Deep Learning Phase Recovery: Data-driven or Physics-driven?

Kaiqiang Wang and Edmund Y. Lam
Department of Electrical and Electronic Engineering
The University of Hong Kong, Pokfulam, Hong Kong SAR, China

2024 Photonics & Electromagnetics Research Symposium (PIERS) | 979-8-3503-7590-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/PIERS62282.2024.10618233

Abstract— Phase recovery is the process of computing the phase of a light field from its intensity measurements. This technique is vital in areas such as quantitative phase imaging, coherent diffraction imaging, and adaptive optics, as it enables the reconstruction of an object's refractive index distribution or topography, as well as the correction of imaging system aberrations. In recent years, deep learning has offered unparalleled assistance in addressing phase recovery problems via supervised or un/self-supervised learning modes. The data-driven supervised approach guides the convergence of neural networks through paired hologram-phase datasets, while the physics-driven un/self-supervised approach teaches neural networks through physical models to learn the mapping relationship between the hologram and phase. In this paper, we compare the data-driven and physics-driven approaches via simulation and analyze the advantages and disadvantages of each.

1. INTRODUCTION

As an electromagnetic field, light consists of amplitude and phase. Optical detectors, which typically depend on the conversion of photons to electrons (such as charge-coupled device sensors and the human eye), determine the intensity that corresponds to the square of the light field's amplitude but lose the phase. As the light field propagates, the phase delay also induces alterations in the amplitude distribution. Thus, we can capture the amplitude of the propagated light field and subsequently compute the relevant phase, the so-called phase recovery [1]. Some main parameters of the object can be further deduced from the recovered phase, such as thickness, refractive index distribution or surface profile of the sample, and aberrated wavefront caused by atmospheric turbulence.

Due to the loss of light field phase information on the recording plane, it is ill-posed to directly compute the light field phase or complex amplitude at the object plane from its intensity measurement. On the one hand, the light field phase can be retrieved bit by bit from intensity measurements through iterative optimization, namely, phase retrieval [2]. On the other hand, extra information can be added to convert it into a well-posed problem and be solved directly, such as holography [3] or interferometry by introducing reference light, Shack-Hartmann wavefront sensing [4] by introducing micro-lens array, and transport of intensity equation [5] by introducing multiple through-focus amplitudes.

In deep-learning-based computational imaging [6], people train an artificial neural network with dataset-based or physical-model-based prior information to let it learn the mapping relationship from the measurement domain to the objective domain. This idea can also be used to perform phase recovery, i.e., training a neural network to learn the mapping relationship from the measured intensity to the phase of the light field. There are two main strategies to train/drive the neural network: data-driven (DD) [1] and physics-driven (PD) [1]. DD trains the neural network by paired input-label (hologram-phase [7–9], intensity-phase [10], distortion-phase [11]) datasets as the implicit prior. In PD, the neural network is trained by physical models as the explicit prior [12, 13]. In the following part, we introduce and compare these two strategies by simulation.

2. DATA-DRIVEN (DD) DEEP LEARNING PHASE RECOVERY

Data-driven deep learning phase recovery methods operate in supervised learning mode, relying on a large number of paired input-label datasets [1]. Typically, a considerable amount of intensity images, such as diffraction images or holograms, must be experimentally collected as input. Conventional methods are then used to calculate the corresponding phase as ground truth. The crucial aspect of this paired dataset is its implicit mapping relationship from intensity to phase. An untrained or initialized neural network is trained iteratively using the paired dataset as an *implicit prior*. The gradient of the loss function propagates through the neural network, updating its parameters. Once trained, the network serves as an end-to-end mapping tool, inferring the phase from intensity, as shown in Figure 1(a). For physical processes that can be well modeled, the paired datasets can also be generated by numerical simulation.

3. PHYSICS-DRIVEN (PD) DEEP LEARNING PHASE RECOVERY

In contrast to the dataset-driven approach that utilizes paired input-label datasets as an *implicit prior* for neural network training, the physics-driven (PD) approach employs physical models, such as numerical propagation, as an *explicit prior* to drive the inference or training of neural networks [1]. This method only requires sample measurements as an input-only dataset, making it an un/self-supervised learning mode.

The *explicit prior* can be used in two ways [1]. On the one hand, it can iteratively optimize an untrained neural network to infer the corresponding phase and amplitude from the measured intensity image as input. This process is known as the untrained PD (uPD) strategy, as shown in Figure 1(b). On the other hand, the *explicit prior* can train an untrained neural network using a large number of intensity images as input. The trained network then infers the corresponding phase from unseen intensity images. This approach is called the trained PD (tPD) strategy, as shown in Figure 1(c). In addition, we can refine the tested sample on a trained neural network, i.e., tPD with refinement (tPDr), as shown in Figure 1(d). The tPDr can be thought of as a combined version of uPD and tPD.

It should be pointed out that there is a method derived from deep image prior [14] that is very similar to uPD, which uses a fixed random vector or matrix as input to the neural network, called the Network in Physics (NiP) strategy [1]. In the NiP, the neural network simply acts as a generator, rather than establishing a mapping relationship between diffraction intensity and phase. Therefore, the NiP cannot pre-train the neural network like the PD, and has to start from an initialized neural network in every inference [1].

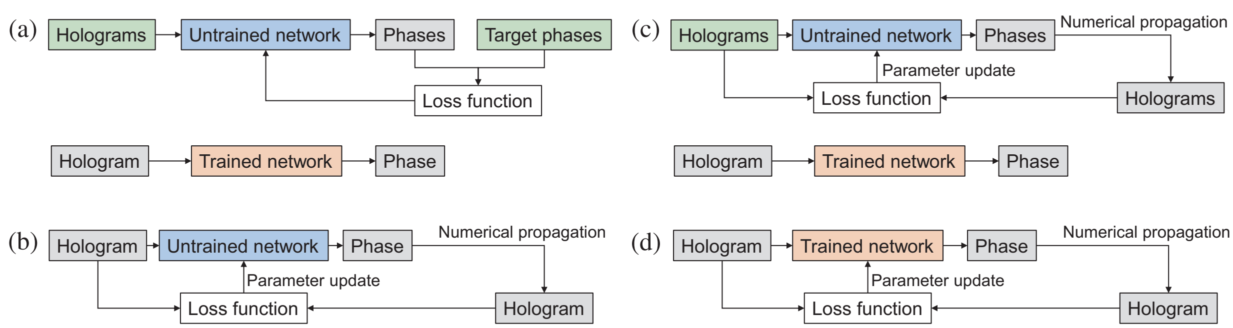

Figure 1: Description of dataset-driven and physics-driven deep learning phase recovery methods. (a) Network training and inference for the DD. (b) Network inference for the uPD. (c) Network training and inference for the tPD. (d) Network training and inference for the tPDr. Green represents the dataset, blue represents the untrained network, and orange represents the trained network.

4. COMPARISON OF DD AND PD

Here we only consider recovering the phase of the light field from its coherent diffraction intensity. We use a publicly available dataset (ImageNet) as the sample phase with values in the range of [0, 1] rad and generate the corresponding diffraction images using free-space propagation (20 mm). In this way, 10,100 pairs of datasets are generated, of which 10,000 pairs are used as training data, and the other 100 pairs are used as testing data. We use a previously proposed Res-UNet [15, 16] as the neural network frame. We set the neural network training parameters uniformly for all methods, e.g., the loss function: mean-square error and the optimizer: Adam. All diffraction and phase images are 256×256. All calculations are performed on the same GPU (NVIDIA RTX 3090).

For the 100 tested samples, we summarize the average inference time, average peak signal-to-noise ratio (PSNR) and average structural similarity index measure (SSIM) of all methods in Table 1, and show some example results in Figure 2.

We analyze as follows. The inference time of DD and tPD is the fastest because the tested diffraction image only needs to be run through the neural network once during the inference process of these two strategies. In addition, the inference process of uPD and tPDr requires multiple iterations on untrained or trained neural networks, making it more time-consuming. Since the initial neural network used for inference is pre-trained, tPDr can achieve higher accuracy than uPD with the same inference cycles and time. Under the same training conditions (same training datasets, same training parameter settings), the inference accuracy of tPD is basically the same as that of DD (Table 1), but the inference results of tPD show more detailed information (Figure 2). It should be noted that there is a small probability that uPD will collapse, i.e., the neural network outputs the wrong phase, even though the hologram corresponding to that phase is the same as the input hologram. This collapse does not occur in tPDr, because the pre-trained neural network can already provide an initial solution for iterative refinement that is very close to the target.

Table 1: Training setting and inference evaluation of DD, uPD, tPD, and tPDr.

Strategy	**training datasets**	**Inference cycles**	**Inference time**	**PSNR ↑**	**SSIM ↑**
DD	10,000 pairs	1	$\sim$ 0.02 seconds	19.6	0.68
uPD	0	1,000	$\sim$ 80 seconds	21.8	0.90
tPD	10,000 inputs	1	$\sim$ 0.02 seconds	18.0	0.69
tPDr	10,000 inputs	1,000	$\sim$ 80 seconds	25.9	0.93

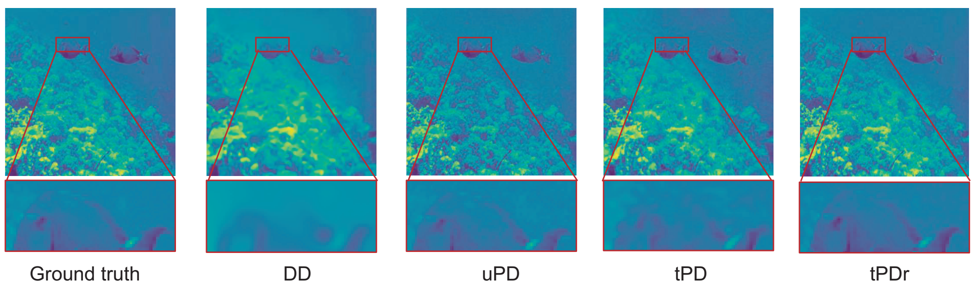

Figure 2: Inference results of DD, uPD, tPD, and tPDr.

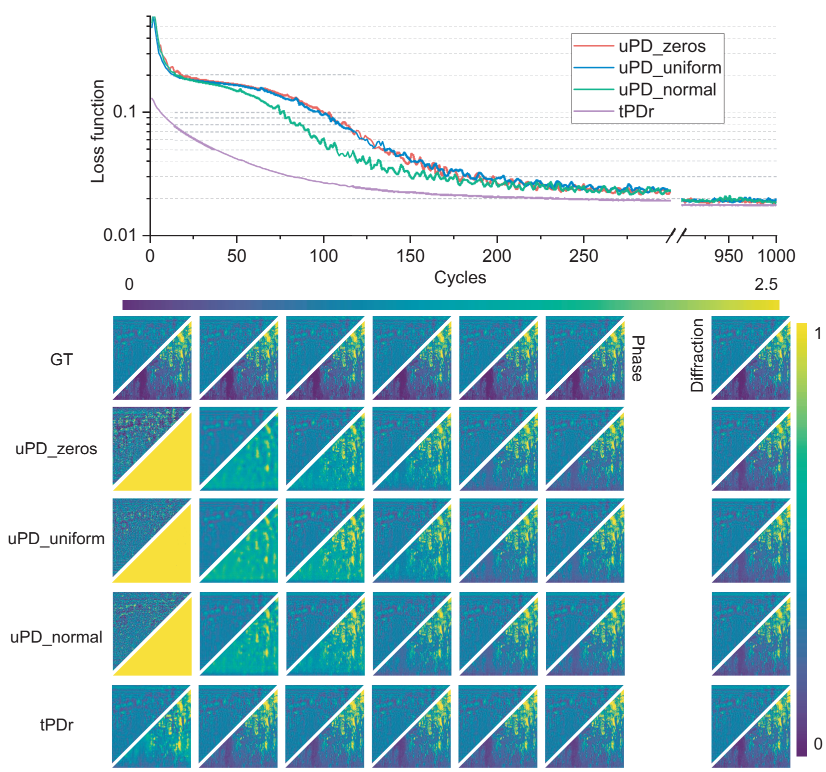

Figure 3: Visualization of uPD and tPD inference processes.

In addition, we compared tPDr and uPD under uniform distribution initialization, normal distribution initialization, and zero matrix initialization. Figure 3 shows the loss functions of these neural networks as they are circularly inferred, as well as the output phases and corresponding diffraction of the neural networks at some epochs. These specific results only show the range of values consistent with ground truth (GT). In Figure 3, all the values in the first result of uPD are equal to 1, which means that they are outside the maximum of GT. For uPD, no matter what initialization method is used, the network eventually converges to the same accuracy. uPD initialized with normal distribution converges faster than that with other methods. Due to the additional pre-training, the initial results of tPD have been very close to the real solution, the convergence speed is much faster than uPD, and the final accuracy is greater than uPD.

5. CONCLUSION

For deep learning phase recovery methods, we compared the characteristics of DD, uPD, tPD, and tPDr in terms of training settings, inference time, and inference accuracy. More in-depth comparisons will be done in the future.

ACKNOWLEDGMENT

The work was supported in part by the Research Grants Council of Hong Kong (GRF 17200321, GRF 17201822, RIF R7003-21).

REFERENCES

1. Wang, K., L. Song, C. Wang, Z. Ren, G. Zhao, J. Dou, J. Di, G. Barbastathis, R. Zhou, J. Zhao, and E. Y., Lam, "On the use of deep learning for phase recovery," *Light: Science and Applications*, Vol. 13, 4, 2024.
2. Song, L. and E. Y. Lam, "Fast and robust phase retrieval for masked coherent diffractive imaging," *Photonics Research*, Vol. 10, 758–768, 2022.
3. Zeng, T., Y. Zhu, and E. Y. Lam, "Deep learning for digital holography: A review," *Optics Express*, Vol. 29, 40572–40593, 2021.
4. Wang, C., Z. Ge, S. Zhu, P. Zhang, and E. Y. Lam, "Tracking the Shack-Hartmann spots using neuromorphic motion compensation," *Optica Topical Meeting in Computational Optical Sensing and Imaging*, CTu2B.5, 2023.
5. Lu, L., J. Li, Y. Shu, J. Sun, J. Zhou, E. Y. Lam, Q. Chen, and C. Zuo, "Hybrid brightfield and darkfield transport of intensity approach for high-throughput quantitative phase microscopy," *Advanced Photonics*, Vol. 4, 056002, 2022.
6. Lam, E. Y., "Artificial intelligence in biophotonics and imaging: Advancing computational reconstruction and inference," *Optoelectronics and Communications Conference*, T4A.2, 2021.
7. Ren, Z., Z. Xu, and E. Y. Lam, "Learning-based nonparametric autofocusing for digital holography," *Optica*, Vol. 5, 337–344, 2018.
8. Wang, K., J. Dou, Q. Kemao, J. Di, and J. Zhao, "Y-Net: A one-to-two deep learning framework for digital holographic reconstruction," *Optics Letters*, Vol. 44, 4765–4768, 2019.
9. Wang, K., Q. Kemao, J. Di, and J. Zhao, "Y4-Net: A deep learning solution to one-shot dual-wavelength digital holographic reconstruction," *Optics Letters*, Vol. 45, 4220–4223, 2020.
10. Wang, K., J. Di, Y. Li, Z. Ren, Q. Kemao, and J. Zhao, "Transport of intensity equation from a single intensity image via deep learning," *Optics and Lasers in Engineering*, Vol. 134, 106233, 2020.
11. Wang, K., M. Zhang, J. Tang, L. Wang, L. Hu, X. Wu, W. Li, J. Di, G. Liu, and J. Zhao, "Deep learning wavefront sensing and aberration correction in atmospheric turbulence," *PhotoniX*, Vol. 2, 8, 2021.
12. Wang, F., Y. Bian, H. Wang, M. Lyu, G. Pedrini, W. Osten, G. Barbastathis, and G. Situ, "Phase imaging with an untrained neural network," *Light: Science and Applications*, Vol. 9, 77, 2020.
13. Huang, L., H. Chen, T. Liu, and A. Ozcan, "Self-supervised learning of hologram reconstruction using physics consistency," *Nature Machine Intenlligence*, Vol. 5, 895–907, 2023.
14. Ulyanov, D., A. Vedaldi, and V. Lempitsky, "Deep image prior," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 9446–9454, 2018.
15. Wang, K., Q. Kemao, J. Di, and J. Zhao, "Deep learning spatial phase unwrapping: A comparative review," *Advanced Photonics Nexus*, Vol. 1, 014001, 2022.
16. https://github.com/kqwang/Phase_unwrapping_by_U-Net/blob/main/Network.py.