

A neural lens for super-resolution biological imaging

James A. Grant-Jacob,^{1,*} Benita S. Mackay,¹ James A. G. Baker,² Yunhui Xie,¹ Daniel J. Heath,¹ Matthew Loxham,³ Robert W. Eason,¹ and Ben Mills¹

¹Optoelectronics Research Centre, University of Southampton, Southampton, SO17 1BJ, UK

²Department of Physics and Astronomy, University of Southampton, Southampton, SO17 1BJ, UK

³Faculty of Medicine, University of Southampton, Southampton, SO17 1BJ, UK

*J.A.Grant-Jacob@soton.ac.uk

Abstract: Visualizing structures smaller than the eye can see has been a driving force in scientific research since the invention of the optical microscope. Here, we use a network of neural networks to create a neural lens that has the ability to transform 20x optical microscope images into a resolution comparable to a 1500x scanning electron microscope image. In addition to magnification, the neural lens simultaneously identifies the types of objects present, and hence can label, colour-enhance and remove specific types of objects in the magnified image. The neural lens was used for the imaging of *Iva xanthiifolia* and *Galanthus* pollen grains, showing the potential for low cost, non-destructive, high-resolution microscopy with automatic image processing.

1. Introduction

There are many approaches to microscopy, each with distinct advantages and disadvantages. The optical microscope is convenient, and relatively low-cost, but has a resolution limited by the diffractive properties of light. Scanning near-field optical microscopy [1] can achieve tens of nanometre resolution, but requires long scan times and has the physical restriction of requiring the placing of probes within the evanescent field of the sample. Photo-activated localization microscopy [2] and stimulated emission depletion [3] can achieve resolutions ten times that of optical microscopes, but require samples to be labelled with fluorescent molecules. Super-oscillatory lenses [4] have been shown to achieve a resolution better than $\lambda/6$ but require working distances on the order of microns. A scanning electron microscope (SEM) can achieve resolution well below the diffraction limit of visible light [5], but this approach is expensive and often requires samples to be placed under vacuum conditions. Atomic force microscopy [6] and scanning tunnelling microscopy [7] can produce sub-nanometre resolution, but only over small areas in a workable time period.

While parametric modelling approaches have been applied in order to enhance medical imaging [8], recent work using deep learning approaches has also shown the potential for the application of neural networks (NNs) for enhancing image resolution. Examples for such an application include, a 4x upscaling on photographic images [9], optical microscopy (improving the resolution from 40x to 100x) [10], dental imaging [11], phase imaging [12], fluorescence microscopy [13], magnetic resonance imaging [14], SEM imaging [15,16], positron-emission tomography [17], stochastic optical reconstruction microscopy [18], and ultrasound imaging [19]. NNs are also well-suited to the classification of objects in images, and accordingly the classification of biological, pollution and colloidal particles from images and scattering patterns has also been demonstrated [20–24].

Such approaches generally use convolutional neural networks (CNNs) [25–27] or conditional generative adversarial networks (cGANs) [28–30], where CNNs are based on the application of a single NN, and cGANs are based on two NNs that are trained via an adversarial methodology [31–35].

In this work, we demonstrate a novel microscopy technique via deep learning, referred to here as a neural lens, which offers the low-cost and convenience of an optical microscope, but with a resolution not limited by the optical diffraction limit, whilst also providing the capability for automatic labelling and enhancement of desired features. Specifically, as well identifying pollen grains from 20x optical microscope images, we demonstrate the magnification of 20x optical microscope images of biological particles, namely, *Iva xanthiifolia* and *Galanthus* pollen grains, into 1500x resolution images, where the features in the generated 1500x images were the statistically most likely to occur. The neural lens was formed as a network of NNs in order to provide the flexibility for multiple outputs.

2. Concept

The concept of the neural lens is shown in Fig. 1, which shows the processing of an experimental 20x microscope image of two pollen grains of different types. For this input image, the neural lens produced a range of different outputs, namely a 1500x image, a sample map showing the location of the different pollen types, a colour-labelled 1500x image, and a coloured 1500x image filtered for only one specific type. The neural lens is fully deterministic, as the randomness from the latent variable was removed after training. The colour labelling for the *Iva xanthiifolia* and the *Galanthus* pollen grains were red and blue, respectively. No post-processing was applied to any of the input or output images in this manuscript. The 20x microscope image used in Fig. 1 was not used during the training of the neural lens and hence this result demonstrates the application of the neural lens to a previously unseen microscope image.

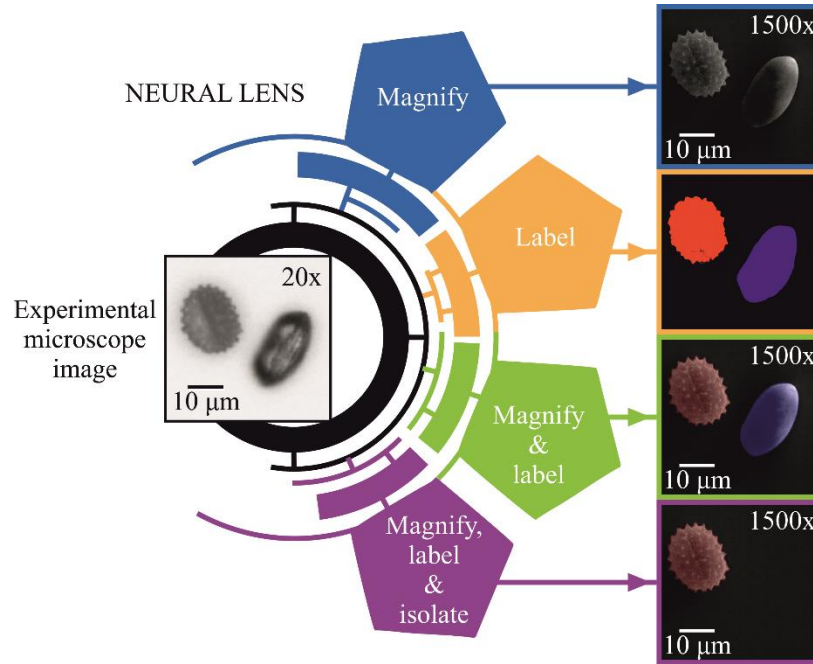


Figure 1. Concept of the neural lens. An experimental 20x microscope image was processed by the neural lens, which automatically produced an array of different output images. Some of the capabilities of the neural lens are shown, including the ability to magnify, label and isolate images of pollen grains. The two types of pollen, *Iva xanthiifolia*, and *Galanthus*, are identified and labelled by the neural lens as red and blue, respectively.

The training data in this case was a dataset containing pairs of low resolution microscope images and associated SEM images. During training, each input image was sequentially processed by the NN, which then generated a higher-magnification image. This generated

output image was compared with the SEM image from the training dataset, and the error between the images was determined. An automatic optimisation process, known as backpropagation [36], was then used in order to reduce the error for subsequent iterations.

3. Experimental methods

3.1 Sample fabrication and imaging

Pollen grains from *Iva xanthiifolia* (supplied by Sigma Aldrich) and *Galanthus* (collected by the authors) were deposited on a silica slide via a pipette, and were therefore randomly orientated and positioned on the silica slide. The particles were imaged using a Nikon microscope with an image magnification of 20x (Nikon, LE Plan, NA=0.4, WD=3.5mm) and CMOS camera (Thorlabs DCC1645C). The SEM images of the pollen were recorded using a Zeiss Evo SEM to produce 2048 pixels by 1536 pixels sized images, which were subsequently cropped to 1024 by 1024 for ease of computing. Microscope images and associated SEM images were paired via scaling and cropping of the microscope images, where the offset and scaling factor was calculated via phase correlation.

3.2 Conditional generative adversarial network

The neural lens consisted of four interconnected NNs that were each trained using a cGAN architecture. The cGANs each consisted of two NNs, referred to as a *generator* and a *discriminator*, which operated in tandem, where the generator transformed input images into *generated* output images and the discriminator would be fed either the experimental SEM image *or* the generated image and learn to discriminate between the two. During training, both the generator and discriminator improved in effectiveness, and hence each cGAN became more effective at encoding the required image transformation. At the start of training, the neuron weightings for the generator and discriminator were randomly initialised, and hence they encoded zero information about the training data. After each of the cGANs were trained, the associated generators were extracted and combined to form the neural lens architecture. The cGAN framework, known as pix2pix [37], was used as the base implementation for the paper, with an extension to a 10-layer generator architecture in order to enable an image resolution of 1024 by 1024 pixels. The network was trained using a learning rate of 0.0002 and drop-out of 0.5.

4. Neural lens architecture and training

Within the neural lens, there were three *types* of NN. Firstly, a *labelling* NN was trained in order to identify the specific type of pollen, hence transforming a 20x image into a coloured sample map. Secondly, a *magnifying* NN was trained to transform a 20x image into a 1500x image with a black background. In this case, there was a *magnifying* NN for each of the pollen types. Thirdly, a *shadowing* NN was trained to generate the shadowing effect on any 1500x image that had a black background. Although the labelling processes demonstrated here could have been achieved via a neural network designed specifically for this purpose, for example SegNet [38], here we chose to use the same architecture at every step of the pipeline for elegance. Fig. 2 shows a schematic of the internal network of the neural lens, illustrating how each type of NN interacted and combined with the others. To connect the NNs inside the neural lens, some basic mathematical functionality, such as the addition or pixel-wise multiplication of images, was included. However, in practice, this basic functionality could also have been implemented as a NN. In the neural lens, the only input data was a single 20x microscope image, and hence all subsequent images were generated automatically.

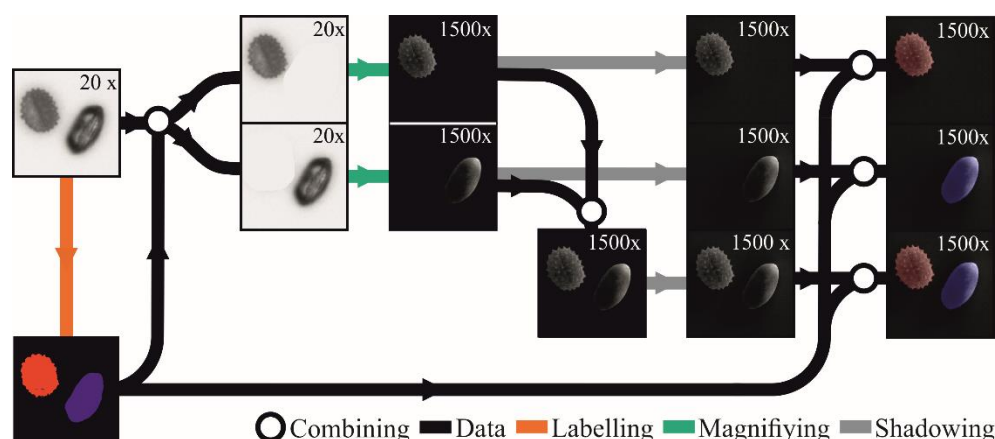


Figure 2. Schematic of the internal network of the neural lens. Through a series of interconnected NNs, the input 20x microscope image was magnified to a 1500x image, whilst the imaged pollen grains were simultaneously identified, labelled, colour-enhanced and isolated. A labelling NN transformed a 20x microscope image into a colour-labelled sample map. A magnifying NN transformed a 20x microscope of a single type into the equivalent 1500x image on a black background. A shadowing NN added shadows to an image of objects on a black background. The 20x microscope image was the only required input, and all other images were generated automatically.

The labelling NN was trained on 20x microscope images (input) and red/blue sample maps (output), where the sample maps were created using Canny-edge detection [39] and human-identification on the associated SEM images. A separate magnification NN was used for each of the pollen types, where each was trained on microscope images of a single type (input) and the associated SEM images on a black background (output). The SEM images with a black background were created by combining the SEM images with the associated red/blue sample map. The shadowing NN was trained on SEM images with a black background (input) and SEM images with shadows (output).

The training dataset for the neural lens consisted of fifty 20x microscope images and associated 1500x SEM images, red/blue colour maps and black background SEM images, where the images covered both single pollen grains of *Iva xanthiifolia* and *Galanthus* as well as mixtures of *Iva xanthiifolia* and *Galanthus* pollen grains. The neuron weightings in each of the NNs were randomly initialised, and hence before training the neural lens had zero encoding of the training data. Each of the NNs inside the neural lens was trained until the training errors reached a minimum, after approximately 200 epochs (where a single epoch is defined as the processing of all training data exactly once), taking approximately 3 days each. The training data was augmented from 50 images to 2500 images via random translation of the training image pairs across the 1024 by 1024 input size. The processing time for the neural lens for a single input microscope image was approximately 10 seconds. Details of the code of the neural network are available in the DOI link at the end of this manuscript.

5. Experimental results and discussion

Demonstration of using the neural lens on data that was not present in the training data set (unseen data) is shown in Fig. 3, which displays a comparison between an experimental 20x image and associated experimental 1500x, along with the generated 1500x image produced using the neural lens. The neural lens correctly determined the positions and orientations for the two pollen grains, and also generated an appropriate lighting and shadowing effect. For the *Iva xanthiifolia* pollen grain, the number, distribution and shapes of the spines, along with the sulcus (groove), was accurately created. Whilst the shape of the *Galanthus* pollen grain was correct and a nano-scale surface structure was generated, the sulcus was missing, which was attributed to the complexity of quantifying the external features from an image of a translucent

object. The structures observable in the 20x image of the *Galanthus* pollen grain corresponded predominantly to internal structures, and hence these features were not present in the SEM image. Therefore, the transformation of the *Galanthus* pollen from 20x to 1500x was particularly challenging. With additional training data, which would provide additional information regarding the appearance of internal and external features, it is anticipated that the sulcus would have been generated.

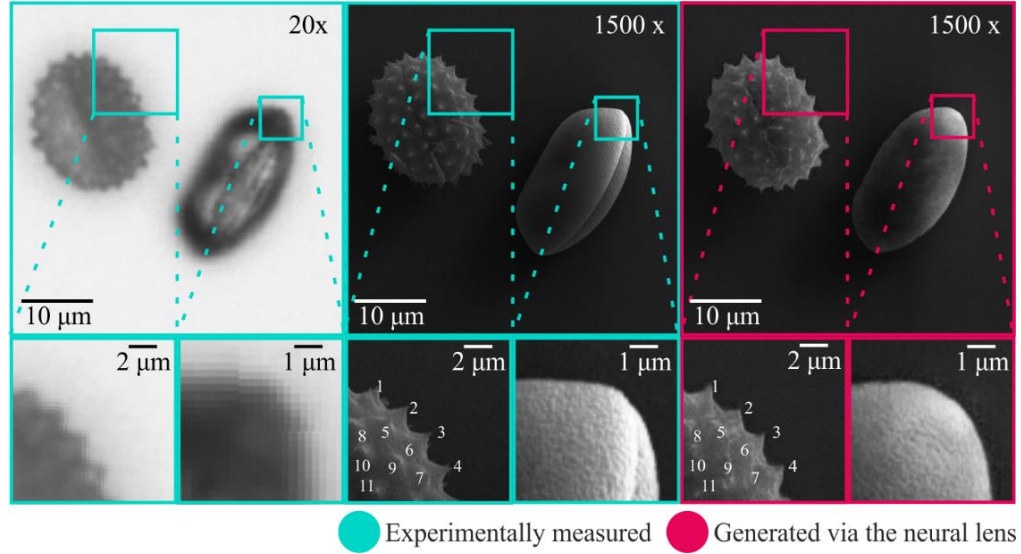


Figure 3. Evaluation of the neural lens, showing a 20x microscope image along with the associated experimental and generated 1500x images. The neural lens correctly determined the position and orientation of the two pollen grains, and produced a realistic lighting and shadowing effect. The generated surface texture on the *Galanthus* closely matched the experimental image. The spikes are labelled in the experimental and generated 1500x images of *Iva xanthiifolia*.

The neural lens enabled a resolution that was considerably superior to the resolution of the original 20x microscope images. It is important to realise that the features in the generated 1500x images were the statistically most likely feature, given the information contained in the 20x image. In other words, the effectiveness of the neural lens resided in the encoding of the statistical correlation between the occurrence of features in the 20x and the associated 1500x images that was learnt during the training process. Therefore, for the *Iva xanthiifolia* pollen grain, the spines that can be observed in the 20x images, for example around the perimeter of the grain, along with those that are less clear in the 20x image, such as those on the top surface of the grain, are generated realistically in the 1500x images. This effect can be similarly observed when comparing the surface texture of the generated and experimental 1500x images for the *Galanthus*. However, despite the statistical nature of the neural lens, the generated outputs are comparable. For larger training datasets, the accuracy would almost certainly be further improved.

Table 1 shows the relative position of each spike of *Iva xanthiifolia* labelled in figure 3. The standard deviation of the positional error for the 11 spikes, was 5.0 pixels (~ 415 nm) in X, and 4.3 pixels (~ 360 nm) in Y.

Table 1. Position of experimental and generated spikes labelled in figure 3, in number of pixels relative to bottom-left of the images, along with the difference between them (the error) in pixels.

Spike	Experimental position (pixels)		Generated position (pixels)		X Error (Pixels)	Y Error (Pixels)
	X	Y	X	Y		
1	33	112	29	115	4	-3
2	58	92	58	89	0	3
3	81	68	79	65	2	3
4	101	34	98	34	3	0
5	43	80	43	70	0	10
6	72	58	69	54	3	4
7	81	28	85	30	-4	-2
8	14	72	10	71	4	1
9	53	43	47	44	6	-1
10	20	50	14	56	6	-6
11	33	13	44	15	-11	-2

Figure 4 shows the neural lens magnification of a single *Iva xanthiifolia* pollen grain. In the image, it is clear that on comparing the experimental microscope image and the experimental SEM image, many of the spikes, for example those labelled “2 spikes” and “3 spikes” on the figure, are not resolvable in the microscope image. However, the neural lens has been able to generate the spikes, despite not being observable in the microscope image.

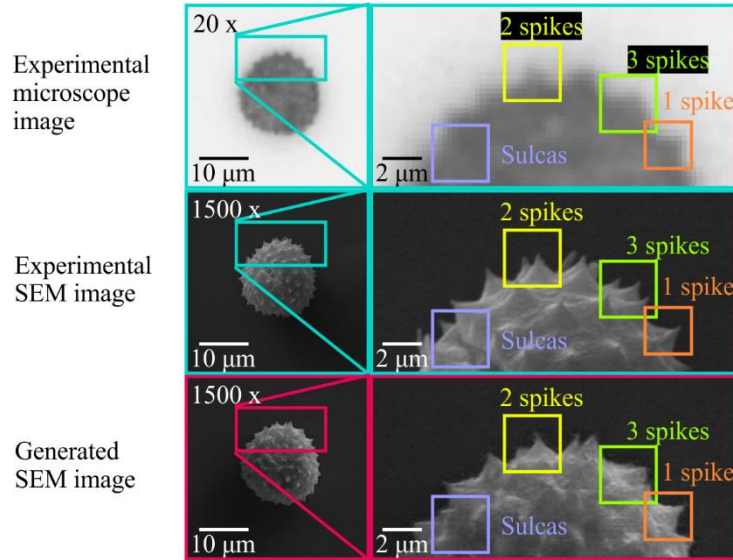


Figure 4. Evaluation of the capability of the neural lens, showing a 20x microscope image along with the associated experimental and generated 1500x images. The generated SEM image shows features that were not resolvable in the experimental microscope image.

6. Conclusion

In conclusion, we have demonstrated a new microscopy technique, formed from a network of NNs, which we have referred to here as a neural lens. The neural lens was shown to transform 20x microscope images of pollen grains into 1500x images with a resolution comparable to an

SEM. Once trained, the power of the neural lens was shown to reside in its capability to perform magnification whilst simultaneously identifying, labelling, enhancing and even removing specific features in the images. The neural lens provides the flexibility and low-cost of optical microscopy, the resolution of electron microscopy, and the functionality of automated image processing. The neural lens learns via the observation of experimental data, and hence there is no requirement for understanding the physical laws of light propagation and light-matter interaction. Whilst two types of pollen grains were shown here, the modular design of the neural lens means that this approach could easily be extended to detect additional object types and include more complex image processing features. Computationally, the modular design means the neural lens was not constrained to executing on a single graphics-processing unit, and hence offers significant potential for parallelisation. The neural lens could be applied to real-time, non-destructive imaging and characterising a variety of objects, such as living biological matter.

Funding

BM was supported by an EPSRC Early Career Fellowship (EP/N03368X/1). DH was supported by an EPSRC Doctoral Prize (EP/N509747/1). ML was supported by a BBSRC Future Leader Fellowship (BB/P011365/1). The RDM data for this paper can be found at DOI: 10.5258/SOTON/D0442

Acknowledgments

We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan X GPU used for this research, donated through NVIDIA GPU Grant Program.

References

1. E. Betzig and R. J. Chichester, "Single molecules observed by near-field scanning optical microscopy," *Science* **262**(5138), 1422–1425 (1993).
2. S.-H. Lee, J. Y. Shin, A. Lee, and C. Bustamante, "Counting single photoactivatable fluorescent molecules by photoactivated localization microscopy (PALM)," *Proc. Natl. Acad. Sci.* **109**(43), 17436–17441 (2012).
3. S. W. Hell and J. Wichmann, "Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy," *Opt. Lett.* **19**(11), 780–782 (1994).
4. E. T. F. Rogers, J. Lindberg, T. Roy, S. Savo, J. E. Chad, M. R. Dennis, and N. I. Zheludev, "A super-oscillatory lens optical microscope for subwavelength imaging," *Nat. Mater.* **11**(5), 432 (2012).
5. T. Gericke, P. Würtz, D. Reitz, T. Langen, and H. Ott, "High-resolution scanning electron microscopy of an ultracold quantum gas," *Nat. Phys.* **4**(12), 949 (2008).
6. F. J. Giessibl, "Atomic resolution of the silicon (111)-(7×7) surface by atomic force microscopy," *Science* **267**(5194), 68–71 (1995).
7. G. Binnig, H. Rohrer, C. Gerber, and E. Weibel, "Surface studies by scanning tunneling microscopy," *Phys. Rev. Lett.* **49**(1), 57 (1982).
8. D. Kouame and M. Ploquin, "Super-resolution in medical imaging: An illustrative approach through ultrasound," in *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro* (2009), pp. 249–252.
9. C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and others, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2017), pp. 4681–4690.
10. Y. Rivenson, Z. Göröcs, H. Günaydin, Y. Zhang, H. Wang, and A. Ozcan, "Deep learning microscopy," *Optica* **4**(11), 1437–1443 (2017).
11. J. Hatvani, A. Horváth, J. Michetti, A. Basarab, D. Kouamé, and M. Gyöngy, "Deep learning-based super-resolution applied to dental computed tomography," *IEEE Trans. Radiat. Plasma Med. Sci.* **3**(2), 120–128 (2019).
12. M. Kellman, E. Bostan, N. Repina, and L. Waller, "Physics-based Learned Design: Optimized Coded-Illumination for Quantitative Phase Imaging," *IEEE Trans. Comput. Imaging* (2019).
13. H. Wang, Y. Rivenson, Y. Jin, Z. Wei, R. Gao, H. Günaydin, L. A. Bentolila, C. Kural, and A. Ozcan, "Deep learning enables cross-modality super-resolution in fluorescence microscopy," *Nat. Methods* **16**103–110 (2019).

14. K. H. Kim, W.-J. Do, and S.-H. Park, "Improving resolution of MR images with an adversarial network incorporating images with different contrast," *Med. Phys.* **45**(7), 3120–3131 (2018).
15. K. Hagita, T. Higuchi, and H. Jinnai, "Super-resolution for asymmetric resolution of FIB-SEM 3D imaging using AI with deep learning," *Sci. Rep.* **8**(1), 5877 (2018).
16. L. Heinrich, J. A. Bogovic, and S. Saalfeld, "Deep learning for isotropic super-resolution from non-isotropic 3D electron microscopy," in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (2017), pp. 135–143.
17. Y. Wang, B. Yu, L. Wang, C. Zu, D. S. Lalush, W. Lin, X. Wu, J. Zhou, D. Shen, and L. Zhou, "3D conditional generative adversarial networks for high-quality PET image estimation at low dose," *Neuroimage* **174**550–562 (2018).
18. E. Nehme, L. E. Weiss, T. Michaeli, and Y. Shechtman, "Deep-STORM: super-resolution single-molecule microscopy by deep learning," *Optica* **5**(4), 458–464 (2018).
19. F. Tom and D. Sheet, "Simulating patho-realistic ultrasound images using deep generative networks with adversarial learning," in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)* (2018), pp. 1174–1177.
20. J. A. Grant-Jacob, B. S. Mackay, J. A. G. Baker, D. J. Heath, Y. Xie, M. Loxham, R. W. Eason, and B. Mills, "Real-time particle pollution sensing using machine learning," *Opt. Express* **26**(21), 27237–27246 (2018).
21. Y. Wu, Y. Rivenson, Y. Zhang, Z. Wei, H. Günaydin, X. Lin, and A. Ozcan, "Extended depth-of-field in holographic imaging using deep-learning-based autofocusing and phase recovery," *Optica* **5**(6), 704–710 (2018).
22. A. Yevick, M. Hannel, and D. G. Grier, "Machine-learning approach to holographic particle characterization," *Opt. Express* **22**(22), 26884–26890 (2014).
23. K. Dong, Y. Feng, K. M. Jacobs, J. Q. Lu, R. S. Brock, L. V Yang, F. E. Bertrand, M. A. Farwell, and X.-H. Hu, "Label-free classification of cultured cells through diffraction imaging," *Biomed. Opt. Express* **2**(6), 1717–1726 (2011).
24. C. L. Chen, A. Mahjoubfar, L.-C. Tai, I. K. Blaby, A. Huang, K. R. Niazi, and B. Jalali, "Deep Learning in Label-free Cell Classification," *Sci. Rep.* **6**21471 (2016).
25. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems* (2012), pp. 1097–1105.
26. A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and L. Fei-Fei, "Large-scale video classification with convolutional neural networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2014), pp. 1725–1732.
27. H. Li, Z. Lin, X. Shen, J. Brandt, and G. Hua, "A convolutional neural network cascade for face detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2015), pp. 5325–5334.
28. M. Mirza and S. Osindero, "Conditional generative adversarial nets," *arXiv Prepr. arXiv1411.1784* (2014).
29. G. Antipov, M. Baccouche, and J.-L. Dugelay, "Face aging with conditional generative adversarial networks," in *2017 IEEE International Conference on Image Processing (ICIP)* (2017), pp. 2089–2093.
30. G. Douzas and F. Bacao, "Effective data generation for imbalanced learning using conditional generative adversarial networks," *Expert Syst. Appl.* **91**464–471 (2018).
31. A. Odena, C. Olah, and J. Shlens, "Conditional image synthesis with auxiliary classifier gans," in *Proceedings of the 34th International Conference on Machine Learning-Volume 70* (2017), pp. 2642–2651.
32. T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of gans for improved quality, stability, and variation," *arXiv Prepr. arXiv1710.10196* (2017).
33. T. Miyato, T. Kataoka, M. Koyama, and Y. Yoshida, "Spectral normalization for generative adversarial networks," *arXiv Prepr. arXiv1802.05957* (2018).
34. H. Zhang, T. Xu, H. Li, S. Zhang, X. Wang, X. Huang, and D. N. Metaxas, "Stackgan: Text to photo-realistic image synthesis with stacked generative adversarial networks," in *Proceedings of the IEEE International Conference on Computer Vision* (2017), pp. 5907–5915.
35. T.-C. Wang, M.-Y. Liu, J.-Y. Zhu, A. Tao, J. Kautz, and B. Catanzaro, "High-resolution image synthesis and semantic manipulation with conditional gans," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2018), pp. 8798–8807.
36. M. Riedmiller and H. Braun, "A direct adaptive method for faster backpropagation learning: The RPROP algorithm," in *Neural Networks, 1993., IEEE International Conference On* (1993), pp. 586–591.
37. P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," *arXiv Prepr.* (2017).
38. V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.* **39**(12), 2481–2495 (2017).
39. J. Canny, "A computational approach to edge detection," in *Readings in Computer Vision* (Elsevier, 1987), pp. 184–203.