





# Neural network-based surrogate model for inverse design of metasurfaces

GUOQING JING,<sup>1,†</sup> PEIPEI WANG,<sup>1,†</sup> HAISHENG WU,<sup>1,†</sup> JIANJUN REN,<sup>1</sup> ZHIQIANG XIE,<sup>1</sup> JUNMIN LIU,<sup>2</sup>   
 HUAPENG YE,<sup>3</sup>  YING LI,<sup>1</sup>  DIANYUAN FAN,<sup>1</sup> AND SHUQING CHEN<sup>1,\*</sup> 

<sup>1</sup>International Collaborative Laboratory of 2D Materials for Optoelectronics Science and Technology, Institute of Microscale Optoelectronics, Shenzhen University, Shenzhen 518060, China

<sup>2</sup>College of New Materials and New Energies, Shenzhen Technology University, Shenzhen 518118, China

<sup>3</sup>Guangdong Provincial Key Laboratory of Optical Information Materials and Technology and Institute of Electronic Paper Displays, South China Academy of Advanced Optoelectronics, South China Normal University, Guangzhou 510006, China

\*Corresponding author: shuqingchen@szu.edu.cn

Received 6 December 2021; revised 24 April 2022; accepted 24 April 2022; posted 25 April 2022 (Doc. ID 450564); published 20 May 2022

Metasurfaces composed of spatially arranged ultrathin subwavelength elements are promising photonic devices for manipulating optical wavefronts, with potential applications in holography, metalens, and multiplexing communications. Finding microstructures that meet light modulation requirements is always a challenge in designing metasurfaces, where parameter sweep, gradient-based inverse design, and topology optimization are the most commonly used design methods in which the massive electromagnetic iterations require the design computational cost and are sometimes prohibitive. Herein, we propose a fast inverse design method that combines a physics-based neural network surrogate model (NNSM) with an optimization algorithm. The NNSM, which can generate an accurate electromagnetic response from the geometric topologies of the meta-atoms, is constructed for electromagnetic iterations, and the optimization algorithm is used to search for the on-demand meta-atoms from the phase library established by the NNSM to realize an inverse design. This method addresses two important problems in metasurface design: fast and accurate electromagnetic wave phase prediction and inverse design through a single phase-shift value. As a proof-of-concept, we designed an orbital angular momentum (de)multiplexer based on a phase-type metasurface, and 200 Gbit/s quadrature-phase shift-keying signals were successfully transmitted with a bit error rate approaching  $1.67 \times 10^{-6}$ . Because the design is mainly based on an optimization algorithm, it can address the “one-to-many” inverse problem in other micro/nano devices such as integrated photonic circuits, waveguides, and nano-antennas. © 2022 Chinese Laser Press

<https://doi.org/10.1364/PRJ.450564>

## 1. INTRODUCTION

Metasurfaces composed of spatially arranged micro/nano antennas introduce abrupt phase changes to electromagnetic waves through strong resonances of the meta-atoms [1–4], and they have been applied to metalenses [5–8], holograms [9–12], and multiplexing communication [13–16]. One of the most significant challenges in metasurface design comes from how to find meta-atoms with desired electromagnetic response [17–20]. The traditional design process mainly includes three steps: selecting the appropriate structure of meta-atoms, sweeping the electromagnetic response of meta-atoms to obtain the target geometrical parameters, and using meta-atoms arranged in arrays to form a metasurface [21–25]. These processes require numerous trial and error sessions, and the scale and range of the sweeping drastically increase with the design parameters [26]. For the sweeping, the target meta-atoms need to fall within the sweep range of the initial structure, causing

the selection to be heavily dependent on the design parameters and resulting in possible design limitations for specific electromagnetic waves [27,28].

To address these problems, various inverse design methods have been proposed to predict the target meta-atoms according to the desired electromagnetic responses [29]. The inverse design methods provide a new design strategy for generating non-intuitive geometric structures and significantly improve the design efficiency. These inverse design methods can be divided as follows: adjoint gradient inverse optimization [30,31] and deep learning inverse prediction [32–34]. The former strategy utilizes the gradient descent approach to optimize and modify the structural parameters according to the returned adjoint field variables [35]. The implementation of this approach is limited by the expensive full-wave simulation requirements, slow convergence, and manufacturing difficulties owing to the massive electromagnetic iteration, small step size, and non-intuitive

structures. By contrast, the deep learning methods can significantly reduce the computational cost by replacing the electromagnetic calculations with neural networks [36]. Using the target electromagnetic response (such as the transmittance spectrum or reflectance spectrum) as the input, the trained neural network model can predict the structures of the meta-atoms. However, the neural network model usually yields the shape and patterns of the meta-atoms as the output, which shows some prediction errors in the inverse design [27], and the predicted meta-atoms are always irregular and require further treatment for practical engineering. Moreover, there are two additional technical challenges in the inverse designing of phase-type metasurfaces [37]. (1) The phase value introduces wavelength-dependent periodic changes, because of which the phase spectrum usually contains irregular jumping discontinuities, leading to difficulty in the convergence of the neural network model during training. (2) It is difficult to obtain the required phase spectrum of the meta-atoms in advance because the modulation of the metasurface relies on the abrupt phase change at a specific wavelength, and only a single phase-shift value is involved in the design.

Combining the physics-based neural network and optimization algorithm, we constructed a physics-based neural network surrogate model (NNSM) to predict the electromagnetic responses of arbitrary manufacturable meta-atoms by training small dataset meta-structure samples among the design variables. This NNSM can replace the electromagnetic calculations of the geometric topologies of the meta-atoms and their electromagnetic response, and significantly reduce the computational cost. Benefiting from the trained NNSM model, the predicted results of the meta-atoms matched well with those of the electromagnetic calculations outside the training set. Therefore, the NNSM established a meta-atom library as a fast pattern-searching dictionary, which provides an efficient tool for deriving the corresponding meta-atoms through the predicted results. The optimization algorithm is used to search for specific meta-structures that meet the light modulation requirements from the library established by NNSM. The meta-atom geometric topology is randomly generated by the computer and sent to the neural network for prediction, and the prediction results are screened following the rules of natural evolution, where the results close to the design requirements are easier to retain. After multiple iterations, the optimization algorithm can search for on-demand meta-atoms, realizing a fast and accurate inverse design. We show that the inverse design process takes only a few seconds, and the two important goals in designing phase-type metasurfaces are accomplished: (1) the predicted meta-atoms can be tolerated and directly applied to the metasurface without further treatment, and (2) the on-demand meta-atoms can be inversely designed using only one phase-shift value. As a proof-of-concept, we designed an orbital angular momentum (OAM) (de)multiplexer based on a phase-type metasurface, and 200 Gbit/s quadrature-phase shift-keying (QPSK) signals were successfully transmitted with the bit error rate (BER) approaching  $1.67 \times 10^{-6}$  at the received optical power of  $-20.5$  dBm. This indicates that this physical-optimization algorithm presents an effective solution for the “one-to-many” problem

(the same optical response can be derived from multiple geometrical parameters) in the inverse design of a metasurface, and it also has the ability to optimize the structure in multi-parameter micro/nano devices.

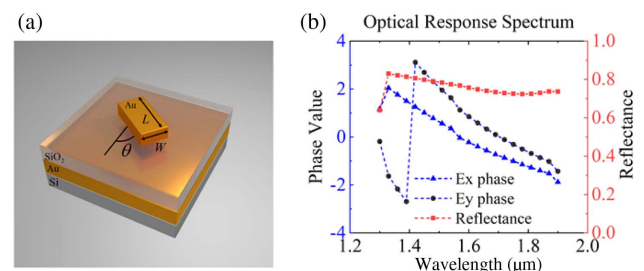
## 2. PRINCIPLES AND METHODS

### A. Geometry of Meta-Atoms and Optical Response Spectrum

To collect sufficient training samples, the generic geometry of the subwavelength structure on the top of a reflective substrate was selected as the unit cell structure for the investigation. The unit cells were composed of multiple layers. Here, we used gold as the meta-atom, silica as the spacer, and gold above the silicon wafer as the substrate. When the unit cells are illuminated with light, the meta-atoms with different structures generate different optical responses [38]. For the feasibility of manufacturing, we fixed the height of the meta-atoms and selected the simplest cuboid as the basic structure. As shown in Fig. 1(a), the design parameters are defined as the length  $L$ , width  $W$ , and rotation angle  $\theta$  of the meta-atoms. By changing these design parameters, we can control the phase-shift modulation values of the meta-atoms to modulate the light field further. To monitor the optical response of the unit cell comprehensively, we measured its reflectance spectrum and phase spectra in  $x$  and  $y$  polarization directions with a bandwidth from 1300 nm to 1900 nm using the finite-difference time-domain (FDTD) method. Considering that, in some applications, optical modulation is only required at a specific wavelength, we discretized these spectra and selected 21 points on each spectrum for the investigations. Therefore, the optical response spectrum data consisted of three parts: phase spectrum in  $x$  polarization, phase spectrum in  $y$  polarization, and reflectance spectrum, with a total of 63 data points, as shown in Fig. 1(b).

### B. Physics-Based NNSM

The designed NNSM is shown in Fig. 2, and its basic framework is a U-net convolutional neural network (CNN), which contains five convolutional layers, four transposed convolutional layers, and three fully connected hidden layers with 2304, 512, and 128 neurons, respectively. The specific architecture and hyperparameters of the U-net CNN are presented in Appendix A.1. The input layer is an image that matches the geometric topologies of the meta-atoms, and the output layer has 63 neurons, corresponding to the optical response spectrum that needs to be predicted. The data transmission between two adjacent convolution layers satisfies



**Fig. 1.** (a) Unit cell of reflection-type metasurface. (b) Optical response spectrum data simulated using the FDTD method.

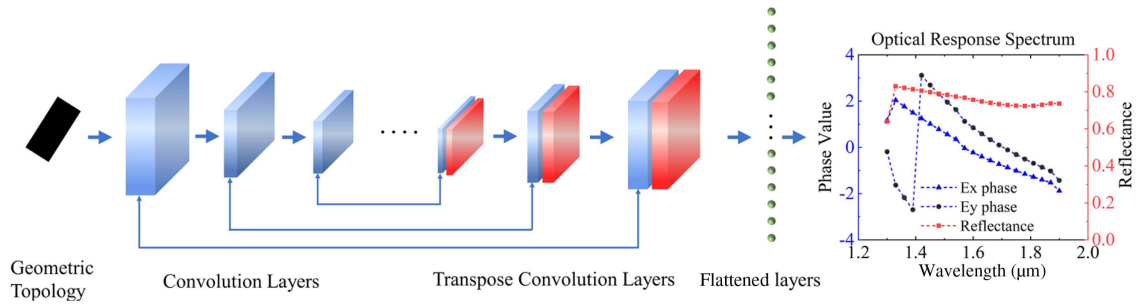


Fig. 2. Diagram of physics-based NNSM.

$$Y_j^m = f\left(b_j^m + \sum_{i \in M_j} X_i^{m-1} \cdot k_{ij}^m\right), \quad (1)$$

where  $M_j$  represents a group of feature images in the  $(m-1)$ th layer. The input  $X_i^{m-1}$  and output  $Y_j^m$  represent the  $i$ th and  $j$ th feature images in the  $(m-1)$ th and  $m$ th layers, respectively. The term  $k_{ij}^m$  denotes the convolution kernel with a size of  $3 \times 3$  associated with  $X_i^{m-1}$  and  $Y_j^m$ . The term  $b_j^m$  denotes the bias value, and  $f$  is the nonlinear activation function. We add a pooling layer after each conventional operation to reduce the dimensions of the feature images and network parameters. With rounds of convolution and pooling operations, the intrinsic features can be learned from the input image. To reduce the number of training samples, we introduced data enhancement into a traditional CNN. When performing the convolution and pooling operations, each output result was recorded. Once the network completes all the convolution operations, the previously recorded convolution results are concatenated to the same size deconvolution layer as the input of the next layer to form a U-net structure. Finally, the U-net is followed by several fully connected layers to obtain the desired outputs. The transmission relationship in the adjacent fully connected layers can be expressed as follows:

$$x_j^{m+1} = f_i^{m+1}\left(\sum_{i=1}^J w_{ij}^{m+1} x_i^m + b_i^{m+1}\right), \quad (2)$$

where  $x_j^m$  denotes the  $j$ th neuron of the  $m$ th layer,  $w_{ij}^{m+1}$  represents the weight from the  $j$ th neuron of the  $m$ th layer to the  $i$ th neuron of the next layer. Here, Leaky-Relu ( $a = 0.1$ ) is used as the activation function to perform the nonlinear activation on the output of the convolutional layer and the transposed convolutional layer. In the fully connected layers (flattened layers), tanh is used as the activation function, which can be expressed as follows:

$$\tanh(y_j) = \frac{1 - \exp(-2y_j)}{\sum_j 1 + \exp(2y_j)}, \quad (3)$$

where  $y_j$  denotes the  $j$ th output neuron. A dropout function with rate of 0.5 is employed on the back convolutional layers to avoid overfitting. The mean squared error (MSE) is used as the loss function, which can be calculated as follows:

$$\text{MSE}[f(X, \theta), Y] = \frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2, \quad (4)$$

where  $y_i$  denotes the ideal output,  $\hat{y}_i$  is the predicted output, and  $m$  is the neuron number of the output layer. In the training process, the adaptive moment estimation (Adam) optimizer is introduced to update the weight and bias parameters to minimize the loss function.

### C. Inverse Design Principle Based on the NNSM and Intelligent Optimization Algorithm

To design the on-demand meta-atoms, we propose an inverse design strategy, the schematic of which is shown in Fig. 3, consisting of an NNSM based on the U-net and an intelligent optimization algorithm. The basic idea of intelligent optimization algorithms comes from traditional heuristic search algorithms, such as genetic algorithms and particle swarm optimization algorithms [39,40]. By imitating the evolution process of organisms in nature, the best design parameters can be found in the search space (the principle and parameters of the optimization algorithm are shown in Appendix A.2). Here, the unit cell of a reflection-type metasurface containing design parameters is known as an individual, which contains three variables: length, width, and rotation angle. Considering the actual manufacturing accuracy and the sensitivity of the device to the parameters,

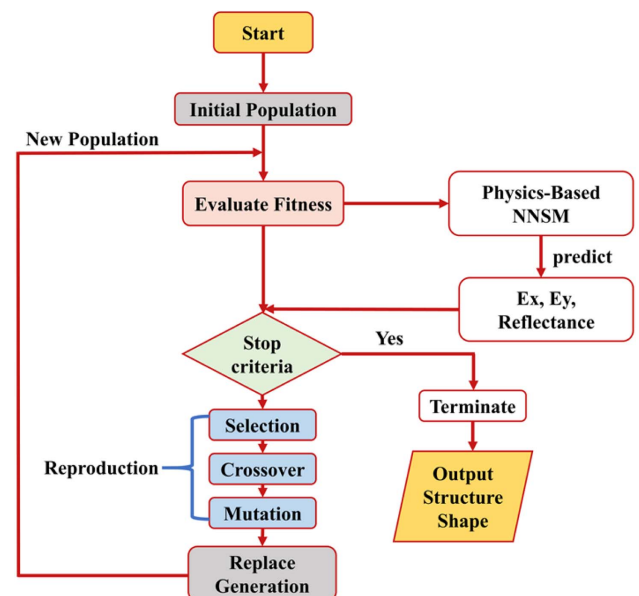


Fig. 3. Schematic of the inverse design process of the intelligent optimization algorithm.

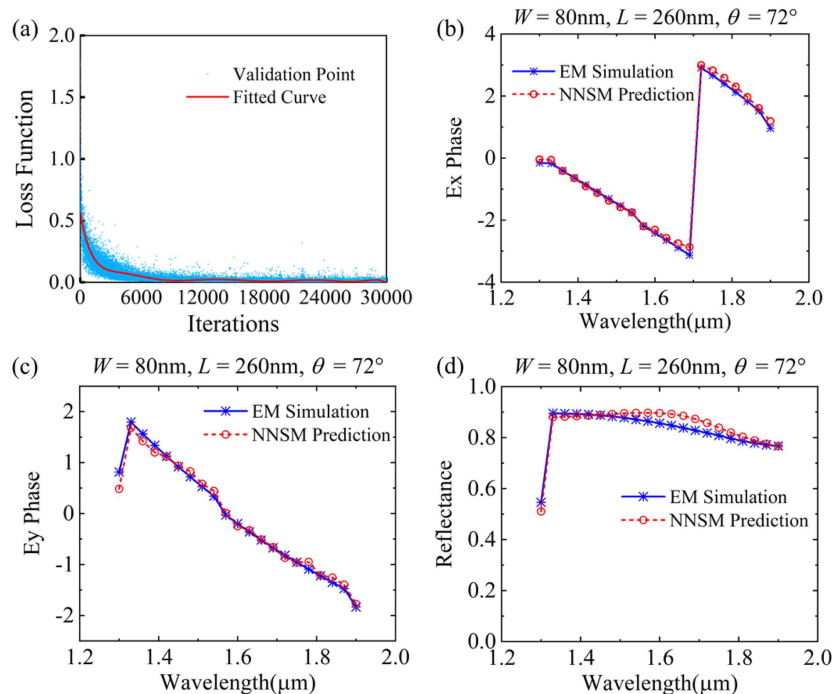
we selected the length  $L$  and width  $W$  to change from 50 nm to 500 nm with a length interval of  $\Delta l = 5$  nm, and the angle changing from  $0^\circ$  to  $180^\circ$  with  $\Delta\theta = 5^\circ$ ; therefore, each individual has a total of  $90 \times 90 \times 36 = 291,600$  different combinations. To evaluate the inverse design performance of the intelligent optimization algorithm, we use the fitness score to calculate the degree of matching between the predicted optical responses and the optimal values of individuals, and the score increases with the similarity. Considering the application background of the meta-device, we hope that the NNSM will put more effort into fitting the phase values of the  $x$  and  $y$  polarizations during the optimization process and will utilize different proportions of weighting to calculate the overall score. In particular, the phases in the  $E_x$  and  $E_y$  directions and reflectivity are weighted by 0.4, 0.4, and 0.2, respectively, and the final score is the sum of these three scores. Here, we select 1000 individuals with random design parameters as a population and calculate the fitness value of each individual in the initial process; thereafter, the NNSM is used to calculate the optical response results. In comparison with complex electromagnetic-based calculations, the trained NNSM can quickly obtain results within a few milliseconds, which saves significant amounts of time and computational cost. After the fitness score evaluation, the discriminator will screen out the individual with the highest fitness score in the population and determine whether it meets the threshold condition. If the fitness score does not meet the threshold, individuals with higher scores will be selected for cross-reproduction and gene mutation to produce the next generation, thereby performing multiple iterations. Here, the crossover and mutation rates are set as 0.8 and 0.1, respectively. According to these steps, we complete the entire selection process using the roulette-wheel selection method. Once the fitness

value meets the set threshold condition, the search optimization process is terminated, and the optimal design parameters are output for further manufacturing.

### 3. RESULTS AND ANALYSIS

#### A. Prediction Results of NNSM

The constructed NNSM model was used as a tool to predict the optical response to accelerate the electromagnetic calculation process. When the design parameters of the meta-atom are input to the neural network as a geometric image, the neural network can promptly provide a predicted optical response spectrum. According to the range of meta-atom design parameters ( $W \in [50, 500]$ ,  $L \in [50, 500]$ ,  $\theta \in [0^\circ, 180^\circ]$ ), we selected 15, 15, and 10 data points for three dimensions, respectively, and a total of 2250 samples were obtained, of which 2000 were used for training and the remaining 250 were used for validation. The training data are used to generate the gradients and optimize weights, and the validation data are used to verify whether the neural network model has learned the physical mapping relationship between the input and ideal output. Figure 4(a) shows the change trend of the loss function of the validation set during the training process. It shows that the predicted spectrum gradually overlaps with the actual optical response spectrum, and the MSE is only 0.02 with iterations reaching  $3 \times 10^4$ , and the average accuracy of the prediction results is 0.952, indicating that the NNSM has an excellent spectrum prediction ability. Furthermore, we randomly selected a meta-structure in the validation set to verify the prediction results of NNSM, as illustrated in Figs. 4(b)–4(d), which almost perfectly coincides with the simulated spectrum. In addition, it can be noticed that the jump discontinuity



**Fig. 4.** Predicted results of NNSM. (a) Loss curve as a function of training iteration. (b) Phase spectrum in  $x$  polarization. (c) Phase spectrum in  $y$  polarization. (d) Reflectance spectrum.

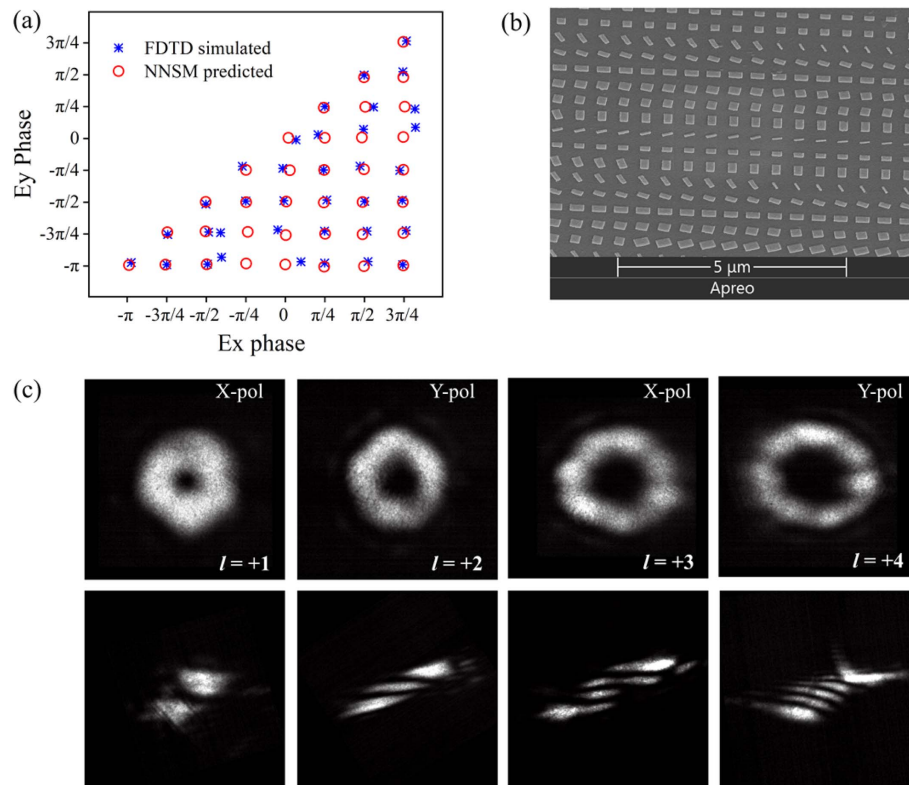
points in the phase spectrum have also been accurately predicted, indicating that the model has high generalizability and has learned the physics-based mapping relationship between input and output.

## B. Meta-Atoms Inverse Design and Fabrication

Using the inverse design model, we can arrange the searched meta-atoms into a metasurface according to the modulation requirements [41]. As a proof-of-concept, we designed an OAM mode modulator, which can be used as a mode (de)multiplexer and employed for OAM communication [42]. In this study, we employed a metasurface with transmission phase modulation to realize OAM (de)multiplexing. For the phase modulation of  $[0, 2\pi]$ , we selected eight points in a phase period, which were  $-3\pi/4, -\pi/2, -\pi/4, 0, \pi/4, \pi/2, 3\pi/4,$  and  $\pi$ . Because the OAM mode is independent of the polarization dimension (the topological charge  $l$  is used to represent the OAM mode), to improve the applicability of the OAM (de)multiplexer, we designed it to independently respond to the  $x$  and  $y$  polarizations simultaneously. For this task, we need to find a total of  $8 \times 8 = 64$  phase points to achieve simultaneous phase control of the  $x$  and  $y$  polarizations. Because the phase points in the  $x$  and  $y$  directions perform a symmetrical distribution, we only need to find 36 different meta-atoms to meet the aforementioned phase modulation requirements. Using the presented inverse design model, we searched the structural parameters of these 36 special meta-atoms, and the results are displayed in

Fig. 5(a). For instance, for a modulation requirement with a phase value of  $\pi/2$  in  $x$  polarization and a phase value of  $\pi/4$  in  $y$  polarization, the optimal design parameters of meta-atoms can be calculated by evaluating the maximum fitness score. Following this procedure, the meta-atoms that meet all the modulation phase requirements can be obtained and arranged into a complete metasurface. Because NNSM is used to replace the FDTD to calculate the optical response of meta-atoms, there are specific errors in the phase values between the inverse model prediction and the FDTD-based calculation [see Fig. 5(a)], but the errors are tolerated in the metasurface-based phase modulation for OAM mode communications.

For the arranged metasurface OAM mode (de)multiplexer, we used the standard electron beam lithography (EBL; EBPG 5150) to fabricate it (see Appendix A.3 for manufacturing details), and the scanning electron microscopy (SEM) image of the fabricated metasurface is shown in Fig. 5(b). The fabricated metasurface can, thereafter, be utilized to modulate the vortex beams for OAM mode (de)multiplexing. For the mode multiplexer, the Gaussian beams with incident angles of  $-10^\circ$  and  $-20^\circ$  in the  $x$  polarization, and  $+10^\circ$  and  $+20^\circ$  in the  $y$  polarization projected on the metasurface, can generate coaxial vortex beams with OAM modes of  $+1, +3, +2,$  and  $+4$ , respectively. From the far-field transmission results shown in Fig. 5(c), the generated vortex beams are consistent with the theory. In addition, the measurement results of these four vortex beams after passing through the cylindrical lens (C-lens) show that the



**Fig. 5.** (a) Comparison between the predicted results by the inverse design model and the ideal results calculated by FDTD method. (b) SEM image of the fabricated metasurfaces. (c) First line represents the intensity distributions of vortex beams with different OAM modes generated by Gaussian beam with different incident angles and polarization states, and the second line represents the intensity distributions of vortex beams after passing through the C-lens.

fringe number is equal to the carried OAM mode, indicating that the proposed inverse design model can accurately design the corresponding metasurface and provide an effective scheme for the design of high-performance micro/nano photonic devices. For the demultiplexer, the coaxial input vortex beams can be diffracted to different spatial positions by the same metasurface, and the OAM modes are degenerated into a Gaussian mode, thereby realizing OAM mode demultiplexing.

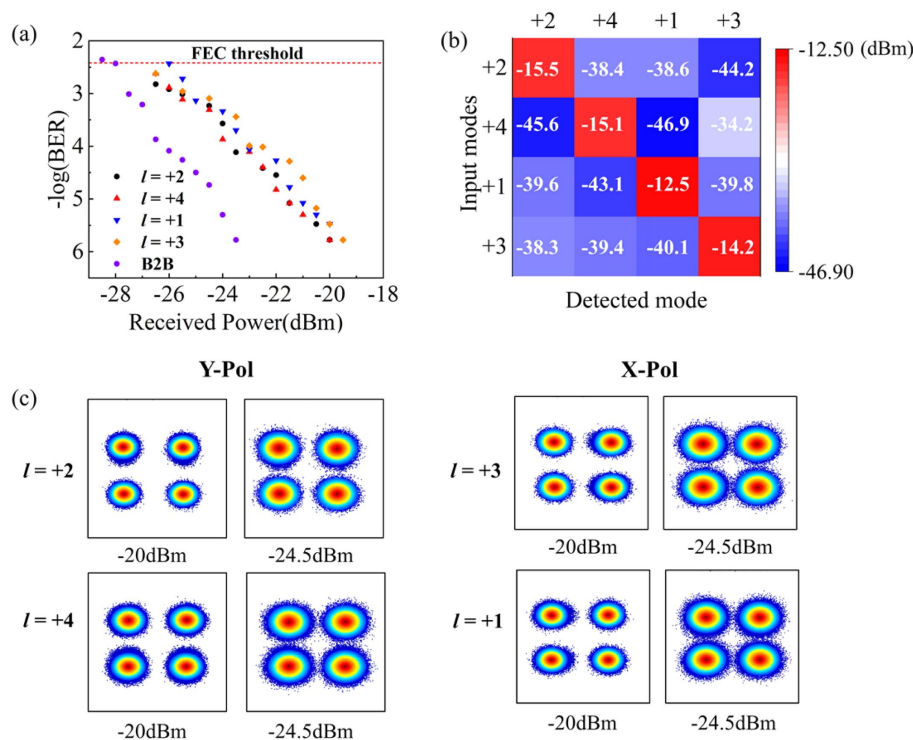
### C. OAM Multiplexing Communication

To confirm the feasibility of the designed metasurface, we constructed an OAM mode-division multiplexing link and employed the designed OAM mode multiplexer/demultiplexer to perform channel multiplexing/demultiplexing (details of the communication link are provided in Appendix A.4). Figure 6 shows the experimental results. Here, the transmitter generates four OAM mode channels with modes of +1, +2, +3, and +4 and realizes coaxial multiplexing. After 1 m free-space transmission, the four vortex beams are demultiplexed by the demultiplexer and converted to original digital signals. By using a real-time digital signal analyzer (DSA) and processing it offline in a computer, the four output electrical waveforms can be recorded, and the information can be restored. The measured BERs at different optical received powers are shown in Fig. 6(a). It is observed that the BERs of all the OAM channels are below the forward-error-correction (FEC) threshold ( $3.8 \times 10^{-3}$ ) for the received optical power of  $-20$  dBm, indicating that the OAM multiplexing link possesses low signal cross talk. Figure 6(b) shows the signal and noise powers in this OAM multiplexing link, where the signal power is listed

on the diagonal of the matrix, and the other OAM mode channels are the noise sources of the signal channel. After calculation, the optical signal-to-noise ratios corresponding to the OAM channels with  $l = +2, +4, +1,$  and  $+3$  are  $-19.95, -20.01, -23.41,$  and  $-18.62$  dB, demonstrating that the OAM beam multiplexing link possesses a low mode cross talk. Constellations of the recovered QPSK signals at different received powers are shown in Fig. 6(c); when the received power is  $20$  dBm, the constellation points converge very well, and there is almost no cross talk. As the received power decreases, the constellation points are more likely to overlap and induce the BERs to increase. However, even if the received power is reduced to  $-24.5$  dBm, the BERs of the corresponding channels are still greater than the FEC, implying that the inverse-designed metasurface has satisfactory multiplexing and demultiplexing ability.

### 4. DISCUSSION

The inverse prediction of the meta-atom phase is always a significant issue that needs to be addressed in metasurface design. Because the phase changes periodically with the working wavelength, the corresponding phase spectrum curve will have abrupt spikes. In comparison with traditional inverse design methods that utilize smooth data distribution (transmittance spectrum or reflectance spectrum) as training samples, the neural network model trained by the phase spectrum is more difficult to converge. In the previous neural network-based phase inverse design, the inputs are generally the geometric parameters of meta-atoms, and the intrinsic geometric information is

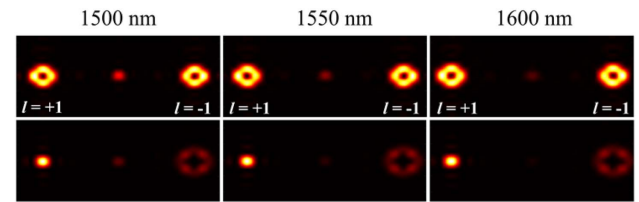


**Fig. 6.** Results of metasurfaces-based four channels OAM multiplexing communication. (a) BER as a function of received power for different OAM modes. (b) Signal and noise powers corresponding to OAM multiplexing. (c) Constellations of different channels at received optical powers of  $20$  dBm and  $24.5$  dBm.

easily lost because only a few geometric parameters participate in the training process. Here, the input is set as the meta-atom geometry images that contain more feature information, and this can make the prediction results easier to match with the truth spectrum. By designing an NNSM to train the samples, the model has a good generalization capacity with less training samples. Generally, as a data-driven simulator, the prediction performance increases with the number of sample data, which will become increasingly accurate in finding the implicit inherent connections between meta-atom structures and their optical spectrum response. However, more samples require more collection time; thus, there is a trade-off when selecting the number of samples.

The proposed inverse design model consists of two parts, optimization algorithm and NNSM, where the optimization algorithm is used to screen the optimal structure, and the NNSM is used as a full-wave simulator. Here, we use the genetic algorithm for structural optimization, which can be replaced by other algorithms, such as the ant colony algorithm, particle swarm algorithm, and gradient-based algorithm. Among them, the gradient-based algorithm only needs a few iterations to obtain a better nanostructure [35], which can be used as a potential alternative to shorten the iterative calculation time. But it is worth noting that, in these optimization algorithms, complex full-wave simulations are usually required, and the constructed NNSM can still be used as a simulator to further reduce the computation time. For NNSM, we select CNN as the basic architecture. Compared with other full-wave simulation networks, such as the predicting neural network (PNN) [33], the U-net CNN-based NNSM designed in this work has two main advantages. (1) Benefiting from the connection between the convolution layer and deconvolution layer, the NNSM can extract deeper feature information from the nanostructure patterns, making it well predict the spectrum with phase mutations. (2) By splicing the output information of the forward convolutional layers into the input of deconvolutional layers in the later part of the model, the NNSM can perform data enhancement on feature information, and it still has strong robustness and generalization ability under the training of less samples.

To further demonstrate the powerful phase prediction ability of NNSM at multiple wavelengths, we use the proposed inverse design model to optimize the structure of meta-atoms with constant phase response in the wavelength range from 1500 nm to 1600 nm, and we further design a broadband OAM generator. By setting the optimization constraint that the phase response of meta-atoms in the wavelength range from 1500 nm to 1600 nm satisfies  $\varphi_x - \varphi_y = \pi$ , a meta-atom with length  $L = 260$  nm, width  $W = 410$  nm, and rotation angle  $\Delta\theta = 0^\circ$  is obtained. Under the geometric phase modulation [13], by rotating this meta-atom to different degrees and then arranging them, an OAM generator with broadband response can be obtained, which can be used for dense wavelength division multiplexing to further improve the capacity of OAM communications. As a proof-of-principle, we simulated a binary vortex grating that can simultaneously generate OAM modes with topological charges of +1 and -1, where the pixel size and area of the transverse simulation area are  $64 \times 64$  and



**Fig. 7.** Test results of the metasurface at different wavelengths (1500 nm, 1550 nm, 1600 nm). Top row, far-field light intensity distributions of right circularly polarized plane waves. Bottom row, detection results of projecting a vortex beam with  $l = -1$  on the OAM generator.

$51.2 \mu\text{m} \times 51.2 \mu\text{m}$ . Figure 7 shows the test results of broadband response characteristics. The top row shows the far-field light intensity distributions of right circularly polarized plane waves with different working wavelengths (1500 nm, 1550 nm, 1600 nm) after passing the OAM generator. Further, we detect the generated two OAM modes by projecting a vortex beam with  $l = -1$ , and the results are shown in the second row. It can be seen that the vortex beam with  $l = +1$  on the left is restored to a Gaussian beam, and the light spot on the right is transformed into a vortex beam with  $l = -2$ . At the same time, since the grating has different diffraction angles for different wavelengths, there are slightly differences in the positions of the far-field diffraction light spots between the three wavelengths, which also verifies that the NNSM has powerful phase prediction ability for multiple wavelengths.

The inverse design model is based on the accurate and fast prediction of NNSM. Unlike previous methods that directly predict the meta-atom structure via the ideal phase response, our model relies on the constant interaction between the NNSM and the optimization algorithm. A large number of meta-atom images with design parameters were randomly generated and sent to the NNSM for prediction. After the selectivity of the optimization algorithm, one meta-atom with a spectral response close to the ideal values can be obtained by setting the appropriate fitness. Because a specific phase-shift value can be selected as the optimization object, one or multiple phase values can be set as the optimization target, and the inverse prediction of the meta-atom structure can be realized. Therefore, it presents an effective solution for the “one-to-many” problem in the inverse design of micro/nano devices.

## 5. CONCLUSION

In this study, we proposed a fast and accurate inverse design method to design a metasurface. Utilizing the output of the convolutional layers to concatenate the input of the deconvolutional layers for data enhancement, we realized the direct prediction of the abrupt phase spectrum. Based on the interaction between the NNSM and the optimization algorithm, we designed a novel inverse design model that can inversely predict on-demand meta-atoms using a single phase-shift value. With this inverse model, we designed an OAM (de)multiplexer based on a phase-type metasurface, and 200 Gbit/s QPSK signals were successfully transmitted with a BER approaching  $1.67 \times 10^{-6}$  at a received optical power of  $-20$  dBm. This

indicates that, using this method, we can effectively realize the inverse design of a meta-atom structure from a single on-demand phase shift. Such an inversely designed meta-atom could be applied to the design of other complex electromagnetic modulators, such as integrated photonic devices and optical antennas.

## APPENDIX A: METHODS

### 1. Architecture of NNSM

The basic structure of NNSM is shown in Fig. 8, where the orange arrows represent convolution operations and the red arrows represent transpose convolution operations. After performing each convolution operation, we employ the leaky-Relu function for nonlinear activation. In the process of full connection, we employed the tanh function for nonlinear activation.

### 2. Principle of Optimization Algorithm

The basic idea of an intelligent optimization algorithm (also known as genetic algorithm) is to imitate the evolution process of organisms in nature. Each combination can be denoted as an individual, represented by a binary code, and one individual represents one combination. Therefore, each combination has its own code, and decoding the individual can obtain its specific design parameters. These binary codes are known as “chromosomes,” where the binary fragment of each design parameter is known as a gene. The individuals of the next generation are created by crossover and mutation of the parent chromosomes. During selection, individuals with lower fitness values will be eliminated, whereas individuals with higher fitness values will survive and reproduce more offspring. Therefore, genes with higher fitness values will have a greater probability of being passed on to the next generation. Generally, the average fitness value of the offspring is higher than that of the parents. During crossover, the parent chromosomes are intermixed and create new chromosomes constituting the population of the children to replace the parent populations. Notably, the crossover point was randomly selected. Assuming that the parent 1 is “11|0100|10” and parent 2 is “00|0010|00,” after crossover, we can get child 1 “11|0010|10” and child 2 “00|0100|00.” During mutation, “genes” in newly generated chromosomes will mutate with a specific probability, that is, a change from 0 to 1, or from 1 to 0. Here, the crossover rate and mutation rate are set as 0.8 and 0.1, respectively. According to these steps, we complete the entire selection process using the roulette-wheel selection method.

### 3. Fabrication of the Designed Metasurface

We used standard EBL (EBPG 5150) to fabricate the designed meta-atoms. First, a gold background layer (220 nm) and a SiO<sub>2</sub> spacer (160 nm) were deposited onto a 15 mm × 15 mm silicon substrate using an electron beam evaporator (ASB-EPI-C6). Notably, a very thin layer of titanium (Ti) exists between the SiO<sub>2</sub> substrate and gold (Au) background layer (approximately 2 nm), which is used to increase the adhesion between SiO<sub>2</sub> and the Au background layer to prevent the Au from falling off during the subsequent degumming process. Thereafter, the positive polymethyl methacrylate (PMMA, 950 K) resist film was spin-coated on the SiO<sub>2</sub> spacer layer. The sample

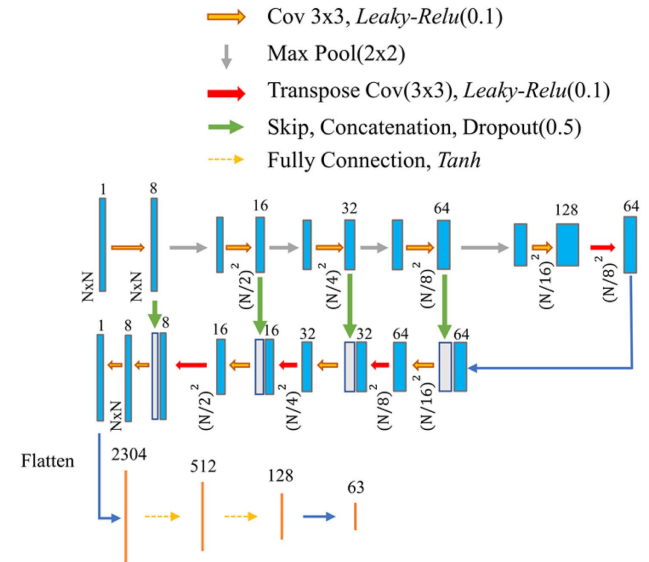


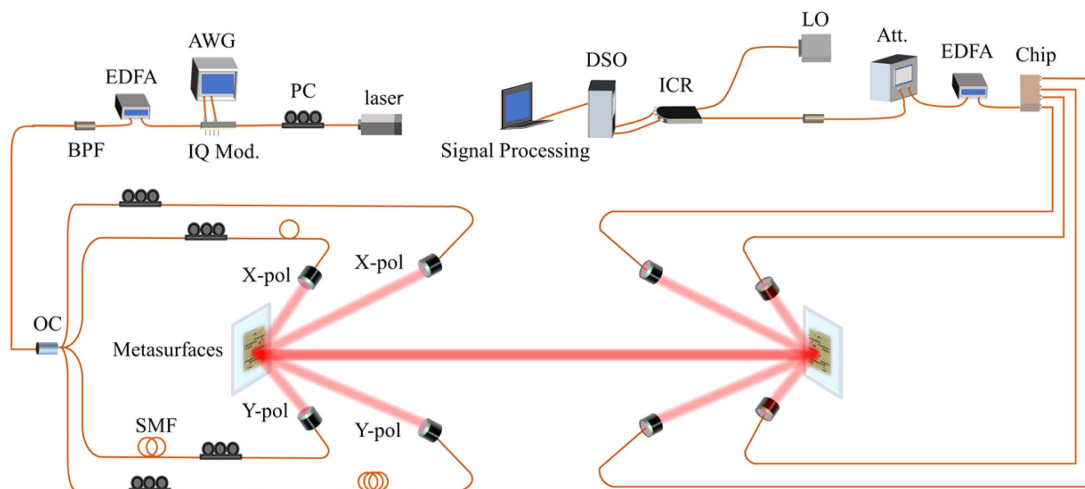
Fig. 8. Specific architecture of physics-based NNSM.

was placed into the homogenizer, and a speed of 4000 r/min was selected for homogenization for 1 min; baking was performed at 180°C for 1 min and 30 s, and a 150-nm-thick PMMA layer was obtained. Because the height of our gold nano-antennas is 50 nm, the thickness of PMMA must be higher than 150 nm (triple the height of gold nano-antennas). Thereafter, the nanostructures were etched on the PMMA film by EBL, and the area of a single metasurface was 640 μm × 640 μm. After etching by EBL, we need to develop and fix the sample, which can be completed by placing the sample in the developer for 1 min, then putting it in the stopper solution for 30 s, and finally drying the remaining stopper solution with nitrogen. Next, a 50 nm gold film was deposited on the sample via thermal evaporation. Finally, the sample was placed in a 30% acetone solution and allowed to stand for 6 h. The beaker was, thereafter, shaken to remove the excess PMMA and Au to obtain the metasurface structure. The presented design can ensure large resonance and high reflectivity at a working wavelength of 1550 nm.

### 4. OAM Multiplexing Communication System

An OAM mode-division multiplexing link was constructed to confirm the feasibility of the inverse-designed metasurface. As shown in Fig. 9, a laser with a working wavelength of 1550 nm was launched into the polarization controller (PC), and QPSK signals with a transmission rate of 50 Gbit/s were loaded using an arbitrary waveform generator (AWG 7122C) and an IQ modulator. After amplification by an erbium-doped fiber amplifier (EDFA) and filtration using a bandpass filter (BPF), the signal light was divided into four subchannels through an optical coupler (OC). Thereafter, these four subchannels are polarization-modulated by PCs, where two channels are *x* polarizations and the other two channels are *y* polarizations. To realize channel multiplexing, four Gaussian signal lights are decorrelated by single-mode fibers (SMFs) of different lengths, carrying different signals. The four Gaussian beams, each of which carried QPSK signals, were incident on the metasurface





**Fig. 9.** Metasurfaces-based four channels OAM multiplexing communication link. PC, polarization controller; IQ Mod., in-phase/quadrature modulator; EDFA, erbium-doped fiber amplifier; AWG, arbitrary waveform generator; BPF, bandpass filter; OC, optical coupler; SMF, single-mode fiber; Att., attenuator; LO, local oscillator; ICR, integrated coherent receiver; DSO, digital sampling oscilloscope.

at different angles to generate four coaxially vortex beams with OAM modes ( $l = +1, +2, +3, +4$ ) and realize multiplexing. After collimation and transmitting 1 m in free space, the coaxial vortex beams were demultiplexed by another identical metasurface. Thereafter, the demultiplexed four channels were received and converted into the original digital signals one by one. In particular, at the receiver, the demultiplexed signals were amplified by the EDFA and filtered by a BPF to improve the receiver sensitivity, and the attenuator was used to adjust the received optical power. After being detected by the integrated coherent optical receiver (ICR), the four output electrical waveforms were recorded using a real-time DSA and processed offline in a computer.

**Funding.** Shenzhen Peacock Plan (20180521645C, 20180921273B); China Postdoctoral Science Foundation (2020M682867); Shenzhen Excellent Scientific and Technological Innovative Talent Training Program (RCBS20200714114818094); Shenzhen Universities Stabilization Support Program (SZWD2021013); Science and Technology Project of Shenzhen (GJHZ20180928160407303); Shenzhen Fundamental Research Program (JCYJ20210324095611030, JCYJ20210324095610027); Basic and Applied Basic Research Foundation of Guangdong Province (2019A1515111153, 2020A1515011392, 2020A1515110572, 2021A1515011762); National Natural Science Foundation of China (12047539, 61805149, 62101334).

**Acknowledgment.** G. J. and P. W. conceived the idea of this research. G. J. performed the simulations and built a neural network model. G. J. wrote the paper. H. W. provided assistance for fabrication of metasurface. J. R., Z. X., J. M., and S. C. shared their insights and contributed to discussions on the results. H. Y., Y. L., D. F., and S. C. supervised the project.

**Disclosures.** The authors declare no competing financial interests.

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

<sup>†</sup>These authors contributed equally to this paper.

## REFERENCES

1. 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).
2. X. G. Luo, "Principles of electromagnetic waves in metasurfaces," *Sci. China Phys. Mech. Astron.* **45**, 1–18 (2015).
3. S. Chen, Z. Li, Y. Zhang, H. Cheng, and J. Tian, "Phase manipulation of electromagnetic waves with metasurfaces and its applications in nanophotonics," *Adv. Opt. Mater.* **6**, 1800104 (2018).
4. Y. Ke, S. Chen, W. Shu, and H. Luo, "Generation of perfect vector beams based on the combined modulation of dynamic and geometric phases," *Opt. Commun.* **446**, 191–195 (2019).
5. W. T. Chen, A. Y. Zhu, V. Sanjeev, M. Khorasaninejad, Z. Shi, E. Lee, and F. Capasso, "A broadband achromatic metalens for focusing and imaging in the visible," *Nat. Nanotechnol.* **13**, 220–226 (2018).
6. I. Tanriover and H. V. Demir, "Broad-band polarization-insensitive all-dielectric metalens enabled by intentional off-resonance waveguiding at mid-wave infrared," *Appl. Phys. Lett.* **114**, 043105 (2019).
7. M. Khorasaninejad, Z. Shi, A. Y. Zhu, W. T. Chen, V. Sanjeev, A. Zaidi, and F. Capasso, "Achromatic metalens over 60 nm bandwidth in the visible and metalens with reverse chromatic dispersion," *Nano Lett.* **17**, 1819–1824 (2017).
8. M. Khorasaninejad, W. T. Chen, R. C. Devlin, J. Oh, A. Y. Zhu, and F. Capasso, "Metalenses at visible wavelengths: diffraction-limited focusing and subwavelength resolution imaging," *Science* **352**, 1190–1194 (2016).
9. G. Zheng, H. Mühlenbernd, M. Kenney, G. Li, T. Zentgraf, and S. Zhang, "Metasurface holograms reaching 80% efficiency," *Nat. Nanotechnol.* **10**, 308–312 (2015).
10. W. Ye, F. Zeuner, X. Li, B. Reineke, S. He, C. W. Qiu, J. Liu, Y. Wang, S. Zhang, and T. Zentgraf, "Spin and wavelength multiplexed nonlinear metasurface holography," *Nat. Commun.* **7**, 11930 (2016).
11. X. Li, L. Chen, Y. Li, X. Zhang, M. Pu, Z. Zhao, X. Ma, Y. Wang, M. Hong, and X. Luo, "Multicolor 3D meta-holography by broadband plasmonic modulation," *Sci. Adv.* **2**, e1601102 (2016).

12. Q. Xiao, Q. Ma, T. Yan, L. W. Wu, C. Liu, Z. X. Wang, X. Wan, Q. Cheng, and T. J. Cui, "Orbital-angular-momentum-encrypted holography based on coding information metasurface," *Adv. Opt. Mater.* **9**, 2002155 (2021).
13. H. Tan, J. Deng, R. Zhao, X. Wu, G. Li, L. Huang, J. Liu, and X. Cai, "A free-space orbital angular momentum multiplexing communication system based on a metasurface," *Laser Photon. Rev.* **13**, 1800278 (2019).
14. Q. Mai, C. Wang, X. Wang, S. Cheng, M. Cheng, Y. He, J. Xiao, H. Ye, D. Fan, Y. Li, and S. Chen, "Metasurface based optical orbital angular momentum multiplexing for 100 GHz radio-over-fiber communication," *J. Lightwave Technol.* **39**, 6159–6166 (2021).
15. Z. Jin, D. Janoschka, J. Deng, L. Ge, P. Dreher, B. Frank, G. Hu, J. Ni, Y. Yang, J. Li, C. Yu, D. Lei, G. Li, S. Xiao, S. Mei, H. Giessen, F. zu Heringdorf, and C. Qiu, "Phyllotaxis-inspired nanosieves with multiplexed orbital angular momentum," *eLight* **1**, 5 (2021).
16. J. Wang, J. Y. Yang, I. M. Fazal, N. Ahmed, Y. Yan, H. Huang, Y. Ren, Y. Yue, S. Dolinar, M. Tur, and A. E. Willner, "Terabit free-space data transmission employing orbital angular momentum multiplexing," *Nat. Photonics* **6**, 488–496 (2012).
17. J. A. Fan, "Freeform metasurface design based on topology optimization," *MRS Bull.* **45**, 196–201 (2020).
18. W. Cai, D. Zhu, Z. Liu, L. Raju, and A. S. Kim, "Building multifunctional metasystems via algorithmic construction," *ACS Nano* **15**, 2318–2326 (2021).
19. C. Sitawarin, W. Jin, Z. Lin, and A. W. Rodriguez, "Inverse-designed photonic fibers and metasurfaces for nonlinear frequency conversion [Invited]," *Photon. Res.* **6**, B82–B89 (2018).
20. K. Wang, J. Zhao, Q. Cheng, D. S. Dong, and T. J. Cui, "Broadband and broad-angle low-scattering metasurface based on hybrid optimization algorithm," *Sci. Rep.* **4**, 5935 (2014).
21. A. C. Overvig, S. Shrestha, S. C. Malek, M. Lu, A. Stein, C. Zheng, and N. Yu, "Dielectric metasurfaces for complete and independent control of the optical amplitude and phase," *Light Sci. Appl.* **8**, 92 (2019).
22. A. Arbabi, Y. Horie, M. Bagheri, and A. Faraon, "Dielectric metasurfaces for complete control of phase and polarization with subwavelength spatial resolution and high transmission," *Nat. Nanotechnol.* **10**, 937–943 (2015).
23. J. Tang, Z. Li, S. Wan, Z. Wang, C. Wan, C. Dai, and Z. Li, "Angular multiplexing nanoprinting with independent amplitude encryption based on visible-frequency metasurfaces," *ACS Appl. Mater. Interfaces* **13**, 38623–38628 (2021).
24. H. Feng, Q. Li, W. Wan, J. H. Song, Q. Gong, M. L. Brongersma, and Y. Li, "Spin-switched three-dimensional full-color scenes based on a dielectric meta-hologram," *ACS Photon.* **6**, 2910–2916 (2019).
25. J. Jang, G. Y. Lee, J. Sung, and B. Lee, "Independent multichannel wavefront modulation for angle multiplexed meta-holograms," *Adv. Opt. Mater.* **9**, 2100678 (2021).
26. X. Shi, T. Qiu, J. Wang, X. Zhao, and S. Qu, "Metasurface inverse design using machine learning approaches," *J. Phys. D* **53**, 275105 (2020).
27. Z. Liu, D. Zhu, S. P. Rodrigues, K. T. Lee, and W. Cai, "Generative model for the inverse design of metasurfaces," *Nano Lett.* **18**, 6570–6576 (2018).
28. R. Zhu, T. Qiu, J. Wang, S. Sui, C. Hao, T. Liu, Y. Li, M. Feng, A. Zhang, C. W. Qiu, and S. Qu, "Phase-to-pattern inverse design paradigm for fast realization of functional metasurfaces via transfer learning," *Nat. Commun.* **12**, 2974 (2021).
29. D. Xu, Y. Luo, J. Luo, M. Pu, Y. Zhang, Y. Ha, and X. Luo, "Efficient design of a dielectric metasurface with transfer learning and genetic algorithm," *Opt. Mater. Express* **11**, 1852–1862 (2021).
30. M. Minkov, I. A. D. Williamson, L. C. Andreani, D. Gerace, B. Lou, A. Y. Song, T. W. Hughes, and S. Fan, "Inverse design of photonic crystals through automatic differentiation," *ACS Photon.* **7**, 1729–1741 (2020).
31. S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vucković, and A. W. Rodriguez, "Inverse design in nanophotonics," *Nat. Photonics* **12**, 659–670 (2018).
32. L. Jiang, X. Li, Q. Wu, L. Wang, and L. Gao, "Neural network enabled metasurface design for phase manipulation," *Opt. Express* **29**, 2521–2528 (2021).
33. S. An, C. Fowler, B. Zheng, M. Y. Shalaginov, H. Tang, H. Li, L. Zhou, J. Ding, A. M. Agarwal, C. Rivero-Baleine, K. A. Richardson, T. Gu, J. Hu, and H. Zhang, "A deep learning approach for objective-driven all-dielectric metasurface design," *ACS Photon.* **6**, 3196–3207 (2019).
34. R. Zhu, T. Qiu, J. Wang, S. Sui, Y. Li, M. Feng, H. Ma, and S. Qu, "Multiplexing the aperture of a metasurface: inverse design via deep-learning-forward genetic algorithm," *J. Phys. D* **53**, 455002 (2020).
35. W. He, M. Tong, Z. Xu, Y. Hu, X. Cheng, and T. Jiang, "Ultrafast all-optical terahertz modulation based on an inverse-designed metasurface," *Photon. Res.* **9**, 1099–1108 (2021).
36. Q. Zhang, H. Yu, M. Barbiero, B. Wang, and M. Gu, "Artificial neural networks enabled by nanophotonics," *Light Sci. Appl.* **8**, 42 (2019).
37. I. Tanriover, W. Hadibrata, and K. Aydin, "Physics-based approach for a neural networks enabled design of all-dielectric metasurfaces," *ACS Photon.* **7**, 1957–1964 (2020).
38. P. Xu, H. W. Tian, W. X. Jiang, Z. Z. Chen, T. Cao, C. W. Qiu, and T. J. Cui, "Phase and polarization modulations using radiation-type metasurfaces," *Adv. Opt. Mater.* **9**, 2100159 (2021).
39. X. Jiang, H. Yuan, D. Chen, Z. Zhang, T. Du, H. Ma, and J. Yang, "Metasurface based on inverse design for maximizing solar spectral absorption," *Adv. Opt. Mater.* **9**, 2100575 (2021).
40. C. Lu, Z. Liu, Y. Wu, Z. Xiao, D. Yu, H. Zhang, C. Wang, X. Hu, Y. C. Liu, X. Liu, and X. Zhang, "Nanophotonic polarization routers based on an intelligent algorithm," *Adv. Opt. Mater.* **8**, 1902018 (2020).
41. S. Chen, Z. Xie, H. Ye, X. Wang, Z. Guo, Y. He, Y. Li, X. Yuan, and D. Fan, "Cylindrical vector beam multiplexer/demultiplexer using off-axis polarization control," *Light Sci. Appl.* **10**, 22 (2021).
42. W. Shu, Y. Liu, Y. Ke, X. Ling, Z. Liu, B. Huang, H. Luo, and X. Yin, "Propagation model for vector beams generated by metasurfaces," *Opt. Express* **24**, 21177–21189 (2016).