

# Applications, promises, and pitfalls of deep learning for fluorescence image reconstruction

Chinmay Belthangady and Loic A. Royer  \*

**Deep learning is becoming an increasingly important tool for image reconstruction in fluorescence microscopy. We review state-of-the-art applications such as image restoration and super-resolution imaging, and discuss how the latest deep learning research could be applied to other image reconstruction tasks. Despite its successes, deep learning also poses substantial challenges and has limits. We discuss key questions, including how to obtain training data, whether discovery of unknown structures is possible, and the danger of inferring unsubstantiated image details.**

Fluorescence microscopy is an indispensable tool in the biologist's arsenal that has enabled the systematic spatiotemporal dissection of life's molecular machines<sup>1</sup>. In all microscopy modalities, the resolution and quality of the images obtained is fundamentally limited by the optics, the photophysics of molecular probes, and sensor technology. Recent developments such as super-resolution microscopy can allow researchers to circumvent these limits by means of clever experimental strategies followed by image reconstruction. In the case of photo-activated localization microscopy (PALM)<sup>2</sup> and stochastic optical reconstruction microscopy (STORM)<sup>3</sup>, hundreds of images need to be processed to reconstruct a super-resolved image<sup>4</sup>. Computation is thus increasingly becoming an essential component of the imaging process. Any substantial improvements in the algorithms used for reconstruction would not only improve image quality but also open the door to the development of new imaging modalities.

In recent years, deep learning<sup>5</sup> and, in particular, deep convolutional neural networks (CNNs) have had stunning successes, surpassing human performance for hard problems such as visual<sup>6</sup> and speech recognition<sup>7</sup> (Box 1). Almost every discipline of science and engineering has been impacted, from astronomy<sup>8</sup> and biology<sup>9</sup> to high-energy physics<sup>10</sup>. CNNs are particularly relevant in fields that produce or interpret images. For example, medical imaging picked up the trend early and applied deep CNNs to image processing of computed tomography and magnetic resonance images<sup>11,12</sup>. Fluorescence microscopy, too, has recently benefited from these advances, with deep CNNs applied to image restoration<sup>13</sup>, deconvolution<sup>14,15</sup>, super-resolution<sup>16–18</sup>, translation between label-free and fluorescence images<sup>19,20</sup>, and 'virtual' staining<sup>21</sup>, as well as image segmentation, classification, and phenotyping<sup>22</sup>.

## Deep learning for image reconstruction

A common denominator between applications mentioned above is the formulation of image reconstruction as a transformation between acquired and reconstructed images. Classical approaches to image reconstruction handcraft these transformations from first principles (Box 2). For example, in deconvolution, this requires precise understanding of the optics and well-characterized noise statistics. This has led, for example, to the design of popular algorithms such as Richardson–Lucy deconvolution, which require knowledge of the point spread function of the microscope and assume Poisson noise statistics<sup>23</sup>. However, such handcrafted algorithms are limited

by the accuracy of their assumptions. They do not capture the full statistical complexity of microscopy images. Hence, data-driven approaches may in some cases be more broadly applicable and better suited than analytical ones to solve image reconstruction problems. Ideally, scientists want generic reconstruction algorithms that can be trained from exemplary data obtained either experimentally or from simulation. This is how deep learning comes into play. Deep neural networks can learn end-to-end image transformations from data without the need for explicit analytical modeling (Box 3).

## Fluorescence microscopy image reconstruction

Image reconstruction in fluorescence microscopy encompasses image denoising, deconvolution, registration, stitching, fusion, and super-resolution, as well as challenges brought by more exotic acquisition modalities such as tomographic, light-field, lens-less, and deep-tissue imaging. Here we review existing applications of deep learning in fluorescence microscopy and suggest promising new ones.

**Image denoising.** Image noise in fluorescence imaging is typically caused by the combined effect of Poisson statistics from photon-counting, thermal and readout noise of the sensor, and background autofluorescence. Early on, denoising algorithms were recognized as an opportunity to speed up acquisition by using shorter exposures and reduce photodamage by decreasing illumination intensity<sup>24</sup>. Algorithms were specially developed for fluorescence images that leveraged compressed sensing<sup>25</sup>, image self-similarity<sup>24,26</sup>, and thresholding after wavelet transform<sup>27</sup>. Today, state-of-the-art classical denoising algorithms in computer vision research use sparse coding techniques<sup>28</sup>, low-rank decomposition<sup>29</sup>, or image self-similarity<sup>30,31</sup>. The lesson learned over time is that to effectively denoise an image, it is better to look beyond small patches of pixels and gather information across multiple image regions and ideally learn from large sets of images. In 2009, Jain et al.<sup>32</sup> trained a five-layer CNN to denoise natural images, achieving a lower error rate and processing two orders of magnitude faster than that of previous approaches. Since then, numerous deep CNN architectures have been proposed for denoising natural images<sup>33</sup>. Recently, Weigert et al.<sup>13</sup> introduced content-aware image restoration (CARE), a method that trains networks with low signal-to-noise ratio (input) and high signal-to-noise ratio (target) image pairs. They demonstrate effective denoising of fluorescence microscopy images

**Box 1 | The deep learning revolution**

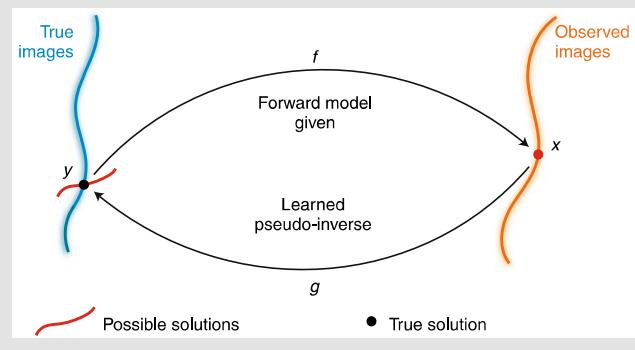
In 2012, Krizhevsky, Sutskever, and Hinton published their now famous AlexNet paper<sup>89</sup>. In that work, they showed that a deep convolutional neural network (CNN) could win the ImageNet Large-scale Visual Recognition Challenge (ILSVRC) by an unprecedented margin. AlexNet was trained on 1.2 million high-resolution training examples to classify images into 1,000 object categories. The double-digit decrease in error rate achieved by Hinton's team astounded the computer vision community. One year later, most competing teams used CNNs for what was considered at the time to be the most challenging machine learning competition (ILSVRC). Since 2012, deep CNNs have continued to break records, going from an error rate of 15.4% (ref. <sup>89</sup>) in 2012 to just 3.08% (ref. <sup>6</sup>) in 2018, surpassing human accuracy. Now, deep CNNs are the state of the art for major machine learning challenges on object<sup>6</sup> and speech recognition<sup>7</sup>. Recent notable artificial intelligence successes critically depended on deep CNNs, but also on other deep learning models such as deep reinforcement learning, as with DeepMind's AlphaGo victory against the second best human go player<sup>84</sup>. This has contributed to the enthusiasm, but also hype, surrounding deep learning.

acquired under conditions with low light or short camera exposures (Fig. 1a). CARE maintains image quality while decreasing light exposure 60-fold and outperforms classical denoising such as nonlocal means<sup>30</sup> or block-matching and three-dimensional (3D) filtering<sup>31</sup>. Such deep-learning-based restoration can be useful when specimens are light sensitive, when fluorophore bleaching is of concern, or to study fast dynamics in live samples. Yet a major obstacle to applying deep learning for image denoising is the need for matched pairs of high-noise and low-noise training images. This requirement can be relaxed, as shown by recent developments. Self-supervision approaches such as noise2noise<sup>34,35</sup>, noise2self<sup>36</sup>, and noise2void<sup>37,38</sup> leverage statistical independence of noise across pixels. Another approach, deep image prior<sup>39</sup>, uses a generative model trained per image. Also promising are new, unpaired image transformation schemes such as cycle generative adversarial network (cycleGAN)<sup>40</sup> (Fig. 2f). We expect that these advances will broaden the applicability and facilitate the adoption of CNN-based denoising in fluorescence microscopy.

**Spatial deconvolution.** The resolution of images obtained by fluorescence microscopy is fundamentally limited by the optical parameters of the microscope used. Image deconvolution can improve image quality by enhancing high-frequency details. In the case of 3D fluorescence microscopy, this is particularly useful for reducing axial blur and out-of-focus light<sup>41,42</sup>. Classic approaches such as Richardson–Lucy deconvolution<sup>23</sup> are still used extensively and have been extended to multi-view light-sheet microscopy<sup>43</sup>. Recent open-source packages such as DeconvolutionLab2<sup>44</sup> have helped to democratize advanced methods for deconvolution in fluorescence microscopy. Challenges remain: when imaging large samples, the point spread function (PSF) varies in space and across views, and even possibly over time. Shajkofci et al.<sup>15</sup> recently showed how a CNN can be used to estimate optical aberrations and improve image quality compared with that achieved with other deconvolution algorithms. It would be interesting to investigate whether the same approach can be applied to volumetric imaging. Another interesting opportunity would be to address challenges that are opto-mechanical in nature: in light-sheet microscopy there is a limited choice of illumination and detection objectives that can be placed orthogonal to each other. To solve this problem, one could use lower-numerical-aperture objectives that mechanically fit and apply deep learning to computationally increase the detection numerical aperture<sup>45</sup>.

**Box 2 | Image reconstruction, inverse problems, and deep learning**

In most imaging modalities, there is a well-founded understanding of the physics behind the image formation process. People can write formulas that describe, and code that simulates, the way light propagates through optical elements to eventually form an image on a detector; this is called the forward model  $f$  (see figure). This model is a transformation from the desired ideal image  $y$  to observed images  $x$ . Observed images  $x$  are incomplete, degraded, or convoluted compared with  $y$ . For example, image noise, low-pass filtering, pixel-value quantization, and subsampling cause partial and often irrecoverable loss of information. In the opposite direction, reconstructing the true image  $y$  from the observed images  $x$  is a challenging inverse problem. Finding the inverse of  $x$  necessitates a pseudo-inverse function  $g$  that is classically implemented as an iterative algorithm. One obstacle to inversion is that forward processes are often stochastic. Another fundamental difficulty is that the forward model maps many images to the same observation, and so there is no well-defined inverse but instead multiple confounding solutions (red line in the figure). However, in most cases, there is a priori knowledge about the validity of solutions, which makes it possible to select one inverse. Until recently, this prior information about the solution had to be mathematically formulated by experts and engineered into algorithms. These handcrafted priors typically prescribe, for example, that correct solutions should be smooth (Tikhonov's and total variation regularizations) or admit a concise representation in some basis (wavelet basis, dictionary)<sup>108</sup>. However, the statistical structure of real microscopy images is far more complex, thus limiting the quality of images reconstructed with such priors. This led to the advent of data-driven inversion schemes that learn a pseudo-inverse  $f$  from large datasets of  $(x, y)$  pairs. These pairs can be obtained by simulation, that is, by computation of  $f$ , or can be obtained empirically by means of clever experimental strategies that produce both  $x$  and  $y$  (ref. <sup>13</sup>). The learning procedure learns not only the function  $g$  but also, implicitly, the prior knowledge about 'good' solutions. Deep learning is one promising data-driven approach for solving inverse problems and, by extension, image reconstruction tasks<sup>109</sup>.



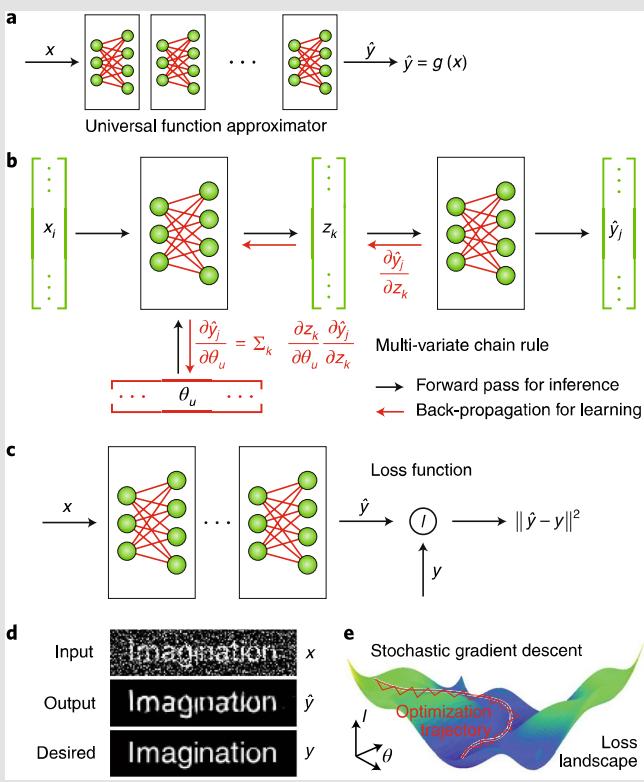
cal-aperture objectives that mechanically fit and apply deep learning to computationally increase the detection numerical aperture<sup>45</sup>.

**Axial inpainting.** A pervasive problem in microscopy is poor axial sampling and resolution that leads to 3D image anisotropy, that is, lower resolution along the objective axis ( $z$ ) than in the transverse ( $x-y$ ) planes. To address this, Weigert et al.<sup>13,14</sup> recently demonstrated a self-supervised training strategy, Isonet, which can deconvolve axial images ( $xz$  and  $yz$ ) to restore isotropic resolution and substantially outperforms the popular Richardson–Lucy algorithm.

**Box 3 | Deep learning as differential learning**

In deep learning, nonlinear parameterized processing modules are combined to progressively transform an input  $x$  into the desired output  $y$ , typically attaining only approximations  $\hat{y}$  (panels **a** and **d**). The adjective “deep” refers to the large number of stacked modules required for building a universal function approximator  $g$  (panel **a**). The learning process begins with a random choice for the parameters  $\theta$  of the modules, which are iteratively updated to improve the accuracy of the reconstruction. This is possible because each module is differentiable, meaning that it is known exactly how changes in parameter values  $\theta_k$  cause changes in output values (panel **b**). More formally, the partial derivatives  $\partial\hat{y}/\partial\theta$  of the output  $\hat{y}$  are known with respect to any parameter  $\theta$ . Similarly, one can compute the partial derivatives  $\partial z_k/\partial\theta$  for intermediate values  $z_k$  (panel **b**). Artificial neural networks such as CNNs are the most popular implementation of such modules. Networks with a sufficient number of parameters, and at least three layers, can theoretically approximate any function<sup>110</sup>.

The back-propagation algorithm<sup>111</sup> uses the chain rule to efficiently compute all partial derivatives, or gradients, with just one forward pass through the network followed by a backward pass. The discrepancy between the desired output  $y$  and the actual output  $\hat{y}$  is quantified using a loss function  $l(\hat{y}, y)$  (panel **c**). The parameters  $\theta$  are updated to minimize this loss using stochastic gradient descent (panel **e**). The speed provided by computation of the updates on small batches of data, in parallel, on specialized hardware, allows one to fit networks with millions of parameters on datasets with millions of observations.



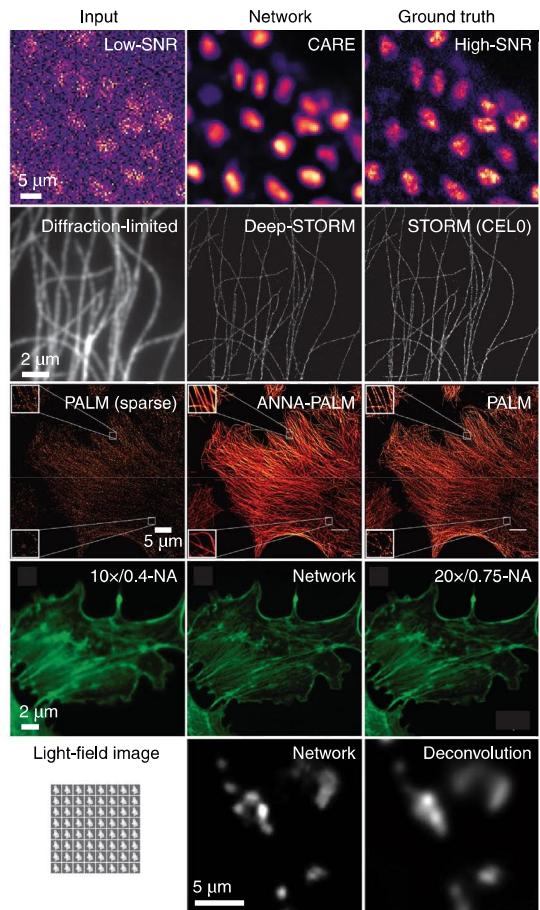
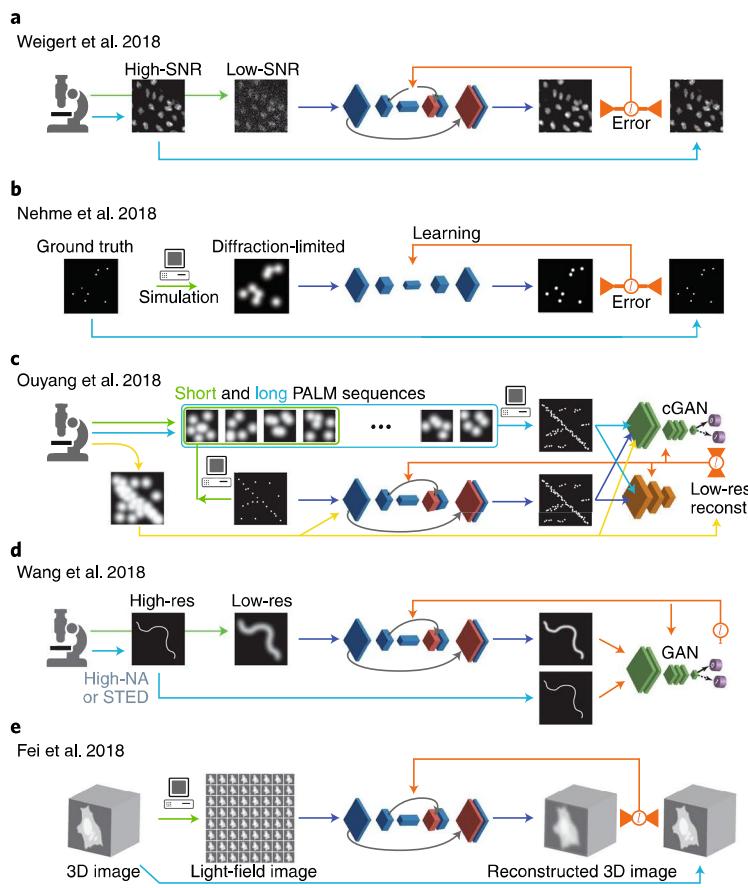
**Super-resolution.** PALM<sup>2</sup> and STORM<sup>3</sup> achieve super-resolution by localizing sparse fluorophore emissions in thousands of images. These images are an intermediate product that must be processed to reconstruct a super-resolved image of fluorophore locations<sup>46,47</sup>. A technique called deep-STORM<sup>18</sup> uses an encoder-decoder network

(Box 4) to localize emitters (Fig. 1b). The authors used simulated pairs of diffraction-limited and super-resolved images to train their network and obtain high emitter detection efficiency on both real and simulated images. Instead of predicting high-resolution images, DeepLoco<sup>48</sup> demonstrates direct prediction of emitter coordinates. In another application of deep learning to super-resolution microscopy, Ouyang et al.<sup>17</sup> trained a U-Net (Box 4) fed with sparsely sampled PALM images to predict their densely sampled counterparts (Fig. 1c). In contrast with the mean-squared-error training loss of deep-STORM, their artificial neural network accelerated (ANNA)-PALM method uses an ingenious combination of  $L_1$ , structural similarity (SSIM) and conditional adversarial losses<sup>49,50</sup>. Using a similar network architecture and training loss, Wang et al.<sup>16</sup> recently demonstrated that the resolution of diffraction-limited confocal images can be enhanced (Fig. 1d). This work shows that use of structured losses, whether explicit (SSIM) or adversarially learned (conditional GAN (cGAN); Box 4), result in output images that are sharper and of better perceptual quality<sup>50</sup>. For super-resolved fluorescence imaging on live cells, algorithms must cope with high emission densities, which has led to new approaches capable of handling higher densities, such as super-resolution radial fluctuations (SRRF)<sup>51</sup>. We expect that the use of CNNs in this challenging regime could improve image quality and temporal resolution for live cell imaging.

**Structured illumination.** Another super-resolution approach is structured illumination microscopy (SIM). It can surpass the diffraction limit twofold by computational synthesis of images acquired by shifted illumination patterns<sup>52,53</sup>. However, current algorithms require precise knowledge of the effective pattern and are thus sensitive to distortions or attenuation caused by the sample. Although classical algorithms exist that can handle low-intensity<sup>54</sup>, unknown<sup>55</sup>, or distorted<sup>56</sup> illumination patterns, it would be interesting to compare these with a CNN-based SIM reconstruction algorithm (Fig. 2a), possibly leveraging ideas similar to magnetic resonance imaging reconstruction with automated transform by manifold approximation (AUTOMAP)<sup>57</sup>. The training data could be generated by simulations that would consider all relevant degradations and distortions. We believe that such an algorithm could help scientists obtain better SIM reconstructions under challenging signal-to-noise conditions.

**Spectral deconvolution.** The trend in fluorescence imaging is for faster acquisition with more colors. However, fluorescent proteins and dyes used for labeling have broad, overlapping emission spectra, which leads to channel mixing that complicates interpretation. One solution that sacrifices speed is to acquire the channels sequentially. Another solution is to acquire the channels simultaneously, but this requires the separation of channels with spectral unmixing algorithms. To address this problem, algorithms such as linear unmixing<sup>58</sup> and phasor approaches<sup>59</sup> have been devised to deconvolve these spectra into separate channels corresponding to distinct labels. In some sense, the unmixing of multicolor acquisitions is a deconvolution problem in the spectral dimension that is difficult in low signal-to-noise conditions. Could CNNs be applied to deconvolve more wavelengths in poor signal-to-noise conditions? One straightforward approach to obtain training data would be to acquire each channel sequentially and then simultaneously (Fig. 2b), or to use physics-based models to generate synthetic training data. Other, more creative uses of deep learning for multicolor microscopy can be envisioned: Hershko et al.<sup>60</sup> recently showed how CNNs can leverage chromatic effects on the PSF to determine the color of single emitters from a grayscale image.

**Corrective uneven illumination and detection.** Unfortunately, in most microscopes, illumination and detection efficiency are not uniform over the entire field of view. In widefield and spinning disk confocal microscopes<sup>61</sup>, uneven detection efficiency



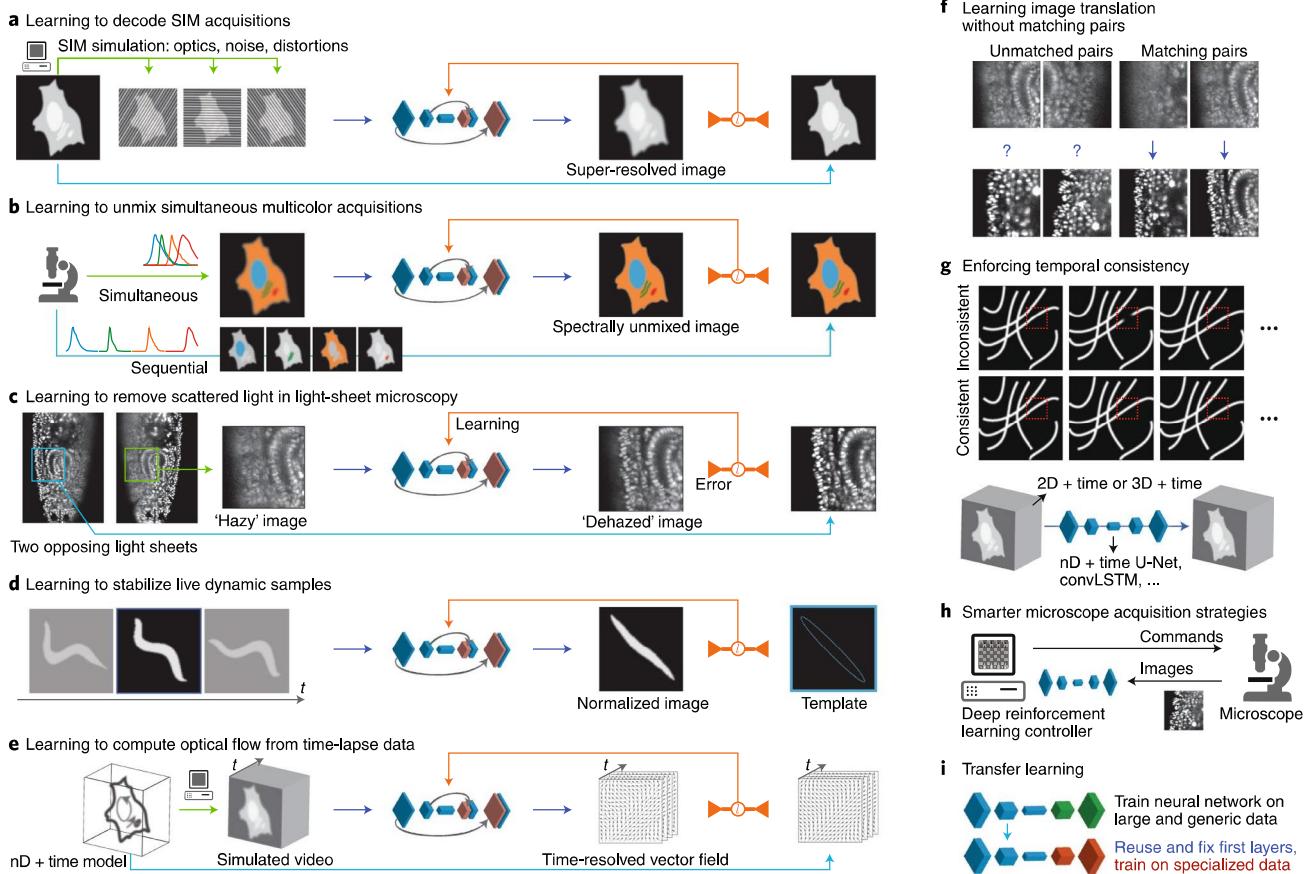
**Fig. 1 | Review of current applications of deep learning to fluorescence microscopy.** **a**, CARE for denoising, super-resolution, and axial deconvolution<sup>13,14</sup>. **b**, Deep-STORM for STORM super-resolution<sup>18</sup> compared with a non-deep-learning technique (CELO<sup>17</sup>). **c**, ANNA-PALM for accelerated PALM super-resolution<sup>17</sup>. **d**, Computational fluorescence super-resolution by learning to translate from low- to high-numerical-aperture (NA) images<sup>16</sup>. **e**, Fluorescence light-field reconstruction for high-speed, high-resolution volumetric imaging<sup>73</sup>.

often requires nonuniformity correction. In the case of mesoscale light-sheet microscopy, the problem is far more complex because illumination shadows are caused by sample-induced occlusion<sup>62</sup>. We believe that CNNs are well positioned to address this problem. Promising CNN-based techniques exist in computer vision for context-based inpainting of missing image regions<sup>63,64</sup>. Fortunately, in light-sheet microscopy, information is not completely absent because the occlusion is often only partial. Moreover, in multiview imaging, training data are readily available in the form of images acquired from different detection arms or with light sheets of distinct propagation angles.

**Scattered light haze reduction.** Scattered light is a cause of substantial image-quality deterioration in live fluorescence imaging. For example, in light-sheet microscopy, background induced by scattered light is often suppressed to facilitate subsequent image processing steps such as image stitching and fusion<sup>65</sup>. Yet many challenges remain because scattering is depth sensitive and sample dependent. Recently, powerful hybrid approaches have been developed that combine CNNs and optical modeling for dehazing natural scenes<sup>66</sup>. Could such ideas be applied to microscopy to further improve the quality of images? In multiview light-sheet microscopy, an obvious way to obtain training data is to collect low- and high-haze image pairs from overlapping regions obtained from distinct views (Fig. 2c). In some cases, matching image pairs cannot be obtained, such as when one is imaging deep within a sample using

a spinning-disk confocal microscope. In such instances, recent advances in unpaired image-to-image translation<sup>40</sup> could be used to learn to dehaze images using images at the surface and deeper in the sample (Fig. 2f).

**Registration, stitching, and fusion.** Because of the limited field of view of objectives and large size of some specimens, it is often necessary to break down an acquisition into 2D or 3D tiles<sup>43</sup>. Reconstructing a complete image requires stitching the tiles together after registration to a common reference frame. State-of-the-art algorithms<sup>43,67</sup> rely on iterative schemes that leverage direct image cross-correlation or matching fiducial markers. One obvious application of deep learning to image registration is to teach CNNs to find image correspondences<sup>68</sup>, and thus have a deep learning component be part of a larger algorithm. However, one key promise of deep learning is its ability to train complex models solving a sequence of tasks in an end-to-end fashion. Recently, Rohe et al.<sup>69</sup> showed that pairs of large 3D medical images can be nonrigidly registered in less than 30 ms with a fully convolutional CNN that directly outputs a deformation vector field. Their method is more accurate and faster than traditional approaches and does not require iterations of an optimizer, but it does require a large amount of training data. Similarly, Nguyen et al.<sup>70</sup> demonstrated an unsupervised approach for nonrigid transformation estimation that is faster and more accurate than classical correlation-based or fiducial-based approaches. We expect that these ideas will eventually percolate into the field of



**Fig. 2 | Potential applications of deep learning in fluorescence microscopy and key concepts.** **a**, Learning to reconstruct super-resolved images from structured illumination (SIM) acquisitions. **b**, Learning spectral unmixing of simultaneous multi-color acquisitions. **c**, Learning to reduce scattered-light ‘haze’ in light-sheet microscopy. **d**, Learning to stabilize live dynamic samples with weak supervision against a template shape. **e**, Fluorescence optical flow with deep learning. Computing vector fields from fluorescence spatiotemporal fluctuations with simulations. **f**, Unsupervised learning to reduce or obviate the need for curation and ground truth (for example, CycleGANs<sup>40</sup>). **g**, Leveraging the time dimension by enforcing temporal consistency in predictions to further improve reconstruction. **h**, Deep reinforcement learning to drive smart microscopes capable of conducting complex imaging experiments. **i**, Transfer learning to facilitate training and reduce the need for specialized curated datasets.

fluorescence imaging. An upcoming and exciting challenge is imaging of live moving animals. For example, *Caenorhabditis elegans* larvae twist incessantly before hatching, making the analysis of time-lapse data challenging. This requires specialized stabilization algorithms capable of unfolding, straightening, and registering worm images<sup>71</sup>. Can a CNN learn to stabilize 3D time-lapse images of highly dynamic samples against a fixed template shape (Fig. 2d)?

**Light-field, lens-less imaging and tomography.** Recently, exotic forms of fluorescence imaging have emerged that pose novel computational challenges. For example, light-field imaging can multiplex an entire volume into a single 2D image acquisition, thus enabling rapid functional imaging at single-neuron resolution<sup>72</sup>. Preliminary work by Fei et al.<sup>73</sup> revealed results on CNN-based light-field 3D image reconstruction of live *C. elegans* worms (Fig. 1e). Similarly, lens-less imaging schemes such as DiffuserCam achieve one-shot 3D imaging by replacing the objective lens with a diffusing element placed in front of the sensor<sup>74</sup>. Tomography is another indirect approach to obtain 3D images from 2D acquisitions that is particularly relevant for mesoscopic fluorescence imaging<sup>75</sup>. In both cases the final 3D image is computationally reconstructed from a single 2D image. We believe that both lens-less imaging and tomography will benefit from deep-learning-based algorithms, just as has been demonstrated for light-field imaging.

**Temporal consistency.** Applications of deep learning to microscopy typically do not capture temporal patterns: predictions are made independently, frame by frame. Therefore, successive reconstructed images are not necessarily temporally consistent and can exhibit artifacts that are clearly noticeable when examined across time. Explicit modeling of the temporal dimension and training with time-resolved data would lead to more accurate reconstructions. A straightforward approach for augmenting any CNN with temporal consistency is simply to treat time as an additional dimension (Fig. 2h). However, this can become intractable for large networks handling long-term correlations. Instead, a better alternative is to combine CNNs with state-of-the-art recurrent neural networks whose outputs are fed back as inputs and hence are specifically designed for sequence prediction. One obvious candidate is the convolutional long–short term memory (convLSTM) architecture, which combines CNNs for space and LSTMs for time, and has been applied to weather forecasting<sup>76</sup>.

**Beyond fluorescence.** In some cases, fluorescence measurements are a proxy for other quantities. For example, measurement of material flow by fluorescence is an important tool in the emerging field of cell mechanobiology<sup>77</sup>. Unfortunately, current hand-crafted optical flow algorithms typically do not produce smooth flow fields in challenging imaging conditions. CNN-based optical flow estimation algorithms recently developed for self-driving cars<sup>78</sup> could

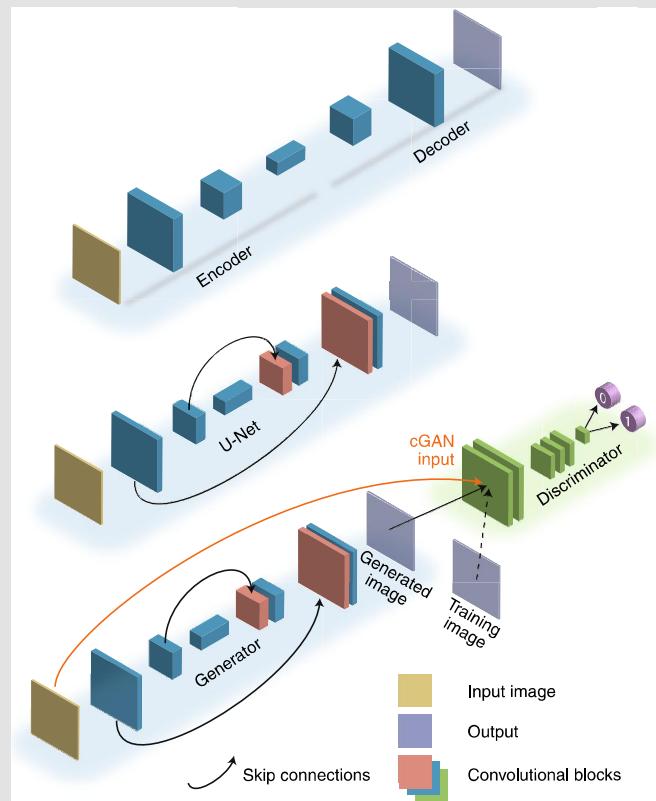
**Box 4 | Architectures for image reconstruction**

Image restoration or reconstruction requires neural networks that map one or several input images to an output image. Input and output images may be 2D or 3D and may even have different dimensions. The encoder–decoder convolutional network architecture<sup>112</sup> is one of the simplest networks capable of image-to-image translation. First, the input image is successively and repeatedly downsampled and convolved, and is also subjected to other normalization and nonlinear operations typical for CNNs. At every step, the image dimensions are reduced while the number of image channels—an additional dimension—is increased. This creates a representation bottleneck that is then reversed with successive and repeated upsampling and convolution steps. The bottleneck forces the network to encode the image into a small set of abstracted variables (often called latent variables) along the channel dimension, which are then decoded by the second half of the network. This ensures that only the most relevant features from the input images are retained and those that are deemed unessential to the reconstruction are discarded.

The U-Net<sup>113,114</sup> is a variation of the encoder–decoder network. It adds shortcut or skip connections between the encoder and decoder branches. These additional connections are beneficial for image restoration tasks because they send fine image details directly to the decoder pass, skipping the bottleneck. For image restoration tasks in which the identity function is a good first guess, residual connections that add the input image to the output image can be used to learn a residual mapping. Overall, skip and residual connections have been shown to help train deeper networks by preventing vanishing gradients<sup>115</sup>.

In the GAN<sup>116</sup> framework, two networks, a generator and a discriminator, are trained together. In applications to microscopy, an improvement on GANs is often used: cGAN. The generator learns to transform an input image into an output image while the discriminator learns to classify images as real (from the training set; label 1) or counterfeit (from the generator; label 0). The input image is provided to the discriminator<sup>50</sup>. Therefore, cGANs can be understood as providing an implicitly learned loss function

defined by the action of the discriminator. In practice, one achieves this behavior by training the generator to minimize the loss function and training the discriminator to maximize it. At the end of training, the discriminator network is dropped and only the generator is used for image reconstruction.



be a source of inspiration for designing more robust algorithms (Fig. 2e). Another interesting challenge is registering electron microscopy (EM) and fluorescence microscopy images. This task often requires the placement of fiducial markers visible in both modalities<sup>79</sup>. Recent work on virtual fluorescence labeling<sup>19–21</sup> suggests that it is possible to automatically translate images across modalities. Unsupervised approaches for learning image correspondence such as cycleGAN<sup>40</sup> could be applied for image registration in correlative imaging<sup>20</sup>.

**Deep learning for adaptive and smart imaging.** State-of-the-art microscopes—in particular light-sheet microscopes—are becoming adaptive machines capable of closed-loop control and real-time image analysis<sup>80,81</sup>. Yet current adaptive microscopes rely on time-consuming and complex iterative schemes<sup>80</sup> or direct measurements<sup>81</sup>. What if one could design a system that quickly and correctly guesses the best imaging parameters from few measurements? This is an inverse problem, because it may be easy to deduce the consequence of a particular choice of parameters but difficult to determine which choice will lead to the desired image. A few limited examples already show, for example, how to (1) train a CNN to focus light deep within turbid media<sup>82</sup>, (2) engineer illumination patterns for optimal imaging<sup>83</sup>, and (3) design PSFs for optimal spectral separation<sup>80</sup>. Yet the true challenge going forward will be to

optimize multistep imaging strategies: the microscope will observe, make moves that change imaging parameters, and thus play a game with the ultimate goal of gaining the most information about the sample (Fig. 2i). Deep reinforcement learning—another key concept behind DeepMind’s AlphaGo success<sup>84</sup> (Box 1)—could be used to automatically learn such strategies. Ultimately, we expect that most of the human decision-making required for imaging optimization will be done automatically, something that will also necessarily require extensive end-to-end robotic automation so that the ‘game’ can be restarted at will.

**Limits and pitfalls**

**Deep learning versus classical methods.** Although classical principled algorithms are often outperformed by deep learning models, they also retain key advantages. First, classical algorithms inherently produce outputs that are consistent with their inputs because they rely exclusively on first principles formulated as explicit analytical models instead of being trained from data (see section “The hallucination problem” and Box 4). Second, classic algorithms generalize to any valid measurement because they are not limited by the adequacy of the training data (see section “The generalization problem”). Third, classic algorithms—in contrast to deep learning models—do not produce widely erratic outputs after very minute tweaks to their inputs (see section “The adversarial fragility problem”).

**The hallucination problem.** The most serious issue when applying deep learning for discovery is that of hallucination. When looking at random patterns such as clouds, human brains can perceive shapes of objects and animals. Similarly, deep learning systems can hallucinate details and make mistakes when provided with inadequate training data (Box 5). These hallucinations are deceptive artifacts that appear highly plausible in the absence of contradictory information and can be challenging, if not impossible, to detect. Some network architectures, such as GANs, are particularly susceptible to hallucination because their explicit goal is to fool a discriminator by forging persuasive details. The cGANs (Box 4) as used by Isola et al.<sup>50</sup> alleviate this problem but may still produce unsubstantiated image details. A possible mitigation strategy is to add additional consistency losses to prevent hallucinations, as done in ANNA-PALM<sup>17</sup>. In general, mistakes made by deep learning models can be highly plausible and more subtle than those of classical algorithms, a potential cause for concern when such models are used for discovery.

**The generalization problem.** Another failure mode of neural networks is overlearning or overfitting: the neural network memorizes the training data exactly and fails to generalize to unseen data. This can be caused by insufficient training data, or by a poor choice of network parameters. Architectural features such as drop-out regularization can help, but the best way to avoid overfitting is to train with data that exceed the memorization capacity of the networks. Overall, current deep learning approaches typically do not generalize well to data obtained from a different microscope or under different conditions<sup>13</sup>.

**The adversarial fragility problem.** Recent research such as DeepFool<sup>85</sup> and the work by Sabour et al.<sup>86</sup> shows that neural networks can be tricked into producing completely different outputs after the application of imperceptible perturbations to their inputs. The modification of even a single pixel can be enough to fool deep neural networks into confusing different classes of object on the CIFAR-10 image classification dataset<sup>87</sup>. Overall, this fragility of deep learning in the face of adversarial attacks is concerning and raises many questions about the nature and brittleness of deep CNN computation (see “Interpreting and trusting the black box” below). Research is currently under way in the deep learning community to better understand adversarial attacks, as well as defenses<sup>88</sup>.

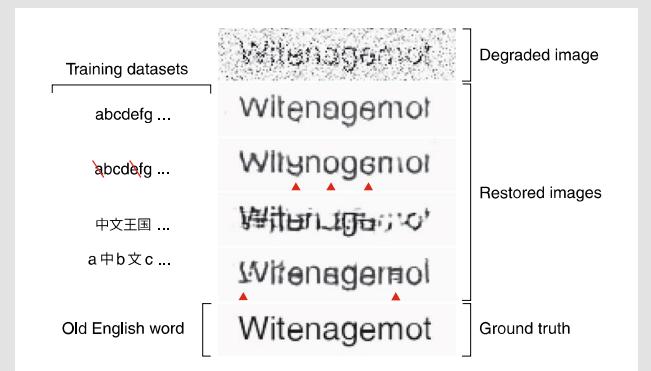
## Challenges

**What amount of training data is needed?** The success of deep learning depends critically on the availability of training data and their application to the correct category of images. If the amount of training data is not sufficient, poor performance will ensue (Box 5). Yet a common misconception is that deep learning always requires a very large number of training examples. For example, Wang et al.<sup>16</sup> used at most 3,000 image pairs for training, but Christiansen et al.<sup>19</sup> needed <100 high-resolution images (at most  $4,600 \times 4,600$  pixels), and Ounkomol et al.<sup>20</sup> used only 40 images ( $1,500 \times 1,500$  pixels). In some cases, the training and inference can be done on the very same 3D image stack, as shown by Weigert et al.<sup>14</sup>. This is in stark contrast with the millions of images needed for the ImageNet challenge<sup>89</sup>. Indeed, far fewer training data are required in fluorescence microscopy than are needed to distinguish between, for example, 800 different kinds of birds in natural images. In fact, the quality of the data and their suitability for the problem are perhaps more important (Box 5). In any case, fluorescence microscopy needs creative experimental and computational strategies to obtain more and better training data.

**Strategies for obtaining training data.** One approach is to perform dedicated experiments that produce the necessary images for training. For example, Weigert et al.<sup>13</sup> acquired pairs of low-quality

## Box 5 | Deep-learning-based discovery and its limits

Does data-driven image reconstruction preclude discovery? Can novel structures and unknown patterns be discovered even if absent from the training data? To shed some light on this subtle but important question, we will use an analogy: let us train a neural network (U-Net) to restore a highly degraded image of an ancient English word: “Witenagemot.” First, we use a training dataset consisting of images of all three-letter substrings occurring in the most common English words. We exclude “Witenagemot” and its variants. As shown in the figure, the word is recovered. We discovered a word that is not present in the training data. Knowing which alphabet to expect does not preclude the discovery of new words, sentences, or ideas. What if we exclude the letters *a* and *e* from the training data? In that case, the network does not recognize these letters, and discovery of the word is hampered. If one uses the wrong alphabet for training—for example, 6,000 Chinese characters—discovery is impossible. The mistakes made are interesting in their own right: the network tries to reconcile the inadequate prior with the information present in the degraded image. Finally, we can rescue the restoration by using instead a 50%–50% mix of Chinese and Latin characters. However, we also see that some characters such as *m* are decorated with artifacts of Chinese provenance.



and high-quality images by varying the exposure and laser power. In cases where the physics of the image-degradation process is well understood, forward-model simulations can be used to generate realistic images<sup>13,18</sup>. Neural networks could also be used to improve the quality of these simulations. Recently, much effort has been invested in building generative models of cells<sup>90</sup> using adversarial generative models<sup>91–93</sup>. The images produced by these models could in turn be used to train restoration or reconstruction algorithms.

**Leveraging available training data.** A classic approach to increasing the number of training examples is data augmentation, which creates variants of existing images through image rotation, scaling, lighting, and other methods. Another way to leverage available data is transfer learning, which relies on the observation that the first layers of a neural network act as detectors of universally applicable textures and patterns<sup>94</sup> (Fig. 2g). Transfer learning consists in pre-training networks on large datasets from other domains, thus accelerating convergence and improving generalization<sup>95</sup>. As an example, Estava et al.<sup>96</sup> recently demonstrated how Google’s Inception network trained on ImageNet images can be retrained to classify dermatological images for melanoma classification. Instead of using ImageNet, could scientists use existing large collections of fluorescence microscopy images? Prime candidates are the cellular and organelle fluorescence images from the Human Protein Atlas<sup>97</sup> or

the Broad Bioimage Benchmark Collection<sup>98</sup>. Also, Sullivan et al.<sup>99</sup> combined crowd-sourced annotations with machine learning on an image-classification task. Could some of these ideas be applied to fluorescence image reconstruction? Or, at least, could such repositories be used for transfer learning?

**Interpreting and trusting the black box.** Another crucial challenge is that deep neural networks are black-box models whose operation is opaque and difficult to interpret. For deep learning to become a trusted component of the microscopy-based discovery pipeline, researchers need conceptual frameworks, and interactive graphical tools, to help explain how a particular result is attained. Fortunately, progress has been made within the computer vision community in this direction, with techniques developed to (1) better understand which parts of the input image are essential for prediction of the output<sup>100</sup>, (2) analyze the function of intermediate layers<sup>94</sup>, (3) tease apart contributions from different network layers with ablation studies<sup>94</sup>, (4) build hierarchical explanatory graphs across layers<sup>101</sup>, and (5) design network architectures that are interpretable by design<sup>102</sup>. These ideas need to be adapted to fluorescence imaging, and tools should be built to facilitate the interpretation of results. One possibility is to devise tools that would explain given predictions. Take, for example, low-contrast images of fluorescently labeled cell membranes that would be restored using a CNN. The user could select pixels along a predicted membrane and ask for an explanation—for example, in the form of a covariance map—that could be presented interactively to show which pixels and patterns in the input image justify the reconstruction and why. Yet it is not enough to explain how a result is attained; it is also necessary to quantify the confidence in these results. Recently, several methods for computing pixel-wise confidence measures and confidence intervals have been proposed that leverage an ensemble of predictors and adversarial training<sup>13,103,104</sup>. Another promising approach is to use internal consistency checks, for example, by comparing the input of the network to the output transformed by the forward model as done during inference in ANNA-PALM<sup>17</sup>.

**Is discovery possible?** An oft-asked question is, “Is it possible to reconstruct structures or patterns that are not present in the training data?” To illustrate this problem, we use the restoration of degraded text as an analogy for image reconstruction in microscopy (Box 5). What are the lessons to be learned? First, data-driven discovery is possible, but it depends critically on the quality and compatibility of the training data. There are also serious hallucination risks when the training data contain less or more information than strictly required. The same is true for microscopy: the recognition of image features does not preclude further analysis of their spatial organization or the study of their temporal dynamics. For example, being able to super-resolve microtubules with fewer raw image acquisitions makes it possible to make observations at higher temporal resolution and consequently opens the door to new discoveries<sup>17,18</sup>. However, mistakes can be made if the training data are inadequate. In light of these results, it is our opinion that deep learning results cannot be blindly trusted, something certainly true of any singular piece of evidence. Although confidence measures are needed<sup>13</sup>, discoveries must be confirmed from multiple lines of evidence gathered from distinct experiments.

**Reproducibility.** Reproducibility is the bedrock of modern science. It is only natural to expect independent validation of published deep learning results. However, a lack of experimental consistency and reporting guidelines sometimes prevents deep learning researchers from replicating each other’s results<sup>105</sup>. Moreover, the statistical significance of measured performance metrics is rarely reported in the published literature because of high costs—in terms of time and resources—associated with

training<sup>106</sup>. To ameliorate this, researchers should adopt guidelines when publishing their results. Such best practices could include, for example, (1) making the source code and trained models freely available, (2) allowing access to the dataset that was used to train the models, (3) full disclosure of all hyperparameters used during training, (4) reporting the statistical significance of observed results, and (5) reporting failure cases in addition to exemplary ones. Such approaches will increase confidence in the results and facilitate widespread adoption of deep learning techniques in fluorescence microscopy.

## Dissemination

To accelerate the adoption of deep learning in microscopy, new software frameworks tailored for biologists are needed to use, adapt, train, validate, and interpret deep neural networks<sup>13</sup>. In particular, tools with better ergonomics, user-friendly graphical interfaces, and smart parameter-free or auto-tuning algorithms would help lower the technical bar to adoption. Another potential obstacle is the cost of hardware for training. Training deep learning models today typically requires specialized hardware such as graphical processing units (GPUs). Using GPUs instead of standard central processing units (CPUs) leads to 100-times-faster training speeds. This typically reduces training times from weeks and days to hours. Whereas buying or assembling desktop computers outfitted with GPUs is relatively affordable, assembling and maintaining such a machine requires nontrivial technical skills. An alternative is pay-as-you-go cloud-based platforms, which are best reserved for short-term high-intensity workloads.

## Conclusion

Deep learning holds many promises for fluorescence microscopy. Some applications will stand the test of time; others will fall short. Some applications will be redefined by deep learning, whereas others will continue to rely on classical methods. In either case, learned algorithms will be key for upcoming computational advances in microscopy. Because of the compositionality of deep learning models, it is possible to train models that directly process acquired raw data and produce the final analysis product, such as segmentation and classifications. Hence, in the future, we expect that deep learning models trained end-to-end will blur the frontier between image reconstruction and image analysis.

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

## Data availability

Source code for the experiment described in Box 5 can be found at <http://github.com/royerlab/DLDiscovery>.

Received: 12 June 2018; Accepted: 22 May 2019;

Published online: 08 July 2019

## References

1. Lichtman, J. W. & Conchello, J.-A. Fluorescence microscopy. *Nat. Methods* **2**, 910–919 (2005).
2. Betzig, E. et al. Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **313**, 1642–1645 (2006).
3. Rust, M. J., Bates, M. & Zhuang, X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (storm). *Nat. Methods* **3**, 793–795 (2006).
4. Schermelleh, L. et al. Super-resolution microscopy demystified. *Nat. Cell Biol.* **21**, 72–84 (2019).
5. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
6. Szegedy, C., Ioffe, S., Vanhoucke, V. & Alemi, A. A. Inception-v4, inception-resnet and the impact of residual connections on learning. *AAAI* **4**, 12 (2017).

7. Zhao, H., Zarar, S., Tashev, I. & Lee, C.-H. Convolutional-recurrent neural networks for speech enhancement. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing* (eds. Hayes, M. et al.) 2401–2405 (IEEE, 2018).
8. Lam, C. & Kipping, D. A machine learns to predict the stability of circumbinary planets. *Mon. Not. R. Astron. Soc.* **476**, 5692–5697 (2018).
9. Ching, T. et al. Opportunities and obstacles for deep learning in biology and medicine. *J. R. Soc. Interface* **15**, 20170387 (2018).
10. Radovic, A. et al. Machine learning at the energy and intensity frontiers of particle physics. *Nature* **560**, 41–48 (2018).
11. Xu, Y. et al. Deep learning of feature representation with multiple instance learning for medical image analysis. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing* (eds. Gini, F. et al.) 1626–1630 (IEEE, 2014).
12. Jin, K. H., McCann, M. T., Froustey, E. & Unser, M. Deep convolutional neural network for inverse problems in imaging. *IEEE Trans. Image Process.* **26**, 4509–4522 (2017).
13. Weigert, M. et al. Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nat. Methods* **15**, 1090–1097 (2018).
14. Weigert, M., Royer, L., Jug, F. & Myers, G. Isotropic reconstruction of 3D fluorescence microscopy images using convolutional neural networks. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2017* (eds. Descoteaux, M. et al.) 126–134 (Springer, 2017).
15. Shajkofci, A. & Liebling, M. Semi-blind spatially-variant deconvolution in optical microscopy with local point spread function estimation by use of convolutional neural networks. In *2018 25th IEEE International Conference on Image Processing (ICIP)* (eds. Nikou, C. et al.) 3818–3822 (IEEE, 2018).
16. Wang, H. et al. Deep learning enables cross-modality super-resolution in fluorescence microscopy. *Nat. Methods* **16**, 103–110 (2019).
17. Ouyang, W., Aristov, A., Lelek, M., Hao, X. & Zimmer, C. Deep learning massively accelerates super-resolution localization microscopy. *Nat. Biotechnol.* **36**, 460–468 (2018).
18. Nehme, E., Weiss, L. E., Michaeli, T. & Shechtman, Y. Deep-STORM: super-resolution single-molecule microscopy by deep learning. *Optica* **5**, 458–464 (2018).
19. Christiansen, E. M. et al. In silico labeling: predicting fluorescent labels in unlabeled images. *Cell* **173**, 792–803 (2018).
20. 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).
21. Rivenson, Y. et al. Deep learning-based virtual histology staining using auto-fluorescence of label-free tissue. Preprint at <https://arxiv.org/abs/1803.11293> (2018).
22. Moen, E. et al. Deep learning for cellular image analysis. *Nat. Methods* <https://doi.org/10.1038/s41592-019-0403-1> (2019).
23. Richardson, W. H. Bayesian-based iterative method of image restoration. *JOSA* **62**, 55–59 (1972).
24. Carlton, P. M. et al. Fast live simultaneous multiwavelength four-dimensional optical microscopy. *Proc. Natl Acad. Sci. USA* **107**, 16016–16022 (2010).
25. Marim, M. M., Angelini, E. D. & Olivo-Marin, J.-C. A compressed sensing approach for biological microscopy image denoising. In *SPARS '09—Signal Processing with Adaptive Sparse Structured Representations* (eds. Gribonval, R. et al.) **inria-00369642** (IEEE, 2009).
26. Boulanger, J. et al. Patch-based nonlocal functional for denoising fluorescence microscopy image sequences. *IEEE Trans. Med. Imaging* **29**, 442–454 (2010).
27. Luisier, F., Blu, T. & Unser, M. Image denoising in mixed Poisson–Gaussian noise. *IEEE Trans. Image Process.* **20**, 696–708 (2011).
28. Xu, J., Zhang, L. & Zhang, D. A trilateral weighted sparse coding scheme for real-world image denoising. In *Computer Vision—ECCV 2018* (eds. Ferrari, V. et al.) 21–38 (Springer, 2018).
29. Yair, N. & Michaeli, T. Multi-scale weighted nuclear norm image restoration. In *Proc. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (eds. Brown, M. S. et al.) 3165–3174 (IEEE, 2018).
30. Buades, A., Coll, B. & Morel, J.-M. A non-local algorithm for image denoising. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)* (eds. Schmid, C., Soatto, S. & Tomasi, C.) 60–65 (IEEE, 2005).
31. Dabov, K., Foi, A., Katkovnik, V. & Egiazarian, K. Image denoising with block-matching and 3D filtering. In *Image Processing: Algorithms and Systems, Neural Networks, and Machine Learning* (eds. Nasrabadi, N. M. et al.) 606414 (International Society for Optics and Photonics, 2006).
32. Jain, V. & Seung, S. Natural image denoising with convolutional networks. In *Advances in Neural Information Processing Systems 21* (eds. Koller, D. et al.) 769–776 (NIPS, 2009).
33. Zhang, K., Zuo, W., Chen, Y., Meng, D. & Zhang, L. Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising. *IEEE Trans. Image Process.* **26**, 3142–3155 (2017).
34. Lehtinen, J. et al. Noise2noise: learning image restoration without clean data. In *Proc. 35th International Conference on Machine Learning*, PMLR (eds. Dy, J. & Krause, A.) 2965–2974 (PMLR, 2018).
35. Buchholz, T.-O., Jordan, M., Pigno, G. & Jug, F. Cryo-care: content-aware image restoration for cryo-transmission electron microscopy data. Preprint at <https://arxiv.org/abs/1810.05420> (2018).
36. Batson, J. & Royer, L. Noise2Self: blind denoising by self-supervision. Preprint at <https://arxiv.org/abs/1901.11365> (2019).
37. Krull, A., Buchholz, T.-O. & Jug, F. Noise2Void—learning denoising from single noisy images. Preprint at <https://arxiv.org/abs/1811.10980> (2018).
38. Laine, S., Lehtinen, J. & Aila, T. Self-supervised deep image denoising. Preprint at <https://arxiv.org/abs/1901.10277v1> (2019).
39. Ulyanov, D., Vedaldi, A. & Lempitsky, V. S. Deep image prior. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (eds. Brown, M. S. et al.) 9446–9454 (IEEE, 2018).
40. Zhu, J., Park, T., Isola, P. & Efros, A. A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *2017 IEEE International Conference on Computer Vision (ICCV)* (eds. Ikeuchi, K. et al.) 2242–2251 (IEEE, 2017).
41. Hom, E. F. et al. AIDA: an adaptive image deconvolution algorithm with application to multiframe and three-dimensional data. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **24**, 1580–1600 (2007).
42. Dey, N. et al. Richardson–Lucy algorithm with total variation regularization for 3D confocal microscope deconvolution. *Microsc. Res. Tech.* **69**, 260–266 (2006).
43. Preibisch, S. et al. Efficient Bayesian-based multiview deconvolution. *Nat. Methods* **11**, 645–648 (2014).
44. Sage, D. et al. DeconvolutionLab2: an open-source software for deconvolution microscopy. *Methods* **115**, 28–41 (2017).
45. Rivenson, Y. et al. Deep learning microscopy: enhancing resolution, field-of-view and depth-of-field of optical microscopy images using neural networks. In *2018 Conference on Lasers and Electro-Optics* (eds. Andersen, P. et al.) 18024028 (IEEE, 2018).
46. Henriques, R. et al. Quickpalm: 3D real-time photoactivation nanoscopy image processing in ImageJ. *Nat. Methods* **7**, 339–340 (2010).
47. Sage, D. et al. Quantitative evaluation of software packages for single-molecule localization microscopy. *Nat. Methods* **12**, 717 (2015).
48. Boyd, N., Jonas, E., Babcock, H. P. & Recht, B. DeepLoco: fast 3D localization microscopy using neural networks. Preprint at <https://www.biorxiv.org/content/10.1101/267096v1> (2018).
49. Goodfellow, I. et al. Generative adversarial nets. In *Proc. 27th International Conference on Neural Information Processing Systems* (eds. Ghahramani, Z. et al.) 2672–2680 (MIT Press, 2014).
50. 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)* (eds. Chellappa, R. et al.) 5967–5976 (IEEE, 2017).
51. Gustafsson, N. et al. Fast live-cell conventional fluorophore nanoscopy with ImageJ through super-resolution radial fluctuations. *Nat. Commun.* **7**, 12471 (2016).
52. Gustafsson, M. G. L. et al. Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination. *Biophys. J.* **94**, 4957–4970 (2008).
53. Li, D. et al. Extended-resolution structured illumination imaging of endocytic and cytoskeletal dynamics. *Science* **349**, aab3500 (2015).
54. Huang, X. et al. Fast, long-term, super-resolution imaging with Hessian structured illumination microscopy. *Nat. Biotechnol.* **36**, 451 (2018).
55. Mudry, E. et al. Structured illumination microscopy using unknown speckle patterns. *Nat. Photon.* **6**, 312 (2012).
56. Ayuk, R. et al. Structured illumination fluorescence microscopy with distorted excitations using a filtered blind-SIM algorithm. *Opt. Lett.* **38**, 4723–4726 (2013).
57. Zhu, B., Liu, J. Z., Cauley, S. F., Rosen, B. R. & Rosen, M. S. Image reconstruction by domain-transform manifold learning. *Nature* **555**, 487–492 (2018).
58. Jahr, W., Schmid, B., Schmied, C., Fahrbach, F. O. & Huisken, J. Hyperspectral light sheet microscopy. *Nat. Commun.* **6**, 7990 (2015).
59. Cutrale, F. et al. Hyperspectral phasor analysis enables multiplexed 5D in vivo imaging. *Nat. Methods* **14**, 149–152 (2017).
60. Hershko, E., Weiss, L. E., Michaeli, T. & Shechtman, Y. Multicolor localization microscopy and point-spread-function engineering by deep learning. *Opt. Express* **27**, 6158–6183 (2019).
61. Blasse, C. et al. Premosa: extracting 2D surfaces from 3D microscopy mosaics. *Bioinformatics* **33**, 2563–2569 (2017).

62. Mayer, J., Robert-Moreno, A., Sharpe, J. & Swoger, J. Attenuation artifacts in light sheet fluorescence microscopy corrected by OPTiSPIM. *Light Sci. Appl.* **7**, 70 (2018).
63. Pathak, D., Krahenbuhl, P., Donahue, J., Darrell, T. & Efros, A. A. Context encoders: feature learning by inpainting. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (eds. Tuytelaars, T. et al.) 2536–2544 (IEEE, 2016).
64. Liu, G. et al. Image inpainting for irregular holes using partial convolutions. In *Computer Vision—ECCV 2018* (eds. Ferrari, V. et al.) 89–105 (Springer, 2018).
65. Amat, F. et al. Efficient processing and analysis of large-scale light-sheet microscopy data. *Nat. Protoc.* **10**, 1679–1696 (2015).
66. Cai, B., Xu, X., Jia, K., Qing, C. & Tao, D. DehazeNet: an end-to-end system for single image haze removal. *IEEE Trans., Image Process.* **25**, 5187–5198 (2016).
67. Saalfeld, S., Fetter, R., Cardona, A. & Tomancak, P. Elastic volume reconstruction from series of ultra-thin microscopy sections. *Nat. Methods* **9**, 717–720 (2012).
68. Zbontar, J. & LeCun, Y. Computing the stereo matching cost with a convolutional neural network. In *Proc. 28th IEEE Conference on Computer Vision and Pattern Recognition* (eds. Bischof, H. et al.) 1592–1599 (2015).
69. Rohé, M.-M., Datar, M., Heimann, T., Sermesant, M. & Pennec, X. SVF-Net: learning deformable image registration using shape matching. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2017* (eds. Descoteaux, M. et al.) 266–274 (Springer, 2017).
70. Nguyen, T., Chen, S. W., Skandan, S., Taylor, C. J. & Kumar, V. Unsupervised deep homography: a fast and robust homography estimation model. *IEEE Robot. Autom. Lett.* **3**, 2346–2353 (2018).
71. Christensen, R. P. et al. Untwisting the *Caenorhabditis elegans* embryo. *eLife* **4**, e10070 (2015).
72. Prevedel, R. et al. Simultaneous whole-animal 3D imaging of neuronal activity using light-field microscopy. *Nat. Methods* **11**, 727–730 (2014).
73. Fei, P. et al. Deep learning light field microscopy for rapid four-dimensional imaging of behaving animals. Preprint at <https://www.biorxiv.org/content/10.1101/432807v1> (2018).
74. Antipa, N. et al. Diffusercam: lensless single-exposure 3D imaging. *Optica* **5**, 1–9 (2018).
75. Vinegoni, C., Pitsouli, C., Razansky, D., Perrimon, N. & Ntziachristos, V. In vivo imaging of *Drosophila melanogaster* pupae with mesoscopic fluorescence tomography. *Nat. Methods* **5**, 45–47 (2008).
76. Xingjian, S. et al. Convolutional LSTM network: a machine learning approach for precipitation nowcasting. In *Advances in Neural Information Processing Systems 28* (eds. Cortes, C. et al.) 802–810 (NIPS, 2015).
77. Naganathan, S. R., Frithauer, S., Nishikawa, M., Jlicher, F. & Grill, S. W. Active torque generation by the actomyosin cell cortex drives left-right symmetry breaking. *eLife* **3**, e04165 (2014).
78. Meister, S., Hur, J. & Roth, S. Unflow: unsupervised learning of optical flow with a bidirectional census loss. In *Thirty-second AAAI Conference on Artificial Intelligence* (eds. McIlraith, S. & Weinberger, K.) 7251–7259 (AAAI Press, 2018).
79. Haring, M. T. et al. Automated sub-5nm image registration in integrated correlative fluorescence and electron microscopy using cathodoluminescence pointers. *Sci. Rep.* **7**, 43621 (2017).
80. Royer, L. A. et al. Adaptive light-sheet microscopy for long-term, high-resolution imaging in living organisms. *Nat. Biotechnol.* **34**, 1267–1278 (2016).
81. Liu, T.-L. et al. Observing the cell in its native state: imaging subcellular dynamics in multicellular organisms. *Science* **360**, eaq1392 (2018).
82. Turpin, A., Vishniakou, I. & Seelig, J. D. Light scattering control with neural networks in transmission and reflection. Preprint at <https://arxiv.org/abs/1805.05602> (2018).
83. Horstmeyer, R., Chen, R. Y., Kappes, B. & Judkewitz, B. Convolutional neural networks that teach microscopes how to image. Preprint at <https://arxiv.org/abs/1709.07223> (2017).
84. Silver, D. et al. Mastering the game of go with deep neural networks and tree search. *Nature* **529**, 484–489 (2016).
85. Moosavi-Dezfooli, S.-M., Fawzi, A. & Frossard, P. DeepFool: a simple and accurate method to fool deep neural networks. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (eds. Tuytelaars, T. et al.) 2574–2582 (IEEE, 2016).
86. Sabour, S., Cao, Y., Faghri, F. & Fleet, D. J. Adversarial manipulation of deep representations. Preprint at <https://arxiv.org/abs/1511.05122> (2015).
87. Su, J., Vargas, D. V. & Sakurai, K. One pixel attack for fooling deep neural networks. *IEEE Trans. Evol. Comput.* <https://doi.org/10.1109/TEVC.2019.2890858> (2019).
88. Madry, A., Makelov, A., Schmidt, L., Tsipras, D. & Vladu, A. Towards deep learning models resistant to adversarial attacks. Preprint at <https://arxiv.org/abs/1706.06083v1> (2017).
89. Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems* (eds. Pereira, F. et al.) 1097–1105 (Curran Associates, 2012).
90. Johnson, G. R., Donovan-Maiye, R. M. & Maleckar, M. M. Generative modeling with conditional autoencoders: building an integrated cell. Preprint at <https://arxiv.org/abs/1705.00092> (2017).
91. Osokin, A., Chessel, A., Salas, R. E. C. & Vaggi, F. Gans for biological image synthesis. In *2017 IEEE International Conference on Computer Vision (ICCV)* (eds. Ikeuchi, K. et al.) 2252–2261 (IEEE, 2017).
92. Goldsborough, P., Pawlowski, N., Caicedo, J. C., Singh, S. & Carpenter, A. CytoGAN: generative modeling of cell images. Preprint at <https://www.biorxiv.org/content/10.1101/227645v1> (2017).
93. Yuan, H. et al. Computational modeling of cellular structures using conditional deep generative networks. *Bioinformatics* <https://doi.org/10.1093/bioinformatics/bty923> (2018).
94. Zeiler, M. D. & Fergus, R. Visualizing and understanding convolutional networks. In *Computer Vision—ECCV 2014* (eds. Fleet, D. et al.) 818–833 (Springer, 2014).
95. Yosinski, J., Clune, J., Bengio, Y. & Lipson, H. How transferable are features in deep neural networks? In *Advances in Neural Information Processing Systems 27* (eds. Ghahramani, Z. et al.) 3320–3328 (Curran Associates, 2014).
96. Esteva, A. et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **542**, 115 (2017).
97. Thul, P. J. et al. A subcellular map of the human proteome. *Science* **356**, eaal3321 (2017).
98. Ljosa, V., Sokolnicki, K. L. & Carpenter, A. E. Annotated high-throughput microscopy image sets for validation. *Nat. Methods* **9**, 637 (2012).
99. Sullivan, D. P. et al. Deep learning is combined with massive-scale citizen science to improve large-scale image classification. *Nat. Biotechnol.* **36**, 820 (2018).
100. Simonyan, K., Vedaldi, A. & Zisserman, A. Deep inside convolutional networks: visualising image classification models and saliency maps. Preprint at <https://arxiv.org/abs/1312.6034> (2013).
101. Zhang, Q., Cao, R., Shi, F., Wu, Y. N. & Zhu, S.-C. Interpreting CNN knowledge via an explanatory graph. In *Thirty-second AAAI Conference on Artificial Intelligence* (eds. McIlraith, S. & Weinberger, K.) 4454–4463 (AAAI Press, 2018).
102. Zhang, Q., Wu, Y. N. & Zhu, S.-C. Interpretable convolutional neural networks. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (eds. Brown, M. S. et al.) 8827–8836 (IEEE, 2018).
103. Lakshminarayanan, B., Pritzel, A. & Blundell, C. Simple and scalable predictive uncertainty estimation using deep ensembles. In *Advances in Neural Information Processing Systems 30* (eds. Guyon, I. et al.) 6402–6413 (NIPS, 2017).
104. Kendall, A. & Gal, Y. What uncertainties do we need in Bayesian deep learning for computer vision? In *Advances in Neural Information Processing Systems 30* (eds. Guyon, I. et al.) 5580–5590 (NIPS, 2017).
105. Hutson, M. Artificial intelligence faces reproducibility crisis. *Science* **359**, 725–726 (2018).
106. Henderson, P. et al. Deep reinforcement learning that matters. In *Thirty-Second AAAI Conference on Artificial Intelligence* (eds. McIlraith, S. & Weinberger, K.) 3207–3214 (AAAI Press, 2018).
107. Gazagnes, S., Soubies, E. & Blanc-Féraud, L. High density molecule localization for super-resolution microscopy using CEL0 based sparse approximation. In *IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)* (eds. Egan, G. et al.) 28–31 (IEEE, 2017).
108. McCann, M. T., Jin, K. H. & Unser, M. Convolutional neural networks for inverse problems in imaging: a review. *IEEE Signal Process. Mag.* **34**, 85–95 (2017).
109. Lucas, A., Iliadis, M., Molina, R. & Katsaggelos, A. K. Using deep neural networks for inverse problems in imaging: beyond analytical methods. *IEEE Signal Process. Mag.* **35**, 20–36 (2018).
110. Hornik, K., Stinchcombe, M. & White, H. Multilayer feedforward networks are universal approximators. *Neural Netw.* **2**, 359–366 (1989).
111. Rumelhart, D. E., Hinton, G. E. & Williams, R. J. Learning representations by backpropagating errors. *Nature* **323**, 533–536 (1986).
112. Masci, J., Meier, U., Cireşan, D. & Schmidhuber, J. Stacked convolutional auto-encoders for hierarchical feature extraction. In *Artificial Neural Networks and Machine Learning—ICANN 2011* (eds. Honkela, T. et al.) 52–59 (Springer, 2011).
113. Ronneberger, O., Fischer, P. & Brox, T. U-Net: convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-assisted Intervention—MICCAI 2015* (eds. Navab, N. et al.) 234–241 (Springer, 2015).
114. Falk, T. et al. U-Net: deep learning for cell counting, detection, and morphometry. *Nat. Methods* **16**, 67 (2019).

115. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (eds. Tuytelaars, T. et al.) 770–778 (IEEE, 2016).

### Acknowledgements

We thank our colleagues at the Chan Zuckerberg Biohub: A. Krishnan, B. Chhun, M. Leonetti, S. Mehta, J. Batson, and R. Gomez Sjoberg for insightful discussions, feedback, and review of the manuscript. We thank J. Zou for advice; M. Weigert for reviewing the manuscript and for innumerable discussions on the topics of deep learning and microscopy; and W. Ouyang, R. Prevedel, F. Jug, and others for feedback on the first version of the preprint. We thank the Chan Zuckerberg Biohub and its donors for funding this work.

### Author contributions

L.A.R. conceived the piece. C.B. and L.A.R. wrote the manuscript and designed the figures.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41592-019-0458-z>.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Correspondence** should be addressed to L.A.R.

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

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© Springer Nature America, Inc. 2019