PHOTONICS Research

Intelligent coding metasurface holograms by physics-assisted unsupervised generative adversarial network

Che Liu,[1,2] Wen Ming Yu,[1,2] Qian Ma,[1,2] Lianlin Li,[3] and Tie Jun Cui[1,2,*]

[1]Institute of Electromagnetic Space, Southeast University, Nanjing 210096, China
[2]State Key Laboratory of Millimeter Wave, Southeast University, Nanjing 210096, China
[3]School of Electronic Engineering and Computer Sciences, Peking University, Beijing 100871, China
*Corresponding author: tjcui@seu.edu.cn

Received 30 November 2020; revised 4 February 2021; accepted 6 February 2021; posted 8 February 2021 (Doc. ID 416287); published 31 March 2021

Intelligent coding metasurface is a kind of information-carrying metasurface that can manipulate electromagnetic waves and associate digital information simultaneously in a smart way. One of its widely explored applications is to develop advanced schemes of dynamic holographic imaging. By now, the controlling coding sequences of the metasurface are usually designed by performing iterative approaches, including the Gerchberg–Saxton (GS) algorithm and stochastic optimization algorithm, which set a large barrier on the deployment of the intelligent coding metasurface in many practical scenarios with strong demands on high efficiency and capability. Here, we propose an efficient non-iterative algorithm for designing intelligent coding metasurface holograms in the context of unsupervised conditional generative adversarial networks (cGANs), which is referred to as physics-driven variational auto-encoder (VAE) cGAN (VAE-cGAN). Sharply different from the conventional cGAN with a harsh requirement on a large amount of manual-marked training data, the proposed VAE-cGAN behaves in a physics-driving way and thus can fundamentally remove the difficulties in the conventional cGAN. Specifically, the physical operation mechanism between the electric-field distribution and metasurface is introduced to model the VAE decoding module of the developed VAE-cGAN. Selected simulation and experimental results have been provided to demonstrate the state-of-the-art reliability and high efficiency of our VAE-cGAN. It could be faithfully expected that smart holograms could be developed by deploying our VAE-cGAN on neural network chips, finding more valuable applications in communication, microscopy, and so on.

https://doi.org/10.1364/PRJ.416287

1. INTRODUCTION

Electromagnetic (EM) metasurfaces use periodic or quasi-periodic macroscopic basic units to simulate the atoms or molecules on the microscale of traditional material science. These macroscopic basic units could interact with external electric fields via resonance effects and express unique EM properties [1–6] to realize various novel functional devices such as cloaks [7–10], concentrators [11], illusion optics devices [12,13], special lenses [14,15], and diffuse reflections [16]. Recently, a series of special EM metasurfaces called coding, digital, and programmable metasurfaces [17] ('coding metasurfaces' for short in the remainder of this article) have gained more attention due to the unique methods for manipulating the EM waves. By encoding the phase responses of the metasurfaces as digital numbers '0' or '1', the EM property of each meta-unit in the coding metasurfaces could be switched in real time when controlled by a field-programmable gate array (FPGA), which allows people to design metasurfaces in digital space rather than the analog domain. Further, the digital representation of meta-units could link the EM space and digital world and has been widely used in manipulating amplitude [18,19], polarization [20–22], and orbital angular momentum [23–25], yielding the concept of information metasurfaces [26–29] and many other novel applications [30–33].

With the development of communication technology, people need to deal with more and more information in daily life, which adds the burden of the whole society and promotes the development of artificial intelligence (AI) to help people deal with various information handling tasks such as speech recognition [34–36], image recognition [37–39], automatic translation [40–42], and robot control [43–45]. The growth of AI has also brought changes to the design of EM metasurfaces, especially coding metasurfaces. The digital representation of the coding metasurfaces makes it more convenient to put AI technology into the design of the meta-unit states [46–48] and has

already sprung out various interesting applications such as the smart system [49], high-resolution imager [47,50], and recognizer [51].

Despite the wide usage of AI in coding metasurfaces, the design of meta-unit states for holographic imaging by coding metasurfaces still rests on iteration optimization algorithms like the Gerchberg–Saxton (GS) algorithm [52,53] or greedy algorithm [54]. Recently, a valuable trail has been made for using supervised deep neural networks (DNNs) to recover holograms from the generated speckles distorted by a thin diffuser [55]. However, the demand for thousands of training samples acquired by real measurements remarkably increases the cost in the usage of this supervised DNN. Here, we propose a new method based on AI or deep learning that could rapidly generate the coding pattern (metasurface hologram) of the 1 bit coding metasurface when a target holographic image is given. The deep learning structure composed of unsupervised variational auto-encoder (VAE) and conditional generative adversarial networks (cGANs) is presented in our method. The use of VAE [56], as an unsupervised algorithm, could remarkedly reduce the time consumption for the preparation of training data, and the EM propagation model described by the rigorous dyadic Green's function (DGF) [53] is used to make the VAE structure possible while keeping the physical interpretability. We merge the Wasserstein distance [57–59] [by the form of Wasserstein GAN (WGAN) [60,61]] and mean square error (MSE) in the design of loss function to make the distributions of generated holographic images and target holographic images closer. The structure of cGAN [62–64] is also used to avoid the confusion of the generative network. We design a 40 × 40 one bit coding metasurface working at the frequency of 35 GHz to validate our intelligent method, and the simulation and experimental results both validate the efficiency and reliability of the proposed approach.

2. THEORY AND METHOD

A. EM Propagation Model

Figure 1(a) shows the sketch of our hologram system, which consists of a 40 × 40 one bit coding metasurface controlled by FPGA and a feed antenna radiating the EM waves in the frequency range from 34 to 36 GHz. Each meta-unit on the coding metasurface could be switched between two opposite states, radiating the EM waves with opposite phases of 0 or π, which are represented by codes '0' and '1', respectively. The 1 bit coding metasurface is the imager that loads the metasurface hologram, which can project the holographic image at the imaging plane 30 cm away from it. The size of the metasurface hologram and imaging plane is $L_1 = 152$ mm and $L_2 = 200$ mm. Figure 1(b) shows the sketch of the meta-unit's structure, whose parameters are given by $a = 3.8$ mm, $b = 2.2$ mm, $c = 1.4$ mm, $h_1 = 0.813$ mm, and $h_2 = 0.435$ mm. The substrate material of the metasurface is Rogers RO 4003C ($\varepsilon_r = 3.55$, $\tan\delta = 0.0027$). A PIN diode is integrated on the top layer of the metasurface element, connecting with two metal patches, which offer the bias voltage control through two via-holes. When the on/off state of diode is changed by specific bias voltage, the surface electric field alters distinctively to generate the reflected phase response with 180° difference.

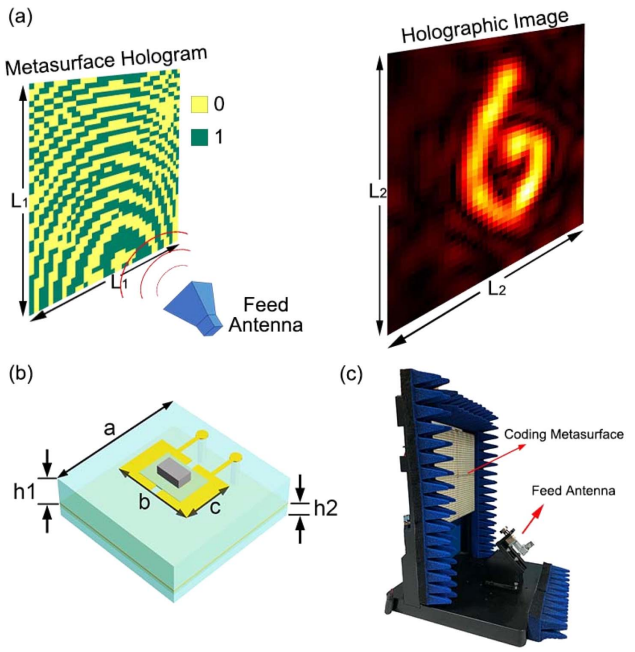

Fig. 1. (a) Sketch of our hologram system that consists of a 1 bit coding metasurface loading the hologram and a feed antenna. The distance between the holographic imaging plane and coding metasurface is 30 cm. (b) The meta-unit of the 1 bit coding metasurface. (c) The side-looking photograph of our hologram system. The interval between meta-units on the coding metasurface is 3.8 mm, and the feed antenna radiates the EM waves with frequencies from 34 GHz to 36 GHz.

The photograph of the hologram system is illustrated in Fig. 1(c), which is a well-integrated system with the coding metasurface controlled by the FPGA via an ethernet communication interface.

We accept rigorous DGF [53] as the basic computational kernel of the forward propagation from the source currents of meta-units to the near-field EM distribution:

$$\vec{G}(\mathbf{r},\mathbf{r}') = \left(\vec{\mathbf{I}} + \frac{\nabla\nabla}{k^2}\right) g(\mathbf{r},\mathbf{r}'), \tag{1}$$

where $\vec{\mathbf{I}}$ is a 3 × 3 dyadic identity matrix, $\mathbf{r}$ and $\mathbf{r}'$ are source and field points, respectively, and

$$g(\mathbf{r},\mathbf{r}') = \frac{e^{-jkR}}{4\pi R} \tag{2}$$

is the free-space Green's function, in which $R = |\mathbf{r} - \mathbf{r}'|$. The forward propagation formula could be represented as

$$E(\mathbf{r}) = -j\omega\mu \int_V \mathrm{d}\mathbf{r}' \vec{G}(\mathbf{r},\mathbf{r}') \cdot J(\mathbf{r}'), \tag{3}$$

where $E(\mathbf{r})$ and $J(\mathbf{r}')$ represent the electric field at the field point and the current at the source point, respectively. Owing to the discrete array form of the coding meta-units, the source currents can also be expressed in a discrete form. For convenience, we use a current element to represent a coding meta-unit and discrete the EM field into M points at the same time. Then, Eq. (3) could reduce to

$$E(\mathbf{r}_m) = -j\omega\mu \sum_{n=1}^{N} \vec{G}(\mathbf{r}_m, \mathbf{r}'_n) \cdot J(\mathbf{r}'_n), \quad m = 1, \ldots, M, \tag{4}$$

where N is the number of coding meta-units. Further, we organize the scalar components of $E(r_m)$ and $J(r'_n)$ in Eq. (4) for all m and n into column vectors $\mathbf{E}$ and $\mathbf{J}$, respectively. As we can see from Eq. (4), the relationship between $E(r_m)$ and $J(r'_n)$ is linear, and, hence, they could be connected with a complex-value coefficient matrix. Then, Eq. (4) could ultimately reduce to

$$\mathbf{E} = \mathbf{W} \cdot \mathbf{J}, \tag{5}$$

where $\mathbf{W}$ is the coefficient matrix that links the source $\mathbf{J}$ and field $\mathbf{E}$. We only care about the vertical polarization EM waves and the relative amplitude of the near-field EM distribution. Hence, the current vector $\mathbf{J}$ could be described as

$$\mathbf{J} = \mathbf{J}^r \odot \mathbf{J}^\varphi, \tag{6}$$

where $\mathbf{J}^r$ is a complex-value vector and represents the current part that is directly caused by the incident EM waves, which is proportional to the incident electric-field values; $\odot$ means element-wise multiplication; and $\mathbf{J}^\varphi$ represents the current part that is controlled by each coding meta-unit, which is a real-value vector whose elements are '1' or '−1' to represent the phase of 0 or π, corresponding to code '0' or '1'. Then, Eq. (5) could be rewritten as

$$\begin{aligned} \mathbf{E} &= \mathbf{W} \cdot \mathbf{J}^r \odot \mathbf{J}^\varphi = \mathbf{W} \cdot \mathrm{diag}(\mathbf{J}^r) \cdot \mathbf{J}^\varphi = \mathbf{W}^r \cdot \mathbf{J}^\varphi, \\ \mathbf{W}^r &= \mathbf{W} \cdot \mathrm{diag}(\mathbf{J}^r). \end{aligned} \tag{7}$$

Thus, the forward propagation process could be represented as a form of matrix multiplication from a real-value vector $\mathbf{J}^\varphi$ to the near-field electric-field distribution $\mathbf{E}$. The design objective of VAE-cGAN is to get the current vector $\mathbf{J}^\varphi$ to generate a target near-field electric-field distribution $\mathbf{E}$, whose amplitude distribution represents a target holographic image.

B. Network Structure

Figure 2 demonstrates the network structure of the proposed VAE-cGAN. It maily consists of two modules: generator and discriminator. The generator is responsible for generating the current vector $\mathbf{J}^\varphi$ when a target holographic image is given. The binarization target holographic image is first input to the Resnet34 [65] network of the generator, whose output is activated by the hyperbolic tangent function and finally becomes a 1 × 1600 current vector $\mathbf{J}^\varphi$, whose element value is near '1' or '−1', representing the code '0' or '1' in a 40 × 40 one bit coding metasuface. We replace the batch normalization [66] with the instance normalization [67] in the Resnet34 network to prevent interaction among batches. Then, the holographic image generated by $\mathbf{J}^\varphi$ is calculated by Eq. (7) and divided by its maximum value to keep its value in the range [0,1]; next, it is sent into the discriminator to calculate the Wasserstein distance [57] between itself and the target holographic image. The MSE is also calculated at the same time as another index, indicating the similarity between the generated holographic image and target holographic image. So, the generator together with the EM forward propagation process makes up the structure of the VAE [55], whose optimization target is making the output image as similar as possible to the input image. The usage of the unsupervised structure VAE gets us out of the trouble of making the marked training dataset, whose data must be organized in pairs with the target holographic image and its corresponding coding pattern (metasurface hologram) generated by iterative algorithms like the GS algorithm. Besides, the multiplicity-solution property of the iterative algorithm will cause one-to-many problems in the marked training dataset and makes the training process difficult to converge.

The discriminator adopts the structure of cGAN [63] and WGAN [61], which is responsible for calculating the approximate Wasserstein distance between the generated holographic image and the target holographic image. It will be trained in the adversarial process towards the generator and improve the imaging quality of the generated holographic image.

C. Backpropagation Process

The mainstream deep learning platform such as TensorFlow and Pytorch could not directly deal with the complex values, which will cause difficulty in the backpropagation process from loss function to $\mathbf{J}^\varphi$. Thus, we need to derive the backpropagation partial derivative equation from the generated holographic image (the amplitude of $\mathbf{E}$) to $\mathbf{J}^\varphi$ to get rid of the calculation of complex values, which is expressed as (using numerator layout)

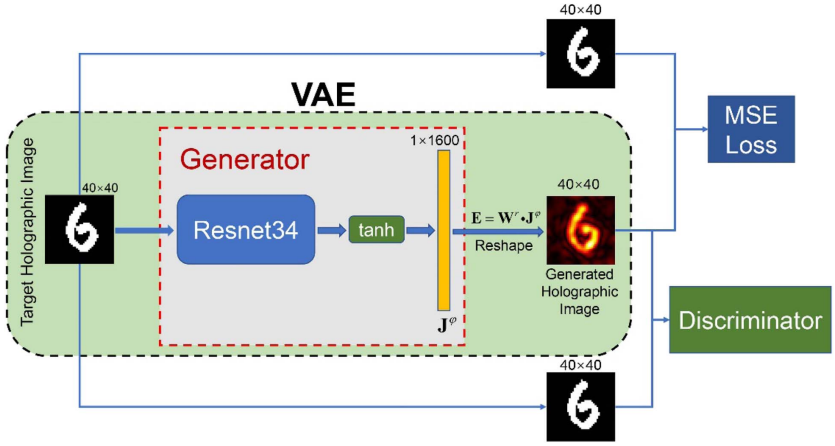

Fig. 2. Schematic diagram of the proposed VAE-cWGAN. The generator together with the EM propagation process makes up the VAE structure. Two kinds of distance criteria (MSE and Wasserstein distance) are used to improve the imaging quality of the generator.

$$\frac{\partial|\mathbf{E}|}{\partial \mathbf{J}^{\varphi}} = |\mathbf{E}|^{\odot -1} \odot \mathrm{Re}\left[\mathbf{E}^{*} \odot \frac{\partial \mathbf{E}}{\partial \mathbf{J}^{\varphi}}\right] = |\mathbf{E}|^{\odot -1} \odot \mathrm{Re}[\mathbf{E}^{*} \odot \mathbf{W}^{r}], \quad (8)$$

where the superscript notation $\odot -1$ means taking the reciprocal element-wisely and $*$ means taking the conjugate matrix. Then, the backpropagation process from loss function to $\mathbf{J}^{\varphi}$ can be calculated by real values,

$$\frac{\partial \mathrm{Loss}}{\partial \mathbf{J}^{\varphi}} = \frac{\partial \mathrm{Loss}}{\partial |\mathbf{E}|} \cdot \frac{\partial |\mathbf{E}|}{\partial \mathbf{J}^{\varphi}}. \quad (9)$$

D. Design of Loss Function

We merge the MSE loss and Wasserstein distance as our final loss function. The MSE loss and Wasserstein distance are both criterions to indicate the distance between two distributions. Although using the Wasserstein distance evaluation function alone could also reach an equally good result, we still add the MSE evaluation to the loss function of the generator because the MSE evaluation could act as the 'lubricant' for the training process to accelerate and stabilize the convergence of the generator and discriminator. The discriminator cannot tell fake samples apart from real ones at the beginning of the training process; thus, it should be trained to learn a K-Lipschitz continuous function [60,61] so as to compute the Wasserstein distance. Therefore, at the beginning of training process, the generator is not able to get any effective guides from the discriminator for the updates of parameters, which would raise the risk of 'mode collapse'. Luckily, the MSE evaluation function could help to guide the updates of the generator before the discriminator is well trained so as to prevent the generator from the 'mode collapse'. This is the reason why we add it into the final loss function of the generator.

The MSE loss is described as

$$\mathrm{MSE} = \frac{1}{N}\mathrm{sum}[(|\mathbf{E}| - |\mathbf{E}_t|)^{\odot 2}],$$
$$\frac{\partial \mathrm{MSE}}{\partial |\mathbf{E}|} = \frac{2}{N}[(|\mathbf{E}| - |\mathbf{E}_t|)]^{T}, \quad (10)$$

where $|\mathbf{E}_t|$ represents the target holographic image and superscript $\odot 2$ means element-wise square.

Wasserstein distance [59] indicates the minimum movement for changing one distribution P_1 to another distribution P_2:

$$W(P_1, P_2) = \inf_{\gamma \sim \Pi(P_1, P_2)} \mathbb{E}_{(x,y)\sim\gamma}(\|x - y\|_p), \quad (11)$$

where $\Pi(P_1, P_2)$ represents the set of all of the possible joint distributions between P_1 and P_2, and notation $\|\cdot\|_p$ means p-norm. The Wasserstein distance is a more appropriate criterion to indicate the distance between two distributions. Figure 3(a) shows the three sets consisting of discrete sequences sampled from distributions P_1, P_2, and P_3, respectively, defined by

$$\forall \text{ sequence} \in P_1, \quad \text{sequence} = [x, 0, 0, 0, 0, 0, 0, 0] \quad \text{and} \quad x \sim U(0, 1),$$
$$\forall \text{ sequence} \in P_2, \quad \text{sequence} = [0, x, 0, 0, 0, 0, 0, 0] \quad \text{and} \quad x \sim U(0, 1),$$
$$\forall \text{ sequence} \in P_3, \quad \text{sequence} = [0, 0, 0, 0, 0, 0, x, 0] \quad \text{and} \quad x \sim U(0, 1).$$

The mathematical expectation of p-norm distance between the discrete sequences sampled from distributions P_1 and P_2 is the same as that between P_1 and P_3. However, it is obvious that

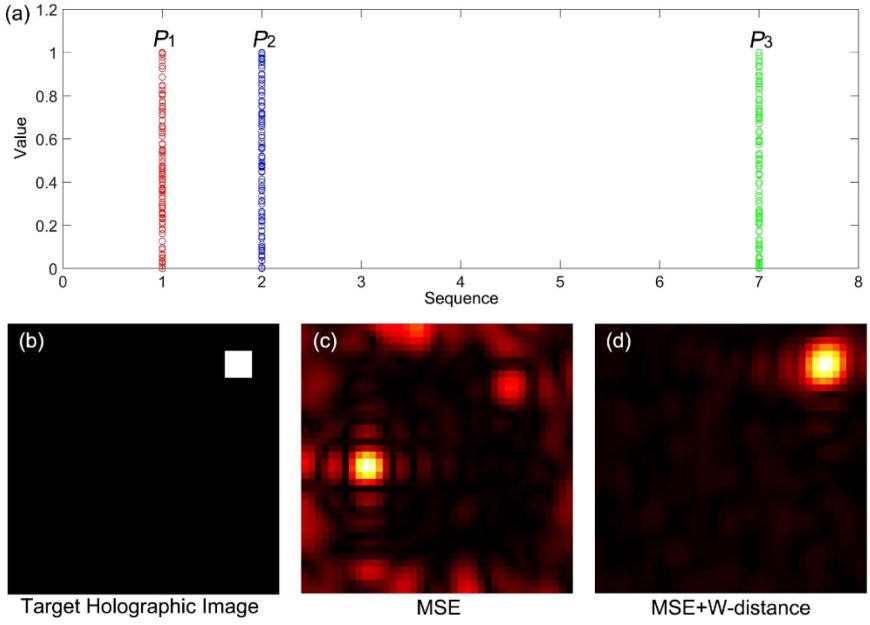

Fig. 3. (a) Three discrete-sequence sets sampled from distributions P_1, P_2, and P_3, respectively. (b) The target holographic image is input to the trained generator. (c) The generated holographic image [corresponding to Fig. 3(b)] output by a generator trained only using the MSE loss. (d) The generated holographic image output by a generator trained using MSE loss and Wasserstein distance simultaneously.

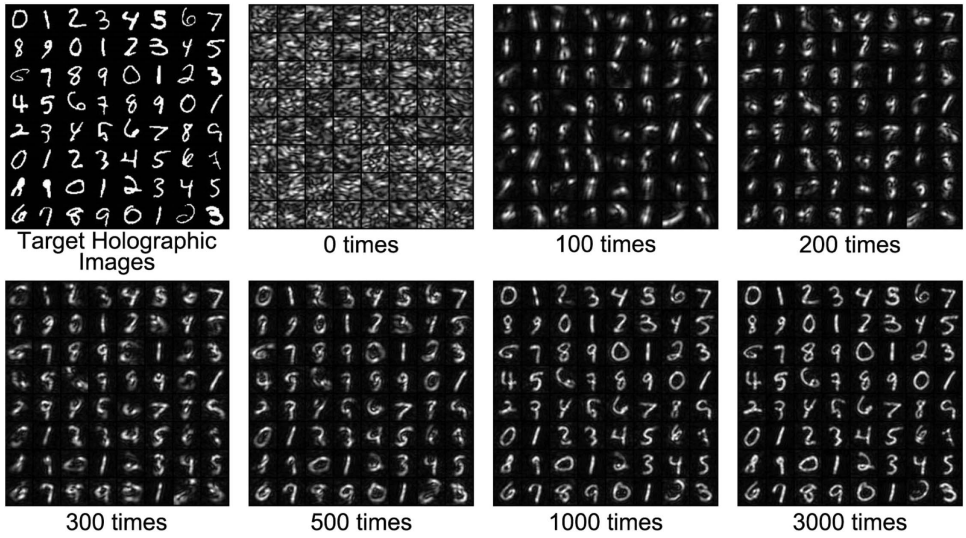

Fig. 4. Generated holographic images at each training time corresponding to the valid target holographic images. One time of training is when training generator has three iterations after the training discriminator has one iteration.

P_2 is more visually similar to P_1 than P_3. Thus, the p-norm criterion may fail to indicate the distance between two sparse distributions, just as demonstrated in Figs. 3(b) and 3(c), in which the target holographic image [Fig. 3(b)] can be expressed as a sparse matrix, making it difficult to find the direction of optimization and fall into local minimum with the generator trained using only MSE loss, and eventually output a wrong holographic image [Fig. 3(c)]. Luckily, the Wasserstein distance could help denote this distribution difference, and the generator trained by the Wasserstein distance could generate the correct holographic image [Fig. 3(d)].

For the loss function of the discriminator, we use the loss function of WGAN [61] to simulate the calculation of Wasserstein distance and introduce the concept of cGAN [63] to match the target holographic image with the generated holographic image:

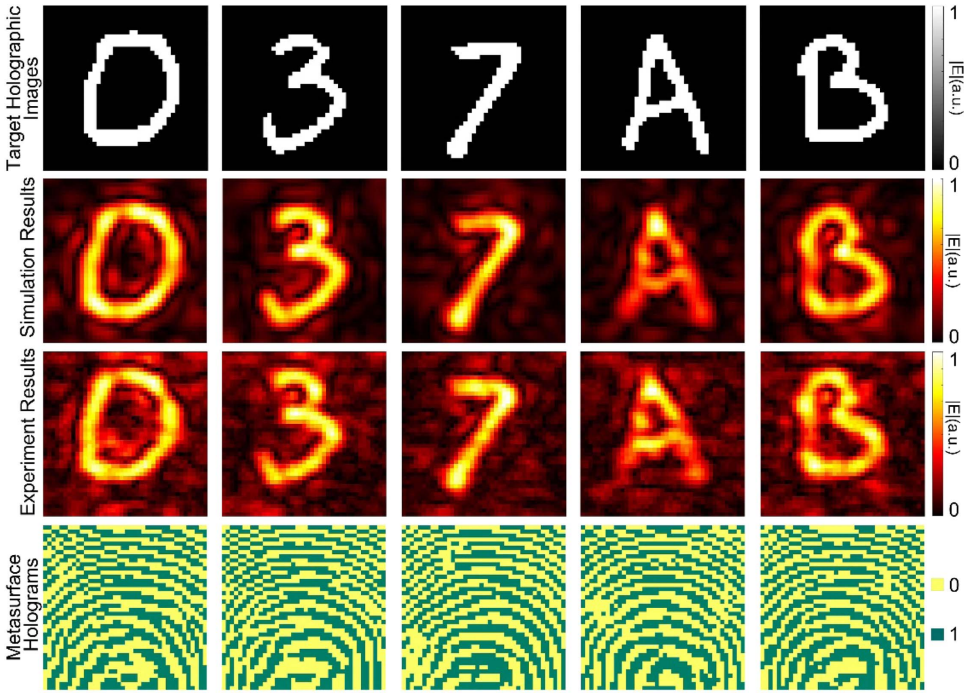

Fig. 5. Testing results of our proposed intelligent metasurface hologram system. The target holographic images are randomly chosen from the testing MNIST dataset or images of handwritten letters. The simulation holographic images are calculated by Eq. (7) with the binarized current vectors $\mathbf{J}^{\varphi}$ output by the generator. The experimental holographic images are radiated by our 1 bit coding metasurface configured with the corresponding metasurface holograms generated by the binarized current vectors $\mathbf{J}^{\varphi}$ from the generator.

$$\text{Loss\_}D = \mathop{\mathbb{E}}_{\tilde{x}\sim P_g|P_r}[D(\tilde{x})] - \mathop{\mathbb{E}}_{x\sim P_r|P_r}[D(x)] + \lambda \mathop{\mathbb{E}}_{\hat{x}\sim P_{\hat{x}}}[(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2], \quad (12)$$

where P_g and P_r represent the distributions of the generated holographic images and target holographic images, respectively. Meanwhile, by adding the MSE loss in Eq. (10), our loss function for the generator could be expressed as

$$\text{Loss\_}G = -\mathop{\mathbb{E}}_{\tilde{x}\sim P_g|P_r}[D(\tilde{x})] + \text{MSE}. \quad (13)$$

Figure 3(d) shows the generated holographic image output by the joint optimization of MSE loss and Wasserstein distance, which matches well with the target holographic image [Fig. 3(b)] and demonstrates the necessity of the mixed distance criterion.

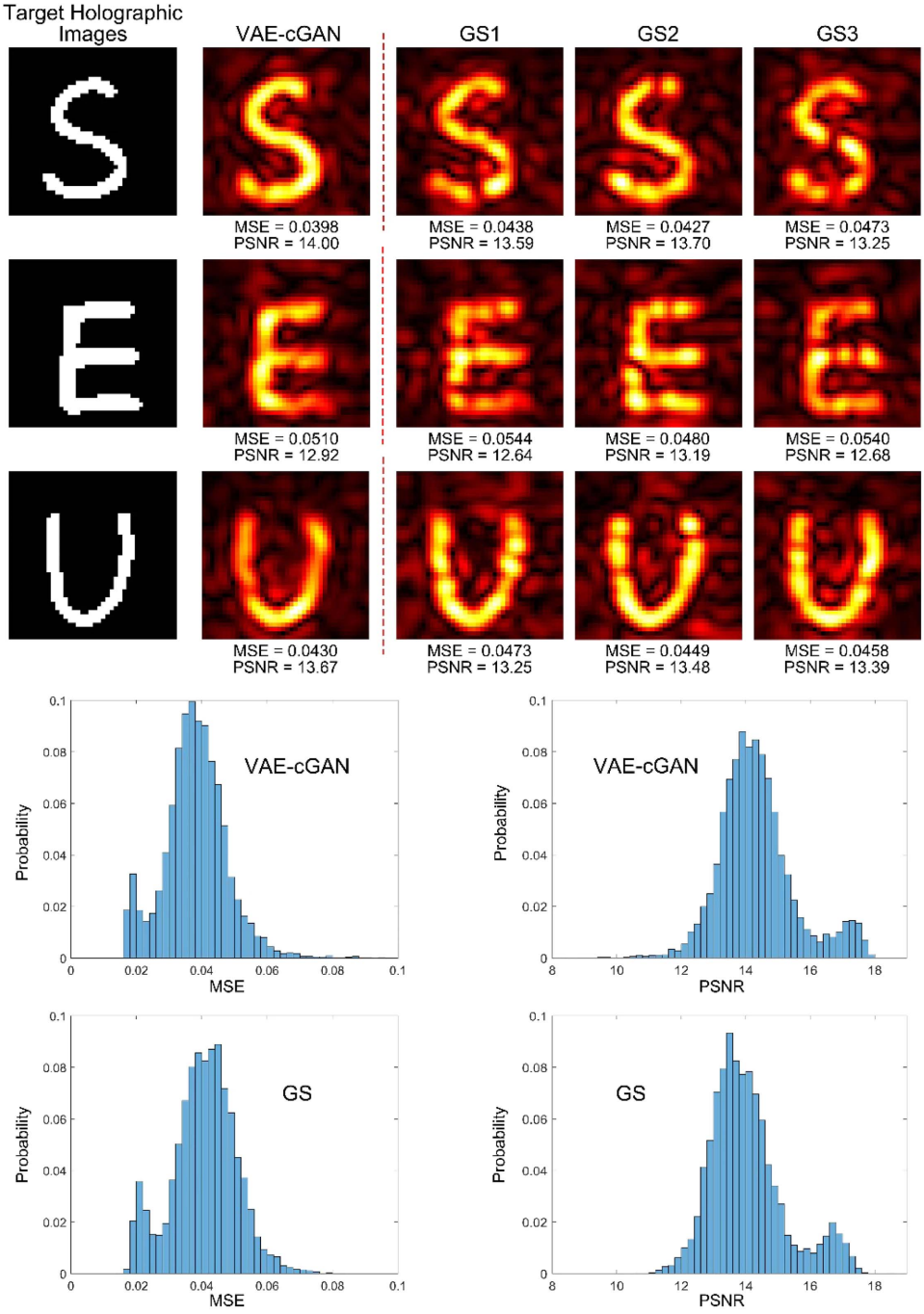

Fig. 6. Comparison results between the VAE-cGAN and GS algorithms. The top half presents the simulated results of the generated holographic images radiated by metasurface holograms designed by our VAE-cGAN and GS algorithms, respectively. Here, GS1, GS2, and GS3 are the generated holographic images obtained by running the GS algorithm three times in sequence. The values of MSE and PSNR evaluations are marked below the corresponding holographic images. The bottom half of this figure illustrates the statistical frequency histograms of the holographic image quality in terms of MSE and PSNR.

3. RESULTS AND DISCUSSION

We use the images from the Modified National Institute of Standards and Technology (MNIST) data sets of handwritten digits [68] as the target holographic images to train our VAE-cGAN. In accordance with tradition, we divide the images of the MNIST database into three parts: training, validating, and testing datasets, with 35,000, 5000, and 10,000 samples respectively. The images from the MNIST database are reshaped to 40 × 40 and binarized before being input to the generator or discriminator. Because the discriminator is much more easily trained than the generator, we train the generator three times before training the discriminator. The Adam optimizer is used with an initial learning rate of 3×10^{-4}, and the size of batch is 64. We use the Pytorch deep learning platform to construct our VAE-cGAN and run the program on a graphics processing unit (GPU) (Nvidia Quadro RTX 8000). The generator converges and becomes stable after about 9000 training iterations, and the total time consumption is about 2 h. The generated holographic images at each training time when giving the same 64 valid target holographic images are shown in Fig. 4, demonstrating that the generated holographic images successfully converge to the target holographic images when the training process goes on.

For the experimental process, we input a testing target holographic image into the trained generator and get its output current vector $\mathbf{J}^{\varphi}$. Then, the generated $\mathbf{J}^{\varphi}$ is binarized to '1' or '−1', which corresponds to the reflection phase of 0 or π (the coding meta-unit state of 0 or 1), respectively. The simulated holographic images are calculated by Eq. (7) with binarized current vectors $\mathbf{J}^{\varphi}$. The coding metasurface [Fig. 1(b)] is set using an FPGA with the coding meta-unit states generated from $\mathbf{J}^{\varphi}$. Last, we measure the holographic images radiated by the digital coding metasurface holograms in a standard microwave chamber.

Figure 5 shows some of the measured results. The target holographic images of handwritten digits '0', '3', and '7' in Fig. 5 are randomly chosen from the testing MNIST dataset, which means that they have not participated in the training process of the generator. The high imaging quality of the simulation and experimental holographic images indicates that the generator possesses the abilities of generalization. Furthermore, we use extra testing holographic images of handwritten letters, which have different features from the handwritten digits, to further verify the generalization ability of the generator. The randomly selected testing results corresponding to the target holographic images of handwritten letters 'A' and 'B' are also presented in Fig. 5, indicating that the generator could handle well these handwritten letters and proving the generality of our intelligent metasurface hologram system.

The comparison results between VAE-cGAN and GS algorithms are provided in Fig. 6, in which the top half presents the simulated results of the generated holographic images radiated by the metasurface holograms designed by VAE-cGAN and GS algorithms, respectively. Since the results of the GS algorithm are inconsistent due to its random parameters initialization, for one target holographic image, we run the GS algorithm three times and get three holographic images indicated by GS1, GS2, and GS3, respectively. The values of MSE and peak signal-to-noise ratio (PSNR) evaluation are marked below the corresponding holographic images, showing that the generated holographic images from our VAE-cGAN generally have smaller MSE and larger PSNR evaluation values than those from the GS algorithm. The bottom half of Fig. 6 illustrates the statistical frequency histograms of the holographic image quality in terms of MSE and PSNR, which are counted using the generated holographic images corresponding to 10,000 target images from the testing dataset of handwritten digits. Considering the inconsistency of the GS algorithm, for each target holographic image, the GS algorithm is run three times, and the smallest MSE value and the largest PSNR value are recorded as the final statistical results.

The mean MSE and PSNR values of our VAE-cGAN evaluated with the whole testing dataset are 0.0382 and 14.33, respectively. Compared with 0.0407 (mean-MSE) and 14.05 (mean-PSNR) of the GS algorithm, our VAE-cGAN shows a better capability in searching the global optimum than the GS algorithm.

4. CONCLUSION

We propose a new intelligent non-iterative approach (VAE-cGAN) based on deep learning methods for metasurface holograms. The usage of an unsupervised VAE structure makes our system easily trained from scratch, and the introduction of Wasserstein distance criterion improves the imaging quality of holographic images. After the VAE-cGAN is well trained, we just need to use the generator part to generate the coding patterns (metasurface holograms) corresponding to the target holographic images. The non-iterative structure of the generator enables the realization of holographic imaging with high quality and high efficiency, which are validated by both simulation and experimental results.

When deploying our trained generator on a neural network chip, our system could become a real-time holographical imager to rapidly generate the desired holographic images. It could be expected that our intelligent metasurface hologram system could become an efficient tool for microwave or even optical holograms [69], and more valuable applications could be explored in wireless communications [32], smart EM environment, health monitoring [51,70], and so on.

Funding. National Key Research and Development Program of China (2017YFA0700201, 2017YFA0700202, 2017YFA0700203, 2018YFA0701900).

Disclosures. The authors declare no conflicts of interest.

REFERENCES

1. V. G. Veselago, "Electrodynamics of substances with simultaneously negative values of sigma and μ," Sov. Phys. Usp. **10**, 509–514 (1968).
2. R. Zhao, L. Huang, and Y. Wang, "Recent advances in multi-dimensional metasurfaces holographic technologies," PhotoniX **1**, 20 (2020).
3. X. Ding, Z. Wang, G. Hu, J. Liu, K. Zhang, H. Li, B. Ratni, S. N. Burokur, Q. Wu, J. Tan, and C.-W. Qiu, "Metasurface holographic image projection based on mathematical properties of Fourier transform," PhotoniX **1**, 16 (2020).

4. Q. Ma, Q. R. Hong, X. X. Gao, H. B. Jing, C. Liu, G. D. Bai, Q. Cheng, and T. J. Cui, "Smart sensing metasurface with self-defined functions in dual polarizations," Nanophotonics **9**, 3271–3278 (2020).
5. X. Ni, N. K. Emani, A. V. Kildishev, A. Boltasseva, and V. M. Shalaev, "Broadband light bending with plasmonic nanoantennas," Science **335**, 427 (2012).
6. N. Yu, P. Genevet, M. A. Kats, F. Aieta, J.-P. Tetienne, F. Capasso, and Z. Gaburro, "Light propagation with phase discontinuities: generalized laws of reflection and refraction," Science **334**, 333–337 (2011).
7. D. Schurig, J. J. Mock, B. J. Justice, S. A. Cummer, J. B. Pendry, A. F. Starr, and D. R. Smith, "Metamaterial electromagnetic cloak at microwave frequencies," Science **314**, 977–980 (2006).
8. J. Li and J. B. Pendry, "Hiding under the carpet: a new strategy for cloaking," Phys. Rev. Lett. **101**, 203901 (2008).
9. R. Liu, C. Ji, J. J. Mock, J. Y. Chin, T. J. Cui, and D. R. Smith, "Broadband ground-plane cloak," Science **323**, 366–369 (2009).
10. Q. Ma, Z. L. Mei, S. K. Zhu, T. Y. Jin, and T. J. Cui, "Experiments on active cloaking and illusion for Laplace equation," Phys. Rev. Lett. **111**, 173901 (2013).
11. W. X. Jiang, T. J. Cui, Q. Cheng, J. Y. Chin, X. M. Yang, R. Liu, and D. R. Smith, "Design of arbitrarily shaped concentrators based on conformally optical transformation of nonuniform rational B-spline surfaces," Appl. Phys. Lett. **92**, 264101 (2008).
12. Y. Lai, J. Ng, H. Chen, D. Han, J. Xiao, Z.-Q. Zhang, and C. T. Chan, "Illusion optics: the optical transformation of an object into another object," Phys. Rev. Lett. **102**, 253902 (2009).
13. L. Chen, Q. Ma, Q. F. Nie, Q. R. Hong, H. Y. Cui, Y. Ruan, and T. J. Cui, "Dual-polarization programmable metasurface modulator for near-field information encoding and transmission," Photon. Res. **9**, 116–124 (2021).
14. N. Kundtz and D. R. Smith, "Extreme-angle broadband metamaterial lens," Nat. Mater. **9**, 129–132 (2010).
15. W. X. Jiang, C.-W. Qiu, T. C. Han, Q. Cheng, H. F. Ma, S. Zhang, and T. J. Cui, "Broadband all-dielectric magnifying lens for far-field high-resolution imaging," Adv. Mater. **25**, 6963–6968 (2013).
16. X. M. Yang, X. Y. Zhou, Q. Cheng, H. F. Ma, and T. J. Cui, "Diffuse reflections by randomly gradient index metamaterials," Opt. Lett. **35**, 808–810 (2010).
17. T. J. Cui, M. Q. Qi, X. Wan, J. Zhao, and Q. Cheng, "Coding metamaterials, digital metamaterials and programmable metamaterials," Light Sci. Appl. **3**, e218 (2014).
18. J. Li, Y. Zhang, J. Li, X. Yan, L. Liang, Z. Zhang, J. Huang, J. Li, Y. Yang, and J. Yao, "Amplitude modulation of anomalously reflected terahertz beams using all-optical active Pancharatnam–Berry coding metasurfaces," Nanoscale **11**, 5746–5753 (2019).
19. R. Y. Wu, L. Zhang, L. Bao, L. W. Wu, Q. Ma, G. D. Bai, H. T. Wu, and T. J. Cui, "Digital metasurface with phase code and reflection-transmission amplitude code for flexible full-space electromagnetic manipulations," Adv. Opt. Mater. **7**, 1801429 (2019).
20. Q. Ma, Q. R. Hong, G. D. Bai, H. B. Jing, R. Y. Wu, L. Bao, Q. Cheng, and T. J. Cui, "Editing arbitrarily linear polarizations using programmable metasurface," Phys. Rev. Appl. **13**, 021003 (2020).
21. Q. Ma, C. B. Shi, G. D. Bai, T. Y. Chen, A. Noor, and T. J. Cui, "Beam-editing coding metasurfaces based on polarization bit and orbital-angular-momentum-mode bit," Adv. Opt. Mater. **5**, 1700548 (2017).
22. Q. Ma, L. Chen, H. B. Jing, Q. R. Hong, H. Y. Cui, Y. Liu, L. Li, and T. J. Cui, "Controllable and programmable nonreciprocity based on detachable digital coding metasurface," Adv. Opt. Mater. **7**, 1901285 (2019).
23. G. Ding, K. Chen, X. Luo, J. Zhao, T. Jiang, and Y. Feng, "Dual-helicity decoupled coding metasurface for independent spin-to-orbital angular momentum conversion," Phys. Rev. Appl. **11**, 044043 (2019).
24. J. Han, L. Li, H. Yi, and Y. Shi, "1-bit digital orbital angular momentum vortex beam generator based on a coding reflective metasurface," Opt. Mater. Express **8**, 3470–3478 (2018).
25. Q. Zheng, Y. Li, Y. Han, M. Feng, Y. Pang, J. Wang, H. Ma, S. Qu, and J. Zhang, "Efficient orbital angular momentum vortex beam generation by generalized coding metasurface," Appl. Phys. A **125**, 136 (2019).
26. T. J. Cui, L. Li, S. Liu, Q. Ma, and Q. Cheng, "Information metamaterial systems," iScience **23**, 101403 (2020).
27. L. Li and T. J. Cui, "Information metamaterials - from effective media to real-time information processing systems," Nanophotonics **8**, 703–724 (2019).
28. T. J. Cui, S. Liu, and L. Zhang, "Information metamaterials and metasurfaces," J. Mater. Chem. C **5**, 3644–3668 (2017).
29. Q. Ma and T. J. Cui, "Information metamaterials: bridging the physical world and digital world," PhotoniX **1**, 1 (2020).
30. L. Zhang, X. Q. Chen, S. Liu, Q. Zhang, J. Zhao, J. Y. Dai, G. D. Bai, X. Wan, Q. Cheng, G. Castaldi, V. Galdi, and T. J. Cui, "Space-time-coding digital metasurfaces," Nat. Commun. **9**, 4334 (2018).
31. T. J. Cui, S. Liu, G. D. Bai, and Q. Ma, "Direct transmission of digital message via programmable coding metasurface," Research **2019**, 2584509 (2019).
32. H. Zhao, Y. Shuang, M. Wei, T. J. Cui, P. D. Hougne, and L. Li, "Metasurface-assisted massive backscatter wireless communication with commodity Wi-Fi signals," Nat. Commun. **11**, 3926 (2020).
33. Y. Shuang, H. Zhao, W. Ji, T. J. Cui, and L. Li, "Programmable high-order OAM-carrying beams for direct-modulation wireless communications," IEEE J. Emerg. Sel. Top. Circuits Syst. **10**, 29–37 (2020).
34. G. Hinton, L. Deng, D. Yu, G. E. Dahl, A. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. N. Sainath, and B. Kingsbury, "Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups," IEEE Signal Process. Mag. **29**, 82–97 (2012).
35. G. Pironkov, S. U. N. Wood, and S. Dupont, "Hybrid-task learning for robust automatic speech recognition," Comput. Speech Lang. **64**, 101103 (2020).
36. A. Graves, A. R. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in *IEEE International Conference on Acoustics, Speech and Signal Processing* (2013), pp. 6645–6649.
37. K. Lu, "Intelligent recognition system for high precision image significant features in large data background," in *Cyber Security Intelligence and Analytics* (2020), pp. 1056–1062.
38. T. Tong, X. Mu, L. Zhang, Z. Yi, and P. Hu, "MBVCNN: joint convolutional neural networks method for image recognition," AIP Conf. Proc. **1839**, 020091 (2017).
39. H. Sun, Q. Zhang, H. Wang, and A. Li, "Research of images recognition method based on RBF neural network," in *7th International Conference on System of Systems Engineering* (2012), pp. 24–26.
40. K. Cho, B. Van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. J. C. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," arXiv:1406.1078v3 (2014).
41. C. Escolano, M. R. Costa-Jussa, and J. A. R. Fonollosa, "From bilingual to multilingual neural-based machine translation by incremental training," in *57th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop* (2019), pp. 236–242.
42. S. Kwon, B. H. Go, and J. H. Lee, "A text-based visual context modulation neural model for multimodal machine translation," Pattern Recogn. Lett. **136**, 212–218 (2020).
43. C. He, Y. Wan, Y. Gu, and F. L. Lewis, "Integral reinforcement learning-based multi-robot minimum time-energy path planning subject to collision avoidance and unknown environmental disturbances," IEEE Control Syst. Lett. **5**, 983–988 (2021).
44. B. Sangiovanni, G. P. Incremona, M. Piastra, and A. Ferrara, "Self-configuring robot path planning with obstacle avoidance via deep reinforcement learning," IEEE Control Syst. Lett. **5**, 397–402 (2021).
45. J. Yoo, D. Jang, H. J. Kim, and K. H. Johansson, "Hybrid reinforcement learning control for a micro quadrotor flight," IEEE Control Syst. Lett. **5**, 505–510 (2021).
46. Q. Zhang, X. Wan, S. Liu, J. Y. Yin, L. Zhang, and T. J. Cui, "Shaping electromagnetic waves using software-automatically-designed metasurfaces," Sci. Rep. **7**, 3588 (2017).
47. L. L. Li, H. X. Ruan, C. Liu, Y. Li, Y. Shuang, A. Alu, C. W. Qiu, and T. J. Cui, "Machine-learning reprogrammable metasurface imager," Nat. Commun. **10**, 1082 (2019).
48. T. Qiu, X. Shi, J. Wang, Y. Li, S. Qu, Q. Cheng, T. Cui, and S. Sui, "Deep learning: a rapid and efficient route to automatic metasurface design," Adv. Sci. **6**, 1900128 (2019).

49. Q. Ma, G. D. Bai, H. B. Jing, C. Yang, L. Li, and T. J. Cui, "Smart metasurface with self-adaptively reprogrammable functions," Light Sci. Appl. **8**, 98 (2019).
50. H. Li, H. Zhao, M. Wei, H. Ruan, Y. Shuang, T. J. Cui, P. del Hougne, and L. Li, "Intelligent electromagnetic sensing with learnable data acquisition and processing," Patterns **1**, 100006 (2020).
51. L. Li, Y. Shuang, Q. Ma, H. Li, H. Zhao, M. Wei, C. Liu, C. Hao, C.-W. Qiu, and T. J. Cui, "Intelligent metasurface imager and recognizer," Light Sci. Appl. **8**, 97 (2019).
52. L. Li, T. J. Cui, W. Ji, S. Liu, J. Ding, X. Wan, Y. B. Li, M. Jiang, C.-W. Qiu, and S. Zhang, "Electromagnetic reprogrammable coding-metasurface holograms," Nat. Commun. **8**, 197 (2017).
53. J. Wu, Z. Wang, L. Zhang, Q. Cheng, S. Liu, S. Zhang, J. Song, and C. T. Jun, "Anisotropic metasurface holography in 3D space with high resolution and efficiency," IEEE Trans. Antennas Propag. **69**, 302–316 (2020).
54. C. Liu, Q. Ma, L. Li, and T. J. Cui, "Work in progress: intelligent metasurface holograms," in *1st ACM International Workshop on Nanoscale Computing, Communication, and Applications* (2020), pp. 45–48.
55. Z. Y. Zhou, J. Xia, J. Wu, C. L. Chang, X. Ye, S. G. Li, B. T. Du, H. Zhang, and G. D. Tong, "Learning-based phase imaging using a low-bit-depth pattern," Photon. Res. **8**, 1624–1633 (2020).
56. R. Lopez, J. Regier, M. I. Jordan, and N. Yosef, "Information constraints on auto-encoding variational Bayes," in *32nd Conference on Neural Information Processing Systems (NeurIPS)* (2018), pp. 6114–6125.
57. I. Deshpande, Z. Zhang, and A. Schwing, "Generative modeling using the sliced Wasserstein distance," in *IEEE Conference on Computer Vision and Pattern Recognition* (2018), pp. 3483–3491.
58. M. Arjovsky and L. Bottou, "Towards principled methods for training generative adversarial networks," in *5th International Conference on Learning Representations (ICLR)* (2017), pp. 1–17.
59. V. M. Panaretos and Y. Zemely, "Statistical aspects of Wasserstein distances," arXiv:1806.05500v3 (2019).
60. M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein generative adversarial networks," in *34th International Conference on Machine Learning (ICML)* (2017), pp. 214–223.
61. I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville, "Improved training of wasserstein GANs," in *31st Annual Conference on Neural Information Processing Systems (NIPS)* (2017), pp. 5768–5778.
62. M. Mirza, "Conditional generative adversarial nets," arXiv: 1411.1784v1 (2014).
63. P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2017), pp. 5967–5976.
64. T.-C. Wang, M.-Y. Liu, J.-Y. Zhu, A. Tao, J. Kautz, and B. Catanzaro, "High-resolution image synthesis and semantic manipulation with conditional GANs," in *31st Meeting of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2018), pp. 8798–8807.
65. K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *21st ACM Conference on Computer and Communications Security (CCS)* (2016), pp. 630–645.
66. S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," in *32nd International Conference on Machine Learning (ICML)* (2015), pp. 448–456.
67. D. Ulyanov and A. Vedaldi, "Instance normalization: the missing ingredient for fast stylization," arXiv:1607.08022v3 (2017).
68. Y. LeCun, C. Cortes, and C. J. C. Burges, "The MNIST database of handwritten digits," http://yann.lecun.com/exdb/mnist/ (2012).
69. X. Zou, G. Zheng, Q. Yuan, W. Zang, R. Chen, T. Li, L. Li, S. Wang, Z. Wang, and S. Zhu, "Imaging based on metalenses," PhotoniX **1**, 2 (2020).
70. A. Ahad, M. Tahir, and K. A. Yau, "5G-based smart healthcare network: architecture, taxonomy, challenges and future research directions," IEEE Access **7**, 100747 (2019).