cs.CV] (12 March 2018).

60. Y. Wang, H. Pinkard, E. Khwaja, S. Zhou, L. Waller, B. Huang, Image denoising for fluorescence microscopy by supervised to self-supervised transfer learning. *Opt. Express* **29**, 41303–41312 (2021).

61. L. von Chamier, R. F. Laine, J. Jukkala, C. Spahn, D. Krentzel, E. Nehme, M. Lerche, S. Hernández-Pérez, P. K. Mattila, E. Karinou, S. Holden, A. C. Solak, A. Krull, T.-O. Buchholz, M. L. Jones, L. A. Royer, C. Leterrier, Y. Shechtman, F. Jug, M. Heilemann, G. Jacquemet, R. Henriques, Democratising deep learning for microscopy with ZeroCostDL4Mic. *Nat. Commun.* **12**, 2276 (2021).

62. R. F. Laine, G. Jacquemet, A. Krull, Imaging in focus: An introduction to denoising bioimages in the era of deep learning. *Int. J. Biochem. Cell Biol.* **140**, 106077 (2021).

63. V. Mannam, Y. Zhang, Y. Zhu, E. Nichols, Q. Wang, V. Sundaresan, S. Zhang, C. Smith, P. W. Bohn, S. S. Howard, Real-time image denoising of mixed Poisson–Gaussian noise in fluorescence microscopy images using ImageJ. *Optica* **9**, 335–345 (2022).

64. C. Belthangady, L. A. Royer, Applications, promises, and pitfalls of deep learning for fluorescence image reconstruction. *Nat. Methods* **16**, 1215–1225 (2019).

65. E. Bostan, R. Heckel, M. Chen, M. Kellman, L. Waller, Deep phase decoder: Self-calibrating phase microscopy with an untrained deep neural network. *Optica* **7**, 559–562 (2020).

66. Y. Li, Y. Su, M. Guo, X. Han, J. Liu, H. D. Vishwasrao, X. Li, R. Christensen, T. Sengupta, M. W. Moyle, J. Chen, T. B. Usdin, D. Colón-Ramos, H. Liu, Y. Wu, H. Shroff, Incorporating the image formation process into deep learning improves network performance in deconvolution applications. bioRxiv 2022.03.05.483139 [**Preprint**]. 6 March 2022. https://doi.org/10.1101/2022.03.05.483139.

67. H. Hermessi, O. Mourali, E. Zagrouba, Convolutional neural network-based multimodal image fusion via similarity learning in the shearlet domain. *Neural. Comput. Appl.* **30**, 2029–2045 (2018).

68. H. Andrade-Loarca, G. Kutyniok, tfShearlab: The TensorFlow Digital Shearlet Transform for Deep Learning. arXiv:2006.04591 [eess.IV] (8 June 2020).

69. R. Diekmann, J. Deschamps, Y. Li, T. Deguchi, A. Tschanz, M. Kahnwald, U. Matti, J. Ries, Photon-free (s)CMOS camera characterization for artifact reduction in high- and super-resolution microscopy. *Nat. Commun.* **13**, 3362 (2022).

70. S. I. Djomehri, B. Burman, M. E. Gonzalez, S. Takayama, C. G. Kleer, A reproducible scaffold-free 3D organoid model to study neoplastic progression in breast cancer. *J. Cell Commun. Signal* **13**, 129–143 (2019).

71. S. Lee, J. Chang, S. M. Kang, E. Parigoris, J.-H. Lee, Y. S. Huh, S. Takayama, High-throughput formation and image-based analysis of basal-in mammary organoids in 384-well plates. *Sci. Rep.* **12**, 317 (2022).

72. E. Parigoris, S. Lee, D. Mertz, M. Turner, A. Y. Liu, J. Sentosa, S. Djomehri, H. C. Chang, K. Luker, G. Luker, C. G. Kleer, S. Takayama, Cancer cell invasion of mammary organoids with basal-In phenotype. *Adv. Healthc. Mater.* **10**, e2000810 (2021).

73. M. Müller, V. Mönkemöller, S. Hennig, W. Hübner, T. Huser, Open-source image reconstruction of super-resolution structured illumination microscopy data in ImageJ. *Nat. Commun.* **7**, 10980 (2016).

74. R. Jha, Q. Wu, M. Singh, M. K. Preininger, P. Han, G. Ding, H. C. Cho, H. Jo, K. O. Maher, M. B. Wagner, C. Xu, Simulated microgravity and 3D culture enhance induction, viability, proliferation and differentiation of cardiac progenitors from human pluripotent stem cells. *Sci. Rep.* **6**, 30956 (2016).

75. S. J. Smith, P. Ataliotis, S. Kotecha, N. Towers, D. B. Sparrow, T. J. Mohun, The MLC1v gene provides a transgenic marker of myocardium formation within developing chambers of the Xenopus heart. *Dev. Dyn.* **232**, 1003–1012 (2005).

76. J. R. Janesick, Photon transfer noise sources, in *Photon Transfer*, J. R. Janesick, Ed. (SPIE, 2009), pp. 21–34.

77. J. R. Janesick, Photon transfer theory, in *Photon Transfer*, J. R. Janesick, Ed. (SPIE, 2009), pp. 35–48.

78. A. Foi, M. Trimeche, V. Katkovnik, K. Egiazarian, Practical Poissonian-Gaussian noise modeling and fitting for single-image raw-data. *IEEE Trans. Image Process.* **17**, 1737–1754 (2008).

79. H. T. Beier, B. L. Ibey, Experimental comparison of the high-speed imaging performance of an EM-CCD and sCMOS camera in a dynamic live-cell imaging test case. *PLOS ONE* **9**, e84614 (2014).

80. B. Moomaw. Camera technologies for low light imaging: Overview and relative advantages. *Methods Cell Biol.* **114**, 243–283 (2013).

81. G. D. Boreman, Transfer function techniques, in *Handbook of Optics*, M. Bass, Ed. (McGraw-Hill, 1995), pp. 1–10.

82. W. K. Pratt, *Introduction to Digital Image Processing* (CRC Press, 2018).

83. M. Colom, A. Buades, J.-M. Morel, Nonparametric noise estimation method for raw images. *J. Opt. Soc. Am. A* **31**, 863–871 (2014).

84. G. R. Easley, D. Labate, W. Q. Lim, Optimally sparse image representations using shearlets. *Conference Record - Asilomar Conference on Signals, Systems and Computers* (2006), pp. 974–978.

85. V. M. Patel, G. R. Easley, D. M. Healy, Shearlet-based deconvolution. *EEE Trans Image Process.* **18**, 2673–2685 (2009).

86. G. Kutyniok, W.-Q. Lim, Image separation using wavelets and shearlets, in *Curves and Surfaces*, J.-D. Boissonnat, P. Chenin, A. Cohen, C. Gout, T. Lyche, M.-L. Mazure, L. Schumaker, Eds. (Springer, 2012), vol. 6920, pp. 416–430.

87. G. Easley, D. Labate, W.-Q. Lim, Sparse directional image representations using the discrete shearlet transform. *Appl. Comput. Harmon. Anal.* **25**, 25–46 (2008).

88. A. Descloux, K. S. Grußmayer, A. Radenovic, Parameter-free image resolution estimation based on decorrelation analysis. *Nat. Methods* **16**, 918–924 (2019).

89. K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, Image restoration by sparse 3D transform-domain collaborative filtering. *Image Processing: Algorithms and Systems VI* **6812**, 681207 (2008).

90. A. Danielyan, V. Katkovnik, K. Egiazarian, Deblurring of Poissonian images using BM3D frames. *Wavelets and Sparsity XIV* **8138**, 813812 (2011).

91. P. M. Carlton, J. Boulanger, C. Kervrann, J. B. Sibarita, J. Salamero, S. Gordon-Messer, D. Bressan, J. E. Haber, S. Haase, L. Shao, L. Winoto, A. Matsuda, P. Kner, S. Uzawa, M. Gustafsson, Z. Kam, D. A. Agard, J. W. Sedat, Fast live simultaneous multiwavelength four-dimensional optical microscopy. *Proc. Natl. Acad. Sci.* **107**, 16016–16022 (2010).

92. A. Buades, B. Coll, J. M. Morel, A non-local algorithm for image denoising, in *Proceedings - 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005* **II**, (2005), pp. 60–65.

93. V. Duval, J.-F. Aujol, Y. Gousseau, A Bias-variance approach for the nonlocal means. *SIAM J. Imaging Sci.* **4**, 760–788 (2011).

94. I. Frosio, J. Kautz, "On Nearest Neighbors in Non Local Means Denoising. *31st Conference on Neural Information Processing Systems (NIPS 2017),* arXiv: 1711.07568 [cs.CV] (20 November 2017).

95. O. Lotan, M. Irani, Needle-match: Reliable patch matching under high uncertainty, in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2016-Decem*, (2016), pp. 439–448.

96. M. Maggioni, G. Boracchi, A. Foi, K. Egiazarian, Video denoising, deblocking, and enhancement through separable 4-D nonlocal spatiotemporal transforms. *IEEE Trans. Image Process.* **21**, 3952–3966 (2012).

97. K. Dabov, A. Foi, K. Egiazarian, Video denoising by sparse 3D transform-domain collaborative filtering, in *15th European Signal Processing Conference* (2007), pp. 145–149.

98. S. Zhu, K.-K. Ma, A new diamond search algorithm for fast block-matching motion estimation. *IEEE Trans. Image Process.* **9**, 287–290 (2000).

99. R. Heintzmann, Band limit and appropriate sampling in microscopy. *Cell Biol.* **3**, 29–36 (2006).

100. M. Mäkitalo, A. Foi, Poisson-Gaussian denoising using the exact unbiased inverse of the generalized anscombe transformation, in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings* (2012), pp. 1081–1084.

101. K. Zhang, W. Zuo, Y. Chen, D. Meng, L. Zhang, Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. *IEEE Trans. Image Process.* **26**, 3142–3155 (2018).

102. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2016), pp. 770–778.

103. B. Eismann, T. G. Krieger, J. Beneke, R. Bulkescher, L. Adam, H. Erfle, C. Herrmann, R. Eils, C. Conrad, Automated 3D light-sheet screening with high spatiotemporal resolution reveals mitotic phenotypes. *J. Cell Sci.* **133**, jcs245043 (2020).

104. T. Falk, D. Mai, R. Bensch, Ö. Çiçek, A. Abdulkadir, Y. Marrakchi, A. Böhm, J. Deubner, Z. Jäckel, K. Seiwald, A. Dovzhenko, O. Tietz, C. D. Bosco, S. Walsh, D. Saltukoglu, T. L. Tay, M. Prinz, K. Palme, M. Simons, I. Diester, T. Brox, O. Ronneberger, U-Net: Deep learning for cell counting, detection, and morphometry. *Nat. Methods* **16**, 67–70 (2019).

105. M. G. Haberl, C. Churas, L. Tindall, D. Boassa, S. Phan, E. A. Bushong, M. Madany, R. Akay, T. J. Deerinck, S. T. Peltier, M. H. Ellisman, *CDeep3M*—Plug-and-Play cloud-based deep learning for image segmentation. *Nat. Methods* **15**, 677–680 (2018).

106. M. Brady, Computational approaches to image understanding. *ACM Computing Surveys* **14**, 3–71 (1982).

107. T. A. Bubba, G. Kutyniok, M. Lassas, M. Mrz, W. Samek, S. Siltanen, V. Srinivasan, Learning the invisible: A hybrid deep learning-shearlet framework for limited angle computed tomography. *Inverse Probl.* **35**, 064002 (2019).

108. Z. Lyu, C. Zhang, M. Han, DSTnet: A new discrete shearlet transform-based CNN model for image denoising. *Multimed. Syst.* **27**, 1165–1177 (2021).