Analysis of non-iterative phase retrieval based on machine learning

Yohei Nishizaki1,2 · Ryoichi Horisaki1,3 · Katsuhisa Kitaguchi2 · Mamoru Saito2 · Jun Tanida1

Received: 4 September 2019 / Accepted: 24 December 2019 / Published online: 9 January 2020
© The Author(s) 2020

Abstract
In this paper, we analyze a machine-learning-based non-iterative phase retrieval method. Phase retrieval and its applications have been attractive research topics in optics and photonics, for example, in biomedical imaging, astronomical imaging, and so on. Most conventional phase retrieval methods have used iterative processes to recover phase information; however, the calculation speed and convergence with these methods are serious issues in real-time monitoring applications. Machine-learning-based methods are promising for addressing these issues. Here, we numerically compare conventional methods and a machine-learning-based method in which a convolutional neural network is employed. Simulations with several conditions show that the machine-learning-based method realizes fast and robust phase recovery compared with the conventional methods. We also numerically demonstrate machine-learning-based phase retrieval from noisy measurements with a noisy training data set for improving the noise robustness. The machine-learning-based approach used in this study may increase the impact of phase retrieval, which is useful in various fields, where phase retrieval has been used as a fundamental tool.

Keywords Phase retrieval · Machine learning · Deep learning · Convolutional neural network · Inverse problem

1 Introduction

Optical phenomena are described using waves in wave optics [1]. However, image sensors detect only the intensity of a light wave and disregard its phase. Interferometric methods, such as digital holography, have been widely used to observe both the amplitude and phase of a light wave [2–4]. Such methods have been applied to label-free biomedical sensing, which is an example of quantitative phase imaging [5, 6]. A disadvantage of interferometric methods, however, is the need for bulky optics to introduce the reference light.

Diffraction imaging is an alternative method to measure the complex amplitude (amplitude and phase) of a light wave with an intensity pattern of the diffracting field and without reference light [7–10].

The inverse problem of recovering the phase from an intensity image is known as phase retrieval [11–16]. Diffraction imaging has been used for X-ray imaging because imaging optics and highly coherent light sources that work in this spectral regime are difficult to fabricate [17, 18]. Recently, phase retrieval techniques have also been introduced in the visible regime for lensless imaging and speckle-correlation imaging [19–24].

Phase retrieval algorithms basically employ an iterative process between the object and sensor domains to recover the phase from the intensity, and they require many iterations to achieve sufficient convergence [14, 15]. Machine learning, such as deep learning, has recently been used for robust and fast phase retrieval. Methods for machine-learning-based phase retrieval may be categorized into two approaches. In the first approach, a deep neural network (DNN) is used as a denoiser in the iterative phase retrieval process [25, 26]. This approach improves the stability during the iterations. In the second approach, a DNN is used to calculate the inverse function in the phase retrieval problem [27, 28]. The second approach enables faster non-iterative phase retrieval than the conventional methods, and it has been used for real-time diffraction imaging, imaging through scattering...
media, computer-generated holograms, wavefront sensing, and pulse measurement [27–34]. Also, such DNN-based inversion has been introduced to optical sensing methods other than phase retrieval [35–38].

In this paper, we numerically analyze and compare conventional phase retrieval and machine-learning-based non-iterative phase retrieval in terms of the noise robustness and the calculation time. We also demonstrate enhancement of the noise robustness in the machine-learning-based phase retrieval by use of a noisy training data set. Here, for simplicity while maintaining versatility, we assume positive, real objects, and phase retrieval from Fourier intensity measurements [14–16]. The fast, noise-robust phase retrieval based on machine learning demonstrated in this paper will contribute to various fields, including biomedicine, security, and astronomy.

2 Method

In the optical setup assumed in this study, an object field \( g \) propagates towards a sensor plane located in the far field and is captured as a single intensity image, as shown in Fig. 1. This imaging process is modeled as follows:

\[
I = |G|^2 = |\mathcal{F}[g]|^2, \tag{1}
\]

where \( \mathcal{F}[\cdot] \) is the Fourier transform, \( G \) is the Fourier spectrum of the object field via Fraunhofer diffraction, and \( I \) is the captured intensity image. Here we assume that the object is non-negative and real, which are general assumptions in the fields of astronomy and crystallography [14–16]. These assumptions have been used for enhancing the convergence and uniqueness of the solution in conventional iterative phase retrieval.

The inverse function of (1) is written as

\[
g = \mathcal{H}[I], \tag{2}
\]

where \( \mathcal{H}[\cdot] \) is the inverse function for phase retrieval. The inverse problem has been recently solved non-iteratively using machine-learning-based approaches [26–31, 33]. In the present research, we use a convolutional residual network called ResNet [27, 31, 39] for calculating the inverse function in (2) non-iteratively. ResNet is a known practical network architecture, and it has been used in various applications. ResNet utilizes residual learning by skip connections to prevent vanishing/exploding gradients during the training stage. In the case of non-iterative machine-learning-based phase retrieval, \( \mathcal{H} \) is regressed with ResNet by using a training data set. Two types of networks with different depths, as shown in Fig. 2, are investigated to compare their calculation costs and noise-robustness. The first network, which is called ResNet1 here, has one down-and-up-sampling process, as shown in Fig. 2a. The second network, which is called ResNet2 here, has two down-and-up-sampling processes, as shown in Fig. 2b. Here, “D” is a down-sampling block, as shown in Fig. 2c, “U” is an up-sampling block, as shown in Fig. 2d, and “S” is a convolutional block for a skip convolutional connection, as shown in Fig. 2e. The definitions of each layer are as follows: “BatchNorm” is batch normalization [40], and “ReLU” is a rectified linear unit [41]. “Conv(\( s \), \( l \))” and “TConv(\( s \), \( l \))” are, respectively, a 2D convolution and the transposed 2D convolution with a filter size \( s \) and a stride \( l \).

In this paper, the error reduction (ER) method and the hybrid input-output (HIO) method are employed as the baseline of the conventional iterative phase retrieval, as shown in Fig. 3 [14]. In this case, the inverse function \( \mathcal{H} \) is iteratively processed as follows:

1. Initial estimation of the Fourier spectrum \( G_n \) : \( G_n = A \exp(j\Phi_n) \), where \( A \) is the amplitude which is set as \( \sqrt{I} \), \( \Phi_n \) is the phase distribution, which is initially set randomly, and the subscript \( n \) is a counter of the iteration, which is initially set to one.
2. Calculation of the intermediately

![Flow of imaging and reconstruction](image)

**Fig. 1** Flow of imaging and reconstruction

![Diagram of a non-iterative machine-learning-based phase retrieval method. a ResNet1, b ResNet2, c D-block, d U-block, and e S-block](image)

![Diagram of a non-iterative machine-learning-based phase retrieval method. a ResNet1, b ResNet2, c D-block, d U-block, and e S-block](image)
estimated object field $g_n^\prime = \mathcal{F}^{-1}[G_n^\prime]$, where $\mathcal{F}^{-1}$ is the inverse Fourier transform. (3) Update of the object field $g_n^\prime$ : The estimated object field is refined with constraints on the object domain, which are mentioned in the next paragraph. (4) Calculation of the intermediately estimated Fourier spectrum $G_n^\prime$: $G_n^\prime = \mathcal{F}[g_n^\prime]$. (5) Update of the Fourier spectrum $G_n$: The amplitude of the Fourier spectrum is replaced by $\sqrt{I}$, and the counter $n$ is incremented by one. Steps (2)–(5) are iterated until the object field converges.

In the case of the ER method, the update rule at Step (3) in the $n$th iteration is written as

$$g_n^\prime(x, y) = \begin{cases} g_n^\prime(x, y), & (x, y \notin \eta), \\ 0, & (x, y \in \eta), \end{cases}$$

(3)

where $\eta$ is the set of all spatial positions that violate the constraints, and $x$ and $y$ are the lateral coordinates on the object plane. The update process of the HIO is written as

$$g_n(x, y) = \begin{cases} g_n^\prime(x, y), & (x, y \notin \eta), \\ g_{n-1}(x, y) - \beta g_n^\prime(x, y), & (x, y \in \eta), \end{cases}$$

(4)

where $\beta$ is a feedback parameter. In all numerical experiments in this study, the constraints on the object field are realness and non-negativity [14–16].

### 3 Analysis

The analyses were carried out by numerical simulations. The networks used here have different depths, as shown in Fig. 2. In addition, training data sets without and with noise were used for each of the networks to verify that their performance levels depended on the noise in the training data set. In the noisy training data set, white Gaussian noise was added to each of the captured images. The noise level was randomly set between 10 and 30 dB.

The object images were handwritten numbers randomly taken from the EMNIST database, as shown in Fig. 4a [42]. The pixel count of the original and captured images was 28 × 28. The training data set was composed of 200,000 pairs of the object and captured images, and the test data set was composed of 1000 pairs of the object and captured images, without any overlapping. A learning algorithm called Adam was used for optimizing the network with an initial learning ratio of $1 \times 10^{-5}$, a batch size of 100, and a number of epochs of 100 [43]. The loss function of the optimization was the mean squared error. The number of iterations in ER and HIO to achieve sufficient convergence was 200 [13]. The feedback parameter $\beta$ in (4) was 0.9. The code was implemented in Python and Keras and was executed on a computer with an Intel Xeon 6134 CPU running at 3.2 GHz, with 192 GB of RAM, and an NVIDIA Tesla V100 GPU with 16 GB of VRAM.

The phase retrievals with the two networks when using the noisy and noiseless training data sets and the ER and HIO methods were compared under different measurement noise levels, namely, signal-to-noise ratios (SNRs) of 10, 20, 30 and $\infty$ dB. The reconstructed results are shown in Fig. 4b. In the cases of ER and HIO, the ambiguity of the spatial shift and flip was compensated with a cross-correlation process. Some artifacts clearly appeared in the reconstructions of ResNet1, ER, and HIO. The reconstruction fidelity was evaluated with the normalized mean squared error (NMSE) between the original and estimated images as follows [44]:

$$\text{NMSE} = \frac{\sum_{x,y} (I(x,y) - \hat{I}(x,y))^2}{\sum_{x,y} \hat{I}(x,y)^2}.$$
where \( \hat{g} \) is the estimated object image obtained through phase retrieval. Here, \( M \) and \( N \) are the number of elements along the \( x \)- and \( y \)-axes, respectively. The average NMSEs with the whole test data set are shown in Fig. 5. The results in Figs. 4 and 5 show that phase retrieval with ResNet2 realized accurate and robust reconstructions, even for noisy measurements, compared with ER and HIO. Also, the deep network (ResNet2) gave better estimated results than the shallow one (ResNet1). The NMSEs of the reconstructions with ResNet2 were about 1.6-times smaller than those of ResNet1, ER, and HIO, as shown in Fig. 5. Furthermore, the results show that the noisy training data set enhanced the noise robustness of the networks. In addition, ResNet2 with a noisy training data set of the handwritten numbers was applied to measurements simulated from the handwritten alphabetic characters in the EMNIST database [42]. The measurement SNR was 10 dB. In this case, the reconstruction NMSE was 0.55, which shows that the network was over-fitted to the handwritten numbers, although it was comparable to ER and HIO.

A plot depicting the influence of the number of training pairs on the estimation accuracy obtained with ResNet2 using the noisy training data set is shown in Fig. 6. This result shows that a larger number of training pairs trivially reduces the estimation error. In addition, using the noisy training data set, it was found that there is a high noise-robustness with ResNet2 at each number of training pairs. Figure 7 summarizes the calculation times of ResNet1, ResNet2, ER, and HIO. The results show that the non-iterative machine-learning-based phase retrieval method was about thirty-times faster than the conventional methods. Here, it should be noted that the evaluations of ER and HIO were performed with the CPU using NumPy in Python for a fair comparison. In these cases, the calculation time of the Fourier transform with the GPU using CuPy in Python was about ten-times longer than that with the CPU. This was not caused by a latency of data transfer between the CPU and GPU, but by an inefficient parallelism of the GPU depending on the image size as shown in Ref. [45].

4 Conclusion

We numerically analyzed non-iterative phase retrieval algorithms based on machine learning in comparison with conventional iterative algorithms. The machine-learning-based
algorithms used convolutional neural networks called ResNet with different depths. In the numerical comparisons, the deeper network (ResNet2) realized better reconstructions than those realized by the shallower network (ResNet1) and conventional iterative algorithms, which were ER and HIO. Furthermore, an improvement in the noise robustness of these networks was verified by learning with noisy training data sets. A large training data set, for example, where the number of training pairs was 200,000, reduced the reconstruction error in the machine-learning-based algorithms. The machine-learning-based algorithms were one order of magnitude faster than the conventional iterative ones.

As demonstrated in this paper, deep convolutional neural networks are promising solutions for the phase retrieval problem, in terms of accurate reconstruction, noise robustness, and calculation speed. Phase retrieval has a long history, and the range of its applications is significant and still growing. The machine-learning-based approach studied here may solidify the impact of phase retrieval in various fields, such as life sciences and materials science.

**Funding** This work was supported by JSPS KAKENHI Grant Numbers JP17H02799 and JP17K00233, JST PRESTO Grant Number JPMJPR17PB.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

**References**

1. Goodman, J.W.: Introduction to Fourier Optics. McGraw-Hill, New York (1996)
2. Cuche, E., Bevilacqua, F., Depeursinge, C.: Digital holography for quantitative phase-contrast imaging. Opt. Lett. **24**, 291–3 (1999)
3. Mann, C.J., Yu, L., Lo, C.-M., Kim, M.K.: High-resolution quantitative phase-contrast microscopy by digital holography. Opt. Express **13**, 8693 (2005)
4. Xia, P., Shimozato, Y., Ito, Y., Tahara, T., Kakue, T., Awatsuji, Y., Nishio, K., Ura, S., Kubota, T., Matoba, O.: Improvement of color reproduction in color digital holography by using spectral estimation technique. Appl. Opt. **50**, H177 (2011)
5. Shaffer, E., Moratal, C., Magistretti, P., Marquet, P., Depeursinge, C.: Label-free second-harmonic phase imaging of biological specimen by digital holographic microscopy. Opt. Lett. **35**, 4102 (2010)
6. Belashov, A.V., Zhikhoreva, A.A., Belyaev, T.N., Kornilova, E.S., Petrov, N.V., Salova, A.V., Semenova, I.V., Vasyutinskii, O.S.: Digital holographic microscopy in label-free analysis of cultured cells’ response to photodynamic treatment. Opt. Lett. **41**, 5035 (2016)
7. Maleki, M.H., Devaney, A.J.: Phase-retrieval and intensity-only reconstruction algorithms for optical diffraction tomography. J. Opt. Soc. Am. A **10**, 1086–1092 (1993)
8. Marathe, S., Kim, S.S., Kim, S.N., Kim, C., Kang, H.C., Nickles, P.V., Noh, D.Y.: Coherent diffraction surface imaging in reflection geometry. Opt. Express **18**, 7253 (2010)
9. Burvall, A., Lundström, U., Takman, P.A.C., Larsson, D.H., Hertz, H.M.: Phase retrieval in X-ray phase-contrast imaging suitable for tomography. Opt. Express **19**, 10359 (2011)
10. Latychevskaia, T., Longchamp, J.-N., Fink, H.-W.: When holography meets coherent diffraction imaging. Opt. Express **20**, 28871 (2012)
11. Sayre, D.: Some implications of a theorem due to Shannon. Acta Crystallogr. **5**, 843–843 (1952)
12. Misell, D.L.: An examination of an iterative method for the solution of the phase problem in optics and electron optics: II. Sources of error. J. Phys. D Appl. Phys. **6**, 2217–2225 (1973)
13. Fienup, J.R.: Reconstruction of an object from the modulus of its Fourier transform. Opt. Lett. **3**, 27–9 (1978)
14. Fienup, J.R.: Phase retrieval algorithms: a comparison. Appl. Opt. **21**, 2758–2769 (1982)
15. Fienup, J.R., Wackerman, C.C.: Phase-retrieval stagnation problems and solutions. J. Opt. Soc. Am. A **3**, 1897 (1986)
16. Millane, R.P.: Phase retrieval in crystallography and optics. J. Opt. Soc. Am. A **7**, 394–411 (1990)
17. Miao, J., Sayre, D., Chapman, H.N.: Phase retrieval from the magnitude of the Fourier transforms of nonperiodic objects. J. Opt. Soc. Am. A **15**, 1662 (1998)
18. Miao, J., Charalambous, P., Kirz, J., Sayre, D.: Extending the methodology of X-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens. Nature **400**, 342–344 (1999)
19. Katkovnik, V., Shevkunov, I., Petrov, N.V., Egiazarian, K.: Computational super-resolution phase retrieval from multiple phase-coded diffraction patterns: simulation study and experiments. Optica **4**, 786 (2017)
20. Shevkunov, I., Katkovnik, V., Petrov, N.V., Egiazarian, K.: Super-resolution microscopy for biological specimens: lensless phase retrieval in noisy conditions. Biomed. Opt. Express **9**, 5511 (2018)
21. Bertolotti, J., Van Putten, E.G., Blum, C., Lagendijk, A., Vos, W.L., Mosk, A.P.: Non-invasive imaging through opaque scattering layers. Nature **491**, 232–234 (2012)
22. Katz, O., Heidmann, P., Fink, M., Gigan, S.: Non-invasive single-shot imaging through scattering layers and around corners via speckle correlations. Nat. Photon. **8**, 784–790 (2014)
23. Okamoto, Y., Horisaki, R., Tanida, J.: Noninvasive three-dimensional imaging through scattering media by three-dimensional speckle correlation. Opt. Lett. **44**, 2526 (2019)
24. Horisaki, R., Okamoto, Y., Tanida, J.: Single-shot noninvasive three-dimensional imaging through scattering media. Opt. Lett. **44**, 4032–4035 (2019)
25. Metzler, C., Schniter, P., Veeraraghavan, A., baraniuk, R.: prDeep: robust phase retrieval with a flexible deep network. In: Proceedings of the 35th international conference on machine learning (ICML’18) (JMLR, 2018), vol. 80, pp. 3501–3510
26. Isil, Ç., Öktem, F.S., Köc, A.: Deep iterative reconstruction for phase retrieval. Appl. Opt. **58**, 5422–5431 (2019)
27. Sinha, A., Lee, J., Li, S., Barbastathis, G.: Lensless computational imaging through deep learning. Optica **4**, 1117–1125 (2017)
28. Cherukara, M.J., Nashed, Y.S., Harder, R.J.: Real-time coherent diffraction inversion using deep generative networks. Sci. Rep. 8, 1–8 (2018)
29. Horisaki, R., Takagi, R., Tanida, J.: Learning-based imaging through scattering media. Opt. Express 24, 13738 (2016)
30. Horisaki, R., Takagi, R., Tanida, J.: Learning-based focusing through scattering media. Appl. Opt. 56, 4358 (2017)
31. Horisaki, R., Takagi, R., Tanida, J.: Deep-learning-generated holography. Appl. Opt. 57, 3859–3863 (2018)
32. Paine, S.W., Fienup, J.R.: Machine learning for improved image-based wavefront sensing. Opt. Lett. 43, 1235 (2018)
33. Nishizaki, Y., Valdivia, M., Horisaki, R., Kitaguchi, K., Saito, M., Tanida, J., Vera, E.: Deep learning wavefront sensing. Opt. Express 27, 240 (2019)
34. White, J., Chang, Z.: Attosecond streaking phase retrieval with neural network. Opt. Express 27, 4799 (2019)
35. Yuan, X., Pu, Y.: Parallel lensless compressive imaging via deep convolutional neural networks. Opt. Express 26, 1962 (2018)
36. Kürüm, U., Wiecha, P.R., French, R., Muskens, O.L.: Deep learning enabled real-time speckle recognition and hyperspectral imaging using a multimode fiber array. Opt. Express 27, 20965 (2019)
37. Van der Jeught, S., Dirckx, J.J.J.: Deep neural networks for single shot structured light profilometry. Opt. Express 27, 17091 (2019)
38. Shimobaba, T., Takahashi, T., Yamamoto, Y., Endo, Y., Shiraki, A., Nishitsuji, T., Hoshikawa, N., Kakue, T., Ito, T.: Digital holographic particle volume reconstruction using a deep neural network. Appl. Opt. 58, 1900 (2019)
39. 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), pp. 770–778 (2016)
40. Ioffe, S., Szegedy, C.: Batch normalization: accelerating deep network training by reducing internal covariate shift. In: Proceedings of the 32nd International Conference on Machine Learning (ICML’15) (JMLR, 2015), vol. 37, pp. 448–456
41. Nair, V., Hinton, G.E.: Rectified linear units improve restricted Boltzmann machines. In: Proceedings of the 27th International Conference on International Conference on Machine Learning, (Omnipress, USA, 2010), ICML’10, pp. 807–814
42. Cohen, G., Afshar, S., Tapson, J., Van Schaik, A.: EMNIST: extending MNIST to handwritten letters. In: Proceedings of the International Joint Conference on Neural Networks, pp. 2921–2926 (2017)
43. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization. In: CoRR (2015)
44. Fienup, J.R.: Invariant error metrics for image reconstruction. Appl. Opt. 36, 8352 (1997)
45. Peternier, A., Boncori, J.P.M., Pasquali, P.: Near-real-time focusing of ENVISAT ASAR Stripmap and Sentinel-1 TOPS imagery exploiting OpenCL GPGPU technology. Remote Sens. Environ. 202, 45–53 (2017)

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