ADVANCED OPTICAL MATERIALS

Supporting Information

for *Adv. Optical Mater.*, DOI: 10.1002/adom.202001433

Multifunctional Metasurface Design with a Generative Adversarial Network

Sensong An, Bowen Zheng, Hong Tang, Mikhail Y. Shalaginov, Li Zhou, Hang Li, Myungkoo Kang, Kathleen A. Richardson, Tian Gu, Juejun Hu, Clayton Fowler, and Hualiang Zhang**

Supporting Information

Title: Multifunctional Metasurface Design with a Generative Adversarial Network

Sensong An[1], Bowen Zheng[1], Hong Tang[1], Mikhail Y. Shalaginov[2], Li Zhou[1], Hang Li[1], Myungkoo Kang[3], Kathleen A. Richardson[3], Tian Gu[2], Juejun Hu[2], Clayton Fowler[1,], Hualiang Zhang[1,*]*

1. Detailed network architecture

Fig. S1 illustrates the network architecture of the proposed neural network, with the data flow details included. Input conditions (1×1×N), including amplitude and phase responses, either frequency-dependent or polarization-dependent, were firstly concatenated with the randomly generated latent vector (1×1×100) to form a conditioned noise (1×1×(N+100)). This conditioned noise was then passed to the generator in order to produce fake meta-atom samples (64×64×1) that resemble real ones (64×64×1) well enough to confuse the Discriminator. Subsequently, the input conditions (1×1×N) were spatially tiled into 3D matrixes (64×64×N) and concatenated with real meta-atom samples (64×64×1) or fake meta-atom samples (64×64×1). The Wasserstein distance between the real and fake conditioned meta-atom samples (64×64×(N+1)) were later measured by the discriminator. During the training process, the discriminator can learn by maximizing the Wasserstein distance between the fake and real samples, while the generator aims to fool the discriminator by generating fake designs with minimized Wasserstein distance comparing to the real ones. The trainings of the generator and the discriminator were executed in turns, meaning that the parameters in the generator were fixed during the training of the discriminator, and vice versa. To stabilize the training, parameters in the discriminator were updated 3 times in each iteration, while the generator was trained once. Both modules approach the real data distribution through parameter tuning during this adversarial process. The training ended when losses for both the Generator and the Discriminator were stabilized. More detailed architectures of the generator and discriminator are shown in Fig. S1b and S1c, respectively.

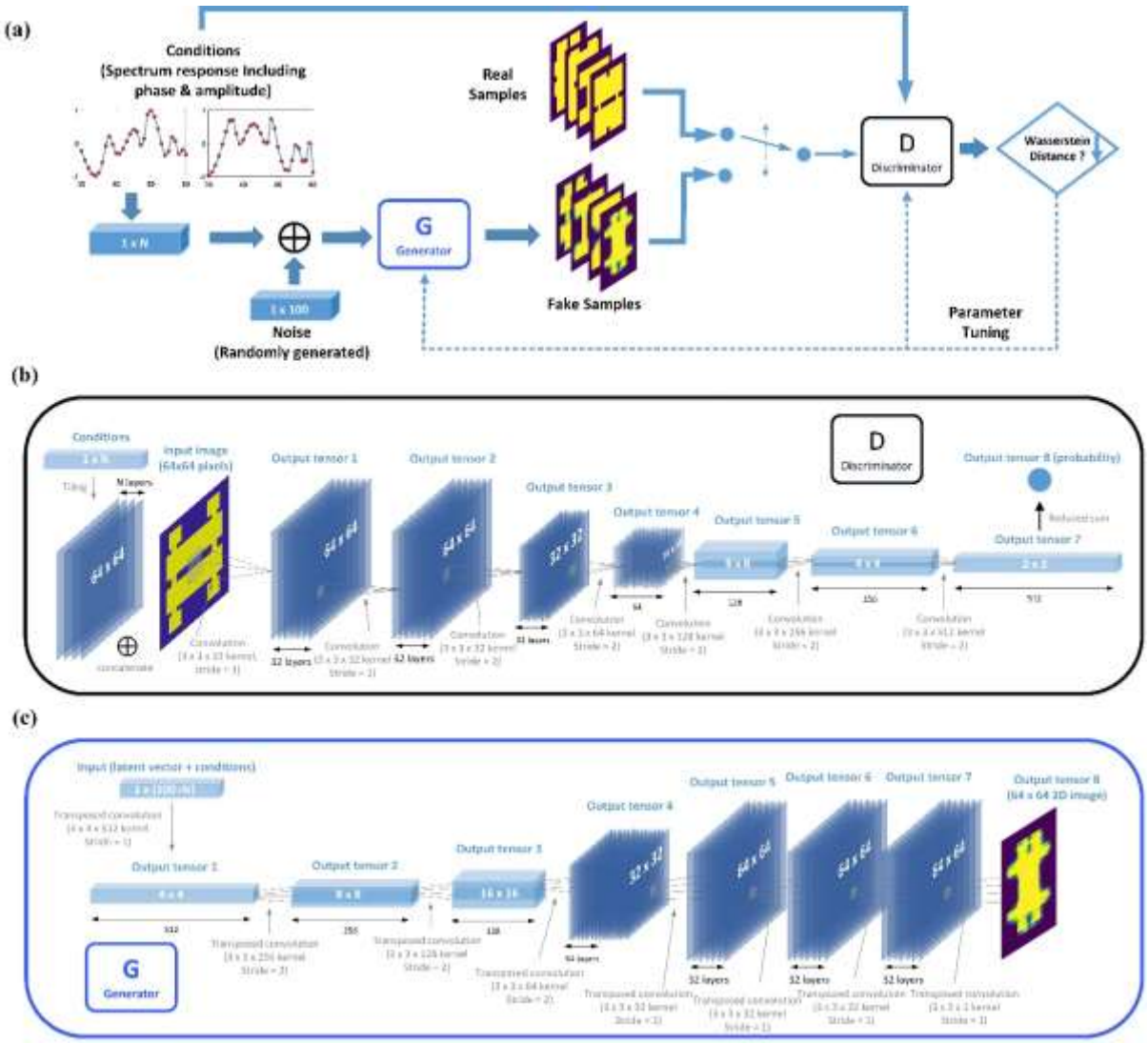


Figure S1. Detailed network architectures of the GAN. (a) Illustration of the training process of the GAN. (b) Detailed network structure of the discriminator. The discriminator consists of seven consecutive convolution layers. The output of each layer is batch-normalized and passed through a ReLU activation function before being passed on to the next layer. The conditions are tiled into a 3D vector (64×64×N) and then concatenated with the 2D image (64×64×1). The initially planar dimensions of the combined input (64×64×(N+1)) are decreased while the depths are increased via (2, 2) stride convolutional layers. The output tensor of layer #7 is flattened into a 1D array and the reduced sum is calculated to represent its Wasserstein distance. (c) Detailed network structure of the generator. Conversely, the generator consists of eight consecutive transposed convolution layers for which the depth of an output tensor is decreased while gradually being flattened into a 2D meta-atom image. The output of each layer is batch-normalized and passed through a Leaky ReLU activation function. After the last transposed convolutional layer, a *tanh* activation function generates a 2D image representing the meta-atom design. Details of each output tensor, shapes of the (transposed) convolutional kernels and strides used during convolutions are given in the figure.

For more design DOF, despite that the original input images (with 28 × 28 pixels resolution) that are used to sketch the 2D shape of the meta-atoms from the training dataset, the input images were all rescaled into 64 × 64 pixel images before they were processed using convolutional layers. The

generator and the discriminator were trained alternately: after the discriminator was trained with 3 different batches of training data, values of the hidden neurons in each convolutional layer are updated and fixed. The generator was then trained with a new batch of training data with the help of the most up-to-date discriminator, and so on, until training completes. Details, including dimensions of each output tensor, shapes of the convolutional kernels and strides used during convolutions are given in Fig. S1b and S1c.

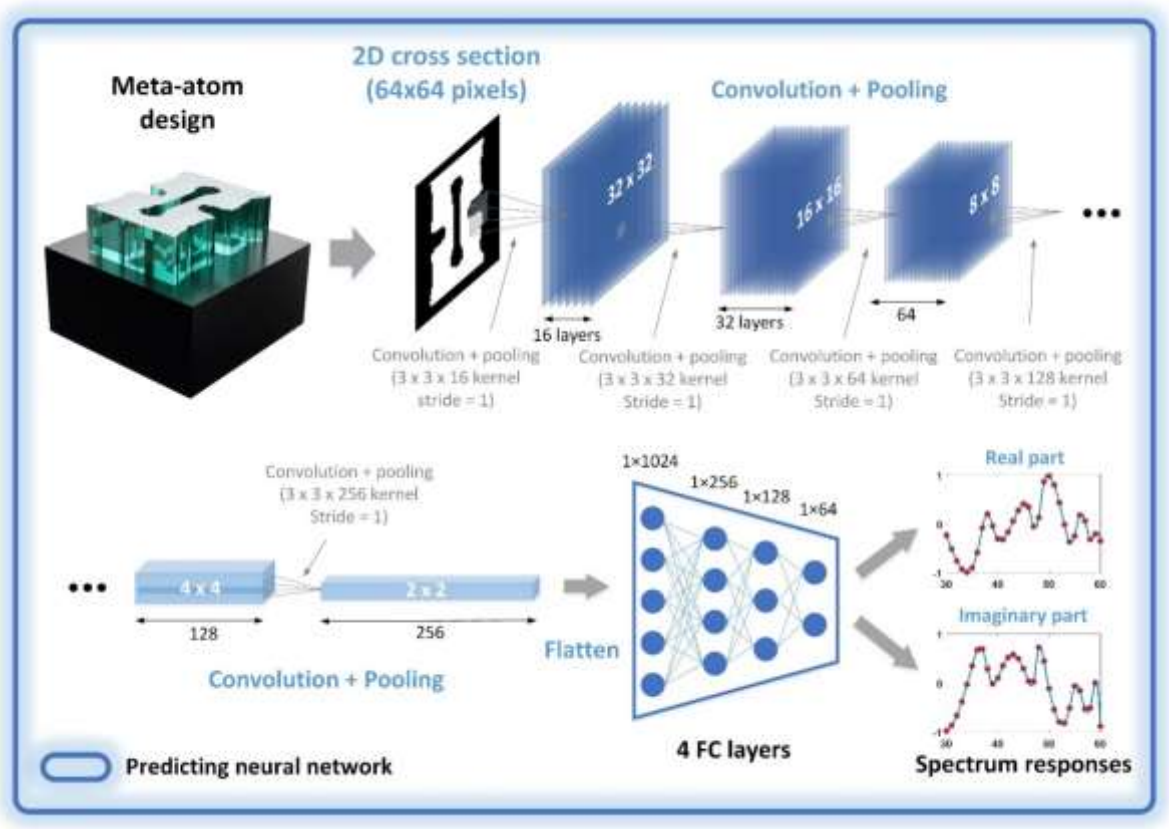


Figure S2. Detailed network architectures of the predicting neural network (PNN). Details including dimensions of each output tensors, shapes of the (transposed) convolutional kernels and strides used during convolutions are given in the figure.

Following the approaches introduced in [57], we constructed two PNNs (Fig. S2) to predict the real and imaginary parts of the transmission spectrum, respectively. Transmissive amplitude and phase are then derived using the predicted real and imaginary parts. Detailed network architecture of the PNN cascaded to the GAN is shown in Fig. S2. The PNN was constructed based on a convolutional neural network (CNN) architecture. It functions as a critic and examines the performance of the designs generated by the proposed GAN. Specifically, the PNN is able to precisely predict the transmission spectrum of free-form meta-atom designs within the frequency range of 30 to 60 THz. In contrast to

full-wave simulation tools, the PNN characterizes the meta-atoms on a one-time calculation basis and, thus, significantly speeds up the whole design process.

2. Training data collection

Without loss of generality, the all-dielectric meta-atom consists of a 1 μm thick dielectric component (preferably with a high refractive index, n_1. In this case n_1 = 5) sitting on a dielectric substrate (preferably with a low refractive index, n_2. In this case n_2 = 1.4) with a unit cell size of 2.8 × 2.8 μm^2 (Fig. S3a). The 2D pattern of each meta-atom was generated with the "needle drop" approach using the numerical computing tool MATLAB. Several (3 to 7) rectangular bars, with a minimum generative resolution of 0.1 μm, were randomly generated and placed together within a square lattice to form random patterns (Fig. S3b). To minimize inter-cell coupling, a minimum spacing of 0.4 μm was applied between adjacent meta-atoms. To speed up the data-collection process, the all-dielectric components are only generated in the top left quadrant of each unit cell and then symmetrically replicated along *x* and *y* axes to form the whole pattern. A set of meta-atoms generated in this manner is guaranteed to possess polarization-diverse performance.

The full-wave electromagnetic simulations were performed using a commercial FEM simulation tool CST. For each meta-atom, perfect electric conducting surface (E_t = 0) and perfect magnetic conducting surface (H_t = 0) boundary conditions were employed to calculate the transmission and phase shift of a square lattice structure. Open boundaries are applied along both the negative and positive *z* directions, while an *x*-polarized plane wave was illuminated from the substrate side for each meta-atom. To further accelerate the full-wave simulations, E_t = 0 and H_t = 0 symmetry planes were applied in the center *y*-*z* plane and *x*-*z* plane for each meta-atom, respectively. The frequency range of interest is set to be between 30 to 60 THz. A total number of 69,000 meta-atoms with different shapes were generated and simulated to find their wide-spectrum phase and amplitude responses. These simulations were performed on eight servers running in parallel. The data collection process was completed in 3 days. After removing similar patterns (to speed up the training), 29,000 meta-atom structures were selected and documented for further training.

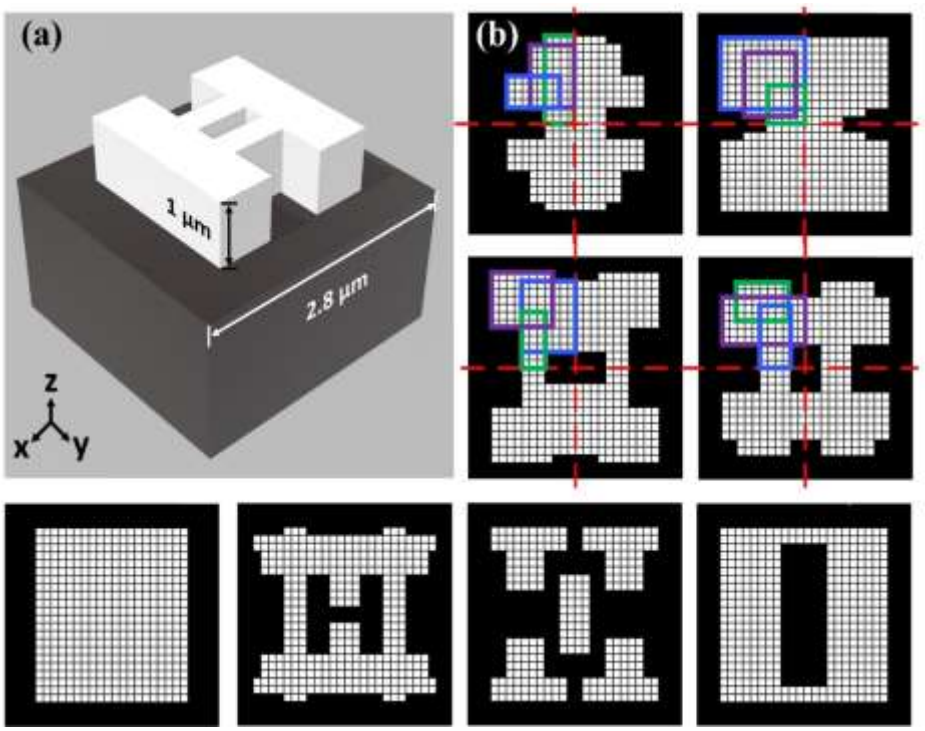


Figure S3. Training data collection process. (a) 3D view of a generated meta-atom of arbitrary shape. Lattice constant for each meta-atom cell is 2.8 μm, meta-atom height is fixed to be 1 μm. White and black colors represent high- and low-index (substrate) dielectric components. **(b)** Demonstration of the pattern generation process. 2D patterns in x-y plane are meshed, each mesh pixel has a dimension of 0.1 by 0.1 μm^2. Rectangles outlined in different colors represent distinct high-index "needles" that were randomly generated and dropped on the top-left quadrant of the substrate canvas. Patterns were completed by mirroring the pattern along the *x* and *y* axes.

3. Customized gradient-penalty method

The Wasserstein distance is only accurate when the discriminator is a 1-Lipschitz function.[45] To enforce this constraint, the original WGAN applied a simple, but rough, value clipping to restrict the maximum weight value in each layer of the discriminator. Instead, WGAN-GP uses a gradient penalty term to ensure that the norm of its gradients is equal to 1 almost everywhere [46] so that the discriminator is 1-Lipschitz. Conditional-Wasserstein distance with a gradient penalty term can be represented as:

$$W(P_{data}, P_G) \approx \sup_{\|D\|_L \leq 1} \Big\{ \mathbb{E}_{y \sim P_{data}}[D(y|x)] - \mathbb{E}_{y' \sim P_G}[D(y'|x)] - \lambda \mathbb{E}_{y \sim P_{penalty}}[\max(0, \|\nabla_y (D(y))\| - 1)] \Big\} \quad \text{(S1)}$$

Traditional WGAN-GP randomly interpolates between network generated patterns $P_{generator}$ and real sample patterns P_{data} to generate gradient penalty samples $P_{penalty}$, as shown in Fig. S4a. The

interpolation method has an important advantage: as the training progresses and $P_{generator}$ approaches P_{data}, the gradient norm of this more widespread distribution $P_{penalty}$, instead of the real sample distribution P_{data}, satisfies the Lipschitz constraint, and we can thus conclude that the discriminator is 1-Lipschitz almost everywhere within P_{data}.

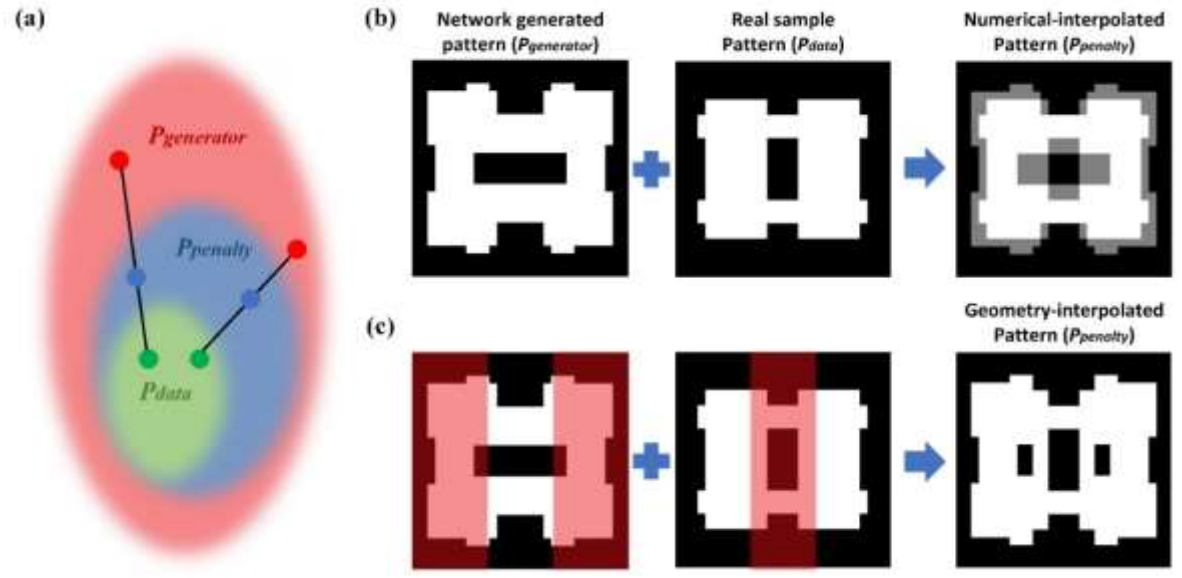


Figure S4. A novel interpolation method for customized gradient penalty. (a) Schematic diagram of the random interpolation process in WGAN-GP. The network randomly interpolates between $P_{generator}$ and P_{data} to get $P_{penalty}$. (b) Numerical interpolation methods employed by traditional WGAN-GPs. The meta-atom patterns are binarized, such that white represents "1" and black represents "0". (c) The proposed novel geometry interpolation method. Random proportions (marked in red) were taken from both fake samples and real samples and later combined into a new pattern.

In our case, with meta-atom patterns as target design goals, the generated outputs can be converted into binary images consisting of 1's that represent the dielectric material and 0's that represent voids. The conventional numerical interpolation process is not applicable for the meta-atom discriminator, because generated values between "0" and "1" don't correspond to any physical structures (Fig. S4b). As a result, the discriminator that is trained to satisfy the Lipschitz constraint for this $P_{penalty}$ is intuitively challenged to yield stable Wasserstein distance results for real samples from P_{data} during the training. We therefore employed a novel geometry interpolation method that combines random geometry portions from both $P_{generator}$ and P_{data} to form the sample in $P_{penalty}$ (Fig. S4c). The interpolated results obtained in this manner fully characterize the samples between $P_{generator}$ and P_{data}. This unique interpolation method also allows the generator to extrapolate and explore the

ground truth distribution when the training data is insufficient to cover the whole design space. Training experiments which validate the training stability, design accuracy and extrapolation capability of the proposed gradient-penalty method are presented in Section 4.

4. Hyperparameters and training curves

Table S1. Hyperparameters used in the training of GANs and PNNs.

Hyperparameters	**Meta-atom design network**	**Dual-polarization meta-atom design network**	**PNN for real part**	**PNN for imaginary part**
Training set size	*29,000*	*29,000*	*69,000*	*69,000*
Optimizer (learning rate)	Adam *(1e-4)*	Adam *(1e-4)*	Adam *(1e-4)*	Adam *(1e-4)*
Batch size	*64*	*64*	*256*	*256*
Batch Norm.	Yes	Yes	No	No
Nonlinear activations	ReLU for the discriminator, Leaky ReLU (alpha = 0.2) for the generator, *tanh* for the generator's last layer		ReLU	
Penalty coefficient	*10*	*10*	NA	NA
Iterations (time)	*3,000 (72* h*)*	*3,000 (72* h*)*	*10,000 (6* h*)*	*10,000 (6* h*)*

Hyperparameters used during training are shown in Table S1. Training curves for the meta-atom design network, dual-polarization meta-atom design network, and PNN are shown in Fig. S5. Despite the slightly different structures and training data fed to these networks, their generator losses all decrease gradually while discriminator losses remain constant. As shown in Fig. S5 (a-b), after approximately 3,000 epochs of training, each network converged to a point that both generator and discriminator loss stabilized, which means that the generator is able to generate samples that are close enough to the real samples that the discriminator is unable to differentiate between real and fake.

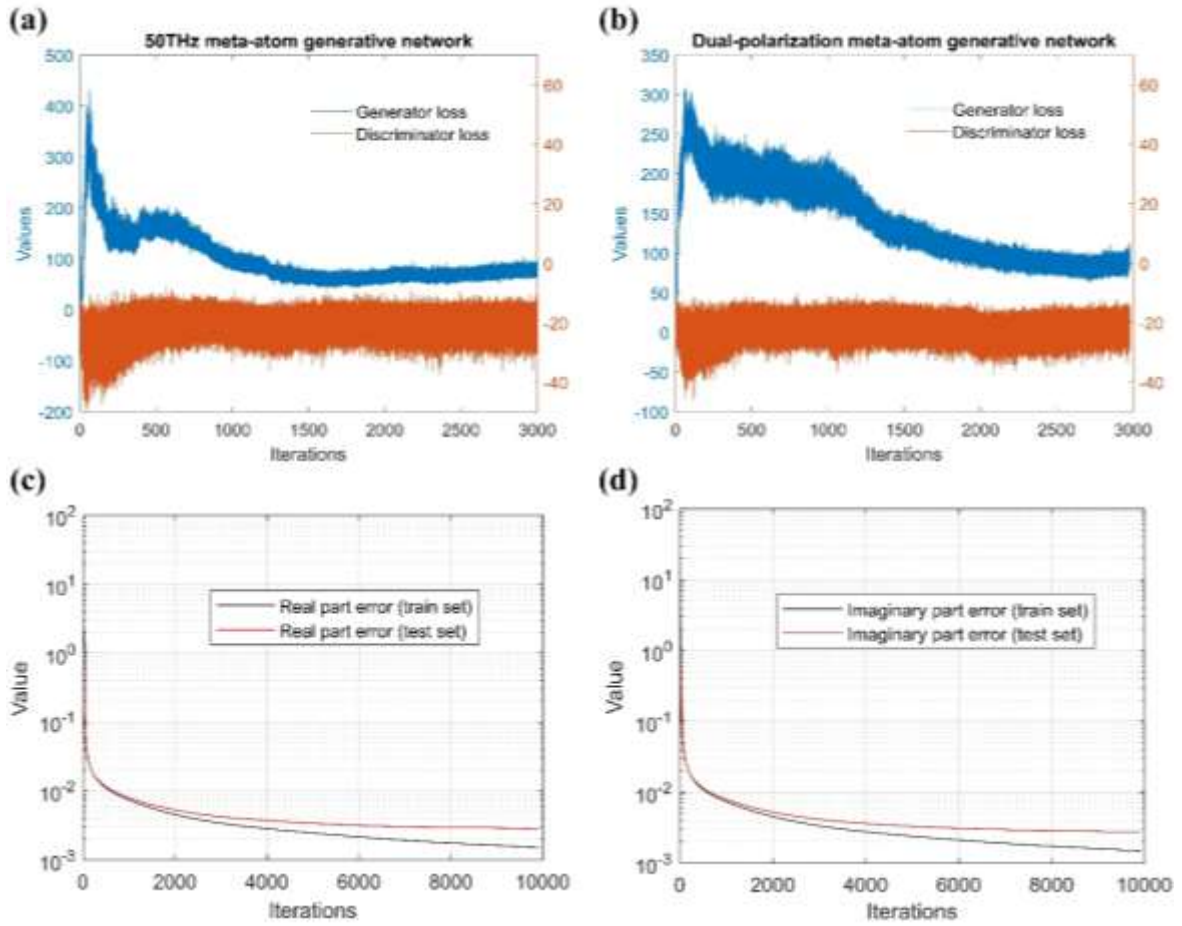


Figure S5. Training losses of the two GANs and the PNN proposed in the paper. (a) The meta-atom generative network operating at 50 THz. (b) The dual-polarization meta-atom generative network working at 55 THz. (c) PNN real part prediction. (d) PNN imaginary part prediction.

Training curves for the two constructed PNNs were included in Fig. S5(c-d). Both networks are trained with the same 69,000 groups of training data collected for the training of the GANs. Both networks converged well after 10,000 iterations.

To better visualize how the network learned the meta-atom design principles and actually "evolved" during the training process, we recorded the network models during the training process and employed several half-trained models to design the same bifocal metalens, presented in Fig. 3, and tested their performance by numerical simulations. Four different sets of bifocal metalens designs based on GAN models derived after 1, 2, 100 and 3,000 training iterations are presented. The designed metalenses, along with their full-wave simulated E-fields, are plotted in Fig. S6. Interestingly, at the beginning of the training process, the generated meta-atoms designs have similar shapes with large volumes and unclear boundaries (Fig. S6a). As the training proceeds, the GAN models start to generate meta-atom

patterns with more diverse shapes and refined details (Fig. S6b-d). The corresponding E-field distributions also gradually converged to two sharp focal spots (Fig. S6d), as desired, which firmly confirm the increasing learning capability of the proposed GAN model during the training process.

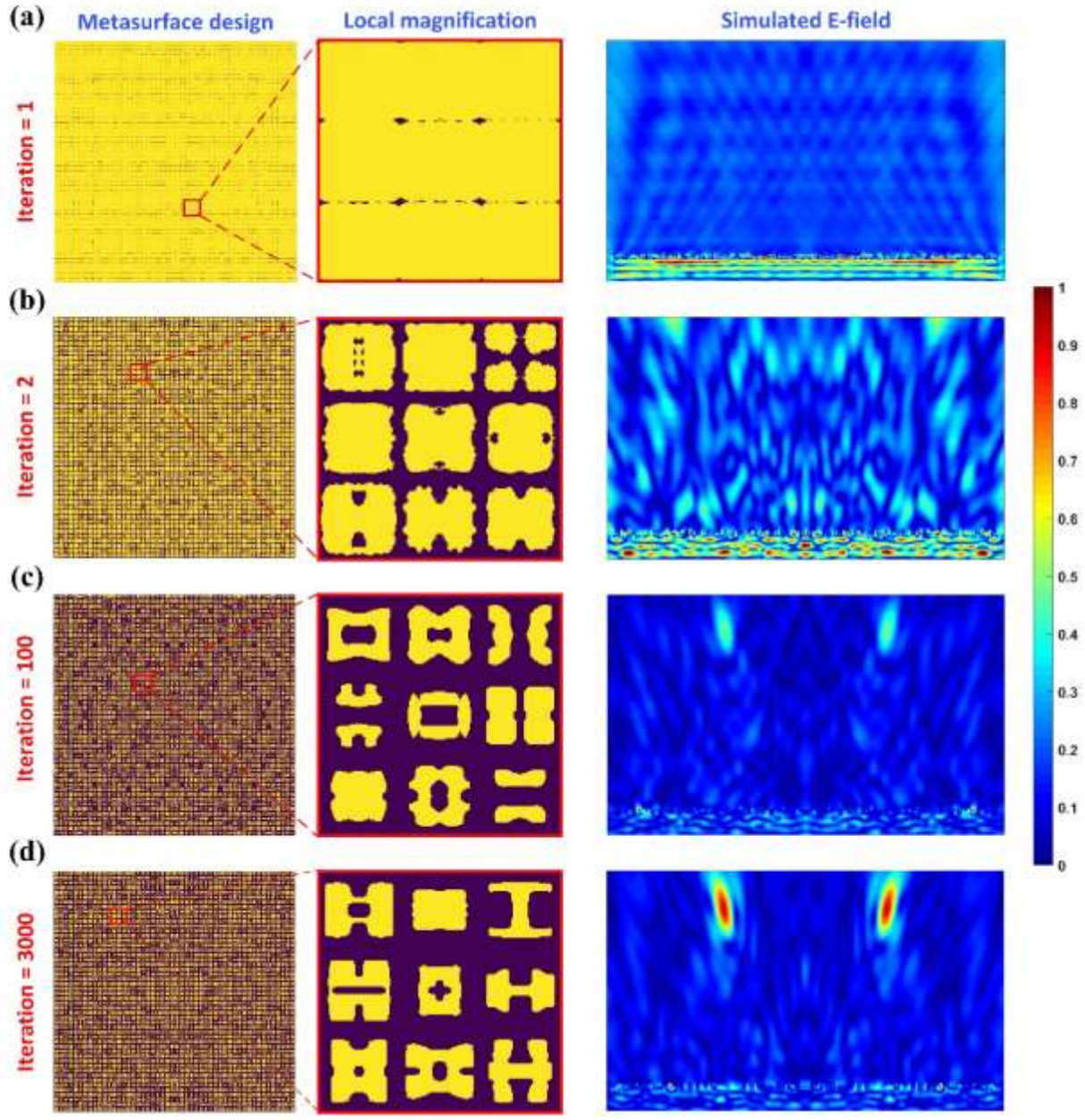


Figure S6. Visualization of the GAN training process. Bifocal metalenses with target amplitude and phase maps shown in Fig. 3a and 3b, designed using GAN models trained for (a) 1 iteration, (b) 2 iterations, (c) 100 iterations and (d) 3,000 iterations. Several meta-atoms from each metalens design are magnified for a clear view. Numerically simulated E-field of each metalens were performed using CST and plotted on the right side of each subplot.

5. Polarization-independent metalens design

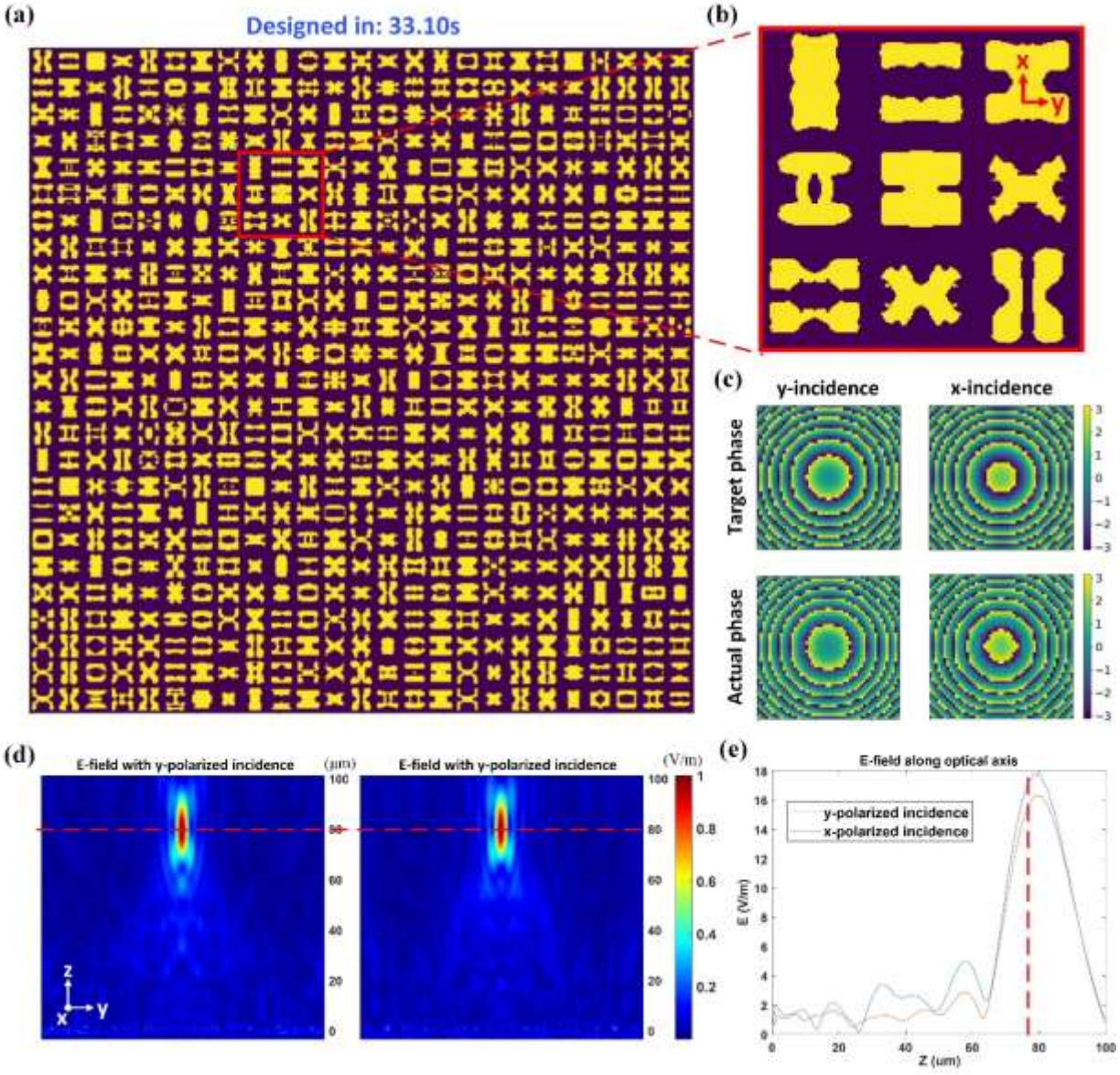


Figure S7. A polarization-independent focusing lens designed with the dual-polarization meta-atom generative network. (a) Metasurface pattern designed with the dual-polarization meta-atom generative network. Part of the metasurface device circled in red lines was enlarged in (b) for a clear view. (c) Element-wise phase and amplitude responses of each designed meta-atom under y and x polarized incidences. (d) Full wave simulated amplitude of E-field in y-z plane under two orthogonal polarization incidences. Focal length remained 80 μm while polarization direction was switched. (e) Full-wave simulated E-field along optical axis under two orthogonal polarization directions.

To further explore the versatility of the multifunctional meta-atom design network, we utilized our dual-polarization meta-atom generative network to design a polarization-insensitive transmissive focal lens with an equal focal length of 80 μm for both polarizations. One way to achieve this goal is enforcing that the phase shifts for *x* and *y* polarization are identical, which primarily results in

structures that feature 4-fold rotational symmetry, in accordance with their polarization-insensitive nature. The 4-fold rotational symmetry requirement can be relaxed by considering that the relative phase difference between two polarization states need not be zero, but simply maintained constant. Following this approach, with the help of dual-polarization meta-atom generative network, a 90 degree constant phase bias (difference) was added to the lens' phase mask under x-polarized plane wave incidence (versus that under y-polarized incidence). For each single cell in the metasurface lens, its target phase profiles under both polarization were designated as input of the network and one qualified design was generated to assemble the metalens (Fig. S7a). The full wave simulated electric fields in y-z plane for the whole lens in Fig. S7a are plotted in Fig. S7d. The E-fields along the optical axis in both cases share the same focal length of 80 μm with near-equal magnitude (Fig. S7e), validating the efficacy of proposed design approach.

6. Meta-filter design network

Wideband metasurface filters normally require distinct transmission performances at different frequency points, which are non-intuitive and thus their design largely relies on trial-and-error approaches. Meanwhile, design complexity of the proposed deep learning-enabled metasurface/ metamaterial inverse design network is not affected by the size of or correlation between the inputs, which makes it a perfect solution to these devices. The network structure proposed in this paper was also slightly modified to build a meta-filter generative network. The training curve of this GAN is shown in Fig. S8.

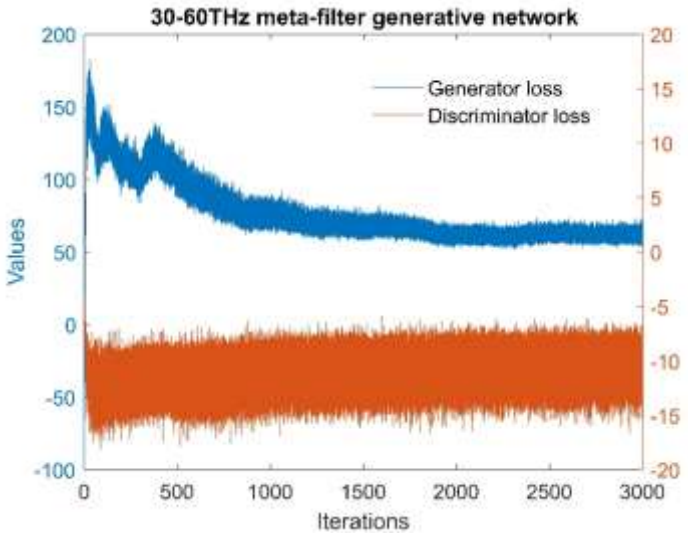


Figure S8. Training losses of the meta-filter generative networks.

The condition vector x, is denoted as: $x = [T_{\omega 1}(y), T_{\omega 2}(y), \cdots, T_{\omega N}(y)]$, where N represents the number of the frequency points used to sample the whole spectrum. We consider here a relatively wideband spectrum covering from 30 THz to 60 THz, which corresponds to a 5-10 μm spectrum in wavelength. A total number of 31 frequency points, with a spacing of 1THz, were chosen to sample the spectrum. Loss function for this network is the Euclidean distance between target spectrum responses (S) and responses of the designed meta-atoms ($T(y)$), defined as:

$$L_f = \frac{1}{N}\sum_{1}^{N}(T_{\omega i}(y) - S_{\omega i})^2 \tag{S2}$$

After the training fully converged, several sample patterns were created with different target transmission conditions (red curves) using the trained generative model (Fig. S9), along with the corresponding electromagnetic responses (blue curves) obtained through full wave simulations.

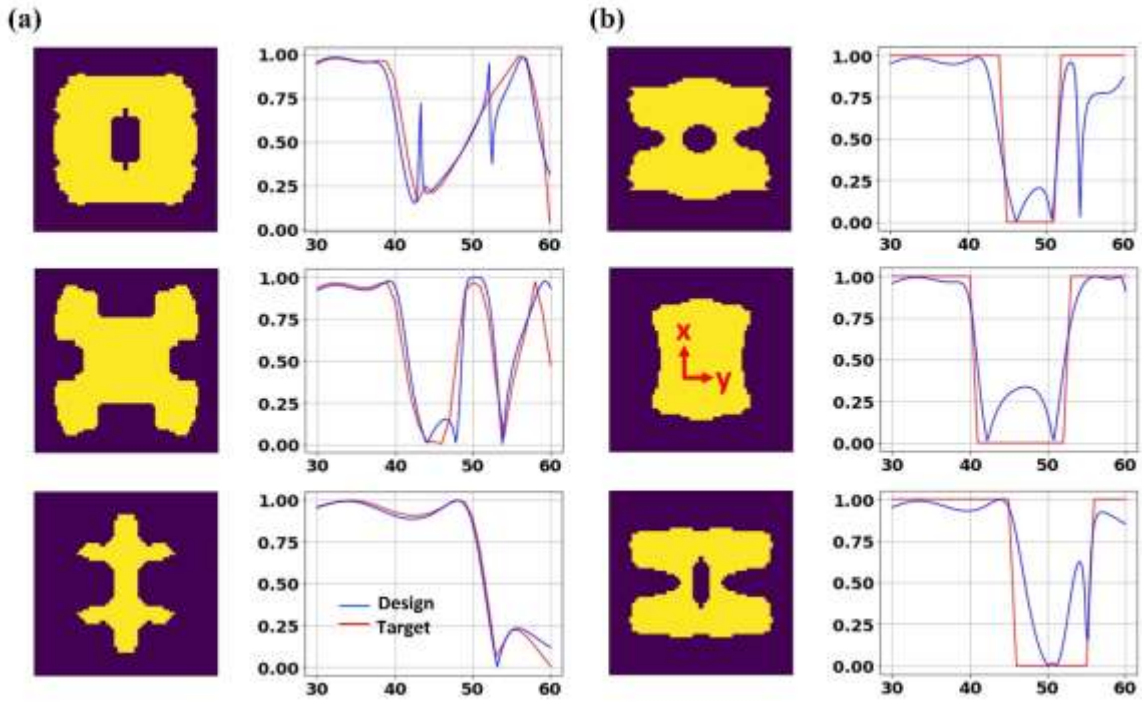


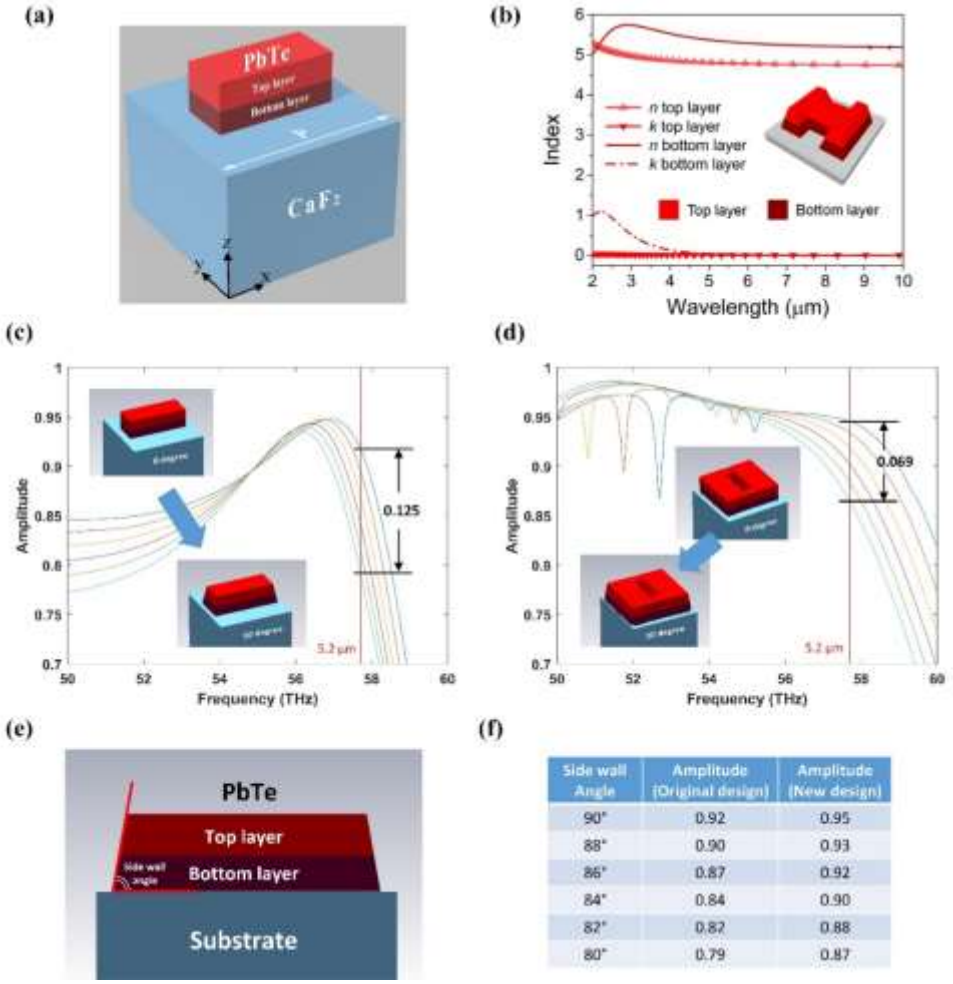
Figure S9. A wideband transmissive metasurface filters designed with the meta-filter generative network. (a) Design examples with targeted spectral responses from training dataset. (b) Design examples with user-defined targets. Red curves represent targeted spectrum while blue curves represent FEM simulated spectral responses of the generated designs. An x-polarized plane wave was illuminated from the substrate side for each meta-atom.

From Fig. S9, we can conclude that the transmission spectra of network-generated meta-atoms agree well with the targeted spectral responses in the cases when target responses are from the training datasets (Fig. S9a) and user-defined (Fig. S9b). This indicates the extraordinary "extrapolating" capability of the proposed network, which isn't necessarily limited by the training data distribution P_{data}. With the results as shown above, we feel the proposed network is able to handle inverse design tasks with larger size conditional vectors.

7. Meta-atom robustness analysis

Fabricated nano-structures often exhibit a non-vertical sidewall angle, which leads to resonance peak/dip shifts and potential efficiency drops. Normally, distinct meta-atoms designs perform differently under perturbations, which makes some of the designs more robust under perturbed side wall angles. Due to the generative nature of the proposed network, the well-trained model can create various designs that allows us to select the ones which are insensitive to this fabrication defects.

To make this case more convincing, we consider the metasurface fabrication platform with the material combination of PbTe and CaF2 (experimentally demonstrated in ref. [6], where a non-vertical sidewall angle was reported) to perform our verification. With a high refractive index exceeding 5, PbTe is considered ideally suited for creating dielectric meta-atoms supporting multiple Mie resonances. The material pair of PbTe and CaF2 also exhibits low optical attenuation in the mid-IR spectrum, allowing high transmission efficiency meta-atom designs. In reference to the experimental setups in [6], all meta-atoms under consideration are composed of a PbTe block sitting on a CaF2 substrate (Fig. S10a). The period, P, is set to be 2.5 μm along both axes. Total thickness of the PbTe block is fixed at 650 nm. The material refractive index n and extinction coefficient k of the PbTe material reported are employed, in which the PbTe material was modeled using a phenomenological two-layer structure in order to characterize the slight spatial composition variation caused by noncongruent vaporization. The thicknesses of top layer and bottom layer are set to be 0.35 μm and 0.3 μm, respectively, while their corresponding n and k values are plotted in Fig.S10b. Two selected meta-atoms, shown as the insets in Fig. S10c and S10d, have almost identical amplitude and phase responses at the working frequency (57.7THz). However, after applying a non-vertical side wall angle (Fig. S10e) to both cases, we find out that these two meta-atoms perform differently under perturbations. As the side wall angle is decreased from 90-degree (vertical) to 80-degree, the amplitude of the original design (as shown in Fig. S10c) is dropped from 0.92 to 0.79, while that of the new meta-atom designed by the generative network stays at 0.87 or above (Fig. S10d), indicating that it's a more robust design under perturbed side wall angles.



(f)

Side wall Angle	Amplitude (Original design)	Amplitude (New design)
90°	0.92	0.95
88°	0.90	0.93
86°	0.87	0.92
84°	0.84	0.90
82°	0.82	0.88
80°	0.79	0.87

Figure S10. Robustness analysis on meta-atoms' side wall angles. (a) Schematic front-view of a PbTe-CaF2 meta-atom structure. (b) Corresponding refractive index n and extinction coefficient k of the PbTe material fitted to a two-layer model. (c, d) Wide band amplitude responses of (c) original designs and (d) new designs with 90, 88, 86, 84, 82 and 80 degrees side wall angles. Designs with vertical side walls and incline side walls are included as insets. (e) Side wall angle is defined as the angle between high-index meta-atom sidewall and the substrate interface. (f) Simulated amplitude responses of both meta-atoms fabricated with different sidewall angles.

Figure (b) Reproduced with permission. [6] 2018, Nature communications.

The wideband spectrum responses as shown in Fig. S10c and Fig. S10d further reveal the difference between these two meta-atoms. As shown in Fig. S10c, at the target wavelength (or frequency), the original bar-shaped meta-atom is operating at the edge of a resonance dipole (the region where the amplitude is sensitive subject to any perturbation). The varying sidewall angle shifts this resonance towards the lower frequencies, which causes significant amplitude drop. In contrast, as shown in Fig. S10d, the new design is not operating at the edge of the resonance dipole, making it less sensitive to the spectral shift induced by non-vertical sidewall angles. To conclude, benefiting from the superior design efficiency and accuracy, the proposed network can readily generate numerous meta-atom candidates with qualified performance at low time cost, where robust and preferable designs can be identified and selected afterwards, which is highly desirable for metasurface fabrication and application.