Analyzing complex single-molecule emission patterns with deep learning

Peiyi Zhang1,5, Sheng Liu1,5, Abhishek Chaurasia2, Donghan Ma1, Michael J. Mlodzianoski1, Eugenio Culurciello1,2 and Fang Huang1,3,4*

A fluorescent emitter simultaneously transmits its identity, location, and cellular context through its emission pattern. We developed smNet, a deep neural network for multiplexed single-molecule analysis to retrieve such information with high accuracy. We demonstrate that smNet can extract three-dimensional molecule location, orientation, and wavefront distortion with precision approaching the theoretical limit, and therefore will allow multiplexed measurements through the emission pattern of a single molecule.

Analysis of single-molecule emission patterns is critical for the retrieval of structural and physiological information about molecules’ tagged targets and, further, a clear understanding of their interactions and cellular context. These emission patterns of tiny light sources (i.e., point spread functions (PSFs)) encode information such as a molecule’s location, orientation, and environment within the specimen, as well as the path the emitted photons took before reaching the camera. Detection and tracking of single fluorescent probes on the basis of their emission patterns lay the foundation for modern single-molecule-based imaging methods. These methods allow the investigation of cellular dynamics, transcriptional regulation, and protein interaction, and permit single-molecule-based super-resolution imaging in both fixed and living specimens.

One key in single-molecule studies is understanding how the features of the PSF encode the properties of a single molecule, that is, the analysis of single-molecule emission patterns. Conventionally, such analysis focuses on dominant and recognizable features of PSFs, such as rotation of a double-helix PSF, and on modeling of feature changes through a simplified mathematical form, such as a Gaussian PSF model. Most recently, interpolation-based numerical methods, such as splines, have been demonstrated for 3D localization of single molecules. However, the number of parameters required in the interpolation scales exponentially with PSF complexity and measurement dimensions, which makes it challenging to use such methods to retrieve information carried by high-dimensional PSFs (Supplementary Note 1). Furthermore, single-molecule emission patterns carry multiple classes of molecular information simultaneously. Because of the difficulty of perceiving and summarizing a comprehensive PSF model, the retrieval of multiplexed information beyond 3D position from complex or high-dimensional single-molecule data remains challenging.

Deep neural networks (DNNs) extract features from the input and learn its connection to the output automatically. A basic neural network (one or two fully connected layers) has been used to identify fluorophore species in single-molecule fluorescence lifetime experiments, as well as to speed up the estimation of dipole orientation from an analytical approximation of the dipole PSF. Through the deep architecture of DNNs, the complex mapping between input and output is extracted from different levels of features hierarchically. The inference precision, instead of depending on domain expertise (e.g., feature recognition), depends mainly on the design of the network architecture.

We applied deep learning to extract multiplexed information carried by single-molecule patterns, skipping conventional steps such as feature recognition, model simplification, and iterative regression methods. For high-dimensional single-molecule data, we designed the network to tackle each inference task independently, thereby allowing complexities in each dimension to be added instead of multiplied. We show that a well-designed DNN architecture can be trained to efficiently extract both molecular and specimen information, such as molecule location, dipole orientation, and wavefront distortion, from complex and subtle features of PSFs, which otherwise are considered too complex for established algorithms.

The general principle of our DNN for single-molecule studies (referred to as smNet) is illustrated in Fig. 1. smNet is a deep network of 27–36 layers (Supplementary Figs. 1–3 and Supplementary Table 1) consisting of convolutional layers, residual blocks, and fully connected layers together with batch normalization and parametric rectified linear unit. The complex and subtle features of the PSF lie in the photon distribution within a small subregion (2–10µm²). To fully utilize the information contained in the spatial domain, we used a large kernel size in the beginning layers, and stacked a number of convolutional layers and ‘bottleneck’ residual blocks to capture as many levels of features as possible. This architecture helped smNet learn different levels of features and generate an optimal mapping from the input to the output (Supplementary Notes 2 and 3). smNet can be trained on either simulated or measured PSFs. We designed the training cost function to measure the relative difference between the measurement error of a specific single-molecule property and the information limit calculated by the Cramér–Rao lower bound (CRLB) for each training image (Supplementary Notes 4–6). With this design, the training process of smNet tunes the parameters to achieve the specific CRLB set by the inherent information content of each image. This allows smNet to extract information close to the theoretical limit for a large range of detected photons and background levels simultaneously. We found that it was sufficient to train smNet with ~1 million PSF patterns for each inference task with a reasonably large range of measured properties (Supplementary Fig. 4 and Supplementary Table 2).

*Weldon School of Biomedical Engineering, Purdue University, West Lafayette, IN, USA. 1School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, USA. 2Purdue Institute for Integrative Neuroscience, Purdue University, West Lafayette, IN, USA. 3Purdue Institute of Inflammation, Immunology and Infectious Disease, Purdue University, West Lafayette, IN, USA. 4These authors contributed equally: Peiyi Zhang and Sheng Liu. *e-mail: fanghuang@purdue.edu
To test the performance of smNet, we first evaluated the precision and accuracy when pinpointing single-molecule centers in both simulated and experimentally obtained single-molecule patterns. We found that smNet localized single molecules at a precision matching the theoretical information limit given by the CRLB with a small or ignorable bias, despite the considerable amount of aberrations and PSF complexity (2.35±2.71 nm (bias±s.d.) for astigmatism, 1±0.84 nm (bias±s.d.) for double-helix PSFs, 2±1.8 nm (bias±s.d.) for simulated complex PSFs, and 102.1±7 nm (mean±s.d.) for experimental complex PSFs; Supplementary Figs. 5–7). This performance was achieved consistently for various conditions such as molecule locations, intensity, and background levels (Fig. 2e,f, Supplementary Figs. 8–10, Supplementary Table 3, and Supplementary Notes 7–9 including results of smNet testing on various conditions). We further demonstrated smNet in 3D single-molecule switching nanoscopy (SMSN) experiments. SMSN relies on the localization of millions of PSFs down to a precision of 10–25 nm, which together with localization accuracy is essential for the successful reconstruction of SMSN images. Using smNet, we reconstructed 3D SMSN volumetric images of the fluorescently labeled mitochondrial protein TOM20 in COS-7 cells (Methods), imaged either at the bottom coverslip or through an ~12-µm-thick sample cavity. In fact, smNet learned to build a deep network from PSF images generated from an experimentally retrieved pupil, containing measured optical aberrations, modeled by 64 Zernike polynomials (Supplementary Note 10). This allowed smNet to retrieve the correct molecular positions despite the substantial amount of aberrations (Fig. 2d, Supplementary Figs. 1b and 5). We found that the x–z cross-sections of the mitochondrial outer membrane showed substantial artifacts with conventional Gaussian-based methods, whereas smNet accurately reconstructed the surface contours of the subcellular organelles despite the imaging depth (Fig. 2a–c, Supplementary Fig. 11, and Supplementary Notes 11–16).

Single-molecule emission patterns can be designed to evolve and encode molecular properties such as 3D positions, probe spectra, identities, and orientations. However, when two or more classes of information are encoded in the emission patterns, their dimensionality increases, which can make traditional decoding processes such as feature recognition and regression challenging.

smNet learns to recognize PSF features to extract the desired measurement through the information-limit weighted cost function (Supplementary Note 3). During this process, smNet optimizes its parameters specifically for a certain measurement task, ignoring other irrelevant features. We found that smNet allowed independent and therefore parallel inference of the spatial location together with the polar and azimuthal angles from a single-molecule dipole-emission pattern with little change to its architecture (Supplementary Table 4 and Supplementary Note 17). The inference precisions for all five dimensions closely approached the information limits in a large parameter range, and degeneracies and wrappings of both polar and azimuthal angles could also be correctly predicted (Supplementary Fig. 12).

The possibility of extracting multiplexed information from emission patterns with smNet inspired us to use it to measure wavefront distortions within a specimen. Although the emission patterns in a single-molecule dataset originate from different locations within the detected region of the specimen, they share a similar wavefront distortion induced by the inhomogeneous refractive indices of cell and tissue structures. smNet, designed to obtain its measurements (e.g., the amount of horizontal coma) from the common features of PSF patterns, directly extracted the shared wavefront distortion from a small collection of detected emission patterns without any additional information (Supplementary Note 11). We found that smNet was capable of simultaneously measuring amplitudes of 12–21 Zernike polynomials (Wyant order), representing wavefront shapes, while achieving a residual wavefront error of <30 nm (Fig. 3a,b, Supplementary Fig. 13, Supplementary Video 1, and Supplementary Note 8). We found that the resulting wavefront shape measured by smNet was in close agreement with that obtained via the phase retrieval method using beads on a coverslip surface (Supplementary Fig. 14). Notably, smNet was able to...
measure wavefront distortion without using a guide star\textsuperscript{20} or scanning a bead\textsuperscript{18} sample, which restricts the wavefront measurement from the actual imaging volume.

Using smNet, we were able to extract sample induced aberration through the raw single-molecule blinking data itself, which allowed wavefront measurement deep into the specimen. As a demonstration, we measured the evolution of 12 Zernike aberration modes through 11 consecutive optical sections through an immunolabeled specimen (TOM20 in COS-7 cells) on a custom-built biplane setup. We found that first-order spherical aberration evolved continuously while aberrations such as diagonal astigmatism decreased with increasing depth (Fig. 3c,d and Supplementary Video 2).

We found that these wavefront measurements made with smNet stabilized after averaging of 100–300 subregions, or 20–60 raw data frames (depending on emitter density). This fast response time is useful for tracking dynamic wavefront distortions during continuous data acquisition. To demonstrate this, we used smNet to capture sudden aberration changes by introducing multiple cycles of controlled wavefront distortion with a deformable mirror during continuous acquisition. We found that our input-voltage amplitude for the deformable mirror, which resembled normalized Zernike polynomials, could be rapidly and consistently captured by smNet (Fig. 3e, Supplementary Fig. 15, Supplementary Videos 3–5, and Supplementary Data). We expect that further development could allow smNet to provide continuous feedback to a wavefront-control element during SMSN imaging of a living specimen.

In summary, we developed smNet, a DNN for complex and high-dimensional analysis of single-molecule emission patterns. As demonstrated through both computer-generated and experimentally obtained datasets, both general and subtle features of single-molecule emission patterns can be learned close to the information limit of the data for tasks including determination of 3D position, orientation, and measurement of wavefront distortion. The designed architecture and network depth ensure smNet’s performance in
precision, accuracy, and speed. Furthermore, smNet decouples high-dimensional single-molecule measurement from limitations in human-based feature recognition, model simplification, and regression, and therefore could further allow encoding and extraction of highly multiplexed physical and physiological information from the emission pattern of a single molecule.

Online content
Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41592-018-0153-5.

Received: 26 January 2018; Accepted: 9 August 2018; Published online: 30 October 2018

References
1. Moerner, W. E. & Fromm, D. P. Rev. Sci. Instrum. 74, 3597–3619 (2003).
2. von Dziedzic, A., Shechtman, Y. & Moerner, W. E. Chem. Rev. 117, 7244–7275 (2017).
3. Backlund, M. P., Lew, M. D., Backer, A. S., Sahl, S. J. & Moerner, W. E. ChemPhysChem 15, 587–599 (2014).
4. Moon, S. et al. J. Am. Chem. Soc. 139, 10944–10947 (2017).
5. Burke, D., Patton, B., Huang, F., Bewersdorf, J. & Booth, M. J. Optica 2, 177–185 (2015).
6. Andrews, N. L. et al. Nat. Cell Biol. 10, 955–963 (2008).
7. Femino, A. M., Fay, F. S., Fogarty, K. & Singer, R. H. Science 280, 585–590 (1998).
8. Ha, T. et al. Proc. Natl. Acad. Sci. USA 93, 6264–6268 (1996).
9. Baddeley, D. & Bewersdorf, J. Annu. Rev. Biochem. 87, 965–989 (2018).
10. Sage, D. et al. Nat. Methods 12, 717–724 (2015).
11. Pavani, S. R. et al. Proc. Natl. Acad. Sci. USA 106, 2995–2999 (2009).
12. Babcock, H. P. & Zhuang, X. Sci. Rep. 7, 552 (2017).
13. Li, Y. et al. Nat. Methods 15, 367–369 (2018).
14. LeCun, Y., Bengio, Y. & Hinton, G. Nature 521, 436–444 (2015).
15. Bowen, B. P., Scruggs, A., Enderlein, J., Sauer, M. & Woodbury, N. J. Phys. Chem. A 108, 4799–4804 (2004).
16. Zhang, Y. et al. Protein Cell 15, 598–606 (2013).
17. He, K., Zhang, X., Ren, S. & Sun, J. in Proc. 29th IEEE Conference on Computer Vision and Pattern Recognition: CVPR 2016 (eds Agapito, L. et al.) 770–778 (IEEE, Piscataway, NJ, 2016).
18. Liu, S., Kromann, E. B., Krueger, W. D., Bewersdorf, J. & Lidke, K. A. Opt. Express 21, 29462–29487 (2013).
19. Cutler, F. J. et al. PLoS One 8, e64320 (2013).
20. Ji, N. Nat. Methods 14, 378–380 (2017).

Acknowledgements
We thank A.J. Schaber for help with initial data acquisition and the Bindley Bioscience Center at Purdue, a core facility of the NIH-funded Indiana Clinical and Translational Sciences Institute. We thank K.A. Lidke and M.J. Wester from the University of New Mexico for an initial contribution to the PSF toolbox software. We thank E.B. Kromann (Technical University of Denmark) for sharing the phase unwrapping code. We thank C. Haig (Hamamatsu Photonics K.K.) and K.F. Ziegler for their help on the project, and P.-M. Ivey, Y. Li, and F. Xu for suggestions on the manuscript. P.Z., S.L., M.J.M., D.M., and E.H. were supported by the NIH (grant 1R01GM119785) and DARPA (grant D14AP00093).

Author contributions
P.Z., S.L., and A.C. developed the algorithm. P.Z. and S.L. wrote the software, performed the experiments, and analyzed the data. P.Z., S.L., and D.M. prepared the specimens. M.J.M. and D.M. constructed the imaging systems. P.Z. and E.H. conceived the study. E.C. and E.H. supervised the study. All authors wrote the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41592-018-0153-5.
Reprints and permissions information is available at www.nature.com/reprints.
Correspondence and requests for materials should be addressed to F.H.

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. 2018
Methods

Optical setup. All experimental data (except for complex PSFs and wavefront estimation data) were recorded on a custom-built SMSN setup built around an Olympus IX-73 microscope (Olympus America Inc., Waltham, MA) with a 100× CLS-NA 1.35 (with 0.03 NA aperture) silicone oil-immersion objective lens (IFU-27814; Olympus America Inc.) and 405-nm (DL-405-100; CrystaLaser, Reno, NV) and 642-nm (2RU-VFL-P-2000-642-B1R; MPB Communications Inc.) lasers for activation and excitation, respectively. The filter turret contained a dichroic mirror (DM03-R045/488/S615/635-11; Semrock Inc.). A deformable mirror (MultDM-3.5; Boston Micromachines, Cambridge, MA) placed at the conjugated pupil plane was used to correct systematic aberrations and introduce astigmatism for 3D SMSN. Collected fluorescence emission was passed through a band-pass filter (FF01-731/137-25; Semrock Inc.) placed just before the camera. The fluorescence signal was recorded on an EMCCD (electron-multiplying charge-coupled device) camera (C9100-23B; Hamamatsu, Tokyo, Japan). The overall system magnification was ~53×, resulting in an effective pixel size of 113 nm.

For wavefront distortion measurements, the fluorescence emission after the imaging lens was split into two beam paths by a 50/50 beam splitter (BS016; Thorlabs). A small difference in the optical path length was introduced between the two paths to create a dual-focal-plane configuration, resulting in a plane separation of 430 nm at the sample plane. The two beams were then combined by a right angle prism (47065; Edmund Optics) and received by an sCMOS (scientific complimentary metal-oxide semiconductor) camera (Orca-Flash4.0v3; Hamamatsu). The overall system magnification was ~53×, resulting in an effective pixel size of 122 nm. A 100×/1.4 NA oil-immersion objective (ULP1SAPO100X; Olympus America Inc., Waltham, MA) was used for wavefront distortion measurement. A pair of multi-plane sensors were inserted in wavefront distortion measurement to avoid degeneracies between aberration modes.

smNet architecture. smNet is composed of 3–5 convolutional layers, 7–11 residual blocks, and 0–2 fully connected layers (Supplementary Note 2). Each convolutional layer is followed by batch normalization and a parametric rectified linear unit (PReLU) (Supplementary Note 2), except for the last convolutional layer in M3 (Supplementary Table 1). The first fully connected layer is followed by a PReLU, and the last fully connected layer is followed by a HardTanh (https://github.com/torch/nn/blob/master/doc/transfer.md). Detailed information about smNet architecture and its variations is shown in Supplementary Table 1. Because this input image has a small size and the features of the PSF span a small number of pixels, it is essential to fully utilize the information contained in the spatial domain. To achieve this, we used larger kernels in early layers than in the later layers of our neural network. We used 64 kernels with a size of 7 × 7 pixels in the first layer, and then 128 kernels with a size of 5 × 5 pixels. After this, we focused on capturing as many rich features as possible. Stacking a large number of convolutional layers helped us achieve this; however, this approach often makes neural networks untrainable. To avoid this, we used a stack of 7–11 residual blocks in our architecture. Each residual block used the “bottleneck structure”, where the number of features is first squeezed and then expanded. This design helps not only to reduce the number of training parameters but also to encourage learning of more relevant features.

We assumed that in later layers there was much less spatial information left to be learned by smNet, so we used 1 × 1 convolutional layers. Finally, we followed these by fully connected layers. We found that reducing the number of both fully connected layers and 1 × 1 convolutional layers helped to improve the accuracy of the wavefront distortion estimation.

In our study, the output of smNet was a vector of 12 or 21 elements representing the amplitudes of 12 or 21 Zernike modes; a vector of 2 elements representing x and y coordinates; or a scalar representing the z position, polar angle (α), or azimuthal angle (β). Because x, y positions are based on the emitter’s location, we subjugated the z position, polar and azimuthal angles, and wavefront distortions are based on the shape information or a combination between shape and position information of the emitter, we decided to construct separate networks (with the same architecture) to perform these different tasks. We did not use any subsampling or pooling methods in smNet for position and angular estimation. Instead, we found a stride of 4 in the fourth residual block for estimation of the amplitudes of 12 Zernike modes (from astigmatism to second spherical), and a stride of 4 in both the fourth and eighth residual blocks for estimation of 21 Zernike modes (from astigmatism to third spherical).

Sample preparation. Immediately before SMSN imaging, a round coverslip (25-mm diameter) containing immunostained COS-7 cells was placed on a custom-made sample holder, and 150 µl of imaging buffer (10% (w/v) glucose in 50 mM Tris (JT4109-02; VWR), 50 mM NaCl (S271-500; Fisher Scientific), 10 mM sodium borohydride (452882-25 G; Sigma-Aldrich) in PBS). After fluorescence quenching, cells were washed three times with PBS and treated for 10 min with 10 mM Tris, pH 7.3 (JT4109-02; VWR). Cells were then rinsed three times with deionized water and mounted in an Attofluor cell chamber (A7816; Life Technologies). Then 600 µl of imaging buffer (as described above) was added to the chamber and covered with mineral oil.

Immunoﬂuorescence labelling. COS-7 cells (CRL-1651; ATCC) were seeded on 25-mm-diameter coverslips (CHSP-N1.5-25; Bioscience Tools, San Diego, CA) 1–2 d before immunoﬂuorescence labelling. Cells were first rinsed three times with PBS and prewarmed at (37 °C) 5% CO2 (95% air) before use. Then cells were rinsed for 15 min at RT with prewarmed at (37 °C) 3% paraformaldehyde (75219-500G; Sigma-Aldrich) for 2 h. The diﬀusion was prepared from a stock solution that we made by dissolving a small amount of Alexa Fluor 647 (A20006; Life Technologies) powder in DMSO (27985-1000 ML; Sigma-Aldrich); the color was dark blue. The sample was then rinsed three times with deionized water and mounted in an Attofluor cell chamber (A7816; Life Technologies). Then 600 µl of imaging buffer (as described above) was added to the chamber and covered with mineral oil.

Data acquisition. The experimental complex PSFs (Supplementary Note 10) were collected on a custom-built microscope (W-WP4SMSS; custom-designed from a previous design). The bead sample was excited with a 642-nm laser (2RU-VFL-P-2000-642-B1R; MPB Communications Inc., Canada) at an excitation intensity of 12 W/cm². We adjusted the sample’s z position by moving a piezo-driven nano-stage (P-541.ZCD; Physik Instrumente, Canada). We generated the complex PSF shape by applying a distorted wavefront at the pupil plane using a deformable mirror (MultDM-5.5; Boston Micromachines). The wavefront consisted of a combination of mirror mode 6 (resembling trefoil in Zernike polynomial) and Zernike polynomial 5 (Wyant ordering, astigmatism) and their amplitudes of 0.48 and 0.32 (unit: µm), respectively. Data for the generation of training PSFs were acquired at 2 positions ranging from –1.5 µm to 1.5 µm, with a step size of 0.1 µm, and one frame rate of 50 Hz. For testing, data were acquired at z positions from –1 µm to 1 µm, with a step size of 100 nm, a frame rate of 10 Hz, and 20 frames per z position. We used different fluorescent beads in the same sample to acquire data for training and testing.

The PSF images used for phase retrieval were collected on a custom-built microscope. The bead sample was excited with a 642-nm laser (2RU-VFL-P-2000-642-B1R; MPB Communications Inc., Canada) at an excitation intensity of 55 W/cm². We adjusted the sample’s z position by moving a PIFOC objective positioner (ND72Z2LAQ; Physik Instrumente, Germany). Data were acquired at z positions from –1 µm to 1 µm, with a step size of 100 nm, a frame rate of 50 Hz, and 100 frames per z position. Acquisition of a single dataset of experimental PSFs normally took from 1 to 5 min. We therefore did not expect significant drift in our SMSN system. We note that if sample drift were significant during experimental PSF calibration, the experimental PSF would tilt and subsequently worsen the accuracy of single-molecule analysis. Therefore, we recommend that the experimental PSF be
obtained within a short acquisition window. For systems that exhibit substantial drift (typically SMSN systems have much less drift), it is therefore mandatory to have a closed-loop drift-compensation mechanism enabled during acquisition (for example, with fiduciary markers).

For beads imaged at the top coverslip (Supplementary Note 10), we measured the distance between the two coverslips by first recording the piezo stage position when the dusts on the bottom coverslip were in focus, then recording a second piezo stage position when the beads were in focus. We then estimated the distance as the difference between the two recorded positions.

The SMSN data of COS-7 cells were collected on a custom-built microscope. The sample was first excited with a 642-nm laser (2RU-VFL-P-2000-642-B1R; MPB Communications Inc., Canada) at a low intensity of 55 W/cm² to find a region of interest. Then the blinking data were collected at a laser intensity of 6–9 kW/cm² and a frame rate of 71 Hz. During data acquisition, a 405-nm laser (DL-405-100; CrystaLaser) was used as an activation laser and was gradually adjusted from 0 to 37 W/cm². For data with a single optical section, the mitochondria structure was imaged at around 1 μm from the coverslip surface (Supplementary Fig. 11). For data with multiple optical sections, the mitochondria structure was imaged from the top coverslip surface (Fig. 2a–c) to the top of the cell, with a step size of 400 nm in the axial direction. Typically, 90,000–180,000 frames were collected for each dataset.

The blinking data for aberration measurement were collected in the same manner as described above. For TOMM20 in COS-7 cells, the mitochondria structure was imaged from the top of the cell to the top coverslip surface at a step size of 1 μm and a frame rate of 50 Hz. For dyes immobilized on the coverslip, we collected data by applying a single aberration type using the deformable mirror every 600–700 frames, with a frame rate of 50 Hz.

**Data simulation.** A 3D Gaussian-PSF model was used to generate the localization result of aberrated astigmatism PSFs in Fig. 2d (see Supplementary Table 3 for simulation parameters and Supplementary Note 5 for simulation details). The aberrated astigmatism PSFs and double-helix PSFs in Fig. 2e,f were simulated on the basis of scalar diffraction theory, and their pupil functions were modified with either index mismatch aberration at a depth of 1 μm or the transfer function of propagation-invariant wave fields (Supplementary Table 3 and Supplementary Note 5). The error radius at each photon/background pair was calculated from the average over all the errors of the 11,000 (or 21,000) localizations from smNet. Each dashed red circle in Fig. 2e,f represents the averaged error radius generated from Monte Carlo simulation based on the CRLB (Supplementary Note 7), which simulates a localization result by sampling from a Gaussian distribution with its mean equal to the true position and a variance equal to the CRLB.

The test data for Fig. 3a,b were simulated from Fourier transform of pupil functions (wavefront distortion) composed of 12 or 21 Zernike aberration modes (from astigmatism to second spherical). Each Zernike mode was simulated with random amplitude (in the range of –159.1549 to 159.1549 m). Supplemented Table 3 and Supplementary Fig. 13 for simulation parameters and Supplementary Note 5 for simulation details).

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

**Code availability.** LuaJIT scripts for training smNet, Matlab script for generating various PSF models and corresponding calculation of the CRLB, Matlab script for phase retrieval, and Matlab script for estimation of total photon background photon counts are available as Supplementary Software, and further updates will be made available at https://github.com/HuanglabPurdue/smNet.

**Data availability**

The data that support the findings of this study are available from the corresponding author upon request. Example data are available in the Supplementary Data and Supplementary Software packages.

**References**

21. Lecun, Y., Bottou, L., Bengio, Y. & Haffner, P. Proc. IEEE Inst. Electr. Electron. Eng. 86, 2278–2244 (1998).
22. Nielsen, M. A. Neural Networks and Deep Learning. (Determination Press, San Francisco, 2015).
23. Ioffe, S. & Szegedy, C. in Proc. 32nd International Conference on Machine Learning Vol 37 (eds Bach, F. & Leli, D.) 448–456 (JMLR/Microtome Publishing, Cambridge, MA, 2015).
24. He, K., Zhang, X., Ren, S. & Sun, J. in Proc. 2015 IEEE International Conference on Computer Vision (ICCV) (eds Bajcsy, R. et al.) 1026–1034 (IEEE, Piscataway, NJ, 2015).
25. Huang, F. et al. Cell 166, 1028–1040 (2016).
26. Wang, B. & Booth, M. J. Opt. Commun. 282, 4467–4474 (2009).
27. Wyant, J. C. & Creath, K. in Applied Optics and Optical Engineering Vol XI (eds Shannon, R. R. & Wyant, J. C.) 1–53 (Academic, New York, 1992).
28. Smith, C. S., Joseph, N., Rieger, B. & Lidke, K. A. Nat. Methods 7, 373–375 (2010).
29. Hanser, B. M., Gustafsson, M. G. L., Agard, D. A. & Sedat, J. W. J. Microsc. 216, 32–48 (2004).
30. Piestun, R., Schechner, Y. Y. & Shamir, J. J. Opt. Soc. Am. A Opt. Image Sci. Vis. 17, 294–303 (2000).
31. Ober, R. J., Ram, S. & Ward, E. S. Biophys. J. 86, 1185–1200 (2004).
32. Liu, S. & Lidke, K. A. ChemPhysChem 15, 696–704 (2014).
Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size
   Describe how sample size was determined.
   For biological data, the number of sub-regions analyzed is determined by the number of emission events obtained from single molecule switching nanoscopy experiments. For simulated data, each photon-background case for characterizing smNet performance consists 1,000 to 1,000,000 emission pattern to allow accurate performance quantification (Supplementary Table 3).

2. Data exclusions
   Describe any data exclusions.
   No data were excluded from this study

3. Replication
   Describe whether the experimental findings were reliably reproduced.
   All attempts at replication were successful

4. Randomization
   Describe how samples/organisms/participants were allocated into experimental groups.
   Training of smNet used uniform random distributed properties such as location of single molecule, orientation and aberrations in terms of Zernike modes. Testing datasets are also drawn from uniform random distribution from the tested range (see supplementary Table 2-4 for more details on ranges used in training and testing).

5. Blinding
   Describe whether the investigators were blinded to group allocation during data collection and/or analysis.
   Training and testing of the developed algorithm were performed based uniform random distributed measurement properties generated by random number generator.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.
6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

| n/a | Confirmed |
|-----|-----------|
| ☑️ | The exact sample size \( n \) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.) |
| ☑️ | A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☑️ | A statement indicating how many times each experiment was replicated |
| ☑️ | The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section) |
| ☑️ | A description of any assumptions or corrections, such as an adjustment for multiple comparisons |
| ☑️ | The test results (e.g. \( P \) values) given as exact values whenever possible and with confidence intervals noted |
| ☑️ | A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range) |
| ☑️ | Clearly defined error bars |

See the web collection on statistics for biologists for further resources and guidance.

7. Software

Describe the software used to analyze the data in this study.

The developed 'smNet' LuaJIT software package, PSF toolbox package and Phase retrieval package. The source code is compressed as a zip file and included as supplementary software. It includes scripts and functions written in LuaJIT to load, train and evaluate example training and testing datasets (also included in the zip file) and MatLab code to calculate PSF models in the manuscript and perform phase retrieval. pdf files describing the usage of the software are also included.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). Nature Methods guidance for providing algorithms and software for publication provides further information on this topic.

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

No unique materials were used

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

anti-TOM20 primary antibody is a commercially available antibody (Catalogue # sc-11415, Santa Cruz Biotechnology) and this antibody has been validated and reported previously (PMID: 27397506, https://doi.org/10.1016/j.cell.2016.06.016).

10. Eukaryotic cell lines

   a. State the source of each eukaryotic cell line used.
   
   COS-7 cell from Yale culture facility
   
   b. Describe the method of cell line authentication used.
   
   Cells were not authenticated
   
   c. Report whether the cell lines were tested for mycoplasma contamination.
   
   Cell line was test negative for mycoplasma contamination
   
   d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.
   
   No commonly misidentified cell lines were used.
Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals
   Provide details on animals and/or animal-derived materials used in the study.
   No animals were used

Policy information about studies involving human research participants

12. Description of human research participants
   Describe the covariate-relevant population characteristics of the human research participants.
   The study did not involve human research participants