applied optics

Focus prediction in digital holographic microscopy using deep convolutional neural networks

Tomi Pitkäaho,[1] Aki Manninen,[2] and Thomas J. Naughton[1,*]

[1]Department of Computer Science, Maynooth University–National University of Ireland Maynooth, Maynooth, County Kildare, Ireland
[2]Faculty of Biochemistry and Molecular Medicine, Biocenter Oulu, Oulu Center for Cell-Matrix Research, University of Oulu, Oulu, Finland
*Corresponding author: tomn@cs.nuim.ie

Received 18 September 2018; revised 7 December 2018; accepted 8 December 2018; posted 10 December 2018 (Doc. ID 346136); published 16 January 2019

Deep artificial neural network learning is an emerging tool in image analysis. We demonstrate its potential in the field of digital holographic microscopy by addressing the challenging problem of determining the in-focus reconstruction depth of Madin–Darby canine kidney cell clusters encoded in digital holograms. A deep convolutional neural network learns the in-focus depths from half a million hologram amplitude images. The trained network correctly determines the in-focus depth of new holograms with high probability, without performing numerical propagation. This paper reports on extensions to preliminary work published earlier as one of the first applications of deep learning in the field of digital holographic microscopy.

https://doi.org/10.1364/AO.58.00A202

1. INTRODUCTION

Deep learning [1] is a technique used to solve hitherto open problems in image analysis and other fields that is starting to have an impact in the field of biomedical optics; for example, OCT [2–4] and other forms of microscopy [5–9]. Image-based applications of deep learning [10] are characterized by neural networks with at least eight hidden layers, at least tens of thousands of images, at least hundreds of images per class, at least millions of learned parameters, and training times of at least weeks if run on a single-processor personal computer. This type of network has been used successfully in various different visual object recognition and object detection applications [11,12].

A digital hologram is an efficient encoding of a diffraction volume. It would be desirable for digital holography researchers to be able to edit the hologram directly, to effect some semantic change in the diffraction volume, such as three-dimensional (3D) segmentation, or even more simply, to analyze the hologram directly to understand the 3D scene. Unfortunately, in general, this ability has eluded researchers in digital holography. Researchers are limited to sampling the reconstruction volume (i.e., using numerical propagation to reconstruct from the digital hologram at a plurality of depths) before they can understand the encoded 3D scene. A handful of notable exceptions exist, such as landmark papers by Vikram and Billet [13], and Onural and Özgen [14], and subsequently others over the past decade [15–21], whose work allows one to determine the size and position of individual particles based on an analysis of the hologram directly. However, these approaches are limited to the special case of idealized spherical particles. Here, we consider significantly more complicated multicellular partially transparent objects.

Holography has a history as an enabling technology for artificial neural networks [22]. Conventional artificial neural networks have been applied before in the fields of digital holographic microscopy [21,23] and, more generally, digital holography [24,25]. Recently, deep learning has been applied in different holographic applications [26–38].

Deep convolutional neural networks (CNNs), which are one form of deep learning, were first defined by LeCun *et al.* [39]. A CNN is a feed-forward artificial neural network that typically comprises separate layers for feature extraction and classification. Typically, each feature extractor layer performs separate convolution operations on its input with a bank of convolution kernels.

The values in the convolution kernels are the learned weights in the layer. The number of learned kernels is determined as part of the network architecture design. The classification layers are typically fully connected layers well-known from conventional multilayer perceptron neural networks.

This paper is organized in four sections. Section 2 contains a summary of digital holographic microscopy and existing autofocusing methods. Section 3 contains the proposed solution with descriptions of the convolutional neural network architectures used and the experimental results, and Section 4 concludes the paper.

2. DIGITAL HOLOGRAPHIC MICROSCOPY

Digital holographic microscopy (DHM) is a label-free, single-shot technique well suited for imaging living 3D microbiological samples [40]. One open problem in DHM is how to efficiently determine the appropriate in-focus plane for an arbitrary semitransparent microscopic object encoded in a digital hologram. If the sample contains more complicated objects than idealized spherical particles [14], researchers must perform a plurality of numerical propagation steps, which is inefficient. In this paper, we demonstrate that it is possible to design a deep convolutional neural network to predict the in-focus distance of a living cell cluster from only the digital hologram plane amplitude. With deep learning, we propose that DHM researchers now have a tool at their disposal that is a major step toward removing the need to perform any propagation steps to determine the in-focus distance.

DHM overcomes the shallow depth-of-field problem in optical microscopes, by allowing one to change the in-focus plane after hologram capture. A digital hologram $H(x,y)$ can be propagated to any depth z using the Fresnel approximation [41],

$$U(x,y;z) = \frac{-i}{\lambda z}\exp(\mathrm{i}kz)H(x,y) \otimes \exp\left(\mathrm{i}\pi\frac{x^2+y^2}{\lambda z}\right), \quad (1)$$

where λ is the wavelength of the light, $\otimes$ denotes a convolution operation, and $k = 2\pi/\lambda$. The amplitude and phase components of the complex-valued reconstruction are defined as

$$A(x,y;z) = \{\mathrm{Re}[U(x,y;z)]^2 + \mathrm{Im}[U(x,y;z)]^2\}^{0.5}, \quad (2)$$

$$\phi(x,y;z) = \arctan\{\mathrm{Im}[U(x,y;z)]/\mathrm{Re}[U(x,y;z)]\}, \quad (3)$$

respectively. However, since each in-focus plane has a narrow depth of field, the object of interest is in focus only at a small range of reconstruction depths. Determining the most appropriate in-focus depth is essential for applications such as autofocusing, extended focus imaging, and 3D object segmentation/recognition.

The critical importance of this problem to digital holography researchers is evidenced by the regularity of newly proposed focus metrics to apply to Eqs. (2) and (3) [42–52], among others. We highlight some of the proposals below.

Gillespie and King were the first to automate in-focus plane detection of digital holograms [42]. They applied entropy as a focus metric. Ferraro *et al.* detected phase changes and used this information as a focus metric [43]. Liebling and Unser used a wavelet approach to measure image energy as a focus metric [44]. A reconstruction in a stack that contains high levels of image energy in as few coefficients as possible is considered to be the in-focus plane. Dubois *et al.* discovered that integrated amplitude reconstruction is at a minimum for pure amplitude objects and at a maximum for pure phase objects [45]. McElhinney *et al.* used the maximum of the gray level variance of a amplitude reconstruction stack to find in-focus objects in a scene [46]. Langehanenberg *et al.* compared power spectra, gray level variance, edge detection, and Laplacian filtering of pure phase objects and concluded that power spectra and edge detection are equally well suited for pure phase objects [47]. Memmolo *et al.* showed that Tamura coefficient estimation on pure phase objects is at minimum on a reconstruction stack for the in-focus plane [48]. Dohet-Eraly *et al.* used multiwavelength DHM and developed a focus metric that works on Fourier domain phase. They showed that their metric yielded a minimum at the in-focus plane of a reconstruction stack [49]. He *et al.* calculated cosine score on an amplitude reconstruction stack. By finding the minimum value, they were able to discover the in-focus plane [50]. An approach of Ren *et al.* using structure tensors and their eigenvalues was shown to work with overlapping objects [51]. By finding maximum peak(s) in their focus metric, applied on a reconstructions stack, they discovered in-focus plane(s). Lyu *et al.* calculated the sum of subtracted neighboring reconstructions in a stack, and by finding the maximum of amplitude objects and the minimum for phase objects, were able to find the in-focus plane [52].

Despite a great number of publications tackling this challenging problem, they all have the same drawback; namely, that a stack of reconstructed images must be computed, and the focus metric must be applied to each reconstruction. This time-consuming drawback is compounded by the fact that the whole procedure must be applied to each new hologram. There is a trade-off between computational search time and accuracy. By sampling the diffraction volume at too low a resolution longitudinally, one may miss the in-focus plane. Or, sampling the volume with a fine resolution could lead to an increased computation time.

The greatest benefit of the deep learning method outlined in this paper is that after training, the in-focus depth can be obtained from the hologram plane intensity directly, in constant time, without any numerical propagation. It should be noted that this paper extends results reported in [26,27] by rigorously verifying the performance of the approach in a regression context. Ren *et al.* [30] reported a deep-learning-based approach for both amplitude and phase objects. An interesting deep learning approach where the network learns to propagate a hologram of a transparent object to an in-focus plane and perform a phase unwrapping operation was reported by Rivenson *et al.* [29]. In contrast to these papers, the phase objects reported in this paper are living biological 3D multicellular samples with all of the noise and inhomogeneties of living organisms.

3. EXPERIMENTAL RESULTS

A. Training

Two different deep convolutional neural networks were used (see Fig. 1). The first was based on AlexNet [11], which won the classification and localization tasks in the Large Scale Visual Recognition Challenge 2012. It has five convolution layers, three fully connected layers, and uses convolutional filters up to 11×11 pixels in size. The second, VGG16, was based on a variant of the VGG [12] architecture that was successful in the classification task in the same challenge in 2014. VGG16 has 13 convolution layers, three fully connected layers, and learns smaller features in the hologram plane with its 3×3 pixel convolutional filters.

Both of these network architectures are well-known and highly accepted models for a wide range of real-world, image-based applications. In this paper, we show that these standardized models can be used as-is with holographic data from

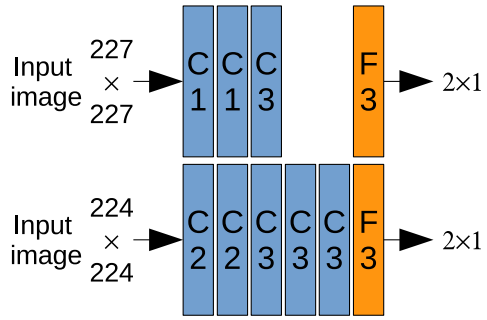

Fig. 1. Network architectures based on AlexNet (above) and VGG16 (below): C, convolution block; F, fully connected block; input size, 227 × 227 pixels. Numbers show amount of layers in each block. Each convolution block is followed by a max-pooling layer (with kernel size of 3 and stride of 2 in AlexNet, and with kernel size of 2 and stride of 2 in VGG16).

biological samples, and that it is not necessary to invent a custom architecture for each new application.

A set of 494 holograms of semitransparent Madin–Darby canine kidney (MDCK) epithelial cell clusters was captured using an off-axis Mach–Zehnder digital holographic microscope (Lyncée Tec T1000, Lyncée Tec SA, Lausanne, Switzerland). The microscope comprises a 660 nm laser source, a 1024 × 1024 pixel CCD camera with 6.45 μm square pixels, and a 40X microscope objective with 0.7 numerical aperture (Leica HCX PL Fluotar). Each hologram was preprocessed automatically by removing the zero order and twin terms [53]. To obtain the ground truth data, Tomi Pitkäaho, one of the authors of this paper, manually determined the in-focus depth z for each hologram in the training, validation, and test datasets. To do this, holograms were reconstructed [using Eq. (1)] at multiple depths, and the particular depth z that visually displayed sharpest outer border was chosen as the in-focus depth.

The holograms were used to generate a database of images. An amplitude reconstruction was obtained from each hologram at each of 21 depths distributed equally over the range ±100 mm centered on the in-focus plane (Fig. 2). (Note: We follow a convention [43,47] of ignoring the effect of the microscope objective in our reconstruction distances.) Through different combinations of rescaling and cropping, each reconstruction was used to generate six similar but distinct 227 × 227 pixel rescaled and cropped images. Each such image was augmented with 90, 180, and 270 deg rotations, and each resulting image was further augmented through horizontal mirroring.

This formed a database of 497 952 images. From this database, all augmented images from the twelve hand-picked holograms shown in Fig. 3 (comprising 12 096 images, 2.4% of the set) were set aside as test data. The remaining images were partitioned randomly into training (87.8%, 437 271) and validation (9.8%, 48 585) data. Finally, a mean image (calculated from the training data only) was subtracted from each image.

The actual training used Nvidia's Deep Learning GPU Training System (DIGITS) software with two Nvidia GTX 1080 graphics cards. Learning rates were fixed at 1×10^{-5} and 1×10^{-6} for AlexNet and VGG, respectively. AlexNet was trained for 100 epochs (16 h) and VGG16 for 80 epochs

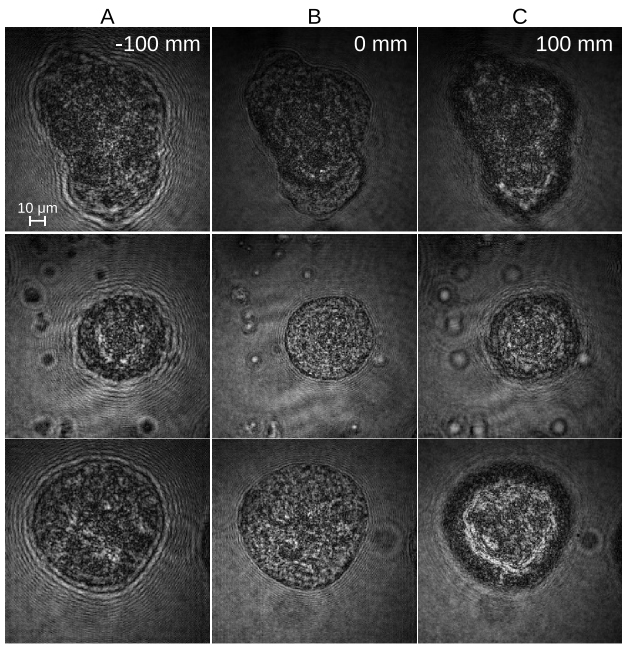

Fig. 2. Example of training images: Each row shows amplitude reconstructions from one hologram (at the in-focus plane and at distances ±100 mm from the in-focus plane).

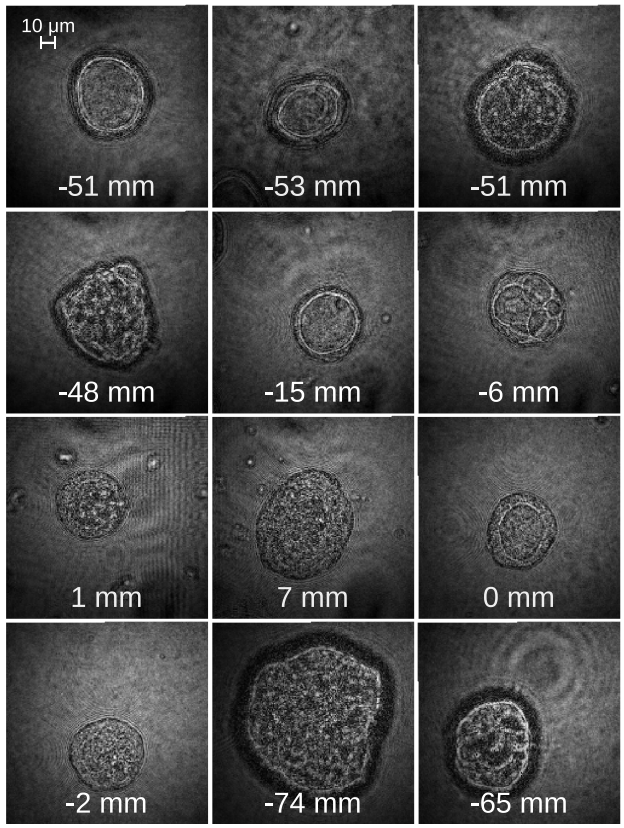

Fig. 3. Amplitudes of the twelve dc- and twin-free holograms used for testing (with ground truth in-focus distance shown).

(210 h) with a stochastic gradient descent solver. The loss function was mean square error and each network returned a single value for the predicted in-focus depth. The minibatch size was set to 150 and 25 for AlexNet and VGG16, respectively. The learned filters from each of the first convolution layers are shown in Fig. 4, allowing one to infer the basic features that each network learned to extract from an image for analysis in subsequent layers. When the training was finished, the training loss was 2.73 and 4.47 for AlexNet and VGG16, respectively (Fig. 5).

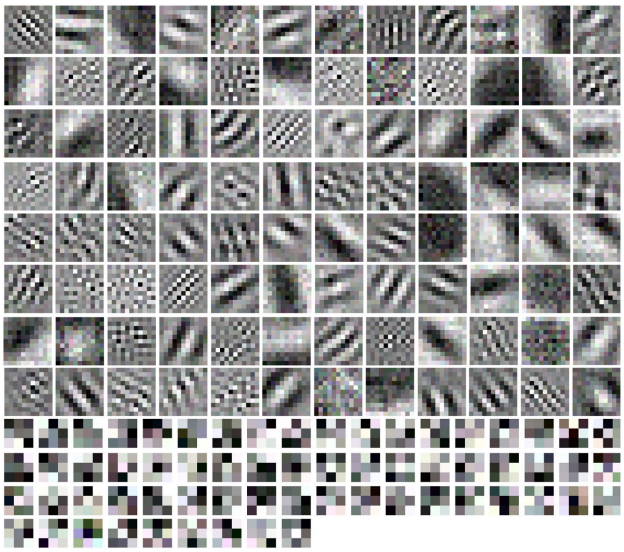

Fig. 4. Learned filters from the first convolution layer: 9611 × 11 pixel filters from AlexNet (top), 633 × 3 pixel filters from VGG16 (bottom).

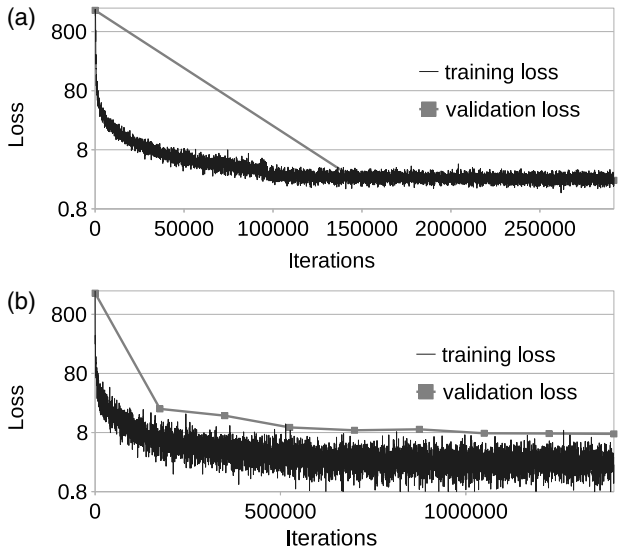

Fig. 5. Loss functions (square mean error, logarithmic scale) for (a) AlexNet, (b) VGG16. Possibly, VGG16 would have benefited from a greater minibatch size (as evidenced by the fluctuation in the loss function).

B. Testing

Testing was performed on a separate computer to demonstrate the portability of deep learning. The trained models (with sizes 227.4 MB and 662.9 MB for AlexNet and VGG16, respectively) were imported into the Caffe [54] deep learning framework using the general-purpose Python programming language. The run times (mean of 200 holograms) were 247 ms (AlexNet) and 680 ms (VGG16) on a 2014-era Dell Latitude Ultrabook (Intel Core i5 processor) without a discrete graphics card. For comparison, on the same computer, and using a hologram of the same dimensions, a single Tamura coefficient calculation (mean of 200) requires 932 ms (aberration removal 380 ms, reconstruction 318 ms, phase unwrapping 231 ms, Tamura coefficient calculation 3 ms).

Testing was performed by classifying holograms that were not used in training or validation, as explained. Of the 12 096 test images, 100% and 99.9% were within an acceptable error margin of 20 mm for AlexNet and VGG16, respectively (Figs. 6 and 7). The rms errors were 6.37 mm and 6.49 mm for AlexNet and VGG16, respectively. The error margin of 20 mm was determined experimentally, and there was no significant visual difference in amplitude reconstruction quality within this range. All required morphological information about cell clusters can be extracted within this range with a good accuracy. This is supported by extracting a single morphological feature, size, with test holograms (Fig. 8).

To examine how the networks responded to holograms that may have an in-focus distance not a multiple of the 10 mm discretization used in training, the holograms from Fig. 3 were used directly (Table 1). The mean absolute error over the 12 holograms was 3.62 mm and 4.25 mm for AlexNet and VGG16, respectively.

Systematic testing was then performed with the holograms from Fig. 3 over the range ±100 mm centered on the in-focus depth, but this time with a finer depth resolution of 1 mm. For a system to generalize well outside the discrete set of 21 in-focus depths with which it was trained, the shape of the scatter plot should be linear. Both networks generalized well with each test hologram, and typical examples are shown in Figs. 9(a) and 9(b).

To test the extrapolation capability of the models beyond depths not used in the training, the test holograms were reconstructed at distances in the range ±200 mm from the manually estimated in-focus plane, with steps of 1 mm [Fig. 9(c)]. It can be observed that beyond depths that were not used in

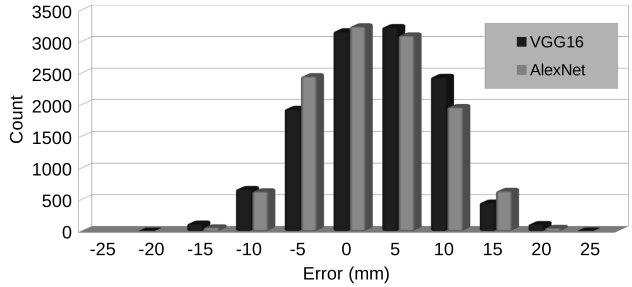

Fig. 6. Absolute estimation errors with the testing data. In 88.7% and 88.9% of test cases, for AlexNet and VGG, respectively, the absolute depth estimation error is ≤ 10 mm; in 100% and 99.9% of test cases, respectively, the absolute error is ≤ 20 mm.

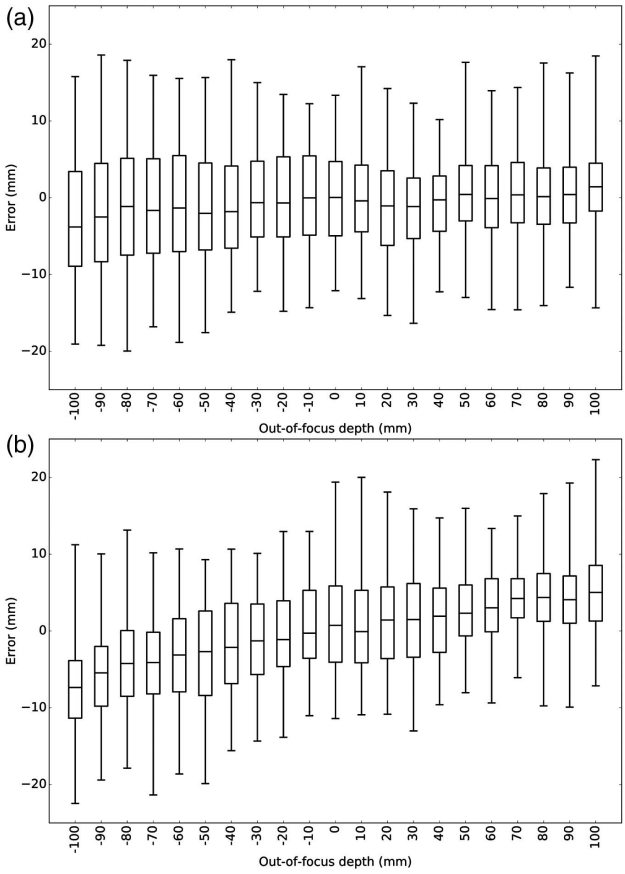

Fig. 7. Box plots (a) AlexNet and (b) VGG16 showing distributions of in-focus depth prediction errors for different out-of-focus depths.

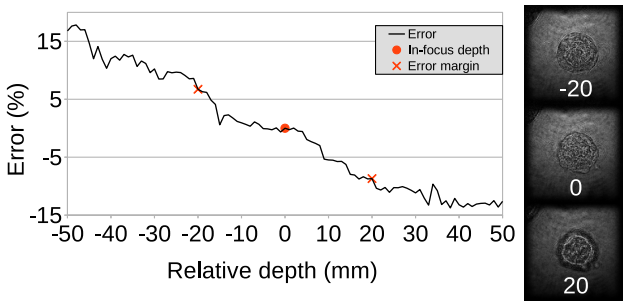

Fig. 8. Area calculated at multiple depths on one of the test holograms. In addition, showing reconstructions at −20 mm, 0 mm, and 20 mm from the in-focus plane.

the training, the error increases. This is expected because neural networks in general are not good in extrapolation. As the volume size is known before the training, the system should be trained within the full depth of the volume so that there is no need for extrapolation during prediction.

To test the generalization capability and push the networks past their designed capabilities, the networks were tested with a human cell line captured with the same DHM hardware. An example with 1 mm reconstruction steps is shown in Fig. 10. It can be observed that, in general, the CNN is able

Table 1. Test Results Using the 12 Holograms from Fig. 3, Showing Regression Result from Each Network (G, Ground Truth; A, AlexNet; AE, Absolute Error; V, VGG16)

G (mm)	A (mm)	AE (mm)	V (mm)	AE (mm)
−51	−46.0	4.9	−52.5	1.5
−53	−52.4	0.6	−55.3	2.3
−51	−49.8	1.2	−57.7	6.7
−48	−41.1	6.9	−42.9	5.1
−15	−15.2	0.2	−15.5	0.5
−6	−5.1	0.9	−13.4	7.4
1	1.4	0.4	−4.3	5.3
7	0.0	7	5.1	1.9
0	−4.4	4.4	−4.7	4.7
−2	−1.0	1.0	2.7	4.7
−74	−83.0	9.0	−82.1	8.1
−65	−60.1	4.9	−67.8	2.8

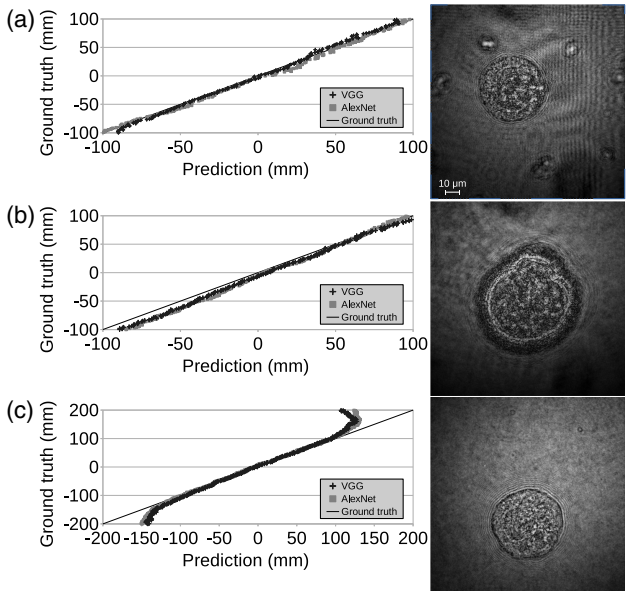

Fig. 9. Fine (1 mm) depth-resolution test results with different sized cell clusters. A linear result indicates perfect estimation. (a) and (b) Two holograms reconstructed ±100 mm around the in-focus plane and (c) hologram reconstructed ±200 mm around the in-focus plane.

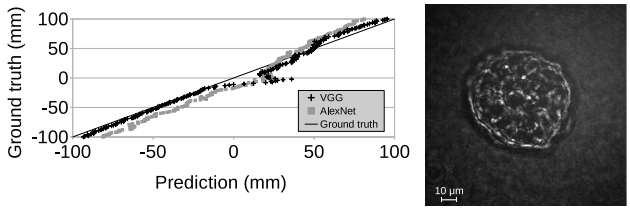

Fig. 10. Fine (1 mm) depth-resolution test results with a human cell line sample.

to determine in-focus depths correctly. However, the result displays greater error at the middle region. Typically, with new samples that differ greatly from a training set of a trained

model, one needs to retrain a model with new data. However, transfer learning [55] can be used to speed up the training process.

4. CONCLUSIONS

In a rigorous treatment of our preliminary results [26,27], one of the first applications of deep learning to digital holographic microscopy, we show that a deep artificial neural network can be designed to learn the appropriate in-focus depth of an arbitrary MDCK cell cluster encoded in a digital hologram. Its greatest benefit is that the in-focus depth can be obtained from the hologram plane intensity only, and in constant time, without any numerical propagation. It generalizes well to in-focus depths differing from those in its training set.

Funding. Science Foundation Ireland (SFI) (13/CDA/2224); Irish Research Council Postgraduate Scholarship.

REFERENCES

1. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature **521**, 436–444 (2015).
2. P. Prentašić, M. Heisler, Z. Mammo, S. Lee, A. Merkur, E. Navajas, M. F. Beg, M. Šarunić, and S. Lončarić, "Segmentation of the foveal microvasculature using deep learning networks," J. Biomed. Opt. **21**, 075008 (2016).
3. A. Abdolmanafi, L. Duong, N. Dahdah, and F. Cheriet, "Deep feature learning for automatic tissue classification of coronary artery using optical coherence tomography," Biomed. Opt. Express **8**, 1203–1220 (2017).
4. S. P. K. Karri, D. Chakraborty, and J. Chatterjee, "Transfer learning based classification of optical coherence tomography images with diabetic macular edema and dry age-related macular degeneration," Biomed. Opt. Express **8**, 579–592 (2017).
5. D. Cireşan, A. Giusti, L. M. Gambardella, and J. Schmidhuber, "Deep neural networks segment neuronal membranes in electron microscopy images," Adv. Neural Inf. Process. Syst. **25**, 2843–2851 (2012).
6. H. Wang, A. Cruz-Roa, A. Basavanhally, H. Gilmore, N. Shih, M. Feldman, J. Tomaszewski, F. Gonzalez, and A. Madabhushi, "Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features," J. Med. Imaging **1**, 034003 (2014).
7. H. Rezaeilouyeh, A. Mollahosseini, and M. H. Mahoor, "Microscopic medical image classification framework via deep learning and shearlet transform," J. Med. Imaging **3**, 044501 (2016).
8. G. Gopakumar, K. H. Babu, D. Mishra, S. S. Gorthi, and G. R. K. S. Subrahmanyam, "Cytopathological image analysis using deep-learning networks in microfluidic microscopy," J. Opt. Soc. Am. A **34**, 111–121 (2017).
9. Y. Rivenson, Z. Göröcs, H. Günaydin, Y. Zhang, H. Wang, and A. Ozcan, "Deep learning microscopy," Optica **4**, 1437–1443 (2017).
10. O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet large scale visual recognition challenge," Int. J. Comput. Vision **115**, 211–252 (2015).
11. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Adv. Neural Inf. Process. Syst. **25**, 1097–1105 (2012).
12. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv:1409.1556 (2014).
13. C. S. Vikram and M. L. Billet, "Far-field holography at non-image planes for size analysis of small particles," Appl. Phys. B **33**, 149–153 (1984).
14. L. Onural and M. T. Özgen, "Extraction of three-dimensional object-location information directly from in-line holograms using Wigner analysis," J. Opt. Soc. Am. A **9**, 252–260 (1992).
15. F. Soulez, L. Denis, C. Fournier, É. Thiébaut, and C. Goepfert, "Inverse-problem approach for particle digital holography: accurate location based on local optimization," J. Opt. Soc. Am. A **24**, 1164–1171 (2007).
16. S.-H. Lee, Y. Roichman, G.-R. Yi, S.-H. Kim, S.-M. Yang, A. van Blaaderen, P. van Oostrum, and D. G. Grier, "Characterizing and tracking single colloidal particles with video holographic microscopy," Opt. Express **15**, 18275–18282 (2007).
17. F. C. Cheong, B. J. Krishnatreya, and D. G. Grier, "Strategies for three-dimensional particle tracking with holographic video microscopy," Opt. Express **18**, 13563–13573 (2010).
18. J. Fung, K. E. Martin, R. W. Perry, D. M. Kaz, R. McGorty, and V. N. Manoharan, "Measuring translational, rotational, and vibrational dynamics in colloids with digital holographic microscopy," Opt. Express **19**, 8051–8065 (2011).
19. M. Seifi, L. Denis, and C. Fournier, "Fast and accurate 3D object recognition directly from digital holograms," J. Opt. Soc. Am. A **30**, 2216–2224 (2013).
20. A. Yevick, M. Hannel, and D. G. Grier, "Machine-learning approach to holographic particle characterization," Opt. Express **22**, 26884–26890 (2014).
21. B. Schneider, J. Dambre, and P. Bienstman, "Fast particle characterization using digital holography and neural networks," Appl. Opt. **55**, 133–139 (2016).
22. D. Psaltis and N. Farhat, "Optical information processing based on an associative-memory model of neural nets with thresholding and feedback," Opt. Lett. **10**, 98–100 (1985).
23. U. S. Kamilov, I. N. Papadopoulos, M. H. Shoreh, A. Goy, C. Vonesch, M. Unser, and D. Psaltis, "Learning approach to optical tomography," Optica **2**, 517–622 (2015).
24. Y. Frauel and B. Javidi, "Neural network for three-dimensional object recognition based on digital holography," Opt. Lett. **26**, 1478–1480 (2001).
25. A. E. Shortt, T. J. Naughton, and B. Javidi, "Compression of optically encrypted digital holograms using artificial neural networks," J. Display Technol. **2**, 401–410 (2006).
26. T. Pitkäaho, A. Manninen, and T. J. Naughton, "Focus classification in digital holographic microscopy using deep convolutional neural networks," in *European Conference on Biomedical Optics* (Optical Society of America, 2017), paper 104140K.
27. T. Pitkäaho, A. Manninen, and T. J. Naughton, "Performance of autofocus capability of deep convolutional neural networks in digital holographic microscopy," in *Digital Holography and Three-Dimensional Imaging* (Optical Society of America, 2017), paper W2A-5.
28. T. Nguyen, V. Bui, V. Lam, C. B. Raub, L.-C. Chang, and G. Nehmetallah, "Automatic phase aberration compensation for digital holographic microscopy based on deep learning background detection," Opt. Express **25**, 15043–15057 (2017).
29. Y. Rivenson, Y. Zhang, H. Günaydn, D. Teng, and A. Ozcan, "Phase recovery and holographic image reconstruction using deep learning in neural networks," Light: Sci. Appl. **7**, 17141 (2018).
30. Z. Ren, Z. Xu, and E. Y. Lam, "Learning-based nonparametric autofocusing for digital holography," Optica **5**, 337–344 (2018).
31. R. Horisaki, R. Takagi, and J. Tanida, "Deep-learning-generated holography," Appl. Opt. **57**, 3859–3863 (2018).
32. 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**, 704–710 (2018).
33. M. D. Hannel, A. Abdulali, M. O'Brien, and D. G. Grier, "Machine-learning techniques for fast and accurate feature localization in holograms of colloidal particles," Opt. Express **26**, 15221–15231 (2018).
34. G. Zhang, T. Guan, Z. Shen, X. Wang, T. Hu, D. Wang, Y. He, and N. Xie, "Fast phase retrieval in off-axis digital holographic microscopy through deep learning," Opt. Express **26**, 19388–19405 (2018).
35. L. Jae-Sung, "Autofocusing using deep learning in off-axis digital holography," in *Digital Holography and Three-Dimensional Imaging* (Optical Society of America, 2018), paper DTh1C-4.
36. T. Pitkäaho, A. Manninen, and T. J. Naughton, "Classification of digital holograms with deep learning and hand-crafted features," in *Digital*

Holography and Three-Dimensional Imaging (Optical Society of America, 2018), paper DW2F-3.
37. T. Nguyen, V. Bui, and G. Nehmetallah, "3D optical diffraction tomography using deep learning," in *Digital Holography and Three-Dimensional Imaging* (Optical Society of America, 2018), paper DW2F-4.
38. H. Wang, M. Lyu, and G. Situ, "eHoloNet: a learning-based end-to-end approach for in-line digital holographic reconstruction," Opt. Express **26**, 22603–22614 (2018).
39. Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," Neural Comput. **1**, 541–551 (1989).
40. E. Cuche, F. Bevilacqua, and C. Depeursinge, "Digital holography for quantitative phase-contrast imaging," Opt. Lett. **24**, 291–293 (1999).
41. J. W. Goodman, *Introduction to Fourier Optics* (Roberts and Company, 2005).
42. J. Gillespie and R. A. King, "The use of self-entropy as a focus measure in digital holography," Pattern Recogn. Lett. **9**, 19–25 (1989).
43. P. Ferraro, G. Coppola, S. D. Nicola, A. Finizio, and G. Pierattini, "Digital holographic microscope with automatic focus tracking by detecting sample displacement in real time," Opt. Lett. **28**, 1257–1259 (2003).
44. M. Liebling and M. Unser, "Autofocus for digital Fresnel holograms by use of a Fresnelet-sparsity criterion," J. Opt. Soc. Am. A **21**, 2424–2430 (2004).
45. F. Dubois, C. Schockaert, N. Callens, and C. Yourassowsky, "Focus plane detection criteria in digital holography microscopy by amplitude analysis," Opt. Express **14**, 5895–5908 (2006).
46. C. P. McElhinney, J. B. McDonald, A. Castro, Y. Frauel, B. Javidi, and T. J. Naughton, "Depth-independent segmentation of macroscopic three-dimensional objects encoded in single perspectives of digital holograms," Opt. Lett. **32**, 1229–1231 (2007).
47. P. Langehanenberg, B. Kemper, D. Dirksen, and G. von Bally, "Autofocusing in digital holographic phase contrast microscopy on pure phase objects for live cell imaging," Appl. Opt. **47**, D176–D182 (2008).
48. P. Memmolo, M. Iannone, M. Ventre, P. A. Netti, A. Finizio, M. Paturzo, and P. Ferraro, "On the holographic 3D tracking of in vitro cells characterized by a highly-morphological change," Opt. Express **20**, 28485–28493 (2012).
49. J. Dohet-Eraly, C. Yourassowsky, and F. Dubois, "Fast numerical autofocus of multispectral complex fields in digital holographic microscopy with a criterion based on the phase in the Fourier domain," Opt. Lett. **41**, 4071–4074 (2016).
50. G. He, W. Xiao, and F. Pan, "Automatic focus determination through cosine and modified cosine score in digital holography," Opt. Eng. **56**, 034103 (2017).
51. Z. Ren, N. Chen, and E. Y. Lam, "Automatic focusing for multisectional objects in digital holography using the structure tensor," Opt. Lett. **42**, 1720–1723 (2017).
52. M. Lyu, C. Yuan, D. Li, and G. Situ, "Fast autofocusing in digital holography using the magnitude differential," Appl. Opt. **56**, F152–F157 (2017).
53. E. Cuche, P. Marquet, and C. Depeursinge, "Spatial filtering for zero-order and twin-image elimination in digital off-axis holography," Appl. Opt. **39**, 4070–4075 (2000).
54. Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: convolutional architecture for fast feature embedding," arXiv:1408.5093 (2014).
55. S. J. Pan and Q. Yang, "A survey on transfer learning," IEEE Trans. Knowl. Data Eng. **22**, 1345–1359 (2010).