



PHOTONICS Research

Deep learning in nano-photonics: inverse design and beyond

PETER R. WIECHA,^{1,*} ARNAUD ARBOUET,^{2,4} CHRISTIAN GIRARD,^{2,5} AND OTTO L. MUSKENS^{3,6}

¹LAAS, Université de Toulouse, CNRS, Toulouse, France

²CEMES, Université de Toulouse, CNRS, Toulouse, France

³Physics and Astronomy, Faculty of Engineering and Physical Sciences, University of Southampton, Southampton, UK

⁴e-mail: arbouet@cemes.fr

⁵e-mail: girard@cemes.fr

⁶e-mail: o.muskens@soton.ac.uk

*Corresponding author: pwiecha@laas.fr

Received 2 December 2020; revised 11 January 2021; accepted 27 January 2021; posted 29 January 2021 (Doc. ID 415960); published 14 April 2021

Deep learning in the context of nano-photonics is mostly discussed in terms of its potential for inverse design of photonic devices or nano-structures. Many of the recent works on machine-learning inverse design are highly specific, and the drawbacks of the respective approaches are often not immediately clear. In this review we want therefore to provide a critical review on the capabilities of deep learning for inverse design and the progress which has been made so far. We classify the different deep-learning-based inverse design approaches at a higher level as well as by the context of their respective applications and critically discuss their strengths and weaknesses. While a significant part of the community's attention lies on nano-photonics inverse design, deep learning has evolved as a tool for a large variety of applications. The second part of the review will focus therefore on machine learning research in nano-photonics “beyond inverse design.” This spans from physics-informed neural networks for tremendous acceleration of photonics simulations, over sparse data reconstruction, imaging and “knowledge discovery” to experimental applications.

Published by The Optical Society under the terms of the [Creative Commons Attribution 4.0 License](#). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

<https://doi.org/10.1364/PRJ.415960>

1. INTRODUCTION

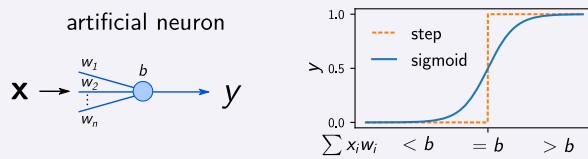
Light-matter interaction at sub-wavelength dimensions can lead to astonishing effects such as localized surface plasmon resonances which concentrate light to deeply sub-wavelength volumes [1], the appearance of optical magnetic resonances in otherwise non-magnetic media [2], the possibility to shape optical near-fields with sub-wavelength structure [3], the emergence of non-linear optical phenomena [4], or strong enhancement of quantum emitter luminescence [5], to name just a few. Those nano-scale optical effects can be exploited for a broad variety of applications, for instance in integrated quantum optics [6], for metamaterials [7], and in this context specifically for metasurfaces like flat lenses [8]. It is, for example, even possible to create all-optical devices which use light to solve integral equations or perform other analog optical computing tasks [9–11].

Still, ever since the advent of nano-optics with the invention of near-field microscopy [12–14] the numerical description of many problems continues to be challenging [15]. An example is the rational design of nano-photonic structures for specific

tasks, which remains a general problem that often involves brute force “forward” calculations or solving inverse scattering problems. Other challenges in nano-optics are related to experimental limitations such as the stochastic nature of single-photon emitters, fluctuating nano-scale force fields such as Brownian motion, and the diffraction limit blocking access to sub-wavelength information. Such effects often complicate the interpretation of nano-optics experiments and require the use of more sophisticated techniques for data analysis, for example, combining data with prior knowledge or sparsity constraints. All these obstacles are about to be pushed significantly further by the emerging computational methods around machine learning. In particular, “deep learning,” a sub-field of machine learning which uses complex ANNs with millions of ANs, recently emerged as a versatile and powerful numerical tool [16,17]. Deep learning techniques have proven to be particularly good at the categorization of huge and complex datasets, a task that they perform radically differently compared to classical algorithms. Following a rather “intuitive” approach, ANNs mimic the working principle of biological neurons

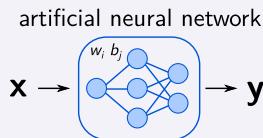
Box 1. Artificial neurons, neural networks, and their training

An artificial neuron (AN) is simply a mathematical function which mimics the behavior of a biological neuron.



The step-like behavior of neuronal activation, which starts to fire once a threshold stimulation is exceeded, can be implemented by various non-linear mathematical functions. A popular example is the logistic function (also called “sigmoid”), shown in the above sketch. If the scalar product of an input vector \mathbf{x} and the neuron-intrinsic weight parameters w_i are larger than the neuron’s bias parameter b , the output y is “high” (the artificial neuron fires). Else it is “low.”

An artificial neural network (ANN) is composed of several of such ANs, usually arranged in “layers.” The output value of a neuron is fed into a succeeding layer of neurons. The final layer is the network output \mathbf{y} . For instance, in a so-called fully connected ANN, every neuron of one layer is connected to every neuron of the following layer.



Hence, an ANN represents a vectorial function $f(\mathbf{x}) = \mathbf{y}$ characterized by a large number of parameters w_i and b_j .

Training an ANN is done via a numerical minimization of a loss function L , which describes the error of the

and the human brain. A brief overview of the basic concepts is given in Box 1.

Research in medicine is often of statistical nature, for which data-driven analysis methods such as deep learning are particularly interesting. Consequently, one of the first scientific fields to which deep learning methods have been extensively applied was medical research. In medical diagnostics, especially medical imaging such as radiology, the use of machine learning techniques for analysis and interpretation has literally exploded in the recent past, which has led to extraordinary successes with diagnostic classification accuracies often far beyond human performance [18,19].

In nano-optics and photonics, machine learning started to emerge a little later, but recently celebrated some remarkable breakthroughs, enabling the analysis, categorization, and interpretation of data which seemed formerly impossible. While already back in the 1990s simple ANNs had been discussed and used for applications in spectroscopy or for automated instrumental control, for instance, to counteract drifts in microscopy [20], it took two decades before the available computational

power reached a level that deep ANNs with millions or even up to hundreds of billions of free parameters [21] could be successfully trained on formerly unsolved problems. Today, deep learning models have evolved to an extent that they readily outperform humans on specialized tasks such as image recognition [16,22]. This progress was possible especially thanks to the rapid development of massive parallel computing architectures in modern graphics processing units (GPUs), and lately of specific “tensor cores,” integrated logic circuits optimized for the mathematical matrix operation tasks required for neural network training. Even all-optical implementations of artificial neural networks have been subject to recent research; however, their performance is still limited by the lack of energy-efficient all-optical non-linear units [23–25].

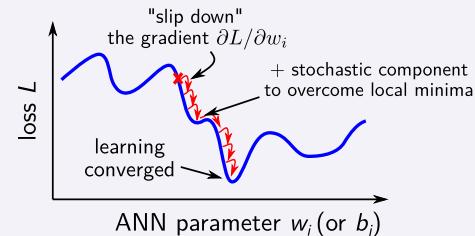
Several review articles have been published recently, which categorize in great detail the latest developments of deep learning applications in photonics and nano-optics. For an exhaustive overview we therefore invite the reader to consult these articles [26–30]. Also a few thematically more distantly related review articles have been published recently, which we want to

$$L(w_i, b_j) = \frac{1}{N} \sum_{l=1}^N [\mathbf{y}_{\text{train},l} - \mathbf{y}_{\text{ANN}}(\mathbf{x}_{\text{train},l})]^2, \quad (1)$$

where $\mathbf{x}_{\text{train}}$ and $\mathbf{y}_{\text{train}}$ are N random samples from the training data, and \mathbf{y}_{ANN} are the network predictions corresponding to $\mathbf{x}_{\text{train}}$. N is called the batch size.

The term “learning” refers to optimizing the parameters w_i and b_j describing the ANN, with the goal to minimize the loss L . A small loss means that the network output approximates well the training data, ideally by learning to “understand” the underlying correlations. L is numerically minimized by “slipping down” its gradient with respect to the parameters w_i and b_j .

training: minimize loss function $L(w_i, b_j)$



Training on small batches composed of random subsets of N training samples helps to “jump” out of local minima by adding a stochastic component to the procedure. One of the most common training algorithms is stochastic gradient descent [17].

indicate to the interested reader. They cover, for example, conventional inverse design and optimization methods for metasurfaces [31] and nano-photonics [32], but also a few more general reviews on artificial intelligence in nano-technology, photonics, and for light–matter interaction have been published [33–36]. Finally, for the sake of conciseness of this review, we intentionally ignore the vast and very active research field on hardware implementations of artificial neural networks, which includes—but is not limited to—research efforts on photonics platforms [23,37,38].

In this mini-review we focus on selected key results that have recently led to breakthrough advancements in the research on inverse design of photonic nano-structures and metasurfaces. Rather than compiling an exhaustive catalog of every single publication, we provide an overview of milestone concepts for improving deep learning inverse design fidelity, which recently allowed to bring ANNs closer to the performance of conventional optimization methods. We believe that such a summary of concepts is of particular interest for researchers in the field. We dedicate the second part of the review to an overview of original applications of deep learning in nano-photonics beyond structural inverse design. Specifically, we summarize recent developments around physics-informed neural networks in optics, on deep learning for knowledge discovery and explainable machine learning, as well as on applications of ANNs to nano-photonics experiments.

2. DEEP-LEARNING-BASED NANO-PHOTONICS INVERSE DESIGN

The first part of this mini-review is dedicated to deep-learning-based inverse design techniques as well as to concepts to improve the inverse design model fidelity. As stated before, we do not aim to provide an exhaustive list of applications. An up-to-date and very complete overview of possible optimization targets can be found, for instance, in the recent reviews by Ma *et al.* [27] or by Jiang *et al.* [29].

A. “Conventional” Inverse Design Methods

Before the recent rise of deep learning methods, inverse design of nano-photonic structures was often based on intuitive considerations and systematic fine-tuning (see, e.g., Refs. [39,40]). A more systematic alternative was the combination of numerical simulation methods with gradient-based or heuristic optimization algorithms, such as stimulated annealing, topology optimization, and genetic algorithms [32,41–44]. Such methods led to some remarkable success for instance in the optimization of plasmonic optical antennas [45,46], dielectric multi-functional nano-structures [47], and metasurfaces [31,48]. A great advantage of such methods is the possibility to include fabrication constraints or robustness conditions in the optimization procedure [47,49].

However, heuristics coupled to numerical simulation techniques is slow and computationally expensive. Furthermore, for each new optimization target, the parameter space needs to be searched from scratch, implying hundreds to thousands of numerical simulations. The recent advent of data-driven techniques such as deep learning holds promise to accelerate the computation by many orders of magnitude and quite some re-

markable progress has been made in the past few years. One can distinguish two types of approach that have gained traction. The first one replaces the forward simulation in an iterative optimization with an ANN, while the second aims to build an inverse ANN that solves the problem directly. Below we critically discuss the two approaches as well as efforts at improving the quality of results.

B. Surrogate-Model-Based Inverse Design

Deep learning models are particularly strong in predicting approximate solutions to direct problems such as the optical response of photonic structures. A possible approach to accelerate inverse design is therefore to use a “forward neural network” as an ultra-fast predictor together with an optimization technique. In such a case the ANN acts as a so-called surrogate model, taking the place of the much slower conventional simulation method.

1. Deep Learning Forward Solver

ANNS have been successfully trained on the prediction of various physical quantities in nano-photonics. Early works have proposed ANNs to create phenomenological models of non-linear optical effects or of optical ionization using experimental training data [50,51]. Recently, the idea has been picked up and it has been shown, for instance, that scattering and extinction spectra can be predicted with high accuracy [52] and also that the phase can be included in the predictions [53], which is important for nano-structures in metasurfaces. The prediction of far-field observables can also be extended to include proximity effects in a dense metasurface, beyond the local phase approximation. The latter has been demonstrated by including the near-field interactions with the nearest neighbor structures in the training data [54]. The prediction of physical effects is not limited to extinction, transmission, or other far-field effects. It has been shown that also near-field effects can be approximated accurately, for instance, around nano-wires of complex shape [55].

While networks that predict an observable such as the scattering cross section perform usually very well within the range of their training data, such models often generalize rather poorly to cases outside the parameter range covered by the training data. The ANN acts then as universal function approximator, but it does not develop a deeper “understanding” of the underlying physics. In order to alleviate this problem, it turned out to be helpful to provide the network with pre-processed data. For instance, instead of training an ANN with pure optical extinction spectra, So *et al.* [56] trained their model using a decomposition in multiple electric and magnetic dipole resonances to predict the optical response of multi-material multi-shell nano-spheres. The approach is illustrated in Fig. 1(a) and has also been used to inverse design multi-shell spheres for Kerker-type directional scattering. Using a metallic grating as a model example, Blanchard-Dionne and Martin demonstrated that a neural network that learns light–matter interaction through a representation as multiple Lorentz oscillators generalizes about an order of magnitude better outside the training data range, compared to a predictor network based on the raw optical spectrum [57] [see Fig. 1(b)]. Instead of predicting specific physical observables such as the extinction cross

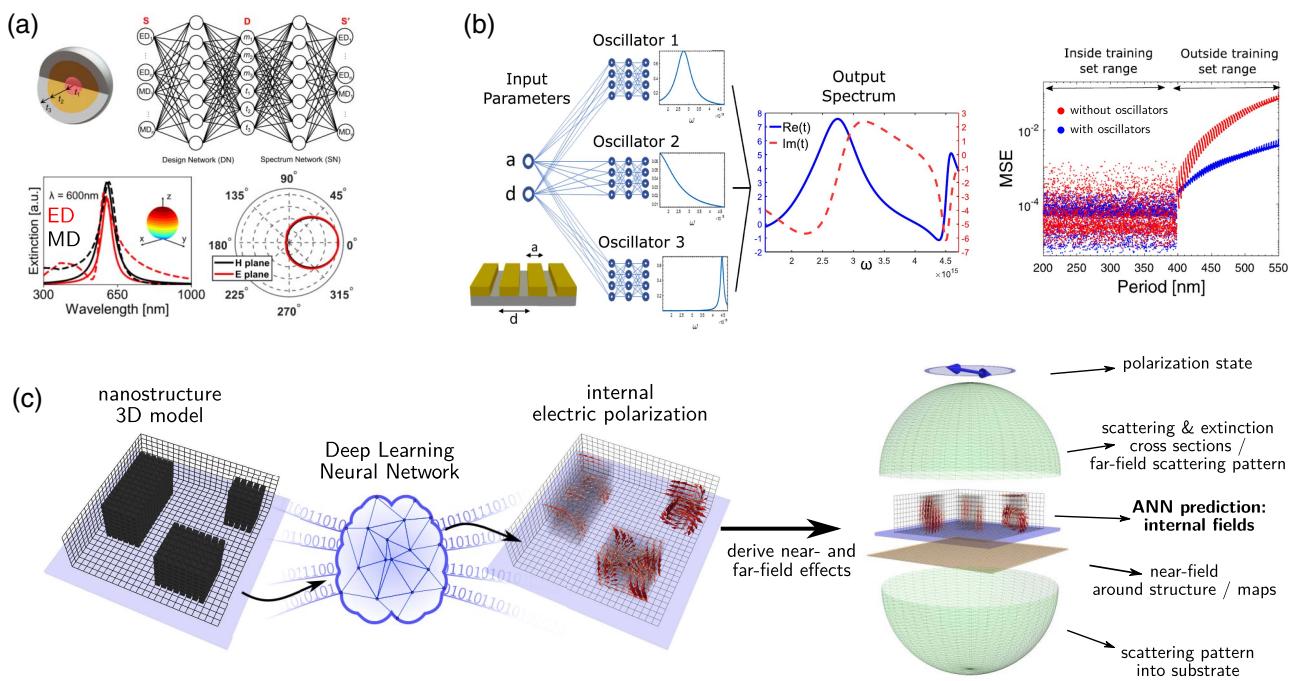


Fig. 1. Deep-learning-based forward solvers for ultra-fast physics predictions. (a) Simultaneous electric and magnetic dipole resonance prediction and inverse design in multi-layer nano-spheres. Adapted with permission from [56], copyright (2019) American Chemical Society. (b) Nano-optics solver network, which predicts the optical response of a grating based on multiple Lorentz oscillators. As shown in the right panel, the physics-based data representation allows the network to generalize well outside the range of the training data (blue points). Adapted with permission from [57], copyright (2020) Optical Society of America. (c) Internal electric polarization density predictor network. The results can be used in a coupled dipole approximation framework to calculate a large number of secondary near- and far-field effects. Adapted with permission from [58], copyright (2020) American Chemical Society.

section, Wiecha *et al.* demonstrated that a network can learn a discrete dipole approximation of the electric polarization density inside a 3D nano-structure of arbitrary shape [58]. The concept is depicted in Fig. 1(c) and allows accurate derivation of manifold secondary quantities in the near and far fields from a single generalized predictor neural network.

2. Forward Predictor Networks + Evolutionary Optimization
In general, the greatest advantage of deep learning techniques as surrogate models for physics simulations is their tremendous evaluation speed. Once trained, an ANN delivers its prediction within fractions of milliseconds, which is usually orders of magnitude faster than a numerical simulation. Therefore, replacing conventional physics simulations by surrogate ANNs is a natural solution to speed-up the inverse design of photonic nano-structures via global optimization heuristics [59,60]. This concept has recently been applied by several groups to the design of individual photonic nano-structures or metasurfaces [61–66].

However, while the approach can significantly accelerate heuristics-based inverse design, it remains an iterative approach requiring thousands of calls to the surrogate model as well as intermediate computation steps. Furthermore, the surrogate model represents only an approximation to the physical reality, introducing a systematic error. And even worse than that, it cannot be guaranteed that the surrogate model does not contain singular points of totally false solutions [67], to which the optimization algorithm may converge in the worst-case scenario.

Robust implementations therefore require a simulation-based fine-tuning procedure subsequent to the surrogate-based optimization run, which often relativizes the gain in speed [68,69]. The same problem also holds, of course, for the here-after discussed ANN-only inverse design methods.

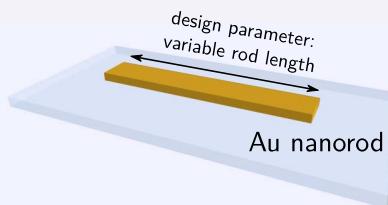
C. Direct Neural Network Inverse Design

As mentioned above, using forward ANNs as surrogate models for evolutionary optimization is computationally not the most efficient technique and bears the risk of converging to singular points of the surrogate model. In the recent past tremendous efforts have therefore been dedicated to the development of exclusively ANN-based inverse design schemes. The main obstacle which needs to be circumvented is the so-called “one-to-many” problem, which describes the fact that most inverse design problems are ambiguous, and hence several non-unique solutions exist for the same design target. In consequence a naive inversion of the ANN layout usually fails [70], but several solutions have been developed to tackle the one-to-many problem. One possibility is the above-described technique to use a forward network as surrogate model, coupled to a global optimization algorithm. In this section we give a brief overview of pure neural network models to solve non-unique inverse problems. The different concepts are also schematized in Box 2.

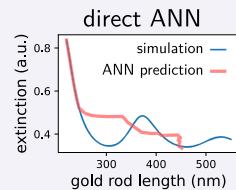
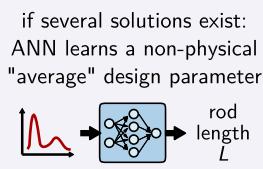
A popular type of a stable inverse design network is the so-called tandem network architecture [52,56,70–72]. In a tandem ANN a forward solver network is trained in a first step.

Box 2. Inverse design: The one-to-many problem

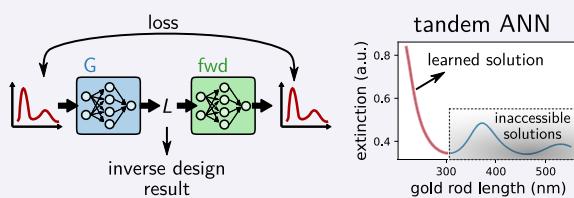
Let us assume a simple toy problem. Under fixed wavelength illumination, we want to tailor the extinction coefficient of a gold nano-rod by varying its length.



Already this simple problem is ambiguous: several rod lengths can lead to the same extinction, which makes a naive ANN implementation fail in those cases.



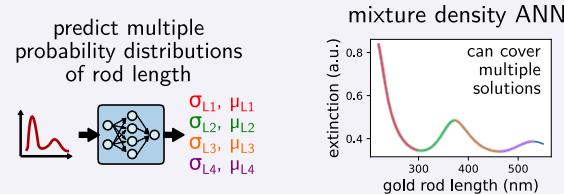
The “tandem neural network” can stabilize the generator (G) via a physics loss, based on a pre-trained forward model (fwd). This approach, however, limits the inverse design to one solution per design target, rendering inaccessible possible multiple solutions to a given problem.



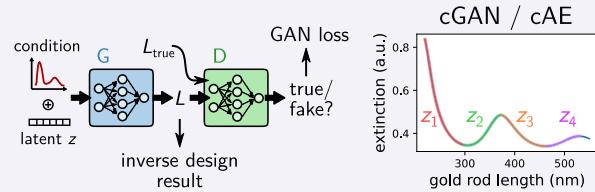
The training of the actual inverse design network (the generator) subsequently uses the fixed pre-trained forward model as a physics predictor to evaluate the inverse design output. In consequence, the loss function does not compare ambiguous design layouts but operates in the physics domain (comparing, e.g., the extinction efficiency rather than the design parameters). In this way, different design parameters which lead to a similar physical response no longer confuse the ANN, and all correct solutions to a given design problem yield a positive training feedback.

Another model that circumvents the one-to-many problem is the cGAN [68,73–76]. A cGAN takes as input not only the design target but also an additional “latent vector,” which is a normally distributed sequence of random values. The network then learns to use different values of the latent vector to address the distinct non-unique solutions. In addition to the introduction of a latent vector, a further peculiarity of cGANs is their loss function, which is a discriminator network that tries to distinguish generated solutions from real ones, and which is also subject to training. During training, the cGAN loss function

Mixture density networks predict several possible solutions at a time including their respective importance as Gaussian distributions. A disadvantage is that the maximum number of possible simultaneous solutions needs to be known.



Conditional generative adversarial networks (cGANs) or conditional (adversarial) autoencoders add a normally distributed latent vector (usually “z”) to the design target (“condition”), which encodes dynamically multiple possible solutions. Adversarial models furthermore use a trained loss, a so-called discriminator network (D), which tries to distinguish generated (fake) from training samples (true).

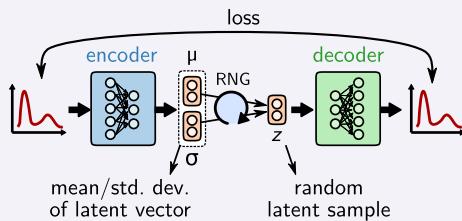


hence evolves together with the ANN, which allows ideally a better convergence. It is worth noting that it is a delicate task to tune the network and training hyperparameters in GANs such that the learning converges. The training of both the generator network and discriminator network needs to evolve in a balanced way for the adversarial loss function to work efficiently.

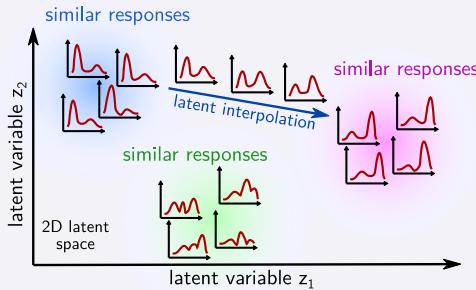
A further type of one-to-many solving networks is conditional adversarial or conditional variational autoencoders [66,77–80]. Those are usually symmetric models that take the physical response as input, which they try to identically reconstruct at their output layer. In a conditional autoencoder, a bottleneck layer is placed in the ANN center. This bottleneck contains the design parameters on the one hand (as it is the case in a tandem network), but on the other hand an additional latent vector is appended to the design parameters. Like in the cGAN, the latent vector can be used by the ANN to address potential multiple solutions. Unlike in the tandem network the forward model is trained simultaneously with the generator. Conditional autoencoders can be seen as a mixture of a tandem

Box 3. Variational autoencoders and the latent space

Variational autoencoders (VAEs) learn to compress information in a lower-dimensional latent space, by being trained on a reconstruction task.



In a VAE, forward propagation uses a random number generator (RNG) to draw samples z with mean value μ and standard deviation σ . The random component ensures that the learned latent variables z follow a normal distribution. However, gradient descent training requires analytical gradients, which cannot be backpropagated through the RNG. This is why a re-parametrization into deterministic layers of μ and σ is necessary [81].



By forcing the latent variables on a normal distribution, the trained VAE clusters similar inputs in the latent space. Furthermore, transitions between solutions in latent space are smooth, which allows, for example, interpolation operations.

network and a cGAN. For a short explanation of the basic idea behind VAEs and the meaning of the latent space, see also Box 3.

For completeness we want to mention also work on reinforcement learning for iterative design optimization, where the neural network learns to behave as an iterative optimization algorithm. The expectation is that the ANN can adapt its optimization strategy specifically to the given problem and hence outperform conventional heuristic algorithms [82,83].

The above discussed models have been quite successfully used for manifold inverse design problems in nano-photonics. Figure 2(a) shows an example of multi-mode interference devices (MMIs) designed by a tandem ANN. MMIs are large waveguides that support many modes and that can have several inputs and outputs (here 3×3). The here shown MMIs are patterned with small perturbations in order to obtain specific light-routing properties. The tandem ANN has been trained to design perturbation patterns which produce arbitrary transmission states. This allows, for instance, to define MMI patterns which swap a pair of the 3×3 input and output paths, while one of the transmission channels remains constant, as demonstrated in panels (i) and (ii) in Fig. 2(a) [84]. Figure 2(b) shows a metasurface which acts as a flat lens with two focal spots, designed by a variant of a cGAN network [73]. Other examples are the design of chiral plasmonic structures [71,85], dielectric structures [86], multi-shell nano-spheres [56,87], invisibility cloaks [88–90], or metasurface design [91–93].

D. Strategies to Improve Neural Network Inverse Design

Data-driven inverse design has the important drawback that the accuracy of the model is first of all limited by the quality of the data and an interpolation error between the data samples is introduced by the ANN. Early works on inverse design therefore reported rather qualitative agreement, but relatively large quantitative inaccuracies. Therefore, in the recent past remarkable

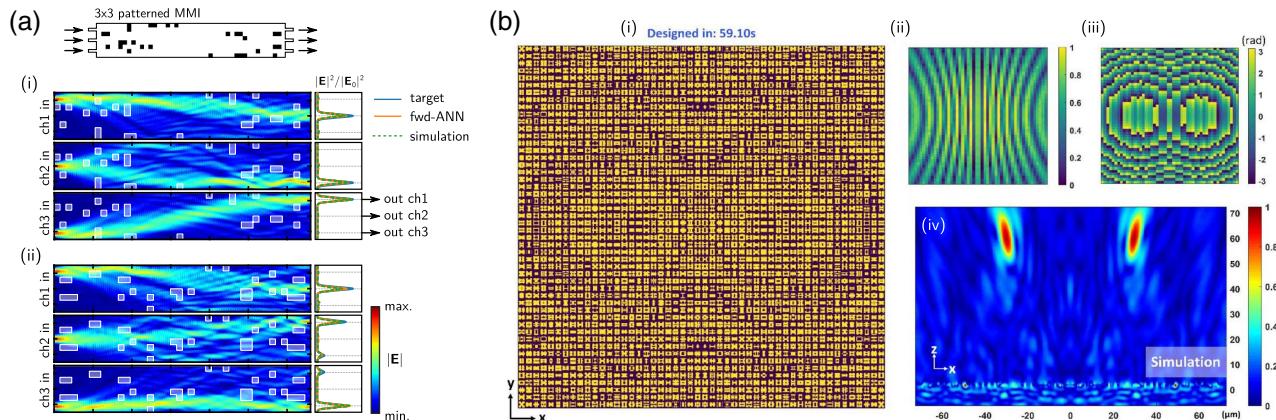


Fig. 2. Examples of devices inverse designed by ML algorithms. (a) Encoder-decoder type tandem inverse network used to design perturbation patterns for 3×3 MMIs as arbitrary transmission matrix elements. The light routing behavior of the second and the third input channels is interchanged between cases (i) and (ii), while the first input channel keeps routing light to the second output. Adapted with permission from [84], copyright (2021) American Chemical Society. (b) Double-focus flat lens designed by a conditional WGAN inverse network. (i) shows the dielectric metasurface, (ii) the corresponding amplitude, and (iii) the phase mask. (iv) shows a numerical simulation of the field intensity to test the ANN design. Adapted with permission from [73].

efforts have been put in developing methods to improve neural network inverse design. In this section we want to provide an overview over the most successful concepts. In general, two main constituents offer the largest potential for optimization: the training data and the neural network model.

1. Improving the Data Quality

As mentioned before, many ANN models do actually generalize relatively poorly to cases outside the parameter range of the training data. They act mainly as generalized function approximators, and hence they interpolate very efficiently to fill the gaps in the training data, while their extrapolation capability remains limited. But also, the interpolation risks may be unsatisfactory if the physical model underlying the training data has sharp features such as high quality factor resonances. If the training data does not contain a sufficient number of such resonant cases, there is a high risk that those features will be very poorly approximated by an ANN.

To tackle this problem, training data can be generated using an optimization algorithm to produce specific responses for the dataset [84,88]. In the case of many free parameters this procedure is time-consuming. Therefore, recently the idea of iterative training data generation has emerged [64,66,84,88,89,94]. The principal idea is depicted in Fig. 3(a). An initial dataset is generated traditionally via a randomized procedure, on which an inverse design ANN is trained. This network is subsequently used to construct devices based on realistic design targets, but these designs are likely to be rather mediocre as the initial ANN performs relatively poor. Now, the true physical response of these mediocre ANN designs is calculated in another run of numerical simulations, and these samples are appended to the training data. The generator ANN is then trained again on the now extended training data and the generative cycle is repeated. In this way, the neural network can literally learn from its previous mistakes and its performance on the specific design task will significantly improve. Figure 3(a) shows the example of an optical cloak design problem, for which the inverse design accuracy could be improved by more than 1 order of magnitude thanks to iterative training [89]. To visualize the evolution of the training data quality, Fig. 3(b) shows the statistical distributions of resonator quality factors in a fully random dataset of photonic crystal cavities (left) compared to a dataset after one iteration of iterative training (right) [94]. The lack of resonant geometries in the randomly generated dataset is evident. Despite those solutions not being present in the initial dataset, the ANN managed to conjecture a certain amount of resonant cases, improving the training data for the second iteration. By repeating the procedure, the training data increasingly contains resonant geometries, which consequently allows the ANN to inverse design close-to-optimum solutions. Another positive side effect specifically in tandem networks is that iterative training simultaneously improves the accuracy of the forward network [84]. A recent work showed that an even better design performance can be achieved by iteratively increasing the network complexity together with a successive augmentation of the training data, as depicted in Fig. 3(c) [95].

An obvious drawback of iterative procedures is their computational cost. Data generation is usually slow, and the expensive network training needs to be repeated several times on

increasing amounts of training samples. Several suggestions have been made to accelerate the convergence of iterative data generation in order to reduce the number of cycles. For instance, by training several networks, the statistics from multiple predictions can be used to assess the quality and the uncertainty of the ANN output (“wisdom of the many” [96]; see also Box 4). This information can be exploited to choose only the best new solutions for re-simulation and insertion into the expanded training data, which reduces the number of expensive physics simulations [64]. Similarly, an evolutionary optimization algorithm might be coupled to a generative ANN in the iterative cycle to further specialize the training data with regards to the anticipated optimization target [66]. A drawback of such training-data optimization strategies is a risk of over-specializing the network to optimum cases and losing its capability to generalize to arbitrary situations. Therefore, care needs to be taken that the training data remains sufficiently diverse.

2. Physics-Model-Based Loss Function

A similar, yet somehow more radical concept is to not use a fixed set of training data at all but instead to implement a loss function based on a physical model within the framework of the machine learning toolkit. Such an approach has been illustrated recently by the example of inverse designing multi-layer thin-film stacks for specific reflection and transmission spectra [97]. As highlighted by a red box on the right in Fig. 3(f), a transfer matrix method (TMM) has been implemented directly in the deep learning toolkit as a loss function. In consequence, error backpropagation is possible through the TMM solver, and the network can be trained without an explicit dataset. The loss function in this so-called “GLOnet” is used to optimize the transmission and reflection spectra of a multi-layer stack with respect to a design target. It is worth mentioning that the GLOnet learns to optimize a single design target, and hence in principle the training of the network takes the place of a conventional global optimization algorithm run (hence its name “GLOnet”). The authors of Ref. [97] claim that the training dynamics allow their GLOnet to ideally adapt its optimization scheme to each problem, resulting in better and faster convergence compared to hard-coded optimizers. The same authors have generalized their concept to a somehow more flexible inverse network called “conditional GLOnet,” using an iterative training scheme instead of a fully differentiable physics loss function. For the training, gradients of the design efficiency are calculated via adjoint simulations and re-injected for back-propagation through the network [98]. The conditional GLOnet is conceptually similar to a Pareto optimization in which a set of optimum solutions for a multi-objective problem is calculated [99]. While the specific solving of a single problem is intentional in Refs. [97,98], as already mentioned before over-specialization is an inherent danger of all iterative data-generation methods.

Another concept to replace the dataset by a direct evaluation of a physics model has been demonstrated for the Helmholtz equation, by developing a loss function which directly evaluates this partial differential equation (PDE). Such an ANN model is called a “physics-informed neural network” (PINN). In the case of a Helmholtz-PINN, the network learns to directly solve the wave equation in the frequency domain. The inverse design

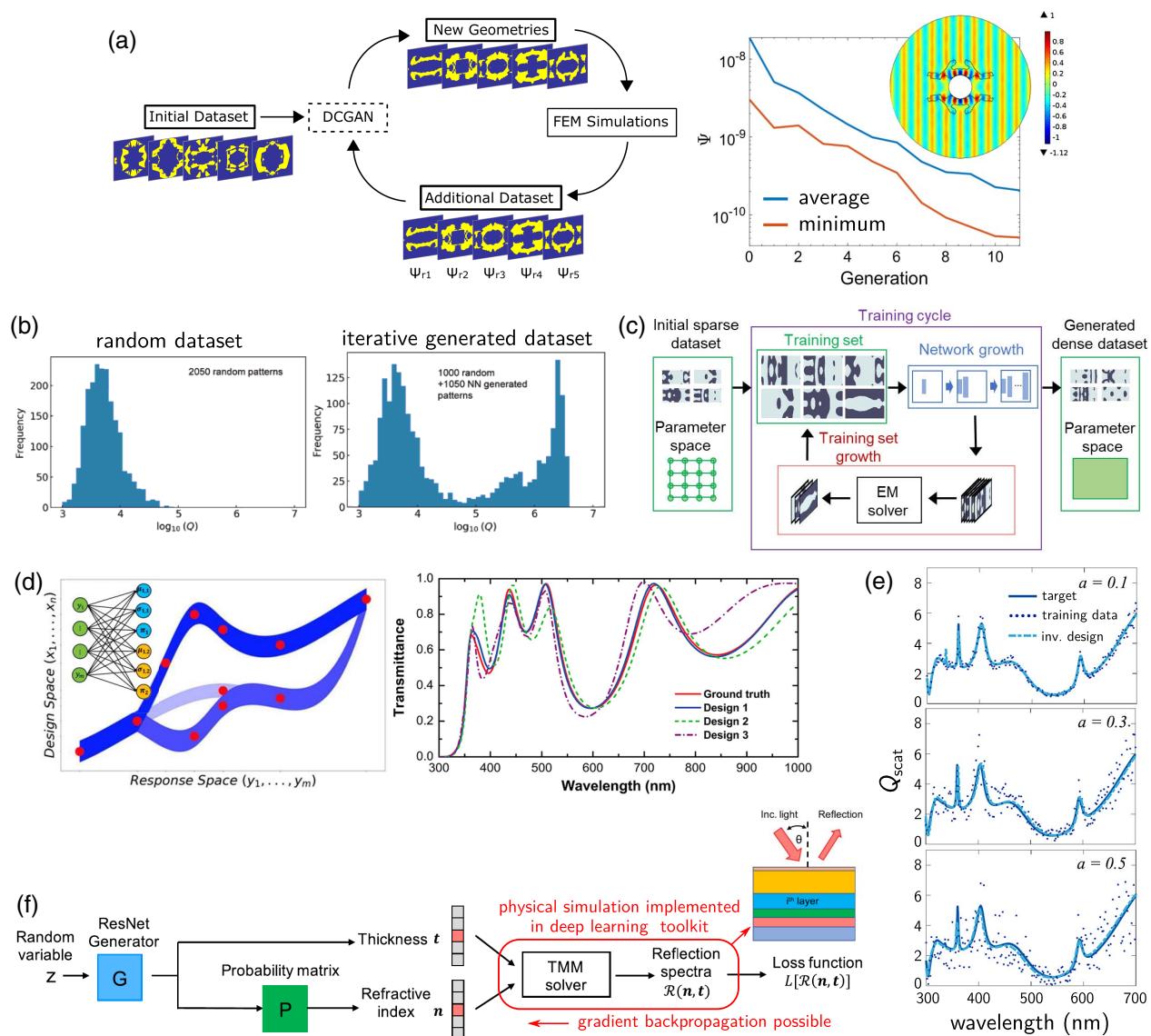


Fig. 3. Concepts to improve common shortcomings of inverse design ANNs. (a) Iterative training data generation, in which a network learns from its own errors, here applied to the inverse design of an invisibility cloak device. Adapted from [89], copyright (2021) Optical Society of America. (b) Comparison of the Q -factors for photonic crystal cavities in a random dataset (left) and in an iteratively generated dataset after the first iteration (right). Adapted from [94], copyright (2019) de Gruyter. (c) Together with the training data, the network complexity can be progressively growing, allowing even better performance by successive learning of smaller features. Reprinted with permission from [95], copyright (2020) American Chemical Society. (d) Mixture density ANN which represents multiple solutions with Gaussian probability distributions to find several non-unique solutions to ambiguous problems. The shown example deals with the spectral design of a multi-layer stack. Adapted with permission from [100], copyright (2020) American Chemical Society. (e) De-noising inverse ANN as robust approach for training on noisy data (noise parameter a increasing from top to bottom). Adapted from [101], copyright (2019) Optical Society of America. (f) “GLOnet”: inverse design ANN using a transfer-matrix model loss for reflectivity and transmission spectra optimization of multi-layer stacks. Adapted from [97], copyright (2020) de Gruyter.

target is then implemented as a boundary condition matching problem [90,102]. As in the GLOnet case, also such a PINN inverse design requires a new training run for each optimization target. PINNs will be discussed in more detail later in this review.

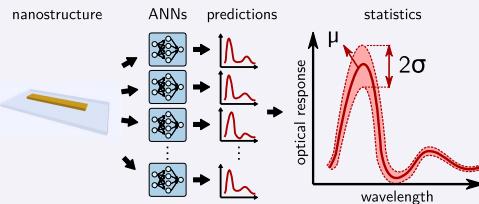
3. Sophisticated ANN Models

The second main lever allowing for performance optimization of inverse design ANNs is the neural network model itself.

It has been proven helpful to adopt recent findings in the research on optimum network layout for deep learning. For instance, if applicable the “U-Net” architecture [103] offers much better training convergence and generalization capacity than standard convolutional neural networks—even in cases where its particularly efficient segmentation capabilities are not required [58,104]. Furthermore, so-called residual blocks, or ResNets [22], should be adopted whenever possible. Residual blocks are characterized by their skip connections which avoid

Box 4. “Wisdom of the many”

Wisdom of the many or also *wisdom of the crowd* denotes the procedure of training multiple neural networks on the same data, each ANN with random initialization. We illustrate the idea by the example of an optical spectrum predictor network.



While this approach adds a significant computational cost (training several networks), the mean μ of N independent predictions provides a \sqrt{N} times smaller statistical error compared to using a single ANN. Furthermore, the standard deviation σ of multiple predictions can be used to assess the credence of the ANN output.

the vanishing gradient problem, allowing the training of very deep network layouts.

In addition to the application of general “best-practice” ANN design rules, problem-specific tailoring of the network layout can be very favorable for optimum inverse design performance. For instance, to tackle the one-to-many problem, “multi-branch” or “mixture density” ANNs can be applied in addition to the above-named network architectures. The concept is based on representing the design parameters in a “modal” representation as multiple Gaussian distributions,

where each of the Gaussian distributions describes a possible solution to an ambiguous problem (see also Box 2). This concept was proposed some time ago for microwave device inverse design [105,106] and was recently adapted to nano-photonics [100,107] [see also Fig. 3(d)]. The advantage is that the network can in principle deliver all possible solutions together with a weight for their respective priorities. A drawback of the approach is that the approximate number of non-unique solutions needs to be known in advance.

Another recent proposition to optimize inverse networks specifically for noisy situations like in experiments is the implementation of concepts from machine-learning-based image denoising [108]. As shown in Fig. 3(e), Hu *et al.* added artificial noise on training data and could demonstrate that a denoising network-based inverse ANN offers a very robust performance even when trained on very noisy data [101]. This opens promising perspectives for experimental applications.

4. Reformatting the Input Data

Apart from optimizing the network model and generating training data of high quality, the format of the inputs and outputs of a neural network can play a decisive role in whether the ANN manages to “understand” the data or not. An example is illustrated in Fig. 4(a), where a physical problem is to be solved on a non-Cartesian coordinate domain. On 2D problems such as the one here shown, typically convolutional neural networks (CNNs) are most efficient. However, as can be seen in the left-most panel, the imposed discretization on a square mesh is very poor. This holds in particular for the domain borders. Gao *et al.* [109] proposed therefore to apply a transformation of the coordinate system from the physics domain to the CNN reference domain prior training. As illustrated in Fig. 4(b) by the example of solving the heat equation, this additional pre-processing allows to successfully apply ANNs to very complex non-uniform physical domains.

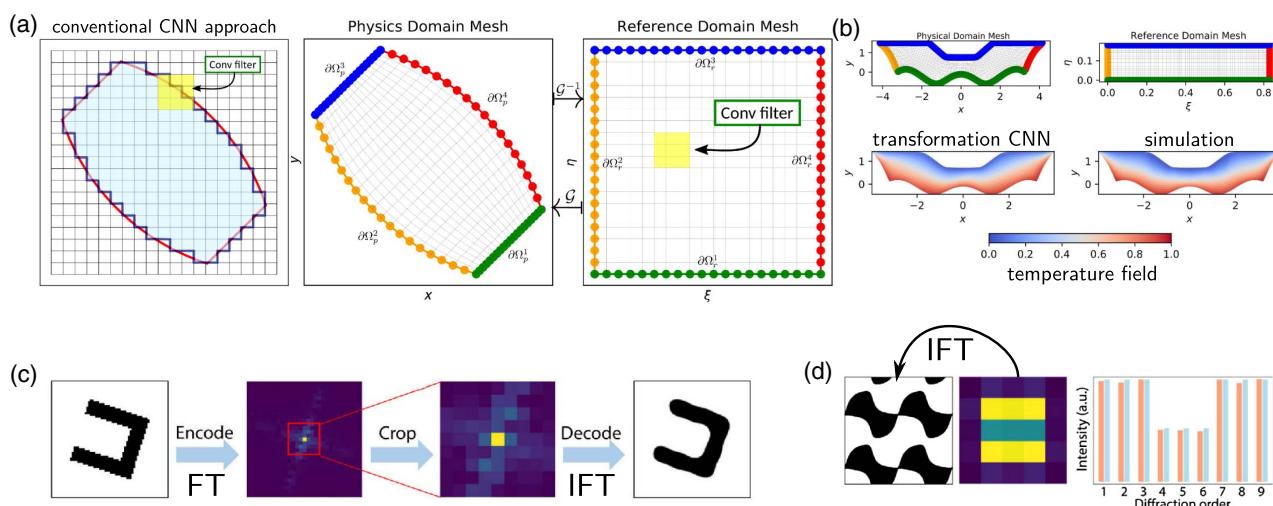


Fig. 4. Examples of input data pre-processing for optimized physics domain representation. (a),(b) Deep learning on irregular grids via coordinate transform (a) which is implemented within the deep learning toolkit to allow fast gradient calculations through the coordinate system transformation. (b) The transformation allows to efficiently train networks on complex-shaped physical domains. Adapted with permission from [109], copyright (2020) Elsevier. (c) Data encoding and compression using a topology description based on low-frequency Fourier components, which allows data-efficient treatment of complex shapes, here for example a free-form metagrating. Adapted from [110], copyright (2020) Optical Society of America.

The problem of discretization can also be alleviated by applying a topology encoding procedure, for instance via Fourier transformation [110]. The idea is illustrated in Figs. 4(c) and 4(d). Such encoding can allow not only to describe geometries with odd shapes without restrictions due to discretization, but it allows furthermore to condense the information to a low-dimensional space, which is helpful to reduce ANN complexity and furthermore advantageous in preventing overfitting.

5. Other Concepts

Further possibilities to improve the quality of ANN-based inverse design are to use the ANN only as a first step for a rough estimate and apply a conventional iterative approach in a subsequent refinement step. Heuristic optimization algorithms usually benefit strongly from a good initial guess [68]. Another recent proposition is to use a forward neural network purely as an ultra-fast physics predictor to construct a huge lookup table [111]. Using a well-trained forward network, a lookup table can be created which covers the entire parameter space at a very fine resolution, impossible to achieve with conventional numerical methods. Appropriate solutions to specific problems can subsequently be searched in this database. Transfer learning has also been recently applied to nano-optics problems to improve ANN performance if only small amounts of data exist [112]. For instance, experimental data is often expensive, but the situation can be improved by training an ANN first on simulated data, and subsequently specializing the pre-trained network via transfer learning on the experimental dataset [113].

E. Heuristics versus Deep Learning—A Critical Comparison

It is of utmost importance to emphasize that a data-driven inverse design technique can never outperform an iterative method if it is based on the same simulation model used for training data generation. At least not if no time constraint is set for the iterative optimization. Well-trained and optimized data-driven ANNs usually produce errors of the order of a few percent [55,58]. Furthermore, it is virtually impossible to completely suppress outliers in the network predictions [67]. At the singular points the error of the ANN can be orders of magnitude higher. It is thus a delicate task to assess whether a prediction is valid or rather the result of a singularity in the ANN.

While recently some sophisticated training techniques were presented that are capable to train ANNs for performances similar to conventional inverse optimization, they are either still considerably constrained or the high accuracy has a severe impact on the computational cost. Examples are physics-loss based inverse ANNs or networks based on progressive-complexity training schemes [95,97]. The model described in Ref. [97], for example, is constrained to a simple transfer-matrix description of a multi-layer system as well as to the inverse design of a single optimization target.

The fact that ANNs always introduce an additional error is inherent to the data-driven nature of machine learning (ML), which implies that an ML model can never outperform the accuracy of the simulations used to create the dataset or the model defining the training loss. On the other hand, once trained ANN techniques can offer extreme speed-up of the inverse design,

generally many orders of magnitude faster than iterative approaches based on numerical simulations, it is not unusual that milliseconds stand against hours or even days. This is a marvelous advantage and often well worth it to accept the reduced accuracy of ANN-based techniques. In daily applications a few percent error might actually not even matter too much, in particular when compared to the typical magnitude of inaccuracies in fabrication.

On the other hand, concerning the inverse design speed it is important to remember that the ultra-fast predictions require a fully trained neural network. This implies the computationally highly demanding data generation as well as the very expensive training of the ANN. In many situations, conventional global optimization is in sum actually computationally cheaper. In conclusion, deep-learning-based inverse design is mainly interesting for applications which require a large number of repetitions of similar design tasks, or that rely on ultimate speed for the design generation.

3. BEYOND INVERSE DESIGN

The second part of this review is dedicated to applications of deep learning in nano-photonics “beyond inverse design.” We give an overview on physics-informed neural networks; we present recent work on ANNs for physics interpretation and knowledge discovery as well as experimental applications.

A. Physics-Informed Neural Networks: Solving PDEs

Most machine learning applications in physics aim to predict derived observables such as transmittance or extinction cross sections. In contrast, the idea of PINNs is to train an ANN to directly predict the solution of a PDE. While this would be also possible using a dataset of pre-calculated solutions, the particularity of PINNs is that instead of using a loss function for data comparison like MSE, the PINN-loss implements an explicit evaluation of the PDE. In consequence, no pre-calculated training data is required for training. For the PINN-loss, the PDE derivatives of the ANN-predicted observables are directly implemented in the respective deep learning toolkit. Thus, the PINN-loss can be seen as a consistency check for the predicted solution. Because modern deep learning toolkits offer powerful automatic differentiation functionalities, error backpropagation through the PINN-loss remains possible and the ANN can be efficiently trained without data.

This concept was first proposed in 2019 by Raissi *et al.* [114] and has since then attracted a great deal of attention across countless research communities in physics, such as fluid mechanics [114,115], thermodynamics [109], and geophysics [116]. Compared to data-based ANNs, the accuracy of PINNs is in general significantly higher. On the other hand, because PINNs evaluate the underlying PDE “point by point,” they are usually slower than conventional data-based models. Since the latter work on physical observables, it is easier to predict higher-dimensional data structures at a time, making better use of the massive parallel computing architectures of modern GPUs. Nevertheless, PINNs are usually orders of magnitude faster than numerical PDE solvers.

Applications to nano-photonics are still scarce. Recently Moseley *et al.* demonstrated that PINNs are capable of

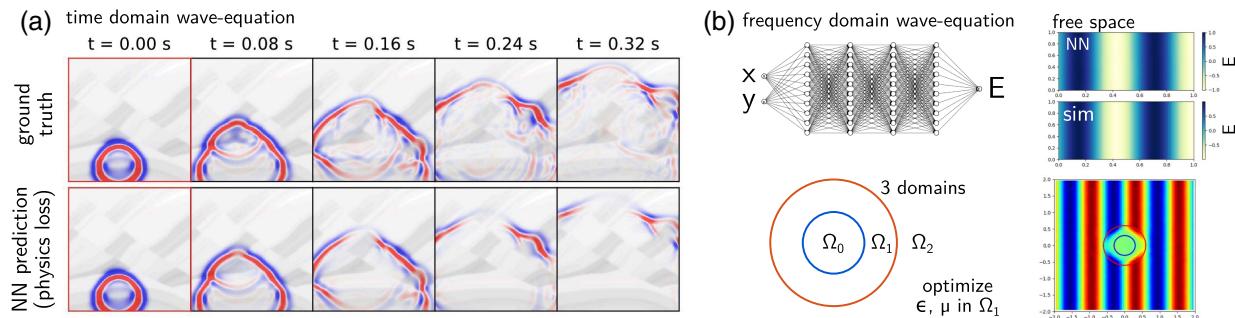


Fig. 5. Physics-informed neural networks (PINNs) for nano-optics. (a) PINN for solving the wave equation in the time domain. Adapted with permission from [116]. (b) Top: solving the Helmholtz equation (frequency domain); bottom: using the PINN for inverse design of the permittivity distribution in domain Ω_1 for an invisibility cloak application. Adapted with permission from [102], copyright (2020) IEEE.

accurately solving the wave equation in the time domain [116]. An example is shown in Fig. 5(a), demonstrating seismic wave propagation through an inhomogeneous medium at successive snapshots in time. As can be seen, the PINN is capable of predicting the evolution of the wave propagation even in a complex environment. While Ref. [116] treats shock waves in geophysics, the problem is conceptually identical to the wave equation in electrodynamics.

Depicted in Fig. 5(b), Fang and Zhan recently demonstrated that a PINN can accurately solve the Helmholtz equation, describing wave propagation in the frequency domain [102]. They found that sinusoidal activation functions are the most adequate choice to solve a differential equation with time-harmonic solutions. By formulating the inverse design as a boundary condition matching problem, it was possible to use the Helmholtz-PINN for the design of an optical cloak, as illustrated in the bottom of Fig. 5(b). A similar frequency-domain PINN has been proposed for the homogenization of optical metamaterials [90]. A disadvantage of PINNs is that the environment needs to be defined at the training stage and hence a new network needs to be trained if the boundary conditions change. Each PINN-based inverse design therefore involves a new training procedure, comparable with conventional iterative techniques, which is evidently much slower than “direct” inverse ANN models. Conceptually related to PINNs is also the so-called “GLOnet,” which is discussed in more detail above [97] [see also Fig. 3(f)].

B. Interpretation of Physical Properties

In this section we will review recent approaches to extract information and correlations from deep learning models in order to reveal physical insights.

There is on the one hand the possibility to use deep learning models for dimensionality reduction. Figure 6(a) shows a work of Kiarashinejad *et al.* in which the number of discrete values in reflectance spectra from a set of electrodynamical simulations is reduced from 200 to 2 via an unsupervised autoencoder ANN. In a second step, the non-convex hull of the compressed responses is calculated, which represents the region in the 2D compressed space containing all encoded points. This region allows us to assess the range of accessible physical responses within the allowed design parameters, and the method is hence helpful to identify the physical limitations of specific

nano-structure models. Note that the full physical response of any point in the reduced dimensionality space can be reconstructed using the decoder part of the autoencoder, also those points that were not present in the training data. This means that feasible and non-feasible responses can be analyzed in the original response space (under the assumption that the neural network generalizes well to out-of-training situations). Note that autoencoders are unsupervised ANN models, which are known to require relatively few data for training. This facilitates the application of the technique to experimental data.

In a similar approach, the impact of variations of individual design parameters on the latent space can be studied. Those parameters whose variations have large (respectively little) impact on the latent space contribute strongly (respectively weakly) to the optical response [117–119]. The latent space is indicated by yellow highlighted neurons in Fig. 6(b), top right. The impact of physical parameters on these weights is illustrated in the bottom right of Fig. 6(b). By varying the size of the bottleneck (i.e., reducing the latent space dimension), it is furthermore possible to extract something like the number of principal components of the response, as shown in the left column of Fig. 6(b). Iten *et al.* [120] extended the encoder-decoder ANN for interpretable physics via an approach inspired by humans’ interpretation and modeling of physical observations. The concept is depicted in Fig. 6(c), where the motion of a mass is observed as a function of time $x(t)$. To implement this concept in an ANN the authors append a condition to the latent vector at the bottleneck of an encoder-decoder ANN [see Fig. 6(d)]. This condition is here called a question; the example in Fig. 6(c) uses the time t' for which the ANN shall predict the position of the moving mass (= the answer). In the context of nano-photonics the question could be an optical spectrum of a nano-structure. The “answer” returned by the ANN might then be the material or the size of the nano-structure, or a wavelength or laser polarization state. This kind of ANN is conceptually very similar to inverse design ANNs (in particular to the cGAN or cAE models), but instead of using it for the design of nano-structures, it is here used to understand causal correlations imposed by the implicit physics in the training data.

A more direct approach to extract physical knowledge from ANNs consists in using the ultra-fast approximation capability

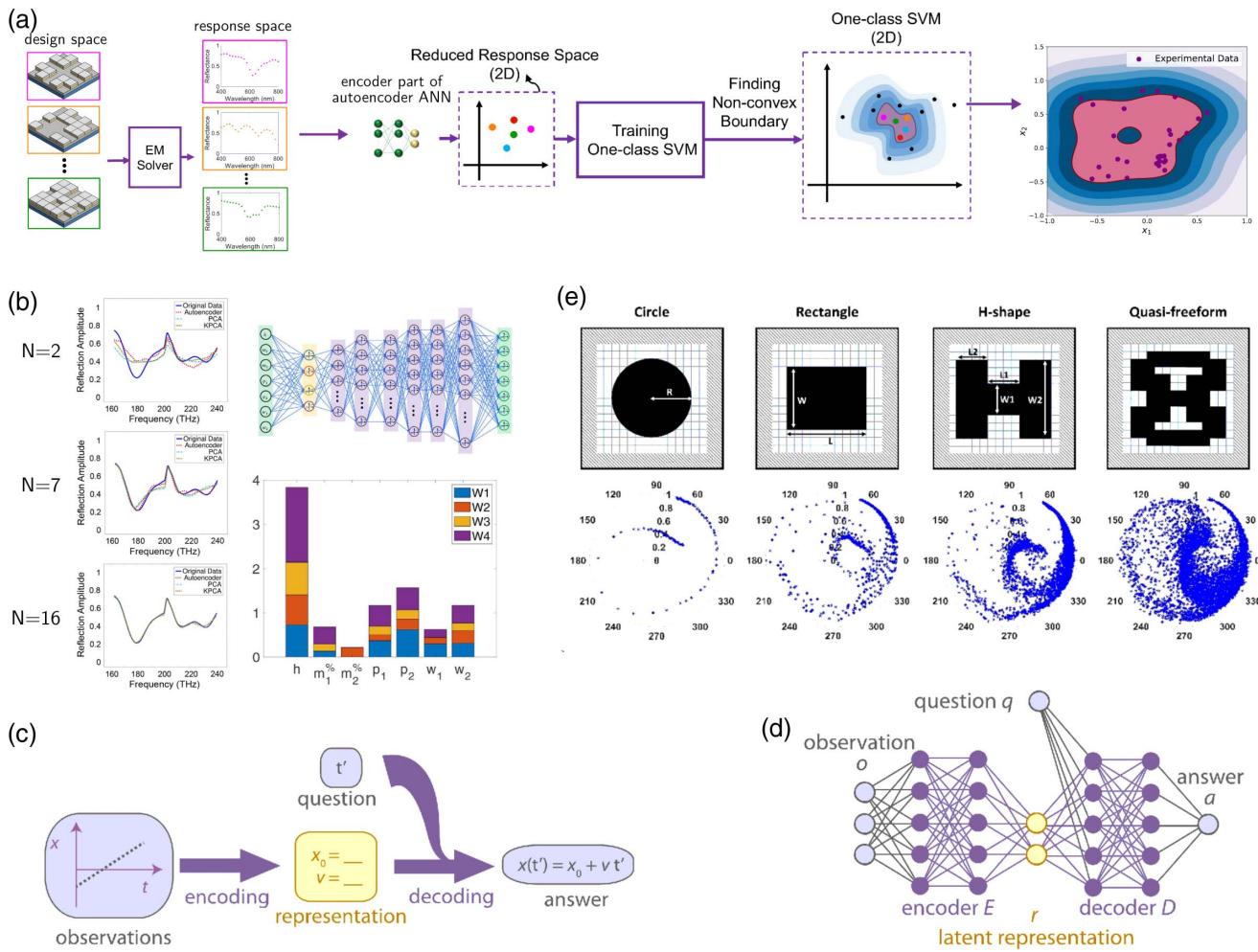


Fig. 6. Examples of “knowledge discovery” through machine learning. (a) The feasibility of a physical response by a defined geometric model can be assessed by a dimensionality reduction through an autoencoder neural network and subsequent non-convex hull determination. Adapted from [121], copyright (2019) the authors. (b) Study of the impact of the number of bottleneck neurons N (left spectra) as well as of nano-structure design variations on the activation of the bottleneck neurons (W_1 – W_4 in case $N = 4$, yellow neurons in the top right panel). This analysis allows to assess the physical importance of individual design parameters and reveals information about the complexity of the optical response. Adapted with permission from [117], copyright (2019) John Wiley and Sons. (c), (d) By mimicking the human approach of interpreting and modeling physical observations (c), a conditional encoder–decoder network (d) can be used to discover implicit physics concepts from data. Reprinted with permission from [120], copyright (2020) APS. (e) Exploiting the high speed of a physics predictor network permits a systematic analysis of the achievable phase and intensity variations in metasurface constituent design. Adapted from [122], copyright (2020) Optical Society of America.

of deep learning surrogate models. Through a systematic scan of the whole parameter space it is, for example, possible to assess the accessible optical responses with a specific nano-structure model. In this way, accessible phase and intensity values for metasurface elements have been classified systematically by An *et al.* [122]. The logical conclusion of the study was that allowing more complex shapes for the meta-atoms leads to a larger accessible range for the phase and intensity, as depicted in Fig. 6(e). From left to right are shown increasingly complex geometric models (top row) and their accessible scattering phase and intensity range (bottom row).

As already mentioned before, another way to gain insight in physical processes through a machine learning analysis is to use a physical parametrization of the training data, such that the neural network explicitly returns a physical quantity. As shown

in Figs. 1(a) and 1(b), extinction spectra can, for example, be pre-processed in a modal decomposition, such as a superposition of electric and magnetic dipole resonances [56] or as a decomposition in Lorentzian resonance profiles [57]. Once trained, the respective neural networks deliver an explicit interpretation of the predicted spectra.

In another recent work, so-called explainable machine learning has been used to assess the importance of constituent parts of a nano-structure with respect to its optical response, as well as to identify those parts of the structure that contribute only weakly to the light–matter interaction [123]. Such information is important for the design of fabrication-robust nano-structures, but also for applications in which sub-constituents of high impact on the nano-structure’s optical response need to be identified, e.g., for switchable optical antennas. Another

recent work proposes interpretable machine learning models like decision trees and random forests to understand the physical mechanisms behind inverse design results [124].

C. Deep Learning for Interpretation of Photonics Experiments

The last section of this review is dedicated to recent applications of deep learning in nano-photonics experiments.

Deep learning has proven to enable unprecedented statistical evaluation of large and complicated data, which was formerly impossible with conventional methods. It has been demonstrated, for instance, that ANN models can learn from huge microscopy datasets to optically characterize 2D materials such as graphene or transition-metal dichalcogenides [125] or to automatically localize and classify nano-scale defects [126] or to track particles in 3D space using holographic microscopy [127]. Deep learning was also successfully applied for the ultra-fast analysis of single-molecule emission patterns [128] as well as for the experimental reconstruction of quantum states for quantum optics tomography [129].

By training an ANN on large amounts of experimental optical scattering spectra from complex photonic nano-structures, recently an optical information storage concept has been proposed, able to push the data density beyond the optical diffraction limit [130]. The principle is depicted in Fig. 7(a). Digital information is encoded in silicon nano-structures, which are designed such that each nano-structure encoding a specific bit-sequence possesses a unique scattering spectrum. Visible light scattering is subsequently interpreted by an artificial neural network [Fig. 7(b)]. Training on experimental data renders this readout robust against fabrication imperfections and instrumental noise. Therein, the ANN is the key ingredient to allow high readout accuracies from distorted data [Fig. 7(c)]. Deep learning can be used for various further experimental classification tasks in nano-optics. For instance, as depicted in Fig. 7(d), it has been recently demonstrated that an ANN can learn to classify different species of anthrax spores from holographic microscopy images [131]. The confusion rates in the individual classes [Fig. 7(e)] allow furthermore to assess similarities and differences between the different anthrax species. Similar recent deep-learning-based holographic image classification tasks include analysis of colloidal dispersions [132] or the real-time determination of size and refractive index of sub-wavelength small particles [133].

Deep learning is particularly strong at the interpretation of sparse, undersampled data. In a recent example, Argun *et al.* used a deep neural network for force field calibration in microscopy, by monitoring and interpreting Brownian particle motion [134]. As depicted in Fig. 7(f), complex trapping potentials (top left) can be reconstructed efficiently from few experimental samples (top right). In contrast to a conventional method (bottom right), the ANN (bottom left) reconstructs the correct potential with high accuracy also from little data [using only the dark part in the top right panel of Fig. 7(f)]. Similarly, machine learning has been used for real-time particle tracking [135–137]. Recently ANNs have also been successfully trained on simulated data to efficiently predict the optical forces in complex particle trapping situations [138]. Moreover, deep learning has been found to be very powerful in solving inverse

problems occurring in imaging experiments. In this context often sparsity assumptions are required to enable deconstruction of undersampled data, which demands computationally complex inverse solving techniques such as compressive sensing. Corresponding imaging applications include phase recovery [139,140], image reconstruction or enhancement [141–144], super-resolution microscopy [145–149], and coherent diffractive imaging [150,151]. In the context of photonics, it has been demonstrated that speckle patterns which occur after light transmission through complex media can be deconstructed very efficiently with deep learning methods [104,152–156]. While such speckles appear as if they were random patterns, they are actually the result of deterministic multiple scattering events. Therefore, a fixed correlation between input and output before and after the complex medium can be established, which is classically done by constructing a transmission matrix [157], involving complex regularization schemes, inversion procedures, or computationally expensive compressive sensing techniques [158]. While speckle-based methods allow, for instance, imaging through opaque media or the reconstruction of spectral information, the aforementioned computational burden usually prohibits real-time applications. ANN models, on the other hand, can be trained to solve the implicit inverse problem in speckle deconstruction very efficiently, which recently enabled use of complex media such as multi-mode fibers for real-time applications in imaging [104,153–155,159], spectral reconstruction [156], or both (hyper-spectral imaging) [152]. Figure 7(g) illustrates a setup for such speckle-based hyper-spectral imaging. An image is formed via an intensity spatial light modulator, spectrally shaped using an acousto-optic tunable filter, and focused on the aperture of a multi-core multi-mode fiber bundle. The fiber cores act as pixels of the image, whose individual speckle patterns encode the spectral information. Kürüm *et al.* [152] demonstrated that even under noisy conditions and in the undersampling regime, an ANN can reconstruct the spectral information of several thousand fibers with a speed of a few frames per second. In contrast, conventional compressive sensing algorithms require tens of minutes for the same task with similar reconstruction fidelity [158].

In the context of sparse data reconstruction, deep learning has recently been used in quantum optics applications for the reconstruction of statistical distributions from experiments with weak photon counts, as schematized in Fig. 7(h). For instance, Cortes *et al.* [160] demonstrated the successful reconstruction of time-dependent data from few photon events using statistical learning. In this procedure a machine learning algorithm learns to predict the statistical distribution of the data. A similar approach has been applied to assess whether a nano-diamond contains a single or several nitrogen vacancy photon emitters [161]. Another work demonstrated a machine learning model capable of differentiating between coherent and thermal light sources via a statistical analysis of the temporal distribution of a very low number of photons [162]. These learning-based statistical analysis methods are capable of outperforming conventional data fitting techniques thanks to their capacity to learn the most probable statistical distributions from the actual data. Essentially, the machine learning model learns

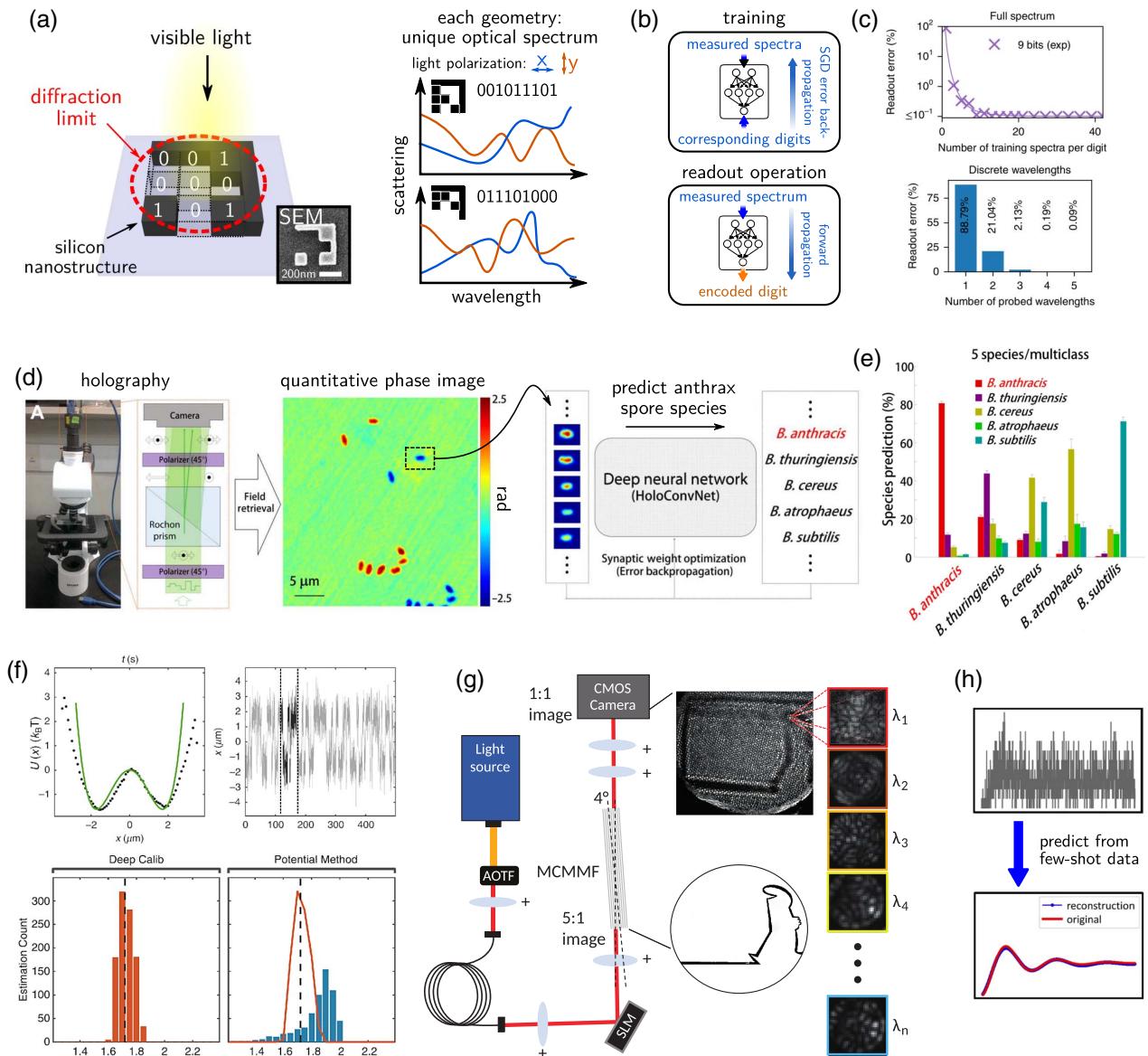


Fig. 7. Examples of ML applications in experimental data interpretation. (a)–(c) ANN used to decode information from optical information storage via a spectral scattering analysis from sub-diffraction small nano-structures. (a) Each bit sequence is encoded by a specific geometry which is designed such that it possesses a unique scattering spectrum. (b) A neural network is trained on a large amount of spectra such that it learns to decode noisy spectra of formerly not seen structures. (c) Even if only few wavelengths are probed, the readout accuracy of the network is excellent. Adapted with permission from [130], copyright (2019) Springer Nature. (d), (e) Holographic anthrax spore classification via holography microscopy. A machine learning algorithm is trained on phase images of different spore species, as depicted in (d). The neural network is capable to classify five different anthrax species with a very high accuracy. Adapted from [131], copyright (2017) the authors. (f) Microscopy force field calibration (top left, green line: trapping potential; dots: reconstructed potential). Evaluation of $U(x)$ via ANN-based analysis of Brownian motion from undersampled statistical data (top right). Comparison of reconstruction fidelity of ANN (bottom left) and conventional method (bottom right). Ground truth is indicated by a black dashed line. Adapted from [134], copyright (2020) the authors. (g) ANN enabled real-time hyper-spectral image reconstruction from speckle patterns produced by a multi-core multi-mode fiber bundle (MCMMF). The technique exploits the wavelength dependence of the speckle patterns. Adapted from [152], copyright (2019) Optical Society of America. (h) Scheme depicting the use of machine learning for statistics reconstruction of few-shot data acquisitions. Reprinted from [160], with the permission of AIP Publishing.

to “focus” on the important regions in the data (comparable to adaptive fitting weights). Conventional data fitting algorithms on the other hand tend to attach too much importance to “flat” areas, to the detriment of the accuracy in the relevant regions. Just as with accidentally over-specialized inverse networks, care must be taken when interpreting the ANN

reconstructions. Since data-driven approaches always bear the risk of being biased toward the training data, a neural network might, for instance, detect a learned statistical distribution even in pure noise.

Deep learning can be applied not only to data analysis but is also increasingly used to control real-time experimental

feedback systems. Recent examples touching the field of nanophotonics are mainly found in AI-stabilized microscopy. ANNs can be applied, for instance, to real-time image enhancement [163], microscopy stabilizing feedback systems [20,164], or to conduct sparse data acquisition schemes for the acceleration of scanning microscopy systems via compressive sensing [165]. ANNs have been also applied to controlling laser mode-locking stabilization systems [166–168]. So far, the direct application of ANNs to experimental hardware for nano-photonics is still scarce, but the research is in an early stage. A recent work proposed, for instance, to calibrate and control electrically reconfigurable photonic circuits by deep learning algorithms [169]. Another example is a pioneering work of Selle *et al.* [51] which proposed to use ANNs coupled to a femtosecond laser pulse shaper for real-time control of the light-matter interaction in nano-structures or molecules. We expect a very rapid development of applications in this direction in the near future; in particular, real-time critical applications such as sensing [170] will hugely benefit from the tremendous acceleration potential of ANNs.

4. CONCLUSIONS AND PERSPECTIVES

In conclusion, in this mini-review we discussed the most recent developments in deep learning methods applied to nanophotonics. In the first section we focused on ANN-driven nano-photonic inverse design methods and discussed concepts to improve the design quality of inverse ANNs in comparison with conventional optimization techniques. In the second part we discussed applications of deep learning in nano-photonics “beyond inverse design,” spanning from physics-informed neural networks over ANNs for physical knowledge extraction to data interpretation and experimental applications.

We would like to emphasize that despite their latest remarkable success and their undeniable great potential, artificial neural networks are “black boxes.” It is extremely hard, mostly even impossible, to understand how a neural network generates its predictions. It has been demonstrated on many occasions that even the most sophisticated ANNs, trained on the most carefully assembled datasets, contain singular points at which their predictions diverge. Another noteworthy danger of data-driven techniques is that they bear a considerable risk to be biased with respect to their training data, such as an incident where Google’s image-tagging algorithm learned implicit racism from its training data [171]. We therefore appeal to the reader to keep in mind that, simply speaking, “what you put in is what you get out.” In consequence the ANN models are only the second most important ingredient to deep learning. The essential element is first of all the training data. Unfortunately, it is often understated and not discussed with sufficient emphasis that high-quality training data is of the utmost importance. By reviewing techniques that aim at improving the training data quality, we tried to arouse some awareness in this respect. Another important aspect in this context is the amount of training data required to train a well-performing and generalizing ANN. Unfortunately, in many problems which would be naturally suited for deep learning applications, training data is scarce or very expensive to generate. Additionally, the more general a problem for an ANN is, the more training data is

usually required for a good prediction fidelity. Last but not least, adapting an ANN model to a new problem often requires the entire training data to be generated from scratch, which might even be the case for minor modifications. These aspects can create considerable computational barriers for broad and flexible applications of ANNs.

Deep learning techniques in the context of nano-photonics have experienced a tremendous amount of attention in the past few years and research has literally exploded. ANNs have enabled manifold applications which formerly seemed strictly impossible. As discussed above, a prominent example is data-driven ultra-fast solvers for various inverse problems, for which conventional methods are computationally extremely expensive and slow. We expect that further groundbreaking applications will be developed in the near future. For instance, very promising progress has been made in the field of quantum machine learning [172], which aims at using deep learning concepts to push the capabilities and interpretability of quantum computing systems. In this context, machine learning algorithms recently have autonomously proposed designs for non-trivial quantum optics experiments [173–175]. We expect that deep learning will continue to produce exciting pioneering results. We also anticipate that deep learning techniques will become a common numerical tool, regularly employed for the daily use.

Funding. CALMIP Toulouse (p20010); Engineering and Physical Sciences Research Council (EP/M009122/1); Deutsche Forschungsgemeinschaft (WI 5261/1-1).

Acknowledgment. We thank the NVIDIA Corporation for the donation of a Quadro P6000 GPU used for this research. This work was supported by the German Research Foundation (DFG) through a research fellowship. The authors acknowledge the CALMIP computing facility. OM acknowledges support through EPSRC.

Disclosures. The authors declare no conflicts of interest.

REFERENCES

1. P. Mühlischlegel, H.-J. Eisler, O. J. F. Martin, B. Hecht, and D. W. Pohl, “Resonant optical antennas,” *Science* **308**, 1607–1609 (2005).
2. A. I. Kuznetsov, A. E. Miroshnichenko, M. L. Brongersma, Y. S. Kivshar, and B. Luk'yanchuk, “Optically resonant dielectric nanostructures,” *Science* **354**, aag2472 (2016).
3. C. Girard, “Near fields in nanostructures,” *Rep. Prog. Phys.* **68**, 1883–1933 (2005).
4. M. Kauranen and A. V. Zayats, “Nonlinear plasmonics,” *Nat. Photonics* **6**, 737–748 (2012).
5. G. C. des Francs, J. Barthes, A. Bouhelier, J. C. Weeber, A. Dereux, A. Cuche, and C. Girard, “Plasmonic Purcell factor and coupling efficiency to surface plasmons. Implications for addressing and controlling optical nanosources,” *J. Opt.* **18**, 094005 (2016).
6. J. Wang, F. Sciarrino, A. Laing, and M. G. Thompson, “Integrated photonic quantum technologies,” *Nat. Photonics* **14**, 273–284 (2020).
7. J. B. Pendry, “Negative refraction makes a perfect lens,” *Phys. Rev. Lett.* **85**, 3966–3969 (2000).
8. P. Genevet, F. Capasso, F. Aieta, M. Khorasaninejad, and R. Devlin, “Recent advances in planar optics: from plasmonic to dielectric metasurfaces,” *Optica* **4**, 139–152 (2017).

9. N. M. Estakhri, B. Edwards, and N. Engheta, "Inverse-designed metastructures that solve equations," *Science* **363**, 1333–1338 (2019).
10. W. R. Clements, P. C. Humphreys, B. J. Metcalf, W. S. Kolthammer, and I. A. Walmsley, "Optimal design for universal multiport interferometers," *Optica* **3**, 1460–1465 (2016).
11. F. Zangeneh-Nejad, D. L. Sounas, A. Alù, and R. Fleury, "Analogue computing with metamaterials," *Nat. Rev. Mater.* **6**, 207–225 (2021).
12. E. A. Ash and G. Nicholls, "Super-resolution aperture scanning microscope," *Nature* **237**, 510–512 (1972).
13. D. W. Pohl, W. Denk, and M. Lanz, "Optical stethoscopy: image recording with resolution $\lambda/20$," *Appl. Phys. Lett.* **44**, 651–653 (1984).
14. E. Betzig, A. Harooshian, A. Lewis, and M. Isaacson, "Near-field diffraction by a slit: implications for superresolution microscopy," *Appl. Opt.* **25**, 1890–1900 (1986).
15. B. Gallinet, J. Butet, and O. J. F. Martin, "Numerical methods for nanophotonics: standard problems and future challenges," *Laser Photonics Rev.* **9**, 577–603 (2015).
16. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature* **521**, 436–444 (2015).
17. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT, 2016).
18. S. Chan and E. L. Siegel, "Will machine learning end the viability of radiology as a thriving medical specialty?" *Br. J. Radiol.* **92**, 20180416 (2018).
19. A. S. Lundervold and A. Lundervold, "An overview of deep learning in medical imaging focusing on MRI," *Z. Med. Phys.* **29**, 102–127 (2019).
20. D. A. Cirovic, "Feed-forward artificial neural networks: applications to spectroscopy," *TRAC Trends Anal. Chem.* **16**, 148–155 (1997).
21. T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. M. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, and D. Amodei, "Language models are few-shot learners," in *Proceedings of Advances in Neural Information Processing System* (2020), pp. 1877–1901.
22. C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, inception-ResNet and the impact of residual connections on learning," in *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence* (2016), pp. 4278–4284.
23. X. Lin, Y. Rivenson, N. T. Yardimci, M. Veli, Y. Luo, M. Jarrahi, and A. Ozcan, "All-optical machine learning using diffractive deep neural networks," *Science* **361**, 1004–1008 (2018).
24. T. W. Hughes, I. A. D. Williamson, M. Minkov, and S. Fan, "Wave physics as an analog recurrent neural network," *Sci. Adv.* **5**, eaay6946 (2019).
25. D. Mengu, Y. Rivenson, and A. Ozcan, "Scale-, shift- and rotation-invariant diffractive optical networks," *ACS Photon.* **8**, 324–334 (2021).
26. R. S. Hegde, "Deep learning: a new tool for photonic nanostructure design," *Nanoscale Adv.* **2**, 1007–1023 (2020).
27. W. Ma, Z. Liu, Z. A. Kudyshev, A. Boltasseva, W. Cai, and Y. Liu, "Deep learning for the design of photonic structures," *Nat. Photonics* **15**, 77–90 (2020).
28. S. So, T. Badloe, J. Noh, J. Bravo-Abad, and J. Rho, "Deep learning enabled inverse design in nanophotonics," *Nanophotonics* **9**, 1041–1057 (2020).
29. J. Jiang, M. Chen, and J. A. Fan, "Deep neural networks for the evaluation and design of photonic devices," arXiv:2007.00084 (2020).
30. L. Huang, L. Xu, and A. E. Miroshnichenko, "Deep learning enabled nanophotonics," in *Advances in Deep Learning* (InTech, 2020).
31. M. M. R. Elsayed, S. Lanteri, R. Duvigneau, J. A. Fan, and P. Genevet, "Numerical optimization methods for metasurfaces," *Laser Photonics Rev.* **14**, 1900445 (2020).
32. S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vucković, and A. W. Rodriguez, "Inverse design in nanophotonics," *Nat. Photonics* **12**, 659–670 (2018).
33. G. M. Sacha and P. Varona, "Artificial intelligence in nanotechnology," *Nanotechnology* **24**, 452002 (2013).
34. K. Yao, R. Unni, and Y. Zheng, "Intelligent nanophotonics: merging photonics and artificial intelligence at the nanoscale," *Nanophotonics* **8**, 339–366 (2019).
35. J. Zhou, B. Huang, Z. Yan, and J.-C. G. Bünzli, "Emerging role of machine learning in light-matter interaction," *Light Sci. Appl.* **8**, 1 (2019).
36. D. Piccinotti, K. F. MacDonald, S. Gregory, I. Youngs, and N. I. Zheludev, "Artificial intelligence for photonics and photonic materials," *Rep. Prog. Phys.* **84**, 012401 (2020).
37. J. Moughames, X. Porte, M. Thiel, G. Ulliac, L. Larger, M. Jacquot, M. Kadic, and D. Brunner, "Three-dimensional waveguide interconnects for scalable integration of photonic neural networks," *Optica* **7**, 640–646 (2020).
38. X. Porte, A. Skalli, N. Haghghi, S. Reitzenstein, J. A. Lott, and D. Brunner, "A complete, parallel and autonomous photonic neural network in a semiconductor multimode laser," arXiv:2012.11153 (2020).
39. L.-J. Black, Y. Wang, C. H. de Groot, A. Arbouet, and O. L. Muskens, "Optimal polarization conversion in coupled dimer plasmonic nanoantennas for metasurfaces," *ACS Nano* **8**, 6390–6399 (2014).
40. M. Celebrano, X. Wu, M. Baselli, S. Großmann, P. Biagioni, A. Locatelli, C. De Angelis, G. Cerullo, R. Osellame, B. Hecht, L. Duò, F. Cicaccini, and M. Finazzi, "Mode matching in multiresonant plasmonic nanoantennas for enhanced second harmonic generation," *Nat. Nanotechnol.* **10**, 412–417 (2015).
41. J. S. Jensen and O. Sigmund, "Topology optimization for nanophotonics," *Laser Photonics Rev.* **5**, 308–321 (2011).
42. S. D. Campbell, D. Sell, R. P. Jenkins, E. B. Whiting, J. A. Fan, and D. H. Werner, "Review of numerical optimization techniques for meta-device design [Invited]," *Opt. Mater. Express* **9**, 1842–1863 (2019).
43. F. Meng, X. Huang, and B. Jia, "Bi-directional evolutionary optimization for photonic band gap structures," *J. Comput. Phys.* **302**, 393–404 (2015).
44. S. Osher and J. A. Sethian, "Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations," *J. Comput. Phys.* **79**, 12–49 (1988).
45. T. Feichtner, O. Selig, M. Kiunke, and B. Hecht, "Evolutionary optimization of optical antennas," *Phys. Rev. Lett.* **109**, 127701 (2012).
46. P. R. Wiecha, P. R. Wiecha, C. Majorel, C. Girard, A. Cuche, V. Paillard, O. L. Muskens, and A. Arbouet, "Design of plasmonic directional antennas via evolutionary optimization," *Opt. Express* **27**, 29069–29081 (2019).
47. P. R. Wiecha, A. Arbouet, C. Girard, A. Lecestre, G. Larrieu, and V. Paillard, "Evolutionary multi-objective optimization of colour pixels based on dielectric nanoantennas," *Nat. Nanotechnol.* **12**, 163–169 (2017).
48. D. Z. Zhu, E. B. Whiting, S. D. Campbell, D. B. Burckel, and D. H. Werner, "Optimal high efficiency 3D plasmonic metasurface elements revealed by lazy ants," *ACS Photonics* **6**, 2741–2748 (2019).
49. Y. Augenstein and C. Rockstuhl, "Inverse design of nanophotonic devices with structural integrity," *ACS Photonics* **7**, 2190–2196 (2020).
50. R. Selle, G. Vogt, T. Brixner, G. Gerber, R. Metzler, and W. Kinzel, "Modeling of light-matter interactions with neural networks," *Phys. Rev. A* **76**, 023810 (2007).
51. R. Selle, T. Brixner, T. Bayer, M. Wollenhaupt, and T. Baumert, "Modelling of ultrafast coherent strong-field dynamics in potassium with neural networks," *J. Phys. B* **41**, 074019 (2008).
52. I. Malkiel, M. Mrejen, A. Nagler, U. Arieli, L. Wolf, and H. Suchowski, "Plasmonic nanostructure design and characterization via deep learning," *Light Sci. Appl.* **7**, 60 (2018).
53. S. An, C. Fowler, B. Zheng, M. Y. Shalaginov, H. Tang, H. Li, L. Zhou, J. Ding, A. M. Agarwal, C. Rivero-Baleine, K. A. Richardson, T. Gu, J. Hu, and H. Zhang, "A deep learning approach for objective-driven all-dielectric metasurface design," *ACS Photonics* **6**, 3196–3207 (2019).
54. M. V. Zhelyeznyakov, S. L. Brunton, and A. Majumdar, "Deep learning to accelerate Maxwell's equations for inverse design of dielectric metasurfaces," arXiv:2008.10632 (2020).

55. Y. Li, Y. Wang, S. Qi, Q. Ren, L. Kang, S. D. Campbell, P. L. Werner, and D. H. Werner, "Predicting scattering from complex nanostructures via deep learning," *IEEE Access* **8**, 139983 (2020).
56. S. So, J. Mun, and J. Rho, "Simultaneous inverse design of materials and structures via deep learning: demonstration of dipole resonance engineering using core-shell nanoparticles," *ACS Appl. Mater. Interfaces* **11**, 24264–24268 (2019).
57. A.-P. Blanchard-Dionne and O. J. F. Martin, "Teaching optics to a machine learning network," *Opt. Lett.* **45**, 2922–2925 (2020).
58. P. R. Wiecha and O. L. Muskens, "Deep learning meets nanophotonics: a generalized accurate predictor for near fields and far fields of arbitrary 3D nanostructures," *Nano Lett.* **20**, 329–338 (2020).
59. Y. Zhu, N. Zabaras, P.-S. Koutsourelakis, and P. Perdikaris, "Physics-constrained deep learning for high-dimensional surrogate modeling and uncertainty quantification without labeled data," *J. Comput. Phys.* **394**, 56–81 (2019).
60. T. Chugh, C. Sun, H. Wang, and Y. Jin, "Surrogate-assisted evolutionary optimization of large problems," in *High-Performance Simulation-Based Optimization*, T. Bartz-Beielstein, B. Filipič, P. Korošec, and E.-G. Talbi, eds. (Springer, 2020), pp. 165–187.
61. S. D. Campbell, D. Z. Zhu, E. B. Whiting, J. Nagar, D. H. Werner, and P. L. Werner, "Advanced multi-objective and surrogate-assisted optimization of topologically diverse metasurface architectures," *Proc. SPIE* **10719**, 107190U (2018).
62. V. Kalt, A. K. González-Alcalde, S. Es-Saidi, R. Salas-Montiel, S. Blaize, and D. Macías, "Metamodeling of high-contrast-index gratings for color reproduction," *J. Opt. Soc. Am. A* **36**, 79–88 (2019).
63. A. K. González-Alcalde, R. Salas-Montiel, V. Kalt, S. Blaize, and D. Macías, "Engineering colors in all-dielectric metasurfaces: metamodelling approach," *Opt. Lett.* **45**, 89–92 (2020).
64. R. Pestourie, Y. Mroueh, T. V. Nguyen, P. Das, and S. G. Johnson, "Active learning of deep surrogates for PDEs: application to metasurface design," arXiv:2008.12649 (2020).
65. R. S. Hegde, "Photonics inverse design: pairing deep neural networks with evolutionary algorithms," *IEEE J. Sel. Top. Quantum Electron.* **26**, 7700908 (2020).
66. Z. A. Kudyshev, A. V. Kildishev, V. M. Shalaev, and A. Boltasseva, "Machine learning assisted global optimization of photonic devices," *Nanophotonics* **10**, 371–383 (2020).
67. J. Su, D. V. Vargas, and S. Kouichi, "One pixel attack for fooling deep neural networks," *IEEE Trans. Evol. Comput.* **23**, 828–841 (2019).
68. J. Jiang, D. Sell, S. Hoyer, J. Hickey, J. Yang, and J. A. Fan, "Freeform diffractive metagrating design based on generative adversarial networks," *ACS Nano* **13**, 8872–8878 (2019).
69. R. Trivedi, L. Su, J. Lu, M. F. Schubert, and J. Vuckovic, "Data-driven acceleration of photonic simulations," *Sci. Rep.* **9**, 19728 (2019).
70. D. Liu, Y. Tan, E. Khoram, and Z. Yu, "Training deep neural networks for the inverse design of nanophotonic structures," *ACS Photonics* **5**, 1365–1369 (2018).
71. W. Ma, F. Cheng, and Y. Liu, "Deep-learning-enabled on-demand design of chiral metamaterials," *ACS Nano* **12**, 6326–6334 (2018).
72. L. Gao, X. Li, D. Liu, L. Wang, and Z. Yu, "A bidirectional deep neural network for accurate silicon color design," *Adv. Mater.* **31**, 1905467 (2019).
73. S. An, B. Zheng, H. Tang, M. Y. Shalaginov, L. Zhou, H. Li, T. Gu, J. Hu, C. Fowler, and H. Zhang, "Multifunctional metasurface design with a generative adversarial network," arXiv:1908.04851 (2020).
74. Z. Liu, D. Zhu, S. P. Rodrigues, K.-T. Lee, and W. Cai, "Generative model for the inverse design of metasurfaces," *Nano Lett.* **18**, 6570–6576 (2018).
75. S. So and J. Rho, "Designing nanophotonic structures using conditional deep convolutional generative adversarial networks," *Nanophotonics* **8**, 1255–1261 (2019).
76. A. Mall, A. Patil, D. Tamboli, A. Sethi, and A. Kumar, "Fast design of plasmonic metasurfaces enabled by deep learning," *J. Phys. D* **53**, 49LT01 (2020).
77. Z. Liu, L. Raju, D. Zhu, and W. Cai, "A hybrid strategy for the discovery and design of photonic structures," *IEEE J. Emerging Sel. Top. Circuits Syst.* **10**, 126–135 (2020).
78. W. Ma, F. Cheng, Y. Xu, Q. Wen, and Y. Liu, "Probabilistic representation and inverse design of metamaterials based on a deep generative model with semi-supervised learning strategy," *Adv. Mater.* **31**, 1901111 (2019).
79. X. Shi, T. Qiu, J. Wang, X. Zhao, and S. Qu, "Metasurface inverse design using machine learning approaches," *J. Phys. D* **53**, 275105 (2020).
80. W. Ma and Y. Liu, "A data-efficient self-supervised deep learning model for design and characterization of nanophotonic structures," *Sci. China Phys. Mech. Astron.* **63**, 284212 (2020).
81. D. P. Kingma and M. Welling, "An introduction to variational autoencoders," *Found. Trends Mach. Learn.* **12**, 307–392 (2019).
82. T. Badloe, I. Kim, and J. Rho, "Biomimetic ultra-broadband perfect absorbers optimised with reinforcement learning," *Phys. Chem. Chem. Phys.* **22**, 2337–2342 (2020).
83. H. Wang, Z. Zheng, C. Ji, and L. J. Guo, "Automated multi-layer optical design via deep reinforcement learning," *Mach. Learn. Sci. Technol.* (2020).
84. H. Wang, Z. Zheng, C. Ji, and L. J. Guo, "Automated multi-layer optical design via deep reinforcement learning," *Mach. Learn. Sci. Technol.* **2**, 025013 (2021).
85. E. Ashalley, K. Acheampong, L. V. Besteiro, L. V. Besteiro, P. Yu, A. Neogi, A. O. Govorov, A. O. Govorov, and Z. M. Wang, "Multitask deep-learning-based design of chiral plasmonic metamaterials," *Photon. Res.* **8**, 1213–1225 (2020).
86. J. Trisno, H. Wang, H. T. Wang, R. J. H. Ng, S. D. Rezaei, and J. K. W. Yang, "Applying machine learning to the optics of dielectric nanoblobs," *Adv. Photonics Res.* **1**, 2000068 (2020).
87. J. Peurifoy, Y. Shen, L. Jing, Y. Yang, F. Cano-Renteria, B. G. DeLacy, J. D. Joannopoulos, M. Tegmark, and M. Soljačić, "Nanophotonic particle simulation and inverse design using artificial neural networks," *Sci. Adv.* **4**, eaar4206 (2018).
88. A. Sheverdin, F. Monticone, and C. Valagiannopoulos, "Photonic inverse design with neural networks: the case of invisibility in the visible," *Phys. Rev. Appl.* **14**, 024054 (2020).
89. A.-P. Blanchard-Dionne and O. J. F. Martin, "Successive training of a generative adversarial network for the design of an optical cloak," *OSA Contin.* **4**, 87–95 (2021).
90. Y. Chen, L. Lu, G. E. Karniadakis, and L. D. Negro, "Physics-informed neural networks for inverse problems in nano-optics and metamaterials," *Opt Express* **28**, 11618–11633 (2020).
91. I. Sajedian, H. Lee, and J. Rho, "Double-deep Q-learning to increase the efficiency of metasurface holograms," *Sci. Rep.* **9**, 10899 (2019).
92. A. D. Phan, C. V. Nguyen, P. T. Linh, T. V. Huynh, V. D. Lam, A.-T. Le, and K. Wakabayashi, "Deep learning for the inverse design of mid-infrared graphene plasmons," *Crystals* **10**, 125 (2020).
93. C. Yeung, J.-M. Tsai, B. King, B. Pham, J. Liang, D. Ho, M. W. Knight, and A. P. Raman, "Designing multiplexed supercell metasurfaces with tandem neural networks," *Nanophotonics* **10**, 1133–1143 (2021).
94. T. Asano and S. Noda, "Iterative optimization of photonic crystal nanocavity designs by using deep neural networks," *Nanophotonics* **8**, 2243–2256 (2019).
95. F. Wen, J. Jiang, and J. A. Fan, "Robust freeform metasurface design based on progressively growing generative networks," *ACS Photonics* **7**, 2098–2104 (2020).
96. S. Wang, K. Fan, N. Luo, Y. Cao, F. Wu, C. Zhang, K. A. Heller, and L. You, "Massive computational acceleration by using neural networks to emulate mechanism-based biological models," *Nat. Commun.* **10**, 4354 (2019).
97. J. Jiang and J. A. Fan, "Multiobjective and categorical global optimization of photonic structures based on ResNet generative neural networks," *Nanophotonics* **10**, 361–369 (2020).
98. J. Jiang and J. A. Fan, "Global optimization of dielectric metasurfaces using a physics-driven neural network," *Nano Lett.* **19**, 5366–5372 (2019).
99. K. Deb, *Multi-Objective Optimization Using Evolutionary Algorithms* (Wiley, 2001), Vol. **16**.
100. R. Unni, K. Yao, and Y. Zheng, "Deep convolutional mixture density network for inverse design of layered photonic structures," *ACS Photonics* **7**, 2703–2712 (2020).

101. B. Hu, B. Wu, D. Tan, J. Xu, J. Xu, Y. Chen, and Y. Chen, "Robust inverse-design of scattering spectrum in core-shell structure using modified denoising autoencoder neural network," *Opt. Express* **27**, 36276–36285 (2019).
102. Z. Fang and J. Zhan, "Deep physical informed neural networks for metamaterial design," *IEEE Access* **8**, 24506–24513 (2020).
103. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," arXiv:1505.04597 (2015).
104. N. Borhani, E. Kakkava, C. Moser, and D. Psaltis, "Learning to see through multimode fibers," *Optica* **5**, 960–966 (2018).
105. H. Kabir, Y. Wang, M. Yu, and Q.-J. Zhang, "Neural network inverse modeling and applications to microwave filter design," *IEEE Trans. Microwave Theory Tech.* **56**, 867–879 (2008).
106. C. Zhang, J. Jin, W. Na, Q.-J. Zhang, and M. Yu, "Multivalued neural network inverse modeling and applications to microwave filters," *IEEE Trans. Microwave Theory Tech.* **66**, 3781–3797 (2018).
107. Y.-T. Luo, P.-Q. Li, D.-T. Li, Y.-G. Peng, Z.-G. Geng, S.-H. Xie, Y. Li, A. Alù, J. Zhu, and X.-F. Zhu, "Probability-density-based deep learning paradigm for the fuzzy design of functional metasurfaces," *Research* **2020**, 8757403 (2020).
108. J. Xie, L. Xu, and E. Chen, "Image denoising and inpainting with deep neural networks," in *Advances in Neural Information Processing Systems*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, eds. (Curran Associates, 2012), Vol. **25**, pp. 341–349.
109. H. Gao, L. Sun, and J.-X. Wang, "PhyGeoNet: physics-informed geometry-adaptive convolutional neural networks for solving parameterized steady-state PDEs on irregular domain," *J. Comput. Phys.* **428**, 110079 (2020).
110. Z. Liu, Z. Liu, Z. Zhu, and W. Cai, "Topological encoding method for data-driven photonics inverse design," *Opt. Express* **28**, 4825–4835 (2020).
111. C. C. Nadell, B. Huang, J. M. Malof, and W. J. Padilla, "Deep learning for accelerated all-dielectric metasurface design," *Opt. Express* **27**, 27523–27535 (2019).
112. Y. Qu, L. Jing, Y. Shen, M. Qiu, and M. Soljačić, "Migrating knowledge between physical scenarios based on artificial neural networks," *ACS Photonics* **6**, 1168–1174 (2019).
113. M. Närhi, L. Salmenla, J. Toivonen, C. Billet, J. M. Dudley, and G. Genty, "Machine learning analysis of extreme events in optical fibre modulation instability," *Nat. Commun.* **9**, 4923 (2018).
114. M. Raissi, P. Perdikaris, and G. E. Karniadakis, "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).
115. M. Raissi, A. Yazdani, and G. E. Karniadakis, "Hidden fluid mechanics: learning velocity and pressure fields from flow visualizations," *Science* **367**, 1026–1030 (2020).
116. B. Moseley, A. Markham, and T. Nissen-Meyer, "Solving the wave equation with physics-informed deep learning," arXiv:2006.11894 (2020).
117. Y. Kiarashinejad, S. Abdollahramezani, M. Zandehshahvar, O. Hemmatyar, and A. Adibi, "Deep learning reveals underlying physics of light–matter interactions in nanophotonic devices," *Adv. Theor. Simul.* **2**, 1900088 (2019).
118. Y. Kiarashinejad, S. Abdollahramezani, and A. Adibi, "Deep learning approach based on dimensionality reduction for designing electromagnetic nanostructures," arXiv:1902.03865 (2019).
119. C. Yeung, J.-M. Tsai, B. King, Y. Kawagoe, D. Ho, M. Knight, and A. P. Raman, "Elucidating the behavior of nanophotonic structures through explainable machine learning algorithms," *ACS Photonics* **7**, 2309–2318 (2020).
120. R. Iten, T. Metger, H. Wilming, L. del Rio, and R. Renner, "Discovering physical concepts with neural networks," *Phys. Rev. Lett.* **124**, 010508 (2020).
121. Y. Kiarashinejad, M. Zandehshahvar, S. Abdollahramezani, O. Hemmatyar, R. Pourabolghasem, and A. Adibi, "Knowledge discovery in nanophotonics using geometric deep learning," *Adv. Intell. Syst.* **2**, 1900132 (2020).
122. S. An, B. Zheng, M. Y. Shalaginov, H. Tang, H. Li, L. Zhou, J. Ding, A. M. Agarwal, A. M. Agarwal, C. Rivero-Baleine, M. Kang, K. A. Richardson, T. Gu, J. Hu, C. Fowler, C. Fowler, H. Zhang, and H. Zhang, "Deep learning modeling approach for metasurfaces with high degrees of freedom," *Opt. Express* **28**, 31932–31942 (2020).
123. C. Yeung, J.-M. Tsai, Y. Kawagoe, B. King, D. Ho, and A. P. Raman, "Elucidating the design and behavior of nanophotonic structures through explainable convolutional neural networks," arXiv:2003.06075 (2020).
124. M. Elzouka, C. Yang, A. Albert, S. Lubner, and R. S. Prasher, "Interpretable inverse design of particle spectral emissivity using machine learning," arXiv:2002.04223 (2020).
125. B. Han, Y. Lin, Y. Yang, N. Mao, W. Li, H. Wang, V. Fatemi, L. Zhou, J. I.-J. Wang, Q. Ma, Y. Cao, D. Rodan-Legrain, Y.-Q. Bie, E. Navarro-Moratalla, D. Klein, D. MacNeill, S. Wu, W. S. Leong, H. Kitadai, X. Ling, P. Jarillo-Herrero, T. Palacios, J. Yin, and J. Kong, "Deep learning enabled fast optical characterization of two-dimensional materials," arXiv:1906.11220 (2019).
126. M. Ziatdinov, O. Dyck, A. Maksov, X. Li, X. Sang, K. Xiao, R. R. Unocic, R. Vasudevan, S. Jesse, and S. V. Kalinin, "Deep learning of atomically resolved scanning transmission electron microscopy images: chemical identification and tracking local transformations," *ACS Nano* **11**, 12742–12752 (2017).
127. S. Shao, S. Shao, K. Mallory, K. Mallory, S. S. Kumar, S. S. Kumar, J. Hong, and J. Hong, "Machine learning holography for 3D particle field imaging," *Opt. Express* **28**, 2987–2999 (2020).
128. P. Zhang, S. Liu, A. Chaurasia, D. Ma, M. J. Mlodzianoski, E. Culurciello, and F. Huang, "Analyzing complex single-molecule emission patterns with deep learning," *Nat. Methods* **15**, 913–916 (2018).
129. A. M. Palmieri, E. Kovlakov, F. Bianchi, D. Yudin, S. Straupe, J. D. Biamonte, and S. Kulik, "Experimental neural network enhanced quantum tomography," *npj Quantum Inf.* **6**, 20 (2020).
130. P. R. Wiecha, A. Lecestre, N. Mallet, and G. Larrieu, "Pushing the limits of optical information storage using deep learning," *Nat. Nanotechnol.* **14**, 237–244 (2019).
131. Y. Jo, S. Park, J. Jung, J. Yoon, H. Joo, M.-H. Kim, S.-J. Kang, M. C. Choi, S. Y. Lee, and Y. Park, "Holographic deep learning for rapid optical screening of anthrax spores," *Sci. Adv.* **3**, e1700606 (2017).
132. A. Yevick, M. Hannel, and D. G. Grier, "Machine-learning approach to holographic particle characterization," *Opt. Express* **22**, 26884–26890 (2014).
133. B. Midtvedt, E. Olsén, F. Eklund, F. Höök, C. B. Adiels, G. Volpe, and D. Midtvedt, "Holographic characterisation of subwavelength particles enhanced by deep learning," arXiv:2006.11154 (2020).
134. A. Argun, T. Thalheim, S. Bo, F. Cichos, and G. Volpe, "Enhanced force-field calibration via machine learning," *Appl. Phys. Rev.* **7**, 041404 (2020).
135. M. D. Hannel, A. Abdulali, M. O'Brien, and D. G. Grier, "Machine-learning techniques for fast and accurate feature localization in holograms of colloidal particles," *Opt. Express* **26**, 15221–15231 (2018).
136. J. M. Newby, A. M. Schaefer, P. T. Lee, M. G. Forest, and S. K. Lai, "Convolutional neural networks automate detection for tracking of submicron-scale particles in 2D and 3D," *Proc. Natl. Acad. Sci. USA* **115**, 9026–9031 (2018).
137. S. Helgadottir, A. Argun, and G. Volpe, "Digital video microscopy enhanced by deep learning," *Optica* **6**, 506–513 (2019).
138. I. C. D. Lenton, G. Volpe, A. B. Stilgoe, T. A. Nieminen, and H. Rubinsztein-Dunlop, "Machine learning reveals complex behaviours in optically trapped particles," *Mach. Learn. Sci. Technol.* **1**, 045009 (2020).
139. Y. Rivenson, Y. Zhang, H. Günaydn, D. Teng, and A. Ozcan, "Phase recovery and holographic image reconstruction using deep learning in neural networks," *Light Sci. Appl.* **7**, 17141 (2018).
140. Y. Nishizaki, R. Horisaki, K. Kitaguchi, M. Saito, and J. Tanida, "Analysis of non-iterative phase retrieval based on machine learning," *Opt. Rev.* **27**, 136–141 (2020).
141. K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, "Deep convolutional neural network for inverse problems in imaging," *IEEE Trans. Image Process.* **26**, 4509–4522 (2017).

142. G. Ongie, A. Jalal, C. A. Metzler, R. G. Baraniuk, A. G. Dimakis, and R. Willett, "Deep learning techniques for inverse problems in imaging," *IEEE J. Sel. Areas Inform. Theor.* **1**, 39–56 (2020).
143. G. Barbastathis, A. Ozcan, and G. Situ, "On the use of deep learning for computational imaging," *Optica* **6**, 921–943 (2019).
144. Y. Rivenson, Z. Göröcs, H. Günaydin, Y. Zhang, H. Wang, and A. Ozcan, "Deep learning microscopy," *Optica* **4**, 1437–1443 (2017).
145. E. Nehme, L. E. Weiss, T. Michaeli, and Y. Shechtman, "Deep-STORM: super-resolution single-molecule microscopy by deep learning," *Optica* **5**, 458–464 (2018).
146. W. Ouyang, A. Aristov, M. Lelek, X. Hao, and C. Zimmer, "Deep learning massively accelerates super-resolution localization microscopy," *Nat. Biotechnol.* **36**, 460–468 (2018).
147. E. Nehme, D. Freedman, R. Gordon, B. Ferdinand, L. E. Weiss, O. Alalouf, T. Naor, R. Orange, T. Michaeli, and Y. Shechtman, "DeepSTORM3D: dense 3D localization microscopy and PSF design by deep learning," *Nat. Methods* **17**, 734–740 (2020).
148. T. Pu, J.-Y. Ou, V. Savinov, G. Yuan, N. Papasimakis, and N. Zheludev, "Unlabeled far-field deeply subwavelength topological microscopy (DSTM)," *Adv. Sci.* **8**, 2002886 (2020).
149. T. Pu, J. Y. Ou, N. Papasimakis, and N. I. Zheludev, "Label-free deeply subwavelength optical microscopy," *Appl. Phys. Lett.* **116**, 131105 (2020).
150. D. Bouchet, J. Seifert, and A. P. Mosk, "Optimizing illumination for precise multi-parameter estimations in coherent diffractive imaging," *Opt. Lett.* **46**, 254–257 (2021).
151. A. Ghosh, D. J. Roth, L. H. Nicholls, W. P. Wardley, A. V. Zayats, and V. A. Podolskiy, "Machine learning—based diffractive imaging with subwavelength resolution," arXiv:2005.03595 (2020).
152. U. Kürüm, P. R. Wiecha, R. French, and O. L. Muskens, "Deep learning enabled real time speckle recognition and hyperspectral imaging using a multimode fiber array," *Opt. Express* **27**, 20965–20979 (2019).
153. R. Horisaki, R. Takagi, and J. Tanida, "Learning-based imaging through scattering media," *Opt. Express* **24**, 13738–13743 (2016).
154. L. Yunzhe, X. Yujia, and T. Lei, "Deep speckle correlation: a deep learning approach toward scalable imaging through scattering media," *Optica* **5**, 1181–11819 (2018).
155. B. Rahmani, D. Loterie, G. Konstantinou, D. Psaltis, and C. Moser, "Multimode optical fiber transmission with a deep learning network," *Light Sci. Appl.* **7**, 69 (2018).
156. G. D. Bruce, L. O'Donnell, M. Chen, M. Facchin, and K. Dholakia, "Femtometer-resolved simultaneous measurement of multiple laser wavelengths in a speckle wavemeter," *Opt. Lett.* **45**, 1926–1929 (2020).
157. S. Popoff, G. Leroosey, M. Fink, A. C. Boccara, and S. Gigan, "Image transmission through an opaque material," *Nat. Commun.* **1**, 81 (2010).
158. R. French, S. Gigan, and O. L. Muskens, "Snapshot fiber spectral imaging using speckle correlations and compressive sensing," *Opt. Express* **26**, 32302–32316 (2018).
159. H. Pinkard, Z. Phillips, A. Babakhani, D. A. Fletcher, and L. Waller, "Deep learning for single-shot autofocus microscopy," *Optica* **6**, 794–797 (2019).
160. C. L. Cortes, S. Adhikari, X. Ma, and S. K. Gray, "Accelerating quantum optics experiments with statistical learning," *Appl. Phys. Lett.* **116**, 184003 (2020).
161. Z. A. Kudyshev, S. I. Bogdanov, T. Isacsson, A. V. Kildishev, A. Boltasseva, and V. M. Shalaev, "Rapid classification of quantum sources enabled by machine learning," *Adv. Quantum Technol.* **3**, 2000067 (2020).
162. C. You, M. A. Quiroz-Juárez, A. Lambert, N. Bhusal, C. Dong, A. Perez-Leija, A. Javaid, R. de. J. León-Montiel, and O. S. Magaña-Loaiza, "Identification of light sources using machine learning," *Appl. Phys. Rev.* **7**, 021404 (2020).
163. Y. Rivenson, H. Ceylan Koydemir, H. Wang, Z. Wei, Z. Ren, H. Günaydin, Y. Zhang, Z. Göröcs, K. Liang, D. Tseng, and A. Ozcan, "Deep learning enhanced mobile-phone microscopy," *ACS Photonics* **5**, 2354–2364 (2018).
164. X. Li, J. Dong, B. Li, Y. Zhang, Y. Zhang, A. Veeraraghavan, and X. Ji, "Fast confocal microscopy imaging based on deep learning," in *IEEE International Conference on Computational Photography (ICCP)* (2020), pp. 1–12.
165. J. M. Ede and R. Beanland, "Partial scanning transmission electron microscopy with deep learning," *Sci. Rep.* **10**, 8332 (2020).
166. S. L. Brunton, X. Fu, and J. N. Kutz, "Self-tuning fiber lasers," *IEEE J. Sel. Top. Quantum Electron.* **20**, 464–471 (2014).
167. J. N. Kutz and S. L. Brunton, "Intelligent systems for stabilizing mode-locked lasers and frequency combs: machine learning and equation-free control paradigms for self-tuning optics," *Nanophotonics* **4**, 459–471 (2015).
168. T. Baumeister, S. L. Brunton, and J. N. Kutz, "Deep learning and model predictive control for self-tuning mode-locked lasers," *J. Opt. Soc. Am. B* **35**, 617–626 (2018).
169. A. Youssry, R. J. Chapman, A. Peruzzo, C. Ferrie, and M. Tomamichel, "Modeling and control of a reconfigurable photonic circuit using deep learning," *Quantum Sci. Technol.* **5**, 025001 (2020).
170. B. Wang, J. C. Cancilla, J. S. Torrecilla, and H. Haick, "Artificial sensing intelligence with silicon nanowires for ultraselective detection in the gas phase," *Nano Lett.* **14**, 933–938 (2014).
171. "Google says sorry for racist auto-tag in photo app," <https://www.theguardian.com/technology/2015/jul/01/google-sorry-racist-auto-tag-photo-app> (2015).
172. M. Schuld, I. Sinayskiy, and F. Petruccione, "An introduction to quantum machine learning," *Contemp. Phys.* **56**, 172–185 (2015).
173. M. Krenn, M. Malik, R. Fickler, R. Lapkiewicz, and A. Zeilinger, "Automated search for new quantum experiments," *Phys. Rev. Lett.* **116**, 090405 (2016).
174. A. A. Melnikov, H. P. Nautrup, M. Krenn, V. Dunjko, M. Tiersch, A. Zeilinger, and H. J. Briegel, "Active learning machine learns to create new quantum experiments," *Proc. Natl. Acad. Sci. USA* **115**, 1221–1226 (2018).
175. M. Krenn, M. Erhard, and A. Zeilinger, "Computer-inspired quantum experiments," *Nat. Rev. Phys.* **2**, 649–661 (2020).