Analyzing complex single molecule emission patterns with deep learning

Peiyi Zhang[1,†], **Sheng Liu**[1,†], **Abhishek Chaurasia**[2], **Donghan Ma**[1], **Michael J. Mlodzianoski**[1], **Eugenio Culurciello**[1,2], and **Fang Huang**[1,3,4,*]

[1]Weldon School of Biomedical Engineering, Purdue University, West Lafayette, Indiana, USA

[2]School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana, USA

[3]Purdue Institute for Integrative Neuroscience, Purdue University, West Lafayette, Indiana, USA

[4]Purdue Institute of Inflammation, Immunology and Infectious Disease, Purdue University, West Lafayette, Indiana, USA

Abstract

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 enable retrieving 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.

Editor's summary

The deep neural network smNet enables extraction of multiplexed parameters such as 3D position, orientation and wavefront distortion from emission patterns of single molecules.

Analyzing single-molecule emission patterns plays a critical role in retrieving the structural and physiological information of their tagged targets and, further, understanding their interactions and cellular context[1]. These emission patterns of tiny light sources (i.e. point spread functions, PSFs) encode information such as molecule's location[2], orientation[3], environment within the specimen[4] and the path the emitted photons took before reaching the camera[5]. Detecting and tracking single-fluorescent probes through their emission patterns lay the foundation of modern single-molecule based imaging methods[2]. These methods allow interrogation of cellular dynamics[6], transcriptional regulation[7] and protein interaction[8]

*To whom correspondence should be addressed: Fang Huang, fanghuang@purdue.edu.

and further enable single-molecule based super-resolution imaging in both fixed and living specimens[9].

One key in single molecule studies is to understand how the features of the PSF encode the properties of a single molecule, i.e., the analysis of single molecule emission pattern[2]. Conventionally, the analysis focuses on dominant and recognizable features of the PSFs[10], such as rotation of the double helix PSF[11], and on modeling feature changes through a simplified mathematical form, e.g., Gaussian PSF model. Most recently, interpolation based numerical methods, such as splines, have been demonstrated for three dimensional localization of single molecules[12,13]. However, the number of parameters required in the interpolation scales exponentially with increasing PSF complexities and measurement dimensions[12], challenging its application in retrieving information carried by high-dimensional PSFs (Supplementary Note 1). Furthermore, single-molecule emission pattern carries multiple classes of molecular information simultaneously. Due to difficulties in perceiving and summarizing a comprehensive PSF model, retrieving multiplexed information beyond 3D position from complex or high-dimensional single-molecule data remains challenging.

Deep neural networks (DNN) extract features from the input and learns its connection to the output automatically[14]. A neural network in its basic form (1-2 fully-connected layers) has been utilized to identify fluorophore species in single-molecule fluorescence lifetime experiments[15] as well as to speed up dipole orientation estimation from an analytical approximation of the dipole PSF[16]. Through the deep architecture of DNN, the complex mapping between input and output is extracted from different levels of features hierarchically[14]. The inference precision, instead of depending on domain expertise (e.g. feature recognition), now mainly depends on the design of the network architecture[14].

We have applied deep learning to extract multiplexed information carried by single-molecule patterns skipping conventional steps such as feature recognition, model simplification, and the iterative regression methods. For high-dimensional single-molecule data, we designed the network to tackle each inference task independently allowing complexities in each dimension to add instead of multiplying. 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 distortions from complex and subtle features of the PSFs, which otherwise are considered too complex for established algorithms.

The general principle of our DNN for single-molecule studies is illustrated in Fig. 1 (referred as 'smNet'). smNet is a deep network of 27 to 36 layers (Supplementary Figs. 1–3 and Supplementary Table 1) consisted of convolutional layers, residual blocks and fully connected layers together with batch normalization and parametric rectified linear unit (PReLU). The complex and subtle features within the PSF lie in the photon distribution within a small sub-region (2-10 μm^2). To fully utilize the information contained in the spatial domain, large kernel size was used in beginning layers, and a number of convolutional layers and 'bottleneck' residual blocks[17] were stacked to capture as many levels of features as possible. This architecture helped smNet to learn different levels of

features and to generate an optimal mapping from the input to the output (Supplementary Notes 2 and 3). smNet can be trained using either simulated or measured PSFs. Importantly, 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 at a large range of detected photons and background levels simultaneously. We found that it is sufficient to train smNet with ~1 million PSF patterns for each inference task with a reasonably large range of their measured properties. (Supplementary Fig. 4 and Supplementary Table 2)

To test the performance of smNet, we first evaluated the precision and accuracy when pin-pointing single-molecule centers on both simulated and experimentally obtained single-molecule patterns. We found that smNet localizes single molecules at a precision matching the theoretical information limit given by the CRLB with a small or ignorable bias despite the significant amount of aberrations and PSF complexity (2.35 ± 2.71 nm (bias ± s.t.d.) for astigmatism, 1 ± 0.84 nm for double helix PSFs, 2 ± 1.8 nm for simulated complex PSFs, and 102.1 ± 7 nm (mean ± s.t.d.) for experimental complex PSFs, Supplementary Figs. 5–7). This performance is consistently achieved at various conditions such as molecule locations, intensity and background levels (Fig. 2e,f, Supplementary Fig. 8–10, Supplementary Table 3 and Supplementary Notes 7–9 including testing results of smNet on various conditions). We further demonstrated smNet in three-dimensional single-molecule switching nanoscopy (SMSN) experiments. SMSN relies on localization of millions of PSFs down to a precision of 10-25 nm, which together with localization accuracy are essential to successfully reconstruct 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 a ~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[18] (Supplementary Note 10). This allows smNet to retrieve the correct molecular positions despite the significant amount of aberrations (Fig. 2d, Supplementary Figs. 1B and 5). We found that the x-z cross sections of the mitochondria outer membrane show significant artifacts from conventional Gaussian-based methods, while smNet accurately reconstruct 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, three-dimensional positions[2], probe spectra[19], identities[15] and orientations[3]. However, encoding two or more classes of information in the emission patterns will increase their dimensionality, which challenges the traditional decoding processes such as feature recognition and regression.

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 allows 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 approach the information limits in a large parameter range while degeneracies and wrappings of both polar and azimuthal angles can also be correctly predicted (Supplementary Fig. 12).

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

Using smNet, we are able to extract sample induced aberration through the raw single-molecule blinking data itself allowing wavefront measurement deep into the specimen. As a demonstration, we measured the evolution of 12 Zernike aberration modes through 11 consecutive optical sections through a immunolabeled specimen (TOM20 in COS-7 cells) on a custom-built biplane setup. We found that 1st order sepherical aberration evolves continuously while aberrations such as diagonal astigmatism decreases with increasing depth (Fig. 3c,d and Supplementary Video 2).

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

We developed smNet, a deep neural network for complex and high-dimensional analysis of single-molecule emission patterns. 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 as well as 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 extracting highly multiplexed physical and physiological information through the emission pattern of a single molecule.

Online Methods

Optical Setup

All experimental data (except for complex PSFs and wavefront estimation data) were recorded on a custom-built single molecule switching nanoscopy (SMSN) setup built around an Olympus IX-73 microscope stand (IX-73, Olympus America Inc., Waltham, MA) with a 100×/1.35 NA silicone oil-immersion objective lens (FV-U2B714, Olympus America Inc.), a 405 nm laser (DL-405-100, CrystaLaser, Reno, NV) and a 642 nm laser (2RU-VFL-P-2000-642-B1R, MPB Communications Inc.) for activation and excitation, respectively. The filter turret contains a dichroic mirror (Di03-R405/488/561/635-t1, Semrock Inc.). A deformable mirror (MultiDM-3.5, Boston Micromachines, Cambridge, MA) placed at the conjugated pupil plane is used for correcting systematic aberrations and introducing astigmatism for 3D SMSN. Collected fluorescence emission passed through a bandpass filter (FF01-731/137-25, Semrock Inc.) placed just before the camera. The fluorescence signal was recorded on an EMCCD camera (C9100-23B, Hamamastu, Tokyo, Japan). The overall system magnification was ~141×, 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 optical path length difference 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 mirror (47005, Edmund Optics) and received by a sCMOS camera (Orca-Flash4.0v3, Hamamastu). The overall system magnification was ~53×, resulting in an effective pixel size of 122 nm. A 100x/1.4 NA oil immersion objective (UPLSAPO 100XO, Olumpus America Inc., Waltham, MA) was used for wavefront distortion measurements. Biplane or multi-plane setup is preferred in wavefront distortion measurement to avoid degeneracies between aberration modes.

smNet architecture

smNet is composed of 3 to 5 convolutional layers[21] (Supplementary Note 2.1), 7 to 11 residual blocks[17] (Supplementary Note 2.2) and 0 to 2 fully connected layers[22]. Each convolutional layer is followed by batch normalization[23] (Supplementary Note 2.4) and PReLU[24] (Supplementary Note 2.3), except for the last convolutional layer in M3 (Supplementary Table 1). The first fully connected layer (FC) is followed by a PReLU and the last FC is followed by a HardTanh (https://github.com/torch/nn/blob/master/doc/

transfer.md). The detailed information about smNet architecture and its variations are shown Supplementary Table 1.

Since the input image has a small size and the features of PSF span across a small number of pixels, it is imperative that we fully utilize the information contained in the spatial domain. To achieve this, we used larger kernels in beginning layers as compared to later layers of our neural network. We started with 64 kernels with a size of 7 by 7 pixels in the first layer followed by 128 kernels with a size of 5 by 5 pixels.

After this, we focused on capturing as many rich features as possible. Stacking large number of convolutional layers help us to achieve this, however, they often make neural networks untrainable[17]. To avoid this, we used a stack of 7 to 11 residual blocks in our architecture. Each residual block utilized the 'bottleneck' structure[17], where the number of features is first squeezed and then expanded. This design not only helps in reducing the number of training parameters but also in learning more relevant features.

In later layers, we assume that there is much less spatial information left to be learnt by smNet, we used 1×1 convolutional layers. Finally, they are followed by fully connected layers. We found that reducing the number of both fully connected layers and 1×1 convolutional layers helps in improving the accuracy in wavefront distortion estimation.

In our study, the output of smNet is a vector of 12 or 21 elements representing the amplitudes of 12 or 21 Zernike modes, or a vector of 2 elements representing x and y coordinates, or a scalar representing the z position, polar angle (α) or azimuthal angle (β). Since, x, y positions are based on the emitter's location in the sub-region, and the axial 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 didn't use any subsampling and pooling methods in smNet for position and angle estimations. However, we found it helpful to add a stride of 4 in the 4th residual block for estimating the amplitudes of 12 Zernike modes (from astigmatism to 2nd spherical), and stride of 4 in both 4th and 8th residual block for estimating 21 Zernike modes (from astigmatism to 3rd spherical).

Sample Preparation

Immediately before SMSN imaging, round coverslip (25 mm diameter) containing immune-stained COS-7 cells was placed on a custom-made sample holder, and 150 μL imaging buffer (10% (w/v) glucose in 50 mM Tris (JT4109-02, VWR), 50 mM NaCl (S271-500, Fisher Scientific), 10 mM MEA (M6500-25G, Sigma-Aldrich), 50 mM BME (M3148-25ML, Sigma-Aldrich), 2 mM COT (138924-1G, Sigma-Aldrich), 2.5 mM PCA (37580-25G-F, Sigma-Aldrich) and 50 nM PCD (P8279-25UN, Sigma-Aldrich), pH 8.0) was added on top of the coverslip. Then a cleaned coverslip of the same size was carefully placed on top of it and the excessive buffer was removed. The sample was sealed with melted Valap. Samples with cells on the top coverslip were prepared in a similar manner by

placing the cleaned coverslip at the bottom of sample holder and the coverslip with cells on top of it (the surface with cells facing down).

To obtain experimental complex PSFs, 200 nm diameter fluorescent beads (F8806, Invitrogen) with a dilution of $1:10^4$ in deionized water were used. For experiments using phase retrieval method, 100 nm diameter beads (crimson, custom-designed, Invitrogen) with a dilution of $1:10^6$ in deionized water were used. After diluting the desired type of beads, 200 μL poly-L-lysine (P4707-50ML, Sigma-Aldrich) was added to the center of a round coverslip (25 mm diameter) placed on a custom-made sample holder and was incubated for 20 minutes. Then the sample was rinsed once with deionized water and 200 μL diluted bead solution was added and subsequently incubated for 20 minutes at room temperature. Then the sample was rinsed with deionized water, drained and added with 10 μL deionized water (or 97% TDE (166782-500G, Sigma-Aldrich) for bead samples used in wavefront distortion measurements). Subsequently, a second pre-cleaned coverslip was placed on top and the sample was sealed with two-component silicone sealant (Picodent Twinsil, Picodent, Germany).

The dye coated coverslip sample were prepared by first adding 200 μL poly-L-lysine (P4707-50ML, Sigma-Aldrich) onto a 25 mm round coverslip and was incubated for 1 hour. Then the sample was rinsed once with deionized water and was incubated with 200 μL $1:10^6$ dye dilution in 0.1 M sodium bicarbonate (792519-500G, Sigma-Aldrich) for 2 hours. The dye dilution was prepared from a stock solution that was made by dissolving a small amount of Alexa Fluor 647 (A20006, Life Technologies) powder in DMSO (276855-100ML, Sigma-Aldrich), the color was dark blue. The sample was then rinsed three times with deionized water and mounted in a Attofluor cell chamber (A7816, Life Technologies). Then 600 μL imaging buffer (as described above) was added to the chamber and was covered with mineral oil on top.

Immunofluorescence labeling

COS-7 cells (CRL-1651, ATCC) were seeded on 25 mm diameter coverslips (CSHP-No1.5-25, Bioscience Tools, San Diego, CA) 1~2 days before immunofluorescence labeling. Cells were first rinsed three times with pre-warmed (at 37 ℃) phosphate buffered saline (PBS, 806552-500ML, Sigma-Aldrich) and then fixed for 15 minutes at room temperature (RT) with pre-warmed (at 37 ℃) 3% paraformaldehyde (PFA, 15710, Electron Microscopy Sciences, Hatfield, PA) and 0.1% glutaraldehyde (GA, Electron Microscopy Sciences, 16019, Hatfield, PA) in PBS. Cells were then washed twice with PBS and treated for 7 minutes in freshly-prepared fluorescence quenching buffer (0.1% sodium borohydride (452882-25G, Sigma-Aldrich) in PBS). After fluorescence quenching, cells were washed three times with PBS and treated for 10 minutes with 10 mM Tris (pH 7.3, JT4109-02, VWR). Cells were then rinsed three times with PBS and permeabilized with blocking buffer (3% bovine serum albumin (BSA, 001-000-162, Jackson ImmunoResearch) and 0.2% Triton X-100 (X100, Sigma-Aldrich) in PBS) for 30 minutes, gently rocking at RT. After blocking, cells were incubated with anti-TOMM20 primary antibody (sc-11415, Santa Cruz Biotechnology), diluted to 1:500 in 1% BSA and 0.2% Triton X-100 in PBS, at RT for 12 hours. Cells were then washed three times each time for 5 minutes with wash buffer (0.05%

Triton X-100 in PBS) and incubated with secondary antibody conjugated with Alexa Fluor 647 (A21245, Life Technologies, Grand Island, NY), diluted to 1:500 in 1% BSA and 0.2% Triton X-100 in PBS, at RT for 4 hours. After incubation with secondary antibody, cells were washed three times each time for 5 minutes with wash buffer. And then cells were post-fixed with 4% PFA in PBS for 10 minutes. After post-fixation, cells were rinsed three times with PBS and stored in PBS at 4 ℃ until they were imaged.

Data Acquisition

The experimental complex PSFs (Figure SS13) were collected on a custom-built microscope (W-4PiSMSN constructed from the previous design[25]). 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/cm^2. The sample's z position was adjusted by moving a piezo-driven nano-stage (P-541.ZCD, Physik Instrumente, Germany). The complex PSF shape was generated by applying a distorted wave front at the pupil plane of the emission path using a deformable mirror (MultiDM-5.5, Boston Micromachines). The wave front consisted of a combination of the mirror mode[5,26] 6 (resembling trefoil in Zernike polynomial) and Zernike polynomial[27] 5 (Wyant ordering, astigmatism) and their amplitude of 0.48 and 0.32 (unit: λ) respectively. Data for generating training PSFs were acquired at z positions ranging from −1.5 μm to 1.5 μm, with a step size of 10 nm, a frame rate of 1 Hz and taking one frame per axial position. Data for testing 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 taking 20 frames per z position. Data for training and testing were acquired using different fluorescent beads in the same sample.

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/cm^2. The sample's z position was adjusted 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 taking 100 frames per z position.

One dataset of experimental PSF was normally acquired from 1 to 5 minutes. We therefore, does not expecting a significant drift in our SMSN system. We would like to note that if sample drift significantly during experimental PSF calibration, the experimental PSF will tilt and subsequently worsen the accuracy of single molecule analysis. Therefore, we recommend experimental PSF is taken within a short acquisition window. For systems exhibits large drifts (typically SMSN systems have much smaller drift), it is therefore mandatory to have close loop drift compensating mechanism during acquisition (for example, with fiduciary markers).

For beads imaged at the top coverslip (Figure SS13). The distance between the two coverslips was measured 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. The distance was then estimated 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 low intensity of 55 W/cm^2, to find a region of interest (ROI). Then the blinking data were collected at a laser intensity of 6~9 kW/cm^2, and a frame rate of 71 Hz. During data acquisition, a 405 nm laser (DL-405-100, CrystaLaser) was used as an activate laser and was gradually adjusted from 0 to 37 W/cm^2. 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 axial. Typically, 90,000 to 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 coverslip, the data were collected by applying a single aberration type using the DM every 600 to 700 frames, with a frame rate of 50 Hz.

Data simulation

A 3D Gaussian-PSF model[28] 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[18,29] and double-helix PSFs[11] in Fig. 2e,f were simulated based on scalar diffraction theory and their pupil functions were modified either with index mismatch aberration at 12 μm depth or the transfer function of propagation-invariant wave fields[30] (Supplementary Table 3 and Supplementary Note 5). The error radius at each photon/background pair (blue circle) was calculated from the average over all the errors of the 11000 (or 21000) localizations from smNet. Each dashed red circle represents the averaged error-radius generated from Monte-Carlo simulation based on CRLB[31,32] (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[27] (from Astigmatism to 2nd Spherical). Each Zernike mode was simulated with random amplitude (in the range of −159.1549 to 159.1549 mλ, see Supplementary Table 3 and Supplementary Fig. 13 for simulation parameters and Supplementary Note 5 for simulation details).

Data Availability Statement

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

Code Availability Statement

LuaJIT scripts for training smNet, Matlab script for generating various PSF models and corresponding calculation of CRLB, Matlab script for phase retrieval, and Matlab script for estimation of total photon background photon counts are available in Supplementary

Software and further updates will be made available at https://github.com/HuanglabPurdue/smNet.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

References:

1. Moerner WE, Fromm DP, Methods of single-molecule fluorescence spectroscopy and microscopy. Review of Scientific Instruments 74, 3597–3619 (2003).

2. von Diezmann A, Shechtman Y, Moerner WE, Three-Dimensional Localization of Single Molecules for Super-Resolution Imaging and Single-Particle Tracking. Chem. Rev 117, 7244–7275 (2017). [PubMed: 28151646]

3. Backlund MP, Lew MD, Backer AS, Sahl SJ, Moerner WE, The role of molecular dipole orientation in single-molecule fluorescence microscopy and implications for super-resolution imaging. ChemPhysChem 15, 587–599 (2014). [PubMed: 24382708]

4. Moon S et al., Spectrally resolved, functional super-resolution microscopy reveals nanoscale compositional heterogeneity in live-cell membranes. J. Am. Chem. Soc 139, 10944–10947 (2017). [PubMed: 28774176]

5. Burke D, Patton B, Huang F, Bewersdorf J, Booth MJ, Adaptive optics correction of specimen-induced aberrations in single-molecule switching microscopy. Optica 2, 177–185 (2015).

6. Andrews NL et al., Actin restricts FcepsilonRI diffusion and facilitates antigen-induced receptor immobilization. Nat. Cell Biol 10, 955–963 (2008). [PubMed: 18641640]

7. Femino AM, Fay FS, Fogarty K, Singer RH, Visualization of single RNA transcripts in situ. Science 280, 585–590 (1998). [PubMed: 9554849]

8. Ha T et al., Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. Proc. Natl. Acad. Sci 93, 6264–6268 (1996). [PubMed: 8692803]

9. Baddeley D, Bewersdorf J, Biological Insight from Super-Resolution Microscopy: What We Can Learn from Localization-Based Images. Annu. Rev. Biochem 87, 965–989 (2018). [PubMed: 29272143]

10. Sage D, et al. Quantitative evaluation of software packages for single-molecule localization microscopy. Nat. Methods 12, 717–724 (2015). [PubMed: 26076424]

11. S. R. Pavani, 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). [PubMed: 19211795]

12. Babcock HP, Zhuang X, Analyzing single molecule localization microscopy data using cubic splines. Sci. Rep 7, 552 (2017). [PubMed: 28373678]

13. Li Y et al., Real-time 3D single-molecule localization using experimental point spread functions. Nat. Methods 15, 367–369 (2018). [PubMed: 29630062]

14. LeCun Y, Bengio Y, Hinton G, Deep learning. Nature 521, 436–444 (2015).

15. Bowen BP, Scruggs A, Enderlein J, Sauer M, Woodbury N, Implementation of neural networks for the identification of single molecules. J. Phys. Chem. A 108, 4799–4804 (2004).

16. Zhang Y, et al., Ultrafast, accurate, and robust localization of anisotropic dipoles. Protein & Cell 4, 598–606 (2013). [PubMed: 23744341]

17. He K, Zhang X, Ren S, Sun J, Deep residual learning for image recognition Proc. of the IEEE conference on computer vision and pattern recognition 770–778 (2016).

18. Liu S, Kromann EB, Krueger WD, Bewersdorf J, Lidke KA, Three dimensional single molecule localization using a phase retrieved pupil function. Opt. Express 21, 29462 (2013). [PubMed: 24514501]

19. Cutler PJ et al., Multi-color quantum dot tracking using a high-speed hyperspectral line-scanning microscope. PLoS One 8, e64320 (2013). [PubMed: 23717596]

20. Ji N, Adaptive optical fluorescence microscopy. Nat. Methods 14, 374–380 (2017). [PubMed: 28362438]

Reference for Online Methods:

21. Lecun Y, Bottou L, Bengio Y, Haffner P, Gradient-based learning applied to document recognition. Proc. IEEE 86, 2278–2324 (1998).

22. Nielsen MA, Neural Networks and Deep Learning (Determination Press, 2015).

23. Szegedy C, Batch normalization: accelerating deep network training by reducing internal covariate shift. International Conference on Machine Learning 448–456 (2015).

24. He K, Zhang X, Ren S, Sun J, Delving deep into rectifiers: surpassing human-level performance on ImageNet classification. Proc. of the IEEE international conference on computer vision 1026–1034 (2015).

25. Huang F et al., Ultra-high resolution 3D imaging of whole cells. Cell 166, 1028–1040 (2016). [PubMed: 27397506]

26. Wang B, Booth MJ, Optimum deformable mirror modes for sensorless adaptive optics. Opt. Commun 282, 4467–4474 (2009).

27. Wyant JC, Creath K, "Basic wavefront aberration theory for optical metrology" in Applied Optics and Optical Engineering, (Academic, New York, 1992), Vol. XI, pp. 1–53.

28. Smith CS, Joseph N, Rieger B, Lidke KA, Fast, single-molecule localization that achieves theoretically minimum uncertainty. Nat. Methods 7, 373–375 (2010). [PubMed: 20364146]

29. Hanser BM, Gustafsson MGL, Agard DA, Sedat JW, Phase-retrieved pupil functions in wide-field fluorescence microscopy. J. Microsc 216, 32–48 (2004). [PubMed: 15369481]

30. Piestun R, Schechner YY, Shamir J, Propagation-invariant wave fields with finite energy. J. Opt. Soc. Am. A 17, 294 (2000).

31. Ober RJ, Ram S, Ward ES, Localization accuracy in single-molecule microscopy. Biophys. J 86, 1185–1200 (2004). [PubMed: 14747353]

32. Liu S, Lidke KA, A multiemitter localization comparison of 3D superresolution imaging modalities. ChemPhysChem. 15, 696–704 (2014). [PubMed: 24281982]
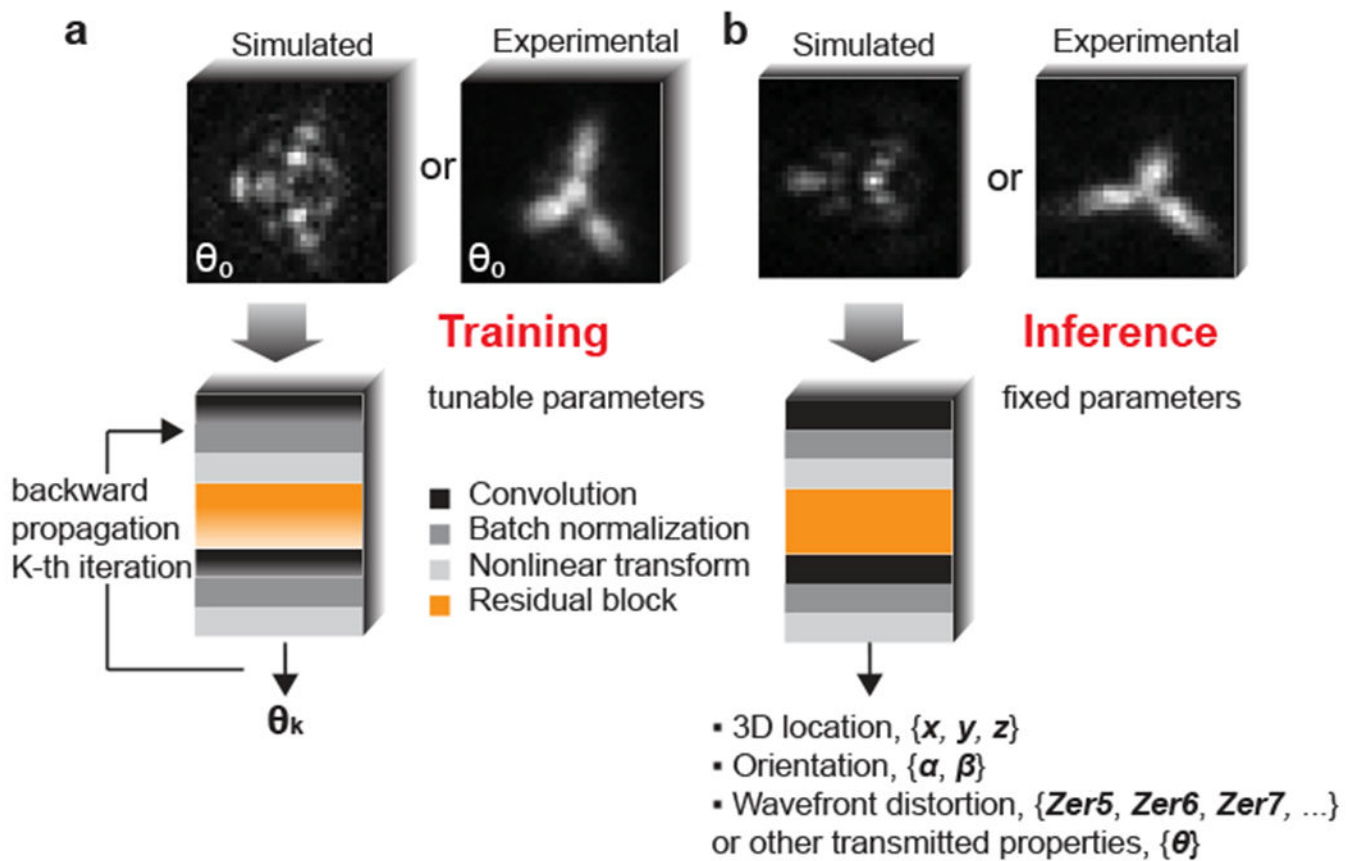
Figure 1: Concept of training and inference with smNet and description of basic architecture.
(**a**) smNet training process. The example input training data set consists of PSF images and the underlying true values of the measurements θ_0. smNet takes in the training images and outputs estimations θ_k for k[th] iteration. Iteratively, through backpropagation, the parameters in the network are updated by minimizing the designed information-limit weighted cost-function. (**b**) The estimation process with smNet. Once trained, smNet takes in a stack of PSF images and outputs the desired estimations. The simplified illustration shows four major building blocks of smNet: convolutional layer, nonlinear transformation, batch normalization and residual block.
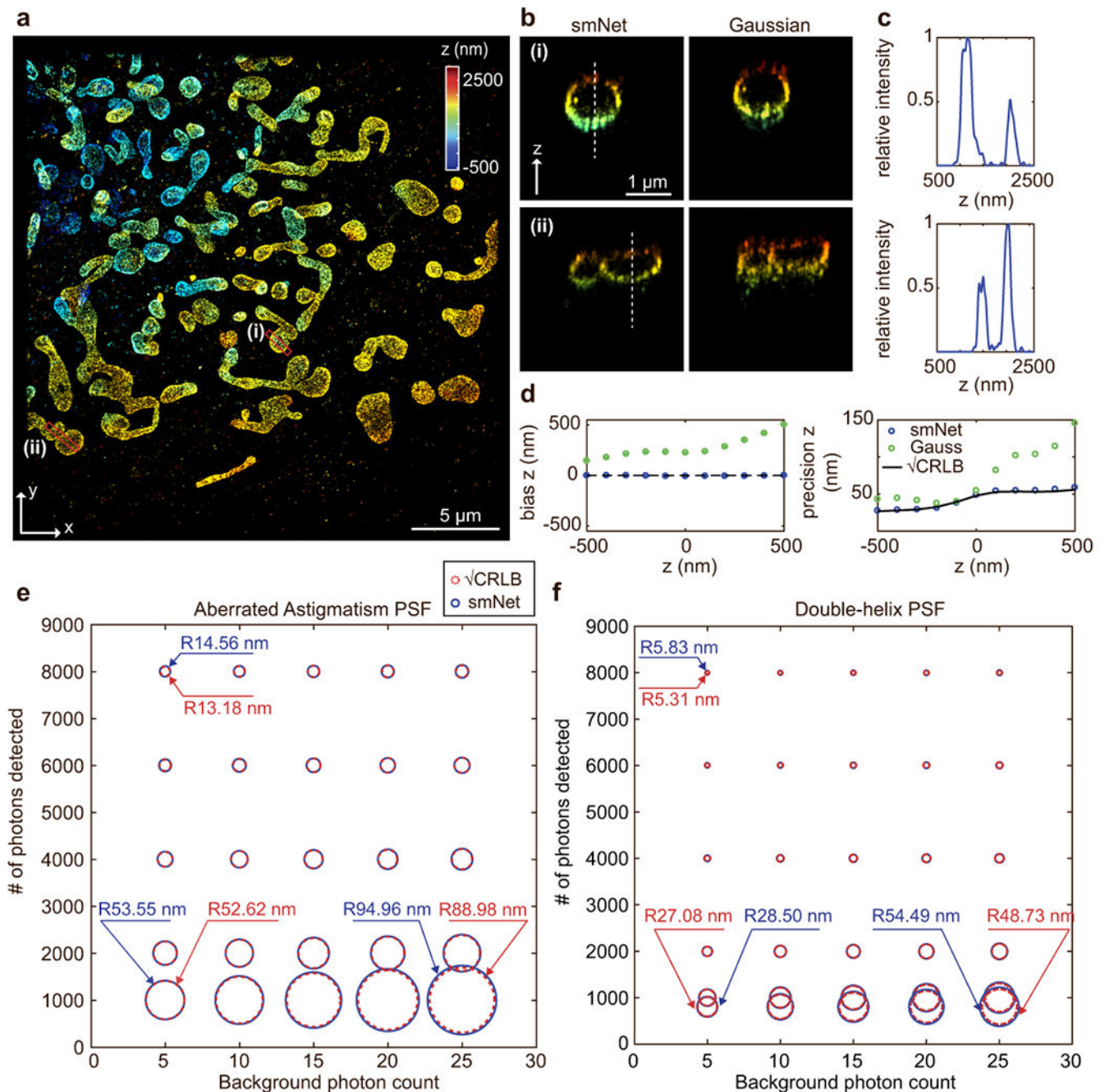
Figure 2: 3D-SMSN reconstruction using smNet and its error radii comparison with the CRLB for aberrated astigmatism and double helix PSFs.

(**a**) 3D-SMSN reconstruction using smNet. The sample is TOM 20 in a COS-7 cell at a depth of 12 μm from coverslip surface. Color represents the relative axial position of the localized molecules. The image is representative of 9 independently repeated experiments. (**b**) Comparison of cross sections of selected regions (red boxes in a) from both smNet reconstruction and Gaussian-based reconstruction (Supplementary Note 12). The data are representative of 6 independently repeated experiments (**c**) Intensity profiles along the dashed lines in b. The data are representative of 6 independently repeated experiments. (**d**)

Comparison of precision and accuracy for axial localization using smNet and Gaussian-based method. Data were generated from an experimentally retrieved pupil function including aberrations at 12 μm above the coverslip. (**e,f**) Error radii achieved at various total photon and background levels for aberrated astigmatism PSFs (same as in d) and double-helix PSFs.
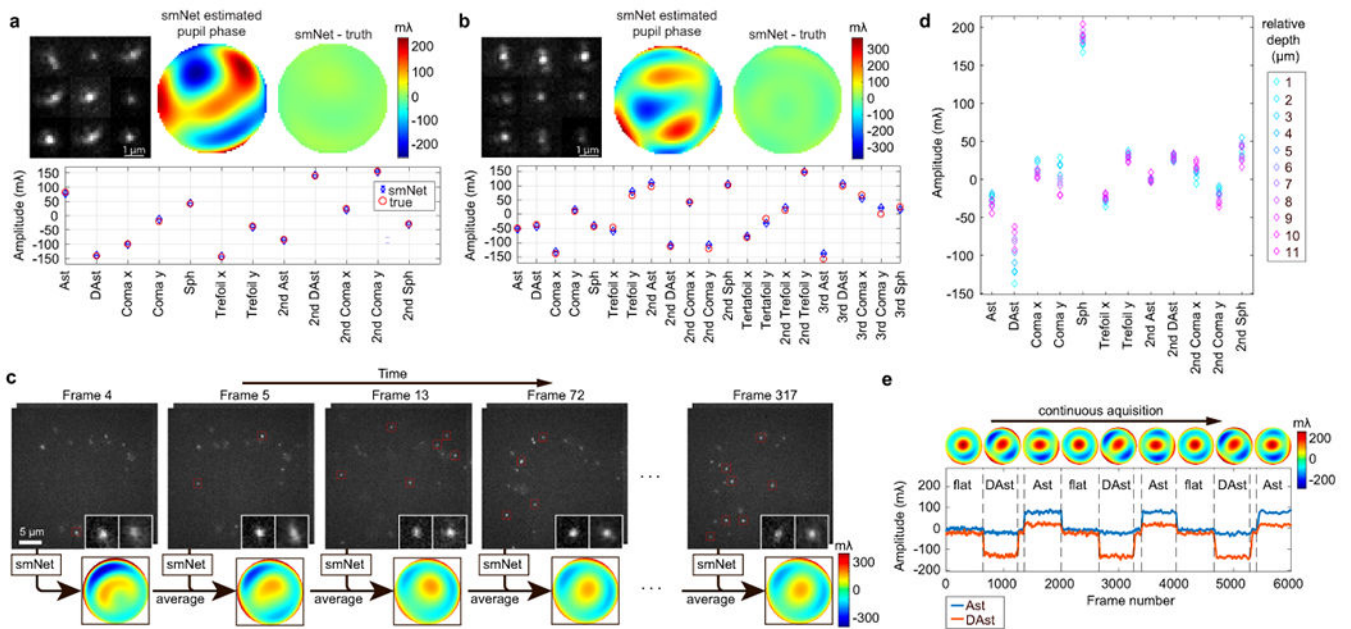
Figure 3: smNet measures wavefront distortions.
(**a,b**) smNet measurement of wavefront shape composed of 12 or 21 Zernike modes through simulated PSFs. The amplitude of each mode is randomly sampled from −159.15 to 159.15 mλ. The nine example PSFs (upper left) are randomly picked from the 1st planes of the 1000 biplane PSFs. The centers and error bars in (a, b) are mean and s.e.m respectively (n = 1000). (**c**) Illustration of smNet's wavefront measurement based on single-molecule blinking data during SMSN imaging. Sub-regions are enlarged to show PSF shape in the biplane setup. (**d**) Measurement of wavefront distortion along 11 optical sections during SMSN imaging of a whole cell. (**e**) Tracking of dynamic wavefront changes introduced by a deformable mirror (DM) using smNet. Each dashed line represents a change of the input voltage to the deformable mirror (DM) shape which resembles normalized Zernike polynomials. Eight experiments were repeated independently with similar result and additional tracking results are included in Supplementary Fig. 15 and Supplementary Data.