Deep learning for dense single-molecule localization microscopy

Elias Nehme[1,2], Lucien E. Weiss[2], Eran Hershko[1], Daniel Freedman[3], Racheli Gordon[2], Boris Ferdman[2], Tomer Michaeli[1], and Yoav Shechtman[2]

[1]Department of Electrical Engineering, Technion, 32000 Haifa, Israel
[2]Department of Biomedical Engineering, Technion, 32000 Haifa, Israel
[3]Google Research, Google, Haifa, Israel

In conventional microscopy, the spatial resolution of an image is bounded by Abbe's diffraction limit, corresponding to approximately half the optical wavelength. Over the last decade, super resolution methods have revolutionized biological imaging, enabling the observation of cellular structures at the nanoscale. These include the popular localization microscopy methods, like photo-activated localization microscopy ((F)PALM) [1, 2] and stochastic optical reconstruction microscopy (STORM) [3]. However, despite the great advancement, existing localization microscopy methods are still limited in their acquisition and post-processing speeds, and in their ability to extract 3D and multicolor properties of the imaged samples.

1 Deep localization microscopy

Localization microscopy relies on acquiring a sequence of diffraction-limited images of blinking fluorophores (Fig. 1), so that each frame contains a sparse set of point-spread functions (PSFs). Using e.g. Gaussian fitting, the emitters can be localized with very high precision, and by combining the recovered emitter positions from all frames, a super-resolved image is produced with resolution typically an order of magnitude better than the diffraction limit (down to tens of nanometers).

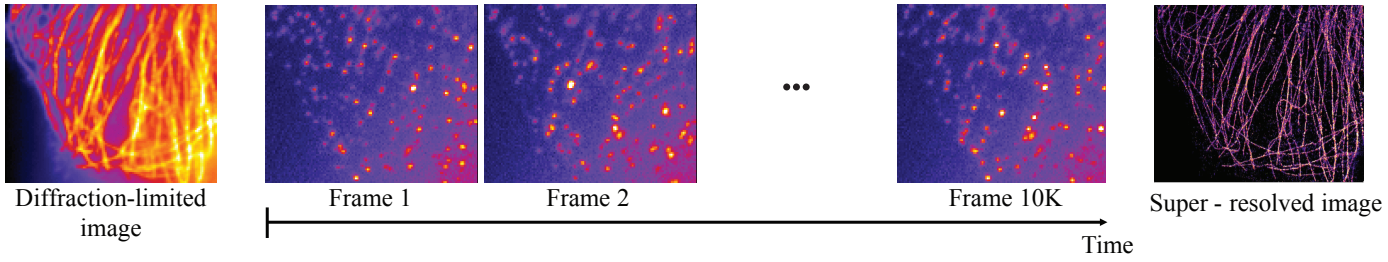

Fig. 1. Super resolution microscopy via single-molecule localization. By acquiring a sequence of diffraction-limited images and super-localizing single emitters via Gaussian fitting, the resolution of an optical microscope can be improved by a factor of ~ 10.

Most existing reconstruction methods for localization microscopy require that the emitters in each frame be sufficiently sparse, which leads to a long acquisition time (seconds to minutes), and limits the ability to capture fast dynamics of sub-wavelength processes within live cells. Various algorithms have been developed to handle overlapping PSFs. While successful localization of densely-spaced emitters has been demonstrated, all existing methods suffer from two fundamental drawbacks: data-processing time and sample-dependent parameter tuning.

Recently we demonstrated precise, fast, parameter-free, super-resolution image reconstruction by harnessing Deep-Learning [4] (highlighted also in Nature methods [5]). In this work, we employed a fully convolutional encoder-decoder architecture for super-resolution image reconstruction from dense fields of overlapping emitters. Our method, dubbed Deep-STORM, does not explicitly localize emitters. Instead, it directly creates a super-resolved image from the raw data. To train the network we generated simulated training examples with matching experimental conditions, namely, camera specifications, PSF model, approximate signal-to-noise ratio (SNR), and the expected emitter density.

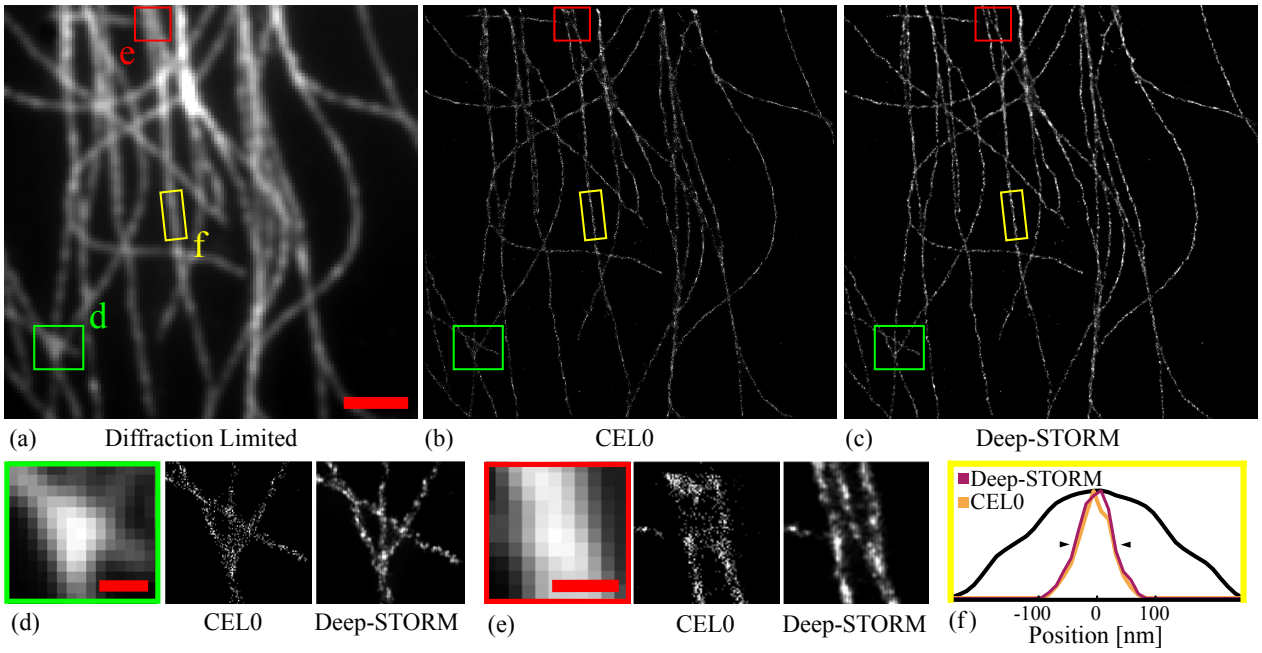

Fig. 2. Reconstruction of experimentally measured microtubules. (a) Sum of the acquisition stack (361 dense frames). Scale bar is 2 μm. (b) Reconstruction by CEL0 [6]. (c) Reconstruction by Deep-STORM. (d)-(e) Magnified views of two selected regions. Scale bars are 0.5 μm. (f) The width projection of the highlighted yellow region. The attained FWHM (black triangles) for CEL0 was 61 nm and 67 nm for Deep-STORM. The black line shows the diffraction-limited projection.

We validated Deep-STORM on super-resolution data, and benchmarked against a high-performance multi-emitter fitting algorithm (CEL0 [6]). Specifically, we tested the result of Deep-STORM on experimental high-density data obtained from Sage et al. [7], training solely on simulated data with similar experimental conditions - namely, SNR and emitter density. Deep-STORM resolves nearby lines and fine structures, and produces more continuous shapes compared to the output of CEL0 (Fig. 2). Deep-STORM not only yields image reconstruction results that are comparable to or better than leading algorithms, but also does so $\sim 1-3$ orders of magnitude faster.

Next, motivated by the success in solving the 2D high-density emitter-fitting problem, we extended our approach to 3D [8]. Various methods proposed ways to enable the imaging of three-dimensional (3D) objects, a prominent one being point spread function (PSF) engineering for microscopic particle localization [9]. In PSF engineering, the PSF is modified by additional optical elements, e.g. a phase mask, inserted into the detection path of a standard microscope (Fig. 3 (a)). This enables designing PSFs whose shapes change distinctively as a function of depth, thus allowing to recover the axial position of each emitter from a single 2D image.

As in 2D localization microscopy, here as well the number of emitters per frame (and thereby their density) determines the temporal resolution of the method. However, overlapping emitters pose an even more severe algorithmic challenge in the 3D case since encoding the axial position of an emitter introduces additional complexity into the PSF shape and increases its size, especially for large axial ranges (>3 μm) using the Tetrapod PSF [9, 10] (Fig. 3 (b)). To solve the high-density localization problem in 3D, we trained a CNN that receives a 2D image of overlapping Tetrapod PSFs spanning an axial range of 4 μm, and outputs a 3D vacancy grid which is post-processed to compile a list of localizations (Fig. 3 (c)). We compared the CNN performance to the leading state-of-the-art methods and found it to be superior by a wide margin [8].

2 Deep optics design

The Tetrapod PSF was optimized for the single emitter case, designed by maximizing the Fisher Information of the system [9, 10]. However, when considering the multiple-emitter case, an intriguing question arises: What is the optimal PSF for high density 3D localization over a large axial range? To answer this question we need to rethink the design metric; extending the Fisher Information criterion [9] to account for emitter density is not-trivial, and while it is intuitive that a smaller-footprint PSF would be preferable for dense emitters, it is not clear how to mathematically balance this

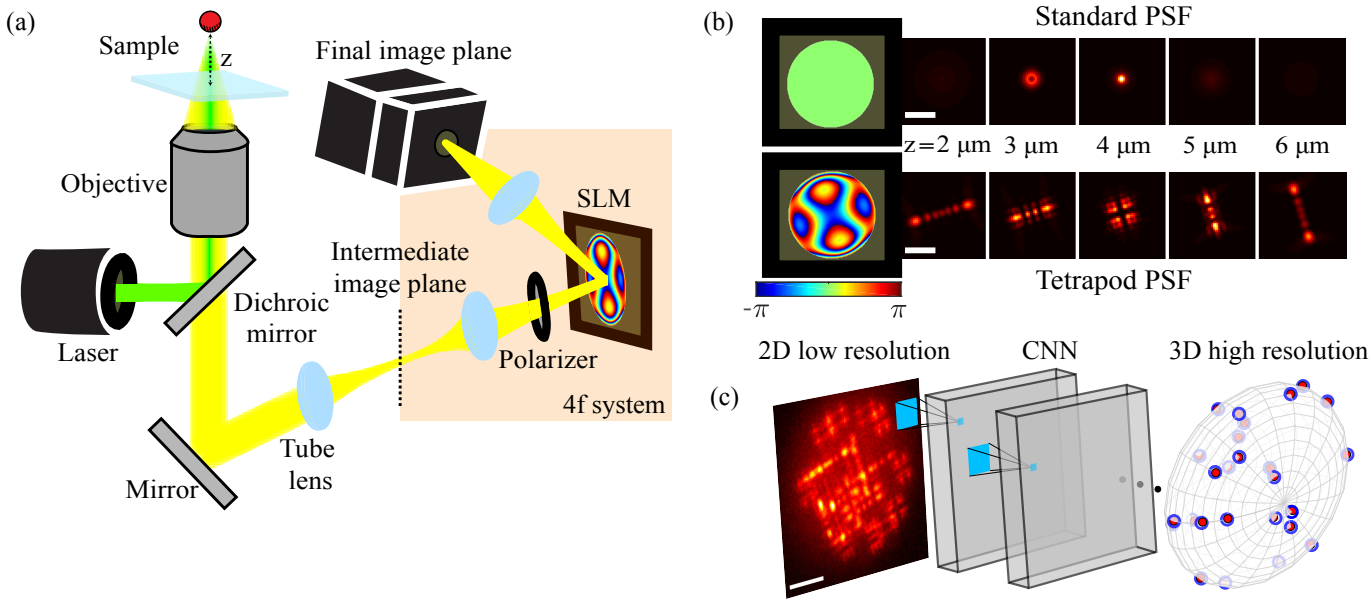

Fig. 3. Optical setup for 3D localization microscopy. (a) The light emitted from a fluorescent microscopic particle is collected by the objective and focused through the tube lens into an image at the intermediate image-plane. This plane is extended using a 4f system with an SLM placed at the Fourier plane in between the 2 4f lenses. (b) The phase mask implemented on the SLM dictates the PSF's axial profile measured on the final image plane. (c) After training, our CNN receives a 2D low resolution image of overlapping Tetrapod PSFs and outputs a 3D high-resolution volume which is translated to a list of 3D localizations. Blue empty spheres denote simulated GT positions along the surface of an ellipsoid. Red spheres denote CNN detections. Scale bar is 3 μm.

demand with the requirement for high localization precision per emitter.

Our PSF-design logic is based on the following: since we have already established that a CNN yields superior reconstruction for high-density 3D localization, *we are interested in a PSF (encoder) that would be optimally localized by a CNN (decoder)*. Therefore, we adopt a co-design approach (Fig. 4 (a)). To jointly optimize the PSF and the localization CNN, we introduce a differentiable physical simulation layer, which is parametrized by a phase mask that dictates the microscope's PSF. This layer encodes 3D point sources to their respective low-resolution 2D image. Then, this image is fed to the localization CNN which decodes it and recovers the underlying 3D source positions. During training, the net is presented with simulated point sources at random locations and, using the difference between the CNN recovery and the simulated 3D positions, we optimize both the phase mask and the localization CNN parameters in an end-to-end fashion. The learned PSF (Fig. 4 (b)) has a small lateral footprint, which is critical for minimizing overlaps at high densities. Moreover, the learned phase mask twists in a spiral trajectory causing the PSF to rapidly rotate throughout the axial range, a trait that was previously shown to be valuable for encoding depth [11].

Finally, of particular importance for biological imaging is real-time, correlative information between multiple species in a sample. Typically, this is achieved by attaching spectrally-distinctive fluorophores to molecules of interest, thus necessitating multicolor imaging. Using an RGB camera is not practical for low-signal applications where photons are precious. Therefore, we explored a similar co-design approach in the context of multicolor localization microscopy [12]. For this application, we learned a phase mask in order to optimize the net's ability to distinguish the colors of single emitters from grayscale images and demonstrated that it is possible to come close to perfect color classification from a grayscale image, even in challenging SNR regimes [12].

References

[1] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, "Imaging intracellular fluorescent proteins at nanometer resolution," *Science*, 2006.

[2] S. T. Hess, T. P. Girirajan, and M. D. Mason, "Ultra-high resolution imaging by fluorescence photoactivation localization microscopy," *Biophysical Journal*, vol. 91, no. 11, pp. 4258 – 4272, 2006.

[3] M. J. Rust, M. Bates, and X. Zhuang, "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)," *Nature Methods*, vol. 3, no. 10, pp. 793–795, 2006.

[4] E. Nehme, L. E. Weiss, T. Michaeli, and Y. Shechtman, "Deep-storm: super-resolution single-molecule microscopy by deep learning," *Optica*, vol. 5, pp. 458–464, Apr 2018.

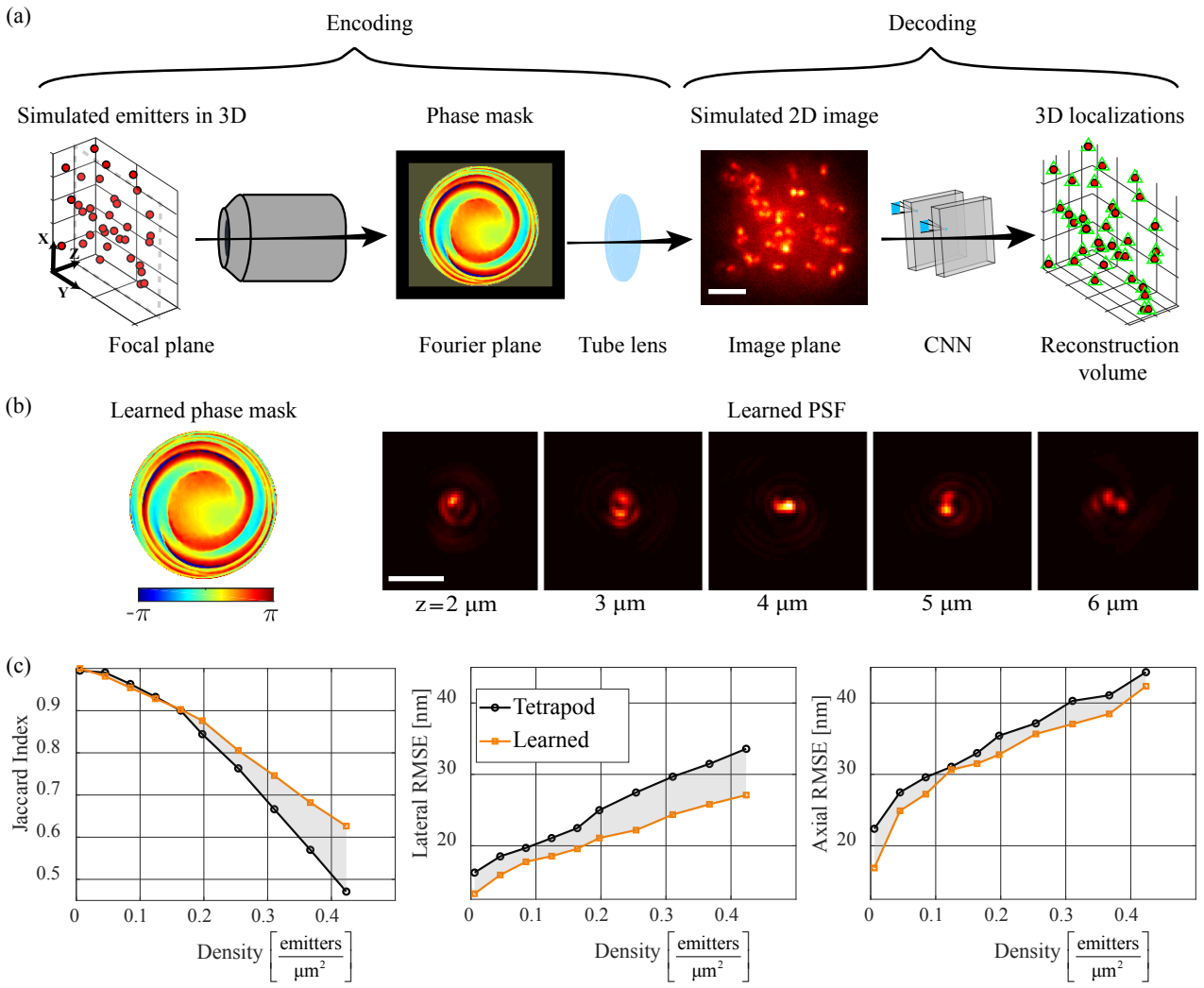

Fig. 4. PSF Learning. **a** Simulated 3D emitter positions are fed to the image formation model to simulate their low resolution CCD image (Encoding). Next, this image is fed to a CNN that tries to recover the simulated emitter positions (Decoding). The difference between the simulated positions and the positions recovered by the CNN is used to jointly optimize the phase mask at the Fourier plane, and the recovery CNN parameters. (b) Learned phase mask (left), with a simulation of the learned PSF as function of the emitter axial position (right). (c) Jaccard index and lateral \axial RMSE comparison between two CNNs with the same architecture, one trained to recover 3D positions from 2D images of the Tetrapod PSF (black), and the second trained to recover 3D positions from 2D images of the learned PSF (orange). Scale bar is 3 μm.

[5] R. Strack, "Deep learning advances super-resolution imaging," *Nature Methods*, vol. 15, p. 403, Jun 2018.

[6] S. Gazagnes, E. Soubies, and L. Blanc-Féraud, "High density molecule localization for super-resolution microscopy using CEL0 based sparse approximation," *ISBI 2017-IEEE International Symposium on Biomedical Imaging*, no. 1, p. 4, 2017.

[7] D. Sage, H. Kirshner, T. Pengo, N. Stuurman, J. Min, S. Manley, and M. Unser, "Quantitative evaluation of software packages for single-molecule localization microscopy," *Nature Methods*, vol. 12, no. 8, pp. 717–724, 2015.

[8] E. Nehme, D. Freedman, R. Gordon, B. Ferdman, T. Michaeli, and Y. Shechtman, "Dense three dimensional localization microscopy by deep learning," *arXiv preprint arXiv:1906.09957*, 2019.

[9] Y. Shechtman, S. J. Sahl, A. S. Backer, and W. Moerner, "Optimal point spread function design for 3d imaging," *Physical review letters*, vol. 113, no. 13, p. 133902, 2014.

[10] Y. Shechtman, L. E. Weiss, A. S. Backer, S. J. Sahl, and W. Moerner, "Precise three-dimensional scan-free multiple-particle tracking over large axial ranges with tetrapod point spread functions," *Nano letters*, vol. 15, no. 6, pp. 4194–4199, 2015.

[11] S. R. P. Pavani, M. A. Thompson, J. S. Biteen, S. J. Lord, N. Liu, R. J. Twieg, R. Piestun, and W. Moerner, "Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function," *Proceedings of the National Academy of Sciences*, vol. 106, no. 9, pp. 2995–2999, 2009.

[12] E. Hershko, L. E. Weiss, T. Michaeli, and Y. Shechtman, "Multicolor localization microscopy and point-spread-function engineering by deep learning," *Optics express*, vol. 27, no. 5, pp. 6158–6183, 2019.