Deep convolutional neural network-based lensless quantitative phase retrieval

Shevkunov, Igor, Kilpeläinen, Jarkko, Eguiazarian, Karen

Igor Shevkunov, Jarkko Kilpeläinen, Karen Eguiazarian, "Deep convolutional neural network-based lensless quantitative phase retrieval," Proc. SPIE 11653, Quantitative Phase Imaging VII, 116531F (5 March 2021); doi: 10.1117/12.2581428

SPIE.

Event: SPIE BiOS, 2021, Online Only

Deep convolutional neural network-based lensless quantitative phase retrieval.

Igor Shevkunov, Jarkko Kilpeläinen, and Karen Egiazarian

Faculty of Information Technology and Communication Sciences, Tampere University, P.O. Box 553, FI-33101 Tampere, Finland

ABSTRACT

In this paper, we employ a deep convolutional neural network for the solution of the phase retrieval problem in a lensless optical system from a single observation. We utilize U-net-like structured DCNN to reconstruct phase from the amplitude images at the sensor plane, and after applying computational backpropagation, the complex objects' amplitude is reconstructed at the object plane. Results are demonstrated by simulation experiments.

1. INTRODUCTION

Phase problem is a problem of phase loss during wavefront sensing/registration in the ultraviolet to visible range illumination. The loss is happening due to the high frequency of the light in the given range, and currently existing electronic devices (cameras) are not fast enough to register the phase change of the light field, therefore they register only the squared amplitude of the field (intensity), but not phase. To overcome this problem, there are two solutions. The first one is holography, which utilizes the interference of disturbed by an object and undisturbed wavefronts.[1] It is a well-known and widespread technique, however, its main disadvantage is the need of undisturbed reference wavefront which makes the optical system complex. The second one is the phase retrieval technique, which typically utilizes a simpler optical system, but the phase reconstruction needs a set of different observations and sophisticated computational processing.[2] It is because phase retrieval is an ill-posed problem and the exact solution cannot be found analytically.

Nowadays, for solving ill-posed inverse problems, deep convolutional neural networks (DCNN) are utilized widely.[3] In different configurations they successfully applied in the tasks of super-resolution,[4] imaging through thick scattering media,[5] diffractive optical elements creation,[6,7] etc. For the phase problem solution, there are several possible applications. According to a review paper in 2019,[8] the first solution to a phase problem by CNN was done in 2017 by Sinha et al.[9] In the study, they developed and trained an end-to-end residual convolutional neural net (ResNet) to do phase retrieval and reconstruction of objects from observed diffraction patterns in a lensless system. The results show ResNet can perform phase retrieval inside (with controlled settings) and outside the training dataset. Another approach is to do phase retrieval through regularization by denoising (prDeep).[10] Rivenson with co-authors[11] utilized DCNN for phase reconstruction at object plane, with an initial backpropagation of the registered diffraction pattern to the object plane, where information from an object, inverse object, and zero-order are mixed. To filter amplitude and phase images, they exploited DCNN with several down-sampling steps to allow the network to learn how to suppress the appearing artifacts created by objects with different feature sizes. The same approach of suppression artifacts at the object plane is demonstrated in Ref.12 with an extension to bigger simulated training data sets. U-net structured DCNN[13] is used in Ref. 14 for phase reconstruction and uncertainty prediction, however, for a single reconstruction they need 5 images. Besides, the U-net is used in the iterative phase retrieval algorithm in Ref.15.

DCNN is a very powerful instrument that might solve a lot of complicated problems based only on heuristics and therefore with no guarantees to be optimal. Hence, to optimize the solution, we must be careful in the formulation of the problem and exclude from the learning the problems that are known and can be solved without the use of DCNN. In this manner, as the contribution of this paper, we are describing a novel approach to phase retrieval by phase reconstruction at the sensor plane by DCNN from the observed intensity, and after

Further author information: (Send correspondence to)
Igor Shevkunov: E-mail: igor.shevkunov@tuni.fi

Quantitative Phase Imaging VII, edited by Yang Liu, Gabriel Popescu, YongKeun Park, Proc. of SPIE Vol. 11653, 116531F · © 2021 SPIE · CCC code: 1605-7422/21/$21 · doi: 10.1117/12.2581428

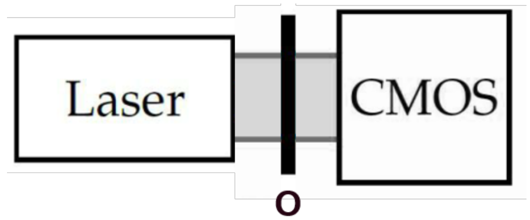

Figure 1. Proposed optical setup.'O' is an object and 'CMOS' is a lensless registration camera.

by wavefront backpropagation to the object plane, we reconstruct a full complex amplitude of the object. Thus, we exclude wavefront backpropagation from the learning. This approach is an optimization of the approach proposed in,[9] where phase at object plane was reconstructed by DCNN directly from the observed intensity at the sensor plane.

2. PROBLEM FORMULATION

In a lensless optical system configuration, the complex-valued light wavefront of an object is expressed by $u_0 = A \cdot \exp(i \cdot \varphi)$ and the phase retrieval problem is to find the phase φ from the observed intensity distribution I registered at the sensor plane:

$$\mathrm{I} = |P_d\{u_o\}|^2, \tag{1}$$

where $u_o \in \mathbb{C}^{N\times N}$ is the $N \times N$ complex-valued object wavefront (wavefront just after the object), $P_d : \mathbb{C}^{N\times N} \mapsto \mathbb{C}^{N\times N}$ stands for the free space propagation operator on the distance d.

The optical setup for the proposed problem is drawn in Fig.1, it is an as-simple-as-possible system with a coherent laser source, object 'o', and a sensor 'CMOS' camera. Light wavefront emitted by the laser propagates through the object and brings the information about the object structure to the sensor, where the intensity of the wavefront is registered. The simpler the optical setup, the harder the reconstruction, therefore all the burden of the reconstruction lies on the computational algorithm, in the frame of this paper, DCNN.

In earlier investigations in phase retrieval (e.g.[9]), DCNN used as a tool for direct reconstruction of the object phase from observation I, which includes both phase reconstruction and wavefront backpropagation. However, in our opinion, it is an overcomplicated task since the propagation model operators are well-known and directly correspond to physical wavefront propagation (as the solution of the Fresnel–Kirchhoff integral).[16] With the known amplitude and phase at the sensor plane, we can calculate the amplitude and phase at the object plane, making the reconstruction by backpropagation. Therefore, we suggest employing DCNN only for the phase reconstruction at the sensor plane, which will ease the DCNN learning process. Therefore, the solution for the object wavefront u_o is the result of the backpropagation of the wavefront reconstructed at the sensor plane, where the amplitude is calculated as a square root of the observed intensity, I, and the phase is reconstructed by DCNN, as it is written in the equation:

$$u_o = P_{-d}\{\sqrt{I} \cdot \exp(i \cdot \mathrm{DCNN}\{\sqrt{I}\})\}. \tag{2}$$

3. DCNN

We use U-net[13] DCNN structure as it is well applied in various imaging tasks and especially in phase retrieval.[15] The U-net architecture contains batch normalization layers and skip connection between the input and output, which helps the net to better learn the residual between the input amplitude and output phase images. The structure of the net is presented in Fig.2. It comprises 57 layers, the input image is 128×128 pixels amplitude, each convolutional layer kernel is 3×3, and the depth is changed gradually from 64 on the upper layer until 1024 at the bottom ones. Down-sampling is performed by max-pooling layers and up-sampling - by transpose convolutional layer with stride 2×2.

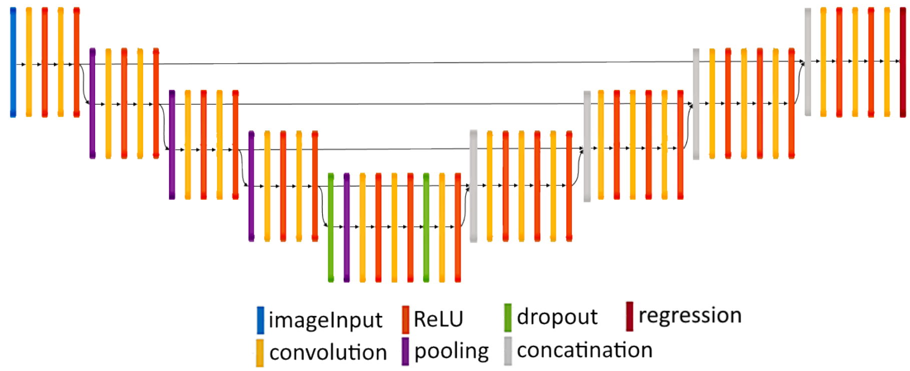

Figure 2. DCNN structure.

3.1 Training

For training data sets, we have calculated intensity distributions, I, according to (1) at the sensor plane and the corresponding phase distributions. For that purpose, we created a set of phase-only objects $u_o = A \cdot \exp(i\varphi)$, which means that amplitudes, A, are uniform and equal to 1; phases, φ, are images taken from the data set MNIST,[17] converted to gray scale, and scaled to a phase range of $[0, \pi]$. We took 49000 images for learning, 1000 for validation, and 10000 for testing.

For the learning process control, we used $l2$ distance between DCNN output and true phase images with the adaptive moment estimation optimizer with a gradient decay factor of 0.9, and a squared gradient factor of 0.99. The learning rate was step-wise starting from the value of 0.0001 and multiplied by 0.1 for each 10 epoch, batch size 64, a number of the epoch was equal to 30. Intensity distributions were modeled as free space propagation (angular spectrum) from the object to the sensor plane with wavelength 600 nm, the distance was taken equal to 2 mm, and camera pixel size was 3.45 μm. For proper propagation modeling, zero-padding of initial 28×28 objects to 128×128 was applied. The learning pairs for DCNN were 128×128 amplitudes (square root of intensity) at the sensor plane and corresponding phases.

For training procedure, we utilized Matlab 2020a on a Windows 10 XEON(r) W-2145 CPU computer with 64 Gb of RAM and GPU NVIDIA GeForce RTX 2080 Ti.

4. RESULTS AND DISCUSSION

We provide a comparison of two reconstruction approaches, which are based on different learning data sets. The first is the proposed one with phase reconstruction by DCNN at the sensor plane and backpropagation of the sensor wavefront to the object plane. Here the learning pairs are composed of amplitudes and phases, which are both from the sensor plane. The second is direct phase reconstruction by the DCNN at the object plane. In this case, the learning pairs are composed of amplitudes from the sensor plane but phases from the object plane.

In Figure 3 we show the randomly picked up central 28×28 pixels part of the reconstruction results for the images from the test MNIST dataset. Columns present different randomly picked objects from the test data set. Rows from up to the bottom present: observations (row 1), phases reconstructed by DCNN(row 2) and true phases(row 3) at the sensor plane; phases reconstructed by backpropagation (row 4) and true phases(row 5) at the object plane; row 6 is for the phases, which are reconstructed directly by DCNN from observations, without the use of backpropagation. Blue square emphasis results related to the proposed approach, and green square - related to the direct reconstruction. For better visualization, all presented images are normalized to a range of [0, 1] . Red numbers in the images correspond to a relative root-mean-square error (RRMSE) for the given reconstructed phase. RRMSE defines the accuracy of phase reconstruction by a criterion:

$$\mathrm{RRMSE}_{\varphi} = \frac{\sqrt{||\hat{\varphi}_{est} - \varphi_{true}||_2^2}}{\sqrt{||\varphi_{true}||_2^2}}, \tag{3}$$

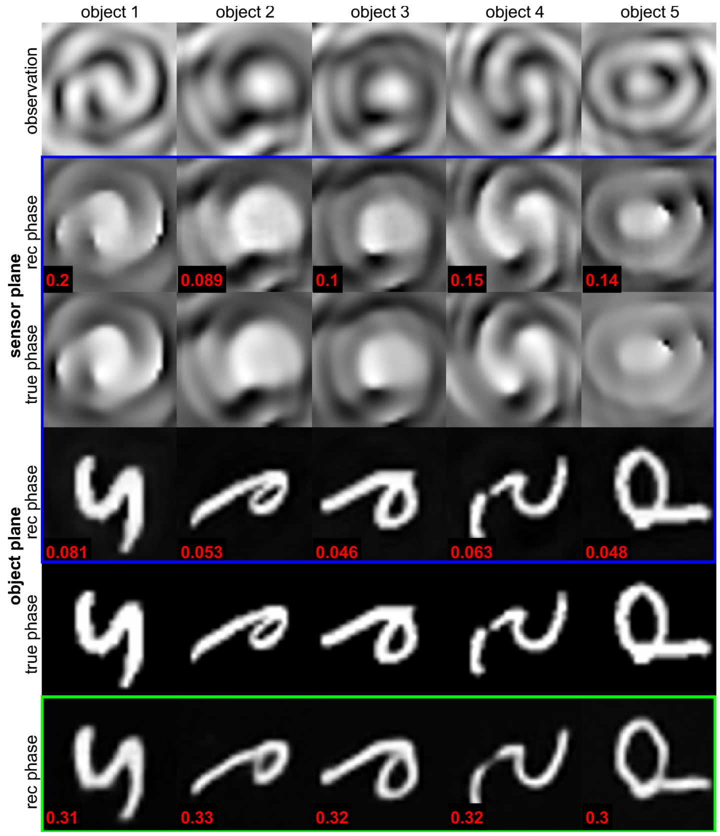

Figure 3. Images of the reconstruction results for five randomly picked objects from the test data set. Columns present images related to one object. The top row is for the observations, row 2 is for phases reconstructed by DCNN and row 3 is for true phases at the sensor plane, row 4 is for phases reconstructed by backpropagation and row 5 is for true phases at the object plane. Rows 2 and 4 are for the proposed approach. Row 6 represents reconstruction results for the direct phase reconstruction from the observations.

where $\hat{\varphi}_{est}$ is the reconstructed phase and φ_{true} is the noiseless true phase, $||.||$ stays for the Frobenius norm. RRMSE value less than 0.1 corresponds to a good quality phase reconstruction.

As it is seen, for the proposed approach, at the sensor plane, RRMSE values are bigger than at the object plane, which means that the backpropagation from the sensor plane to the object plane provides an enhancement in the reconstruction quality. It is explained using the amplitude as the square root of the observation (2), which adds a true value to the reconstruction. Both trained DCNNs demonstrate visually good reconstruction results, however the RRMSE values are low only for the proposed algorithm, indicating a good reconstruction of phase

Figure 4. Errors maps of phase reconstruction provided by the proposed algorithm, top row, and direct reconstruction, bottom row, for the same objects from Fig.3.

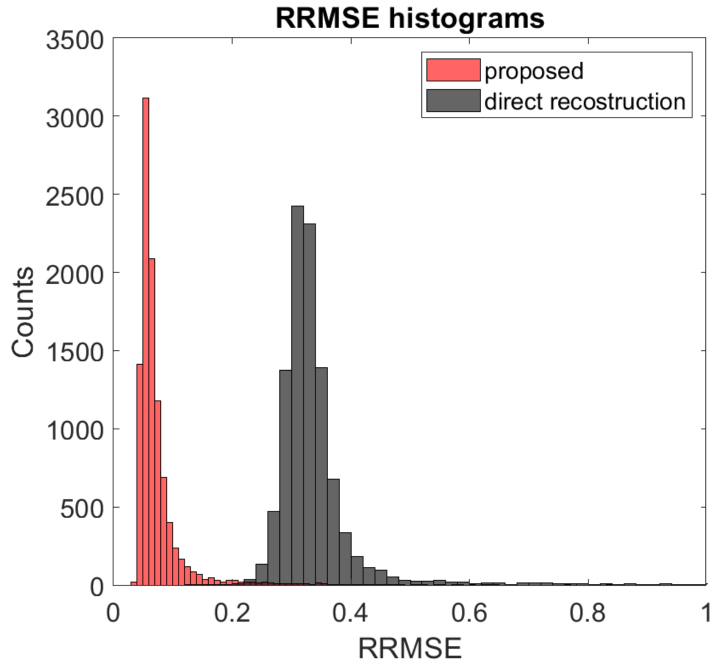

Figure 5. RRMSE histograms for phase test dataset reconstructions by proposed (red) and direct (gray) approaches.

values. RRMSE values bigger than 0.1 mean that the approach of direct reconstruction provides erroneous phase value estimation. In Fig.4 we provide reconstruction error maps which are calculated as the difference between true and reconstructed phases for both approaches for the same five objects. The top row is for the proposed approach and the bottom is for the direct reconstruction. It is seen that the errors in direct reconstruction are mainly caused by blurring of the edges, which means that high frequencies are lost during reconstruction.

For statistical analysis, we provide in Fig.5 RRMSE histograms of phase reconstruction by proposed approach (red color) and approach with direct reconstruction (gray) for the test data set. It is shown that for the given DCNN structure, the proposed approach significantly outperforms the direct reconstruction method with median RRMSE values of 0.062 and 0.32, respectively. Even though direct reconstruction provides visually reliable quality, it is only qualitative imaging, but the proposed approach provides quantitative precision for phase imaging.

In manner of phase imaging, taken MNIST training set provides relatively simple phase distributions with binary values, see the provided 5 images in Fig. 3. For the given setup parameters, at the modeled distance of 2 mm from the object to the sensor, the diffraction effect is strong and the observations do not visually correspond to objects (see rows 1 and 5 in Fig.3), therefore, the task of sharp phase image reconstruction becomes complex. From the other side, the phase reconstruction at the sensor plane seems simpler since the spatial distributions of the observations and phases are similar, see Fig.3 row 1 (observations) and row 3 (true phase at sensor plane), which additionally justify the proposed approach contrary to the direct reconstruction.

5. CONCLUSION

We have demonstrated the successful application of DCNN with the U-net structure for phase retrieval at a sensor plane from the observed intensity distribution. Lensless configuration of the proposed optical scheme makes phase reconstruction challenging, however, DCNN provides reliable and promising results. As further work, we plan to investigate learning on more complex datasets with 8-Bit intensity gradation, to utilize different net structures(e.g., ResNet[9]), and to improve nets by the exploitation of generative adversarial networks.[18]

ACKNOWLEDGMENTS

Jane and Aatos Erkko Foundation and Finland Centennial Foundation funded Computational Imaging without Lens (CIWIL) project.

REFERENCES

[1] Gabor and Gross, "Interfrence Microscope with total Wavefront Reconstruction," *Journal of the Optical Society of America* **56**(7) (1966).

[2] Gerchberg, R. W. and Saxton, W. O., "A Practical Algorithm for the Determination of Phase from Image and Diffraction Plane Pictures," *Phys. E. ppl. Opt. OPTIK* **2**(352), 237–246 (1969).

[3] Jin, K. H., McCann, M. T., Froustey, E., and Unser, M., "Deep Convolutional Neural Network for Inverse Problems in Imaging," *IEEE Transactions on Image Processing* **26**, 4509–4522 (9 2017).

[4] Kim, J., Lee, J. K., and Lee, K. M., "Accurate image super-resolution using very deep convolutional networks," in [*Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*], **2016-Decem**, 1646–1654 (2016).

[5] Lyu, M., Wang, H., Li, G., Zheng, S., and Situ, G., "Learning-based lensless imaging through optically thick scattering media," *Advanced Photonics* **1**(03), 1 (2019).

[6] Nadell, C. C., Huang, B., Malof, J. M., and Padilla, W. J., "Deep learning for accelerated all-dielectric metasurface design," *Optics Express* **27**, 27523 (9 2019).

[7] Cheremkhin, P., Evtikhiev, N., Krasnov, V., Rodin, V., Rymov, D., and Starikov, R., "Machine learning methods for digital holography and diffractive optics," *Procedia Computer Science* **169**, 440–444 (2020).

[8] Barbastathis, G., Ozcan, A., and Situ, G., "On the use of deep learning for computational imaging," *Optica* **6**, 921–943 (Aug 2019).

[9] Sinha, A., Lee, J., Li, S., and Barbastathis, G., "Lensless computational imaging through deep learning," *Optica* **4**(9), 1117 (2017).

[10] Metzler, C. A., Schniter, P., Veeraraghavan, A., and Baraniuk, R. G., "prDeep: Robust Phase Retrieval with a Flexible Deep Network," *35th International Conference on Machine Learning, ICML 2018* **8**, 5654–5663 (2 2018).

[11] Rivenson, Y., Zhang, Y., Günaydın, H., Teng, D., and Ozcan, A., "Phase recovery and holographic image reconstruction using deep learning in neural networks," *Light: Science & Applications* **7**, 17141–17141 (2 2018).

[12] Kemp, Z. D. C., "Propagation based phase retrieval of simulated intensity measurements using artificial neural networks," *Journal of Optics* **20**, 045606 (4 2018).

[13] Ronneberger, O., Fischer, P., and Brox, T., "U-Net: Convolutional Networks for Biomedical Image Segmentation," in [*Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*], **9351**, 234–241 (2015).

[14] Xue, Y., Cheng, S., Li, Y., and Tian, L., "Reliable deep-learning-based phase imaging with uncertainty quantification," *Optica* **6**, 618 (5 2019).

[15] Isıl, C., Oktem, F. S., and Koç, A., "Deep Iterative Reconstruction for Phase Retrieval," *Applied Optics* **58**, 5422 (4 2019).

[16] Goodman, J., [*Introduction to Fourier optics*], Roberts & Co (2005).

[17] LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P., "Gradient-based learning applied to document recognition," *Proceedings of the IEEE* **86**(11), 2278–2324 (1998).

[18] Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y., "Generative adversarial networks," *arXiv preprint arXiv:1406.2661* (2014).