Learned deconvolution using physics priors for light-sheet microscopy with propagation-invariant beams.

Philip Wijesinghe[1], Stella Corsetti[1], Darren J.X. Chow[2,3,4], Shuzo Sakata[5], Kylie R. Dunning[2,3,4], and Kishan Dholakia[1,6,7]

[1]SUPA, School of Physics and Astronomy, University of St Andrews, North Haugh, St Andrews, Fife, KY16 9SS, UK
[2]Robinson Research Institute, Adelaide Medical School, The University of Adelaide, Adelaide, SA, Australia
[3]Australian Research Council Centre of Excellence for Nanoscale BioPhotonics, The University of Adelaide, Adelaide, SA, Australia
[4]Institute for Photonics and Advanced Sensing, The University of Adelaide, Adelaide, SA, Australia
[5]Strathclyde Institute of Pharmacy and Biomedical Sciences, University of Strathclyde, Glasgow, G4 0RE, UK
[6]School of Biological Sciences, The University of Adelaide, Adelaide, South Australia, Australia
[7]Department of Physics, College of Science, Yonsei University, Seoul 03722, South Korea

Deconvolution is a challenging inverse problem, particularly in techniques that employ complex engineered point-spread functions, such as microscopy with propagation-invariant beams. Here, we present a deep learning method that uses known physics priors, in lieu of any ground truths, to rapidly and robustly deconvolve and super-resolve these microscopy images. Specifically, we constrain a generative adversarial network with images generated with the known point-spread function of the system, and combine this with unpaired experimental data that preserves perceptual content. Our network overcomes a major present barrier to deep learning adoption, namely, the need for 1,000–10,000s of paired training images. It is experimentally unsupervised and can be trained solely on a few hundred regions of interest. We demonstrate its performance on light-sheet microscopy with propagation-invariant Airy beams, including in calibration beads, oocytes, preimplantation embryos, and excised brain tissue, as well as illustrate its utility for Bessel-beam LSM. This method aims to democratise learned methods for deconvolution, as it does not require data acquisition outwith the conventional imaging protocol.

deep learning | deconvolution | physics-informed learning | light-sheet microscopy | propagation-invariant beams | Airy beam

Correspondence: ***pw64@st-andrews.ac.uk; kd1@st-andrews.ac.uk***

Introduction

Deconvolution is a well-known problem in many fields of engineering and science (1–3). Specifically in computational imaging, the problem of deconvolution is one of the accurate reconstruction of the imaged object from data encoded by the imaging system. This encoding, for instance through the use of structured light fields and engineered point-spread functions (PSF), can convey much greater information about the imaged object than is possible with conventional imaging (4). Propagation-invariant beams, in particular, have extended the field of view and resolution (5, 6), and have demonstrated deeper penetration into scattering tissues (7, 8). These methods have pushed the envelope of many important trade-offs, such as those between the resolution, contrast, field of view, imaging speed and photo-damage.

Specifically, Airy and Bessel beams have enabled rapid multiscale imaging at exceptional resolution in light-sheet microscopy (LSM) (5, 9). These beam shapes can be generated using accessible optical components and implemented in simple optical geometries (5, 10, 11). Despite their superior performance (5, 9), they are yet to be broadly adopted by the microscopy community due to the practical challenges of deconvolution. This is because these PSFs are often burdened with side-lobe structures, which challenge contemporary deconvolution algorithms, especially in the presence of poor signal-to-noise (SNR), scattering and speckle (3). Further, deconvolution can often be time demanding and computationally intensive.

To overcome these major barriers, in this paper, we introduce deconvolution based on deep learning that is capable of robustly and rapidly deconvolving and super-resolving microscopy images with arbitrary PSFs. Recently, deep learning has emerged as a method to solve a variety of inverse problems in microscopy (4, 12). Deep learning has enabled super-resolution from single images (13, 14), which may be considered as a form of deconvolution of blurring by the diffraction limit. Whilst achieving impressive results, a major hurdle to these data-driven approaches is that they require several 1,000s to 10,000s (13, 14) of experimentally acquired images, which additionally have to be co-registered to high-resolution experimental ground truths for supervised paired training. This often requires the user to have access to additional costly super-resolution imaging systems and hardware. Further, many methods learn, non-discriminately, the inversion of the physical system and the properties of the samples contained in the training set (4). This leads to poor generalisation and, thus, poor performance when a trained network is applied to samples outwith the training parameters (15). Thus, the user is required train bespoke networks for each system and sample imaged, which is a present and major barrier to broad adoption.

Instead of a data-driven approach, we propose that a network can be trained with good performance without any measurement or estimate of a ground truth, solely when informed by physics priors, *i.e.,* a priori knowledge of the physics of the imaging system. This physics-informed learning has emerged to reduce the need for experimental training data and to direct training towards generalisation that is agnostic of the samples being imaged (4, 12). In this paper, we follow the inspiration of such networks and make use of three key priors: (1) we make use of the known PSF of the imaging system to generate simulated paired data that is consistent with images that can be generated by the experiment; (2) the generated images are controlled in their spatial and spatial-frequency content, thus, provide control over the sparsity ex-

pected in the network outputs; and (3) we use experimental images to guide reconstruction towards the perceptual quality and power spectral content expected in biological samples.

Importantly, our method is entirely experimentally unsupervised, *i.e.,* it does not require experimental ground truths, and may be trained rapidly on a single light-sheet volume, including the one that is desired to be deconvolved. Our method reduces the data burden from thousands of sub-pixel registered image pairs (13, 14), to a few hundred unpaired low-resolution image regions. We validate our method experimentally on LSM using an Airy beam in calibration beads, autofluorescent mouse oocytes and embryos, and excised mouse brain tissue. We supplement this validation with demonstrations in multiphoton LSM with the commonly used Bessel beam for developmental biology. The method generalises well, learning primarily the inversion of the imaging operator, which we demonstrate by training only a single network for each beam shape, and applying it to all of the samples we subsequently image. We show that deep learning is superior to conventional deconvolution in its ability to achieve a more symmetric deconvolved PSF, a 3–5 dB% increased peak signal-to-noise ratio, and can be performed rapidly at 0.2 s per widefield image. Further, we illustrate that our method is robust to noise and is 3-fold more resistant to model mismatch than conventional deconvolution, which is advantageous in the presence of misalignment or aberrations. Our approach to deep learning makes the use of structured light fields and deconvolution highly accessible as it does not require acquisition outwith the conventional imaging protocol. Towards this, we illustrate its use in an openly available dataset, and provide the deep learning code as open source to facilitate broad evaluation of this approach among the imaging community.

Results

Network design. Figure 1 illustrates the architecture of our method under training in the context of the LSM setup. The network architecture is described in detail in the Methods Section, and briefly summarised here. We perform deep learning based on a generative adversarial network (GAN) (16) and several tuned training loss functions (TLFs). The inverse problem, *i.e.,* the transformation between an image encoded with a structured PSF (LR) to a high-resolution estimate of the ground truth, is learned using a Generator (G) network, based on a 16-layer residual network (ResNet) (17) (Supplementary Notes 1 and 4). Concurrently, another Discriminator (D) network (PatchGAN (18)) is trained to identify network outputs (DL) and real ground truths (GT). We have found that this adversarial training leads to the learning of strong image features that are consistent with the GT and that are difficult to capture using pixel-wise TLFs (Supplementary Note 3), which is consistent with previous investigations (19).

Figure 1(a) illustrates the LSM geometry (see Methods for details). An Airy light field is generated using a nematic spatial light modulator (SLM) in diffractive mode, encoding the appropriate cubic phase profile (the Fourier transform of the Airy function). The SLM enables precise control of the field intensity and phase in illumination. The desired pupil phase function (Fig. 1(c)), *i.e.,* the spatial Fourier transform of the beam cross-section at focus, is projected onto the SLM (Supplementary Note 13). In this way, it is trivial to estimate the PSF of the LSM by combining the Fourier transform of the pupil, evaluated at the spatial coordinates established by the various lenses, and the diffraction-limited Gaussian spot size of the detection. Here, we train and perform deconvolution in 2D, in the cross-sectional plane of the illumination and detection objective (Fig. 1(a), inset (1)). This plane includes the PSF structure of the propagation-invariant light fields, and corresponds to the lowest resolution plane of the system. In this geometry, training and deconvolution in 2D is a close approximant to a 3D model, whilst remaining computationally accessible to consumer hardware (Supplementary Note 7).

During a single training iteration, a batch of simulated images consistent with the LSM is fed into G. These images comprise two domains: a sparse collection of points, and speckle (Fig. 1(b)). Network inputs and their ground truths are formed, respectively, by convolving the images with the system PSF and a Gaussian spot set to $1/\sqrt{2}$ of the diffraction limit of the objectives. The sparse points provide strong gradients for the network to learn the PSF shape. Speckle is a Fourier-transform pair to a uniformly distributed random phase and a Gaussian spectral amplitude. Thus, speckle images encode a wide spatial-frequency content with low sparsity. This pushes the network to prioritise physical learning (*i.e.,* the transformation by the PSF) over content learning (*i.e.,* the expectation of what microscopy images should look like), which benefits generalisation. During training, the network outputs are compared to the ground truths using the Discriminator (adversarial loss), and using L1-norm error (pixel loss). Concurrently, a random selection of experimental images are processed by G. The inputs and outputs of the network are compared using a pre-trained perceptual loss network (VGG-16 (20)), V (Fig. 1(b)), which focuses on the conservation of salient image content (21) (Supplementary Note 2). The addition of this TLF leads to the preservation of the power spectral density of experimental images, which is an important prior (22), and is effective at eliminating network artefacts. The combined loss of G is the linear combination of the adversarial, pixel and perceptual losses. The Discriminator is trained using conventional means (18), and training is alternated evenly between G and D.

Our method and architecture results in a robust and generalised model. Importantly, and in contrast with most learned image enhancement methods (12), we demonstrate this by training only a single network for each beam shape using only one experimental low-resolution volume of mouse embryos ($\sim$200 images) and no experimental ground truths. We apply this network, non-discriminately, to a wide array of samples, namely fluorescent beads samples, other embryos, oocytes, turbid brain tissue and zebrafish embryos.

Physics-informed learning. Most recent advances in deep learning for image enhancement rely on a careful match of ground truths to low-resolution inputs, often requiring

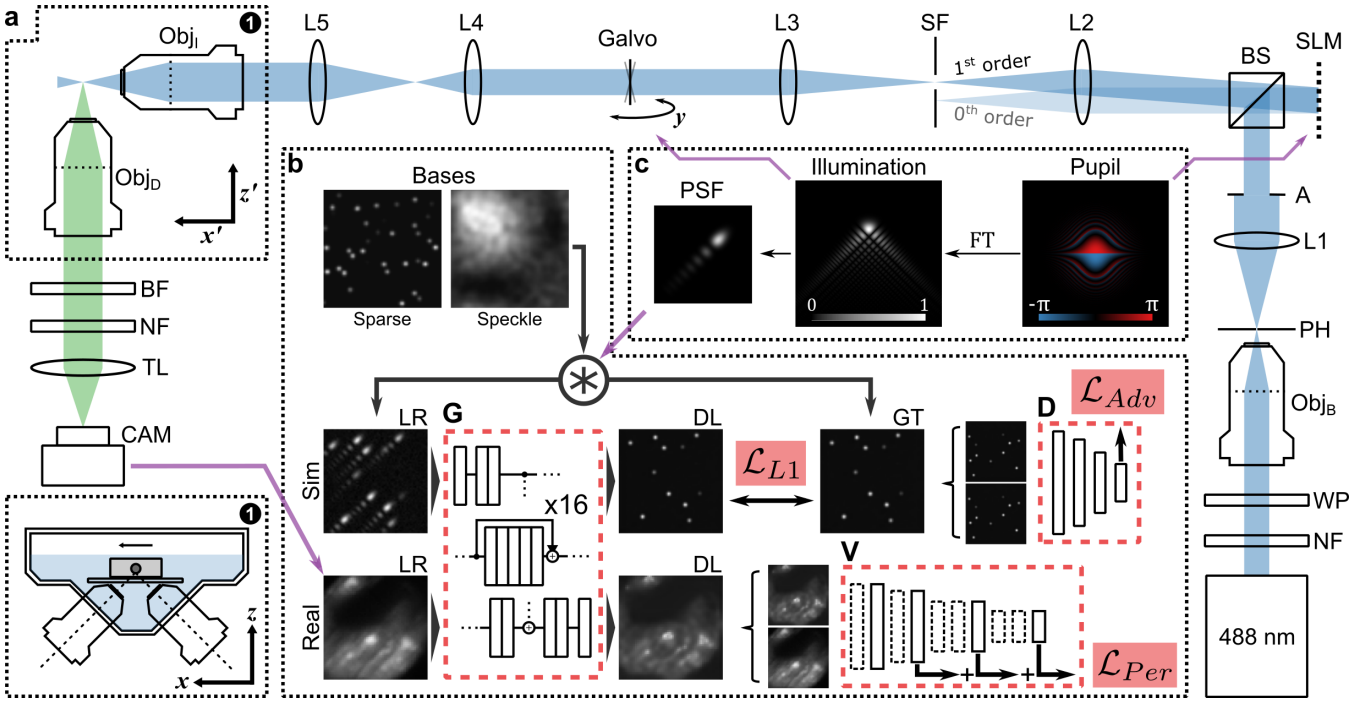

Fig. 1. Light sheet microscopy system and physics-informed learned deconvolution training. (a) Optical geometry of the system. Insets marked by (1) show the optical and physical positioning of the sample, which are denoted using the $x'z'$ and xz coordinates, respectively. (b) Deep learning method, consistent with (c) the physics of beam shaping using a spatial light modulator (SLM). Obj B, I, D: objectives for beam expansion, illumination and detection respectively. A 488 nm laser source is filtered using a notch filter (NF), half wave plate (WP), pinhole (PH) and aperture (A). L1-5: lenses. BS: beamsplitter. The SLM shapes the illumination field in diffraction mode to match the pupil function in (c), which is filtered using a spatial filter (SF) and scanned as a light sheet using a galvonometer driven mirror. Sample fluorescence filtered using a bandpass filter (BF) and a notch filter (NF), and imaged using a tube lens (TL) onto the camera (CAM). Deep learning learns from simulated images generated from sparse and speckle bases and the known PSF, and the real images captured by the CAM. During training, the generator (G) learns from a combination of L_1 pixel loss, adversarial loss from the discriminator (D) and the perceptual loss from the perceptual loss network (V).

1,000–10,000s image pairs (15). Early demonstrations, such as content-aware reconstruction (CARE) in microscopy (13) and cross-modality super-resolution (CMSR) (14), have shown that the choice and quality of training data led to effective image enhancement, tailored to tasks such as super-resolution (CARE Deconvolve), denoising in low-SNR imaging (CARE Denoise), or a reconstruction of an isotropic resolution, whilst using off-the-shelf U-Net network architectures (13, 14) (Supplementary Note 12). More recent advances have improved performance through more tailored networks, such as with a three-dimensional residual channel attention network (RCAN) (23) or with a dual-stage processing network (DSP-Net) (24), based on dense residual block architectures. Recent networks have also addressed the onus of extensive ground truth matching. For instance, in DSP-Net, high-resolution ground truths were degraded such that subsequent undersampled acquisitions may be reconstructed close to the original image quality (DSP-Net Upscale) (24), enabling thigh-throughput imaging. Alternatively, super-resolution ground truth could be generated computationally using SRRF (DSP-Net Content) (24). In similar approaches, the output of deconvolution using a back projection algorithm was used to train a network to speed up processing (25), and high-resolution xy cross-sections in microscopy have been used as a ground truth target to deconvolve xz sections to achieve isotropic volumes (26). In these methods, deep learning has provided an exceptional speed-up and has reduced data requirements, however, these approaches are unlikely to outperform the original numerical algorithms or the imaging system used to generate the ground truths.

In contrast to these methods, we propose that an image enhancement network can be trained without any measurement or estimate of a ground truth. Figure 2 provides a visual comparison of methods available in the instance where only a single volume of a low-resolution image is available and is sought to be deconvolved. Specifically, we demonstrate outputs of DSP-Net (24), CARE (13), RCAN (23) and CMSR (14), pre-trained on biological ground truths. To accomodate an Airy PSFs, we perform two-step reconstruction by performing Richardson-Lucy (RL) deconvolution on the Airy LSM image (LR) first and then processing the result using the relevant pre-trained or 'demo' network. Outputs are then compared to the same area images with a Gaussian beam as the ground truth (GT), and quantified using simulation in Table 1. It is possible to re-train the above networks with the same physics-consistent image pairs and loss functions proposed in this work. However, since many of the networks share very similar architectures, with performance is strongly dependent on training data, the output quality would be similar. Here, instead of comparing the utility of individual building blocks of neural networks, we seek to illustrate the main differences between physics-informed and content learning, which we will discuss presently. Our implementation of DL is similar to DSP-Net and RCAN in the use of a residual architecture. We also present outputs of our DL method trained using an alternative U-Net architecture (18) (DL (U-Net))

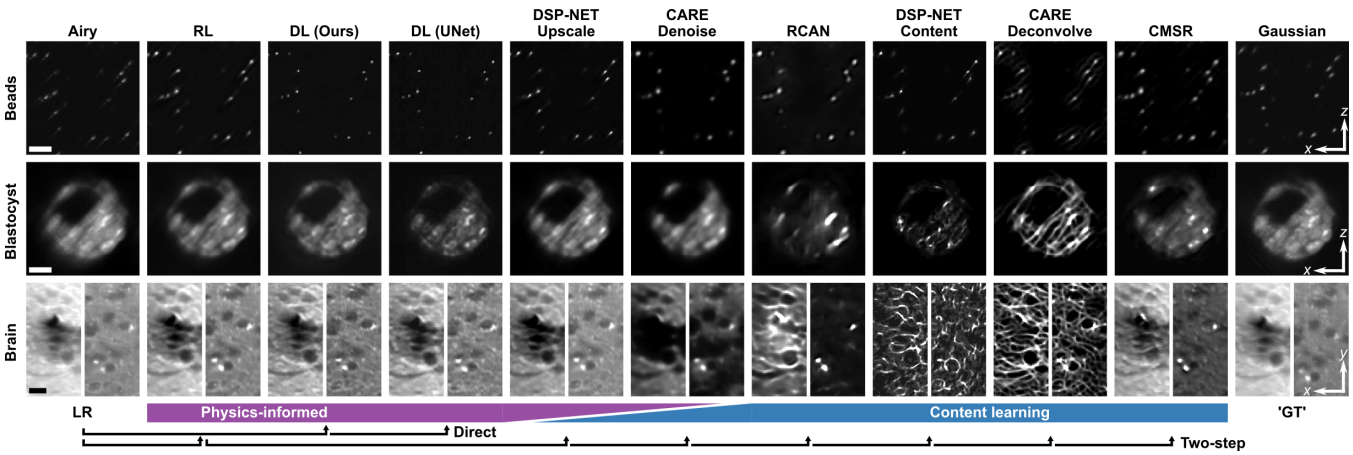

Fig. 2. Illustration of physics-informed learning in the context of contemporary deep learning microscopy methods. Scale bar: 20 μm.

	RL	DL (Ours)	DL (UNet)	DSP Upscale	CARE Denoise	RCAN	DSP Content	CARE Deconv.	CMSR	GT
PSNR Bead (dB)	33.0	**38.6**	37.3	21.2	27.5	25.4	22.4	29.1	30.8	
Res. Bead (um)	3.1	2.0	**1.9**	2.8	4.1	5.7	2.1	3.2	2.4	2.6
PSNR Bio (dB)	30.9	**34.0**	28.7	30.3	27.2	21.2	20.4	23.3	25.0	
Res. Bio (um)	3.7	2.8	2.6	2.8	14.6	4.4	**2.0**	3.2	2.6	4.0

Table 1. Comparison of image enhancement methods using the peak signal-to-noise ratio (PSNR) and resolution (Res.) metrics.

that underlies CMSR, CARE denoise and CARE deconvolve (Supplementary Note 4). We discuss and compare each network in detail in Supplementary Note 12.

There is a finite amount of information that may be carried by a detected signal or image based on the digitisation and SNR. Many image restoration methods rely on enhancing information carried by an image, which is naturally an underdetermined inverse problem. As such, constraints must be placed on inversion to achieve a unique and accurate result (27). For example, regularisation schemes, such as total variation or L_2-norm minimisation, have been used with deconvolution. Deep learning achieves this exceptionally well using feature learning by generating outputs that are most similar to expected images, whilst excluding those solutions that are unlikely to be encountered in microscopy. Specifically, we can consider deep learning to comprise physical and content learning (4, 28). Physical learning encourages consistency with the transformation by the imaging system, for instance, the blur from the finite bandwidth of optical elements or the noise in detectors. Content learning encourages output images to resemble features observed in the ground truths related to, for instance, the type of samples or spectral contrast expected. The greatest constraint through the combination of physical and content learning, for example by training on a single sample type and instrument, has the potential to lead to the greatest performance gain whilst compromising generality. For instance, networks trained to super-resolve microtubules may demonstrate a high resolution and contrast enhancement, but will likely struggle to reconstruct images of other samples (15). Despite the performance of content learning, it is often desirable to achieve a generalised network that learns the physical inversion of the imaging system agnostic of the sample, such that one may be confident that the output is faithful to the imaged data. Physical learning may be encouraged by incorporating imaging physics in the network architecture or the loss functions, or by providing training images that emphasise physics transformations (12).

Table 1 quantifies the performance of each method using simulated data based on beads and biological images (Supplementary Note 12) based on peak signal-to-noise ratio (PSNR) and resolution (Supplementary Note 9). It is immediately evident that those networks that have learned indiscriminately the physics and content of paired images (marked as content learning in Fig. 2) have vastly varied performance across sample type that is often inconsistent with the expected GT. Specifically, RCAN was trained on labelled endoplasmic reticulum in U2OS cells, DSP-Net Content on labelled microtubules in U2OS cells, and CARE Deconvolve on microtubules in INS-1 cells. In these networks, outputs resemble the style of the training set regardless of the input image. DSP-Net Content outputs present the highest resolution, but are visually inconsistent with the GT. The CMSR network was trained on TxRed labelled endothelial cells and presents improved generality across sample types. However, the relative intensity in the images emphasise bright features over the GT. We note that this is common to U-Net architectures (Supplementary Note 4), which can lead to greater potential for improved resolution at the expense of a reduction in PSNR in dense tissues, as evident by comparing the performance of our DL method with DL (U-Net). The CARE Denoise network was trained on high- and low-SNR acquisitions of the sample sample and system. This training favoured system noise reduction instead of image content transformation, which is evident by the low resolution of its outputs. However, when provided with good SNR images, the network has attenuated regions of low image intensity, which is especially evident in the brain. DSP-Net Upscale was trained on high-resolution GTs that have been digitally downsampled with

added noise. Thus, the training favoured the physical transformation of upscaling and denoising. While DSP-Net Upscale demonstrates outputs consistent with the GT, the outputs offer little quantitative improvement over RL deconvolution. In contrast, our method (DL), which is designed for deconvolution, offers superior visual contrast and resolution in all sample types. When we retrained our method using a U-Net (DL U-Net), while presenting improved resolution, the PSNR of dense biological tissues was substantially lower. This is likely due to spatial sparsity being favoured by U-Nets because of its autoencoder design (Supplementary Note 4). We employ the ResNet DL architecture for the demonstrations in this work because of its high PSNR and low computational demand (discussed in Supplementary Note 12). Independent from the network architecture, however, the major distinction of our method to those published previously is that our learning approach is free from any experimental ground truth.

Learned deconvolution. Here, we illustrate that DL can achieve superior performance to numerical methods, specifically the well-known Richardson-Lucy (RL) algorithm (27, 29). This is because, firstly, we forego the common assumption that noise is a uniform and randomly distributed Poisson process (3). Secondly, neural networks naturally use sparsity in images to improve reconstruction (30). Finally, as we will demonstrate, our network is less prescriptive than RL deconvolution and can tolerate greater model mismatch, for instance, when the theoretical and experimental PSFs differ. This means that our network is likely to preserve its good performance even if the microscope is not perfectly aligned.

First, we quantify performance with simulated data on several propagation-invariant beam types. Figure 3(a) evaluates deconvolution in images simulated by convolving a blastocyst embryo GT with PSFs generated by the Airy, Bessel and Gaussian beams. The networks were trained solely on a collection of sparse and speckle bases with no knowledge of the blastocyst images. The generation of the various beam shapes is described in the Methods section and in Supplementary Note 13. Briefly, the Airy beams are generated by a cubic phase in the pupil plane, with a scale-invariant α parameter that corresponds to the gradient of the cubic function. The Bessel beams are formed by a ring in the pupil plane, parameterised by a ring's radius and cross-sectional width. We later focus on the Airy and Gaussian beams in the experiment as they are readily generated by our LSM configuration.

From Figure 3(a) it is evident that when using propagation-invariant beams, light-sheet images (LR) exhibit substantial side lobes. Here, the LR images include randomly distributed Gaussian (thermal) and Poisson (shot) noise. The conventional RL algorithm is able to reduce side lobes and present a spatially consistent image to the GT. However, RL-deconvolved images fail to present a visual quality matching that of the GT or the high-resolution Gaussian beam. Despite losing the structure, side lobes persist as elevated background intensities. We further include a common implementation of RL-deconvolution with total variation regularisation (TV), which exhibits a similar performance to RL but with

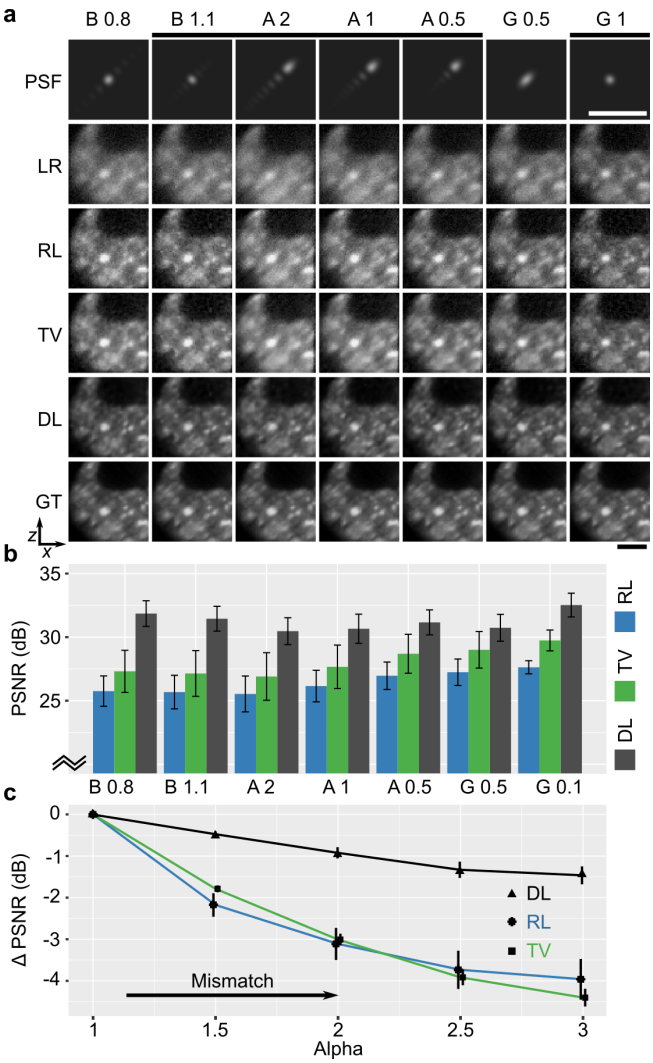

Fig. 3. Comparison of deep learning (DL), Richardson-Lucy (RL) deconvolution and RL deconvolution with total variation (TV) regularisation in simulated data. (a) Deconvolution of low-resolution noisy images (LR) simulated using beams (B: Bessel; A: Airy; G: Gaussian) visualised by their point-spread function (PSF) (parameters listed in Methods) using RL, TV and DL algorithms and compared to the ground truth (GT). Scale bar is 20 μm. (b) Mean peak signal-to-noise ratio (PSNR) of RL, TV and DL deconvolution for each beam shape evaluated in 240 simulated images (Error bar is standard deviation). (c) Relative change in PSNR in deconvolution of an Airy beam due to model mismatch. RL, TV and DL assume that the Airy $\alpha = 1$; however, LR images were simulated with α in [1, 3].

reduced high-frequency noise.

Beam shapes marked by a black bar in Figure 3(a) carry similar spatial resolution. DL is capable of deconvolving LR images much closer to the GT, eliminating the side lobes and presenting a Gaussian-like PSF. This is especially clear in simulated images of beads in Supplementary Note 5. Figure 3(b) quantifies the performance of these methods using the peak signal-to-noise ratio (PSNR) over 240 simulated images. It is evident that DL substantially outperforms RL and TV deconvolution, achieving a 30–33 dB PSNR, whilst RL and TV range in 25–30 dB. We note that a 3 dB change in PSNR is approximately a two-fold change in noise. DL further outperforms RL consistently with increasing noise and with added background fluorescence (Supplementary Note 6). It is important to note that convergence of RL iter-

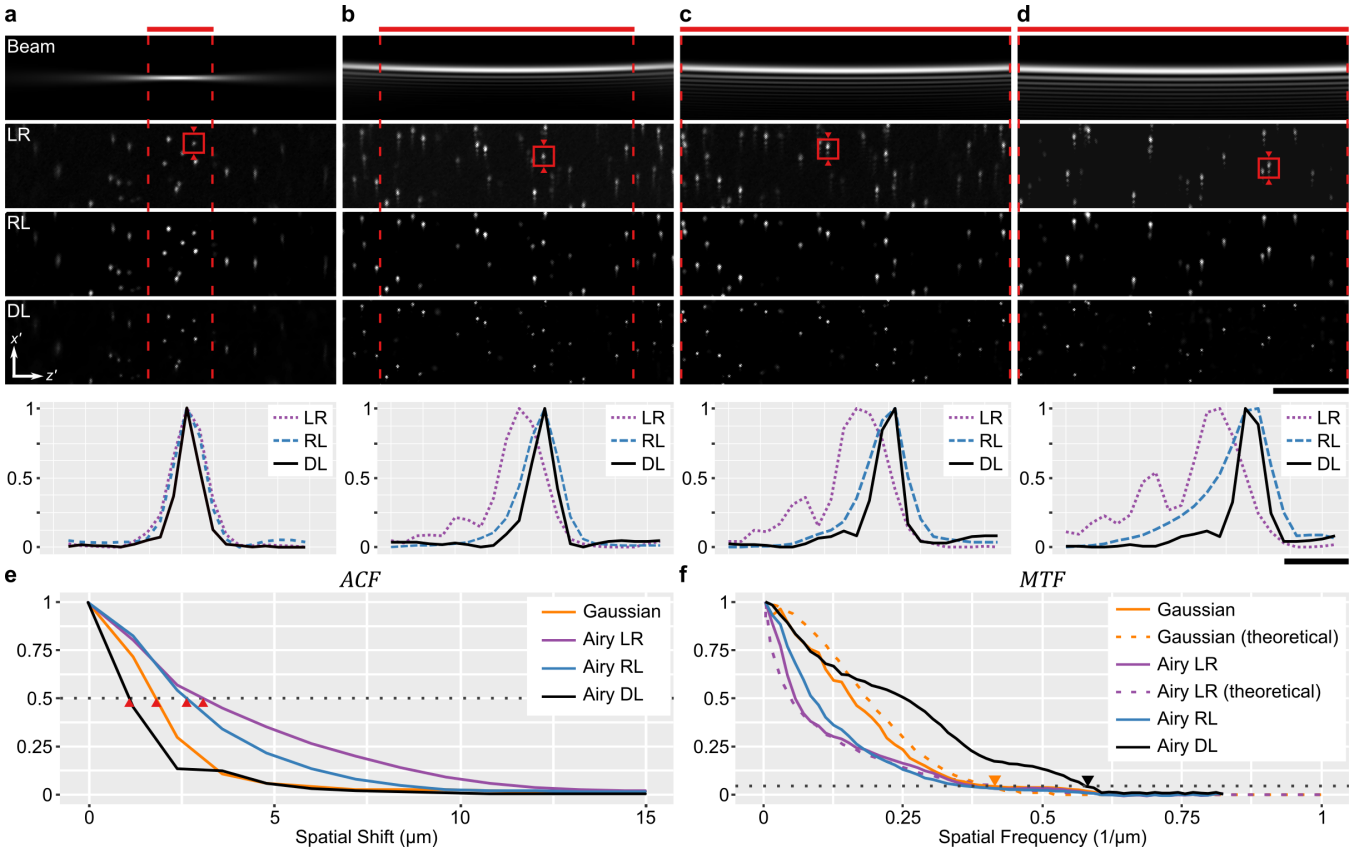

Fig. 4. Learned deconvolution in beads using a (a) Gaussian, (b) Airy ($\alpha = 0.5$), (c) Airy ($\alpha = 1$), and (d) Airy ($\alpha = 2$) beam. (a–d) Show the illumination beam and the corresponding acquired LSM image (LR), which is deconvolved using Richardson-Lucy (RL) and deep learning (DL) methods. Profiles correspond to the bead marked by a red square. Red dashed lines mark the Rayleigh range of each beam. Scale bars are 50 μm for images and 5 μm for plots. (e) Autocorrelation function (ACF) and (f) Modulation transfer function (MTF) with respect to depth of the Gaussian and Airy ($\alpha = 1$) LSM images.

ations leads to high image noise and early truncation is often used as a spatial regulariser. Here, and for all further RL deconvolution, we have truncated RL to 6 iterations specifically to maximise the PSNR parameter for a best-case comparison. Further, we used a TV regularisation parameter of 0.001 for the most visually compelling recovery.

A major challenge in deconvolution and in deep learning with priors is model mismatch (4). Here, this arises when the theoretical PSF differs from that of the imaging system, for instance, due to misalignment or aberrations. Figure 3(c) illustrates the robustness of our method to model mismatch. Specifically, we have generated LR images corresponding to Airy beams with α values in the range of 1 to 3. However, to deconvolve the images, we have used a network trained on an Airy beam with an $\alpha = 1$. Similarly, we have used a single Airy ($\alpha = 1$) PSF as a reference for RL and TV deconvolution. The matched training and testing case of $\alpha = 1$ was used as a reference. Figure 3(c) shows the relative degradation in the PSNR (to the known GT) as greater mismatch is introduced. It is clear that DL tolerates much greater mismatch than RL and TV deconvolution, on the order of 3-fold improvement (2.7–3.8 times) to the error metrics. While blind deconvolution methods may overcome some issues of mismatch, we have observed that blind deconvolution struggles to converge when the PSF is structured (*i.e.,* it is not monotonically decreasing from a central peak) and when images are not spatially sparse, which is consistent with previous observations (3).

Figure 4 extends the demonstration of our DL method to experimental data of 200-nm beads. Using the SLM, we readily generate a Gaussian illumination (Fig. 4(a)), and compare it to the Airy beam with α parameters of 0.5, 1 and 2, respectively (Figs. 4(b–d)). The red bar indicates the Rayleigh range of each beam shape, which corresponds to the field of view available for LSM. The acquired LSM images are labelled as LR, and are deconvolved using DL and RL deconvolution. Here, we exclude TV regularisation for clarity due to its emphasis on sharp discontinuities. For each beam shape, the network was trained on one acquisition of tissue data (see Methods) and applied to all experimental data included in this work. The 200-nm beads are sub-diffraction limit, thus, represent the system response or the PSF. We can see that DL offers superior image quality to RL deconvolution. This is clear in the transverse profile of the beads marked by the red box. The Airy function can be seen in the LR profiles. The RL deconvolution removes the structure of the side lobes, but fails to substantially reduce the envelope, leading to an asymmetric PSF. The superior performance of DL is evident by the sharp and more Gaussian-like PSF.

In addition to decoding the structure, DL leads to a super-resolution of individual beads beyond the diffraction limit of the light field. We can quantify this performance over

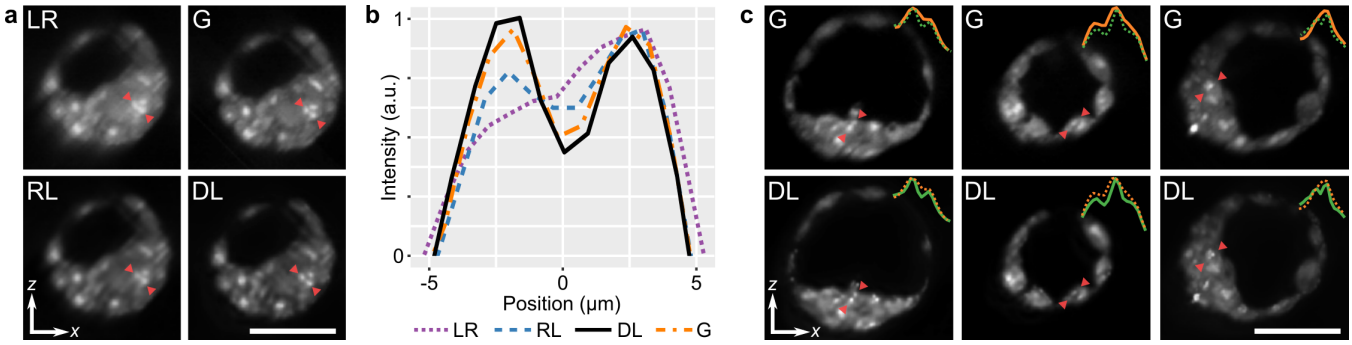

Fig. 5. Cross-sectional images of mouse blastocyst-stage embryos. (a) Airy LSM image (LR) is deconvolved with Richardson-Lucy (RL) and deep learning (DL) deconvolution, and compared with a Gaussian LSM (G). (b) Intensity profiles marked by the red arrows in (a). (c) Cross-sections of blastocyst-stage embryos imaged with a Gaussian LSM (G) and compared to an Airy LSM with deep learning deconvolution (DL). Profiles marked by the red triangles are displayed as insets. Scale bar is 50 μm.

many beads in the field of view using the autocorrelation function (Fig. 4(e)), which represents the shift invariance and relates to the depth or axial LSM resolution. From the Wiener–Khinchin theorem, we know that the ACF of intensity is a Fourier pair to the power spectral density, *i.e.,* the square of the modulation transfer function (MTF). Figure 4(f) shows the MTF of the Gaussian LSM and the Airy ($\alpha = 1$) beams compared to their theoretical MTF, as well as the RL and DL deconvolutions. We can see that the MTF in the acquired images match theory well (dashed lines). The dotted line marks the 5% threshold of the MTF, which corresponds approximately to the full-width at half-maximum (FWHM) of the PSF. We note that the intersection of the Gaussian and Airy beam MTFs at the 5% threshold indicates their capacity to carry similar high-frequency content. However, this information is not well conveyed by the Airy images in the spatial coordinate space as illustrated by the widened ACF. As expected, RL deconvolution decodes Airy images (LR) to improve the ACF width with little benefit seen in the MTF. Remarkably, DL expands the MTF beyond that of the Gaussian beam, and extends the crossing of the 5% threshold by a factor of $\sqrt{2}$. The improvement in the resolution is consistent with the simulation of the HR training data at $\sqrt{2}$ of the LR images. In principle, it is possible to tailor the resolution improvement based on the choice of the priors. However, larger resolution improvements are met with lowered contrast at lower spatial frequencies. Our choice of a $\sqrt{2}$ improvement was motivated empirically by visually inspecting the network outputs. This improved performance is also evident by the rapidly decreasing ACF. The FWHM resolution was quantified as 2.6 μm in the Gaussian and Airy images, and 1.7 μm with the use of DL. This exemplifies the power of deep learning as a method to solve inverse problems that leverages sparsity, learned noise statistics and the imaging operator to both decode information and extend the bandwidth limit.

DL has provided the Airy LSM a $\sqrt{2}$ enhancement in resolution over a substantially wider FOV of up to 372 μm vs. the 27-μm FOV of the Gaussian beam, a greater than 13 fold improvement. The accessible and rapid inference of DL, the wide FOV and the ease of generating propagation-invariant beams confers exemplary potential to multiscale imaging applications. To demonstrate this, we explore our DL method in mouse oocytes and embryos, and in excised mouse brain tissue.

Oocytes and embryos. Rapid minimally invasive imaging of embryos is important for diagnostics and may aid in the monitoring of embryo development following in vitro fertilisation (IVF) (31, 32). LSM is particularly promising in this area due to its rapid volumetric imaging, reduced photodamage and low cost compared to confocal microscopy (31). Our DL-assisted Airy LSM is capable of performing cross-sectional imaging of the entire depth of an embryo ($\sim$180 μm) in one shot, *i.e.,* within a single acquisition on the camera. Further, it is able to convey high-resolution morphological and functional information from autofluorescence.

Figure 5 demonstrates the performance of DL deconvolution in mouse blastocyst-stage embryos. The blastocyst is the final stage of preimplantation embryo development, and comprises two sub-populations of cells: the inner cell mass (ICM) which forms the fetus, and the trophectoderm (TE) which leads to the development of all extra embryonic tissues, including the placenta (33). Figure 5(a) shows a cross-section through a blastocyst imaged with an Airy ($\alpha = 1$) beam and deconvolved using RL and DL methods. The same region was imaged using a Gaussian beam (G). We can see that DL readily decodes the Airy PSF from the LR image, and matches and exceeds the image quality presented by the Gaussian LSM. We illustrate this further by taking a profile marked by the red triangles. In Figure 5(b), we can see that quantitatively the DL profile matches that of the Gaussian GT, while the RL fails to match the relative intensity profile. It is important to mention that the LR profile does not match that of G. This is because the Airy PSF structure (here, extending top-right to bottom-left in the images) has distinct oscillating side lobes. Thus, the local intensity varies non-monotonically with distance from the scattering object. Deconvolution must be performed on such images to decode the underlying quantitative fluorescence distribution, which in our case is done well with DL compared to RL. This can be further evidenced in similar cross-section in Supplementary Note 8. Using Fourier analysis (Supplementary Note 9), the spatial resolution was quantified as 5.0 μm in LR, 4.9 μm in RL, 3.8 μm in G, and 3.4 μm in DL images.

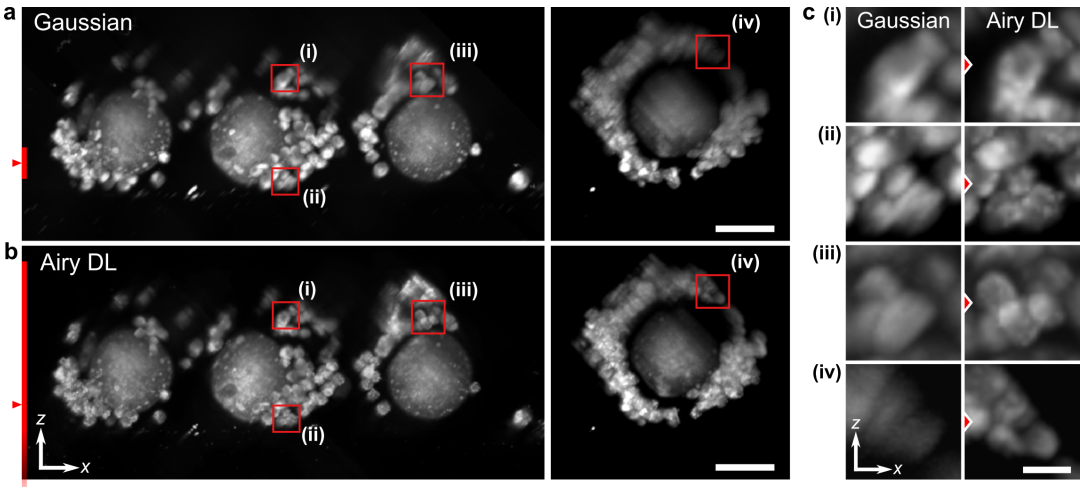

Fig. 6. Cross-sectional LSM images of mouse cumulus oocyte complexes (COCs) with a Gaussian and Airy beams with deep learning deconvolution. (a,b) Widefield cross-sectional maximum intensity projections (range: 50 μm). (c) Zoomed-in regions of cumulus cells (i–iv) emphasising the enhancement of Airy DL over Gaussian LSM from corresponding regions in (a,b) marked by red rectangles. The red line on the left of (a,b) shows the theoretical DOF of the Gaussian and Airy beams. The red triangle marks the focal position. Scale bars are (a,b) 50 μm and (c) 10 μm.

We now focus on demonstrating the advantage of using the Airy beam LSM with deep learning deconvolution compared to the conventional Gaussian LSM, which represents the most popular embodiment of LSM. Figure 5(c) shows several cross-sections from different blastocyst-stage embryos imaged with a Gaussian LSM (G) compared to matching regions imaged with an Airy LSM with DL, hereafter referred to as 'Airy DL' for succinctness. Visually, we see an improvement in the contrast and resolution in the DL images. Line profiles marked by the red triangles and inset into the figures further illustrate that DL is able to distinguish autofluorescent features with a higher spatial resolution. We attribute these bright features to active mitochondria performing high levels of metabolic activity and, thus, generating higher amount of intracellular fluorophores, such as flavin adenine dinucleotide (FAD) (34). This is consistent with the excitation and emission filters in our LSM setup. Resolution, intensity and count of these features may be an important metric for IVF success (32). It is clear that the added resolution may assist the counting of individual sharp features, and the FOV in depth enabled doing so in a single camera snapshot.

Figure 6 compares a Gaussian LSM to Airy LSM with DL in mouse cumulus oocyte complexes (COCs). COCs comprise of an oocyte at the centre, surrounded by much smaller cumulus cells. Imaging of the COC may provide an opportunity to assess the health of the oocyte prior to IVF (34). The larger field of view encompassing multiple COCs emphasises the major advantage of using an Airy beam, namely, the extended depth of field (DOF). Figure 6(a) and (b) shows the widefield cross-sections and Fig. 6(c) emphasises the image enhancement of Airy DL over conventional Gaussian LSM in the identification and inspection of cumulus cells in COCs. These sections are maximum intensity projections over a 50 μm range to emphasise COC morphology. Figure 6(a) and (b) mark the focus and DOF of the illumination beams by the red triangle and bar, respectively. It is clear that much of the Gaussian cross-sections are out of focus. Airy DL is able to maintain the COC volume in the focal region and demonstrates clear morphological features and enhanced resolution, both in and out of focus. In the Figure 6(c), we can see the clear demarcation of cumulus cells, even in regions well beyond the focus available to the Gaussian beam. In focus, we have quantified the resolution (Supplementary Note 9) to be 4.3 μm for LR, 3.9 μm for Gaussian, and 3.1 μm for Airy DL images.

The volumetric and widefield imaging capacity of Airy DL LSM is further detailed in Supplementary Note 10. It shows enface (xy) sections corresponding to Fig. 6 and illustrates an oocyte arrested at the metaphase II stage of meiosis. During metaphase, the chromosomes are aligned and held in place by the meiotic spindle along a plane termed the metaphase plate. This can be seen by the white line that separates two darker spherical regions that depict the spindle barrel. This morphology is particularly prominent in the Airy DL images, and can be confirmed for all oocytes by inspecting the volume (see Data Availability). This is an essential event for an oocyte undergoing nuclear maturation, necessary for fertilisation. As described for the blastocyst-stage embryo, the small bright regions within the oocyte are indicative of metabolic activity of mitochondria through FAD autofluorescence. The multitude of morphological features enhanced by Airy DL over a wide FOV and depth range underlines its potential for label-free and low-phototoxic imaging for embryo health and IVF success. Revealing the source and implications of these markers would be of considerable interest for future bioimaging studies.

Brain. LSM is a powerful emerging technique for neuroscience (35). It is well positioned to address multiscale imaging at depth, combining imaging at the resolutions and timescales of synapses, over tailored fields-of-view corresponding to whole brains of model organisms or regions of pathogenesis. However, penetration depth is a present challenge, especially in single-photon modalities. The Airy light field has several properties that are of great benefit in this area. Beyond its propagation invariance, Airy light fields

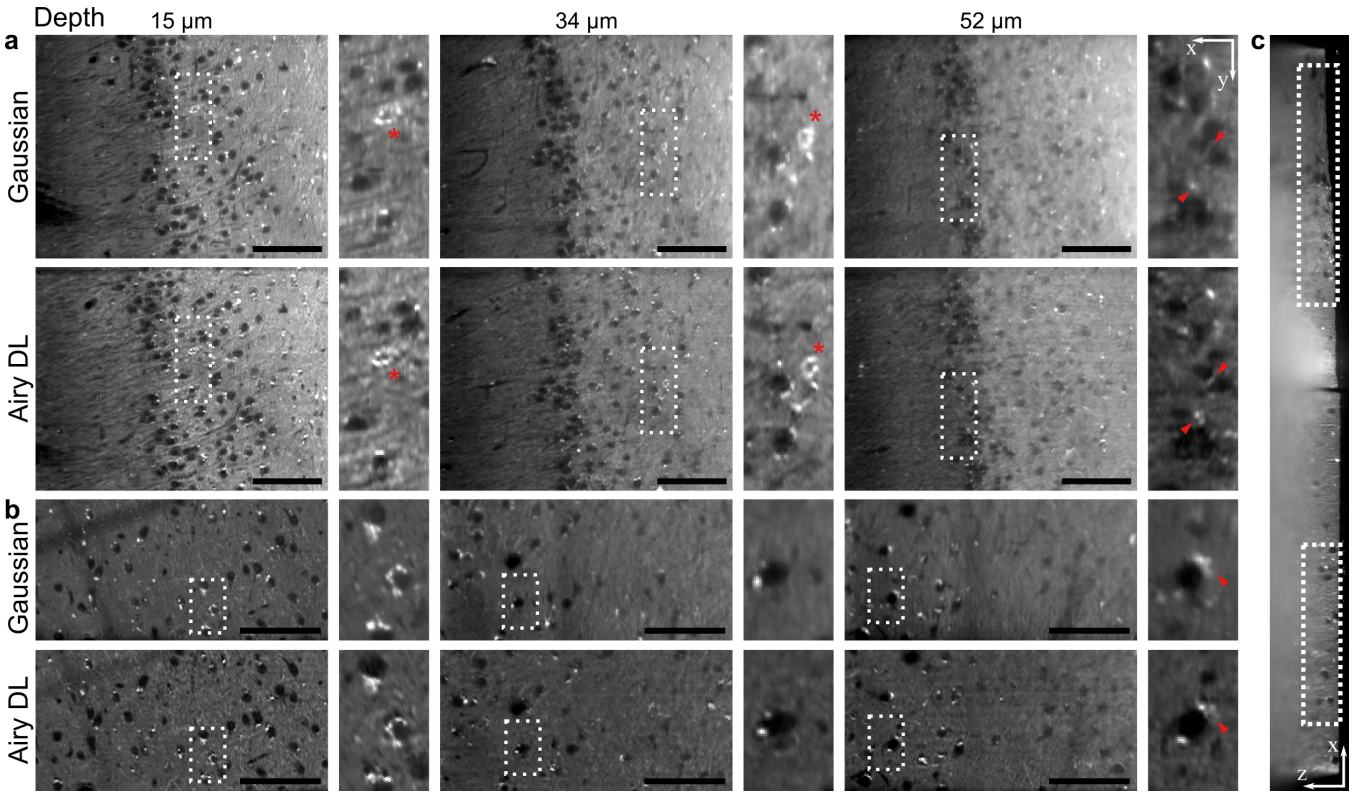

Fig. 7. Excised brain tissue from a mouse expressing enhanced YFP in parvalbumin-positive (PV+) neurons. (a, b) Enface sections in depth imaged using a Gaussian and Airy LSM with deep learning deconvolution. Sections correspond to the (a) hippocampus and (b) deeper cortical regions. Insets to the right correspond to the regions marked by a dashed rectangle. Red asterisks mark PV+ neurons. Red triangles mark regions that illustrate enhanced performance from the combination of an Airy beam and deep learning at depth. (c) Cross-sectional image (z is depth) showing the regions displayed in (a) and (b). Scale bars are 100 μm.

are self-healing. If a portion of the field is blocked or scattered, the beam cross-section will reform into its expected Airy function. This has led to several observations that Airy light fields penetrates on the order of 30% deeper into tissues (7, 36).

Towards this, we demonstrate our technique in excised mouse brain that expresses enhanced YFP in parvalbumin-positive (PV+) neurons. Figure 7(a) and (b) show enface (xy) sections with depth taken from cortical areas, including the hippocampus. These regions were selected from a larger acquired volume, indicated in the cross-sectional xz image (Fig. 7(c)). To the right of each xy image are zoomed-in regions corresponding to the areas marked by the white dashed rectangles. The xy sections visualise the imaging performance in depth, and focus on regions of interest selected by the plane of brain dissection. Because PV+ neurons provide somatic inhibition in the hippocampus (37) and the neocortex (38), we hypothesise that axonal boutons in perisomatic regions (around the cell body) clearly mark PV+ neurons in Airy DL images as bright rings (red asterisks). While PV+ neurons are clearly seen, axonal terminals (bright small spots) around the cell bodies (large dark circles) are also visible in other neurons in lower number. Close to the surface, the enhancement in contrast and resolution of the Airy DL, for instance when inspecting the areas marked by the red asterisk, matches that seen in the blastocysts (Fig. 5). We have quantified the resolution (Supplementary Note 9) to be 7.1 μm for LR, 5.3 μm for Gaussian, and 4.5 μm for Airy DL images. Image quality in the Gaussian LSM rapidly deteriorates with depth, showing substantial blurring and loss of contrast. However, image quality is preserved in depth with Airy DL LSM, distinguishing sharper fluorescent features. This is particularly evident at 52 μm depth in the regions marked by the red triangles. The capacity to image optically thick sections of the brain with high resolution and contrast enables the inspection of axonal connections with minimal damage, in contrast with, for instance, microtome sectioning needed for conventional fluorescence microscopy. Airy DL LSM provides attractive opportunities for detailed studies of morphology, neurodegeneration and broader pathogenesis in neuroscience.

Multiphoton LSM with a Bessel beam. We have demonstrated our DL method with Airy LSM because of the facile control of the PSF, enabling Gaussian LSM ground truths. However, our method is also readily adaptable to deconvolution of other PSFs, such as the Bessel beam which is common to many micropscopy methods (6, 8) and underpins lattice LSM methods (9). Towards this, we briefly demonstrate the application of DL to a multiphoton LSM system using a Bessel beam, generated using an axicon lens (11). The setup and data acquisition methods are presented in detail in (39). Figure 8 illustrates the low-resolution Bessel images (LR) deconvolved with a DL network trained on a manually matched Bessel PSF. Figure 8(a) clearly illustrates the capacity to remove the side lobe structures generated by the Bessel beam.

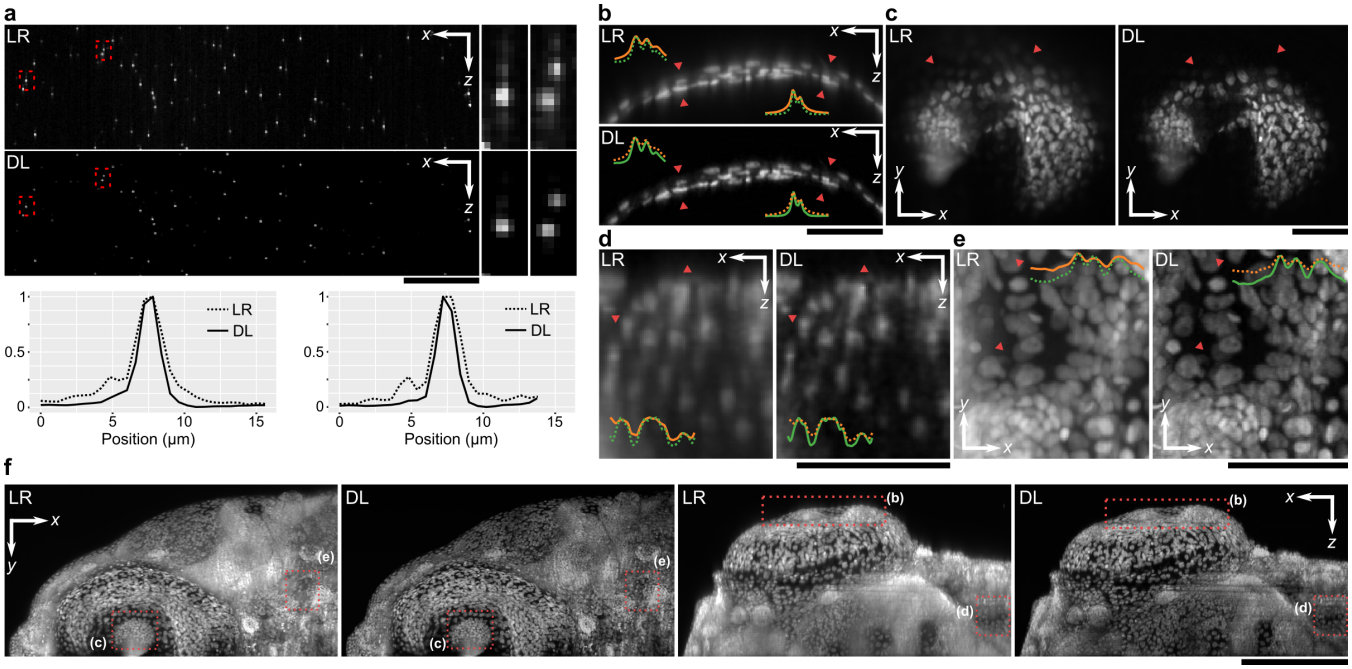

Fig. 8. Learned deconvolution in a multiphoton LSM with a Bessel beam. (a) Fluorescent beads encoded by a Bessel PSF (LR) deconvolved using deep learning (DL). Zoomed in regions correspond to areas marked by the red rectangle. Line plots are representative of the axial intensity profile, showing deconvolution of side lobes. (b–f) Zebrafish embryo sections: (b) cross-sections of the iris; (c) en-face sections of the pupil; (d) cross-sections of the body; (e) en-face maximum intensity projections (20 μm) of the body. Line insets correspond to intensity profiles along the regions marked by red triangles. In (c), red triangles illustrate out-of-plane defocus. (f) Maximum intensity projections of the entire sample in the cross-section and en-face planes. Red rectangles mark locations corresponding to (b–e). Scale bars are 50 μm in (a–e) and 200 μm in (f).

The resolution improvement in Fig. 8(a) can be quantified as the FWHM of the PSF profiles. Resolution improved from 2.3 μm in LR to 1.6 μm in DL images, which is consistent with the improvements in the Airy LSM (Fig. 4).

Figures 8(b–f) illustrate DL deconvolution in images of a Zebrafish embryo, which is a common model organism for developmental and cardiovascular studies (39). Figures 8(b–e) show the performance of DL in different sections of the Zebrafish corresponding to regions marked in Fig. 8(f). Image intensities were normalised to the background intensity. It is evident that DL images present sharper contrast and importantly a reduction in haze from the side lobe structures generated by the Bessel beam (e.g., Fig. 8(c)) in both the zoomed-in regions and the maximum intensity projections in Fig. 8(f) The spatial resolution was improved from 2.6 μm in LR to 1.9 μm in DL images (Supplementary Note 9).

In Supplementary Note 11, we further demonstrate the broad applicability of our method by deconvolving data made openly available by others. Specifically, timelapses of a developing Medaka larvae imaged using multiphoton LSM with a Bessel beam, provided by Takabezawa *et al.,* (40) as open source. The full timelapse videos are available as Supplementary Movies S1 and S2.

Discussion

We have demonstrated that learned deconvolution using physics priors in lieu of experimental ground truths can achieve powerful and generalised deconvolution. This simplifies the requirements to solely images that would be collected in a standard experiment and, importantly, does not require investment in new imaging hardware solely for training the network. This physics-informed learning is a marked contrast to data-driven approaches (13, 14, 23), which derive their performance, in part, from sample-specific priors, thus, requiring training for each class of sample. The use of physics theory to train neural networks has emerged in recent works (12), including in correcting for scattering (41) and enabling deconvolution of spatially varying PSFs for diffuser-based imaging (42). It is likely that more complex physics simulations, such as those including the simulation of aberration and scattering, will lead to better-performing networks (12). Accurate physics simulations have already demonstrated substantial improvements in computational methods in light-field microscopies (43, 44). Further, aberrations from alignment incorporated in a network via Zernike polynomials have improved phase retrieval (45). More complex priors may also be matched with the capacity for deeper or volumetric network training, or training with multi-step networks to address individual image enhancement tasks, as demonstrated by RCAN (23) and DSP-Net (24) networks.

While physics-informed learning simplifies training and enables broad generalisation (4, 12), the addition of content learning though accurate experimental ground truths offers the greatest constraints to optimisation and inversion tasks. Thus, content learning will likely still offer the greatest performance, which is evident in cross-modality restoration (14, 23). Though, care must be taken to ensure that accurate image content is recovered through adequate generalisation for a particular type of sample (15). The challenge

in paired data acquisition, however, limits the broad uptake of such methods, which may become only feasible for routine imaging of samples with fixed optical instruments, for instance, with commercially available microscopes. It may be possible to augment data-driven approaches with physics priors to reduce the volume of data required and optimise the balance in physical vs. content learning (12).

Our DL method is particularly useful for LSM with propagation-invariant Airy and Bessel beams, decoding image content with a high contrast, over an extended field of view and additionally provide super-resolution capacity. Specifically, we show that DL is superior to conventional deconvolution in its ability to achieve a more symmetric deconvolved PSF, a PSNR of 30–33 dB compared to 25–30 dB for RL deconvolution, and a 3-fold improvement in the tolerance to a mismatched PSF. Importantly, we also show that the combined Airy DL method exceeds the performance of a conventional Gaussian LSM, exceeding both the resolution and the contrast, as illustrated by the MTF (Fig. 5(f) and Supplementary Note 9). It is important to note that the use of propagation-invariant beams leads to a reduction in contrast due to the distribution of power into the side lobes (5), establishing a trade-off in resolution, contrast and DOF. Scattering through turbid biological tissues and background fluorescence can lead to a reduction in resolution and contrast (Supplementary Note 9). In these cases, Airy DL still extends on both resolution and contrast compared to the Gaussian LSM, which offers new opportunities for rapid multiscale bioimaging, such as in embryology and neuroscience.

Such performance is rarely afforded by numerical solutions to the inverse problem of imaging. This is because many algorithms are designed to optimise the solution with minimal consideration of content or physics priors. For instance, in the case of RL, a maximum-likelihood solution is reached for a general case (3). DL, however, by its nature (30) and careful design, takes advantage of learned priors to preform more challenging underdetermined inversions (12). DL leverages sparsity and other common image features that are present in natural images and the estimate of the forward operator of LSM imaging to exclude those solutions that cannot be practical in our imaging configuration. We note that priors can be utilised in conventional deconvolution algorithms, such as TV regularisation, especially in those seeking super-resolution (3). However, complexity in these methods is often met with longer computation time, and ill-posed nature often requires high signal-to-noise in the experiment, for instance, in SRRF (46) which can require more than 300 images in one location. DL provides an accessible means to deconvolution with learned priors with rapid feed-forward inference (here, 0.2 s per widefield image). Further, the support of DL by contemporary hardware, such as the GPU, promise ultrafast and even real-time processing with a low barrier-to-entry. Airy and Bessel beams can be generated readily using cheap optical components (5, 10, 11). The combination of DL with compact and inexpensive LSM may provide a powerful and accessible tool for bioimaging. This may further democratise other precision LSM setups, such as the lattice light sheet (9) and more broadly other microscopic approaches that employ deconvolution.

Conclusion

We have demonstrated learned deconvolution of structured PSFs in LSM, leading to the recovery of high-resolution and high-contrast information over extended fields of view. Our method is distinguished from previous microscopy restoration methods in that it is experimentally unsupervised and requires no high-resolution ground truths, which has led to good performance and generality. The application to embryology and neuroscience show attractive potential across a range of bioimaging studies. Our method, which is provided to the reader as open source, may enable ready implementation accessible to many imaging setups.

Materials and Methods

Network. Let us consider the LSM imaging process described by $y = H(x) + \epsilon$, where x is the ground truth (GT), y is the recorded image (LR), H is the imaging operator that we seek to invert, and ϵ is some noise. The goal of deep learning is to train a network, G, such that is provides a good estimate of the ground truth (DL) from LR images, $\hat{x} = G(y)$. This is achieved by minimising a training loss function (TLF), $\mathcal{L}_G$, *i.e.,* $\text{argmin}_G \{\mathcal{L}_G\}$. We implement this process using a generative adversarial network (GAN) (16), where G is trained concurrently with an adversarial discriminator network, D, that seeks to maximise its ability to discriminate between x and $\hat{x}$. D features in $\mathcal{L}_G$, and leads G towards optimising the adversarial min-max game (16).

The discriminator, D, is trained to minimise the following TLF based on least-squares (47):

$$\mathcal{L}_D = \frac{1}{2}\mathbb{E}[(D(x) - \delta_T)^2] + \frac{1}{2}\mathbb{E}[(D(G(y)) - \delta_F)^2],$$

where δ_T and δ_F are the 'true' and 'false' labels, and $\mathbb{E}$ is the expected value evaluated over the entire training batch. We utilise one sided smooth labels, *i.e.,* $\delta_T = 0.9$ and $\delta_F = 0$ to prevent D from over-saturating learning from a particular strong feature. D is based on the Patch GAN architecture (18) (Supplementary Note 1).

The Generator, G, is trained to minimise a linear combination of three TLFs:

$$\mathcal{L}_G = \mathcal{L}_{Adv} + \lambda\mathcal{L}_{L1} + \gamma\mathcal{L}_{Per}$$

which are the adversarial, L_1-norm and perceptual losses, respectively. Weights λ and γ are chosen such that $\mathcal{L}_{Adv}$ and $\mathcal{L}_{L1}$ have approximately equal contribution to $\mathcal{L}_G$ at the end of training (here, $\lambda = 30$). Perceptual loss acts as a regulariser, and was manually chosen as 1×10^{-4} (Supplementary Note 2). G is based on a 16-layer residual network (ResNet) (17), inspired by SRGAN (19) (Supplementary Note 1).

We further use multi-objective training (48), *i.e.,* we train two independent discriminators, D_1 and D_2. Using each discriminator, we evaluate a least-squares adversarial loss (47):

$$\mathcal{L}_{Adv_{1,2}} = \mathbb{E}[(D_{1,2}(G(y)) - \delta_T)^2].$$

The minimum of the two is used to update the gradients of G:

$$\mathcal{L}_{Adv} = \min\{\mathcal{L}_{Adv_1}, \mathcal{L}_{Adv_2}\}\,.$$

Multi-objective training prevents a singe D and G being locked into a min-max game over a single strong feature. In such a case, the second discriminator becomes momentarily free of the min-max game, and is able to reestablish a good training gradient.

Pixel loss using L_1-norm establishes spatial consistency between DL ($G(y)$) and GT (x) images, and is given as:

$$\mathcal{L}_{L1} = \mathbb{E}[||G(y) - x||_1]\,.$$

A further perceptual loss is provided with a pre-trained and publicly available VGG-16 (20) network (V). Inference of an image with V leads to its layers possessing some quantification of image content that is less dependent on its precise pixel values (21). Specifically, we compare LR and DL image content by extracting intermediate features of their V inference:

$$\mathcal{L}_{Per} = \mathbb{E}\left[\sum_i (V^i(G(y')) - V^i(y'))^2\right]\,,$$

where i are the selected layers of V.

It is important to note that $\mathcal{L}_{Adv}$ and $\mathcal{L}_{L1}$ are calculated using the DL and GT images that are simulated based on the known light propagation in the LSM (Fig. 1). Real acquired LR images feature solely in $\mathcal{L}_{Per}$. Perceptual loss, in this instance, plays an important role in directing G towards good performance in real tissues. If a pixel-wise loss was to be used instead, it would act to make DL images look like LR images, which is counter-productive. Instead, perceptual loss provides two important features (Supplementary Note 2). First, it focuses on preserving content that is present in the LR images, namely, the distribution of spectral density, contrast and brightness, and the low-resolution content that should be unaffected by the deconvolution process. Second, it flags network artefacts, such as fixed pattern noise, with strong increases in loss. This makes it a valuable regularisation parameter to bridge mismatch between the content that can be efficiently simulated and real samples. Perceptual loss has previously been valuable for phase-retrieval in microscopy (49).

Training and Inference. Training was performed using simulated LR images and their GT, and real LR images of tissue. All inputs and DL outputs were 64×64 pixel grayscale images. LR images were simulated by convolving a sparse or speckle bases (Fig. 1(b)) with a known PSF. The PSF was evaluated by superimposing a Gaussian detection beam orthogonal to an illumination beam. The beam propagation models are described in the subsequent sections. GT images were simulated by convolving the same basis used for LR with a Gaussian spot at $1/\sqrt{2}$ the diffraction limit. Real LR images were retrieved from an LSM volume of a blastocyst in an automated fashion by identifying the brightest region of interest in each cross-section and applying a random spatial shift. Simulated LR and GT images were normalised to a range of (0.1, 0.9) to avoid clipping. The real LR images were scaled to the minimum and maximum intensity of the entire LSM volume.

Training for each beam type was performed using a total of 512 sparse and 512 speckle simulations, and 215 LR images from a blastocyst volume. Of these, a random selection of 32 sparse and 32 speckle LR images were used for validation. Due to the adversarial nature of GAN training, no early stopping was used. This is because adversarial loss is not an indication of convergence. We used a batch size of 8 and trained for 300 epochs. Gradient-based optimisation was performed using the Adam optimiser with learning rates of 1×10^{-4} and 4×10^{-4} for G and D, respectively, and hyper-parameters $\beta_1 = 0.5$ and $\beta_2 = 0.999$. Training was performed using PyTorch and an NVidia RTX 2060, and took approximately 2.5 hours. Inference on widefield images was performed by a 64×64 sliding window with an 8-pixel overlap to avoid boundary issues. Inference took 3 ms for a single LR image, which corresponds to approximately 0.5 s for a widefield LSM image. We replicated training and inference on an NVidia RTX 3090, which reduced the total training time to 1.2 hours, and inference time to 1 ms and 0.2 s for single and widefield sections, respectively. This is comparable to RL deconvolution, which took 0.3 s for a widefield image or 0.5 s with TV regularisation.

Richardson-Lucy deconvolution. Richardson-Lucy deconvolution was performed in 2D (and 3D in Supplementary Note 7) using a Python interface to ImageJ, which also includes common implementations, such as using total variation regularisation, as part of its Ops framework.

Imaging system. The imaging setup is illustrated in Figure 1(a). A 488-nm laser source (TOPTICA, Germany) was used for illumination. The beam was expanded with an objective (Obj_B, RMS10X, 0.25NA, 10X, Olympus) and lens (L1, f:150 mm), and apodised using an aperture. The beam was filtered spectrally using a notch filter (λ_c:480 nm, $\Delta\lambda$:17 nm) and spatially using a pinhole (15-μm diameter) at the objective focus. A half wave plate controlled polarisation to maximise the efficiency of the spatial light modulator (SLM, X10468, Hamamatsu, Japan). Phase and amplitude was controlled by using the SLM in diffraction mode. The desired phase was projected onto the SLM and the desired intensity was modulated by a phase ramp, realising a blazed grating. The first-order diffraction was spatially filtered using an aperture (SF). The beam was relayed by lenses (L2–L5, f:250, 100, 50, 75 mm) and the illumination objective (Obj_I, 54-10-12, 0.367NA 4X, Navitar). The light sheet was generated by a galvo scanner (Galvo, Thorlabs, NJ) in the Fourier plane. Fluorescence was collected by the detection objective (Obj_D, 54-10-12, 0.367NA 4X, Navitar) and relayed to the camera (CAM, Iris 15, Teledyne Photometrics, AZ) by a tube lens (TL, f:200 mm). Fluorescence was filtered using a bandpass filter (BP, λ_c:532 nm, $\Delta\lambda$:50 nm) and a notch filter (NF, λ_c:488 nm).

Acquisition. LSM volumes were acquired by placing the sample at the focus of the objectives and laterally scanning using a motorised transducer (M-230.10, Physik Instrumente Ltd). Scanning position was synchronised to the camera acquisition, and images were acquired in 0.5-μm steps. Due to the geometry illustrated in the inset in Figure 1(a), lateral scanning was at 45° to the detection plane (also detailed in (36)). Due to this shearing, acquired image stacks were transformed to the physical Cartesian coordinate space using in-house built software (MATLAB, MA) based on interpolation. All LSM images presented in this paper were sampled with an isotropic 0.85 μm pixel size.

Structured light fields. The description of the beams used in this work is provided in Supplementary Note 13. Briefly, all beam shapes are described in the pupil plane. The Gaussian beam is described by its $1/e^2$ waist size (w_0). The Airy beam is described by the $1/e^2$ waist size (w_0) of a Gaussian intensity and a scale-invariant α parameter that represents the cubic phase modulation rate. The Bessel beam is described by a ring with a $1/e^2$ waist size (w_0) and a ring radius (rr). Table 2 lists the beam parameters used in this work.

Beam	Parameters (mm)	Res. (μm)	DOF (μm)
Gaussian 1 (G1)	w_0: 1.02	2.5	27
Gaussian 0.5 (G0.5)	w_0: 0.510	5.3	107
Airy 0.5 (A0.5)	w_0: 1.38; α: 0.5	2.6	169
Airy 1 (A1)	w_0: 1.53; α: 1	2.5	261
Airy 2 (A2)	w_0: 1.73; α: 2	2.5	372
Bessel 0.8 (B0.8)	w_0: 0.04; rr: 0.8	3.7	510
Bessel 1.1 (B1.1)	w_0: 0.055; rr: 1.1	2.8	276

Table 2. Illumination beam properties. Res: axial resolution. DOF: depth of focus.

Phantoms. Phantoms were utilised to characterise the system's PSF and demonstrate DL performance. Phantoms were fabricated from 200-nm diameter green fluorescent microspheres (G200, Duke Scientific, CA) manually mixed with 1.5% agarose. The samples were pipetted into and cured in a 3D-printed sample holder. The sample holder featured a thin fluorinated ethylene propylene (FEP) window for imaging and index matching (RI: 1.344).

Mouse oocytes and embryos. Female (21-23 days) CBA x C57BL/6 first filial (CBAF1) generation mice were obtained from Laboratory Animal Services (University of Adelaide, Australia) and maintained on a 12 h light: 12 h dark cycle with rodent chow and water provided *ad libitum*. All studies were approved by the University of Adelaide's Animal Ethics Committee and were conducted in accordance with the Australian Code of Practice for the Care and Use of Animals for Scientific Purposes.

Female mice were administered intraperitoneally (i.p.) with 5 IU of equine chorionic gonadotropin (eCG; Folligon, Braeside, VIC, Australia), followed by 5 IU human chorionic gonadotrophin (hCG, i.p.; Kilsyth, VIC, Australia) 48 h later. Mice were culled by cervical dislocation 14 h post-hCG administration and the oviducts carefully removed. Cumulus oocyte complexes (COCs) were harvested by gently puncturing the ampullae of the oviduct using a 29-gauge insulin syringe with needle (Terumo Australia Pty Ltd, Australia). These COCs were either fixed immediately in 4% paraformaldehye (PFA) or co-incubated with sperm for IVF. In vitro fertilisation (IVF) occurred through co-incubation of matured COCs with sperm in Research Fertilisation medium (ARTLab Solutions, Australia) for 4 h at 37^oC. Resulting presumptive zygotes were transferred into Research Cleave medium (ARTLab Solutions, Australia) under paraffin oil and allowed to develop to the blastocyst-stage (96 h post-IVF) at 37°C in 5% O_2.

Cumulus oocyte complexes (COCs) and blastocyst-stage embryos were fixed in 4% PFA for 30 mins at room temperature, followed by washes in PBV (phosphate buffer saline; PBS containing 0.3 mg/ml of polyvinyl alcohol; PVA). Fixed COCs and blastocyst-stage embryos were mounted under oil on a 3D-printed sample holder with an FEP window for imaging.

Mouse brain tissue. All animal experiments in Figure 7 were performed in accordance with the United Kingdom Animals (Scientific Procedures) Act of 1986 Home Office regulates and approved by the Home Office (PPL70/8883). Detailed procedures are described elsewhere (50). Briefly, a female adult (8.4 months old) PV-Cre::Ai32 mouse (PV-Cre, JAX008069; Ai32, JAX012569) was used. This mouse expressed channelrhodopsin-2 (ChR2) tagged with enhanced YFP (EYFP) in parvalbumin-positive (PV+) neurons. The mouse was deeply anaesthetised with a mixture of pentobarbital and lidocaine, and perfused transcardially with physiological saline followed by 4% paraformaldehyde/0.1 M phosphate buffer, pH 7.4. After an overnight post-fixation in the same fixative, the brain was immersed in 30% sucrose in phosphate buffered saline (PBS) at 4°C for cryoprotection. The brain was cut into 1-mm thick coronal sections and kept in PBS until imaged. The imaging was focused on the cortex and hippocampus regions.

Multiphoton LSM. The data and setup used for multiphoton LSM experiments on the zebrafish were used with permission from our previous work (39).

Acknowledgements. We would like to acknowledge Federico Gasparoli for early support in constructing the imaging system and providing multiphoton data, and Mirna Merkler for preparing the excised mouse brain section. We acknowledge Erik Linder-Norén's implementations of GANs in PyTorch (https://github.com/eriklindernoren/PyTorch-GAN) that instantiated the code developed in this project.

Data Availability. Code used to simulate light-sheet microscopy and perform deep learning is available publicly at https://github.com/philipwijesinghe/learned-deconvolution. All data underpinning this study is available at https://doi.org/10.17630/bf92bc18-0b81-41f7-bd44-d74040af7cf0.

Funding. This project was funded by the UK Engineering and Physical Sciences Research Council (grants EP/P030017/1 and EP/R004854/1), and has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement (EC-GA 871212) and H2020 FETOPEN project "Dynamic" (EC-GA 863203). PW was supported by the 1851 Research Fellowship from the Royal Commission. KRD was supported by a Mid-Career Fellowship from the Hospital Research Foundation (C-MCF-58-2019). SS was funded by BBSRC (BB/M00905X/1).

Author Contribution. PW and KD initiated the project. PW designed the network and imaging theory. SC built the setup and performed imaging. PW performed image processing, network training and image analysis. DJXC and KRD prepared the embryo samples, and provided training and embryo image analysis. SS provided brain tissue, training and analysed brain data. PW wrote the manuscript with input from all authors. All authors reviewed and approved the work. KD supervised the project.

Conflicts of Interest. The authors declare no conflicts of interest.

Bibliography

1. Sibarita, J.-B. Deconvolution Microscopy. In Rietdorf, J. (ed.) *Microscopy Techniques: -/-*, Advances in Biochemical Engineering, 201–243 (Springer, Berlin, Heidelberg, 2005). URL https://doi.org/10.1007/b102215.
2. McNally, J. G., Karpova, T., Cooper, J. & Conchello, J. A. Three-Dimensional Imaging by Deconvolution Microscopy. *Methods* **19**, 373–385 (1999). URL https://www.sciencedirect.com/science/article/pii/S1046202399908733.
3. Sage, D. *et al.* DeconvolutionLab2: An open-source software for deconvolution microscopy. *Methods* **115**, 28–41 (2017). URL https://www.sciencedirect.com/science/article/pii/S1046202316305096.
4. Barbastathis, G., Ozcan, A. & Situ, G. On the use of deep learning for computational imaging. *Optica* **6**, 921–943 (2019). URL https://www.osapublishing.org/optica/abstract.cfm?uri=optica-6-8-921.
5. Vettenburg, T. *et al.* Light-sheet microscopy using an Airy beam. *Nature Methods* **11**, 541–544 (2014). URL http://www.nature.com/nmeth/journal/v11/n5/abs/nmeth.2922.html.
6. Planchon, T. A. *et al.* Rapid three-dimensional isotropic imaging of living cells using Bessel beam plane illumination. *Nature Methods* **8**, 417–423 (2011). URL https://www.nature.com/articles/nmeth.1586.
7. Nylk, J., McCluskey, K., Aggarwal, S., Tello, J. A. & Dholakia, K. Enhancement of image quality and imaging depth with Airy light-sheet microscopy in cleared and non-cleared neural tissue. *Biomedical Optics Express* **7**, 4021–4033 (2016). URL https://www.osapublishing.org/abstract.cfm?uri=boe-7-10-4021.
8. Fahrbach, F. O., Simon, P. & Rohrbach, A. Microscopy with self-reconstructing beams. *Nature Photonics* **4**, 780–785 (2010). URL https://www.nature.com/articles/nphoton.2010.204.
9. Chen, B.-C. *et al.* Lattice light-sheet microscopy: Imaging molecules to embryos at high spatiotemporal resolution. *Science* **346**, 1257998 (2014). URL http://science.sciencemag.org/content/346/6208/1257998.
10. Yang, Z. *et al.* A compact Airy beam light sheet microscope with a tilted cylindrical lens. *Biomedical Optics Express* **5**, 3434–3442 (2014). URL https://www.osapublishing.org/abstract.cfm?uri=boe-5-10-3434.
11. Arlt, J. & Dholakia, K. Generation of high-order Bessel beams by use of an axicon. *Optics Communications* **177**, 297–301 (2000). URL https://www.sciencedirect.com/science/article/pii/S0030401800005721.
12. Wijesinghe, P. & Dholakia, K. Emergent physics-informed design of deep learning for microscopy. *Journal of Physics: Photonics* (2021). URL http://iopscience.iop.org/article/10.1088/2515-7647/abf02c.
13. Weigert, M. *et al.* Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nature Methods* **15**, 1090–1097 (2018). URL https://www.nature.com/articles/s41592-018-0216-7.
14. Wang, H. *et al.* Deep learning enables cross-modality super-resolution in fluorescence microscopy. *Nature Methods* **16**, 103 (2019). URL https://www.nature.com/articles/s41592-018-0239-0.
15. Belthangady, C. & Royer, L. A. Applications, promises, and pitfalls of deep learning for fluorescence image reconstruction. *Nature Methods* **16**, 1215–1225 (2019). URL https://www.nature.com/articles/s41592-019-0458-z.
16. Goodfellow, I. *et al.* Generative Adversarial Nets. In Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N. D. & Weinberger, K. Q. (eds.) *Advances in Neural Information Processing Systems 27*, 2672–2680 (Curran Associates, Inc., 2014). URL http://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf.
17. He, K., Zhang, X., Ren, S. & Sun, J. Deep Residual Learning for Image Recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778 (2016). URL https://openaccess.thecvf.com/content_cvpr_2016/html/He_Deep_Residual_Learning_CVPR_2016_paper.html.
18. Isola, P., Zhu, J.-Y., Zhou, T. & Efros, A. A. Image-to-Image Translation with Conditional Adversarial Networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 5967–5976 (2017).
19. Ledig, C. *et al.* Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. *arXiv:1609.04802 [cs, stat]* (2017). URL http://arxiv.org/abs/1609.04802.
20. Simonyan, K. & Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv:1409.1556 [cs]* (2015). URL http://arxiv.org/abs/1409.1556.
21. Johnson, J., Alahi, A. & Fei-Fei, L. Perceptual Losses for Real-Time Style Transfer and Super-Resolution. In Leibe, B., Matas, J., Sebe, N. & Welling, M. (eds.) *Computer Vision – ECCV 2016*, Lecture Notes in Computer Science, 694–711 (Springer International Publishing, Cham, 2016).
22. Li, S. & Barbastathis, G. Spectral pre-modulation of training examples enhances the spatial resolution of the phase extraction neural network (PhENN). *Optics Express* **26**, 29340–29352 (2018). URL https://www.osapublishing.org/oe/abstract.cfm?uri=oe-26-22-29340.
23. Chen, J. *et al.* Three-dimensional residual channel attention networks denoise and sharpen fluorescence microscopy image volumes. *Nature Methods* **18**, 678–687 (2021). URL https://www.nature.com/articles/s41592-021-01155-x.
24. Zhang, H. *et al.* Exceeding the limits of 3D fluorescence microscopy using a dual-stage-processing network. *Optica* **7**, 1627–1640 (2020). URL https://www.osapublishing.org/optica/abstract.cfm?uri=optica-7-11-1627.
25. Guo, M. *et al.* Rapid image deconvolution and multiview fusion for optical microscopy. *Nature Biotechnology* **38**, 1337–1346 (2020). URL https://www.nature.com/articles/s41587-020-0560-x.
26. Weigert, M., Royer, L., Jug, F. & Myers, G. Isotropic reconstruction of 3D fluorescence microscopy images using convolutional neural networks. *arXiv:1704.01510 [cs]* (2017). URL http://arxiv.org/abs/1704.01510.
27. Richardson, W. H. Bayesian-Based Iterative Method of Image Restoration*. *JOSA* **62**, 55–59 (1972). URL https://www.osapublishing.org/josa/abstract.cfm?uri=josa-62-1-55.
28. Deng, M. *et al.* On the interplay between physical and content priors in deep learning for computational imaging. *Optics Express* **28**, 24152–24170 (2020). URL https://www.osapublishing.org/oe/abstract.cfm?uri=oe-28-16-24152.
29. Lucy, L. B. An iterative technique for the rectification of observed distributions. *The Astronomical Journal* **79**, 745 (1974). URL http://adsabs.harvard.edu/abs/1974AJ.....79..745L.
30. Ulyanov, D., Vedaldi, A. & Lempitsky, V. Deep Image Prior. *International Journal of Computer Vision* **128**, 1867–1888 (2020). URL https://doi.org/10.1007/s11263-020-01303-4.
31. Huisken, J. & Stainier, D. Y. R. Selective plane illumination microscopy techniques in developmental biology. *Development* **136**, 1963–1975 (2009). URL http://dev.biologists.org/content/136/12/1963.
32. Gosnell, M. E. *et al.* Quantitative non-invasive cell characterisation and discrimination based on multispectral autofluorescence features. *Scientific Reports* **6**, 23453 (2016). URL https://www.nature.com/articles/srep23453.
33. Xenopoulos, P., Kang, M. & Hadjantonakis, A.-K. Cell lineage allocation within the inner cell mass of the mouse blastocyst. *Results and Problems in Cell Differentiation* **55**, 185–202 (2012).
34. Richani, D., Dunning, K. R., Thompson, J. G. & Gilchrist, R. B. Metabolic co-dependence of the oocyte and cumulus cells: essential role in determining oocyte developmental competence. *Human Reproduction Update* **27**, 27–47 (2021). URL https://doi.org/10.1093/humupd/dmaa043.
35. Corsetti, S., Gunn-Moore, F. & Dholakia, K. Light sheet fluorescence microscopy for neuroscience. *Journal of Neuroscience Methods* (2018). URL http://www.sciencedirect.com/science/article/pii/S0165027018302188.
36. Corsetti, S. *et al.* Widefield light sheet microscopy using an Airy beam combined with deep-learning super-resolution. *OSA Continuum* **3**, 1068–1083 (2020). URL https://www.osapublishing.org/osac/abstract.cfm?uri=osac-3-4-1068.
37. Freund, T. F. & Buzsáki, G. Interneurons of the hippocampus. *Hippocampus* **6**, 347–470 (1996). URL https://onlinelibrary.wiley.com/doi/

abs/10.1002/%28SICI%291098-1063%281996%296%3A4%3C347%3A%3AAID-HIPO1%3E3.0.CO%3B2-I.
38. Kawaguchi, Y. & Kubota, Y. GABAergic cell subtypes and their synaptic connections in rat frontal cortex. *Cerebral Cortex* **7**, 476–486 (1997). URL https://doi.org/10.1093/cercor/7.6.476.
39. Gasparoli, F. M. *et al.* Is laser repetition rate important for two-photon light sheet microscopy? *OSA Continuum* **3**, 2935–2942 (2020). URL https://opg.optica.org/osac/abstract.cfm?uri=osac-3-10-2935.
40. Takanezawa, S., Saitou, T. & Imamura, T. Wide field light-sheet microscopy with lens-axicon controlled two-photon Bessel beam illumination. *Nature Communications* **12**, 2979 (2021). URL http://www.nature.com/articles/s41467-021-23249-y.
41. Xiao, L. *et al.* Network-enabled efficient image restoration for 3D microscopy of turbid biological specimens. *bioRxiv* 2020.05.27.118869 (2020). URL https://www.biorxiv.org/content/10.1101/2020.05.27.118869v1.
42. Yanny, K. *et al.* Deep learning for fast spatially varying deconvolution. *Optica* **9**, 96–99 (2022). URL https://opg.optica.org/optica/abstract.cfm?uri=optica-9-1-96.
43. Zhang, Y. *et al.* Computational optical sectioning with an incoherent multiscale scattering model for light-field microscopy. *Nature Communications* **12**, 6391 (2021). URL http://www.nature.com/articles/s41467-021-26730-w.
44. Wu, J. *et al.* Iterative tomography with digital adaptive optics permits hour-long intravital observation of 3D subcellular dynamics at millisecond scale. *Cell* **184**, 3318–3332.e17 (2021). URL https://www.sciencedirect.com/science/article/pii/S0092867421005328.
45. Bostan, E., Heckel, R., Chen, M., Kellman, M. & Waller, L. Deep phase decoder: self-calibrating phase microscopy with an untrained deep neural network. *Optica* **7**, 559–562 (2020). URL https://www.osapublishing.org/optica/abstract.cfm?uri=optica-7-6-559.
46. Gustafsson, N. *et al.* Fast live-cell conventional fluorophore nanoscopy with ImageJ through super-resolution radial fluctuations. *Nature Communications* **7**, 12471 (2016). URL https://www.nature.com/articles/ncomms12471.
47. Mao, X. *et al.* Least Squares Generative Adversarial Networks. *arXiv:1611.04076 [cs]* (2017). URL http://arxiv.org/abs/1611.04076.
48. Albuquerque, I. *et al.* Multi-objective training of Generative Adversarial Networks with multiple discriminators. *arXiv:1901.08680 [cs, stat]* (2019). URL http://arxiv.org/abs/1901.08680.
49. Deng, M., Goy, A., Li, S., Arthur, K. & Barbastathis, G. Low Photon Budget Phase Retrieval with Perceptual Loss Trained Deep Neural Networks. *arXiv:1906.05687 [eess]* (2019). URL http://arxiv.org/abs/1906.05687.
50. Lyngholm, D. & Sakata, S. Cre-Dependent Optogenetic Transgenic Mice Without Early Age-Related Hearing Loss. *Frontiers in Aging Neuroscience* **11** (2019). URL https://www.frontiersin.org/articles/10.3389/fnagi.2019.00029/full.