A Deep-Learning Approach for Low-Spatial-Coherence Imaging in Computer-Generated Holography

Xin Tong, Renjun Xu,* Ke Liu, Liangliang Zhao, Weilai Zhu, and Daomu Zhao*

The low-spatial-coherence imaging capability of computer-generated holography (CGH) is a key to high-resolution displays, virtual reality, augmented reality, and holographic microscopy. The low spatial coherence caused by complex disturbances can damage the image quality irreversibly. The optical field with low spatial coherence has large fluctuations, making it difficult to be quantified and modeled directly. To tackle these challenges, a deep neural network-based model U-residual dense network (U-RDN) is proposed, which obtains the optimal solution under the low-spatial-coherence condition. The large-scale images are generated using optical experiments, with analysis and restoration by deep learning. Extensive experiments demonstrate the strong out-of-distribution robustness of U-RDN, which is generalizable to unseen classes in unseen domains. The learning-based approach and low-spatial-coherence dataset open a new path toward the next generation of CGH.

1. Introduction

Computer-generated holography (CGH) is extensively used in modern optical imaging, which was first proposed by Brown and Lohmann in 1966. They utilized three different forms of circuitous phase encoding to generate computational holograms.[1,2] In 1967, Paris applied the fast-Fourier transform algorithm to calculate the hologram, which greatly reduced the calculation time.[3] Compared with conventional optical holography, CGH discards the complex and precise optical structure design. It records phase information through computer coding, which is not limited to actual objects and has the advantages of convenient fabrication and simple optical path realization.

Therefore, it has been widely studied and applied in 3D imaging, wavefront control, interferometry,[4–6] etc. Computational holograms include amplitude holograms and phase holograms.[7] Phase holograms have higher contrast and do not generate conjugated images. The current mainstream algorithms for phase holograms mainly include the Gerchberg–Saxton (GS) algorithm,[8] Yang–Gu algorithm,[9] transport of intensity equation,[10] deep learning,[11] etc. These algorithms are designed to optimize holographic fringe patterns which in turn improve imaging quality.[12] However, these front-end algorithms are often limited by the coherence of the light source. Until today, it remains as an unsolved challenge for CGH to present high images quality with low spatial coherence.

The coherence of the light source plays a crucial role in the imaging system. High-coherence laser will generate severe speckles due to interference, which seriously damages the image quality and even causes damage to the human eyes.[13] Therefore, reducing the coherence of the light source is an effective way to eliminate the speckle and improve imaging quality.[14,15] In 1938, Zernike researched the coherence optical field.[16] In the 1950s, Wolf introduced the transmission law for the correlation function in the space–time domain and the space–frequency domain,[17,18] and in 1978 he introduced the Gaussian Schell mode beam.[19] These theories are based on partially destroyed coherence, which is defined as partial coherence. In subsequent studies, partially coherent beams have been proved to resist atmospheric turbulence, making them the main choice in complex imaging environments.[20,21] Scholars have recently been experimenting with the partially coherent beam for CGH imaging. This technique can weaken the speckle effect and improve image quality, but sacrifice sharpness and details, resulting in unnecessary blur and contrast reduction.[22,23] It is essential for CGH imaging that the light source has high coherence in both horizontal and vertical directions; otherwise, the image quality will be irreversibly deteriorated.[24] Because of high fluctuations, low-spatial-coherence imaging is difficult to be described quantitatively by physical models. There is a consensus that CGH imaging cannot work in low spatial coherence. In previous studies, few scholars studied how to reconstruct images at low spatial coherence, because it is an irreversible problem. Therefore, it is crucial to deal with these difficulties, which can help us manage the challenges posed by complex environments. Here we try to use a deep learning method to learn and build this model and calculate its optimal solution.

X. Tong, L. Zhao, W. Zhu, D. Zhao
Zhejiang Province Key Laboratory of Quantum Technology and Device
School of Physics
Zhejiang University
Hangzhou 310027, P. R. China
E-mail: optics@zju.edu.cn

X. Tong, R. Xu, K. Liu
Center for Data Science
Zhejiang University
Hangzhou 310027, P. R. China
E-mail: rux@zju.edu.cn

In recent years, the application of deep learning in the optical imaging field has attracted enough attention. Models based on neural networks (NN) and generative adversarial networks (GAN) have shown higher computational efficiency and accuracy than traditional methods and have achieved great success in biomedical imaging, super-resolution, scattering imaging, phase reconstruction, image masking,[25–32] etc. Deep learning is a data-driven method that relies on learning a large amount of data to find an optimal solution. It has strong flexibility and adaptability when dealing with complex models and environments. The way of simulating and computing in deep learning opens up a new direction for optical imaging, because it can optimize or even eliminate some irresistible errors by a well-trained NN. There are different layers in the computing architecture of NN, and each layer contains different activation functions and is interconnected. The purpose is to learn the mapping relationship between input and output. Finally, backpropagation updates network parameters to minimize the discrepancy between the output vector and the target vector.

We propose U-residual dense network (U-RDN) which nests the residual neural network[33] and Densenet[34] into the U-net[35] framework. Note that deep neural networks have been criticized for lacking out-of-distribution (OOD) robustness in AI community.[36] Besides physical significance, out-of-domain robustness is of crucial importance. Therefore, our model is based on the physics informed of end-to-end engines which combine physics priors and network models perfectly. The combination of physical hardware and neural network helps improve the OOD robustness of the model. Physics informed means that the input data of NN comes from physical models and experiments rather than purely relying on computer simulations, so it can be applied in practice rather than just in theoretical simulation. We use the controllable partially coherent beam (Gaussian Schell mode) as the light source in the optical design. During the model training phase, we used the ImageNet[37] dataset containing natural images, as its images cover a wide range and are suitable for generalization. For the purpose of demonstrating the generalization of the network, we select data that is completely disjoint from the training set as the test set, which has never been trained on the network. Extensive experiments demonstrate that the proposed model performs well in unseen test datasets and can effectively improve the imaging quality at low spatial coherence.

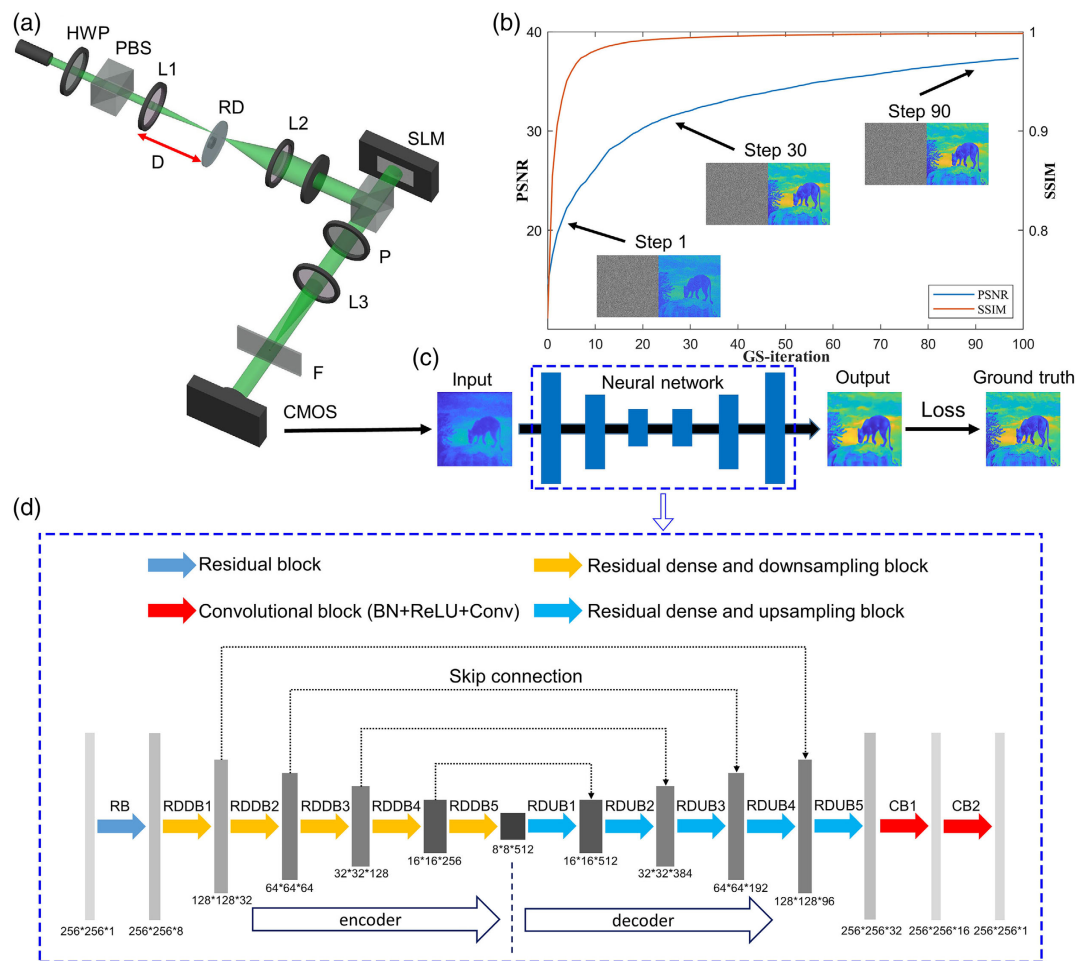## 2. Experimental Section

2.1. Optical Experiment

The proposed architecture of the optical equipment and the associated neural network is illustrated in **Figure 1**. We generated partially coherent beams by rotating diffuser (RD). It is worth mentioning that Peng[38] recently used superluminescent light-emitting diodes (sLED) to improve the imaging quality. The sLED coupled into a single-mode fiber belongs to partially spatial coherence, and we cannot change its spatial coherence. The RD can be used to generate spatial coherent beams with controllable coherence, so it can help us investigate imaging questions at different spatial coherence. In the optical equipment, a quasimonochromatic high-coherent beam from the solid-state laser

(CNI, MLL-FN, 532 nm) first passed through a half-wave plate (HWP) and a polarized beam splitter (PBS). As one aspect, the beam intensity can be adjusted, while as another, the polarization direction of the beam can be changed to horizontal, which was used to match the modulation mode of the spatial light modulator (SLM, PLUTO). The distance between the converging lens L1 ($f_1 = 150$ mm) and the diffuser (DHC, GCL-201) was D, which was used to control the incident beam spot size. The DC motor drove the diffuser to rotate, and the speed can be controlled by changing the input voltage. After collimating the scattered light from the diffuser by the collimating lens L2 ($f_2 = 150$ mm), it became the light source for computational holographic imaging. The light source was partially spatially coherent, belonging to the Schell mode. After adjusting the beam radius with the aperture, it was incident to the SLM through the beam splitter (BS). As the beam modulated by SLM passes through BS, it found its way to the linear polarizer and Fourier lens L3 ($f_3 = 150$ mm). Finally, the formed images were captured by Complementary metal–oxide–semiconductor (CMOS) (Sony, E3ISPM). Between L3 and CMOS, we placed a high-pass filter at a specific location to remove the unmodulated zeroth-order diffracted light.

2.2. Data Acquisition

The modulation mode and performance of the SLM depend on the polarization properties of the incident beam and the internal response curve. Specifically, we set the incident beam as the horizontal linear polarization and set the polarization direction of the polarizer to 30°. Experiments verify that this setup can achieve approximately pure-phase modulation of the SLM and satisfy the linear relationship between phase and grayscale (8-bit) (see Supporting Information). The SLM resolution was 1920×1080, and the pixel pitch was 6.4 μm. The center 512×512 was used as the effective area, and the radius of the incident beam was set to 5 mm by the aperture. If the beam radius is too large, the generation of zero-order diffracted light will be intensified. The kinoform was generated by the iterative GS algorithm. Figure 1b illustrates the curve of PSNR and SSIM with the number of iterations. The reconstruction performance was improved with more iterations. Due to the modulation efficiency problem in the internal structure of the SLM, we added the spherical wave phase to the original kinoform, $\phi_s(x, y) = -\frac{k}{2r}(x^2 + y^2)$, where k is the wavenumber and r is the distance from the SLM to the Fourier lens L3. Here we set $r = 200$ mm, the Fourier reconstruction plane and the focal plane were separated along the optical axis. A high-pass filter was placed at the focal plane of the Fourier lens L3, blocking the zero-order diffracted light that was not modulated. According to the equations of the literature,[39] the distance between the reconstruction plane and the focal plane was 60 mm. Images captured by CMOS at the reconstruction plane were used as input of the U-RDN and then processed and calculated on the Intel I7 CPU and NVIDIA GTX3080Ti graphics card.

The light spot size on the diffuser can control the spatial coherence of the light source (see "Experimental Section" for the Theory of Coherence). We measured the spatial coherence

Figure 1. Experimental equipment and data processing. a) The CGH imaging system. HWP, half-wave plate; PBS, polarized beam splitter; L, lens; RD, rotating diffuser; SLM, spatial light modulator; P, polarizer; F, filter. b) Reconstruction quality curves of kinoforms with different iterations. c) The general process of network training. d) Detailed architecture of U-RDN. Arrows of different colors indicate different operation steps.

of three light sources (see Supporting Information). $D = f_1$, the spatial coherence was properly destroyed, and the coherence was about 0.48. $D = 1.5f_1$, the spatial coherence was about 0.3. $D = 2f_1$, the spatial coherence was severely damaged, and the coherence was only about 0.14. The grit size and rotational speed of the diffuser slightly affected the coherence.[40] To ensure a single variable in the experiment, the speed of RD was controlled at 200 rpm, and the grit size of the diffuser was 600 grid. Spatial coherence can be controlled by changing the distance D between L1 and RD. Therefore, we can acquire the training data and test data with different spatial coherence. For the training set, we only selected ImageNet as the original data. The 8000 ImageNet samples were converted into kinoforms by the GS algorithm, which were input into the modulated SLM. The training data consisted of images captured by CMOS. The captured images and original samples were resized to a suitable size by an adjacent pixel averaging algorithm, those were used as the input and ground truth of the network respectively. For the test set, we selected five different datasets, CelebFaces attributes (CelebA),[41] microsoft common objects in context (COCO),[42] labeled faces in the wild (LFW),[43] ImageNet, PASCAL visual object classes challenge

(VOC),[44] and thepreprocessing method was the same as the training set. All data were processed in the same way.

2.3. U-RDN Architecture

The proposed U-RDN is based on the U-net architecture, and the transfer between layers relies on residual dense blocks (RDB) to achieve. RDB consists of residual and dense networks, making full use of the hierarchical features of all convolutional layers. It adaptively learns hierarchical features through local feature fusion, thereby stabilizing the training. After sufficient dense local features are obtained, local residual fusion is used to combine and adaptively learn hierarchical features in a holistic manner. A continuous memory mechanism is formed between adjacent RDBs, which can connect the state of the previous block to all levels of the current RDB.[45] Therefore, RDB has better-learned features and stronger generalization ability.

Figure 1d details the U-RDN internal architecture, which follows the encoder–decoder framework. The image captured and processed by CMOS was used as the input of the network. After preprocessing a standard residual block (RB) with a convolution

kernel of 3×3, it went through five residual dense downsampling blocks (RDDB) and five residual dense upsampling blocks (RDUB), respectively. Finally, the network estimation output was generated by two standard convolutional blocks (CB). The residual dense structure fully fused the features of the previous layer and extracted feature maps through downsampling transition blocks to encode the image. In the same way, the residual dense structure and the upsampling transition block performed pixel regression on the previous layer to decode the image. In addition, we used skip connections to connect layers of different spatial scales to achieve retention and transfer of high-frequency information. More details about all blocks and working mechanisms are described in Supporting Information.

A recent study demonstrated that the NPCC can effectively promote sparse prediction.[46] After testing and comparison, it is found that NPCC can improve the quality of image restoration under low spatial coherence (see Supporting Information). Therefore, we designed different loss functions in different scenarios.

$$\text{Loss} = \begin{cases} \text{MSE} & D = f_1, D = 1.5f_1 \\ \text{MSE} + \text{NPCC} & D = 2f_1 \end{cases}, \tag{1}$$

$$\text{MSE}(O, G) = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} (O(i,j) - G(i,j))^2 \tag{2}$$

$$\text{NPCC}(O, G) =$$
$$- \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} (O(i,j) - \tilde{O})(G(i,j) - \tilde{G})}{\sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} (O(i,j) - \tilde{O})^2} \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} (G(i,j) - \tilde{G})^2}} \tag{3}$$

where m and n are the width and height of the image. O means the output of the network, G denotes the ground truth. $\tilde{O}$ and $\tilde{G}$ are the mean value of output and ground truth, respectively. All training and testing stages are placed on the GPU.

2.4. Theory of Coherence

In optical experiments, we utilized RDs to generate spatial coherent light. After the quasimonochromatic high-coherence laser entered the diffuser, it was randomly scattered and lost its original amplitude and phase distribution. When the diffuser was rotated, the amplitude and phase of the beam were changed randomly, which made the beam completely incoherent light at the output surface of the diffuser. According to Zernike's theorem, the partially coherent beam will be generated during the subsequent transmission. Here we used a lens with a focal length of 150 mm (L2) to collimate it. From the RD to the lens, the kernel function $H(r, \rho)$ was regarded as the response function between the incident surface and the exit surface.

$$H(r, \rho) = -\frac{i}{\lambda f} \exp \left[\frac{i\pi}{\lambda f} (r^2 - 2\rho r) \right] \tag{4}$$

Where r and ρ represent the vector coordinates of the incident and exit surfaces, respectively. For quasimonochromatic light, we usually use the cross-spectral density function to characterize the second-order statistical properties of the optical field.

$$W(\rho_1, \rho_2) = \int I(r) H_0^*(\rho_1, r) H_0(\rho_2, r) d^2 r \tag{5}$$

where $I(r)$ means the intensity distribution at the incident surface. By substituting the above formula, the cross-spectral density function of the exit surface can be obtained.

$$W(\rho_1, \rho_2) = \int I(r) \exp \left[-\frac{i2\pi r}{\lambda f} (\rho_1 - \rho_2) \right] d^2 r \tag{6}$$

We know that the spectral coherence between two points

$$\mu(\rho_1, \rho_2) = \frac{W(\rho_1, \rho_2)}{\sqrt{W(\rho_1, \rho_1)} \sqrt{W(\rho_2, \rho_2)}} \tag{7}$$

Therefore, we can change the spatial coherence of the beam by adjusting the intensity distribution of the beam on the diffuser.
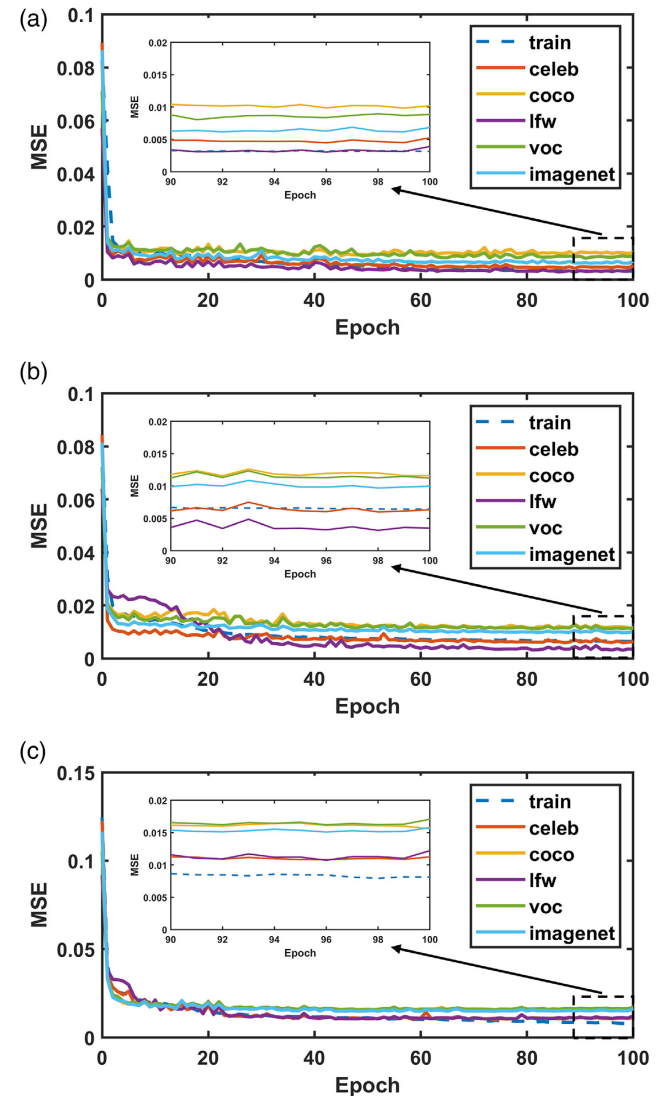
Figure 2. Error curves of training and test sets versus training epochs with a) $D = f_1$, b) $D = 1.5f_1$, and c) $D = 2f_1$. D is the distance between the converging lens L1 ($f_1 = 150$ mm) and the diffuser.

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
PHOTONICS
RESEARCH**
Open Access

www.adpr-journal.com

In addition, the wavelength of the incident beam and the focal length of the collimating lens enables control over the scaling of the coherence distribution.
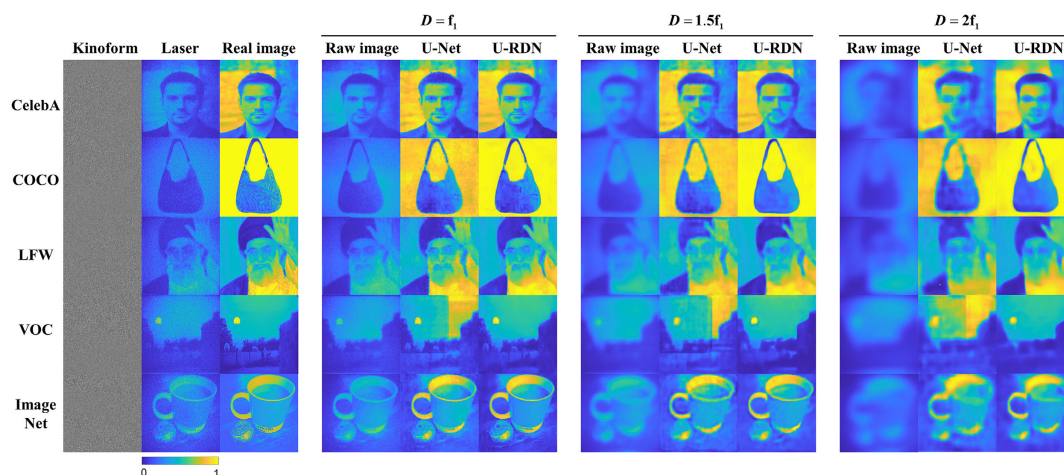
3. Results

3.1. Experimental Evaluation

The captured images and ground truth are first uniformly downscaled to 256×256 pixels, which is consistent with the input and output size of the network and reduces the network parameters and computing memory. Here we have prepared three types of data, representing different spatial coherence in the CGH imaging. These data will be used to train the network parameters. The training process was optimized by the loss function between the network output and the ground truth. The mean-squared loss (MSE) results for different training epochs are shown in **Figure 2**. Every training epoch consists of 8000 ImageNet data, and the network is tested with five different datasets that have never been trained. It can be seen that the error of the network on the training sets and the test sets simultaneously decrease as the number of training increases, indicating that it eventually converges. There is no overfitting state during the entire training process. We found that the LFW dataset performed better than the training set. This is simply because the training set contains other classes with poor performance, resulting in poor average precision. Instructions and training logs about other parameters, peak signal-to-noise ratio (PSNR), structural similarity index (SSIM), and negative Pearson correlation coefficient (NPCC), are described in the Supporting Information.

The qualitative output results of each test set with different spatial coherence are shown in **Figure 3**. In addition, we also show the reconstructed results of U-Net,[35] where all training and test data are identical to the U-RDN. Note that we modified the size of the input and output layer of the U-net model to match the dimensions of the dataset. It can be seen obviously

from Figure 3 that the image quality is irreversibly destroyed as spatial coherence continues to decrease; even we cannot recognize the details in the original image. Compared with the results reconstructed by the U-Net, our U-RDN model can not only reconstruct deteriorative images with low spatial coherence but also remove noise and make the image smoother, which further improves image quality. When the coherence of the light source is not seriously damaged, the recovered image quality has been greatly improved, and the blur and speckle effects have disappeared, almost identical to the real picture. In contrast, the U-Net fails to reconstruct images with low spatial coherence and suffers from severe intensity distortion when reconstructing the VOC sample. To quantitatively evaluate the quality of reconstruction images with different spatial coherence, **Table 1** shows the average quality evaluation index of the original image and the reconstructed image compared to the ground truth: MSE and SSIM. The evaluation results for other parameters (PSNR and NPCC) are shown in the supplementary material. Based on the detailed data in Table 1 and the quantitative results in Figure 3, we fully demonstrate that our U-RDN model can disrupt the limitations of low-coherence imaging. The average values of four evaluation indices for each test set with different spatial coherence are shown in **Figure 4**. Here we prepared all original test datasets called "None" as a control group in advance. It can be seen that each evaluation index of images has risen significantly compared to the control group after being processed by U-RDN, which means that image quality has improved dramatically. For example, the facial features are recovered more distinct in the face test set than original images. Even image features with poor sparsity can be recovered to the basic contour. It is worth mentioning that we did not select a dataset with large sparsity like MNIST for testing. Our goal is to achieve image recovery with low spatial coherence under more general conditions. When faced with unseen images, both qualitative and quantitative results show that our model can perform well and generalize across different domains, not just single-pattern matching.


Figure 3. Qualitative results of U-RDN and U-Net outputs with different spatial coherence. The kinoforms obtained by the GS algorithm are imported to SLM, and the iteration step is 100. Each row in the group includes original images captured by CMOS and reconstructed results by different methods at different special coherence. Where $D = f_1$, coherence is 0.48. $D = 1.5f_1$, coherence is 0.3. $D = 2f_1$, coherence is 0.14. These test data are never been trained by the network. D is the distance between the converging lens L1($f_1 = 150$ mm) and the diffuser.

Table 1. Quantitative results on the evaluation indicators MSE and SSIM for each original and reconstructed image compared to the ground truth with different spatial coherence. Co means the coherence of the light source.

Evaluation index[a]	MSE					SSIM				
Methods\Example	CelebA	COCO	Image Net	LFW	VOC	CelebA	COCO	Image Net	LFW	VOC
Laser	0.1110	0.4266	0.0224	0.0919	0.0184	0.3849	0.1958	0.4297	0.3509	0.4570
$D=f_1$, Co = 0.48	0.0917	0.3935	0.0212	0.0641	0.0201	0.6200	0.3548	0.6000	0.6151	0.6439
f_1-output(U-RDN)	**0.0010**	**0.0022**	**0.0011**	**0.0014**	**0.0013**	**0.8994**	**0.8207**	**0.8637**	**0.8528**	**0.8424**
f_1-output(U-Net)	0.0031	0.0227	0.0049	0.0034	0.0170	0.8297	0.6966	0.7788	0.7835	0.6814
$D=1.5f_1$, Co = 0.3	0.0866	0.3714	0.0177	0.0714	0.0139	0.6131	0.4200	0.6326	0.5909	0.7030
$1.5f_1$-output(U-RDN)	**0.0017**	**0.0037**	**0.0024**	**0.0018**	**0.0020**	**0.8483**	**0.7907**	**0.7866**	**0.8089**	**0.8116**
$1.5f_1$-output(U-Net)	0.0054	0.0252	0.0052	0.0089	0.0156	0.7614	0.7218	0.7001	0.6713	0.6884
$D=2f_1$, Co = 0.18	0.1299	0.4732	0.0425	0.1036	0.0268	0.5000	0.3130	0.5264	0.5044	0.6734
$2f_1$-output(U-RDN)	**0.0038**	**0.0066**	**0.0043**	**0.0055**	**0.0027**	**0.7679**	**0.7724**	**0.7335**	**0.7479**	**0.7761**
$2f_1$-output(U-Net)	0.0150	0.0384	0.0103	0.0148	0.0133	0.6627	0.6995	0.6491	0.6844	0.7139
Ground truth	0	0	0	0	0	1	1	1	1	1

[a] Bold values indicate the highest quantitative metrics.
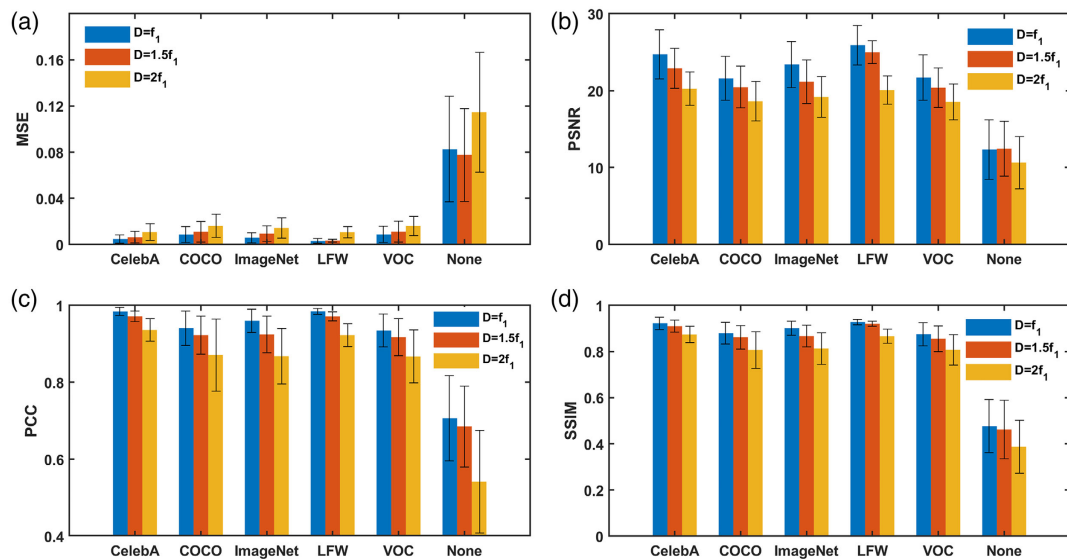
Figure 4. Average evaluation index of U-RDN output results on different test sets. These test sets are never trained. None represents all original test datasets that have not been trained by the network.

3.2. Resolution and Robustness Evaluations

In this section, we design special experiments to test the spatial resolution of the trained U-RDN. The image quality and resolution decrease rapidly as the spatial coherence of the light source decreases. After being processed by trained U-RDN, the image quality and resolution are much improved. To evaluate the spatial resolution of the network, we designed special parallel line pairs. The line pairs are made up of rectangles whose dimensions are 10×400 px, and their distance is measured in units of 5 px, gradually increasing from 10 to 50 px. The subsequent processing steps for the fringe pattern are the same as those described in

Section 2. Stripe images with different spatial coherences are captured by CMOS, and then we input them to the trained U-RDN to generate restored images (Note that the stripe images are never been trained by network). The results are shown in **Figure 5**. Row(a) shows the resolution test results with a moderate reduction in the spatial coherence of the light source. Compared with the results of laser imaging, it can be concluded that the spatial partially coherent light source can effectively reduce the noise effect, and after network training, the imaging sharpness can be further improved. It also can be seen from the intensity curve that the curve has changed from rough (laser) to smooth and even close to the true value. Rows (a)–(c) show the
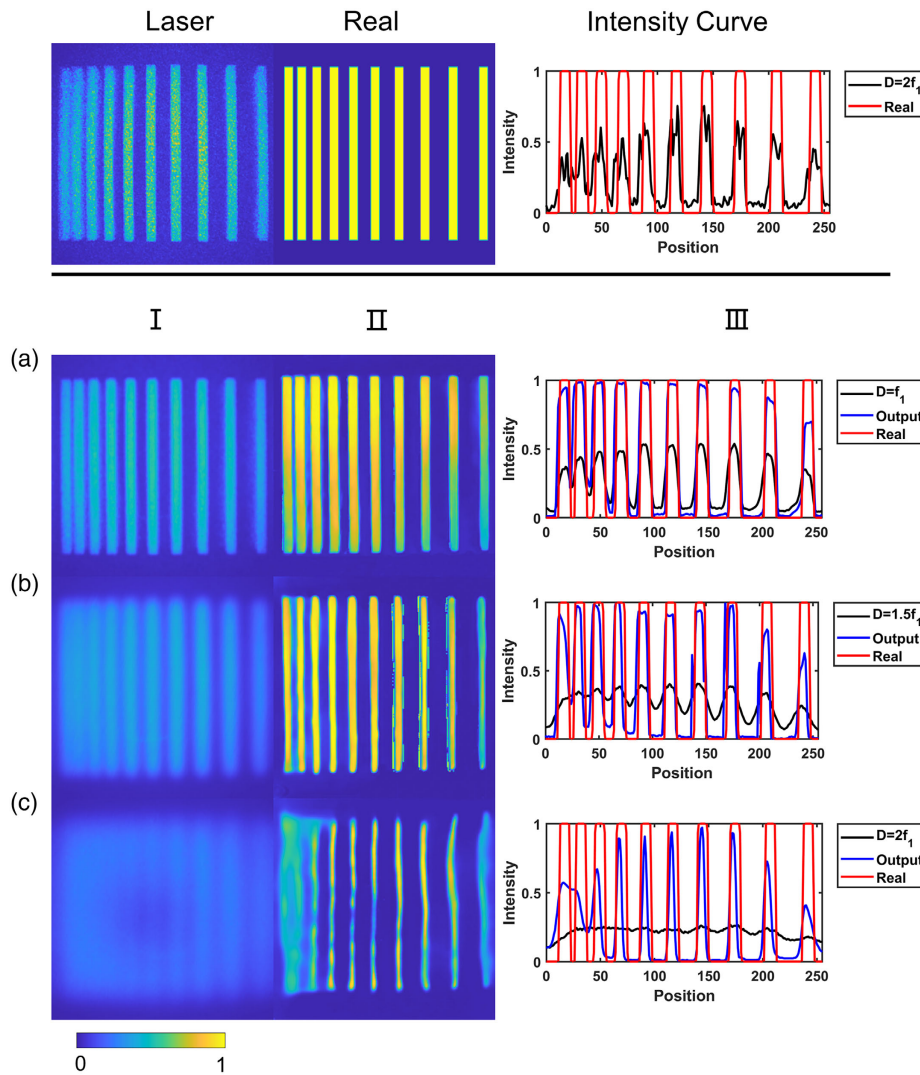
Figure 5. Experimental resolution test result for U-RDN with different special coherence. a) $D = f_1$, coherence is 0.48. b) $D = 1.5f_1$, coherence is 0.3. c) $D = 2f_1$, coherence is 0.14. (I) Original images captured by CMOS. (II) The output results of U-RDN. (III) The lateral intensity curve of the original image, U-RDN output, and ground truth. D is the distance between the converging lens L1 ($f_1 = 150$ mm) and the diffuser.

resolution test results with different spatial coherence. It can be clearly seen that as the coherence decreases, the resolution decreases continuously. In the case of $D = 2f_1$, it is almost impossible to distinguish the adjacent fringes. After U-RDN processing, the resolution has been improved, and the line pairs can be seen clearly.

To quantitatively analyze the resolution enhancement after using U-RDN, here we use the concept of contrast and propose the mean line pairs contrast (MLPC), which is given by

$$\text{MLPC} = \frac{1}{N} \sum_{A_i \in \text{line pairs}} \frac{I_{\max}^{A_i} - I_{\min}^{A_i}}{I_{\max}^{A_i} + I_{\min}^{A_i}}. \tag{8}$$

where N is the total number of line pairs, A_i means the area of ith line pairs, and $I_{\max}^{A_i}, I_{\min}^{A_i}$ represent the intensity Max and intensity Min in A_i. Obviously, the closer MLPC is to 1, the higher the contrast of the line pairs, which means the higher the resolution.

Table 2 summarizes the results of each line pairs contrast and MLPC under different spatial coherence. Considering the difference in human eye recognition, we refer to the Rayleigh criterion and proposed the line pairs resolution coefficient (LPRC), LPRC $= I_{\min}^{A_i} / \max(I_{\max}^{A_i}, I_{\max}^{A_j})$. If LPRC > 0.8, the human eye can hardly distinguish line pairs. If $0.5 \leq$ LPRC ≤ 0.8, we can distinguish them fairly. When LPRC < 0.5, these line pairs can be easily distinguished by the human eye. **Table 3** shows the LPRC under different spatial coherence. Note that we recorded extremum rather than average. From the intensity curve in Figure 5, it is not difficult to find that the intensity change from dark streaks to bright streaks is continuous rather than abrupt like real case. The average will weaken the characteristics of these data, so we use extreme values for calculation. In Table 2 and 3, the performance data of the line pairs after U-RDN processing are very close to the ideal case. Both qualitative and quantitative results prove that the network is capable of

Table 2. The contrast of each line pairs area and MLPC with different spatial coherence. Real means the ideal case. Co means the coherence of light source.

Method\A_i[a]	Area of ith line pairs										MLPC
	1th	2th	3th	4th	5th	6th	7th	8th	9th	10th	
Laser	0.2822	0.6136	0.8531	0.8173	0.8984	0.8978	0.8922	0.8674	0.8932	0.8047	0.7820
$D=f_1$,Co$=0.48$	0.2537	0.4598	0.6343	0.7547	0.7700	0.7793	0.7796	0.8090	0.8163	0.7813	0.6838
f_1-output	**0.6256**	**0.6258**	**0.9286**	**0.9789**	**0.9813**	**0.9793**	**0.9893**	**0.9847**	**0.9771**	**0.9790**	**0.9052**
$D=1.5f_1$,Co$=0.3$	0.0161	0.0722	0.1287	0.2275	0.2806	0.3754	0.4898	0.5631	0.6551	0.5782	0.3387
$1.5f_1$-output	**0.5645**	**0.8382**	**0.8404**	**0.9244**	**0.9593**	**0.9788**	**0.9786**	**0.9823**	**0.9813**	**0.9894**	**0.9037**
$D=2f_1$,Co$=0.18$	0.0241	0.0071	0.0423	0.0474	0.8555	0.7755	0.1038	0.1534	0.2205	0.1172	0.0879
$2f_1$-output	**0.0462**	**0.4164**	**0.9225**	**0.9802**	**0.9834**	**0.9889**	**0.9842**	**0.9751**	**0.9358**	**0.7035**	**0.7936**
Real	1	1	1	1	1	1	1	1	1	1	1

[a]Bold values indicate the performance indicators of U-RDN output.

Table 3. The LPRC results of each line pair area with different spatial coherence. LPRC > 0.8, hard to distinguish; LPRC < 0.5, easy to distinguish. Co means the coherence of light source.

Method\A_i[a]	Area of ith line pairs								
	1th	2th	3th	4th	5th	6th	7th	8th	9th
Laser	0.4457	0.2098	0.7926	0.0878	0.0448	0.0538	0.0569	0.0710	0.0564
$D=f_1$,Co$=0.48$	0.5010	0.3400	0.2215	0.1268	0.1299	0.1222	0.1236	0.1055	0.1011
f_1-output	**0.2209**	**0.2289**	**0.0370**	**0.0106**	**0.0085**	**0.0104**	**0.0053**	**0.0077**	**0.0115**
$D=1.5f_1$,Co$=0.3$	0.8900	0.8174	0.7412	0.6117	0.5511	0.4502	0.3424	0.2795	0.2084
$1.5f_1$-output	**0.2617**	**0.0872**	**0.0853**	**0.03927**	**0.0207**	**0.0098**	**0.0067**	**0.0091**	**0.0179**
$D=2f_1$,Co$=0.18$	0.8838	0.9551	0.9187	0.8944	0.8423	0.7951	0.7923	0.7339	0.6386
$2f_1$-output	**0.8116**	**0.3179**	**0.0302**	**0.0098**	**0.0080**	**0.0053**	**0.0079**	**0.0126**	**0.0331**
Real	0	0	0	0	0	0	0	0	0

[a]Bold values indicate the performance indicators of U-RDN output.

enhancing image resolution that was previously destroyed by low spatial coherence, showing that our framework is capable of restoring images with high-level semantic understanding of the imaging process.
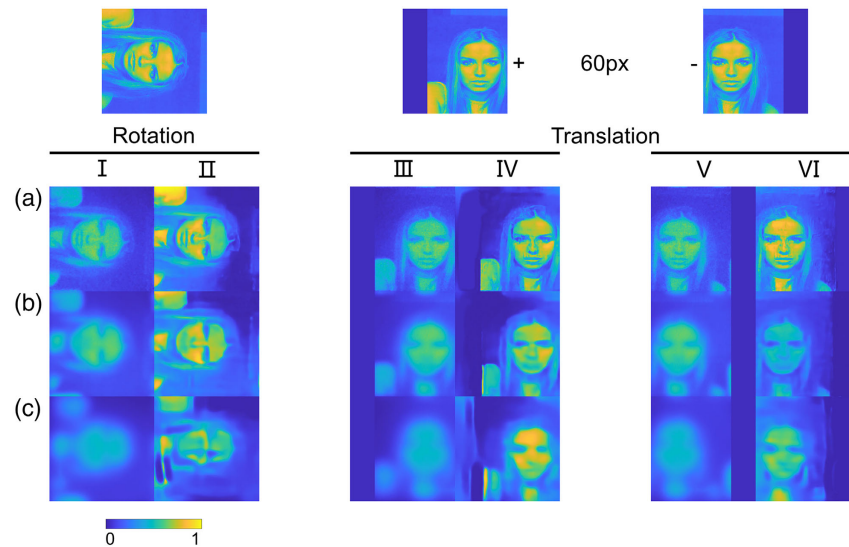
We also test the rotation and translation robustness of the U-RDN. It is entered into the trained network after the original image was rotated and translated. The data we selected belongs to the OOD sample, which has never been trained on the network. **Figure 6** shows the output results of the network. Results for other angles and translation distances are shown in the Supporting Information. When the coherence is not seriously damaged, the network shows better robustness. However, when the coherence is severely damaged, the robustness of the U-RDN is affected to a certain extent. Especially the rotation robustness, although the reconstruction result can identify the rotation direction, the reconstruction quality is very low. This is because the pixel features of the training data tend to be spatially consistent. When the network blindly reconstructs the rotated image, it incorrectly identifies some rotated pixel regions as unrotated, which leads to the failure of reconstruction. Compared to the rotational robustness, the translational robustness performs better, but the overall strength is reduced. From the above two experiences, we can conclude that the robustness of the network will be minor damaged if the spatial coherence is extremely poor.

4. Discussion

In summary, we have deliberately developed an architecture for deep neural networks that can be effectively used to recover and improve the imaging quality of CGH with low spatial coherence. The whole research includes designing optical experiments to obtain imaging results under different spatial coherence, inputting images into U-RDN for training and testing, and finally realizing the prediction and recovery of images generalizable to unseen classes and unseen domains. Our joint design of physical hardware combined with the neural network is proved to have strong OOD robustness, generalizable to any unseen classes as well as any unseen domains, with various experiments. Ultimately, our research results are expected to expand the applications of CGH. For example, when the external environment is disturbed, we can achieve high-definition holographic displays for virtual and augmented reality, high-resolution medical micro-imaging, and real-time simulation for holographic microscopy.

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
PHOTONICS
RESEARCH**
Open Access

www.adpr-journal.com

Figure 6. Rotation and translation robustness test results for U-RDN with different special coherence. a) $D = f_1$, coherence is 0.48. b) $D = 1.5f_1$, coherence is 0.3. c) $D = 2f_1$, coherence is 0.14. (I) (III) (V) The inputs of U-RDN. (II) (IV) (VI) The outputs of U-RDN.

However, in the face of ultralow spatial coherence, the restoration effect of our proposed scheme still suffers from loss of image details and unclearness. These deficiencies are also reflected in the resolution and robustness tests. Therefore, better methods to obtain high-quality imaging in ultralow spatial coherence are challenging.

Deep learning has been integrated into various fields in physics. It relies on big data to produce approximate outputs and achieve remarkable results in engineering applications. "Approximation" is a very important concept in physics, and deep learning is an approximate method that relies on big data. We aim to use this strategy to solve problems that cannot currently be described by the known physical model. As machine learning algorithms have developed in recent years, convolutional neural networks are no longer the only method for imaging. In the field of computer vision, for example, various GANs[47,48] and vision transformer architectures[49,50] have shown strong performance. In future studies, we plan to explore their potential in the field of low coherence and CGH. Finally, we believe that our approach benefits CGH in low spatial coherence, making holographic displays a viable technology, particularly for imaging through random scattering media, virtual reality, and augmented reality applications.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

X.T.: conceived the idea. D.Z. and R.X.: supervised the project. X.T.: designed the experiments, built the procedures, and analyzed most of the data. K.L. and X.T.: wrote programs. X.T., R.X., L.Z., and W.Z.: process the data. D.Z., X.T., R.X., and K.L.: wrote the manuscript. All authors discussed and commented on the manuscript.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

[1] B. R. Brown, A. W. Lohmann, *IBM J. Res. Dev.* **1969**, *13*, 160.
[2] B. R. Brown, A. W. Lohmann, *Appl. Opt.* **1966**, *5*, 967.
[3] A. W. Lohmann, D. P. Paris, *Appl. Opt.* **1967**, *6*, 1739.
[4] C. Slinger, C. Cameron, M. Stanley, *Computer* **2005**, *38*, 46.
[5] J. P. Cui, N. Zhang, J. Liu, D. L. Wu, H. Xu, D. Y. Yan, P. Ma, *Opt. Express* **2018**, *26*, 28067.
[6] Y. Hu, S. Fu, S. Wang, W. Zhang, P. O. Kwok, in: *9th Int. Symp. on Advanced Optical Manufacturing and Testing Technologies: Meta-Surface-Wave and Planar Optics*, Chengdu, China **2019**, pp. 125.

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
PHOTONICS
RESEARCH**
Open Access

www.adpr-journal.com

[7] L. B. Lesem, P. M. Hirsch, J. A. Jordan, *IBM J. Res. Dev.* **1969**, *13*, 150.

[8] R. W. Gerchberg, A. Saxton, *Optik* **1972**, *35*, 237.

[9] G. Z. Yang, B. Y. Gu, B. Z. Dong, *Int. J. Mod. Phys. B* **1993**, *7*, 3153.

[10] N. Streibl, *Opt. Commun.* **1984**, *49*, 6.

[11] M. Eybposh, N. W. Caira, M. Atisa, P. Chakravarthula, N. C. Pégard, *Opt. Express* **2020**, *28*, 26636.

[12] Z. He, X. Sui, G. Jin, D. Chu, L. Cao, *Opt. Express* **2021**, *29*, 119.

[13] K. V. Chellappan, E. Erden, H. Urey, *Appl. Opt.* **2010**, *49*, 79.

[14] M. Nixon, B. Redding, A. A. Friesem, H. Cao, N. Davidson, *Opt. Lett.* **2013**, *38*, 3858.

[15] Z. Cui, A. T. Wang, Z. Wang, S. L. Wang, C. Gu, H. Ming, C. Q. Xu, *J. Disp. Technol.* **2015**, *11*, 330.

[16] F. Zernike, *Physica* **1938**, *5*, 785.

[17] E. Wolf, *Il Nuovo Cimento* **1954**, *12*, 884.

[18] E. Wolf, *Nature* **1953**, *172*, 535.

[19] E. Collett, E. Wolf, *Opt. Lett.* **1978**, *2*, 27.

[20] J. C. Ricklin, F. M. Davidson, *J. Opt. Soc. Am. A* **2003**, *20*, 856.

[21] J. Wu, A. D. Boardman, *J. Mod. Opt.* **1991**, *38*, 1355.

[22] J. Liu, X. Duan, Z. B. Zhang, in *AOPC 2019: Display Technology and Optical Storage*, Beijing, China **2019**, p. 1133502.

[23] X. Duan, J. Liu, J. Hao, B. Zhao, T. Gao, in *AOPC 2019: Display Technology and Optical Storage*, Beijing, China **2019**, p. 1133503.

[24] K. A. Chen, B. Abbey, B. Dilanian, R. Balaur, E. Van Riessen, G. Junker, M. Nugent, *Phys. Rev. B* **2012**, *86*, 235401.

[25] L. Huang, T. Liu, X. Yang, Y. Luo, Y. Rivenson, A. Ozcan, *ACS Photonics* **2021**, *8*, 1763.

[26] C. Saharia, J. Ho, W. Chan, T. Salimans, D. J. Fleet, M. Norouzi, ArXiv Preprint ArXiv:2104.07636, **2021**.

[27] Y. Li, Y. Xue, L. Tian, *Optica* **2018**, *5*, 1181.

[28] R. Horisaki, R. Takagi, J. Tanida, *Opt. Express* **2016**, *24*, 13738.

[29] Y. Rivenson, Y. Zhang, H. Gunaydin, T. Da, A. Ozcan, *Light: Sci. Appl.* **2017**, *7*, 17141.

[30] A. T. Sinha, J. Lee, S. Li, G. Barbastathis, *Optica* **2017**, *4*, 1117.

[31] A. Goy, K. Arthur, S. Li, G. Barbastathis, *Phys. Rev. Lett.* **2018**, *121*, 243902.

[32] K. He, X. Chen, S. Xie, Y. Li, P. Dollár, R. Girshick, ArXiv Preprint ArXiv:2111.06377, **2021**.

[33] K. He, X. Zhang, S. Ren, J. Sun, in, *2016 IEEE Conf. on Computer Vision and Pattern Recognition* **2016**, pp. 770–778.

[34] G. Huang, Z. Liu, V. Laurens, K. Q. Weinberger, in: IEEE Computer Society, **2016**, pp. 4700–4708.

[35] O. Ronneberger, P. Fischer, T. Brox, in *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention*, Munich, Germany **2015**, pp. 234–241.

[36] Y. Bengio, Y. Lecun, G. Hinton, *Commun. ACM* **2021**, *64*, 58.

[37] D. Jia, D. Wei, R. Socher, L. J. Li, L. Kai, F. F. Li, in *2009 IEEE Conf. on Computer Vision and Pattern Recognition*, IEEE, Piscataway, NJ **2009**, pp. 248–255.

[38] Y. Peng, S. Choi, J. Kim, G. Wetzstein, *Sci. Adv.* **2021**, *7*, 5040.

[39] H. Zhang, J. Xie, J. Liu, Y. Wang, *Appl. Opt.* **2009**, *48*, 5834.

[40] Z. Chen, M. Mitchell, M. Segev, T. H. Coskun, D. N. Christodoulides, *Phys. Rev. Lett.* **1996**, *77*, 490.

[41] Z. Liu, L. Ping, X. Wang, X. Tang, in *IEEE Int. Conf. on Computer Vision*, IEEE, Piscataway, NJ **2016**, pp. 3730–3738.

[42] T. Y. Lin, M. Maire, S. Belongie, J. Hays, C. L. Zitnick, in Springer International Publishing, **2014**, pp. 740–755.

[43] Y. Srivastava, V. Murali, S. R. Dubey, in: *National Conf. on Computer Vision, Pattern Recognition, Image Processing, and Graphics*, Computer Vision, Pattern Recognition, Image Processing, and Graphics, Dharwad, India **2020**, pp. 322–332.

[44] M. Everingham, J. Winn, *Pattern Anal. Stat. Modell. Comput. Learn. Tech. Rep.* **2011**, *8*, 5.

[45] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, Y. Fu, in *2018 IEEE Conf. on Computer Vision and Pattern Recognition*, **2018**, pp. 2472–2481.

[46] S. Li, M. Deng, J. Lee, A. Sinha, G. Barbastathis, *Optica* **2018**, *5*, 803.

[47] H. Sun, J. Plawinski, S. Subramaniam, A. Jamaludin, T. Kadir, A. Readie, G. Ligozio, D. Ohlssen, M. Baillie, T. Coroller, ArXiv Preprint ArXiv:2106.13199, **2021**.

[48] T. Miyato, M. Koyama, ArXiv Preprint ArXiv:1802.05637, **2018**.

[49] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, N. Houlsby, ArXiv Preprint ArXiv:2010.11929, **2020**.

[50] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, B. Guo, in *Proc. of the IEEE/CVF Int. Conf. on Computer Vision*, Montreal, BC, Canada **2021**, pp. 10012–10022.

2200264 (10 of 10)