Machine-learning-based inverse design of wide-band metasurfaces with interdependent unit cells

PAPER • OPEN ACCESS

Machine-learning-based inverse design of wide-band metasurfaces with interdependent unit cells

IOP Publishing Mach. Learn.: Sci. Technol. 7 (2026) 015024 https://doi.org/10.1088/2632-2153/ae365e

MACHINE LEARNING
Science and Technology

OPEN ACCESS

RECEIVED
24 June 2025

REVISED
13 November 2025

ACCEPTED FOR PUBLICATION
9 January 2026

PUBLISHED
2 February 2026

PAPER

Machine-learning-based inverse design of wide-band metasurfaces with interdependent unit cells

Timo Gahlmann[1], **Tobias Wenger**[2] **and Philippe Tassin**[1,*]

[1] Department of Physics, Chalmers University of Technology, SE-41296 Göteborg, Sweden
[2] Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, United States of America
[*] Author to whom any correspondence should be addressed.

E-mail: zernikephaseplate@outlook.com

Keywords: nanophotonics, metasurface, machine learning, Zernike phase mask

Supplementary material for this article is available online

Abstract

Metasurfaces are thin optical materials consisting of tiny meta-atoms that locally change the phase, amplitude, and polarization of light. Such metasurfaces have been shown to exhibit optical responses beyond what is possible with natural materials, e.g. in thin achromatic lenses, holograms, reflectors, and the generation of structured light. A considerable challenge lies in designing the sometimes thousands of meta-atoms that are needed to locally modify the wavefronts of light. Here we use a conditional generative adversarial network to demonstrate for the first time a way to design metasurfaces where the scattering responses of different meta-atoms are interdependent, i.e. some of the individual scattering parameters are unimportant, and it is the relative values between different meta-atoms that matter. We illustrate our approach with a design of a broadband phase mask for a Zernike wavefront sensor, with complex requirements on the phase relationship between the scattering parameters of an inner disk and the surrounding region. Our inverse design method offers increased design freedom, since it does not require specifying absolute scattering parameters, and allows for metasurfaces with large bandwidths.

1. Introduction

Optical metasurfaces are a new class of artificial materials that can manipulate light in unprecedented ways [1–5]. They consist of arrays of nanostructures—also called meta-atoms—that are much smaller than the wavelength of light and can locally modify the phase, amplitude, and polarization of incident light (see figure 1 for an illustration). Optical metasurfaces have many advantages over conventional optical elements, such as ultra-small thickness, high efficiency, and design flexibility. They have enabled the realization of various novel optical devices and functionalities, such as polarization detection [6] and manipulation [7, 8], image processing [9, 10], wavefront control [11, 12], beam steering [13] and focusing [14–16], and the generation and manipulation of beams with optical angular momentum [17]. Optical metasurfaces are expected to have a wide range of applications in optics, photonics, information technology, and biomedicine.

Metasurfaces often act as a phase or amplitude mask, where the phase or amplitude of each meta-atom is related. For example, a quadratic phase profile creates a lens focusing a light beam [18, 19]. The absolute phase of each meta-atom is unimportant; only their relative phase differences matter. In this paper, we will present a way to generate meta-atoms with interdependent properties without fixing the value of phase or amplitude in transmission (or reflection) of any individual meta-atom. This opens up substantial design freedom, which is impossible to achieve with human intuition. However, it is a challenge to determine the shape of the meta-atoms needed for a given desired optical response. Often, thousands of different meta-atoms must be designed and this is especially difficult for multiobjective metasurfaces, as there are no simple forward design techniques available [20–22]. To achieve such inverse design of metasurfaces, we use deep neural networks (DNNs), which are a powerful machine learning

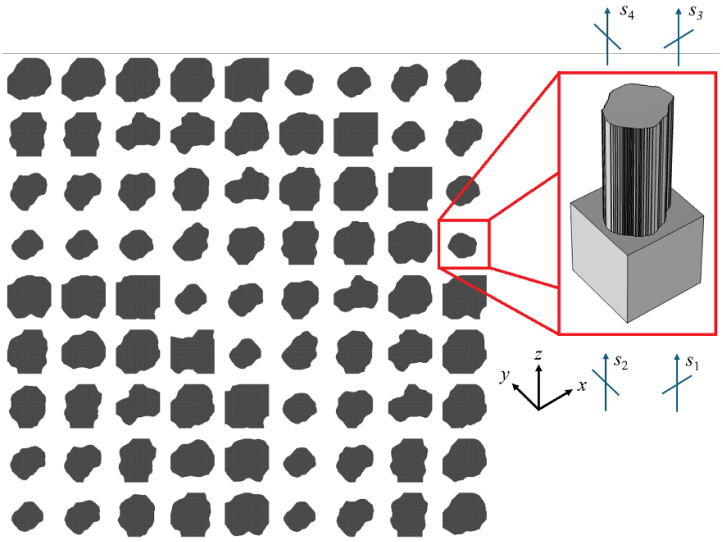

Figure 1. Illustration of a metasurface, consisting of a substrate with a large number of meta-atoms. The inset shows a 3D illustration of a meta-atom, where the cross section is a free-form shape to be designed. The waves determining the $\mathcal{S}$ parameters, and their corresponding polarizations, are indicated. The S parameters are defined by $\mathcal{S}_{ij} = s_j/s_i$ and are the optical properties we aim to design.

architecture [23–25] that can be used to learn complex patterns and functions from data, including to learn laws of physics [26, 27]. They have been widely used for various tasks such as image recognition [28, 29], natural language processing [30], and recommendation systems [31]. More recently, neural-network models have been used to achieve inverse design. This is often implemented by constructing a surrogate model that predicts physical properties for a device with arbitrary geometry [32–34], potentially aided by dimensionality reduction [35], manifold learning [36], or physics-informed networks [37–39]. Such models can then be inverted using a tandem network [40–42], an autoencoder [43, 44], or a conditional generative adversarial network (CGAN) [45–49]. A key advantage of DNNs in the design of physical devices is their ability to quickly generate designs for devices with any desired physical response. For example, in an optical metasurface we can design meta-atoms with any desired transmission amplitude and phase in the subset of physically realizable scattering parameters once trained, eliminating the need to run an optimization for each meta-atom. The training of neural networks requires a large number of simulations of the forward problem, but these simulations are independent, meaning they can be run simultaneously on high-performance computer clusters.

In this paper, we study a novel conditional generative adversarial network (CGAN) graph configuration for the inverse design of metasurface unit cells with interdependent properties, i.e. with design requirements including relative differences between scattering parameters of different meta-atoms. This implies that we must co-design all the meta-atoms, often thousands of meta-atoms, simultaneously. This provides several advantages over other inverse design techniques: (1) We get more design freedom, since we do not need to arbitrarily fix an absolute value of the $\mathcal{S}$ parameters at a specific location, which is important, e.g. for lenses where only relative phase differences are really required. (2) It allows us to design metasurfaces with large bandwidths, which we will demonstrate for a Zernike phase mask. (3) We can find designs that were untractable until now.

We will first present a graph for a CGAN scheme that can co-design all meta-atoms. Subsequently, we will discuss the training of the CGAN model. After we have presented our CGAN model, we will illustrate our CGAN on a problem that was untractable until now: a Zernike phase mask. This problem requires relative phase differences between meta-atoms in two different regions and also has a complex set of design requirements that cannot be solved by a simple parameter map or by human intuition.

2. Results

Our graph is based on a CGAN for the inverse design of individual meta-atoms [47], shown in figure 2, which is then adapted to enable the codesign of many meta-atoms. The machine learning method works as follows. Initially, a forward model $\mathcal{F}$ is established to approximate the optical properties of meta-atoms as a function of the grid representing the lithographic mask. This forward model aids in accelerating the assessment of optical properties compared to direct full-wave solutions. A binary grid is utilized

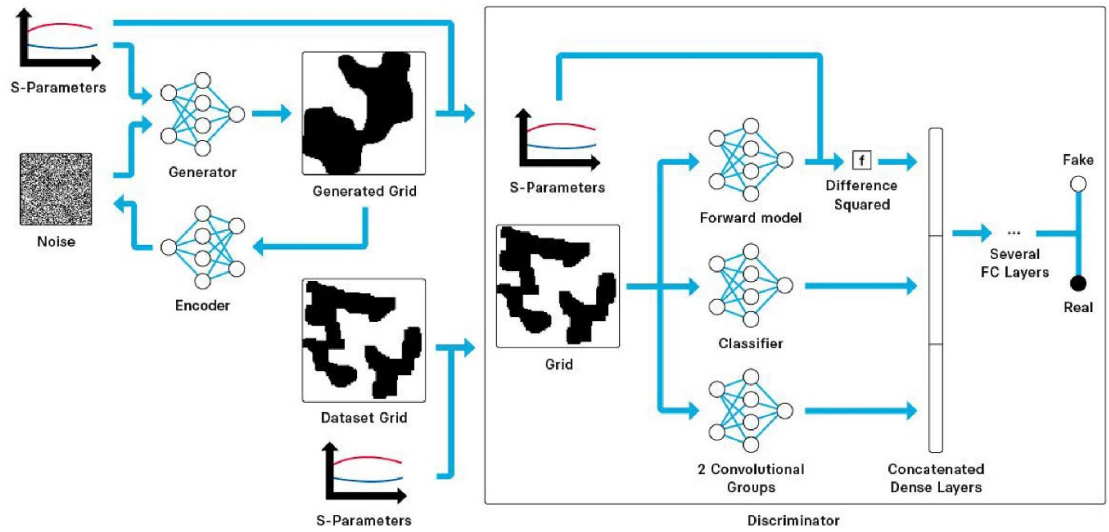

Figure 2. The conditional generative adversarial network model. The generator creates meta-atom grids from noise, $\mathcal{S}$ parameters and aspect ratio, while the encoder prevents mode collapse. Two expert networks (forward and classifier model) are integrated inside the discriminator. Reprinted figure with permission from [47], Copyright 2022 by the American Physical Society.

to represent the lithographic mask U of a meta-atom, indicating where the top layer is etched away and where not. Due to similarities with pixelated images, a DenseNet [50] is used to model the scattering parameters. This model is trained with about 80 000 semi-random grids created and loaded into a finite-element simulator to obtain the scattering parameters. We also use a classifier network $\mathcal{C}$ to estimate the fabrication feasibility as described in reference [47]. The two expert models ($\mathcal{F}$ and $\mathcal{C}$) are embedded into the discriminator of a CGAN by feeding the unit cell U to the two networks and their output to a hidden layer of the discriminator after computing the metrics of the model, such as evaluating the difference and difference squared of the desired/ground truth $\mathcal{S}$ parameters and the by the forward model approximated $\mathcal{S}$ parameters $\mathcal{F}(U)$.

The graph setup of this CGAN model is given by

$$\begin{aligned} \mathcal{G}[\mathcal{N},\mathcal{S},\mathcal{A}] &= U_{\mathcal{G}}, \\ \mathcal{D}[U,\mathcal{S},\mathcal{A}] &= \mathcal{D}[U, f(\mathcal{F}(U),\mathcal{S}), t(\mathcal{C}(U),\mathcal{A})] = O, \end{aligned} \tag{1}$$

where $\mathcal{G}$ is the generator, $\mathcal{D}$ is the discriminator, $\mathcal{S}$ are the scattering parameters (optical response), $\mathcal{A}$ is the aspect ratio of the meta-atom, $\mathcal{N}$ is noise, U is the lithographic mask of the meta-atom, and O is the discriminator output. The loss functions of the discriminator and generator are given by the usual cross-entropy loss:

$$\begin{aligned} \mathcal{E}_{\mathcal{D}}(U_D,\mathcal{S},\mathcal{A},\mathcal{N}) &= -\ln(\mathcal{D}(U_D,\mathcal{S},\mathcal{A})) - \ln(1-\mathcal{D}(\mathcal{G}(\mathcal{N},\mathcal{S},\mathcal{A}),\mathcal{S},\mathcal{A})), \\ \mathcal{E}_{\mathcal{G}}(\mathcal{S},\mathcal{A},\mathcal{N}) &= \ln(1-\mathcal{D}(\mathcal{G}(\mathcal{N},\mathcal{S},\mathcal{A}),\mathcal{S},\mathcal{A})), \end{aligned} \tag{2}$$

where U_D is a unit cell from the data set.

2.1. Interdependent $\mathcal{S}$ parameters

We now discuss how the above neural-network model for inverse design can be adapted to not only create one meta-atom U_0 with desired scattering parameters $\mathcal{S}_0$, but also $N \in \mathbb{N}$ additional meta-atoms $U_{1\ldots N}$ with interdependent $\mathcal{S}$ parameters $\mathcal{S}_{1\ldots N}$, where $\mathcal{S}_n = \phi_n(\mathcal{S}_0)$ with $n \in [1,N]$ (e.g. $\phi_n(\mathcal{S}_0)$ could be $\phi_1^n(\mathcal{S}_0)$, a phase rotation with constant amplitude, for beam steering). Generally, $\mathcal{S}_0$ has to be in the manifold $\mathbb{S}$ of $\mathcal{S}$ parameters spanned by the training data (used to train the discriminator and the forward model) to facilitate that the output of the generator is not a hallucination and the generated unit cell U_0 actually approximates the desired $\mathcal{S}$ parameters ($\mathcal{S}_0$) sufficiently well. For the same reason, the function ϕ_n must be a transformation that projects the set of desired $\mathcal{S}$ parameters $\mathcal{S}_0$ back onto $\mathbb{S}$ ($\phi_n(\mathcal{S}_0) \in \mathbb{S}$). This is not a strict condition in the sense that the generator will still provide output if $\phi_n(\mathcal{S}_0) \notin \mathbb{S}$, but, in the same way that the output U_0 is most likely a hallucination when $\mathcal{S}_0 \notin \mathbb{S}$, the generated unit cells U_n will not approximate the projected $\mathcal{S}$ parameters $\mathcal{S}_n$ if ϕ_n is not an endomorphism on $\mathbb{S}$. Nevertheless, this can be guaranteed by constructing a suitable training data set.

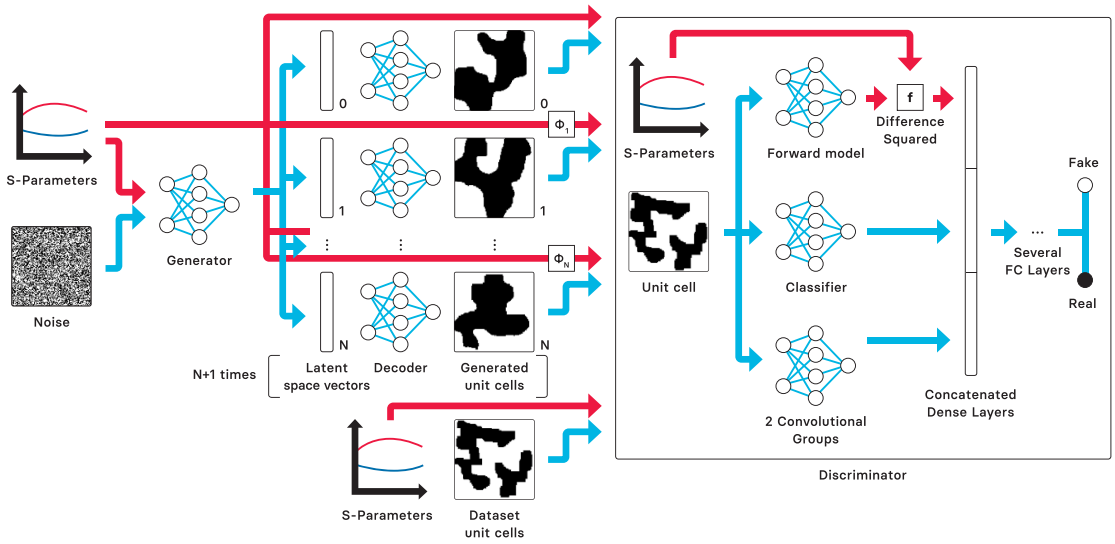

Figure 3. Neural network graph of the conditional generative network (CGAN) for the creation of unit cells with interdependent scattering parameters. The red arrows highlight the $\mathcal{S}$ parameter related connections of the graph and the blue arrows depict the rest of the graph.

In order to output multiple unit cells U_N with interdependent properties, the original generator architecture is split into two networks

$$\begin{aligned} \mathcal{G}_l\left[\mathcal{S}_0, \mathcal{A}, \mathcal{N}\right] &= \left[l_0, \ldots, l_N\right], \\ \mathcal{L}\left[l_n\right] &= U_n, \\ \mathcal{L}\left[\mathcal{G}_l\left[\mathcal{S}_0, \mathcal{A}, \mathcal{N}\right]\right] &= \left[U_0, \ldots, U_N\right], \end{aligned} \tag{3}$$

where $\mathcal{G}_l$ is a multiple latent vector generator, which outputs $N+1$ latent vectors and $\mathcal{L}$ is a latent vector decoder used to decode each latent vector l_n into a unit cell U_n. All these unit cells get individually fed to the discriminator with ϕ_n-transformed $\mathcal{S}$ parameters (ϕ_0 is the identity operation):

$$\mathcal{D}\left[U_n, \phi_n\left(\mathcal{S}_0\right), \mathcal{A}\right] = O_n. \tag{4}$$

The neural network graph of this adapted CGAN set up is visualized in figure 3. Subsequently, a weighted sum of the cross entropy losses of the O_n outputs is used in the loss function:

$$\begin{aligned} \mathcal{E}_{\mathcal{D}}\left(U_D, \mathcal{S}_0, \mathcal{A}, \mathcal{N}\right) &= -\ln\left(\mathcal{D}\left(U_D, \mathcal{S}_0, \mathcal{A}\right)\right) - \frac{\sum_{n=0}^{N} a_n \ln\left(1 - O_n\right)}{\sum_{n=0}^{N} a_n}, \\ \mathcal{E}_{\mathcal{G}}\left(\mathcal{S}_0, \mathcal{A}, \mathcal{N}\right) &= \frac{\sum_{n=0}^{N} a_n \ln\left(1 - O_n\right)}{\sum_{n=0}^{N} a_n}, \end{aligned} \tag{5}$$

where the factors a_n are the weights of the generator output U_n (and the respective discriminator output O_n) in the loss function.

2.2. Model training

While $a_0 = 1$ for all epochs, a_1 to a_N start out at 0.001 at epoch 0. Thereby, the generator outputs U_1 to U_N play almost no role in the early stages of training, making this effectively a standard inverse design CGAN setup (comparable to the graph in equation (2)). This improves the generator output U_0 and, in the process, minimizes the $\mathcal{S}$ parameter error $|\mathcal{S}_0 - \mathcal{F}[U_0]|$. Once the $\mathcal{S}$ parameter error sinks under a certain threshold, a_1 to a_N increase as the $\mathcal{S}$ parameter error of U_0 decreases (and approach $0.5/N$ in the limit of the error going to 0). In this way, the generator first learns to output l_0. Additionally, the latent decoder is trained to output U_0 from l_0, while the discriminator learns the wide range of characteristics that have to be fulfilled for a meta-atom mask to be a suitable design. Once this works so well that most of a set of training instances would converge to a useful generator, the weight factors a_1 to a_N are increased in order to gradually shift the focus and learning of the generator to output l_1 to l_N (note that the same latent decoder network $\mathcal{L}$ is used to produce U_0 to U_N from l_0 to l_N). More details about the model training is provided in the Methods section.

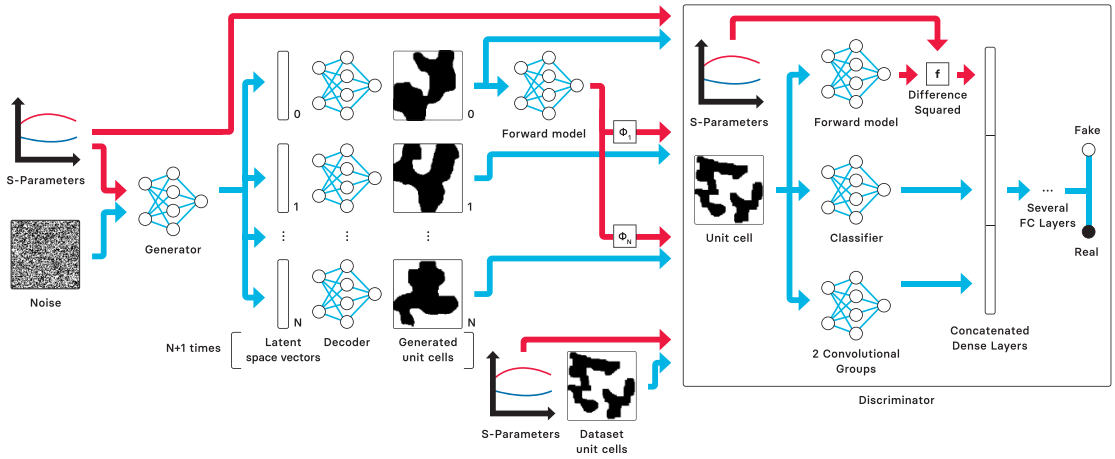

Figure 4. Graph of the second version of the conditional generative network (CGAN) for the creation of unit cells with interdependent scattering parameters using a more adaptive strategy. The red arrows highlight the $\mathcal{S}$ parameter related connections of the graph and the blue arrows depict the rest of the graph.

The unit cells U_1 to U_n will generally approximate the $\mathcal{S}$ parameters $\phi_1(S_0)$ to $\phi_N(S_0)$, because these are fed to the discriminator, which is punishing deviations. Thereby, the following sum is minimized

$$|\mathcal{S}_0 - \mathcal{F}[U_0]| + \sum_{n=1}^{N} |\phi_n(\mathcal{S}_0) - \mathcal{F}[U_n]|. \tag{6}$$

2.3. Model without reference phase/amplitude

In the neural network graph above, we still fix the scattering parameters $\mathcal{S}_0$ of one of the meta-atoms. This is often undesirable as it unnecessarily restricts the design freedom, e.g. for a lens, the transmission phase of the center meta-atom can be arbitrary as long as the relative phase compared to the center meta-atom has a quadratic phase as a function of the distance to the center meta-atom. If the forward model $\mathcal{F}$ has the necessary accuracy, then we can modify equation (4) by replacing $\phi_n(\mathcal{S})$ with $\phi_n(\mathcal{F}[U_0])$ for $n > 0$:

$$\begin{aligned} \mathcal{D}[U_0, \mathcal{S}_0, \mathcal{A}] &= O_0 \\ \mathcal{D}[U_n, \phi_n(\mathcal{F}[U_0]), \mathcal{A}] &= O_n \ \forall \ n > 0. \end{aligned} \tag{7}$$

The network graph of this setup is shown in figure 4. In this way, it is possible to give more design freedom to the generator to find a set of unit cells where

$$|\mathcal{S}_0 - \mathcal{F}[U_0]| + \sum_{n=1}^{N} |\phi_n(\mathcal{F}[U_0]) - \mathcal{F}[U_n]| \tag{8}$$

is minimized. In the training process, we will generally feed $\mathcal{S}_0$ parameters to the generator and discriminator, although, by utilizing an input mask ($\mathcal{S}_m$) for the $\mathcal{S}_0$ parameters, it is possible to feed only a subset of all parameters to the network or none at all. Feeding a wide variety of $\mathcal{S}_0\,(\in \mathbb{S})$ parameters to the networks helps generalize their capabilities, and feeding none ($\mathcal{S}_m = 0 \ \rightarrow |\mathcal{S}_m(\mathcal{S}_0 - \mathcal{F}[g_0])| = 0$) enables the maximal design freedom of the generator to minimize

$$\sum_{n=1}^{N} |\phi_n(\mathcal{F}[U_0]) - \mathcal{F}[U_n]|. \tag{9}$$

2.4. Zernike phase masks

We subsequently use our trained CGAN for the design of a broadband phase mask for a Zernike wavefront sensor (ZWFS) [51–54]. A ZWFS can be used to detect deviations of a detected wavefront from a plane wavefront, which can then be used to correct the incoming light signal. Therefore, it is a very powerful technique for ground-based telescopes, where the incoming light is distorted by the Earth's atmosphere, and eliminating this effect can significantly improve a telescope's ability to view distant

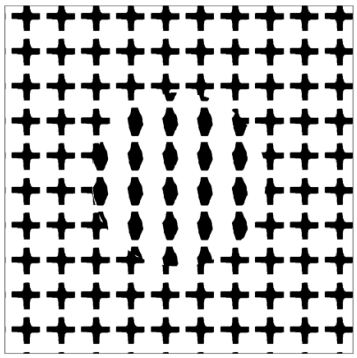

Figure 5. A sketch of the 2D cross section of the ZWFS metasurface (meta-atoms not to scale), unit cells from figure 7(a). A Zernike phase mask has a central circular region and a surrounding region; there are phase relationships between the S parameters of both regions.

objects. In addition, segmented aperture telescopes, e.g. at the Keck observatory or the recently launched James Webb Space telescope, require accurate phasing between the different segments in order to form the precise images required. Recent work at the Keck II telescope demonstrated the utility of metasurface masks by improving the observed Strehl ratios of imaged distant stars [55].

A ZWFS implements the phase contrast technique, where light is passed through a focal plane mask consisting of two regions, imparting different phases on impinging light. The light then interferes at an exit pupil to create intensity variations that can be readily imaged with a camera. A phase mask can, for instance, be created by etching a shallow dimple into a piece of glass to create a difference in optical path length between the etched and non-etched regions. However, such an implementation leads to a limited dynamic range in terms of measurable phase errors. The dynamic range can be improved by using polarization as a degree of freedom and simultaneously impart two different phases on two orthogonal polarization states, which can later be separated to create two images containing different phase information. Here, we utilize our CGAN to design metasurfaces that not only impart different phases on two polarization states, but also do so in a highly achromatic manner. This achromatic behavior is important for broadband functionality of metasurface devices in general, and in the ZWFS context this enables the use of more of the incident light signal as compared to a highly chromatic device.

We design a Zernike phase mask that consists of two regions, as illustrated in figure 5, with the following interdependent $\mathcal{S}$ parameters over the entire frequency band of interest:

- a relative phase difference of $\pi/2$ between the $\mathcal{S}_{31}$ parameters in both regions;
- a relative phase difference of $-\pi/2$ between the $\mathcal{S}_{42}$ parameters in both regions;
- high and approximately constant amplitudes of $\mathcal{S}_{31}$ and $\mathcal{S}_{42}$ in both regions;

where ports 1 and 2 are two orthogonal polarization states at the input plane and ports 3 and 4 are two orthogonal modes at the output plane. Therefore, the transformation ϕ, mapping $\mathcal{S}$ parameters from the central region to the surrounding region, is given by

$$\phi_1(\mathcal{S}_{31}) = \mathcal{S}_{31}\ e^{\frac{\pi}{2}}, \phi_1(\mathcal{S}_{42}) = \mathcal{S}_{42}\ e^{-\frac{\pi}{2}} \tag{10}$$

in our example of a Zernike phase mask.

Utilizing two polarization states with opposite phase difference between the regions, it is possible to improve the dynamic range of Zernike wavefront sensing [56]. Designing broadband meta-atoms for such a Zernike phase mask is challenging, because there are requirements on phase and amplitude simultaneously, for two different polarization states. It can be achieved in the circular polarization basis with Berry–Pancharatnam meta-atoms [57–59], but such an implementation suffers from leakage which ultimately affects the Zernike sensing. Furthermore, subsequent polarization splitting in a circular basis requires additional optical components, which complicates the system. As a reference, we plot in figure 6 the efficiency of a Zernike phase plate with Berry–Pancharatnam meta-atoms; we see that the efficiency is high at the center frequency, but then gradually drops towards the edges of the operational bandwidth, which is a fundamental limitation of this type of metasurfaces [60]. We now ask our CGAN for interdependent meta-atoms to design the metasurface for a Zernike phase plate and, in figures 7(a) and (b), we plot the meta-atom pairs for two different designs returned by the CGAN. The resulting etching masks, shown in figures 7(a) and (b), show that our CGAN can find designs in a much larger design

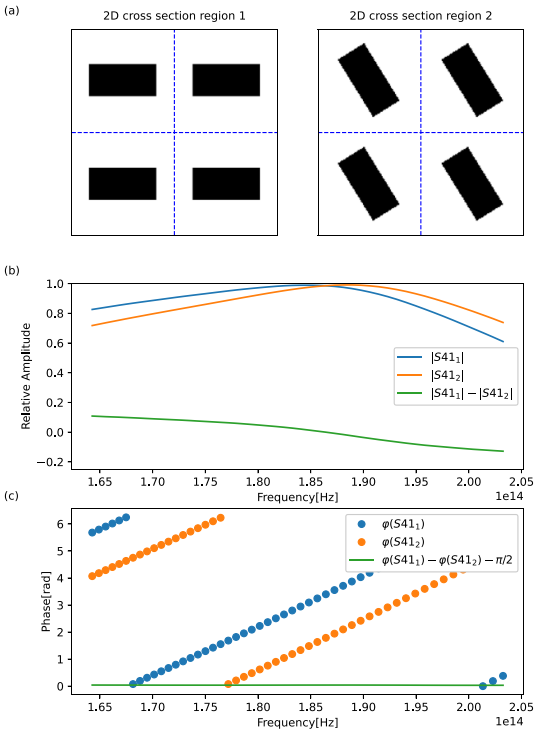

Figure 6. Berry–Pancharatnam meta-atoms in the two regions of a Zernike phase mask, converting a circular polarization (port 1) to the orthogonal polarization state in transmission (port 4). (a) 2 × 2 repetitions of the unit cells are shown for both regions (center disk and region outside the center disk). (b) The conversion amplitudes in the two difference regions, as well as their difference. (c) The phases for the scattering amplitudes in both regions, as well as the deviations from design target. The subscript in the $\mathcal{S}$ parameters indicates the region of the Zernike phase mask.

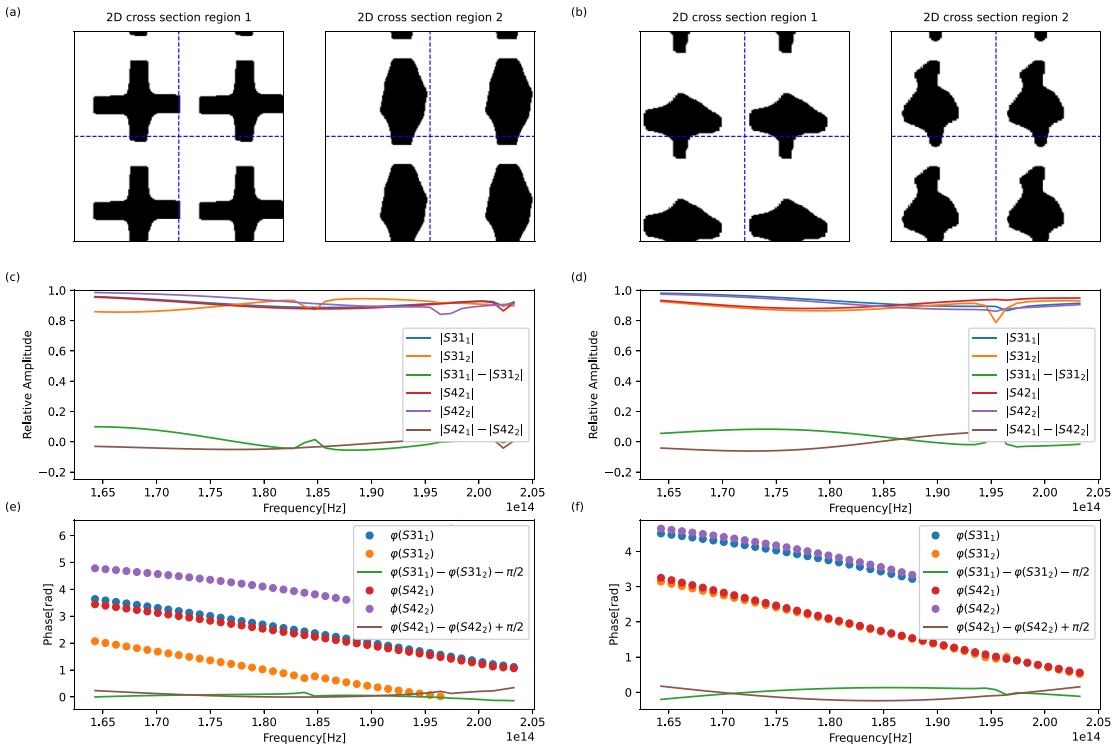

Figure 7. Two example designs of meta-atoms for the two phase-altering regions of a Zernike phase mask, as created by our CGAN variation with interdependent $\mathcal{S}$ parameter relations across the complete frequency band. The metasurfaces work in the linear polarization basis. (a), (b) 2 × 2 repetitions of the unit cells are shown for both regions. (c), (d) The transmission amplitudes for the two orthogonal polarization states, as well as the differences between the polarization states. (e), (f) The transmission phases for the two orthogonal polarization in both regions, as well as the deviations from design target. The subscript in the $\mathcal{S}$ parameters indicates the region of the Zernike phase mask. All curves shown are exact results from full-wave simulations (not part of the ground truth data.).

space, with designs that are difficult to find with traditional approaches. In figures 7(c) and (d), we plot the transmission efficiency of the meta-atoms in both regions. The CGAN-derived designs have efficiencies that are constant over a very large bandwidth; they are slightly fluctuating, but do not systematically drop off towards the edges of the operational bandwidth. In figures 7(e) and (f), we observe that the transmission phases satisfy the requirement to a very high degree, also constant over the very large bandwidth. We point out that the desired phase is here imprinted for both polarization states in linear polarization basis ($\mathcal{S}_{31}$ and $\mathcal{S}_{42}$), eliminating the need for additional optical components needed for the classical designs using Berry–Pancharatnam metasurfaces. We plot also the other $\mathcal{S}$ parameters in the supplementary material, showing amongst others that the cross polarization leakage is negligible. This confirms that CGAN-based inverse design can find designs beyond the current state of the art for meta-atoms with interdependent optical properties.

3. Discussion

In this article, we have demonstrated a CGAN-based inverse design method for metasurfaces consisting of meta-atoms with interdependent optical properties. This is a very common situation in optics, since an overall phase shift of the wavefronts is often unimportant, but leads to intractable designs without the use of machine learning. We demonstrated our CGAN with a metasurface for the phase mask of a ZWFS, which contains two regions of meta-atoms with multiple, relative requirements (phase and amplitude) on multiple $\mathcal{S}$ parameters of the meta-atom pairs. The CGAN is able to design etching masks for such phase plates with much lower leakage than existing solutions. In addition, our CGAN has built-in functionality to enforce that its resultant designs are fabricationally feasible. The training time of our CGAN is about 12 to 56 h on 2 GPUs. The inference time is a few milliseconds. In figure 8, we estimate the learning ability by plotting the mean square error (MSE) of every pair of meta-atoms from our training set (blue bars), which is a measure of how a Zernike phase plate with random meta-atoms chosen from the training data set perform, and we compare it with the MSE for designs created with different generators (orange bars). In this plot, the training data set pairs already include the samples that were generated by iterative improvement by using the CGAN to find training samples in sparsely populated regions of the design space. We see that the CGAN is able to find designs for meta-atom pairs with a significantly lower MSE than present in the training data set. Our Zernike phase masks demonstrate that machine learning can help with optical design of devices that cannot be designed with conventional methods.

Machine-learning-based designs always come at the cost of a large number of training data samples that need to be simulated, but this cost can be acceptable if no other design method is known. The cost is also reduced if the CGAN can be reused for multiple designs, e.g. in gradient-phase metasurfaces where thousands of different meta-atoms are needed. For example, another possible application of our CGAN for metasurfaces with interdependent unit cells could be a metasurface lens without chromatic aberrations. Since the phase profile of a lens is wavelength-dependent, the individual unit cells need to have interdependent $\mathcal{S}$ parameters—the required phase profile would be given by $\varphi(r,\lambda) = -\frac{2\pi}{\lambda}(\sqrt{r^2+f^2}-f)$, where f is the focal point distance of the lens and r is the distance of a unit cell from the center of the lens. Consequently, the shift in phase from unit cell to unit cell also depends on the wavelength. Hence, the $\mathcal{S}$ parameters (at each and every frequency) of the individual unit cells in a metasurface lens without chromatic aberration must obey interdependent relations. If the network is trained with a sufficiently diverse training set, it can be reused for other applications using the same materials platform; otherwise the networks will need to be retrained. Another possibility is to reuse the networks in a transfer learning setting, where the trained networks are reused with additional layers at the beginning and end of the the neural-network structure. We therefore anticipate that our CGAN-based inverse design method for meta-atoms with interdependent optical properties will lead to new designs for a multitude of optical devices. While it is a black box model that allows to find solutions beyond human intuition on physical insight, it can also be combined using recent techniques to incorporate the physics in the surrogate/forward model [61]. Furthermore, since our method is transparent to the physics (the physics is contained in the training data), our method can be straightforwardly transferred to systems in other paradigms of physics, such as acoustics, mechanics, and fluid dynamics.

4. Methods

4.1. Training of neural networks

The models were trained on NVIDIA Tesla V100 SXM2 GPUs with 32 GB RAM with the following optimizers and training durations:

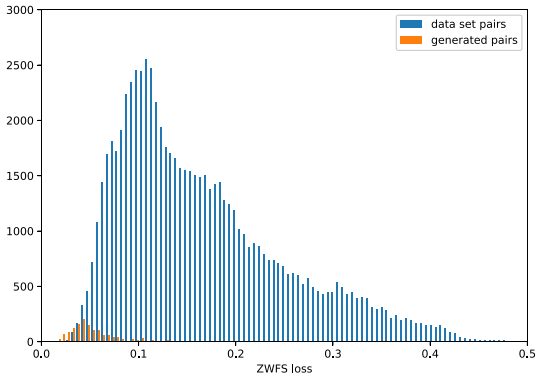

Figure 8. Minimum MSE loss for all data points in the data set compared with every other $\mathcal{S}$ parameter data point (blue bars; lowest MSE = 0.0216), including the data points generated with iterative improvement. MSE loss of the CGAN generated meta-atom pairs (orange bars; lowest MSE loss = 0.0152).

- forward network
  - optimizer: Adamax(learning rate = 0.001, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-7}$)
  - typical training time: 200 epochs in 4 h
- classifier
  - optimizer: Adamax(learning rate = 0.001, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-7}$)
  - typical training time: 150 epochs in 5 h
- generator optimizer: Adamax(learning rate = 0.001, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-7}$)
- discriminator optimizer: Adam(learning rate = 0.000 002, $\beta_1 = 0.5$, $\beta_2 = 0.999$, $\epsilon = 10^{-7}$)
- encoder optimizer: Adamax(learning rate = 0.001, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-7}$)
- typical CGAN training time: 80 to 400 epochs in 12 to 56 h (2 GPUs).

The following python packages were used:

- tensorflow
- numpy
- pandas
- multiprocessing.

Validation loss curves for the forward network, the classifier network, the generator, and the discriminator can be found in the supplementary material. The generator and the discriminator are competing in a zero-sum game and, therefore, their validation loss curves have information on the stability of the training process rather than the quality of the outputs.

4.2. Initial training data generation

Our training data set consists of a set of 100×100 binary grids, where 0s mark the areas in which the material is etched away and 1s mark the areas where the lithographic mask prohibits the etching process and consequently the material stays intact. In order to create a data set of semi-random pixelated images representing the lithographic masks of experimentally feasible structures, the binary grids were grown from random noise by filling too small holes or erasing too small islands of increasing sizes, since too small holes and islands cannot be realized by current lithographic techniques. In this way, random noise was transformed successively from iteration to iteration into structures with courser features until the fabrication constraints were met. As a result of this process, the allowed geometries of our meta-atoms are only limited by the resolution of the grid (in practice our grid pixels are so much smaller than the wavelength of light and slightly smaller than the estimated fabrication error that further reducing the pixel size is pointless) and by fabrication feasibility—and not by any other constraints arising from the computational method.

4.3. Training strategy and iterative improvement of training data

To improve the quality of the training data set, we have iteratively improved the training data using the following algorithm:

1. Expert models training: We train the classifier and forward networks on the current set of grids, labels, and $\mathcal{S}$ parameters.
2. CGAN training: We train several CGAN generators with the classifier and forward models trained in the previous step.
3. Classifier data enhancement: From all generators of this iteration, we produce grids for various different kinds of target $\mathcal{S}$ parameters and noise, round them to values of 0 and 1, save and label them for the classifier, and then add the data to the classifier training data.
4. Forward model data enhancement: The quality and accuracy of grids generated by the generator is directly dependent on the accuracy of the forward model in the regime of interest. Since the number of randomly generated grid pairs that satisfy the design criteria for a Zernike phase mask is rather low, we increase the size of the raining set iteratively. To this end, a diverse selection of grids in the regime of interest is created by the generators of this iteration, corrected for fabricationability if necessary, simulated by COMSOL, and then added to the training data of the forward model. In this way, we gradually create more and more training data pairs that are relevant for the design of a Zernike phase mask.

By applying this iterative process, we are able to increase the accuracy of all models, reduce the number of corrections to the grids generated by the generator necessary to satisfy feasibility constraints, and obtain $\mathcal{S}$ parameters of the generated grids that better approximate the target parameters.

4.4. Electromagnetic simulations

The electromagnetic simulations of the meta-atoms (training sample set) are performed using the finite-element method with the package COMSOL. Scattering parameters are calculated using absorbing boundary conditions (ports) for a plane wave incident from the top of the metasurface. The errors on the $\mathcal{S}$ parameters in the finite-element simulations are approximately 1%.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Supplementary material available at https://doi.org/10.1088/2632-2153/ae365e/data1.

Acknowledgments

We acknowledge support from Chalmers' Area of Advance 'Nano' and from the Swedish Research Council under Grant No. 2020-05284. The training data generation and ANN training were performed on resources provided by the Swedish National Infrastructure for Computing (NAISS), at the Chalmers/C3SE and KTH/PDC sites, partially funded by the Swedish Research Council under Grant No. 2022-06725. The work was partially performed in the framework of the Excellence Center META-PIX. Part of the research presented in this article was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004). We thank Jonas Rose for help with the graphical design of figures 3 and 4.

ORCID iD

Timo Gahlmann 0000-0001-5618-8987

References

[1] Yu N, Genevet P, Kats M A, Aieta F, Tetienne J-P, Capasso F and Gaburro Z 2011 Light propagation with phase discontinuities: generalized laws of reflection and refraction *Science* **334** 333–7
[2] Aieta F , Genevet P, Kats M A, Yu N, Blanchard R, Gaburro Z and Capasso F 2012 Aberration-free ultrathin flat lenses and axicons at telecom wavelengths based on plasmonic metasurfaces *Nano Lett.* **12** 4932–6
[3] Lin D, Fan P, Hasman E and Brongersma M L 2014 Dielectric gradient metasurface optical elements *Science* **345** 298–302
[4] Ding F, Pors A and Bozhevolnyi S I 2017 Gradient metasurfaces: a review of fundamentals and applications *Rep. Prog. Phys.* **81** 026401
[5] Hsiao H-H, Chu C H and Tsai D P 2017 Fundamentals and applications of metasurfaces *Small Methods* **1** 1600064
[6] Shen B, Wang P, Polson R and Menon R 2015 An integrated-nanophotonics polarization beamsplitter with 2.4× 2.4 μm^2 footprint *Nat. Photon.* **9** 378–82
[7] Black L-J, Wang Y, De Groot C, Arbouet A and Muskens O L 2014 Optimal polarization conversion in coupled dimer plasmonic nanoantennas for metasurfaces *ACS Nano* **8** 6390–9

[8] Desiatov B, Mazurski N, Fainman Y and Levy U 2015 Polarization selective beam shaping using nanoscale dielectric metasurfaces *Opt. Express* **23** 22611–8
[9] Zhang X , Zhou Y, Zheng H, Linares A E, Ugwu F C, Li D, Sun H-B, Bai B and Valentine J G 2021 Reconfigurable metasurface for image processing *Nano Lett.* **21** 8715–22
[10] Kwon H, Cordaro A, Sounas D, Polman A and Alù A 2020 Dual-polarization analog 2D image processing with nonlocal metasurfaces *ACS Photonics* **7** 1799–805
[11] Suzuki T and Kondoh S 2018 Negative refractive index metasurface in the 2.0-THz band *Opt. Mater. Express* **8** 1916–25
[12] Kamali S M, Arbabi E, Arbabi A and Faraon A 2018 A review of dielectric optical metasurfaces for wavefront control *Nanophotonics* **7** 1041–68
[13] Wei Z , Cao Y, Su X, Gong Z, Long Y and Li H 2013 Highly efficient beam steering with a transparent metasurface *Opt. Express* **21** 10739–45
[14] Paniagua-Domínguez R *et al* 2018 A metalens with a near-unity numerical aperture *Nano Lett.* **18** 2124–32
[15] Andrén D, Martínez-Llinàs J, Tassin P, Käll M and Verre R 2020 Large-scale metasurfaces made by an exposed resist *ACS Photonics* **7** 885–92
[16] Wenger T, Muller R, Wilson D, Gunapala S D and Soibel A 2021 Large metasurface-based optical concentrators for infrared photodetectors *AIP Adv.* **11** 085221
[17] Karimi E, Schulz S A, De Leon I, Qassim H, Upham J and Boyd R W 2014 Generating optical orbital angular momentum at visible wavelengths using a plasmonic metasurface *Light: Sci. Appl.* **3** e167
[18] Khorasaninejad M and Capasso F 2017 Metalenses: versatile multifunctional photonic components *Science* **358** eaam8100
[19] Martínez-Llinàs J *et al* 2019 A Gaussian reflective metasurface for advanced wavefront manipulation *Opt. Express* **27** 21069–82
[20] Campbell S D *et al* 2018 Advanced multi-objective and surrogate-assisted optimization of topologically diverse metasurface architectures *Proc. SPIE* **10719** 43–48
[21] Jafar-Zanjani S, Inampudi S and Mosallaei H 2018 Adaptive genetic algorithm for optical metasurfaces design *Sci. Rep.* **8** 11040
[22] Allen K W, Dykes D J, Reid D R and Lee R T 2020 Multi-objective genetic algorithm optimization of frequency selective metasurfaces to engineer Ku-passband filter responses *Prog. Electromagn. Res.* **167** 19–30
[23] Mehlig B 2021 *Machine Learning With Neural Networks: An Introduction for Scientists and Engineers* (Cambridge University Press)
[24] Da Silva I N *et al* 2017 *Artificial Neural Network Architectures and Training Processes* (Springer) pp 21–28
[25] Gupta A, Anpalagan A, Guan L and Khwaja A S 2021 Deep learning for object detection and scene perception in self-driving cars: survey, challenges and open issues *Array* **10** 100057
[26] Van Nieuwenburg E P, Liu Y-H and Huber S D 2017 Learning phase transitions by confusion *Nat. Phys.* **13** 435–9
[27] Iten R, Metger T, Wilming H, Del Rio L and Renner R 2020 Discovering physical concepts with neural networks *Phys. Rev. Lett.* **124** 010508
[28] Khan S *et al* 2018 *A Guide to Convolutional Neural Networks for Computer Vision* vol 8 (Springer)
[29] Wang P, Fan E and Wang P 2021 Comparative analysis of image classification algorithms based on traditional machine learning and deep learning *Pattern Recognit. Lett.* **141** 61–67
[30] Liu Y *et al* 2023 Summary of ChatGPT-related research and perspective towards the future of large language models *Meta-Radiol.* **1** 100017
[31] Portugal I, Alencar P and Cowan D 2018 The use of machine learning algorithms in recommender systems: a systematic review *Expert Syst. Appl.* **97** 205–27
[32] So S, Mun J and Rho J 2019 Simultaneous inverse design of materials and structures via deep learning: demonstration of dipole resonance engineering using core–shell nanoparticles *ACS Appl. Mater. Interfaces* **11** 24264–8
[33] Pestourie R, Mroueh Y, Nguyen T V, Das P and Johnson S G 2020 Active learning of deep surrogates for PDEs: application to metasurface design *npj Comput. Mater.* **6** 1–7
[34] González-Alcalde A K, Salas-Montiel R, Kalt V, Blaize S and Macías D 2020 Engineering colors in all-dielectric metasurfaces: metamodeling approach *Opt. Lette.* **45** 89–92
[35] Zandehshahvar M, Kiarashi Y, Chen M, Barton R and Adibi A 2021 Inverse design of photonic nanostructures using dimensionality reduction: reducing the computational complexity *Opt. Lette.* **46** 2634–7
[36] Zandehshahvar M , Kiarashinejad Y, Zhu M, Maleki H, Brown T and Adibi A 2022 Manifold learning for knowledge discovery and intelligent inverse design of photonic nanostructures: breaking the geometric complexity *ACS Photonics* **9** 714–21
[37] Lu L *et al* 2021 Physics-informed neural networks with hard constraints for inverse design *SIAM J. Sci. Comput.* **43** B1105–32
[38] Khatib O, Ren S, Malof J and Padilla W J 2022 Learning the physics of all-dielectric metamaterials with deep Lorentz neural networks *Adv. Opt. Mater.* **10** 2200097
[39] Lim J and Psaltis D 2022 MaxwellNet: physics-driven deep neural network training based on Maxwell's equations *APL Photonics* **7** 011301
[40] Liu D, Tan Y, Khoram E and Yu Z 2018 Training deep neural networks for the inverse design of nanophotonic structures *ACS Photonics* **5** 1365–9
[41] Xu X , Sun C, Li Y, Zhao J, Han J and Huang W 2021 An improved tandem neural network for the inverse design of nanophotonics devices *Opt. Commun.* **481** 126513
[42] Unni R, Yao K, Han X, Zhou M and Zheng Y 2021 A mixture-density-based tandem optimization network for on-demand inverse design of thin-film high reflectors *Nanophotonics* **10** 4057–65
[43] Shi X, Qiu T, Wang J, Zhao X and Qu S 2020 Metasurface inverse design using machine learning approaches *J. Phys. D: Appl. Phys.* **53** 275105
[44] Zhu R *et al* 2023 Direct field-to-pattern monolithic design of holographic metasurface via residual encoder-decoder convolutional neural network *Opto-Electron. Adv.* **6** 220148–1
[45] Liu Z, Zhu D, Rodrigues S P, Lee K-T and Cai W 2018 Generative model for the inverse design of metasurfaces *Nano Lett.* **18** 6570–6
[46] Ma W, Cheng F, Xu Y, Wen Q and Liu Y 2019 Probabilistic representation and inverse design of metamaterials based on a deep generative model with semi-supervised learning strategy *Adv. Mater.* **31** 1901111
[47] Gahlmann T and Tassin P 2022 Deep neural networks for the prediction of the optical properties and the free-form inverse design of metamaterials *Phys. Rev. B* **106** 085408
[48] An S *et al* 2021 Multifunctional metasurface design with a generative adversarial network *Adv. Opt. Mater.* **9** 2001433
[49] Jiang J , Sell D, Hoyer S, Hickey J, Yang J and Fan J A 2019 Free-form diffractive metagrating design based on generative adversarial networks *ACS Nano* **13** 8872–8

[50] Huang G, Liu Z, Van Der Maaten L and Weinberger K Q 2017 Densely connected convolutional networks *Proc. IEEE Conf. on Computer Vision and Pattern Recognition* pp 4700–8
[51] Doelman D S, Auer F F, Escuti M J and Snik F 2019 Simultaneous phase and amplitude aberration sensing with a liquid-crystal vector-Zernike phase mask *Opt. Lett.* **44** 17–20
[52] Steeves J, Wallace J K, Kettenbeil C and Jewell J 2020 Picometer wavefront sensing using the phase-contrast technique *Optica* **7** 1267–74
[53] Wallace J K, Rao S, Jensen-Clem R M and Serabyn G 2011 Phase-shifting Zernike interferometer wavefront sensor *Proc. SPIE* **8126** 110–20
[54] Jensen-Clem R, Wallace J K and Serabyn E 2012 *Characterization of the Phase-Shifting Zernike Wavefront Sensor for Telescope Applications* pp 1–7
[55] Salama M *et al* 2024 Keck primary mirror closed-loop segment control using a vector-Zernike wavefront sensor *Astrophys. J.* **967** 171
[56] Wenger T and Wallace J K 2025 Polarization-dependent metasurface for vector Zernike wavefront sensing with increased dynamic range *Opt. Lett.* **50** 726–9
[57] Lalanne P, Astilean S, Chavel P, Cambril E and Launois H 1998 Blazed binary subwavelength gratings with efficiencies larger than those of conventional échelette gratings *Opt. Lett.* **23** 1081–3
[58] Huang L *et al* 2012 Dispersionless phase discontinuities for controlling light propagation *Light: Sci. Appl.* **12** 5750–5
[59] Zheng G , Mühlenbernd H, Kenney M, Li G, Zentgraf T and Zhang S 2015 Metasurface holograms reaching 80% efficiency *Nat. Nanotechnol.* **10** 308–12
[60] Overvig A C , Shrestha S, Malek S C, Lu M, Stein A, Zheng C and Yu N 2023 Dielectric metasurfaces for complete and independent control of the optical amplitude and phase *Light: Sci. Appl.* **8** 92
[61] Lilja V A, Svärdsby A J, Gahlmann T and Tassin P 2025 A general framework for knowledge integration in machine learning for electromagnetic scattering using quasinormal modes (arXiv:2509.06130)