Review

Sunae So[a], Trevon Badloe[a], Jaebum Noh[a], Jorge Bravo-Abad* and Junsuk Rho*

Deep learning enabled inverse design in nanophotonics

https://doi.org/10.1515/nanoph-2019-0474
Received November 20, 2019; revised January 2, 2020; accepted January 13, 2020

Abstract: Deep learning has become the dominant approach in artificial intelligence to solve complex data-driven problems. Originally applied almost exclusively in computer-science areas such as image analysis and nature language processing, deep learning has rapidly entered a wide variety of scientific fields including physics, chemistry and material science. Very recently, deep neural networks have been introduced in the field of nanophotonics as a powerful way of obtaining the nonlinear mapping between the topology and composition of arbitrary nanophotonic structures and their associated functional properties. In this paper, we have discussed the recent progress in the application of deep learning to the inverse design of nanophotonic devices, mainly focusing on the three existing learning paradigms of supervised-, unsupervised-, and reinforcement learning. Deep learning forward modelling i.e. how artificial intelligence learns how to solve Maxwell's equations, is also discussed, along with an outlook of this rapidly evolving research area.

Keywords: artificial intelligence; nanophotonics; machine learning; inverse design; forward modelling; metamaterials.

[a]**Sunae So, Trevon Badloe and Jaebum Noh:** These authors contributed equally to this work.
***Corresponding authors: Junsuk Rho,** Department of Mechanical Engineering, Pohang University of Science and Technology (POSTECH), Pohang 37673, Republic of Korea; and Department of Chemical Engineering, Pohang University of Science and Technology (POSTECH), Pohang 37673, Republic of Korea, e-mail: jsrho@postech.ac.kr. https://orcid.org/0000-0002-2179-2890; and **Jorge Bravo-Abad,** Department Física Teórica de la Materia Condensada and Condensed Matter Physics Center (IFIMAC), Universidad Autónoma de Madrid, 28049-Madrid, Spain, e-mail: jorge.bravo@uam.es
Sunae So, Trevon Badloe and Jaebum Noh: Department of Mechanical Engineering, Pohang University of Science and Technology (POSTECH), Pohang 37673, Republic of Korea. https://orcid.org/0000-0001-8606-2234 (S. So)

1 Introduction

Nanophotonics is devoted to the study of light-matter interaction at the subwavelength scale [1]. During the last few decades, important fundamental advances combined with the spectacular progress of nanoscale fabrication methods [2–4] have led to a broad range of innovations in nanophotonics [5, 6], largely based on tailoring periodically structured materials to create 2D and 3D metasurfaces [7, 8] or metamaterials [9] that exhibit extraordinary properties that cannot be found in nature. This includes advances in the fields of plasmonics [10, 11], holography [12, 13], artificial chirality [14, 15] and topological photonics [16, 17].

Remarkably, most of these breakthroughs are mainly based on human intuition. By way of illustration, let us consider a fundamental photonic problem: the scattering of light with a simple dielectric object. It is well-known that for obtaining a polarization insensitive optical response, we should use a symmetric design such as a circular rod [18, 19]. Then, if we want to add a different response for each polarization state, we can elongate the rod in one direction to create an oval shape, as we know from previous experience and reasonable physical arguments that the increased amount of material along the elongated dimension will create a different optical response for each incident polarization [20, 21]. However, there is no reason to believe that our intuition has led us to an optimal design with the highest possible performance. The limiting character of human-based intuition for the design of improved nanoscale devices shown in this simple example is apparent from a more general standpoint when considering the tremendous control over topology and composition of nanophotonic structures allowed by state-of-the-art nanofabrication techniques.

Inverse design has received increasing attention as a powerful approach to go beyond human-intuition based devices [22, 23]. The conventional design process usually starts from the known library of designs that have been proven to work for the given task. Then, computational optimization techniques [24–30] are used to find the optimal design. Although these techniques do indeed

Open Access. © 2020 Jorge Bravo-Abad, Junsuk Rho et al., published by De Gruyter. This work is licensed under the Creative Commons Attribution 4.0 Public License.

represent a key tool in current nanophotonic research, the process is extremely time and computationally intensive, leading to a singular "best" design for each optimization goal and parameter space under study (a modification of that optimization goal requires a new inverse optimization process to be run again from scratch). In addition, when fabricating the obtained design, there could be additional practical constraints preventing certain design parameters to be met exactly. When that occurs the corresponding optimization process has to be often carried out again. As nanophotonic designs become more intricate and fabrication techniques allow for more complex three dimensional designs [31–33], this process becomes even more resource intensive.

In parallel, the unprecedented development of artificial intelligence (AI) that has taken place during the last few years has remarkably accelerated the pace of technological disruptive advances in multitude contexts. Deep learning (DL) in particular, a branch of AI originally inspired by the biological neural networks of animal brains, has largely benefited from the availability of large datasets and recent advances in architectures, algorithms, and computational hardware. This, in turn, has led to impressive new applications that we could not have been imagined a few years ago, from improved computer vision [34] paving the way for driverless cars [35], enhanced speech recognition [36] that allows us to interact verbally with devices, to new stock management systems that provide us with next day delivery services all over the world [37].

The combination of the above two areas (i.e. inverse design in nanophotonics and DL) is nowadays emerging as a fundamentally new approach that offers the promise of solving some of the key challenges faced in nanophotonic inverse design. Conventional optimization algorithms [26, 38] are usually programmed within a specific set of boundaries, with a figure of merit used to optimize the output. A DL model on the other hand, is trained through non-linear activation functions and back propagation [39] to intelligently learn the nonlinear relationships between the input and output values over a large dataset. In this way, a model is able to effectively "learn" Maxwell's equations and how to solve them, without explicitly knowing them. This, in turn, allows for the possibility of the discovery of solutions outside of the boundaries of the training data, and also the ability to transfer knowledge between problems, a method known as "transfer learning" [40]. This approach represents a complete change of paradigm of how nanophotonics research have been understood until now, and it is expected to lead to an equally disruptive series of novel findings in nanophotonics.

While DL has revolutionized many fields over recent years, it is still very much so in its infancy in the field of nanophotonics. The inherent weaknesses of DL in all fields are, of course, also present in nanophotonics. In particular, the large datasets that allow for facial recognition and image classification are made up from millions of users on platforms around the world, whereas for a single problem in nanophotonics, the dataset generally needs to be made specifically for the task. While a single run of other inverse designs may take a few hundred simulations to reach a desired optimization, DL has a much higher up-front computational cost (the number of data that a single DL model should learn easily exceeds the number of simulations required for other inverse design techniques). In addition, the large amount of data required for a DL approach could not be easily accessible and is usually created using simulation methods such as RCWA, FEM and FDTD, which are time and computationally expensive. The results of DL are also sensitive to the dataset, so care needs to be taken to ensure that we are allowing the network to learn from good data. This includes normalizing, standardizing and cleaning the dataset to increase its practicality. Finally, the hyperparameters for the network and the DL algorithm also need to be optimized, which requires an extensive study to find the optimal network for the task at hand. In this article we have presented a review of the recent activity in the rapidly evolving area of DL assisted inverse design of nanophotonic devices. This review is organized as follows. We start by introducing the recent progress of DL in nanophotonics regarding forward modelling. DL has been extremely successful in this area, with various examples of models that can instantaneously and successfully predict the optical properties of nanophotonic design. DL-based forward modelling also represents a key concept to understand subsequent advances on inverse design. After that, we have presented a throughout discussion of DL enabled inverse design in nanophotonics, considering the current three paradigms of machine learning, namely, supervised learning, unsupervised learning and reinforcement learning (RL). Finally, we end up with a set of conclusions of this work and outlook on the bright future of DL in nanophotonics.

2 Deep learning for forward nanophotonic modelling

In essence, forward modelling in nanophotonics consists in predicting the optical properties of photonic structures featuring subwavelength-scale complex features. To do

that, conventional approaches solve, either analytically [41] or numerically [42], the corresponding Maxwell's equations governing light propagation in such complex photonic environments. For instance, the transfer matrix method analytically describes the light propagation in a stratified medium by obtaining closed expressions for the complex amplitude of reflection and transmission [43, 44]. Similarly, the rigorous coupled wave analysis (RCWA) method offers a semi-analytic approach [45, 46] that can be used to obtain the optical responses of periodic structures in the Fourier domain. Despite the obvious value of these methods, as the complexity of the photonic environment increases, it becomes more and more difficult to obtain analytic or semi-analytic solutions that can accurately capture all the physical ingredients featured by the considered problems (this is particularly the case when going from 1 and 2 dimensions to a 3 dimensional system). For arbitrarily complex nanophotonic structures, fully numerical simulations (such as the finite-element method or the finite difference time domain method [47, 48]) are employed to obtain the associated optical responses. These highly sophisticated approaches are essentially based on discretizing the studied system and solving Maxwell's equations at each spatial location. This is precisely what, on one hand, gives these techniques their general character, but on the other hand, makes them computationally expensive, especially as the complexity of the design grows and an increasingly finer spatial discretization is required [42]. Plasmonic systems [11], in which the deep-subwalength scale of plasmonic resonances is often combined with much larger length scales, are perhaps the canonical case that illustrates this aspect.

Recently, data driven approaches of DL have been introduced as a new, powerful, versatile approach for forward modelling in nanophotonics [49–56]. The underlying idea is based on predicting the optical response of a given photonic system by approximating Maxwell's equations, therefore allowing the response to be obtained without explicitly solving them. That, in turn, removes the need of computationally intensive numerical simulations from the picture. In the pioneering work, Peurifoy et al. predicted the scattering cross section of a silicon dioxide (SiO_2)/titanium dioxide (TiO_2) multi-layered, core-shell nanoparticle using a deep neural network (DNN) [49]. Specifically, they used a fully connected DNN featuring several hidden layers to approximate the scattering cross section of the core-shell nanoparticle for the given inputs of thickness for each layer (Figure 1A). The network was

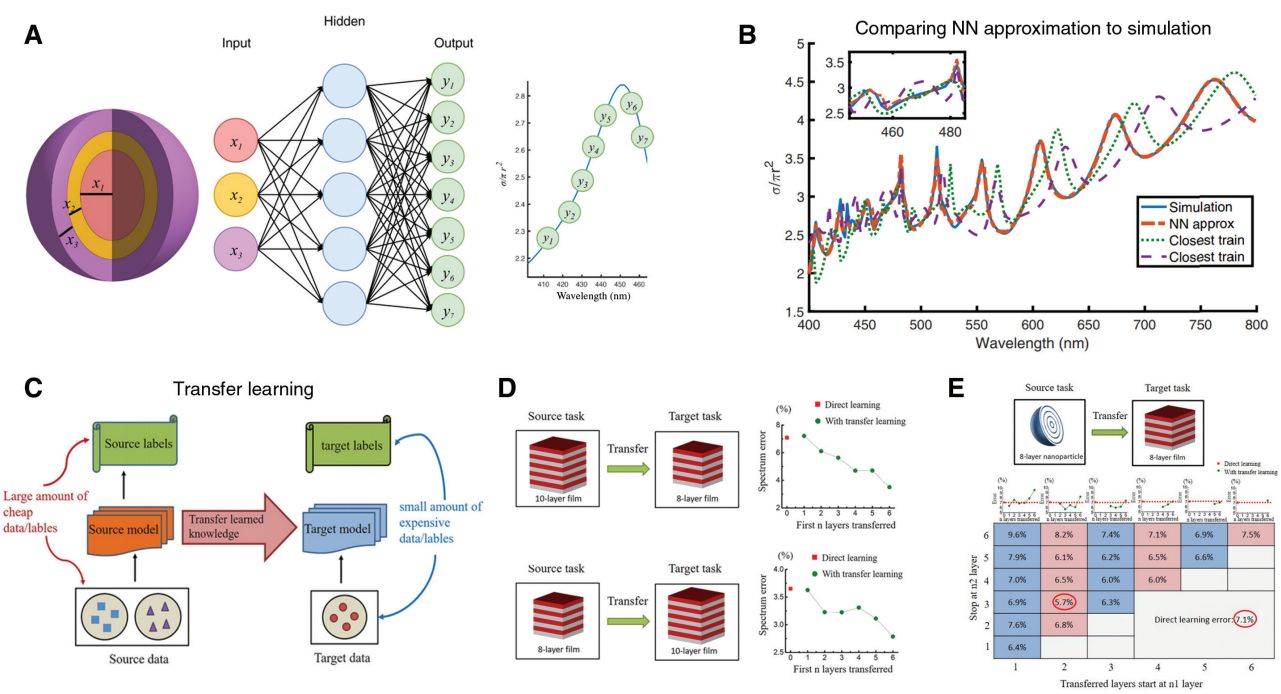

Figure 1: Forward modelling enabled by DL.
(A) A schematic of DNN to predict scattering cross section of the core-shell nanoparticle. (B) Comparison of the approximated spectrum by NN to the simulation results, and closest data. (C) A schematic of transfer learning. Knowledge is transferred from the source data to target data. (D), (E) Results of the transferred learning between (D) similar and (E) different physical systems. (A) and (B) are adapted with permission from Ref. [49] (CC BY 4.0). (C–E) are reprinted with permission from Ref. [50], Copyright 2019 American Chemical Society.

trained with a set of previously obtained 50,000 scattering cross section spectra, as calculated for different instances of the system generated varying randomly the thickness of each layer within a given experimentally-accessible interval of values (note that the transfer method was used to efficiently generate this large amount training spectra). After training, a test multilayer core-shell particle, which had never been seen in the previous training steps, was used to validate the network. The predicted optical responses of the scattering cross section as a function of wavelength agreed well with the target responses of the given structural parameters (Figure 1B). Moreover, a comparison between the predicted result and the closest training data showed that the network had uncovered some underlying pattern between the input and output data, rather than simply interpolating or averaging the closest data points (Figure 1B).

Another remarkable demonstration of how DL algorithms can indeed learn complex optical behavior from structural parameters is found in Ref. [50]. In that work, Qu et al. demonstrated the possibility of predicting optical properties in a given physical scenario with the help of the knowledge obtained from a different, but related, physical problem. They used the "transfer learning" [39, 40], which essentially consists in reusing a previously trained model for a different, but related task (Figure 1C). In particular, in Ref. [50] a DNN was first trained with transmission data corresponding to 8-layer multilayer system, and then was reused to predict the transmission response of 10-layer system (Figure 1D). The underlying physics extracted from the 8-layer system was therefore transferred to the case of the 10-layer system, for which efficient learning took place even with an insufficient amount of training data. They also showed that this approach can be extended to two different tasks, namely, the scattering spectrum of a multilayer nanoparticle system and the transmission of multilayer films (Figure 1E). With the help of the transferred knowledge, the error of the network was significantly reduced. The underlying idea of why this approach actually working well is that, although two different physical structures are being considered, both structures share common physical rules, and, in that way, DL indeed discovers the underlying physics from data, rather than simply interpolating or regularizing it. Importantly, this method provides a particularly efficient learning method when the dataset is too small or hard to acquire (i.e. a method for avoiding the overfitting problem on a small dataset). In this context, we also point out that recent research also has shown that DL can be employed to discover knowledge about the physics of light-matter interactions by finding the range of feasible responses in the latent space [53].

The above examples demonstrate how DL can be used to predict optical properties from given structural parameters, where the inputs are limited to a few geometric parameters. Recently, attempts have been made to predict optical properties from any given arbitrary shapes [54, 55]. Inampudi and Mosallaei firstly developed a metagrating antenna with arbitrary shapes of a sixteen-sided polygon [55]. The shapes of antenna are defined by the radius coordinates of sixteen vertices (Figure 2A), which can represent any arbitrary shapes. The DNN takes sixteen coordinates as inputs and predicts the diffraction efficiency of 13 diffraction orders. The results showed that the network can predict diffraction efficiency given the vertices coordinate information (Figure 2B). Thus, this research work extended for the first time the use of DL for the prediction of optical properties from any structures using coordinate information. Later, in Ref. [54] Sajedian et al. generalized that approach by using a 2D image as the input. Specifically, 2D cross-sectional images of plasmonic structures were used as inputs, and the neural network predicted the corresponding absorption spectra. Input images were composed of arbitrary structures, where black and white images represent the existence (or not) of a silver structure at certain locations. For the DL process, they used a convolutional neural network (CNN) [57], which has proved to be an efficient implementation for extracting key features from images [35]. In addition, a recurrent neural network (RNN) [58] was used to find the correlation within the data (Figure 2C). Interestingly, the results showed that the network was able to predict the absorption spectra from the given input structural images (Figure 2D). Thus, this research work extended for the first time the use of DL for the prediction of optical properties from any structure using images.

As discussed above, recent work on forward modelling enabled by DL has shown the ability of AI algorithms to learn the complex relations between nanophotonic structures and their associated optical responses. A natural extension of this concept consists in assuming that DL can also solve complex inverse design problems, i.e. the inverse process of forward modelling [49]. However, unlike forward modelling (where there is a one-to-one mapping between one physical system and its corresponding response), inverse design has to tackle the possible problem of non-uniqueness (i.e. several different designs can produce the same optical response, which makes the whole problem significantly more challenging) [59]. One of the most common approaches to overcome these issues relies on adding a forward modelling network into the inverse design DNN architecture [59–61], so that an additional tandem network can be trained

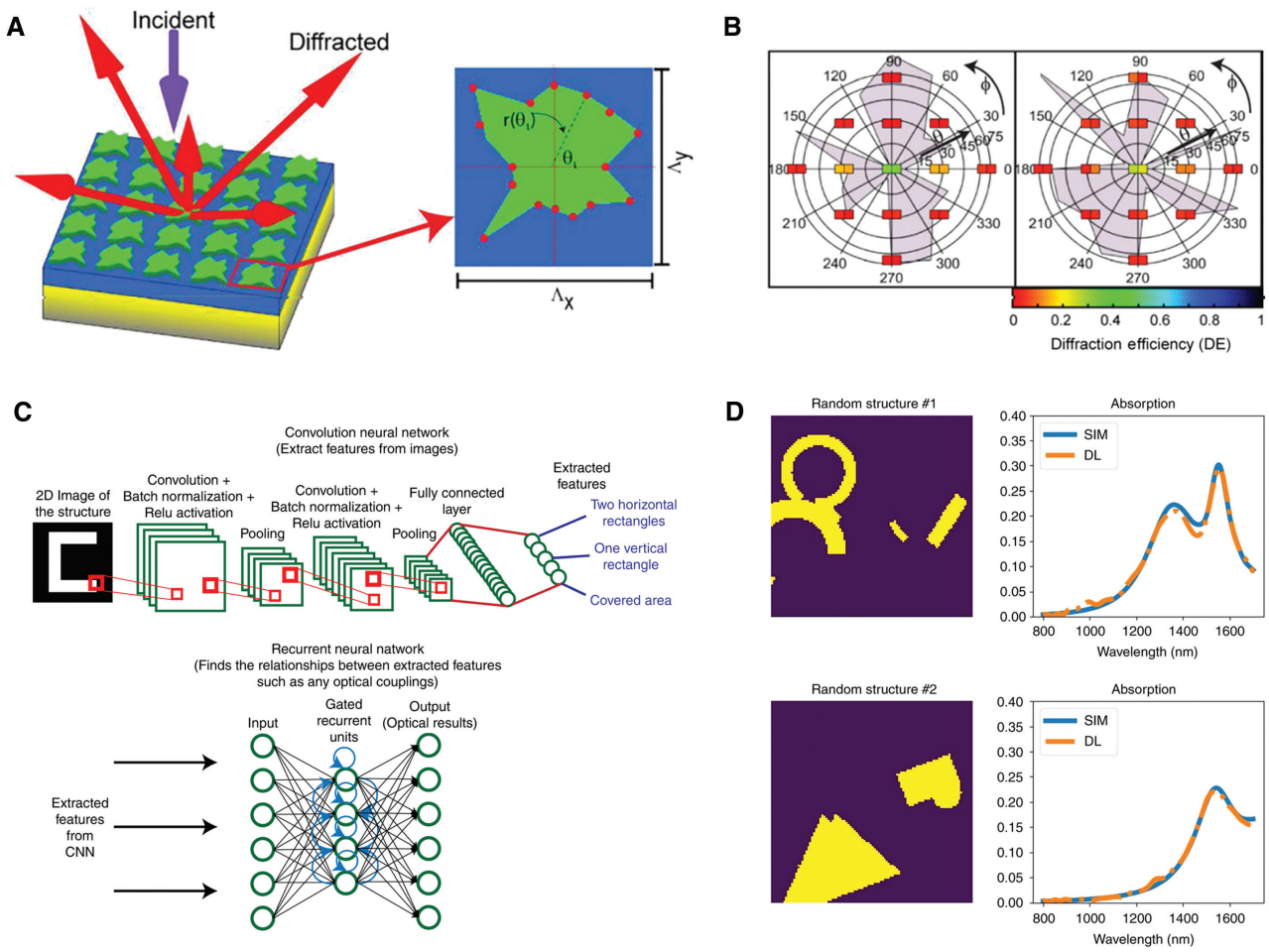

Figure 2: Image-based forward modelling via DL.
(A) A schematic of a metagrating antenna with arbitrary shape. A sixteen-sided polygon is defined by 16 radius coordinates of the corresponding vertices. (B) Examples of the prediction results. In each panel, the background grey images display the shapes of metagrating, whereas the double square boxes correspond to target (left) and predicted (right) diffraction efficiency at each diffraction angle. (C) A network structure combined with CNN and RNN. A CNN is used to extract important features from the 2D images, and an RNN is used to find the correlations in the data. (D) Examples of the results. 2D images (lefts) are fed as inputs, and the network provides absorption spectra (rights) as outputs. (A) and (B) are adapted with permission from Ref. [55] and (C) and (D) are reprinted with permission from Ref. [54] (CC BY 4.0).

simultaneously to find optimal designs. This point is discussed in more detail in the following sections.

3 Deep learning nanophotonic inverse design

3.1 Supervised learning in inverse design

Supervised learning can be defined as the task of finding the complex (in general non-linear) relationships between two sets of pre-labelled data [62]. Because in this case the network learns mappings with explicit instances of input-output pairs, the supervised learning method features an excellent performance with dealing with well-defined problems. In nanophotonic inverse design, this learning method has been applied to design structural parameters of the pre-defined structural shapes [49, 59, 61, 63–65]. The data are prepared with explicit pairs of design parameters and their associated optical properties, and the network is trained iteratively to provide appropriate structural parameters for given input optical properties.

In the work of Peurifoy et al. [49] (already discussed in Section 2), the authors used their forward DL model in reverse to design a multi-layered core-shell nanoparticle. Specifically, after training their forward modelling

network, they used it in reverse to infer the best design parameters for a "random" spectrum. In this first example of inverse design, the team froze the weights of the DNN and fixed the output to a specific spectrum. Rather than retraining the network, they iterated through possible input values for the network to find the combination that gave the closest result. Notably, while conventional optimization methods typically get stuck in a local minimum, the proposed DNN avoided that fate and lead to the most optimized result.

An improved method for using DL in inverse design was demonstrated by Liu et al. [59], by employing a tandem DNN that combined the inverse network with a pretrained forward model (Figure 3A). In that work, the authors designed an SiO_2 and Si_3N_4 multilayer structure (Figure 3B), where the thickness of each layer made up the design parameters. After training the forward model that links the design parameters to the transmission spectra, they fixed the weights and added the inverse design network to the front. Their training dataset was made up of 500,000 labelled pairs of data, with 50,000 more for testing. The desired spectrum was used as the input, and the network minimized the loss between the desired spectrum and the recovered spectrum, while the design parameters were extracted from the intermediate layer. As mentioned at the end of the previous section, this particular architecture has the important advantage of overcoming the issues of non-uniqueness, as the solution does not require the design to be specific for a set of design parameters (it instead requires that the loss between the desired and predicted output spectra is small). In addition, in that work the authors demonstrated that the same method can be used to design 2D structures able to modulate transmission phase delay at three specific wavelengths.

Despite the success in the inverse design using DL, the above described methods have been applied to design a few structural parameters, while the different types of the component materials are fixed. In Ref. [61] So et al. took a step forward and investigated the design of a core-shell nanoparticle by combining into the same implementation both regression and classification for the design parameters and the material of each layer, respectively (Figure 3C). In this context, let us recall in passing that, depending on the characteristics of the output data, machine learning problems can be generally divided into two main categories: regression and classification. Regression requires the prediction of continuous quantities, such as a time series, whereas classification focuses on allocating the data into discrete classes. To solve the combined regression and classification problems, in Ref. [61] the loss function has to incorporate both the regression, for the

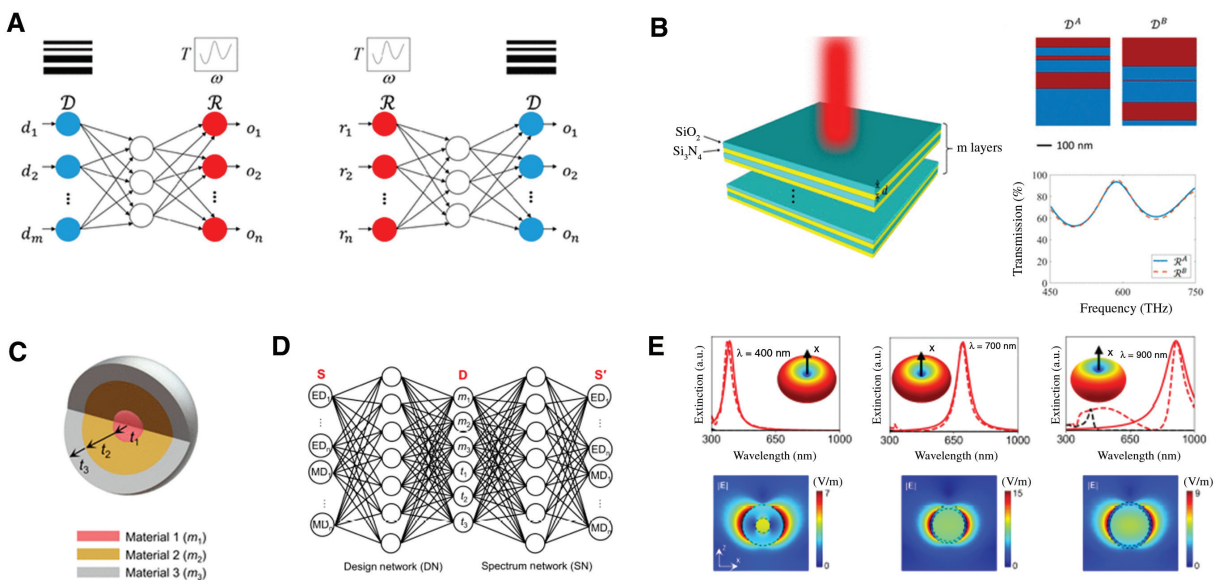

Figure 3: Inverse design in nanophotonics via DL.
(A) Schematics of forward model and inverse model. (B) A schematic of the SiO_2 and Si_3N_4 multilayer structure and two designs that have the same transmission response. (C) A schematic of Core-shell nanoparticle and (D) corresponding DL model to simultaneously design materials and thicknesses of the particle. (E) Examples of practical inverse design of Core-shell nanoparticle for spectral tuning ED resonances. (A), (B) are reprinted with permission from Ref. [59], Copyright 2018 American Chemical Society. (C–E) are adapted with permission from Ref. [61], Copyright 2019 American Chemical Society from [59].

continuous values of core-shell layer thickness, and classification, to choose the most appropriate material of each layer (Figure 3D). In that work a dataset of 18,000 labelled data was used, with 80% used for training and 10% each for validation and testing. The obtained network was able to determine the structural parameters and materials to give the required electric dipole (ED) and magnetic dipole (MD) extinction spectra. After being trained, the model was tested by using hand drawn Lorentzian functions with peaks at specific locations (Figure 3E). The model was successful in determining a design that gave a response that is extremely similar to the input. Remarkably, this allows for the design of core-shell nanoparticles that have ED and MD resonances at specific, user determined locations, within fractions of a second.

Recent research works were on utilizing DL for the inverse design for more practical applications. For example, DL has been applied in the field of topological photonics [66] by Pilozzi et al. [67]. They created a photonic topological insulator with an array of layers, modelled by the transfer matrix technique (Figure 4A). In the proposed inverse design process, the model was required to find a target edge-state with a specific frequency, which was then used as the input. In this application, discontinuities in the features space led the authors to use multiple independent DNNs for each specific variable (Figure 4B). To ensure that the solutions provided by the inverse network were indeed viable and physical solutions, Pilozzi et al. also validated the obtained solutions by using them as inputs to their forward modelling network. This allowed them to check that any multivalued degeneracy was effectively removed.

Meanwhile, Baxter et al. [52] applied DL to the field of plasmonic color. Using both experimental and simulated data, they predicted laser parameters and the geometry of nanoparticles to create desired colors (Figure 4C). Their unique inverse design method trained n DNNs for each design, where n is the number of

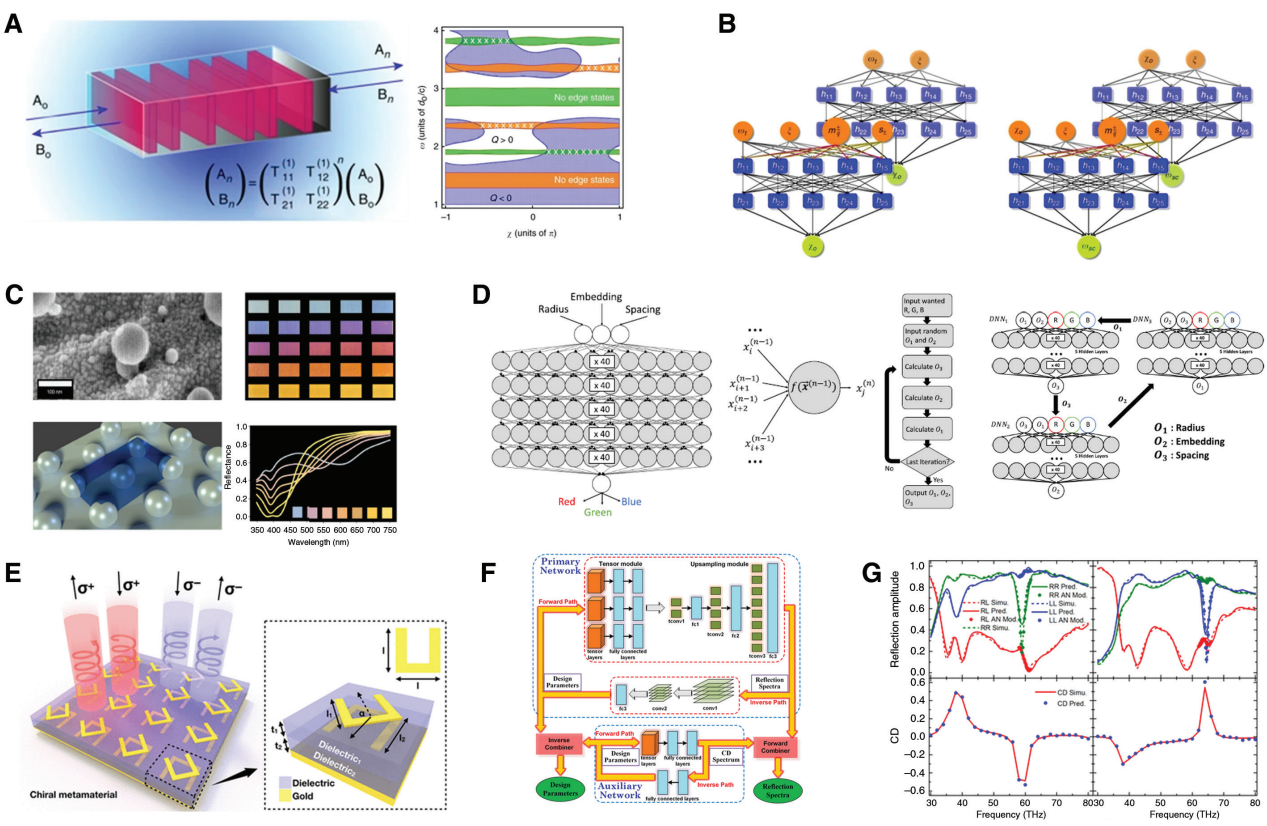

Figure 4: Practical applications of DL for inverse design in nanophotonic structures.
(A) A schematic of the topological insulator and its edge state obtained from transfer matrix. (B) Direct and inverse networks used to find the specific edge state frequency. (C) Laser machined silver surface and its color representation. (D) Iterative DNN inverse design. (E) A schematic of the two-layer split-ring resonator. (F) The DNN design, including the forward modelling and inverse design. (G) An example of the predicted spectra for a given CD response with two resonance points. (A) and (B) are reprinted with permission from Ref. [67] (CC BY 4.0). (C–D) are adapted with permission from Ref. [52] (CC BY 4.0). (E–G) are reprinted with permission from Ref. [68], Copyright 2018 American Chemical Society.

outputs, i.e. design parameters (Figure 4D). Each DNN takes the input color and all but one design parameter as inputs, to produce one output value. The next DNN is then trained with the updated design parameter, and this process is repeated until a solution is found. In that work, the authors chose to initialize the starting design parameters as the mean values from their dataset. This process generally took 10–20 iterations to converge to a final solution, but did not guarantee that the optimal solution was found, as it depended on the initial values. To overcome this issue, Baxter et al. put forward that multiple initialized inputs can be used to produce multiple designs, and the one with the least error can be selected.

On a different note, Ma et al. implemented two bidirectional DNNs for the on-demand inverse design of chiral metamaterials [68]. Since one resonant feature of a chiral response can be approximated by a Lorentzian function, the authors hand-drew a desired spectrum with two resonance peaks and predicted the five design parameters for a two-layered split ring resonator that displayed that spectrum (Figure 4E). With a dataset of 30,000 samples, Ma et al. were able to simultaneously solve the forward modelling and inverse design problems (Figure 4F). From the obtained results, they were able to discover designs that leading high circular dichroism responses in structures that were almost symmetric. Importantly, this finding goes against human intuition, which highlights the nonlinear relationship between geometric chirality and chiroptical response (Figure 4G).

3.2 Unsupervised learning in inverse design

In contrast to supervised learning, unsupervised learning handles data without explicit instructive labels; that is, the network infers important patterns from the data without a desired or correct answer [69, 70]. This may cause the DNN to define problems less clearly and therefore be more difficult to solve. However, since unsupervised systems learn by themselves without a specific goal, they are superior to supervised learning systems at discovering new patterns in completely new data. In nanophotonics, the strength of unsupervised learning methods has been used to solve inverse design problems for arbitrary shapes [60, 71–74]. In this class of problems, the network is trained with certain types of geometric images and their optical properties, then infers the nonlinear mapping to design arbitrary shapes. In this context, 2D cross sectional images were used to represent arbitrary geometries of the structures.

In Ref. [60], Liu et al. demonstrated for the first time the possibility of performing inverse design of nanophotonic structures using an unsupervised learning system. More specifically, they used a generative adversarial network (GAN) [75] to design arbitrary geometries of metasurfaces. GANs are rather a recently developed machine learning algorithm, but have become one of the most interesting unsupervised learning methods [75]. Due to their relevance, let us briefly account for how a GAN works. A GAN consists of two networks, a generator and a discriminator. The two networks compete in a zero-sum game and simultaneously learn. The generator takes random noise and generates structural images that should have the desired optical properties, while the discriminator judges whether the generated images are from the structural geometric data of interest. The objective of the generator network is to deceive the discriminator network by generating realistic structural images. Therefore, after training, the generator network is able to create designs that resemble images in the actual geometric data, i.e. the generator network infers important patterns from the data through the feedback from the discriminator network. In addition to the GAN model, in Ref. [60], the authors also added a simulator network to approximate the optical properties from the generated design images (Figure 5A). A total of 6,500 sets of data were prepared for the DL algorithm, each entry consisting of a binary pixel image of the antenna metasurface and a complete set of transmission spectra for each polarization. After training, the network was capable of providing structural images for given transmission spectra. Both test data and random user-drawn spectra were used to evaluate the trained network, and the results showed that the network can provide structures that have desired optical properties (Figure 5B). The network generates arbitrary structural patterns, which allows an insight into new structures that are beyond human intuition built on experience and knowledge. However, as GANs solve the unstable Nash equilibrium problem, they provide unstable solutions. Recent research has introduced deep convolutional GANs [76] into nanophotonic inverse design problems to solve the zero-sum problem in a more stable manner [71].

Recently, Ma et al. [77] have introduced a semi-supervised learning strategy, where both labelled- and unlabelled-data are used for training to improve the model performance [72]. They used a probabilistic model of a variational auto-encoder (VAE) [78] for inverse design. VAEs are another class of powerful generative models [79], which reconstruct the input after being compressed into a few latent variables. Unlike other inverse design approaches, the network takes input geometry, and

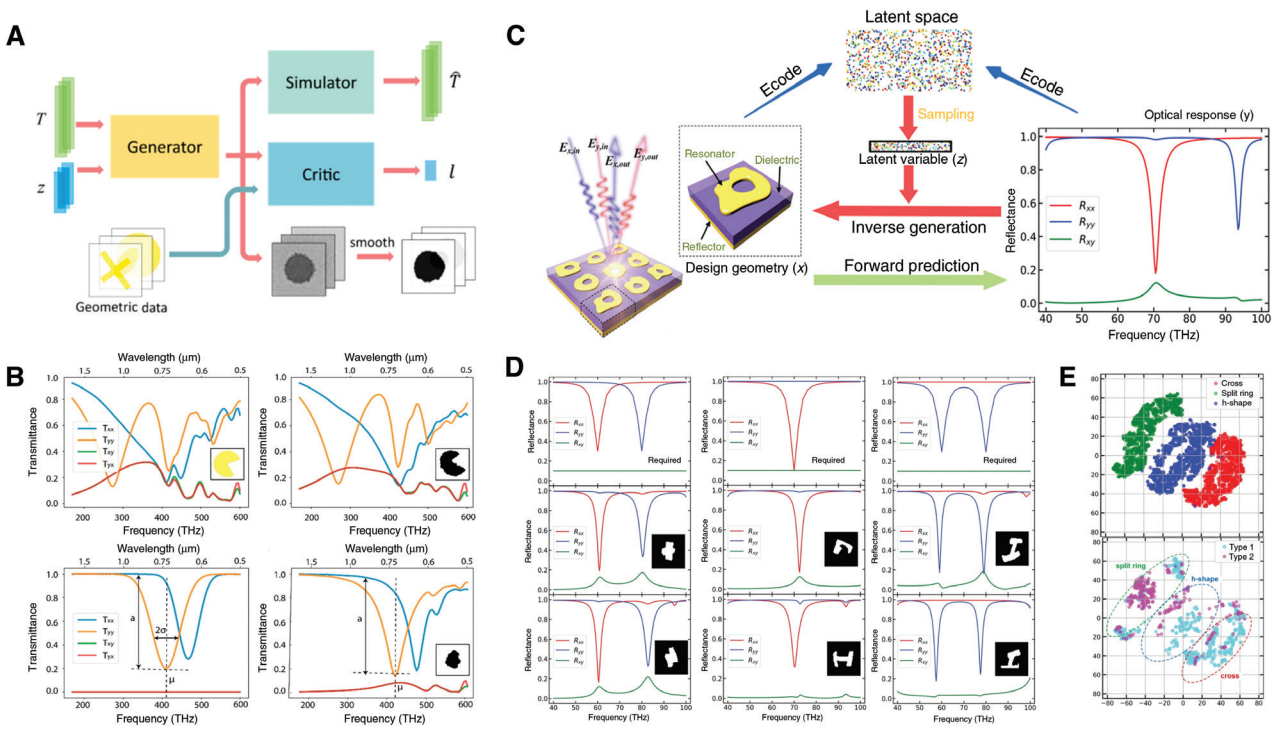

Figure 5: Inverse design using unsupervised learning methods.
(A) A GAN model to inverse design structural images. (B) Examples of the results of the inverse design. Full sets of transmission spectra (left) of the inputs (right) obtained from the designed structures. Insets are the associated structural images. (C) A schematic of inverse design using VAE. (D) Examples of the results. The top spectra are target spectra, and the middles and bottoms are two examples of the retrieved structural designs (insets) and their associated reflection spectra. (E) Visualization of the latent space. (A) and (B) are reprinted with permission from Ref. [60], Copyright 2018 American Chemical Society. (C–E) are adapted with permission from Ref. [72], Copyright 2019 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

encodes the structural design and optical responses into the latent variables with a predefined distribution (Figure 5C). In the case of Ref. [77], the important information of 64×64 structural designs and their optical properties were compactly encoded in 20 latent variables. Then, these latent variables were stochastically sampled from the latent space and decoded to reconstruct the original structural geometry. Accordingly, the decoding process can be used to solve the inverse design problem. In addition, the sampling process provides diversity in the outputs, allowing the production of many candidates for the inverse design (Figure 5D). This is closely connected with the physical insight of the non-uniqueness problem inherent in inverse design. Indeed, the network was able to provide very different meta-atom geometries with optical properties that are quite similar to the desired input spectrum (Figure 5D). Moreover, the latent space was explored for interpretability, and it could be observed that three different geometry groups (cross, split ring, and h-shaped antennas) emerged from the analysis (Figure 5E). This clearly showed that the network automatically learns to distinguish the different geometry groups without any specific labels or instructions. Finally, Ma et al. demonstrated that the slight change in the sampling latent variables caused the network to generate patterns with different geometries, which could provide a more comprehensive generation mechanism from the network.

The combination of unsupervised learning methods of DL, in particular GANs, with a physics-informed inverse design method has been also recently proposed as an approach to find optimal nanophotonic designs. In particular, Jiang et al. combined the idea of a GAN with a topology optimization method to create a novel inverse design method [73]. This method exploits the generating property of GANs and uses it to generate sufficient quantities of training data of the metasurfaces (Figure 6A). The goal was to design the optimal topology of a metasurface for high diffraction efficiency at a target wavelength and incident angle. The DL network captures the important features of the metasurfaces with high efficiency and generates possible candidates that contain those features. After that, the topology optimization is

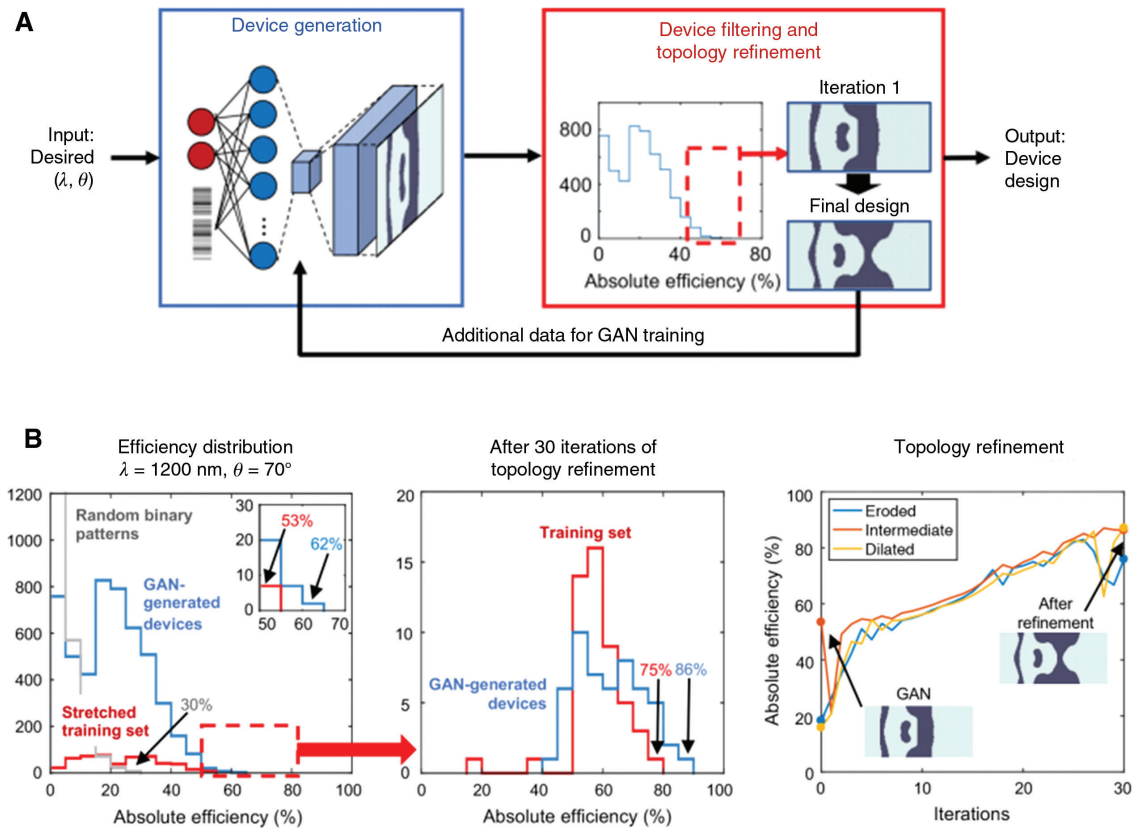

Figure 6: DL-assisted inverse design method combined with topology optimization.
(A) A schematic of the inverse design where a GAN generates training data. The generated designs are further optimized using topology optimization. (B) Metagrating generation and their topology optimization results. Efficiency histogram of the randomly generated pattern (left) and the refined pattern (middle), and the diffraction efficiency changes with the iterations (right). (A) and (B) are reprinted with permission from Ref. [73], Copyright 2019 American Chemical Society.

used to further improve the device efficiencies (Figure 6B) [73]. This method combines the advantages of both DL and topology optimization and allows for the inverse design of high-performance metasurfaces at a moderate computational cost. In addition, this strategy works well with relatively less training data, because the algorithms focus on only extracting the important features from the data, rather than predicting their associated optical properties. Accordingly, it substantially reduces the burden of creating huge training datasets. In this context, other recent work by Jiang and Fan [80] has shown that the direct incorporation of adjoint variable calculations into GANs enables finding the global optimal solution for high diffraction efficiency with relatively low computational cost. They introduced a physics-driven loss function, where adjoint variable calculations are incorporated. The gradient of the loss ultimately leads to find the maximum device efficiency. The proposed global optimization method was compared with a conventional adjoint-based topology optimization method [26], and through a statistical analysis found that the proposed method is able to create designs with higher diffraction efficiencies. Therefore, we can conclude that the incorporation of physics information has indeed extended the practical utilization of DL as an inverse design method.

3.3 Reinforcement learning in inverse design

The last three paradigms of machine learning is RL [81–83]. DeepMind's AlphaZero [84] and AlphaStar [85] are popular examples of this class of goal-oriented machine learning approaches. Those algorithms are able to play popular games such as chess, shogi and go, and have been even expanded to learn games just of them limited information, such in the case of AlphaStar [85]. Remarkably, after a few hours of training by playing games against itself, the agents were able to achieve a human level of competency, while only being told about the rules of the game [86].

The main idea of RL is based on training an agent to learn about the parameter space of an environment through its own experience, by means of combining exploration and exploitation with the maximization a given cumulative reward. This can be understood as, for instance, analogous to humans eating quality food, for a short-term reward we can enjoy the taste, and for a long-term reward we stay healthy. Short-term rewards can also be negative to discourage certain choices, akin to the bad taste of low-quality food. Interestingly, in contrast to some of the algorithms discussed in the previous sections, RL does not require the creation of an extensive dataset to train on, as the policy is learned through the experience of the rewards received by doing certain actions in certain states. The decisions are made sequentially using Markov decision processes [87]. Markovian approaches are ubiquitous in physics and can be summarized by this simple statement "Future is independent of the past given the present', i.e. the current state includes all the information that has been learned from the past states.

The key components of how RL works are summarized in Figure 7. We have provided a brief description of each of these components of other important concepts underlying RL and their connection to nanophotonic design problems:

- *Agent*: An agent is a component that take actions on the environment.
- *Actions*: Actions (A) are a set of possible ways that the agent can interact with the environment. In the inverse design in nanophotonics, they often correspond to changes in a physical parameter of the system (such as a geometrical parameter of material forming some part of the system). The actions are defined within the environment and can be limited in states where physical limits could be exceeded.
- *Environment*: The environment is the parameter space that the agent explores and learns about. This could be a set of physical dimensions, materials or incident angles, to name just a few examples.
- *State*: The state (S) is the situation in which the agent exists at a specific moment in time. In nanophotonics, this can be understood as the current set of parameters that describe a given design (such as the material, height and radius of a nanorod in a metasurface for example).
- *Reward*: The reward (R) is the feedback that the agent receives for taking a specific action in a specific state. These rewards are a way of evaluating the action taken by the agent in the given state. A good example of a reward would be the optical properties of the specific design, such as reflection, transmission or absorption.
- *Policy:* The policy (π) is the strategy that the agent learns about the environment. The agent uses it to determine what its next action will be.
- *Discount factor*: The discount factor (usually denoted as γ) is a real number between 0 and 1, which is multiplied by future rewards make those future rewards less fulfilling than immediate one. A discount factor of 1 would give future rewards the same worth as immediate ones, whereas a discount factor or 0 would only consider the immediate rewards. This is a hyperparameter of the algorithm that should be tuned for each application.
- *Value*: The value (V) is the expected long-term reward (including the discount factor) for the current state

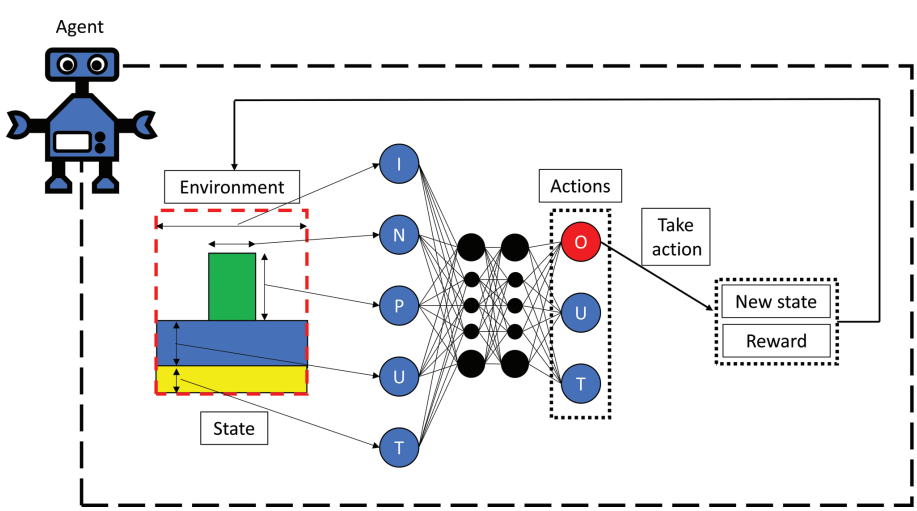

Figure 7: A schematic of the main features of reinforcement learning for nanostructure inverse design.

while using the policy π (it is usually denoted by $V\pi(s)$).

– *Q-value*: Similar to V, the Q-value takes a chosen action (a) an extra parameter into account. Specifically, the Q-value ($Q\pi(s, a)$) takes the current state (s) and the chosen action (a) under the policy π and maps the state-action pairs to rewards.

The Q-learning algorithm is able to handle problems featuring stochastic transitions and rewards, and was proven to converge to the optimum action-values [88]. In deep Q-learning (DQN), the Q values are approximated by a nonlinear function, such as the one of DNNs. However, this choice caused the learning stage to be unstable and even divergent in some cases. This issue mainly arose due to the associations between previous states and the sequences they had visited, for which a small update of the Q value could significantly change the policy and the links between the Q values and targets [89–91]. To overcome this important drawback, DeepMind introduced a technique known as "experience replay" [92]. In this a method a random sample of previous actions and states is used to proceed, instead of using the most recent ones, thereby removing correlations in the sequence of observations. To further reduce the correlations, the target values are not constantly updated, but periodically. A problem of deep Q-learning is that the chosen action is evaluated using the same policy as the Q-value estimation. This yields an overestimation of the action values, which hinders learning. To avoid this, double deep Q-learning (DDQN) was introduced in Ref. [93], and, due to its superior performance, became the basis of the RL applications in nanophotonics. In DDQN, two different models are used, one which evaluates the expected Q-value and another that chooses the next action. This increases the stability as a recent change in the policy will not affect the next chosen action. If the agent was to choose actions all the time, it would never be able to explore the parameter space and learn anything new, so to add a stochastic element to the choice of action, an epsilon greedy policy is usually used. A random number is generated and compared with the value of epsilon at that step. If the number is greater than the epsilon value at that step, then the agent uses the network to choose an action, otherwise an action is chosen at random [94].

In the first application of DDQN to nanophotonics put forward by Sajedian et al. [94], a previously reported color filter was further optimized to produce red, green, and blue (RGB) primary colors closer to the pure RGB colors than those previously reported in Ref. [95] (Figure 8). From a possible ~36.5 million states, the agent devised by Sajedian et al. was able to produce results superior to those in Ref. [95] within 9,000 steps. In particular, in their implementation the agent had a choice of 9 actions, related to changing one of the design parameters of the nanorod or the antireflective layer. The used reward system is shown in Figure 8C. As the goal was finding the closest representations of RGB, a simulation of the reflectance was

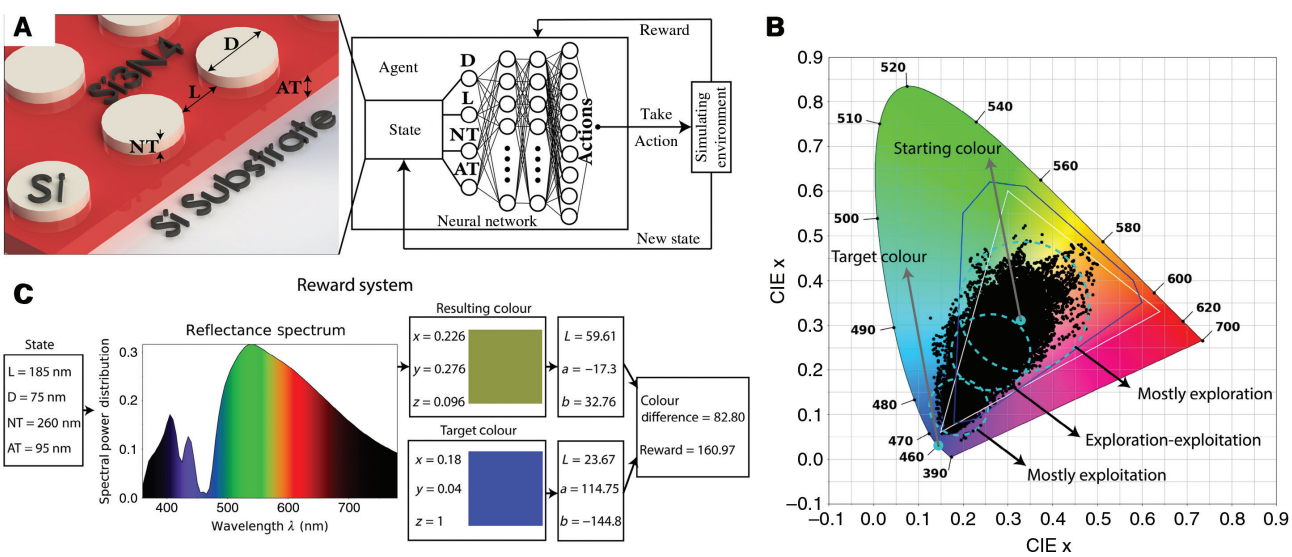

Figure 8: Reinforcement learning used to optimise the parameters of a dielectric color filter.
(A) A schematic of the nanostructure and DNN design. (B) The data points as the agent traverses the parameter space and reaches the goal of finding the bluest possible color. (C) The reward system based on the Lab color space color difference. Figures are adapted with permission from Ref. [95], The Optical Society (OSA).

undertaken for each state and action pair, and the color was calculated from the resulting spectrum. After converting the XYZ color values to the Lab color space, the reward was obtained from the color difference between the resulting color and the target. Notably, the applied approach produced colors with comparable or smaller color differences than those human researchers were able to achieve [96].

In the next example of RL performed by the same group, the authors used the same DDQN algorithm to design a highly efficient, transmission-type, polarisation-independent hologram [97]. In this case, the agent had 16 actions within the parameter space, which despite having over 5 billion possible states, was able to converge to a significant result within 2169 steps. Here, changing the material of the design was available in the parameters for the agent. This was done by including actions that can cycle through materials. The initial design included options for a thin film and a grating structure, but the agent determined that a structure with no grating and no film was deemed to be the most suitable, by choosing a design that set those parameters to 0. The reward system was structured in a way that prioritized rewards for generating the required phase properties, while giving smaller rewards for high efficiency, as a highly efficient device that does not cover the whole range of required phases cannot be used for a hologram. A terminal state was defined for a structure that produced a particular value of the reward of 700. This resulted in a hologram with 32% efficiency and a high-quality output, shown in Figure 9.

In the final example of the DDQN algorithm being used in nanophotonics, Badloe et al. [98] optimized the parameters of a moth-eye structure, a perfect absorber for a variety of materials (Figure 10A). Much in the same way a similar algorithm can be used to complete different Atari games, here each different materials of the structure are the counterparts of different games in the Atari games example. A common problem in RL is reward hacking by the agent [99]. In essence, this issue arises when the agent finds a way to exploit the reward system, usually by taking advantage of an unseen loophole to gain a definite cumulative reward much larger than that obtained by exploring more the parameter space. To stop this kind of behavior, the researchers introduced a reward system shaped in a way that the agent would get higher rewards for being closer to the target, as shown in Figure 10B. The target was finding a structure with an absorption of 90% as quickly as possible. To encourage this, at each time step the agent was given a negative reward of –10, with smaller negative rewards for structures with absorption between 85%

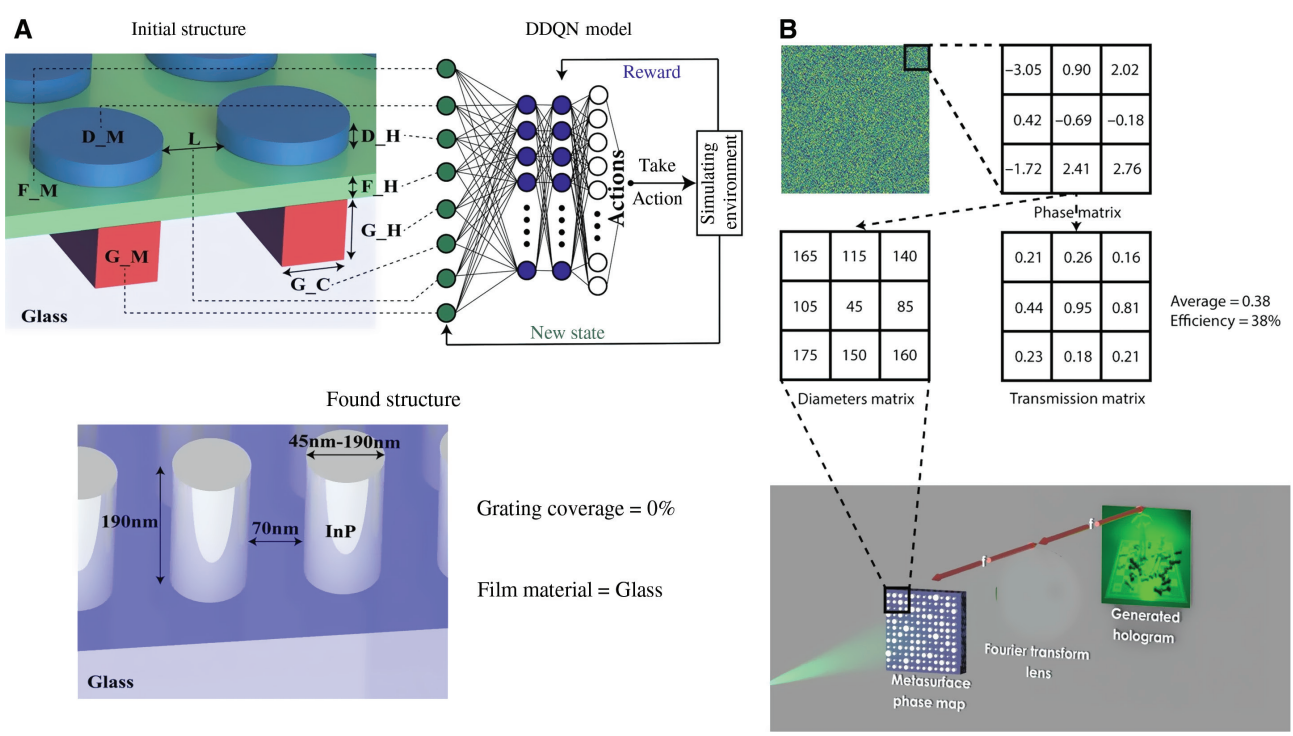

Figure 9: An example of reinforcement learning applied to optimizing a metahologram.
(A) A schematic of the initial and optimised structures. (B) The simulated results of the high efficiency metahologram. Figures are reprinted from [97] with permission (CC BY 4.0).

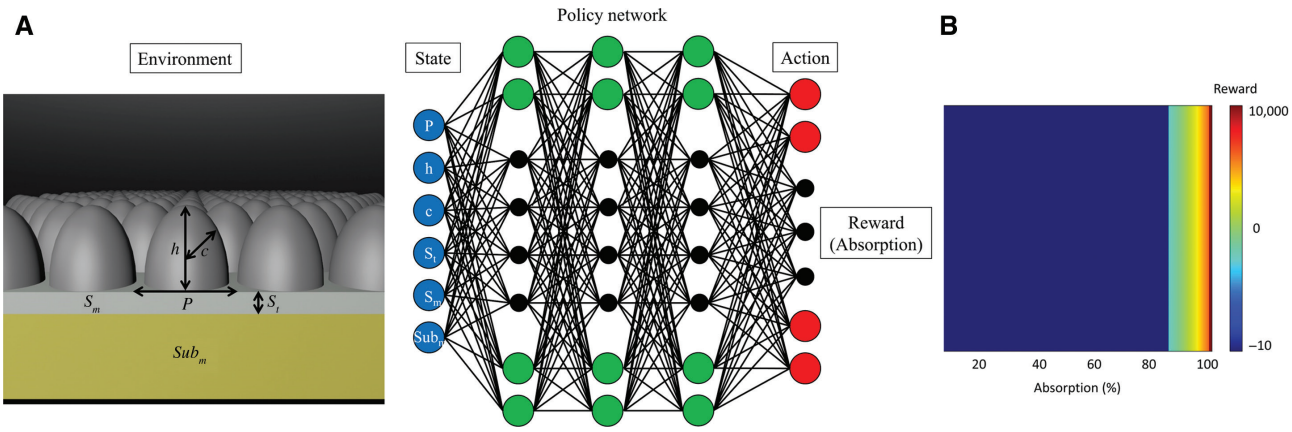

Figure 10: Reinforcement learning applied to designing ultra-broadband absorbers for a variety of transition metals. (A) A schematic of the moth-eye perfect absorber design and the policy network for the DDQN. (B) A visual representation of the shaped reward system. Figures are reprinted from Ref. [98] with permission (CC BY 4.0).

and 90%, and positive rewards for absorption over 90%. After being trained in an initial environment with a chromium (Cr) moth-eye structure, the same agent was used to optimize the other materials. This was done by setting the epsilon greedy policy to have a constant value of 0.1, which means that the agent would use the DNN to choose the action 90% of the time. From a parameter space of over 1 billion possible states, using this implementation, the agent was able to find structures with absorption over 90% for 6 different transition metals within just 100–200 steps.

4 Conclusions and outlook

In this review, we have summarized the recent progress of DL-assisted inverse design in nanophotonics. First, we discussed DL-based forward modelling, which shows how artificial intelligence can learn to solve Maxwell's equations without explicitly being informed about them. Then, we discussed state-of-the-art inverse design problems being solved by DL, categorizing them by the three different classes of learning methods: supervised, unsupervised, and RL. As discussed in this work, DL provides a new platform not only for approximating Maxwell's equations, but also for the inverse design of various nanophotonic devices that can by far exceed human capability. Although it has only recently been introduced into the field of nanophotonics, the fundamental change of paradigm introduced by DL, along with the tremendous potential it offers for the discovery of new nanophotonic devices and functionalities, is drawing increasing attention from a growing community of researchers worldwide.

There are, however, several important issues that must still be faced in this emerging area. Firstly, the solution given by a DL model for an inverse design problem is not guaranteed to be the most optimized, or the global solution (note that this fact could, on the other hand, be seen as a strength of DL-based inverse design, as it allows the rapid generation of number of solution candidates for the same problem). Secondly, DL inverse design methods feature some degree of dependence on human intuition, mainly because the basic considered shapes and problem settings are made based on previous known designs and physics. Thirdly, the design of "on-demand" structures is not always possible, especially if the design space of the training data is limited. Finally, a challenge that is actually general in the application of DL to scientific problems, is the fact that the learning mechanisms of DL are in most cases operating as black boxes, which in turn makes it difficult to exploit the trained network for further analysis. There is an emerging trend in artificial intelligence research trying to overcome this fundamental issue, but there is still a great deal of work to be done in this regard.

Overall, we envision that further advances in the above-mentioned directions could unleash the true potential of DL to nanophotonics, allowing for this approach to eventually become the main driver of the next-generation of significant discoveries in this field. We expect these novel findings will be beyond human intuition and imagination, and they will perhaps open a whole new perspective in our understanding of how nanophotonic research is carried out.

Funding: This work is financially supported by the National Research Foundation (NRF) grants (NRF-2019R1A2C3003129, CAMM-2019M3A6B3030637, NRF-2019R1A5A8080290 and NRF-2018M3D1A1058998) funded by the Ministry of Science and ICT, Republic of Korea. S.S acknowledges the global Ph.D. fellowship (NRF-2017H1A2A1043322) from the NRF-MSIT, Republic of Korea. J.B.A. acknowledges financial support from Ministerio de Ciencia, Innovacion y Universidades (RTI2018-098452-B-I00).

References

[1] Shen Y, Friend CS, Jiang Y, Jakubczyk D, Swiatkiewicz J, Prasad PN. Nanophotonics: interactions, materials, applications. J Phys Chem B 2000;104:7577–87.
[2] Tseng A A, Kuan C, Chen CD, Ma KJ. Electron beam lithography in nanoscale fabrication: recent development. IEEE Trans Elec Pack Manufac 2003;26:141–9.
[3] Yoon G, Kim I, So S, Mun J, Kim M, Rho J. Fabrication of three-dimensional suspended, interlayered and hierarchical nanostructures by accuracy-improved electron beam lithography overlay. Sci Rep 2017;7:6668.
[4] Oran D, Rodriques SG, Gao R, et al. 3D nanofabrication by volumetric deposition and controlled shrinkage of patterned scaffolds. Science 2018;362:1281.
[5] Prasad PN. Nanophotonics. John Wiley & Sons, 2004.
[6] Zhang Q, Yu H, Barbiero M, Wang B, Gu M. Artificial neural networks enabled by nanophotonics. Light Sci Appl 2019;8:42.
[7] Yu N, Genevet P, Kats MA, et al. Light propagation with phase discontinuities: generalized laws of reflection and refraction. Science 2011;334:333.
[8] Yu N, Capasso F. Flat optics with designer metasurfaces. Nat Mat 2014;13:139.
[9] Liu Y, Zhang X. Metamaterials: a new frontier of science and technology. Chem Soc Rev 2011;40:2494–507.
[10] Nguyen DM, Lee D, Rho J. Control of light absorbance using plasmonic grating based perfect absorber at visible and near-infrared wavelengths. Sci Rep 2017;7:2611.
[11] Maier SA. Plasmonics: fundamentals and applications. Berlin, Springer Science & Business Media, 2007.
[12] Wu Y, Luo Y, Chaudhari G, et al. Bright-field holography: cross-modality deep learning enables snapshot 3D imaging with bright-field contrast using a single hologram. Light Sci Appl 2019;8:25.
[13] Yoon G, Lee D, Nam KT, Rho J. Pragmatic metasurface hologram at visible wavelength: the balance between diffraction efficiency and fabrication compatibility. ACS Photonics 2018;5:1643–7.
[14] Ma Z, Li Y, Li Y, Gong Y, Maier SA, Hong M. All-dielectric planar chiral metasurface with gradient geometric phase. Opt Express 2018;26:6067.
[15] Lee H-E, Ahn H-Y, Mun J, et al. Amino-acid- and peptide-directed synthesis of chiral plasmonic gold nanoparticles. Nature 2018;556:360–5.
[16] Khanikaev AB, Hossein Mousavi S, Tse W-K, Kargarian M, MacDonald AH, Shvets G. Photonic topological insulators. Nat Mater 2012;12:233.
[17] Gao W, Lawrence M, Yang B, et al. Topological photonic phase in chiral hyperbolic metamaterials. Phys Rev Lett 2015;114:037402.
[18] Kim I, So S, Rana Ahsan S, Mehmood Muhammad Q, Rho J. Thermally robust ring-shaped chromium perfect absorber of visible light. Nanophotonics 2018;7:1827.
[19] Nguyen TT, Lim S. Wide incidence angle-insensitive metamaterial absorber for both TE and TM polarization using eight-circular-sector. Sci Rep 2017;7:3204.
[20] Zhang L, Zhou P, Lu H, Chen H, Xie J, Deng L. Ultra-thin reflective metamaterial polarization rotator based on multiple plasmon resonances. IEEE Antenn Wireless Propag Lett 2015;14:1157–60.
[21] Grady NK, Heyes JE, Chowdhury DR, et al. Terahertz metamaterials for linear polarization conversion and anomalous refraction. Science 2013;340:1304.
[22] Molesky S, Lin Z, Piggott AY, Jin W, Vucković J, Rodriguez AW. Inverse design in nanophotonics. Nat Photon 2018;12:659–70.
[23] Yao K, Unni R, Zheng Y. Intelligent nanophotonics: merging photonics and artificial intelligence at the nanoscale. Nanophotonics 2019;8:339.
[24] Sanchis L, Håkansson A, López-Zanón D, Bravo-Abad J, Sánchez-Dehesa J. Integrated optical devices design by genetic algorithm. Appl Phys Lett 2004;84:4460–2.
[25] Huntington MD, Lauhon LJ, Odom TW. Subwavelength lattice optics by evolutionary design. Nano Lett 2014;14:7195–200.
[26] Jensen JS, Sigmund O. Topology optimization for nano-photonics. Laser Photon Rev 2011;5:308–21.
[27] Hughes TW, Minkov M, Williamson IAD, Fan S. Adjoint method and inverse design for nonlinear nanophotonic devices. ACS Photon 2018;5:4781–7.
[28] Lalau-Keraly CM, Bhargava S, Miller OD, Yablonovitch E. Adjoint shape optimization applied to electromagnetic design. Opt Exp 2013;21:21693–701.
[29] Jafar-Zanjani S, Inampudi S, Mosallaei H. Adaptive genetic algorithm for optical metasurfaces design. Sci Rep 2018;8:11040.
[30] Cheng J, Inampudi S, Mosallaei H. Optimization-based dielectric metasurfaces for angle-selective multifunctional beam deflection. Sci Rep 2017;7:12228.
[31] Bodaghi M, Damanpack AR, Hu GF, Liao WH. Large deformations of soft metamaterials fabricated by 3D printing. Mater Des 2017;131:81–91.
[32] Sadeqi A, Rezaei Nejad H, Owyeung RE, Sonkusale S. Three dimensional printing of metamaterial embedded geometrical optics (MEGO). Microsys Nanoeng 2019;5:16.
[33] Reeves JB, Jayne RK, Barrett L, White AE, Bishop DJ. Fabrication of multi-material 3D structures by the integration of direct laser writing and MEMS stencil patterning. Nanoscale 2019;11:3261–7.
[34] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing System 25. Lake Tahoe, NV, USA, NIPS, 2012:1097–105.
[35] Shalev-Shwartz S, Shammah S, Shashua A. Safe, multi-agent, reinforcement learning for autonomous driving, 2016. Preprint arXiv:1610.03295.

[36] Deng L, Li X. Machine learning paradigms for speech recognition: an overview. IEEE Trans Audio Speech Lang Process 2013;21:1060–89.
[37] Min H. Artificial intelligence in supply chain management: theory and applications. Int J Log Res Appl 2010;13:13–39.
[38] Davis L. Handbook of genetic algorithms. New York, Van Nostrand Reinhold, 1991.
[39] Yosinski J, Clune J, Bengio Y, Lipson H. How transferable are features in deep neural networks? In: Advances in Neural Information Processing Systems 27. Montreal, Canada, NIPS, 2014:3320–8.
[40] Torrey L, Shavlik J. Transfer learning. In: Torrey L, Shavlik J, eds. Handbook of research on machine learning applications and trends: algorithms, methods, and techniques. Hershey, PA, USA, IGI Global, 2010:242–64.
[41] Petschulat J, Menzel C, Chipouline A, et al. Multipole approach to metamaterials. Phys Rev A 2008;78:043811.
[42] Gallinet B, Butet J, Martin OJF. Numerical methods for nanophotonics: standard problems and future challenges. Laser Photon Rev 2015;9:577–603.
[43] Katsidis CC, Siapkas DI. General transfer-matrix method for optical multilayer systems with coherent, partially coherent, and incoherent interference. Appl Opt 2002;41:3978–87.
[44] Troparevsky MC, Sabau AS, Lupini AR, Zhang Z. Transfer-matrix formalism for the calculation of optical response in multilayer systems: from coherent to incoherent interference. Opt Exp 2010;18:24715–21.
[45] Moharam MG, Gaylord TK. Rigorous coupled-wave analysis of planar-grating diffraction. J Opt Soc Am 1981;71:811–8.
[46] Moharam MG, Grann EB, Pommet DA, Gaylord TK. Formulation for stable and efficient implementation of the rigorous coupled-wave analysis of binary gratings. J Opt Soc Am A 1995;12:1068–76.
[47] Monk P. Finite element methods for Maxwell's equations. Oxford, Oxford University Press, 2003.
[48] Joseph RM, Taflove A. FDTD Maxwell's equations models for nonlinear electrodynamics and optics. IEEE Trans Antenn Propag 1997;45:364–74.
[49] Peurifoy J, Shen Y, Jing L, et al. Nanophotonic particle simulation and inverse design using artificial neural networks. Sci Adv 2018;4:eaar4206.
[50] Qu Y, Jing L, Shen Y, Qiu M, Soljačić M. Migrating knowledge between physical scenarios based on artificial neural networks. ACS Photon 2019;6:1168–74.
[51] Balin I, Garmider V, Long Y, Abdulhalim I. Training artificial neural network for optimization of nanostructured VO2-based smart window performance. Opt Exp 2019;27:A1030–40.
[52] Baxter J, Calà Lesina A, Guay J-M, Weck A, Berini P, Ramunno L. Plasmonic colors predicted by deep learning. Sci Rep 2019;9:8074.
[53] Kiarashinejad Y, Zandehshahvar M, Abdollahramezani S, Hemmatyar O, Pourabolghasem R, Adibi A. Knowledge discovery in nanophotonics using geometric deep learning. Adv Intell Syst 2019;1900132. Available at: https://doi.org/10.1002/aisy.201900132.
[54] 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. Microsys Nanoeng 2019;5:27.
[55] Inampudi S, Mosallaei H. Neural network based design of metagratings. Appl Phys Lett 2018;112:241102.
[56] Nadell CC, Huang B, Malof JM, Padilla WJ. Deep learning for accelerated all-dielectric metasurface design. Opt Exp 2019;27:27523–35.
[57] Lawrence S, Giles CL, Tsoi AC, Back AD. Face recognition: a convolutional neural-network approach. IEEE Trans Neural Networks 1997;8:98–113.
[58] Mikolov T, Karafiát M, Burget L, Černocký J, Khudanpur S. Recurrent neural network based language model. In: Eleventh annual conference of the international speech communication association 11, Makuhari, Japan, INTERSPEECH, 2010:1045–8.
[59] Liu D, Tan Y, Khoram E, Yu Z. Training deep neural networks for the inverse design of nanophotonic structures. ACS Photon 2018;5:1365–9.
[60] Liu Z, Zhu D, Rodrigues SP, Lee K-T, Cai W. Generative model for the inverse design of metasurfaces. Nano Lett 2018;18:6570–6.
[61] So S, Mun J, Rho J. Simultaneous inverse design of materials and structures via deep learning: demonstration of dipole resonance engineering using core–shell nanoparticles. ACS Appl Mater Interf 2019;11:24264–8.
[62] Caruana R, Niculescu-Mizil A. An empirical comparison of supervised learning algorithms. In: Proceedings of the 23rd International Conference on Machine Learning. New York, NY, USA, ACM, 2006:161–8.
[63] Malkiel I, Mrejen M, Nagler A, Arieli U, Wolf L, Suchowski H. Plasmonic nanostructure design and characterization via deep learning. Light Sci Appl 2018;7:60.
[64] Asano T, Noda S. Optimization of photonic crystal nanocavities based on deep learning. Opt Exp 2018;26:32704–17.
[65] Hemmatyar O, Abdollahramezani S, Kiarashinejad Y, Zandehshahvar M, Adibi A. Full color generation with Fano-type resonant HfO2 nanopillars designed by a deep-learning approach. Nanoscale 2019;11:21266–74.
[66] Wu B, Ding K, Chan C, Chen Y. Machine prediction of topological transitions in photonic crystals, 2019. Preprint arXiv:1907.07996.
[67] Pilozzi L, Farrelly FA, Marcucci G, Conti C. Machine learning inverse problem for topological photonics. Commun Phys 2018;1:57.
[68] Ma W, Cheng F, Liu Y. Deep-learning-enabled on-demand design of chiral metamaterials. ACS Nano 2018;12:6326–34.
[69] Barlow HB. Unsupervised learning. Neural Comput 1989;1:295–311.
[70] Hofmann T. Unsupervised learning by probabilistic latent semantic analysis. Mach Learn 2001;42:177–96.
[71] So S, Rho J. Designing nanophotonic structures using conditional deep convolutional generative adversarial networks. Nanophotonics 2019;8:1255–61.
[72] 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 2019;31:1901111.
[73] Jiang J, Sell D, Hoyer S, Hickey J, Yang J, Fan JA. Free-form diffractive metagrating design based on generative adversarial networks. ACS Nano 2019;13:8872–8.
[74] An S, Zheng B, Tang H, et al. Generative multi-functional metaatom and metasurface design networks, 2019. Preprint arXiv:1908.04851.
[75] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets. In: Advances in Neural Information Processing Systems 27, Montreal, Canada, NIPS, 2014:2672–80.

[76] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks, 2015. Preprint arXiv:1511.06434.
[77] Chapelle O, Scholkopf B, Zien A. Semi-supervised learning (Chapelle, O. et al., eds.; 2006)[book reviews]. IEEE Trans Neural Netw 2009;20:542.
[78] Doersch C. Tutorial on variational autoencoders, 2016. Preprint arXiv:1606.05908.
[79] Higgins I, Matthey L, Pal A, et al. beta-VAE: Learning basic visual concepts with a constrained variational framework. In: 5th International conference on learning representataions, Vol. 2, Toulon, France, ICLR, 2017:6.
[80] Jiang J, Fan JA. Global optimization of dielectric metasurfaces using a physics-driven neural network. Nano Lett 2019;19:5366–72.
[81] Sutton RS, Barto AG. Introduction to reinforcement learning. Cambridge: MIT Press, 1998.
[82] François-Lavet V, Henderson P, Islam R, Bellemare MG, Pineau J. An introduction to deep reinforcement learning. Found Trends Mach Learn 2018;11:219–354.
[83] van Otterlo M, Wiering M. Reinforcement learning and Markov decision processes. In: van Otterlo M, Wiering M, editors, Reinforcement learning. Springer, 2012;3–42.
[84] Silver D, Hubert T, Schrittwieser J, et al. A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. Science 2018;362:1140.
[85] Vinyals O, Babuschkin I, Czarnecki WM, et al. Grandmaster level in StarCraft II using multi-agent reinforcement learning. Nature 2019;575:350–4.
[86] Mnih V, Kavukcuoglu K, Silver D, et al. Human-level control through deep reinforcement learning. Nature 2015;518: 529–33.
[87] Puterman ML. Markov decision processes: discrete stochastic dynamic programming. Hoboken, NJ, John Wiley & Sons, 2014.
[88] Watkins CJ, Dayan P. Q-learning. Mach Learn 1992;8:279–92.
[89] Baird L. Residual algorithms: reinforcement learning with function approximation. In: Baird L, editor, Machine Learning Proceedings. Elsevier, 1995;30–7.
[90] Henderson P, Islam R, Bachman P, Pineau J, Precup D, Meger D. Deep reinforcement learning that matters. In: Thirty-Second AAAI Conference on Artificial Intelligence. New Orleans, LA, USA, AAAI, 2018:3207–14.
[91] Nikishin E, Izmailov P, Athiwaratkun B, et al. Improving stability in deep reinforcement learning with weight averaging. In: Uncertainty in artificial intelligence workshop on uncertainty in Deep learning, 2018. (Accessed January 29, 2020, at https://izmailovpavel.github.io/files/swa_rl/paper.pdf).
[92] Mnih V, Kavukcuoglu K, Silver D, et al. Playing atari with deep reinforcement learning, 2013. Preprint arXiv:1312.5602.
[93] Van Hasselt H, Guez A, Silver D. Deep reinforcement learning with double q-learning. In: Thirtieth AAAI Conference on artificial intelligence. Phoenix, AZ, USA, AAAI, 2016:2094–100.
[94] Kuleshov V, Precup D. Algorithms for multi-armed bandit problems, 2014. Preprint arXiv:1402.6028.
[95] Sajedian I, Badloe T, Rho J. Optimisation of color generation from dielectric nanostructures using reinforcement learning. Opt Exp 2019;27:5874–83.
[96] Dong Z, Ho J, Yu YF, et al. Printing beyond sRGB color Gamut by mimicking silicon nanostructures in free-space. Nano Lett 2017;17:7620–8.
[97] Sajedian I, Lee H, Rho J. Double-deep Q-learning to increase the efficiency of metasurface holograms. Sci Rep 2019;9:10899.
[98] Badloe T, Kim I, Rho J. Biomimetic ultra-broadband perfect absorbers optimised with reinforcement learning. Phys Chem Chem Phys 2020;22:2337–42.
[99] Mallozzi P, Pardo R, Duplessis V, Pelliccione P, Schneider G. MoVEMo: a structured approach for engineering reward functions. In: 2018 Second IEEE International Conference on Robotic Computing. Laguna Hills, CA, USA, IRC, 2018:250–7.