This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.

Xu, Zhimin; Zuo, Si; Lam, Edmund Y.

End-to-end learning for digital hologram reconstruction

Published in:
High-Speed Biomedical Imaging and Spectroscopy III

DOI:
[10.1117/12.2288141](https://doi.org/10.1117/12.2288141)

Published: 01/01/2018

Document Version
Publisher's PDF, also known as Version of record

Please cite the original version:
Xu, Z., Zuo, S., & Lam, E. Y. (2018). End-to-end learning for digital hologram reconstruction. In *High-Speed Biomedical Imaging and Spectroscopy III: Toward Big Data Instrumentation and Management* (Vol. 10505). [1050510] SPIE. <https://doi.org/10.1117/12.2288141>

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

PROCEEDINGS OF SPIE

SPIDigitalLibrary.org/conference-proceedings-of-spie

End-to-end learning for digital hologram reconstruction

Zhimin Xu, Si Zuo, Edmund Y. Lam

Zhimin Xu, Si Zuo, Edmund Y. Lam, "End-to-end learning for digital hologram reconstruction," Proc. SPIE 10505, High-Speed Biomedical Imaging and Spectroscopy III: Toward Big Data Instrumentation and Management, 1050510 (20 February 2018); doi: 10.1117/12.2288141

SPIE.

Event: SPIE BiOS, 2018, San Francisco, California, United States

End-to-End Learning for Digital Hologram Reconstruction

Zhimin Xu^a, Si Zuo^b, and Edmund Y. Lam^c

^aSharpSight Limited, Hong Kong

^bSchool of Electrical Engineering, Aalto University, Espoo, Finland

^cDepartment of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong

ABSTRACT

Digital holography is a well-known method to perform three-dimensional imaging by recording the light wavefront information originating from the object. Not only the intensity, but also the phase distribution of the wavefront can then be computed from the recorded hologram in the numerical reconstruction process. However, the reconstructions via the traditional methods suffer from various artifacts caused by twin-image, zero-order term, and noise from image sensors. Here we demonstrate that an end-to-end deep neural network (DNN) can learn to perform both intensity and phase recovery directly from an intensity-only hologram. We experimentally show that the artifacts can be effectively suppressed. Meanwhile, our network doesn't need any preprocessing for initialization, and is comparably fast to train and test, in comparison with the recently published learning-based method. In addition, we validate that the performance improvement can be achieved by introducing a prior on sparsity.

Keywords: Digital holography, Deep neural network, Deep learning, Phase retrieval, Computational imaging

1. INTRODUCTION

Digital holography (DH) has become a technique utilized widely for three dimensional imaging. The capability of recovering the complex amplitude distribution scattered by the sample permits post-processing, like numerical refocus and a quantitative measurement of the sample phase.¹ Due to the benefit of such special characteristics, DH is used in many applications in different fields of science and technology.²⁻⁴

The basic pipeline in digital hologram reconstruction is firstly to locate the object, then to reconstruct complex distribution of the wavefront, and finally phase unwrapping to recover the phase. As each step in the reconstruction process may suffer from noise or artifacts, the accumulated results become worse. Numerous algorithms have been proposed. However, they only focus on one specific problem in the pipeline, so the whole process results are still unsatisfied.

Deep learning is a rapidly developing area and has been proved to be successful in many applications.⁵⁻⁷ Recently, several works have tried to use this technique in digital holography.⁸⁻¹² However, these combination of digital holography and deep learning still focus on solving the problems encountered in traditional pipeline.

In this paper, we propose and demonstrate a deep learning network trained end-to-end, pixel-to-pixel on digital holograms with the knowledge of the corresponding ground truth. The major contribution is that the proposed network is capable of reconstructing both intensity and phase directly from a hologram without following the traditional reconstruction pipeline. As seen from the real experimental results, our well-trained network shows remarkable performance visually and quantitatively. The artifacts, such as zero-order, twin images and various types of noise, have been significantly suppressed.

Further author information: (Send correspondence to Zhimin Xu)
Zhimin Xu: E-mail: eezmxu@gmail.com

Figure 1: The structure of the proposed network. From left to middle is the encoding path while the rest is the decoding path. A skip connection is used to connect two residual blocks from the encoding path and decoding path respectively which own the same shapes. Each residual block contains a residual skip connection within the same path of its own.

2. METHODS

2.1 Network Structure

Fig. 1 shows our network structure. It mainly consists of two parts, the encoding part and the decoding part. The encoding part extracts features of interest in different scales while the symmetric decoding part performs the reconstruction. Four basic blocks are used to construct our network. The gray block is a traditional convolution layer followed by parametric rectified linear unit activation and batch normalization. Following the convolution layer in encoding path, a max-pooling layer performs downsampling so as to reduce features numbers as well as parameters. Upsampling blocks appear in decoding path to upsample input data with interpolations. The yellow block represents a residual layer consisting of three convolution layers and one skip connection. The detail network is shown in Table 1.

The number of kernel starts from sixteen. In the encoding path, number of kernels are doubled after max-pooling are performed. A bridge layer is used to increase the number of feature maps. Following the bridge layer, decoding path starts to upsample input data with halving feature maps to construct a symmetric network. A long skip connection links the encoding block and decoding block with the same shape so as to add the feature maps. For each residual block a short skip connection links the input and the output of the block. The outputs of network are two channel feature maps, that is because we not only reconstruct the intensity but also the phase information from an intensity-only hologram.

The mean squared error (MSE) between the groundtruth images and network output images in pixel level is chosen as our loss function. During the training phase, loss is optimized by the Adaptive Moment Estimation (ADAM) optimizer¹³ given optimization parameters *decay steps* and *learning rate decay factor* in every batch.

We also tried to add a regularizer of the sparsity of network outputs to loss function with different weights. It is shown to be a way for improving training results, reducing computational load and fast convergence. Such a modification with proper weight somehow improves our training results when other parameters stay the same.

Table 1: Our network structure in details. The outputs of network are two channel feature maps, one for reconstructed image intensity and the other for reconstructed image phase.

Block Type	Detail	Size of feature map
inputs		512x640x1
down 1	conv + res + conv + maxpooling	512x640x16 256x320x16
down 2	conv + res + conv + maxpooling	256x320x32 128x160x32
down 3	conv + res + conv + maxpooling	128x160x64 64x80x64
down 4	conv + res + conv + maxpooling	64x80x128 32x40x128
bridge	conv + res + conv	32x40x256
up 4	deconv + merge + conv + res + conv	64x80x128 64x80x128
up 3	deconv + merge + conv + res + conv	128x160x64 128x160x64
up 2	deconv + merge + conv + res + conv	256x320x32 256x320x32
up 1	deconv + merge + conv + res + conv	512x640x16 512x640x16
outputs	conv	512x640x2

3. EXPERIMENTAL RESULTS

3.1 Optical System

As shown in Figure 2, the optical setup used in this paper is based on a Mach-Zehnder interferometer working in the off-axis architecture. Specifically, SF is spatial filter. L is collimation lens. BE is beam expander. HWP1

Figure 2: The schematic diagram of the experimental DH system.

Figure 3: Example of hologram inputs to the network

and HWP2 are half-waveplates. PBS and BS are polarization and non-polarization beam splitters. M1 and M2 are mirrors. OBJ is the object. PD is the camera.

3.2 Experimental Details

The object in our experiments is a small region of a negative USAF 1951 resolution chart. We capture the holograms at 10 different distances to the camera with equal interval. Each distance has 1,000 holograms. That is 10,000 in total with 10 different labels. We use 75% of the data as the training set, 15% for validation and 10% for testing.

Our network is implemented using Tensorflow.¹⁴ The deployment and training of deep learning networks are realized on a server equipped with two NVIDIA GTX Geforce 1080 GPUs. We train the network with a initial learning rate of 0.001 and a decay factor of 0.9 after 20 epochs. Batch size is set to 16 and the total epoch is 100.

3.3 Results

As a comparison, we also implement the latest published network in Ref. 10. Fig. 3 is an example of the input hologram. Fig. 4 shows the ground truth as well as the outputs from different networks. As seen from the reconstructed results, including both intensity and phase, our proposed network visually obtains almost the same results as the ground truth. There is no obvious artifacts caused by twin-image, zero-order term, and noise from the image sensor. However, when it comes to the results from the network in Ref. 10 shown in Fig. 3(c) and (f), blur appears on character borders. Meanwhile, the details are not well preserved.

To test the robustness of our proposed network, we add further Gaussian noise in different levels to the input holograms and take the network in Ref. 10 as a comparison. As shown in Table 2, when the standard deviation of Gaussian noise increases from 0 to 15, the loss value of the comparison network is more than two times as large as that of our network. Especially when noise level is greater than 15, the loss value of comparison network increases significantly, which shows our proposed network is more robust.

4. CONCLUSION

In this paper, we demonstrate an end-to-end deep neural network architecture for the digital hologram reconstruction. The experimental results on the real data show that our network is able to learn to perform both intensity and phase recovery directly from an intensity-only hologram. There are no obvious artifacts which may appear when using the traditional methods as well as the other learning-based methods.

Figure 4: The reconstruction results with different methods.

Table 2: Average loss on test data. Gaussian noise is added to the test holograms with different standard deviations $\sigma = 0, 5, 15, 20, 30$, which is shown in the first column of table.

Noise level	Comparison network	Proposed network
0	0.00573	0.00225
5	0.00575	0.00205
15	0.00657	0.00236
20	0.00762	0.00251
30	0.01224	0.00296

REFERENCES

- [1] Schnars, U., Falldorf, C., Watson, J., and Jüptner, W., “Digital holography,” in *[Digital Holography and Wavefront Sensing]*, 39–68, Springer (2015).
- [2] Doblas, A., Sánchez-Ortiga, E., Martínez-Corral, M., Saavedra, G., and Garcia-Sucerquia, J., “Accurate single-shot quantitative phase imaging of biological specimens with telecentric digital holographic microscopy,” *J. Biomed. Opt.* **19**(4), 046022–046022 (2014).
- [3] Pourvais, Y., Asgari, P., Abdollahi, P., Khamedi, R., and Moradi, A.-R., “Microstructural surface characterization of stainless and plain carbon steel using digital holographic microscopy,” *J. Opt. Soc. Am. B* **34**(5), B36–B41 (2017).
- [4] Cuhe, E., Marquet, P., and Depeursinge, C., “Simultaneous amplitude-contrast and quantitative phase-contrast microscopy by numerical reconstruction of fresnel off-axis holograms,” *Appl. Opt.* **38**(34), 6994–7001 (1999).

- [5] Krizhevsky, A., Sutskever, I., and Hinton, G. E., “Imagenet classification with deep convolutional neural networks,” in [*Advances in Neural Information Processing Systems*], 1097–1105 (2012).
- [6] Collobert, R. and Weston, J., “A unified architecture for natural language processing: Deep neural networks with multitask learning,” in [*International Conference on Machine Learning*], 160–167, ACM (2008).
- [7] Xu, Y., Mo, T., Feng, Q., Zhong, P., Lai, M., Eric, I., and Chang, C., “Deep learning of feature representation with multiple instance learning for medical image analysis,” in [*International Conference on Acoustics, Speech and Signal Processing (ICASSP)*], 1626–1630, IEEE (2014).
- [8] Yevick, A., Hannel, M., and Grier, D. G., “Machine-learning approach to holographic particle characterization,” *Opt. Express* **22**(22), 26884–26890 (2014).
- [9] Park, H. S., Rinehart, M. T., Walzer, K. A., Chi, J.-T. A., and Wax, A., “Automated detection of p. falciparum using machine learning algorithms with quantitative phase images of unstained cells,” *PloS One* **11**(9), e0163045 (2016).
- [10] Rivenson, Y., Zhang, Y., Gunaydin, H., Teng, D., and Ozcan, A., “Phase recovery and holographic image reconstruction using deep learning in neural networks,” *Light: Science & Applications* **141**, doi: 10.1038 (2017).
- [11] Nguyen, T., Bui, V., Lam, V., Raub, C. B., Chang, L.-C., and Nehmetallah, G., “Automatic phase aberration compensation for digital holographic microscopy based on deep learning background detection,” *Opt. Express* **25**(13), 15043–15057 (2017).
- [12] Sinha, A., Lee, J., Li, S., and Barbastathis, G., “Lensless computational imaging through deep learning,” *arXiv:1702.08516* (2017).
- [13] Kingma, D. and Ba, J., “Adam: A method for stochastic optimization,” *arXiv:1412.6980* (2014).
- [14] Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., et al., “Tensorflow: Large-scale machine learning on heterogeneous distributed systems,” *arXiv:1603.04467* (2016).