Review

Qizhou Wang, Maksim Makarenko, Arturo Burguete Lopez, Fedor Getman
and Andrea Fratalocchi*

Advancing statistical learning and artificial intelligence in nanophotonics inverse design

<https://doi.org/10.1515/nanoph-2021-0660>

Received November 1, 2021; accepted December 7, 2021;

published online December 22, 2021

Abstract: Nanophotonics inverse design is a rapidly expanding research field whose goal is to focus users on defining complex, high-level optical functionalities while leveraging machines to search for the required material and geometry configurations in sub-wavelength structures. The journey of inverse design begins with traditional optimization tools such as topology optimization and heuristics methods, including simulated annealing, swarm optimization, and genetic algorithms. Recently, the blossoming of deep learning in various areas of data-driven science and engineering has begun to permeate nanophotonics inverse design intensely. This review discusses state-of-the-art optimizations methods, deep learning, and more recent hybrid techniques, analyzing the advantages, challenges, and perspectives of inverse design both as a science and an engineering.

Keywords: deep learning; inverse design; metamaterials; nanophotonics; optimization.

1 Introduction

Throughout the past decade, the complexity of nanophotonics circuitry increased exponentially, mixing

non-linearity with dense nanoscale integration, advanced material manufacturing, and broad-band engineering functionalities [1–3]. Traditional intuition-driven design based on electromagnetic (EM) simulations manifests several weaknesses for modern device engineering that could efficiently incorporate all the factors mentioned above. While lacking clear demonstration that one preferred design is optimal against all required constraints [4], direct design often relies on brute-force searches, leading to either high computational costs or oversimplifications on design domains [5]. Inverse design, a paradigm in which the user provides the desired output, and the machine finds the required geometry, promises to address these bottlenecks by furnishing an automated platform for nanophotonic systems engineering [6].

Initial applications of inverse design (Figure 1) rely on classical numerical optimization algorithms [7, 8], employing suitably defined cost functions that the scheme minimizes after successive iterations. The minimum of the cost function defines the design's objective and identifies a set of parameters that specify the design sought. The main drawback of this type of inverse design is the lack of generalization ability: all the information learned during the search of one design is typically lost and not used in the future, requiring the designer to rerun the optimization. The recent development in artificial intelligence and deep neural networks (NN) [9, 10] is significantly reshaping this field, implementing new schemes that take advantage of the universal learning and prediction abilities of NN.

NN are statistical learning systems composed of networks of sequentially layered neurons, with each neuron processing a weighted sum of forehead layers through a non-linear activation function [11]. In theory, sufficiently deep NN layers have universal approximation abilities: they can learn a user-defined function from the analysis of input–output data sequences and perform predictions on future trends [12]. Such generalization ability originated new inverse design schemes that can predict novel device configurations by forwarding queries on the NN, with a generalized framework that easily transfers between similar design tasks [13]. A challenge lies in the chaotic

*Corresponding author: Andrea Fratalocchi, PRIMALIGHT, Faculty of Electrical Engineering, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia, E-mail: andrea.fratalocchi@kaust.edu.sa. <https://orcid.org/0000-0001-6769-4439>

Qizhou Wang, Maksim Makarenko, Arturo Burguete Lopez and Fedor Getman, PRIMALIGHT, Faculty of Electrical Engineering, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia, E-mail: qizhou.wang@kaust.edu.sa (Q. Wang), maksim.makarenko@kaust.edu.sa (M. Makarenko), arturo.burguetelopez@kaust.edu.sa (A. Burguete Lopez), fedor.getman@kaust.edu.sa (F. Getman). <https://orcid.org/0000-0003-4883-1068> (Q. Wang), <https://orcid.org/0000-0002-6761-8695> (M. Makarenko), <https://orcid.org/0000-0002-2318-1727> (A. Burguete Lopez), <https://orcid.org/0000-0002-6588-1668> (F. Getman)

Figure 1: Schematic introduction of inverse design. The figure compares four main sub-fields of inverse design in terms of generalization ability and dimensionality of design space. From the bottom to top, the introduced approaches are respectively: **Topology optimization** which utilizes gradient-based numerical iteration tools on the device shape, often represented by the spatial distribution of permittivity/refractive index. The exhibited applications include: (A) an amorphous silicon metalens. Adapted from [34]. (B) A broadband blazed metagrating. Adapted from [29]. (C) A wavelength demultiplexer operation at 1300 nm and 1550 nm. Adapted from [27]. **Heuristics** which imitates natural phenomenon and solves non-gradient based optimization problems. The exhibited applications include: (D) a metasurface with on-demand focal length composed of lattice opto-materials. Adapted from [132]. (E) A high-NA nanophotonic lens for GaAs nanowires. Adapted from [133]. **Deep learning** which trains neural networks to model the dual relationship between design parameters and corresponding optical responses. The exhibited applications include: (F) on-demand design of chiral metamaterials. Adapted from [134]. (G) Prediction of scattering coefficients of eight-shell nanoparticles. Adapted from [135]. (H) Design of reflective silver nanostructures. Adapted from [110]. (I) Hybrid metasurfaces composed of plasmonic structures and phase-change materials. Adapted from [118]. **Hybrid methods** which combines both advantage from deep learning and optimization-based methods, resulting in accelerated iteration and less requirement of the dataset. The exhibited applications include: (J) vertical grating couplers working in the C band. Adapted from [16]. (K) Flexible metasurfaces for arbitrary resonance control. Adapted from [19]. (L) Metallic metamolecules for polarization rotation. Adapted from [17].

instabilities of NN models, which sometimes arise because of the model complexity [14].

Recently, different research groups have proposed hybrid strategies combining both deep learning and optimization methods to overcome this issue [15–19]. These methods leverage sophisticated deconstruction approaches when changing design parameters during the iterative trial-and-error search process [18]. Hybrid schemes, the most recent trend in inverse design, are generally capable of training large spaces of design parameters while preserving the productivity among transferred tasks, thus retaining a strong generalization ability within highly complex design spaces.

2 Inverse design by optimization

These inverse design approaches leverage classical optimization techniques to explore the design space of possible solutions efficiently, ideally converging to the desired result exponentially faster if compared to a direct search [4]. The main idea of these methods is to define a suitable figure of merit (FOM), or equivalently a cost function whose minimum defines the structure sought to engineer. The cost function is progressively optimized with an intelligent exploration of its landscape by either constructing approximate models or exploring neighboring manifolds in the design space. The main difference between these

two approaches is merely the possibility to differentiate the cost function with various choice of design variables. When the parameters are distributed in continuous ranges, such as width, thickness, refractive index, and hole sizes, the gradient is computationally available, and topology optimization [20] provides a fast and robust search with a mathematically guaranteed rate of convergence in a local minimum of the FOM. If the problem considered is intrinsically not differentiable, for example, considering discrete material categories with characteristics selected from a library of candidates, heuristic optimization [21] will be a preferred choice. Although sometimes lacking a precise theory of convergence, these nature-inspired approaches can identify global minima of the cost function, even in large designs spaces where the objective function contains an exponential number of local, sub-optimal solutions.

2.1 Topology optimization

Gradient-based inverse designs exploit the first-order derivative (gradient) of the cost function to build approximate models that are minimized at each iteration [22–38]. Traditional topology optimization implements the gradient-descent optimization scheme [39, 40], while advanced techniques exploit convex optimization such as the trust-region [41], and moving asymptotes [42, 43]. In topology optimization based on density variables, the design parameters β are typically the permittivity $\epsilon_r(\mathbf{x})$ probed at discrete spatial lattice points $\mathbf{x}$. The resulting design space is a 2D binary image, in which the pixel density measures the distribution of different classes of materials [27–34]. Other topology approaches express β following a re-parameterization strategy such as, e.g., the level-set method [36–38]. In two-dimensional material profiles, the level-set method represents design variables via horizontal slices of a 3D surface, defined as the level-set function. During the optimization, the continuous movement of a cross-sectional plane simulates the variation of design variables. In contrast to explicit curve representation such as analytic equations, the level-set method traces topology changes, including merging, splitting, generating, and vanishing different shapes.

In topology optimization, the iterative calculation of the gradient is the most time-consuming operation. Yablonovitch et al. [44] developed a fast numerical implementation of this operation based on the adjoint method [45, 46]. This approach performs a forward simulation and an adjoint simulation, distinguished by the applied incident source. In the simple example of a density-based optimization task, with the FOM represented by the field intensity $|\mathbf{E}(\mathbf{x}_0)|^2$ measured at a certain diffraction order,

the forward simulation calculates the electric field $\mathbf{E}^{\text{fwd}*}(\mathbf{x})$ at one grid point $\mathbf{x}$ under normal incident light. The adjoint simulation, on the other hand, calculates the electric field $\mathbf{E}^{\text{adj}}(\mathbf{x})$ under incident light from the target diffraction order. This formulation allows expressing the gradient with respect to the permittivity at the grid point $\partial\text{FOM}/\partial\epsilon_r(\mathbf{x})$ as [44]:

$$\frac{\partial\text{FOM}}{\partial\epsilon_r(\mathbf{x})} = \text{Re}\{\mathbf{E}^{\text{fwd}*}(\mathbf{x})\mathbf{E}^{\text{adj}}(\mathbf{x})\} \quad (1)$$

with $\mathbf{E}^{\text{fwd}*}(\mathbf{x})$ representing the complex conjugate of electric field from forward simulation. Equation (1) is rapidly evaluated by using only the local values of the electric field, and significantly speeds up calculations.

Figure 2A reports a group of broad-band photonic devices designed by density-based topology optimization. In the work [33], Hammond et al. also investigated foundry design-rule constraints, including both systematic limitations and fabrication errors. These are introduced as inequalities in the form of $g_k \leq G_k$. For each k , G_k is a positive constant indicating the constraint for a certain metric g_k , including line width, line spacing, minimum area, and minimum enclosed area [47, 48]. To keep into account constraints of under and over-etching, which unavoidably occur at each fabrication step, Hammond et al. applied a conic image filter [49] on the design pattern. The modified design $\bar{\beta}$ satisfying all design-rule constraints listed above simulates the real-world fabrication scheme, including dilation and erosion. The topology design utilizes a classical binary representations of a silicon structure on SiO_2 substrate. At each design point, the algorithm interpolates the permittivity as: $\epsilon_r(\bar{\beta}) = \epsilon_{\min} + \bar{\beta}(\epsilon_{\max} - \epsilon_{\min})$, where $\epsilon_{\min}$ is the permittivity of SiO_2 and $\epsilon_{\max}$ is the permittivity of silicon. Constraining the parameter $\bar{\beta}$ in $[0,1]$ allows continuously varying the permittivity across the design area and, during the initial stage of the optimization, it facilitates convergence by exploring all possible continuous distributions of design parameters. The optimization then introduces regularization terms that penalize all parameters that are different from zero or one, which does not represent silicon or SiO_2 . After convergence, the algorithm produces a discrete material distribution satisfying fabrication's constraints. The above study contains three independent mirror, bend, and T-splitter designs working at 1500–1600 nm. The devices share a minimum line width of 90 nm, a minimum line spacing of 90 nm, a minimum area (for islands) of $0.08 \mu\text{m}^2$ and a minimum enclosed area (for holes) of $0.2 \mu\text{m}^2$. An increasing threshold coefficient that weights the penalization term ensures clear boundaries at the end of optimization. The algorithm iterates for 210 iterations and incorporates fabrication

Figure 2: Topology optimization in nanophotonic inverse design (A) design-rule-constrained nanophotonic devices optimized with foundry design rule checks. The panels depict the optimization process and efficiency of top: mirror, middle: bend and bottom: T-splitter. Adapted from [33]. (B) High-NA, high-efficiency metalenses designed with linearized topology optimization. Top left: SEM images of the fabricated metalenses. Top right: experimental and simulated efficiencies of metalenses with NAs of 0.2, 0.5, and 0.8, respectively. Middle row: measured intensity at the focal planes of the designed metalenses. Bottom row: transmission efficiencies in the wavelength range 580–700 nm of the designed metalenses. Adapted from [31]. (C) Two dielectric metalenses with NA = 0.78 and 0.94, respectively (left) and the corresponding focusing efficiencies during the optimization process (right). Adapted from [34].

constraints during the last 70 iterations. Figure 2A shows the resulting devices, which achieve a simulated transmission efficiency above 94% within a processing time of 6–8 h. In 2020, Augenstein et al. [50] introduced mechanical loads into the optimization of nanophotonic devices, enabling additive manufacturing and further enriching the library of design-rule constraints.

Topology optimization reported also the implementation of metalenses at ultraviolet [51], visible [52, 53], infrared [54, 55], and millimeter [56] wavelengths with high numerical aperture (NA). Unlike conventional refractive lenses, metalenses focus incident light with sub-wavelength elements, which simulate the required surface curvature [57]. Researchers investigated non-periodic

metalenses where the optical path varies among different designs to optimize the FOM, that is, the focusing efficiency at the focal length. Both relative (focused power/transmitted power) and absolute (focused power/incident power) efficiencies are accepted as part of the standard metric for metalenses. The optical properties, including deflection angle and phase response, are optimized at specific spatial points along the lens radius. As a result of this procedure, the numerical complexity of this class of designs is notably higher than the periodic metasurfaces mentioned above. To address this challenge, Phan et al. [31] studied a computationally efficient approach for the design of large-area metalens. The idea is to decompose the desired phase response along the radius

into wavelength-scale segments and apply topology optimization to design isolated metasurfaces with scattering properties that linearly fit the corresponding segments. By integrating all discrete parts, the design converges to a functional metalens, with a time complexity that reduces from $O(L^{2.4})$ to $O(L)$ for L being the degree of freedom in one-dimensional design space. Figure 2B depicts a series of 200 μm -wide metalenses designed to focus incident plane wave with NAs ranging from 0.2 to 0.8. The algorithm performs adjoint-based topology optimization on each 2 μm linear segment. Experimental results reveal efficiencies above 89% for all NAs in a wavelength range from 580 nm to 700 nm. The concept of linearization vastly reduces the system complexity compared to monolithic optimization. However, since the parallel tasks use perfectly matched layer (PML) boundary conditions [58, 59], the algorithm does not consider mode couplings between metasurface sections. This factor limits the linearization approach for an increasing number of segments, introducing strong couplings that finally obstruct convergence to an actual structure. Thus, the scheme searches for the best compromise between computational efficiency and linear fitting errors.

Recently, Mansouree et al. [34] reported an efficient design for high-NA metalenses with rectangular parameterization. Rather than directly modeling the permittivity at all design areas [60], the proposed method defines design variables as the widths of rectangular sub-wavelength structures. The design optimizes two metalenses with desired NAs of 0.78 and 0.94, formed by a total of $N = 19,200$ of 430 nm-thick amorphous Si bars. The nanostructures are symmetric in all four quadrants of the space, reducing the number of design variables to $N = 4800$. To calculate the gradient of the width variables, the authors implement an adjoint method with a perfect electric conductor (PEC) boundary condition. Figure 2C shows the resulting focal planes, located at 20 μm and 8.33 μm above the device surface for the metalens with NA = 0.78 and NA = 0.94, respectively. The rectangular parameterization inherits both advantages from periodic unit cells and monolithic design. It reduces the redundancy of bulk density variables while keeping the design diversity for complex, non-periodic nanostructures. The optimized metalenses with NAs of 0.78 and 0.94 achieve focusing efficiencies of 78% and 55%, respectively. This work assesses the performances of the proposed parameterization technique by designing two unit-cell metalenses as a control group under identical conditions. By comparing the simulated focusing efficiencies, the work reports a 10% performance improvement.

2.2 Heuristics and meta-heuristics

At variance with classical optimizers, heuristics and meta-heuristics exploit randomness in the searching process, imitating the behavior of different types of natural systems and phenomena. The primary example of heuristic methods is simulated annealing (SA), which originates in the context of non-differentiable, combinatorial problems, such as, e.g., the traveling salesman [61]. Problems of this type are mathematical ‘hard’ in the sense that the straightforward solution increases in time with a non-polynomial function, becoming quickly impossible to solve by any classical hardware.

Simulated annealing attempts to replicate the behavior of metals whose temperature decreases slowly, allowing the medium to converge towards a solid material, representing the global minimum of its potential energy. SA optimization imitates this process by setting the inverse design cost function $C(\beta)$ as the potential energy of an equivalent molecular dynamics system, with each molecular particle described by a position β_i representing a design parameter that we intend to find. The annealing then proceeds to the global minimization of the particles’ positions $\beta = [\beta_1, \dots, \beta_n]$ by progressively reducing the temperature with a suitably defined cooling scheme. If the temperature decreases sufficiently slowly, the probability distribution of the sequences of β found at each iteration follows the Boltzmann probability density $\sim e^{-C(\beta)/T}$ of a classical thermodynamic system [62]. At $T \rightarrow 0$, the molecular system converges to the set positions β representing the global minima of the cost function, solving the inverse design problem.

Zhao et al. [63] implemented this strategy to design diffusion metasurfaces that scatter the incident light uniformly in all directions. By exploring phase changing materials, this work designs a 2.4 $\mu\text{m} \times 2.4 \mu\text{m}$ reconfigurable metasurface [64] composed of 36 rectangular blocks as basic elements. The experimental results show a minimized directional reflection, with a 10 dB radar-cross-section (RCS) [65] reduction under normal incidence for both TE and TM polarizations. Another recent study by Xie et al. [66] proposed a magnetic resonator assembled by disordered nanostructures. The applied SA algorithm optimizes a five-by-five array of all-dielectric sub-wavelength scale resonators. The converged configuration reaches an experimentally measured peak magnetic field enhancement factor of 16.51 at 5.8675 GHz.

The main challenge of SA is that the convergence towards the optimal minimum lies in the heuristic idea of recreating a Boltzmann-type thermodynamic machine, which is not mathematically guaranteed to occur in every case. Ideally, addressing this problem requires sampling

the probability space with different SA runs launched with diverse (random) input conditions. This operation, which could be time-consuming, has a drawback: it does not employ the knowledge acquired during the past since different SA runs proceed independently. Meta-heuristics schemes such as particle swarm optimization (PSO) provide a possible approach to address this issue. These optimizations take inspiration from the social behaviors of crowded biological systems, including ant colonies, fish, and birds flocks. Despite the apparent simplicity of their social interactions, these systems are highly efficient in exploring a given terrain when looking for food, flowers, and other equivalents 'solutions' to their search problem. In computational PSO, the system is initialized with sufficiently large swarms of randomly chosen *particles* $\beta_1, \dots, \beta_n$, with each β_i representing a candidate solution to the problem. During each iteration, particles generate new configurations of β with the information from the best position found by each particle and neighboring particles. As in SA, the equivalent temperature of the system progressively decreases; unlike SA, however, temperature rescaling does not follow a pre-imposed annealing schedule but relies on the information acquired during the search [67].

Chen et al. [68] reported recent progress on PSO designed polarization beam splitters. The devices strengthen its compatibility to complementary metal-oxide-semiconductor (CMOS) by optimizing the silicon-on-insulator (SOI) structures. As shown in Figure 3A, the dashed area constrains the design space, which completes the functionality as a counter-tapered coupler. To reduce computational complexity, the authors split the coupler into ten silicon stripes within a total coupling length of 5 μm . The algorithm includes particles with positional vectors ps_n represented by widths of the segmented stripes and velocity vectors ve_n employed to update the former. The FOM is defined based on the polarization extinction ratio between output channels. As TE and TM waves are separated into the Bar and Cross port, respectively, the FOM is calculated by $P_{\text{TE_Bar}}/P_{\text{TE_Cross}} + P_{\text{TM_Cross}}/P_{\text{TM_Bar}}$, where P represents the transmitted power. After initialization and EM simulation, each particle records its best position bp_n and the global best position gp_n among the swarm. The update formula for ps_n and ve_n is as follows:

$$\begin{aligned} ve_{n+1} &= w_1 \times ve_n + r_1 \times \text{rand}() \times (bp_n - ps_n) \\ &\quad + r_2 \times \text{rand}() \times (gp_n - ps_n) \\ ps_{n+1} &= ps_n + ve_n \end{aligned} \quad (2)$$

where w_1 is the inertial weight, describing the momentum of previous movements. $0 \leq r_1 \leq 1$ and $0 \leq r_2 \leq 1$ are

the cognitive and social rate, representing the weights of individual bp_n and social memory gp_n of past best configurations. Figure 3B shows experimental results revealing high polarization extinction ratios over 16 dB for both TE-type and TM-type grating couplers obtained with PSO. Recent PSO-based researches also provide solutions across a wide range of nanophotonic devices including power splitters [69], solar cells [70], achromatic metalenses [71] and phase changing antennas [72].

The main issue in PSO is the speed of the algorithm's convergence, which follows the set of worst-performing particles. Evolutionary schemes, such as, e.g., genetic algorithms, try to address this issue by selectively mixing individual information arising from the design population [73]. Spuhler et al. [74] first proposed a genetic algorithm for the design of fiber-waveguide coupler. By splitting the SiO_2/SiON functional area into segments with identical lengths and flexible widths, a population of design candidates containing a distinct width sequence evolves with random mutations, cross-overs, and die-out. A non-intuitive structure differing from any existing design emerges after 1132 optimization steps, with a 2 dB reduction on coupling loss compared to direct butt-coupling devices.

Following this initial genetic approach, a series of photonic devices have been proposed with increasing design complexity during the past decade [75]. Figure 3C reports a cluster of dielectric nano-antennas for polarization-encoded color display obtained by genetic algorithms. The work exploits sub-wavelength nanostructures that alter their optical properties for different light polarizations. To address this design task, Wiecha et al. [76] proposed an evolutionary multi-objective optimization approach, which maximizes simultaneously the scattering efficiency at demanded wavelength, as well as both TE and TM polarizations. The periodic unit cell is a set of four silicon blocks, with horizontal sizes ranging from 60 nm to 160 nm. The entire design space is a 600 nm $\times$ 600 nm area, with available parameter combinations larger than 1×10^{15} . The algorithm defines a fitness function that integrates polarization-dependent metrics and evaluates all candidate designs in each iteration. The selected high-performance designs reproduce a new generation by cross-over and random mutation and finally lead to a convergence indicated by the Pareto front, a standard metric for multi-object programming [77]. In this specific task, the Pareto front is the set of design parameters whose modifications enhancing the scattering at one polarization cause the efficiency to drop for the other polarization. The Pareto front keeps expanding as the optimization processes until the device reaches saturation, indicating

Figure 3: Heuristic optimization in nanophotonic inverse design (A) a silicon polarization beam splitter placed on SiO_2 substrate. Bar and Cross are two polarization-related output channels. (B) The measured spectra of the fabricated beam splitters in (A). (A) and (B) are adapted from [68]. (C) Multi-resonant dielectric nano-antennas with fixed resonance at 550 nm for one polarization X, and ranging resonance peak from 400 nm to 650 nm for the other polarization Y. Adapted from [76]. (D) Plasmonic gold antennas with maximum scattering on the desired direction. Each panel shows the fittest solution at different iterations, starting from initialization to convergence. (E) Optimized examples of directional antennas for various target angles and scattering environments. Left: backscattering in vacuum. Right: scattering towards glass substrate. (D) and (E) are adapted from [78].

the best performance. As depicted in the bottom panel of Figure 3C, the X-polarization (TE or TM) scattering peak is at 550 nm. In contrast, the Y-polarization (TM or TE, respectively) scattering peak is evolved gradually from 400 nm to 650 nm, all with consistent efficiency (within $\pm 20\%$) in both polarizations. Compared to direct search, the evolutionary algorithm rapidly explored the design space within 100 iterations from a population of 20 candidate designs.

The work in [78] exemplifies a light-directional antenna, which maximizes the ratio of direct scattering at a specific direction to the residual scattering directions. Inheriting similar optimization strategies from [76], the

authors leveraged a binary-shaped design space containing $40 \text{ nm} \times 40 \text{ nm} \times 40 \text{ nm}$ gold blocks (represented by “1”) and air (represented by “0”). The total amount of all possible arrangements extends to 10^{11} . As shown in Figure 3D, the algorithm starts with an initial population size $N_p = 500$ and reaches convergence after 50 generations for fixed azimuthal angle $\varphi = 45^\circ$. The scattering antenna evolves spontaneously to form a structure resembling a Yagi–Uda RF antenna [79] with one driving feed, one reflector, and one director. Figure 3E reports examples of scattering antennas optimized under different conditions (e.g., propagation media, diffraction windows) from this work.

3 Inverse design by deep learning

Recent trends in inverse design take advantage of the large body of research available in the context of deep learning [80]. Deep learning offers advantages over classical optimization schemes, such as the ability to predict design output given a set of design parameters at the input, without the need of iteratively minimizing the FOM function. Deep learning encompasses a large variety of schemes and diverse taxonomies of classification. Traditionally, a standard classification divides deep learning into supervised, unsupervised, and reinforcement learning schemes. Supervised learning [80] uses labeled datasets to train the network to learn the required input–output tasks, declaring the demand for high-quality, large-scale datasets. This condition is relaxed in unsupervised schemes [81], which do not use any pre-assigned information. Reinforcement learning, on the contrary, trains an automatic agent to control the design variables, with possible actions such as increasing/decreasing thickness, changing material types, increasing/decreasing unit diameters [82].

Due to the increased complexity of nanophotonic devices, modern inverse design schemes often integrate these different learning models into a single approach [83]. Therefore, state-of-the-art inverse design techniques have blurred boundaries between these three learning domains. Another possible classification follows the recent progress in the computer vision community [84], which discusses deep learning schemes in terms of discriminative and generative models. In the specific context of nanophotonic inverse design, discriminative models learn the one-to-one mapping between optical responses and material configuration layouts. Generative models, on the contrary, learn the statistical distribution of potential designs that minimize the FOM, achieving a one-to-many mapping, i.e., a single desired output and many possible resulting designs.

3.1 Discriminative models

Discriminative models bypass the repeated EM input–output simulations, often the bottlenecks of the optimization schemes reviewed in the previous section. Discriminative models are essentially unique configurations of deep learning architectures that learn the relationship between designs in the parameter space and their optical responses (reflection, transmission, amplitude or phase) [85–91]. Based on the study of plasmonic

nanostructures, Malkiel et al. [92] introduced a bidirectional network composed of geometry-predicting-network (GPN) and spectra-predicting-network (SPN). The authors train the GPN in a supervised manner with a dataset composed of 15,000 randomly generated device layouts containing eight geometrical parameters, including width, height, rotation angle, and the presence of positional elements in the H-shaped framework. Additional to the device data, finite elements simulations (FEM) calculate the corresponding transmission spectra. The GPN predicts the geometrical design parameters by processing 86 sample points from both TE and TM spectra and dispersive material coefficients. After training, the SPN, constructed in a cascade with the GPN output layer, retrieves the spectra from the eight geometrical parameters launched at the input. The network efficiently leverages the limited dataset by using unseen shapes generated by GPN as the training set of SPN. The loss function and gradient propagate to each layer in both networks. The end-to-end training of the network converges within 2 h, after which the time required to complete queries for potential nanostructures lies in the millisecond range. As exemplified in this work, the GPN retrieves classic nanophotonic structures such as nanobars, L-shaped, and split-ring resonators. To efficiently extract the hidden features from nanostructures, the training of a forward model (device-to-response) and an inverse model (response-to-device) is beneficial. Similar to the GPN and SPN introduced above, the two models can be either constructed and trained independently [93] or with inner connections, such as the generative models that will be introduced in the next section.

Other discriminative models leverage the generalization ability of NN in two main directions [94, 95]. On the one hand, they exploit the fact that a trained system predicts novel results for an arbitrarily large number of queries; on the other, they use the information acquired by the NN during training and transplant it via transfer learning [96] in a different task. Work proposed by Qu et al. [89] embedded and demonstrated transfer learning in a nanophotonic inverse design scheme. The approach includes the precise prediction of transmission spectra for an 8-layer thin film and a 10-layer thin film. Figure 4B shows an overall schematic of the transfer learning process. A 7-layer fully connected network (FCN), acting as a base network, is trained on a complete dataset containing 50,000 simulation examples. After the training, a second FCN acting as the transferred network replaces the top n layers of the base network, leaving the rest $7-n$ layers sharing weights from the trained one. The second network demonstrates

Figure 4: Discriminative deep learning model for spectra predicting and design parameter retrieval (A) bidirectional network for plasmonic nanostructure design. Left: geometry predicting network with TE&TM spectra and material properties as input. Right: spectra predicting network, concatenated to the first network in a cascaded structure. Adapted from [92]. (B) Knowledge transfer between the design of 8-layered and 10-layered transmissive films. As the shared network layers keep increasing, both transfer schemes (8-layered to 10-layered and the inverse) show better prediction precision compared with direct training. (C) Example of predicted spectra compared to real spectra from the 8-to-10-layer transfer learning tasks. Left: direct learning. Right: transfer learning. (B) and (C) are adapted from [89]. (D) Results of the discriminative design on the plasmonic silver absorber. The device structure is encoded as 2D images, with absorption spectra predicted by the RNN and FCN layers. Adapted from [88].

the advantage of transfer learning by training 500 examples sliced from simulation results in a different design task, which are insufficient for training a model from scratch. As a quantitative example, Figure 4B illustrates the

transfer learning schemes where the source task and target task switch between 8-layer and 10-layer films. By increasing the number of shared layers, the mean square error (MSE) reduces by 50.5% and by 23.7% in the

8-to-10-layer and 10-to-8-layer transfer model, respectively. The improved performance reveals that the knowledge from the complete dataset is preserved and leveraged in training the sliced dataset. Figure 4C shows a test example, comparing the performance of direct learning strategy and transfer learning in a working bandwidth between 400 nm and 800 nm. This work demonstrates that while comprehensive modeling of physical effects could be difficult to achieve, the implicit knowledge carried out by pre-trained NN is a practical approach to mitigate this issue, especially when simulations are time-expensive.

The examples discussed above include different FCNs to grasp the physical relation between characteristic parameters of a design and the output spectral response furnished by the system. A recent trend of nanophotonic inverse design [97] introduces a high-level hierarchical representation of data for fast image processing and sequential analysis. Sajedian et al. [88] exemplified a design of periodic silver nanoparticles based on a combination of convolutional neural networks (CNN) and recurrent neural networks (RNN). Two-dimensional images of 100×100 pixels 2D encode the top view of three-dimensional structures with a definite constant thickness along the Z-axis. The authors use a training dataset with random images comprising a large variety of geometries, width, length, position, and rotation. Finite-difference time-domain (FDTD) simulations [98] computes the absorption spectra of 100,000 structures in the dataset. The applied CNN is ResNet [99], which is a powerful architecture in computer vision for pattern recognition problems. ResNet contains shortcut connections that link non-adjacent network layers, providing a straight path for identity mapping [100]. This feature prevents the gradient vanishing and gradient exploding issue in training deep CNNs and grants the model a more robust capability to extract information from the 2D structure.

The output arising from ResNet follows a time-distributed RNN layer connected to a final predictive FCN stage. Due to the increased complexity of this network, the reported training requires approximately one week to converge. The trained network predicts the absorption spectra of silver structures. As a proof of concept, the authors test the model on a validation dataset containing 1000 sample points on the spectra between 800 nm and 100 nm, which reaches an $\text{MSE} = 4.2591 \times 10^{-5}$. As shown in Figure 4D, vital characteristics such as peaks and valleys are fitted well with the simulation results. This work extends the design space from geometrical parameters to arbitrary 2D shapes by applying image processing networks.

An empirical principle in machine learning dictates that the higher the NN model complexity, the larger the dataset needed to obtain good prediction accuracy on untrained data [101]. This condition also applies to and challenges NN-based inverse nanophotonic design. To train a more sophisticated NN with linearly expanding layers, the size of datasets can increase drastically, requiring vast amounts of EM simulations. From this point of view, the computational cost of training deep NN scales up by requiring more resources than optimization-based methods. However, the main strength of the deep learning-based inverse design is transfer learning: transferring the knowledge acquired by the progressive solution of many numerical simulations into different tasks. In the future of NN-based inverse design with increased model complexity, we expect a more decisive role of transfer learning techniques in formulating the network model.

3.2 Generative models

An issue with discriminative models is that one trained network can only provide a single, one-to-one mapping between the design variables and the spectral response at the output. However, in optics, various distinct structures can achieve a specific response. Generative models address this issue by computing distributions of possible structures that achieve analogous target responses. The models start the training by sampling a ‘latent space’ constructed from known distributions of random vectors. In most cases, these probabilities are multi-dimensional Gaussian distributions, with enough randomness provided for the model to learn a complex projection from the latent space to the distribution of functional nanophotonic devices. In general, the generative models decode the sampled vector to produce design parameters. Different generative models implement diverse strategies for learning these complex projections.

Generative adversarial networks (GANs) [102] represent the most prominent class of deep learning generative methods developed during the past decade in various research fields, including image translation [103, 104], privacy protection [105], natural language processing [106]. We here summarize a generic schematic of conditional GAN (cGAN) [107] architecture implemented in nanophotonic inverse designs. The essential parts of a GAN comprise a generator **G** network, which provides the design, and a discriminator **D**, also known as the critic that evaluates the provided designs. In the context of inverse design, an additional simulator network **S** serves as an external unit for predicting the optical response from the

generated design and measuring the pre-defined FOM corresponding to specific tasks. The models train the generator $\mathbf{G}$ and discriminator $\mathbf{D}$, simultaneously and adversarially. As depicted in Figure 5A, the pioneering work of Liu et al. [108] defines $\mathbf{z}$ as the random vector sampled from latent space, and $\mathbf{T}$ as the conditional variables that indicate the target response, which is the transmission spectra measured across TE and TM polarizations in this work. Under the goal of modeling and designing photonic devices, the generator $\mathbf{G}$ produces structural designs $\mathbf{G}(\mathbf{z}, \mathbf{T})$ resembling actual structures $\mathbf{X}$ in the provided dataset. In contrast, the discriminator $\mathbf{D}$ is trained to distinguish $\mathbf{G}(\mathbf{z}, \mathbf{T})$ from $\mathbf{X}$. To carry out this task, the discriminator $\mathbf{D}$ predicts a $[0,1]$ ranged value l , which represents the probability of each structure being real (existing in the dataset) or fake (generated by $\mathbf{G}$). The authors train the discriminator $\mathbf{D}$ to minimize the distance between l and the actual categories, while the objective

of the generator $\mathbf{G}$ is to maximize this distance. These adversarial training goals finally converge to a Nash equilibrium [109], representing a balance between the productivity of $\mathbf{G}$ and the discriminative ability of $\mathbf{D}$. At this stage, the generated design images $\mathbf{G}(\mathbf{z}, \mathbf{T})$ share similar geometrical features with $\mathbf{X}$ in the training set, while $\mathbf{D}$ fails to identify them from the actual structures $\mathbf{X}$. According to various design tasks, the simulator $\mathbf{S}$ is pre-trained on labeled datasets, including device structures and the corresponding simulation results (transmission/reflection/absorption spectra, scattering coefficient) in different polarizations and wavelength ranges. It calculates the FOM by predicting the optical response on the generated structure $\mathbf{G}(\mathbf{z}, \mathbf{T})$. The algorithm then integrates the optimization of FOM into backpropagation and gradient descent of $\mathbf{G}$ layers. The generator can produce optimized designs at convergence with all necessary structural constraints and topologies shared in the dataset.

Figure 5: Generative adversarial network (GAN) implemented in nanophotonic inverse design (A) GAN's generic schematic applied in nanophotonic design, which includes a generator, discriminator(also known as critic), and simulator. The generated device pattern is further optimized for fabrication by a smoothing procedure. Adapted from [108]. (B) A GAN-based silver structure design under arbitrary reflection spectra. The panel shows nine examples of device binary images generated by GAN and the corresponding reflection spectra as condition (solid line) and predicted response (dash line). Adapted from [110]. (C) Basic training strategy of GAN in the design of optical cloaks. The panel shows a successive data augmentation technique, where FEM simulations evaluate the newly generated geometries to complement the initial dataset. (D) Left: average and minimal scattering coefficient among the top-1000 configurations generated by the network. Right: the scattered field in an optimized 2-dimensional optical cloak. (C) and (D) are adapted from [112].

Rho et al. [110] proposed an alternative deep cGAN model for predicting reflection spectra of silver nano-antennas. The cross-sectional design space of the surface reflectors is limited in a square of $500 \text{ nm} \times 500 \text{ nm}$, represented by 64×64 binary images. FDTD simulations provide reflection spectra with 200 frequency points sampled from 250–500 THz. This work further optimizes the generator network architecture by constructing two separate convolutional blocks for the feature extraction of the random vector $\mathbf{z}$ and conditional variables, specifically, user-defined reflection spectra. Then the model fuses the output of two convolutional blocks into one 1024-channel tensor and eventually produces the design image. Instead of building an independent simulator model shown in Figure 5A, the FOM is measured directly from FDTD simulation on generator-suggested designs. In the early stage of training, the generator produces chaotic distributions of pixel values. As the training proceeds, the output images gradually show precise geometric shapes without additional regularization. Figure 5B shows examples of generated designs. The solid black spectra are the design objectives corresponding to the dataset's structural images (black). Moreover, the structures drawn in red are the GAN-optimized designs, which share a similar topology yet comprise deformations such as erosion, dilation, and mirroring. The mean absolute error (MAE) between the reconstructed reflection spectra (red dash line) and simulated spectra is 0.0322. This result reveals GAN's strength in modeling one-to-many mappings. Furthermore, the authors test the model on user-defined reflection spectra generated by curve functions. Despite the non-existing nature of such reflection patterns, the model achieves a <5% MAE between the reconstructed and the user-defined spectra, revealing a good approximation. An advantage of GANs is to explore the latent space implicitly, with seed vectors sampled from pre-defined probability distributions [111]. The generated randomness helps the networks find equivalent structures that differ from the training set, further enriching the existing dataset. Recently, Blanchard et al. [112] proposed a successive training strategy employing GAN-based data augmentation for the design of dielectric optical cloaks. The design space is binary images of a split ring resonator (SRR) with a central PEC circular boundary. When the incident light propagates through the cloak, the wavefront remains identical. This work defines a relative scattering coefficient to measure the difference between background field and scattered field. Similar to the architecture described in Figure 5A, the proposed approach comprises a generator, a discriminator, and a

forward simulator that is pre-trained to predict scalar scattering coefficients. The authors freeze the weights of the forward network during GAN training while its output participates in the backpropagation. Minimizing the relative scattering coefficient demands the forward loss to be calculated simply by the distance between the predicted scalar value and constant 0. A total number of 13,000 FEM simulation results initialized the dataset, where the cloaking performance varies according to the shell geometries.

Figure 5C illustrates the unsupervised data augmentation technique applied every 60 iterations, where the forward simulator first evaluates newly generated configurations. The 1000 best geometric structures with the lowest scattering coefficients are simulated using FEM and added to the dataset. The model then utilizes the updated dataset to refine the prediction accuracy of the forward simulator and subsequently train the GAN. As a consequence of such a joint training strategy, GAN-generated geometries converge to a stage with minimum scattering coefficients by backpropagation, with high-quality samples increasingly dominating the dataset. Figure 5D exemplifies the learning curve of the average and the best scattering suppression at every 60 epochs. The right panel visualizes the normalized H_z field of the optimal configuration, on which the ratio of the scattered field to background field reaches 0.0089, implying almost uniform propagation.

Ma et al. [113] introduced the concept of conditional variational autoencoder (cVAE) [114, 115] in the inverse design of reflective metasurfaces. Depicted in Figure 6A, the cVAE contains three correlated sub-models, respectively, recognition model, generation model, and prediction model. The authors train the prediction model on simulated data. After convergence, the model can predict reflection spectra that cycle into the other two models as additional information. Acting as the 'encoder', the recognition model projects the input geometry to a 20-dimensional space by assigning mean value μ and variance σ to the pre-defined Gaussian distribution. A reparameterization technique is applied to sample latent vectors from this probability density function, described as follows:

$$\mathbf{z} = \sigma \mathbf{i} + \mu \quad (3)$$

where $\mathbf{i}$ follows a standard Gaussian distribution. The latent vector $\mathbf{z}$ is then decoded by the generator model to reconstruct the input geometry, given the reflection spectra as conditional variables similarly to the GAN training scheme mentioned above. The training proceeds to show the formation of a clear unit cell pattern with accurate optical response in the range 40–100 THz. VAE is

Figure 6: Variational autoencoder (VAE) assisted nanophotonic design (A) a c-VAE based metamaterial design system, including a prediction model, recognition model, and generation model. (B) Visualization of the reduced latent space, where three types of geometries and two-types of optical responses are depicted on the plot. (C) Examples of on-demand metamaterial reflector design. Two equivalent structures are generated under the given spectra condition. (A), (B) and (C) are adapted from [113]. (D) Examples of broadband power splitters with $2.25 \times 2.25 \mu\text{m}^2$ footprint. A 550 nm working bandwidth is achieved for all devices with arbitrary splitting ratios. Adapted from [116].

a highly structured model, where adjacent points in the latent space correspond to continuous change on device structures. Exploring the converged Gaussian distribution reveals three groups of devices in the scatter plot of Figure 6B, respectively, cross, split-ring, and H-shaped nanostructures. The authors further characterize the reflection properties by locating the most robust resonance and dividing the nanostructures into two types with resonance frequency under and above 60 THz. As depicted in the right panel of Figure 6B, different types of spectra distribute among all clusters. This result demonstrates that prior-defined device form factors can be exploited in the design while each form sufficiently covers a wide range of optical responses. Figure 6C shows an example of generated structures under given conditions, which include TE to TE (R_{xx}), TM to TM (R_{yy}), and TE to TM (R_{xy}) reflection spectra. The required response contains two resonance valleys on R_{xx} and R_{yy} , respectively. Two generated structures show in the below panels, where their numerical simulations show major agreement with the required spectra at resonant points.

In the second example of cVAE-assisted inverse design, Tang et al. [116] developed a set of nano-patterned power splitters with user-defined split ratios. The devices are silicon-based square couplers with a footprint of $2.25\ \mu\text{m} \times 2.25\ \mu\text{m}$, which distribute 400 independent etching hole positions across the square area. The configurations of etching holes on the coupler is a 20×20 vector scaled in the range [0,1], each variable indicating the absence (<0.3) or a rescaled diameter size (>0.3) of one etching hole. The dataset contains 15,000 simulated transmittances of randomly generated hole vectors. Two ideally flat transmission spectra constitute the dependent variables to achieve the desired split ratio of the output ports. Above the original c-VAE architecture with a latent space, the authors introduced the adversarial block [117], an alternative branch of NN that validates the sampled latent vector $\mathbf{z}$. This network estimates a diversity of optical responses from points across the latent space distribution, following the fact that the actual response of a power splitter is commonly irregular and not smooth as the ideal condition. The network obtains the highest performance by forcing the latent space distribution to generate various yet effective samples that correspond to the defined split ratio. Figure 6D exemplified two prototypes generated by the model, with splitting ratios of 6:4 and 8:2, respectively. The optimized design achieved over 87% total transmission efficiency across the working bandwidth from 1250 nm to 1800 nm. The training in [116] also implements an active learning method, where data

augmentation based on FDTD simulations operates after each stage of training. Besides VAE, other autoencoder-like models exclude the random sampling process by directly reducing the input dimensionality to a fixed-length hidden vector. This lightweight model also accelerates nanophotonic applications with low computational burden, such as the design of reflective metasurface [118] and photonic topological insulators [119].

Illustrated in the above examples, the main advantage of generative versus discriminative models is producing novel data. Generative models learn the representation of the data structure and can supplement and refine the training set. Among various application fields such as text generation [120], texture filling [121], and image translation [104], the representation-learning of generative models also show significant improvements for inverse design tasks. Initially trained on a dataset containing nanostructures and optical properties, the converged network produces an optimized design distribution given certain conditions. The typical approach of external validation is EM simulation, which labels the generated design with the corresponding spectral response and eventually contributes to enhancing the original dataset. Besides this, direct experimental measurement on the fabricated devices also serves as an effective data generation method, compared to the intense computational demand of numerical simulations.

4 Merging deep learning with optimization techniques

While deep learning models benefit from strong predicting abilities, they often require vast initial training resources. Efficient training of NN to learn to express user-defined functions is presently still a subject of intense research. Even with the latest generation of training schemes, instabilities and non-convergence problems [122] persist, especially when the network architecture becomes sufficiently large [123]. On the contrary, optimization-based inverse design requires the knowledge of a single resource, the FOM, to tackle multiple design objectives and independent tasks at one time. Optimization schemes provide a more straightforward yet solid evolution framework for the design parameters. Blending deep learning and optimization methods is a current trend in nanophotonic inverse design. Such a hybrid framework leverages fast gradient retrievals and low-cost evolution from optimization while inheriting the generalization ability of deep learning models to reduce the design dimensionality and accelerate convergence towards the optimal design.

Figure 7A shows a hybrid design model based on GAN and topology optimization. In this work, Jiang et al. [15] demonstrated that the adjoint method could reduce the computational cost for training deep NN. The proposed model, named GLOnet, replaces the discriminator in standard GAN architectures with EM simulations. The design divides the diffractive metagratings equally into 256 segments, of which a 256-dimensional vector represents the refractive indexes. The desired deflection angle and working wavelength are the conditional variables to the generator. GLOnet implements an image processing model concatenated with a Gaussian filter to eliminate small discrete features on the structure to fulfill fabrication constraints. As described in Section 2, the gradient of FOM is directly calculated from the forward and adjoint

simulation results, while all network layers backpropagate this gradient to update weights. The proposed method combines fast gradient calculations from topology optimization with the generation capability of GAN. This work has no dependence on the initial dataset, as all generated patterns are simulated dynamically at each iteration. Like topology optimization, GLOnet introduces an additional penalization term to discretize the device structure by forcing the refractive indexes to approach one of the two given materials (silicon or air). An increasing trend of design structures with higher transmission efficiency and binarization of refractive indexes emerge during the training.

In the following work [124], the authors further compared GLOnet with pure topology optimization

Figure 7: Hybrid design methods with deep learning models performing complex projection on design space (A) schematic of global optimization network (GLOnet) as a combination of GAN architecture and topology optimization approaches. The generated design candidates are utilized to calculate the gradient for the FOM by adjoint methods. Adapted from [15]. (B) Deflection efficiency of devices optimized by gradient topology optimization and the proposed GLOnet-based approach. The working wavelength varies from 800 nm to 1200 nm and the deflection angle varies from 40° to 70°. Adapted from [124]. (C) The PCA-based dimensionality reduction of design space for vertical grating couplers. Exhaustive mapping is conducted at the reduced sub-space to generate the optimized design. (D) Top: coupling efficiency of the optimized devices measured from the principle hyperplane. Bottom: back-reflection measured from the principle hyperplane. (C) and (D) are adapted from [16].

methods. Figure 7B visualizes the optimized device performances under working wavelength varying from 800 nm to 1200 nm and deflection angle from 40° to 70° . The results reveal a performance improvement that 90% of the GAN-generated devices achieve higher efficiency than topology-optimized ones. The authors explain this performance by considering that GLOnet searches the entire design space via deep NN, while topology optimization performs a one-path search for every independent task. Crosstalk among different GAN-generated samples accelerates the global convergence by exploring possible high-quality devices, granting them a higher impact factor in the final probability density. A following up work [125] expands the model to a multi-objective flat design where the network produces both material types and device layouts.

In the previous examples, deep NN architectures, including GAN and VAE, model the design space. Daniele et al. [16] proposed an alternative path supported by principal component analysis (PCA) in the design of vertical grating couplers. Five geometrical parameters constitute the dimensions of the design space. An effective Fourier-type eigenmode expansion simulator [126] calculates the coupling efficiency from the device to a vertically placed optical fiber. An auxiliary NN predictor is trained on this dataset to accelerate the first-stage optimization. As shown in Figure 7C, the initial optimization conducted by a local search algorithm generates a sparse collection of rough designs, whose proximity in the design space represents the potential manifold of high-performance devices. Then, PCA reduces the dimensionality from 5 to 2 orthogonal principal axes, with essential information conserved during the process. The linear combination of the basis vectors composes a continuous 2D space, where exhaustive mapping rapidly characterizes all possible designs. Figure 7D depicts the mapping results with coupling efficiency larger than 0.7 and back reflection below 15 dB. The authors then select the outranking designs and evaluate them with FDTD full-spectra simulation as a second-stage optimization. All designs perform with 1 dB bandwidth covering the telecommunication C-band ranging from 1530 nm to 1565 nm. The authors demonstrated that the search complexity is reduced by orders of magnitudes, resulting in 400 times less computation time during the optimization.

Nanophotonic inverse design can also help address partial differential equation (PDE)-constrained optimization problems that rely on physical principles. As discussed in the previous section, in recent discriminative inverse design approaches, this idea applies to train networks solving the PDE [127] and approximating the optical response

[87, 92]. Another approach utilizes traditional optimization methods as the core unit that produces varying design samples. Getman et al. [18] implemented a hybrid inverse design framework integrating PSO and pre-trained predictors. Focusing on the design and applications surrounding flat-optics surfaces, the authors demonstrated the universal expressivity of these devices from physical principles. The work demonstrates that in suitably engineered nanoresonators, linear propagation and resonances blend, defining a system response $H(w)$ that can approximate any user-defined function [128]. The design problem of the entire device then reduces to characterizing a combination of nanoresonator shapes. The authors designed and implemented an Autonomous Learning Framework for Rule-based Evolutionary Design (ALFRED), a parallel software exploring a large design space consisting of multimodal nanoresonators described by 2D binary images and discrete thickness values. As depicted in Figure 8A, a parallel swarm optimizer performs the collective search among inertia, social, and memorial elements of designs. Complementing this, an NN reduces the high computational cost of running first-principle EM simulations to a forward propagation within milliseconds during the iterations. A CNN-based, switch-connected simulator network predicts the individual optical responses of devices with various shapes and thicknesses. Two popular example devices, a polarization beam splitter, and a dichroic mirror are fabricated and measured experimentally, showing over 95% transmission efficiency across the visible range.

By using ALFRED, the authors reported a new type of two-subpixel metasurface color display. A conventional LCD color display employs a linear polarizer (LP) sandwich structure, wrapping liquid crystal (LC), and three color filters. The proposed design uses only one LP and LC layer, enabling the free rotation of backlight polarization. Figure 8B shows the schematic, where the unpolarized backlight sources are first polarized by the LP and then rotated by the following LC cell. Two polarization-sensitive metasurfaces filter the output of the LC cell, and through the use of two different chromaticities, compose a displayed color. With this approach, a suitably designed pair of metasurfaces reconstruct a high-fidelity color space similar to RGB displays by manipulating four input variables (i.e., polarization and intensities for two pixels). As shown in Figure 8C, the simulation results and the measured response cover a large gamut area in the chromatic space, indicating a wide color range for displays. In follow-up works [19], the design space expands to combinations of arbitrary polygons, ellipses, and cuboids.

Figure 8: Hybrid design methods leveraging deep learning models as fast EM solver (A) An autonomous learning framework for rule-based evolutionary design, comprising a fast CNN predictor and a parallel PSO algorithm. (B) Generalized two-subpixel color display designed with a polarization-sensitive response, where backlight (BL), linear polarizer (LP), and liquid crystal (LC) layers alter the incident polarization to the flat-optics subpixels. (C) Left: chromaticity gamut of the subpixels in the color space. Right: SEM images of the fabricated subpixels. (A), (B) and (C) are adapted from [18]. (D) A hybrid system for the design of diatomic polarizers including cooperative coevolution algorithm and a pre-trained compositional pattern-producing network. (E) Left: SEM image of two optimized structure with rotation angles of 30° and 45°. Right: incident polarization (blue dash line), desired rotation angle (green/red solid line) and experimental results (green/red dash line). (D) and (E) are adapted from [17].

Additionally, the training introduces random perturbations on the structure shape to build prediction robustness against fabrication errors. The authors also implemented a t-SNE dimensionality reduction technique and K-means clustering to analyze the predicted responses dynamically. The training applies the data augmentation technique to

build robust predictions by absorbing new samples from the high-prediction-error clusters.

The work of Liu et al. [17] also explained the advantage of using NN as a fast EM solver. To design and fabricate diatomic polarizers, the authors developed a hybrid system including a compositional pattern-producing network

Table 1: Comparison of inverse design approaches.

	Differentiable	Transferable	Group optimization	Model constructing cost	Production cost
Topology optimization	Yes	No	No	Low	High
Heuristics	No	No	Yes	Medium	High
Discriminative model	Yes	Yes	No	High	Low
Generative model	Yes	Yes	Yes	High	Low
Hybrid model	Yes	Yes	Yes	High	Medium

(CPPN) and a pre-trained simulator network. The CPPN uses convolutional layers, which translate a set of input variables into the pixel value of a device pattern. The generated pattern includes two adequately isolated nanoparticles, namely, meta-atoms as the unit cell of the polarizer. The authors train the simulator network to predict the far-field response as the superposition of responses from two independent meta-atoms. As depicted in Figure 8D, this work applies cooperative coevolution (CC) [129] algorithm to optimize the device structure. Due to the design species represented by the dimensionality-reduced parameterized vectors, the evolution processes fewer variables, thus saving computational efforts. A group of designed meta-atoms output different polarization angles between 15° and 60° with linearly polarized incident lights. Each specific design converges within 20 s on a single-GPU workstation. As a final verification, FEM provides a full-spectra simulation of the optimized design. Figure 8E exemplifies two devices with orientation angles of 30° and 45° . The blue dashed curves indicate the incident polarizations, while the solid line and dashed line in red and green represent the desired rotation angle and measured value, respectively. The fabricated devices show a good approximation to the design object, with angular errors below $\pm 1.3\%$.

5 Conclusions and outlook

In this review, we discussed the start-of-art advances in the field of nanophotonic inverse design. These techniques overcome the limitation of intuition-based designs, boosting the implementation of non-intuitive optical devices, ranging from beam splitters to metalenses, optical cloaks, reflective metasurfaces, and polarization-sensitive displays. These applications have broadened the horizon of how light interacts with complex media, enabling significant leaps towards the miniaturization of optical components and systems for manipulating light.

The first discussed optimization methods represent the oldest and perhaps the most intuitive applications of inverse design. Binary representations and re-parameterization of design structures in topology optimization and nature-based solutions in heuristic methods iterate successive FOM minimization until reaching the optimal solution. Both techniques present different advantages. Topology optimization utilizes straightforward update rules based on fast gradient computation to rapidly converge in local optima of material structures. Heuristics methods, on the contrary, can navigate complex energy landscapes and search for global optima of design structures. The recent exploding interest in artificial intelligence fuels research attention towards inverse designs schemes that leverage data-driven science and engineering. Rather than considering only a particular FOM, different flavors of deep learning methods model complex physical relationships between material properties and measured responses, allowing higher-level FOMs to be optimized, simultaneously. Discriminative learning models serve as universal predictors in the design space, where integrated structures hold the ability to approximate arbitrary optical responses. After being trained on sufficiently large datasets, these methods allow efficient forward simulation and fast design parameters convergence in inverse design schemes. The non-uniqueness problem that commonly exists in such single-model approaches can be solved by subdividing the parameter space and training a series of NN [130]. Generative models further expand the previous architectures' generalization ability by exploiting a latent space whose projection contains all dominant design factors discovered by the network. Properly trained generative models can produce distributions of optimal designs, sampling various structures and analyzing them concurrently.

A novel trend that is attracting considerable attention is the fusion of optimization schemes with deep learning methods. The precise information flow in topology optimization improves the training stability of NN [125], while the deep learning predictors accelerate the convergence speed of optimization methods significantly [19]. The use

of dimensionality reduction methods, including PCA [16], autoencoder-like models [119] and Gaussian mixture models (GMM) [131] facilitates the data-augmentation strategy and overcomes the problems in creating huge datasets. Hybrid methods, in general, incorporate superior features from both optimization and deep learning fields, enabling accurate, robust, and fast-converging design within a computationally efficient framework. We here provide a qualitative comparison in Table 1 of all the mentioned design methods. The table discusses five aspects involving (1) whether the model is differentiable; (2) whether the model is transferable; (3) whether the inverse design is conducted on a distribution of candidates or a single device; (4) the cost of constructing (training/initializing) the model; (5) the cost of producing demanded design (forward prediction/numerical optimization).

Nanophotonics inverse design represents a young and exciting research field that proceeds fast, together with the rapid development of resources in artificial intelligence. Future directions in inverse design can occur in implementing new types of large-scale nanophotonics systems in areas where intuition-based design fails to provide efficient solutions. Inverse design techniques can also open the door to implementing structures that direct design typically avoids due to their inherent complexity, either geometrical or material-based. Artificial intelligence can significantly widen this horizon, thanks to the ability to grasp complex input–output relationships from simple data sequences.

A legitimate question is whether the strong permeation of nanophotonics with AI will lessen our understanding of fundamental physical phenomena, as we will rely more and more on approximate models built automatically via machine learning. The evidence gathered in this review supports a positive outcome. In nearly all cases reviewed, the introduction of AI resources accompanies the formulations of novel theories for describing light's properties or the realization of new hierarchically complex systems that previous work did not explore. As it often happens in science, answering this question will provide sufficient research material for scientists in the present and coming generations.

Author contribution: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

Research funding: We acknowledge funding from KAUST (Award REI/1/4811-16-01).

Conflict of interest statement: The authors declare no conflicts of interest regarding this article.

References

- [1] M. Kim, I. Kim, J. Jang, D. Lee, K. T. Nam, and J. Rho, "Active color control in a metasurface by polarization rotation," *Appl. Sci.*, vol. 8, no. 6, p. 982, 2018.
- [2] M. L. Solomon, A. A. E. Saleh, L. V. Poulikakos, J. M. Abendroth, L. F. Tadesse, and J. A. Dionne, "Nanophotonic platforms for chiral sensing and separation," *Acc. Chem. Res.*, vol. 53, no. 3, pp. 588–598, 2020.
- [3] Y. Zhou, H. Zheng, I. I. Kravchenko, and J. Valentine, "Flat optics for image differentiation," *Nat. Photonics*, vol. 14, no. 5, pp. 316–323, 2020.
- [4] S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vucković, and A. W. Rodriguez, "Inverse design in nanophotonics," *Nat. Photonics*, vol. 12, no. 11, pp. 659–670, 2018.
- [5] W. Ma, Z. Liu, Z. A. Kudyshev, A. Boltasseva, W. Cai, and Y. Liu, "Deep learning for the design of photonic structures," *Nat. Photonics*, vol. 15, no. 2, pp. 77–90, 2021.
- [6] T. W. Liao and G. Li, "Metaheuristic-based inverse design of materials – a survey," *J. Materiomics*, vol. 6, no. 2, pp. 414–430, 2020.
- [7] O. Sigmund and K. Maute, "Topology optimization approaches," *Struct. Multidiscip. Optim.*, vol. 48, no. 6, pp. 1031–1055, 2013.
- [8] J. N. Hooker, *Integrated Methods for Optimization*, vol. 170, New York, Springer, 2012.
- [9] A. Vaswani, N. Shazeer, N. Parmar, et al., "Attention is all you need," in *Advances in Neural Information Processing Systems*, 2017, pp. 5998–6008.
- [10] A. W. Senior, R. Evans, J. Jumper, et al., "Improved protein structure prediction using potentials from deep learning," *Nature*, vol. 577, no. 7792, pp. 706–710, 2020.
- [11] M. Mohri, A. Rostamizadeh, and A. Talwalkar, *Foundations of Machine Learning*, Cambridge, Massachusetts, MIT press, 2018.
- [12] S. Sonoda and N. Murata, "Neural network with unbounded activation functions is universal approximator," *Appl. Comput. Harmon. Anal.*, vol. 43, no. 2, pp. 233–268, 2017.
- [13] C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, and C. Liu, "A survey on deep transfer learning," in *International Conference on Artificial Neural Networks*, Springer, 2018, pp. 270–279.
- [14] S. So, T. Badloe, J. Noh, J. Bravo-Abad, and J. Rho, "Deep learning enabled inverse design in nanophotonics," *Nanophotonics*, vol. 9, no. 5, pp. 1041–1057, 2020.
- [15] J. Jiang and J. A. Fan, "Global optimization of dielectric metasurfaces using a physics-driven neural network," *Nano Lett.*, vol. 19, no. 8, pp. 5366–5372, 2019.
- [16] D. Melati, Y. Grinberg, M. K. Dezfouli, et al., "Mapping the global design space of nanophotonic components using machine learning pattern recognition," *Nat. Commun.*, vol. 10, no. 1, pp. 1–9, 2019.
- [17] Z. Liu, D. Zhu, K.-T. Lee, A. S. Kim, L. Raju, and W. Cai, "Compounding meta-atoms into metamolecules with hybrid artificial intelligence techniques," *Adv. Mater.*, vol. 32, no. 6, p. 1904790, 2020.
- [18] F. Getman, M. Makarenko, A. Burguete-Lopez, and A. Fratallocchi, "Broadband vectorial ultrathin optics with

- experimental efficiency up to 99% in the visible region via universal approximators,” *Light Sci. Appl.*, vol. 10, no. 1, pp. 1–14, 2021.
- [19] M. Makarenko, Q. Wang, A. Burguete-Lopez, F. Getman, and A. Fratalocchi, “Robust and scalable flat-optics on flexible substrates via evolutionary neural networks,” *Adv. Intell. Syst.*, vol. 3, no. 11, p. 2100105, 2021.
 - [20] J. S. Jensen and O. Sigmund, “Topology optimization for nano-photonics,” *Laser Photon. Rev.*, vol. 5, no. 2, pp. 308–321, 2011.
 - [21] Z. Beheshti and S. M. H. Shamsuddin, “A review of population-based meta-heuristic algorithms,” *Int. J. Adv. Soft Comput. Appl.*, vol. 5, no. 1, pp. 1–35, 2013.
 - [22] P. I. Borel, A. Harpøth, L. H. Frandsen, et al., “Topology optimization and fabrication of photonic crystal structures,” *Opt. Express*, vol. 12, no. 9, pp. 1996–2001, 2004.
 - [23] J. S. Jensen and O. Sigmund, “Topology optimization of photonic crystal structures: a high-bandwidth low-loss t-junction waveguide,” *J. Opt. Soc. Am. B*, vol. 22, no. 6, pp. 1191–1198, 2005.
 - [24] M. Gerken and D. A. B. Miller, “Multilayer thin-film structures with high spatial dispersion,” *Appl. Opt.*, vol. 42, no. 7, pp. 1330–1345, 2003.
 - [25] Y. Tsuji and K. Hirayama, “Design of optical circuit devices using topology optimization method with function-expansion-based refractive index distribution,” *IEEE Photon. Technol. Lett.*, vol. 20, no. 12, pp. 982–984, 2008.
 - [26] H. Men, K. Y. K. Lee, R. M. Freund, J. Peraire, and S. G. Johnson, “Robust topology optimization of three-dimensional photonic-crystal band-gap structures,” *Opt. Express*, vol. 22, no. 19, pp. 22632–22648, 2014.
 - [27] A. Y. Piggott, J. Lu, K. G. Lagoudakis, J. Petykiewicz, T. M. Babinec, and J. Vučković, “Inverse design and demonstration of a compact and broadband on-chip wavelength demultiplexer,” *Nat. Photonics*, vol. 9, no. 6, pp. 374–377, 2015.
 - [28] L. F. Frellsen, Y. Ding, O. Sigmund, and L. H. Frandsen, “Topology optimized mode multiplexing in silicon-on-insulator photonic wire waveguides,” *Opt. Express*, vol. 24, no. 15, pp. 16866–16873, 2016.
 - [29] D. Sell, J. Yang, S. Doshay, and J. A. Fan, “Periodic dielectric metasurfaces with high-efficiency, multiwavelength functionalities,” *Adv. Opt. Mater.*, vol. 5, no. 23, p. 1700645, 2017.
 - [30] Y. Chen, F. Meng, G. Li, and X. Huang, “Topology optimization of photonic crystals with exotic properties resulting from Dirac-like cones,” *Acta Mater.*, vol. 164, pp. 377–389, 2019.
 - [31] T. Phan, D. Sell, E. W. Wang, et al., “High-efficiency, large-area, topology-optimized metasurfaces,” *Light Sci. Appl.*, vol. 8, no. 1, pp. 1–9, 2019.
 - [32] J. Rong and W. Ye, “Multifunctional elastic metasurface design with topology optimization,” *Acta Mater.*, vol. 185, pp. 382–399, 2020.
 - [33] A. M. Hammond, A. Oskooi, S. G. Johnson, and S. E. Ralph, “Photonic topology optimization with semiconductor-foundry design-rule constraints,” *Opt. Express*, vol. 29, no. 15, pp. 23916–23938, 2021.
 - [34] M. Mansouree, A. McClung, S. Samudrala, and A. Arbabi, “Large-scale parametrized metasurface design using adjoint optimization,” *ACS Photonics*, vol. 8, no. 2, pp. 455–463, 2021.
 - [35] M. Burger, “A framework for the construction of level set methods for shape optimization and reconstruction,” *Interfaces Free Boundaries*, vol. 5, no. 3, pp. 301–329, 2003.
 - [36] M. Yu. Wang, X. Wang, and D. Guo, “A level set method for structural topology optimization,” *Comput. Methods Appl. Mech. Eng.*, vol. 192, nos. 1–2, pp. 227–246, 2003.
 - [37] M. Mansouree and A. Arbabi, “Metasurface design using level-set and gradient descent optimization techniques,” in *2019 International Applied Computational Electromagnetics Society Symposium (ACES)*, IEEE, 2019, pp. 1–2.
 - [38] N. Lebbe, C. Dapogny, E. Oudet, K. Hassan, and A. Gliere, “Robust shape and topology optimization of nanophotonic devices using the level set method,” *J. Comput. Phys.*, vol. 395, pp. 710–746, 2019.
 - [39] C. D. Freeman and J. Bruna, *Topology and geometry of half-rectified network optimization*, 2016, arXiv preprint arXiv:1611.01540.
 - [40] J. A. Fan, “Freeform metasurface design based on topology optimization,” *MRS Bull.*, vol. 45, no. 3, pp. 196–201, 2020.
 - [41] F. A. M. Gomes and T. A. Senne, “An slp algorithm and its application to topology optimization,” *Comput. Appl. Math.*, vol. 30, pp. 53–89, 2011.
 - [42] N. Aage and B. S. Lazarov, “Parallel framework for topology optimization using the method of moving asymptotes,” *Struct. Multidiscip. Optim.*, vol. 47, no. 4, pp. 493–505, 2013.
 - [43] N. Aage, E. Andreassen, and B. S. Lazarov, “Topology optimization using petsc: an easy-to-use, fully parallel, open source topology optimization framework,” *Struct. Multidiscip. Optim.*, vol. 51, no. 3, pp. 565–572, 2015.
 - [44] C. M. Lalau-Keraly, S. Bhargava, O. D. Miller, and E. Yablonovitch, “Adjoint shape optimization applied to electromagnetic design,” *Opt. Express*, vol. 21, no. 18, pp. 21693–21701, 2013.
 - [45] M. B. Giles and N. A. Pierce, “An introduction to the adjoint approach to design,” *Flow, Turbul. Combust.*, vol. 65, no. 3, pp. 393–415, 2000.
 - [46] T. W. Hughes, M. Minkov, I. A. D. Williamson, and S. Fan, “Adjoint method and inverse design for nonlinear nanophotonic devices,” *ACS Photonics*, vol. 5, no. 12, pp. 4781–4787, 2018.
 - [47] M. Zhou, B. S. Lazarov, F. Wang, and O. Sigmund, “Minimum length scale in topology optimization by geometric constraints,” *Comput. Methods Appl. Mech. Eng.*, vol. 293, pp. 266–282, 2015.
 - [48] M. Liehr, M. Baier, G. Hoefler, et al., “Foundry capabilities for photonic integrated circuits,” in *Optical Fiber Telecommunications VII*, Elsevier, 2020, pp. 143–193.
 - [49] F. Wang, J. S. Jensen, and O. Sigmund, “Robust topology optimization of photonic crystal waveguides with tailored dispersion properties,” *J. Opt. Soc. Am. B*, vol. 28, no. 3, pp. 387–397, 2011.
 - [50] Y. Augenstein and C. Rockstuhl, “Inverse design of nanophotonic devices with structural integrity,” *ACS Photonics*, vol. 7, no. 8, pp. 2190–2196, 2020.

- [51] L. Guo, S. Xu, R. Wan, et al., “Design of aluminum nitride metalens in the ultraviolet spectrum,” *J. Nanophotonics*, vol. 12, no. 4, p. 043513, 2018.
- [52] S. Wang, P. C. Wu, V.-C. Su, et al., “A broadband achromatic metalens in the visible,” *Nat. Nanotechnol.*, vol. 13, no. 3, pp. 227–232, 2018.
- [53] P. Lalanne and P. Chavel, “Metalenses at visible wavelengths: past, present, perspectives,” *Laser Photon. Rev.*, vol. 11, no. 3, p. 1600295, 2017.
- [54] S. Zhang, A. Soibel, S. A. Keo, et al., “Solid-immersion metalenses for infrared focal plane arrays,” *Appl. Phys. Lett.*, vol. 113, no. 11, p. 111104, 2018.
- [55] H. Zuo, D.-Y. Choi, X. Gai, et al., “High-efficiency all-dielectric metalenses for mid-infrared imaging,” *Adv. Opt. Mater.*, vol. 5, no. 23, p. 1700585, 2017.
- [56] H. Yasuda and S. Nishiwaki, “A design method of broadband metalens using time-domain topology optimization,” *AIP Adv.*, vol. 11, no. 5, p. 055116, 2021.
- [57] M. Khorasaninejad and F. Capasso, “Metalenses: versatile multifunctional photonic components,” *Science*, vol. 358, pp. 6367, 2017.
- [58] P. Lalanne and G. M. Morris, “Highly improved convergence of the coupled-wave method for tm polarization,” *J. Opt. Soc. Am. A*, vol. 13, no. 4, pp. 779–784, 1996.
- [59] L. Li, “Use of Fourier series in the analysis of discontinuous periodic structures,” *J. Opt. Soc. Am. A*, vol. 13, no. 9, pp. 1870–1876, 1996.
- [60] R. Sivapuram, R. Picelli, and Y. M. Xie, “Topology optimization of binary microstructures involving various non-volume constraints,” *Comput. Mater. Sci.*, vol. 154, pp. 405–425, 2018.
- [61] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, “Optimization by simulated annealing,” *Science*, vol. 220, no. 4598, pp. 671–680, 1983.
- [62] E. Aarts and J. Korst, *Simulated Annealing and Boltzmann Machines: A Stochastic Approach to Combinatorial Optimization and Neural Computing*, Hoboken, New Jersey, John Wiley & Sons, Inc., 1989.
- [63] Yi. Zhao, X. Cao, J. Gao, et al., “Broadband diffusion metasurface based on a single anisotropic element and optimized by the simulated annealing algorithm,” *Sci. Rep.*, vol. 6, no. 1, pp. 1–9, 2016.
- [64] L. Li, T. J. Cui, W. Ji, et al., “Electromagnetic reprogrammable coding-metasurface holograms,” *Nat. Commun.*, vol. 8, no. 1, pp. 1–7, 2017.
- [65] E. F. Knott, J. F. Schaeffer, and M. T. Tulley, *Radar Cross Section*, Raleigh, North Carolina, SciTech Publishing, 2004.
- [66] Y. Xie, M. Liu, T. Feng, and Yi. Xu, “Compact disordered magnetic resonators designed by simulated annealing algorithm,” *Nanophotonics*, vol. 9, no. 11, pp. 3629–3636, 2020.
- [67] J. Kennedy and R. Eberhart, “Particle swarm optimization,” in *Proceedings of ICNN’95-International Conference on Neural Networks*, vol. 4, IEEE, 1995, pp. 1942–1948.
- [68] W. Chen, B. Zhang, P. Wang, et al., “Ultra-compact and low-loss silicon polarization beam splitter using a particle-swarm-optimized counter-tapered coupler,” *Opt. Express*, vol. 28, no. 21, pp. 30701–30709, 2020.
- [69] J. C. C. Mak, C. Sideris, J. Jeong, A. Hajimiri, and J. K. S. Poon, “Binary particle swarm optimized 2×2 power splitters in a standard foundry silicon photonic platform,” *Opt. Lett.*, vol. 41, no. 16, pp. 3868–3871, 2016.
- [70] M. Hussein, K. R. Mahmoud, M. F. O. Hameed, and S. S. A. Obayya, “Optimal design of vertical silicon nanowires solar cell using hybrid optimization algorithm,” *J. Photon. Energy*, vol. 8, no. 2, p. 022502, 2017.
- [71] Y. Wang, Q. Chen, W. Yang, et al., “High-efficiency broadband achromatic metalens for near-ir biological imaging window,” *Nat. Commun.*, vol. 12, no. 1, pp. 1–7, 2021.
- [72] X. He, T. Dong, J. He, and Y. Xu, “Design of an optical phased array with low side-lobe level and wide-angle steering range based on particle swarm optimization,” in *Asia Communications and Photonics Conference M4A–138*, Optical Society of America, 2020.
- [73] R. Sarker, J. Kamruzzaman, and C. Newton, “Evolutionary optimization (evopt): a brief review and analysis,” *Int. J. Comput. Intell. Appl.*, vol. 3, no. 04, pp. 311–330, 2003.
- [74] B. J. Offrein, G.-L. Bona, R. Germann, I. Massarek, D. Erni, and M. M. Spuhler, “A very short planar silica spot-size converter using a nonperiodic segmented waveguide,” *J. Lightwave Technol.*, vol. 16, no. 9, p. 1680, 1998.
- [75] A. Zunger, “Inverse design in search of materials with target functionalities,” *Nat. Rev. Chem.*, vol. 2, no. 4, pp. 1–16, 2018.
- [76] P. R. Wiecha, A. Arbouet, C. Girard, A. Lecestre, G. Larrieu, and V. Paillard, “Evolutionary multi-objective optimization of colour pixels based on dielectric nanoantennas,” *Nat. Nanotechnol.*, vol. 12, no. 2, p. 163, 2017.
- [77] D. A. Van Veldhuizen and G. B. Lamont, “Evolutionary computation and convergence to a pareto front,” in *Late Breaking Papers at the Genetic Programming 1998 Conference*, Citeseer, 1998, pp. 221–228.
- [78] P. R. Wiecha, C. Majorel, C. Girard, et al., “Design of plasmonic directional antennas via evolutionary optimization,” *Opt. Express*, vol. 27, no. 20, pp. 29069–29081, 2019.
- [79] T. Kosako, Y. Kadoya, and H. F. Hofmann, “Directional control of light by a nano-optical Yagi–Uda antenna,” *Nat. Photonics*, vol. 4, no. 5, pp. 312–315, 2010.
- [80] R. Sathya and A. Abraham, “Comparison of supervised and unsupervised learning algorithms for pattern classification,” *Int. J. Adv. Res. Artif. Intell.*, vol. 2, no. 2, pp. 34–38, 2013.
- [81] Y. Bengio, A. C. Courville, and P. Vincent, “*Unsupervised feature learning and deep learning: a review and new perspectives*,” CoRR, abs/1206.5538, 1:2012, 2012.
- [82] I. Sajedian, T. Badloe, and J. Rho, “Optimisation of colour generation from dielectric nanostructures using reinforcement learning,” *Opt. Express*, vol. 27, no. 4, pp. 5874–5883, 2019.
- [83] R. S. Hegde, “Deep learning: a new tool for photonic nanostructure design,” *Nanoscale Adv.*, vol. 2, no. 3, pp. 1007–1023, 2020.
- [84] T. Jebara, *Machine Learning: Discriminative and Generative*, vol. vol. 755, Berlin/Heidelberg, Germany, Springer Science & Business Media, 2012.
- [85] C. C. Nadell, B. Huang, J. M. Malof, and W. J. Padilla, “Deep learning for accelerated all-dielectric metasurface design,” *Opt. Express*, vol. 27, no. 20, pp. 27523–27535, 2019.

- [86] M. H. Tahersima, K. Kojima, T. Koike-Akino, et al., “Deep neural network inverse design of integrated photonic power splitters,” *Sci. Rep.*, vol. 9, no. 1, pp. 1–9, 2019.
- [87] T. Zhang, J. Wang, Qi. Liu, et al., “Efficient spectrum prediction and inverse design for plasmonic waveguide systems based on artificial neural networks,” *Photon. Res.*, vol. 7, no. 3, pp. 368–380, 2019.
- [88] I. Sajedian, J. Kim, and J. Rho, “Finding the optical properties of plasmonic structures by image processing using a combination of convolutional neural networks and recurrent neural networks,” *Microsyst. Nanoeng.*, vol. 5, no. 1, pp. 1–8, 2019.
- [89] Y. Qu, Li. Jing, Y. Shen, M. Qiu, and M. Soljacic, “Migrating knowledge between physical scenarios based on artificial neural networks,” *ACS Photonics*, vol. 6, no. 5, pp. 1168–1174, 2019.
- [90] J. Lim and D. Psaltis, *Maxwellnet: physics-driven deep neural network training based on Maxwell’s equations*, 2021, arXiv preprint arXiv:2107.06164.
- [91] Z. Liu, D. Zhu, L. Raju, and W. Cai, “Tackling photonic inverse design with machine learning,” *Adv. Sci.*, vol. 8, no. 5, p. 2002923, 2021.
- [92] I. Malkiel, M. Mrejen, A. Nagler, U. Arieli, L. Wolf, and H. Suchowski, “Plasmonic nanostructure design and characterization via deep learning,” *Light Sci. Appl.*, vol. 7, no. 1, pp. 1–8, 2018.
- [93] R. Guo, Z. Lin, T. Shan, et al., “Physics embedded deep neural network for solving full-wave inverse scattering problems,” *IEEE Trans. Antenn. Propag.*, 2021. <https://doi.org/10.1109/tap.2021.3102135>.
- [94] K. Kawaguchi, L. P. Kaelbling, and Y. Bengio, *Generalization in deep learning*, 2017, arXiv preprint arXiv:1710.05468.
- [95] M. V. Zhelyeznyakov, S. Brunton, and A. Majumdar, “Deep learning to accelerate scatterer-to-field mapping for inverse design of dielectric metasurfaces,” *ACS Photonics*, vol. 8, no. 2, pp. 481–488, 2021.
- [96] M. Qiu, “Transfer learning for nanophotonics,” in *2019 IEEE Photonics Society Summer Topical Meeting Series (SUM)*, IEEE, 2019, pp. 1–3.
- [97] K. Yao, R. Unni, and Y. Zheng, “Intelligent nanophotonics: merging photonics and artificial intelligence at the nanoscale,” *Nanophotonics*, vol. 8, no. 3, pp. 339–366, 2019.
- [98] D. M. Sullivan, *Electromagnetic Simulation Using the FDTD Method*, Hoboken, New Jersey, John Wiley & Sons, 2013.
- [99] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.
- [100] I. C. Duta, Li. Liu, F. Zhu, and L. Shao, “Improved residual networks for image and video recognition,” in *2020 25th International Conference on Pattern Recognition (ICPR)*, IEEE, 2021, pp. 9415–9422.
- [101] In. J. Myung, “The importance of complexity in model selection,” *J. Math. Psychol.*, vol. 44, no. 1, pp. 190–204, 2000.
- [102] I. Goodfellow, J. Pouget-Abadie, M. Mirza, et al., “Generative adversarial networks,” ArXiv, abs/1406.2661, 2014.
- [103] Z. Yi, H. Zhang, P. Tan, and M. Gong, “Dualgan: unsupervised dual learning for image-to-image translation,” in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 2849–2857.
- [104] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, “Unpaired image-to-image translation using cycle-consistent adversarial networks,” in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 2223–2232.
- [105] Y. Wu, F. Yang, Y. Xu, and H. Ling, “Privacy-protective-gan for privacy preserving face de-identification,” *J. Comput. Sci. Technol.*, vol. 34, no. 1, pp. 47–60, 2019.
- [106] J. Li, W. Monroe, T. Shi, S. Jean, A. Ritter, and D. Jurafsky, *Adversarial learning for neural dialogue generation*, 2017, arXiv preprint arXiv:1701.06547.
- [107] M. Mirza and S. Osindero, *Conditional generative adversarial nets*, 2014, arXiv preprint arXiv:1411.1784.
- [108] Z. Liu, D. Zhu, S. P. Rodrigues, K.-T. Lee, and W. Cai, “Generative model for the inverse design of metasurfaces,” *Nano Lett.*, vol. 18, no. 10, pp. 6570–6576, 2018.
- [109] J. F. Nash, “Equilibrium points in n-person games,” *Proc. Natl. Acad. Sci. USA*, vol. 36, no. 1, pp. 48–49, 1950.
- [110] S. So and J. Rho, “Designing nanophotonic structures using conditional deep convolutional generative adversarial networks,” *Nanophotonics*, vol. 8, no. 7, pp. 1255–1261, 2019.
- [111] P. Bojanowski, A. Joulin, D. Lopez-Paz, and A. Szlam, *Optimizing the latent space of generative networks*, 2017, arXiv preprint arXiv:1707.05776.
- [112] A.-P. Blanchard-Dionne and O. J. F. Martin, “Successive training of a generative adversarial network for the design of an optical cloak,” *Osa Continuum*, vol. 4, no. 1, pp. 87–95, 2021.
- [113] W. Ma, F. Cheng, Y. Xu, Q. Wen, and Y. Liu, “Probabilistic representation and inverse design of metamaterials based on a deep generative model with semi-supervised learning strategy,” *Adv. Mater.*, vol. 31, no. 35, p. 1901111, 2019.
- [114] D. P. Kingma and M. Welling, *Auto-encoding variational Bayes*, 2013, arXiv preprint arXiv:1312.6114.
- [115] K. Sohn, H. Lee, and X. Yan, “Learning structured output representation using deep conditional generative models,” *Adv. Neural Inf. Process. Syst.*, vol. 28, pp. 3483–3491, 2015.
- [116] Y. Tang, K. Kojima, T. Koike-Akino, et al., “Generative deep learning model for inverse design of integrated nanophotonic devices,” *Laser Photon. Rev.*, vol. 14, no. 12, p. 2000287, 2020.
- [117] Ye. Wang, T. Koike-Akino, and D. Erdogmus, *Invariant representations from adversarially censored autoencoders*, 2018, arXiv preprint arXiv:1805.08097.
- [118] Y. Kiarashinejad, S. Abdollahramezani, M. Zandehshahvar, O. Hemmatyar, and A. Adibi, “Deep learning reveals underlying physics of light–matter interactions in nanophotonic devices,” *Adv. Theory Simulat.*, vol. 2, no. 9, p. 1900088, 2019.
- [119] L. Pilozi, F. A. Farrelly, G. Marcucci, and C. Conti, “Topological nanophotonics and artificial neural networks,” *Nanotechnology*, vol. 32, no. 14, p. 142001, 2021.
- [120] L. Yu, W. Zhang, J. Wang, and Y. Yu, “Seqgan: sequence generative adversarial nets with policy gradient,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 31, 2017.

- [121] W. Xian, P. Sangkloy, V. Agrawal, et al., “Texturegan: controlling deep image synthesis with texture patches,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 8456–8465.
- [122] S. Ruder, *An overview of gradient descent optimization algorithms*, 2016, arXiv preprint arXiv:1609.04747.
- [123] S. Gu and L. Rigazio, *Towards deep neural network architectures robust to adversarial examples*, 2014, arXiv preprint arXiv:1412.5068.
- [124] J. Jiang and J. A. Fan, “Simulator-based training of generative neural networks for the inverse design of metasurfaces,” *Nanophotonics*, vol. 9, no. 5, pp. 1059–1069, 2020.
- [125] J. Jiang and J. A. Fan, “Multiobjective and categorical global optimization of photonic structures based on resnet generative neural networks,” *Nanophotonics*, vol. 10, no. 1, pp. 361–369, 2021.
- [126] B. Schölkopf, A. Smola, and K.-R. Müller, “Nonlinear component analysis as a kernel eigenvalue problem,” *Neural Comput.*, vol. 10, no. 5, pp. 1299–1319, 1998.
- [127] H. Gao, L. Sun, and J.-X. Wang, “Phygeonet: physics-informed geometry-adaptive convolutional neural networks for solving parameterized steady-state pdes on irregular domain,” *J. Comput. Phys.*, vol. 428, p. 110079, 2021.
- [128] M. Makarenko, A. Burguete-Lopez, F. Getman, and A. Fratalocchi, “Generalized Maxwell projections for multi-mode network photonics,” *Sci. Rep.*, vol. 10, no. 1, pp. 1–17, 2020.
- [129] Z. Yang, Ke. Tang, and X. Yao, “Large scale evolutionary optimization using cooperative coevolution,” *Inf. Sci.*, vol. 178, no. 15, pp. 2985–2999, 2008.
- [130] T. Repän, R. Venkitakrishnan, and C. Rockstuhl, “Artificial neural networks used to retrieve effective properties of metamaterials,” *Opt. Express*, vol. 29, no. 22, pp. 36072–36085, 2021.
- [131] M. Zandehshahvar, Y. Kiarashi, M. Zhu, H. Maleki, T. Brown, and A. Adibi, “Manifold learning for reducing the design complexity of photonic nanostructures,” in *CLEO: QELS_Fundamental Science*, pages JT3A–115, Optical Society of America, 2021.
- [132] M. D. Huntington, L. J. Lauhon, and T. W. Odom, “Subwavelength lattice optics by evolutionary design,” *Nano Lett.*, vol. 14, no. 12, pp. 7195–7200, 2014.
- [133] E. Johlin, S. A. Mann, S. Kature, A. F. Koenderink, and E. C. Garnett, “Broadband highly directive 3d nanophotonic lenses,” *Nat. Commun.*, vol. 9, no. 1, pp. 1–8, 2018.
- [134] W. Ma, F. Cheng, and Y. Liu, “Deep-learning-enabled on-demand design of chiral metamaterials,” *ACS Nano*, vol. 12, no. 6, pp. 6326–6334, 2018.
- [135] J. Peurifoy, Y. Shen, Li. Jing, et al., “Nanophotonic particle simulation and inverse design using artificial neural networks,” *Sci. Adv.*, vol. 4, no. 6, p. eaar4206, 2018.