nature methods

Deep learning enables cross-modality super-resolution in fluorescence microscopy

Author: Hongda Wang, Yair Rivenson, Yiyin Jin, Zhensong Wei, Ronald Gao, Harun Günaydın, Laurent A. Bentolila, Comert Kural and Aydogan Ozcan

University of California, Los Angeles (UCLA)

Published 17 December 2018

Slides compiled by Mengzhou Li

A successful application of DL in optical SR

Benefits:

- Field of view extended
- Throughput of microscope increased
- Resolution enhancement
- SNR improvement
- Different modality images transformation

A successful application of DL in optical SR

Results:

- Wide-field

  10x/0.4-NA to 20x/0.75-NA

- Confocal to STED

  ~290nm to ~110nm

- TIRF to TIRF-SIM

  resolution enhancement

Wide-field(10x/0.4-NA to 20x/0.75-NA)

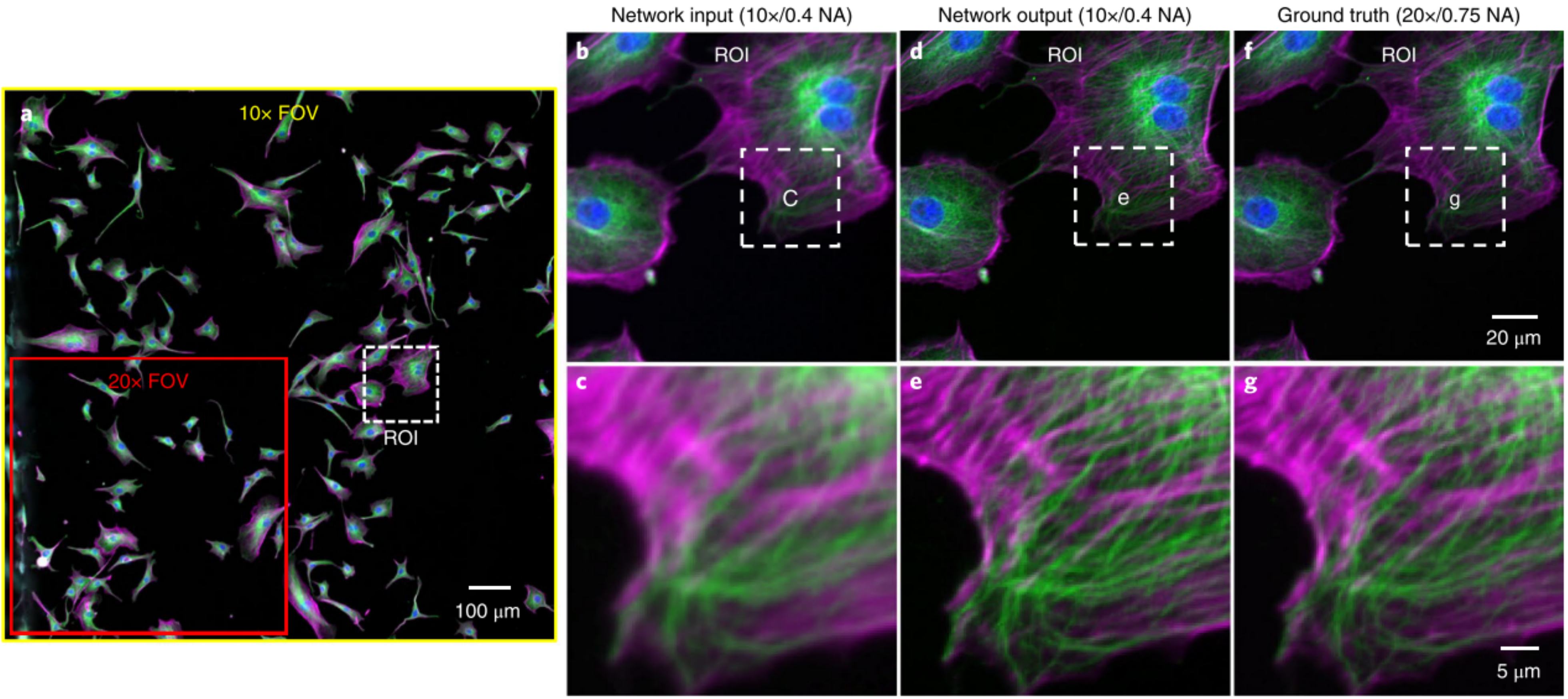


Fig. 1 | Deep-learning-based super-resolved images of bovine pulmonary artery endothelial cells (BPAECs). a, Network input image acquired with a 10×/0.4-NA objective lens. **b**-**g**, Smaller ROIs are magnified and shown in (**b**,**c**) network input, (**d**,**e**) network output, and (**f**,**g**) ground truth (20×/0.75-NA). Experiments were repeated with >250 images, achieving similar results. Color map: magenta for F-actin, green for microtubules, blue for nuclei.

3 gray-level models trained for 3 channels, then combined to get the color image.

3 channels: DAPI, FITC and TxRED

Confocal to STED (~290nm to ~110nm)

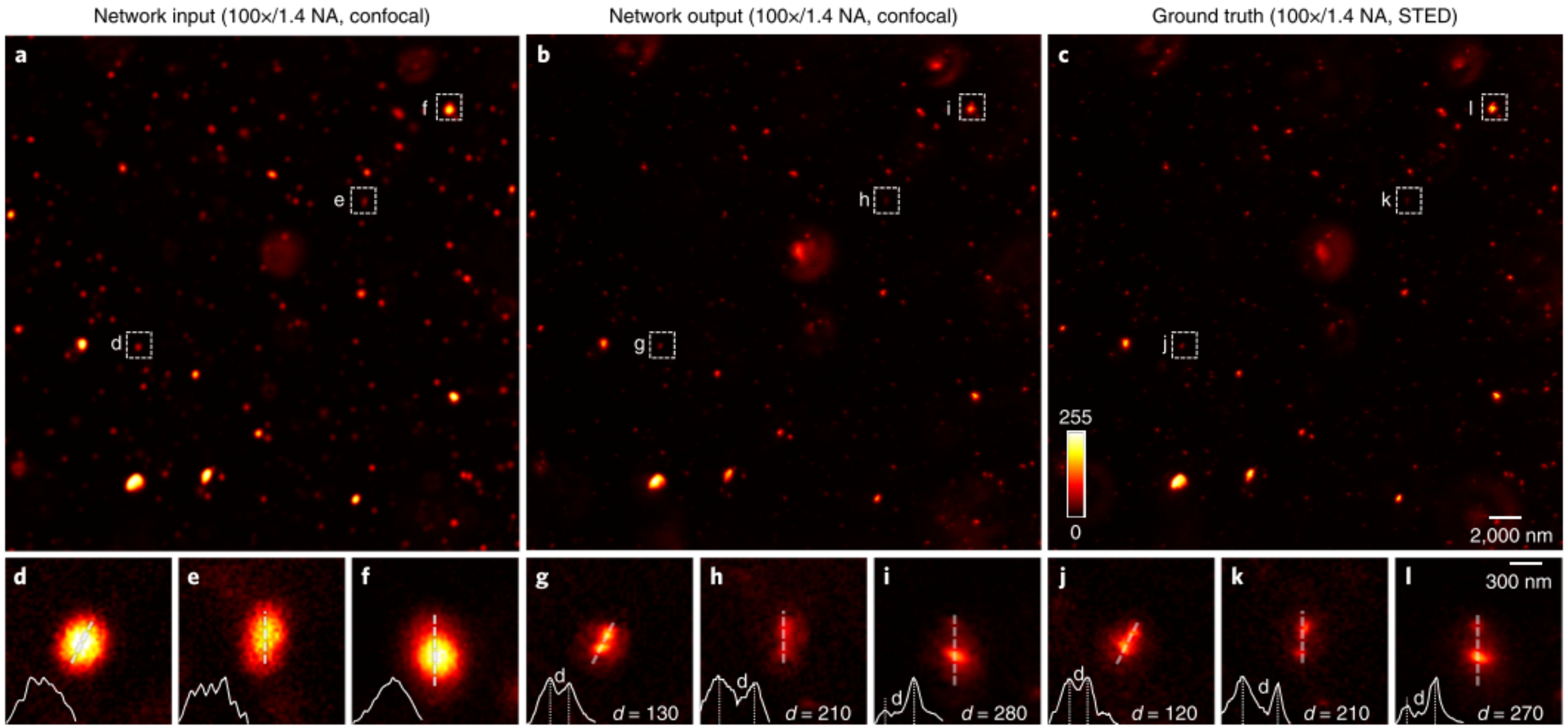


Fig. 3 | Image resolution improvement beyond the diffraction limit: from confocal microscopy to STED. a–c, A diffraction-limited confocal microscope image is used as input to the network and is super-resolved to blindly yield (**b**) the network output, which is comparable to (**c**) a STED image of the same FOV, used as the ground truth. **d–f**, Examples of closely spaced nano-beads that cannot be resolved by confocal microscopy. **g–l**, The trained neural network takes **d–f** as input and resolves the individual beads (**g–i**), very well agreeing with STED microscopy images (**j–l**). The cross-sectional profiles reported in **d–l** are extracted from the original images. Peak-to-peak distance (*d*) in these cross-sectional profiles is reported in nanometers. Also see Fig. 5 for further quantification of the performance of the deep network on confocal images, and its comparison to STED. Experiments were repeated with 75 images, achieving similar results.

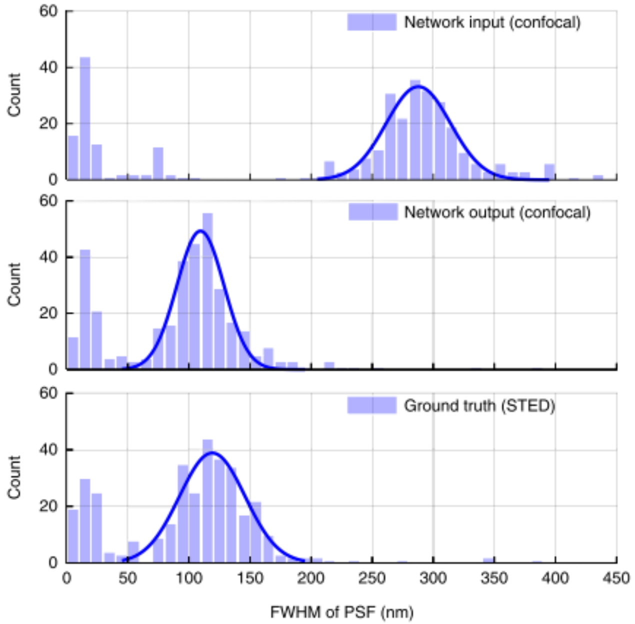
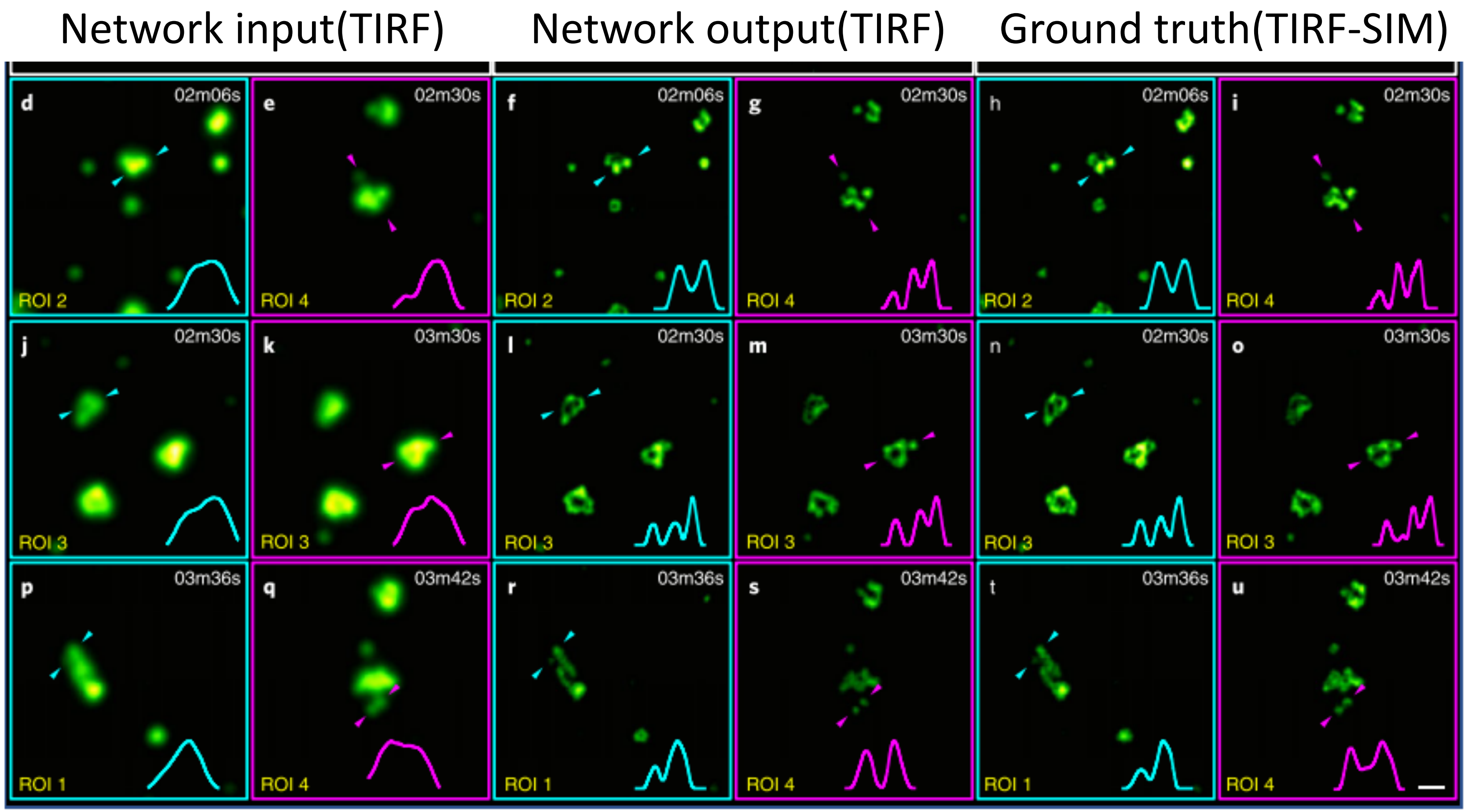
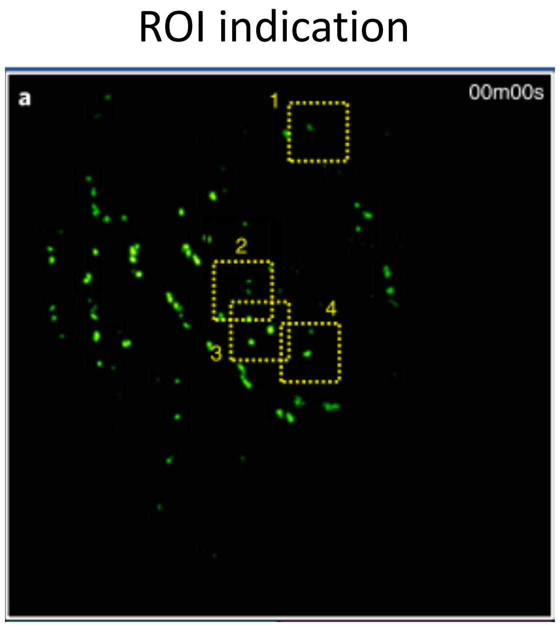


Fig. 4 | PSF characterization, before and after the network, and its comparison to STED. We extracted more than 400 bright spots from the same locations of the network input (confocal), network output (confocal), and the corresponding ground truth (STED) images. Each one of these spots was fit to a 2D Gaussian function, and the corresponding FWHM distributions are shown in each histogram. These results show that the resolution of the network output images is significantly improved from ~290 nm (top row: network input using a confocal microscope) to ~110 nm (middle row: network output), which provides a very good fit to the ground truth STED images of the same nano-particles, summarized in the bottom row.

TIRF to TIRF-SIM





Networks

Generative adversarial network (GAN)

- Loss function:

  adversarial loss + two regularization terms(MSE + SSIM)

$$\mathcal{L}(G;D) = -\log D(G(x)) + \lambda \times \mathrm{MSE}(G(x), y)$$
$$-\nu \times \log[(1 + \mathrm{SSIM}(G(x), y))/2]$$
$$\mathcal{L}(D;G) = -\log D(y) - \log[1 - D(G(x))]$$

The MSE loss and the SSIM loss were accommodated to be ~1-10% of the *L(G;D)*

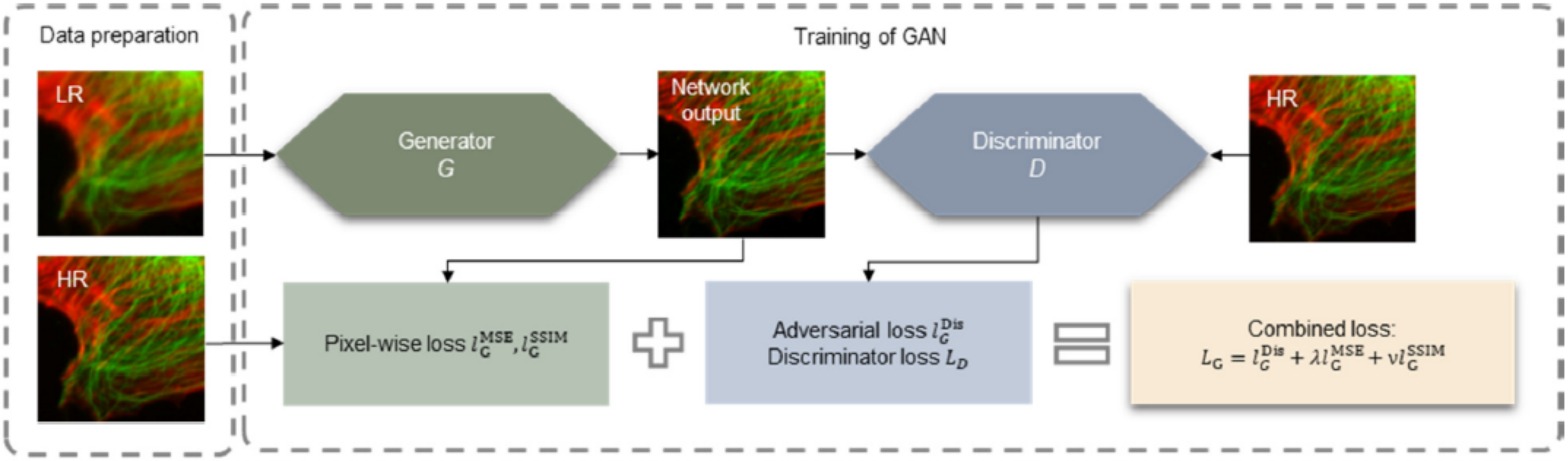
- Generative model (using U-net)

  generator: 4 down-sampling blocks + 4 up-sampling blocks

- Discriminative model

  conv layer + 5 conv blocks + average pooling layer + 2 FC layer

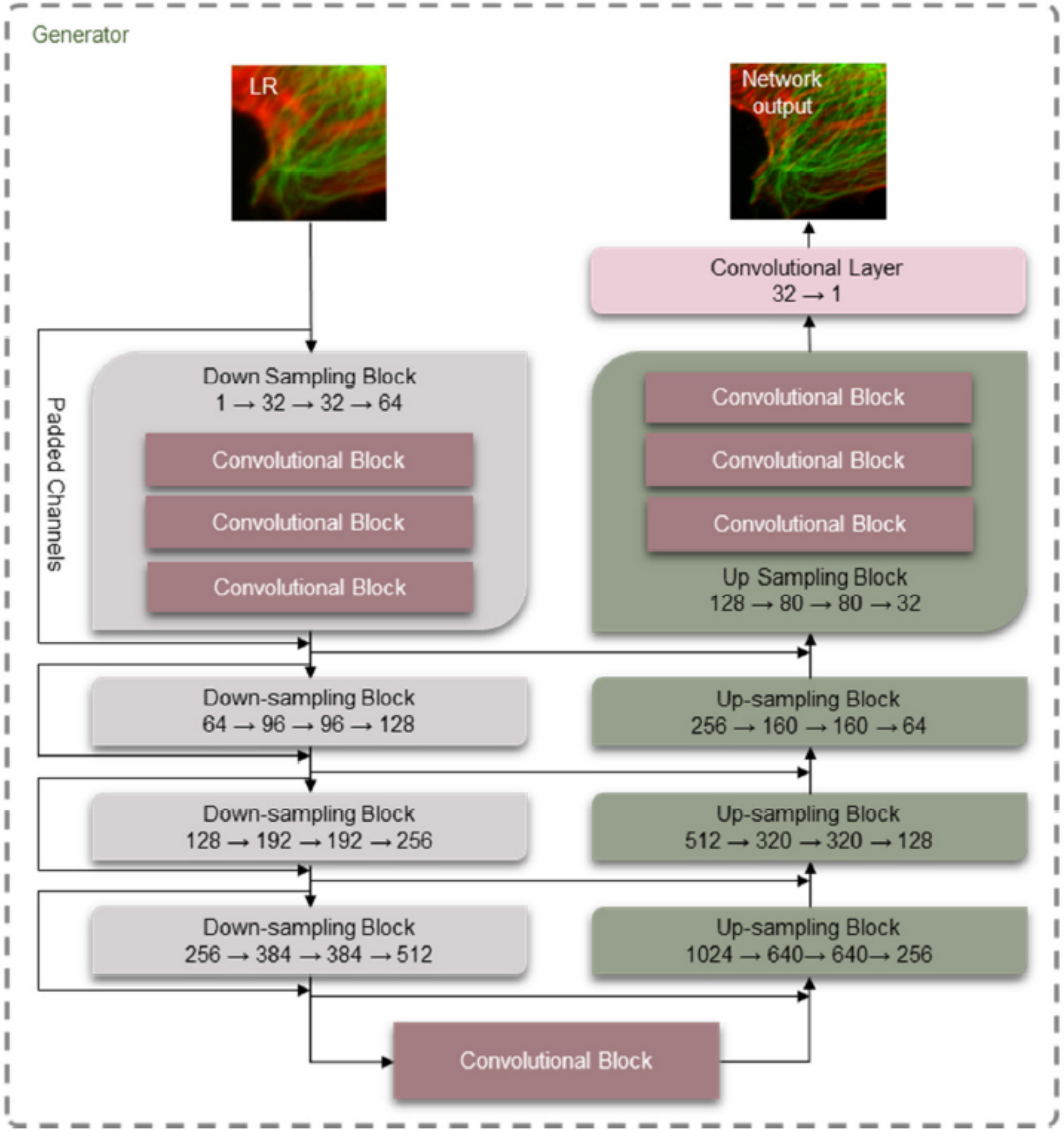
Networks - GAN



$$\mathcal{L}(G;D) = -\log D(G(x)) + \lambda \times \mathrm{MSE}(G(x), y) - \nu \times \log[(1 + \mathrm{SSIM}(G(x), y))/2]$$

$$\mathcal{L}(D;G) = -\log D(y) - \log[1 - D(G(x))]$$

$$\mathrm{SSIM}(x, y) = \frac{(2\mu_x \mu_y + c_1)(2\sigma_{x,y} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$



- Down sampling block

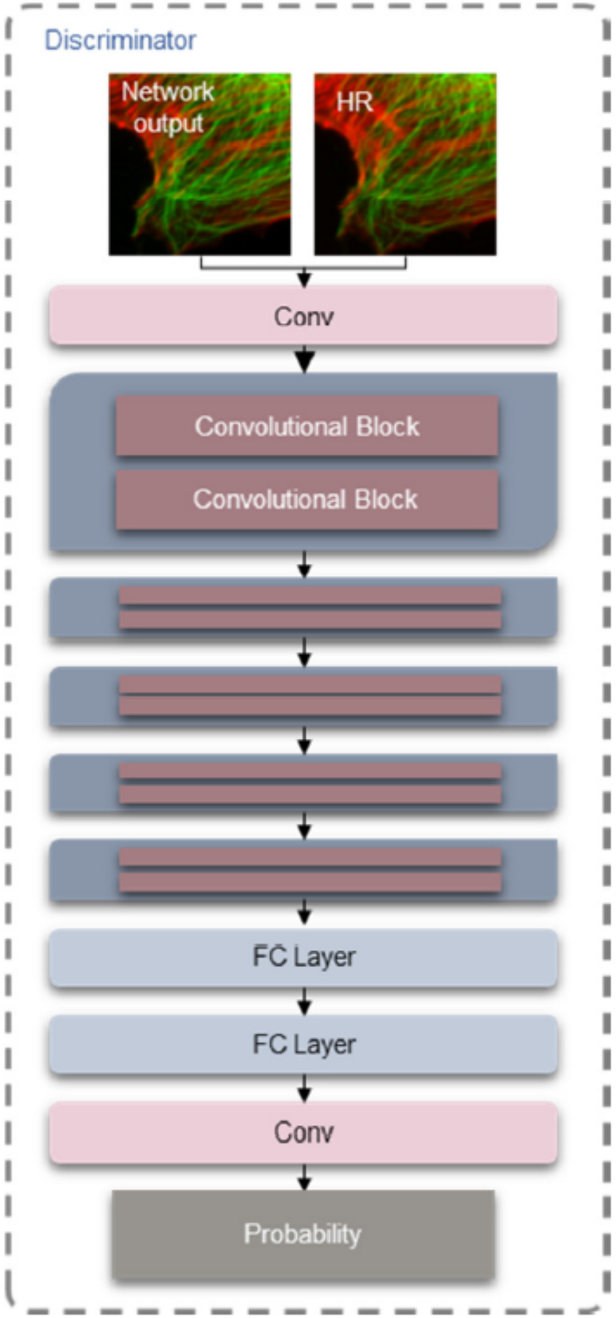
$$\begin{aligned} x_k &= x_{k-1} + \mathrm{LReLU}\left[\mathrm{Conv}\{\mathrm{LReLU}\left[\mathrm{Conv}\{\mathrm{LReLU}\left[\mathrm{Conv}\{x_{k-1}\}\right]\}\right]\}\right], \\ k &= 1, 2, 3, 4 \end{aligned}$$

- Average pooling
- Leaky rectified linear unit activation function

$$\mathrm{LReLU}(x;\alpha) = \max(0, x) - \alpha \times \max(0, -x)$$

- Up sampling block

$$\begin{aligned} y_k &= \mathrm{LReLU} \\ &\quad \left[\mathrm{Conv}\{\mathrm{LReLU}\left[\mathrm{Conv}\{\mathrm{LReLU}\left[\mathrm{Conv}\{\mathrm{Concat}(x_{5-k}, y_{k-1})\}\right]\}\right]\}\right], \\ k &= 1, 2, 3, 4 \end{aligned}$$



- Convolutional blocks

$$z_k = \mathrm{LReLU}\,[\mathrm{Conv}\{\mathrm{LReLU}\,[\mathrm{Conv}\{z_{k-1}\}]\}],\, k = 1, 2, 3, 4, 5$$

- Average pooling layer

- Sigmoid activation function

$$D(z) = \frac{1}{1 + \exp(-z)}$$

Keys to this success

- Optimal data
- Carefully registration and alignment
- One sample type one model
- Insight

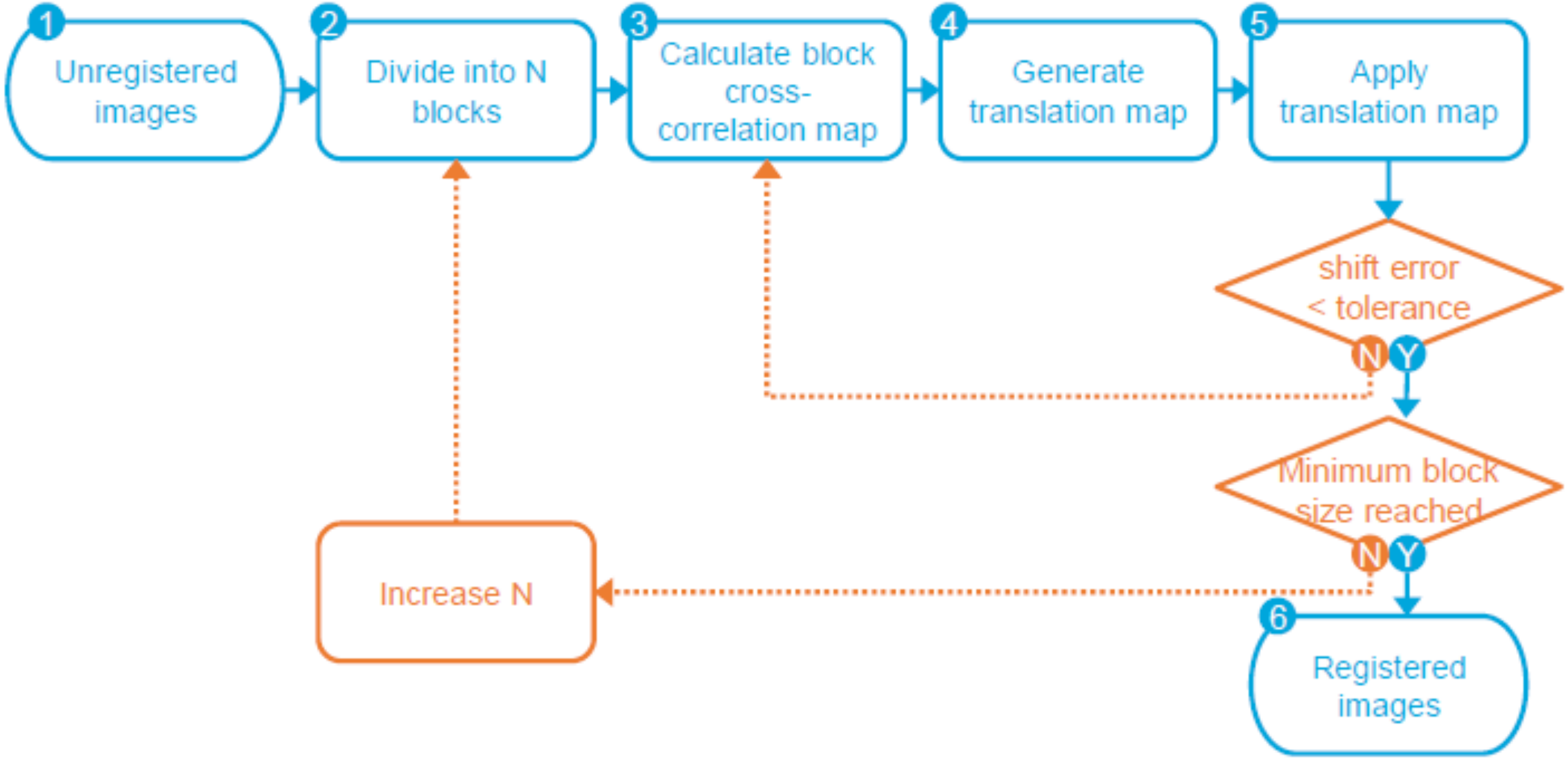
Optimal data

Super-resolution network	Number of training image pairs	Number of validation image pairs	Number of testing image pairs
Wide-field (TxRed)	1945	680	94
Wide-field (FITC)	1945	680	94
Wide-field (DAPI)	1945	680	94
Confocal-STED (nanobeads)	607	75	75
Confocal-STED (transfer learning)	1100	100	30
TIRF-SIM	3003	370	1100

- To increase confocal- STED SNR

16 times line average and 30 times frame average for nano-beads;

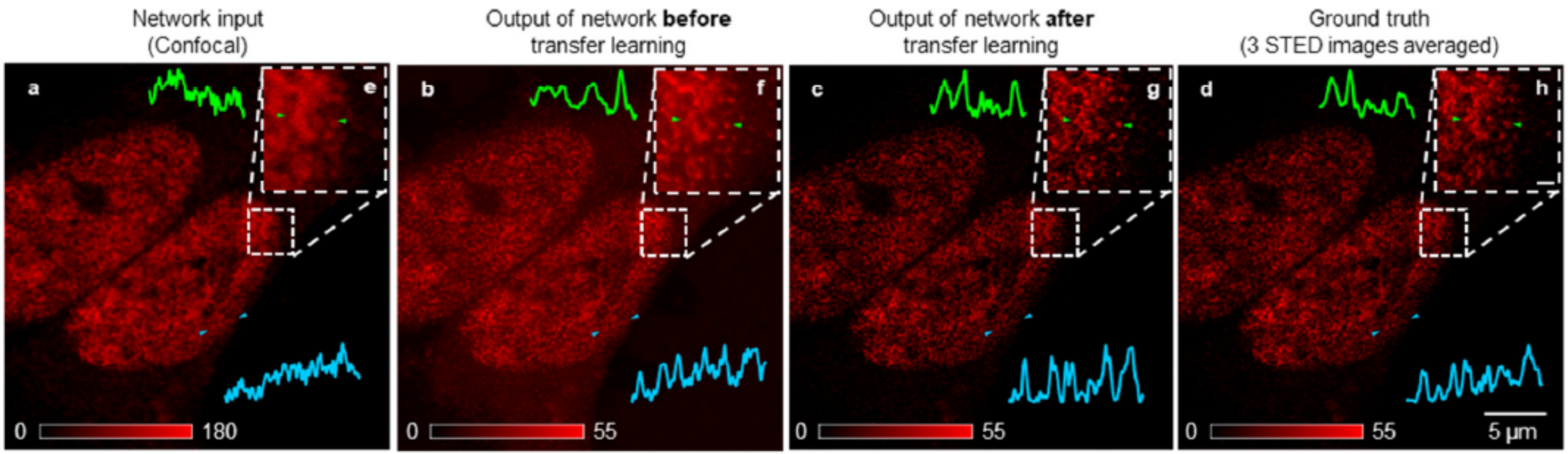
Careful Registration and alignment



- Cross-correlation
- 2D Gaussian function fitting

To achieve 0.1 pixel accuracy

One sample type one model



- Trained only with beads,
- Try imaging with nuclei

- Adding cell nuclei for training
- Image nuclei

Insight

- Taken the PSF as a probability density function

  $P(x, y)$ represents the probability of photons emitted from an ideal point source on the sample to arrive at a certain displacement on the detector;

- Data distribution transformation

  input data distribution $X(P_{\mathrm{LR}}(x, y))$

  output data distribution $Y(P_{\mathrm{HR}}(x, y))$

- Statistically separate out noise patterns to achieve resolution enhancement

Thanks for your attention