nature photonics

FOCUS | REVIEW ARTICLE

https://doi.org/10.1038/s41566-020-0685-y

Deep learning for the design of photonic structures

Wei Ma[1,2], Zhaocheng Liu[3,4], Zhaxylyk A. Kudyshev[5,6,7,8], Alexandra Boltasseva[5,6,7] ✉, Wenshan Cai[3,4] ✉ and Yongmin Liu[1,2] ✉

Innovative approaches and tools play an important role in shaping design, characterization and optimization for the field of photonics. As a subset of machine learning that learns multilevel abstraction of data using hierarchically structured layers, deep learning offers an efficient means to design photonic structures, spawning data-driven approaches complementary to conventional physics- and rule-based methods. Here, we review recent progress in deep-learning-based photonic design by providing the historical background, algorithm fundamentals and key applications, with the emphasis on various model architectures for specific photonic tasks. We also comment on the challenges and perspectives of this emerging research direction.

New photonic structures, materials, devices and systems have been driving forces for transformative technologies, including high-speed optical communication and computing, ultrasensitive biochemical detection, efficient solar energy harvesting and super-resolution imaging, as well as quantum information processing. Over the past three decades, we have witnessed tremendous progress in and success of artificially engineered photonic structures, including photonic crystals, metamaterials and plasmonic nanostructures, with unparalleled capabilities in tailoring light–matter interactions and unlocking new device concepts. Many fundamental laws have been revisited or generalized in these structured media, and consequently they promise a wide range of important applications. For instance, photonic crystals can realize complete photonic bandgaps, so that light can transmit around a sharp bend surrounded by such crystals with near-perfect efficiency[1]. Metamaterials demonstrate exceptional properties through rational structural designs, exemplified by negative refractive indices: these enable light to be refracted to the negative direction, in contrast to normal refraction based on Snell's law[2–4]. By using metallic nanoparticles with different sizes, geometries and compositions, plasmonics can break the classical diffraction limit, offering the opportunity to control light emission at the single-molecule level[5].

Whether we are discussing individual plasmonic nanostructures, or metamaterials and photonic crystals composed of arrays of dielectric or metallic building blocks, structural designs play a central role. So far, there are two main design approaches. First, we can resort to physics-based methods, such as simplified analytical models, knowledge obtained from prior or related practice, and scientific intuition. For example, dielectric and metallic nanoparticles with simple geometries (such as spheres, cylinders and core–shell particles) can be accurately modelled by Mie theory[6]. The scattering, absorption and extinction responses of the particles arise from their electric and magnetic multipolar resonances. The initial idea of split-ring resonators, widely used in the metamaterials community to produce effective magnetism, was based on electromagnetics and electrical circuit theory[7]. Specifically, varying external magnetic fields induce a current loop and thus a magnetic dipole, which is greatly enhanced around the resonance frequency determined by the internal capacitance and inductance of the resonator. For photonic crystals, the intuition originated from the successful understanding of electron transport in the periodic potential well of a solid material[8]. In an analogue, light transmission through periodically modulated refractive indices can be greatly modified, enabling 'photonic semiconductors' with complete bandgaps that disallow the propagation of light in the 'forbidden band'. Although these physics-based approaches offer important guidelines, it is not trivial to find the right structures to realize the desired photonic properties, especially when the geometry and spatial arrangements of the structure become complicated.

Therefore, we have to rely on the second approach: electromagnetic modelling based on numerical simulation methods such as the finite-difference time-domain method, the finite-element method, the finite integration technique or the method of moments, with or without optimization algorithms. Generally, starting from certain initial and boundary conditions, these computational electromagnetics simulations solve the design problem by discretizing Maxwell's equations spatially and temporally. By setting up sufficient meshes and iteration steps, we can accurately calculate the optical properties of a given structure. Nevertheless, we often need to fine-tune the geometry and iteratively perform simulations to gradually approach the targeted responses. This procedure largely relies on past experience of the design templates, and owing to constraints on simulation power and time, only limited design parameters are adjusted in searching for the optimal structure.

The inverse design problem, meaning the direct retrieval of the proper structure for the desired optical performance, requires exploration of a much larger degree of freedom in the design space, and hence is even more challenging. To search the formidably large design space efficiently, the inverse design procedure is usually guided by optimization algorithms: either gradient-based approaches (for example, topology optimization, adjoint method or level-set method) or evolutionary approaches (such as genetic algorithms or particle swarm algorithms). Such inverse design algorithms enable one to find non-intuitive, irregularly shaped photonic structures that outperform empirically designed structures in many applications, such as silicon photonic components, photonic crystals and metamaterials. Fundamentally, these algorithms are rule-based approaches containing iterative searching steps in a case-by-case manner, often relying on numerical simulations in each step to produce intermediate results that help to modify the searching strategy. Such stochastic algorithms are limited by their random-search nature and hence are insufficient for complex design

[1]Department of Mechanical and Industrial Engineering, Northeastern University, Boston, MA, USA. [2]Department of Electrical and Computer Engineering, Northeastern University, Boston, MA, USA. [3]School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA. [4]School of Materials Science and Engineering, Georgia Institute of Technology, Atlanta, GA, USA. [5]School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, USA. [6]Birck Nanotechnology Center, Purdue University, West Lafayette, IN, USA. [7]Purdue Quantum Science and Engineering Institute, Purdue University, West Lafayette, IN, USA. [8]Center for Science of Information, Purdue University, West Lafayette, IN, USA. ✉e-mail: aeb@purdue.edu; wcai@gatech.edu; y.liu@northeastern.edu

in a multi-constrained problem. Readers interested in inverse design in photonics can refer to recent reviews on this topic[9–11].

Deep learning allows a computational model composed of multiple layers of processing units to learn multiple levels of abstraction in given data[12]. In light of its exceptional success in domains related to computer science and engineering, including computer vision[13], natural language processing[14], speech recognition[15], knowledge graphs[16] and decision making[17], deep learning has attracted increasing attention from researchers in other disciplines, including materials science[18], chemistry[19], laser physics[20], particle physics[21], quantum mechanics[22], computational imaging[23] and microscopy[24], demonstrating potential to circumvent the drawbacks of traditional methods and create unprecedented opportunities in these areas. The unique advantages of deep learning lie in its data-driven methodology, which allows the model to discover useful information automatically from a huge amount of data, in sharp contrast to physics- or rule-based approaches. Over the past few years, deep learning has become a radically new approach in the context of photonic design.

In this Review, we draw attention to a collection of recent results that showcase the power of deep learning in the design of photonic structures, materials and devices, where empirical or traditional approaches are infeasible or inefficient. We begin by providing background on the deep-learning model for photonics and highlighting the formulation, development and advantages of deep neural networks. Then we discuss several major model architectures, from the basic multilayer perceptron (MLP) and advanced deep neural networks to hybrid models with other optimization methods, emphasizing their potential to design photonic crystals, metamaterials, plasmonic nanostructures and integrated silicon photonic devices (Fig. 1). The models can map design parameters (such as geometry, material, topology and spatial arrangement) and optical characteristics (such as polarization, phase, wavelength and orbital angular momentum), enabling both forward prediction and inverse design. Finally, we comment on the challenges and perspectives of this emerging interdisciplinary research area, with its potential to create a new science and engineering paradigm in which photonics and artificial intelligence (AI) are interfaced with each other.

Background

Historically, deep learning can be traced back to the 1940s, and it went through many different names before becoming popularly known by this term[25]. Originally, some of the learning algorithms were intended to computationally model the process of biological learning—that is, to model how learning happens in human brains. Therefore, deep learning was known by the term artificial neural networks (ANNs) since the 1980s, accompanied by the second wave of AI research, which largely emerged through a movement called connectionism[26]. During this period, the milestone was the famous paper written by David Rumelhart, Geoffrey Hinton and Ronald Williams in 1986[27], in which the authors modified and reiterated the importance of the original back-propagation algorithm, making it much faster than earlier approaches to learning and allowing it to solve previously insoluble problems. The modern era of neural network research began with a breakthrough in 2006: Hinton, who later coined the term deep learning, showed that deep neural networks could be efficiently trained using a strategy called greedy layer-wise pretraining[28]. Deep learning then gained broad popularity, as researchers were able to train deeper neural networks than had been possible before, and the importance of depth of the model architecture was theoretically realized. Aided by the ever-increasing scale of available data, deep learning still leads the current boom in AI research, with the performance surpassing previous models by a large margin, or even beating humans[29].

In the context of photonics research, the introduction of deep learning (or more appropriately ANNs, at that time) dated back to the 1990s during the second prevalence period of AI research. This interdisciplinary field primarily featured extensive work from the microwave community, where ANNs were used as a computer-aided design tool for fast prototyping of microwave devices and radio-frequency circuits[30]. The success of ANN-aided design was attributed to the revival of the MLP, the simplest form of a feedforward neural network that comprises several fully connected layers. Each neuron in a layer is connected to all neurons in the next layer with its own unique weight. Therefore, this is also called a fully connected neural network or a dense neural network. At this early stage, the application of ANNs was straightforward. The design problem was often transformed to training an ANN that linked an input port with an output port. According to specific tasks, different variables, such as circuit or device parameters (for example, geometry, physical property, bias or frequency) or performance parameters (for example, S-parameters, voltage, current or power) were carefully chosen to feed to each port. Based on this scheme, many microwave design problems were readily solved by ANNs, including transmission lines[31], vias[32], filters[33], amplifiers[34] and antennas[35].

The use of ANNs in microwave and other photonic designs follows a standard supervised learning paradigm, which aims to find a mapping between input variable X and output variable Y. The training process is to optimize the model parameters to make it an accurate representation of the desired mapping, on a previously collected training dataset containing examples of X and its corresponding label Y. Starting with random initialization, this data-driven learning approach modifies the model parameters iteratively on the training dataset according to a specific loss function evaluated on true labels Y, until a convergence occurs and the generalized model is able to predict unseen data. It is very different from conventional optimization approaches, either gradient-based or gradient-free, in which no such pre-collected input–output pairs are available, and the optimization of given targets is guided by certain rules in a case-by-case manner. To efficiently train an ANN model, the back-propagation algorithm was developed. Detailed explanations of the basic ANN structure and the back-propagation algorithm are presented in Box 1 and Box 2, respectively.

The interplay between microwave research and ANNs, as the first attempt to combine electromagnetics with AI, was not a coincidence. On the one hand, most designs of microwave devices can be decomposed into optimizing some parameters for a given target, both of which can be conveniently represented by a few variables. With well-understood physics, mature simulation tools and thus little difficulty in data acquisition, finding the relations between these variables is less complicated for an MLP model. On the other hand, more complex photonic designs such as plasmonic nanostructures, metamaterials, photonic crystals or silicon photonic devices were still in their infancy, with limited understanding in the community. These structures and devices, featuring deep sub-wavelength dimensions, large design flexibility, extreme dispersions or complicated performance characteristics, were not tractable to the shallow ANN model at that time. Related research then stagnated until the recent epoch-making development of deep learning.

The past decade has witnessed the rise of deep learning with unprecedented impact on a plethora of research topics. With the invention of new training and regularization techniques such as 'ReLU' (rectified linear units) activation[36], dropout[37] and batch normalization[38], it is now feasible to design and train deeper and deeper neural networks that can exploit larger datasets with better performance. Meanwhile, advanced model architectures have been proposed or improved to solve tasks in specific fields of machine learning and pattern recognition, such as the convolutional neural network (CNN)[13] for image recognition, recurrent neural network (RNN)[39] for natural language processing, generative adversarial network (GAN)[40] and variational autoencoder (VAE)[41] for image generation. The photonics community also benefited from the rapid

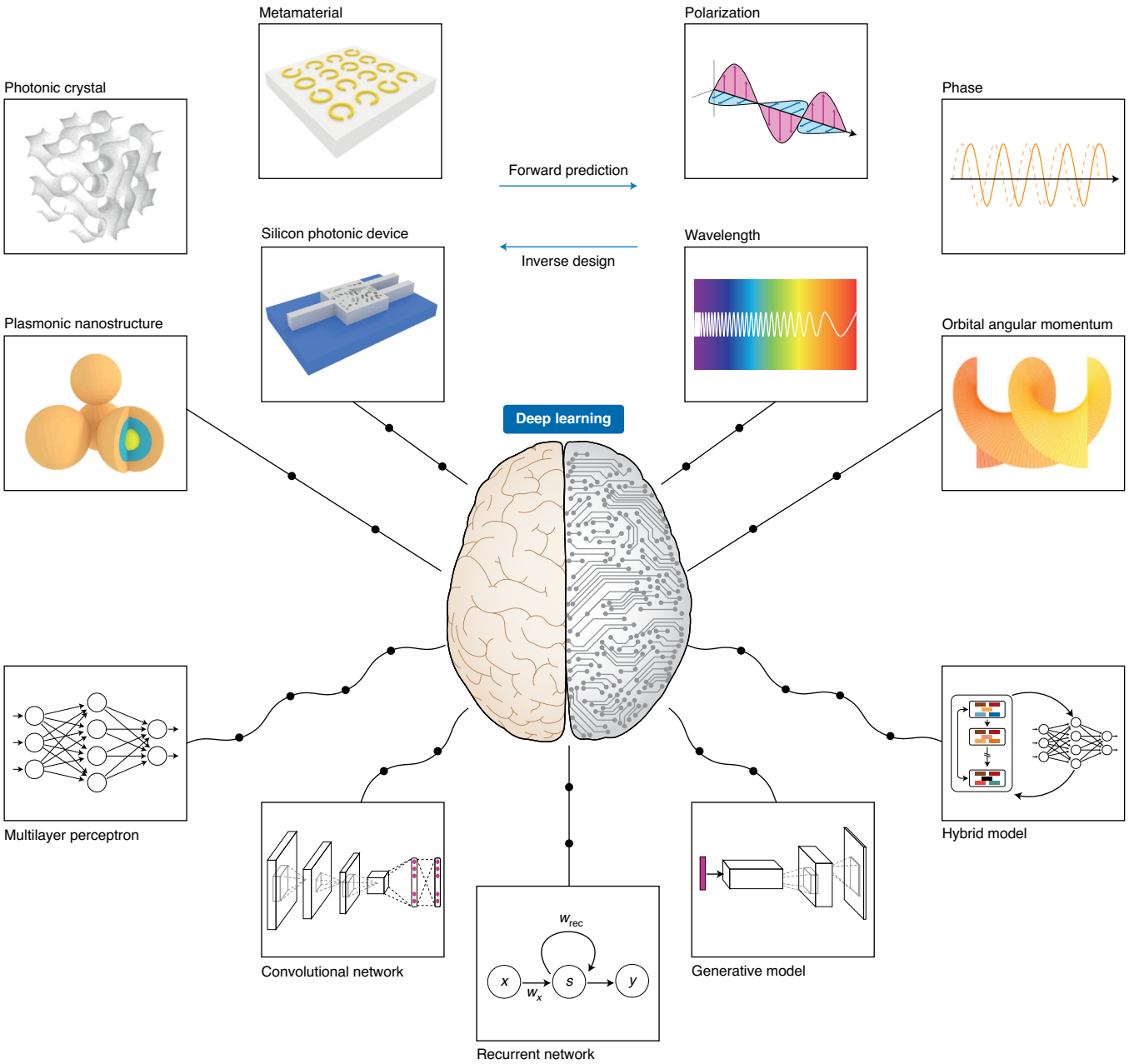


Fig. 1 | Applying deep learning to solve photonic design problems. Linked by the hub of data-driven methodology, deep learning associates various model architectures (for example, multilayer perceptron, convolutional network, recurrent network, generative model or hybrid model) with specific photonic design tasks (for example, photonic crystal, metamaterial, plasmonic nanostructure or silicon photonic device). The modelling process takes into consideration both optical characteristics (such as polarization, phase, wavelength and orbital angular momentum) and design parameters (such as geometry, material, topology and spatial arrangement). *x*, input variable; *s*, recurrent neuron; *y*, output variable; w_{rec}, weights for feedback connection; w_x, weights for input variables.

advances in deep-learning techniques. Unlike rule-guided optimization that explores the design space by following certain strategies case by case, deep learning, as a data-driven method, aims to describe the design space holistically, using the training data as samples. Therefore, with generalization ability within a given design space, deep learning can produce fast and accurate designs without the need for case-by-case, time-consuming numerical calculations. Well-trained deep-learning models can directly set up a mapping from design to optical properties of target photonic devices, and vice versa. In addition, deep learning can interact with traditional optimization methods to improve the algorithm performance.

Specific and representative examples

In this section, we will discuss some relevant model architectures and their applications in solving photonics problems.

Multilayer perceptron. Although the architecture is rather simple, the MLP model, as illustrated in Fig. 1, has been theoretically proven as a universal approximator that is capable of fitting any continuous functions with a finite number of neurons[42]. In modern deep-learning models with sophisticated and task-specific architectures, MLP often serves as a bottleneck layer to extract meaningful features as a compact representation of high-dimensional data such

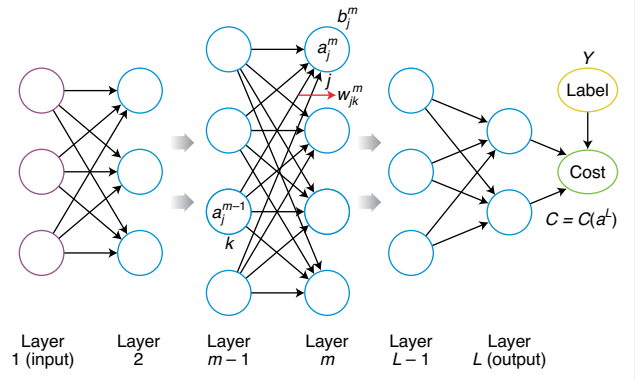
Box 1 | Fully connected neural network

A fully connected neural network with L layers consists of one input layer, one output layer and $L-2$ hidden layers. Each neuron in a given layer is connected to every neuron in the next layer, but without interlayer connections. The neurons process the input data layer by layer with learnable parameters, namely weights and biases. Specifically, let us denote the weight from the kth neuron in the $(m-1)$th layer to the jth neuron in the mth layer as w_{jk}^m, and the bias of the jth neuron in the mth layer as b_j^m. Then, the output of the jth neuron in the mth layer, denoted as a_j^m, can be computed as a weighted summation of the outputs from the previous layer z_j^m followed by a nonlinear activation function $\sigma(\cdot)$, that is

$$a_j^m = \sigma\left(\sum\nolimits_k w_{jk}^m a_k^{m-1} + b_j^m\right) = \sigma\left(w^m a^{m-1} + b^m\right) = \sigma\left(z_j^m\right) \quad (1)$$

In a matrix and vector notation, the learnable model parameters in layer m include a weight matrix $w^m = \left(w_{jk}^m\right)|_{1\le j\le |j|, 1\le k\le |k|}$ and a bias vector $b^m = \left(b_j^m\right)|_{1\le j\le |j|}$. The nonlinear function $\sigma(\cdot)$, also known as an activation function, is crucial as it enables the neural network to tackle highly complex data representation instead of degrading to a simple linear mapping. Popular choices of activation function include the sigmoid, hyperbolic tangents or rectified linear units. Consequently, in the forward pass, the input data flow from one layer to the next, being linearly summed and nonlinearly activated at each stage, similar to the electrical signal flowing in biological neurons and synapses (shown as black arrows).

To estimate the performance of the neural network, a cost function $C(\cdot)$ should be defined to measure the discrepancy between the network output and the desired output. The cost function takes the outputs of the last layer and the ground-truth label as input and returns a scalar value as the error to be minimized. To adjust the model parameters properly, the learning algorithm computes the gradient of the cost function with respect to each weight and bias, indicating how much the error would increase or decrease if the parameter is increased by a tiny amount. Then each parameter is adjusted in the opposite direction to its gradient.



as images. Even before the astounding success of deep learning in computer vision[13], researchers explored MLP in photonics-related domains following the design schemes devised by the microwave community. The potency of these models was largely limited by immature training strategy, lack of data and thus shallow model architecture.

In 2018, Itzik Malkiel and co-workers reported a bidirectional MLP-based deep-learning model that could be used to design plasmonic nanostructures[43]. In an advance on earlier work, the researchers managed to model the intricate physical relationship between photonic structures and their optical characteristics by resorting to a 'deeper' network configuration. As shown in Fig. 2a, the structure of interest is an H-shaped metallic structure represented by eight parameters: three continuous parameters (the length and rotation angle of arms) and five binary parameters (existence of certain arms). The design targets are two reflection spectra under the illumination of light polarized along the horizontal or vertical direction. Each spectrum is discretized into 43 data points, and the material properties (such as permittivity of the indium tin oxide adhesion layer and hosting materials) are represented as a vector of 25 parameters. A geometry-predicting network of eight group layers and a spectrum-predicting network of six layers are trained on 18,000 samples so that the model can simultaneously function as a fast simulator and an inverse design tool. The scale of the dataset, input complexity and the depth of the model show that, from a data-driven view, deeper models armed with larger datasets are a potential alternative approach to solve photonic design tasks.

Meanwhile, to deal with specific photonic problems, adaptation of the MLP-based model burgeoned, with improvements in the model architecture, training strategy and application. Wei Ma and co-workers reported a deep-learning model with two bidirectional neural networks that were integrated by an ensembled learning strategy to achieve on-demand design of chiral metamaterials, which exhibit distinct responses when incident light is left-circularly polarized (LCP) or right-circularly polarized (RCP)[44]. The chiral metamaterial comprises two twisted split-ring resonators placed on a metallic back-reflector (left panel in Fig. 2b). Two bidirectional neural networks, termed the primary network and auxiliary network respectively, are constructed to model the interconnection of three physical quantities, namely reflection spectra, circular dichroism spectra and design parameters. By introducing the auxiliary network, the accuracy of the forward prediction of spectra around resonances (middle panel in Fig. 2b), as well as the inverse retrieval of the design parameters, is substantially improved. The model functionality is also extended, enabling retrieval of possible metamaterial design from simple requirements on several parameters of desired chiroptical response, or direct prediction of the circular dichroism spectrum from design parameters (right panel in Fig. 2b).

When trained on large datasets, MLP-based models for inverse design often fail to converge, since there exist multiple candidates satisfying similar requirements but with very different designs. To solve this, Dianjing Liu and colleagues proposed a tandem training method for inverse design[45]. The idea is to first train a forward modelling network, mapping the design to optical responses. This pretrained forward network is then connected to the output of the inverse design network, with the forward prediction error serving as the supervision signal. By indirectly training in this tandem configuration, the inverse retrieval outputs are forced to converge to only one possibility guided by the forward model, efficiently solving the data inconsistency issue that arises from the fundamental property of non-uniqueness in the inverse design problem.

Instead of focusing solely on the inputs and outputs, MLP-based models, as universal approximators of functions, also allow gradients to be calculated analytically. John Peurifoy and colleagues used a single neural network to approximate light scattering by multilayer nanoparticles, where the analytical gradient obtained from the model ewas used for structural optimization given specific requirements in the spectra[46]. They showed that, guided by the analytical gradient from an MLP model, the single-band high scattering effect in core–shell nanoparticles can be efficiently optimized (left panel in Fig. 2c). They also made a quantitative comparison between a

Box 2 | Back-propagation algorithm

A back-propagation algorithm is used to train a neural network by only one backward pass from the output layer to the input layer, which is a practical application of the chain rule for derivatives of a multivariate function. To better illustrate the procedure of back-propagation, we introduce an intermediate error vector, $\delta^l = \left(\partial C/\partial z_1^l, \partial C/\partial z_2^l, \partial C/\partial z_3^l \ldots\right)$, which represents the partial derivative of cost C with respect to the weighted input in layer l. Then naturally at the last layer (layer L), we have

$$\delta^L = (\nabla_a C) \odot \sigma'\left(z^L\right) \tag{2}$$

where the Hadamard product $\odot$ denotes an elementwise product of two vectors. Note that both the cost function $C(a^L)$ and activation function $\sigma(z^L)$ have an analytical form, so the error value δ^L at the last layer can be directly obtained. Similarly, by applying the chain rule of partial derivatives, the error vector of layer l (other than the last layer), δ^l, can be calculated from the errors of the next layer δ^{l+1} by

$$\delta^l = \left[\left(w^{l+1}\right)^T \delta^{l+1}\right] \odot \sigma'\left(z^l\right) \tag{3}$$

Equations (2) and (3) represent the central idea of the back-propagation procedure, where the cost, initially generated from the discrepancy between network output and target output, backflows from the last layer to the first layer, with the intermediate error of each layer retained. Then the quantities of real interest, $\partial C/\partial w_{jk}^l$ and $\partial C/\partial b_j^l$, are simply related to those errors as follows

$$\frac{\partial C}{\partial w_{jk}^l} = a_k^{l-1}\delta_j^l \tag{4}$$

$$\frac{\partial C}{\partial b_j^l} = \delta_j^l \tag{5}$$

With the partial derivatives of the cost function C with respect to all of the learnable parameters in the neural network — that is, the weight matrices w^l and bias vectors b^l for all layers ($l = 1, 2, \ldots, L$) — the model can then be trained by stochastic gradient descent (SGD). In the SGD algorithm, the training data are shuffled and then divided into several batches, each containing only a small portion of all data. The training step consists of feeding the network with a batch of data, computing the outputs and the errors, back-propagating to compute the average gradient for this batch, and adjusting the weights and biases accordingly.

Concretely, we need to define a hyper-parameter called the learning rate η, to control how much the parameters are modified as a portion of their gradients with respect to the cost in this batch. For a batch containing M training data, according to equations (4) and (5), we have the following update rules for weights and biases:

$$w^l \to w^l - \frac{\eta}{M}\sum\nolimits_x \delta^{x,l}\left(a^{x,l-1}\right)^T \tag{6}$$

$$b^l \to b^l - \frac{\eta}{M}\sum\nolimits_x \delta^{x,l} \tag{7}$$

The training step is repeated many times until the value of the cost stops decreasing. After training, the performance of the model is evaluated on a different test set, containing samples not seen by the model during training, to test its generalization ability.

deep neural network and interior-point methods, a numerical inverse design algorithm suitable for nanoparticles. Under the same error threshold during inverse design, the runtime of the neural network is two orders of magnitude shorter than numerical methods. Even more notably, as illustrated in the right panel of Fig. 2c, with the increase of design complexity (that is, number of nanoparticle layers), the neural network shows a linear increase in runtime, in contrast to a polynomial increase in runtime (with an exponent of 4.5) for numerical methods. These results indicate that within the same accuracy standard, neural networks, once fully trained, beat conventional numerical inverse design methods by a large margin in speed.

Following these first forays into applying data-driven methodology in photonics research, deep MLP models and their variants have been extensively studied for many photonic design tasks, including topological photonics[47], integrated silicon photonics[48], colour generation from nanostructures[49,50], metamaterials[51–53], plasmonic structures[54] and photonic crystals[55]. For instance, Fig. 2d illustrates topological photonic structures whose edge-state dispersion is well reconstructed by an MLP model[47]. In work shown in Fig. 2e, the researchers propose to use deep learning to design silicon-on-insulator-based 1×2 power splitters with various target transmission ratios at two ports (T_1 and T_2), in which the device structure is represented by a matrix of binary variables, indicating whether the silicon is etched or not at certain coordinates[48]. In Fig. 2f, the MLP approach is applied in colour-generation metasurfaces, and colour pixels composed of HfO_2 nanopillars are optimized by the model to form a wide colour gamut[49].

In addition to design problems, researchers have also developed sophisticated MLP models to tackle other problems in photonics, such as modal classification and effective refractive index retrieval in waveguides[56,57], dimension reduction to reveal underlying physics[58] and near-field manipulation of plasmonic nanoantennas[59]. As the plain form of the deep-learning model, MLP requires vectorized input and output. By parameterizing model inputs and outputs into vectors composed of several discrete and dimensionless elements, the MLP model can be adapted to a broad range of photonic applications regardless of the underlying physics.

Advanced deep-learning techniques. Although MLP models offer an effective approach for many photonic tasks, the simple connections in the model still pose some difficulties when the intrinsic structure, target response and design space are either multimodal or hard to parameterize. Therefore, application-specific model architectures have been invented for photonic research or transplanted from other fields, including CNNs specialized to input with local semantic correlation, RNNs for time-dependent inputs and deep generative models to deal with structured-output conditions. Such advanced models can extend the functionality and improve performance in many photonic design problems, which are discussed below.

Convolutional neural networks. In deep learning, a CNN is a class of deep networks that extracts the features of the inputs by using convolution operations on the output of each layer. Owing to the nature of the convolution operation, a CNN can capture the local correlation of spatial information in images. In photonics research, a CNN is an ideal candidate to process data represented in high-dimensional spaces, such as photonic patterns represented as images, and spectral responses of given photonic devices. As a counterpart of CNN, transposed CNN (TCNN) uses the convolution in 'reversed' order, enabling the generation of high-dimensional data from low-dimensional vectors. TCNNs are commonly implemented as a part of generative models for the discovery and design of photonic structures with large degrees of freedom. The basic architectures of CNN- and TCNN-based generative models that transform

data between high-dimensional images and low-dimensional feature vectors are illustrated in the lower panels of Fig. 1 (second and fourth insets, respectively).

CNNs have been used in various optics and photonics problems, such as the inverse scattering problem[60], wavefront correction[61], digital coding metasurfaces[62,63], and prediction of optical properties in complex photonic and materials systems[64,65]. The convolution operation in a CNN is translationally symmetric, making it suitable to model periodic photonic structures such as photonic crystals or metamaterials. These structures, usually represented by two-dimensional (2D) images, inherently satisfy translation invariance when the periodic boundary condition is applied in the numerical simulations. Recently, Takashi Asano and co-workers used a neural network consisting of CNNs to approximate the *Q*-factor of photonic crystals[55]. Back-propagation was used to optimize the positions of nanocavities to greatly improve the *Q*-factor from 3.8×10^8 to 1.6×10^9, as shown in Fig. 3a. Besides the basic architectures of CNNs, advanced network structures including residual networks[66] and inception networks[67] have been introduced to enhance the performance and capability of the network. These network structures can be used to train surrogate models for the simulation of complex photonic structures with large degrees of freedom[68].

Recurrent neural networks. RNNs tackle problems associated with sequential data such as sentences and audio signals. As shown in Fig. 1 (lower panels, third inset), the network receives sequential data one at a time and incrementally generates new data series. For photonic design, RNNs are suitable to model optical signals or spectra in the time domain with a specific line shape originated from various modes of the resonance. RNNs have been implemented to analyse optical signals and equalize noises in high-speed fibre transmission[69]. In combination with CNNs, RNNs were also used to improve the approximation of the optical responses of nanostructures represented in images[68]. Figure 3b (left panel) presents random silver nanostructures to be simulated. With the help of an RNN, the network is able to predict the absorption of the structure from 800 nm to 1,700 nm, showing excellent agreement with full wave simulation (Fig. 3b, right panel). The performance of RNNs can be enhanced by adopting advanced varieties of RNNs, such as long short-term memory[39] and gated recurrent units[14]. Network systems hybridizing CNNs and RNNs are promising techniques to model and design photonic devices that manipulate unconventional spatiotemporal properties of light.

Deep generative models. Instead of determining the conditional distribution and thus decision boundaries, generative models describe the joint distribution of the input and output to optimize a certain objective in a probabilistically generative manner. Unlike their deterministic counterparts, such models can handle one-to-many mapping and produce structured outputs given prescribed requirements, as illustrated in Fig. 1 (lower panels, fourth inset). Empowered by deep-learning algorithms, deep generative models can produce data that replicate in the same way as or are similar to the training dataset. An autoencoder (AE) consists of an encoder and a decoder. The encoder maps the training data into a reduced-dimensional space, that is, latent space, while the decoder reconstructs the variable in the latent space into the training data. AEs are used to reduce the dimensionality of the design space and optical response features of photonic devices, providing essential information on the light–matter interaction for device optimization[58].

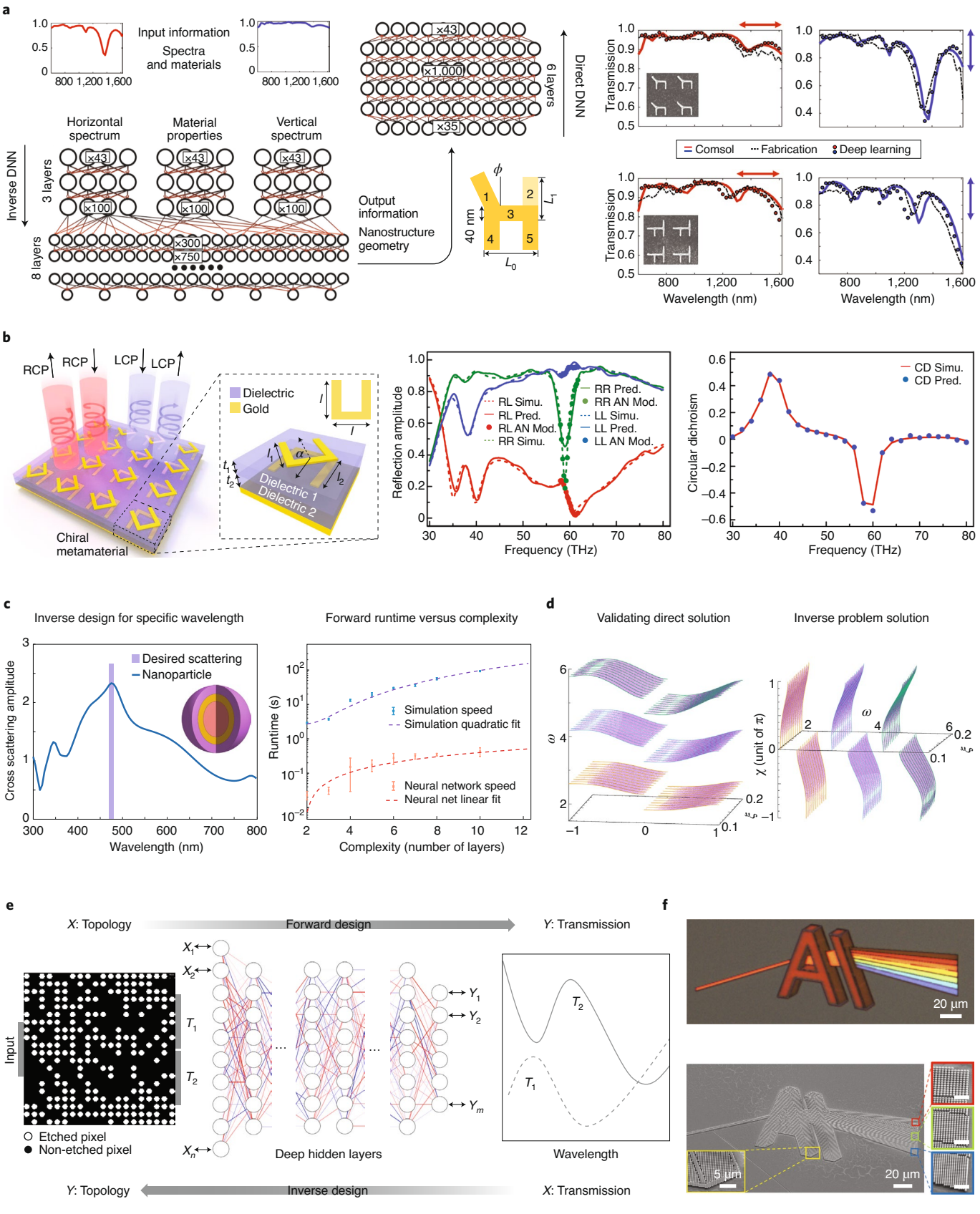
Similar to an AE, a variational autoencoder (VAE) is constructed by adding probabilistic perturbation in the latent space of the AE[41]. The structure of a VAE is illustrated in Fig. 3c (top panel). Because the latent space is a continuous representation of the training data, new designs can be constructed by sampling the latent space[70]. Leveraging the reduced dimension and continuity of the latent space, Wei Ma and colleagues used a VAE-based generative model to encode the meta-atoms of double-layered chiral metamaterials and their optical responses, enabling the investigation of the complex relationship between structure and performance without extensive data collection[71]. Figure 3d (left panel) shows the required circular dichroism spectrum with a sharp peak at 60 THz. The VAE-based framework can identify a double-layered metamaterial with a circular dichroism spectrum replicating the desired one (right panel). More importantly, the model can solve the one-to-many mapping inverse problem and come up with distinct metamaterial structures to satisfy the same requirement of chiroptical responses.

The GAN, as shown in Fig. 3c (bottom panel), is another class of deep generative model constructed with a generator and a discriminator[40]. The generator is trained in an adversarial manner to create samples that, ideally, form a distribution indistinguishable from the training dataset. With the ability to generate massive nanostructures within short time, GANs have been used in the design and optimization of dielectric and metallic metasurfaces in a stochastic manner[72]. Recently, Zhaocheng Liu and co-workers have proposed a framework leveraging GANs to inversely design metasurface nanostructures that match on-demand design objectives[73]. A GAN and a pretrained simulator jointly identify the topology of nanostructures from a user-defined geometric dataset. In Fig. 3e, given a user-defined transmittance spectrum for the inverse design problem (left panel), the framework can identify a nanostructure with a spectral behaviour that matches the input one with high fidelity (right panel).

Besides the standard architecture of the GAN and VAE, some other varieties of generative models can also be implemented for the discovery and design of photonic structures. For instance, compositional pattern-producing networks (CPPNs) have been reported as serving as a generator in the GAN framework to produce high-quality nanostructure patterns for the inverse design of

Fig. 2 | Photonic designs enabled by an MLP model. a, A bidirectional MLP model for design of H-shaped plasmonic nanostructures (left) and the predicted (by deep-learning model), simulated (by Comsol software) and measured spectra of retrieved H-shaped structures under different polarization conditions (right). The numbers in boxes denote the number of neurons in each specific layer of the neural network. DNN, deep neural network. **b**, A deep-learning model for chiral metamaterial design. AN Mod., modified by auxiliary network (AN); Pred., predicted; Simu., simulated; RL, left circularly polarized (LCP) input and right circularly polarized (RCP) output; RR, RCP input and RCP output; LL, LCP input and LCP output; CD, circular dichroism. **c**, Deep-learning model as an optimization tool by providing analytical gradients, which can realize single-peak high scattering by a multilayer nanoparticle (left) with a speed two orders of magnitude faster than that of the conventional numerical optimization (right). **d**, Reconstruction of edge-states dispersion by MLP models. Left: direct problem solution. Right: inverse problem solution. **e**, Deep neural network to model silicon-on-insulator-based 1 × 2 power splitters. T_1 and T_2 denote transmission ratio of port 1 and port 2, respectively. **f**, Colour generation from nanostructures designed by deep learning. Top: AI-prism logo captured by an optical microscope. Bottom: scanning electron micrograph of the corresponding colour pixels. Figure reproduced with permission from: **a**, ref. [43], **d**, ref. [47], **e**, ref. [48], under a Creative Commons licence (http://creativecommons.org/licenses/by/4.0/); **b**, ref. [44], American Chemical Society; **c**, ref. [46], © The Authors, some rights reserved; exclusive licensee AAAS. Distributed under a Creative Commons licence (http://creativecommons.org/licenses/by-nc/4.0/). Reprinted with permission of AAAS; **f**, ref. [49], RSC.

metasurfaces with multiple meta-atoms in a unit cell[74]. Figure 3f (left panel) presents the designed gradient metasurface, generated from a CPPN-GAN, that partially converts LCP incident light to its cross polarization with a constant phase gradient. Figure 3f (right panel) presents the polarization states of incident and diffracted light measured after a quarter waveplate, confirming the switch



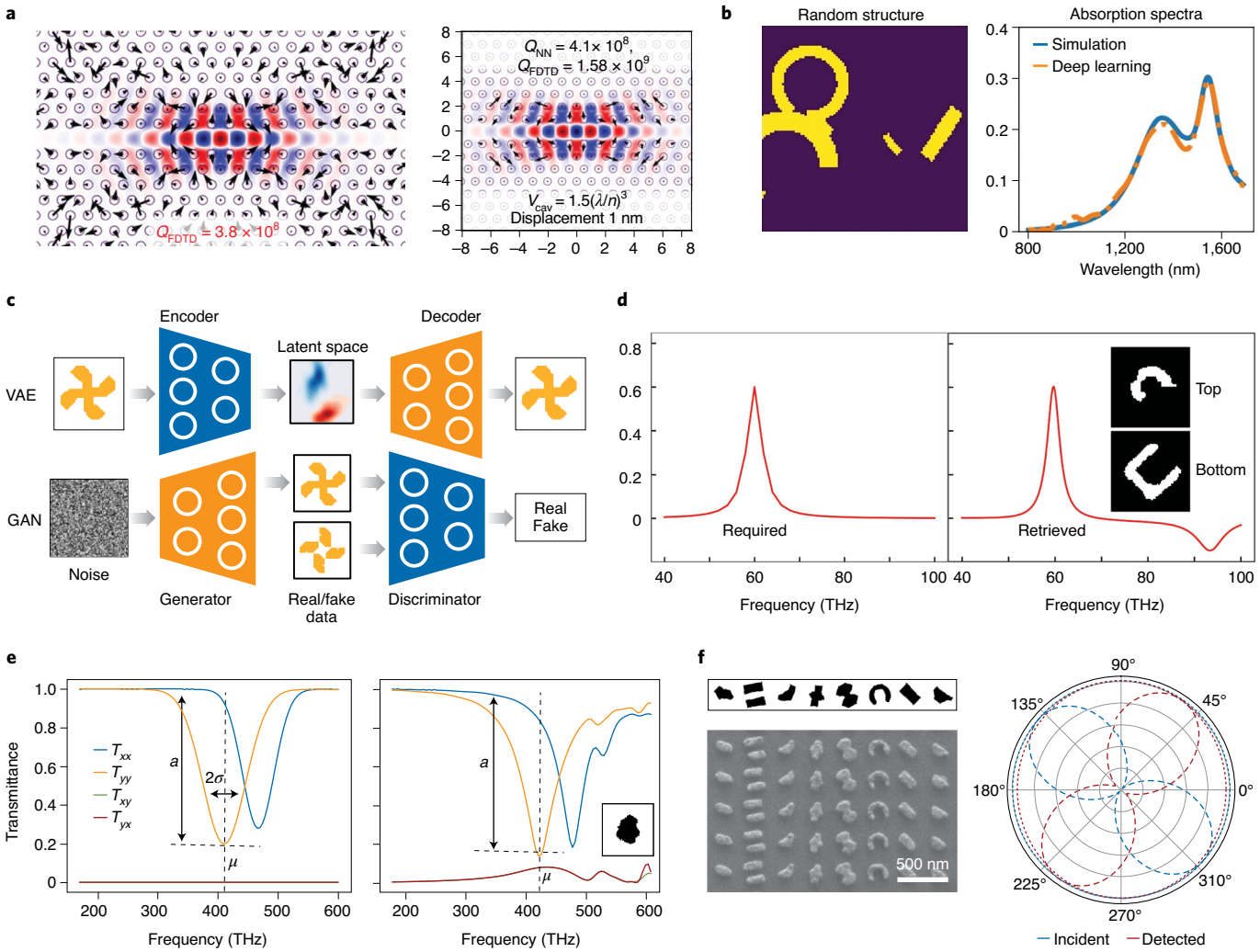


Fig. 3 | Advanced deep-learning frameworks in optics and photonics. a, Optimization of *Q*-factors of photonic crystals with CNNs. Left: initial photonic crystals with a *Q*-factor of 3.8×10^8. Right: optimized photonic crystal with an improved *Q*-factor of 1.58×10^9. Q_{FDTD}, quality factor calculated by finite-difference time-domain (FDTD) simulation; Q_{NN}, quality factor predicted by neural network; V_{cav}, modal volume. **b**, Prediction of absorption of nanostructures assisted by RNNs. Left: randomly generated silver nanostructures (yellow region). Right: absorption approximated by a consolidation of CNN and RNN (orange), and the actual absorption of the nanostructures (blue). **c**, Illustrative representation of the architectures of an VAE and GAN shown in the top and bottom panels, respectively. **d**, Design of a chiral metamaterial by an VAE and a semi-supervised learning strategy. Left: required circular dichroism spectrum. Right: retrieved patterns in the top and bottom layer of the chiral meta-atom (insets), as well as the corresponding spectrum. **e**, Identification of metasurface nanostructures by combining a GAN and a pretrained simulator network. Left: custom-defined Gaussian-like transmission spectra. Right: identified unit cell of the metasurface (lower right corner) and the simulated spectra. **f**, GAN-assisted design of spatially varied metamolecules. Left: designed and fabricated meta-atoms that are able to convert left-circularly polarized light to its cross polarization. Right: measured polarization states. Figure reproduced with permission from: **a**, ref. [55], The Optical Society; **b**, ref. [68], under a Creative Commons licence (http://creativecommons.org/licenses/by/4.0/); **d**, ref. [71], Wiley; **e**, ref. [73], American Chemical Society. Figure adapted with permission from: **f**, ref. [74], Wiley.

of polarization. Other examples of generative models applied in photonics research have also been reported. For example, U-net, a deep neural network architecture originally designed for image segmentation, has been used to accurately predict the near-field optical responses of arbitrary 3D nanostructures, enabling the approximation of various far-field responses of nanostructures[75]. By combining traditional optimization techniques and a generative model, global optimization of a 1D meta-grating can be achieved with high efficiency for various wavelengths and deflection angles[76].

Optimization assisted by deep-learning algorithms. Recent advances in material optimization, along with the progress of nanofabrication techniques allow global challenges in quantum information[77], energy[78], space exploration[79] and secure communication[80] to be addressed with nanophotonic devices. However, such inherently complex problems require highly constrained, multiparametric optimization of the device. Conventionally, adjoint, topology and genetic optimization methods have long been used to tackle a wide range of inverse problems in photonics. These methods usually require a lot of computation power and time, which scale up even further with the increasing dimension of the optimization parametric space and the number of constraints. Such resource heaviness substantially limits the applicability of the conventional techniques to the aforementioned problems, which demand optimization over extended parametric space including both geometrical optimization sub-space and material parametric domain. By hybridization of conventional optimization methods with

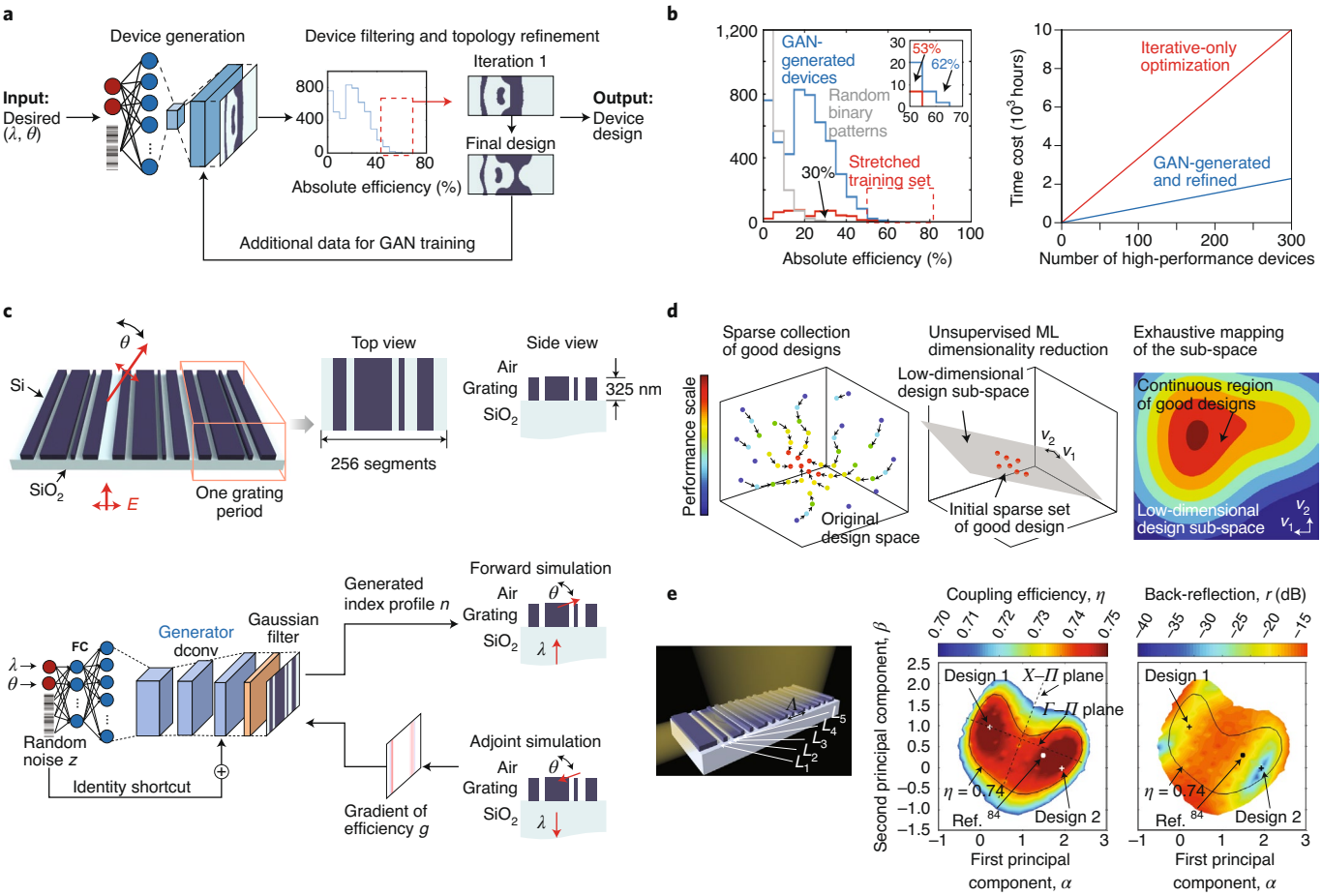


Fig. 4 | Deep-learning-assisted optimization methods. a, GAN-based development of metasurface design. Conditional GAN, which consists of generator and discriminator, trains on topologically optimized designs. **b**, Left: efficiency histograms of meta-gratings produced from the trained GAN generator, geometrically stretched training set, and randomly generated binary patterns. The inset shows the magnified view of the histogram outlined by the dashed red box. Right: time cost of generating devices using iterative-only optimization (red line) and GAN generation and refinement (blue line). **c**, Global optimization driven by a generative neural network. Top: schematic of a silicon meta-grating design that needs to be optimized. Bottom: optimization is done by coupling a generative network with direct electromagnetic solver. FC, fully connected layers; dconv, deconvolution layers. **d**, High-parametric space analysis based on dimensionality reduction. **e**, Left: schematic representation of the grating coupler structure. Right: reducing the design parameters from five-dimensional parametric space to the two principal component coefficients α and β makes the exhaustive mapping of the sub-space of good designs achievable with modest computation resources. Two designs with comparable coupling efficiency are marked along with the design reported in ref. [84]. Γ-Π and X-Π are two orthogonal hyperplanes that are both orthogonal to the first two principal components α and β. Figure reproduced with permission from: **a**,**b**, ref. [81], **c**, ref. [76], American Chemical Society; **d**,**e**, ref. [83], under a Creative Commons licence (http://creativecommons.org/licenses/by/4.0/).

advanced deep-learning algorithms, it is possible to exploit the best aspects of both approaches and greatly reduce the required computational resources.

Coupling deep generative networks (such as GANs and different types of AE) with an adjoint optimization framework has been explored for advancing both speed and performance. This concept, as schematically shown in Fig. 1 (lower panels, fifth inset), exploits the strong correlation between the device topology (geometrical shape) and its optical response. During the training phase, a generative network learns characteristic geometrical features of the pre-optimized device designs and uses this knowledge to construct a compressed representation of the design space. Recently, it has been demonstrated that conditional GANs can be efficiently trained on labelled topology optimized meta-grating designs for the rapid generation of a large family of highly efficient device designs[81] (Fig. 4a). Particularly, it has been shown that the GAN is able to learn and generalize the main features of the meta-gratings from the training set. This allows generation of devices with operating parameters (wavelength and deflection angle) beyond those used during the training. As shown in Fig. 4b, while GAN-generated designs show similar performance to those obtained from the direct topology optimization, GAN-assisted optimization generates five times as many high-performance devices in the same time.

Moreover, it has been demonstrated that generative networks can be directly hybridized with the topology optimization method by substituting the discriminator with a direct electromagnetic solver[76]. By taking topology optimization of a dielectric meta-grating as a showcase example (Fig. 4c), Jiaqi Jiang and Jonathan Fan have demonstrated that this approach allows them to continuously optimize the device distribution until it converges to a cluster of high-efficiency devices[76]. Physics-based gradients, calculated based on forward and adjoint electromagnetic simulations, are used for error back-propagation during the training phase of the generator, ensuring the direct connection of the compressed space formation of the generative network with enhancing device efficiency. Moreover, such an approach avoids the process of

generating the training set, since the generative network learns the physical relationship between device geometry and optical response directly through electromagnetic simulations. The loss function is constructed such that it gives higher weight to devices with better efficiency and decreases the impact of low-efficiency designs that are potentially locked in local optima. As a result, one can narrow down the compressed space representation of the generative network to the sub-domain of highly efficient designs. The proposed approach only needs 10% of the computational cost of direct adjoint-based topology optimization calculations.

More recently, an adversarial autoencoder (AAE), consisting of three neural networks (encoder, decoder and discriminator), has been coupled with a topology optimization framework for the development of a thermal emitter design[82]. The AAE-based approach ensures almost ideal (98%) thermal emission reshaping efficiency and requires a third of the computation time of direct topology optimization. By interfacing an AAE-based approach with a pre-trained CNN network, one can achieve up to a 4,900-fold speed-up in comparison with direct topology optimization. Along with efficient optimization search, such an approach promises unparalleled control over the data distribution within the compressed design space. The latter fact opens up the possibility of performing global optimization searches directly within the high-dimensional compressed design space, avoiding time-consuming post-processing of the generated designs.

Optimization within high-dimensional parametric space can be efficiently realized by adapting one of the dimension-reduction machine-learning algorithms. Recently, linear principal component analysis has been used to map and characterize a multiparameter design space of a photonic system[83]. As shown in Fig. 4d, the developed approach consists of three main steps: (1) construction of a sparse collection of pre-optimized designs by applying conventional optimization methods; (2) determination of a lower-dimensional sub-space of highly efficient designs by applying a dimensionality reduction algorithm to the high-dimensional parametric space of pre-optimized designs; and (3) exhaustive mapping of the design sub-space to determine the continuous region of highly efficient designs. This approach has been applied for the multiparametric optimization of a vertical fibre grating coupler (Fig. 4e). It has been shown that by applying the dimensionality algorithm, the complexity of the problem can be exponentially scaled down and that the continuous sub-space of grating couplers with comparable fibre coupling efficiencies can be mapped[84]. Such global analysis ensures efficient multiparametric optimization and reveals the limitations of performance or structure of the given design configuration.

Researchers have developed another approach to address one of the major bottlenecks of the conventional optimization techniques, namely intensive direct electromagnetic simulations. This approach uses data-driven methods for accelerating Maxwell's equation solver, which is involved in the optimization process. The feasibility of using data-driven methods for solving partial differential equations has been investigated very recently, with the demonstration of learning an 'optimal' approximation of derivatives needed for time-domain simulations[85] or using neural networks for solving partial differential equations[86,87]. Particularly, it has been demonstrated that data-driven methods allow nonlinear partial differential equations to be solved in the time domain at resolutions 4 to 8 times coarser than required by standard finite-difference methods. By using the CNN, the generalized minimal residual algorithm for the solution of frequency-domain Maxwell's equations has been accelerated[88]. This approach realizes an order-of-magnitude reduction in the number of iterations required for solving frequency-domain wave equations when trained on a dataset of wavelength-splitting gratings. The proposed approach can be directly applied within gradient-descent-based optimization of a photonic device, where many simulations with similar material distribution are required.

Discussions and perspectives

In this section, we will comment on the directions, challenges and perspectives of interfacing deep learning with photonics.

Deep learning for complete control of light. It is believed that deep learning can serve as a powerful tool in finding complex relations between a structure and its optical responses. Nevertheless, at present, most literature related to the topic of the present review is limited to the study of the optical properties of a structure in terms of the optical spectra (reflection, transmission, scattering, absorption and so on) under the illustration of linear or circular polarization. Spectra can be conveniently discretized into vectors[43] (single spectrum) or matrices[44] (multiple spectra), and thus can be readily incorporated into the deep-learning model. To aid in the complete control of light, one immediate task that we need to work on is to enhance the capability of deep-learning models with more degrees of freedom, such as the phase[74], angular momentum, trajectory, nonlinearity, topology and near-field distributions[75]. This will lead to new devices including but not limited to multidimensional meta-holograms showing distinct images at different wavelengths, topological photonic crystals for robust light transport that is immune to defects, and controllable nanoscale hotspots to enhance the emission of single quantum emitters.

Unlike a direct discretization, some pre-processing steps or modification of the model input–output may be required accordingly. For instance, one could consider using a periodic output activation for phase retrieval or using dimension reduction[83] when dealing with near-field characteristics. Transfer-learning techniques[89] may also help to construct versatile models for complete light control from base models trained with spectral responses. Reinforcement learning is another approach worthy of further exploration, to expand the photonic design capability[50]. Unlike supervised learning or unsupervised learning, reinforcement learning manages to maximize a cumulative reward by taking actions in the current environment. This indirect optimization process balances the current knowledge and unknown domains, and could potentially investigate more optical properties of photonic design tasks to realize complete control of light.

The burden of data collection. Since deep neural networks contain thousands to millions of learnable parameters, a gigantic amount of labelled data is inevitable to train the network. However, generating data requires physical simulations or experimental measurements. Collecting a massive dataset is not always practical. In these circumstances, unsupervised and semi-supervised learning strategies can be leveraged to alleviate the burden of data collection. Unlike supervised learning, unsupervised learning algorithms, such as principal component analysis, and AEs/VAEs/GANs, require only a small number of labels, usually serving as tools for data clustering and dimensionality reduction. The processed data from unsupervised learning strategies can be analysed to unveil the structural parameters of the photonic devices that have most impact, assisting the neural network to recognize the essential information of the design without redundant data collection[90,91].

Massive data collection can also be mitigated by merging the deep-learning models with underlying physics. For example, instead of training a model with enormous datasets to approximate the bidirectional relationship between physical structures and their responses, deep learning can be adopted as intermediate steps to efficiently solve partial differential equations that describe the physical systems[87,88,92]. Alternatively, learning the governing laws behind the physical phenomena can also reduce the dependence on data[86,93]. Deep learning in this strategy is usually used to extract features from limited data for the regression of a few parameters that parametrizes the models or equations of the system. Furthermore, it is possible to generate a solution that satisfies certain partial

differential equations by physics-informed deep-learning models[94]. In other circumstances where a neural network is well trained for the approximation of certain physical processes, transfer learning can be applied to migrate the knowledge of the neural network to other similar simulation scenarios with substantially reduced data collection[89].

Comparative advantages of deep learning over conventional optimization methods. From the viewpoint of efficiency, a well-trained deep-learning model is faster by orders of magnitude than traditional optimization algorithms[46], on the premise that enough training data are collected in advance. However, collecting data does not conflict with other traditional optimization strategies[9], especially global optimization algorithms that extensively explore the solution space. The difference between data collection for deep learning and that for numerical optimization algorithms lies in how the models make use of data. For deep learning, the model actively accumulates training data that holistically describe the problem under investigation, so that the model is driven by data in the sense that the quantity and quality of pre-collected training data contribute to its performance. In contrast, traditional numerical algorithms passively produce data in the iterative optimization steps, where data, more like a by-product, are mainly used to check how far the optimization has proceeded and when the strategy needs to be changed. Usually, for a specific task, the quantity of data required by deep learning is much more than produced during a conventional optimization. However, data collection for deep learning is a one-time cost, as opposed to an ongoing cost as in the case of ab initio numerical optimization. This feature makes deep learning especially competitive in photonic design tasks of periodic structures such as metamaterials or photonic crystals, where a huge number of unit cells need to be optimized to constitute a device or a system.

Meanwhile, balancing training time and run time for a deep-learning model requires efficient exploitation of data, probably in an online manner where data are continuously provided as the model evolves according to different tasks. Traditional optimization methods explore the design space guided by gradient information (for example, level-set methods) or evolution process (for example, genetic algorithm). To some extent, deep learning has similarities with gradient-based optimization methods, because the optimization of the deep-learning model is also directed by the average gradient of the loss function with respect to all the training data. On the other hand, similar to gradient-free optimization that introduces more stochasticity to enlarge the possible design space, deep learning uses a large amount of pre-collected training data to extend the scope of design varieties far beyond what is achievable with optimization methods.

In conjunction with traditional numerical optimizations, a fast deep-learning approximator for the following optimization steps can be constructed by saving the simulation data right from the beginning of the optimization[95,96]. In these cases, the deep neural network is also capturing the gradients of the objective functions with respect to the input parameters. If a network is sufficiently reliable, performing local gradient descent optimization with gradients calculated from the surrogate network is equivalent to the traditional adjoint methods. On the other hand, deep learning can help to identify a global minimum in various ways. For example, fast inference speed enables global search algorithms for the design of highly complex structures. Latent variable models and generative models are able to capture the underlying physics so that unnecessary exploration of the solution space can be avoided. In more practical cases, traditional optimization and data-driven methods can be jointly incorporated into the same design pipeline, where deep learning serves to identify solutions near global minimum and traditional optimization refines the performance to the extreme. Even if a true global optimum can hardly be found in a complex photonic design task, these optimization methods, including deep learning, always yield fairly good results beyond empirical designs. In some cases in which only a small amount of data can be collected in a reasonable time, incorporating statistical learning and other machine-learning algorithms to the design strategy is also advantageous. However, because deep-learning and machine-learning models always require sufficient training datasets, data-driven approaches may not be an effective starting point in extreme instances where retrieving a single data point takes from hours to days. Instead, direct optimization such as Bayesian optimization could be considered before deep-learning models are applied.

Enlarging design space for 'global' photonic designs. Addressing modern multidisciplinary fundamental science and engineering challenges, such as those related to energy, security, data processing and storage, as well as emerging quantum technology, requires multifunctional devices with highly constrained electromagnetic, chemical, thermal and mechanical properties. Thus far, optical technologies have mainly used photonic optimization in a restricted design space that is largely limited to structural topology (geometry) and shape. Such optimization often omits in-depth feedback from other critical design layers including the evolution of electromagnetic properties with time or in harsh environments, as well as fabrication and characterization constraints. This represents a fundamental limitation, which prevents current optimization methods from providing efficient, 'globally' optimized solutions. For example, global optimization would be critical for multiparametric and multiscale optimization of structures based on nanophotonic, phase-change, nonlinear and gain material platforms, in which optical response depends strongly on material composition and the fabrication process. Further development of advanced hybrid optimization schemes will be a key step to address such optimization challenges and seek potential answers for the ultimate achievable performance of a photonic device.

By integrating available material knowledge into the training process, latent space could be extended to include a rich constituent material domain into the compressed representation. For instance, owing to temperature-dependent dielectric permittivity of the refractory material platforms[97], it is challenging to design metasurface-based thermal emitters for thermophotovoltaics that operate efficiently in a wide temperature range. If temperature-dependent dielectric permittivity functions of the constituent material platform are incorporated into the training process, a 'globally' optimized design with robust emissivity response over a wide range of operating temperatures may be determined. Global optimization search within the extended compressed space will not only ensure optimization of the design topology but will also provide optimal material properties and composition, or even guide the fabrication process to achieve the best possible performance of the targeted photonic device, architecture and system. This approach could be used to address various nonlinear problems in photonics, such as quantum entangled-state generation using spontaneous parametric down-conversion based on metastructures[98]. To be able to realize efficient entangled-state generation, it is important to optimize not only the optical resonance of the nanostructure but also the nonlinear properties of the constituent material. By integrating the dependencies of the linear and nonlinear optical properties on the deposition parameters of the material into the training process, it would be possible to simultaneously optimize the topology or shape of metastructures and retrieve the optimal deposition parameters of the host materials.

We envisage that the global optimization framework could become an essential part of a more generic, hierarchical machine-learning framework based on a multistep strategy. As an example applicable to photonic systems, the first step could be to

define the main target functionality of the device and determine the proper photonic concept that would give the optimal performance (for example, gratings, multilayers, photonic crystals or metasurfaces). The second step would then be to choose the right material platform. This would require creating a wide database of known optical materials together with their properties and their evolution under changing temperature and other harsh environments. By using the available databases on chosen material platform properties, machine-learning-driven global optimization could then be used as the third step to obtain ultimate-efficiency device designs and to determine suitable fabrication (growth condition, doping level, stoichiometry and so on) and integration schemes.

Neuromorphic photonics. While the rapid advances in deep learning provide radically new approaches to solve photonic design problems, this assistance is not one-way but interactive[99]. 'Big data', the catalyst of deep learning, continues to revolutionize AI research with record-breaking improvements in various domains, while at the same time it causes an exponential increase in computational power and energy consumption. At present, most deep-learning algorithms are deployed in conventional computers with von Neumann architecture, whose serial nature is an intrinsic barrier to efficient support for neural networks. Even with some application-specific integrated circuits that are deliberately optimized to run deep-learning models, most industrial-grade models for practical applications still take a formidable time and cost to train. Superior to their electronic counterpart in both speed and power consumption, photonic platforms for deep learning, including nanophotonic scatterers[100], integrated silicon photonic chips[101–103] and 3D-printed diffractive layers[104], are under active investigation.

The tremendous acceleration of deep-learning models on an optical platform depends on the capability of parallel signal processing of light, which makes matrix–vector multiplications in constant time with respect to the matrix dimension, in contrast to the quadratic time complexity on a digital processor. Besides matrix multiplication, nonlinear activation functions play a key role in ANNs, enabling them to learn complex mappings when cascading multiple linear layers. Such nonlinear activation functions for an optical platform can be realized by using saturable absorption of 2D materials[105], nonlinear electro-optic modulation in silicon[106], or simply an external digital processor[101], with implementation in a fixed or reprogrammable manner[107].

Instead of implementing mathematical operations such as matrix multiplication and nonlinear activation in an ANN, optical components allow a biological neural system to be mimicked in a more analogous way. Optical spiking neural networks naturally emulate the basic integrate-and-fire functionality of a biological neuron, using on-chip optical components such as waveguides, wavelength-division multiplexers and ring resonators[102,108,109]. In particular, Wolfram Pernice's group fabricated an all-optical spiking neuron circuit consisting of 4 neurons, 60 synapses and 140 optical elements in total, and successfully demonstrated its function of letter recognition[104]. Training and learning in the system can be implemented in either a supervised or an unsupervised manner. These neuromorphic photonic platforms can make use of the overwhelming advantages of light in speed and parallelism when processing information. The interactions between new photonic structures and deep learning may overcome the limitation of current computing approaches and systems, and potentially lead AI research to new horizons. Indeed, it has been demonstrated that inverse-designed metastructures can solve integral equations using electromagnetic fields[110]. On the other side, wave physics can also be regarded as an analogue recurrent neural network[111]. The potential of new photonic structures, some of which may be designed by deep learning, to enable unconventional computing and AI techniques is worthy of further exploration.

Summary

The concept of deep learning has gone far beyond being a computational analogy of biological neural systems and has emerged as a powerful tool that solves extremely complex problems by building up multilevel abstraction of massive data. For the photonics community, deep learning and other AI techniques are currently transforming the areas of optical design, integration and measurements. Deep-learning techniques have already demonstrated their tremendous potential for photonic structure design, optimization of architecture, materials and entire optical systems, and will continue to uncover new ways to unparalleled speed-up of optical measurements and even unlocking new optical effects.

In this Review, we have surveyed various model structures for photonic design, ranging from individual plasmonic nanoparticles to metamaterials comprising an array of meta-atoms, and to integrated photonic devices. All of these remarkable developments have been demonstrated within the past few years, and further advances are expected as researchers with different backgrounds contribute to this emerging field. Deep learning and AI researchers should team up with optical scientists to develop unorthodox, physics-driven algorithms and networks that are not only robust, generative and interpretable while using fewer data, but also provide unconventional ways to realize unparalleled optical functionalities. Such cross-disciplinary approaches merging AI, photonics and materials platforms will allow for large-scale photonic designs with unique functionalities as well as new methods for optical characterization, paving the way to high-speed super-resolution imaging, real-time detection and manipulation, efficient energy conversion systems and transformative advances in the area of quantum measurements and metrology. On this path, the photonics community should ultimately build an 'optical structures and materials genome' to construct a comprehensive dataset of photonic concepts, architectures, components and photonic materials to enable hierarchical machine-learning algorithms that could provide ultimate-efficiency devices. This effort should be extended to realize all-optical platforms to perform deep learning and other AI algorithms at the speed of light, ushering in an even brighter AI era.

Received: 8 January 2020; Accepted: 27 July 2020;
Published online: 5 October 2020

References

1. Joannopoulos, J. D., Johnson, S. G., Winn, J. N. & Meade, R. D. *Photonic Crystals: Molding the Flow of Light* (Princeton Univ. Press, 2011).
2. Smith, D., Pendry, J. & Wiltshire, M. Metamaterials and negative refractive index. *Science* **305**, 788–792 (2004).
3. Liu, Y. & Zhang, X. Metamaterials: a new frontier of science and technology. *Chem. Soc. Rev.* **40**, 2494–2507 (2011).
4. Cai, W. & Shalaev, V. M. *Optical Metamaterials: Fundamentals and Applications* (Springer, 2010).
5. Maier, S. A. *Plasmonics: Fundamentals and Applications* (Springer, 2007).
6. Bohren, C. F. & Huffman, D. R. *Absorption and Scattering of Light by Small Particles* (Wiley, 2008).
7. Pendry, J. B., Holden, A., Robbins, D. & Stewart, W. Magnetism from conductors and enhanced nonlinear phenomena. *IEEE Trans. Microw. Theory Tech.* **47**, 2075–2084 (1999).
8. John, S. Strong localization of photons in certain disordered dielectric superlattices. *Phys. Rev. Lett.* **58**, 2486–2489 (1987).
9. Molesky, S. et al. Inverse design in nanophotonics. *Nat. Photon.* **12**, 659–670 (2018).
10. Li, W., Meng, F., Chen, Y., Li, Y. & Huang, X. Topology optimization of photonic and phononic crystals and metamaterials: a review. *Adv. Theory Simul.* **2**, 1900017 (2019).
11. Campbell, S. D. et al. Review of numerical optimization techniques for meta-device design. *Opt. Mater. Express* **9**, 1842–1863 (2019).
12. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
13. Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet classification with deep convolutional neural networks. In *Proc. 25th Int. Conf. Neural Information Processing Systems* 1097–1105 (NIPS, 2012).

14. Cho, K. et al. Learning phrase representations using RNN encoder–decoder for statistical machine translation. In *Proc. 2014 Conf. Empirical Methods in Natural Language Processing (EMNLP)* 1724–1734 (2014).
15. Hinton, G. et al. Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups. *IEEE Signal Proc. Mag.* **29**, 82–97 (2012).
16. Socher, R., Chen, D., Manning, C. D. & Ng, A. Reasoning with neural tensor networks for knowledge base completion. In *NIPS'13: Proc. 26th Int. Conf. Neural Information Processing Systems* 926–934 (NIPS, 2013).
17. Silver, D. et al. Mastering the game of Go with deep neural networks and tree search. *Nature* **529**, 484–489 (2016).
18. Sanchez-Lengeling, B. & Aspuru-Guzik, A. Inverse molecular design using machine learning: generative models for matter engineering. *Science* **361**, 360–365 (2018).
19. Goh, G. B., Hodas, N. O. & Vishnu, A. Deep learning for computational chemistry. *J. Comput. Chem.* **38**, 1291–1307 (2017).
20. Zahavy, T. et al. Deep learning reconstruction of ultrashort pulses. *Optica* **5**, 666–673 (2018).
21. Baldi, P., Sadowski, P. & Whiteson, D. Searching for exotic particles in high-energy physics with deep learning. *Nat. Commun.* **5**, 4308 (2014).
22. Carrasquilla, J. & Melko, R. G. Machine learning phases of matter. *Nat. Phys.* **13**, 431–434 (2017).
23. White, A., Khial, P., Salehi, F., Hassibi, B. & Hajimiri, A. A silicon photonics computational lensless active-flat-optics imaging system. *Sci. Rep.* **10**, 1869 (2020).
24. Rivenson, Y. et al. Deep learning microscopy. *Optica* **4**, 1437–1443 (2017).
25. Goodfellow, I., Bengio, Y. & Courville, A. *Deep Learning* (MIT Press, 2016).
26. McClelland, J. L., McNaughton, B. L. & O'Reilly, R. C. Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol. Rev.* **102**, 419–457 (1995).
27. Rumelhart, D. E., Hinton, G. E. & Williams, R. J. Learning representations by back-propagating errors. *Nature* **323**, 533–536 (1986).
28. Hinton, G. E., Osindero, S. & Teh, Y.-W. A fast learning algorithm for deep belief nets. *Neural Comput.* **18**, 1527–1554 (2006).
29. Russakovsky, O. et al. ImageNet Large Scale Visual Recognition Challenge. *Int. J. Comput. Vis.* **115**, 211–252 (2015).
30. Zhang, Q.-J., Gupta, K. C. & Devabhaktuni, V. K. Artificial neural networks for RF and microwave design-from theory to practice. *IEEE Trans. Microw. Theory Tech.* **51**, 1339–1350 (2003).
31. Patnaik, A., Mishra, R., Patra, G. & Dash, S. An artificial neural network model for effective dielectric constant of microstrip line. *IEEE Trans. Antennas Propag.* **45**, 1697 (1997).
32. Watson, P. M. & Gupta, K. C. EM-ANN models for microstrip vias and interconnects in dataset circuits. *IEEE Trans. Microw. Theory Tech.* **44**, 2495–2503 (1996).
33. Kabir, H., Wang, Y., Yu, M. & Zhang, Q.-J. Neural network inverse modeling and applications to microwave filter design. *IEEE Trans. Microw. Theory Tech.* **56**, 867–879 (2008).
34. Zaabab, A. H., Zhang, Q.-J. & Nakhla, M. A neural network modeling approach to circuit optimization and statistical design. *IEEE Trans. Microw. Theory Tech.* **43**, 1349–1358 (1995).
35. Southall, H. L., Simmers, J. A. & O'Donnell, T. H. Direction finding in phased arrays with a neural network beamformer. *IEEE Trans. Antennas Propag.* **43**, 1369–1374 (1995).
36. Nair, V. & Hinton, G. E. Rectified linear units improve restricted Boltzmann machines. In *Proc. 27th Int. Conf. Machine Learning (ICML-10)* 807–814 (2010).
37. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R. Dropout: a simple way to prevent neural networks from overfitting. *J. Machine Learning Res.* **15**, 1929–1958 (2014).
38. Ioffe, S. & Szegedy, C. Batch normalization: accelerating deep network training by reducing internal covariate shift. In *Proc. 32nd Int. Conf. Machine Learning* 448–456 (PMLR, 2015).
39. Hochreiter, S. & Schmidhuber, J. Long short-term memory. *Neural Comput.* **9**, 1735–1780 (1997).
40. Goodfellow, I. et al. Generative adversarial nets. In *NIPS'14: Proc. 27th Int. Conf. Neural Information Processing Systems* 2672–2680 (NIPS, 2014).
41. Kingma, D. P. & Welling, M. Auto-encoding variational Bayes. In *Proc. 2nd Int. Conf. Learning Representations* (ICLR, 2014); preprint at https://arxiv.org/abs/1312.6114.
42. Haykin, S. S. *Neural Networks and Learning Machines* (Prentice Hall, 2009).
43. Malkiel, I. et al. Plasmonic nanostructure design and characterization via deep learning. *Light Sci. Appl.* **7**, 60 (2018).
44. Ma, W., Cheng, F. & Liu, Y. Deep-learning-enabled on-demand design of chiral metamaterials. *ACS Nano* **6**, 6326–6334 (2018).
45. Liu, D., Tan, Y., Khoram, E. & Yu, Z. Training deep neural networks for the inverse design of nanophotonic structures. *ACS Photonics* **5**, 1365–1369 (2018).
46. Peurifoy, J. et al. Nanophotonic particle simulation and inverse design using artificial neural networks. *Sci. Adv.* **4**, eaar4206 (2018).
47. Pilozzi, L., Farrelly, F. A., Marcucci, G. & Conti, C. Machine learning inverse problem for topological photonics. *Commun. Phys.* **1**, 57 (2018).
48. Tahersima, M. H. et al. Deep neural network inverse design of integrated photonic power splitters. *Sci. Rep.* **9**, 1368 (2019).
49. Hemmatyar, O., Abdollahramezani, S., Kiarashinejad, Y., Zandehshahvar, M. & Adibi, A. Full color generation with Fano-type resonant HfO_2 nanopillars designed by a deep-learning approach. *Nanoscale* **11**, 21266–21274 (2019).
50. Sajedian, I., Badloe, T. & Rho, J. Optimisation of colour generation from dielectric nanostructures using reinforcement learning. *Opt. Express* **27**, 5874–5883 (2019).
51. Chen, Y., Zhu, J., Xie, Y., Feng, N. & Liu, Q. H. Smart inverse design of graphene-based photonic metamaterials by an adaptive artificial neural network. *Nanoscale* **11**, 9749–9755 (2019).
52. Nadell, C. C., Huang, B., Malof, J. M. & Padilla, W. J. Deep learning for accelerated all-dielectric metasurface design. *Opt. Express* **27**, 27523–27535 (2019).
53. Qiu, T. et al. Deep learning: a rapid and efficient route to automatic metasurface design. *Adv. Sci.* **6**, 1900128 (2019).
54. Zhang, T. et al. Efficient spectrum prediction and inverse design for plasmonic waveguide systems based on artificial neural networks. *Photonics Res.* **7**, 368–380 (2019).
55. Asano, T. & Noda, S. Optimization of photonic crystal nanocavities based on deep learning. *Opt. Express* **26**, 32704–32717 (2018).
56. Alagappan, G. & Png, C. E. Modal classification in optical waveguides using deep learning. *J. Mod. Opt.* **66**, 557–561 (2019).
57. Alagappan, G. & Png, C. E. Deep learning models for effective refractive indices in silicon nitride waveguides. *J. Opt.* **21**, 035801 (2019).
58. Kiarashinejad, Y., Abdollahramezani, S., Zandehshahvar, M., Hemmatyar, O. & Adibi, A. Deep learning reveals underlying physics of light-matter interactions in nanophotonic devices. *Adv. Theory Simul.* **2**, 1900088 (2019).
59. Comin, A. & Hartschuh, A. Efficient optimization of SHG hotspot switching in plasmonic nanoantennas using phase-shaped laser pulses controlled by neural networks. *Opt. Express* **26**, 33678–33686 (2018).
60. Li, L. et al. DeepNIS: deep neural network for nonlinear electromagnetic inverse scattering. *IEEE Trans. Antennas Propag.* **67**, 1819–1825 (2018).
61. Turpin, A., Vishniakou, I. & Seelig, J. D. Light scattering control in transmission and reflection with neural networks. *Opt. Express* **26**, 30911–30929 (2018).
62. Zhang, Q. et al. Machine-learning designs of anisotropic digital coding metasurfaces. *Adv. Theory Simul.* **2**, 1800132 (2019).
63. Li, L. et al. Machine-learning reprogrammable metasurface imager. *Nat. Commun.* **10**, 1082 (2019).
64. Maksov, A. et al. Deep learning analysis of defect and phase evolution during electron beam-induced transformations in WS_2. *npj Comput. Mater.* **5**, 12 (2019).
65. Xie, T. & Grossman, J. C. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Phys. Rev. Lett.* **120**, 145301 (2018).
66. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. In *Proc. 2016 IEEE Conf. Computer Vision and Pattern Recognition* https://doi.org/10.1109/CVPR.2016.90 (IEEE, 2016).
67. Szegedy, C. et al. Going deeper with convolutions. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition* https://doi.org/10.1109/CVPR.2015.7298594 (IEEE, 2015).
68. Sajedian, I., Kim, J. & Rho, J. Finding the optical properties of plasmonic structures by image processing using a combination of convolutional neural networks and recurrent neural networks. *Microsyst. Nanoeng.* **5**, 27 (2019).
69. Zhou, Q., Yang, C., Liang, A., Zheng, X. & Chen, Z. Low computationally complex recurrent neural network for high speed optical fiber transmission. *Opt. Commun.* **441**, 121–126 (2019).
70. Liu, Z., Raju, L., Zhu, D. & Cai, W. A hybrid strategy for the discovery and design of photonic nanostructures. *IEEE Trans. Emerg. Sel. Top. Circuits Systems* **10**, 126–135 (2020).
71. Ma, W., Cheng, F., Xu, Y., Wen, Q. & Liu, Y. Probabilistic representation and inverse design of metamaterials based on a deep generative model with semi-supervised learning strategy. *Adv. Mater.* **31**, 201901111 (2019).
72. So, S. & Rho, J. Designing nanophotonic structures using conditional deep convolutional generative adversarial networks. *Nanophotonics* **8**, 1255–1261 (2019).
73. Liu, Z., Zhu, D., Rodrigues, S. P., Lee, K.-T. & Cai, W. Generative model for the inverse design of metasurfaces. *Nano Lett.* **18**, 6570–6576 (2018).
74. Liu, Z. et al. Compounding meta-atoms into meta-molecules with hybrid artificial intelligence techniques. *Adv. Mater.* **32**, 1904790 (2019).
75. Wiecha, P. R. & Muskens, O. L. Deep learning meets nanophotonics: a generalized accurate predictor for near fields and far fields of arbitrary 3D nanostructures. *Nano Lett.* **20**, 329–338 (2019).

76. Jiang, J. & Fan, J. A. Global optimization of dielectric metasurfaces using a physics-driven neural network. *Nano Lett.* **19**, 5366–5372 (2019).
77. Bogdanov, S. I., Boltasseva, A. & Shalaev, V. M. Overcoming quantum decoherence with plasmonics. *Science* **364**, 532–533 (2019).
78. Linic, S., Christopher, P. & Ingram, D. B. Plasmonic-metal nanostructures for efficient conversion of solar to chemical energy. *Nat. Mater.* **10**, 911–921 (2011).
79. Ilic, O. & Atwater, H. A. Self-stabilizing photonic levitation and propulsion of nanostructured macroscopic objects. *Nat. Photon.* **13**, 289–295 (2019).
80. Li, J. et al. Addressable metasurfaces for dynamic holography and optical information encryption. *Sci. Adv.* **4**, eaar6768 (2018).
81. Jiang, J. et al. Free-form diffractive metagrating design based on generative adversarial networks. *ACS Nano* **13**, 8872–8878 (2019).
82. Kudyshev, Z. A., Kildishev, A. V., Shalaev, V. M. & Boltasseva, A. Machine-learning-assisted metasurface design for high-efficiency thermal emitter optimization. *Appl. Phys. Rev.* **7**, 021407 (2020).
83. Melati, D. et al. Mapping the global design space of nanophotonic components using machine learning pattern recognition. *Nat. Commun.* **10**, 4775 (2019).
84. Watanabe, T., Ayata, M., Koch, U., Fedoryshyn, Y. & Leuthold, J. Perpendicular grating coupler based on a blazed antiback-reflection structure. *J. Light. Technol.* **35**, 4663–4669 (2017).
85. Bar-Sinai, Y., Hoyer, S., Hickey, J. & Brenner, M. P. Learning data-driven discretizations for partial differential equations. *Proc. Natl Acad. Sci. USA* **116**, 15344–15349 (2019).
86. Rudy, S. H., Brunton, S. L., Proctor, J. L. & Kutz, J. N. Data-driven discovery of partial differential equations. *Sci. Adv.* **3**, e1602614 (2017).
87. Han, J., Jentzen, A. & Weinan, E. Solving high-dimensional partial differential equations using deep learning. *Proc. Natl Acad. Sci. USA* **115**, 8505–8510 (2018).
88. Trivedi, R., Su, L., Lu, J., Schubert, M. F. & Vuckovic, J. Data-driven acceleration of photonic simulations. *Sci. Rep.* **9**, 19728 (2019).
89. Qu, Y., Jing, L., Shen, Y., Qiu, M. & Soljacic, M. Migrating knowledge between physical scenarios based on artificial neural networks. *ACS Photonics* **6**, 1168–1174 (2019).
90. Liu, C.-X., Yu, G.-L. & Zhao, G.-Y. Neural networks for inverse design of phononic crystals. *AIP Adv.* **9**, 085223 (2019).
91. Ma, W. & Liu, Y. M. A data-efficient self-supervised deep learning model for design and characterization of nanophotonic structures. *Sci. China Phys. Mech. Astron.* **63**, 284212 (2020).
92. Sirignano, J. & Spiliopoulos, K. DGM: a deep learning algorithm for solving partial differential equations. *J. Comput. Phys.* **375**, 1339–1364 (2018).
93. Raissi, M. & Karniadakis, G. E. Hidden physics models: machine learning of nonlinear partial differential equations. *J. Comput. Phys.* **357**, 125–141 (2018).
94. Raissi, M., Perdikaris, P. & Karniadakis, G. E. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **378**, 686–707 (2019).
95. Hansen, N., Müller, S. D. & Koumoutsakos, P. Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES). *Evol. Comput.* **11**, 1–18 (2003).
96. Hegde, R. S. Photonics inverse design: pairing deep neural networks with evolutionary algorithms. *IEEE J. Sel. Top. Quant. Electron.* **26**, https://doi.org/10.1109/JSTQE.2019.2933796 (2019).
97. Guler, U., Boltasseva, A. & Shalaev, V. M. Refractory plasmonics. *Science* **344**, 263–264 (2014).
98. Marino, G. et al. Spontaneous photon-pair generation from a dielectric nanoantenna. *Optica* **6**, 1416–1422 (2019).
99. Zhang, Q., Yu, H., Barbiero, M., Wang, B. & Gu, M. Artificial neural networks enabled by nanophotonics. *Light Sci. Appl.* **8**, 42 (2019).
100. Khoram, E. et al. Nanophotonic media for artificial neural inference. *Photonics Res.* **7**, 823–827 (2019).
101. Shen, Y. et al. Deep learning with coherent nanophotonic circuits. *Nat. Photon.* **11**, 441–446 (2017).
102. Feldmann, J., Youngblood, N., Wright, C., Bhaskaran, H. & Pernice, W. All-optical spiking neurosynaptic networks with self-learning capabilities. *Nature* **569**, 208–214 (2019).
103. Hughes, T. W., Minkov, M., Shi, Y. & Fan, S. Training of photonic neural networks through in situ backpropagation and gradient measurement. *Optica* **5**, 864–871 (2018).
104. Lin, X. et al. All-optical machine learning using diffractive deep neural networks. *Science* **361**, 1004–1008 (2018).
105. Bao, Q. et al. Monolayer graphene as a saturable absorber in a mode-locked laser. *Nano Res.* **4**, 297–307 (2011).
106. Tait, A. N. et al. Silicon photonic modulator neuron. *Phys. Rev. Appl.* **11**, 064043 (2019).
107. Williamson, I. A. et al. Reprogrammable electro-optic nonlinear activation functions for optical neural networks. *IEEE J. Sel. Top. Quant. Electron.* **26**, https://doi.org/10.1109/JSTQE.2019.2930455 (2019).
108. Shastri, B. J. et al. Spike processing with a graphene excitable laser. *Sci. Rep.* **6**, 19126 (2016).
109. Tait, A. N. et al. Neuromorphic photonic networks using silicon photonic weight banks. *Sci. Rep.* **7**, 7430 (2017).
110. Estakhri, N. M., Edwards, B. & Engheta, N. Inverse-designed metastructures that solve equations. *Science* **363**, 1333–1338 (2019).
111. Hughes, T. W., Williamson, I. A., Minkov, M. & Fan, S. Wave physics as an analog recurrent neural network. *Sci. Adv.* **5**, eaay6946 (2019).

Acknowledgements

Y.L. acknowledges financial support from the US National Science Foundation (NSF) (ECCS-1916839) and the Office of Naval Research (N00014-16-1-2409). W.C. acknowledges support from the Office of Naval Research (N00014-17-1-2555) and the NSF (DMR-2004749). The Purdue team acknowledges financial support from DARPA/DSO (HR00111720032, Z.A.K.), the US National Science Foundation (ECCS-2029553, A.B.) and the Air Force Office of Scientific Research (AFOSR) (FA9550-20-1-0124, A.B.).

Competing interests

The authors declare no competing interests.

Additional information

Correspondence should be addressed to A.B., W.C. or Y.L.