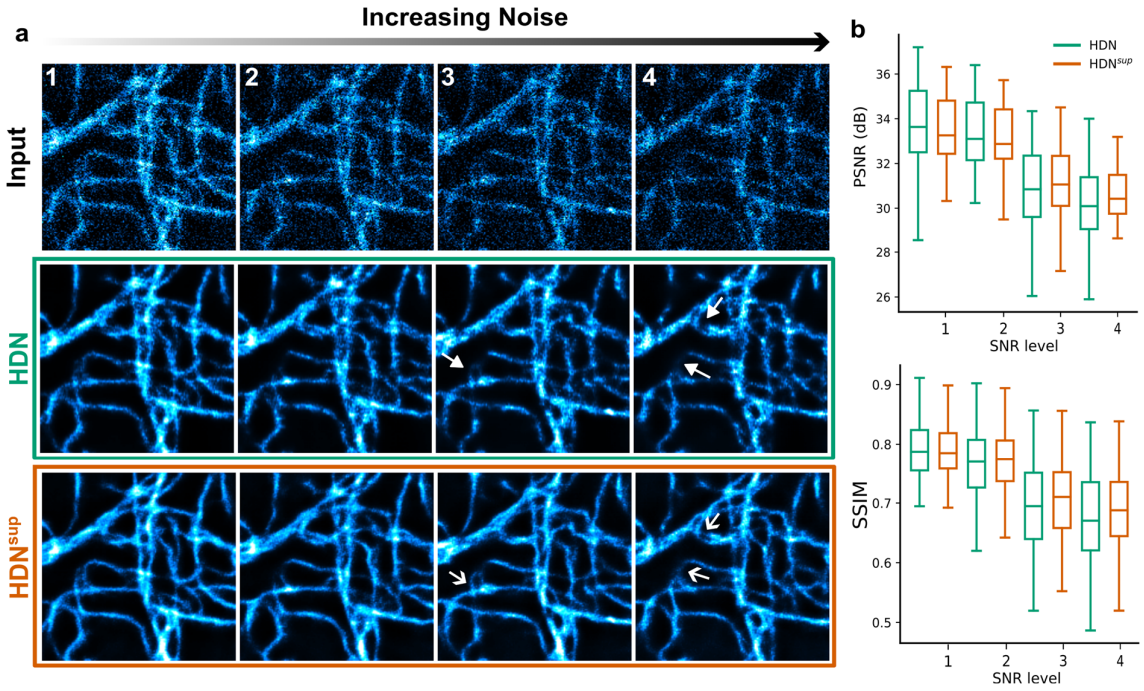
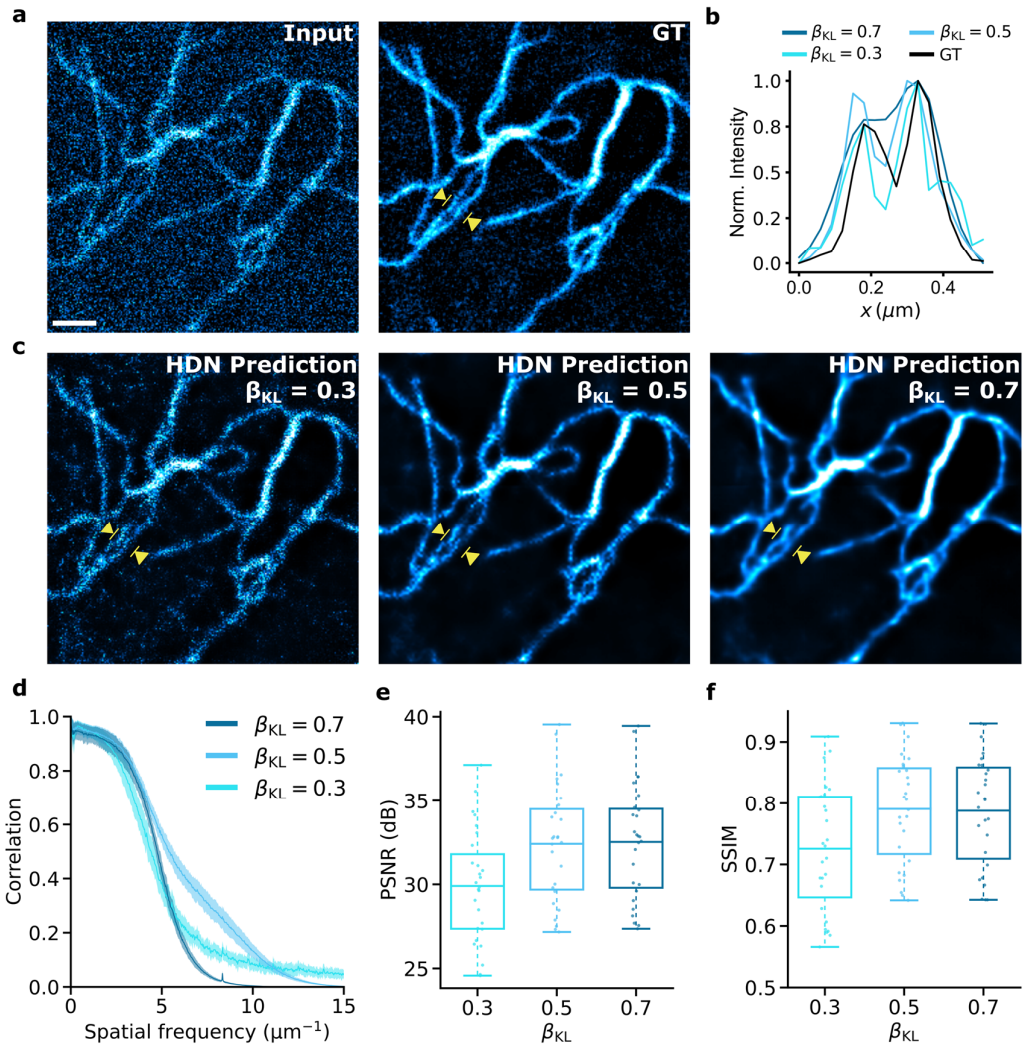


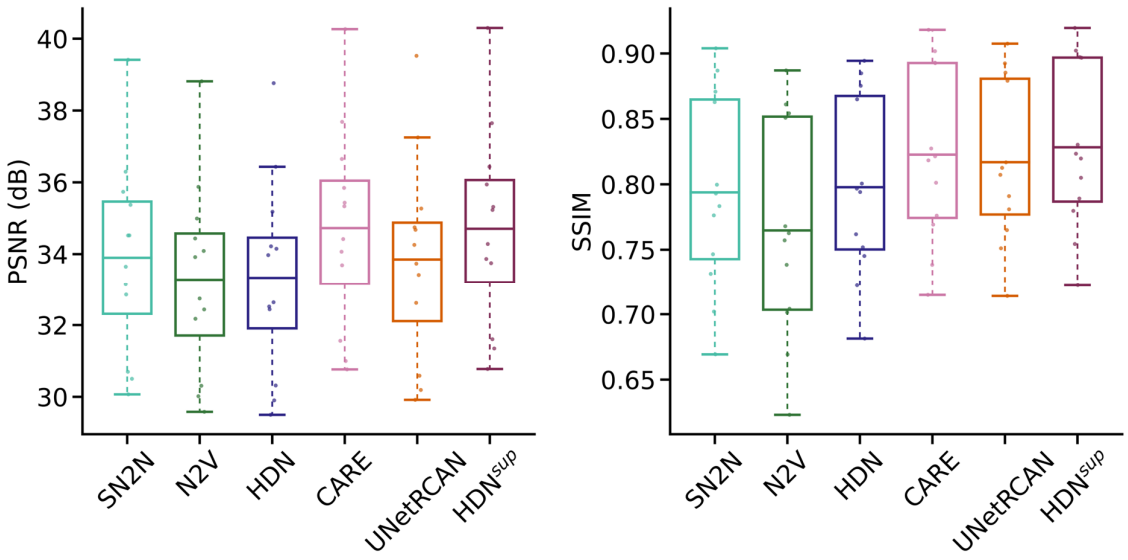
Supplementary Figure S1: HDN vs. HDN^{sup} for different SNR levels.

a Prediction example for HDN and HDN^{sup} for 4 different SNR input levels, on the fixed vimentin-rsEGFP2 dataset. Different SNR have been obtained by letting the sample bleach over multiple acquisitions. Arrows show structures better recovered by HDN^{sup} for low SNR levels. **b** PSNR and SSIM over the whole test dataset (N = 3 cells, cut in 600x600 images for more statistics). HDN^{sup} gives superior performances over HDN for lower SNR levels 3 and 4 for both metrics.



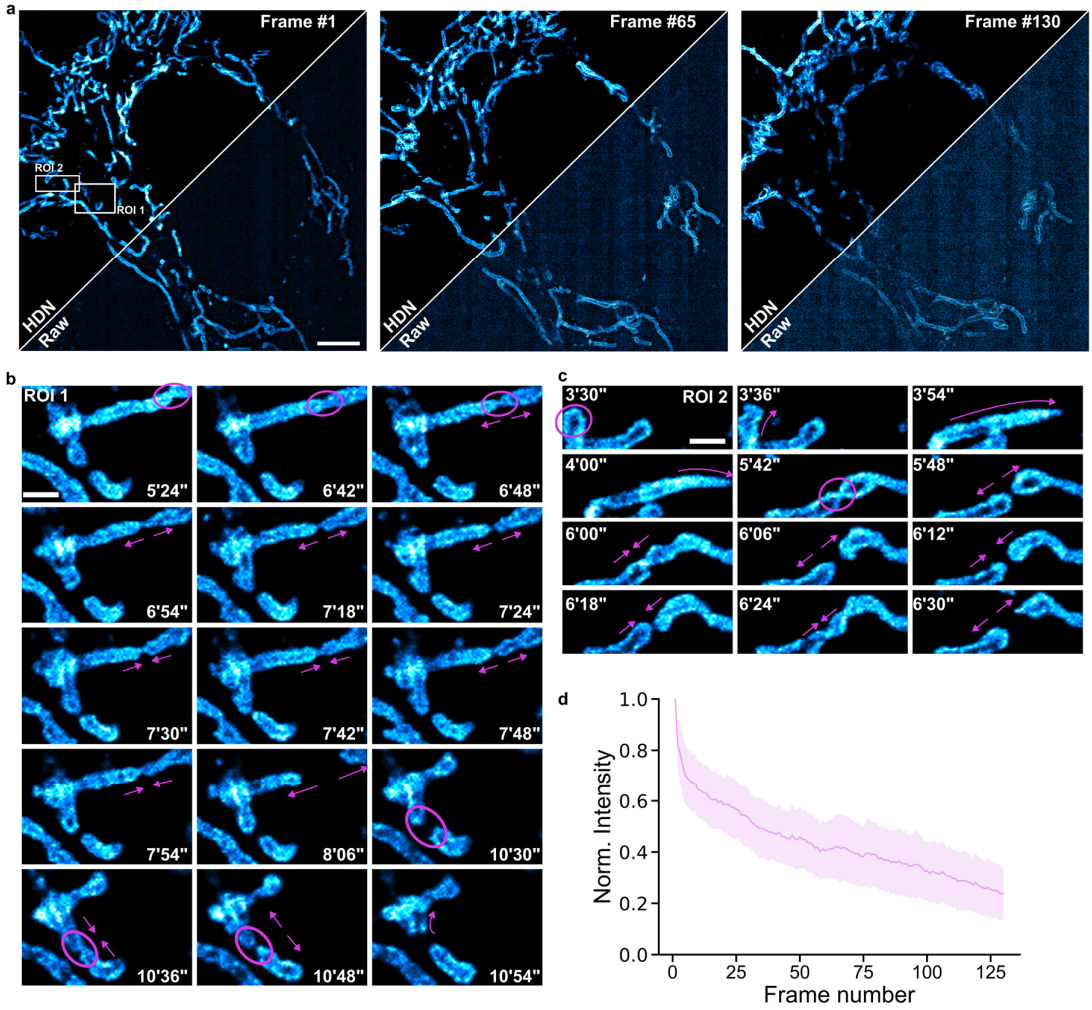
Supplementary Figure S2: Fine-tuning of the KL-loss weight β_{KL} for HDN training

The weight β_{KL} must be tuned to denoise enough without over smoothing. Here an example of this fine-tuning is shown for HDN training on the vimentin-rsEGFP2 dataset, with 3 weight values: too low (β_{KL} = 0.3), too high (β_{KL} = 0.7) and the selected value (β_{KL} = 0.5) which allows the best compromise between denoising and maintaining resolution. **a,c** GT, input and predictions results for the 3 β_{KL} values. Scale bar, 1 μm. **b** Line profile showing two close-by filaments in the GT (yellow arrows in **a**, **c**). Note that β_{KL} = 0.7 does not allow to distinguish the two structures. **d** FRC curves calculated over the whole training dataset (mean ± std, N=20 cells), showing a clear loss of resolution for β_{KL} = 0.7. **e** PSNR and SSIM over the whole training dataset (N=20 cells), showing that β_{KL} = 0.3 is too low for optimal SNR restoration.



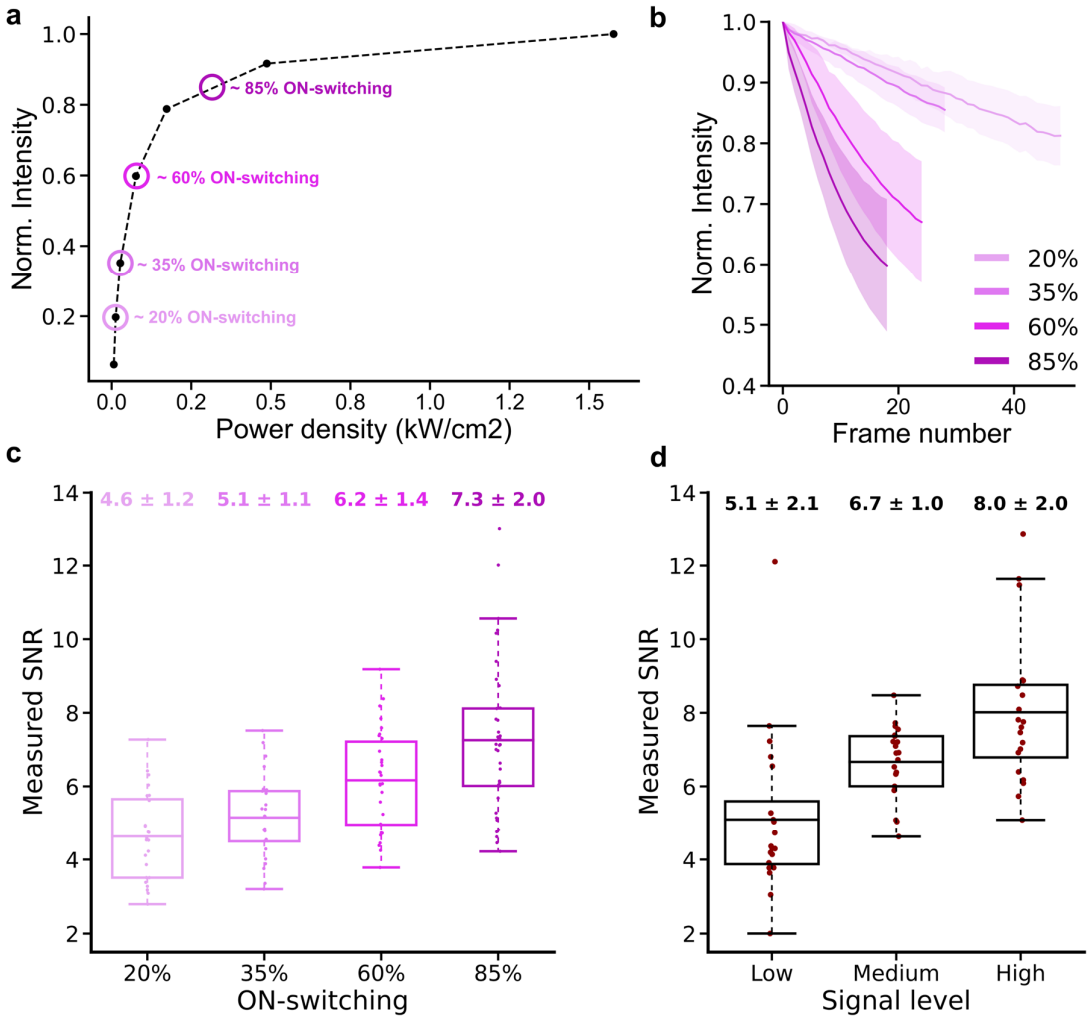
Supplementary Figure S3: PSNR and SSIM for all denoising methods

PSNR and SSIM metrics for all denoising methods, calculated over the test dataset only (N = 3 cells, cut in 600 × 600 images for more statistics).



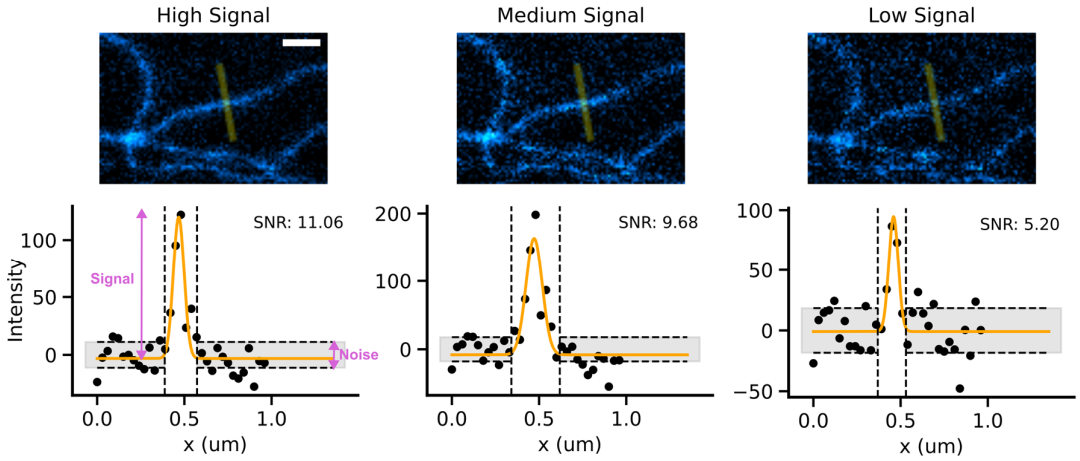
Supplementary Figure S4: Prolonged imaging OMP-rsEGFP2 in U20S

MoNaLISA + HDN allows to image mitochondria in U20S cells for 130 frames, ~ 4 times more than regular MoNaLISA. Frame rate: 6 s. **a** Full FOV examples of raw and denoised for frame #1, #65 and #130. Scale bar, 5 µm. **b** Zoom-in over ROI1, showing first a mitochondria fission, followed by 2 back-and-forth contacts before definite separation happening (t = 5'24" – t = 8'06"), and another quick contact of 2 mitochondria between (t = 10'30" – t = 10'54"). Pink arrows highlight the movement of the mitochondria. Scale bar, 1µm. **c** Zoom-in over ROI2, showing one mitochondrion first quickly extending, then later splitting into 2, and contacting again before definite fission. Scale bar, 1 µm. **d** Normalized fluorescence intensity over the 130 frames (on raw data, mean ± std, N = 4 cells).



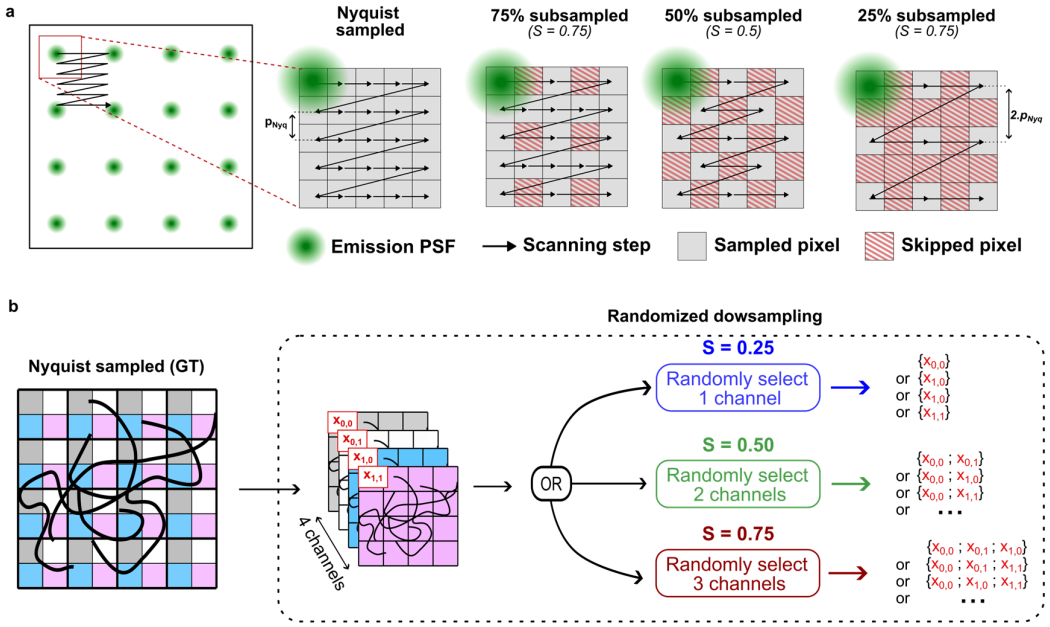
Supplementary Figure S5: ON-switching, bleaching and SNR on vimentin-rsEGFP2.

a Experimentally obtained ON-switching curve. The ON-switching levels chosen for bleaching measurements are indicated with circles. **b** Bleaching rate depending on the ON-switching (mean ± std, N = 5 cells for each ON-switching level). **c** Measured SNR for each ON-switching level, on live-cells, on the first frame of the timelapse. Mean ± std is indicated at the top of each box. SNR measurements are done from line profiles (see supplementary Figure 6), using N = 25 – 30 filaments per ON-switching level. **d** Measured SNR on the fixed cells dataset, where the different signal levels (High, Medium, Low) have been obtained by letting the sample bleach, using N = 20 filaments per signal level.



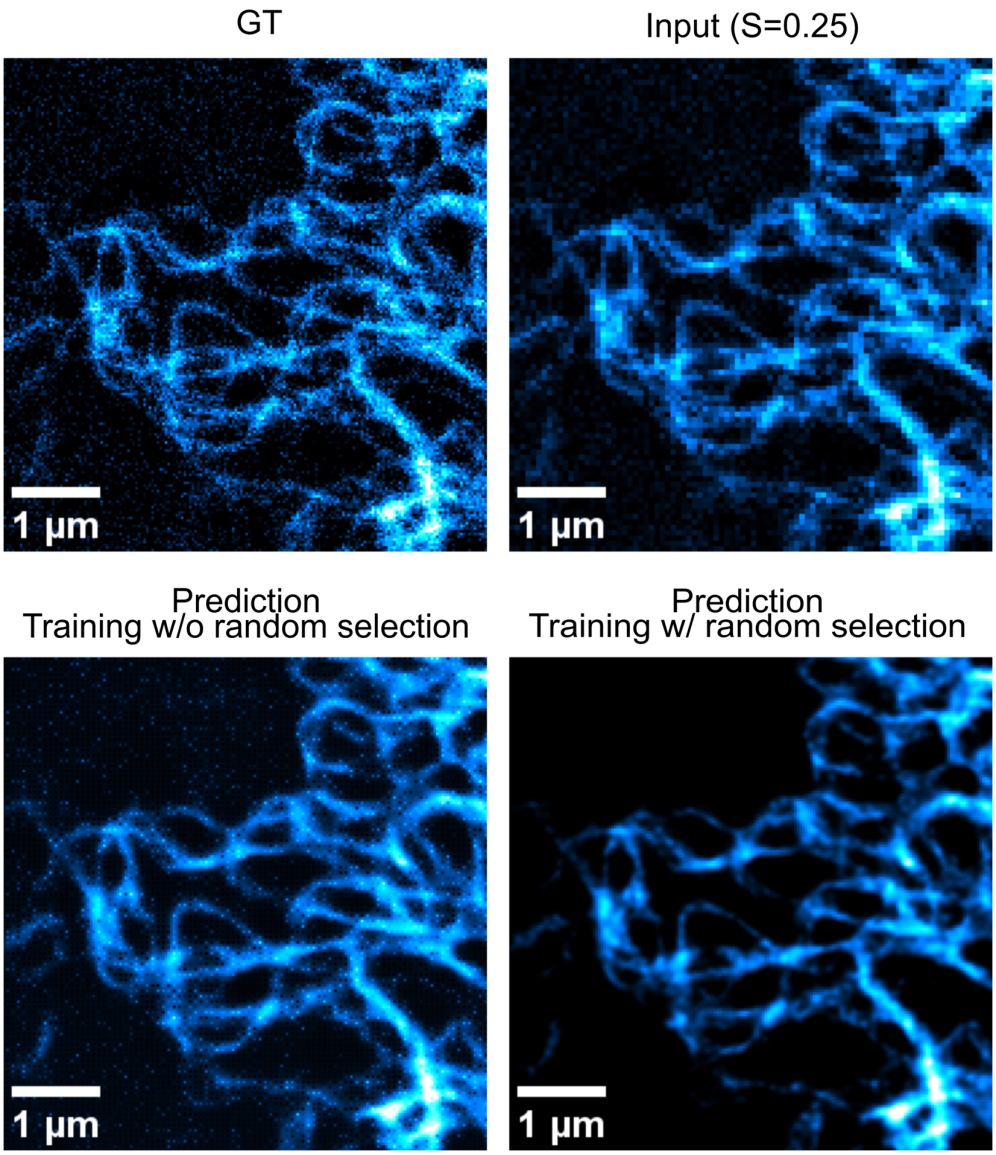
Supplementary Figure S6: Example of SNR calculation.

We calculate SNR as a ratio between the signal and the std of the background. For that, we use the line profiles of filaments. The signal is taken as the peak value of the fitted gaussian, and values outside of gaussian are considered as background. Here we show an example of such line profiles on the fixed vimentin-rsEGFP2 dataset for the 3 different signal levels, giving SNR values of 11.06 (High), 9.68 (Medium), and 5.20 (Low).



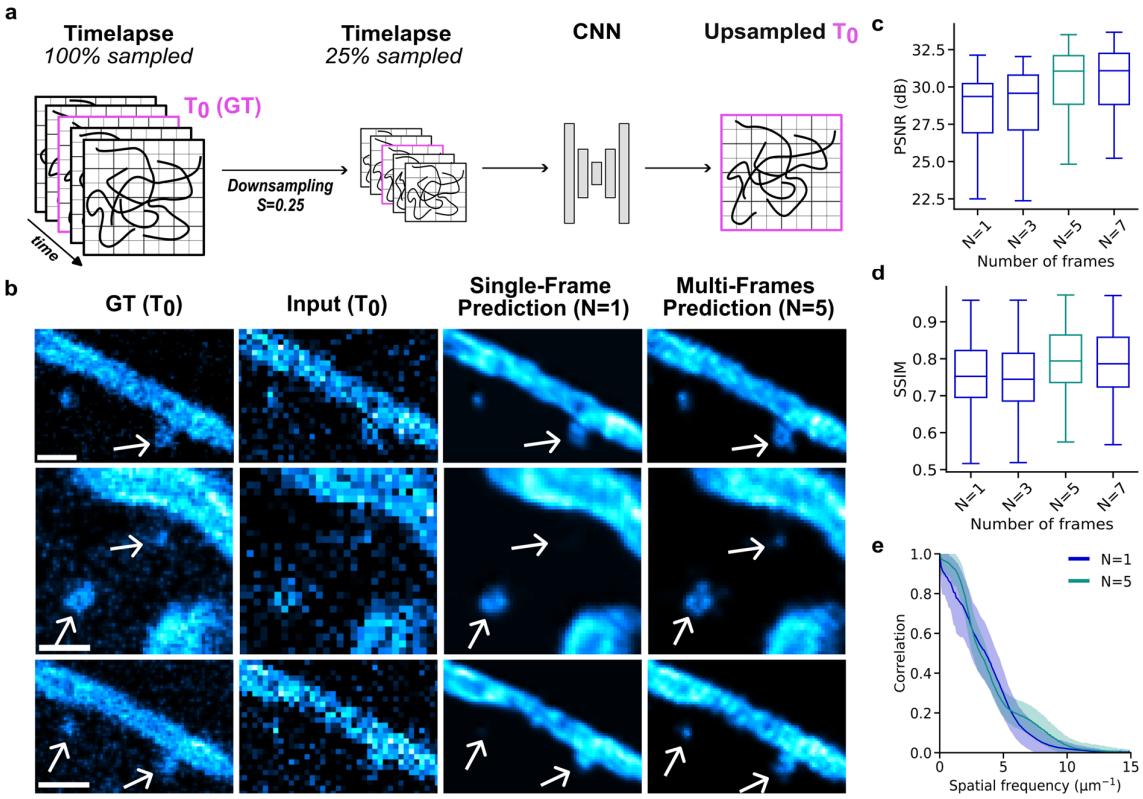
Supplementary Figure S7: Sub-sampling strategies and randomized downsampling

a In a parallelized system, we use a Raster Scan in between 2 emitting PSFs. A Nyquist sampling corresponds to a scan step of $\mathbf{p_{NYQ}}$, i.e half of the expected resolution, in both XY direction. Doubling the scan step to $\mathbf{2p_{NYQ}}$ in both directions will remove 75 % of the information, creating an image sampled at only 25 % (S = 0.25). By skipping only a quarter or half of the scans steps, we can have intermediate sub-sampling images: 75 % sampling (S = 0.75) and 50 % sampling (S = 0.5). **b** Randomized downsampling strategy for S = 0.25, S = 0.5 and S = 0.75. From a Nyquist sampled GT, we first create 4 downsampled images, each one corresponding to a doubled pixel size, but with a different subset of pixel. From this stack of 4 channels, we can select 1,2, or 3 depending on the desired sampling ratio S. For a given S, the channel selection is randomized.



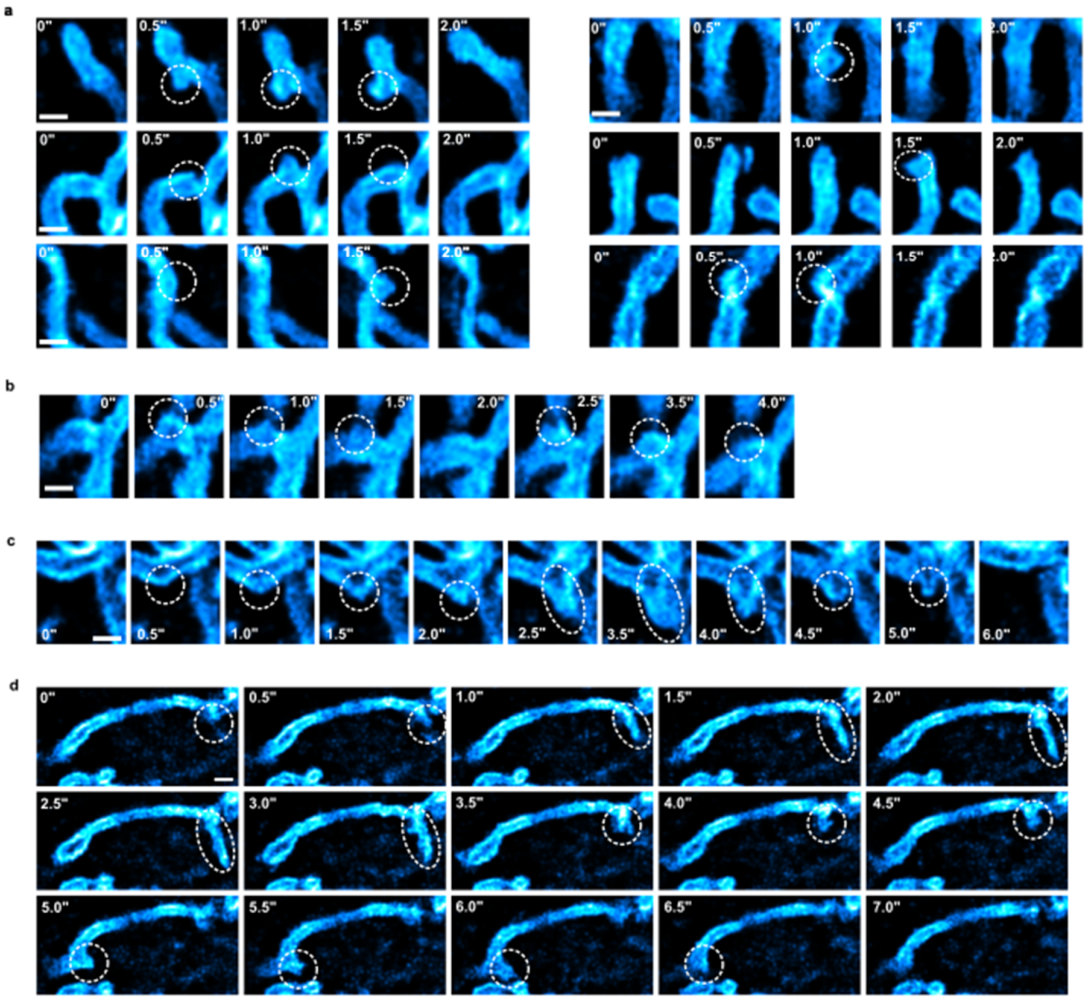
Supplementary Figure S8: Random vs deterministic pixel selection.

With a deterministic pixel selection during the training, the pixels in the input are always coming from the same positions from the GT. Therefore, it is easy for the network to learn that it should keep those exact pixels value to minimize the loss between prediction and GT. The GT's noise present in those pixels will thus be kept in the prediction. The other pixels that the network have to learn how to predict, however, will be necessarily more smoothed (the network cannot predict noise), creating this patterned noise effect. On the contrary, if the downsampling process does a random pixel selection, the network cannot know from where the input pixels are coming from. It must then converge towards a smooth output for all pixels.



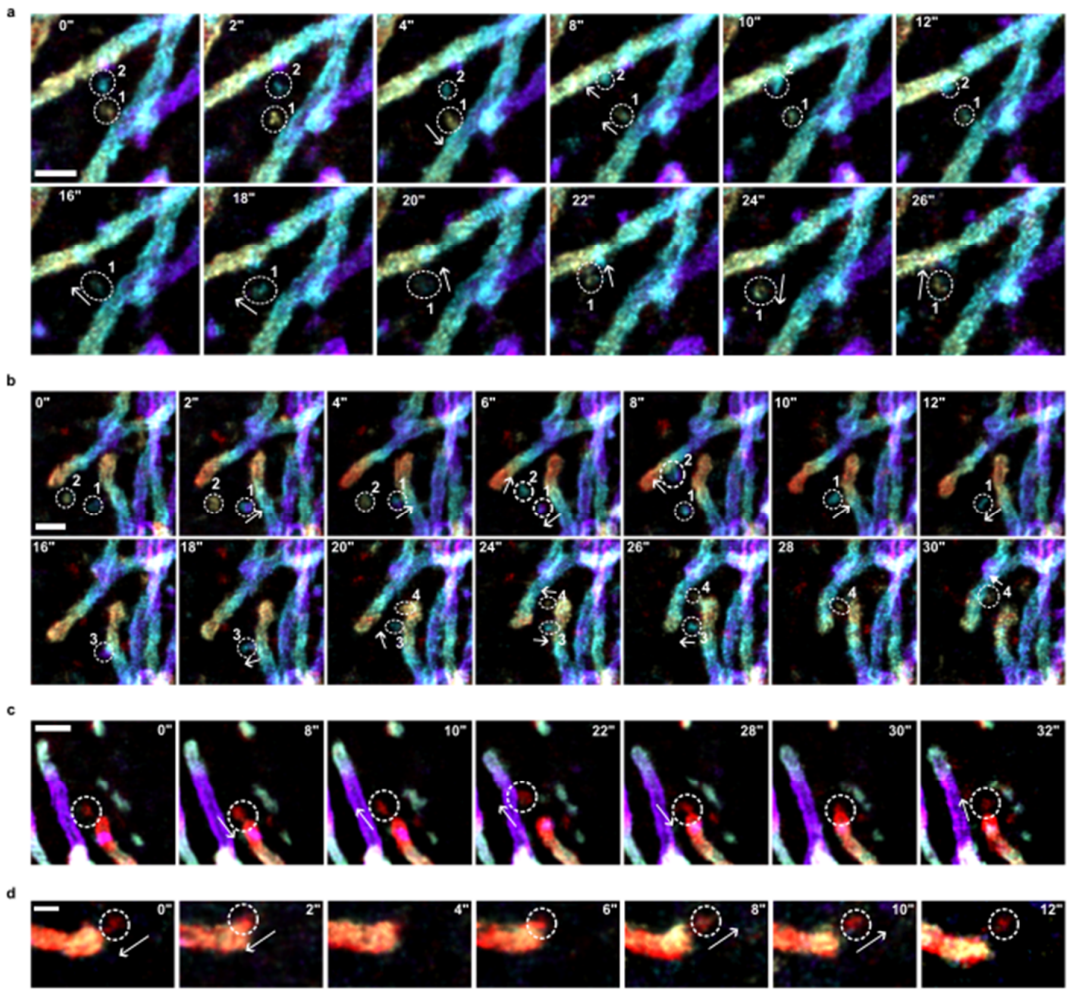
Supplementary Figure S9: Multi-Frame Training and Results on Mitochondria

a A multi-frames training pair is obtained by downsampling a fully-sampled timelapse with a sampling ratio of S = 0.25. The sub-sampled timelapse is feeded into the CNN, which is trained to predict the fully-sampled middle timepoint, $\mathbf{T_0}$. **b** Representative examples of the improvement of Multi-Frames training (with N=5 frames) over the Single-Frame (N = 1 frame). White arrows with a solid line points to vesicles better resolved with the Multi-Frames approach. Arrows with a dotted line shows small vesicles that disappear in the Single-Frame prediction (considered as noise by the network), but accurately kept by the Multi-Frames.**c,d** PSNR and SSIM results for different number of frames N in the input timelapse, over the whole test dataset. We chose N = 5 frames for all following Multi-Frames training. **e** FRC measurements over the whole test dataset for Single-Frame and Multi-Frames, confirming that the Multi-Frame gives a slight edge in contrast for higher frequencies (mean ± std).



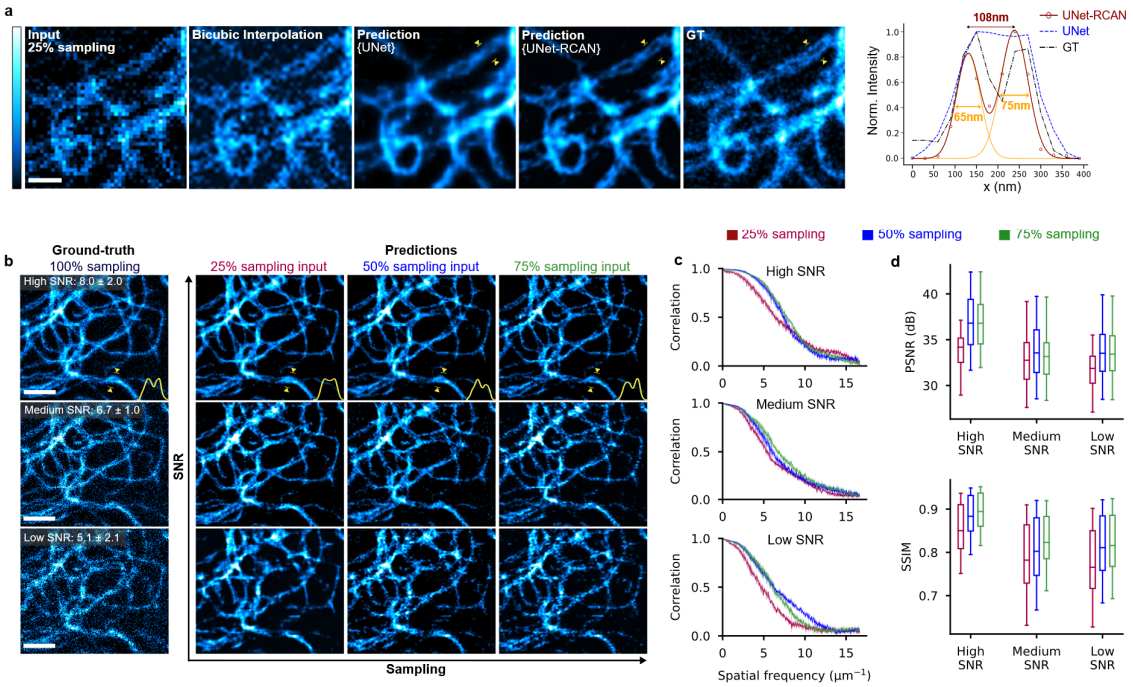
Supplementary Figure S10: Mitochondria fast protrusions observed with Fast-MoNaLISA

a 6 examples of fast and small mitochondria protrusion, elongating and retracting in 0.5 – 2". **b** Example of 2 protrusions in a row. **c,d** Examples of longer, slower protrusions. All scale bars are 500 nm. Frame rate is 2 Hz.



Supplementary Figure S11: MDV-mitochondria interactions examples

Multiple examples of Mitochondria Derived Vesicles (MDVs) budding out, fusing in, or touching multiple times mitochondria. Dotted circles shows the MDVs of interest, numbered when there are multiple ones in the same ROI. White arrows highlight the movement of the mitochondria. **a**, **b**, **c** Scale bar 1 µm, **d** scale bar 500 nm.



Supplementary Figure S12: Upsampling predictions for different sampling and SNR levels (Single-Frame training)

a From left to right: 25 % sampling input, Bicubic Interpolation, UNet prediction, UNet-RCAN prediction, GT, and line profile comparison showing UNet-RCAN superior ability to resolve close-by filaments. **b** Prediction examples for 3 different SNR levels and 3 different sampling. Scale bar, 1 μm. **c** Relative FRC with GT for all SNR and sampling levels (N = 3 cells). **d** PSNR and SSIM calculation for all SNR and sampling levels. (N = 3 cells).

Task	Dataset	Figure	Network	Patch size	# Patches	Learning Rate	Epochs	β_{KL} (when relevant)
Denoising	Vimentin	Main Fig.1,2 S1, S2, S3, S12	CARE	128	1613	1×10^{-3}	200	-
			UNet-RCAN	64	5467	1×10^{-4}	200	-
			N2V	256	2419	4×10^{-4}	200	-
			SN2N	128	7074	2×10^{-4}	200	-
			HDN	64	5467	1×10^{-4}	100	5×10^{-1}
			HDN ^{sup}	64	5467	1×10^{-4}	100	1×10^{-2}
	Actin	Main Fig.2	HDN ^{sup}	64	9953	1×10^{-4}	200	1×10^{-3}
	Mitochondria	S4	HDN	64	6918	1×10^{-4}	50	2×10^{-1}
Upsampling	Vimentin (Single-Frame)	Main Fig. 3 S8	UNet-RCAN	64	5467	1×10^{-4}	100	-
	Mitochondria (Multi-Frames)	Main Fig. 4 S9,10,11	UNet-RCAN	96	5922	5×10^{-5}	500	-

Supplementary Table 1: Training parameters

Legend of supplementary movies

Video 1: Prolonged mitochondrial imaging

Timelapse of 130 frames, 6sec frame rate, denoised with HDN (unsupervised), of U20S cells expressing OMP-rsEGFP2. Zoomed in region shows multiple fissions and multiple "back-and-forth contacts", highlighted by white arrows.

Video 2: Timelapse of actin meshwork dynamics

Timelapse of 50 frames, 5sec frame rate, denoised with HDN^{sup} of U20S cells expressing LifeAct-rsEGFP2. Zoomed in region shows that we can follow the actin meshwork dynamics for ~50 frames.

Video 3: Mitochondria motion artefacts example in regular MoNaLISA

Timelapse acquired with regular MoNaLISA, on a reduced FOV to achieve 1Hz frame rate, in U20S cells expressing OMP-rsEGFP2. Motion artefacts are highlighted by white circles.

Video 4: Mitochondria fast dynamics observed with 2D Fast-MoNaLISA

Fast-MoNaLISA timelapse (2Hz, full FOV) of U20S cells expressing OMP-rsEGFP2. First zoomed-in region highlights very small and fast protrusions elongating and retracting. Second zoomed-in region shows other varieties of events observed.

Video 5: Fast volumetric imaging and MDV budding-out and budding-in

Volumetric rendering of a Fast-MoNaLISA volumetric acquisition of U20S cells expressing OMP-rsEGFP2. Zoom-in region shows 2 Mitochondria Derived Vesicles (MDVs) budding-out and fusing with another mitochondrion.