Deep learning assisted variational Hilbert quantitative phase imaging

Zhuoshi Li, Jiasong Sun, Yao Fan, Yanbo Jin, Qian Shen, Maciej Trusiak, Maria Cywińska, Peng Gao, Qian Chen and Chao Zuo

---

Related articles

Deep-learning-based ciphertext-only attack on optical double random phase encryption
Meihua Liao, Shanshan Zheng, Shuixin Pan, Dajiang Lu, Wenqi He, Guohai Situ, Xiang Peng
Opto-Electronic Advances   2021 **4**, 200016   doi: 10.29026/oea.2021.200016

Deep-learning-enabled dual-frequency composite fringe projection profilometry for single-shot absolute 3D shape measurement
Yixuan Li, Jiaming Qian, Shijie Feng, Qian Chen, Chao Zuo
Opto-Electronic Advances   2022 **5**, 210021   doi: 10.29026/oea.2022.210021

Benchmarking deep learning-based models on nanophotonic inverse design problems
Taigao Ma, Mustafa Tobah, Haozhu Wang, L. Jay Guo
Opto-Electronic Science   2022 **1**, 210012   doi: 10.29026/oes.2022.210012

Single-shot mid-infrared incoherent holography using Lucy-Richardson-Rosen algorithm
Vijayakumar Anand, Molong Han, Jovan Maksimovic, Soon Hock Ng, Tomas Katkus, Annaleise Klein, Keith Bambery, Mark J. Tobin, Jitraporn Vongsvivut, Saulius Juodkazis
Opto-Electronic Science   2022 **1**, 210006   doi: 10.29026/oes.2022.210006

More related article in Opto-Electron Journals Group website

OES | Opto-Electronic Science

OE_Journal   Website

Deep learning assisted variational Hilbert quantitative phase imaging

Zhuoshi Li[1,2,3], Jiasong Sun[1,2,3], Yao Fan[1,2,3], Yanbo Jin[1,2,3], Qian Shen[1,2,3], Maciej Trusiak[4], Maria Cywińska[4], Peng Gao[5]*, Qian Chen[3]* and Chao Zuo[1,2,3]*

We propose a high-accuracy artifacts-free single-frame digital holographic phase demodulation scheme for relatively low-carrier frequency holograms — deep learning assisted variational Hilbert quantitative phase imaging (DL-VHQPI). The method, incorporating a conventional deep neural network into a complete physical model utilizing the idea of residual compensation, reliably and robustly recovers the quantitative phase information of the test objects. It can significantly alleviate spectrum-overlapping-caused phase artifacts under the slightly off-axis digital holographic system. Compared to the conventional end-to-end networks (without a physical model), the proposed method can reduce the dataset size dramatically while maintaining the imaging quality and model generalization. The DL-VHQPI is quantitatively studied by numerical simulation. The live-cell experiment is designed to demonstrate the method's practicality in biological research. The proposed idea of the deep learning-assisted physical model might be extended to diverse computational imaging techniques.

Keywords: quantitative phase imaging; digital holography; deep learning; high-throughput imaging

Introduction

Quantitative phase imaging (QPI), as a powerful label-free imaging technique, enables dynamic 2D and 3D non-destructive imaging of completely transparent structures[1−3]. It uses the refractive index as an endogenous contrast agent to generate subcellular-specific quantitative maps of analyzed live bio-structure[4,5]. QPI solutions based on digital holographic microscopy (DHM) encode a complex wavefront information into intensity modulations by the interference of a scattered sample wave and a reference wave[6−9]. And it can robustly perform the quantitative analysis of wave-matter interactions by decoding phase delay from a hologram. DHM has emerged as a valuable means in the biomedical fields, such as measurements for stain-free biological cells[3,10], optical metrology of nanostructures[11−14], and drug release monitoring *in vitro*[15].

Regarding the phase demodulation strategy employed,

[1]Smart Computational Imaging Laboratory (SCILab), School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China; [2]Smart Computational Imaging Research Institute (SCIRI) of Nanjing University of Science and Technology, Nanjing 210094, China; [3]Jiangsu Key Laboratory of Spectral Imaging and Intelligent Sense, Nanjing 210094, China; [4]Institute of Micromechanics and Photonics, Warsaw University of Technology, 8 Sw. A. Boboli St., Warsaw 02-525, Poland; [5]School of Physics, Xidian University, Xi'an 710126, China.

*Correspondence: P Gao, E-mail: peng.gao@xidian.edu.cn; Q Chen, E-mail: chenqian@njust.edu.cn; C Zuo, E-mail: zuochao@njust.edu.cn

there are two main configurations for holographic wavefront acquisition in DHM, *i.e.*, in-line and off-axis digital holography (DH). In-line DH records complete wavefront information by the interference of the object light and the reference light on the same optical axis, which can realize full detector-bandwidth phase reconstruction. However, due to the superimposed twin image, the phase retrieval results of samples are severely impacted by imaging artifacts. It always needs to be processed via iterative phase retrieval[16,17] or noniterative phase-shifting methods[18–20], which dramatically sacrifices the temporal resolution. Therefore, it is difficult for the in-line DH, which is vulnerable to external disturbance and vibration, to be applied to dynamic measurement. Alternatively, off-axis DH implements twin-image separation by introducing a slight angle between the object beam and reference beam and recovers the complex wavefront of the sample from the single-frame off-axis hologram. Whereas, for achieving the separation of autocorrelation and cross-correlation terms in the spatial frequency domain (SFD), the off-axis DH needs to provide a sufficiently high carrier frequency at the expense of the space-bandwidth product (SBP) of the imaging system[21]. The slightly off-axis DH regime, as a single-frame high-SBP DH imaging solutions, is therefore proposed[22–24]. It optimizes SBP through full spectral separation of conjugated object lobes while leaving the autocorrelation term partially overlapped with information-carrying cross-correlation terms. Under this configuration, the inevitable spectrum overlapping causes phase artifacts, which greatly degrades the imaging quality and impairs the practicality of the slightly off-axis DH configuration.

High-accuracy artifacts-free phase recovery from the low-carrier frequency holograms is the key to slightly off-axis DH application. This process is presently implemented by suppressing autocorrelation term iteratively[25], utilizing dual-frame decoding scheme[26,27], employing second wavelength assistance[28] and performing the 1D limited processing[29,30]. With inspiration from the theory of "*cepstrum*" and homomorphic filtering[31], a slightly off-axis DH demodulation scheme based on the Kramers-Kronig (KK) relations is proposed, which utilizes the half-space bandwidth of the sensor to achieve high-SBP imaging[32,33]. Although it is able to increase the SBP of full complex field recovery significantly, it inevitably requires intensity restrictions on the object and reference beams and the separation of the cross-correlation terms of the interferogram in the extended SFD. Noteworthily,

an exquisite low-carrier frequency fringe demodulation approach has been presented recently, namely variational Hilbert quantitative phase imaging (VHQPI)[34]. The VHQPI, as an end-to-end pure numerical add-on module, deploys the merger of tailored variational image decomposition[35] and enhanced Hilbert spiral transform[36] to achieve quantitative phase recovery. It adaptively alleviates the overlapped-spectrum problem and robustly demodulates high-quality phase information, performing excellent practicality in biological applications.

Although VHQPI has demonstrated excellent low-carrier frequency fringe demodulation capability, the algorithm-inherent limitations (*e.g.*, parameter robustness and iterative stability) still cause non-sufficient image frequency component extraction, resulting in imaging artifacts in the phase reconstruction results. Deep learning (DL), as a subfield of machine learning, has currently gained extensive attention in the field of optical metrology and demonstrated great potential in solving optical metrology tasks[37–46]. When sufficient training data is collected in an environment that reproduces real experimental conditions, the trained model may have advantages over physics-model-based approaches on some issues (e.g., computing speed, parameter adaptivity, algorithm complexity)[37]. Specifically, in terms of a series of ill-posed inverse phase retrieval problems, the traditional physical model tends to exhibit higher physics complexity and time consumption. Driven by a large dataset, the deep neural network (DNN) can directly and efficiently reconstruct the phase and amplitude images of the objects from the captured holograms[47–49]. Nevertheless, in DL-based phase recovery tasks, it is pretty tricky and laborious to capture massive datasets and generate the corresponding ground truth, especially when applied to bio-samples. Deep image prior (DIP) applies an untrained network to the solution of several inverse problems without a massive training dataset and ground truth, which can fit a randomly initialized DNN to a single corrupted image[50]. Inspired by the DIP, an untrained network model named "PhysenNet" is proposed, which incorporates a complete physical model into the conventional DNN to achieve phase retrieval from a single intensity image[51].

Inspired by the successful application of the interplay between DNN and the physical model, in this work, we propose a DL-assisted variational Hilbert quantitative phase imaging approach (DL-VHQPI). Unlike the massive-data-driven DL training model, DL-VHQPI,

which utilizes DNN to compensate and optimize the possible solutions of the physics-driven model, can achieve high-precision artifacts-free phase recovery using only a small fraction datasets. Specifically, VHQPI, as the underlying physical model, can complete the preliminary extraction of the background components of the fringes to provide a physical prior for the deep learning model. The DNN compensates for the image frequencies that cannot be extracted by the physical model using the idea of residual compensation. Due to the physical model reducing the information entropy of the dataset, the DL-VHQPI performs higher reconstruction accuracy utilizing less than one-tenth of the dataset of the conventional end-to-end model (without the physical model). The simulation experiments quantitatively demonstrate that the proposed method can achieve high-accuracy artifacts-free quantitative phase imaging from single-frame low-carrier frequency holograms. And the results of live-cell experiments demonstrate the practicality of the method in biological research.

Principle of VHQPI

The VHQPI, as the physical model of the DL-VHQPI, adaptively and effectively completes the low-carrier frequency fringe demodulation employing the unsupervised variational image decomposition (uVID) and enhanced Hilbert spiral transform (HST). This section will focus on describing the process details and physical limitations of this method. In the DH wavefront recording, the interferogram containing the required object information is constructed upon the coherent superimposition of the object and reference beams. The intensity distribution of the recorded hologram can be expressed as:

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos(\theta) + n = a + b\cos(\theta) + n . \quad (1)$$

It consists of a sum of three fundamental intensity components: background (a, incoherent sum of intensities I_1 and I_2 of interfering beams), high-frequency noise (n), and coherent interference fringes term comprised by a cosine function modulated in phase (θ) and amplitude (b, $2\sqrt{I_1 I_2}$). Acquiring the accurate fringes term from the three components is the prerequisite of high-precision artifacts-free phase recovery. The uVID approach achieves image frequency components extraction, which is based on the notion of the classical variational image decomposition to separate the information components of the image with two steps in terms of methodology[52,53]: 1) A block-matching 3D (BM3D) algorithm is employed

to remove noise with remarkable efficiency[54]; 2) Background-fringes differentiation is performed using modified Chambolle projection algorithm with an automatic stopping criterion to set the number of projections, and there is no need to pre-set any parameter values[35]. The based-on uVID image frequency components extraction process is shown in Step 1 of Fig. 1. Although the uVID provides a robust and automatic one-stop-shop solution for single-frame fringe pattern analysis, there are physical limitations in the process of frequency component extraction, i.e., iterative instability and parameter robustness, which directly cause non-sufficient background term removal and then impair phase recovery accuracy and artifacts-suppression effect[52].

To recover the phase information of the object, the uVID-filtered noise-free zero-mean-valued interferogram is then analyzed using the HST algorithm[36], as shown in Step 2 of Fig. 1. The HST is the two-dimensional variant of the Hilbert transform (HT), in which the complex analytic signal can be constructed, whereas several requirements must be fulfilled. First, the processed interferogram must be of zero mean value, which is satisfied based on background term removal using the uVID approach. And the amplitude term (b in Eq. (1)) has to be a slowly varying function. This is the so-called Bedrosian theorem which can be applied to general pure-phase objects at relatively low carrier frequencies[55]. The complex analytic signal constructed by HST can be expressed as

$$AFP = 2\sqrt{I_1 I_2}\cos(\theta) - i\exp(-i\beta)$$
$$\cdot \mathcal{F}^{-1}\{SPF * \mathcal{F}[2\sqrt{I_1 I_2}\cos(\theta)]\} , \quad (2)$$

where, AFP denotes the analytic fringe pattern and SPF is the spiral phase function; $\mathcal{F}$ and $\mathcal{F}^{-1}$ denote Fourier transform (FT) and inverse Fourier transform (IFT) operator respectively. It is important to emphasize that carrier-free single-shot interferogram analysis is a fully 2D phase demodulation problem, whereas carrier-based FT phase demodulation is a 1D simplification of the HT analytic relation. The HST, therefore, requires the local fringe direction map (β, modulo 2π)[56]. The modulus value and angle of the 2D complex analytic signal constitute the intensity and phase in QPI, respectively. SPF is defined as

$$SPF(u, v) = \frac{u + i \cdot v}{\sqrt{u^2 + v^2}} = \exp[i \cdot \phi(x, y)] , \quad (3)$$

where (u, v) is the coordinate of (x, y) corresponding to the SFD. $\phi(x, y)$ is the polar coordinate phase expression.
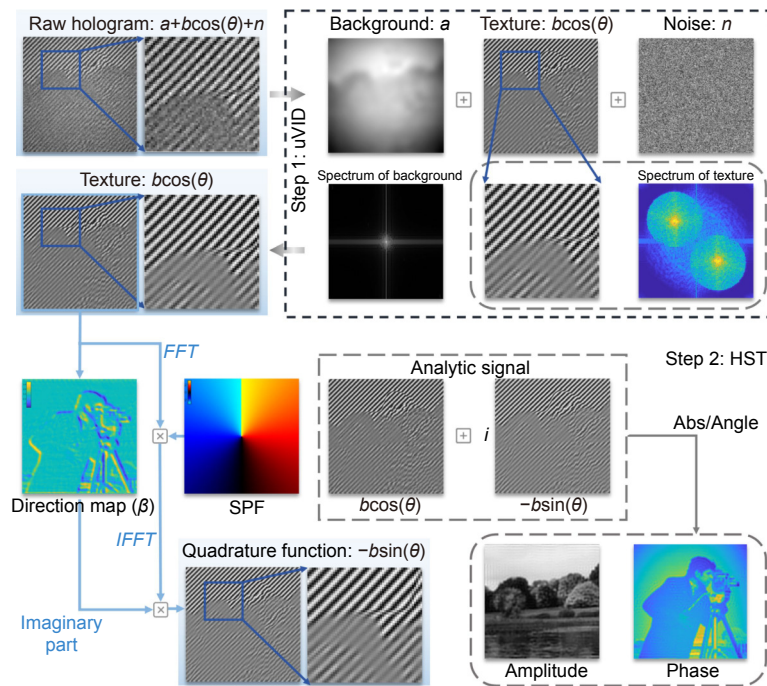
Fig. 1 | Flow chart of slightly off-axis interferometric fringe demodulation based on VHQPI.

Figure 1 specifically showcases the schematic diagram of the VHQPI-based low-carrier interferogram quantitative phase demodulation algorithm.

Deep learning assisted VHQPI model

VHQPI has been proven to have excellent robustness and practicality in low-carrier frequency fringe demodulation issue though[34]. However, the algorithm-inherent iterative instability and parameter robustness restrict the image frequency component extraction capability, which will cause the non-perfect background term removal. DL methods driven by massive datasets provide a new route to address this problem by virtue of their high-powerful image feature extraction characteristics. Whereas, when encountering insufficient training data, which is very common, the DL method based on massive datasets may have a poor effect. A feasible scheme is to train the DNN on a stronger-constrained available standardized dataset[57]. Here, we employ Shannon entropy theory of the images in the dataset for that purpose: the lower the entropy of the datasets is, the more constrained prior information is, giving it a better same-domain generalization ability[58,59]. Therefore, in the proposed DL-assisted VHQPI model, the uVID is utilized to extract the image background term as the physical prior of the network to reduce the dataset's entropy. The first convolutional neural network (CNN1) is used to "learn" the residual terms and assists the physical model to complete the pre-

liminary estimation of the background components of the fringes. Furthermore, to further improve the imaging accuracy, the original hologram and the preliminary estimation background are re-fed into the model (CNN2) for advanced component extraction. Dual-channel input is used because the preliminarily estimated background terms have been very close to the ground truth after the first residual compensation by CNN1. Hence, the preliminary estimated background can be used to provide the network with feature guidance and helps CNN2 achieve the advanced component extraction.

As depicted in Fig. 2, with the original hologram as input, CNN1 completes the preliminary background component extraction by compensating for the residual (ε_1) of the background component acquired by uVID, as shown in Fig. 2(b). With the preliminary estimated background term and the original hologram, the CNN2 (as shown in Fig. 2(c)) uses the two as dual-channel inputs to implement the more advanced background residual (ε_2) compensation. After the high-accuracy fringes terms extraction, the complete complex analytic signal can be constructed by HST. And then the final phase results are recovered by calculating the angle of the 2D complex analytic signal. The whole method flow chart is shown in Fig. 2(a).

Moreover, both CNN1 and CNN2 networks are composed of a convolutional layer (Conv), a group of residual blocks (containing four residual blocks), and two

convolutional layers. Each residual block comprises two sets of Convs stacked one above the other. The network architecture uses Batch Normalization[60] and ReLU activation[61] to accelerate the model convergence. It establishes a shortcut between input and output, which can solve the problem of accuracy decline as the network deepens, thereby easing the training process. The output of the Convs is a 3-D tensor of shape (H, W, C), where H and W are the height and width of pixels of the hologram respectively, and C is the number of channels. The hyperparameters of the two networks, *i.e.*, the weights, bias, and convolutional kernels, are trained using backpropagation on mean-squared errors between the results of the network output and the ground truth. The loss function is computed as

$$Loss(\omega) = \frac{1}{H \times W} \left\| Y_{\text{output}}^{\omega} - G \right\|^2 , \qquad (4)$$

where ω represents the parameter space of the model, Y_{output} is the results predicted by the model, and G is the ground truth.

Experiments and results

In this section, we demonstrate the performance of the proposed DL-VHQPI method over the conventional physics-driven low-carrier frequency fringe demodulation techniques and pure DL approach without a physical model (DL-noPhy) through numerical simulation and live-cell experiment. A rich set of paired training data is the prerequisite for network generalization during DL training. It is challenging to acquire a reliable ground truth in the real-world DH system due to environment-induced instability and system-inherent speckle noise. Consequently, we simulated low-carrier frequency holograms and the corresponding ground truth for training and quantitative analysis. We separately constructed the complex amplitude distributions of the object and reference light waves, and then the holograms can be constructed by solving the square of the modulus of the sum of the two. The sum of the squares of the modulus values of the two was calculated to obtain the background (ground truth) needed for training. The more specific process can be found in Supplementary information Section 1.

In the live-cell experiment, we used the Digital holographic smart computational light microscope (DH-SCLM) developed by SCILab, and turned it to a slightly off-axis state for hologram acquisition[1]. In the DH-
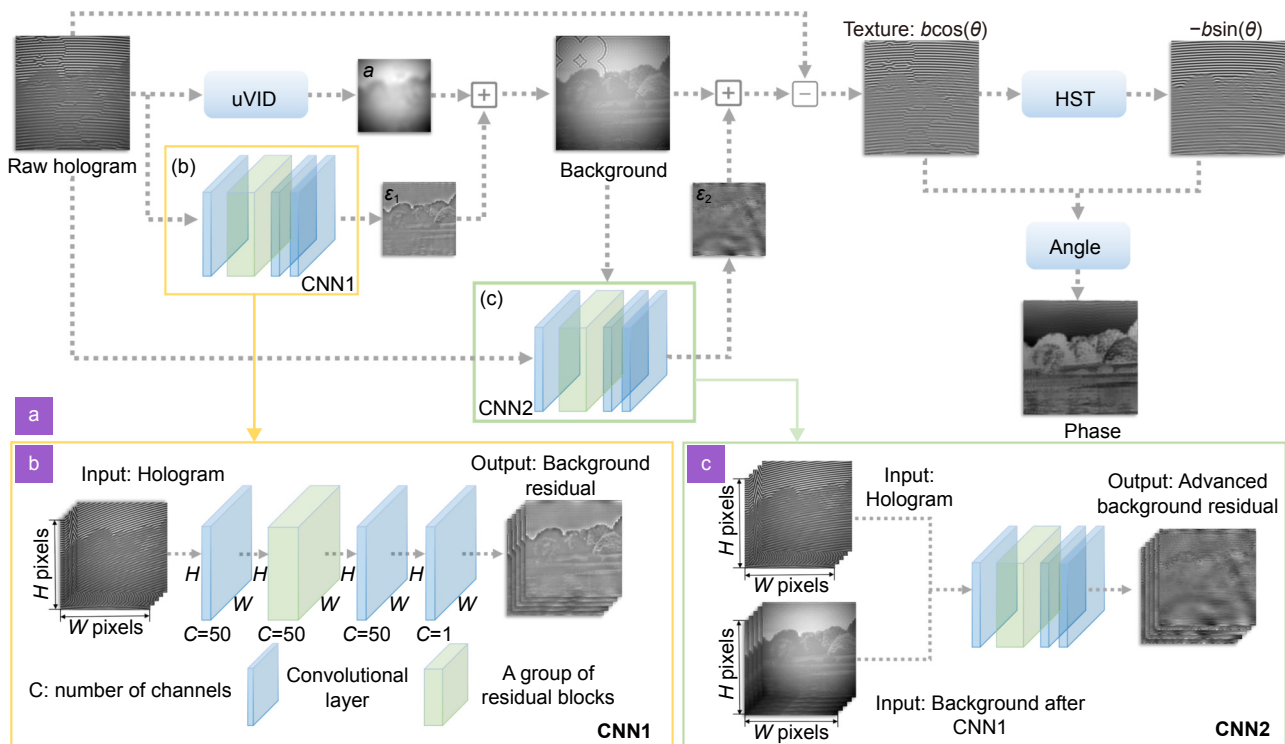

Fig. 2 | Deep learning-assisted VHQPI. (a) Total network structure, combining uVID and HST with CNN respectively for phase reconstruction. **(b)** CNN1 takes a hologram as input and consists of three convolutional layers and a group of residual blocks to achieve compensation of background residuals by learning. **(c)** The CNN2 network structure is the same as CNN1, except that CNN2 combines the original hologram and the result of the first process into a two-channel input for advanced background compensation.

SCLM, the object wave transmitting the objective lens (UPLanSAPO ×20/0.45NA, Olympus, Japan) interferes with the reference light and is recorded by the camera (The Imaging Source DMK 23U274, 1600×1200, 4.4 µm). The central wavelength of the illumination is 532 nm. The used sample is Henrietta Lacks (HeLa) human cervical cancer cells cultured in DMEM medium with 10% fetal bovine serum under standard cell culture conditions (37.2 °C in 5% CO_2 in a humidified incubator). To acquire the ground truth from the configuration, each intensity map of the object and reference light paths needs to be captured separately under the highly stable condition of the holographic system (Refer to Section 2 of the Supplementary Information for detailed processing). The complete training process was implemented using the TensorFlow framework (Google) and was computed on a GTX Titan graphics card (NVIDIA). A fixed learning rate of 0.0001 for the experiment is adopted for the Adam optimizer[62].

Simulation

Figure 3 presents the experimental results under the numerical simulation, demonstrating the quantitative analysis between DL-VHQPI and the conventional single-frame fringe demodulation techniques. Figure 3(a) shows the phase result recovered by the conventional Fourier transform (FT) method. It can be seen that the phase artifacts severely disturb imaging results due to the spectrum-overlapped problem in the SFD. Although reducing the filtering window size can attenuate the phase artifacts, this will sacrifice the SBP of the system while causing blurred imaging. More details about it can be found in Supplementary Section 3. The size of the filtering window used in the FT-based phase reconstruction results shown in Fig. 3(a) is calculated under the simulated numerical aperture (NA), as shown in the red filtering window in Fig. S2(a) of Supplementary Section 3. In VHQPI, the uVID can extract the fringes term from the hologram; however, the non-perfect background term
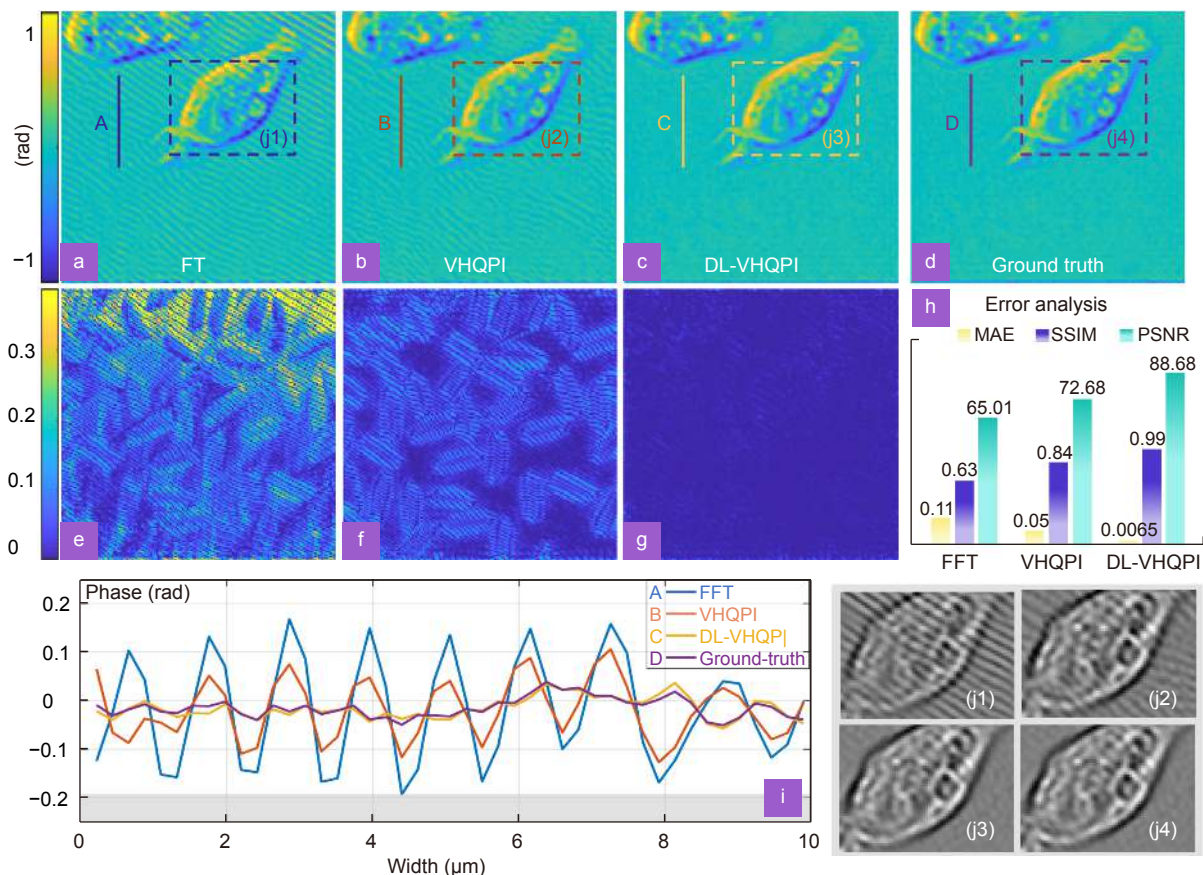

Fig. 3 | The experiment results under the numerical simulation. (**a**) The FT method phase recovery result. (**b**) The phase recovery result of VHQPI. (**c**) The phase result reconstructed by DL-VHQPI. (**d**) The ground truth. (**e–g**) The difference between the phase results of the three methods (*i.e.* FT, VHQPI, DL-VHQPI) and the ground truth. (**h**) Quantitative error analysis of three methods. (**i**) The cross-section of the phase results of FT, VHQPI, DL-VHQPI, and ground truth, and (**j1–j4**) are the DIC views of the partially enlarged views of their corresponding phase maps respectively.

removal still inevitably causes imaging artifacts, as shown in Fig. 3(b). In contrast, as presented in Fig. 3(c), the experimental results demonstrate DL-VHQPI's excellent performance, in terms of artifacts suppression, over two physics-driven methods: FT and VHQPI. Figure 3(d) is the ground truth recovered by phase recovery after theoretical background removal. The magnified views of the corresponding rectangular boxes in Fig. 3(a–d), as shown in Fig. 3(j1–j4), are the phase gradient images by digital differential interference contrast (DIC) for them. To discuss the performance of methods intuitively and quantitatively, we respectively calculated the Mean Absolute Error (MAE), Structural Similarity Index (SSIM), and Peak Signal-to-Noise Ratio (PSNR) between the FT, VHQPI, and DL-VHQPI phase recovery results and the ground truth. Compared with the FT and VHQPI methods, Fig. 3(h) quantitatively demonstrates that DL-VHQPI has an excellent phase recovery accuracy and artifacts-suppression effect (More than 10 times improvement in precision.). The background-part cross-section of the four phase results depicted in Fig. 3(i) shows the phase result reconstructed by DL-VHQPI has a higher similarity to the ground truth, which also demonstrates that it can be more effective in suppressing the fringe-like error of the background part.

In addition, we also designed a comparison experiment with DL-noPhy (The specific network is provided in the Section 4 of Supplementary information) to demonstrate the high-efficiency and high-accuracy characteristics exhibited by the proposed method. Table 1 quantitatively shows the comparison results of the DL-VHQPI and DL-noPhy; DL-VHQPI performs a higher phase reconstruction accuracy while only utilizing one-tenth of the datasets of DL-noPhy. The reason is that DL-VHQPI adopts a physical model (uVID) to the background-component extraction process of the fringe pattern and acquires the residual components for training, which is inherently a process of image entropy reduction. According to the Shannon entropy theory, lower image entropy implies more image constraints, which provides

DNN with a more powerful same-domain generalization ability. The simulated holograms with the size of 160×160 were fed to the network. During the training process, the CNN1 of DL-VHQPI over 150 epochs took 1 hour and 20 minutes, and CNN2 over 150 epochs took 1.5 hours; in contrast, DL-noPhy over 150 epochs took 7 hours and 50 minutes. Fewer training datasets for the same DNN model naturally mean shorter training time, so our method performs higher training efficiency than DL-noPhy while ensuring excellent imaging quality.

Live-cell experiment on HeLa cells

We performed holographic biological experiments on HeLa cells under a ×20/0.45NA lens to demonstrate the application of the method in biological research. The denoised interferogram presented in Fig. 4(a) is of overall low spatial carrier frequency, which results in a spectrum overlapping of cross-correlation and autocorrelation terms, as shown in Fig. 4(b). Figure 4(c) and 4(d) respectively show the phase reconstruction of captured low-carrier frequency holograms utilizing the FT and DL-VHQPI methods for HeLa cells. The field of view (FoV) of Fig. 4(c) and 4(d) is 0.093 mm² (The Imaging Source DMK 23U274, 1600×1200, 4.4 μm), and the SBP of the complex amplitude image is 210000 pixels [the area of the FoV, multiplied by the area of the spatial frequency band, $\pi(NA/\lambda)^2$]. To compare the imaging results of the two methods in detail, we selected two regions of interest (ROI, Area1 and Area2) on the specimens, and their magnified views are shown in Fig. 4(e1, e3, f1, f3). Additionally, Fig. 4(e2, e4, f2, f4) vividly depict the reconstructed phase gradient images by digital DIC. It can be revealed that spectrum-overlapping-caused fringe-like error dramatically degrades the phase recovery quality. The selected regions in the red rectangle box of Fig. 4(c) and 4(d) highlight the artifacts-suppression capability on the phase background. And the enlarged views after DIC processing are shown in Fig. 4(g) and 4(h), respectively. The background part of the FT-based reconstructed phase result features many

Table 1 | The quantitative comparison results of DL-VHQPI and DL-noPhy.

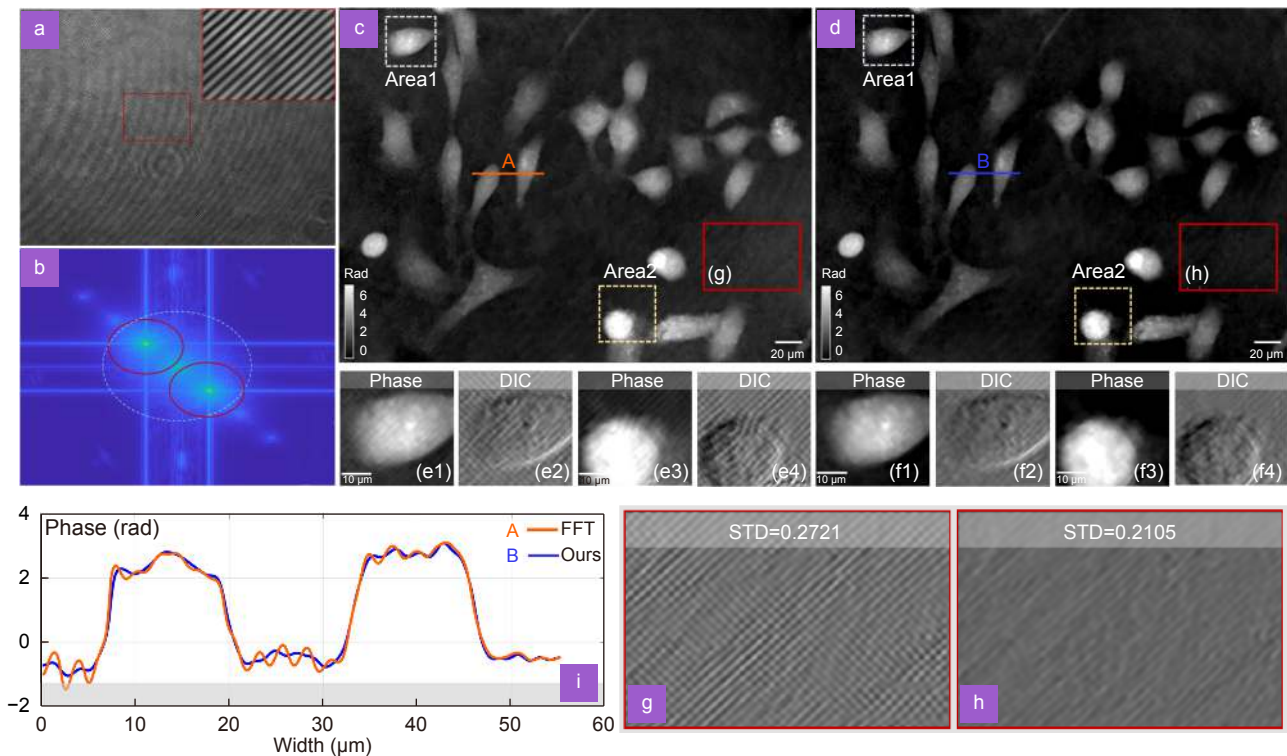| Evaluation Index	Group1		Group2		Group3		Group4	
	DL-noPhy	DL-VHQPI	DL-noPhy	DL-VHQPI	DL-noPhy	DL-VHQPI	DL-noPhy	DL-VHQPI
MAE	0.0105	0.0065	0.0172	0.0142	0.0202	0.0171	0.0191	0.0170
SSIM	0.9914	0.9969	0.9712	0.9781	0.9637	0.9718	0.9646	0.9697
PSNR	84.8196	88.6846	79.7276	81.5757	78.1123	79.9083	79.4679	80.0348
Size of dataset	3600	324	3600	324	3600	324	3600	324

Fig. 4 | Results of holographic experiments on HeLa cells. (**a**) Low-carrier-frequency high-contrast hologram collected by slightly off-axis interferometry system. (**b**) Corresponding spatial frequency spectrum. (**c**) The result of phase recovery by slightly off-axis holography using FT method under ×20 lens. (**d**) The result of phase recovery using DL-VHQPI. (**e1**–**e4**) and (**f1**–**f4**) correspond to the local amplification results of "Area1" and "Area2" for the two samples under different phase recovery methods. Where (e2, e4, f2, f4) are the corresponding DIC views, respectively. (**g**) and (**h**) The DIC views after partial magnification of the phase map in the corresponding red box. (**i**) The numerical distribution of the cross-section and detail-preservation feature of the DL-VHQPI.

coarse diagonal-fringe distributions; in comparison, that of DL-VHQPI is much smoother. The calculated Standard Deviation (STD) quantitatively demonstrates that DL-VHQPI performs a better flatness distribution. As can be readily observed in the cross-section presented in Fig. 4(i), in the FT phase recovery, the reconstruction errors brought by the autocorrelation term will introduce noticeable artifacts to the correct phase result. The results demonstrate that DL-VHQPI can excellently suppress phase artifacts and own the effectiveness and applicability for a practical slightly off-axis DH system.

Indeed, reducing the size of the FT filter window may also be a good way to alleviate artifacts, but this will not fundamentally address the problem of the overlapped spectrum and will cause phase imaging blur. The reason is that reducing the filtering window is at the expense of the system's SBP and the high-frequency information of the object cannot be enclosed in the limited filtering window. In the Section 3 of Supplementary information, we experimentally present the imaging effects under different FT filtering windows for living cells. To verify the

generalization of DL-VHQPI, we supplemented a new group of experimental results for living cells in Supplementary Section 5, in which we added a comparison and discussion with the VHQPI method and the traditional FT method. The results demonstrate that DL-VHQPI still performs the best artifact-suppression ability and generalizability under a new group of biological applications.

Conclusions and discussions

In summary, we proposed a high-accuracy artifacts-free single-frame low-carrier frequency fringe demodulation approach for the slightly off-axis DH system, *i.e.*, a model using the DNN-assisted physical process. When the cross-correlation and autocorrelation are inevitably aliased in the SFD, the phase reconstruction based on the conventional FT method cannot eliminate the effect of phase artifacts caused by zero-order term[6]. Although reducing the size of the FT filter window may alleviate the problem of imaging artifacts, the high-frequency information loss of the object caused by the limited filtering

window will cause imaging blur. The method based on Kramers-Kronig relation is proposed on the basis of the concept of "*cepstrum*" and homomorphic filtering[31], however, this method must depend on the limited condition of the object-reference ratio and need the separation of the high-order terms in the extended SFD[32,33]. Furthermore, the VHQPI implements the background component removal of single-frame hologram utilizing the principle of image frequency components extraction, while it inevitably suffers from the non-sufficient background term removal caused by the physical method[34]. In contrast, DL-VHQPI, a novel DL-assisted physical model method, can better suppress phase artifacts while improving imaging accuracy. The simulation result quantitatively demonstrates that the phase recovery accuracy obtained by DL-VHQPI is greatly superior to that by FT and VHQPI. Moreover, the live-cell experiment results demonstrate that our method is applicable in biological research.

In addition, it is noteworthy that in the classical end-to-end DNN model (without a physical model), massive data pairs are required to train the network model for a higher reconstruction precision. However, it may be prohibitively laborious and time-consuming for the real-world DH system to collect datasets and generate the corresponding ground truth. Conversely, the proposed DL-VHQPI can perform better same-domain generalization ability and image data-feature extraction capability without a large of datasets. Compared to the classical end-to-end DNN model (i.e., DL-noPhy), DL-VHQPI can achieve a higher reconstruction accuracy utilizing only a small fraction of the datasets due to the physical model reducing the information entropy of DL training objects. Meanwhile, fewer datasets mean shorter training time and higher training efficiency.

The significance of our work lies in the multiple possibilities of applying the proposed DL-assisted physical model idea to the QPI. This idea can be applied to many scenarios in which deep learning methods are applied to the QPI, e.g., addressing a series of ill-posed inverse phase retrieval problems and holography-based high-throughput optical diffraction tomography (ODT) problems[63−65]. Specifically, the artifacts-free low-carrier-frequency fringe demodulation capability of the proposed method has application possibilities for ODT imaging of wide-bandwidth objects. In addition, it has also implications for high-throughput studies of high-robust common-path off-axis interferometer systems[66,67]. We envision that the idea presented in this research can be applicable to a diverse range of future computational imaging techniques, not just limited to what we discussed here.

References

1. Fan Y, Li JJ, Lu LP, Sun JS, Hu Y et al. Smart computational light microscopes (SCLMs) of smart computational imaging laboratory (SCILab). *PhotoniX* **2**, 19 (2021).

2. Lee K, Kim K, Jung J, Heo J, Cho S et al. Quantitative phase imaging techniques for the study of cell pathophysiology: from principles to applications. *Sensors* **13**, 4170–4191 (2013).

3. Park Y, Depeursinge C, Popescu G. Quantitative phase imaging in biomedicine. *Nat Photonics* **12**, 578–589 (2018).

4. Vicar T, Balvan J, Jaros J, Jug F, Kolar R et al. Cell segmentation methods for label-free contrast microscopy: review and comprehensive comparison. *BMC Bioinformatics* **20**, 360 (2019).

5. Gao P, Wirth R, Lackner J, Sunbul M, Jaeschke A et al. Super-resolution imaging of live cells with genetically encoded silicon rhodamine-binding RNA aptamers. *Biophys J* **118**, 145A (2020).

6. Li ZS, Fan Y, Sun JS, Zuo C, Chen Q. A commercialized digital holographic microscope with complete software supporting. *Proc SPIE* **11571**, 115711C (2020).

7. Kim MK. Principles and techniques of digital holographic microscopy. *SPIE Rev* **1**, 018005 (2010).

8. Kemper B, von Bally G. Digital holographic microscopy for live cell applications and technical inspection. *Appl Opt* **47**, A52–A61 (2008).

9. Gao P, Yuan CJ. Resolution enhancement of digital holographic microscopy via synthetic aperture: a review. *Light Adv Manuf* **3**, 105–120 (2022).

10. Bettenworth D, Lenz P, Krausewitz P, Brückner M, Ketelhut S et al. Quantitative stain-free and continuous multimodal monitoring of wound healing *in vitro* with digital holographic microscopy. *PLoS One* **9**, e107317 (2014).

11. Coppola G, Ferraro P, Iodice M, De Nicola S, Finizio A et al. A digital holographic microscope for complete characterization of microelectromechanical systems. *Meas Sci Technol* **15**, 529–539 (2004).

12. Anand V, Han ML, Maksimovic J, Ng SH, Katkus T et al. Single-shot mid-infrared incoherent holography using Lucy-Richardson-Rosen algorithm. *Opto-Electron Sci* **1**, 210006 (2022).

13. Xu K, Wang X E, Fan X H et al. Meta-holography: from concept to realization. *Opto-Electron Eng* **49**, 220183 (2022).

14. Gao H, Fan XH, Xiong W, Hong MH. Recent advances in optical dynamic meta-holography. *Opto-Electron Adv* **4**, 210030 (2021).

15. Gabai H, Baranes-Zeevi M, Zilberman M, Shaked NT. Continuous wide-field characterization of drug release from skin substitute using off-axis interferometry. *Opt Lett* **38**, 3017–3020 (2013).

16. Huang ZZ, Memmolo P, Ferraro P, Cao LC. Dual-plane coupled phase retrieval for non-prior holographic imaging. *PhotoniX* **3**, 3 (2022).

17. Wu XJ, Sun JS, Zhang JL, Lu LP, Chen R et al. Wavelength-scanning lensfree on-chip microscopy for wide-field pixel-super-resolved quantitative phase imaging. *Opt Lett* **46**, 2023–2026

(2021).

18. Wang HD, Göröcs Z, Luo W, Zhang YB, Rivenson Y et al. Computational out-of-focus imaging increases the space–bandwidth product in lens-based coherent microscopy. *Optica* **3**, 1422–1429 (2016).

19. Micó V, García J, Zalevsky Z, Javidi B. Phase-shifting Gabor holography. *Opt Lett* **34**, 1492–1494 (2009).

20. Poon TC. *Digital Holography and Three-Dimensional Display: Principles and Applications* (Springer, New York, 2006).

21. Claus D, Iliescu D, Bryanston-Cross P. Quantitative space-bandwidth product analysis in digital holography. *Appl Opt* **50**, H116–H127 (2011).

22. Zhong Z, Bai HY, Shan MG, Zhang YB, Guo LL. Fast phase retrieval in slightly off-axis digital holography. *Opt Lasers Eng* **97**, 9–18 (2017).

23. Xue L, Lai JC, Wang SY, Li ZH. Single-shot slightly-off-axis interferometry based Hilbert phase microscopy of red blood cells. *Biomed Opt Express* **2**, 987–995 (2011).

24. Shaked NT, Zhu YZ, Rinehart MT, Wax A. Two-step-only phase-shifting interferometry with optimized detector bandwidth for microscopy of live cells. *Opt Express* **17**, 15585–15591 (2009).

25. Pavillon N, Arfire C, Bergoënd I, Depeursinge C. Iterative method for zero-order suppression in off-axis digital holography. *Opt Express* **18**, 15318–15331 (2010).

26. Trusiak M, Picazo-Bueno JA, Patorski K, Zdankowski P, Mico V. Single-shot two-frame π-shifted spatially multiplexed interference phase microscopy. *J Biomed Opt* **24**, 096004 (2019).

27. León-Rodríguez M, Rayas JA, Cordero RR, Martínez-García A, Martínez-Gonzalez A et al. Dual-plane slightly off-axis digital holography based on a single cube beam splitter. *Appl Opt* **57**, 2727–2735 (2018).

28. Han JH, Gao P, Yao BL, Gu YZ, Huang MJ. Slightly off-axis interferometry for microscopy with second wavelength assistance. *Appl Opt* **50**, 2793–2798 (2011).

29. Ikeda T, Popescu G, Dasari RR, Feld MS. Hilbert phase microscopy for investigating fast dynamics in transparent systems. *Opt Lett* **30**, 1165–1167 (2005).

30. Guo CS, Wang BY, Sha B, Lu YJ, Xu MY. Phase derivative method for reconstruction of slightly off-axis digital holograms. *Opt Express* **22**, 30553–30558 (2014).

31. Pavillon N, Seelamantula CS, Kühn J, Unser M, Depeursinge C. Suppression of the zero-order term in off-axis digital holography through nonlinear filtering. *Appl Opt* **48**, H186–H195 (2009).

32. Baek Y, Lee K, Shin S, Park Y. Kramers–Kronig holographic imaging for high-space-bandwidth product. *Optica* **6**, 45–51 (2019).

33. Baek Y, Park Y. Intensity-based holographic imaging via space-domain Kramers–Kronig relations. *Nat Photonics* **15**, 354–360 (2021).

34. Trusiak M, Cywińska M, Micó V, Picazo-Bueno JÁ, Zuo C et al. Variational Hilbert quantitative phase imaging. *Sci Rep* **10**, 13955 (2020).

35. Cywińska M, Trusiak M, Patorski K. Automatized fringe pattern preprocessing using unsupervised variational image decomposition. *Opt Express* **27**, 22542–22562 (2019).

36. Larkin KG, Bone DJ, Oldfield MA. Natural demodulation of two-dimensional fringe patterns. I. General background of the spiral phase quadrature transform. *J Opt Soc Am A* **18**, 1862–1870 (2001).

37. Zuo C, Qian JM, Feng SJ, Yin W, Li YX et al. Deep learning in optical metrology: a review. *Light Sci Appl* **11**, 39 (2022).

38. Feng SJ, Chen Q, Gu GH, Tao TY, Zhang L et al. Fringe pattern analysis using deep learning. *Adv Photonics* **1**, 025001 (2019).

39. Feng SJ, Zuo C, Hu Y, Li YX, Chen Q. Deep-learning-based fringe-pattern analysis with uncertainty estimation. *Optica* **8**, 1507–1510 (2021).

40. Cywińska M, Brzeski F, Krajnik W, Patorski K, Zuo C et al. DeepDensity: convolutional neural network based estimation of local fringe pattern density. *Opt Lasers Eng* **145**, 106675 (2021).

41. Pan B. Optical metrology embraces deep learning: keeping an open mind. *Light Sci Appl* **11**, 139 (2022).

42. Zuo C, Qian JM, Feng SJ, Yin W, Li YX et al. Correction: deep learning in optical metrology: a review. *Light Sci Appl* **11**, 74 (2022).

43. Feng SJ, Zuo C, Zhang L, Yin W, Chen Q. Generalized framework for non-sinusoidal fringe analysis using deep learning. *Photonics Res* **9**, 1084–1098 (2021).

44. Li YX, Qian JM, Feng SJ, Chen Q, Zuo C. Deep-learning-enabled dual-frequency composite fringe projection profilometry for single-shot absolute 3D shape measurement. *Opto-Electron Adv* **5**, 210021 (2022).

45. Zheng CH, Wang TS, Liu ZQ et al. Deep transfer learning method to identify orbital angular momentum beams. *Opto-Electron Eng* **49**, 210409 (2022).

46. Zheng ZH, Zhu SK, Chen Y, Chen HY, Chen JH. Towards integrated mode-division demultiplexing spectrometer by deep learning. *Opto-Electron Sci* **1**, 220012 (2022).

47. Rivenson Y, Zhang YB, Günaydın H, Teng D, Ozcan A. Phase recovery and holographic image reconstruction using deep learning in neural networks. *Light Sci Appl* **7**, 17141 (2018).

48. Rivenson Y, Wu YC, Ozcan A. Deep learning in holography and coherent imaging. *Light Sci Appl* **8**, 85 (2019).

49. Chen HL, Huang LZ, Liu TR, Ozcan A. Fourier Imager Network (FIN): a deep neural network for hologram reconstruction with superior external generalization. *Light Sci Appl* **11**, 254 (2022).

50. Lempitsky V, Vedaldi A, Ulyanov D. Deep image prior. In *Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* 9446–9454 (IEEE, 2018);http://doi.org/10.1109/CVPR.2018.00984.

51. Wang F, Bian YM, Wang HC, Lyu M, Pedrini G et al. Phase imaging with an untrained neural network. *Light Sci Appl* **9**, 77 (2020).

52. Duran J, Coll B, Sbert C. Chambolle's projection algorithm for total variation denoising. *Image Process Line* **3**, 311–331 (2013).

53. Zhu XJ, Chen ZQ, Tang C. Variational image decomposition for automatic background and noise removal of fringe patterns. *Opt Lett* **38**, 275–277 (2013).

54. Bianco V, Memmolo P, Paturzo M, Finizio A, Javidi B et al. Quasi noise-free digital holography. *Light Sci Appl* **5**, e16142 (2016).

55. Kluver JW. Elimination of slip and instability effects in certain *M*-type electron beams. *Proc IEEE* **51**, 868–868 (1963).

56. Yang X, Yu QF, Fu SH. A combined method for obtaining fringe orientations of ESPI. *Opt Commun* **273**, 60–66 (2007).

57. Deng M, Li S, Zhang ZY, Kang I, Fang NX et al. On the interplay between physical and content priors in deep learning for computational imaging. *Opt Express* **28**, 24152–24170 (2020).

58. Shannon CE. A mathematical theory of communication. *ACM SIGMOBILE Mob Comput Commun Rev* **5**, 3–55 (2001).
59. Cover TM. *Elements of Information Theory*. John Wiley & Sons, 1999).
60. 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* (JMLR. org, 2015).
61. Nair V, Hinton GE. Rectified linear units improve restricted Boltzmann machines. In *Proceedings of the 27th International Conference on Machine Learning* 807–814 (Omnipress, 2010).
62. Kingma DP, Ba J. Adam: a method for stochastic optimization. In *Proceedings of the 3rd International Conference on Learning Representations*. https://arxiv.org/abs/1412.6980 (2015).
63. Choi W, Fang-Yen C, Oh S, Lue N, Dasari RR et al. Tomographic phase microscopy: quantitative 3D-mapping of refractive index in live cells. *Imaging Microsc* **10**, 48–50 (2008).
64. Sung Y, Choi W, Fang-Yen C, Badizadegan K, Dasari RR et al. Optical diffraction tomography for high resolution live cell imaging. *Opt Express* **17**, 266–277 (2009).
65. Li JJ, Matlock AC, Li YZ, Chen Q, Zuo C et al. High-speed *in vitro* intensity diffraction tomography. *Adv Photonics* **1**, 066004 (2019).
66. Mico V, Zalevsky Z, García J. Superresolution optical system by common-path interferometry. *Opt Express* **14**, 5168–5177

(2006).
67. Zhang JW, Dai SQ, Ma CJ, Xi TL, Di JL et al. A review of common-path off-axis digital holography: towards high stable optical instrument manufacturing. *Light Adv Manuf* **2**, 333–349 (2021).

Acknowledgements

Competing interests

The authors declare no competing financial interests.

Supplementary information

Supplementary information for this paper is available.
https://doi.org/10.29026/oes.2023.220023