PHOTONICS Research

Dual-wavelength in-line digital holography with untrained deep neural networks

CHEN BAI,[1,†] ⓘ TONG PENG,[1,2,†] ⓘ JUNWEI MIN,[1] ⓘ RUNZE LI,[1] ⓘ YUAN ZHOU,[1] ⓘ AND BAOLI YAO[1,3,*] ⓘ

[1]*State Key Laboratory of Transient Optics and Photonics, Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, China*
[2]*Xi'an Jiaotong University, Xi'an 710049, China*
[3]*Pilot National Laboratory for Marine Science and Technology (Qingdao), Qingdao 266200, China*
*Corresponding author: yaobl@opt.ac.cn

Dual-wavelength in-line digital holography (DIDH) is one of the popular methods for quantitative phase imaging of objects with non-contact and high-accuracy features. Two technical challenges in the reconstruction of these objects include suppressing the amplified noise and the twin-image that respectively originate from the phase difference and the phase-conjugated wavefronts. In contrast to the conventional methods, the deep learning network has become a powerful tool for estimating phase information in DIDH with the assistance of noise suppressing or twin-image removing ability. However, most of the current deep learning-based methods rely on supervised learning and training instances, thereby resulting in weakness when it comes to applying this training to practical imaging settings. In this paper, a new DIDH network (DIDH-Net) is proposed, which encapsulates the prior image information and the physical imaging process in an untrained deep neural network. The DIDH-Net can effectively suppress the amplified noise and the twin-image of the DIDH simultaneously by automatically adjusting the weights of the network. The obtained results demonstrate that the proposed method with robust phase reconstruction is well suited to improve the imaging performance of DIDH. © 2021 Chinese Laser Press

https://doi.org/10.1364/PRJ.441054

1. INTRODUCTION

Digital holography (DH) can "capture and freeze" the wavefront of an object wave and realize lensless imaging based on interference [1]. The ability to recover phases makes DH widely used in biomedicine and materials science as a means of quantitative phase imaging [2]. Typically, DH employs two major configurations: in-line and off-axis structures [3]. Even though the off-axis technique allows wavefront reconstruction from a single-shot digital hologram, space bandwidth and resolution losses are frequently introduced [4,5]. In comparison, in-line DH, with its relatively simple and compact setup, is often preferred in many microscopic imaging techniques [5]. However, the original phase map reconstructed by DH, in both the off-axis and in-line approaches, is often limited by 2π phase wrapping [6]. This phase wrapping can be subsided by using unwrapping algorithms [7]. However, outstanding performances are hardly achieved due to the influences that are easily caused by many factors. Moreover, such algorithms often fail to measure samples that have a high-aspect-ratio or a rough surface [8].

Compared with single-wavelength DH, recording holograms with dual-wavelengths is another effective method to quantitatively retrieve unwrapped phase information of samples

[9,10]. Dual-wavelength in-line digital holography (DIDH) not only expands the range of the measured optical path difference (OPD) by a synthesized beat wavelength but also achieves high-resolution measurement and fast implementation [10]. Unfortunately, two inherent factors often obfuscate the reconstruction in practical DIDH: (1) the noise signals at each wavelength detection appear in the dual-wavelength hologram at the same time, leading to amplified noise in the phase reconstruction [11], and (2) the well-known twin-image problem that manifests itself as an out-of-focus version of the reconstructed plane should be bounded to in-line DH [5].

To reduce the phase noise and increase the reconstructed accuracy in DH, numerical filtering, a headmost approach, can be easily implemented to remove noises, but the details of the object itself are also often filtered out [12]. Additionally, the phase distribution with suppressed noise can be directly acquired via the linear regression [13], although it is common for the determination of parameters to hardly satisfy the imaging relationship. In contrast, the level of the amplified noise can be reduced to the order of a single wavelength by introducing the guiding phase [11]. In terms of the twin-image problem, existent solutions can be mainly classified into two strategies, i.e., the physical modification in a holographic setup [14] and

the numerical compensation [5,6]. Even though some specific setups have been proven to get rid of the twin-image, the complexity of the DH setup is often unavoidably increased. In contrast, numerical solutions, known as phase retrieval [15], are essentially a class of iterative algorithms that reduce the twin-image and acquire a relatively real phase at each iteration, such as the finite transmission constraint [16] and the Gerchberg–Saxton (GS) algorithm [17]. Moreover, based on Fourier analysis and sparsity, the wave propagation is physically modeled with compressive sensing (CS), which leads to a physics-driven compressive sensing-digital holography (CS-DH) method [18], taking advantage of the significant difference between the twin-image and the existing object. However, most of these traditional frameworks often struggle in the presence of strong noise [19], which becomes more notable when encountering amplified noise in DIDH.

Recently, deep learning (DL) has been successfully utilized for phase retrieval from only one intensity pattern [4], which also converts reconstructions to artifacts-free [20], twin-image-free [21], or noise-free [22] versions. As a powerful machine learning method, DL should be naturally introduced in DIDH. However, most DL-based strategies are data-driven or end-to-end net approaches [22], including derivatives like the regularization by denoising (RED) frame [23], which results in excessive data dependency and limited generalization ability, especially when the reconstructed target is out of the training set [23]. In contrast, an untrained network, as a training-free DL approach, has been investigated to directly reconstruct high-quality image or phase information through self-calibration, typically via a deep image prior (DIP) framework [24,25] or by coupling double DIPs with more decomposed basic components [26]. In these paradigms, a complete physical model that represents the imaging process and the DIP frame can be combined for more practical interpretability [24]. Even though single-wavelength twin-image-free DH has been implemented with DIP, called deep DIH [27], artifacts are caused when measuring phases with relatively strong noises due to the lack of special treatments, thereby hardly transferring to the DIDH directly as the distortions of the phase should be further deteriorated due to the amplified noise. In contrast, concise deep decoder (CDD), as a variant of DIP, provides a much simpler and under parameterized architecture to learn a low-dimensional manifold and a decoding operation of the full image, which can be a relatively more robust and faster convergent than DIP [28]. However, the CDD has hitherto not been incorporated into a complete physical model with practical imaging interpretability.

Enlightened by previous studies, this work demonstrates that it is possible to experimentally recover the phase distribution of a sample with suppressed twin-image and noise from DIDH via untrained neural networks, i.e., DIDH-Net, which is built by combining concise and non-convolutional networks [28] with a real-world imaging model. In terms of DIDH-Net, the incorporation of CDD with a task-specific DIDH model for optical imaging reduces the amount of labeled data required to train the network. Thus, neither additional modification of the setup nor operations (for example, phase shifting, training data) are required, which enables a high-resolution and high-accuracy measurement.

2. METHODS

A. Problem Statement

Suppose that a sample with a certain optical thickness L is illuminated simultaneously by two different wavelengths of λ_1 and λ_2 without color absorption. In most cases, L is larger than λ_1 and λ_2, and thus the corresponding phase diagram Φ_{est}^m ($m = 1, 2$), estimated from the captured hologram I_{cap}^m, can be written as follows:

$$\Phi_{est}^m = \Phi_{pure}^m + 2\pi\varepsilon_m = \varphi_m + 2\pi c_m + 2\pi\varepsilon_m, \quad (1)$$

where Φ_{pure}^m is the noise-free phase, φ_1 and φ_2 are the unwrapped phases at each single-wavelength and belong to $[0, 2\pi]$, c_1 and c_2 are unknown nonnegative integers at a certain point (x, y) of the wrapped phase image, and $2\pi\varepsilon_m$ represents the variance of noise that was introduced by the detection noise into the phase estimation. For the DIDH technique, the optical thickness of sample L can be calculated as follows [6,13]:

$$L = \begin{cases} \Lambda \frac{\Phi_{est}^1(x, y) - \Phi_{est}^2(x, y)}{2\pi}, & \Phi_{est}^1(x, y) - \Phi_{est}^2(x, y) \geq 0, \\ \Lambda \frac{\Phi_{est}^1(x, y) - \Phi_{est}^2(x, y) + 2\pi}{2\pi}, & \Phi_{est}^1(x, y) - \Phi_{est}^2(x, y) < 0, \end{cases}$$

$$(2)$$

where $\Lambda = \lambda_1\lambda_2/|\lambda_1 - \lambda_2|$ is the beat wavelength [6]. Moreover, due to the high wavelength selectivity of the Bayer mosaic filter of the color camera, two single-wavelength holograms, I_{cap}^1 and I_{cap}^2, can be extracted from the single-shot dual-wavelength hologram I_{cap}^{RGB} for subsequent processing without the crosstalk between the two wavelengths [6,29] as follows:

$$\begin{bmatrix} I_{cap}^1 \\ I_{cap}^2 \end{bmatrix} = \begin{bmatrix} a_{1r}^2 + a_{1g}^2 + a_{1b}^2 & a_{1r}a_{2r} + a_{1g}a_{2g} + a_{1b}a_{2b} \\ a_{1r}a_{2r} + a_{1g}a_{2g} + a_{1b}a_{2b} & a_{2r}^2 + a_{2g}^2 + a_{2b}^2 \end{bmatrix}^{-1}$$
$$\times \begin{bmatrix} a_{1r}I_{cap}^{HR} + a_{1g}I_{cap}^{HG} + a_{1b}I_{cap}^{HB} \\ a_{2r}I_{cap}^{HR} + a_{2g}I_{cap}^{HG} + a_{2b}I_{cap}^{HB} \end{bmatrix}, \quad (3)$$

where $(I_{cap}^{HR}, I_{cap}^{HG}, I_{cap}^{HB})$ are the red, green, and blue (RGB) components of the recorded color hologram I_{cap}^{RGB}, and (a_{1r}, a_{1g}, a_{1b}) and (a_{2r}, a_{2g}, a_{2b}) are the RGB components of the single-wavelength hologram at λ_1 and λ_2, respectively, which can be calibrated with the corresponding single-wavelength imaging at the beginning of the experiment.

1. Amplified Noise Obstructs the High-Precision Reconstruction
According to Eqs. (1) and (2), the optical thickness L can also be expressed as follows:

$$L = \begin{cases} \Lambda \frac{\Phi_{pure}^1(x, y) - \Phi_{pure}^2(x, y)}{2\pi} + \Lambda(\varepsilon_1 + \varepsilon_2), & \Phi_{pure}^1(x, y) - \Phi_{pure}^2(x, y) \geq 0, \\ \Lambda \frac{\Phi_{pure}^1(x, y) - \Phi_{pure}^2(x, y) + 2\pi}{2\pi} + \Lambda(\varepsilon_1 + \varepsilon_2), & \Phi_{pure}^1(x, y) - \Phi_{pure}^2(x, y) < 0. \end{cases} \quad (4)$$

In other words, the amplitude of the noise contained in the single-wavelength phase wrapping diagram $\varphi_m(x, y)$ is $2\pi\varepsilon_m$, and the corresponding noise in the optical path change $L_m(x, y)$ is $\lambda_m\varepsilon_m$. In the phase difference distribution $\varphi_1(x, y) - \varphi_2(x, y)$, i.e., the noise level is $2\pi(\varepsilon_1 + \varepsilon_2)$, and the corresponding noise in the finally reconstructed optical thickness L is $\Lambda(\varepsilon_1 + \varepsilon_2)$. Compared with the single wavelength, the noise increases by $\Lambda(\varepsilon_1 + \varepsilon_2)/\lambda_m\varepsilon_m$, which approximately reaches $2\Lambda/\lambda_m$, thereby amplifying the noise in the phase distribution and the reconstructed thickness. A smaller beat wavelength can weaken the problem of amplified noise, but its range in measurement is limited. Thus, the amplified noise can be suppressed under the condition of a large beat wavelength by eliminating the noise at each wavelength.

2. Twin-Image Problem Caused by Interference

To illustrate the twin-image problem of DIDH, U_{obj}^m is used as an object wave in terms of the m-th wavelength, and the non-scattered wave R_{ref}^m is used as the reference wave. Subsequently, the hologram I_{cap}^m exists simultaneously in the interference of the physical symmetry [18] as follows:

$$
\begin{aligned}
I_{cap}^m &= H_{DIDH}^m\{\Phi_{pure}^m\} + E_{cap}^m \\
&= |U_{obj}^m + R_{ref}^m|^2 + E_{cap}^m \\
&= [U_{obj}^m]^* R_{ref}^m + U_{obj}^m[R_{ref}^m]^* + |U_{obj}^m|^2 + |R_{ref}^m|^2 + E_{cap}^m,
\end{aligned}
\tag{5}
$$

where $H_{DIDH}^m\{\cdot\}$ means the forward operator or mapping function, E_{cap}^m is the detection noise, and * represents the conjugate term. Even though the original intention of DH is to record the wavefront U_{obj}^m, the physical conjugate terms $[U_{obj}^m]^* R_{ref}^m$ and $U_{obj}^m[R_{ref}^m]^*$ are recorded together, and the conjugate wave $[U_{obj}^m]^*$ is also recorded as a by-product. As a result, the reconstruction of the virtual image plane is the superposition of the original object and the twin-image. This is usually because the reconstruction of holograms is considered to be a wave rather than a transmission reconstruction problem.

B. Untrained Network-Based DIDH Reconstruction

Since DIDH only depends on intensity measurement, the reconstruction can be regarded as a highly ill-conditioned inverse problem. The DIDH-Net method proposed here only needs one captured intensity hologram, $I_{cap}^{RGB}(z = d)$, which specifically means a diffraction DH pattern formed at the distance of d from the imaging plane located at $z = 0$. As shown in Fig. 1, after extracting holograms $[I_{cap}^1(z = d), I_{cap}^2(z = d)]$ at each wavelength from the single-shot dual-wavelength hologram $I_{cap}^{RGB}(z = d)$ used in Eq. (3), the diffractive DH pattern $I_{cap}^m(z = d)$ is then input to the designed structure for generating the estimated phase of the object $\Phi_{est}(z = 0)$. In a traditional neural network, it is necessary to know the true phase object $\Phi_{pure}(z = 0)$ in the training set and calculate the error between $\Phi_{est}(z = 0)$ and $\Phi_{pure}(z = 0)$ to optimize the weight and deviation. However, the proposed DIDH-Net does not require the true phase $\Phi_{pure}(z = 0)$. Instead, the physical model $H_{DIDH}^{m,z}\{\cdot\}$, a specific mapping process in Eq. (5), is utilized to calculate $\Phi_{est}(z = 0)$ from $I_{est}^m(z = d)$. Then the weight and bias of the net are optimized by the error between $I_{est}^m(z = d)$ and the measured $I_{cap}^m(z = d)$ with gradient descent. In other words, the estimated diffraction hologram $I_{est}^m(z = d)$ is forced to gradually converge to the measurement hologram
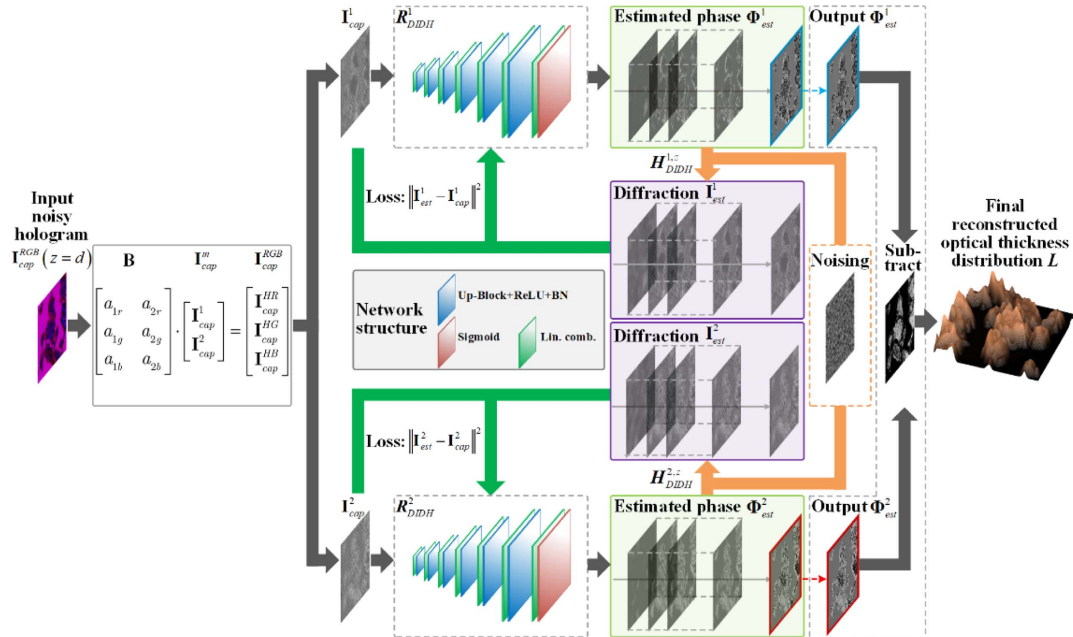

Fig. 1. Schematic of the DIDH-Net imaging system. A captured hologram $I_{cap}^{RGB}(z = d)$ of a phase object is the input to the neural networks after extracting holograms at each wavelength. The output of the neural networks is taken as the estimated phase $\Phi_{est}^m(z = 0)$, which is then numerically propagated to simulate the diffraction and measurement processes $H_{DIDH}^{m,z}\{\cdot\}$ to generate $I_{est}^m(z = d)$. The mean square errors (MSEs) between $I_{cap}^m(z = d)$ and $I_{est}^m(z = d)$ are measured as the loss value to adjust the neural network parameters. The optical thickness distribution L can finally be acquired with the suppressed amplified noises and the free twin-image.

$I_{cap}^m(z = d)$, which continues as an iterative process, as shown in Fig. 1. Finally, the phase search can converge to a feasible solution after the whole iterative process, subsequently reconstructing the optical thickness distribution L.

Specifically, as a phase object $\Phi_{pure}^m(z = 0)$ is illuminated by dual-wavelength coherent plane waves, the diffraction pattern $U_{obj}^m(z = d)$ with a propagation distance $z = d$ can be expressed as follows:

$$U_{obj}^m(z = d) = \iint [U_{obj}^m(z = 0)]^{\mathcal{F}} F_{m,z}$$
$$\times \exp[i2\pi(f_x x + f_y y)]df_x df_y$$
$$= G_{m,z}[\Phi_{pure}^m(z = 0)], \qquad (6)$$

where $F_{m,z} = \exp\left[ikz\sqrt{1 - (\lambda_m f_x)^2 - (\lambda_m f_y)^2}\right]$ is the transfer function, $[U_{obj}^m(z = 0)]^{\mathcal{F}}$ is the 2D Fourier transform of $U_{obj}^m(z = 0) = A\exp[i\Phi_{pure}^m(z = 0)]$, A is the amplitude that can be normalized in the subsequent process, f_x and f_y are the spatial frequencies, and $G_{m,z}[\cdot]$ is a transform operator expressing from $\Phi_{pure}^m(z = 0)$ to $U_{obj}^m(z = d)$. As mentioned before, the sensor not only records $U_{obj}^m(z = d)$ but also records its conjugation, $[U_{obj}^m(z = d)]^*$. According to Eq. (5), note that the term $|R_{ref}^m(z = d)|^2$ is simply a constant, and hence the effect of $|R_{ref}^m(z = d)|^2$ can be removed by eliminating the DC term. Also, the term $|U_{obj}^m(z = d)|^2$ can be regarded as the noise term $E_{cap}^m(z = d)$. In addition, $R_{ref}^m(z = d)$ can be assumed as one without loss of generality [30,31]. Then, the complete mapping from the object phase to the hologram can be expressed as follows:

$$I_{cap}^m(z = d) \approx 2\,\mathrm{Re}\{U_{obj}^m(z = d)\} + E_{cap}^m(z = d)$$
$$= 2\,\mathrm{Re}\{G_{m,z}[\Phi_{pure}^m(z = 0)]\} + E_{cap}^m(z = d)$$
$$= H_{DIDH}^{m,z}\{\Phi_{pure}^m(z = 0)\}. \qquad (7)$$

It should be noted that $H_{DIDH}^{m,z}\{\cdot\}$ not only includes the physical imaging procedure of the generated diffraction pattern but also encapsulates and emphasizes the noising process in the imaging. Then, the typical method for phase imaging, which is a highly ill-posed problem, is to solve the minimization problem as follows:

$$\Phi_{est}^m(z = 0) = \arg\min_{\Phi}\|H_{DIDH}^{m,z}\{\Phi_{pure}^m(z = 0)\} - I_{cap}^m(z = d)\|_2^2$$
$$+ r[\Phi_{pure}^m(z = 0)], \qquad (8)$$

where *a priori* term $r[\cdot]$ is designed artificially or has the characteristics of a dictionary, which can capture the general regularity of objects $\Phi_{est}^m(z = 0)$. The idea of the optimization of Eq. (8) is the core in most of the numerical approaches of the phase retrieval.

In contrast, a typical DL-based approach, i.e., end-to-end net [4,20,21], tries to extract a large number Q of the labeled data $[\Phi_{pure}^m(z = 0), I_{cap}^m(z = d)]$, $q = 1, 2, \ldots,$ and Q for each wavelength $m = 1, 2$, that come from the corresponding training set $ST_m = \{[\Phi_{pure}^m(z = 0), I_{cap}^m(z = d)], m = 1, 2, q = 1, 2, \ldots, Q\}$ to learn the mapping function of the neural network R_{typ}^m, which can be expressed as follows:

$$R_{typ}^m = \arg\min_{\theta \in \Theta}\|R_{typ}^m[I_{cap}^m(z = d)] - \Phi_{pure}^m(z = 0)\|^2,$$
$$\forall\,[\Phi_{pure}^m, I_{cap}^m] \in ST_m. \qquad (9)$$

Here, R_{typ}^m is defined by a set of weights and deviations $\theta \in \Theta$. The training process produces a feasible mapping function R_{typ}^m, which can map the diffraction hologram $I_{cap}^m(z = d)$ that is not in the ST_m back to the corresponding phase, i.e., $\Phi_{est}^m(z = 0) = R_{typ}^m[I_{cap}^m(z = d)]$. In a typical DL application, the size Q of the training set ST_m can be thousands or even larger. It is experimentally time-consuming to collect such a large group of diffraction holograms $I_{cap}^m(z = d)$ and their corresponding original phases $\Phi_{pure}^m(z = 0)$. This usually requires maintaining mechanical and environmental stability in the process of data acquisition for several hours. Although the training set can be created by numerical modeling of the imaging physical process, the mapping function learned in this case is only applicable to those similar test images in the training set. Only the object set with the same priority used in the training process can obtain a good generalization effect [24].

On the contrary, in the proposed DIDH-Net model, the phase recovery formula is as follows:

$$R_{DIDH}^m = \arg\min_{\theta \in \Theta}\|H_{DIDH}^{m,z}\{R_{DIDH}^m[I_{cap}^m(z = d)]\} - I_{cap}^m(z = d)\|^2. \qquad (10)$$

In this objective function, there is no real phase on the ground, which means that the DIDH-Net does not need to be trained in the basic truth stage. The interaction between $H_{DIDH}^{m,z}\{\cdot\}$ and R_{DIDH}^m makes the artificial neural network capture the prior information of $I_{cap}^m(z = d)$. After optimization, the phase can be reconstructed using the obtained mapping function R_{DIDH}^m as follows:

$$\Phi_{est}^m(z = 0) = R_{DIDH}^m[I_{cap}^m(z = d)]. \qquad (11)$$

It should be noted that there is no limit to the network architecture that can choose to implement R_{DIDH}^m. In this study, CDD was used, which only consists of a simple combination of a few building blocks, thereby achieving a relatively outstanding performance. Typically, the network structure includes an input with a diffraction hologram and a decoder path outputting a predicted phase pattern. Specifically, five main modules were used to connect the input and the output [28]: batch normalization (BN), rectified linear unit (ReLU) nonlinearity, up block, sigmoid, and the pixel-wise linear combination of channels, respectively. The neural network is based on the platform of PyTorch 1.8.0, and it was implemented with Python 3.7.6. An Adam optimizer was used with a learning rate of 0.01 to optimize the weight. In our study, the size of the input image $I_{cap}^{RGB}(z = d)$ was 512×512 pixels. The network usually required 7000 periods to find a very good estimation in less than 10 min on a computer with an Intel Xeon CPU e5-2696 V3 processor with 256 GB of RAM and a double NVIDIA Titan V GPUs.

3. RESULTS AND DISCUSSION

A. Evaluation with the Simulated Test Dataset

Simulations were first performed to compare the performance of the proposed DIDH-Net with the backpropagation (BP), CS-DH, end-to-end net, deep DIH, and RED frameworks.

A phase object, 512×512 pixels in size, was used in the transmission, which is shown in Fig. 2(a). The simulated illumination wavelengths were $\lambda_1 = 647$ nm and $\lambda_2 = 485$ nm. The maximum value of L was set to be 1.6 µm, which was a little smaller than the beat wavelength of 1.94 µm but much larger than each illumination wavelength. The distance $z = d$ between the image plane and the image sensor was set to 10 mm to generate a pure noise-free diffraction hologram $I_{\text{pure}}^{\text{RGB}}(z = d)$. The pixel size was set to 5.5 µm. Since the model mismatch and the shot noise are the main sources of noise in many phase retrieval applications, the robustness of several frameworks was evaluated with the noise process [32] as follows:

$$I_{\text{cap}}^{\text{RGB}}(z = d) = I_{\text{pure}}^{\text{RGB}}(z = d) + E_{\text{cap}}^{\text{RGB}}(z = d) = \xi Z + B,$$

(12)

where $Z \sim [I_{\text{pure}}^{\text{RGB}}(z = d)/\xi]$ follows a Poisson distribution, and $B \sim N(0, \sigma^2)$ follows a Gaussian distribution, which is the dominant noise with a noise level of σ^2. The simulated single-shot DIDH hologram $I_{\text{cap}}^{\text{RGB}}(z = d)$ is shown in Fig. 2(b), with $\sigma = 0.30$ and $\xi = 0.02$. Figures 2(c) and 2(d) show the

extracted single-wavelength holograms $I_{\text{cap}}^1(z = d)$ and $I_{\text{cap}}^2(z = d)$ at λ_1 and λ_2, respectively.

The corresponding reconstruction results are shown in Fig. 3. All the methods used only the single-shot diffraction hologram to reconstruct the phase, accompanied with a quantitative assessment through the peak-signal-to-noise-ratio (PSNR) [33] and the structural similarity (SSIM) [34] computed against the ground truth, which are shown in Table 1. Intuitively, the DIDH-Net presented an obvious improvement compared to the backpropagation and the CS-DH results. The end-to-end net (the same neural network structures composed of U-Net [35] without a physical model of DH for each wavelength) with fitting of the training set (400 images were used for training [22]) was used to obtain a training model that mapped the intensity pattern to the phase image. As expected, the PSNR and SSIM of the end-to-end DL method slightly increased to 22.49 dB and 0.77, respectively, since the recovery quality was unpredictably reduced and caused by the similarity error between the test object and the training object. In contrast, directly applying deep DIH for each wavelength relatively enhanced the suppression of the twin-image compared to the
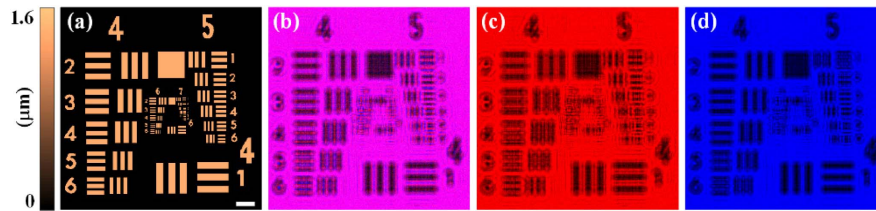

Fig. 2. Simulation results of the numerical phase target for the single-shot DIDH. (a) The simulated optical thickness distribution of the object. (b) The simulated single-shot recorded dual-wavelength in-line hologram calculated at $z = 10$ mm. (c) and (d) are the extracted single-wavelength holograms from (b). The white scale bar measures 200 µm.
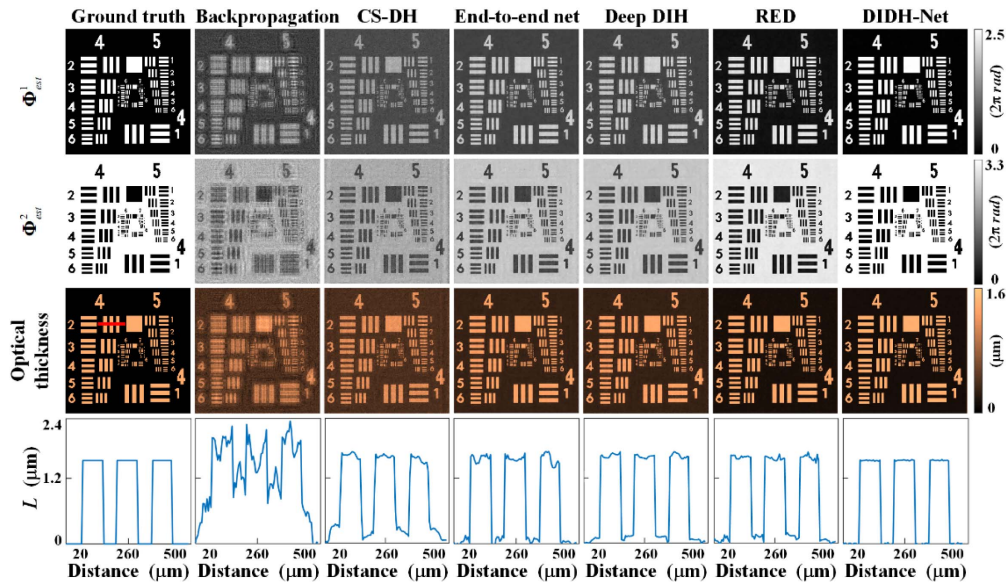

Fig. 3. Comparison of the different phase retrieval methods (from left column to right column): the ground-truth images for intuitive comparison, the phase maps reconstructed by means of direct reconstruction via backpropagation, the CS-DH method, the end-to-end net with the pre-trained network, the deep DIH, the RED frame, and the DIDH-Net. The cross-section optical thickness profiles (along the red line) of each optical thickness map were also measured and are shown in the last row.

Table 1.   Quantitative Results on Imaging with Different Methods

	Backpropagation	CS-DH	End-to-End Net	Deep DIH	RED	DIDH-Net
PSNR/dB	18.59	19.32	22.49	24.32	27.50	31.36
SSIM	0.68	0.73	0.77	0.81	0.85	0.91

CS-DH method, but the exact optical thickness distributions were still stranded by amplified noise. In terms of the RED method, the denoising convolutional neural network (DnCNN) [22] was employed to fit the training dataset to obtain the denoiser for the RED frame. The imaging result is shown in Fig. 3, where the PSNR and SSIM were 27.50 dB and 0.85, respectively. Even though the RED method could improve the imaging performance, small artifacts and blurring still existed. In contrast, the DIDH-Net avoided the predicated errors introduced by traditional training and showed better robustness than other methods. Moreover, the red line is selected in Fig. 3 for optical thickness measurement. For backpropagation, due to the influence of the twin phase and the expanding noise, the optical thickness information was seriously distorted. However, the DIDH-Net had the best suppression effect on the twin phase and expanding noise. Therefore, accurate phase and optical thickness information were obtained. As a matter of fact, the noise and twin image can seriously affect the reconstruction [18], i.e., from the accurately recovered phase information, the sharp optical thickness of the target can be easily obtained [27]. As a result, employing neural network of the DIDH-Net to eliminate the amplified noise and twin image in holography is self-consistent and can become an effective and powerful approach, which is mainly beneficial from the interplay between a handcrafted network structure and a physical image formation model.

Additionally, the effect of the diffraction distance z on the quality of the reconstructed image was also numerically analyzed. Four diffraction distances $z = 1$, 10, 25, and 50 mm were selected as examples to test their properties, where the maximum value of L was set to 1.6 µm. The results are shown in Fig. 4. As shown, in all cases, the DIDH-Net successfully reconstructed the phase from the corresponding diffractive DH pattern. Compared with the true value phase image in Fig. 4(f), the SSIM values of the reconstructed optical thickness distributions from Figs. 4(a)–4(d) were 0.92, 0.91, 0.90, and 0.88, respectively. This observation is also consistent with the convergence of mean square errors (MSEs), which gradually decreased in all the cases as the number of epochs increased in Fig. 4(e). Even though this simulation demonstrated that the DIDH-Net could accurately reconstruct the phase and optical thickness information of samples despite different diffraction distances, the distance between the target and the image sensor should be adjusted by considering the adequacy of the diffraction patterns and the image size acquired by the imaging sensor.
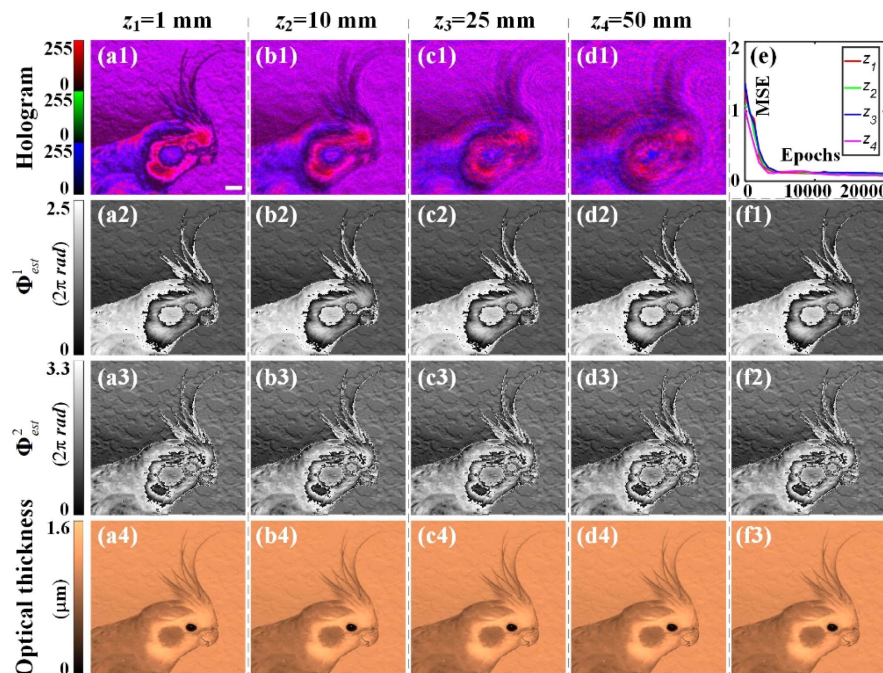

Fig. 4.   Effect of the diffraction distance z on the quality of the reconstructed image. The diffraction holograms [top row (a1)–(d1)] were calculated at z values of 1 mm, 10 mm, 25 mm, and 50 mm, each of which followed their DIDH-Net reconstructed single-wavelength phase images and optical thickness maps in the corresponding rows. The ground truths of images are listed [far right column (f1)–(f3)] under the evolution of the MSE with an increasing number of epochs [top right corner (e)]. The scale bar measures 200 µm.

The captured hologram can be noisy due to the real recording procedure. Consequently, the performance of the DIDH-Net under different noise conditions was investigated. Specifically, the diffraction hologram at $z = 10$ mm with noise levels σ of 0.22, 0.30, and 0.38, respectively, is shown in Fig. 5. The optical thickness maps could be successfully reconstructed from the corresponding holograms by the proposed method, despite the large noises slightly decreasing the imaging quality.

B. Experimental Results of Different Samples

1. Experimental Setup

Experiments were carried out to verify the effectiveness and feasibility of the proposed method in practice, and the experimental setup is shown in Fig. 6. Two semiconductor lasers with wavelengths of $\lambda_1 = 647$ nm and $\lambda_2 = 485$ nm (OBIS, Coherent, Inc., USA) were used as illumination sources. The two laser beams were combined by the dichroic mirror (DM, Thorlabs Inc., USA) before passing through a reversed telescope (L1, $f = 50$ mm and L2, $f = 150$ mm) for beam expansion. After passing through a polarizer (P, Thorlabs Inc.,

USA) to ensure consistent polarization and through the non-polarizing cube beam splitter 1 (NPBS1, Thorlabs Inc., USA), the beams were split into object waves and reference waves, respectively. The collinear object waves illuminated the specimen, and the transmitted waves were collected by the objective ($20 \times$ /NA0.55, Nikon Inc., Japan) that was aligned in a $4f$ configuration with tube lens L4 ($f = 200$ mm) to image the sample onto the image plane. To ensure the equal optical path, the reference waves passed through an identical objective ($20 \times$ /NA0.55, Nikon Inc., Japan) and tube lens L3 ($f = 200$ mm). The thickness of the specimens was smaller than the beat wavelength but larger than each illumination wavelength. The object wave and reference wave interfered after the NPBS2, and the dual-wavelength in-line hologram was recorded by a color CMOS camera (UI-3370CP-C-HQ, 2048×2048 pixels and 5.5 μm pixel size, IDS GmbH, Germany). At first, the CMOS camera was fixed on a moving stage (travel range of 25 mm, displacement accuracy of 0.05 μm, KMTS25E/M, Thorlabs, USA) and placed in the image plane to make sure the target was in focus. Then, the
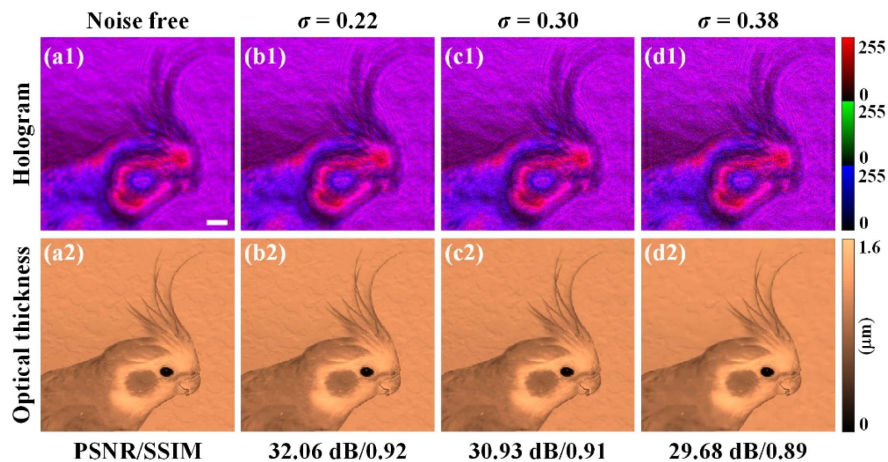

Fig. 5. Reconstructions for the different noise levels: (a1) and (a2) the noise-free hologram at $z = 10$ mm. The DIDH-Net reconstructed optical thickness maps along with the corresponding results with noise levels of (b1) and (b2) $\sigma = 0.22$, (c1) and (c2) $\sigma = 0.30$, and (d1) and (d2) $\sigma = 0.38$. The PSNRs and SSIMs of the optical thickness maps computed against the noise-free ones were also evaluated. The scale bar measures 200 μm.
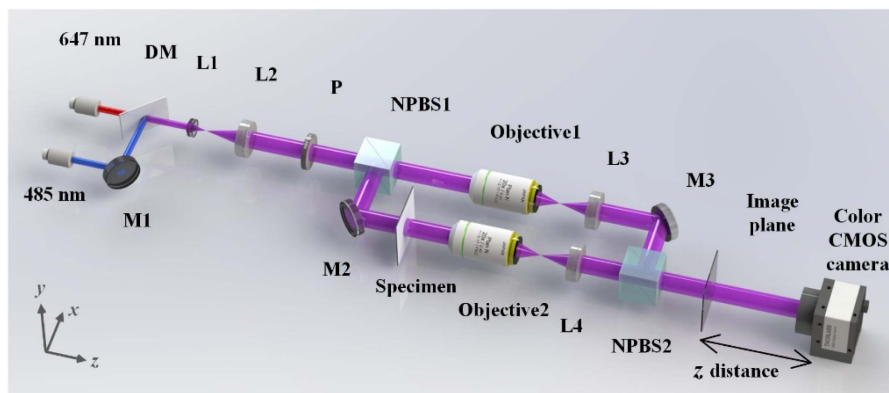

Fig. 6. Schematic of the experimental setup of the DIDH.

CMOS camera was controlled by the moving stage to translate along the z axis. With an accurate and achievable movement, the actual propagating distance could be acquired with high precision. Ultimately, the phase was reconstructed from the diffraction hologram recorded at a distance from the image plane, where the imaging distance z was empirically set to 10 mm by considering the data adequacy and pixel size of the imaging sensor. Additionally, the optical system used the achromatic lenses to eliminate the chromatic aberration.

2. Imaging Results with Different Samples

First, a 180-nm-thick rectangular phase step was selected for the imaging experiment, and the dual-wavelength diffraction hologram was recorded and is shown in Fig. 7(a). The proposed DIDH-Net took the diffraction hologram as its only input and generated the output phase diagram, where the noise level could also be estimated accurately [36]. Compared with the other methods, an envelope curve representing the optical thickness was smoother and in line with the actual situation, which shows the advantages of the proposed method in suppressing the twin-image and the expanding noise again. The average optical thickness, calculated according to the envelope curve, was 180.13 nm, which matched the actual value well. Furthermore, a micro-lens with a spherical top optical thickness of 800 nm was also tested and is shown in Fig. 7(b). Like the phase step results, both the phase image and the corresponding optical thickness curve of the DIDH-Net showed the best reconstruction. Among these methods, the robustness of the CS method in practical application was not good, which also verifies that the convergence of the algorithm was greatly affected in the presence of strong interference. Moreover, the measured results were almost in agreement with the nominal values, which convincingly proves that the proposed method could directly reconstruct the quantitative optical thickness distribution of the specimen from one single-shot dual-wavelength in-line hologram at a high accuracy.

Furthermore, imaging experiments on biological specimens (Ascaris eggs and a water flea jumping foot) were also performed, where the corresponding three-dimensional optical thicknesses based on the reconstructed phase information were simultaneously calculated. Specifically, the plane wave was guided to illuminate the samples, which also produced intensity images of a bright field, as shown in Figs. 8(a1) and 8(b1). Only

the images illuminated by a single wavelength are shown for better understanding and comparison with the phase results. To acquire the diffraction hologram, the camera was placed at a distance $z = 10$ mm from the image plane. After that, the phase results were reconstructed using different methods, which are shown in Figs. 8(a3)–8(a6) and 8(b3)–8(b6). Moreover, as the ground truth was unavailable, a no-reference perceptual blur metric (NPBM [37]), spanning from 0 to 1 (lower was better) was introduced to evaluate the results. Similarly, due to the influence of the twin-images, there were many artifacts in the backpropagation method, and the averaged NPBM was 0.58. By using the CS regularization constraint, the twin-images and the expanded noise were suppressed to a certain extent, in which the averaged NPBM reached 0.41, but the signal-to-noise ratio was still low. In contrast, the traditional RED method could obtain relatively satisfactory results in terms of the reconstructed phase and optical thickness information, but a small amount of information was lost after optical thickness information reconstruction, which is reflected in the averaged NPBM of 0.34. The DIDH-Net method was more accurate in the presence of the twin-image and the amplified noise, and it showed good robustness and the best-averaged NPBM of 0.30. It should be noted that all experiments removed the additional phase brought by the transparent substrate or coverslips. Specifically, a hologram of the specimen-free area, which includes the substrate or the substrate and the coverslips, was recorded at the beginning of each experiment to retrieve a specimen-free phase map that was subtracted from the reconstructed phase distributions with the specimen, i.e., the phase subtraction with double exposure [38] was implemented. In addition, the DIDH-Net requires precise modeling of the image formation mechanism, i.e., $H_{\mathrm{DIDH}}^{m,z}\{\cdot\}$ in our study, which means that the more accurate the measurements, the better the reconstruction results. In order to overcome the effects of the slightly imperfect measurements to achieve the accurate reconstruction, such as phase aberrations caused by nonideal beam collimation and propagating distance measurement, double-exposure [38], mentioned above, and auto-focusing [39] could be respectively utilized, thereby bringing some benefits in imaging, whereas this is not the focus of current study.

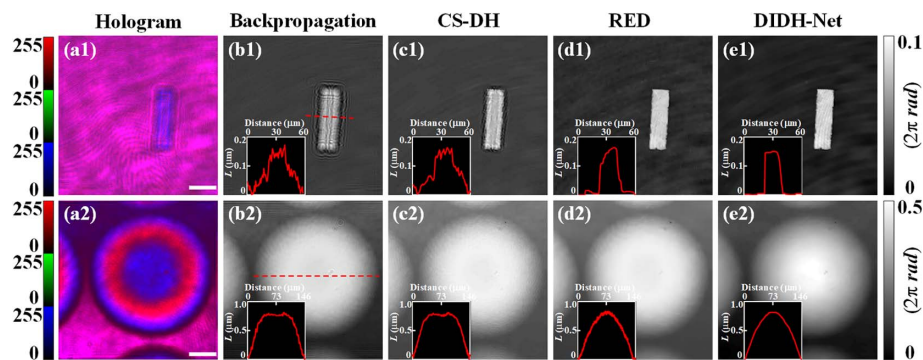Compared with the other numerous approaches (based on GS, CS, and so on) that often make a trade-off between

Fig. 7. Experimental images of the rectangular phase-step [top row (a1)–(e1)] and micro-lens [second row (a2)–(e2)] processed with the backpropagation, the CS, the RED, and the DIDH-Net methods, respectively. The cross-section optical thickness profiles (along the dashed line) were also measured in insets. The scale bars measure 30 μm.
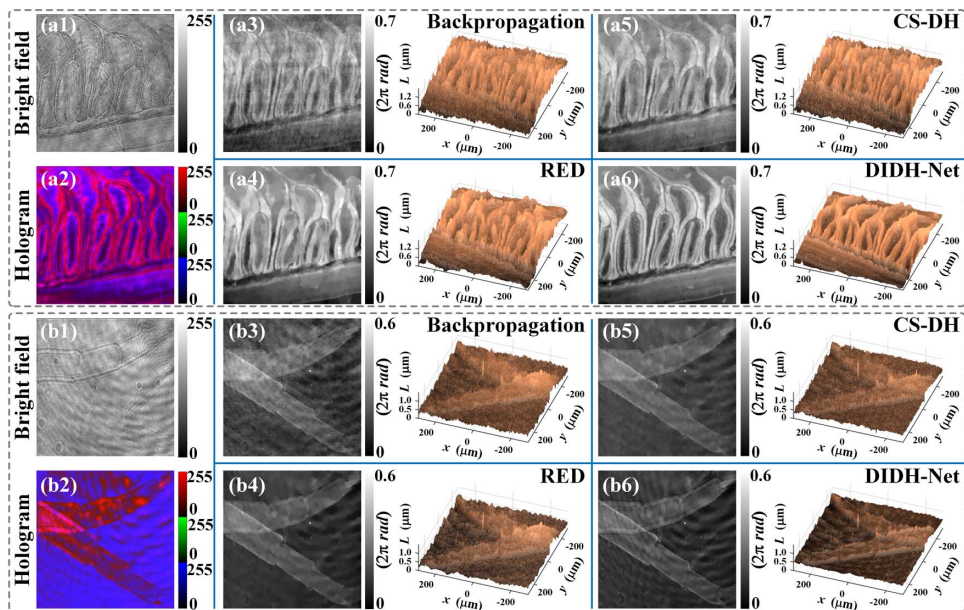
Fig. 8. Imaging results of (a) Ascaris eggs and (b) water flea jumping foot by different methods, including the final reconstructed phase maps and their corresponding optical thickness maps.

accurate phase reconstructions and robustness, the DL-based methods have several advantages in the phase measure of holography [21,40]. However, the traditional end-to-end approaches often learn the mapping function from a set of training data. In fact, when the test data are not fitted with the same set of weights, error is inevitable in the data-driven methods, leading to artifacts and noises in the reconstructed phase. This is more serious under the condition of the amplified noise and the twin-image. In contrast, without any labeled data for training, the DIDH-Net requires relatively accurate modeling of the image forming mechanism. The incorporation of the generated physical model into the traditional deep neural network makes it effective and accurate for reconstructing the phase map of an object with a single hologram of DIDH. However, the calculation errors were enlarged in this current study when the depth of the target exceeded the beat wavelength, which will be addressed in future research.

4. CONCLUSION

In summary, a DL-based technique for overcoming the amplified noise and twin-image problems in DIDH was proposed and verified. In this DIDH-Net, a complete physical model representing the DIDH imaging process is added to the untrained deep neural network to avoid the pre-training of the network and to eliminate the requirement for large amounts of labeled data. By using the interaction of the network and the physical model, the physics-generalization-enhanced method can automatically optimize the network and effectively suppress the amplified noise and the twin-image of DIDH simultaneously, without any additional requirements for data acquisition or illumination conditions. Both simulations and experiments proved the advantages of the method in both accuracy and robustness. Therefore, the proposed DIDH-Net method offers a high-accuracy optical thickness measurement

and a robust phase reconstruction for DIDH. This method can also be extended to other schemes of digital holographic imaging.

Disclosures. The authors declare no conflicts of interest.

Data Availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

[†]These authors contributed equally to this paper.

REFERENCES

1. D. Gabor, "A new microscopic principle," Nature **161**, 777–778 (1948).
2. G. Popescu, T. Ikeda, K. Goda, C. A. Best-Popescu, M. Laposata, S. Manley, R. R. Dasari, K. Badizadegan, and M. S. Feld, "Optical measurement of cell membrane tension," Phys. Rev. Lett. **97**, 218101 (2006).
3. M. Kim, "Digital holographic microscopy," in *Digital Holographic Microscopy: Principles, Techniques, and Application*, Springer Series in Optical Sciences, Vol. **162** (Springer, 2011).
4. 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).
5. J. L. Almeida, E. Comunello, A. Sobieranski, A. M. da R. Fernandes, and G. S. Cardoso, "Twin-image suppression in digital in-line holography based on wave-front filtering," Pattern Anal. Appl. **24**, 907–914 (2021).
6. J. Min, M. Zhou, X. Yuan, K. Wen, X. Yu, T. Peng, and B. Yao, "Optical thickness measurement with single-shot dual-wavelength in-line digital holography," Opt. Lett. **43**, 4469–4472 (2018).

7. J. Nadeau, Y. Park, and G. Popescu, "Methods in quantitative phase imaging in life science," Methods **136**, 1–3 (2018).
8. S. Y. Tong, H. Li, and H. Huang, "Energy extension in three-dimensional atomic imaging by electron emission holography," Phys. Rev. Lett. **67**, 3102–3105 (1991).
9. M. Shan, L. Liu, Z. Zhong, B. Liu, G. Luan, and Y. Zhang, "Single-shot dual-wavelength off-axis quasi-common-path digital holography using polarization-multiplexing," Opt. Express **25**, 26253–26261 (2017).
10. Y. Lee, Y. Ito, T. Tahara, J. Inoue, P. Xia, Y. Awatsuji, K. Nishio, S. Ura, and O. Matoba, "Single-shot dual-wavelength phase unwrapping in parallel phase-shifting digital holography," Opt. Lett. **39**, 2374–2377 (2014).
11. J. Gass, A. Dakoff, and M. K. Kim, "Phase imaging without 2π ambiguity by multiwavelength digital holography," Opt. Lett. **28**, 1141–1143 (2003).
12. D. G. Abdelsalam, R. Magnusson, and D. Kim, "Single-shot dual-wavelength digital holography based on polarizing separation," Appl. Opt. **50**, 3360–3368 (2011).
13. A. Khmaladze, R. L. Matz, C. Zhang, T. Wang, M. M. B. Holl, and Z. Chen, "Dual-wavelength linear regression phase unwrapping in three-dimensional microscopic images of cancer cells," Opt. Lett. **36**, 912–914 (2011).
14. D. G. Abdelsalam and D. Kim, "Two-wavelength in-line phase-shifting interferometry based on polarizing separation for accurate surface profiling," Appl. Opt. **50**, 6153–6161 (2011).
15. Y. Shechtman, Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao, and M. Segev, "Phase retrieval with application to optical imaging: a contemporary overview," IEEE Signal Process. Mag. **32**, 87–109 (2015).
16. T. Latychevskaia and H.-W. Fink, "Solution to the twin image problem in holography," Phys. Rev. Lett. **98**, 233901 (2007).
17. S. M. F. Raupach, "Cascaded adaptive-mask algorithm for twin-image removal and its application to digital holograms of ice crystals," Appl. Opt. **48**, 287–301 (2009).
18. W. Zhang, L. Cao, D. J. Brady, H. Zhang, J. Cang, H. Zhang, and G. Jin, "Twin-image-free holography: a compressive sensing approach," Phys. Rev. Lett. **121**, 93902–93907 (2018).
19. C. Bai, M. Zhou, J. Min, S. Dang, X. Yu, P. Zhang, T. Peng, and B. Yao, "Robust contrast-transfer-function phase retrieval via flexible deep learning networks," Opt. Lett. **44**, 5141–5144 (2019).
20. X. Zhang, Y. Chen, K. Ning, C. Zhou, Y. Han, H. Gong, and J. Yuan, "Deep learning optical-sectioning method," Opt. Express **26**, 30762–30772 (2018).
21. Y. Rivenson, Y. Zhang, H. Günaydın, D. Teng, and A. Ozcan, "Phase recovery and holographic image reconstruction using deep learning in neural networks," Light Sci. Appl. **7**, 17141 (2018).
22. K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising," IEEE Trans. Image Process. **26**, 3142–3155 (2017).
23. Y. Romano, M. Elad, and P. Milanfar, "The little engine that could: regularization by denoising (RED)," SIAM J. Imaging Sci. **10**, 1804–1844 (2017).
24. F. Wang, Y. Bian, H. Wang, M. Lyu, G. Pedrini, W. Osten, G. Barbastathis, and G. Situ, "Phase imaging with an untrained neural network," Light Sci. Appl. **9**, 77 (2020).
25. D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep image prior," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2018), pp. 9446–9454.
26. Y. Gandelsman, A. Shocher, and M. Irani, "'Double-DIP': unsupervised image decomposition via coupled deep-image-priors," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2019). pp. 11026–11035.
27. H. Li, X. Chen, Z. Chi, C. Mann, and A. Razi, "Deep DIH: single-shot digital in-line holography reconstruction by deep learning," IEEE Access **8**, 202648–202659 (2020).
28. R. Heckel and P. Hand, "Deep decoder: concise image representations from untrained non-convolutional networks," arXiv: 1810.03982v2 (2019).
29. J. Min, B. Yao, P. Gao, R. Guo, B. Ma, J. Zheng, M. Lei, S. Yan, D. Dan, and T. Duan, "Dual-wavelength slightly off-axis digital holographic microscopy," Appl. Opt. **51**, 191–196 (2012).
30. D. J. Brady, K. Choi, D. L. Marks, R. Horisaki, and S. Lim, "Compressive holography," Opt. Express **17**, 13040–13049 (2009).
31. H. Zhang, L. Cao, H. Zhang, W. Zhang, G. Jin, and D. J. Brady, "Efficient block-wise algorithm for compressive holography," Opt. Express **25**, 24991–25003 (2017).
32. C. Bai, C. Liu, H. Jia, T. Peng, J. Min, M. Lei, X. Yu, and B. Yao, "Compressed blind deconvolution and denoising for complementary beam subtraction light-sheet fluorescence microscopy," IEEE Trans. Biomed. Eng. **66**, 2979–2989 (2019).
33. Q. Huynh-Thu and M. Ghanbari, "Scope of validity of PSNR in image/video quality assessment," Electron. Lett. **44**, 800–802 (2008).
34. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE Trans. Image Process. **13**, 600–612 (2004).
35. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," in *Conference on Medical Image Computing and Computer-Assisted Intervention* (2015). pp. 234–241.
36. X. Liu, M. Tanaka, and M. Okutomi, "Single-image noise level estimation for blind denoising," IEEE Trans. Image Process. **22**, 5226–5237 (2013).
37. F. Crete and N. Nicolas, "The blur effect: perception and estimation with a new no-reference perceptual blur metric," Proc. SPIE **6492**, 64920I (2007).
38. J. Min, B. Yao, V. Trendafilova, S. Ketelhut, L. Kastl, B. Greve, and B. Kemper, "Quantitative phase imaging of cells in a flow cytometry arrangement utilizing Michelson interferometer-based off-axis digital holographic microscopy," J. Biophoton. **12**, e201900085 (2019).
39. Y. Yao, B. Abidi, N. Doggaz, and M. Abidi, "Evaluation of sharpness measures and search algorithms for the auto focusing of high-magnification images," Proc. SPIE **6246**, 62460G (2006).
40. J. Lim, A. B. Ayoub, D. Psaltis, and M. Abidi, "Three-dimensional tomography of red blood cells using deep learning," Adv. Photonics **2**, 026001 (2020).