

Three-dimensional virtual refocusing of fluorescence microscopy images using deep learning

Yichen Wu^{1,2,3,8}, Yair Rivenson^{1,2,3,8}, Hongda Wang^{1,2,3}, Yilin Luo^{1,2,3}, Eyal Ben-David⁴, Laurent A. Bentolila^{1,3,5}, Christian Pritz⁶ and Aydogan Ozcan^{1,2,3,7*}

We demonstrate that a deep neural network can be trained to virtually refocus a two-dimensional fluorescence image onto user-defined three-dimensional (3D) surfaces within the sample. Using this method, termed Deep-Z, we imaged the neuronal activity of a *Caenorhabditis elegans* worm in 3D using a time sequence of fluorescence images acquired at a single focal plane, digitally increasing the depth-of-field by 20-fold without any axial scanning, additional hardware or a trade-off of imaging resolution and speed. Furthermore, we demonstrate that this approach can correct for sample drift, tilt and other aberrations, all digitally performed after the acquisition of a single fluorescence image. This framework also cross-connects different imaging modalities to each other, enabling 3D refocusing of a single wide-field fluorescence image to match confocal microscopy images acquired at different sample planes. Deep-Z has the potential to improve volumetric imaging speed while reducing challenges relating to sample drift, aberration and defocusing that are associated with standard 3D fluorescence microscopy.

High-throughput volumetric fluorescence imaging is very important in various fields including, for example, biology, life sciences and engineering, and still remains a challenge in microscopy research. 3D fluorescence information is usually acquired through scanning an excitation source through the sample volume to obtain images at multiple planes, forming the basis of volumetric imaging in confocal¹, two-photon², light-sheet^{3–5} and various super-resolution^{6–11} microscopy techniques. However, scanning can limit the imaging speed and throughput, potentially introducing phototoxicity and photobleaching, even with optimized scanning strategies¹ or point-spread-function (PSF) engineering^{3,12}. 3D fluorescence information of a specimen can also be acquired using non-scanning microscopy methods that simultaneously map the axial information onto a two-dimensional (2D) image, such as fluorescence light-field microscopy^{13–18}, Fresnel correlation holography^{19–21} and others^{22,23}. However, these non-scanning 3D fluorescence microscopy approaches require relatively time-consuming iterative algorithms to solve the inverse problem for reconstructing a new image, and use customized optical components and hardware, which increase the complexity of the setup.

There are emerging approaches that use deep learning to solve inverse problems in fluorescence microscopy²⁴, for example, to enhance the lateral^{25–31} and axial^{31–33} resolution using artificial neural networks trained with image data. Such deep-learning-based image reconstruction and enhancement methods take a relatively long time to train; however, this training process is a one-time effort, and after it is complete, each new sample of interest can be rapidly reconstructed through a single forward pass through the trained network, without the need for any iterations or hyperparameter

tuning, which in general forms an important advantage of deep-learning-based solutions to inverse imaging problems³⁴.

Here we introduce a digital image refocusing framework in fluorescence microscopy by training a deep neural network using microscopic image data, enabling 3D imaging of fluorescent samples using a single 2D wide-field image, without the need for any mechanical scanning, additional hardware or parameter estimation. This framework rapidly refocuses a 2D fluorescence image onto user-defined 3D surfaces (such as tilted planes, curved surfaces and others), and can be used to digitally correct for various aberrations caused by the sample and/or the optical system. We term this deep-learning-based approach Deep-Z, and use it to computationally refocus a single 2D wide-field fluorescence image onto 3D surfaces within the sample, without sacrificing the imaging speed, resolution or field of view (FOV) of a standard microscope.

In Deep-Z, an input 2D fluorescence image is first appended with a user-defined digital propagation matrix (DPM) that represents, pixel-by-pixel, the axial distance of the target surface from the plane of the input image (Fig. 1). Deep-Z is trained using a conditional generative adversarial neural network (GAN)^{35,36} with accurately matched pairs of (1) various fluorescence images axially focused at different depths and appended with different DPMs, and (2) the corresponding fluorescence images (that is, the ground-truth labels) captured at the correct (target) focus plane defined by the corresponding DPM. Through this training process that only uses experimental image data, the generator network of the GAN learns to interpret the values of each DPM pixel as an axial refocusing distance, and outputs an equivalent fluorescence image that is digitally refocused within the sample to the 3D surface defined by the user.

¹Electrical and Computer Engineering Department, University of California, Los Angeles, Los Angeles, CA, USA. ²Bioengineering Department, University of California, Los Angeles, Los Angeles, CA, USA. ³California Nano Systems Institute (CNSI), University of California, Los Angeles, Los Angeles, CA, USA. ⁴Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, USA. ⁵Department of Chemistry and Biochemistry, University of California, Los Angeles, Los Angeles, CA, USA. ⁶Department of Genetics, Hebrew University of Jerusalem, Jerusalem, Israel. ⁷Department of Surgery, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, USA. ⁸These authors contributed equally: Yichen Wu, Yair Rivenson. *e-mail: ozcan@ucla.edu

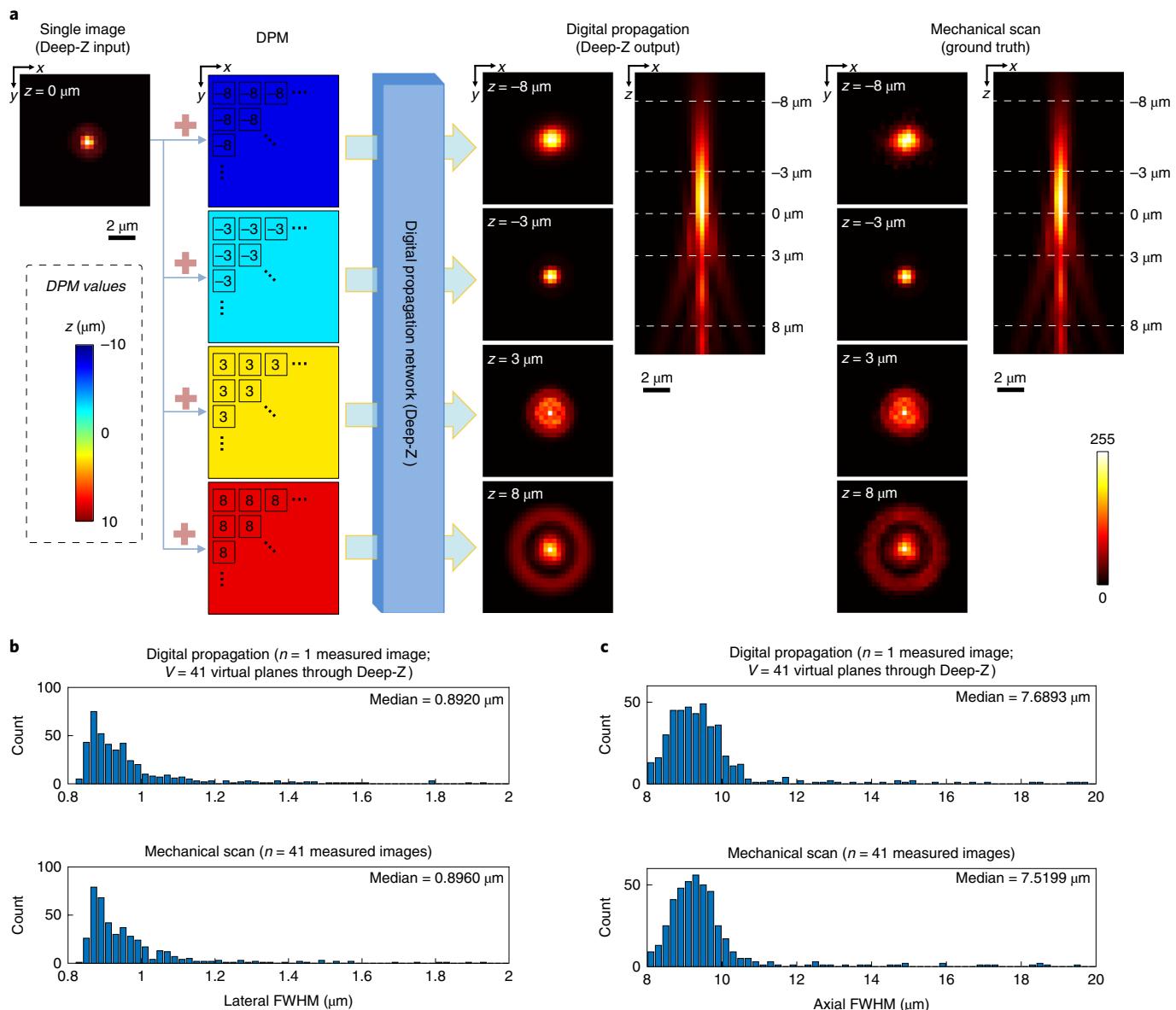


Fig. 1 | Refocusing of fluorescence images using Deep-Z. **a**, Steps involved in using the Deep-Z network. By appending a DPM to a single fluorescence image (left) and passing it through a trained Deep-Z network, refocused images at different planes can be virtually obtained. The PSF generated by Deep-Z (middle) and mechanical scanning (right) are shown for comparison. Color scale indicates intensity. **b**, Lateral FWHM histograms for 461 individual isolated fluorescence nanobeads (300 nm) distributed over $\sim 500 \times 500 \mu\text{m}^2$, measured using Deep-Z inference ($n=1$ captured image) and images obtained using mechanical axial scanning ($n=41$ captured images) match each other very well. **c**, As in **b**, except using the axial FWHM measurements for the same dataset, revealing a very good match between Deep-Z inference results and the axial mechanical scanning results.

Using Deep-Z, we imaged *Caenorhabditis elegans* neurons using a standard wide-field fluorescence microscope and extended the native depth of field (DOF) by ~20-fold. Using Deep-Z, we further demonstrated 3D tracking of the neuron activity of a *C. elegans* worm over an extended DOF using a time sequence of fluorescence images acquired at a single focal plane. Furthermore, we used spatially non-uniform DPMs to refocus a 2D input fluorescence image onto user-defined 3D surfaces to computationally correct for aberrations such as sample drift, tilt and spherical aberrations, all performed after the image acquisition and without any modifications to the optical hardware of a standard fluorescence microscope.

Another important feature of Deep-Z is that it permits cross-modality digital refocusing of fluorescence images, where the GAN is trained with gold-standard label images obtained by a different fluorescence microscopy modality. We term this framework

Deep-Z+. To demonstrate a proof of concept, we trained Deep-Z+ with input and label images that were acquired with a wide-field fluorescence microscope and a confocal microscope, respectively, to blindly generate the output of this cross-modality Deep-Z+: digitally refocused images of an input wide-field fluorescence image that match confocal microscopy images of the same sections.

Results

Digital refocusing of fluorescence images using Deep-Z. Figure 1a demonstrates Deep-Z-based digital refocusing of a single image of a 300-nm fluorescent bead (excitation and emission wavelengths of 538 nm and 584 nm, respectively) to multiple user-defined planes, represented by different DPMs; each one of these propagation matrices represents, pixel-by-pixel, the axial distance of the target surface from the plane of the input image (Fig. 1a). The native

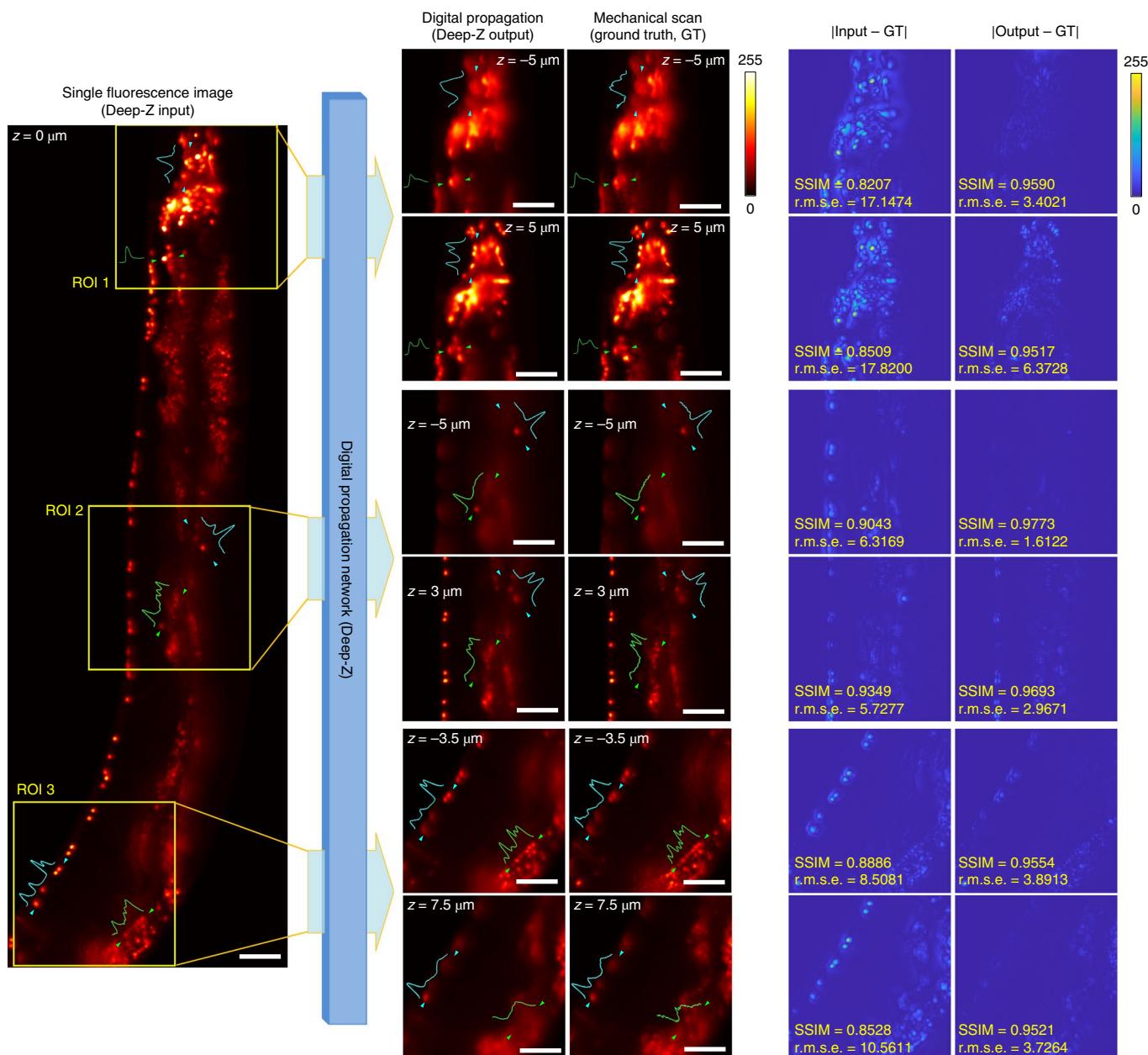


Fig. 2 | 3D imaging of *C. elegans* neuron nuclei using Deep-Z. Different ROIs are digitally refocused using Deep-Z to different planes within the sample volume; the resulting images provide a very good match to the corresponding ground-truth images, acquired using a scanning fluorescence microscope. The absolute difference images of the input and output with respect to the corresponding ground-truth image are also provided on the right, with structural similarity index (SSIM) and root mean square error (r.m.s.e.) values reported, further demonstrating the success of Deep-Z. Lines represent cross-sectional plots and arrowheads indicate where the cross-section was taken; blue and green represent two separate cross-section lateral locations. Scale bars, 25 μm . Experiments were repeated with 20 images, achieving similar results. Color scales indicate intensity.

DOF of the input fluorescence image, as defined by the numerical aperture (NA) of the objective lens ($20\times/0.75$ NA), is $\sim 1 \mu\text{m}$; using Deep-Z, we digitally refocused the image of this fluorescent bead over an axial range of approximately $\pm 10 \mu\text{m}$, matching the corresponding mechanically scanned images of the same region-of-interest (ROI), which form the ground-truth. Note that the PSF in Fig. 1a is asymmetric in the axial direction, which provides directional cues to the neural network regarding the digital propagation of an input image by Deep-Z. Unlike a symmetric Gaussian beam³⁷, such PSF asymmetry along the axial direction is ubiquitous in fluorescence microscopy systems³⁸.

Deep-Z also provides an improved signal-to-noise ratio (SNR) at its output as compared to a fluorescence image of the same

object measured at the corresponding depth (Supplementary Fig. 1). To further quantify Deep-Z output we used PSF analysis; Fig. 1b,c illustrates the histograms of both the lateral and the axial full-width-half-maximum (FWHM) values of 461 individual isolated nanobeads. These histograms agree with each other very well (Fig. 1b,c), confirming the match between Deep-Z output images calculated from a single fluorescence image and the corresponding axially scanned ground-truth images.

Next, we tested Deep-Z by imaging the neurons of a *C. elegans* nematode expressing pan-neuronal tagRFP³⁹. Figure 2 demonstrates our blind-testing results for Deep-Z-based refocusing of different parts of a *C. elegans* worm from a single wide-field fluorescence input image. Using Deep-Z, non-distinguishable fluorescent

neurons in the input image were brought into focus at different depths, while other in-focus neurons in the input image became out-of-focus and smeared into the background, according to their true axial positions in 3D; comparisons to the ground-truth mechanical scans are provided as cross-sections and image difference analyses in Fig. 2 and Supplementary Fig. 2. For optimal performance, this Deep-Z model was specifically trained using *C. elegans* samples, and the axial range of its refocusing capability is determined by the training data range ($\pm 10\text{ }\mu\text{m}$), and fails outside of this training range (Supplementary Figure 3). Using Deep-Z, we also generated (from a single 2D fluorescence image) a virtual 3D stack (Supplementary Video 1) and 3D visualization (Supplementary Video 2) of a *C. elegans* worm, over an axial range of approximately $\pm 10\text{ }\mu\text{m}$. Similar results were also obtained for a *C. elegans* imaged under a $40\times/1.3$ NA objective lens, where Deep-Z successfully refocused the input image over an axial range of approximately $\pm 4\text{ }\mu\text{m}$ (Supplementary Fig. 4).

Next, we captured a video of four moving *C. elegans* worms, where each frame of this fluorescence video was digitally refocused to various depths using Deep-Z. This enabled us to create simultaneously running videos of the same sample, each one focused at a different depth (Supplementary Video 3). Each one of these virtually created videos are temporally synchronized to each other (that is, the frames at different depths have identical timestamps), which is not possible with a scanning-based 3D imaging system owing to the unavoidable time delay between successive measurements of different parts of the sample. Quite importantly, Deep-Z also enables correction for sample-drift-induced defocus after the image capture. Supplementary Video 4 shows a moving *C. elegans* recorded by a fluorescence microscope, where Deep-Z digitally brought the defocused nematode into focus (also see Supplementary Note 1). In addition to 3D imaging of a nematode, Deep-Z also works well to digitally refocus the images of fluorescent samples that are spatially denser such as the mitochondria and F actin structures within bovine pulmonary artery endothelial cells (BPAEC) (Supplementary Fig. 5).

Deep-Z not only substantially boosts the imaging speed, but also reduces photobleaching on the sample. For a wide-field fluorescence microscopy experiment, where an axial image stack is acquired, the illumination excites the fluorophores through the entire specimen, and the total light exposure of a given point within the sample volume is proportional to the number of imaging planes that are acquired during a single-pass z stack. By contrast, Deep-Z only requires a single image acquisition, if its axial training range covers the sample depth. This reduction, enabled by Deep-Z, in the number of axial planes that need to be imaged within a sample directly helps to reduce the photobleaching of samples (Supplementary Fig. 6 and Supplementary Note 2). This reduced light dose is also likely to reduce the phototoxicity associated with volumetric imaging.

So far, the blindly tested samples were inferred with a network that was trained using the same type of sample and the same microscopy system. In Supplementary Notes 3–4, we evaluated the performance of Deep-Z under different scenarios, where a change in the test data distribution is introduced in comparison to the training image set, such as (1) a different type of sample is imaged, (2) a different microscopy system is used for imaging and (3) a different illumination power or SNR is used. Our results (Supplementary Figs. 7–9) reveal the robustness of Deep-Z to these changes; however, to achieve the best performance using Deep-Z, the network should be trained (from scratch or through transfer learning, which expedites the training process) using training images obtained with the same microscope system and the same types of samples as are expected to be used at the testing phase.

As illustrated in Supplementary Notes 5–6, Deep-Z is also robust to changes in the density of the fluorescent objects within the sample (up to a limit, which is a function of the axial refocusing

Table 1 | Neuron segmentation results for a *C. elegans* worm using a watershed-based segmentation algorithm

Watershed-based neuron segmentation			
	n (neurons)	Δz (mean \pm s.d.; μm)	$ \Delta z $ (mean \pm s.d.; μm)
Input image ($M=1$ image)	95	-0.265 ± 1.437	1.156 ± 0.885
Deep-Z output stack ($M=1$ image)	128	-0.575 ± 1.377	0.852 ± 1.223
Merged stack ($M=2$ images)	148	-0.157 ± 0.983	0.639 ± 0.761
Mechanical scan stack ($M=41$ images)	146	Ground truth	Ground truth

The resulting segmented neuron locations were also compared against the ground truth (that is, the corresponding mechanically scanned image stack, $M=41$ images), reporting the axial error, Δz , as well as the absolute axial error, $|\Delta z|$, as mean \pm s.d. in micrometers, respectively (also see Supplementary Note 7).

distance), the exposure time of the input images, as well as the illumination intensity modulation (Supplementary Figs. 10–13 and Supplementary Video 5).

***C. elegans* neuron segmentation.** For *C. elegans* neuron imaging, by virtual refocusing over an extended DOF, Deep-Z helps to segment more neurons and accurately predict their depth location as compared to a single focal plane image. To demonstrate this capability, we show the segmentation results of a *C. elegans* worm (Supplementary Figs. 14d–i) calculated using a watershed segmentation algorithm⁴⁰ from a 2D input image, the corresponding Deep-Z virtual image stack and the mechanically scanned ground-truth image stack (41 depths with $0.5\text{-}\mu\text{m}$ axial spacing); the results are summarized in Table 1. In comparison to the segmentation results obtained from the 2D input image (Supplementary Fig. 14e), the segmentation obtained using the Deep-Z virtual image stack (Supplementary Fig. 14f) detected 33 additional neurons, predicting the correct 3D positions of 128 neurons in total. In comparison to the ground-truth mechanically scanned 3D image stack (Supplementary Fig. 14i), the segmentation algorithm recognized 18 fewer neurons for the Deep-Z generated virtual stack, which were mostly located within the head, where the neurons are much denser and are relatively more challenging to recover and segment. In sparser regions of the worm, the neurons were mostly correctly segmented, matching the results obtained using the mechanically scanned 3D image stack (41 axial scans). The depth locations of the segmented neurons also matched well with the corresponding depths measured using the ground-truth mechanically scanned 3D image stack, with an average depth difference of $\Delta z = -0.575 \pm 1.377\text{ }\mu\text{m}$ (Table 1).

To further improve Deep-Z-based neuron segmentation in denser regions of the sample (such as the head of a worm), images from more than one focal plane can be used as input. In comparison to the mechanically scanned 3D image stack, this is still substantially faster, requiring fewer images to recover the volume of the specimen. For instance, in Supplementary Fig. 14h we demonstrate the segmentation results of merging two virtual image stacks created by Deep-Z (taking the per-pixel maximum), both spanning $-10\text{ }\mu\text{m}$ to $10\text{ }\mu\text{m}$ but generated from two different input images at $z=0\text{ }\mu\text{m}$ and $z=4\text{ }\mu\text{m}$, respectively. The segmentation algorithm in this case identified $n=148$ neurons and the results match better to the ground-truth axial scanning results, $n=146$ (Table 1). To shed more light on this comparison, we also used another segmentation algorithm (TrackMate⁴¹) on the same image dataset; the results of

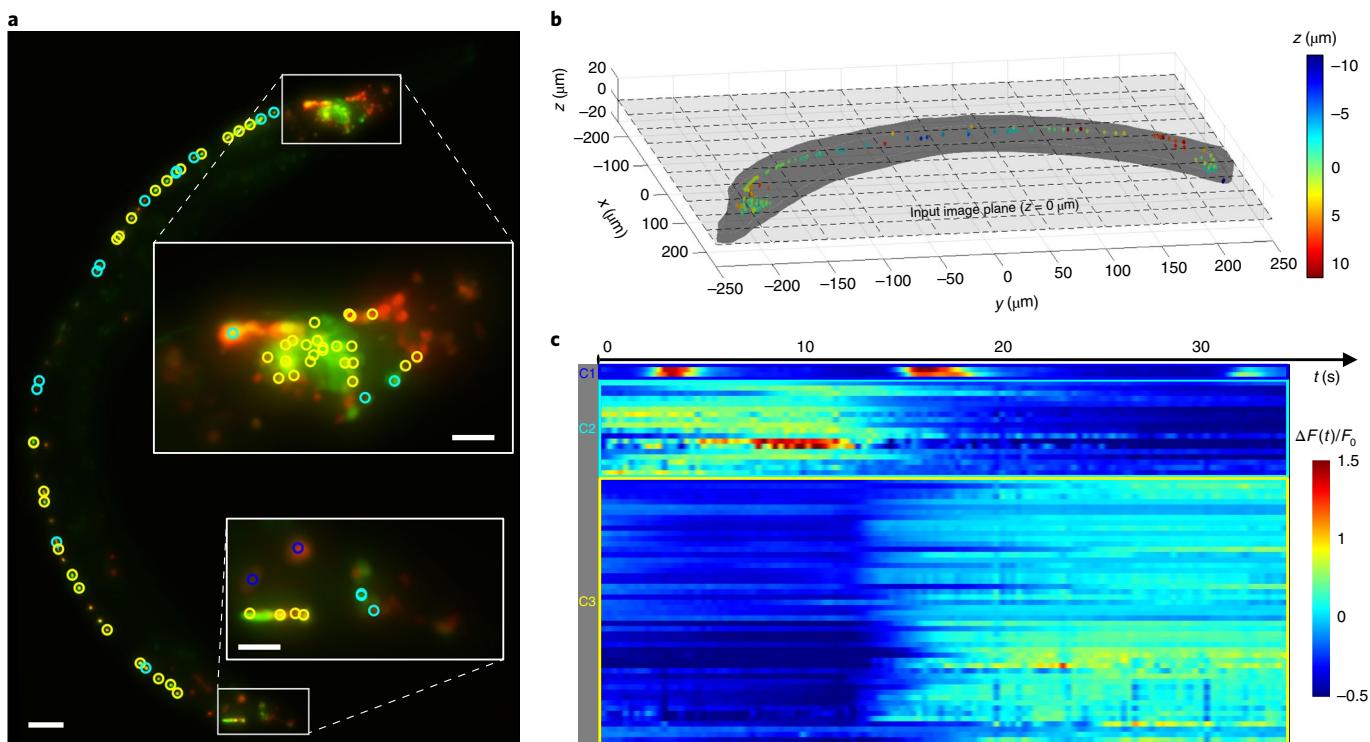


Fig. 3 | *C. elegans* neuron activity tracking in 3D using Deep-Z. **a**, MIP along the axial direction of the median-intensity image taken across the time sequence. The red channel (Texas Red) labels neuron nuclei. The green channel (FITC) labels neuron calcium activity. Scale bar, 25 μm . Scale bars for the expanded regions, 10 μm . **b**, All the 155 localized neurons are shown in 3D, depths are color-coded. **c**, 3D tracking of neuron calcium activity events corresponding to the 70 most active neurons. The neurons were grouped into three clusters (C1–C3) on the basis of similarities in their calcium activity patterns (Methods). The locations of these neurons are marked by the circles in **a**. The colors of the circles in **a** represents different clusters: C1 (blue), C2 (cyan) and C3 (yellow).

this analysis, summarized in Supplementary Note 7, confirmed a similar trend as that shown in Table 1.

3D functional imaging of *C. elegans* using Deep-Z. To highlight the utility of Deep-Z for tracking the activity of neurons in 3D, we recorded the fluorescence video of a *C. elegans* worm at a single focal plane ($z=0 \mu\text{m}$) at $\sim 3.6 \text{ Hz}$ for $\sim 35 \text{ s}$, using a $20\times/0.8 \text{ NA}$ objective lens with two fluorescence channels, FITC for neuron activity and Texas Red for neuron locations. Each frame at each channel of the acquired video was digitally refocused using Deep-Z to a series of planes ($-10 \mu\text{m}$ to $10 \mu\text{m}$, $0.5\text{-}\mu\text{m}$ step size), generating a virtual 3D fluorescence stack for each acquired frame. Supplementary Video 6 shows a comparison of the recorded input video and a video of the maximum intensity projection (MIP) along z for these virtual stacks. As can be seen in this comparison, the neurons that are defocused in the input video can be refocused on demand at the Deep-Z output for both of the fluorescence channels. This enables accurate spatiotemporal tracking of individual neuron activity in 3D from a temporal sequence of 2D fluorescence images, captured at a single focal plane.

Next, we segmented the voxels of each neuron using the Texas Red channel, and tracked the change in the fluorescence intensity, that is, $\Delta F(t) = F(t) - F_0$, in the FITC channel (neuron activity) inside each neuron segment over time, where $F(t)$ is the neuron fluorescence emission intensity and F_0 is its time average (Methods). A total of 155 neurons in 3D were isolated using Deep-Z output images (Fig. 3b, Supplementary Video 7). For comparison, in Supplementary Fig. 14b we report the results of the same segmentation algorithm applied to just the input 2D image, in which 99 neurons were identified, without any depth information.

Figure 3c plots the activities of the 70 most active neurons, which were grouped into clusters C1–C3 on the basis of similarities in their calcium activity patterns (Methods). The activity of all of the 155 neurons inferred using Deep-Z are provided in Supplementary Fig. 15. Figure 3c reports that cluster C3 calcium activities increased at $t = 14 \text{ s}$, whereas the activities of cluster C2 decreased at a similar time point. These neurons very likely correspond to the type A and B motor neurons that promote backward and forward motion, respectively, which typically anticorrelate with each other⁴². Cluster C1 features two cells that were comparatively larger in size, located in the middle of the worm. These cells had three synchronized short spikes at $t = 4, 17$ and 32 s . Their 3D positions and the regularity of their calcium activity pattern suggest that they are either neuronal or muscle cells of the defecation system that initiates defecation in regular intervals in coordination with the locomotion system⁴³. We emphasize that all this 3D-tracked neuron activity was in fact embedded in the input 2D fluorescence image sequence, which was acquired at a single focal plane. Through Deep-Z, the neuron locations and activities were accurately tracked using a 2D microscopic time sequence, without the need for mechanical scanning, additional hardware or a trade-off of resolution or imaging speed.

As Deep-Z generates temporally synchronized virtual image stacks through digital refocusing, it can be used to match the imaging speed to the limit of the camera framerate. To highlight this opportunity, we used the stream mode of the camera of our microscope (Methods) and captured two videos at 100 frames per second to monitor the neuron nuclei (Texas Red) and the neuron calcium activity (FITC) of a moving *C. elegans* over a period of 10 s, and used Deep-Z to generate virtually refocused videos over an axial range of $\pm 10 \mu\text{m}$ (Supplementary Videos 8 and 9).

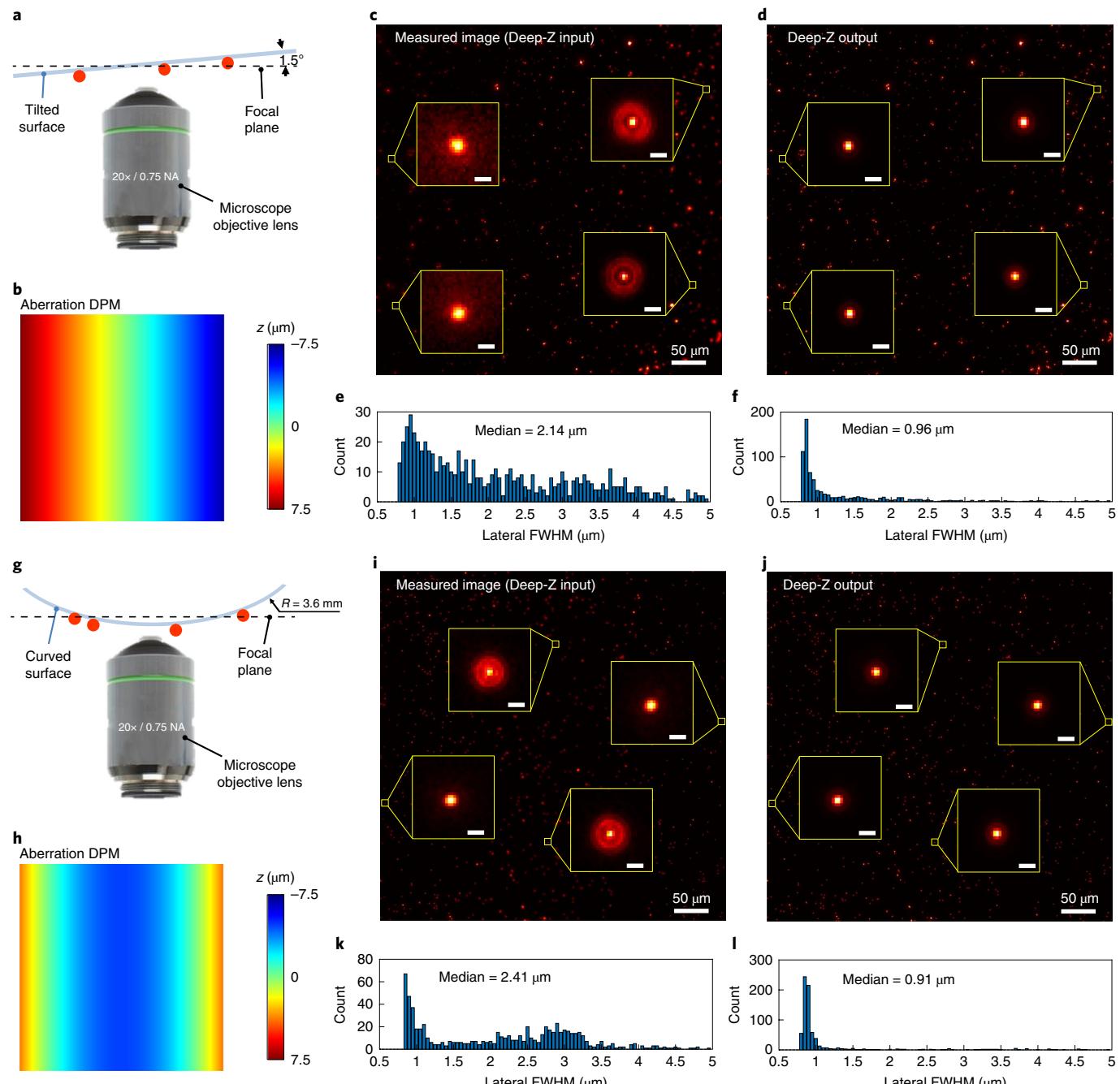


Fig. 4 | Non-uniform DPMs enable digital refocusing of a single fluorescence image onto user-defined 3D surfaces using Deep-Z. **a**, Measurement of a tilted fluorescent sample (300-nm beads). **b**, The corresponding DPM for this tilted plane. **c**, Measured raw fluorescence image; the left and right parts are out of focus in different directions, owing to the sample tilt. **d**, The Deep-Z output rapidly brings all the regions into correct focus. **e,f**, Lateral FWHM values of the nanobeads shown in **c,d**, respectively, clearly demonstrating that Deep-Z with the non-uniform DPM brought the out-of-focus particles into focus. **g**, Measurement of a cylindrical surface with fluorescent beads (300-nm beads). **h**, The corresponding DPM for this curved surface. **i**, Measured raw fluorescence image; the middle region and the edges are out-of-focus owing to the curvature of the sample. **j**, The Deep-Z output rapidly brings all the regions into correct focus. **k,l**, report the lateral FWHM values of the nanobeads shown in **i,j**, respectively, clearly demonstrating that Deep-Z with the non-uniform DPM brought the out-of-focus particles into focus, generating monodispersed distributions with a median of ~0.96 μm and ~0.91 μm, respectively. Experiments were repeated with 32 images, achieving similar results. Scale bars for the expanded regions, 2 μm.

Deep-Z-based aberration correction using spatially non-uniform DPMs. Even though Deep-Z is trained with uniform DPMs, during testing one can also use spatially non-uniform entries as part of a DPM to refocus an input fluorescence image onto user-defined 3D surfaces. Such a unique capability can be useful, among many applications, for simultaneous autofocus of different parts of a

fluorescence image after image capture and measurement or assessment of the aberrations introduced by the optical system (and/or the sample), as well as for correction of such aberrations by applying a desired non-uniform DPM. To exemplify this opportunity, Fig. 4 demonstrates the correction of the planar tilting and cylindrical curvature of two different samples, after the acquisition of a single

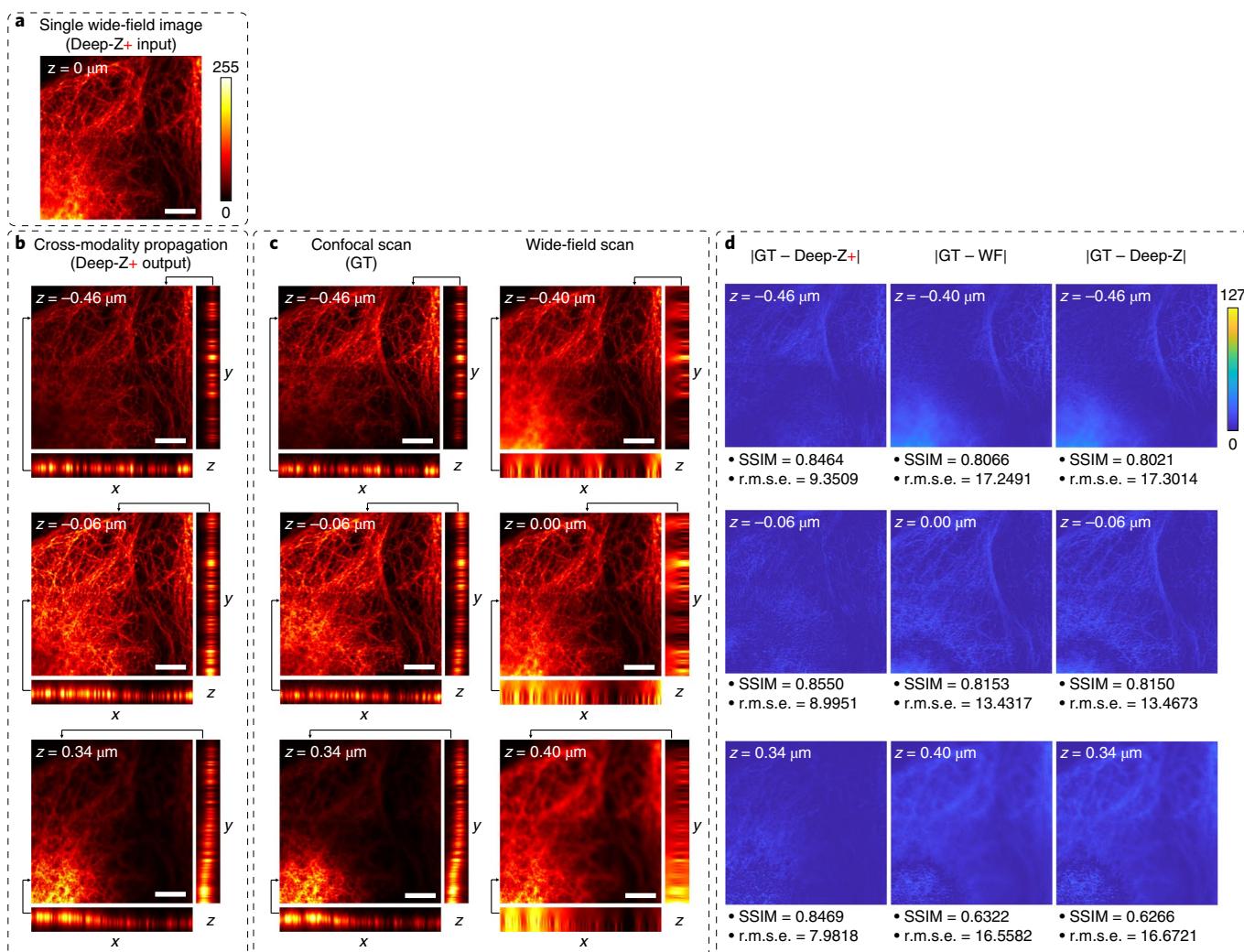


Fig. 5 | Deep-Z+, cross-modality digital refocusing of fluorescence images. **a–c.** A single wide-field fluorescence image (63 \times /1.4 NA objective lens) of BPAEC microtubule structures (**a**) was digitally refocused using Deep-Z+ to different planes in 3D (**b**), matching the images captured by a confocal microscope at the corresponding planes (**c**), retrieving volumetric information from a single input image and performing axial sectioning at the same time. Wide-field (WF) images are also shown in **c** for comparison. The cross-sections (x - z and y - z) of refocused images are shown to demonstrate the match between Deep-Z+ inference and the ground-truth (GT) confocal microscope images of the same planes; the same cross sections (x - z and y - z) are also shown for a wide-field scanning fluorescence microscope, reporting a substantial axial blur in each case. Each cross-sectional expanded image spans 1.6 μm in the z direction (with an axial step size of 0.2 μm) and the dotted arrows mark the locations at which the x - z and y - z cross sections were taken. **d.** The absolute difference images of the Deep-Z+ output with respect to the corresponding confocal images, with SSIM and r.m.s.e. values, further quantifying the performance of Deep-Z+. For comparison, we also show the absolute difference images of the ‘standard’ Deep-Z output images and the scanning wide-field fluorescence microscope images with respect to the corresponding confocal images, both of which report increased error and weaker SSIM as compared to $|\text{GT} - \text{Deep-Z+}|$. The quantitative match between $|\text{GT} - \text{WF}|$ and $|\text{GT} - \text{Deep-Z}|$ also suggests that the impact of 60-nm axial offset between the confocal and wide-field image stacks is negligible. Scale bars, 10 μm . Experiments were repeated with 42 images, achieving similar results. Color scales indicate intensity.

2D fluorescence image per object. Figure 4a illustrates the first measurement, in which the plane of a fluorescent nanobead sample was tilted with respect to the focal plane of the objective lens (Methods). By using a non-uniform DPM (see Fig. 4b), which represents the sample tilt, Deep-Z can act on the blurred input image (Fig. 4c) and accurately bring all the nanobeads into focus (Fig. 4d), even though it was only trained using uniform DPMs. The lateral FWHM values calculated at the network output image became monodispersed, with a median of $\sim 0.96 \mu\text{m}$ (Fig. 4f), as compared to a median of $\sim 2.14 \mu\text{m}$ at the input image (Fig. 4e). Similarly, Fig. 4g illustrates the second measurement, where the nanobeads were distributed on a cylindrical surface with a diameter of $\sim 7.2 \text{ mm}$. Using a non-uniform DPM that defines this cylindrical surface (Fig. 4h), the

aberration in Fig. 4i was corrected using Deep-Z (Fig. 4j); the lateral FWHM values calculated at the network output once again became monodispersed (Fig. 4l). Supplementary Note 8 details an analysis on the 3D surface curvature that a DPM can have without generating artifacts.

Cross-modality digital refocusing of fluorescence images. Deep-Z can also be used to perform cross-modality digital refocusing of an input image, where the generator can be trained using pairs of input and label images captured by two different fluorescence imaging modalities, which we term as Deep-Z+. To demonstrate this capability, we trained a Deep-Z+ network using pairs of wide-field microscopy images (inputs) and confocal microscopy

images at the corresponding planes (ground-truth labels) to perform cross-modality digital refocusing (Methods). Figure 5 demonstrates our blind-testing results for imaging microtubule structures of BPAEC using Deep-Z+. The trained Deep-Z+ network digitally refocused the input wide-field fluorescence image onto different axial distances, while at the same time rejecting some of the defocused spatial features at the refocused planes, matching the confocal images of the corresponding planes, which serve as our ground-truth (Fig. 5). For example, the microtubule structure at the lower left corner of Fig. 5b, which was prominent at a refocusing distance of $z=0.34\text{ }\mu\text{m}$, was digitally rejected by Deep-Z+ at a refocusing distance of $z=-0.46\text{ }\mu\text{m}$ as it became out of focus at this axial distance, matching the corresponding image of the confocal microscope at the same depth (Fig. 5c). Wide-field images are also shown in Fig. 5c for comparison. These scanning wide-field images report the closest heights to the corresponding confocal images, and have an axial offset of 60 nm, as the two image stacks are discretely scanned and digitally aligned to each other. Figure 5 reports x - z and y - z cross sections of Deep-Z+ output images, in which the axial distributions of the microtubule structures are substantially sharper as compared to the axial scanning images of a wide-field fluorescence microscope, providing a very good match to the cross sections obtained with a confocal microscope, matching the aim of its training.

Discussion

We developed a unique framework, termed Deep-Z, powered by deep neural networks, that enables rapid 3D refocusing within a sample using a single 2D fluorescence image as input. This framework is non-iterative and does not require hyperparameter tuning after its training stage. Even though the network is only trained using uniform DPMs, one can still apply various non-uniform DPMs during the inference stage to enable, for example, correction of sample drift, tilt, curvature or other optical aberrations, which might prove useful for longitudinal imaging experiments in biology and life sciences, by digitally recovering valuable information that might otherwise be lost owing to, for example, the sample becoming out of focus or tilted over time. On the basis of these unique features, Deep-Z also has the potential to reduce the photobleaching of samples that is associated with volumetric fluorescence imaging.

Yet another unique feature of this Deep-Z framework is that it permits cross-modality virtual refocusing of fluorescence images, where the network is trained with gold-standard label images obtained by a different fluorescence microscopy modality (for example, confocal) to teach the generator network to digitally refocus an input image (for example, an image acquired by wide-field microscopy) onto another plane within the sample volume, but this time to match the image of the same plane acquired by a different fluorescence imaging modality as compared to the input image. Figure 5 contains an example of wide-field to confocal transformation results.

We also demonstrated the efficacy of Deep-Z for structural and functional imaging of neurons in *C. elegans* nematodes. For neuron segmentation applications, we observed that Deep-Z output images in denser regions of a sample (for example, the head of a *C. elegans*) resulted in under-counting of the segmented neurons, which was improved by merging the Deep-Z output images resulting from two different focal planes as input. In fact, neuron segmentation is in general a challenging task, and not all the neurons in the body of a worm can be accurately identified in each experiment, even using a mechanically scanned image stack with a high NA objective and state-of-the-art neuron-segmentation algorithms⁴⁴. Although not demonstrated here, Deep-Z can potentially be used as a front-end module to jointly optimize future deep-learning-based neuron-segmentation algorithms, which can make use of Deep-Z to reduce the number of images required to accurately and efficiently track neural activity of model organisms. This could also benefit from new

generator architectures that utilize more than one input images, for example, from different focal planes, to more effectively combine additional 3D information acquired at different planes.

Finally, we should note that the retrievable axial range in Deep-Z depends on the SNR of the recorded image, that is, if the depth information carried by the PSF falls below the noise floor, accurate inference will be challenging. To validate the performance of a pre-trained Deep-Z network under variable SNR, we tested the inference of Deep-Z at different exposure conditions (Supplementary Figure 7), revealing the robustness of its inference over a broad range of exposure times that were not included in the training data (Supplementary Note 3). Our results demonstrated an enhancement of $\sim 20\times$ in the DOF of a wide-field fluorescence image using Deep-Z. This refocusing range is in fact not an absolute limit but rather a practical choice for our training data, and it may be further improved through hardware modifications to the optical setup by engineering the PSF in the axial direction^{3,12,45–47}. Supplementary Video 10 shows an experimental demonstration of Deep-Z blind inference for a double-helix PSF.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at <https://doi.org/10.1038/s41592-019-0622-5>.

Received: 25 March 2019; Accepted: 30 September 2019;
Published online: 4 November 2019

References

- Nguyen, J. P. et al. Whole-brain calcium imaging with cellular resolution in freely behaving *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* **113**, E1074–E1081 (2016).
- Schrödel, T., Prevedel, R., Aumayr, K., Zimmer, M. & Vaziri, A. Brain-wide 3D imaging of neuronal activity in *Caenorhabditis elegans* with sculpted light. *Nat. Methods* **10**, 1013–1020 (2013).
- Tomer, R. et al. SPED light sheet microscopy: fast mapping of biological system structure and function. *Cell* **163**, 1796–1806 (2015).
- Siedentopf, H. & Zsigmondy, R. Über sichtbarmachung und größenbestimmung ultramikroskopischer teilchen, mit besonderer anwendung auf goldröhrlinsen. *Ann. Phys.* **315**, 1–39 (1902).
- Lerner, T. N. et al. Intact-brain analyses reveal distinct information carried by SNC dopamine subcircuits. *Cell* **162**, 635–647 (2015).
- Hell, S. W. & Wichmann, J. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Opt. Lett.* **19**, 780–782 (1994).
- Hell, S. W. Far-field optical nanoscopy. *Science* **316**, 1153–1158 (2007).
- Henriques, R. et al. QuickPALM: 3D real-time photoactivation nanoscopy image processing in Image. *J. Nat. Methods* **7**, 339–340 (2010).
- Abraham, A. V., Ram, S., Chao, J., Ward, E. S. & Ober, R. J. Quantitative study of single molecule location estimation techniques. *Opt. Express* **17**, 23352–23373 (2009).
- Dempsey, G. T., Vaughan, J. C., Chen, K. H., Bates, M. & Zhuang, X. Evaluation of fluorophores for optimal performance in localization-based super-resolution imaging. *Nat. Methods* **8**, 1027–1036 (2011).
- Juette, M. F. et al. Three-dimensional sub-100 nm resolution fluorescence microscopy of thick samples. *Nat. Methods* **5**, 527–529 (2008).
- Pavani, S. R. P. et al. Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function. *Proc. Natl Acad. Sci. USA* **106**, 2995–2999 (2009).
- Prevedel, R. et al. Simultaneous whole-animal 3D imaging of neuronal activity using light-field microscopy. *Nat. Methods* **11**, 727–730 (2014).
- Levoy, M., Ng, R., Adams, A., Footer, M. & Horowitz, M. Light Field Microscopy. In *ACM SIGGRAPH 2006 Papers* 924–934 (ACM, 2006).
- Pégard, N. C. et al. Compressive light-field microscopy for 3D neural activity recording. *Optica* **3**, 517–524 (2016).
- Broxton, M. et al. Wave optics theory and 3-D deconvolution for the light field microscope. *Opt. Express* **21**, 25418–25439 (2013).
- Cohen, N. et al. Enhancing the performance of the light field microscope using wavefront coding. *Opt. Express* **22**, 24817–24839 (2014).
- Wagner, N. et al. Instantaneous isotropic volumetric imaging of fast biological processes. *Nat. Methods* **16**, 497–500 (2019).
- Rosen, J. & Brooker, G. Non-scanning motionless fluorescence three-dimensional holographic microscopy. *Nat. Photonics* **2**, 190–195 (2008).

20. Brooker, G. et al. In-line FINCH super resolution digital holographic fluorescence microscopy using a high efficiency transmission liquid crystal GRIN lens. *Opt. Lett.* **38**, 5264–5267 (2013).
21. Siegel, N., Lupashin, V., Storrie, B. & Brooker, G. High-magnification super-resolution FINCH microscopy using birefringent crystal lens interferometers. *Nat. Photonics* **10**, 802–808 (2016).
22. Abrahamsson, S. et al. Fast multicolor 3D imaging using aberration-corrected multifocus microscopy. *Nat. Methods* **10**, 60–63 (2013).
23. Abrahamsson, S. et al. MultiFocus polarization microscope (MF-PolScope) for 3D polarization imaging of up to 25 focal planes simultaneously. *Opt. Express* **23**, 7734–7754 (2015).
24. Belthangady, C. & Royer, L. A. Applications, promises, and pitfalls of deep learning for fluorescence image reconstruction. *Nat. Methods* <https://doi.org/10.1038/s41592-019-0458-z> (2019).
25. Rivenson, Y. et al. Deep learning microscopy. *Optica* **4**, 1437–1443 (2017).
26. Ouyang, W., Aristov, A., Lelek, M., Hao, X. & Zimmer, C. Deep learning massively accelerates super-resolution localization microscopy. *Nat. Biotechnol.* **36**, 460–468 (2018).
27. Nehme, E., Weiss, L. E., Michaeli, T. & Shechtman, Y. Deep-STORM: super-resolution single-molecule microscopy by deep learning. *Optica* **5**, 458–464 (2018).
28. Rivenson, Y. et al. Deep learning enhanced mobile-phone microscopy. *ACS Photonics* **5**, 2354–2364 (2018).
29. Haan, K., de, Ballard, Z. S., Rivenson, Y., Wu, Y. & Ozcan, A. Resolution enhancement in scanning electron microscopy using deep learning. *Sci. Rep.* **9**, 1–7 (2019).
30. Wang, H. et al. Deep learning enables cross-modality super-resolution in fluorescence microscopy. *Nat. Methods* **16**, 103– (2019).
31. Weigert, M. et al. Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nat. Methods* **15**, 1090– (2018).
32. Zhang, X. et al. Deep learning optical-sectioning method. *Opt. Express* **26**, 30762–30772 (2018).
33. Wu, Y. et al. Bright-field holography: cross-modality deep learning enables snapshot 3D imaging with bright-field contrast using a single hologram. *Light Sci. Appl.* **8**, 25 (2019).
34. Barbastathis, G., Ozcan, A. & Situ, G. On the use of deep learning for computational imaging. *Optica* **6**, 921–943 (2019).
35. Goodfellow, I. et al. Generative Adversarial Nets. *Adv. Neural Inf. Process. Syst.* **27**, 2672–2680 (2014).
36. Mirza, M. & Osindero, S. Conditional Generative Adversarial Nets. Preprint at *arXiv* <https://arxiv.org/abs/1411.1784> (2014).
37. Shaw, P. J. & Rawlins, D. J. The point-spread function of a confocal microscope: its measurement and use in deconvolution of 3-D data. *J. Microsc.* **163**, 151–165 (1991).
38. Kirshner, H., Aguet, F., Sage, D. & Unser, M. 3-D PSF fitting for fluorescence microscopy: implementation and localization application. *J. Microsc.* **249**, 13–25 (2013).
39. Nguyen, J. P., Linder, A. N., Plummer, G. S., Shaevitz, J. W. & Leifer, A. M. Automatically tracking neurons in a moving and deforming brain. *PLoS Comput. Biol.* **13**, e1005517 (2017).
40. Gonzalez, R. C., Woods, R. E. & Eddins, S. L. *Digital Image Processing Using MATLAB* (McGraw-Hill, 2004).
41. Tinevez, J.-Y. et al. TrackMate: an open and extensible platform for single-particle tracking. *Methods* **115**, 80–90 (2017).
42. Kato, S. et al. Global brain dynamics embed the motor command sequence of *Caenorhabditis elegans*. *Cell* **163**, 656–669 (2015).
43. Nagy, S., Huang, Y.-C., Alkema, M. J. & Biron, D. *Caenorhabditis elegans* exhibit a coupling between the defecation motor program and directed locomotion. *Sci. Rep.* **5**, 17174 (2015).
44. Toyoshima, Y. et al. Accurate automatic detection of densely distributed cell nuclei in 3D space. *PLoS Comput. Biol.* **12**, e1004970 (2016).
45. Huang, B., Wang, W., Bates, M. & Zhuang, X. Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. *Science* **319**, 810–813 (2008).
46. Antipa, N. et al. DiffuserCam: lensless single-exposure 3D imaging. *Optica* **5**, 1–9 (2018).
47. Shechtman, Y., Sahl, S. J., Backer, A. S. & Moerner, W. E. Optimal point spread function design for 3D imaging. *Phys. Rev. Lett.* **113**, 133902 (2014).

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. 2019

Methods

Sample preparation. The 300-nm red fluorescence nanobeads were purchased from MagSphere (PSF-300NM 0.3 UM RED), diluted 5,000 times with methanol and ultrasonicated for 15 min before and after dilution to break down the clusters. For the fluorescent bead samples on a flat surface and a tilted surface, a number 1 coverslip ($22 \times 22 \text{ mm}^2$, thickness of $\sim 150 \mu\text{m}$) was thoroughly cleaned and plasma treated. Then, a 2.5- μl droplet of the diluted bead sample was pipetted onto the coverslip and dried. For the fluorescent bead sample on a curved (cylindrical) surface, a glass tube (diameter of $\sim 7.2 \text{ mm}$) was thoroughly cleaned and plasma treated. Then, a 2.5- μl droplet of the diluted bead sample was pipetted onto the outer surface of the glass tube and dried.

Structural imaging of *C. elegans* neurons was carried out in strain AML18. AML18 carries the genotype *wtfls3* [rab-3p::NLS::GFP+rab-3p::NLS::tagRFP] and expresses GFP and tagRFP in the nuclei of all the neurons³⁹. For functional imaging, we used the strain AML32, carrying *wtfls5* [rab-3p::NLS::GCaMP6s+rab-3p::NLS::tagRFP]. The strains were acquired from the *Caenorhabditis* Genetics Center. Worms were cultured on nematode growth medium seeded with OP50 bacteria using standard conditions⁴⁸. For imaging, worms were washed off the plates with M9 and anesthetized with 3 mM levamisole⁴⁹. Anesthetized worms were then mounted on slides seeded with 3% agarose. To image moving worms, the levamisole was omitted.

Two slides of multilabeled BPAEC were acquired from Thermo Fisher: FluoCells Prepared Slide 1 and FluoCells Prepared Slide 2. These cells were labeled to express different cell structures and organelles. The first slide uses Texas Red for mitochondria and FITC for F-actin structures. The second slide uses FITC for microtubules.

Fluorescence image acquisition. The fluorescence images of nanobeads, *C. elegans* structure and BPAEC samples were captured by an inverted scanning microscope (IX83, Olympus Life Science) using a $20 \times /0.75 \text{ NA}$ objective lens (UPLSAPO20X, Olympus Life Science). A 130-W fluorescence light source (U-HGLGPS, Olympus Life Science) was used at 100% output power. Two bandpass optical filter sets were used, Texas Red and FITC. The bead samples were captured by placing the coverslip with beads directly on the microscope sample mount. The tilted surface sample was captured by placing the coverslip with beads on a 3D-printed holder, which created a 1.5° tilt with respect to the focal plane. The cylindrical tube surface with fluorescent beads was placed directly on the microscope sample mount. These fluorescent bead samples were imaged using a Texas Red filter set. The *C. elegans* sample slide was placed on the microscope sample mount and imaged using a Texas Red filter set. The BPAEC slide was placed on the microscope sample mount and imaged using Texas Red and FITC filter sets. For all the samples, the scanning microscope had a motorized stage (ProScan XY stage kit for IX73/83) that moved the samples to different FOVs and performed image-contrast-based autofocus at each location. The motorized stage was controlled using MetaMorph microscope automation software (Molecular Devices). At each location, the control software autofocuses the sample on the basis of the s.d. of the image, and a z-stack was taken from $-20 \mu\text{m}$ to $20 \mu\text{m}$ with a step size of $0.5 \mu\text{m}$. The image stack was captured by a monochrome scientific CMOS camera (ORCA-flash4.0 v2, Hamamatsu Photonics K.K), and saved in uncompressed tiff format with 81 planes and $2,048 \times 2,048$ pixels in each plane.

The images of *C. elegans* neuron activity were captured by another scanning wide-field fluorescence microscope (TCS SP8, Leica Microsystems) using a $20 \times /0.8 \text{ NA}$ objective lens (HCPLAPO20x/0.80DRY, Leica Microsystems) and a $40 \times /1.3 \text{ NA}$ objective lens (HC PL APO 40 \times /1.30 oil, Leica Microsystems). Two bandpass optical filter sets were used, Texas Red and FITC. The images were captured by a monochrome scientific CMOS camera (Leica DFC9000GTC-VSC08298). For capturing image stacks of anesthetized worms, the motorized stage controlled by a control software (LAS X, Leica Microsystems) moved the sample slide to different FOVs. At each FOV, the control software took a z stack from $-20 \mu\text{m}$ to $20 \mu\text{m}$ with a step size of $0.5 \mu\text{m}$ for the $20 \times /0.8 \text{ NA}$ objective lens images or a step size of $0.27 \mu\text{m}$ for the $40 \times /1.3 \text{ NA}$ objective lens images, with respect to a middle plane ($z = 0 \mu\text{m}$). Two images were taken at each z plane, for the Texas Red channel and FITC channel, respectively. For capturing 2D videos of dynamic worms, the control software took a time-lapse video that also time-multiplexed the Texas Red and FITC channels at the maximum speed of the system. This resulted in an average framerate of ~ 3.6 frames per second for a maximum camera framerate of 10 frames per second, for imaging both channels.

The BPAEC wide-field and confocal fluorescence images were captured by another inverted scanning microscope (TCS SP5, Leica Microsystems). The images were acquired using a $63 \times /1.4 \text{ NA}$ objective lens (HC PL APO 63 \times /1.40 oil CS2, Leica Microsystems) and a FITC filter set was used. The wide-field images were recorded by a charge-coupled device with $1,380 \times 1,040$ pixels and a 12-bit dynamic range, whereas the confocal images were recorded by a photo-multiplier tube with $1,024 \times 1,024$ pixels and an 8-bit dynamic range. The scanning microscope had a motorized stage that moved the sample to different FOVs and depths. For each location, a stack of 12 images with $0.2 \mu\text{m}$ axial spacing was recorded.

Image preprocessing and preparation of training data. Each captured image stack was first axially aligned using the ImageJ plugin ‘StackReg’⁵⁰, which corrects

the rigid shift and rotation caused by the microscope stage inaccuracy. Then an extended depth of field (EDF) image was generated using the ImageJ plugin ‘Extended Depth of Field’⁵¹. This EDF image was used as a reference image to normalize the whole image stack in the following steps: (1) a triangular threshold⁵² was used on the image to separate the background and foreground pixels; (2) the mean intensity of the background pixels of the EDF image was determined to be the background noise and subtracted; (3) the EDF image intensity was scaled to 0–1, where the scale factor was determined such that 1% of the foreground pixels above the background were greater than one (that is, saturated); and (4) the background level was subtracted from each image in the stack and each image was normalized by the intensity scaling factor. For testing data without an image stack, steps 1–3 were applied on the input image instead of the EDF image.

To prepare the training and validation datasets, on each FOV, a geodesic dilation⁵⁰ with fixed thresholds was applied on fluorescence EDF images to generate a mask that represents the regions containing the sample fluorescence signal above the background. Then, a customized greedy algorithm was used to determine a minimal set of regions with 256×256 pixels that covered this mask, with $\sim 5\%$ area overlap between these training regions. The lateral locations of these regions were used to crop images on each height of the image stack, where the middle plane for each region was set to be the one with the highest s.d. Then, 20 planes above and 20 planes below this middle plane were set to be the range of the stack and an input image plane was generated from each of these 41 planes. Depending on the size of the dataset, around 5–10 of these 41 planes were randomly selected as the corresponding target plane, forming around 150–300 image pairs. For each one of these image pairs, the refocusing distance was determined on the basis of the location of the plane (that is, $0.5 \mu\text{m}$ multiplied by the difference from the input plane to the target plane). By repeating this number, a uniform DPM was generated and appended to the input fluorescence image. The final dataset typically contained $\sim 100,000$ image pairs. This was randomly divided into a training dataset and a validation dataset, which took 85% and 15% of the data, respectively. During the training process, each data point was further augmented five times by flipping or rotating the images by a random multiple of 90° . The validation dataset was not augmented. The testing dataset was cropped from separate measurements with sample FOVs that did not overlap with the FOVs of the training and validation datasets.

Deep-Z network architecture. The Deep-Z network is formed by a least square GAN framework⁵³, and it is composed of two parts: a generator and a discriminator (Supplementary Note 9). The generator is a convolutional neural network inspired by the U-Net⁵⁴, and follows a similar structure as that seen in refs. 33,55. The generator network consists of a downsampling path and a symmetric upsampling path. In the downsampling path, there are five downsampling blocks. Each block contains two convolutional layers that map the input tensor x_k to the output tensor x_{k+1}

$$x_{k+1} = x_k + \text{ReLU}[\text{CONV}_{k_2}\{\text{ReLU}[\text{CONV}_{k_1}\{x_k\}]\}] \quad (1)$$

where $\text{ReLU}[\cdot]$ stands for the rectified linear unit operation and $\text{CONV}[\cdot]$ stands for the convolution operator (including the bias terms). The subscript of CONV denotes the number of channels in the convolutional layer; along the downsampling path we have: $k_1 = 25, 72, 144, 288, 576$ and $k_2 = 48, 96, 192, 384, 768$ for levels $k = 1, 2, 3, 4, 5$, respectively. The ‘+’ sign in equation (1) represents a residual connection. Zero padding was used on the input tensor x_k to compensate for the channel number mismatch between the input and output tensors. The connection between two consecutive downsampling blocks is a 2×2 max-pooling layer with a stride of 2×2 pixels to perform a $2 \times$ downsampling. The fifth downsampling block connects to the upsampling path, which will be detailed next.

In the upsampling path, there are four corresponding upsampling blocks, each of which contains two convolutional layers that map the input tensor y_{k+1} to the output tensor y_k using:

$$y_k = \text{ReLU}[\text{CONV}_{k_4}\{\text{ReLU}[\text{CONV}_{k_3}\{\text{CAT}(x_{k+1}, y_{k+1})\}]\}] \quad (2)$$

where the $\text{CAT}(\cdot)$ operator represents the concatenation of the tensors along the channel direction, that is, $\text{CAT}(x_{k+1}, y_{k+1})$ appends tensor x_{k+1} from the downsampling path to the tensor y_{k+1} in the upsampling path at the corresponding level $k+1$. The number of channels in the convolutional layers, denoted by k_3 and k_4 , are $k_3 = 72, 144, 288, 576$ and $k_4 = 48, 96, 192, 384$ along the upsampling path for $k = 1, 2, 3, 4$, respectively. The connection between consecutive upsampling blocks is an up-convolution (convolution transpose) block that upsamples the image pixels by $2 \times$. The last block is a convolutional layer that maps the 48 channels to one output channel.

The discriminator is a convolutional neural network that consists of six consecutive convolutional blocks, each of which maps the input tensor z_i to the output tensor z_{i+1} , for a given level i

$$z_{i+1} = \text{LReLU}[\text{CONV}_{i_2}\{\text{LReLU}[\text{CONV}_{i_1}\{z_i\}]\}] \quad (3)$$

where the LReLU stands for a leaky ReLU operator with a slope of 0.01. The subscript of the convolutional operator represents its number of channels, which

are $i_1=48, 96, 192, 384, 768, 1,536$ and $i_2=96, 192, 384, 768, 1,536, 3,072$, for the convolution block $i=1, 2, 3, 4, 5, 6$, respectively.

After the last convolutional block, an average pooling layer flattens the output and reduces the number of parameters to 3,072. Subsequently there are fully connected layers of size $3,072 \times 3,072$ with LReLU activation functions, and another fully connected layer of size $3,072 \times 1$ with a sigmoid activation function. The final output represents the discriminator score, which falls within $(0, 1)$, where 0 represents a false and 1 represents a true label.

All the convolutional blocks use a convolutional kernel size of 3×3 pixels and replicate padding of one pixel unless mentioned otherwise. All the convolutions have a stride of 1×1 pixel, except the second convolutions in equation (3), which has a stride of 2×2 pixels to perform a $2 \times$ downsampling in the discriminator path. The weights are initialized using the Xavier initializer⁵⁶ and the biases are initialized to 0.1.

Training and testing of the Deep-Z network. The Deep-Z network learns to use the information given by the appended DPM to digitally refocus the input image to a user-defined plane. In the training phase, the input data of the generator $G(\cdot)$ have the dimensions of $256 \times 256 \times 2$, where the first channel is the fluorescence image and the second channel is the user-defined DPM. The target data of $G(\cdot)$ have the dimensions of 256×256 , which represent the corresponding fluorescence image at a surface specified by the DPM. The input data of the discriminator $D(\cdot)$ have the dimensions of 256×256 , which can be either the generator output or the corresponding target $z^{(i)}$. During the training phase, the network iteratively minimizes the generator loss L_G and discriminator loss L_D , defined as:

$$L_G = \frac{1}{2N} \sum_{i=1}^N \left[D\left(G\left(x^{(i)}\right)\right) - 1 \right]^2 + \alpha \frac{1}{2N} \sum_{i=1}^N \text{MAE}\left(x^{(i)}, z^{(i)}\right) \quad (4)$$

$$L_D = \frac{1}{2N} \sum_{i=1}^N \left[D\left(G\left(x^{(i)}\right)\right) \right]^2 + \frac{1}{2N} \sum_{i=1}^N \left[D\left(z^{(i)}\right) - 1 \right]^2 \quad (5)$$

where N is the number of images used in each batch (for example, $N=20$), $G(x^{(i)})$ is the generator output for the input $x^{(i)}$, $z^{(i)}$ is the corresponding target label, $D(\cdot)$ is the discriminator, and $\text{MAE}(\cdot)$ stands for mean absolute error. α is a regularization parameter for the GAN loss and the MAE loss in L_G . In the training phase, it was chosen as $\alpha=0.02$. For training stability and optimal performance, an adaptive momentum optimizer was used to minimize both L_G and L_D , with a learning rate of 10^{-4} and 3×10^{-5} for L_G and L_D , respectively. In each iteration, six updates of the generator loss and three updates of the discriminator loss were performed. The validation set was tested every 50 iterations and the best network (to be blindly tested) was chosen to be the one with the smallest MAE loss on the validation set.

In the testing phase, once the training is complete, only the generator network is active. Limited by the graphical memory of our GPU, the largest image FOV that we tested was $1,536 \times 1,536$ pixels. Because the image was normalized to be in the range 0–1, whereas the refocusing distance was on the scale of around –10 to 10 (in micrometers), the DPM entries were divided by 10 to be in the range of –1 to 1 before the training and testing of the Deep-Z network to keep the dynamic range of the image and DPM matrices similar to each other.

The network was implemented using TensorFlow⁵⁷, performed on a PC with Intel Core i7-8700K six-core 3.7 GHz CPU and 32 GB RAM, using an Nvidia GeForce 1080Ti GPU. On average, the training takes ~70 h for ~400,000 iterations (equivalent to ~50 epochs). After the training, the network inference time was ~0.2 s for an image with 512×512 pixels and ~1 s for an image with $1,536 \times 1,536$ pixels on the same PC.

Measurement of the lateral and axial FWHM values of the fluorescent beads samples. For characterizing the lateral FWHM of the fluorescent beads samples, a threshold was performed on the image to extract the connected components. Then, individual regions of 30×30 pixels were cropped around the centroid of these connected components. A 2D Gaussian fit was performed on each of these individual regions, which was done using lsqcurvefit (<https://www.mathworks.com/help/optim/ug/lsqcurvefit.html>) in Matlab (MathWorks) to match the function:

$$I(x, y) = A \exp \left[-\frac{(x - x_c)^2}{2\sigma_x^2} - \frac{(y - y_c)^2}{2\sigma_y^2} \right] \quad (6)$$

The lateral FWHM was then calculated as the mean FWHM of x and y directions

$$\text{FWHM}_{\text{lateral}} = 2\sqrt{2 \ln 2} \frac{\sigma_x \Delta_x + \sigma_y \Delta_y}{2} \quad (7)$$

where $\Delta_x = \Delta_y = 0.325 \mu\text{m}$ was the effective pixel size of the fluorescence image on the object plane. A histogram was subsequently generated for the lateral FWHM values for all the thresholded beads (for example, $n=461$ for Fig. 1 and $n>750$ for Fig. 4).

To characterize the axial FWHM values for the bead samples, slices along the $x-z$ direction with 81 steps were cropped at $y=y_c$ for each bead, from either the

digitally refocused or the mechanically scanned axial image stack. Another 2D Gaussian fit was performed on each cropped slice, to match the function

$$I(x, z) = A \exp \left[-\frac{(x - x_c)^2}{2\sigma_x^2} - \frac{(z - z_c)^2}{2\sigma_z^2} \right] \quad (8)$$

The axial FWHM was then calculated as

$$\text{FWHM}_{\text{axial}} = 2\sqrt{2 \ln 2} \sigma_z \Delta_z \quad (9)$$

where $\Delta_z = 0.5 \mu\text{m}$ was the axial step size. A histogram was subsequently generated for the axial FWHM values.

Image quality evaluation. The network output images I^{out} were evaluated with reference to the corresponding ground-truth images I^{GT} using the following five criteria: (1) mean square error (MSE), (2) r.m.s.e., (3) MAE, (4) correlation coefficient and (5) SSIM⁵⁸. The MSE is one of the most widely used error metrics, defined as

$$\text{MSE}(I^{\text{out}}, I^{\text{GT}}) = \frac{1}{N_x N_y} \|I^{\text{out}} - I^{\text{GT}}\|_2^2 \quad (10)$$

where N_x and N_y represent the number of pixels in the x and y directions, respectively. The square root of MSE results in r.m.s.e. In comparison to MSE, MAE uses 1-norm difference (absolute difference) instead of 2-norm difference, which is less sensitive to outlier pixels

$$\text{MAE}(I^{\text{out}}, I^{\text{GT}}) = \frac{1}{N_x N_y} \|I^{\text{out}} - I^{\text{GT}}\|_1 \quad (11)$$

The correlation coefficient is defined as

$$\text{corr}(I^{\text{out}}, I^{\text{GT}}) = \frac{\sum_x \sum_y (I_{xy}^{\text{out}} - \mu_{\text{out}})(I_{xy}^{\text{GT}} - \mu_{\text{GT}})}{\sqrt{\left(\sum_x \sum_y (I_{xy}^{\text{out}} - \mu_{\text{out}})^2\right) \left(\sum_x \sum_y (I_{xy}^{\text{GT}} - \mu_{\text{GT}})^2\right)}} \quad (12)$$

where μ_{out} and μ_{GT} are the mean values of the images I^{out} and I^{GT} respectively.

While these criteria listed above can be used to quantify errors in the network output as compared to the GT, they are not strong indicators of the perceived similarity between two images. SSIM aims to address this shortcoming by evaluating the structural similarity in the images, defined as

$$\text{SSIM}(I^{\text{out}}, I^{\text{GT}}) = \frac{(2\mu_{\text{out}}\mu_{\text{GT}} + C_1)(2\sigma_{\text{out},\text{GT}} + C_2)}{(\mu_{\text{out}}^2 + \mu_{\text{GT}}^2 + C_1)(\sigma_{\text{out}}^2 + \sigma_{\text{GT}}^2 + C_2)} \quad (13)$$

where σ_{out} and σ_{GT} are the standard deviations of I^{out} and I^{GT} respectively, and $\sigma_{\text{out},\text{GT}}$ is the cross-variance between the two images; C_1 and C_2 are constants, used to avoid division by a small denominator.

Tracking and quantification of *C. elegans* neuron activity. The *C. elegans* neuron activity tracking video was captured by time multiplexing the two fluorescence channels (FITC, followed by Texas Red and then FITC and so on). The adjacent frames were combined so that the green color channel was FITC (neuron activity) and the red color channel was Texas Red (neuron nuclei). Subsequent frames were aligned using a feature-based registration toolbox with projective transformation in Matlab (MathWorks) to correct for slight body motion of the worms. Each input video frame was appended with DPMs representing propagation distances from $-10 \mu\text{m}$ to $10 \mu\text{m}$ with a step size of $0.5 \mu\text{m}$, and then tested through a Deep-Z network (specifically trained for this imaging system), which generated a virtual axial image stack for each frame in the video.

To localize individual neurons, the red channel stacks (Texas Red, neuron nuclei) were projected by median-intensity through the time sequence. Local maxima in this projected median-intensity stack marked the centroid of each neuron and the voxels of each neuron was segmented from these centroids by watershed segmentation⁴⁰, which generated a 3D spatial voxel mask for each neuron. A total of 155 neurons were isolated. Then, the average of the 100 brightest voxels in the green channel (FITC, neuron activity) inside each neuron spatial mask was calculated as the calcium activity intensity $F_i(t)$, for each time frame t and each neuron $i=1, 2, \dots, 155$. The differential activity was then calculated, $\Delta F(t)=F(t) - F_0$, for each neuron, where F_0 is the time average of $F(t)$.

By thresholding on the s.d. of each $\Delta F(t)$, we selected the 70 most active cells and performed further clustering on them on the basis of similarities in their calcium activity pattern (Supplementary Figure 15b) using a spectral-clustering algorithm^{59,60}. The calcium activity pattern similarity was defined as

$$S_{ij} = \exp \left(-\frac{\left\| \frac{\Delta F_i(t)}{F_{j0}} - \frac{\Delta F_j(t)}{F_{j0}} \right\|^2}{\sigma^2} \right) \quad (14)$$

for neurons i and j , which results in a similarity matrix S (Supplementary Figure 15c). The s.d. of this Gaussian similarity function is $\sigma=1.5$, which controls the width of the neighbors in the similarity graph. The spectral clustering solves an eigenvalue problem on the graph Laplacian L generated from the similarity matrix S , defined as the difference of weight matrix W and degree matrix D

$$L = D - W \quad (15)$$

where

$$W_{ij} = \begin{cases} S_{ij} & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases} \quad (16)$$

$$D_{ij} = \begin{cases} \sum_j W_{ij} & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \quad (17)$$

The number of clusters was chosen using eigengap heuristics⁵⁹, which was the index of the largest general eigenvalue (by solving the general eigenvalue problem $Lv = \lambda Dv$) before the eigengap, where the eigenvalues jump up, which was determined to be $k=3$ (Supplementary Figure 15d). Then the corresponding first $k=3$ eigenvectors were combined as a matrix, whose rows were clustered using standard k -means clustering⁵⁹, which resulted in the three clusters of calcium activity patterns shown in Supplementary Figure 15e and the rearranged similarity matrix shown in Supplementary Figure 15f.

Cross-modality alignment of wide-field and confocal fluorescence images. After its training, the Deep-Z+ network learns to digitally refocus a single input fluorescence image acquired by a fluorescence microscope to a user-defined target surface in 3D, but the output will match an image of the same sample captured by a different fluorescence imaging modality at the corresponding height (plane). To demonstrate this capability, we trained a Deep-Z+ network using pairs of wide-field microscopy images (inputs) and confocal microscopy images at the corresponding planes, where each stack of the wide field–confocal pair was first self-aligned and normalized using the method described above. Then the individual FOVs were stitched together using the ‘Image Stitching’ plugin in ImageJ (https://imagej.net/Image_Stitching). The stitched wide-field and confocal EDF images were then simultaneously registered using a feature-based registration with projective transformation performed in Matlab (MathWorks; <https://www.mathworks.com/help/vision/ref/detectsurffeatures.html>). Then the stitched confocal EDF images, as well as the stitched stacks, were warped using this estimated transformation to match their wide-field counterparts. The non-overlapping regions of the wide-field and warped confocal images were subsequently deleted. Then the greedy algorithm described above was used to crop non-empty regions of 256×256 pixels from the remaining stitched wide-field images and their corresponding warped confocal images. The same feature-based registration was applied on each pair of cropped regions for fine alignment. This step provides good correspondence between the wide-field image and the corresponding confocal image in the lateral directions (Supplementary Note 10).

Although the axial scanning step size was fixed to be $0.2\mu\text{m}$, the reference zero-point in the axial direction for the wide-field and the confocal stacks needed to be matched. To determine this reference zero-point in the axial direction, the images at each depth were compared with the EDF image of the same region using the SSIM⁵⁷, providing a focus curve. A second-order polynomial fit was performed on four points in this focus curve with highest SSIM values, and the reference zero-point was determined to be the peak of the fit. The heights of wide-field and confocal stacks were then centered by their corresponding reference zero-points in the axial direction. For each wide-field image used as input, four confocal images were randomly selected from the stack as the target, and their DPMs were calculated on the basis of the axial difference of the centered height values of the confocal and the corresponding wide-field images (Supplementary Note 10).

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

Data availability

We declare that all the data supporting the findings of this work are available within the manuscript and its supplementary information.

Code availability

Deep learning models reported in this work used standard libraries and scripts that are publicly available in TensorFlow. Through a custom-written Fiji-based plugin, we provided our trained network models (together with some sample test

images) for the following objective lenses: Leica HC PL APO 20 \times /0.80 NA DRY (two different network models trained on TxRd and FITC channels), Leica HC PL APO 40 \times /1.30 NA oil (trained on TxRd channel), Olympus UPLSAPO20X 0.75 NA (trained on TxRd channel). We made this custom-written plugin and our models publicly available through the following links: <http://bit.ly/deep-z-git> and <http://bit.ly/deep-z>.

References

48. Brenner, S. The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71–94 (1974).
49. Strange, K. (Ed.) *C. elegans: Methods and Applications* (Humana Press, 2006).
50. Thevenaz, P., Ruttimann, U. E. & Unser, M. A pyramid approach to subpixel registration based on intensity. *IEEE Trans. Image Process.* **7**, 27–41 (1998).
51. Forster, B., Van de Ville, D., Berent, J., Sage, D. & Unser, M. Complex wavelets for extended depth-of-field: a new method for the fusion of multichannel microscopy images. *Microsc. Res. Tech.* **65**, 33–42 (2004).
52. Zack, G. W., Rogers, W. E. & Latt, S. A. Automatic measurement of sister chromatid exchange frequency. *J. Histochem. Cytochem.* **25**, 741–753 (1977).
53. Mao, X. et al. Least squares generative adversarial networks. In *Proc. 2017 IEEE International Conference on Computer Vision* 2813–2821 (IEEE, 2017).
54. Ronneberger, O., Fischer, P. & Brox, T. U-Net: convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention 2015* (eds Navab, N. et al.) 234–241 (Springer, 2015).
55. Wu, Y. et al. Extended depth-of-field in holographic imaging using deep-learning-based autofocusing and phase recovery. *Optica* **5**, 704–710 (2018).
56. Glorot, X. & Bengio, Y. Understanding the difficulty of training deep feedforward neural networks. In *Proc. Thirteenth International Conference on Artificial Intelligence and Statistics* 249–256 (2010).
57. Abadi, M. et al. TensorFlow: a system for large-scale machine learning. In *Proc. 12th USENIX Symposium on Operating Systems Design and Implementation* 265–283 (USENIX, 2016).
58. Wang, Z., Bovik, A. C., Sheikh, H. R. & Simoncelli, E. P. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* **13**, 600–612 (2004).
59. Shi, J. & Malik, J. Normalized cuts and image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **22**, 888–905 (2000).
60. von Luxburg, U. A tutorial on spectral clustering. *Stat. Comput.* **17**, 395–416 (2007).

Acknowledgements

The authors acknowledge Y. Luo, X. Tong, T. Liu, H. C. Koydemir and Z. S. Ballard of University of California, Los Angeles (UCLA), as well as Leica Microsystems for their help with some of the experiments. The Ozcan group at UCLA acknowledges the support of Koc Group, National Science Foundation and the Howard Hughes Medical Institute. Y.W. also acknowledges the support of a SPIE John Kiel scholarship. Some of the reported optical microscopy experiments were performed at the Advanced Light Microscopy/Spectroscopy Laboratory and the Leica Microsystems Center of Excellence at the California NanoSystems Institute at UCLA with funding support from National Institutes of Health Shared Instrumentation grant S10OD025017 and National Science Foundation Major Research Instrumentation grant CHE-0722519. We also thank Double Helix Optics for providing their SPINDLE system and DH-PSF phase mask, which was used for engineered PSF data capture, and acknowledge X. Yang and M.P. Lake for their assistance with these engineered PSF experiments and related analysis.

Author contributions

A.O., Y.W. and Y.R. initiated the research. Y.W., Y.R. and H.W. performed the experiments. E.B.-D. cultured and prepared *C. elegans* samples. Y.W. and Y.L. processed the data. L.A.B. and C.P. helped with the experiments and analysis. A.O., Y.W. and Y.R. prepared the manuscript. A.O. supervised the research.

Competing interests

A.O., Y.W. and Y.R. have a pending patent application on the presented framework.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41592-019-0622-5>.

Correspondence and requests for materials should be addressed to A.O.

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

Reprints and permissions information is available at www.nature.com/reprints.

Reporting Summary

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

Statistics

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

n/a Confirmed

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

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

Software and code

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

Data collection

Deep-Z virtual propagation network was trained on wide-field fluorescence images collected from two standard fluorescence microscopes: IX83 (Olympus Corporation) controlled using MetaMorph® microscope automation software (Molecular Devices, LLC), and TCS SP8 controlled using LAS X software (Leica Microsystems) without modifications.
Deep-Z+ virtual cross-modality propagation network was trained on wide-field and confocal fluorescence images of BPAEC collected from an inverted scanning microscope TCS SP5 controlled using LAS X software (Leica Microsystems) without modifications.

Data analysis

Image registrations were all performed using Fiji (ImageJ) plugins "StackReg", "Extended Depth of Field" and "Image Stitching", as well as Matlab R2018a feature-based registration toolbox.
Deep learning models reported in this work used standard libraries and scripts that are publicly available in TensorFlow v1.10 (Google Inc.) interfaced with Python v3.6.
Through a custom-written Fiji based plugin, we provided our trained network models (together with sample test images) for the following objective lenses: Leica HC PL APO 20x/0.80 DRY (two different network models trained on TxRd and FITC channels), Leica HC PL APO 40x/1.30 OIL (trained on TxRd channel), Olympus UPLSAPO20X - 0.75 NA (trained on TxRd channel). We made this custom-written plugin and our models publicly available through the following links:
<http://bit.ly/deep-z-git>
<http://bit.ly/deep-z>

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

Data

Policy information about [availability of data](#)

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

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

We declare that all the data supporting the findings of this work are available within the manuscript and Supplementary Information files.

Field-specific reporting

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

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

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

Life sciences study design

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

Sample size

Four Deep-Z virtual propagation networks with same architecture were trained on 86,000, 110,000, 105,000 and 77,000 pairs of images for four different objective lenses, where each image had 256×256 pixels.
 One Deep-Z+ cross-modality propagation network with same architecture was trained on 40,000 pairs of wide-field and confocal images, each with 256×256 pixels.
 Four additional Deep-Z virtual propagation networks with same architecture were trained using transfer learning on 25,000, 35,000, 18,000 and 110,000 pairs of images, each image with 256×256 pixels.
 A validation dataset (separate and independent from the training and testing datasets) of ~ 20% of the size of the training dataset is used in each training and transfer learning phase; the best network (blindly tested) is chosen to be the one with the smallest MAE loss on the validation dataset.

Data exclusions

During the training of the Deep-Z network for *C. elegans*, the regions of the image stacks that contained a moving worm or a dead worm were manually identified and removed from the training and validation data sets.

Replication

Deep-Z networks were first successfully tested on a total of 120 images, each with 1536×1536 pixels, where each image was virtually refocused to 20 to 40 planes at different depths within the training axial range, including the following:
 -- Tested with 60 images containing 300 nm red fluorescence beads captured under a microscope with a 20× objective and Texas Red filter set, achieving similar results as in Figure 1.
 -- Tested with 20 images containing *C. elegans* worms captured under a microscope with a 20× objective and Texas Red filter set through difference experiments, achieving similar results as in Figure 2 and Supplementary Figure 2.
 -- Tested with 18 images containing *C. elegans* worms captured under a microscope with a 20× objective and FITC filter set through difference experiments, achieving similar results as in Figure 2 and Supplementary Figure 2.
 -- Tested with 12 images containing *C. elegans* worms captured under a microscope with a 40× objective and Texas Red filter set, achieving similar results as in Supplementary Figure 4.
 -- Tested with 5 imaging FOVs (cropped into 180 images with 240×240 pixels) containing BPAEC captured under a microscope with a 20× objective and Texas Red filter set, achieving similar results as in Supplementary Figure 5
 -- Tested with 5 imaging FOVs (cropped into 180 images with 240×240 pixels) containing BPAEC captured under a microscope with a 20× objective and FITC filter set, achieving similar results as in Supplementary Figure 5

The networks were also successfully tested on another 16 time-sequences of fluorescence images, each containing 120 - 1000 frames, with 1536×1536 pixels in each frame, achieving similar results as in Supplementary Videos 3-5, 7, 8.

These networks were further successfully tested on 32 images virtually mapped to 3D surfaces that were not part of the training phase (training only included refocusing to planes), achieving similar results as in Figure 4.

Deep-Z+ network was successfully tested on 42 image patches, each with 512×512 pixels, where each image was virtually refocused to 7 planes at different depths within the training axial range, achieving similar results as in Figure 5.

Randomization

The training, validation and testing data sets were randomly selected.

Blinding

Performances of deep neural networks were blindly tested and reported on images that were not included in the training or validation phases of the network training.

Reporting for specific materials, systems and methods

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

Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	Antibodies
<input checked="" type="checkbox"/>	Eukaryotic cell lines
<input checked="" type="checkbox"/>	Palaeontology
<input type="checkbox"/>	Animals and other organisms
<input checked="" type="checkbox"/>	Human research participants
<input checked="" type="checkbox"/>	Clinical data

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	ChIP-seq
<input checked="" type="checkbox"/>	Flow cytometry
<input checked="" type="checkbox"/>	MRI-based neuroimaging

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals

Caenorhabditis elegans (*C. elegans*) worms [strain AML18 and AML32] were acquired from Caenorhabditis Genetics Center (CGC) and cultured for fluorescence structural and functional imaging of their neurons.

Wild animals

None

Field-collected samples

None

Ethics oversight

No ethical approval or guidance was required because there is no requirement for ethics in worms.

Note that full information on the approval of the study protocol must also be provided in the manuscript.