Received: 15 October 2024 | Accepted: 23 April 2025

DOI: 10.1111/jmi.13420

Journal of Microscopy RMS

ORIGINAL ARTICLE

Model-free machine learning-based 3D single molecule localisation microscopy

Miguel A. Boland[1] | **Jonathan P. E. Lightley**[2] | **Edwin Garcia**[2] | **Sunil Kumar**[2] | **Chris Dunsby**[2] | **Seth Flaxman**[3] | **Mark A. A. Neil**[2] | **Paul M. W. French**[2] | **Edward A. K. Cohen**[1]

[1]Department of Mathematics, Imperial College, London, UK

[2]Department of Physics, Imperial College, London, UK

[3]Department of Computer Science, Oxford University, Oxford, UK

Correspondence
Paul M. W. French, Department of Physics, Imperial College, London, SW7 2AZ, U.K.
Email: paul.french@imperial.ac.uk

Present address
Paul M. W. French,Department of Mathematics, Imperial College London, London, UK

Funding information
Cancer Research UK, Grant/Award Number: A29368; Silicon Valley Community Foundation, Grant/Award Numbers: 2021-234618, 2023-321240; Wellcome Trust, Grant/Award Number: 203799/Z/16/Z

Abstract

Single molecule localisation microscopy (SMLM) can provide two-dimensional super-resolved image data from conventional fluorescence microscopes, while three dimensional (3D) SMLM usually involves a modification of the microscope, for example, to engineer a predictable axial variation in the point spread function. Here we demonstrate a 3D SMLM approach (we call *'easyZloc'*) utilising a lightweight Convolutional Neural Network that is generally applicable, including with 'standard' (unmodified) fluorescence microscopes, and which we consider may be practically useful in a high throughput SMLM workflow. We demonstrate the reconstruction of nuclear pore complexes with comparable performance to previously reported methods but with a significant reduction in computational power and execution time. 3D reconstructions of the nuclear envelope and an actin sample over a larger axial range are also shown.

KEYWORDS

CNN, deep learning, easySTORM, fluorescence, openFrame, single molecule localisation, super-resolved microscopy

1 | INTRODUCTION

Single Molecule Localisation Microscopy (SMLM) is used to describe a range of imaging techniques to achieve spatial resolution beyond the diffraction-limit in fluorescence microscopes.[1–3] These techniques enable super-resolved imaging of fluorescently labelled nanoscale structures with transverse spatial resolution approaching that of electron microscopy. This is achieved by acquiring a sequence of images in which only a sparse subset of fluorophores emit light during each camera acquisition, such that their point spread functions can be distinguished. Thus, the location of each fluorophore in the image plane can be determined, for example, by centroiding the discrete images of individual emitters, to achieve an enhanced precision that depends on the number of photons detected. The most commonly used SMLM techniques include Photo Activated Localisation Microscopy (PALM),[4] Stochastic Optical Reconstruction Microscopy (STORM)[5] – particularly implementations based on direct photoswitching of conventional fluorophores such as dSTORM[6] and GSDIM[7] – and techniques based on transient binding of

diffusing fluorophores such as Points Accumulation for Imaging in Nanoscale Topography (PAINT)[8] and DNA-PAINT.[9] In general, SMLM techniques are relatively straightforward to implement on existing epifluorescence microscopes;[10–12] we and others have demonstrated low-cost SMLM instruments using multimode diode laser excitation with inexpensive CMOS cameras that can help widen access to super-resolved microscopy.

It is also possible to localise sparse emitters in the axial (Z) direction to realise 3D SMLM by analysing the shape of the image of each emitter where the PSF varies with Z. This is most frequently realised in microscopes that have been modified to exhibit significant astigmatism, for example, through the insertion of a cylindrical lens in the imaging path,[13] as previously used to axially localise quantum dots.[14] Alternatively, the microscope PSF can be explicitly engineered, for example, using wavefront-shaping techniques, to provide an extended axial range of Z localisation. The first implementation of this wavefront-shaping approach was the double helix PSF,[15] and this has subsequently been extended to further engineered PSF.[16] A third approach involves imaging the sparse emitters at two or more different focal planes,[17] and comparing the multiple PSFs corresponding to each detected emitter. These and other 3D SMLM techniques entail modification of the fluorescence microscope with varying degrees of complexity and cost; the simplest approach using cylindrical lens-induced astigmatism provides high axial precision but is limited to a relatively small (~500 nm) range.

To further widen access to SMLM and other microscopy techniques, we are developing a cost-effective open-source microscope platform that utilises a modular microscope stand (*'openFrame'*)[18] to implement our low cost *'easySTORM'* approach[10] to SMLM, and we are working towards a practical open-source cost-effective platform for automated SMLM of samples arrayed in multiwell plates.[10,18] As part of our efforts to minimise cost, while maximising access and flexibility in these instruments, we explore utilising machine learning to determine the Z-localisation of emitting fluorophores with no modification of the fluorescence microscope, that is, directly from the image data acquired for 2D SMLM exploiting only the microscope's intrinsic aberration. Our approach is to obtain 'ground truth' training data by acquiring calibration Z-stacks of images of subresolution fluorescent beads in order to map the axial and lateral variation of the microscope's intrinsic PSF across its field of view. This approach, which we term *easyZloc*, makes no assumptions about the instrument and can, in principle, be applied to any microscope configuration, that is, any combination of microscope optics and camera, as long as a 'calibration' Z-stack of subresolution fluorescent beads is acquired before an imaging session.

We note that Z-localisation using the microscope's intrinsic axial variation of experimental PSF has previously been explored using an analytic model-based fitting approach with 2D SMLM data,[19,20] and that several groups have utilised deep learning for 3D SMLM using astigmatic or engineered PSF.[21–24] Most of this prior work uses convolutional neural networks (CNN) to determine the localisation of each emitter in three dimensions and typically requires significant computational resources, often limiting them to relatively small fields of view or requiring significant computation times that would not be practical for some applications, particularly automated multiwell plate SMLM where we aim to acquire and process SMLM data from hundreds of large (>100 μm) fields of view (FOV).

Prior work based on phase-retrieval methods leverage Zernike polynomials to model the microscope PSF, for example, fitting a model to a calibration image of fluorescent beads[25] and generating simulated noise-free samples to fit a Maximum Likelihood Estimate (MLE) of the localisation's depth. In-situ PSF Retrieval (INSPR)[23] extended this approach by constructing the Zernike model directly from experimental data while iteratively localising the dataset. These methods assume a uniform PSF model across the FOV of the microscope and were demonstrated with microscopes engineered to present PSFs with significant axial variation. An alternative paradigm for 3D localisation methodologies relies on empirical models of the PSF that are not constrained by an underlying Zernike polynomial. For example, the use of a 3D cubic spline and MLE to model a PSF was demonstrated on SMLM data from microscopes with astigmatic and saddle point PSFs.[26] This approach yielded an improved localisation accuracy compared to earlier Gaussian based models and should be applicable with any experimental PSF as no assumptions are made concerning a theoretical model of the PSF. This approach was further refined into a GPU-accelerated implementation of many different PSF modalities.[19] However, these approaches also assume a constant PSF across the FOV and could potentially be compromised in microscopes with a significant lateral variation of the PSF aberrations across the FOV, for example, due to spherical or other aberrations.

Further approaches such as *DeepLoc*,[27] *Deep STORM3D*[28] and *DECODE*[22] have also leveraged deep learning for whole image localisation, for example using cubic spline models of an experimental PSF to generate simulated training data.[22] To date, these methods have modelled laterally invariant PSFs,[27] relatively complex (and expensive to implement) tetrapod PSFs[28] or highly astigmatic/double-helix PSFs.[22] To the best of our knowledge only one deep learning-based method, *FD-DeepLoc*,[21] has accounted for lateral variations in PSF aberrations,

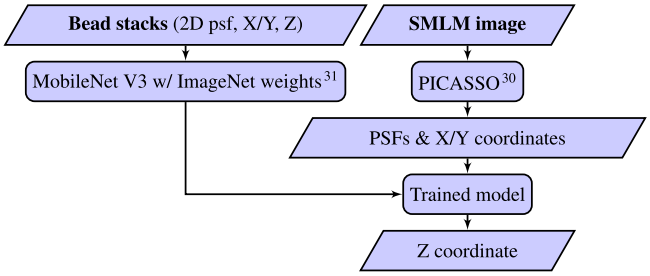

FIGURE 1 Application of *easyZloc* on experimental data.

implementing a vectorial Zernike model fitted to experimental PSF data, but this was demonstrated on optical systems with a strong astigmatism induced by a cylindrical lens or a tetrapod PSF.

Our approach is to realise 3D SMLM with minimal perturbations to the microscope hardware, leveraging existing SMLM workflows and minimal computational requirements. We approach this by using established (2D) SMLM workflows such as *ThunderSTORM*[29] or *PICASSO*[30] for transverse localisation of the emitters and then determining their axial localisation using a rapid CNN-based approach, as illustrated in Figure 1. To the best of our knowledge, this is the first attempt to determine Z-localisation on a 'standard' fluorescence microscope (i.e. with no engineered astigmatism) using deep learning trained with experimental SMLM calibration data acquired on the same instrument. It should be applicable with a wide range of experimental PSF and does not require a theoretical basis for the PSF, for example, from fitting to cubic splines or a Zernike-based PSF model.

We leverage the robustness of MobileNet v3[31] to map the experimental 3D spatially varying PSF, including any optical aberrations, directly from a calibration image of fluorescent subresolution (TetraSpeck) beads (Section 2.2 and Section 2.4). MobileNet v3 was selected as it is the latest version of a computationally efficient high performance convolutional neural network. Lateral single molecule localisation is first undertaken using a separate software package (*PICASSO*)[30] to transform the acquired SMLM data into a 2D localisation table and a set of extracted 2D images of localisations. These are then localised in Z using our trained model. We demonstrate this approach can provide a relatively fast and robust solution to achieve 3D localisation in almost any SMLM microscope, without modification to engineer the PSF and with modest computational requirements. When imaging fluorescently labelled nuclear pore proteins (using *easySTORM* on a standard epifluorescence microscope with HiLo illumination), we were able to resolve structures axially separated by ~50 nm. In a first attempt to benchmark our approach, we also applied *FD-DeepLoc* and *DECODE* to the same SMLM data – albeit noting that these software tools were developed and evaluated using microscopes exhibiting significant astigmatism from engineered PSF – and found that our MobileNet v3-based *easyZloc* approach provided at least comparable reconstructions of a (minimally astigmatic) nuclear pore SMLM data set, while requiring significantly less computational time.

2 | METHODS

2.1 | Single molecule localisation microscope

A commercial epifluorescence microscope frame (Axiovert 200, Zeiss) was used to acquire the *easySTORM* SMLM data presented in this paper. As outlined in Ref. (10), this instrument utilises a vibrating multimode optical fibre to deliver excitation light from multimode laser diodes within a commercial multiline laserbank (Multiline laserbank, Cairn Research Ltd). Using Köhler illumination, this provides uniform illumination over a large (>120 μm diameter) FOV for *easySTORM*.[10] Sample excitation could be switched between HiLo[32] and epi-illumination using an excitation coupling unit positioned in the back port of the commercial frame (OptoTIRF; Cairn Research Ltd). A custom-built hardware-based optical autofocus module[18] was used to maintain focus throughout the SMLM acquisitions. For the data presented in this paper, an sCMOS camera (Photometrics Prime 95B) was used to acquire both the training calibration data and the fluorescent SMLM data. This microscope was fitted with a piezoelectric Z-stage (NanoScan NZ100, Prior) and a motorised XY stage (MS-2000, Applied Scientific Instruments) to control axial and lateral movement. The hardware was all controlled using *μManager* 2.0.[33]

2.2 | Calibration sample preparation and imaging

Bead samples to acquire the calibration training data were prepared as follows: diluted 1:10^7 per millilitre fluorescent polystyrene beads with a diameter of 20 nm, excitation/emission: 625/645 nm (F8782-Invitrogen FluoSpheres^{TM} Carboxylate-Modified Microspheres) were attached to a glass 8-well chamber slide 1.5H (80827 μ-slide8 Ibidi, GmbH) functionalised with poly-L-lysine 0.1% (P8920 Sigma Aldrich) and imaged in 0.4 mL/well oxyrase-based STORM buffer (phosphate buffer saline, 50 mM mercaptamine, 10 mM D-lactate and 10 μL of a debris free SAE0010-5ml EC-Oxyrase Sigma).

These beads were imaged using a 100× 1.46 NA objective (α Plan-Apochromat 100×/1.46 NA Oil, Zeiss) and the Photometrics Prime 95B camera captured a FOV of 120 μm × 120 μm. Image Z-stacks were acquired over a range

of −3 μm to +3 μm in 10 nm steps for 10 different FOVs of fluorescent beads such that in total approximately 1000 beads were imaged. The motorised XY stage was used to move laterally between different FOVs.

The training calibration data was acquired using a range of camera exposure times for the 10 different stacks, varying between either 300 ms, 500 ms or 1000 ms. The power at the sample was varied between settings of either 450 μW, 1.3 mW, and 2 mW. These low excitation powers were chosen as to minimise bleaching of the fluorescent beads during the acquisitions. The sample was excited at a wavelength of 635 nm.

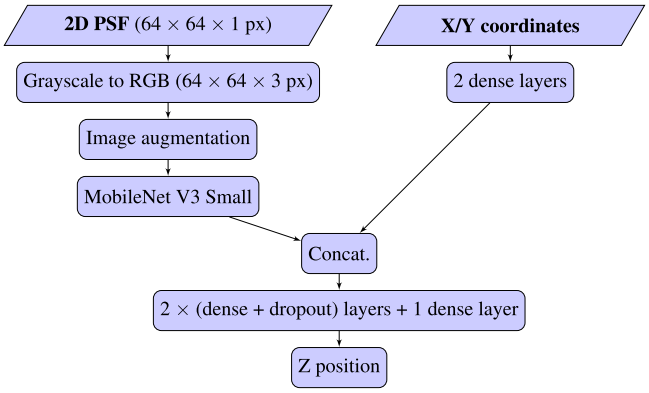

FIGURE 2 Model architecture including MobileNet V3 (small).[31]

2.3 | Calibration data preprocessing

Starting from a full-FOV bead image stack, the most in-focus image frame was identified using the maximum pixel intensity as a function of axial displacement (Z). Individual beads were localised in this image using PICASSO,[30] and 3D image stacks for each bead were retrieved across the full FOV Z-stack. Similar sets of bead ROI Z-stacks from all full FOV image Z-stacks recorded were concatenated to produce a single training dataset.

Beads were excluded from the set of training data for any of the following issues: co-localisation with other beads within the cropped 15×15 pixel frame; poor signal-to-noise ratio (SNR, calculated as the ratio of maximum pixel intensity to mean pixel intensity for the bead's image stack – a minimum value of 2 was used); or excessive noise in the Z or X/Y axes (measured by fitting a 1D or 2D skew Gaussian to the bead ROI Z-stacks in Z or X/Y respectively). The minimum SNR threshold and maximum noise thresholds were chosen through repeated training and evaluation of the model on test beads. Further tuning of the parameters was then informed by the reconstruction quality of nuclear pore samples.

The assignment of modelled Z-positions for each bead in the training data was key to removing bias from the final model, as two closely neighbouring beads with a similar PSF but different Z offsets could cause the model to learn an average of the two Z positions. The relative offset (δZ) of each bead in a FOV was calculated by finding the pixel-wise shift in Z which minimises a mean squared difference of pixel intensity values between any bead and a reference bead (chosen as the bead closest to the centre of the FOV). The overall $Z = 0$ defocus was selected using the frame in the reference bead stack with the largest image sharpness metric, and the relative δZ position of all other frames was calculated using the Z-step size of the acquired image stack (10 nm in this dataset).

The training dataset was formed of the 2D image, X/Y position, and Z position of each bead. Datapoints outside of the desired Z-range (±1 μm in our study) were removed from the dataset; this range can be modified to approximately match the microscope's axial resolution (e.g. axial width of PSF). Beads were divided into training/validation/test sets and the training images were augmented to triple the number of training datapoints by adding Gaussian noise ($\mu = 0$, $\sigma = 0.005$), randomly adjusting the brightness of the image (using a Keras RandomBrightness layer with a factor of 0.075[34]) and adding Poisson noise ($\lambda = 226$). These parameters were found using the hyperparameter discussed in Section 2.4. Images of PSFs were resized to 64 × 64 pixels to allow the use of pretrained weights (see Section 2.4). Pixel intensities were rescaled to a range of [0, 1] by dividing by the maximum possible pixel intensity. X/Y coordinates were rescaled to a range of [−1, 1].

2.4 | Model architecture and training

The CNN model's architecture (see Figure 2) was built around a MobileNet V3, which is a recent version of an efficient and fast CNN.[31] Pretrained ImageNet-21k[35] weights were used to accelerate the training process. Future models could be trained from the weights shared in this paper to further reduce training time.

Images of localisations were passed into the MobileNet V3 for feature extraction. X/Y coordinates were passed through two dense layers. Both input branches of the model were visible in Figure 2. The concatenated embedding of the image and coordinate branches of the network was then passed through a series of dropout and dense layers to output the predicted Z position.

The model was compiled with an AdamW optimiser[36] and a mean squared error loss. Hyperparameter optimisation from a Bayesian-guided search (as implemented in Ref. 37) yielded specific augmentation parameters, dense

layer sizes, learning rates and batch sizes. Trained models and associated training parameters are included in the Supporting Information. Callbacks were used to regularise the training; learning rate reduction was applied if the training loss stagnates, and an early stopping mechanism halts the training if the validation loss ceases to decrease.

The model was trained on an Ubuntu desktop equipped with an Intel Core i7-10700K CPU, a single NVIDIA RTX 3090 24GB graphic card and 128 GB of RAM, although this far exceeds the computational requirements of the small network size (3.4M parameters). Training our final model took approximately 1.5 h, which was significantly faster than for comparable methodologies trained from the same set of bead stacks (6 h for *DECODE* on the machine listed above, 14 h for *FD-DeepLoc* on a Windows machine with an AMD Ryzen Threadripper 3960X 24-Core 3.79 GHz Processor, 64.0 GB of RAM and a single NVIDIA RTX 3080 24GB graphic card). Note that while *DECODE* was trained on the same machine as *easyZloc*, *FD-DeepLoc* includes some Windows dependencies and was therefore trained on the aforementioned Windows machine of comparable hardware specifications. We consider that the difference between the computational hardware would not significantly impact the overall difference in training time.

2.5 | Sample preparation and imaging

The NUP96p proteins of the nuclear pore structure have become an elegant de facto biological standard sample for demonstrating the capabilities of super-resolved microscopes.[38] Figure S1 shows a schematic of the arrangement of these nuclear pore NUP96p proteins in two parallel octagonal rings of 8 clusters of NUP96p proteins where the rings are separated by 46 nm. More information concerning the nuclear pore structure, including electron microscopy images can be found in Ref. (38). To explore performance of easyZloc in comparison with DECODE and FD-DeepLoc, SNAPtag-labelled NUP96 nuclear pore proteins in the U-2OS-CRISPR-NUP96-SNAP cell line clone[38] (CLS Cell lines Services GmbH) that we have labelled with BG-iFluor647[18] were imaged using *easySTORM* implemented on the Axiovert 200 microscope described above. The preparation and imaging followed the protocols shared previously.[10] Briefly, cells were fixated for 1 min with paraformaldehyde 2.4% in TRB buffer (20 μM HEPES, 1 mM EGTA, 10 mM potassium acetate, 10 mM sucrose), washed twice for 5 min each in paraformaldehyde-free-TRB-buffer, PFH quenched 5 min in freshly made 100 mM ammonium chloride in phosphate buffer saline (PBS), washed 3× in PBS, blocked for 10 min with Blocking buffer (3% bovine serum albumin in PBS, 1 μM dithiothreitol), stained overnight with 1 μM SNAPtag in blocking buffer, and washed 3 times prior to imaging in oxyrase buffer. The sample was incubated for 5 min with 1:1000 TetraSpeck beads in PBS followed by 3 washes with PBS to removed unbound beads to mark fiduciary markers. These SNAPtag-labelled nuclear pore samples were imaged on the Axiovert 200 Zeiss microscope and excited at 635 nm, using HiLo illumination[32] for the SMLM acquisition. A power density of 1 kW/cm^2 at the sample plane was used during the acquisition. 20,000 frames were acquired with a camera exposure of 30 ms. Figure S1 presents a range of SMLM reconstructions of nuclear pore structures.

Actin filaments stained with Phalloidin-Atto647N (Sigma Aldrich catalogue 65906) in U2OS cells (Osteosarcome cell line) were also imaged using the same easySTORM microscope. Sample preparation followed the following protocols: osteosarcome U2OS cells were cultured for 48 h in a glass 8-well chamber slide 1.5H (80827 μ-slide8 Ibidi, GmbH). The cells were fixated in 4% paraformaldehyde in PBS (phosphate buffer saline) for 10 min, permeabilised by incubation for 10 min in triton X-100 at 0.05% in PBS (permeabilisation solution) and blocked by incubation for 10 min in PBS supplemented with 0.05% triton X-100 and 1% albumin (blocking solution). The actin filaments were stained with Phalloidin-Atto647N (Sigma Aldrich catalogue 65906) using a 1:1000 dilution in the blocking solution for 1 h at room temperature. The cells were then washed three times for 5 min in permeabilisation solution. To provide fiduciary markers for lateral drift correction in postprocessing, the cells were incubated for 5 min with 1:1000 TetraSpeck beads in PBS, followed by 3 washes with PBS to removed unbound beads. This actin-labelled sample was imaged using epi-illumination on the Axiovert 200 Zeiss microscope, using the same objective lens and camera as for the calibration bead stacks (see Section 2.3). The widefield image was recorded using excitation at 635 nm with a power density at the sample of 0.031 kW/cm^2. For easySTORM, the cells were imaged in 0.4 mL of a STORM oxyrase/mercaptamine buffer. Initially they were illuminated by a combination of two excitation wavelengths for 10 s (635 nm at 1.08 kW/cm^2 and 462 nm at 1.82 kW/cm^2) to excite fluorophores into the dark state and induce blinking. The easySTORM dataset was then acquired using a continuous excitation power density at the sample of 0.0837 kW/cm^2. 50,000 frames were acquired with 30 ms exposure time.

2.6 | Reconstruction of 3D SMLM nuclear pore data

PICASSO[30] was used for lateral single molecule localisation (X-Y) and lateral drift correction was applied using PICASSO via fiduciary markers (0.1 μM T7279 TetraSpeck

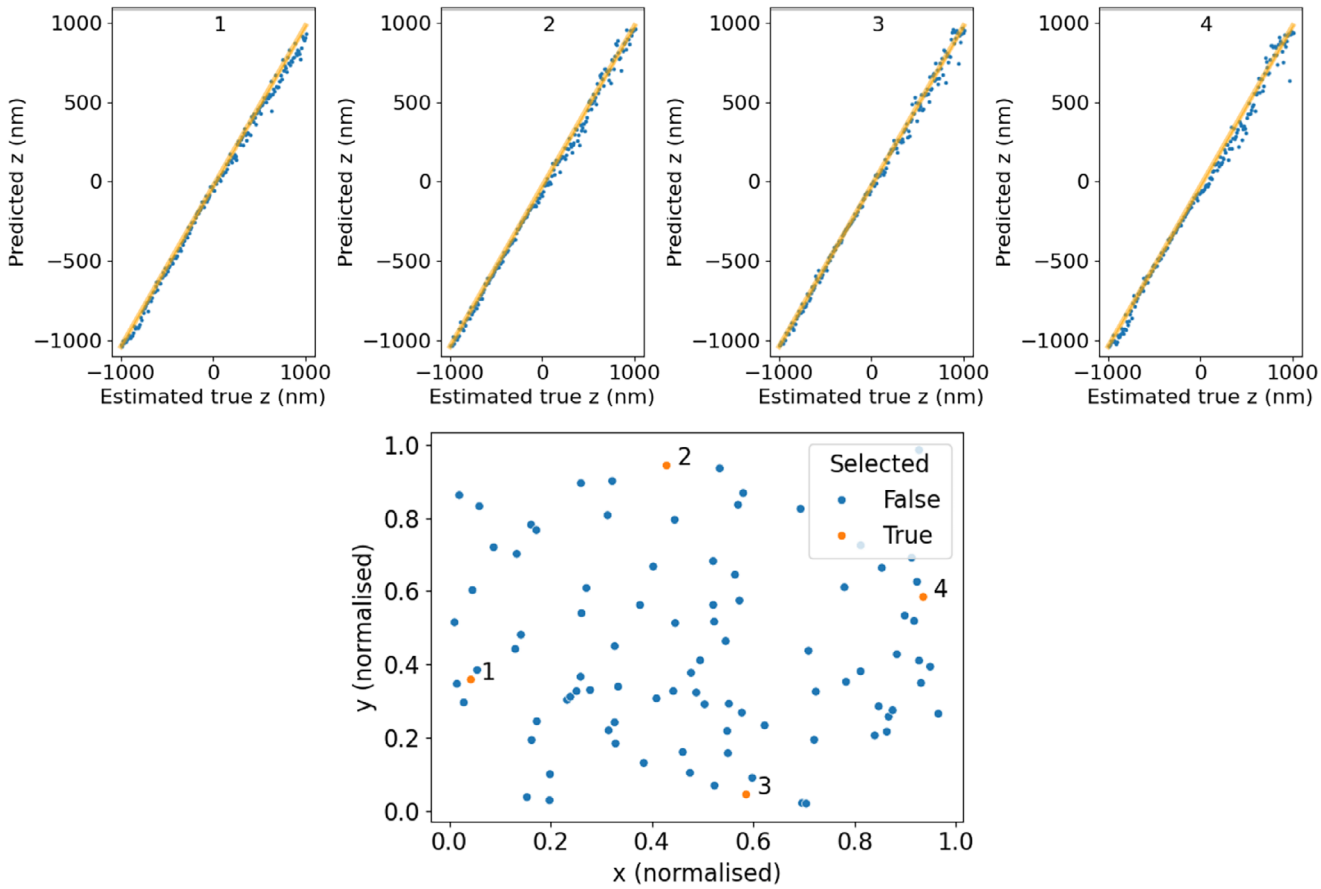

FIGURE 3 Performance of *easyZloc* on test beads. Top: each plot (1–4) was produced from a bead within the Z stack of the training data. The position in Z of each slice in the Z stack was estimated from the peak illumination of the stack and the Z-step of the microscope system (10 nm). The model estimates positions for each slice of the Z stack, and we see a linear correlation (orange line) with the stage movement of the microscope. The model achieves a root mean squared error of 40.93 nm, 39.46 nm, 36.36 nm and 55.35 nm respectively on these relatively noise-free beads. Bottom: orange dots (1–4) show the position of test beads in the FOV, demonstrating the model has learned to map laterally varying aberrations.

Microspheres, Life Technologies). Pixel intensities were rescaled to a range of 0 to 1 by dividing by the maximum pixel intensity (65535, the maximum unsigned 16-bit integer), and X/Y coordinates were rescaled to a range of –1 to 1 in accordance with the FOV of the training data. The preprocessed data was then localised in Z using our trained CNN model to produce a set of 3D localisations. Drift correction in Z was implemented by calculating the mean Z localisation of 500 sequential image frames and fitting a smoothed cubic spline to the mean Z position for all localisations in the frame. The Z positions for localisations in each frame were then corrected by subtracting any offset relative to the mean Z localisation over time.

Nuclear pores were picked in X/Y using PICASSO's Render GUI and analysed individually. Each group of localisations was modelled using a kernel density estimation (KDE) with a bandwidth of 15 nm, selected to image structures on a scale of 50 nm (as expected for the Nup96 protein rings) without overfitting to noise. KDEs smooth the data to minimise the impact of noise and highlight the underlying spatial structure of data. A KDE was applied by evaluating a Gaussian kernel over the range of Z values of localisations in each nuclear pore. In order to quantify the success at resolving nuclear pore structure, the reconstructed data was automatically scanned to identify reconstructed features (i.e., peaks observed with an applied Gaussian KDE, bandwidth = 15 nm, min peak prominence of 0.0001 localisations/nm^2) that could correspond to a nuclear pore's two layers of labelled Nup96 proteins. A nuclear pore was considered to be 'successfully' imaged if the two more prominent peaks are axially separated by 40 to 60 nm, following the criteria in Ref. (21). The visual contrast of the reconstructed image of the nuclear pore layers was enhanced by applying a minimum threshold for the spatial density. This threshold was calculated as the mean of the mean peak height and the local minima.

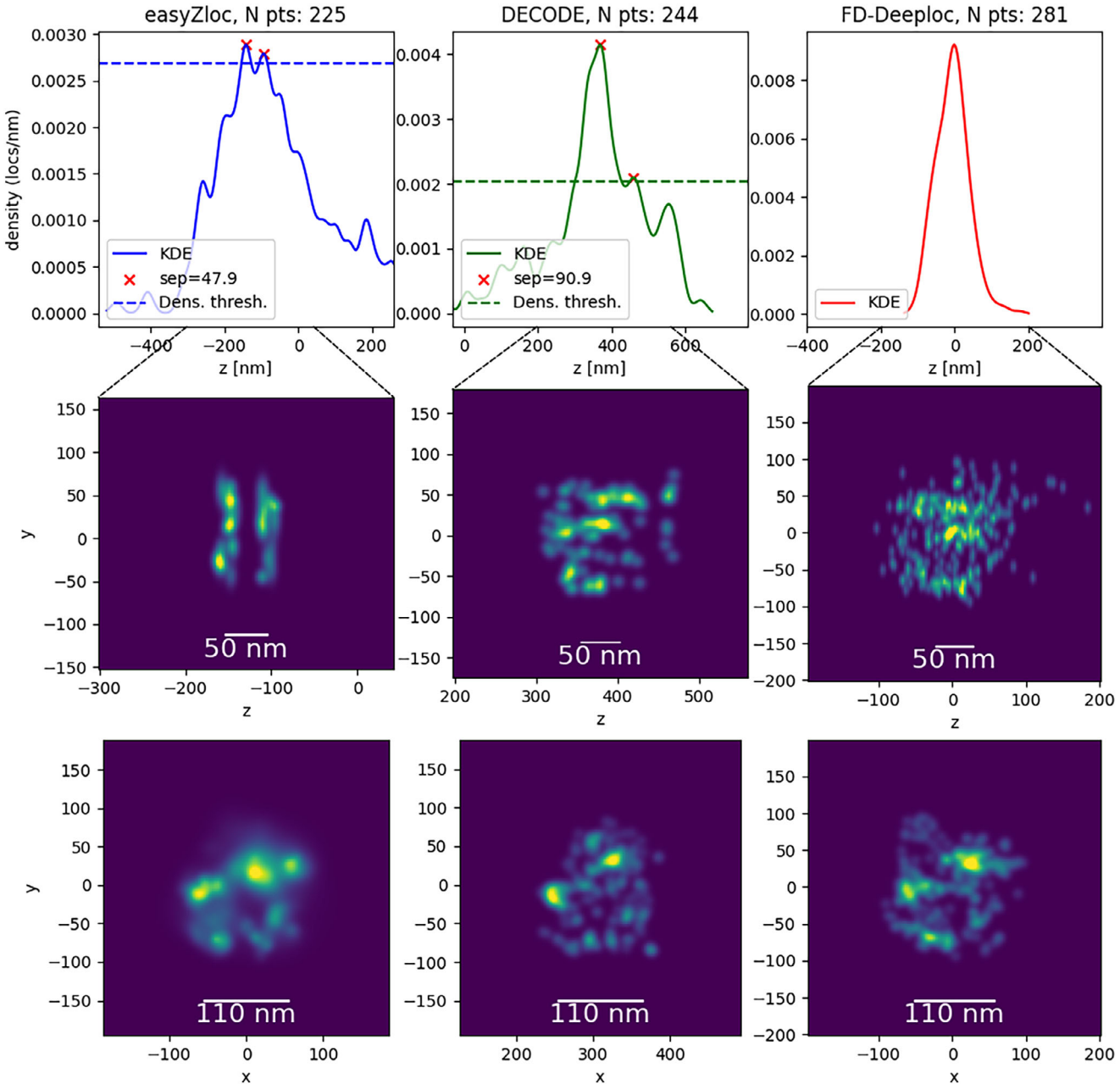

FIGURE 4 Reconstructions of Nup96 proteins in a nuclear pore imaged with a 100× 1.46NA objective lens and analysed using *easyZloc*, *DECODE* and *FD-DeepLoc*. Rendered X/Z images are thresholded by a minimum spatial density to increase the visual contrast of the image, as indicated by the dotted line of each histogram. The threshold density is set automatically following the procedure detailed in Section 2.6. The X/Y view was not thresholded has the contrast was already sufficient to identify the NPC structure. Additional comparative examples of nuclear pore reconstructions can be found in Figure S2, where three 'successful' reconstructions of nuclear pores for each methodology are displayed with the corresponding reconstruction of the same data using the other methodologies. Also provided in the Supplementary Information is a video similarly presenting all the nuclear pores that were 'successfully' reconstructed by at least one methodology.

3 | RESULTS

easyZloc was trained on a bead dataset obtained using the methodology in Section 2.2. The model was evaluated on bead data withheld from the calibration data. Figure 3 illustrates how the *easyZloc* model correctly learned to map field dependent aberrations to the Z position of the bead. The resulting Z localisation accuracy is comparable to those reported by alternative methodologies on the same calibration dataset (root mean squared error of Z localisation, easyZloc: 120 nm, *DECODE*: 148 nm, *FD-DeepLoc*: 156 nm).

The values for *DECODE* and *FD-DeepLoc* were automatically generated by the respective methodology's model training software. We note, however, the different training and test data for each methodology. *easyZloc* directly utilised empirical image Z-stack image data of 20 nm Polystyrene bead (F8782-Invitrogen FluoSpheres) fluorescent beads, whereas *DECODE* and *FD-DeepLoc* were applied using a physical model of the PSF constructed from these 20 nm bead image data that was then used to simulate images of the PSF which were then used for training.

We note that while *DECODE* and *FD-DeepLoc* can function with any modellable PSF modality, they were designed for and demonstrated on microscopes with additional induced astigmatism. The optical system used in our implementation is a more challenging case for Z localisation due to less pronounced changes in the PSF over Z; *DECODE* and *FD-DeepLoc* may therefore provide a lower axial resolution than they would if applied with an engineered PSF.

The trained models were then applied to *easySTORM* SMLM data of Nup96 nuclear pore proteins in the U-2OS-CRISPR-NUP96-SNAP cell line clone[38] (CLS Cell lines Services GmbH) that we have labelled with BG-iFluor647.[18] While *DECODE* and *FD-DeepLoc* directly yield 3D localisation tables from the SMLM data, with our approach the SMLM data was first analysed using PICASSO[30] to obtain the 2D localisation table, and then Z localisation was undertaken using *easyZloc*.

Nuclear pore complexes were extracted from reconstructions of the dataset by each method to explore the relative success rates with which Nup96 layers of the nuclear pore could be distinguished (we used this as an arbitrary definition of a 'successful' reconstruction, as discussed in Section 2.6). Exemplar reconstructed nuclear pore images are shown in Figure 4, and additional examples can be seen in Figure S2.

Figure 5 plots the number of 'successful' 3D nuclear pore images reconstructed from the same *easySTORM* SMLM data set using *easyZloc*, *DECODE* and *FD-DeepLoc*, plotted as a function of the KDE smoothing parameter (see Section 2.6). We note that (MobileNet v3-based) *easyZloc* performs at least as well for KDE bandwidth values above 7 and functions up to a KDE bandwidth >20 – at a significantly lower computational cost (see Section 2.4).

The choice of a KDE bandwidth affects the number of nuclear pores reconstructed. An excessively large bandwidth would blur the data, causing the nuclear pore's layers to be indistinguishable, while a small bandwidth can lead to overfitting of noise. Standard bandwidth tuning methods, as discussed in Ref. (39), resulted in unresolvable reconstructions due to oversmoothing of the data. We note that KDEs have been utilised in previous work[21]

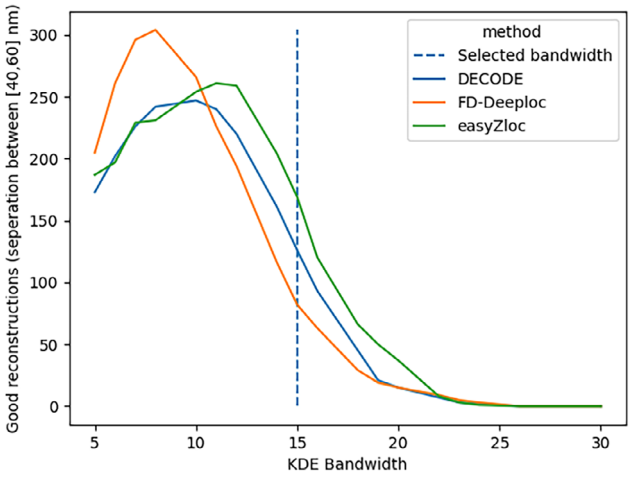

FIGURE 5 Comparison of the number of 'successfully' resolved nuclear pores for each Z localisation methodology applied to the same dataset. The criteria to assess nuclear pore reconstructions are detailed in Section 2.6.

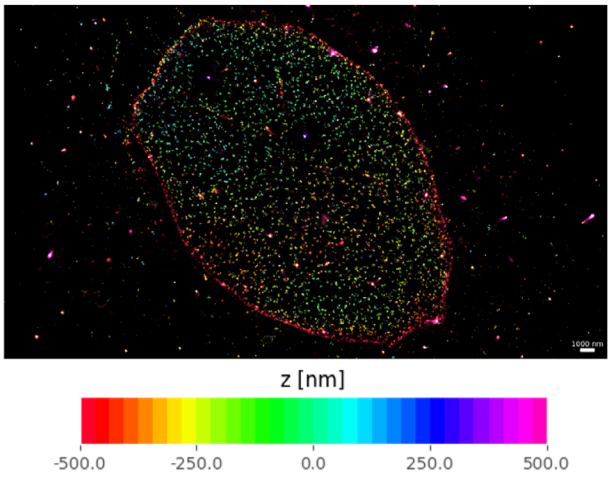

FIGURE 6 3D reconstruction of a cell nucleus imaged with a 1.46 NA objective lens (α Plan-Apochromat 100×/1.46 NA Oil, Zeiss), coloured by depth, as produced by *easyZloc*. Scale bar length is 1 μm.

although the methodology for selecting the value of the KDE bandwidth was not discussed.

The depth colour-coded 3D SMLM reconstruction in Figure 6 illustrates the depth range of *easyZloc* applied to the same *easySTORM* SMLM data (following 2D localisation using PICASSO) of Nup96 nuclear pore proteins within the cellular nuclear envelopes, in a large (>100 μm) FOV. The data processing for this entire FOV required less than 10 min of compute on a Linux computer (hardware details are listed in Section 2.4). The equivalent reconstructions using *DECODE* required significantly more computational time (2.74 h) on the same machine as *easyZloc*,

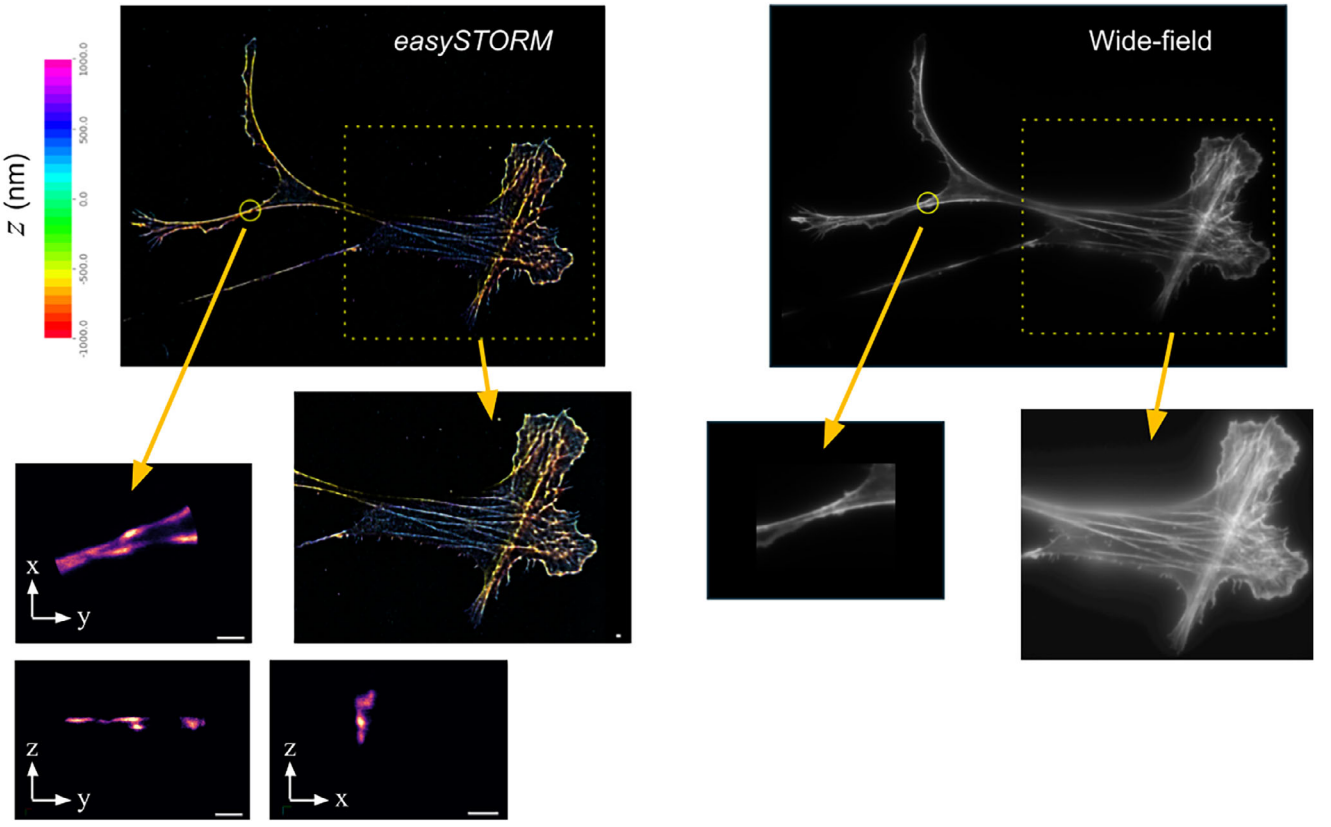

FIGURE 7 3D SMLM reconstruction of actin labelled with Phalloidin-Aexa 647N in osteosarcome U2OS cells imaged with a 1.46 NA objective lens (α Plan-Apochromat 100×/1.46 NA Oil, Zeiss) objective lens using easyZloc following 2D localisation using PICASSO. Scale bar is 1 μm.

while *FD-DeepLoc* required 0.3 h on a similar machine using Windows (see hardware details in Section 2.4).

Figure 7 shows a 3D SMLM reconstruction of easySTORM image data (following 2D localisation using PICASSO) of actin filaments in osteosarcome U2OS cells. Here easyZloc provides depth information over 2 μm, and we can see evidence of overlapping tubulin filaments in the isometric views of the small FOV.

4 | DISCUSSION

We have presented a fast and lightweight machine learning-based methodology for axial localisation of SMLM data, *easyZloc*, which can be used in standard epifluorescence microscopes without specifically engineering an axially varying PSF. Requiring only the additional acquisition of a calibration Z-stack of images of subresolution fluorescent beads, *easyZloc* can be implemented with existing 2D SMLM workflows using established 2D localisation software. Since *easyZloc* performs only a subtask within the overall SMLM localisation task – sequentially identifying the defocus of individual localisations from an extracted FOV centred on that localised emitter – it is less computationally demanding than 3D localisation software tools that identify multiple localisations in their inputs and determine the full 3D localisation of the SMLM dataset. Furthermore, the use of empirical training data circumvents the need for mathematical modelling of the PSF, allowing for the deep learning model to readily account for lateral variation in the microscope. Finally, the use of the latest lightweight deep learning architecture (MobileNet V3 small) means that the methodology requires comparatively modest computational resources and a short execution time. We thus believe this approach has broad potential to widen access to 3D SMLM and may be useful when scaling to higher throughput in automated SMLM instruments.

We note that the choice of the core CNN architecture, MobileNet V3 (small version), was made after considerable hyperparameter exploration using the nuclear pore Nup96 cell line as a test sample. Vision Transformers[40] were originally used but suffered from overfitting issues, while not improving the frequency or clarity of nuclear pore reconstructions. ResNets[41] and VGG[42] architectures presented similar issues. We surmise that these more powerful methodologies may surpass *easyZloc* in speed and precision if given a larger quantity or variety of training

data, and/or when implemented on more powerful deep learning architectures.

ACKNOWLEDGEMENTS

We acknowledge funding from the Cancer Research UK (A29368) and from the Chan Zuckerberg Initiative DAF, an advised fund of the Silicon Valley Community Foundation (grant 2021-234618, 2023-321240). Miguel Boland acknowledges a PhD studentship supported by the Wellcome Trust grant 203799/Z/16/Z.

DATA AVAILABILITY STATEMENT

Data and trained models are available upon request. The codebase supporting this work is available at https://github.com/mb1069/easyZloc.

ORCID

Miguel A. Boland https://orcid.org/0000-0002-2703-8936

REFERENCES

1. Patterson, G., Davidson, M., Manley, S., & Lippincott-Schwartz, J. (2010). Superresolution imaging using single-molecule localization. *Annual Review of Physical Chemistry*, *61*, 345–367.
2. Klein, T., Proppert, S., & Sauer, M. (2014). Eight years of single-molecule localization microscopy. *Histochemistry and Cell Biology*, *141*(6), 561–575.
3. Lelek, M., Gyparaki, M. T., Beliu, G., Schueder, F., Griffié, J., Manley, S., Jungmann, R., Sauer, M., Lakadamyali, M., & Zimmer, C. (2021). Single-molecule localization microscopy. *Nature Reviews Methods Primers*, *1*(1), 39.
4. Betzig, E., Patterson, G. H., Sougrat, R., Lindwasser, O. W., Olenych, S., Bonifacino, J. S., Davidson, M. W., Lippincott-Schwartz, J., & Hess, H. F. (2006). Imaging intracellular fluorescent proteins at nanometer resolution. *Science*, *313*(5793), 1642–1645.
5. Rust, M. J., Bates, M., & Zhuang, X. (2006). Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (storm). *Nature Methods*, *3*(10), 793–796.
6. Heilemann, M., van de Linde, S., Schüttpelz, M., Kasper, R., Seefeldt, B., Mukherjee, A., Tinnefeld, P., & Sauer, M. (2008). Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes. *Angewandte Chemie International Edition*, *47*(33), 6172–6176.
7. Fölling, J., Bossi, M., Bock, H., Medda, R., Wurm, C. A., Hein, B., Jakobs, S., Eggeling, C., & Hell, S. W. (2008). Fluorescence nanoscopy by ground-state depletion and single-molecule return. *Nature Methods*, *5*(11), 943–945.
8. Sharonov, A., & Hochstrasser, R. M. (2006). Wide-field subdiffraction imaging by accumulated binding of diffusing probes. *Proceedings of the National Academy of Sciences*, *103*(50), 18911–18916.
9. Jungmann, R., Steinhauer, C., Scheible, M., Kuzyk, A., Tinnefeld, P., & Simmel, F. C. (2010). Single-molecule kinetics and super-resolution microscopy by fluorescence imaging of transient binding on dna origami. *Nano Letters*, *10*(11), 4756–4761.
10. Kwakwa, K., Savell, A., Davies, T., Munro, I., Parrinello, S., Purbhoo, M. A., Dunsby, C., Neil, M. A., & French, P. M. (2016). easySTORM: a robust, lower-cost approach to localisation and TIRF microscopy. *Journal of Biophotonics*, *9*(9), 948–957.
11. Ma, H., Fu, R., Xu, J., & Liu, Y. (2017). A simple and cost-effective setup for super-resolution localization microscopy. *Scientific Reports*, *7*(1), 1542.
12. Diekmann, R., Till, K., Müller, M., Simonis, M., Schüttpelz, M., & Huser, T. (2017). Characterization of an industry-grade cmos camera well suited for single molecule localization microscopy – high performance super-resolution at low cost. *Scientific Reports*, *7*(1), 14425.
13. Huang, B., Wang, W., Bates, M., & Zhuang, X. (2008). Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. *Science*, *319*(5864), 810–813.
14. Kao, H., & Verkman, A. (1994). Tracking of single fluorescent particles in three dimensions: Use of cylindrical optics to encode particle position. *Biophysical Journal*, *67*(3), 1291–1300.
15. Pavani, S. R. P., Thompson, M. A., Biteen, J. S., Lord, S. J., Liu, N., Twieg, R. J., Piestun, R., & Moerner, W. E. (2009). Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function. *Proceedings of the National Academy of Sciences*, *106*(9), 2995–2999.
16. Shechtman, Y., Sahl, S. J., Backer, A. S., & Moerner, W. E. (2014). Optimal point spread function design for 3D imaging. *Physical Review Letters*, *113*, 133902.
17. Juette, M. F., Gould, T. J., Lessard, M. D., Mlodzianoski, M. J., Nagpure, B. S., Bennett, B. T., Hess, S. T., & Bewersdorf, J. (2008). Three-dimensional sub-100 nm resolution fluorescence microscopy of thick samples. *Nature Methods*, *5*(6), 527–529.
18. Lightley, J., Kumar, S., Lim, M. Q., Garcia, E., Görlitz, F., Alexandrov, Y., Parrado, T., Hollick, C., Steele, E., Roömann, K., Graham, J., Broichhagen, J., McNeish, I. A., Roufosse, C. A., Neil, M. A. A., Dunsby, C., & French, P. M. W. (2023). openFrame: A modular, sustainable, open microscopy platform with single-shot, dual-axis optical autofocus module providing high precision and long range of operation. *Journal of Microscopy*, *292*(2), 64–77.
19. Li, Y., Mund, M., Hoess, P., Deschamps, J., Matti, U., Nijmeijer, B., Sabinina, V. J., Ellenberg, J., Schoen, I., & Ries, J. (2018). Real-time 3D single-molecule localization using experimental point spread functions. *Nature Methods*, *15*(5), 367–369.
20. Li, Y., Wu, Y.-L., Hoess, P., Mund, M., & Ries, J. (2019). Depth-dependent PSF calibration and aberration correction for 3D single-molecule localization. *Biomedical Optics Express*, *10*(6), 2708–2718.
21. Fu, S., Shi, W., Luo, T., He, Y., Zhou, L., Yang, J., Yang, Z., Liu, J., Liu, X., Guo, Z., Yang, C., Liu, C., Huang, Z.-l., Ries, J., Zhang, M., Xi, P., Jin, D., & Li, Y. (2023). Field-dependent deep learning enables high-throughput whole-cell 3D super-resolution imaging. *Nature Methods*, *20*(3), 459–468.
22. Speiser, A., Müller, L.-R., Hoess, P., Matti, U., Obara, C. J., Legant, W. R., Kreshuk, A., Macke, J. H., Ries, J., & Turaga, S. C. (2021). Deep learning enables fast and dense single-molecule localization with high accuracy. *Nature Methods*, *18*(9), 1082–1090.

23. Xu, F., Ma, D., MacPherson, K. P., Liu, S., Bu, Y., Wang, Y., Tang, Y., Bi, C., Kwok, T., Chubykin, A. A., Yin, P., Calve, S., Landreth, G. E., & Huang, F. (2020). Three-dimensional nanoscopy of whole cells and tissues with in situ point spread function retrieval. *Nature Methods*, *17*(5), 531–540.
24. Zelger, P., Kaser, K., Rossboth, B., Velas, L., Schütz, G. J., & Jesacher, A. (2018). Three-dimensional localization microscopy using deep learning. *Optics Express*, *26*(25), 33166–33179.
25. Liu, S., Kromann, E. B., Krueger, W. D., Bewersdorf, J., & Lidke, K. A. (2013). Three dimensional single molecule localization using a phase retrieved pupil function. *Optics Express*, *21*, 29462–29487.
26. Babcock, H. P., & Zhuang, X. (2017). Analyzing single molecule localization microscopy data using cubic splines. *Scientific Reports*, *7*(1), 552.
27. Boyd, N., Jonas, E., Babcock, H., & Recht, B. (2018). Deeploco: Fast 3d localization microscopy using neural networks. *bioRxiv*.
28. Nehme, E., Freedman, D., Gordon, R., Ferdman, B., Weiss, L. E., Alalouf, O., Naor, T., Orange, R., Michaeli, T., & Shechtman, Y. (2020). DeepSTORM3D: Dense 3D localization microscopy and PSF design by deep learning. *Nature Methods*, *17*(7), 734–740.
29. Ovesný, M., Křížek, P., Borkovec, J., Švindrych, Z., & Hagen, G. M. (2014). ThunderSTORM: A comprehensive ImageJ plug-in for PALM and STORM data analysis and super-resolution imaging. *Bioinformatics*, *30*(16), 2389–2390.
30. Schnitzbauer, J., Strauss, M. T., Schlichthaerle, T., Schueder, F., & Jungmann, R. (2017). Super-resolution microscopy with DNA-PAINT. *Nature Protocols*, *12*, 1198–1228.
31. Howard, A., Sandler, M., Chen, B., Wang, W., Chen, L., Tan, M., Chu, G., Vasudevan, V., Zhu, Y., Pang, R., Adam, H., & Le, Q. (2019). Searching for mobilenetv3. In 2019 IEEE/CVF International Conference on Computer Vision (ICCV) (pp. 1314–1324). Los Alamitos, CA, USA: IEEE Computer Society.
32. Tokunaga, M., Imamoto, N., & Sakata-Sogawa, K. (2008). Highly inclined thin illumination enables clear single-molecule imaging in cells. *Nature Methods*, *5*(2), 159–161.
33. Edelstein, A., Amodaj, N., Hoover, K., Vale, R., & Stuurman, N. (2010). Computer control of microscopes using μ manager. *Current Protocols in Molecular Biology*, *92*(1), 14.20.1–14.20.17.
34. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84–90.
35. Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. In 2009 IEEE Conference on Computer Vision and Pattern Recognition (pp. 248–255). IEEE.
36. Loshchilov, I., & Hutter, F. (2019). Decoupled weight decay regularization. In *7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019*. ICLR, OpenReview.net.
37. Biewald, L. (2020). Experiment tracking with weights and biases. https://www.wandb.com/. Software available from wandb.com.
38. Thevathasan, J. V., Kahnwald, M., Cieśliński, K., Hoess, P., Peneti, S. K., Reitberger, M., Heid, D., Kasuba, K. C., Hoerner, S. J., Li, Y., Wu, Y.-L., Mund, M., Matti, U., Pereira, P. M., Henriques, R., Nijmeijer, B., Kueblbeck, M., Sabinina, V. J., Ellenberg, J., & Ries, J. (2019). Nuclear pores as versatile reference standards for quantitative superresolution microscopy. *Nature Methods*, *16*(10), 1045–1053.
39. Silverman, B. W. (1986). *Density estimation for statistics and data analysis*. Chapman & Hall, London.
40. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2021). An image is worth 16x16 words: Transformers for image recognition at scale. ICLR.
41. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 770–778). IEEE.
42. Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. In Y. Bengio, & Y. LeCun (Eds.), *3rd International Conference on Learning Represent ations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings* (pp. 1–14). ICLR.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Boland, M. A., Lightley, J. P. E., Garcia, E., Kumar, S., Dunsby, C., Flaxman, S., Neil, M. A. A., French, P. M. W., & Cohen, E. A. K. (2025). Model-free machine learning-based 3D single molecule localisation microscopy. *Journal of Microscopy*, 1–11. https://doi.org/10.1111/jmi.13420