

SUPPLEMENTARY MATERIAL

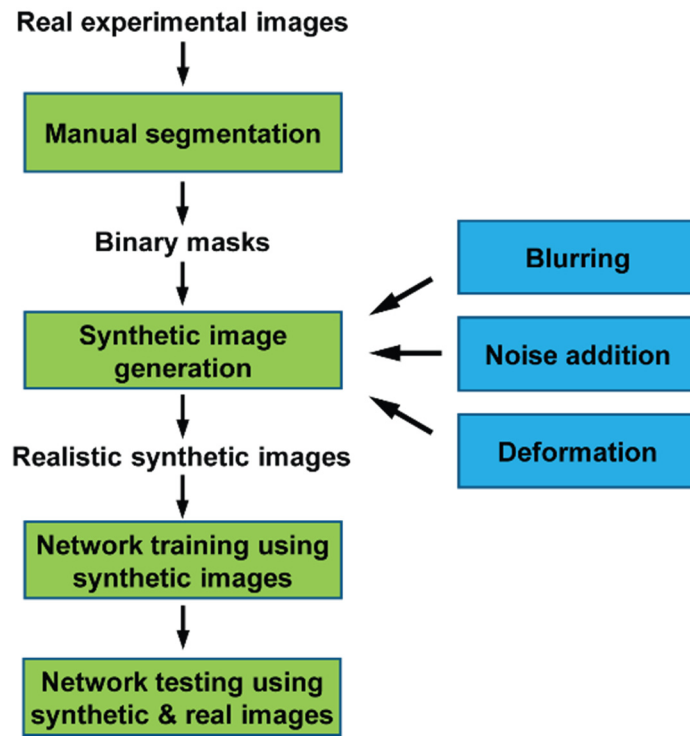


Figure S1. Overview of workflow.

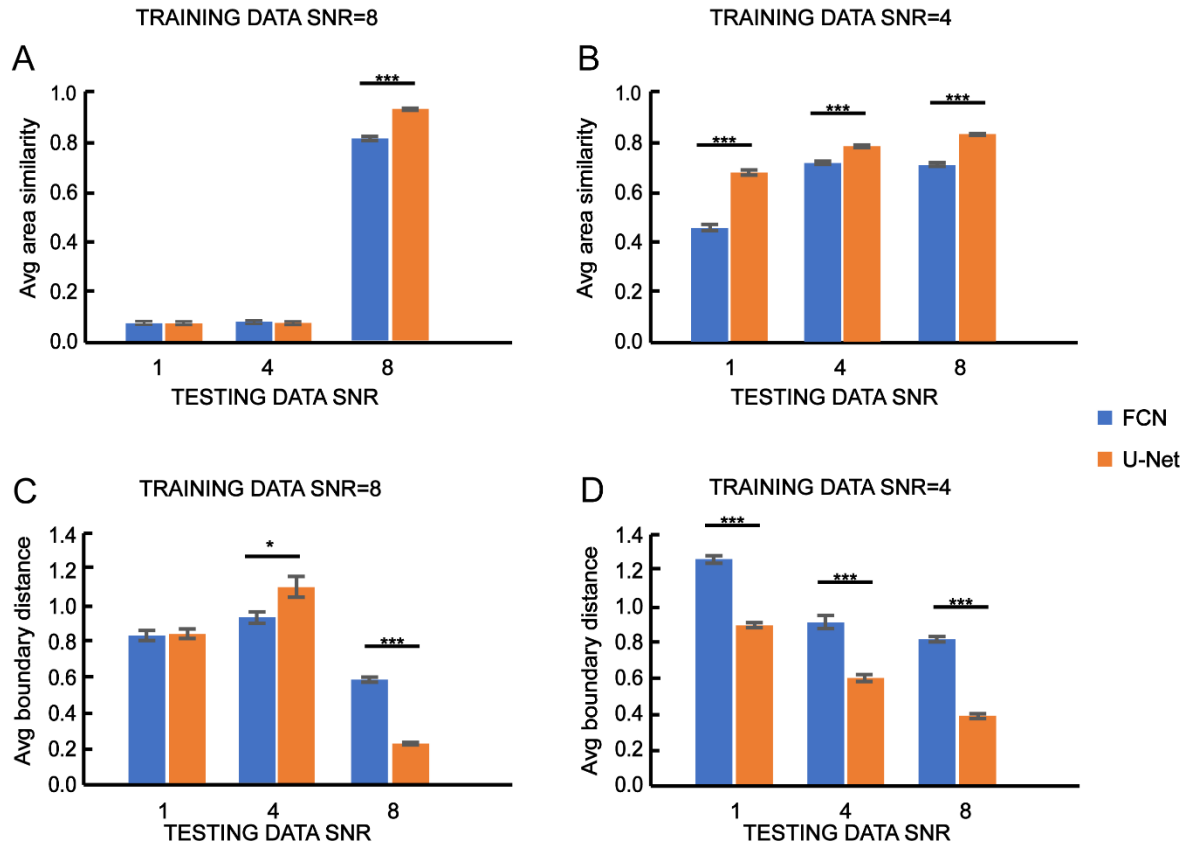


Figure S2. Performance comparison of FCN versus U-Net under match and mismatch of training data SNRs with testing data SNRs. (A) Comparison of average area similarity of FCN versus U-Net, trained with synthetic images generated at SNR = 8 and tested with synthetic images generated at SNR = 1, 4, or 8, respectively. (B) Comparison of average area similarity of FCN versus U-Net, trained with synthetic images generated at SNR = 4 and tested with synthetic images generated at SNR = 1, 4, or 8, respectively. (C) Comparison of average boundary distance of FCN versus U-Net, trained with synthetic images generated at SNR = 8 and tested with synthetic images generated at SNR = 1, 4, or 8, respectively. (D) Comparison of average boundary distance of FCN versus U-Net, trained with synthetic images generated at SNR = 4 and tested with synthetic images generated at SNR = 1, 4, or 8, respectively. For (A-D), see Tables S1, S2, and S3 for detailed lists of p-values of performance comparison. Notation for p-values: ***: $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

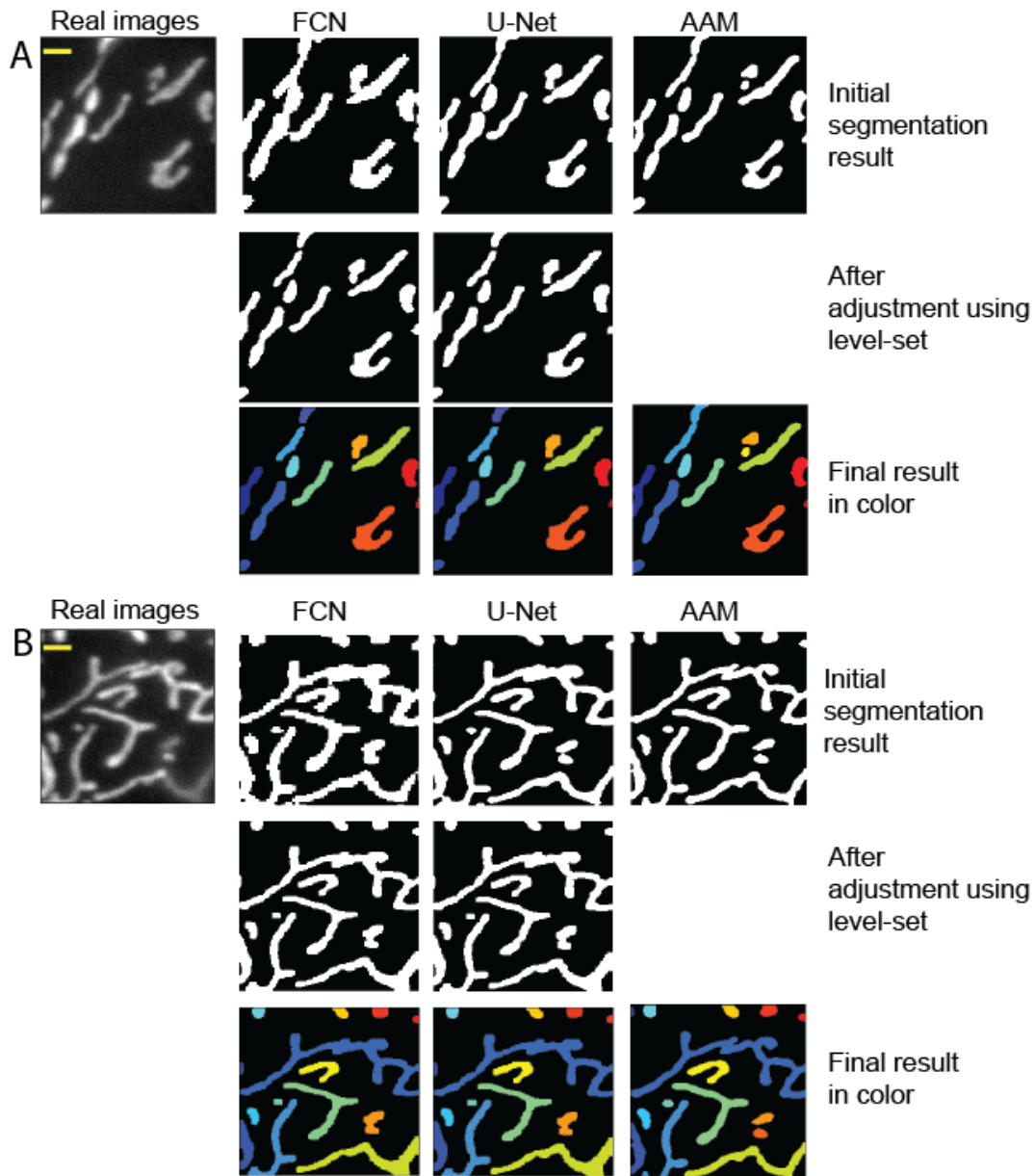


Figure S3. Segmentation of real images using FCN, U-Net, and AAM. (A) An example image of mitochondria and its segmentation results using the three methods. Over-segmentation of the FCN and the U-Net (top row) was reduced by post-processing using a level set segmentation algorithm (middle row). Colors in the bottom row were selected randomly to indicate different mitochondria. (B) Another example, same layout as in (A). (A-B) Scale bars: 2 μm .

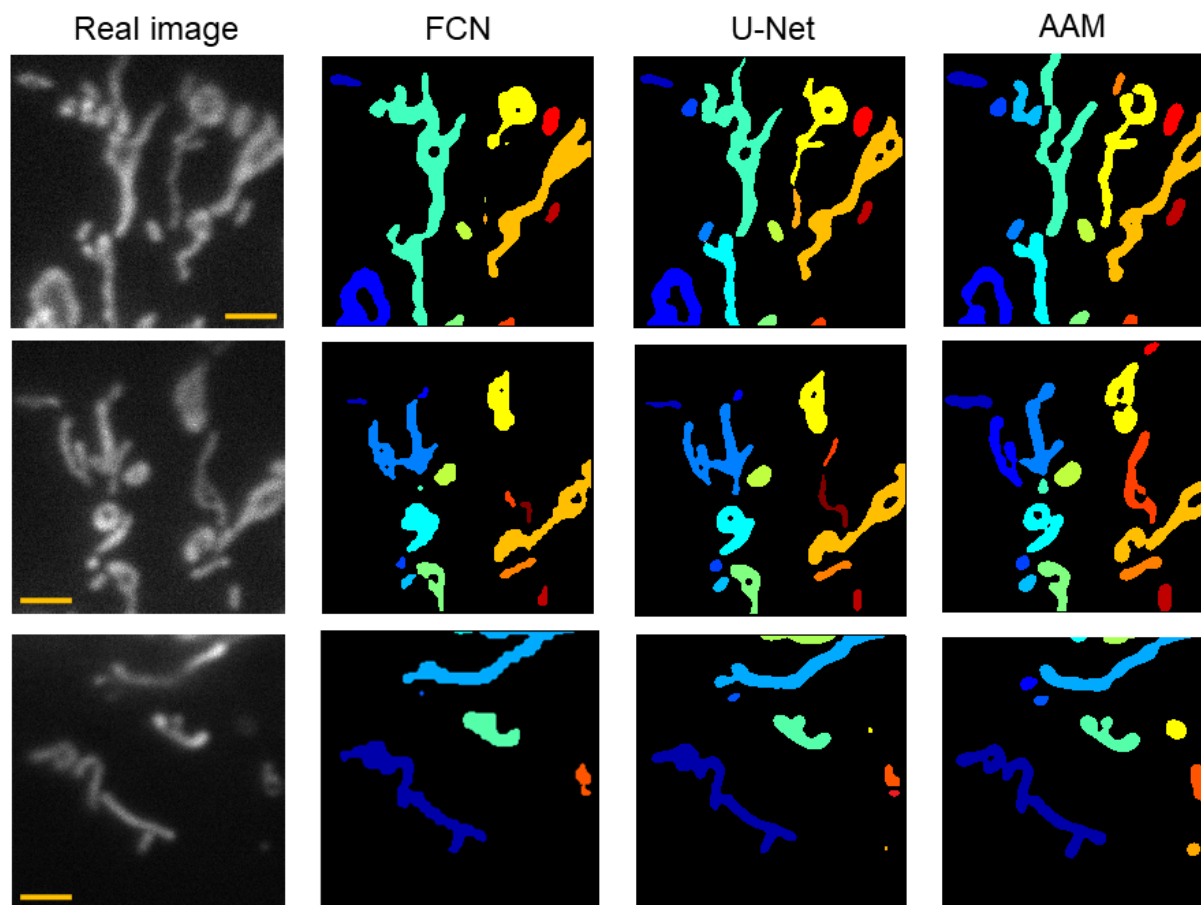


Figure S4. Additional results of segmentation of real images using FCN, U-Net, and AAM without post-processing. Three examples of real images of mitochondria segmented using the three methods. No post-processing was applied. Colors in the bottom row were selected randomly to indicate different mitochondria. Scale bars: 2 μm .

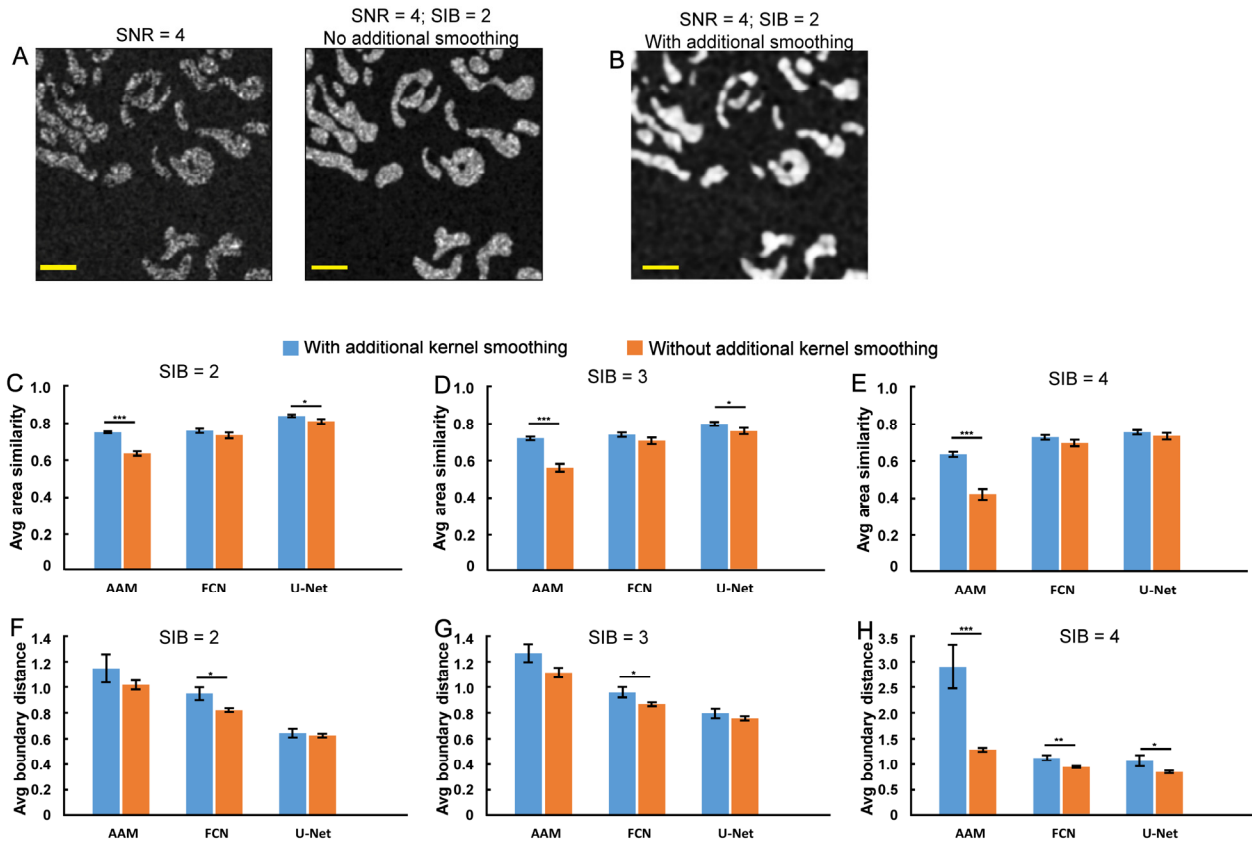


Figure S5. Examining influence of synthetic image generation on performance of FCN and U-Net. (A) Left panel: Synthetic image with an SNR of 4 without additional space-invariant blurring. Right panel: a synthetic image with an SNR of 4 and a space-invariant blurring (SIB) of 2. Both images were generated from the same ground truth in which intensities of individual pixels were set independently. (B) A synthetic image generated with the same ground truth as in (A) but with additional smoothing of pixel intensities inside mitochondria by a Gaussian kernel with $\sigma=1$. (A-B) Scale bars: 2 μm . (C) Comparison of average area similarity of AAM, FCN, and U-Net under SIB = 2; p-values of pairwise comparison: AAM: 1.609×10^{-10} , FCN: 0.202; U-Net: 0.038. (D) Comparison of average area similarity of AAM, FCN, and U-Net under SIB = 3; p-values of pairwise comparison: AAM: 5.918×10^{-9} , FCN: 0.162; U-Net: 0.042. (E) Comparison of average area similarity of AAM, FCN, and U-Net under SIB = 4; p-values of pairwise comparison: AAM: 1.079×10^{-8} , FCN: 0.159; U-Net: 0.282. (F) Comparison of average boundary distance of AAM, FCN, and U-Net under SIB = 2; p-values of pairwise comparison: AAM: 0.276, FCN: 0.017; U-Net: 0.639. (G) Comparison of average boundary distance of AAM, FCN, and U-Net under SIB = 3; p-values of pairwise comparison: AAM: 0.070, FCN: 0.043; U-Net: 0.360. (H) Comparison of average boundary distance of AAM, FCN, and U-Net under SIB = 4; p-values of pairwise comparison: AAM: 4.610×10^{-4} , FCN: 1.132×10^{-3} ; U-Net: 0.034. Notation for p-values: ***: $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

Table S1. p-values of performance comparison of FCN versus U-Net in Figure S2

	Area Similarity (FCN vs U-Net)			Boundary Distance (FCN vs U-Net)		
	TS-SNR = 1	TS-SNR = 4	TS-SNR = 8	TS-SNR = 1	TS-SNR = 4	TS-SNR = 8
TR-SNR = 8	0.997	0.471	4.430×10^{-17}	0.851	0.016	6.375×10^{-29}
TR-SNR = 4	3.300×10^{-15}	1.092×10^{-7}	2.402×10^{-17}	5.377×10^{-21}	6.234×10^{-10}	8.935×10^{-29}

- 1) *TR-SNR*: SNR of training data; *TS-SNR*: SNR of testing data;
- 2) Cells in green indicate a match of TR-SNR with TS-SNR; Cells in brown indicate a mismatch of TR-SNR with TS-SNR;

Table S2. p-values of performance comparison of FCN under different conditions in Figure S2

	Area Similarity (FCN)			Boundary Distance (FCN)		
	TS-SNR = 8 vs TS-SNR = 4	TS-SNR = 8 vs TS-SNR = 1	TS-SNR = 4 vs TS-SNR = 1	TS-SNR = 8 vs TS-SNR = 4	TS-SNR = 8 vs TS-SNR = 1	TS-SNR = 4 vs TS-SNR = 1
TR-SNR = 8	9.610×10^{-54}	6.015×10^{-55}	0.402	5.201×10^{-13}	1.480×10^{-10}	0.0326
TR-SNR = 4	0.468	4.687×10^{-19}	6.772×10^{-21}	0.0198	2.124×10^{-24}	2.630×10^{-11}

- 1) *TR-SNR*: SNR of training data; *TS-SNR*: SNR of testing data;
- 2) Cells in green indicate comparison between a match case and a mismatch case. Cells in light brown indicate comparison between two mismatch cases.

Table S3. p-values for performance comparison of U-NET under different conditions in Figure S2

	Area Similarity (U-Net)			Boundary Distance (U-Net)		
	TS-SNR = 8 vs TS-SNR = 4	TS-SNR = 8 vs TS-SNR = 1	TS-SNR = 4 vs TS-SNR = 1	TS-SNR = 8 vs TS-SNR = 4	TS-SNR = 8 vs TS-SNR = 1	TS-SNR = 4 vs TS-SNR = 1
TR-SNR = 8	2.944×10^{-68}	1.319×10^{-68}	0.903	2.794×10^{-21}	2.828×10^{-29}	0.000221
TR-SNR = 4	2.509×10^{-10}	8.578×10^{-21}	5.193×10^{-13}	6.558×10^{-13}	2.631×10^{-33}	1.131×10^{-17}

- 1) *TR-SNR*: SNR of training data; *TS-SNR*: SNR of testing data;
- 2) Cells in green indicate comparison between a match case and a mismatch case. Cells in light brown indicate comparison between two mismatch cases.