nature photonics

Article

https://doi.org/10.1038/s41566-022-01140-6

# Artificial confocal microscopy for deep label-free imaging

Received: 27 October 2021

Accepted: 1 December 2022

Published online: 12 January 2023

Check for updates

Xi Chen[1,9] ✉, Mikhail E. Kandel[1,10], Shenghua He[2], Chenfei Hu[1,3], Young Jae Lee[1,4], Kathryn Sullivan[5], Gregory Tracy[6], Hee Jung Chung[1,4,6,7], Hyun Joon Kong[1,5,7,8], Mark Anastasio[1,5] & Gabriel Popescu[1,3,5,7]

Wide-field microscopy of optically thick specimens typically features reduced contrast due to spatial cross-talk, in which the signal at each point in the field of view is the result of a superposition from neighbouring points that are simultaneously illuminated. In 1955, Marvin Minsky proposed confocal microscopy as a solution to this problem. Today, laser scanning confocal fluorescence microscopy is broadly used due to its high depth resolution and sensitivity, but comes at the price of photobleaching, chemical and phototoxicity. Here we present artificial confocal microscopy (ACM) to achieve confocal-level depth sectioning, sensitivity and chemical specificity non-destructively on unlabelled specimens. We equipped a commercial laser scanning confocal instrument with a quantitative phase imaging module, which provides optical path-length maps of the specimen in the same field of view as the fluorescence channel. Using pairs of phase and fluorescence images, we trained a convolution neural network to translate the former into the latter. The training to infer a new tag is very practical as the input and ground truth data are intrinsically registered and the data acquisition is automated. The ACM images present much stronger depth sectioning than the input (phase) images, enabling us to recover confocal-like tomographic volumes of microspheres, hippocampal neurons in culture, and three-dimensional liver cancer spheroids. By training on nucleus-specific tags, ACM allows for segmenting individual nuclei within dense spheroids for both cell counting and volume measurements. In summary, ACM can provide quantitative, dynamic data, non-destructively from thick samples while chemical specificity is recovered computationally.

Three-dimensional (3D) cellular systems have been increasingly adopted over 2D cell monolayers to study disease mechanisms[1] and discover drug therapeutics[2], as they more accurately recapitulate the in vivo cellular function and development of extracellular matrices[3]. Three-dimensional cellular structures, including cellular clusters such as organoids and spheroids, have found use in a wide range of applications such as tissue engineering[4], high-throughput toxicology[5] and personalized medicine[6]. A particularly exciting direction of research is engineering multicellular living systems[7–9]. These fields of current scientific interest bring along the urgent need for new methods of investigation to inform on cellular viability and cell cluster proliferation. Such techniques would ideally provide quantitative data with subcellular resolution at arbitrary depths in the cellular system and dynamic information rendered over broad time scales. Importantly, these assays would be completely non-destructive, that is, they would report on the cell cluster without interfering with its viability and function.

A full list of affiliations appears at the end of the paper. ✉e-mail: xc289@cornell.edu

https://doi.org/10.1038/s41566-022-01140-6

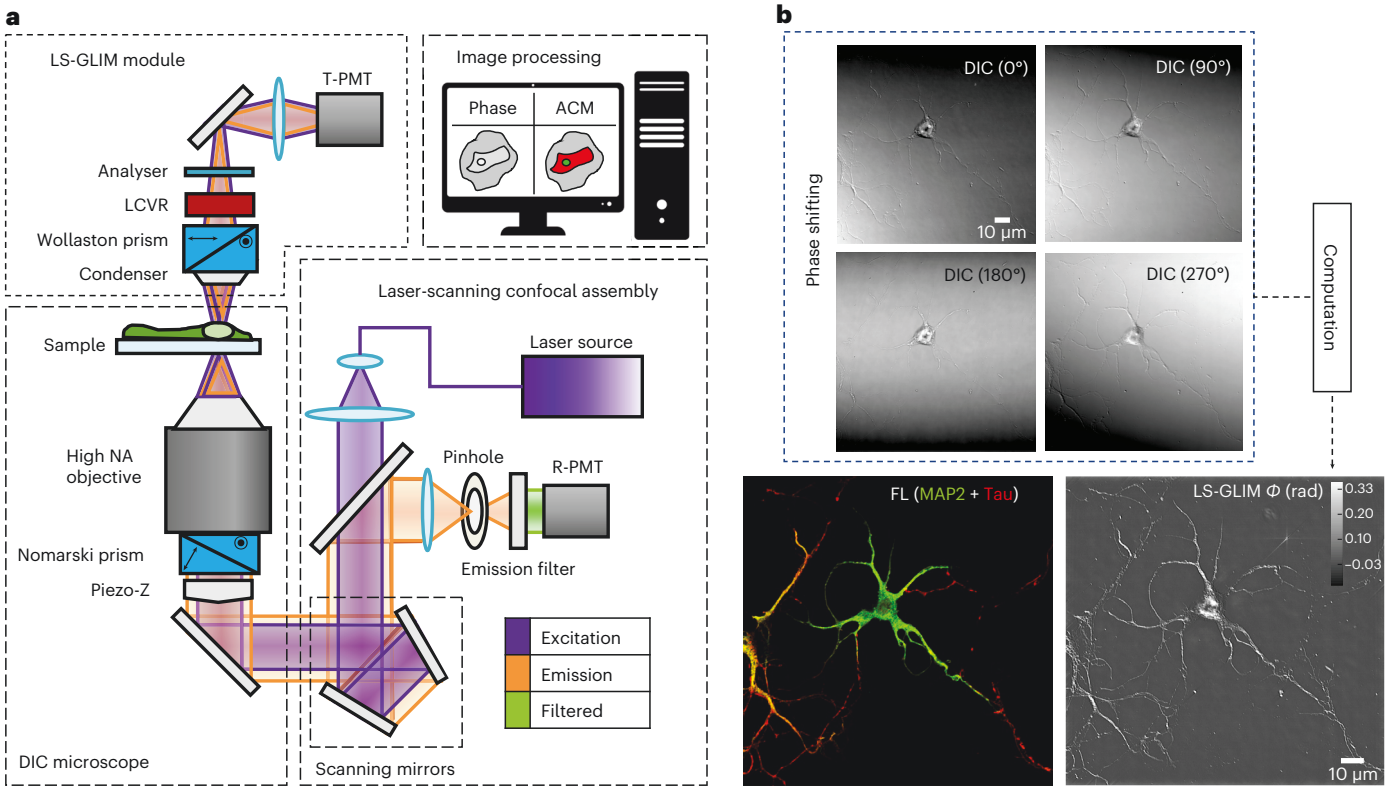



**Fig. 1 | ACM optical path and image processing. a**, The ACM system consists of a laser scanning confocal assembly, a DIC microscope and a LS-GLIM module. QPI was conducted with the green laser line (488 nm) of the confocal excitation. The interferogram was recorded at each point in the scan by the transmission-PMT (T-PMT). The fluorescence images were captured by the reflection-PMT (R-PMT) of the confocal module. **b**, Four phase-shifting frames are recorded and used to reconstruct the quantitative phase image. The confocal fluorescence image (FL) serves as the ground truth, whereas the phase image (LS-GLIM) is the input for the network training. NA, numerical apeture.

Due to visible light wavelengths being comparable in size to subcellular structures, optical methods of investigation are well suited for meeting these requirements; however, considerable challenges exist for the existing optical microscopy techniques when applied to increasingly thick samples. Typical spheroids—ranging from hundreds of micrometres to millimetres in size—are much larger than the scattering mean free path associated with the light wave propagation, generate strong multiple scattering and therefore form optically turbid aggregates that are difficult to analyse at a cellular level[10,11]. As a result, high-throughput investigations often are limited to extracting coarse parameters, such as spheroid diameters, at low-magnification[12].

In 1955, in his pursuit to image 3D biological neuronal networks and mimic their behaviour computationally, Minsky was faced with the challenge of suppressing multiple scattering, which was particularly severe for the wide-field instruments available at the time[13]. In Minsky's own words, "One day it occurred to me that the way to avoid all that scattered light was to never allow any unnecessary light to enter in the first place. An ideal microscope would examine each point of the specimen and measure the amount of light scattered or absorbed by that point"[13]. This first implementation of the confocal scanning microscope was established in a transmission geometry, requiring sample translation. Of course, today's modern confocal instruments take advantage of bright laser sources, use beam scanning and are most often used in a reflection geometry, paired with fluorescence contrast[14]. In time, many other advanced laser scanning techniques have been developed for fluorescence microscopy[15]. Nevertheless, fluorescence imaging is subject to several limitations. Absorption of the excitation light may cause the fluorophore to photobleach, which limits the time interval over which continuous imaging can be performed[16]. The excitation light is typically toxic to cells, a phenomenon referred to as phototoxicity, whereas the exogenous fluorophores themselves can induce chemical toxicity[17]. Although the advancement of green fluorescent protein technology substantially improves the viability of the specimen under investigation, concerns regarding phototoxicity, photobleaching and functional integrity of the cells following genetic engineering still remain[18]. Overcoming these limitations becomes extremely challenging when imaging thick objects over an extended period of time and, for that reason, confocal microscopy is often used on fixed specimens[19,20].

Multiphoton techniques, including harmonic generation and two-photon (intrinsic) fluorescence microscopy, have been established as valuable label-free approaches for deep-tissue imaging with cellular resolution. Multiphoton microscopy uses excitation light with a longer wavelength that penetrates deeper into tissues, whereas the nonlinear process requires a multiphoton interaction that renders 3D localized excitation[21]. However, multiphoton microscopy requires expensive instrumentations such as femtosecond lasers that are less accessible to the broader community, and the higher-order nonlinear excitation is more susceptible to focus aberrations and phototoxicity. Light sheet fluorescence microscopy can acquire 3D tomography of biological specimens in seconds with high optical sectioning and axial resolution due to the sheet-like illumination, minimizing the background fluorescence and photobleaching[22]. However, tomographic reconstruction often requires sample rotation and sophisticated sample mounting. On the other hand, optical coherence tomography—an interferometric label-free method—was reported to detect and count aqueous cells in the anterior chamber of a rodent model of eye inflammation[23] and volumetrically quantify tumour spheroids[24]. Several

https://doi.org/10.1038/s41566-022-01140-6

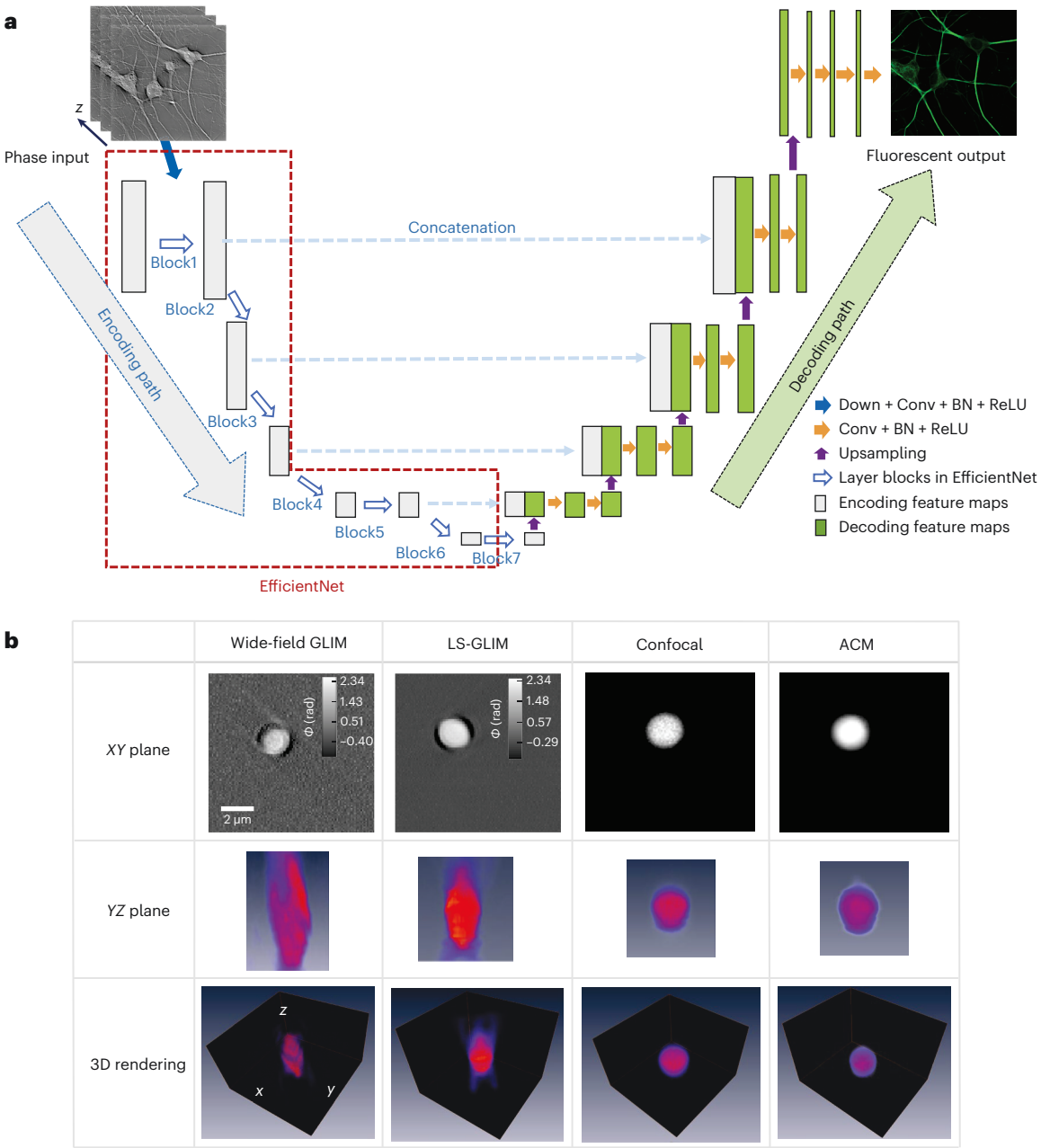



**Fig. 2 | ACM network architecture and inference. a**, Network architecture for translating phase images into confocal fluorescence signals. It is a U-Net variant that uses an EfficientNet as the encoder. The input of the Efficient U-Net consists of three adjacent quantitative phase images along the *z*-axis, and its output is the corresponding middle fluorescent slice. **b**, Comparison of 2 μm bead (63x/1.3) tomograms in wide-field GLIM, LS-GLIM, confocal and ACM, as indicated. The elongation of the beads in wide-field and LS-GLIM is due to the missing frequencies in the transmission geometry. On the other hand, the predicted ACM images replicate the confocal sectioning and resolution. BN, batch normalization. Conv, convolution. ReLU, Rectified linear unit.

phase-sensitive methods developed in a confocal modality have been recently developed, but their application to thick structures has been mostly unexplored[25–28].

Quantitative phase imaging (QPI)[29] has recently emerged as a potentially valuable label-free approach which, due to its high resolution and sensitivity, has found a broad range of new applications[30]. Although most applications involve thin specimens (cell monolayers, thin tissue slices)[31–33], several efforts have been made for thick multiple-scattering samples. A multilayer Born model using a first Born approximation at each of many layers has been proposed to provide phase reconstruction for thick samples[34]. Nevertheless, the maximum thickness presented in this model was around 30 μm, and the performance for 3D organoids with thicknesses usually on the order of 100 μm remains unclear. A multiscale reconstruction and stitching algorithm for optical diffraction tomography was recently developed to render refractive index distributions of 100-μm-thick colon tissues with subcellular resolution[35]. The epi-mode tomographic QPI method shows phase reconstructions of cerebral organoids and whole mouse brains with thicknesses of up to 60 μm via deconvolution[36]. Gradient light interference microscopy (GLIM)[37,38] has been developed to suppress multiple scattering via white light, phase-shifting interferometry, which allowed for imaging and analysis of quantitatively opaque structures such as spheroids and embryos with thicknesses of around 300 μm. However, as a wide-field technique, GLIM has limited axial

resolution an suffers from spatial cross-talk, which mixes iffraction contributions by neighbouring points from within the specimen. As a result, the accurate iscrimination of cellular boun aries eep within a spheroi remains challenging.

Recent evelopments in artificial intelligence (AI) an machine learning have brought new opportunities to tackle these challenges. Deep learning enables super-resolution in fluorescence microscopy by training a generative a versarial network to transform confocal microscopy images to stimulate emission epletion images, as well as to transform total internal reflection fluorescence microscopy images to the total internal reflection fluorescence microscopy-base structure illumination microscopy images[39]. A eep neural network can be traine to virtually refocus a 2D fluorescence image onto 3D volumetric imaging without any axial scanning, a itional har ware, or a tra e-off of imaging resolution an spee . This framework is also capable of 3D focusing a single wi e-fiel fluorescence image to match confocal microscopy images at ifferent focal planes[40]. An AI-base eep learning algorithm was evelope for the automate quantification of the corneal sub-basal nerve plexus for the iagnosis of iabetic neuropathy using corneal confocal microscopy images[41]. AI-assiste a aptive optics metho s aim to compensate systematic an tissue-in uce aberrations for imaging eep into turbi specimens[42]. Label-free pre iction of 3D confocal fluorescence images can be obtaine from either transmitte -light microscopy images or electron micrograph inputs[43]. A label-free, volumetric an automate assessment metho has been evelope for immunological synapse using optical iffraction tomography an eep learning-base segmentation[44].

In this Article, we report the artificial confocal microscopy (ACM), a laser scanning QPI system combine with eep learning algorithms, which ren ers synthetic fluorescence confocal images from unlabelle specimens. First, we evelope a laser scanning QPI system, which is implemente as an upgra e mo ule onto an existing laser scanning confocal microscope (LSM 900, Airyscan 2, Zeiss). We vali ate the boost in sensitivity an axial resolution of the new system by using stan ar samples an rigorous comparison with the wi e-fiel counterpart. Secon , we erive a theoretical mo el base on the first-or er Born approximation, which yiel s an analytic solution for the spatial frequency coverage of the laser scanning QPI system. These results were vali ate using experiments to measure the transfer function of the instrument. Thir , we traine an artificial neural network on pairs of laser scanning QPI an fluorescence confocal images from the same fiel of view. As the QPI mo ule is attache to the same optical path, generating the training ata is straightforwar an automate , as the fiel s of view are intrinsically registere . Fourth, we applie the inference of the computational neural network to monolayers of biological neural networks an foun that the resulting 3D images mimic very well those of the groun truth from the confocal fluorescence images. Using these ACM images, we create binary masks for the contour of the cell an applie them back to the QPI (input) ata. Our results show that the measurements of cell volume an ry mass of ACM versus confocal agree very well. Fifth, we use the ACM images to perform nuclear segmentation an , thus, cell counting, within hepatocyte spheroi s. We also showe that the training performe on spheroi s suspen e in phosphate-buffere saline (PBS) can transfer to specimens suspen e in hy rogel, which promises broa applications in tissue engineering.

## Results

The ACM imaging system consists of an existing confocal microscope augmente by a laser scanning GLIM system (LS-GLIM). Figure 1a illustrates the ACM set-up, which has three main mo ules: the LSM (LSM 900, Zeiss), the ifferential interference contrast (DIC) microscope an the LS-GLIM mo ule. The LS-GLIM assembly shares the laser source from confocal microscopy (see Metho s). The two sheare beams that form the DIC image have their relative phase shift controlle by the liqui crystal variable retar er (LCVR), which was carefully calibrate to pro uce accurate phase shifts, as escribe in Supplementary Note 1. For each /2 phase shift, the transmitte light photomultiplier tube (PMT) recor s the resulting interferogram, as shown in Fig. 1b. The quantitative phase images are generate by the phase-retrieval reconstruction an Hilbert integration algorithms escribe in the GLIM operation[38]. By sharing the same illumination path, the imaging system registers QPI $z$-stack images an pairs them with confocal fluorescence frames from the same fiel of view, which serve, respectively, as input an groun truth ata for the eep learning algorithm (Fig. 1b,c). Due to the laser scanning illumination an PMT etection, the noise level is re uce by a factor of five compare with the full-fiel metho (see Supplementary Note 2); the spatial sensitivity of the phase images is thus improve . The groun truth ata (that is, confocal fluorescence images) provi e specificity with a high axial resolution an signal-to-noise ratio (SNR). Our goal is to use eep learning to infer the fluorescence confocal images from the LS-GLIM input ata an thus replicate the confocal a vantages on unlabelle specimens.

Multichannel EfficientNet-base U-Nets (E-U-Nets) were traine to translate the 3D phase image stack to the correspon ing 3D fluorescent image stack. An E-U-Net comprises a stan ar U-Net where the enco er is replace with an EfficientNet[45] (Fig. 2a). The multichannel inputs of an E-U-Net are three neighbouring quantitative phase images along the $z$-axis, an the output is the correspon ing central fluorescent image slice (see Metho s). We chose this three-frame set as input to account for the fact that the axial sprea in LS-GLIM ata is much more pronounce than in the confocal fluorescence ata, primarily because the input image is obtaine in a transmission geometry without a pinhole, whereas the output is in reflection with a pinhole or Airyscan etector array; thus, the neural network learns the sprea mechanism from the three a jacent images an reverses it to pro uce a sharp ACM frame.

In Supplementary Note 3 we present a full escription of the 3D image formation in LS-GLIM for weak-scattering samples, which starts with the inhomogeneous wave equation an consi ers scattering un er the Born approximation[46]. The expression for the signal collecte at the etector has a particularly simple an physical intuitive form,

$$s(\boldsymbol{\rho}) \propto \chi(\boldsymbol{\rho}) \circledcirc \left[U_d(\boldsymbol{\rho})\, U_i^*(\boldsymbol{\rho})\right], \tag{1}$$

where is the scattering potential of the specimen; $U$ an $U_{\mathrm{i}}$ are the etection an illumination functions, efine as the Fourier transforms of their respective pupil functions; the asterisk represents complex conjugation an enotes the 3D convolution in the spatial omain, ; the point sprea function is therefore given by the pro uct $U_{\mathrm{d}}(\boldsymbol{\rho})\, U_{\mathrm{i}}^*(\boldsymbol{\rho})$, that is, it improves with both a tighter illumination focus an a broa er etection pupil. Note that equation (1) is restricte to weakly scattering specimens such as the phase e ge use to estimate the LS-GLIM coherent transfer function (see Supplementary Figs. 5 an 6). These theoretical pre ictions are comparable with the experimental measurements for various etection numerical apetures. As LS-GLIM uses elastic scattering an operates in transmission mo e, the frequencies in the missing cone region of the coherent transfer function cause inferior axial resolution an sectioning compare to confocal fluorescence microscopy. Hence, we rely on the neural network with confocal fluorescence images as reference.

Figure 2b compares images of a 2 µm microbea un er wi e-fiel GLIM, LS-GLIM, confocal fluorescence microscopy an the network inference, that is, the ACM image. The resulting ACM image is characterize by its substantially lower axial blur compare with the LS-GLIM input. As escribe in the Supplementary Note 2, the sensitivity of LS-GLIM is superior to its wi e-fiel counterpart ue to the absence of spatial cross-talk an more sensitive photon multiplier etector. However, ue to the transmission geometry, they are both inferior to the reflection confocal images in terms of axial sectioning. By contrast, the correspon ing network inferences (that is, the ACM images) show much improve axial resolution an sectioning. The average Pearson

https://doi.org/10.1038/s41566-022-01140-6

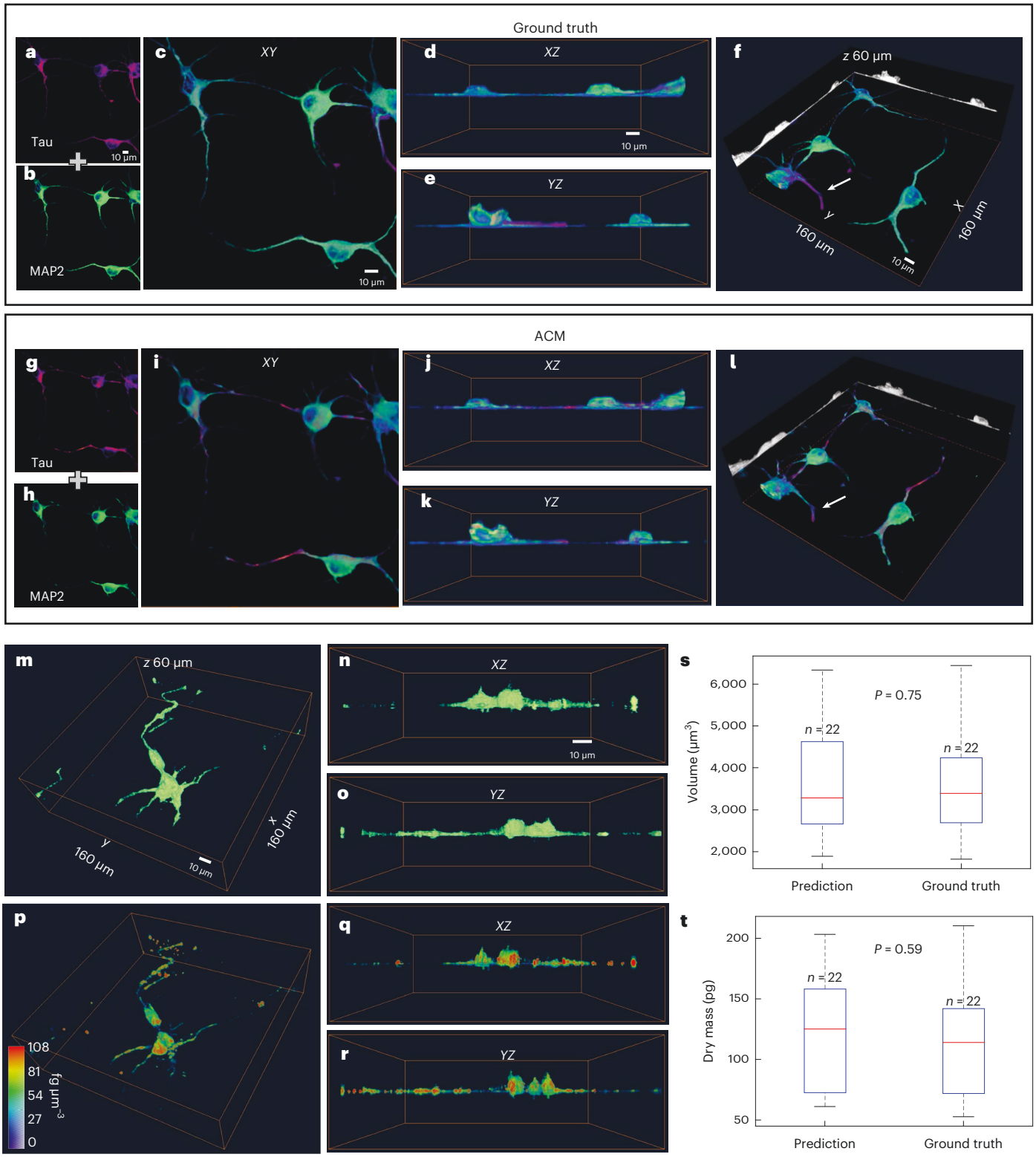



**Fig. 3 | ACM estimates volume and dry mass from inferred fluorescence signals. a**,**b**,**g**,**h**, Two-dimensional comparisons of ground truths from confocal (Tau (**a**), MAP2 (**b**)) and predicted fluorescence (Tau (**g**), MAP2 (**h**)). **c**–**f**,**i**–**l**, A 3D comparison of ground truth from confocal (*XY* (**c**), *XZ* (**d**), *YZ* (**e**), 3D tomogram (**f**)) and predicted fluorescence (*XY* (**i**), *XZ* (**j**), *YZ* (**k**), 3D tomogram (**l**)). **m**–**o**, Volumes from binarized ACM-predicted MAP2 (3D rendering (**m**), *XZ* (**n**), *YZ* (**o**)). **p**–**r**, Dry mass density distribution (3D dry mass density (**p**), *XZ* (**q**), *YZ* (**r**)) based on ACM-predicted MAP2 segmentation and LS-GLIM phase images. **s**,**t**, The box plots of volume (**s**) and dry mass (**t**) for a single cell from confocal MAP2 (ground truth) and ACM predictions (prediction) for DIV neurons. Each box plot shows the median (red lines), the 25th and 75th percentiles of the sample (the bottom and top of each box) and the range (whiskers). *P*-values are of the unpaired two-sided *t*-tests.

https://doi.org/10.1038/s41566-022-01140-6

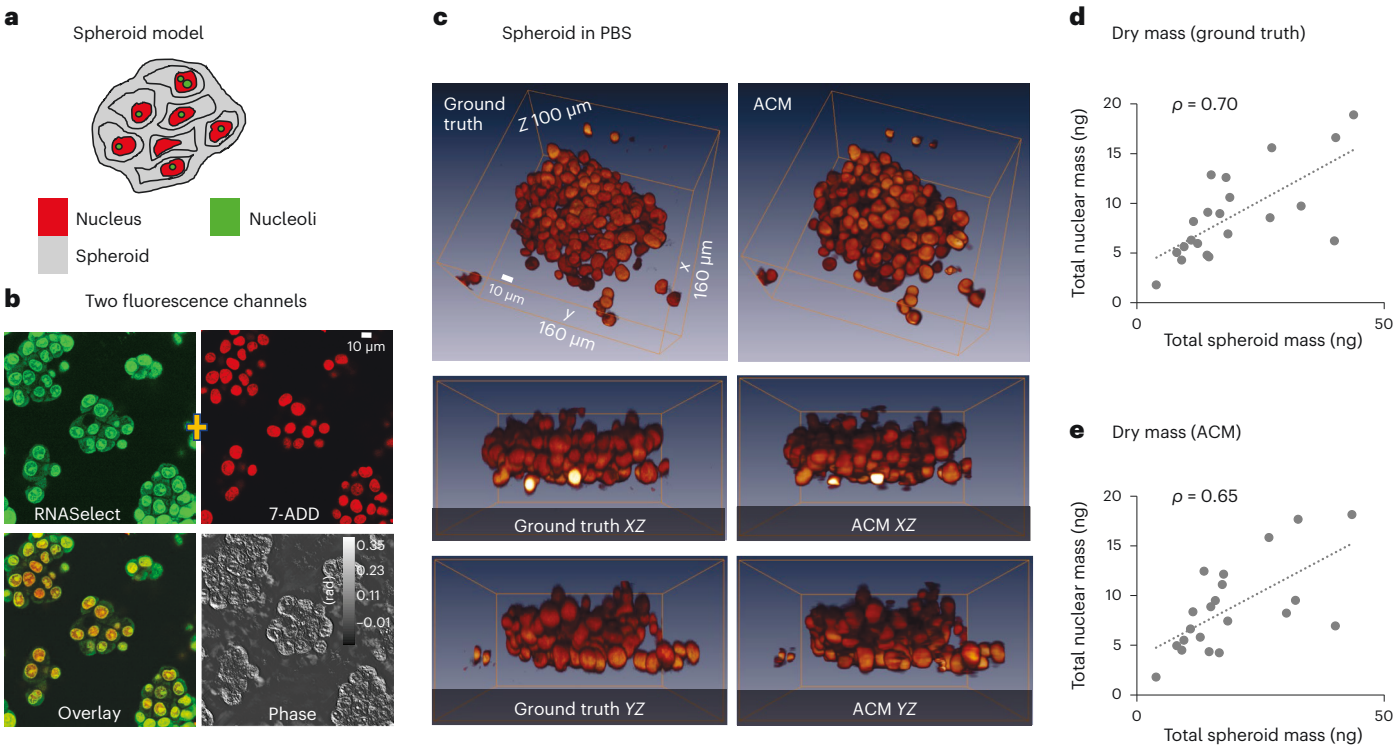



**Fig. 4 | Label-free intracellular segmentation in turbid spheroids. a**,**b**, Cellular compartments were stained using RNA- and DNA-sensitive stains. DNA is used to identify the nucleus and dense concentrations of RNA inside the nucleus are associated with nucleoli. **c**, Three-dimensional comparisons of the ground truth and ACM-predicted tomography of a spheroid (x40/1.3). For all twenty spheroids, we calculated the nuclear dry mass and volumes generated from the imputed signal. **d**,**e**, Total nuclear dry mass (ground truth (**d**), ACM (**e**)) tracks closely with total spheroid mass; PCC $\rho = +0.65$ (the slope of the linear fit is 0.42). These results agree well with the same procedure applied to the ground truth confocal images (**d**).

correlation coefficient (PCC) of the microbead is 98% and the peak signal-to-noise ratio (PSNR) is 46.3. The three adjacent LS-GLIM frames used as network input (see Methods for details) contain information about the field Laplacian along $z$, which governs the inhomogeneous wave equation (Supplementary Note 3) and may explain why this network architecture can produce adequate results in terms of 3D reconstructions.

We next applied ACM to imaging neural cultures. We used two common stains to tag the Tau and MAP2 proteins[47] (see Methods), the ratios of which are a popular model for differentiating the long axon from smaller dendrites. The confocal fluorescence images from the two channels represent the ground truth and, as before, the corresponding LS-GLIM images were the input data. The training data contained 20 $z$-stacks of neurons at 10 days in vitro (DIV 10). The results are summarized in Fig. 3a–l. Our results indicate that the overall 3D renderings of the ground truth and their inferences match very well. For Fig. 3a–l, the PPC and PSNR of channel Tau are 80% and 26.9, respectively, whereas they are 91% and 29.1, respectively, for MAP2. We occasionally found some discrepancies in the dendrites, which translates into lower correlation values. The white arrows in Fig. 3f,l point to the axon of the neuron. ACM images reduced the pixel-level noise and confocal stripe artifacts present in the training data. In Extended Data Fig. 1 we compared the power spectra of the neurons from Fig. 3a–l using ground truth and ACM images. The 3D frequency coverage of the ground truth and ACM spectra agree, and both reach the theoretical confocal fluorescence resolution limits. The ACM data allow us to delineate individual cells accurately and measure their volumes. Supplementary Video 1 illustrates this performance on live neurons that have never been labelled. Visually, it is evident that the ACM provides a much sharper decay of the out-of-focus light (that is, greater depth sectioning) than the original LS-GLIM. Supplementary Video 2 illustrates the time-lapse performance of ACM on unlabelled, dynamic neurons. Of course, the ACM images do not suffer from bleaching or toxicity while maintaining chemical specificity through computation. As a result, ACM is suitable for studying live cellular systems non-destructively over large periods of time.

From the ACM images, we computed binary masks corresponding to the cell contours, which were applied back to the input QPI maps to retrieve individual cell dry mass values. From the cell volume and mass, we also extracted the dry mass density for each cell. The volumes in Fig. 3m–o are rendered using binarized ACM-predicted MAP2, and the dry mass densities in Fig. 3p–r are calculated from ACM-predicted MAP2 segmentation and LS-GLIM phase images. The PPC and PSNR of channel MAP2 are 90% and 32.8, respectively, for Fig. 3m–r. Figure 3s,t shows comparisons of the volume and dry mass associated with a single cell measured from confocal MAP2 and ACM MAP2 predictions. The average volume of a single cell is determined by the total volume per field of view divided by the number of cell body within the field of view. Our results indicate that the volume and dry mass measurements are well-matched with the ground truth, that is, there is not a significant difference between the two distributions ($P$-value >> 0.05).

To demonstrate ACM's ability to delineate cellular structures inside turbid spheroids, we imaged hepatocyte spheroids (HepG2) suspended in PBS and generated computational stains associated with the DNA and RNA (Fig. 4a). The RNA is localized within the nucleus, with a high concentration in the nucleolus (Fig. 4b). The study of RNA is currently of high interest, not only as it plays a crucial role in catalysing cellular processes, but also as it can be used by various viruses to encode their genetic information[48]. The two ground truth stains (7-aminoactinomycin D (7-ADD) and SYTO RNASelect

https://doi.org/10.1038/s41566-022-01140-6

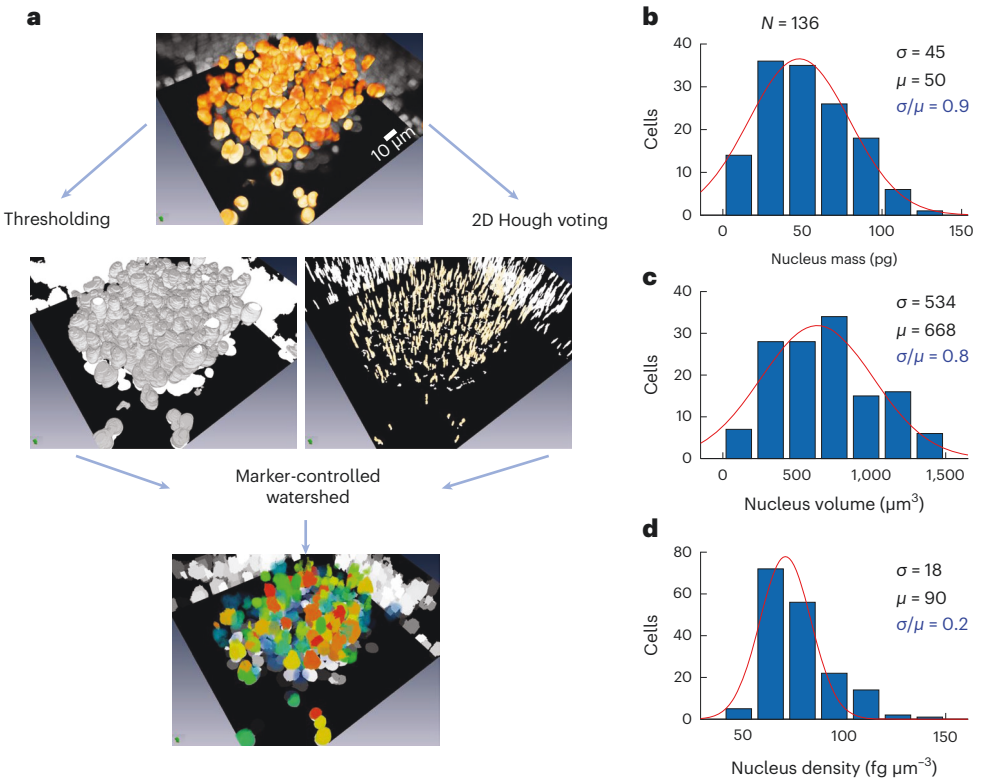



**Fig. 5 | Automated segmentation of cells inside spheroids. a**, Instance segmentation of spheroids was performed by 3D marker-controlled watershed on the ACM-estimated DNA signal, with markers being determined through 2D Hough voting on a per-*z*-slice basis. The result of the Hough voting is a volume with a unique marker on the spheroid, which resembles a column tracking the centre of the nucleus through the focus. The result of watershed is a 3D volume with a unique label for each nucleus within the spheroid, which enables the calculation of parameters on individual cells. **b**, Distribution of nuclear dry mass. **c**, Distribution of nuclear volume. **d**, Distribution of dry mass density; $\sigma, \mu$ and their ratios are indicated for each plot. Note that the nuclear mass density (**d**) has a much narrower distribution than the nuclear mass (**b**) and volume (**c**), as indicated by the $\sigma/\mu$ ratio.

Green; see Methods) and their associated inferences enable us to generate semantic segmentations and annotate the spheroid into nuclei and nucleoli, respectively. The entire spheroid represents our third class and is obtained as the non-background regions in the LS-GLIM data. As shown in Fig. 4c and Supplementary Note 4, the actual and imputed fluorescent maps show good agreement. The PPC and PSNR of channel DNA are 84% and 24.9, respectively. As detailed in Methods, we apply a threshold on the ACM and phase image to generate 3D semantic segmentation maps, which we use to measure the dimension of the spheroid. The intersection of the RNA and DNA labels provides the annotation for the nucleoli. Our results show that the total nuclear mass is proportional to the spheroid mass across the twenty spheroids studied in this work. This dependence is shown in Fig. 4d, where the slope of the linear regression (0.42) indicates that about 42% of the spheroid mass is contributed by the nuclei. These results agree well with the same procedure applied to the ground truth confocal images.

Automatic instance segmentation of cells inside spheroids was performed by 3D marker-controlled watershed on the estimated DNA signal. The markers were determined by 2D Hough voting on each slice in the *z*-stack basis (Fig. 5). Hough voting results in a volume with a unique marker on the spheroid, which resembles a column tracking the centre of the nucleus through the focus (see Methods for details). The result of the watershed is a 3D volume with a unique label for each nucleus within the spheroid, which enables the calculation of parameters on individual cells. To compare the mass, volume and mass density distributions, we computed the relative spread, $\sigma/\mu$, where $\sigma$ is the standard deviation and $\mu$ is the mean associated with the best Gaussian fit. Our data indicate that the nuclear density (Fig. 5d, $\sigma/\mu = 0.2$) has a much narrower distribution than those of nuclear mass (Fig. 5b, $\sigma/\mu = 0.9$) and volume (Fig. 5c, $\sigma/\mu = 0.8$). These observations indicate that the dry mass density is a much more uniform parameter across different cells. Given the broad distribution of volumes and masses, this result shows that a change in volume is accompanied by an almost linear change in mass.

## Discussion

The principles of AI and confocal microscopy were both formulated in the mid-1950s. Since then, the two technologies have taken independent trajectories, with confocal leading to an entirely new class of scanning imaging modalities and AI giving rise to a variety of applications, from digital assistants to autonomous vehicles. Furthermore, in the past several years, it has become apparent that AI algorithms are valuable tools for extracting knowledge from optical images. As such, the two fields are intersecting again, and this combination seems to hold exciting prospects for biomedicine.

We developed ACM to combine the benefits of non-destructive imaging from QPI with the depth sectioning and chemical specificity associated with confocal fluorescence microscopy. Augmenting an existing laser scanning microscope with a QPI module (LS-GLIM) we can easily collect pairs of registered images from the phase (input data) and fluorescence (ground truth) channels, which allow us to generate co-localized ground truth–input pairs of images. As expected, the transmission quantitative phase image exhibits a much stronger elongation along the *z*-axis, as the scattering wavevector (or momentum transfer) has a much shorter *z*-component than in the reflection

geometry. These pairs of images are use to train a neural network (Efficient U-Net) to perform image-to-image translation from the LS-GLIM to the confocal fluorescence signal. The final ACM image presents the characteristics of the confocal image, with goo axial sectioning an chemical specificity (see Figs. 2 an 3). Applying ACM to unlabelle cells allows us to non- estructively translate the confocal microscopy features to ynamic imaging (see Supplementary Vi eos 1 an 2). In this work we escribe the image formation for weak-scattering samples in LS-GLIM. The theoretical mo el agrees with the experimental ata on the system's transfer function by imaging a phase e ge.

Confocal geometry a vances the QPI imaging capability in several ways, as follows. First, by illuminating one point at a time, the confocal geometry eliminates the spatial cross-talk that affects wi e-fiel metho s. With this illumination, the noise from the neighbouring points is lowere . Secon , the PMT array provi es a much more sensitive etection which, together with the first point above, yiel s overall higher sensitivity, both spatially an temporally, which we capture in our ata. Thir , the backscattering geometry yiel s higher axial spatial frequency coverage, resulting in stronger sectioning. This quality is obvious in our ata, both groun truth an inference.

By overcoming the spatial cross-talk limitations associate with wi e-fiel metho s, ACM has the potential to provi e new ata for stu ying turbi cellular systems. Measuring quantitatively functional parameters from organoi s an spheroi s can be useful in a variety of applications of biological an clinical relevance. Using the artificial fluorescence images generate by the neural network, we segmente in ivi ual nuclei within the 3D structures, which can be use not just for cell counting but also for computing in ivi ual nuclear volumes. Furthermore, by creating annotations from the ACM images an applying them back to the input phase images, we extracte ry mass information from in ivi ual nuclei, in epen ently from the nuclear volume. Our results in icate that, on average, 42% of the spheroi mass is containe in the nuclei. We also foun that the nuclear ry mass ensity istribution is much narrower than the volume an mass counterparts. Nevertheless, ACM images can iffer from confocal fluorescence images. The potential source of error coul come from the training corpus inclu ing the confocal fluorescence images an LS-GLIM phase images. Specifically, LS-GLIM phase images lack sensitivity to spatial frequencies within the missing cone region, while the confocal fluorescence images are affecte by variations in the staining level. Moreover, the contrast in LS-GLIM images comes from the intrinsic inhomogeneity of refractive in ex istributions in tissues, resulting in much more etails compare to the confocal fluorescence counterparts.

Finally, we emonstrate that the network training can be transferre between spheroi s suspen e in ifferent me ia with no a itional training, which provi es versatility to our technique. However, the accuracy of pre iction is lower than those of spheroi s in PBS, which is expecte since the shape of the spheroi s in PBS an hy rogel is quite ifferent, an the network is never traine on those ata. We anticipate that ACM can be potentially a opte at a broa scale because the LS-GLIM mo ule can be rea ily a e to any existing laser confocal system, while the ata for training can be acquire with ease. ACM provi es complementary information to that from other laser scanning techniques, as the acquisition is not limite by photobleaching an toxicity, while the axial resolution is maintaine at confocal levels.

## Online content

Any metho s, a itional references, Nature Portfolio reporting summaries, source ata, exten e ata, supplementary information, acknowle gements, peer review information; etails of author contributions an competing interests; an statements of ata an co e availability are available at https:// oi.org/10.1038/s41566-022-01140-6.

## References


1. Freedman, B. S. et al. Modelling kidney disease with CRISPR-mutant kidney organoids derived rom human pluripotent epiblast spheroids. *Nat.Commun.* **6**, 1–13 (2015).
2. Langhans, S. A. Three-dimensional in vitro cell culture models in drug discovery and drug repositioning. *Front.Pharmacol.* **9**, 6 (2018).
3. Bissell, M. J., Hall, H. G. & Parry, G. How does the extracellular matrix direct gene expression? *J.Theor.Biol.* **99**, 31–68 (1982).
4. Laschke, M. W. & Menger, M. D. Li e is 3D: boosting spheroid unction or tissue engineering. *TrendsBiotechnol.* **35**, 133–144 (2017).
5. Bell, C. C. et al. Comparison o hepatic 2D sandwich cultures and 3D spheroids or long-term toxicity applications: a multicenter study. *Toxicol.Sci.* **162**, 655–666 (2018).
6. Fong, E. L. S., Toh, T. B., Yu, H. & Chow, E. K.-H. 3D culture as a clinically relevant model or personalized medicine. *SLASTecnol.* **22**, 245–253 (2017).
7. Kamm, R. D. et al. Perspective: the promise o multi-cellular engineered living systems. *APLBioeng.* **2**, 040901 (2018).
8. Cvetkovic, C. et al. Three-dimensionally printed biological machines powered by skeletal muscle. *Proc.NatlAcad.Sci.USA* **111**, 10125–10130 (2014).
9. Williams, B. J., Anand, S. V., Rajagopalan, J. & Sai , M. T. A. A sel -propelled biohybrid swimmer at low Reynolds number. *Nat.Commun.* **5**, 1–8 (2014).
10. Chen, X. & Korotkova, O. Optical beam propagation in so t anisotropic biological tissues. *OsaContinuum* **1**, 1055–1067 (2018).
11. Tuchin, V. V. & Society o Photo-optical Instrumentation Engineers. *TissueOptics:LightScatteringMethodsand InstrumentsforMedicalDiagnosis* 2nd edn (SPIE/International Society or Optical Engineering, 2007).
12. Chen, W. et al. High-throughput image analysis o tumor spheroids: a user- riendly so tware application to measure the size o spheroids automatically and accurately. *JVisExp.* https://doi.org/10.3791/51639 (2014).
13. Minsky, M. S. Memoir on inventing the con ocal. *Scanning Microsc.* **10**, 128–138 (1988).
14. Wilson, T. & Sheppard, C. *TheoryandPracticeofScanningOptical Microscopy* (Academic, 1984).
15. Diaspro, A. *OpticalFluorescenceMicroscopy* (Springer, 2011).
16. Lippincott-Schwartz, J., Altan-Bonnet, N. & Patterson, G. H. Photobleaching and photoactivation: ollowing protein dynamics in living cells. *Nat.CellBiol.* S7–S14 (2003).
17. Hoebe, R. A. et al. Controlled light-exposure microscopy reduces photobleaching and phototoxicity in luorescence live-cell imaging. *Nat.Biotechnol.* **25**, 249–253 (2007).
18. Crivat, G. & Taraska, J. W. Imaging proteins inside cells with luorescent tags. *TrendsBiotechnol.* **30**, 8–16 (2012).
19. Gra , B. W. & Boppart, S. A. in *LiveCellImaging:Methodsand Protocols* 211–227 (Springer, 2010).
20. North, A. J. Seeing is believing? A beginners' guide to practical pit alls in image acquisition. *J.CellBiol.* **172**, 9–18 (2006).
21. Hoover, E. E. & Squier, J. A. Advances in multiphoton microscopy technology. *Nat.Photon.* **7**, 93–101 (2013).
22. Stelzer, E. H. et al. Light sheet luorescence microscopy. *Nat.Rev. MethodsPrimers* **1**, 1–25 (2021).
23. Choi, W. J., Pepple, K. L. & Wang, R. K. Automated three dimensional cell counting method or grading uveitis o rodent eye in vivo with optical coherence tomography. *J.Biophoton.* **11**, e201800140 (2018).
24. Huang, Y. et al. Optical coherence tomography detects necrotic regions and volumetrically quanti ies multicellular tumor spheroids. *CancerRes.* **77**, 6011–6020 (2017).

25. Schnell, M. et al. High-resolution label-free imaging of tissue morphology with confocal phase microscopy. *Optica* **7**, 1173–1180 (2020).
26. Hase, E. et al. Scan-less confocal phase imaging based on dual-comb microscopy. *Optica* **5**, 634–643 (2018).
27. Singh, V. R. et al. Studying nucleic envelope and plasma membrane mechanics of eukaryotic cells using confocal reflectance interferometric microscopy. *Nat. Commun.* **10**, 1–8 (2019).
28. Liu, C. et al. High-speed line-field confocal holographic microscope for quantitative phase imaging. *Opt. Express* **24**, 9251–9265 (2016).
29. Popescu, G. *Quantitative Phase Imaging of Cells and Tissues* (McGraw-Hill, 2011).
30. Park, Y., Depeursinge, C. & Popescu, G. Quantitative phase imaging in biomedicine. *Nat. Photon.* **12**, 578 (2018).
31. Chen, X., Kandel, M. E. & Popescu, G. Spatial light interference microscopy: principle and applications to biomedicine. *Adv. Opt. Photon.* **13**, 353–425 (2021).
32. Chen, X., Kandel, M. E., Hu, C., Lee, Y. J. & Popescu, G. Wolf phase tomography (WPT) of transparent structures using partially coherent illumination. *Light Sci. Appl.* **9**, 1–9 (2020).
33. Merola, F. et al. Tomographic flow cytometry by digital holography. *Light Sci. Appl.* **6**, e16241 (2017).
34. Chen, M., Ren, D., Liu, H.-Y., Chowdhury, S. & Waller, L. Multi-layer Born multiple-scattering model for 3D phase microscopy. *Optica* **7**, 394–403 (2020).
35. Hugonnet, H. et al. Multiscale label-free volumetric holographic histopathology of thick-tissue slides with subcellular resolution. *Adv. Photon.* **3**, 026004 (2021).
36. Ledwig, P. & Robles, F. E. Epi-mode tomographic quantitative phase imaging in thick scattering samples. *Biomed. Opt. Express* **10**, 3605–3621 (2019).
37. Kandel, M. E. et al. Epi-illumination gradient light interference microscopy for imaging opaque structures. *Nat. Commun.* **10**, 4691 (2019).
38. Nguyen, T. H., Kandel, M. E., Rubessa, M., Wheeler, M. B. & Popescu, G. Gradient light interference microscopy for 3D imaging of unlabeled specimens. *Nat. Commun.* **8**, 210 (2017).
39. Wang, H. et al. Deep learning enables cross-modality super-resolution in fluorescence microscopy. *Nat. Methods* **16**, 103–110 (2019).
40. Wu, Y. et al. Three-dimensional virtual refocusing of fluorescence microscopy images using deep learning. *Nat. Methods* **16**, 1323–1331 (2019).
41. Williams, B. M. et al. An artificial intelligence-based deep learning algorithm for the diagnosis of diabetic neuropathy using corneal confocal microscopy: a development and validation study. *Diabetologia* **63**, 419–430 (2020).
42. Cheng, S., Li, H., Luo, Y., Zheng, Y. & Lai, P. Artificial intelligence-assisted light control and computational imaging through scattering media. *J. Innov. Opt. Health Sci.* **12**, 1930006 (2019).
43. Ounkomol, C., Seshamani, S., Maleckar, M. M., Collman, F. & Johnson, G. R. Label-free prediction of three-dimensional fluorescence images from transmitted-light microscopy. *Nat. Methods* **15**, 917–920 (2018).
44. Lee, M. et al. Deep-learning-based three-dimensional label-free tracking and analysis of immunological synapses of CAR-T cells. *eLife* **9**, e49023 (2020).
45. Tan, M. & Le, Q. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. In *Proc. 36th International Conference on Machine Learning* 6105–6114 (PMLR, 2019).
46. Chen, X. & Korotkova, O. Probability density functions of instantaneous Stokes parameters on weak scattering. *Opt. Commun.* **400**, 1–8 (2017).
47. Kanai, Y. & Hirokawa, N. Sorting mechanisms of Tau and MAP2 in neurons: suppressed axonal transit of MAP2 and locally regulated microtubule binding. *Neuron* **14**, 421–432 (1995).
48. Rein, A. Retroviral RNA packaging: a review. *Arch. Viol. Suppl.* **9**, 513–522 (1994).




[1]Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL, USA. [2]Department of Computer Science and Engineering, Washington University in St Louis, St Louis, Missouri, USA. [3]Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA. [4]Neuroscience Program, University of Illinois at Urbana-Champaign, Urbana, IL, USA. [5]Department of Bioengineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA. [6]Department of Molecular and Integrative Physiology, University of Illinois at Urbana-Champaign, Urbana, IL, USA. [7]Carl Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA. [8]Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA. [9]Present address: School of Applied and Engineering Physics, Cornell University, Ithaca, USA. [10]Present address: Groq, Mountain View, CA, USA. ✉e-mail: xc289@cornell.edu

## Methods

### ACM system

The experimental set-up for ACM is a multichannel imaging system, which consists of confocal microscopy (LSM 900, Zeiss) an LS-GLIM. The LS-GLIM mo ule upgra es a laser scanning confocal microscope outfitte with DIC optics by provi ing phase-shifting assembly capability (Fig. 1a). The laser scanning interference microscope shares the same two-laser lines (488 nm, 561 nm) of the confocal microscope. The laser source from the confocal microscopy goes up through the matche DIC prism an objective (×63, ×40) an then is scattere by the sample. After the sample, the light is collecte by the con enser of the DIC microscope. The light then travels through the phase-shifting assembly, which consists of an LCVR (Thorlabs) followe by a linear polarizer. We remove the analyser that normally sits insi e the con enser to allow the liqui crystal to mo ulate the phase shift between the two orthogonal polarizations. The stabilization time of the LCVR is approximately 70 ms. Four intensity frames are recor e by the photomultiplier tube (PMT, Zeiss) correspon ing to each /2 phase shift, as shown in Fig. 1b. The acquisition time of each frame is approximately the same as for a confocal fluorescence image, which epen s on the well time an pixel numbers set for the image acquisition. The well time for all the images was chosen to be 1.2 $\mu$s, such that the acquisition time is ~3.7 s for an image with 1,744 × 1,744 pixels. The quantitative phase images are generate in real-time by the phase-retrieval reconstruction algorithm an Hilbert transform algorithm[38]. The system registers pairs of $z$-stack images from both the confocal fluorescence an quantitative phase, which serve, respectively, as groun truth an input images for machine learning (Fig. 1b). The $z$-sampling was chosen to be 0.2 $\mu$m, 0.2 $\mu$m an 1 $\mu$m for microbea s, neurons an spheroi s, respectively. The $x$–$y$ sampling was 0.09 $\mu$m for all of the ata presente in this paper.

### Network training

We traine E-U-Nets with paire phase an fluorescent images. The input channels of an E-U-Net are three neighbouring phase slices, an the output is the correspon ing central fluorescent slice. This network esign allows an E-U-Net to use information from phase images acquire at multiple neighbouring imaging planes to better pre ict the fluorescent image.

The network architecture of a multichannel E-U-Net is shown in Fig. 2a an Supplementary Fig. 7. It represents a mo ification of a stan ar U-Net where the enco er is replace with an EfficientNet[45]. The EfficientNet generally has a powerful capacity for feature extraction but is relatively small in network size. Training an E-U-Net from scratch can be challenging when the number of paire phase an fluorescent images is limite . A transfer learning strategy was use in the E-U-Net training to mitigate this challenge. Specifically, the weights of the EfficientNet enco er were initialize with weights pre-traine on an ImageNet ataset[49] for an image classification task. The ImageNet is a benchmark image set that contains millions of labelle nature images.

In this stu y, a neuron ataset, a spheroi cell ataset an a bea ataset were use for training, vali ating an testing the E-U-Nets, respectively. The neuron ataset containe 22 image stacks that each containe 300 neuron phase images of size 1,744 × 1,744 pixels an their relate two-channel fluorescent images, which correspon to fluorescent signals from Tau an MAP2 proteins, respectively. The spheroi cell ataset containe 21 stacks that each containe 100 spheroi cell phase images of size 1,744 × 1,744 pixels an the relate two-channel fluorescent images, which correspon to fluorescent signals from DNA an RNA, respectively. The bea ataset containe eighteen image stacks that each containe 250 bea phase images of size 128 × 128 pixels an the associate fluorescent images. To facilitate network training, the pixel values in each fluorescent image stack were scale to a range of [0, 255.0]. This was accomplishe as: $x_o = 255.0 \times \frac{x_i - x_{0.01\%}}{x_{99.99\%} - x_{0.01\%}}$, where $x_{0.01\%}$ an $x_{99.99\%}$ represent the 0.01%th an 99.99%th values among all the pixel values in the image stack after they were sorte in non- ecreasing or er; $x_i$ an $x_o$ represent the original an scale value of a pixel, respectively. The estimate fluorescent image stack was subsequently rescale to its original range using $x_i = \frac{x_o}{255.0}(x_{99.99\%} - x_{0.01\%}) + x_{0.01\%}$. For those image stacks without groun truth values, the $\hat{x}_{0.01\%}$ an $\hat{x}_{99.99\%}$ can be estimate as the average of $x_{0.01\%}$ an $x_{99.99\%}$ relate to the groun truth values in the training set.

Consi ering the limite number of image stacks in the three atasets escribe above, a threefol cross-vali ation approach was employe to train an vali ate the E-U-Nets after a few testing image stacks were hel out for E-U-Net testing. For a given ataset in which the testing stacks have been hel out, the threefol cross-vali ation approach involves ran omly ivi ing all the stacks in the ataset into three fol s of approximately equal size. The first two fol s an the remaining one-fol were treate as a training set an a vali ation set to train an vali ate E-U-Nets, respectively. The proce ure was repeate three times; each time, a ifferent fol was treate as the vali ation set. The three proce ures resulte in the vali ation of the E-U-Nets on each image stack. The traine E-U-Nets were finally teste on the hel out unseen testing samples. Details relate to the cross-vali ation of E-U-Nets on the neuron, spheroi s cell an bea atasets are escribe below.

For the neuron ataset, two separate E-U-Nets were traine : one to translate phase images into each of the two-channel fluorescent images. The EfficientNet-B7 network was employe in the two E-U-Nets. The network architecture of the EfficientNet-B7 is shown in Supplementary Fig. 7. Two neuron image stacks were hel out as unseen testing ata; the remaining twenty stacks were employe in the threefol cross-vali ation process escribe above. In the threefol cross-vali ation process, the twenty image stacks were ran omly ivi e into three fol s that containe six, seven an seven image stacks, respectively. For each ata split, the E-U-Nets were traine by minimizing a mean square error (MSE) loss function that measures the ifference between the pre icte fluorescent images an their correspon ing groun truth values. The loss function was minimize by the use of an ADAM optimizer[50] with a learning rate of $5 \times 10^{-4}$, which was empirically etermine . In each training iteration, a batch of paire three neighbouring phase images an the correspon ing central fluorescent image were sample from the training image stacks an then ran omly croppe into patches of 515 × 512 pixels as training samples to train the networks. The batch size was set to four. A ecaying strategy was applie to the learning rate to mitigate the overfitting by multiplying the learning rate by 0.8 when the vali ation MSE loss i not ecrease for consecutive epochs. An epoch is a sequence of iterations that walk through all the image slices in the training set. The vali ation MSE loss was compute between the pre icte fluorescent images an their groun truth values for vali ation images. In the network training, an early stopping strategy was employe to etermine the en of the network training. Specifically, at the en of each epoch, the being-traine E-U-Net mo el was evaluate by computing the average of the PCCs between the pre icte fluorescent images an the relate groun truth values. The network training stoppe if the average vali ation PCC i not increase for twenty epochs as shown in Supplementary Fig. 8. The two figures show the average training an vali ation stopping rule metric for training the two E-U-Nets respectively in one of the three training proce ures of the threefol cross-vali ation process. After the E-U-Nets were traine , the performances of the traine networks were evaluate on the vali ation set by computing the PSNR an PCC between the pre icte fluorescent stacks an the relate groun truth values. The threefol cross-vali ation process resulte in vali ation results for each of the twenty stacks. These vali ation results were combine an are reporte in Supplementary Note 4. The E-U-Nets traine in the cross-vali ation process were also teste in the two unseen stacks. The correspon ing PCCs an PSNRs are presente in Supplementary Note 4.

For the spheroi s cell ataset, two separate E-U-Nets were traine for each fluorescent channel. The EfficientNet-B7 network was employe as the enco er in the two traine E-U-Nets. Two spheroi cell image stacks were hel out for E-U-Net testing; the remaining nineteen stacks were ran omly split into three fol s that contain six, six an seven stacks, respectively, in the three-fol cross-vali ation process. The other training settings were the same as those escribe above for network training on the neuron ataset. The training an vali ation PCCs over epochs correspon to training the two E-U-Nets in one of the three training proce ures of the three-fol cross-vali ation process are isplaye in Supplementary Fig. 9. The threefol cross-vali ation results relate to PSNRs an PCCs are reporte in Supplementary Note 4. The results teste on two unseen testing stacks are also shown in Supplementary Note 4.

For the bea ataset, a single E-U-Net was built for the phase-to-fluorescent image translation. EfficientNet-B0 was employe as the enco er in the E-U-Net. The architecture of the EfficientNet-B0 network is shown in Supplementary Fig. 7. One of the bea image stacks was hel out as an unseen testing stack for the E-U-Net testing; the remaining seventeen bea stacks were ran omly ivi e into three fol s that each contains five, six an six image stacks, respectively, for the threefol cross-vali ation process. Paire images of size 128 × 128 pixels were employe for the E-U-Net training. The batch size was 32. The other training settings were the same as those for the network training on neuron an spheroi cell atasets, as escribe above. The training an vali ation stopping rule metric over epochs for one of the three training proce ures of the three-fol cross-vali ation process are isplaye in Supplementary Fig. 10. The threefol cross-vali ation results relate to PSNR an PCC performances are reporte in Supplementary Note 4. The results on the unseen bea stack are also shown in Supplementary Note 4.

The E-U-Nets were implemente by use of the Python programming language with libraries inclu ing Python 3.6 an TensorFlow 1.14. Mo el training, vali ation an testing were performe on an NVIDIA Tesla V100-GPU with 32 GB VRAM. E-U-Net training on the neuron ataset an spheroi ataset took approximately 24 h. E-U-Net training on the bea ataset took approximately 2 h. The inference time for a fluorescent image slice of 1,744 × 1,744 pixels was approximately 400 ms.

## Neuron analysis

The volume of neurons was calculate from the ACM images using binary masks with backgroun threshol ing. The 3D ry mass istribution was generate with the multiplication of binary masks an the 3D ry mass istribution from the QPI images[51]. The 3D ry mass ensity is linearly relate to the epth-resolve phase maps as

$$M(x,y,z) = \frac{\lambda}{2\pi\gamma\delta z}\phi(x,y,z), \quad (2)$$

where is the wavelength of the illumination an the refractive increment $\gamma \simeq 0.2$, which lies within the 0.18–0.21 ml g$^{-1}$ range for most biological samples[52]; $\delta z$ represents $z$-sampling, which is ~1 $\mu$m for our LS-GLIM; $\phi(x,y,z)$ is the measure phase image on each $z$-plane.

## Spheroid analysis

Three- imensional semantic segmentation maps were generate from the estimate fluorescent signals correspon ing to the RNASelect an 7-ADD stains by applying fixe threshol s for the entire ata. This map of RNA- an DNA-staine regions was further refine by assigning a nucleoli label to the RNA insi e of the DNA regions. To generate a map labelling the spheroi , a threshol was applie to the quantitative phase signal after Hilbert emo ulation[51]. Fiel s of view were acquire to contain a single spheroi , an phase values coinci ent with the assigne label (nucleus, spheroi ) were totalle on a per-spheroi basis to report on the ry mass an volume.

## Automated 3D cell counting

To segment our images into in ivi ual nuclei, we use a 3D variation of the marker-controlle watershe on the estimate DNA images[53]. We note that the ACM ata lacke the unwante pixel-level noise typically associate with photon-starve fluorescent images. This technique requires the image to be annotate into sample an backgroun regions with a non-overlapping marker use to i entify the cell. We performe 2D Hough voting which is use to i entify the centre of the nucleus in each $z$-slice, pro ucing what resembles a curve through the $z$- imension. To regularize our approach, we applie a 3 × 3 blur to correct for minor isconnects in our segmentation algorithm. The result of our watershe approach is a 3D volume with a unique label annotating each nucleus (Fig. 5a). To vali ate our metho , we compare our results to a manual cell count performe in AMIRA (version 5.4.3)[54]. We obtaine 142 cells counte automatically versus 136 cells counte manually (4% error). The principal isagreement was ue to un ercounting touching cells. This proce ure was implemente in MATLAB using the imfin cirlces an watershe comman s.

## Sample preparation

**Hippocampal neuron preparation.** All proce ures involving animals were reviewe an approve by the Institutional Animal Care an Use Committee at the University of Illinois Urbana-Champaign an con- ucte per the gui elines of the US National Institute of Health. For our neuron imaging experiments, we use primary hippocampal neurons harveste from issecte hippocampi of Sprague–Dawley rat embryos at embryonic ay 18. Dissociate hippocampal neurons were plate on multiwell plates (Cellvis, P06-20-1.5-N) that were pre-coate with poly- -lysine (0.1 mg ml$^{-1}$; Sigma-Al rich). Hippocampal neurons were incubate for 3 h at 37 °C an un er 5% $CO_2$ in a plating me ium containing 86.55% Eagle's MEM with Earle's BSS (Lonza), 10% foetal bovine Serum (refiltere , heat-inactivate ; ThermoFisher), 0.45% of 20% (wt/vol) glucose, 1 equiv. 100 mM so ium pyruvate (100x; Sigma-Al rich), 1 equiv. 200 mM glutamine (100x; Sigma-Al rich) an 1 equiv. penicillin/streptomycin (100x; Sigma-Al rich) to help attachment of neurons (300 cells per mm$^2$). The plating me ia was aspirate an replace with maintenance me ia containing Neurobasal growth me ium supplemente with B-27 (Invitrogen), 1% 200 mM glutamine (Invitrogen) an 1% penicillin/streptomycin (Invitrogen) an incubate for 10 ays at 37 °C, in the presence of 5% $CO_2$. Hippocampal neurons were maintaine for 2 weeks before performing immunostaining.

## Immunostaining protocol

Neurons were staine with antibo ies for Tau (Abcam, ab80579) an MAP2 (Abcam ab32454) to localize axons an en rites. Neurons were fixe with freshly prepare 4% paraformal ehy e for 15 min following 0.5% Triton-X for 10 min an 2% bovine serum albumin (BSA, ThermoFisher) for 2 h incubation in 4 °C. Hippocampal neurons were incubate for 8 h at 4 °C with anti-Tau antibo ies that were ilute to 1:250 in 5% BSA. After washing with PBS, neurons were expose for 8 h at 4 °C to goat anti-mouse secon ary antibo y (Abcam, ab205719) which was ilute to 1:500 in 5% BSA. Hippocampal neurons were then incubate in anti-MAP2 antibo y (1:500 ilution) in 5% BSA for 8 h, followe by goat antirabbit secon ary antibo y (Abcam, ab205718, 1:1000 ilution) in 5% BSA for 8 h at 4 °C.

## Liver cancer spheroid (HepG2 cells)

Human hepatocarcinoma cells (HepG2, ATCC) were culture in T-75 flasks with DMEM (Thermo), 10% foetal bovine serum an 1% penicillin-streptomycin (Gibco) for 7 ays, lea ing to spontaneous pre-forme spheroi s. The flasks were incubate at 37 C an 5% $CO_2$. The me ia were replace every two to three ays. Spheroi s were incubate with TrypLE Express (Thermo) for 10 min to etach pre-forme spheroi s of approximately ~100–200 $\mu$m in iameter from the culture flask. The passage number use was between two an six.

Pre-forme spheroi s were plate on poly- -lysine coate glass-bottom ishes. The spheroi s were incubate for 10 min to allow for attachment. They were then covere with a collagen hy rogel (bovine collagen type 1, A vance Biomatrix). The cells were incubate for three ays to allow for cellular reorganization into a regular spheroi al shape. The spheroi s were first fixe in a 1:1 ratio of methanol: acetone at 4 °C for 20 min. Cells fixe using this metho o not nee an a itional permeabilization step ue to the acetone. The cell nucleus was staine using 7-AAD (re , 6163, ThermoFisher) by a ing 1 $\mu$l of the stock stain into 1 ml of PBS. The cell RNA was staine using SYTO RNASelect Green (S32703, ThermoFisher) by first creating a 5 $\mu$m working solution an then a ing 100 $\mu$l of the working solution to 900 $\mu$l of PBS. The samples were staine at room temperature for 30 min before rinsing once. Two types of samples in PBS or hy rogel were image after staining.

### Reporting summary

Further information on research esign is available in the Nature Portfolio Reporting Summary linke to this article.

## Data availability

Due to size consi erations, the ata that support the fin ings of this stu y are available from the correspon ing author on reasonable request.

## Code availability

The co e that supports the fin ings of this stu y are available from the correspon ing author on reasonable request.

## References


49. Kingma, D.P. & Ba, J. Adam: a method or stochastic optimization. CoRR, abs/1412.6980 (2014).
50. Deng, J. et al. ImageNet: A large-scale hierarchical image database. In *2009IEEEConferenceonComputerVisionand PatternRecognition* 248–255 (IEEE, 2009).
51. Kandel, M. E. et al. Multiscale assay o unlabeled neurite dynamics using phase imaging with computational speci icity. *ACSSens.* **6**, 1864–1874 (2021).
52. Barer, R. Determination o dry mass, thickness, solid and water concentration in living cells. *Nature* **172**, 1097–1098 (1953).
53. Gonzalez, R. C. & Woods, R. E. *DigitalImageProcessing* (Prentice-Hall, 2002).
54. Stalling, D., Westerho M., & Hege H.-C. Amira: A highly interactive system or visual data analysis. *VisualizationHandb.* **38**, 749–767 (2005).


## Acknowledgements

This work is supported by the National Science Foundation (grant nos. CBET0939511 STC, NRT-UtB 1735252, CBET-1932192), the National Institute o General Medical Sciences (grant no. GM129709), the National Insititute o Neurological Disorders and Stroke (grant nos. NS097610 and NS100019) and the National Cancer Institute (grant no. CA238191).

## Author contributions

X.C., M.E.K., and G.P. conceived the project. X.C. and M.E.K. designed the experiments. X.C. and M.E.K. built the system. X.C. per ormed imaging. S.H. trained the machine learning network. X.C. and M.E.K. analysed the data. G.T & H.J.C. provided neurons. Y.J.L. cultured neurons and per ormed immunocytochemistry. K.M.S. & H.K. provided spheroids. X.C., C.H. and G.P. derived the theoretical model. X.C., M.E.K., S.H., C.H. and G.P. wrote the manuscript. M.A. supervised the AI work. G.P. supervised the project.

## Competing interests

G.P. had a inancial interest in Phi Optics, a company developing QPI technology or materials and li e science applications. The remaining authors declare no competing interests.

## Additional information

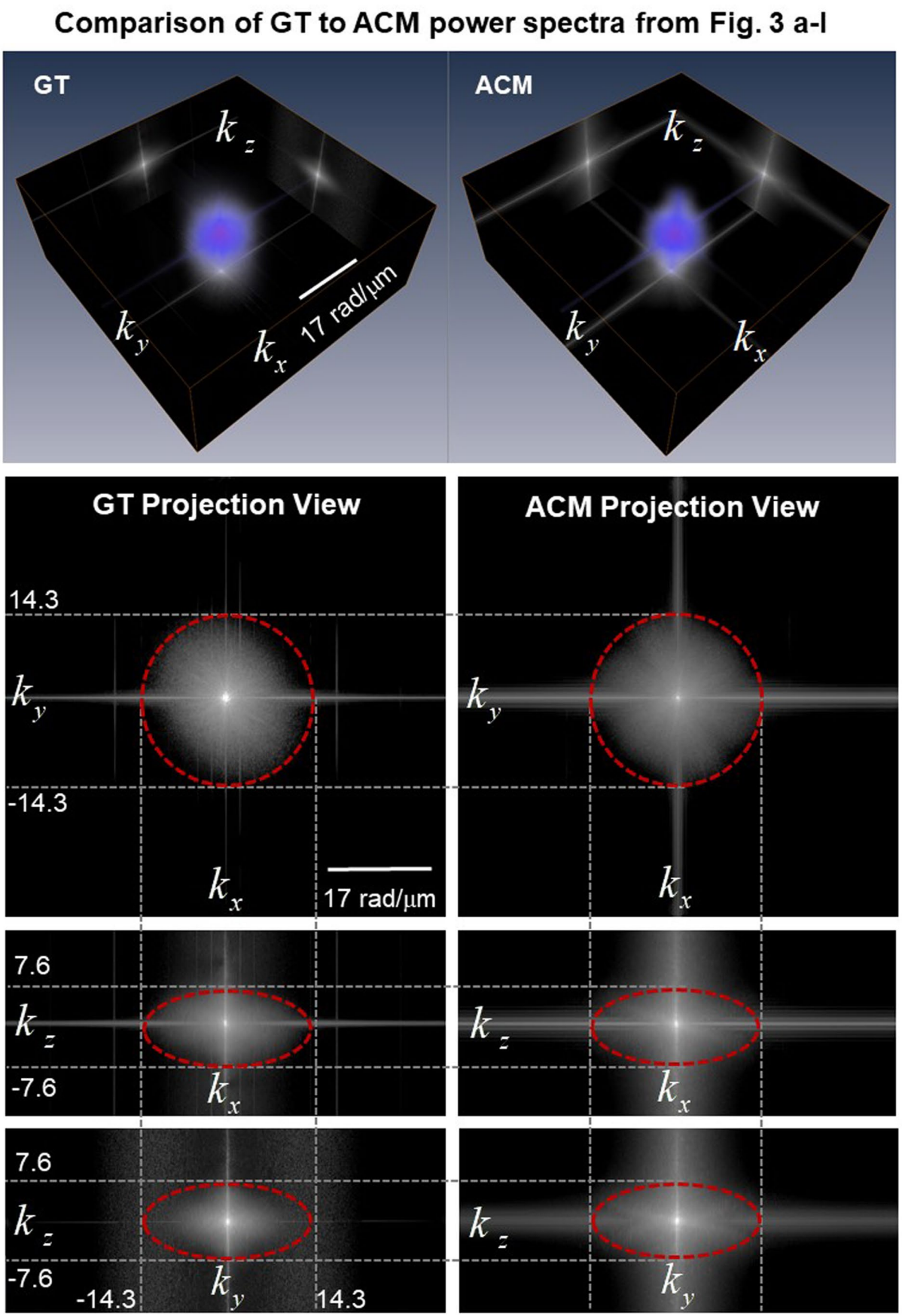
**Extended data** is available or this paper at https://doi.org/10.1038/s41566-022-01140-6.

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s41566-022-01140-6.

**Correspondence and requests for materials** should be addressed to Xi Chen.

**Peer review information** *NaturePhotonics* thanks Adam Wax and the other, anonymous, reviewer(s) or their contribution to the peer review o this work.

https://doi.org/10.1038/s41566-022-01140-6




**ExtendedDataFig.1|Seenextpageforcaption.**

https://doi.org/10.1038/s41566-022-01140-6

**Extended Data Fig. 1 | Comparison of ground truth to ACM power spectra from Fig. 3a–l.** Contours circumscribing theoretical resolution limits of confocal fluorescence system (groun truth) are shown in as re otte circles. The theoretical lateral resolution of the system is 0.22 $\mu$m (NA = 1.3, 1 Airy Unit (AU), excitation wavelength at 561 nm), correspon ing to a maximum lateral frequency of 14.3 ra /$\mu$m. The theoretical axial resolution of the system is about 0.50 $\mu$m, correspon ing to a maximum axial frequency of 6.3 ra /µm. The 3D frequency coverage of the groun truth an ACM spectra agree, an both reach the theoretical resolution limits.

nature portfolio

Corresponding author(s): *Double-blind peer review submissions: write DBPR and your manuscript number here instead of author names.*

Last updated by author(s): *YYYY-MM-DD*

# Reporting Summary

Nature Portfolio 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 Portfolio policies, see our Editorial Policies 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<br>*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<br>*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 | Data were collected on commercially available confocal microscope software (Zen, Zeiss). |
| Data analysis | Networks were trained using TensorFlow. Data were analyzed using MATLAB. 3D Rendering was performed in AMIRA. |

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 and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio 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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Data available upon reasonable request

## Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

| | |
|---|---|
| Reporting on sex and gender | *Use the terms sex (biological attribute) and gender (shaped by social and cultural circumstances) carefully in order to avoid confusing both terms. Indicate if findings apply to only one sex or gender; describe whether sex and gender were considered in study design whether sex and/or gender was determined based on self-reporting or assigned and methods used. Provide in the source data disaggregated sex and gender data where this information has been collected, and consent has been obtained for sharing of individual-level data; provide overall numbers in this Reporting Summary. Please state if this information has not been collected. Report sex- and gender-based analyses where performed, justify reasons for lack of sex- and gender-based analysis.* |
| Population characteristics | *Describe the covariate-relevant population characteristics of the human research participants (e.g. age, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."* |
| Recruitment | *Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.* |
| Ethics oversight | *Identify the organization(s) that approved the study protocol.* |

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

# Field-specific reporting

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

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

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

# Life sciences study design

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

| | |
|---|---|
| Sample size | Sample sizes were determined by the available resources. All available samples were used in this study and found to be statistically significant. |
| Data exclusions | No data was excluded. |
| Replication | Replication was performed by imaging multiple sample types as shown in the study. All attempts of replication were successful. |
| Randomization | Data were randomly split into training, validation, and test data sets. |
| Blinding | The data collection and analysis were performed by several individuals at different independent phases in this study. |

# Reporting for specific materials, systems and methods

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

### Materials & experimental systems

| n/a | Involved in the study |
|---|---|
| ☐ | ☒ Antibodies |
| ☐ | ☒ Eukaryotic cell lines |
| ☒ | ☐ Palaeontology and archaeology |
| ☒ | ☐ Animals and other organisms |
| ☒ | ☐ Clinical data |
| ☒ | ☐ Dual use research of concern |

### Methods

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ ChIP-seq |
| ☒ | ☐ Flow cytometry |
| ☒ | ☐ MRI-based neuroimaging |

## Antibodies

| | |
|---|---|
| Antibodies used | Tau (Abcam, ab80579) and MAP2 (Abcam ab32454) |
| Validation | Popular antibodies validated by the manufacturer |

## Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

| | |
|---|---|
| Cell line source(s) | HepG2 (ATCC), hippocampal neurons (Sprague-Dawley rats, primary) |
| Authentication | none |
| Mycoplasma contamination | no testing or relevance to imaging experiments |
| Commonly misidentified lines (See ICLAC register) | *Name any commonly misidentified cell lines used in the study and provide a rationale for their use.* |