Lensless complex amplitude demodulation based on deep learning in holographic data storage

Jianying Hao, Xiao Lin, Yongkun Lin, Mingyong Chen, Ruixian Chen, Guohai Situ, Hideyoshi Horimai and Xiaodi Tan

---

Related articles

The detection method for grab of portal crane based on deep learning
Zhang Wenming, Liu Xiangyang, Li Haibin, Li Yaqian
Opto-Electronic Engineering   2021 **48**, 200062   doi: 10.12086/oee.2021.200062

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

All-optical computing based on convolutional neural networks
Kun Liao, Ye Chen, Zhongcheng Yu, Xiaoyong Hu, Xingyuan Wang, Cuicui Lu, Hongtao Lin, Qingyang Du, Juejun Hu, Qihuang Gong
Opto-Electronic Advances   2021 **4**, 200060   doi: 10.29026/oea.2021.200060

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

More related article in Opto-Electron Journals Group website

Lensless complex amplitude demodulation based on deep learning in holographic data storage

Jianying Hao[1,3], Xiao Lin[1]*, Yongkun Lin[1], Mingyong Chen[1], Ruixian Chen[1], Guohai Situ[2], Hideyoshi Horimai[3] and Xiaodi Tan[1]*

To increase the storage capacity in holographic data storage (HDS), the information to be stored is encoded into a complex amplitude. Fast and accurate retrieval of amplitude and phase from the reconstructed beam is necessary during data readout in HDS. In this study, we proposed a complex amplitude demodulation method based on deep learning from a single-shot diffraction intensity image and verified it by a non-interferometric lensless experiment demodulating four-level amplitude and four-level phase. By analyzing the correlation between the diffraction intensity features and the amplitude and phase encoding data pages, the inverse problem was decomposed into two backward operators denoted by two convolutional neural networks (CNNs) to demodulate amplitude and phase respectively. The experimental system is simple, stable, and robust, and it only needs a single diffraction image to realize the direct demodulation of both amplitude and phase. To our investigation, this is the first time in HDS that multilevel complex amplitude demodulation is achieved experimentally from one diffraction intensity image without iterations.

Keywords: holographic data storage; complex amplitude demodulation; deep learning; computational imaging

Introduction

According to the prediction of the International Data Cooperation (IDC) in 2018, global data will increase to 175 ZB by 2025 and 2142 ZB by 2035[1]. Handling this enormous increasing volume of data based on the current data storage techniques is challenging. Compared with conventional storage devices such as the flash, hard disk drive (HDD), and magnetic tape, the optical storage techniques and devices such as blue-ray optical discs[2],

optical glass storage[3,4], and HDS[5–8] feature more advantageous properties such as lower energy consumption, stable storage, and longer lifetime that can be leveraged for the long-term storage of big data. HDS handles digital information as a two-dimensional (2D) array called a data page and records a three-dimensional (3D) hologram in light-sensitive media[6,7,9]. The storage density is determined by the hologram size, number of multiplexing recordings, and single data page capacity. Conventional HDS uses a binary amplitude data page[7,9–11], and

[1]College of Photonic and Electronic Engineering, Key Laboratory of Opto-Electronic Science and for Medicine of Ministry of Education, Fujian Provincial Key Laboratory of Photonics Technology, Fujian Provincial Engineering Technology Research Center of Photoelectric Sensing Application, Fujian Normal University, Fuzhou 350117, China; [2]Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Shanghai 201800, China; [3]HolyMine Corporation, 2032-2-301 Ooka, Numazu, Shizuoka 410-0022, Japan.

*Correspondence: X Lin, E-mail: xiaolin@fjnu.edu.cn; XD Tan, E-mail: xtan@fjnu.edu.cn

the capacity of a single page is limited. With the development of optical devices such as spatial light modulators (SLMs), multi-level phase-modulated HDS[12-15] and complex-amplitude-modulated HDS[16-18] that use the phase of light to encode information, have been proposed to increase the amount of information in one data page. Although phase encoding can improve the storage capacity significantly, the phase information cannot be read directly. It needs to be computationally retrieved from the intensity image captured by a detector. To demodulate the complex amplitude quickly and accurately with a simple and stable system is the key for the HDS system to keep the data transmitting rate (phase modulation is considered a special case of complex amplitude modulation with a uniform amplitude of 1).

Some researchers have attempted to decode the HDS complex amplitudes via different methods. Depending on whether there is interference in the retrieval process, HDS data demodulation systems can be divided into two types: interferometric and non-interferometric. Nobukawa proposed a complex-amplitude data page demodulation method based on digital holography and realized two-level amplitude and four-level phase demodulation[16]. Katano used two convolutional neural networks (CNNs) to retrieve the amplitude and phase from four interferometric holograms[17]. Both these methods require a reference beam, and the reading system is complicated and sensitive to environmental vibrations. Moreover, the multi-capture operation also decreases the data transfer rate. Lin proposed non-interferometric phase demodulation based on an iterative Fourier transform algorithm[12,19]. The optical system is simplified, but the data transfer rate decreases because of the iterative calculation in the retrieval process. Bunsen used the TIE algorithm to retrieve the complex amplitude from three diffraction intensity images[18]. It requires three intensity images and iteration calculations that also decrease the transfer rate. Horisaki presented a method for single shot, complex-amplitude imaging that can retrieve the phase and amplitude from a single intensity image directly. However, it is only used in handwritten digit recognition, where the distributions are relatively simple[20]. The data reading of HDS requires a simple system to ensure both stable data transmission and fewer iterations to improve the data transfer rate. In conventional methods, it is difficult to achieve a balance between the transmission speed and stability.

In recent years, we have witnessed the emergence of deep learning that demonstrates great potential in various fields such as computer vision[21], optical encryption[22], and optical computing[23]. Since the pioneering work of Sinha et al.[24] on the recovery of phase directly from a single diffraction intensity image, deep learning networks have been used to reconstruct the amplitude and phase directly from a hologram[25-27] or combined with physics prior to retrieving the phase from a diffraction intensity image[28]. One can refer to a recent survey[29] for more details. In the existing studies, the retrieval of the complex amplitude from the hologram still needs reference beam. The system is complicated and not suitable for HDS. Physical prior-based phase retrieval from diffraction intensity does not need reference beam but requires thousands of iterations to achieve phase demodulation. Inspired by these studies, our previous work proposed a non-iterative lensless phase demodulation method based on deep learning and embedded data used in HDS[30]. In this study, we further propose a complex amplitude encoding and demodulation method. A complex amplitude encoding method with a higher capacity is designed, and a lensless non-interferometric complex amplitude retrieval system is established. The inverse problem is decomposed into two backward operators related to amplitude and phase and represented by two CNNs separately. The four-level complex amplitude can be demodulated directly without interferometry and iteration. The experiment verifies its feasibility and demonstrates the potential of complex amplitude demodulation from a single intensity image. The analysis of the diffraction features and encoded data pages provide guidance for future research on deep learning-based complex amplitude retrieval.

Principle

Complex amplitude modulation for HDS

A schematic of complex amplitude modulated HDS is shown in Fig. 1. Information storage and reading in HDS include four processes: encoding, recording, reading, and decoding. In the encoding process, information is encoded into amplitude and phase data pages to modulate the signal beam. In the recording process, the signal beam is converged by Lens 1, and it subsequently interferes with a reference beam. The interference hologram is recorded in the media. In the reading process, a reconstructed beam is obtained at the focal plane of Lens 2 by using a same reference beam to irradiate the media. The
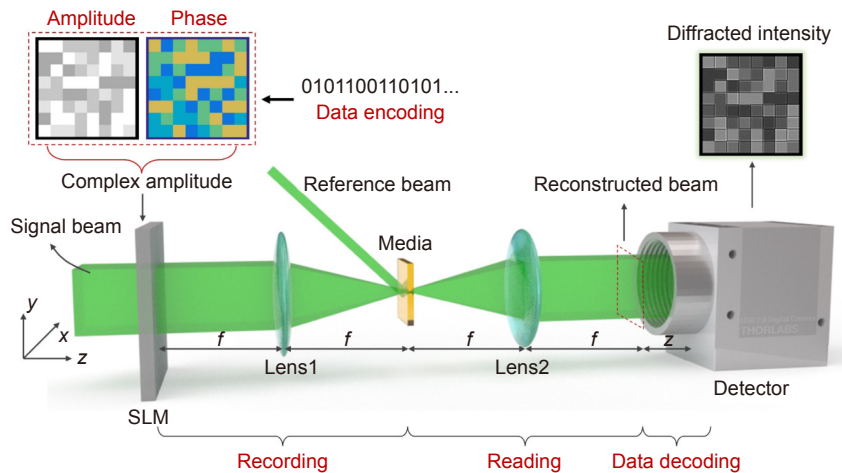
Fig. 1 | Conceptual diagram of HDS with complex-amplitude-modulated data page.

reconstructed beam contains the encoded amplitude and phase information. In the decoding process, a detector is placed at distance z from the reconstructed beam plane to capture the diffraction intensity image. The complex-amplitude information is retrieved from the intensity image by the method proposed in this paper. The diffraction intensity image is used for complex amplitude demodulation, which is captured directly without any lens, so it is called a lensless complex amplitude decoding system.

According to the theory of HDS, the wavefront of the reconstructed beam is exactly the same with the recorded signal beam if the influence of the system and material noise is ignored. Information reading is to obtain the encoded amplitude and phase pages from a single-intensity image captured by the detector. The light field of the reconstructed beam is defined by Eq. (1):

$$U(x, y; 0) = A(x, y; 0) \exp(i\varphi(x, y; 0)) . \quad (1)$$

The position of the reconstructed beam plane (the back focal of Lens 2 in Fig. 1) is defined as $z = 0$. The diffraction field of $U(x, y; z)$ over distance z is given by Eq. (2):

$$U(x, y; z) =$$
$$\int\int_{-\infty}^{\infty} F(f_X, f_Y; 0)\, G \exp(i2\pi(f_X x + f_Y y))\, df_X df_Y , \quad (2)$$

where G is the transfer function $G = \exp\left[i\frac{2\pi}{\lambda}\sqrt{1 - (\lambda f_X)^2 - (\lambda f_Y)^2}\,z\right]$, f_X and f_Y are the spatial frequencies in the x and y directions, respectively, and $F(f_X, f_Y; 0)$ is the angular spectrum of $U(x, y, 0)$, that is the Fourier transform of $U(x, y, 0)$ shown in Eq. (3):

$$F(f_X, f_Y; 0) = \int\int_{-\infty}^{\infty} U(x, y; 0) \exp(i2\pi(f_X x + f_Y y))\, dx dy .$$
$$(3)$$

The intensity of the light field at distance z is expressed by Eq. (4):

$$I(x, y; z) = |U(x, y; z)|^2 = H(U(x, y; 0)) . \quad (4)$$

Because of diffraction, the intensity $I(x, y; z)$ recorded by the detector has a certain pattern, as shown in Fig. 1. The operator $H(\cdot)$ denotes the mapping function that relates the complex amplitude $U(x, y; 0)$ and intensity image $I(x, y, z)$.

Complex amplitude encoding and Fresnel diffraction

According to the Huygens–Fresnel diffraction principle, the constructive and destructive interference of coherent light caused by diffraction produces an intensity pattern in the irradiance result $I(x, y; z)$[31]. A simple theoretical model was developed (Fig. 2) to describe the diffraction process in HDS demodulation. In this model, each data point on the SLM is regarded as a divergent point light source. As light propagates forward, the secondary sources corresponding to different data points are superimposed on each other to form a new light field. The further the propagation, the stronger the diffraction effect. An appropriate distance $z = d$ can be found, where the intensity I related to a data point P is mainly determined by the eight surrounding data points (P_p, $p = 1, \cdots, 8$). Because of the amplitude and phase difference between data points P and P_p, there is a diffraction pattern that has a clear correlation with the encoded amplitude and phase data. Both the original amplitude features and the new feature resulting from the phase difference are
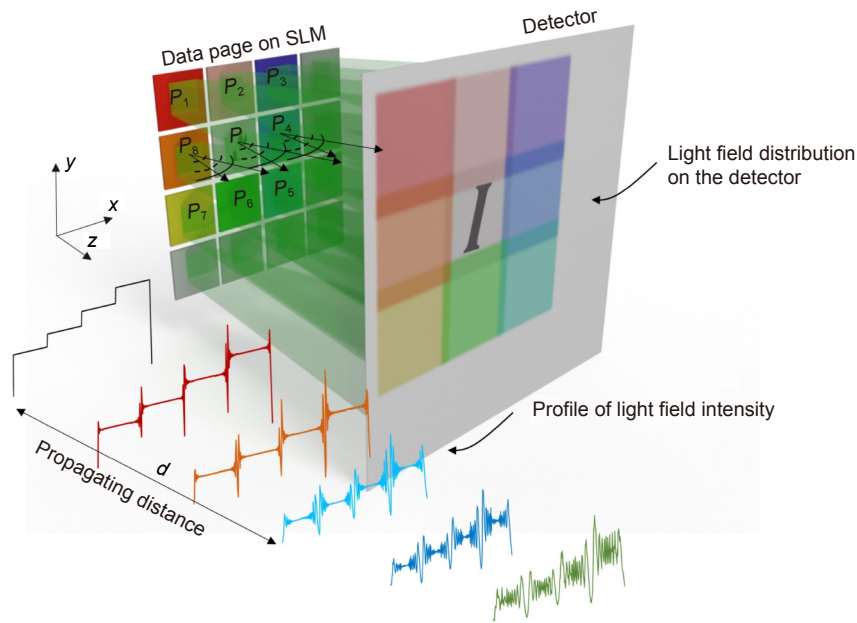
Fig. 2 | Simplified theoretical model of diffraction process in HDS demodulation.

included in this diffraction pattern.

In order to analyze the relationship between the diffraction pattern and the amplitude-phase data numerically, we set the complex amplitude data as four-level phase ($\pi/6$, $2\pi/3$, π, $3\pi/2$) and amplitude (0.7, 0.8, 0.9, 1) randomly. The phase values are randomly selected from $[0, 2\pi]$ at unequal intervals. It is different from the traditional encoding rule in HDS, which will be explained in the second half of this Section. The complex amplitude encoding page and intensity image are shown in Fig. 3(I). Phase-only encoding and amplitude-only encoding are also shown in Fig. 3(II) and Fig. 3(III), respectively, to facilitate the analysis of the diffraction pattern features. The diffraction distance $z = 2$ mm. The cross-section at the position of the red line in the intensity images is shown in Fig. 3(j). The intensity image in Fig. 3(I) shows that the complex-amplitude-based pattern consists of both the amplitude-related and phase-related features. The phase-related feature is manifest in the intensity distribution resulting from light propagation, as shown in Fig. 3(II). The amplitude-related feature in Fig. 3(III) is the original amplitude distribution that is still distinguishable under the condition of this diffraction distance.

Figure 4(a) and Fig. 4(b) is the 2D and 3D contour map of the intensity in Fig. 3(c), respectively. Fig. 4(c) is the cross-section of the red line in Fig. 4(a). Corresponding to the four data points, four diffraction regions are generated. The mean diffracted intensity of each data re-

gion in Fig. 4(c) is consistent with the amplitude distribution. Within each data point region, the intensity variation due to diffraction is a phase-dependent feature. Therefore, it is feasible to detect the amplitude and phase features separately. Here, the amplitude-related features in the diffraction intensity are clearly visible and directly related to the original amplitude. The relationship between the phase-related features and the diffraction pattern requires further analysis.

Based on the simplified theoretical model in Fig. 2, $\{(P, P_p),\ p = 1, 2, \ldots, 8\}$ is defined as a symbol that consists of one central data point $\boldsymbol{P}$ and its eight adjacent data points $\boldsymbol{P}_p$. If only the phase is considered, then the symbol is simplified as $\{(\phi, \phi_p),\ p = 1, 2, \ldots, 8\}$. Several examples of different symbols are shown in Fig. 5(I). Figure 5(II) are the corresponding diffraction intensity images. The general architecture of a symbol consisting of 3×3 data points is shown in Fig. 5(k). Fig. 5(l) is the intensity profile on the red line in the intensity distribution Fig. 5(f–j). Two rules can be observed in Fig. 5. First, the diffraction pattern characteristics are decided by the phase difference $\Delta\phi_p$ between the central data point ϕ and its adjacent points ϕ_p. Second, the diffraction pattern characteristics are not directly decided by the absolute value of the central point ϕ. To build the relationship between the diffraction intensity and central phase value ϕ, the relationship between the diffraction feature and symbol is important.

Therefore, when designing the phase encoding rule for
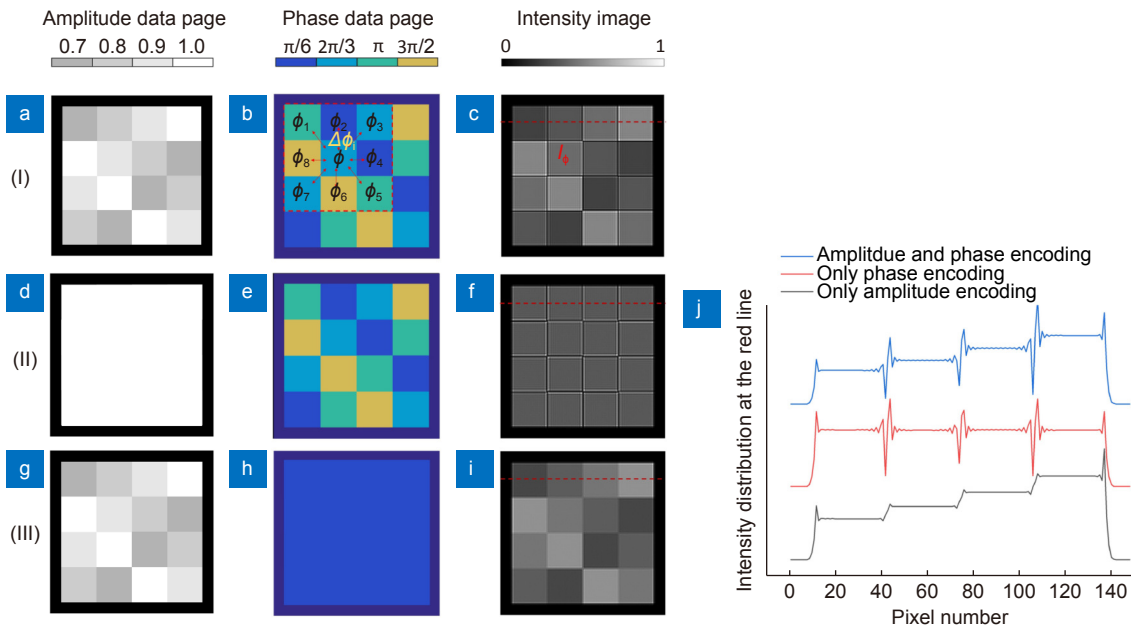
Fig. 3 | Complex amplitude encoding modes and the corresponding diffraction features. (**a**) Four-level amplitude data page. (**b**) Four-level phase data page. (**c**) Diffraction intensity image of the complex-amplitude data page (*z* = 2 mm). (**d**) Uniform amplitude with no encoding. (**e**) Four-level phase data page. (**f**) Diffraction pattern of the phase-only data page (*z* = 2 mm). (**g**) Four-level amplitude data page. (**h**) Uniform phase without encoding. (**i**) Diffraction pattern of the amplitude-only data page (*z* = 2 mm). (**j**) Intensity distribution of the three intensity images in the same horizontal section.
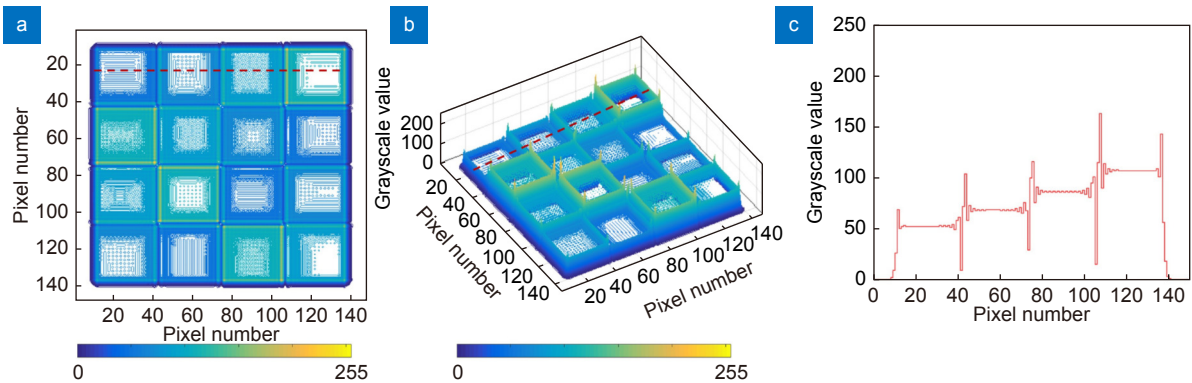

Fig. 4 | (**a**) Contour map of the intensity image in Fig. 3(c). (**b**) 3D contour map of the intensity image in Fig. 3(c). (**c**) Intensity distribution on the red line in (a) and (b).

HDS, the phase difference should be paid more attention to. With n-level phase encoding rule $\{\phi_i, i = 1, 2, \cdots, n\}$, each phase value can be represented by a symbol set $\{S_{\phi=\phi_i}, i = 1, 2, \cdots, n\}$. The symbol set $S_{\phi=\phi_i} = \left\{\left(\Delta\phi_p | (\phi = \phi_i)\right), p = 1, 2, \cdots, 8, i = 1, 2, \cdots, n\right\}$ consists of n^8 symbols with the architecture as shown in Fig. 5(k). The $\phi = \phi_i$ means the central data point value in the symbols. $\Delta\phi_p = \phi_p - \phi$ is the phase difference between the central data point and its adjacent data points. For each phase data point, the diffraction patterns are determined by the structure of the related symbols. To infer the phase ϕ_i of a central data point from

the intensity image, the corresponding symbols in the set $S_{\phi=\phi_i}$ should be distinguishable. This implies that $\Delta\phi_p$ in each symbol set $S_{\phi=\phi_i}$ should be different.

As an example, we analyzed and compared the four-level phase encoding (0, $\pi/2$, π, $3\pi/2$) of conventional phase-modulated HDS[15,16] and ($\pi/6$, $2\pi/3$, π, $3\pi/2$) in this study. The phase difference $\Delta\phi_p$ under these two encoding rules are summarized in Table 1 and Table 3. All values of $\Delta\phi_p$ were set to $[-\pi, \pi]$. Meanwhile, we summarized the number of symbols in each symbol set $S_{\phi=\phi_i}$ and calculated the repetition probability defined as:
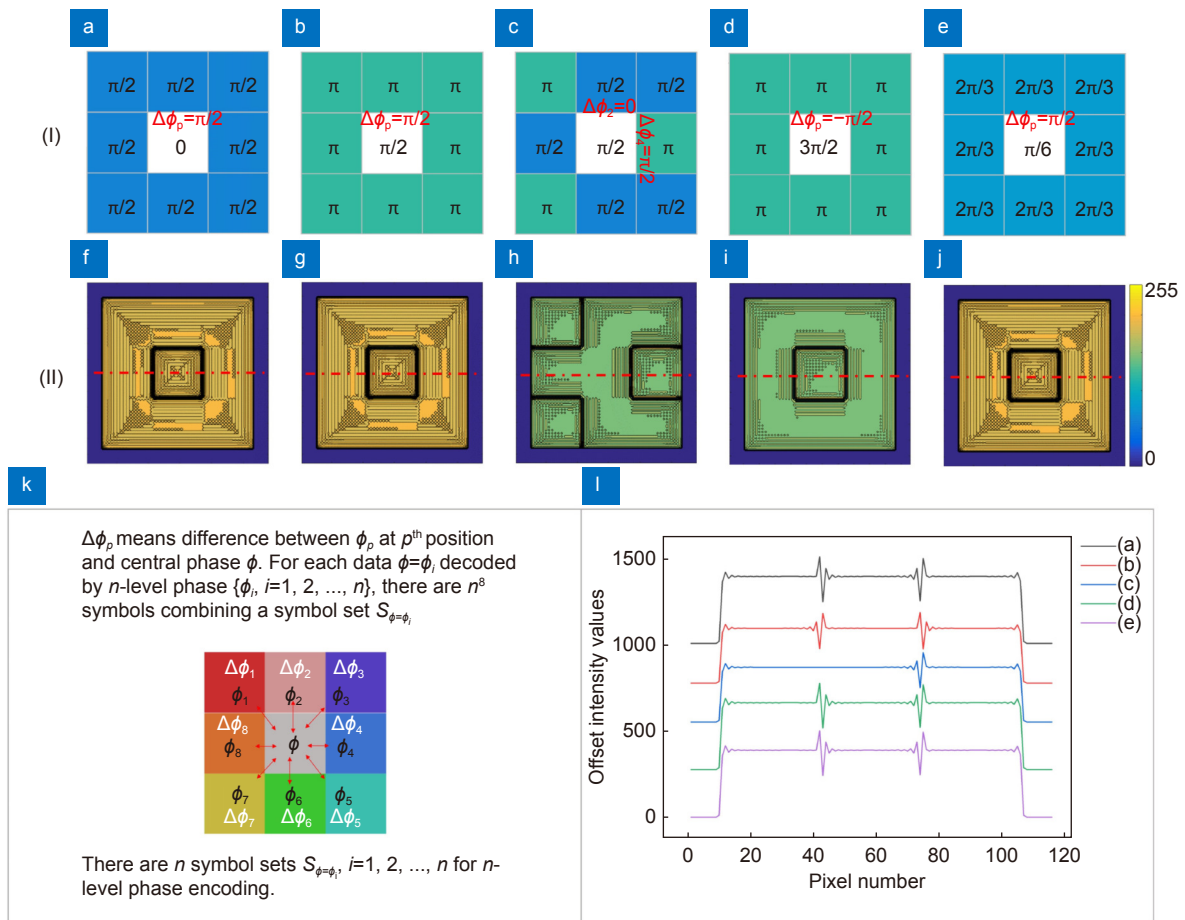
Fig. 5 | (a–e) Symbols with different phase patterns. **(f–j)** Diffraction intensity distributions corresponding to symbols (a)–(e). **(k)** General architecture of the symbol with 3×3 data points. **(l)** Intensity profile on the red line in intensity distribution (f–j).

$$P_{ij} = \frac{N_{ij}}{N_i}, \qquad (5)$$

where N_{ij} is the number of symbols that belong to both $S_{\phi=\phi_i}$ and $S_{\phi=\phi_j}$, and N_i is the total number of symbols in a set, $N_i = n^8$. The repetition probabilities associated with the two encoding rules are shown in Table 2 and Table 4.

From Table 1, it can be seen that for each encoded value $\phi = \phi_i$, the possible generated phase differences $\Delta\phi_p$ are same. This means that under the encoding rule (0, $\pi/2$, π, $3\pi/2$), all symbols in set $S_{\phi=\phi_i} = \left\{ \Delta\phi_p \mid (\phi = \phi_i), p = 1, 2, \ldots, 8, i = 1, 2, \ldots, n \right\}$ of different central phase values $\phi = \phi_i$ are absolutely the same. The repetition probabilities P_{ij} of all symbols in set $S_{\phi=\phi_i}$ and $S_{\phi=\phi_j}$ are 100%, as shown in Table 2. This implies that under the encoding rule (0, $\pi/2$, π, $3\pi/2$), the diffraction patterns corresponding to different ϕ_i are the same, so the phase value ϕ_i cannot be inferred from the pattern of the diffraction intensity. For comparison, under the encoding rule ($\pi/6$, $2\pi/3$, π, $3\pi/2$), the phase difference $\Delta\phi_p$ becomes distinguishable, as shown in Table

3. Therefore, the repetition probabilities of the symbol set in Table 4 are also significantly reduced. It is possible to establish a relationship between the diffraction intensity pattern and the phase value $\phi = \phi_i$.

This assumption was verified by the method introduced in Section *Deep learning-based complex amplitude demodulation*. A simulated experiment was conducted to generate datasets under two encoding rules (0, $\pi/2$, π, $3\pi/2$) and ($\pi/6$, $2\pi/3$, π, $3\pi/2$). Two CNNs were used to detect patterns in diffracted intensity images and establish relationships with amplitude and phase data pages, respectively. The core of the CNN is the convolution kernel, which is used to extract features of different scales in the input image. The hyperparameters of the CNN is set same as the experiment discussed in Section *Experiments*. During the training, the mean square error (MSE) losses represents the gap between the predicted and the true value of the two cases are shown in Fig. 6(a) and Fig. 6(b). The red and black lines represent the losses for the test and training sets, respectively. The smaller the loss function value, the closer the predicted result is to the

ground truth value. A detailed description of the neural network geometry and training is provided in Section *Deep learning-based complex amplitude demodulation* and *Experiments*. As shown in Fig. 6(a), the loss function values of the training and test sets do not converge after a brief plunge as the training progresses. Upon continued training, the training loss began to decrease, whereas the test loss increased, indicating that the model training resulted in overfitting. Under the encoding rule $(0, \pi/2, \pi, 3\pi/2)$, the model failed to establish the relationship between the diffraction intensity and phase data page. However, as shown in Fig. 6(b), as the training progresses, the losses of both the training and test sets drop rapidly and approach zero. This implies that the CNN

Table 1 | Possible phase difference $\Delta\phi_p=\phi_p-\phi$ under encoding rule (0, π/2, π, 3π/2).

Central phase value ϕ	Adjacent phase value ϕ_p			
	0	π/2	π	3π/2
0	$\Delta\phi_p|0 = 0$	$\Delta\phi_p|0 = \pi/2$	$\Delta\phi_p|0 = \pi$	$\Delta\phi_p|0 = -\pi/2$
π/2	$\Delta\phi_p|\pi/2 = -\pi/2$	$\Delta\phi_p|\pi/2 = 0$	$\Delta\phi_p|\pi/2 = \pi/2$	$\Delta\phi_p|\pi/2 = \pi$
π	$\Delta\phi_p|\pi = \pi$	$\Delta\phi_p|\pi/2 = \pi/2$	$\Delta\phi_p|\pi/2 = 0$	$\Delta\phi_p|\pi/2 = \pi/2$
3π/2	$\Delta\phi_p|3\pi/2 = \pi/2$	$\Delta\phi_p|3\pi/2 = \pi$	$\Delta\phi_p|3\pi/2 = -\pi/2$	$\Delta\phi_p|3\pi/2 = 0$

Table 2 | Probability that any two sets $S_{\phi=\phi_i}$ and $S_{\phi=\phi_j}$ have the same symbol under encoding rule (0, π/2, π, 3π/2).

| $\Delta\phi_p|(\phi=\phi_j)$	$\Delta\phi_p|(\phi=\phi_i)$			
	$\Delta\phi_p|0$	$\Delta\phi_p|\pi/2$	$\Delta\phi_p|\pi$	$\Delta\phi_p|3\pi/2$
$\Delta\phi_p|0$	-	100%	100%	100%
$\Delta\phi_p|\pi/2$	100%	-	100%	100%
$\Delta\phi_p|\pi$	100%	100%	-	100%
$\Delta\phi_p|3\pi/2$	100%	100%	100%	-

Table 3 | Possible phase difference $\Delta\phi_p=\phi_p-\phi$ under encoding rule (π/6, 2π/3, π, 3π/2).

Central phase value ϕ	Adjacent phase data ϕ_p			
	π/6	2π/3	π	3π/2
π/6	$\Delta\phi_p\pi/6 = 0$	$\Delta\phi_p|\pi/6 = \pi/2$	$\Delta\phi_p|\pi/6 = 5\pi/6$	$\Delta\phi_p|\pi/6 = -2\pi/3$
2π/3	$\Delta\phi_p|2\pi/3 = -\pi/2$	$\Delta\phi_p|2\pi/3 = 0$	$\Delta\phi_p|2\pi/3 = \pi/3$	$\Delta\phi_p|2\pi/3 = 5\pi/6$
π	$\Delta\phi_p|\pi = -5\pi/6$	$\Delta\phi_p|\pi = -\pi/3$	$\Delta\phi_p|\pi = 0$	$\Delta\phi_p|\pi = \pi/2$
3π/2	$\Delta\phi_p|3\pi/2 = 2\pi/3$	$\Delta\phi_p|3\pi/2 = -5\pi/6$	$\Delta\phi_p|3\pi/2 = -\pi/2$	$\Delta\phi_p|3\pi/2 = 0$

Table 4 | Probability that any two sets $S_{\phi=\phi_i}$ and $S_{\phi=\phi_j}$ have the same symbol under encoding rule (π/6, 2π/3, π, 3π/2).

$\Delta\phi_p|(\phi=\phi_j)$	$\Delta\phi_p|(\phi=\phi_i)$			
	$\Delta\phi_p|\pi/6$	$\Delta\phi_p|2\pi/3$	$\Delta\phi_p|\pi$	$\Delta\phi_p|3\pi/2$
$\Delta\phi_p|\pi/6$	-	0.39%	0.39%	0.0015%
$\Delta\phi_p|2\pi/3$	0.39%	-	0.0015%	0.39%
$\Delta\phi_p|\pi$	0.39%	0.0015%	-	0.39%
$\Delta\phi_p|3\pi/2$	0.0015%	0.39%	0.39%	-
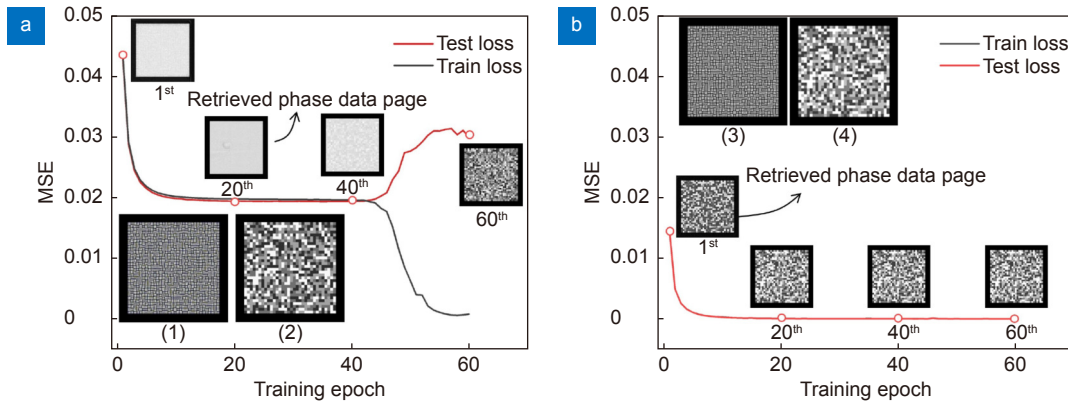
Fig. 6 | Change in the MSE using two different phase encoding datasets with (**a**) (0, π/2, π, 3π/2) and (**b**) (π/6, 2π/3, π, 3π/2) when training the same neural network. (1) and (3) are the same intensity images fed to the CNN and (2) and (4) are the ground truth of the phase data page. The pages at the 1st, 20th, 40th, and 60th epoch are the phase data page retrieved from the intensity image using the trained CNN.

can not only fit the data of the training set appropriately, but can also be generalized to the test set. In other words, the deep learning model can correctly represent the relationship between the diffraction intensity and phase data page under the encoding rule of (π/6, 2π/3, π, 3π/2).

The results in Tables 1–4 and Fig. 6 show the significant difference in phase encoding between the conventional iterative method based on physical processes[15,16,32] and the data-driven deep learning method in this study. In this study, we only compare two encoding rules and explain why the conventional encoding is not suitable for our deep learning method. It is worth emphasizing that the encoding rule (π/6, 2π/3, π, 3π/2) is feasible, but is not necessarily the optimal rule. In fact, the optimal encoding rules for lensless phase retrieval methods require further research which is beyond the scope of the present study.

Deep learning-based complex amplitude demodulation

From the analysis in Section *Complex amplitude encoding and Fresnel diffraction*, the amplitude and phase can be related to different features of the diffraction intensity. Therefore, operator $H^{-1}(\cdot)$ in Eq. (4) can be split into two operators, $H_A^{-1}(\cdot)$ and $H_\varphi^{-1}(\cdot)$, as shown in Eq. (6).

$$A(u, v) = H_A^{-1}(I(u, v)),$$
$$\phi(u, v) = H_\phi^{-1}(I(u, v)), \qquad (6)$$

where $A(u, v)$ is the amplitude data page, and $\phi(u, v)$ is the phase data page. $I(u,v)$ is the intensity matrix captured by the detector. The complex amplitude demodulation problem is decomposed to solve the inverse functions $H_A^{-1}(\cdot)$ and $H_\varphi^{-1}(\cdot)$.

$H_A^{-1}(\cdot)$ and $H_\phi^{-1}(\cdot)$ can be obtained by solving an op-

timization problem of the form shown in Eq. (7):

$$\tilde{C}(u, v) = \operatorname{argmin}_C \| H_C(C) - I \|^2 + \alpha \Phi(C), \qquad (7)$$

where C is used to represent the amplitude $A(u, v)$ or the phase $\phi(u, v)$, depending on whether the $H_A^{-1}(\cdot)$ or $H_\phi^{-1}(\cdot)$ is being optimized. $H_C(\cdot)$ is the forward operator, I is the measurement intensity image, Φ is the regularizer expressing prior information, and α is the regularization parameter that controls the relative strength of the two terms in the optimization function.

The deep learning-based approach to solve this problem involves learning a mapping function $H_C^{-1}(\cdot)$ from a large number of labeled datasets. Two CNNs are used to represent the inverse function operators $H_A^{-1}(\cdot)$ and $H_\phi^{-1}(\cdot)$ and are trained separately using the intensity-amplitude training dataset $\{(I_n, A_n)\, n = 1, \cdots, N\}$ and intensity-phase training dataset $\{(I_n, \varphi_n)\, n = 1, \cdots, N\}$. The main underlying convolution operations of CNN extract features of different scales from the input image. Neural network training optimizes the parameters of operator H_C^{-1}. In this study, the object function is defined by Eq. (8):

$$H_{C,\text{learn}}^{-1} = \operatorname*{argmin}_{H_{C,\theta}^{-1}, \theta \in \Theta} \sum_{n=1}^{N} \mathcal{L}\left(C_n, H_{C,\theta}^{-1}\{I_n\}\right) + \psi(\theta), \quad (8)$$

where H_C^{-1} is defined by a set of parameters that includes the weights and biases, and Θ is the set of all possible parameters in the neural network. $\mathcal{L}(\cdot)$ is the loss function used to measure the error between C_n and $H_{C,\theta}^{-1}\{I_n\}$, and $\psi(\theta)$ is a regularizer employed on the parameters with the aim of avoiding overfitting. The structure and hyperparameter settings of the two networks are the same. After training using datasets (I_n, A_n) and (I_n, ϕ_n) separately, two CNN models with different weights and

biases were obtained. Once the CNNs are trained, the complex-amplitude data page can be demodulated directly from the diffraction intensity image.

In this study, the architecture of the neural networks is set as Unet, as shown in Fig. 7. Unet is a fully convolutional network with an encoder–decoder architecture[33]. The input of the network is the diffraction intensity image $I(u, v)$ and the output is the amplitude data page $A(u, v)$ or phase data page $\phi(u, v)$ depending on whether the model is trained as an intensity-amplitude or intensity-phase model. We use rectified linear units (ReLU), that is, $\text{ReLU}(x) = \max(0, x)$ as the activation function, following the convolutional layers and Sigmoid, that is, $S(x) = (1 + e^{x})^{-1}$ in the output layer[34].

The entire process of the complex amplitude demodulation is shown in Fig. 8. In the training process, the training datasets of the intensity-amplitude and intensity-phase were fed into CNN1 and CNN2, respectively, to optimize the two models. In the test process, the amplitude and phase data pages were retrieved by trained CNN1 and CNN2 from a single diffraction intensity image directly. Subsequently, the corresponding complex-amplitude data were obtained by hard decision from the

complex-amplitude data page.

Experiments

Experimental setup

To verify the proposed method, an experimental optical system was set up, as shown in Fig. 9. A laser beam with a wavelength $\lambda = 532$ nm (MSL-FN-532) irradiated on SLM1 (CAS MICRPSTAR, FSLM-HD70-A/P, pixel pitch 8 μm) after collimation and expansion. P1 and P2 are linear polarizers. P1 is horizontally polarized, and P2 is vertically polarized to achieve amplitude-only modulation for SLM1. The amplitude data page was loaded into SLM1 to modulate the amplitude of the incident light. Subsequently, the amplitude-modulated light passed through two 4-f systems composed of lens L2, L3, L4, and L5, and irradiated on SLM2 (CAS MICRPSTAR, FSLM-2K70-VIS, pixel pitch 8 μm). The phase data page was loaded into SLM2 to modulate the phase. An aperture was used to filter out the unmodulated beam. HWP1 and HWP2 were used to adjust the polarization state of the light to meet the requirements of SLMs. The last 4-f system, composed of lens L6 and L7, is the recording and reading system of the HDS. The media
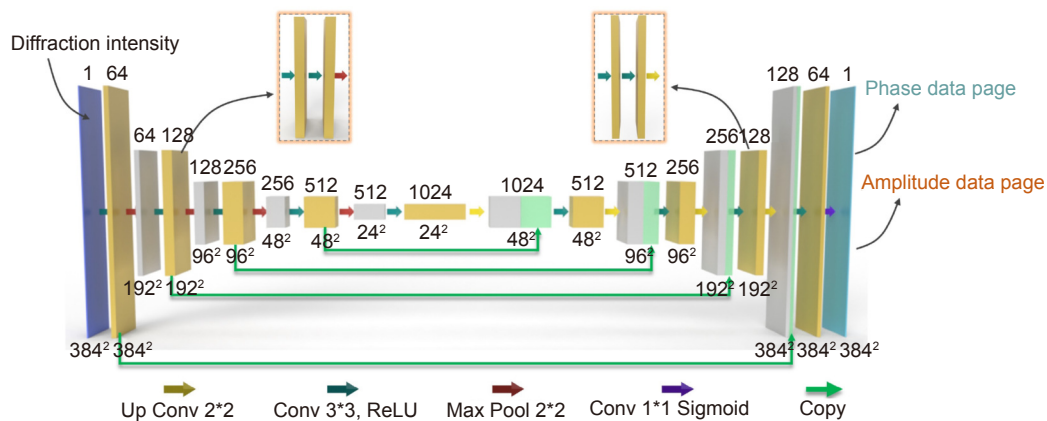

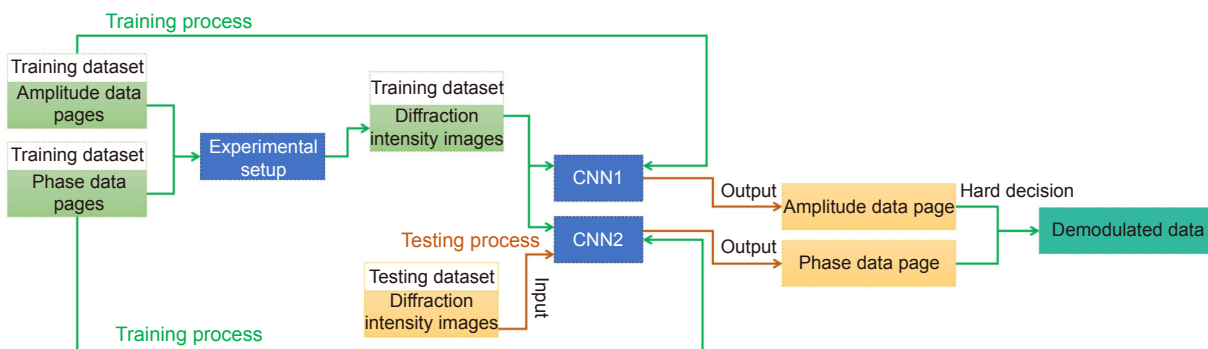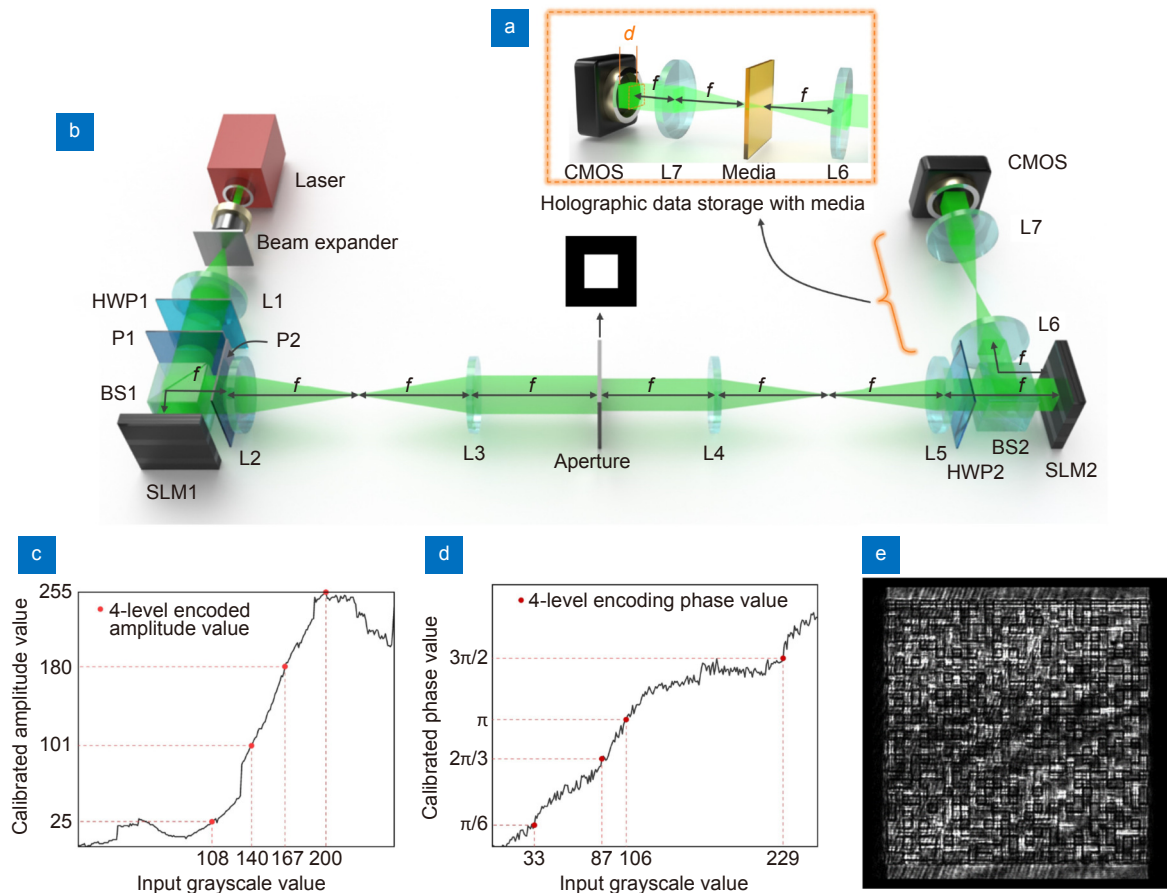**Fig. 7 | Architecture of the convolutional neural network.**

Fig. 8 | Entire process flow. The green and brown lines represent the training and testing processes, respectively.

Fig. 9 | (a) Recording and reading system of HDS. **(b)** Experimental setup. BS: beam splitter; HWP1, HWP2: half-wave plate; L1: collimating lens (f = 300 mm); L2–L7: relay lens (f = 150 mm); SLM1: amplitude-modulated spatial light modulator; SLM2: phase-modulated spatial light modulator; P1, P2: linear polarizer, P1 is horizontally polarized and P2 is vertically polarized. **(c)** Calibration curve of SLM1. **(d)** Calibration curve of SLM2. **(e)** Intensity image captured by the CMOS.

should be placed on the back focal plane of lens L6, as shown in Fig. 9(a). In this experiment, we did not introduce media. We researched the retrieval performance at different diffraction distances through simulated experiments and chose z=2 mm as the diffraction distance for the physical experiment. The reproduced beam obtained at the back focal plane of lens L7 continued to propagate for a distance z = 2 mm and the diffraction intensity was captured by CMOS (CHUM-131M-150, pixel pitch 4 μm). The calibration curves of the amplitude and phase SLM are shown in Fig. 9(c) and Fig. 9(d). The intensity image captured by CMOS is shown in Fig. 9(e).

Data preparation and neural network training

The complex amplitude data page and its corresponding diffraction intensity image are shown in Fig. 10. The data page was a 32 × 32 data matrix. Each data point was represented by 10 × 10 pixels of the SLMs. The pixel matrix of the amplitude and phase data page was 320 × 320. The physical size of the beam on SLM2 was 2.56 × 2.56 mm. Owing to the expansion of the propagating beam caused by diffraction, the linear size of the beam indicated on the CMOS was slightly larger than 2.56 mm, corresponding to the pixel matrix on the CMOS that was larger than 640 × 640. Therefore, we chose 768 × 768 of the diffraction intensity images and shrank it to 384 × 384 by down sampling. The amplitude and phase data pages with 320 × 320 pixels were enlarged to 384 × 384 pixels by padding zero, to maintain consistency with the diffraction intensity image. The amplitude and phase encoded pages were randomly generated and uploaded to the experimental system to obtain the corresponding diffraction intensity images. Approximately 11 h was consumed to capture 9731 intensity images. Among them, 8993 pairs of the intensity-amplitude dataset were used to train CNN1 and 8993 pairs of intensity-phase pages were used to train CNN2. A total of 738 pairs of images were used to test the generalization of the neural networks.
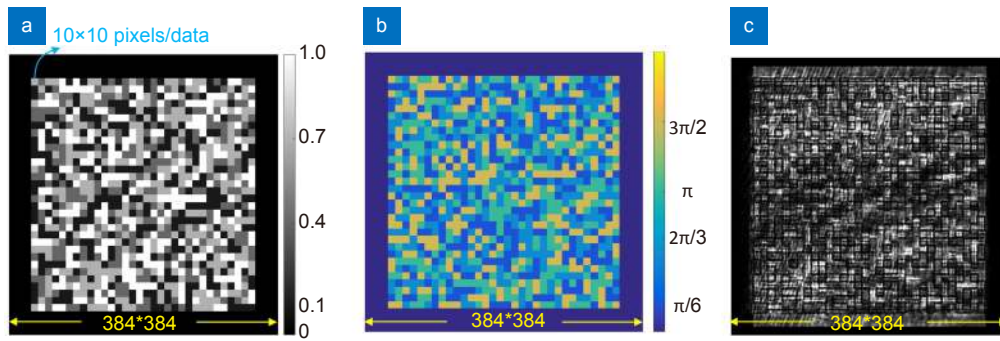
Fig. 10 | Complex-amplitude encoded data page and the intensity image captured by CMOS. (**a**) Data page uploaded on the amplitude SLM1. (**b**) Phase data page uploaded on the phase SLM2. (**c**) Diffraction intensity image captured by the CMOS at a diffraction distance of 2 mm.

The loss function $\mathcal{L}(\cdot)$ in Eq. (8) is defined as MSE, expressed as

$$MSE = \min \frac{1}{JWH} \sum_{j=1}^{J} \sum_{u=1}^{W} \sum_{v=1}^{H} \left(\tilde{C}_j(u,v) - C_j(u,v) \right),$$
(9)

where W and H are the width and height of the data page, respectively, and $J = 4$ is the minibatch size in the stochastic gradient descent (SGD) method[35]. $\tilde{C}(u,v)$ is the amplitude or phase data page predicted from the jth diffraction intensity image I_j, and $C(u,v)$ is the corresponding ground truth. The detailed hyperparameters of the neural network training are shown in Table 5. The program was implemented in Python 3.6 using PyTorch. NVIDIA Quadro RTX5000 was used to accelerate the computation. Approximately 13 h was consumed to optimize a single network.

Experimental results and discussion

Results
A diffraction image was randomly chosen from the test set to retrieve the complex amplitude using the trained CNN1 and CNN2. The predicted data pages and ground truth are shown in Fig. 11. Both the amplitude shown in Fig. 11(c) and the phase shown in Fig. 11(f) were retrieved from the diffraction intensity image in Fig. 11(a). Fig. 11(d) and Fig. 11(g) show the differences between the predicted results and the ground truth (amplitude in

Fig. 11(b) and phase in Fig. 11(e)), respectively. For the pixels with a small difference in the reconstructed image, a hard decision was used to correctly classify them. The points with a larger gap, shown as red data in Fig. 11(d) and Fig. 11(g), will be the error points after hard decision.

The marginal histogram of the total 1024 (32 × 32) complex-amplitude data points decoded from the retrieved amplitude and phase data pages in Fig. 11, is shown in Fig. 12. The data points are divided into 16 complex-amplitude categories (four-level amplitude by four-level phase) by hard decision. The data points in different categories are represented by different colored dots. The coordinates of each point represent the decoded values. Each arrow indicates the direction of error for an erroneous data point from the ground truth position to the retrieved position. For example, the blue arrow indicates that the retrieved value of a data point is (π/6, 0.45) whose ground truth is (π/6, 0.7). It will be categorized as error data after hard decision. The histogram indicates the distribution of the retrieved data. From the data point distribution and marginal histogram, the retrieved data are highly concentrated, and most of the data points are close to the ground truth. Only eight points (0.7%) on the complex-amplitude data page were decoded incorrectly.
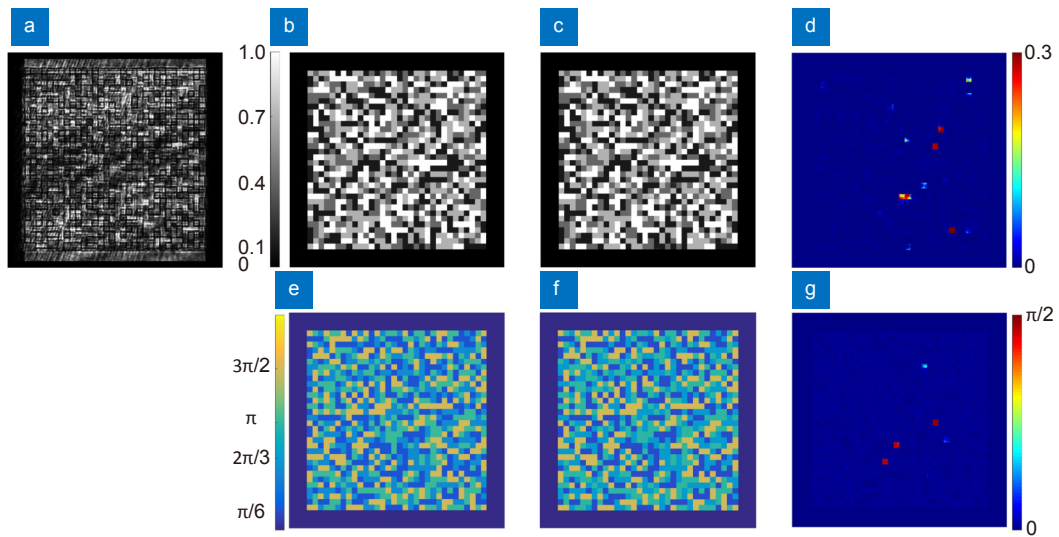
Discussion

Accuracy
To evaluate the accuracy of the proposed method, we

Table 5 | Neural network training hyperparameters.

Loss function	MSE
Training epochs	50
Learning rate	10^{-4}
Batch size	4
Optimizer	Adaptive moment estimation[36]
Active function	ReLU for hidden layer; Sigmoid for output layer[34]

Fig. 11 | Experimental results. (**a**) Diffraction intensity image at *z* = 2 mm. (**b**) Ground truth of amplitude data page. (**c**) Amplitude predicted by CNN1. (**d**) Difference between retrieved amplitude (c) and amplitude ground truth (b). (**e**) Ground truth of the phase data page. (**f**) Phase data page predicted by CNN2. (**g**) Difference between the retrieved phase data page (f) and the phase ground truth (e).
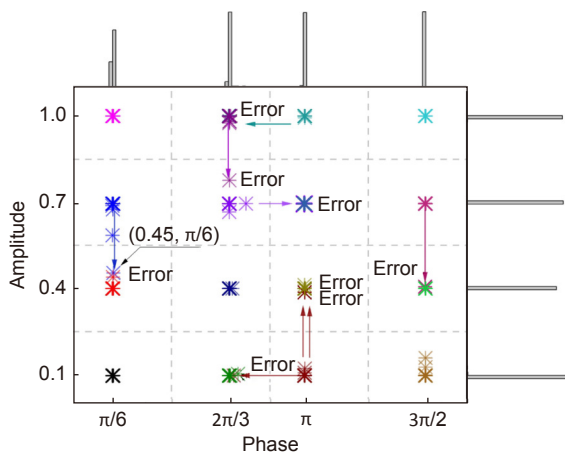
Fig. 12 | Distribution and margin histogram of the demodulated complex-amplitude signals. The horizontal axis represents the phase, and the vertical axis represents the amplitude. The histogram represents the distribution of the decoded data. The 16 colors represent the 16 complex-amplitude classifications. The arrows represent the direction of the wrong data point offset.

used the peak signal-to-noise ratio (PSNR)[37]:

$$PSNR = 10\log_{10}\cdot\left(\frac{(2^n-1)^2}{\frac{1}{WH}\sum_{u=1}^{W}\sum_{v=1}^{H}\left\|\tilde{C}_j(u,v)-C_j(u,v)\right\|^2}\right),$$

(10)

The structural similarity index (SSIM) is given by[38]

$$SSIM = \frac{(2\mu_{\tilde{C}}\mu_C+c_1)(2\sigma_{\tilde{C}C}+c_2)}{(\mu_{\tilde{C}}^2+\mu_C^2+c_1)(\sigma_{\tilde{C}}^2+\sigma_C^2+c_2)},$$

(11)

where $\mu_f, f \in \{C, \tilde{C}\}$, is the mean of the image *f*, and σ_f^2 is the variance; $\sigma_{\tilde{C}C}$ is the covariance of $\tilde{C}$ and C; c_1 and c_2 are the regularization parameters.

PSNR was used to evaluate the quality of the reconstructed image. A high PSNR value is preferred. A PSNR higher than 40 dB indicates that the image quality is considerably close to that of the original image. The SSIM was used to measure the structural similarity of the two images, and the best value was 1. The average PSNR and SSIM between the retrieved amplitude and phase data pages and the ground truth in the training and test sets, respectively, were calculated. The results are shown in Fig. 13. The retrieved image accuracy of the test set was lower compared with that of the training set, as expected based on the deep learning principles. The PSNR and SSIM of the phase test set were 41 and 0.997, respectively, better than the amplitudes of 34 PSNR and 0.993 SSIM.

Furthermore, to verify the robustness of the proposed method, all the 738 intensity images in the test set were demodulated. The calculated bit error rate (BER = number of error data/total number of data*100%) is plotted in Fig. 14. Figure 14(a) and Fig. 14(b) shows the BER distribution of the predicted amplitude and phase, respectively. The results show that the retrieved results have a certain degree of randomness, but the BERs of both the amplitude and phase data were below 3%. The average BER of the amplitude test data was 0.65%, and the phase was 0.5%. The phase-decoded data were relatively more accurate than the amplitude-decoded data.
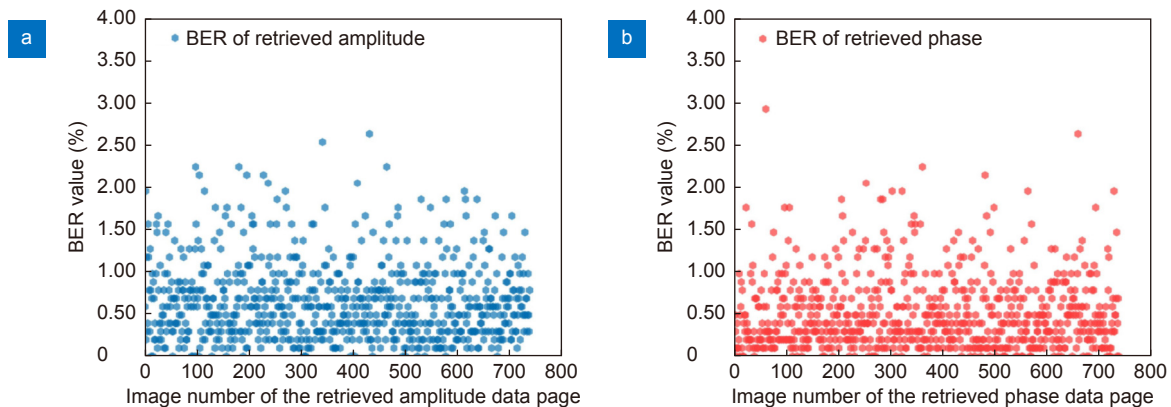
Error analysis

We retrieved all the intensity images in the training and
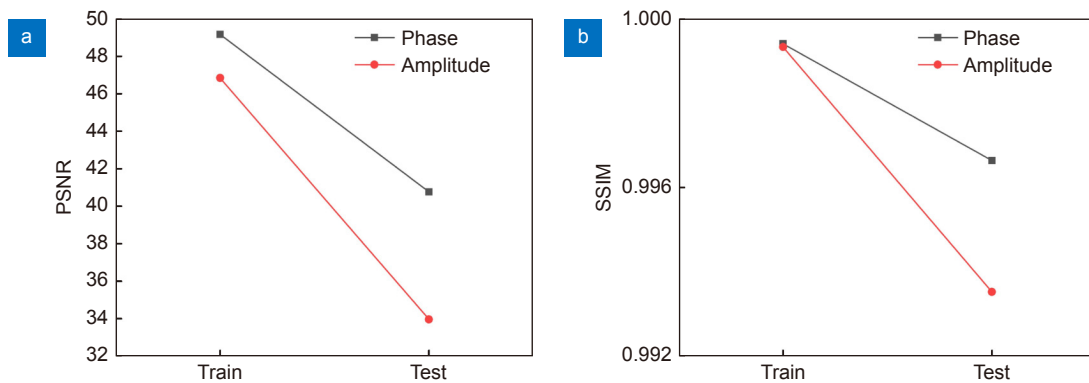
**Fig. 13 |** PSNR and SSIM of the retrieved phase and amplitude data pages in the training and test datasets: (**a**) PSNR and (**b**) SSIM.

Fig. 14 | BER distribution of all the complex-amplitude data pages in the test dataset. (**a**) Amplitude BER. (**b**) Phase BER.

test datasets and superimposed the error map into one image to show the spatial distribution of the error. The results are shown in Fig. 15. Row (I) in Fig. 15 presents the result of the training dataset and row (II) presents the test dataset. The 1st and 3rd columns in Fig. 15 represent the amplitude and phase accumulation of the pixel difference between the retrieved data pages and ground truth, respectively. A region of 3 × 3 points has been randomly selected, enlarged, and displayed. The grayscale distribution of the cross-section of the red line has been drawn. From the enlarged area in Fig. 15(a), within a data point, the error in the edge region of the retrieved amplitude data is larger, whereas the error in the middle region is smaller. The error for the retrieved phase data points in Fig. 15(c) is the opposite—higher in the middle and lower at the periphery. This is because the amplitude is derived from the original amplitude distribution features in the intensity map, whereas the phase is derived from the diffraction features generated by light diffraction. Near field diffraction enhances phase characteristics and weakens amplitude characteristics. This enhancement of phase-related features and the weakening of amplitude-related features mainly occur in the regions adjacent to the data points (refer to Fig. 3 and Fig.

4). The error generated can be explained based on the principle of complex amplitude retrieval proposed in this study and cannot be eliminated. However, it does not generate bit errors after hard decision decoding. The 2nd and 4th columns in Fig. 15 show the error data points distributions of the whole demodulated amplitude and phase page, respectively. From Fig. 15(f) and 15(h), both the errors of the amplitude and phase data are higher in the central area of the data page. The error has a strong relationship with the noise distribution, as observed in the intensity map shown in Fig. 15(e). Owing to the influence of uneven light beams, interference fringes generated by the multi-reflection of the optical system, and thermal noise of the optical devices, the captured intensity map shows uneven brightness and darkness. This noise has a greater influence on amplitude retrieval. This error is due to the experimental noise and can be reduced by improving the experimental accuracy.

Conclusion

In this paper, a complex amplitude demodulation method based on deep learning is proposed that can retrieve both the amplitude and phase from a single-shot near
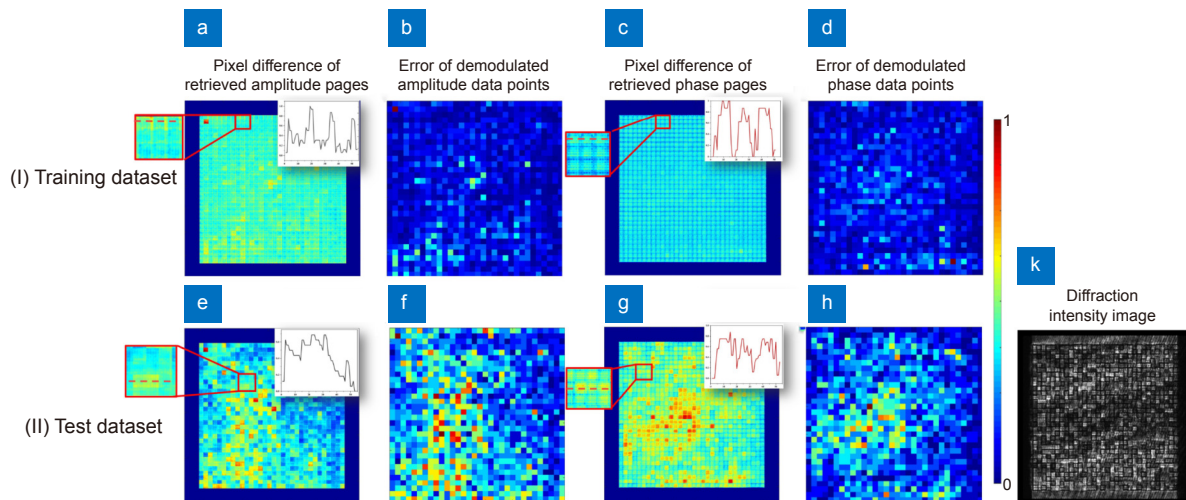
Fig. 15 | Superposition distribution of errors of all retrieved images in the data set. (**a**) Pixel differences of the amplitude training set. (**b**) Error data of the decoded amplitude pages in the training set. (**c**) Pixel differences of the phase training set. (**d**) Error data of decoded phase pages in the training set. (**e**) Pixel differences of the amplitude test set. (**f**) Error data of the decoded amplitude pages in the test set. (**g**) Pixel differences of the phase test set. (**h**) Error data of the decoded phase pages in the test set. (**k**) Intensity image under the influence of noise.

field diffraction intensity image. By analyzing the correlation between the near field diffraction and the encoded data pages, the inverse problem of solving the complex amplitude from the intensity map is decomposed into two inverse problems of retrieving the amplitude and phase and represented by CNN respectively. After the CNNs' training is completed, the amplitude and phase data pages can be directly reconstructed from a diffraction intensity image. To the best of our knowledge, this is the first study to have retrieved the complex amplitude data page from a single intensity image using a non-interferometric system and verified it experimentally, relative to the current demodulation technique in HDS. In addition, the study of diffraction features based on the encoding rule provides a new perspective on the research of deep learning applications in the field of computational imaging. As an end-to-end deep learning method, in actual use, a large amount of experimental data is still needed to be collected to train the neural network. However, we found that by exploiting the relationship of the diffraction pattern to the encoded data, the amount of data required to train the network can be further reduced. In the future, we will continue to study how to train the neural network with a small amount of training data by designing the structure of the encoded data page.

References

1. Reinsel D, Gantz J, Rydning J. *The Digitization of the World from Edge to Core* (International Data Corporation, Framingham, 2018).

2. Flexible, scalable and reliable storage solution. Panasonic Connect. https://panasonic.net/cns/archiver/concept/

3. Anderson P, Black R, Cerkauskaite A, Chatzieleftheriou A, Clegg J et al. Glass: a new media for a new era?. In *Proceedings of the 10th USENIX Workshop on Hot Topics in Storage and File Systems (HotStorage 2018)* (USENIX Association, 2018).

4. Zhang JY, Gecevičius M, Beresna M, Kazansky PG. Seemingly unlimited lifetime data storage in nanostructured glass. *Phys Rev Lett* **112**, 033901 (2014).

5. Dhar L, Curtis K, Fäcke T. Coming of age. *Nat Photonics* **2**, 403–405 (2008).

6. Lin X, Liu JP, Hao JY, Wang K, Zhang YY et al. Collinear holographic data storage technologies. *Opto-Electron Adv* **3**, 190004 (2020).

7. Horimai H, Tan XD, Li J. Collinear holography. *Appl Opt* **44**, 2575–2579 (2005).

8. Project HSD: holographic storage device for the cloud. Microsoft. https://www.microsoft.com/en-us/research/project/hsd/

9. Liu JP, Zhang L, Wu AN, Tanaka Y, Shigaki M et al. High noise margin decoding of holographic data page based on compressed sensing. *Opt Express* **28**, 7139–7151 (2020).

10. Katano Y, Muroi T, Kinoshita N, Ishii N, Hayashi N. Data demodulation using convolutional neural networks for holographic data storage. *Jpn J Appl Phys* **57**, 09SC01 (2018).

11. Shimobaba T, Kuwata N, Homma M, Takahashi T, Nagahama Y et al. Convolutional neural network-based data page classification for holographic memory. *Appl Opt* **56**, 7327–7330 (2017).

12. Lin X, Huang Y, Shimura T, Fujimura R, Tanaka Y et al. Fast non-interferometric iterative phase retrieval for holographic data storage. *Opt Express* **25**, 30905–30915 (2017).

13. Hao JY, Wang K, Zhang YY, Li H, Lin X et al. Collinear non-interferometric phase retrieval for holographic data storage. *Opt Express* **28**, 25795–25805 (2020).

14. Lin X, Hao JY, Wang K, Zhang YY, Li H et al. Frequency expanded non-interferometric phase retrieval for holographic data storage. *Opt Express* **28**, 511–518 (2020).

15. Lin X, Huang Y, Li Y, Liu JY, Liu JP et al. Four-level phase pair

encoding and decoding with single interferometric phase retriev-al for holographic data storage. *Chin Opt Lett* **16**, 032101 (2018).

16. Nobukawa T, Nomura T. Multilevel recording of complex amp-litude data pages in a holographic data storage system using di-gital holography. *Opt Express* **24**, 21001–21011 (2016).

17. Katano Y, Nobukawa T, Muroi T, Kinoshita N, Ishii N. CNN-based demodulation for a complex amplitude modulation code in holographic data storage. *Opt Rev* **28**, 662–672 (2021).

18. Bunsen M, Tateyama S. Detection method for the complex amplitude of a signal beam with intensity and phase modulation using the transport of intensity equation for holographic data storage. *Opt Express* **27**, 24029–24042 (2019).

19. Chen RX, Hao JY, Yu CY, Zheng QJ, Qiu XY et al. Dynamic sampling iterative phase retrieval for holographic data storage. *Opt Express* **29**, 6726–6736 (2021).

20. Horisaki R, Fujii K, Tanida J. Single-shot and lensless complex-amplitude imaging with incoherent light based on machine learning. *Opt Rev* **25**, 593–597 (2018).

21. Zuo HR, Xu ZY, Zhang JL, Jia G. Visual tracking based on transfer learning of deep salience information. *Opto-Electron Adv* **3**, 190018 (2020).

22. Liao MH, Zheng SS, Pan SX, Lu DJ, He WQ et al. Deep-learn-ing-based ciphertext-only attack on optical double random phase encryption. *Opto-Electron Adv* **4**, 200016 (2021).

23. Liao K, Chen Y, Yu ZC, Hu XY, Wang XY et al. All-optical com-puting based on convolutional neural networks. *Opto-Electron Adv* **4**, 200060 (2021).

24. Sinha A, Lee J, Li S, Barbastathis G. Lensless computational imaging through deep learning. *Optica* **4**, 1117–1125 (2017).

25. Wang H, Lyu M, Situ G. eHoloNet: a learning-based end-to-end approach for in-line digital holographic reconstruction. *Opt Ex-press* **26**, 22603–22614 (2018).

26. Wang KQ, Dou JZ, Kemao Q, Di JL, Zhao JL. Y-Net: a one-to-two deep learning framework for digital holographic reconstruc-tion. *Opt Lett* **44**, 4765–4768 (2019).

27. Wang KQ, Kemao Q, Di JL, Zhao JL. Y4-Net: a deep learning solution to one-shot dual-wavelength digital holographic recon-struction. *Opt Lett* **45**, 4220–4223 (2020).

28. 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).

29. Situ G. Deep holography. *Light Adv Manuf* **3**, 278–300 (2022).

30. Hao JY, Lin X, Lin YK, Song HY, Chen RX et al. Lensless phase retrieval based on deep learning used in holographic data stor-age. *Opt Lett* **46**, 4168–4171 (2021).

31. Goodman JW. *Introduction to Fourier Optics* 2nd ed (McGraw-Hill, Singapore, 1996).

32. Tokoro M, Fujimura R. Single-shot detection of four-level phase modulated signals using inter-pixel crosstalk for holographic data storage. *Jpn J Appl Phys* **60**, 022004 (2021).

33. Ronneberger O, Fischer P, Brox T. U-Net: convolutional net-works for biomedical image segmentation. In *Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention* 234–241 (Springer, 2015); http://doi.org/10.1007/978-3-319-24574-4_28.

34. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* **521**, 436–444 (2015).

35. Ferguson TS. An inconsistent maximum likelihood estimate. *J Am Stat Assoc* **77**, 831–834 (1982).

36. Kingma DP, Ba J. Adam: a method for stochastic optimization. In *Proceedings of the 3rd International Conference on Learning Representations* (2015).

37. Korhonen J, You JY. Peak signal-to-noise ratio revisited: is simple beautiful?. In *Proceedings of the Fourth International Workshop on Quality of Multimedia Experience* 37–38 (IEEE, 2012); http://doi.org/10.1109/QoMEX.2012.6263880.

38. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality as-sessment: from error visibility to structural similarity. *IEEE Trans Image Process* **13**, 600–612 (2004).

Acknowledgements

Author contributions

J. Y. Hao, X. Lin and M. Y. Chen conceived the idea; J. Y. Hao, Y. K. Lin and R. X. Chen performed the numerical simulations and conducted the experi-ments, guided by discussions with X. Lin and X. D. Tan; J. Y. Hao, X. Lin, X. D. Tan and G. H. Situ analyzed the results; J. Y. Hao, G. H. Situ, X. Lin and X. D. Tan wrote the manuscript; H. Horimai and X. D. Tan supervised the project.

Competing interests

The authors declare no competing financial interests.