

# DeepSTORM3D: dense 3D localization microscopy and PSF design by deep learning

Elias Nehme<sup>1,2</sup>, Daniel Freedman<sup>3</sup>, Racheli Gordon<sup>2</sup>, Boris Ferdman<sup>2,4</sup>, Lucien E. Weiss<sup>ID 2</sup>, Onit Alalouf<sup>2</sup>, Tal Naor<sup>2</sup>, Reut Orange<sup>2,4</sup>, Tomer Michaeli<sup>1</sup> and Yoav Shechtman<sup>ID 2,4</sup>✉

An outstanding challenge in single-molecule localization microscopy is the accurate and precise localization of individual point emitters in three dimensions in densely labeled samples. One established approach for three-dimensional single-molecule localization is point-spread-function (PSF) engineering, in which the PSF is engineered to vary distinctively with emitter depth using additional optical elements. However, images of dense emitters, which are desirable for improving temporal resolution, pose a challenge for algorithmic localization of engineered PSFs, due to lateral overlap of the emitter PSFs. Here we train a neural network to localize multiple emitters with densely overlapping Tetrapod PSFs over a large axial range. We then use the network to design the optimal PSF for the multi-emitter case. We demonstrate our approach experimentally with super-resolution reconstructions of mitochondria and volumetric imaging of fluorescently labeled telomeres in cells. Our approach, DeepSTORM3D, enables the study of biological processes in whole cells at timescales that are rarely explored in localization microscopy.

**D**etermining the nanoscale positions of point emitters forms the basis of localization microscopy techniques such as single-particle tracking<sup>1,2</sup>, (fluorescence) photoactivated localization microscopy (f)PALM<sup>3,4</sup>, stochastic optical reconstruction microscopy (STORM)<sup>5</sup> and related single-molecule localization microscopy (SMLM) methods. These techniques have revolutionized biological imaging, revealing cellular processes and structures at the nanoscale<sup>6</sup>. Notably, most samples of interest extend in three dimensions, necessitating three-dimensional (3D) localization microscopy<sup>7</sup>.

In a standard microscope, the precise  $z$  position of an emitter is difficult to ascertain because the change of the PSF near the focus is approximately symmetric. Furthermore, outside of this focal range ( $\approx \pm 350$  nm for a high numerical aperture imaging system), the rapid defocusing of the PSF reduces the signal-to-noise ratio (SNR), causing localization precision to quickly degrade. One method to extend the useful  $z$  range and explicitly encode the  $z$  position is PSF engineering<sup>8–10</sup>. Here an additional optical element, for example a phase mask, is placed in the emission path of the microscope, modifying the image formed on the detector<sup>11</sup> (Fig. 1a); the axial position can then be recovered via image processing using a theoretical or experimentally calibrated PSF model<sup>10–16</sup>.

In practically all applications, it is desirable to be able to localize nearby emitters simultaneously. For example, in super-resolution SMLM experiments, the number of emitters localized per frame determines temporal resolution. In tracking applications, PSF overlap from multiple emitters often precludes localization, potentially biasing results in emitter-dense regions. The problem is that localizing overlapping emitters poses a great algorithmic challenge even in two-dimensional (2D) localization and much more so in 3D. Specifically, encoding the axial position of an emitter over large axial ranges ( $> 3 \mu\text{m}$ ) requires the use of laterally large PSFs, for example the Tetrapod<sup>10,17</sup> (Fig. 1b), increasing the possibility of overlap. Consequently, while a variety of methods have been developed to cope with overlapping emitters for the in-focus, standard

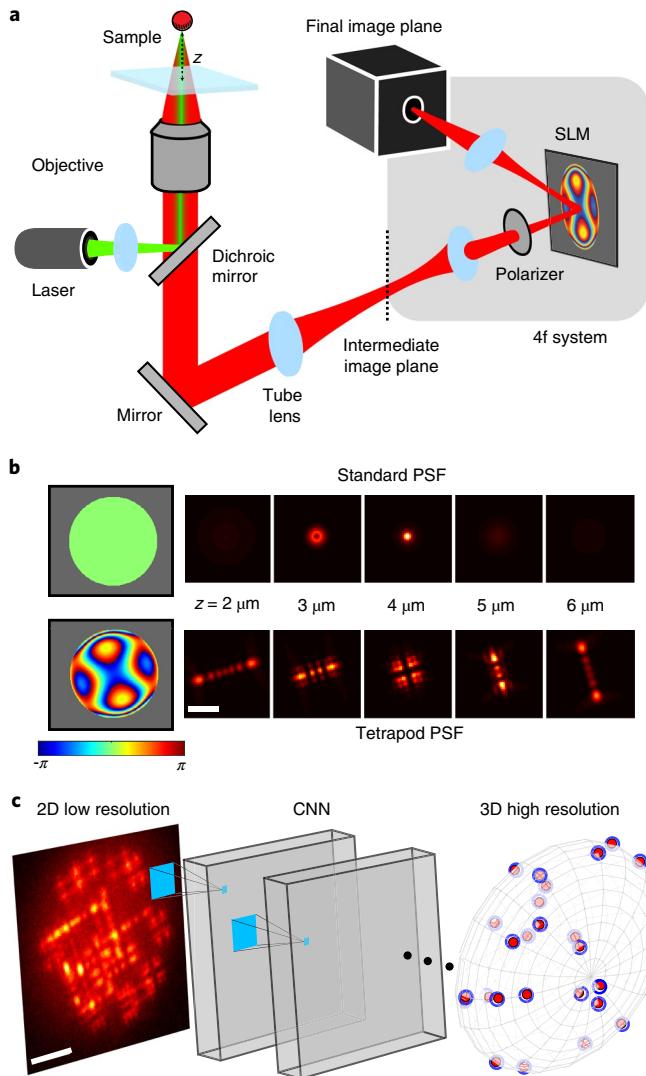
PSF<sup>18–20</sup>, a recent comparison of state-of-the-art software revealed that performance in high-density 3D localization situations is far from satisfactory, even for top-performing algorithms<sup>21</sup>.

Deep learning has proven to be adept at analyzing microscopic data<sup>22–30</sup>, especially for single-molecule localization, handling dense fields of emitters over small axial ranges ( $< 1.5 \mu\text{m}$ )<sup>20,31–38</sup> or sparse emitters spread over larger ranges<sup>39</sup>. Moreover, an emerging application is to jointly design the optical system alongside the data processing algorithm, enabling end-to-end optimization of both components<sup>37,40–47</sup>. Here we present DeepSTORM3D, consisting of two fundamental contributions to high-density 3D localization microscopy over large axial ranges. First, we employ a convolutional neural network (CNN) for analyzing dense fields of overlapping emitters with engineered PSFs, demonstrated with the large-axial-range Tetrapod PSF<sup>10,17</sup>. Second, we design an optimal PSF for 3D localization of dense emitters over a large axial range of  $4 \mu\text{m}$ . By incorporating a physical simulation layer in the CNN with an adjustable phase modulation, we jointly learn the optimal PSF (encoding) and associated localization algorithm (decoding). This approach is highly flexible and easily adapted for any 3D SMLM dataset parameters (emitter density, SNRs and  $z$  range). We quantify the performance of the method by simulation and demonstrate the applicability to 3D biological samples (mitochondria and telomeres).

## Results

**Dense 3D localization with DeepSTORM3D.** 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  $4 \mu\text{m}$  and outputs a 3D grid with a voxel size of  $27.5 \times 27.5 \times 33 \text{ nm}^3$  (Fig. 1c). For architecture details and learning hyper-parameters see Supplementary Notes 2.1 and 4. To compile a list of localizations, we apply simple thresholding, local maximum finding and local averaging on the output 3D grid (Supplementary Note 5).

<sup>1</sup>Department of Electrical Engineering, Technion, Haifa, Israel. <sup>2</sup>Department of Biomedical Engineering, Lorry I. Lokey Center for Life Sciences and Engineering, Technion, Haifa, Israel. <sup>3</sup>Google Research, Haifa, Israel. <sup>4</sup>Russel Berrie Nanotechnology Institute, Technion, Haifa, Israel.  
✉e-mail: [yoavsh@bm.technion.ac.il](mailto:yoavsh@bm.technion.ac.il)



**Fig. 1 | Optical setup and approach overview.** **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 a phase mask placed at the Fourier plane in between the two 4f lenses. **b**, The implemented phase mask (using either a liquid crystal spatial light modulator (LC-SLM) or fabricated fused silica) dictates the shape of the PSF as a function of the emitter's axial position. **c**, After training, our CNN receives a 2D low-resolution image of overlapping PSFs and outputs a 3D high-resolution volume, which is translated to a list of 3D localizations. Blue empty spheres denote simulated ground truth positions along the surface of an ellipsoid. Red spheres denote CNN detections. The Tetrapod PSF is depicted here; however, the approach is applicable to any PSF, including those optimized by the net itself (Fig. 4). Scale bars, 3  $\mu\text{m}$ .

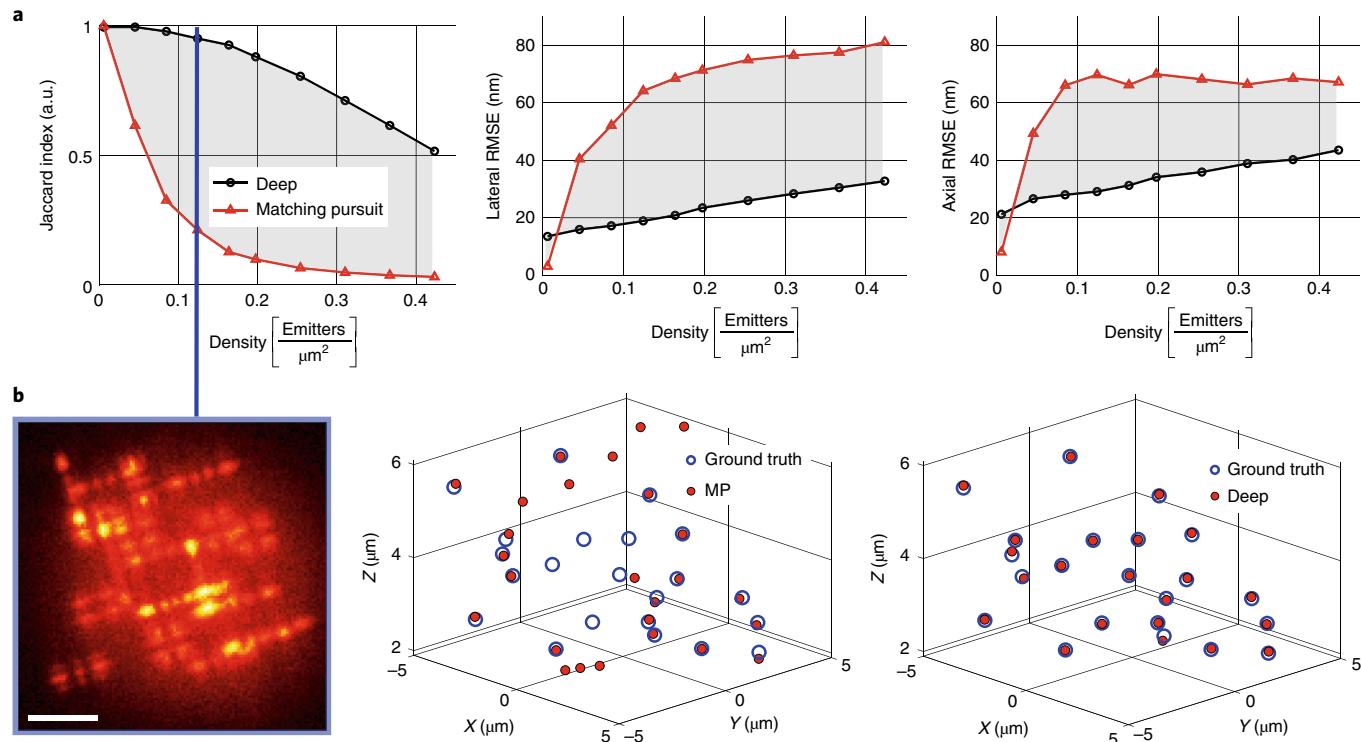
We compare our method to a fit-and-subtract-based matching pursuit (MP) approach<sup>48</sup> (see Supplementary Note 7) as we are unaware of any other methods capable of localizing overlapping Tetrapod PSFs. To quantitatively compare our method with MP solely in terms of density, we simulated emitters with high signal-to-noise ratio (30,000 signal counts, 150 background counts per pixel) at ten different densities ranging from 1 to 75 emitters per  $13 \times 13 \mu\text{m}^2$  field of view (FOV) (for the definition of density see Supplementary Note 1). The results are shown in Fig. 2. As evident in both the Jaccard index<sup>21</sup> (see Supplementary Note 6) and the lateral/axial

root mean square error (RMSE) (Fig. 2a), the CNN achieves remarkable performance in localizing high-density Tetrapods. In the single-emitter (very low density) case, where the performance of the CNN is bounded by the discretization on the 3D grid, the RMSE of the MP localization is lower (better). This is because for a single emitter, MP is equivalent to a continuous maximum likelihood estimator (MLE) (Supplementary Note 7), which is asymptotically optimal<sup>49</sup>, whereas the CNN's precision is bounded by pixilation of the grid (half voxel of 13.75 nm in  $xy$  and 16.5 nm in  $z$ ). However, quickly beyond the single-emitter case, the CNN drastically outperforms MP at both high and low SNR (see Supplementary Note 7.3). A similar result was obtained when compared to a leading single-emitter fitting method<sup>14</sup> applicable also for the multiple-emitter case<sup>21</sup> (see Supplementary Note 8.2). Furthermore, to put our method in context with other existing approaches, we tested DeepSTORM3D on the EPFL Double Helix high-density challenge<sup>21</sup> obtaining favorable results (see Supplementary Note 8.1).

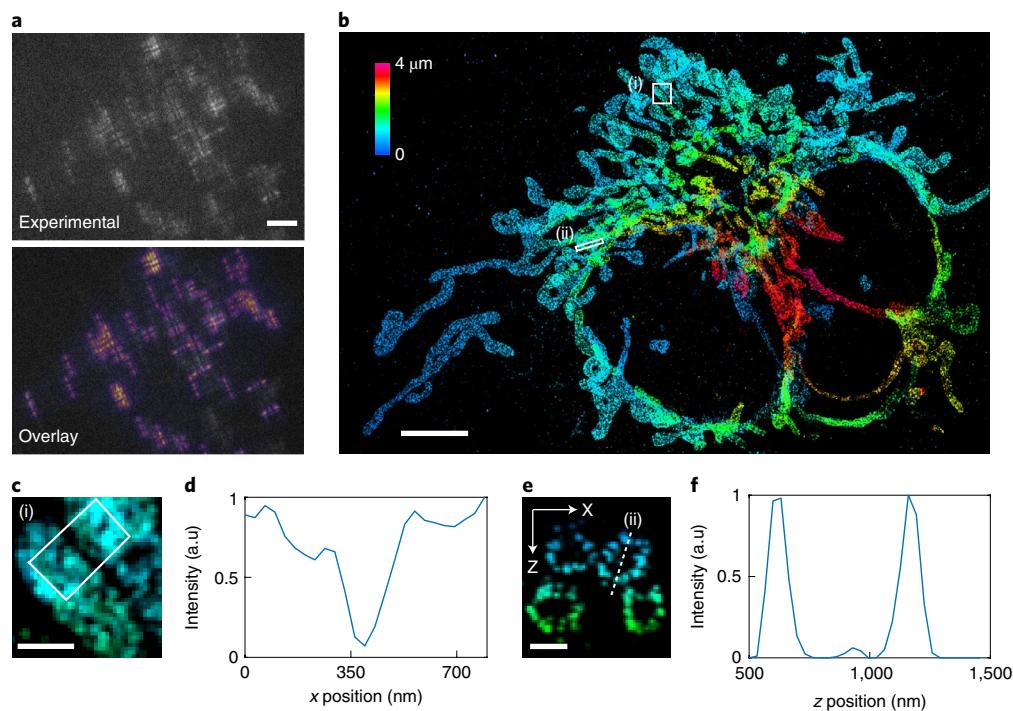
Next, we validated our method for super-resolution imaging of fluorescently labeled mitochondria in COS7 cells (Fig. 3 and Supplementary Videos 1–3). We acquired 20,000 diffraction-limited frames of a  $50 \times 30 \mu\text{m}^2$  FOV and localized them using the CNN in  $\approx 3\text{ h }20\text{ m}$ , resulting in  $\approx 360,000$  localizations. The Tetrapod PSF was implemented using a fabricated fused-silica phase mask (see Supplementary Note 9.1) and the CNN was trained solely on simulations matching the experimental conditions (see Supplementary Note 4.1). The estimated resolution was  $\approx 40$  nm in  $xy$  and  $\approx 50$  nm in  $z$  (see Supplementary Note 9.3). To visually evaluate localization performance in a single frame (Fig. 3a), we regenerated the corresponding 2D low-resolution image and overlaid the recovered image on top of the experimental frame (Fig. 3a and Supplementary Video 1). As seen in the overlay image, the emitter PSFs (3D positions) are faithfully recovered by the CNN. Emitters with an extremely low number of signal photons were ignored. For further acceleration in acquisition time see Supplementary Note 9.2.

**Optimal PSF design for dense 3D imaging.** The Tetrapod is a special PSF that has been optimized for the single-emitter case by Fisher information maximization<sup>10,17</sup>. 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<sup>10</sup> 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 demand with the requirement for high localization precision per emitter.

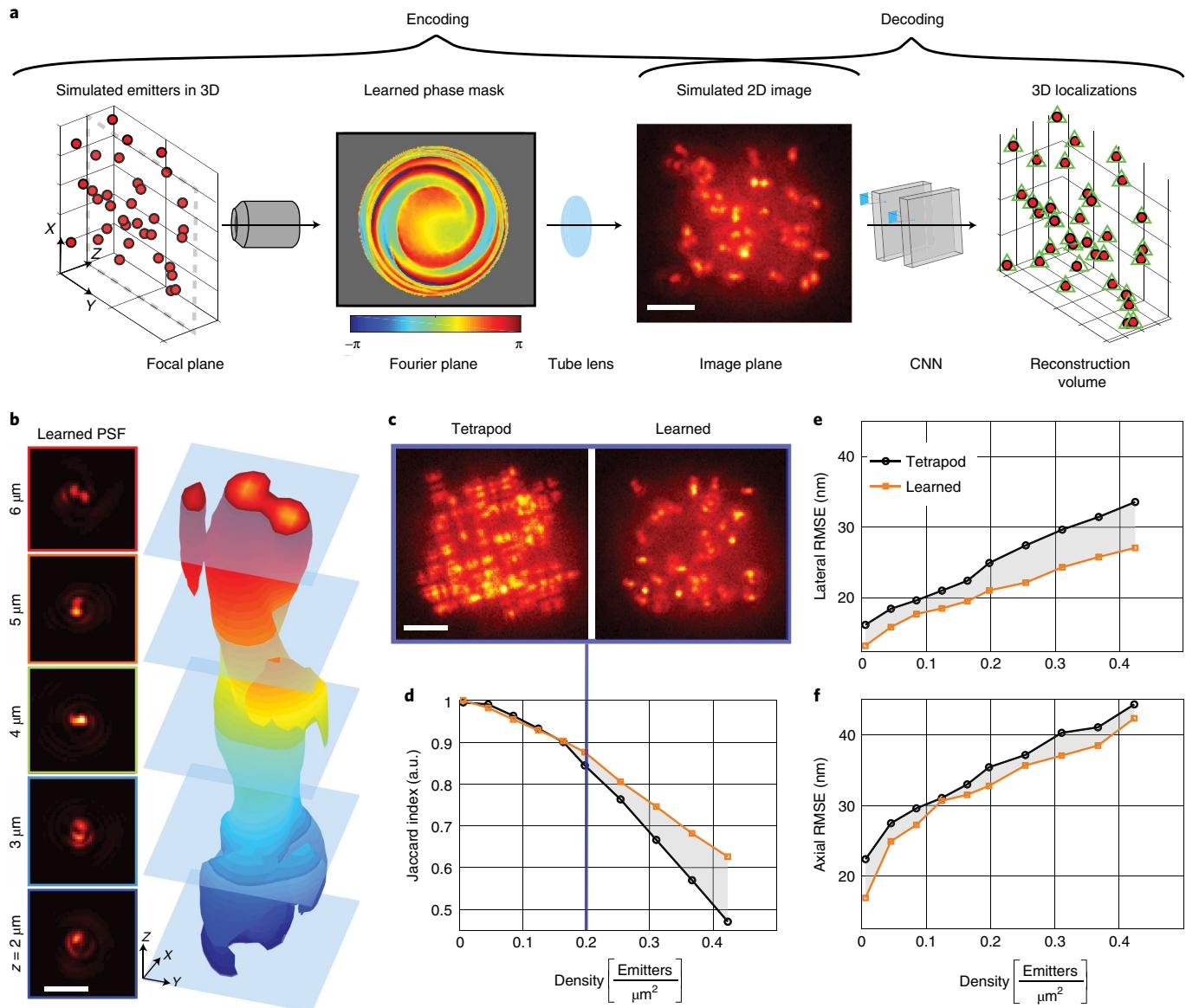
Our PSF design logic is based on the following: as we have already established that a CNN yields superior reconstruction for high-density 3D localization, we are interested in a PSF (encoder) that would be optimally localized by a CNN (decoder). Therefore, in contrast to a sequential paradigm where the PSF and the localization algorithm are optimized separately, we adopt a co-design approach (Fig. 4a). To jointly optimize the PSF and the localization CNN, we introduced 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 (see Supplementary Note 3). This image is then 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 (projected on a fine grid) and using the difference between the CNN recovery and the simulated 3D positions quantified by our loss function (see Supplementary Note 4.2), we optimize both the phase mask and the localization CNN parameters in an end-to-end fashion using the backpropagation algorithm<sup>50</sup> (see Supplementary Note 3.3 and Supplementary Video 4). The learned PSF (Fig. 4b) has a small



**Fig. 2 | Comparison to MP.** **a**, Performance comparison of a trained CNN (black) and the MP approach (red) in both detectability (Jaccard index) and in precision (lateral\axial RMSE). Matching of points was computed with a threshold distance of 150 nm using the Hungarian algorithm. Each data point is an average of  $n=100$  simulated images. Average s.d. in Jaccard index was  $\approx 6\%$  for both methods and average s.d. in precision was  $\approx 6$  nm for the CNN and  $\approx 15$  nm for MP. **b**, Example of a simulated frame of density  $0.124 \frac{\text{emitters}}{\mu\text{m}^2}$  alongside 3D comparisons of the recovered positions by MP (middle) and by the CNN (right). Scale bar, 2  $\mu\text{m}$ .



**Fig. 3 | Super-resolution 3D imaging over a 4  $\mu\text{m}$  z range.** **a**, Representative experimental frame (top) and rendered frame from the 3D recovered positions by the CNN overlaid on top (bottom). Scale bar, 5  $\mu\text{m}$ . **b**, Super-resolved image of mitochondria spanning a  $\approx 4 \mu\text{m}$  z range rendered as a 2D histogram, where z is encoded by color. Scale bar, 5  $\mu\text{m}$ . **c**, Zoom in on white rectangle (i) in **b**. Scale bar, 0.5  $\mu\text{m}$ . **d**, Relative intensity averaged along the shorter side of the white rectangle in **c**. **e**, XZ cross-section of white rectangle (ii) in **b**. Scale bar, 0.5  $\mu\text{m}$ . **f**, Relative intensity along the dashed white line in **e**. The experiment was repeated independently for  $n=3$  cells, twice analyzing 20,000 frames and once analyzing 10,000 frames all leading to similar performance.



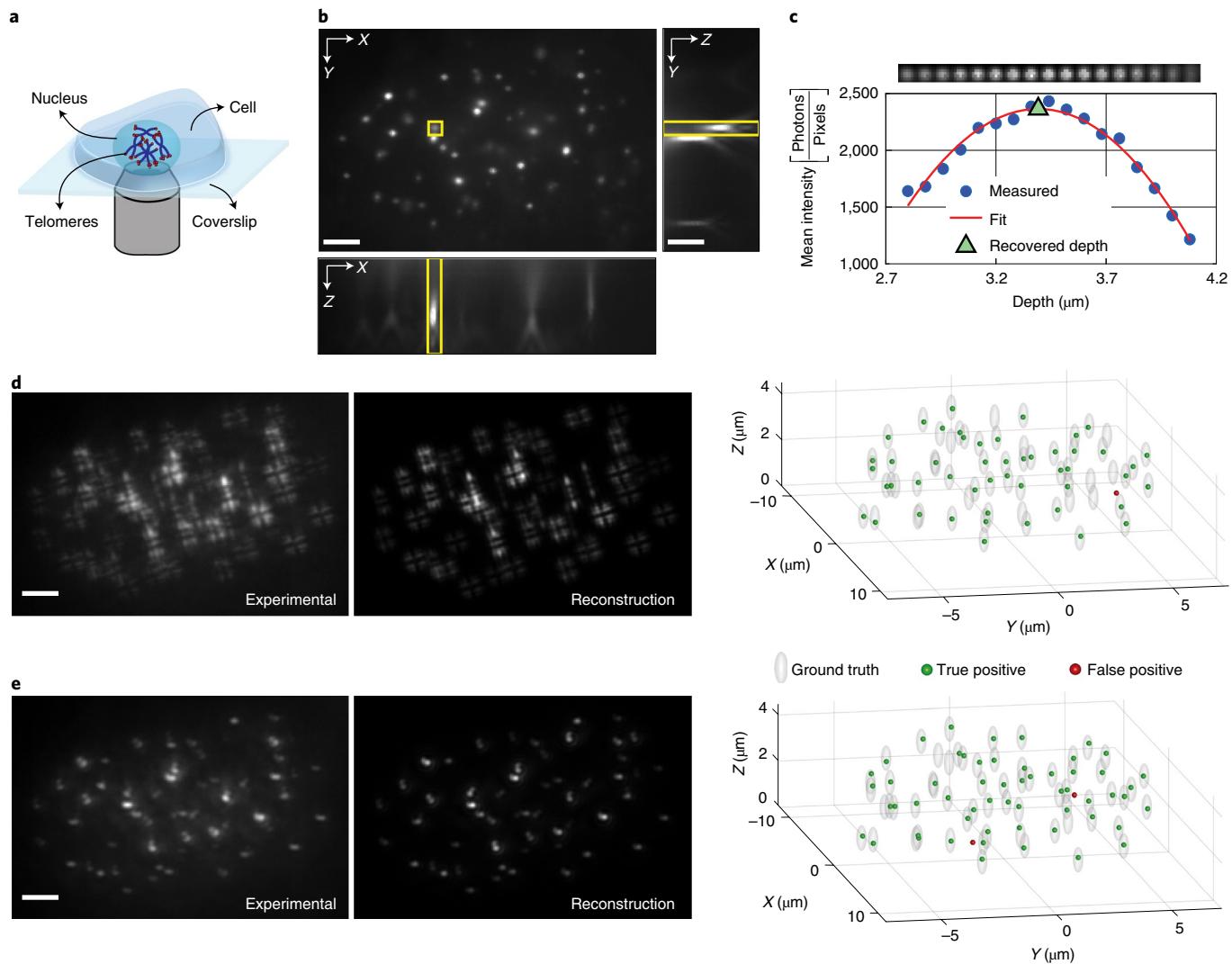
**Fig. 4 | PSF learning for high-density 3D imaging.** **a**, Simulated 3D emitter positions are fed to the image formation model to simulate their low-resolution camera image (encoding). Next, this image is fed to a CNN that tries to recover the simulated emitter positions (decoding). The difference 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**, Simulation of the learned PSF as function of the emitter axial position (left). 3D isosurface rendering of the learned PSF (right). **c**, Example frame of density  $0.197 \left[ \frac{\text{emitters}}{\mu\text{m}^2} \right]$  with the same simulated emitter positions, using the Tetrapod (left) and the learned PSF (right). **d**, Jaccard index comparison between two CNNs with the same architecture, one trained to recover 3D positions from 2D images of Tetrapod PSF (black) and the second trained to recover 3D positions from 2D images of the learned PSF (orange). Each data point is an average of  $n=100$  simulated images. Average s.d. was  $\approx 6\%$  for both PSFs. **e,f**, Lateral and axial RMSE comparison between the same two CNNs from **d**. Average s.d. was  $\approx 6$  nm for both PSFs. Scale bars, 3  $\mu\text{m}$ .

lateral footprint, which is critical for minimizing overlap at high densities. 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<sup>8</sup>.

To quantify the improvement introduced by our new PSF, we first compared it to the Tetrapod PSF in simulations. Specifically, we trained a similar reconstruction net for both the Tetrapod and the learned PSF using a matching training set composed of simulated continuous 3D positions along with their corresponding 2D low-resolution images. The learned PSF performs similarly to the Tetrapod PSF for low emitter densities (Fig. 4d–f). However, as the density goes up (higher than  $\approx 0.2 \left[ \frac{\text{emitters}}{\mu\text{m}^2} \right]$ ) the learned PSF outperforms the Tetrapod PSF in both localization precision (Fig. 4e,f).

and in emitter detectability (Jaccard index) (Fig. 4d). This result is not surprising, as the learned PSF has a smaller spatial footprint and hence it is less likely to overlap than the Tetrapod (Fig. 4c). For further analysis of the learned PSF see Supplementary Note 10.

**Volumetric telomere imaging and tracking.** Next, we demonstrate the superiority of the new PSF experimentally by imaging fluorescently labeled telomeres (DsRed-hTRF1) in fixed U2OS cells. The cell contains tens of telomeres squeezed in the volume of a nucleus with  $\approx 20 \mu\text{m}$  diameter (Fig. 5a,b). From a single snapshot focused inside the nucleus, the CNN outputs a list of 3D positions of telomeres spanning an axial range of  $\approx 3 \mu\text{m}$ . Using the Tetrapod PSF snapshot, the Tetrapod-trained CNN was able to recover 49 out of 62 telomeres with a single false positive, yielding a Jaccard index



**Fig. 5 | Three-dimensional imaging of telomeres in a single snapshot.** **a**, Schematic of imaging fixed U2OS cells with fluorescently labeled telomeres inside their nucleus. **b**, Focus slice with the standard PSF inside a U2OS cell nucleus, obtained via a z scan. The yellow rectangles mark the same emitter in all three orthogonal planes. **c**, Example fit of the mean intensity in sequential axial slices used to estimate approximate emitter axial position. **d**, Experimental snapshot with the Tetrapod PSF (left), rendered image from the 3D recovered positions by the Tetrapod CNN (middle) and a 3D comparison of the recovered positions and the approximate experimental ground truth (right). **e**, Experimental snapshot with the learned PSF (left), rendered image from the 3D recovered positions by the learned PSF CNN (middle) and a 3D comparison of the recovered positions and the approximate experimental ground truth (right). Scale bars, 3  $\mu\text{m}$ . The experiment was repeated independently for  $n=10$  U2OS cells all showing similar characteristics and performance.

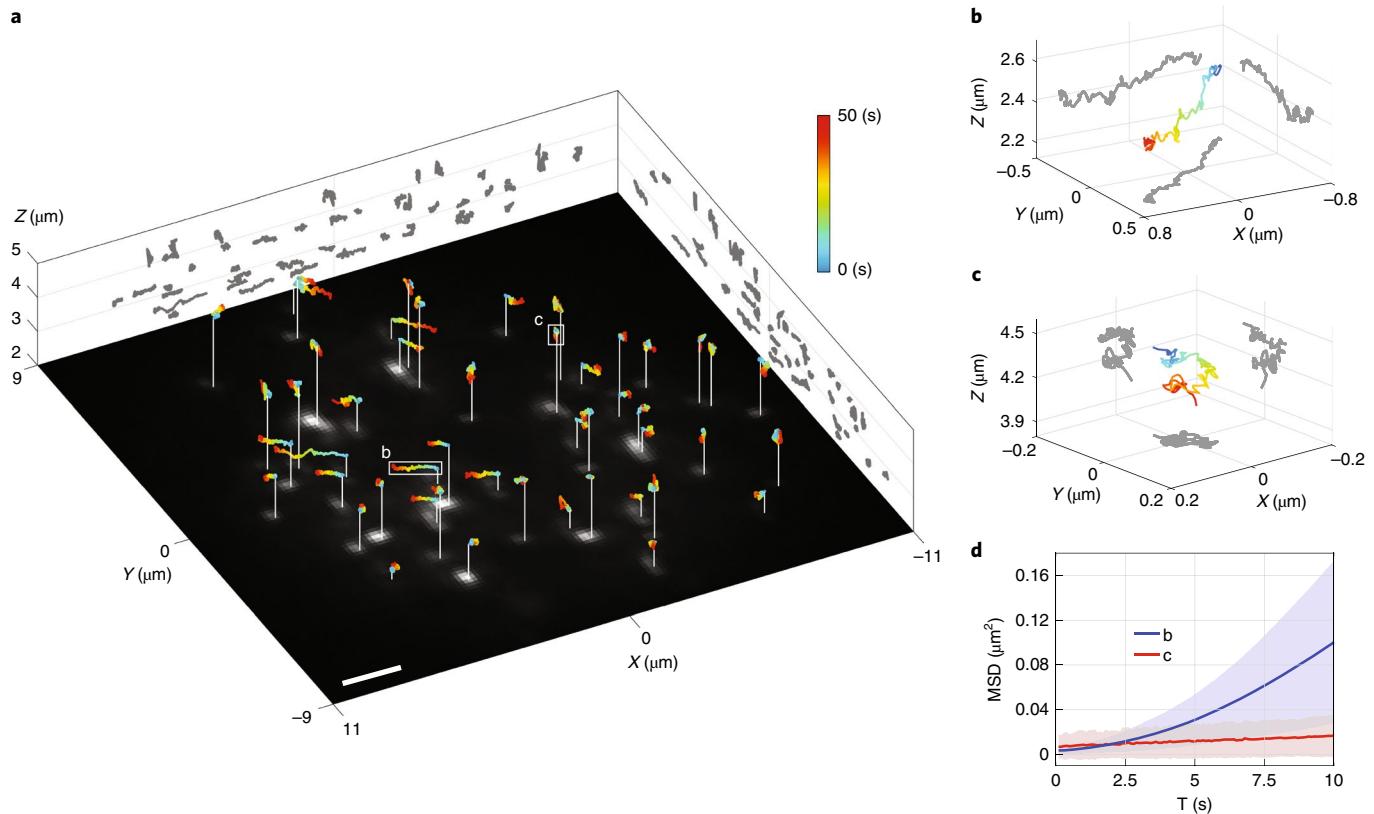
of 0.77 (Fig. 5d). In comparison, using the learned PSF snapshot, the corresponding CNN was able to recover 57 out of the 62 telomeres with only two false positives, yielding a Jaccard index of 0.89 (Fig. 5e). The recovered positions were compared to approximated ground-truth 3D positions (Fig. 5c), obtained by axial scanning and 3D fitting (see Supplementary Note 12 and Supplementary Video 5). The precision of both PSFs was calibrated experimentally using a z scan of a fluorescent microsphere (see Supplementary Note 10.4 and Supplementary Video 6).

To qualitatively compare the recovered list of localizations to the acquired snapshot, we fed this list to the physical simulation layer and generated the matching 2D low-resolution image (Fig. 5d,e). As verified by the regenerated images, the 3D positions of the telomeres are faithfully recovered by the CNNs. Moreover, the misses in both snapshots were either due to local aberrations and/or an extremely low number of signal photons (see Supplementary Note 13.1 for more experimental results).

Finally, a great advantage facilitated by our scan-free learned PSF is increased volumetric temporal resolution. To demonstrate the full capability of our method for dense multiparticle localization, we simultaneously tracked 48 telomeres, diffusing within the volume of a live mouse embryonic fibroblast (MEF) cell, at 10 Hz over 50 s (Fig. 6). Such a measurement can provide information on 3D nuclear rotation (Fig. 6a and Supplementary Video 7) and heterogeneity in motion type (Fig. 6b-d), at timescales that are typically unexplored by volumetric localization microscopy<sup>51</sup>.

## Discussion

In this work we demonstrated 3D localization of dense emitters over a large axial range, both numerically and experimentally. The described network architecture exhibits excellent flexibility in dealing with various experimental challenges, for example low signal-to-noise ratios and optical aberrations. This versatility is facilitated in three ways: (1) the net was trained solely on simulated



**Fig. 6 | Volumetric tracking of telomeres in live MEF cells.** **a**, 3D trajectories of tracked telomeres on top of the first experimental snapshot. White sticks mark the starting point and color encodes time. XZ and YZ projections are plotted in gray. Scale bar, 2  $\mu\text{m}$ . **b**, Example trajectory of a telomere (white box in **a**) with a large mean displacement. **c**, Example trajectory of a telomere (white box in **a**) with a small mean displacement. **d**, Mean squared displacement (MSD) of the trajectories in **b** and **c**. Shaded area marks one s.d. The experiment was repeated independently for  $n=10$  MEF cells all showing similar characteristics and performance.

data, thus producing sufficiently large datasets for optimization; (2) the phase mask that governs the PSF was optimized with respect to the implementation in the imaging system, that is the pixels of the spatial light modular, rather than over a smaller subspace, for example Zernike polynomials<sup>10</sup>; and (3) the CNN localization algorithm was designed in coordination with the development of the PSF, thus the system was optimized for the desired output<sup>37</sup> rather than a proxy.

Attaining a sufficiently large training dataset has thus far been a major limitation for most applications of CNNs. With this limitation in mind, the application of CNNs to single-molecule localization would seemingly be an ideal one, as each emitter's behavior should be approximately the same. This uniformity is broken, however, by spatially varying background, sample density and variable emitter size in biological samples (Supplementary Note 4.1), all of which diversify datasets and necessitate relevant training data. By implementing an accurate simulator (Supplementary Note 3), we have shown that it is possible to build a robust network entirely in silico, generating arbitrarily large, realistic datasets with a known ground truth to optimize nets. This aligns with our previous work in 2D SMLM<sup>20</sup>.

For super-resolution reconstructions using the Tetrapod PSF, the simulator was particularly important due to the highly variable SNR of emitters in the sample. Here, our net was able to selectively localize emitters even in very dense regions by focusing on those with a high SNR (Fig. 3). To optimize a PSF while simultaneously training the net, the simulator was also essential, as it would be prohibitively time consuming to experimentally vary the PSF, while recording and analyzing images to train the net.

A notable aspect of our optimization approach is that the optimized PSF is found by continuously varying the pixels of an initialized mask, while evaluating the output of the localization net, thus the final result represents a local minimum (Fig. 4). By changing the initialization conditions, we have recognized several patterns that indicate how the optimal PSF varies with the experimental conditions, namely, density, axial range and SNR (see Supplementary Note 2.2). Some of the recurrent features are intuitive: for example, in dense fields of emitters with limited SNR, the optimized PSFs have a small footprint over the designed axial range, enabling high density and compacting precious signal photons into as few pixels as possible. What distinguishes the net PSFs over predetermined designs is the utilization of multiple types of depth encoding, namely, simultaneously employing astigmatism, rotation and side lobe movement (Fig. 4), all of which have been conceived of and implemented previously, but never simultaneously.

This work, therefore, triggers many possible questions and research directions regarding its capabilities and limitations. For example, how globally optimal is the resulting PSF? Similarly, how sensitive is the resulting PSF and its performance to different loss functions, CNN architectures, initializations (for example with an existing phase mask) and the sampled training set of locations? Currently, it is unclear how each of these components affects the learning process, although we began to partially answer them in simulations (see Supplementary Notes 2.2 and 10). Finally, the co-design approach employed here paves the way to a wide variety of interesting applications in microscopy, where imaging systems have traditionally been designed separately from the processing algorithm.

**Online content**

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41592-020-0853-5>.

Received: 10 September 2019; Accepted: 6 May 2020;

Published online: 15 June 2020

**References**

- Katayama, Y. et al. Real-time nanomicroscopy via three-dimensional single-particle tracking. *Chem. Phys. Chem.* **10**, 2458–2464 (2009).
- Manzo, C. & Garcia-Parajo, M. F. A review of progress in single particle tracking: from methods to biophysical insights. *Rep. Prog. Phys.* **78**, 124601 (2015).
- Betzig, E. et al. Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **313**, 1642–1645 (2006).
- Hess, S. T., Girirajan, T. P. & Mason, M. D. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophysical J.* **91**, 4258–4272 (2006).
- Rust, M. J., Bates, M. & Zhuang, X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (storm). *Nat. Methods* **3**, 793–796 (2006).
- Sahl, S. J. & Moerner, W. Super-resolution fluorescence imaging with single molecules. *Curr. Opin. Struct. Biol.* **23**, 778–787 (2013).
- von Diezmann, A., Shechtman, Y. & Moerner, W. Three-dimensional localization of single molecules for super-resolution imaging and single-particle tracking. *Chem. Rev.* **117**, 7244–7275 (2017).
- Pavani, S. R. P. et al. Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a Double-Helix point spread function. *Proc. Natl Acad. Sci. USA* **106**, 2995–2999 (2009).
- Huang, B., Wang, W., Bates, M. & Zhuang, X. Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. *Science* **319**, 810–813 (2008).
- Shechtman, Y., Sahl, S. J., Backer, A. S. & Moerner, W. Optimal point spread function design for 3D imaging. *Phys. Rev. Lett.* **113**, 133902 (2014).
- Backer, A. S. & Moerner, W. Extending single-molecule microscopy using optical Fourier processing. *J. Phys. Chem. B* **118**, 8313–8329 (2014).
- Liu, S., Kromann, E. B., Krueger, W. D., Bewersdorf, J. & Lidke, K. A. Three-dimensional single-molecule localization using a phase retrieved pupil function. *Opt. express* **21**, 29462–29487 (2013).
- Babcock, H. P. & Zhuang, X. Analyzing single molecule localization microscopy data using cubic splines. *Sci. Rep.* **7**, 552 (2017).
- Li, Y. et al. Real-time 3D single-molecule localization using experimental point-spread functions. *Nat. Methods* **15**, 367 (2018).
- Aristov, A., Lelandais, B., Rensen, E. & Zimmer, C. Zola-3D allows flexible 3D localization microscopy over an adjustable axial range. *Nat. Commun.* **9**, 2409 (2018).
- Ferdman, B. et al. VIPR: vectorial implementation of phase retrieval for fast and accurate microscopic pixel-wise pupil estimation. *Opt. Express* **28**, 10179–10198 (2020).
- Shechtman, Y., Weiss, L. E., Backer, A. S., Sahl, S. J. & Moerner, W. Precise three-dimensional scan-free multiple-particle tracking over large axial ranges with tetrapod point spread functions. *Nano Lett.* **15**, 4194–4199 (2015).
- Min, J. et al. Falcon: fast and unbiased reconstruction of high-density super-resolution microscopy data. *Sci. Rep.* **4**, 4577 (2014).
- Boyd, N., Schiebinger, G. & Recht, B. The alternating descent conditional gradient method for sparse inverse problems. *SIAM J. Optim.* **27**, 616–639 (2017).
- Nehme, E., Weiss, L. E., Michaeli, T. & Shechtman, Y. Deep-storm: super-resolution single-molecule microscopy by deep learning. *Optica* **5**, 458–464 (2018).
- Sage, D. et al. Super-resolution fight club: assessment of 2D and 3D single-molecule localization microscopy software. *Nat. Methods* **16**, 387 (2019).
- Rivenson, Y., Zhang, Y., Günaydin, H., Teng, D. & Ozcan, A. Phase recovery and holographic image reconstruction using deep learning in neural networks. *Light-Sci. Appl.* **7**, 17141 (2018).
- Nguyen, T., Xue, Y., Li, Y., Tian, L. & Nehmetallah, G. Deep learning approach for Fourier ptychography microscopy. *Opt. Express* **26**, 26470–26484 (2018).
- Weigert, M. et al. Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nat. Methods* **15**, 1090 (2018).
- Ounkomol, C., Seshamani, S., Maleckar, M. M., Collman, F. & Johnson, G. R. Label-free prediction of three-dimensional fluorescence images from transmitted-light microscopy. *Nat. Methods* **15**, 917–920 (2018).
- Krull, A., Buchholz, T.-O. & Jug, F. Noise2void-learning denoising from single noisy images. in *Proc. IEEE Conference on Computer Vision and Pattern Recognition* (eds. Davis, L., Torr, P. & Zhu, S. C.) 2129–2137 (2019).
- Falk, T. et al. U-net: deep learning for cell counting, detection, and morphometry. *Nat. Methods* **16**, 67–70 (2019).
- Rivenson, Y. et al. Phasestain: the digital staining of label-free quantitative phase microscopy images using deep learning. *Light-Sci. Appl.* **8**, 23 (2019).
- Liu, T. et al. Deep learning-based super-resolution in coherent imaging systems. *Sci. Rep.* **9**, 3926 (2019).
- Smith, J. T. et al. Fast fit-free analysis of complex fluorescence lifetime imaging via deep learning. *Proc. Natl Acad. Sci. USA* **116**, 24019–24030 (2019).
- Boyd, N., Jonas, E., Babcock, H. P. & Recht, B. DeepLoco: fast 3D localization microscopy using neural networks. Preprint at *bioRxiv* <https://doi.org/10.1101/267096> (2018).
- Ouyang, W., Aristov, A., Lelek, M., Hao, X. & Zimmer, C. Deep learning massively accelerates super-resolution localization microscopy. *Nat. Biotechnol.* **36**, 460–468 (2018).
- Diederich, B., Then, P., Jügler, A., Förster, R. & Heintzmann, R. cellSTORM: cost-effective super-resolution on a cellphone using dSTORM. *PLoS ONE* **14**, e0209827 (2019).
- Newby, J. M., Schaefer, A. M., Lee, P. T., Forest, M. G. & Lai, S. K. Convolutional neural networks automate detection for tracking of submicron-scale particles in 2D and 3D. *Proc. Natl Acad. Sci. USA* **115**, 9026–9031 (2018).
- Zelger, P. et al. Three-dimensional localization microscopy using deep learning. *Opt. Express* **26**, 33166–33179 (2018).
- Liu, K. et al. Fast 3D cell tracking with wide-field fluorescence microscopy through deep learning. Preprint at <https://arXiv.org/abs/1805.05139> (2018).
- Hershko, E., Weiss, L. E., Michaeli, T. & Shechtman, Y. Multicolor localization microscopy and point-spread-function engineering by deep learning. *Opt. Express* **27**, 6158–6183 (2019).
- Speiser, A., Turaga, S. C. & Macke, J. H. Teaching deep neural networks to localize sources in super-resolution microscopy by combining simulation-based learning and unsupervised learning. Preprint at <https://arXiv.org/abs/1907.00770> (2019).
- Zhang, P. et al. Analyzing complex single-molecule emission patterns with deep learning. *Nat. methods* **15**, 913 (2018).
- Chakrabarti, A. Learning sensor multiplexing design through back-propagation. in *Advances in Neural Information Processing Systems* (eds. Lee, D. D. et al.) 3081–3089 (Curran Associates, 2016).
- Horstmeyer, R., Chen, R. Y., Kappes, B. & Judkewitz, B. Convolutional neural networks that teach microscopes how to image. Preprint at <https://arXiv.org/abs/1709.07223> (2017).
- Turpin, A., Vishniakou, I. & D Seelig, J. Light-scattering control in transmission and reflection with neural networks. *Opt. Express* **26**, 30911–30929 (2018).
- Haim, H., Elmaleh, S., Giryes, R., Bronstein, A. M. & Marom, E. Depth estimation from a single image using deep learned phase coded mask. *IEEE Trans. Comput. Imaging* **4**, 298–310 (2018).
- He, L., Wang, G. & Hu, Z. Learning depth from single images with deep neural network embedding focal length. *IEEE Trans. Image Process.* **27**, 4676–4689 (2018).
- Sitzmann, V. et al. End-to-end optimization of optics and image processing for achromatic extended depth of field and super-resolution imaging. *ACM Trans. Graph.* **37**, 114 (2018).
- Chang, J. & Wetzstein, G. Deep optics for monocular depth estimation and 3D object detection. in *Proc. IEEE International Conference on Computer Vision* (eds. Lee, K. M. et al.) 10193–10202 (2019).
- Wu, Y., Boominathan, V., Chen, H., Sankaranarayanan, A. & Veeraraghavan, A. Phasem3D: learning phase masks for passive single view depth estimation. in *IEEE International Conference on Computational Photography* (ed. Nedevschii, S.) 1–12 (2019).
- Shechtman, Y., Weiss, L. E., Backer, A. S., Lee, M. Y. & Moerner, W. Multicolour localization microscopy by point-spread-function engineering. *Nat. Photonics* **10**, 590 (2016).
- Bickel, P. J. & Doksum, K. A. *Mathematical Statistics: Basic Ideas and Selected Topics, Volumes I-II Package* (Chapman and Hall/CRC, 2015).
- LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
- Bronshtein, I. et al. Loss of lamin A function increases chromatin dynamics in the nuclear interior. *Nat. Commun.* **6**, 8044 (2015).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2020

## Methods

**Sample preparation.** COS7 cells were grown for 24 h on cleaned 22 × 22 mm, 170-μm thick coverslips in a six-well plate in DMEM with 1 g l<sup>-1</sup> D-glucose (low glucose), supplemented with fetal bovine serum, penicillin-streptomycin and glutamine at 37 °C and 5% CO<sub>2</sub>. Cells were fixed with 4% paraformaldehyde and 0.2% glutaraldehyde in PBS (pH 6.2) for 45 min, washed and incubated in 0.3 M glycine/PBS solution for 10 min. The coverslips were transferred into a clean six-well plate and incubated in a blocking solution for 2 h (10% goat serum, 3% BSA, 2.2% glycine and 0.1% Triton-X in PBS, filtered with 0.45-μm PVDF filter unit, Millex). The cells were then immunostained overnight with anti TOMM20-AF647 (Abcam ab209606) 1:230 diluted in blocking buffer and washed five times with PBS. Cover glasses (22 × 22 mm, 170 μm thick) were cleaned in an ultrasonic bath with 5% Decon90 at 60 °C for 30 min, then washed with water, incubated in ethanol absolute for 30 min and sterilized with 70% filtered ethanol for 30 min.

U2OS cells were grown on cleaned 0.18-mm coverslips in a 12-well plate in DMEM with 1 g l<sup>-1</sup> D-glucose (low glucose), supplemented with fetal bovine serum, penicillin-streptomycin and glutamine at 37 °C and 5% CO<sub>2</sub>. The day after cells were transfected with a plasmid encoding the fluorescently tagged telomeric repeat binding factor 1 (DsRed-hTRF1)<sup>51</sup> using Lipofectamine 3000 reagent. At 24 h after transfection, cells were fixed with 4% paraformaldehyde for 20 min, washed three times with PBS and attached to a slide together with mounting medium.

Lamin A double knockout (lmna<sup>-/-</sup>) MEFs were cultured in phenol red-free DMEM/F-12 medium (Gibco, Thermo Fisher Scientific), which was supplemented with 10% fetal bovine serum and 1% penicillin-streptomycin solution (Biological Industries, Bet Ha-emek). Two days before imaging, cells were transferred to 15-mm coverglass-bottom culture plates (Nest scientific). After 24 h, cells were transfected with a plasmid encoding DsRed-hTRF1. The transfection mix was prepared by diluting 8 μg of the plasmid in 50 μl of serum-free DMEM/F-12 and separately diluting 24 μl of transfection reagent Polyjet (SigmaGen Laboratories) in 26 μl of serum-free DMEM/F-12, then immediately adding the diluted Polyjet solution to the DNA mixture and incubating at room temperature for 20 min. The 100 μl DNA-Polyjet mix was then added dropwise to cell plates. Imaging experiments were conducted approximately 26 h after transfection and lasted for 1–2 h.

**Optical setup.** Imaging experiments were performed on the experimental system shown schematically in Fig. 1a. The 4f optical processing system was built alongside the side port of a Nikon Eclipse Ti inverted fluorescence microscope, with a ×100/1.45 NA oil-immersion objective lens (Plan Apo ×100/1.45 NA, Nikon).

**STORM imaging.** For super-resolution imaging, a polydimethylsiloxane chamber was attached to a glass coverslip containing fixed COS7 cells. Blinking buffer (100 mM β-mercaptoethyamine hydrochloride, 20% sodium lactate and 3% OxyFluor (Sigma, SAE0059), modified from Nahidazar et al.<sup>52</sup>, was then added and a glass coverslip was placed on top to prevent evaporation. Low-intensity illumination for recording diffraction-limited images was applied using a Toptica laser (640 nm), on the Nikon TI imaging setup described previously and recorded with an EMCCD (iXon, Andor) in a standard imaging setup. For super-resolution blinking using the Tetrapod PSF, high-intensity (1 W at the back of the objective lens) 640 nm light was applied using a 638 nm 2,000 mW red dot laser module, whose beam shape was cleaned using a 25-μm pinhole (Thorlabs) in coordination with low-intensity (<5 mW) 405 nm light. Emission light was filtered through a 500-nm long pass dichroic and a 650-nm long pass (Chroma), projected through a 4f system containing the dielectric Tetrapod phase mask (see Supplementary Note 9.1) and imaged on a Prime95b Photometrics camera.

**Super-resolution image rendering.** Before rendering the super-resolved image (Fig. 3b), we first corrected for sample drift using the ThunderSTORM<sup>3</sup> ImageJ Fiji plugin<sup>54</sup>. Afterward, we rendered 3D localizations as a 2D average shifted histogram, with color encoding the z position.

**Telomere imaging.** For telomere imaging in fixed cells, the 4f system consisted of two f = 15 cm lenses (Thorlabs), a linear polarizer (Thorlabs) to filter out the light that is polarized in the unmodulated direction of the LC-SLM, a 1,920 × 1,080 pixel LC-SLM (PLUTO-VIS, Holoeye) and a mirror for beam-steering. A sCMOS camera (Prime95B, Photometrics) was used to record the data. The sample was illuminated with 561-nm fiber-coupled laser light source (iChrome MLE, Toptica). The excitation light was reflected up through the microscope objective by a multibandpass dichroic filter (TRF89902-EM-ET-405/488/561/647 nm Laser Quad Band Set, Chroma). Emission light was filtered by the same dichroic and also filtered by another 617-nm band pass filter (FF02-617/73, Semrock).

For volumetric telomere tracking in live cells, images were recorded with an EMCCD camera (Andor iXON), exposure time of 100 ms and EM-gain of 170. The sample was illuminated at ≈2 kW cm<sup>-2</sup> with 561 nm light from a fiber-coupled laser (iChrome MLE, Toptica). All movies were recorded for 50 s (500 frames).

**CNN architecture.** In summary, our localization CNN architecture is composed of three main modules. First, a multiscale context aggregation module processes the input 2D low-resolution image and extracts features with a growing receptive field using dilated convolutions<sup>55</sup>. Second, an upsampling module increases the lateral resolution of the predicted volume by fourfold. Finally, the last module refines the depth and lateral position of the emitters and outputs the predicted vacancy grid. For more details regarding the architecture see Supplementary Note 2.

**Statistics and reproducibility.** The STORM experiment was repeated independently for n = 3 cells, twice analyzing 20,000 frames and once analyzing 10,000 frames, all leading to similar performance. The fixed telomere experiment was repeated independently for n = 10 U2OS cells all showing similar characteristics and performance. The live telomere experiment was repeated independently for n = 10 MEF cells all showing similar characteristics and performance.

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## Code availability

Code is made publicly available at <https://github.com/EliasNehme/DeepSTORM3D>.

## References

52. Nahidazar, L., Agronskaia, A. V., Broertjes, J., van den Broek, B. & Jalink, K. Optimizing imaging conditions for demanding multi-color super resolution localization microscopy. *PLoS ONE* **11**, e0158884 (2016).
53. Ovesný, M., Krížek, P., Borkovec, J., Švindrych, Z. & Hagen, G. M. ThunderSTORM: a comprehensive ImageJ plug-in for PALM and STORM data analysis and super-resolution imaging. *Bioinformatics* **30**, 2389–2390 (2014).
54. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676 (2012).
55. Yu, F. & Koltun, V. Multi-scale context aggregation by dilated convolutions. Preprint at <https://arXiv.org/abs/1511.07122v3> (2016).

## Acknowledgements

We thank the Garini laboratory (Bar-Ilan University) for the U2OS cells, lmna<sup>-/-</sup> MEFs and the plasmid encoding for DsRed-hTRF1. We thank J. Ries for help with the application of SMAP-2018 to Tetrapod PSFs. We gratefully acknowledge the support of the NVIDIA Corporation with the donation of the Titan V GPU used for this research. We thank the staff of the Micro-Nano-Fabrication and Printing Unit at the Technion for their assistance with the phase mask fabrication. We thank Google for the research cloud units provided to accelerate this research. E.N., O.A., B.F. and R.O. are supported by H2020 European Research Council Horizon 2020 (802567); T.M. is supported by the Israel Science Foundation (grant no. 852/17) and by the Ollendorff Foundation; R.G. and O.A. are supported by the Israel Science Foundation (grant no. 450/18); Y.S. is supported by the Technion-Israel Institute of Technology Career Advancement Chairship; L.E.W. and Y.S. are supported by the Zuckerman Foundation. D.F. is supported by Google.

## Author contributions

E.N., D.F., T.M. and Y.S. conceived the approach. E.N. performed the simulations and analyzed the data with contributions from all authors. E.N., R.G., B.F., L.E.W., O.A. and T.N. collected the data. R.O. fabricated the physical phase mask. T.N. prepared MEF cells. E.N., D.F., L.E.W., T.M. and Y.S. wrote the paper with contributions from all authors.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41592-020-0853-5>.

**Correspondence and requests for materials** should be addressed to Y.S.

**Peer review information** Rita Strack was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

Corresponding author(s): Yoav Shechtman

Last updated by author(s): Jun 23, 2020

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give P values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection

Nikon Imaging Software v 5.02.02

Data analysis

Matlab 2017b, Python 3.6. A list of exact python package versions are present on GitHub within the environment.yml file.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

### Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](#)

# Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size was not predetermined based off statistical calculations. The mitochondria experiment was repeated independently for 3 cells, twice analyzing 20K frames and once analyzing 10K frames to ensure reproducibility. For telomere samples we imaged two different types of cells (U2OS and MEF) on two independent setups (different cameras and SLMs) in order to ensure we cover the range of possible biological as well as experimental variations. The variation between different experiments was low.
Data exclusions	No data was excluded.
Replication	All replication attempts were successful.
Randomization	No randomization was used for experimental data; however, simulations systematically test a number of different imaging conditions. No sample randomization was performed in this study.
Blinding	Analyzed data for this study is illustrative of potential applications and was not used to extract specific biological conclusions, thus blinding was not relevant to this study.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems		Methods	
n/a	Involved in the study	n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies	<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines	<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology	<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data		

### Antibodies

Antibodies used	anti TOMM20-AF647 (dilution 1:230, rabbit monoclonal, ab209606, Lot CR3270643-1, Abcam)
Validation	According to the manufacturer anti TOMM20-AF647 antibody (rabbit monoclonal, ab209606, Abcam) is reactive against Homo sapiens (Human) and was validated in HeLa cells for immunocytochemistry and immunofluorescence.

### Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	COS7 cells were a gift of the Elia lab (Ben-Gurion University), U2OS cells were a kind gift of the Garini lab (Bar-Ilan University) used in ref [51], MEF lamin A double knockout cells were also a kind gift of the Garini lab (Bar-Ilan University) used in ref [51].
Authentication	none of the cell lines used were authenticated.
Mycoplasma contamination	Only MEF lamin A double knockout cells were tested and found negative for mycoplasma contamination.
Commonly misidentified lines (See <a href="#">ICLAC</a> register)	No commonly misidentified cell lines were used.