Deep learning in light-matter interactions

N.B. When citing this work, cite the original published paper.

(article starts on next page)

Review

Daniel Midtvedt*, Vasilii Mylnikov, Alexander Stilgoe, Mikael Käll,
Halina Rubinsztein-Dunlop and Giovanni Volpe*

Deep learning in light–matter interactions

Abstract: The deep-learning revolution is providing enticing new opportunities to manipulate and harness light at all scales. By building models of light–matter interactions from large experimental or simulated datasets, deep learning has already improved the design of nanophotonic devices and the acquisition and analysis of experimental data, even in situations where the underlying theory is not sufficiently established or too complex to be of practical use. Beyond these early success stories, deep learning also poses several challenges. Most importantly, deep learning works as a black box, making it difficult to understand and interpret its results and reliability, especially when training on incomplete datasets or dealing with data generated by adversarial approaches. Here, after an overview of how deep learning is currently employed in photonics, we discuss the emerging opportunities and challenges, shining light on how deep learning advances photonics.

Keywords: deep learning; neural networks; optics; photonics.

1 Introduction

The interaction of light with matter at the subwavelength scale constitutes the foundation for many applications, ranging from microscopy and nanosensors to data storage and optical communications [1–3]. To optimize the performance of such applications, the ability to predict and analyze light–matter interactions is crucial. Traditionally, these tasks have been based on algorithmically solving Maxwell's equations for a given setup geometry, whose parameters need to be determined from experimental measurements. However, this approach is often time-consuming, applicable only to relatively simple geometries, and very sensitive to measurement noise [4].

Recently, there has been a surge of interest in employing machine learning, especially deep learning, to tackle the limitations of traditional approaches [5]. Briefly, a deep-learning model is an artificial neural network that converts vectors of input data into vectors of output data through a series of transformations characterized by a large number of trainable parameters [6]. The choice of the structure of the network, i.e., its *architecture*, is still mostly a matter of taste and experience rather than a result of clearly established principles. The network must be sufficiently complex to encode the problem at hand but not so complex as to resist training. Once the architecture is defined, the network is typically trained by employing a set of input data with corresponding desired outputs. Networks with thousands of trainable parameters can be systematically optimized using algorithms such as *stochastic steepest descent* and *error backpropagation* [7] in a reasonable amount of time using commonly available computing resources.

The design of photonic devices by machine learning started back in the 1990s with the optimization of microwave circuit components [8], starting with recurrent neural networks [9] and then transitioning to feed-forward neural networks [10]. Later, in the early 2000s, neural networks were used to design photonic crystal fibers [11, 12]. In the last decade, there has been a tremendous rise of attention to neural networks for the inverse design of photonic and plasmonic components [8], including photonic crystals [13–17], layered photonic structures [18–22], radiation cloaks [23–25], diffractive optical elements [26, 27], metasurface-based devices working in different bandwidths [28–47], and nanoparticles [48–50]. Apart from designing photonic structures, deep learning has also been successfully deployed to interpret light–matter interactions, e.g., for microscopy [51] and optical data

*Corresponding authors: Daniel Midtvedt and Giovanni Volpe,
Department of Physics, University of Gothenburg, Sweden,
E-mail: daniel.midtvedt@physics.gu.se (D. Midtvedt),
giovanni.volpe@physics.gu.se (G. Volpe). https://orcid.org/0000-0001-5057-1846
Vasilii Mylnikov and Mikael Käll, Department of Physics, Chalmers University of Technology, Gothenburg, Sweden.
E-mail: vasilii.mylnikov@chalmers.se (V. Mylnikov),
mikael.kall@chalmers.se (M. Käll). https://orcid.org/0000-0002-0144-3909 (V. Mylnikov)
Alexander Stilgoe and Halina Rubinsztein-Dunlop, School of Mathematics and Physics, University of Queensland, St. Lucia, QLD 4072, Australia, E-mail: stilgoe@physics.uq.edu.au (A. Stilgoe),
halina@physics.uq.edu.au (H. Rubinsztein-Dunlop)

storage [52]. Beyond this, recent works have demonstrated all-optical implementations of deep-learning networks, enabling even faster execution times [53].

Despite this widespread interest, successfully deploying a deep-learning solution is still non-trivial. The strength, and weakness, of deep learning is that the user does not provide the rules connecting the input data to the desired outputs. Instead, the system learns these rules by being fed with ground truth input/output pairs and iteratively adjusting its internal trainable parameters until it reliably provides the intended outputs for the training cases. This enables deep-learning-powered approaches to learn to solve specific problems with the utmost efficiency. However, the lack of user-specified rules also makes it difficult to assess the robustness of the performance of a deep-learning model when presented with data that differ significantly from the training set.

This review provides an overview of the recent success stories and opportunities for applying deep learning in photonics while highlighting the most common challenges and pitfalls encountered when solving a problem using deep learning. In Section 2, we will review some of the most common building blocks and architectures for deep-learning-enabled optics and photonics, as well as the most commonly encountered concepts when working with deep learning. In Section 3, we will review some of the main success stories in applying deep learning to optics and photonics. In Section 4, we will provide an overview of the areas where we believe that photonics and deep learning can work synergistically to offer novel opportunities. In Section 5, we will review the essential challenges and provide simple guidelines for effective deployment of deep-learning-based techniques to study light–matter interactions.

2 Current approaches in deep-learning-enhanced optics and photonics

The basic building block of a neural network is an *artificial neuron* (Figure 1A). The artificial neuron performs a weighted sum of inputs and returns a (typically) nonlinear transformation (*activation function*) of the resulting sum. The weights are trainable parameters that are tuned during the learning process to optimize the output [54].

A variant of the artificial neuron is the *memory gate* (Figure 1B). In contrast to the standard artificial neuron, which transforms and feeds forward the information it receives without keeping any memory of it, the memory gate adapts its internal state in response to previous data [55], which is useful, especially when analysing time series. Neural networks containing multiple memory gates are collectively known as *recurrent neural networks* (RNNs).

The *attention gates* provide a more general way to encode dependencies in the input data (Figure 1C), which
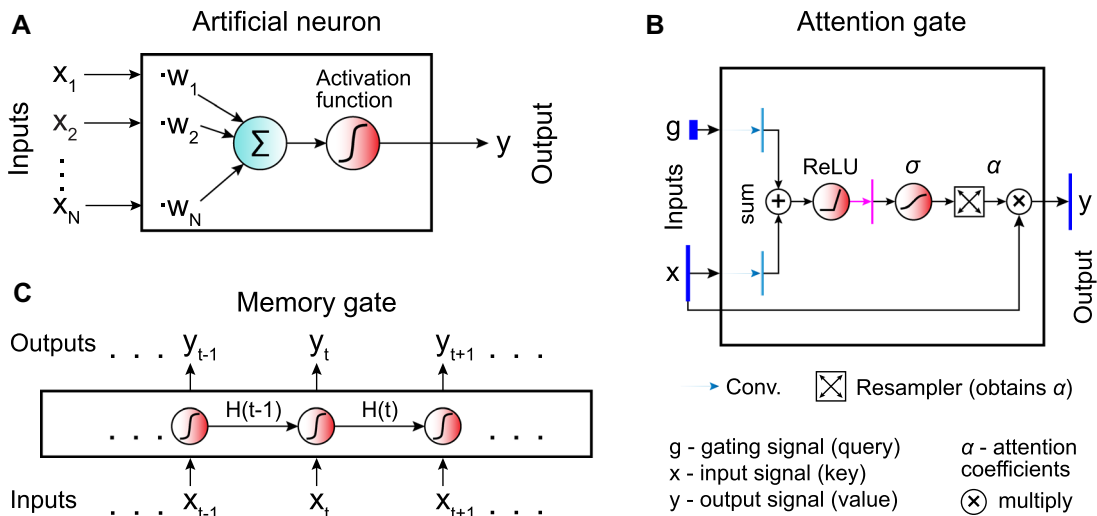

Figure 1: Building blocks of artificial neural networks.
A The basic unit of a neural network is the *artificial neuron*, which performs the sum of its inputs (x_i) weighted by trainable parameters (w_i) and applies an activation function to get its output (y). **B** Temporal information can be encoded by introducing *memory gates*, which retain memory of their history of inputs. The output y_t at the time step t depends on both the current input x_t and the hidden state H_{t-1} obtained in the previous step. Neural networks containing one or multiple memory gates are called *recurrent neural networks* (RNNs). **C** *Attention gates* provide a way to detect long-range temporal or spatial dependencies within the data by guiding the network towards the most relevant parts of the input data.

is often superior to alternative approaches such as RNNs for temporal series and convolutional neural networks for images, particularly for long-range dependencies [56]. The gate receives an input signal x and a gating signal g. These two signals undergo several transformations, which produce the attention coefficients $\alpha \in [0, 1]$ representing the relevance of the elements (e.g., pixels in an image) of the input signal. Then, the resampler associates a unique attention coefficient to each element of the input signal. Finally, the attention gate outputs an element-wise product of the attention coefficients and the input signal. This guides the network to interesting areas of the input data (the neural network pays attention to specific regions).

Dense neural networks (DNNs) consist of connected layers of artificial neurons (Figure 2A). All the nodes in each layer are connected to all the nodes in the neighboring layers (*fully connected network*, also referred to in the literature as a *feed-forward neural network* or *multilayer perceptron-based networks* [57]). Usually, DNNs are utilized for data of small dimension. One could also successfully use DNNs in recognition tasks for small images (e.g., 28 × 28 pixel with 10 predicted classes in the MNIST dataset of handwritten digits [57]). However, when it comes to large images containing thousands of pixels, the number of connections between the neuron layers (and the number of learnable parameters) increases drastically, leading to overfitting. Due to the limited number of neurons in the first hidden layer (which is smaller than the number of pixels in a large input image), the amount of data passed from the input through the ANN is limited.
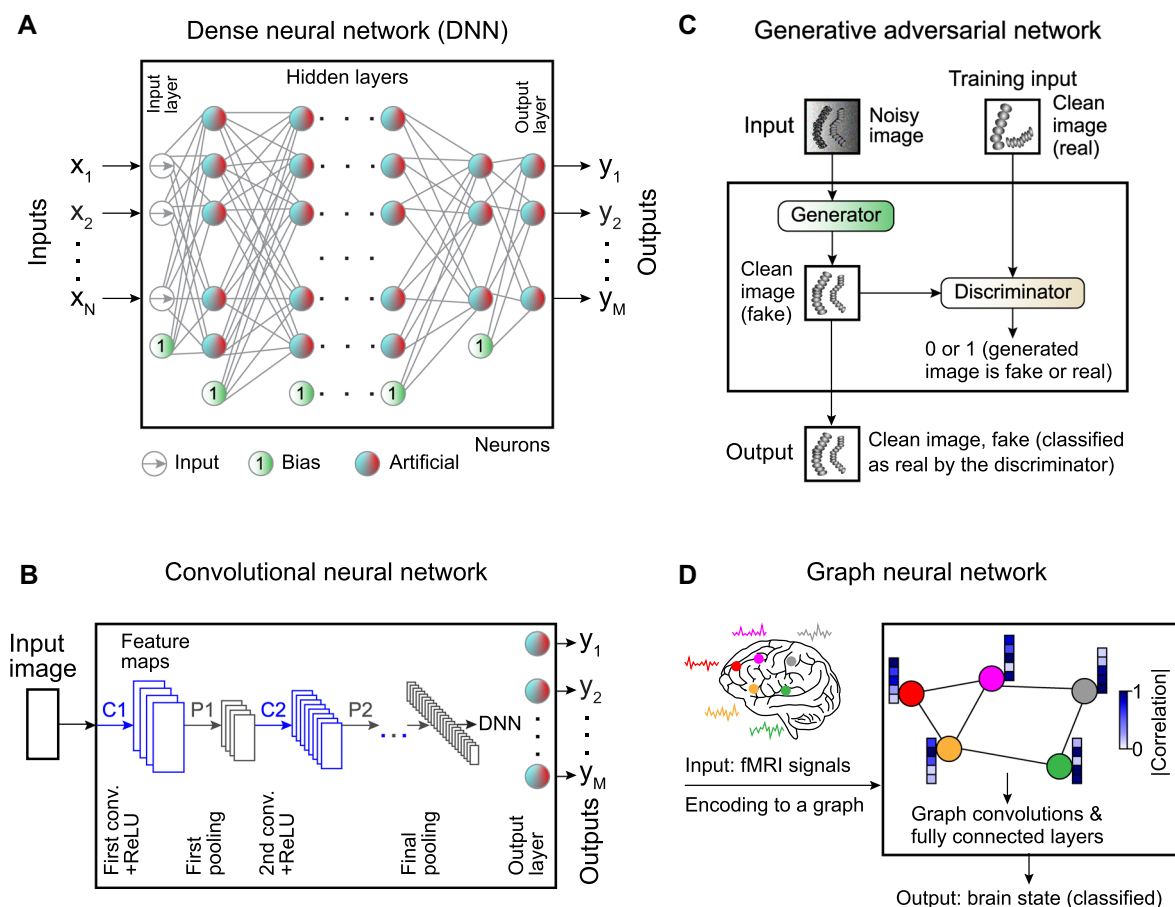

Figure 2: Commonly employed neural networks.
A Artificial neurons can be combined in a *dense neural network* (DNN), where the input layer is connected to the output layer via a set of hidden layers. All the nodes in each layer are connected to all the nodes in the subsequent layer, hence the name "dense." There are also bias neurons that add a constant to the weighted sum. **B** Spatial information can be preserved using a convolutional neural network (CNN) with convolutional kernels whose weights are trainable parameters. **C** *Generative adversarial networks* (GANs) provide a framework to generate synthetic data. Synthetic data is generated by passing input data through a neural network (the *generator*), while a second neural network (the *discriminator*) tries to determine whether the data is real or synthetic. The discriminator's output is used during training to guide the generator to produce more realistic data. **D** *Graph neural networks* (GNNs) provide a powerful method to analyze complex dependencies in the input data using the framework of graph theory.

To circumvent these problems, *convolutional neural networks* (CNNs) [58] use partially connected layers made of 2D arrays of neurons (Figure 2B). Each convolutional kernel (*filter*) uses the same weights across different regions of the image, meaning that the weights are shared. The limited number of connections and weight sharing reduces the number of parameters to be trained, thus, lowering the computational load and the chance of overfitting.

A CNN transforms the input image through a large number of filters [57]. Each filter corresponds to a *feature map*. Typically, one filter detects one feature in the picture, and several filters are used in each *convolutional layer* to detect multiple features. The image is typically downsampled several times to access the information at multiple length scales and reduces the computational load, as schematically shown in Figure 2B. Thus, the image gets smaller while passing through the various layers of a CNN [57]. Sometimes, a dense network is appended to the output of the CNN (called a *dense top*) to produce a final output representing the global information related to the input image, e.g., the coordinates of the position of a particle to be tracked [59].

Deep learning has recently been shown to be able also to generate synthetic data of high quality. Generally, such approaches are helpful to generate outputs that are easily interpretable by humans. Deep-learning-enabled generation of synthetic data is typically achieved through *generative adversarial networks* (GANs) [60] (Figure 2C). GANs are deep-learning models with a unique training scheme called *adversarial training* (training competing ANNs), which is one of the most important recent ideas in machine learning [57]. The idea is that the input data (e.g., a noisy microscopy image) is passed through one neural network called the *generator*, which creates the synthetic data (e.g., the corresponding noise-free image). While training the generator, its output is passed through a second network called the *discriminator*, whose task is to determine whether its input is fake or real data. Usually, both the generator and the discriminator comprise several types of ANN architectures, such as dense neural networks, convolutional neural networks, recurrent neural networks, and graph neural networks. At each training step, the generator's parameters are updated to fool the discriminator. The GAN training consists of many iterations in which the discriminator and generator are both updated in tandem. Within a single training iteration, the GAN update is carried out in two phases [57]. First, the discriminator is trained with the generator's weights fixed. The fake images (labeled as 0, produced by the generator) and the real images (labeled as 1) constitute the training

set in this phase. Second, the generator is trained with the discriminator's weights fixed. The real images are absent in that training phase. Again, the generator creates fake images. In this case, the generator aims to produce data that the discriminator classifies as true images (all 1s). The error vector is being backpropagated through the discriminator to the generator to update its weights. By iterating these two phases multiple times, the generator learns to fool the discriminator, while the discriminator learns to distinguish true images from fake images. In this way, the output of the generator will look more and more realistic.

As a final example, *graph neural networks* (GNNs) provide a powerful method to analyze complex dependencies in the input data in various physics systems where it is necessary to deal with graph data [61] (Figure 2D). A graph is represented by a set of nodes (the data points) interconnected via edges (corresponding to the dependencies in the input data). Depending on the settings, the task of a GNN can be to classify nodes in the graph, predict edges between nodes from incomplete graphs, or generate entire graphs by training on representative data [62]. For example, GNNs have been actively used in the neural science for classification of the neurological disorders (Figure 2D), where they exhibit greater performance than alternative functional magnetic resonance imaging (fMRI) analysis methods [63, 64], and to track microscopic particles, where they have been able to accurately estimate dynamical properties in various biologically-relevant scenarios [65].

3 Success stories

In the past decade, deep learning has found many successful applications within optics and photonics. For most of these applications, the underlying theory connecting the input to the output is unknown or too complex to be of practical use. In such cases, deep learning provides a means to automatize processes that otherwise would require human intervention or large computational resources. In the following sections, we will explore three applications where deep learning has been particularly successful: inverse design of photonic devices, analysis of microscopic and nanoscopic data, and enhancement of microscopy techniques.

3.1 Inverse design

The geometry of a nanostructure corresponding to a desired optical response can be optimized by brute-force

parameter sweeping. For example, in Refs. [66, 67], cylinder-shaped meta-atoms were optimized by sweeping the cylinder diameter D with a fixed height H and recording the optical response as a function of the vacuum wavelength λ to determine the maximum quality factor (Q factor) at the desired wavelength. This was required to lower the lasing threshold and to realize lasing in the smallest nanoparticle possible [67]. Such a brute-force optimization method takes advantage of the scale invariance of Maxwell's equations (i.e., multiplying λ, D, and H by the same constant retains the optical response, permitting the authors to employ dimensionless parameters in sweeping the parameter space). A similar optimization method was implemented in Ref. [68], where a pixelated dielectric metasurface was used to record the absorption fingerprint of a protein. The metasurface was made of meta-atoms arrays (pixels). The geometrical parameters of the meta-atom were linearly scaled by a constant factor to get a different reflectance peak position for each pixel. One hundred pixels were stacked to form a metasurface covering the required bandwidth to map the protein absorption. The brute-force sweeping of the parameter space imposes strict constraints and can easily become computationally cumbersome, especially for geometries with many (>3) parameters.

A powerful alternative is provided by inverse design. Inverse design imposes fewer constraints on the investigated geometries [69], which broadens the solution space and results in more efficient devices. However, the larger design space makes numerical simulations more time-consuming. This demands an efficient method for solution space exploration to lower the simulation time. Such methods can be distinguished into traditional inverse design approaches and inverse design with neural networks.

Traditional inverse design methods explore the solution space iteratively, based on a set of rules. The target is to maximize the fitness function (usually a single number), which is evaluated at every step, after which the system's parameters are adjusted. Stochastic search rules of traditional inverse design methods limit the solution space and the efficiency of the produced devices [69–71]. For example, in Ref. [72], a polarization beam splitter was designed using one of the topology optimization methods known as the *direct-binary search algorithm*. The polarization beam splitter had a square shape (2.4 μm × 2.4 μm), discretized into 20 × 20 square pixels. Thus, each pixel had an area 120 × 120 nm^2 filled either with silicon or air (1 or 0). An initial binary pattern was randomly generated. The state of a random pixel was switched (perturbed), and the fitness function was computed. The fitness function was defined as the average transmission efficiency for TE- and TM-polarized waves. The perturbed pixel state was kept (retracted) if the fitness function increased (decreased). Then, the state of another pixel was perturbed. One iteration consisted of toggling all the pixels sequentially. The procedure was repeated until the fitness function saturated. Since the outcome is sensitive to the initial guess, several initial patterns were considered to achieve the best design. It took more than five days to obtain one design.

In the case of a machine learning-based inverse design, a neural network is trained using many structures with different geometrical parameters (outputs) and the corresponding computed optical responses (inputs). The trained neural network can then be used to obtain the geometry corresponding to a desired optical response. Inverse design with neural networks has the advantage that it is more time-efficient than traditional methods as it does not require case-by-case simulations [70] (once the neural network is trained). Thanks to this feature, the design of new devices is orders of magnitude faster for neural-network-based approaches than for conventional inverse design methods (Figure 3A) [38, 50, 73, 74]. On the other hand, the computational complexity is moved to the generation of the training set. In fact, the quality and size of the training dataset ultimately determine the quality and accuracy of the neural network output [57, 75, 76].

Another advantage of inverse design with neural networks is that it has fewer restrictions on the considered device geometries [78], as compared to conventional methods (e.g., limitation by pre-defined stochastic search rules [72]). Thus, devices and efficiencies unachievable by traditional inverse design approaches can be obtained (Figure 3A and B) [73, 79, 80].

In recent years, neural networks architectures such as DNNs (Figure 2A), CNNs (Figure 2B), GANs (Figure 2C), and hybrid models have been heavily utilized in photonic inverse design [5, 8, 70, 71, 76, 81–86] (Ref. [8] provides an interesting historical perspective). In photonic inverse design, DNNs can be utilized to predict a finite sequence of values. For example, predicting five values of shell thickness in a multilayer particle, given many points of its scattering cross-section spectrum [50]. Besides, DNNs have been successfully used for inverse design of nanoantennas [50, 86, 87], metasurfaces [25, 88–90], and grating couplers [91, 92].

CNNs have been heavily used as parts of generative models for the inverse design of photonic structures [70]. CNNs with a dense top have also been used for inverse
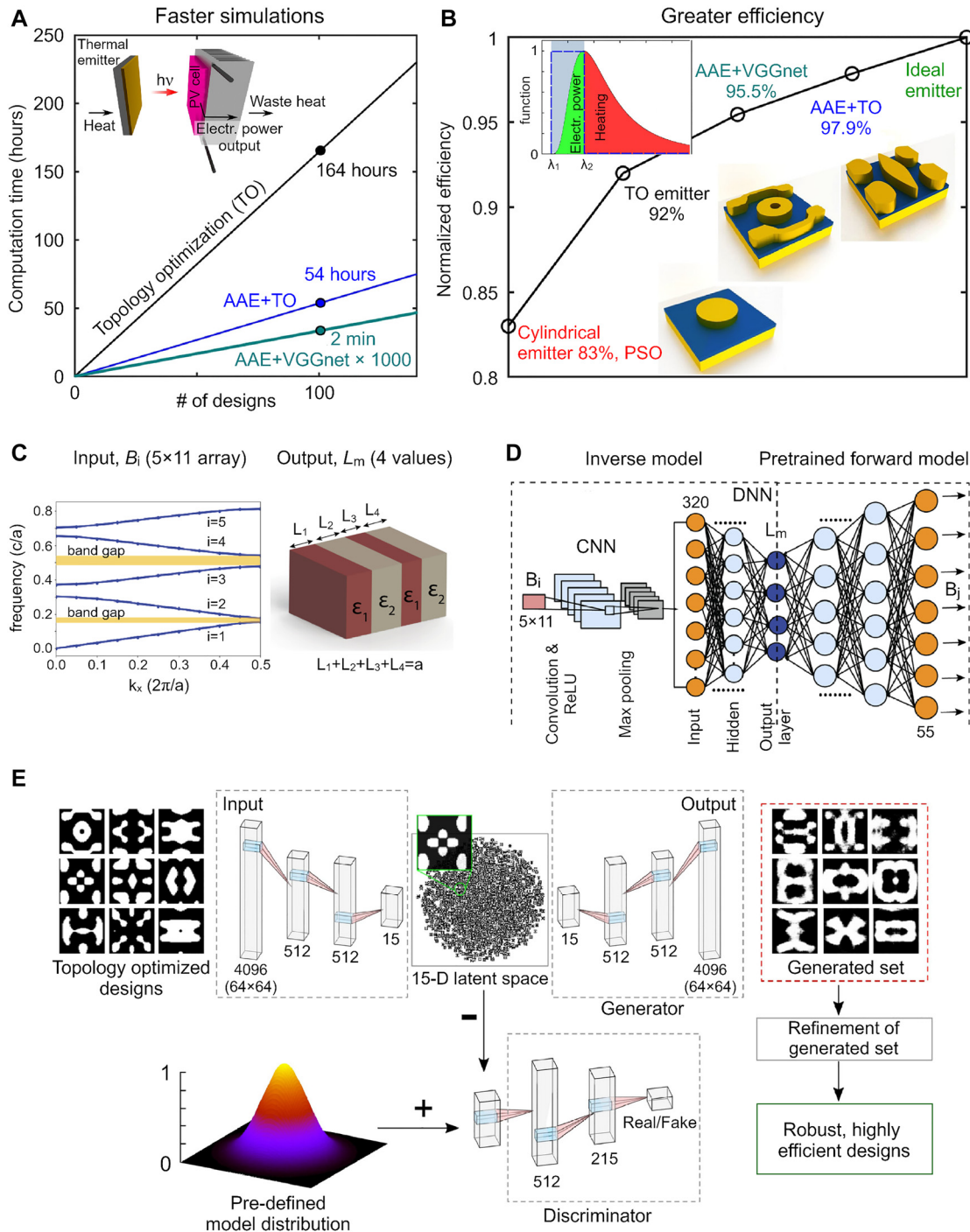
Figure 3: Inverse design of photonic devices by neural networks.
A, B Advantages of neural networks with respect to some conventional simulation methods, namely topology optimization (TO), and particle swarm optimization (PSO). TO is used to generate the training set. TO and adversarial autoencoder (AAE) are compared in the inverse design of the metasurface thermal emitter. TO and CNN VGGnet are employed for structure refinement. **A** Computational time versus number of highly efficient designs. Inset: Schematics of the thermophotovoltaic engine. **B** Efficiencies for the best designs obtained via different simulation methods. Left inset: Normalized spectrum of the black-body radiation; the grey region highlights the photovoltaic cell working band; only the in-band radiation is transformed into electrical power; the blue line depicts the ideal emitter emissivity. Right insets: Finest meta-atom patterns. **C, D** Inverse design of a photonic crystal slab. **C** Band structure (left) used for the inverse design of the 1D photonic crystal slab (right). **D** Architecture of the neural networks implemented in the inverse design. **E** Generative model handling the inverse design of the metasurface thermal emitter. Panels A, B, E are adapted with permission from Ref. [73] (Copyright 2020 American Institute of Physics), and panels C, D from Ref. [77] (Copyright 2020 Optical Society of America).

design [13, 77]. However, DNNs could have been utilized for the latter systems due to moderate number of degrees of freedom. The advantage of using CNNs here is in the detection of more complex patterns in the optical response data, which increases the efficiency of the obtained devices [77]. CNNs in photonics are also utilized for other purposes [70], some of which are discussed later in the text.

In Ref. [77], a CNN with a dense top was used to design a 1D photonic crystal slab, obtaining the geometrical parameters corresponding to a given band structure (Figure 3C). That inverse design has a one-to-many mapping nature, where, e.g., several photonic crystal geometries correspond to a single given band structure (raising some inconsistency issues). This is why some machine learning algorithms fail to converge when handling inverse design problems. The pre-trained forward model with fixed weights was connected to the inverse model to resolve this difficulty (Figure 3D) and the learnable parameters of the inverse model were trained to minimize the cost function defined as an error between the input band structure data B_i and the prediction of the forward model B_j. Training the tandem neural network in such a way circumvents the nonuniqueness issue because the inverse model is not restricted to the production of prior specified designs from the training set [89]. Such an approach has been actively used in photonic inverse design to overcome inconsistency issues [19, 25, 89, 93, 94].

When the degrees of freedom in the design is thousands or greater, it is computationally more efficient to encode the input data to a reduced-dimensional space and reconstruct new designs from it [84]. Deep generative models are capable of doing so and creating new designs similar to the training set but with greater efficiencies. Deep generative models that have been actively used in the design of photonic devices include GANs (Figure 2C) [95–99], variation autoencoders [96, 100–103], and global-topology-optimization networks [8, 104]. For example, in Ref. [73], an adversarial autoencoder was implemented to design a metasurface gap plasmon-based thermal emitter, which was a part of a thermophotovoltaic engine (see inset in Figure 3A), aiming to approach the limit of an ideal emitter (for which the emissivity is equal to one in the desired wavelength range and zero outside, left inset in Figure 3B), essential to reduce the unwanted heating of the photovoltaic cell occurring due to out-of-band radiation. The adversarial autoencoder consisted of three neural networks: an encoder, a decoder/generator, and a discriminator (Figure 3E). The encoder compressed the input 2D meta-atom pattern (64×64 binary image corresponding to a 4096-dimensional vector) into a reduced-dimensional space (15-dimensional latent space) using a DNNs with two hidden layers (512 neurons each). The discriminator was implemented to obtain the latent space within a pre-defined model distribution. The latter was made to represent the solution space continuously by a latent variable (specifically, a continuous Gaussian variable [105]). The decoder generated a 2D binary image of the meta-atom (4096-dimensional vector) out of its input 15-dimensional latent vector. The neural networks learned to obtain a continuous representation of the training data in the reduced-dimensional latent space. After training the adversarial autoencoder, new designs were generated by decoding the sampled latent vector. Then, the generated designs were refined using topology optimization and a CNN VGGnet (Figure 3A and B). The structure refinement smoothed the meta-atom patterns by ruling out sub-30-nm features and keeping the designs with the highest estimated efficiencies. The training set consisted of 8400 samples. Such a large dataset is typically required for adversarial autoencoder training. Generating a training set of such size using topology optimization is time-consuming. Thus, a set of 200 samples was obtained by topology optimization, and the actual training set of 8400 samples was generated by data augmentation. Specifically, the training set of 200 samples was expanded by 20 random lateral translations and a single 90° rotation of the meta-atoms. Thanks to the periodicity and symmetry of the metasurface thermal emitter structure, these perturbations did not affect the emissivity spectra. Such a method allowed for augmentation of the training dataset without additional full-wave simulations.

3.2 Image analysis in microscopy and nanoscopy

The analysis of experimental data is another area where deep learning has been successfully deployed. Such analysis is often time-consuming, requiring human input. With modern data acquisition techniques, the amount of experimental data can easily exceed what is feasible to analyze using conventional methods, making data analysis the limiting factor in many experiments. This is particularly true in microscopy. Every microscopy image may contain millions of pixels, and the designation of a simple rule connecting the individual pixel values to the desired output is non-trivial. Unsurprisingly, in the past decade, deep learning has created a new paradigm for the analysis of microscopy images, making precise, automatized, and

objective data analysis possible at speeds orders of magnitude faster than for conventional methods. In addition, deep-learning-powered approaches have recently demonstrated the capability to extract information beyond the limits of traditional methods, making it possible to analyze data with unprecedented details.

The most well-known example of deep-learning-enhanced image analysis is that of image classification. The task is to classify objects in an image into predefined classes. The input is typically a cropped version of the whole input image, containing only a single object. The neural-network architecture is typically a CNN, which enables extracting object features at multiple scales with reasonable computational cost. As an example, Ref. [106] used a specific CNN architecture called Inceptionv3 to classify and predict mutations from lung cell histopathology slides. The neural network analyses nonoverlapping tiles in the image, providing a single classification score for each tile. The result is a downsampled image where every pixel corresponds to the classification of a specific tile.

As another example, Ref. [107] proposed a deep-learning framework for whole-slide classification for cervical cancer screening. There, each slide can contain tens of thousands of cells. Therefore, the manual identification of lesion cells can be highly time-consuming. The authors propose a three-stage classification scheme: first, a CNN analyses a low-resolution image of the entire slide, indicating suspicious regions. After that, a second CNN analyses high-resolution images of the areas proposed by the first CNN and outputs a probability that the region contains a lesion cell. Finally, the ten highest-scoring regions are analysed by an RNN that provides a final score for the whole slide (Figure 4).

Classification can also be performed on a pixel-by-pixel basis. One such example is image segmentation, where the task is to classify each pixel as either belonging to an object or to the background. This task is challenging using conventional image analysis techniques if the objects do not strongly contrast with the background, which is often the case for biological imaging. In such cases,
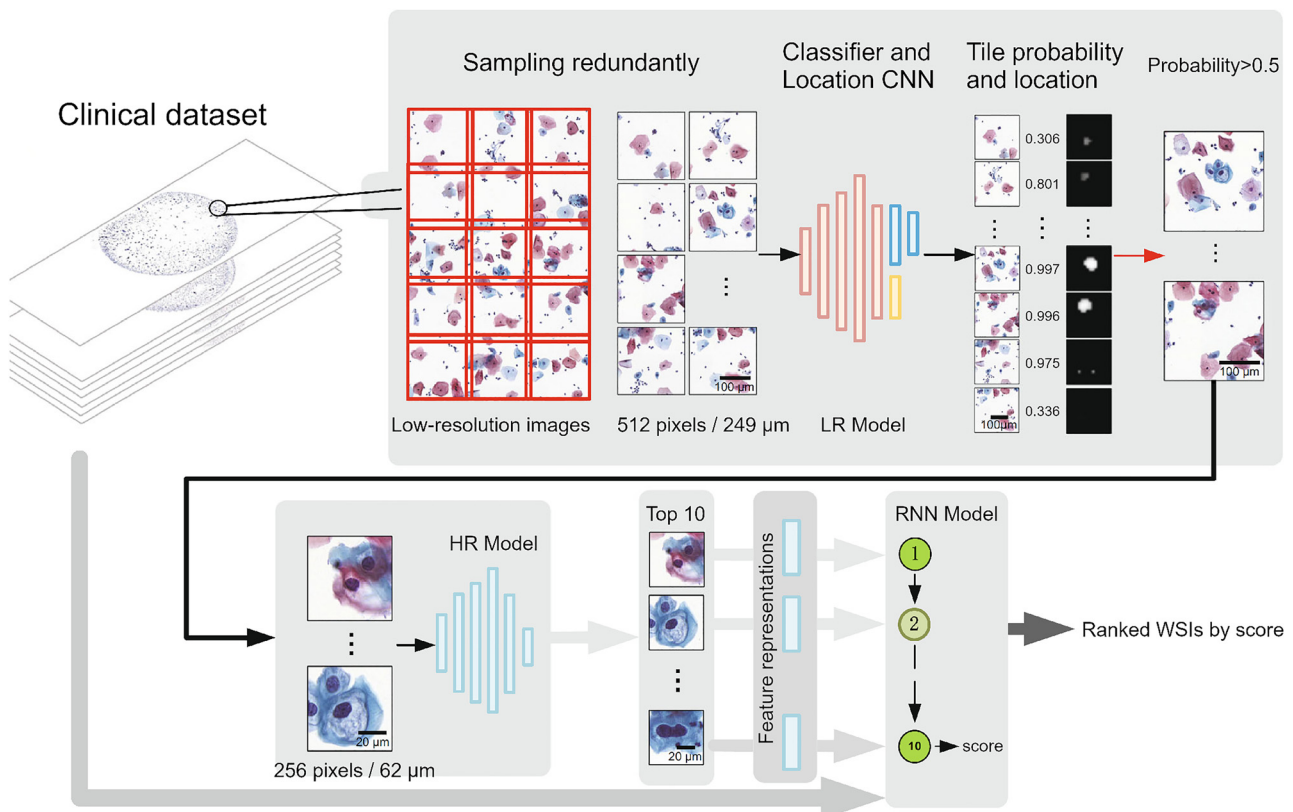

Figure 4: Image analysis for a coarse-grained classification of histopathology slides using a CNN.
A low-resolution image of the entire slide is divided into non-overlapping tiles, which are independently passed through a low-resolution model that locates suspicious lesion regions. These regions are identified at high resolution by the high-resolution model that outputs a probability of the presence of a lesion cell in a tile. This high-resolution model outputs the ten most suspicious lesion tiles. Finally, these ten highest-scoring tiles are analyzed by an RNN to produce a final score for the entire slide. Image reproduced with permission from Ref. [107] (Copyright 2018 Springer-Nature).

the standard approach is the manual segmentation of the images. In 2015, Ref. [108] introduced the *U-Net architecture*, which enables efficient classification of each pixel in the image and automated segmentation of biomedical images [109] (Figure 5A). The U-Net architecture is a special type of CNN that takes an image as input and transforms it into another image. The segmentation of biomedical tissue images [109]) is achieved by first downsampling and then upsampling the image through a series of convolutional layers. Again, downsampling is performed to detect the features in the image and reduce the computational load, while upsampling is used to reconstruct the picture and carry out the segmentation. The downsampled images at the contracting part of the U-Net are concatenated with the images at the corresponding levels of the expanding part. This is performed to preserve local, high-resolution information about the picture. Another study employed a multichannel U-Net model to segment fluorescence images with heterogeneous marker combinations [110]. In particular, it showed that using an attention module makes the resulting model robust to the missing module problem, where only a limited subset of the marker combinations are available.

A further step in the analysis of microscopy images is to quantify the properties of the objects in the field of view. Examples include precisely identifying the location of objects in an image [59, 112], quantification of the scattering properties of objects from microscopic images [51, 113], and analysis of particle motion from sequences of images for characterizing anomalous diffusion [114] or the underlying force field [115]. A prime example of data regression is in object localization. Localizing objects in microscopic images has traditionally been based on thresholding techniques, in which adjacent pixels sharing similar intensities are grouped to form an assessment of where the objects are located in an image. The success of such pixel-by-pixel-based techniques requires the pixelwise intensity of the objects to be well separated from that of the background. In many cases, such as in brightfield imaging or in the presence of noise, that assumption may not hold, making particle localization challenging. In contrast to the pixel-by-pixel-based thresholding analysis, CNNs instead extract features of images at multiple length scales. This enables the network to identify *spatial correlations* in the image data and to learn to exploit these correlations to classify the presence or absence of objects. This approach has been employed in various experimental situations and has enabled automated particle localization in challenging conditions. For example,

in Ref. [59], the authors use a CNN to accurately determine the location of a single particle within a small image region (101×101 pixels). This approach was demonstrated to achieve higher localization accuracy than traditional algorithmic approaches, particularly under poor signal-to-noise ratios (SNR) (Figure 5B). This approach can be helpful for prediction refinement, assuming that some other method has already identified potentially interesting regions in the image. In Ref. [111], the authors instead used a U-Net to directly localize single-molecule emitters in an entire field of view, providing 3D localization as well as an estimate of emitter intensity in a single shot (Figure 5C).

3.3 Multimodal and transfer microscopy

Image classification and regression seek to reduce the information content of an image into a list of numbers. Beyond this, deep learning has proven effective in *image-to-image transformation*, in which the objective is to transform the input image into another image for further data processing. Examples include super-resolution imaging [116, 117], 3D volumetric imaging [118, 119], cross-modality transformation [120–122], and speckle pattern deconstruction [123]. As an example, in Ref. [116], a GAN (Figure 2C) was utilized to artificially enhance the resolution of optical microscopy images. The authors collected low- and high-resolution images of cells and nanoparticles. Notably, the low- and high-resolution images were obtained sequentially on the same field of view of each sample. The task of the neural network was to output the high-resolution image given the low-resolution one as an input. The network outperformed standard image deconvolution algorithms and matched the resolution of the optical method used to acquire the ground truth data. Furthermore, in Ref. [124], the authors demonstrated that by including network layers that analyze the Fourier spectrum of the input images, the details in the obtained super-resolution images were improved (Figure 6A). Deep learning can also be used to transform between different imaging modalities [120–122, 125, 126]. As an example, Ref. [120] used a GAN to transform holographic images into brightfield images, enabling volumetric imaging without the speckle noise typically associated with coherent imaging techniques (Figure 6B).

At this point, a word of caution is required. While deep-learning models can be trained to perform seemingly impossible transformations of the input data, these models cannot learn beyond the information content of
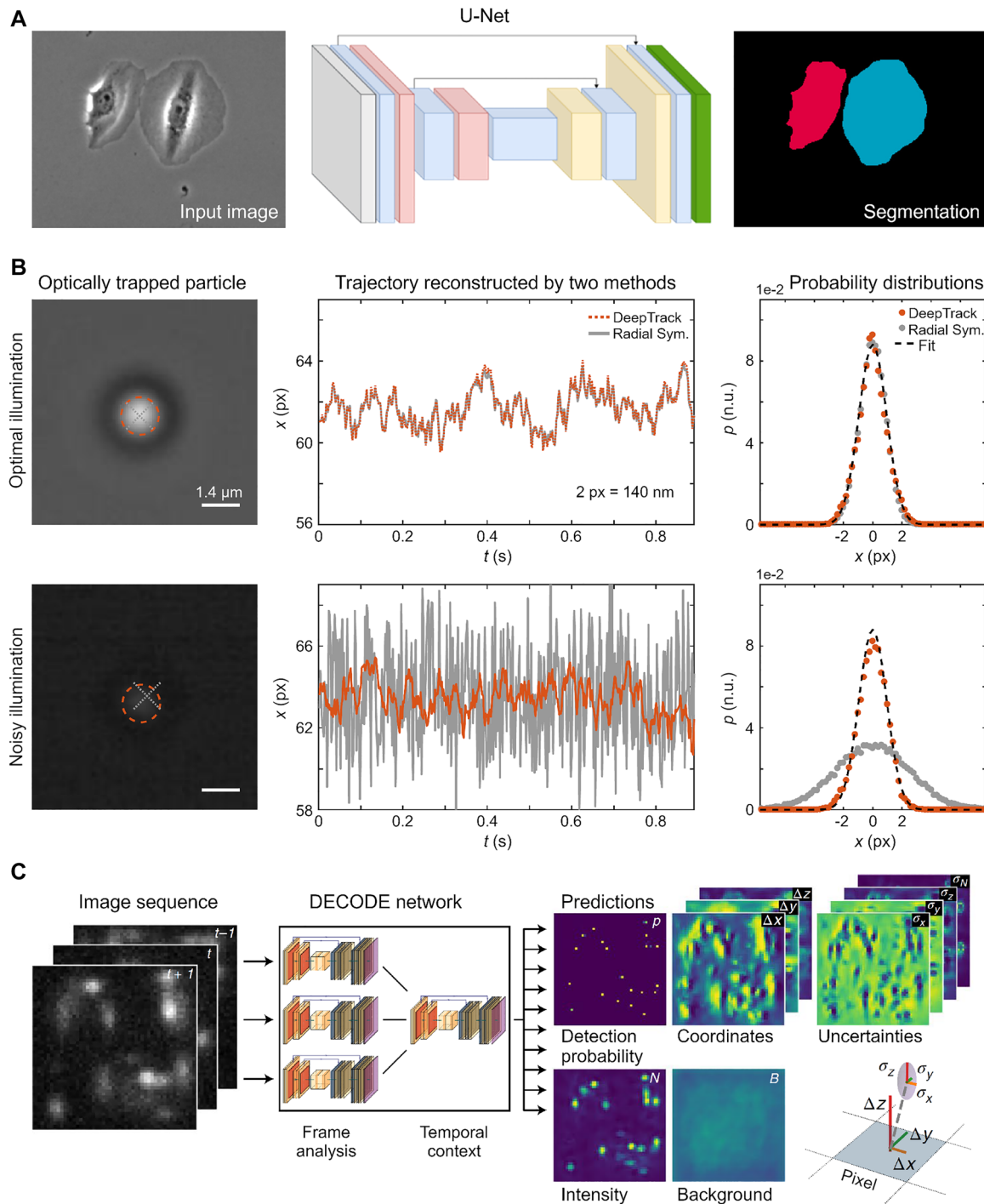
Figure 5: Image analysis by neural networks.
A Pixel-wise classification to segment objects in microscopy images. The method is based on the U-Net architecture and outputs a binary image corresponding to cells (white) and background (black). **B** Particle localization using a CNN. The CNN is trained using simulated images of microscopic particles. The neural network outputs the position of a particle within a region of fixed size. This model (orange lines) outperforms traditional approaches (gray lines) in terms of accuracy, in particular for noisy images (lower panels). **C** Particle localization using a U-Net to identify the position and intensity of particles pixel by pixel. For each pixel, the network predicts the probability that an emitter exists near that pixel, the intensity of that pixel, the three-dimensional vector connecting the pixel to the closest emitter, as well as an estimate of the localization uncertainty. Panel A is reproduced with permission from Ref. [109] (Copyright 2019 Springer-Nature), panel B from Ref. [59] (Copyright 2019 Optica), and panel C from Ref. [111] (Copyright 2021 Springer-Nature).
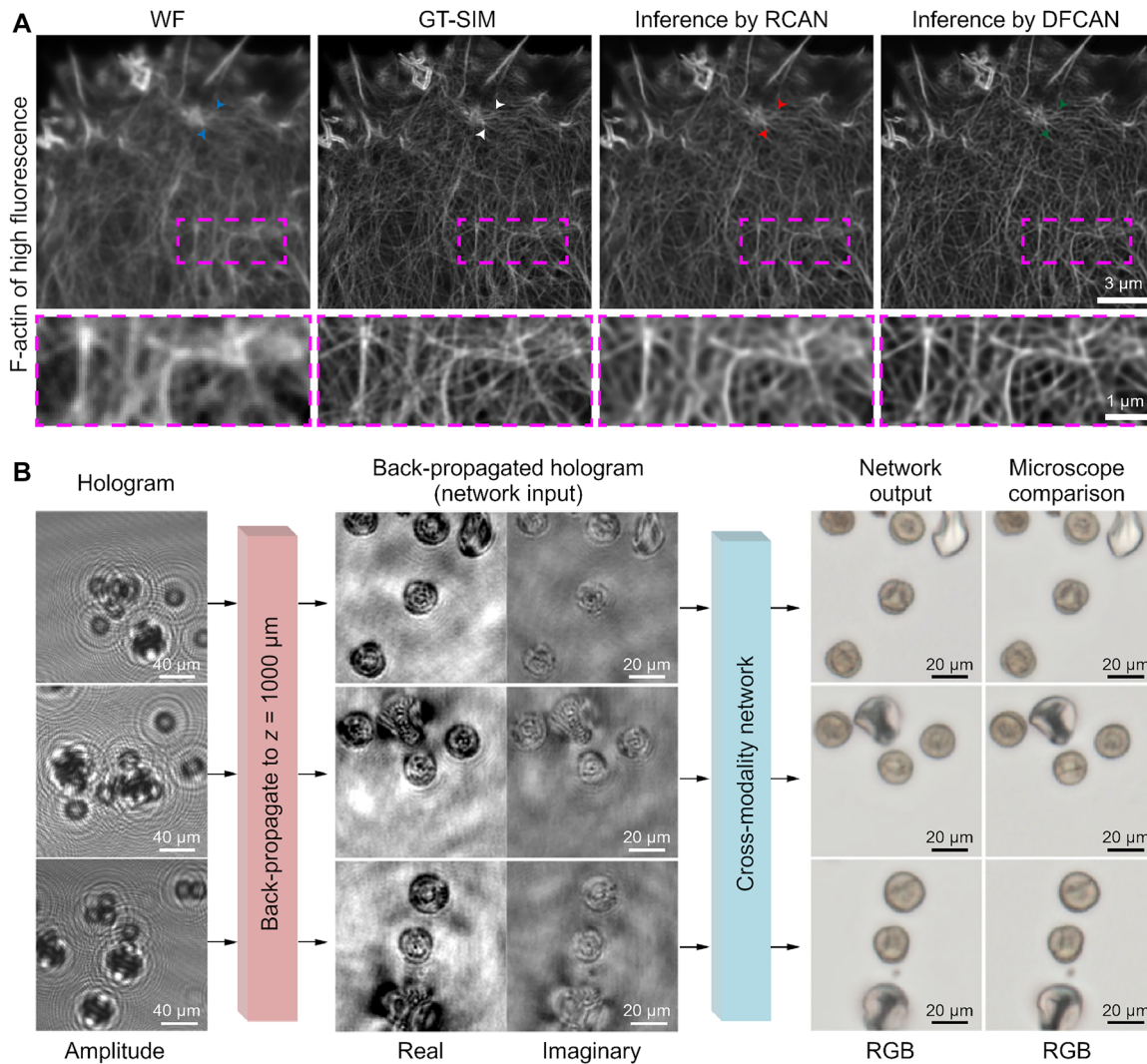
Figure 6: Multimodal and transfer microscopy.
A Deep learning can be used to enhance the resolution of microscopy images. The core idea is to record sets of low-resolution images (the wide-field (WF) image exemplified in the leftmost panel) and high-resolution images (ground truth structured illumination microscopy (GT-SIM) image exemplified in the panel second to the left) of the same field of view, on the same optical setup. The neural network, often a GAN, is then trained to reconstruct the high-resolution image from the low-resolution image as input. The output examples are shown in the two rightmost panels, F-actins inferred by residual channel attention network (RCAN) and deep Fourier channel attention network (DFCAN). **B** A different way to utilize deep learning is by training models to convert between different microscopy modalities. In this example, a GAN was trained to predict brightfield images using inline holographic images as input. Panel A is reproduced with permission from Ref. [124] (Copyright 2021 Springer-Nature), and panel B from Ref. [120] (Copyright 2019 Springer-Nature).

the training data. In other words, deep-learning models can be trained to transform microscopy images from a high-information content image to a low-information content representation, while going in the opposite direction requires the model to extrapolate information based on knowledge it acquires from information in the training data. Utilizing such extrapolated data for further analysis and decision-making is risky, as it relies on information not present in the input data. Depending on the purpose of the analysis, it may be more robust to train a model to explicitly extract relevant features from the input image directly instead of first performing a potentially error-prone modality transformation.

4 Opportunities

Going beyond the success stories presented in the previous sections, deep learning remains underutilized in several areas. This section provides an overview of the areas where we believe that photonics and deep learning can still work synergistically to offer novel opportunities.

4.1 Quantitative data analysis

Deep-learning models can quantitatively measure the properties of their input data. This becomes particularly useful with several microscopy techniques that are today considered qualitative due to the difficulty of extracting quantitative information from light scattering data. For example, brightfield microscopy is arguably the most widely used microscopy technique found in essentially every scientific laboratory. However, due to the low image contrast and incoherent illumination, it is extremely challenging to extract quantitative information from brightfield images, so brightfield microscopy is commonly used as a qualitative technique. However, brightfield images are superpositions of scattering patterns formed by multiple wavelengths of light and contain vast amounts of information. For example, virtual staining of quantitative phase images [126] and brightfield images [122] has recently been demonstrated, transforming such images into synthetic fluorescence images where specific structures have been stained. Importantly, for brightfield imaging, the resulting structures have been shown to *quantitatively* reproduce both the morphologies and fluorescence intensities of the corresponding structures (Figure 7A–C) [122]. This procedure circumvents the limitations of brightfield imaging, namely its low image contrast and challenging data interpretation. Besides, it overcomes the limitations of fluorescence staining, namely potential toxicity, fluorescence bleaching effects, and variability in results between different professionals performing the stain, in this way transforming brightfield into a quantitative microscopy technique.

In other instances, the data itself may be quantitative, but the complexity of the data might make quantitative analysis challenging. For example, in hyperspectral imaging, each pixel contains a wide spectral profile that can be used to characterize an object. The vast amount of data available in each image makes it difficult to analyze it using conventional techniques. Ref. [129] uses a U-Net to predict the location of multiple drugs in a mouse liver from mass spectroscopic images, showing the potential of deep learning for the analysis of such complex data. Furthermore, Ref. [127] demonstrates stimulated Raman spectroscopic imaging of biological samples in the fingerprint region $(400–1800\,\mathrm{cm}^{-1})$. Raman scattering in this spectral region provides "fingerprints" of the chemical composition of a sample and is, therefore, a highly useful characterization technique for many materials. For biological materials, the Raman scattering in this region is very weak, and it has been challenging to perform fingerprint Raman imaging on such samples. Using a U-Net architecture, the authors demonstrate an enhancement of the contrast of Raman spectroscopic images, which enables quantitative measurements of the lipid and protein contents of living cells (Figure 7D–F).

As a different example, the scattering patterns of nano- and microparticles contain information about their size, refractive index, and shapes. However, extracting these properties from experimental scattering patterns is challenging, as it requires solving the inverse problem for Maxwell's equations. Instead, deep learning-powered solutions, which fit the Mie theory to the experimentally obtained scattering patterns, have enabled direct sizing and refractive index determination of particles with radii from about 100 nm to μm [113, 128] (Figure 7G–I).

As a final example, surface plasmon resonance is an optical effect occurring in metallic nanoparticles when they are irradiated. By measuring the absorbance of such nanoparticles as a function of the angle of incidence, the refractive index of the immediate surrounding of the nanoparticles can be determined. That is often used to characterize nanofilms and surfaces. However, when used as an imaging modality, the angle of incidence is kept fixed. The method is then used for qualitative investigations of surfaces rather than as a quantitative characterization tool. Ref. [130] demonstrates quantitative refractive index measurements for surface plasmon resonance imaging through the development of a deep learning-powered phase retrieval algorithm.

4.2 Nanophotonics for deep learning

Various systems are potential platforms for practical neuromorphic computing [53]. Some of the desirable aspects of using photonic-based systems instead of silicon floating-point units are their speed and low power consumption [131, 132]. Furthermore, photonic information processing systems are highly configurable and can process in parallel with amplitude, phase, polarization, and frequency [133], which naturally leads to a wide array of useable techniques. In this section, we will take a closer look at deep learning developments in photonics systems with a focus on systems that have been experimentally realized. They offer a good prospect for implementing neural networks, particularly with photonic chips and 3D printing [62, 134–141]. Our discussion will start with how non-linear activation functions are achieved in photonic systems, followed by the construction of fundamental neural network units such as perceptrons and multiply-and-accumulate (MAC) units, and finally, how these systems can be cascaded (layered) to implement deep learning.
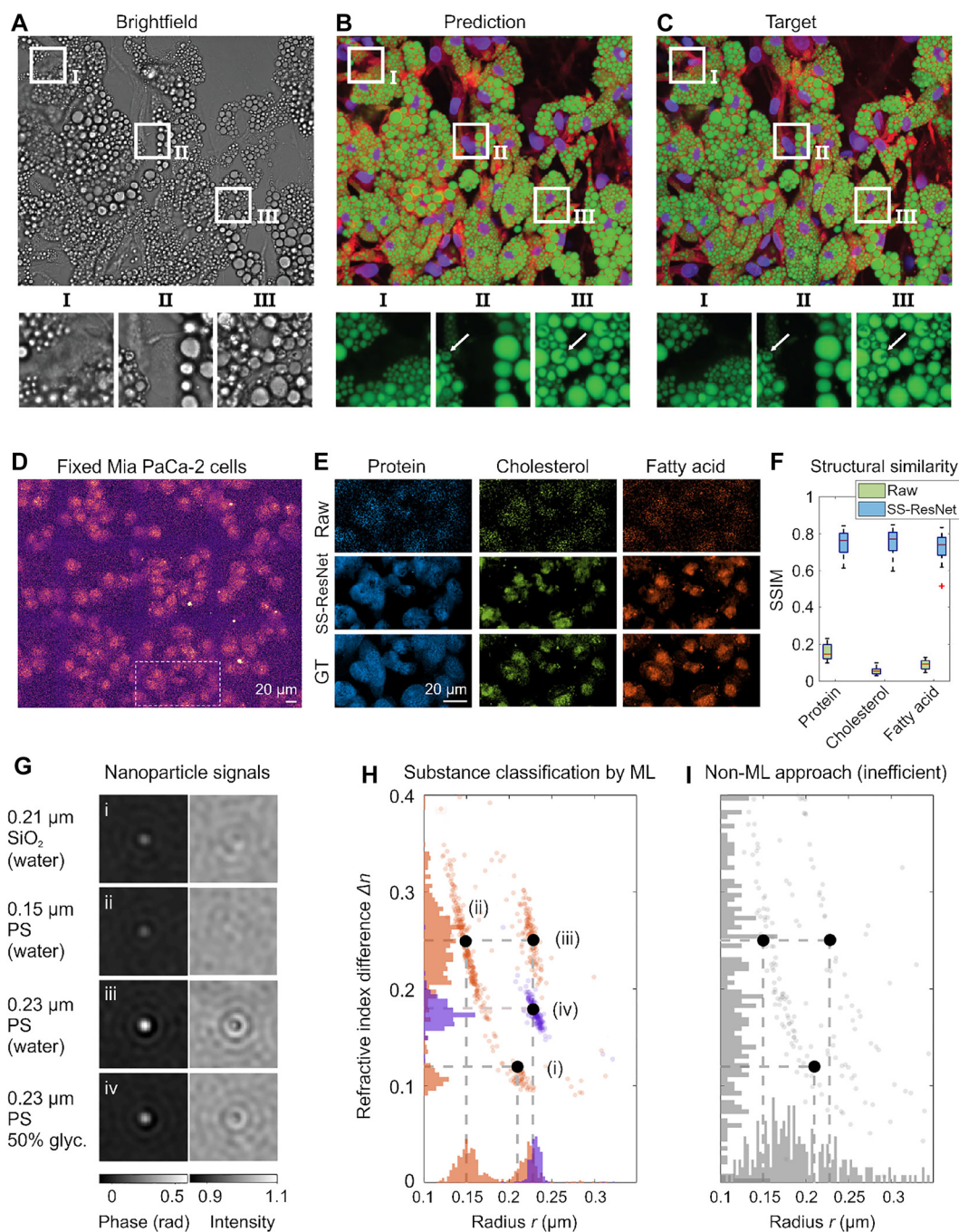
Figure 7: Quantitative imaging with deep learning.
A−C Brightfield images of adipocytes **A** converted into virtually stained fluorescent images **B** and comparison with the relative chemically-stained ground truth **C**. The GAN was trained to reproduce images where lipid droplets (green), cytoplasm (red), and nuclei (violet) were separately stained. The resulting virtually stained structures were shown to quantitatively match the fluorescently stained structures in size and morphology, demonstrating that the network learns to distinguish different cellular structures in brightfield images based on their interaction with the illuminating light. **D−F** Raw Raman scattering signals from biological samples **D** retain quantitative information about the Raman spectrum of the sample, enabling the quantification of lipids and proteins in living cells using a U-Net **E**. **F** Structural similarity (SSIM) index for raw and spatial-spectral residual net (SS-ResNet) normalized by the ground truth of the three chemical channels. **G−I** Holographic images of nanoparticles contain quantitative information about the size and refractive index of the particles. This information can be decoded using deep neural networks to provide much-improved characterization accuracy compared to traditional methods. Panels A−D are adapted with permission from Ref. [122] (Copyright 2021 AIP), panels D−F from Ref. [127] (CC-BY), and panels G−I from Ref. [128] (Copyright 2021 ACS).

A key element of sophisticated classical computing systems is stable and coherent switching behavior between two definite states. Neural networks are a model which incorporates non-linear switching behavior with linear interconnects. The choice of non-linear transfer functions changes from implementation to implementation and from application to application. One of the most common examples is the sigmoid function, which can be realized optically using a deeply saturated differentially biased semiconductor optical amplifier [142]. The pulse train in natural neural networks can be mimicked in systems that resemble the Hodgkin−Huxley circuit model [143]. In this model, the injected current is considered a phase-space parameter. Below the threshold current, the potential difference remains small, and no pulsing behavior is observed. At the threshold current, the dynamics of the circuit switches through a Hopf bifurcation, and periodic solutions can chaotically emerge. Passive optical amplifiers and cavities with parametric instability can also follow this behavior [144]. In general, any kind of bifurcation can be used [145]. These systems are considered for use in artificial neural networks because of support for Hebbian learning models (changes in neuron−neuron association with stimulation in time) in living organisms [146, 147].

The photonics-based schemes can be all-optical [137, 148−151] or electro-optical [140, 152−156] (see also reviews [157−159] for more details). The optical and electro-optical implementations of the activation functions in actual neural networks include phase change materials [137, 160, 161], Fano resonances in nanostructures [162], non-linear states (bifurcation) [145], wavelength-division multiplexing using optical amplifiers [142], electric-optic modulators [152−154], vertical-cavity surface-emitting lasers (VCSELs) [163, 164], passive mode locking with quantum dot [165], chip-based electro-optic feedback circuit [155], and Kerr non-linearity [139]. Only a few of these methods will survive in products in the future due to limits to the level of miniaturization that can be achieved for the corresponding physical process. Compared to the all-optical methods, electro-optical methods can be advantageous due to their relative simplicity in creating activation functions such as rectified linear units (ReLUs). Alas, this comes at the cost of additional waste heat production, limiting miniaturization. A comprehensive and more technical review of these different technologies can be found in Ref. [141].

Summation is an important part of data manipulation and transformation. The core functionality of a perceptron is to calculate an output that is a function of a weighted sum of its inputs. In CNNs [166], the data is multiplied element-wise with some kernel and summed. In RNNs [167], there is an implicit integration in the feedback. The details of how physical implementations of these data structures are made are important because they ultimately determine the speed, accuracy, and efficiency that make a computing system operate. In the single-mode domain, the summation of photons can be performed either by intensity or complex phase (vector/scalar field) [168]. For example, chip-based Mach−Zehnder interferometers have been used for vowel recognition [136] and could conceivably be adapted for general CNN and RNN tasks.

Integrated photonics has the potential to perform vast parallel matrix multiplications. Currently, over a trillion (10^{12}) MACs have been demonstrated with integrated photonics [140, 169]. By virtue of it being based on light (photons), which exhibits weak coupling in dielectrics, and the ease at which different frequency components can be separated, it has two potential advantages over electronic counterparts: such a system can use wavelength-division multiplexing to separate data into discrete streams of information, and it potentially exhibits much lower dispersion than electron-only systems, which enables high modulation rates. The wavelength-division multiplexing scheme and comparison with traditional computing for the same data is shown in Figure 8A−C.

Although neuromorphic computing has been shown to greatly accelerate and reduce the energy requirements of the inference stage of deep learning models, the actual training of such computing systems has proven to be a challenge. One reason for this is that the backpropagation algorithm, which is the pervasive algorithm for training neural networks, cannot be implemented directly in a physical system. This challenge has been recently overcome by combining physical systems with a numerical model which emulates the behavior of the physical system, demonstrating efficient training of physical neural networks using backpropagation [170] (Figure 8D).

In a free-space optical system, convolutions come naturally out of considering the equations of light propagation where the propagation kernel itself is a convolution. This provides an opportunity to produce a deep neural network based on the printing of successive complex amplitude filters, such as demonstrated for a diffractive deep neural network (D2NN) [134] (Figure 8E−G). In this work, a machine learning model of a physical system was first modeled on a computer, and the resulting stacked complex filtering structures were printed using lithographic techniques. This neural network was able to identify objects
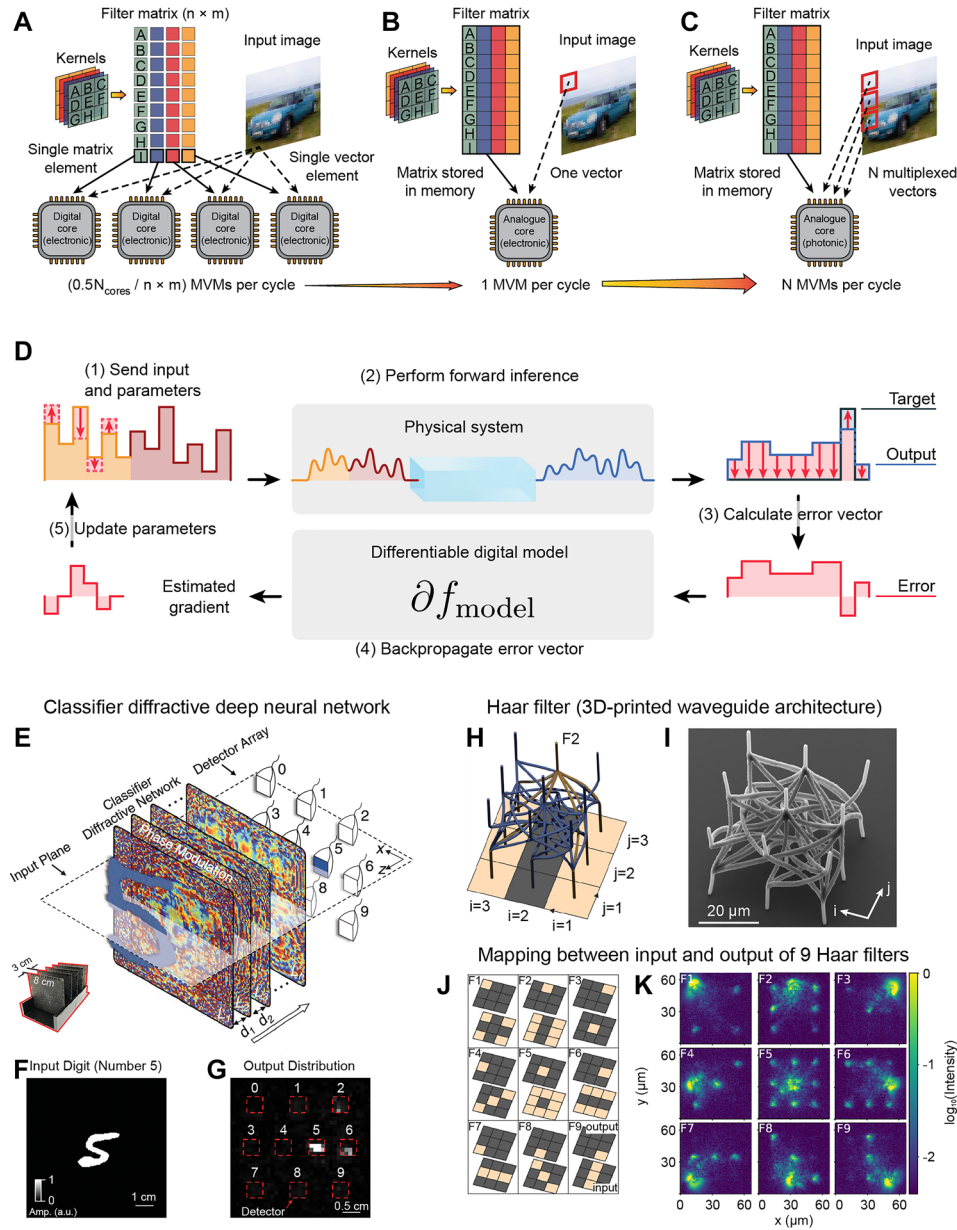
Figure 8: Deep learning with physical systems.
A Digital electronics requires many sequential processing steps distributed across multiple cores to compute convolutional operations on an image. **B** In contrast, an entire matrix-vector multiplication (MVM) can be performed in a single step using analogue electronic in-memory computing. **C** Finally, in photonic in-memory computing, wavelength multiplexing is used to introduce an additional degree of freedom, enabling multiple MVM operations in a single time step (parallel convolutional processing using an integrated photonic tensor core). **D** Recently, it has been demonstrated that efficient training can be achieved by implementing error backpropagation using physics-aware training, where the forward pass in the training step is performed by the physical system, while error backpropagation is performed in a numerical system designed to mimic the response of the physical system. **E** Diffractive Deep Neural Networks comprise multiple transmissive layers, where each point on a given layer acts as a neuron with a complex-valued transmission coefficient. Here, a handwritten digit classifier that classifies **F** input digits (0, 1, ... , 9) based on **G** 10 different detector regions at the output plane of the network, each corresponding to one digit. **H** Design and **I** SEM micrograph of a 3D-printed Haar filter with a kernel of width three. **J** Schematic illustration of the input−output mapping of nine Haar filters (F1−F9). **K** Optical characterization of the filter's connection topology, injection at the output port, and recording the input ports emission. Panels A−C are adapted with permission from Ref. [169] (Copyright 2021 Springer-Nature), panel D from Ref. [170] (Copyright 2022 Springer-Nature), panels E−G from Ref. [134] (Copyright 2018 AAAS), and panels H−K from Ref. [135] (Copyright 2020 Optica).

within images encoded on light projected through the complex photonic structure [134]. Furthermore, convolutional image processing, including classification using deep neural networks, has been demonstrated, e.g., in Refs. [134, 169, 171].

Micro 3D printing can create Cantor-set-like photonic circuits with a 3D structure that allows deeply connected networks to occupy less space than a traditional printed photonic circuit [135] (Figure 8H–K). These structures were shown to exhibit a high degree of convolution in a small space, which provides a potential avenue for extremely compact CNNs.

4.3 Microscopic particles with embodied intelligence

Natural systems have evolved powerful sensing capabilities to gain information about their environments and to communicate [172, 173]. For example, in swarms of midges, schools of fish, and flocks of birds, individuals exchange information as part of their behavior to self-organize into a collective state [174]. Microorganisms have also developed complex strategies to survive and thrive in their environment by integrating sensors, actuators, and information processing. Their biochemical networks and sensory systems are optimized to excel at specific tasks, such as climbing chemical gradients [175], coping with ocean turbulence [176], and efficiently foraging for food [177, 178]. They have also acquired complex strategies to interact with their environment and with other microorganisms, leading to the emergence of macroscopic collective patterns. These patterns are driven by energy conversion from the smallest to the largest scales and permit microorganisms to break free of some of their physical limits. For example, dense systems of bacteria develop "active turbulence" at length scales where only laminar flows are expected from the underlying physical laws [179, 180]. As another example, dense filaments and motor proteins, which are the structural building blocks of cells, develop active nematic structures with new physical properties [181, 182].

On the other hand, synthetic microscopic systems that try to emulate living systems still present many fewer possibilities. Most experimental studies have been constrained to steric, electrostatic, phoretic, or hydrodynamic interactions, which are readily available from physical interactions [5]. Even these simple interactions can lead to interesting complex behaviors and self-organization whose onset is often observed in artificial systems where increased energy input above a threshold density drives a phase transition to an aggregated state.

An example of such behaviors is the formation of "living crystals," which are metastable clusters of active particles [183, 184].

Photonics has the opportunity of making microscopic particles intelligent, providing the tools for artificial microscopic particles to acquire, elaborate, and respond to information from their environment [5]. This can be made by different means. For example, Ref. [185] has recently developed a lithographic fabrication-and-release protocol to build microscopic walking robots activated by light, which provide a new class of voltage-controllable electrochemical actuators that operate at low voltages (200 µV), low power (10 nW) and are completely compatible with silicon processing (Figure 9A). This permits the authors to realize microscopic particles that can be actuated by shining a beam of light on them, as shown in Figure 9B. Also, Figure 9C shows another recent example where microscopic particles are enhanced with metasurfaces that alter the linear and angular momentum of the incident light, therefore, permitting steer the particles [186]. However, these approaches only provide photonic-actuated microrobots, which still require some external control and feedback.

In this context, machine learning can provide new approaches to realize these possibilities. In fact, the early studies presented above have mainly relied on designs based on human intuition, which is now leading to diminishing returns. Designs obtained by machine learning can go beyond what can be simply imagined by human intuition and therefore open new possibilities. For example, machine learning can help to achieve onboard sensing and decision-making, as opposed to external computer-controlled feedback loops. This has been recently demonstrated in a proof-of-principle study by designing reconfigurable organisms [187], as shown in Figure 9D–G. In this work, artificial intelligence methods were employed to automatically design diverse candidate lifeforms in silico to perform some desired function, and transferable designs were then created using a cell-based construction toolkit to realize living systems with the predicted behaviors. In the future, these reconfigurable living organisms can be further enhanced with photonic capabilities such as lasing [188] and light guidance [189].

5 Challenges

In the previous sections, we have shown the potential of deep learning to enhance the study of light–matter interactions beyond the capabilities of conventional methods,
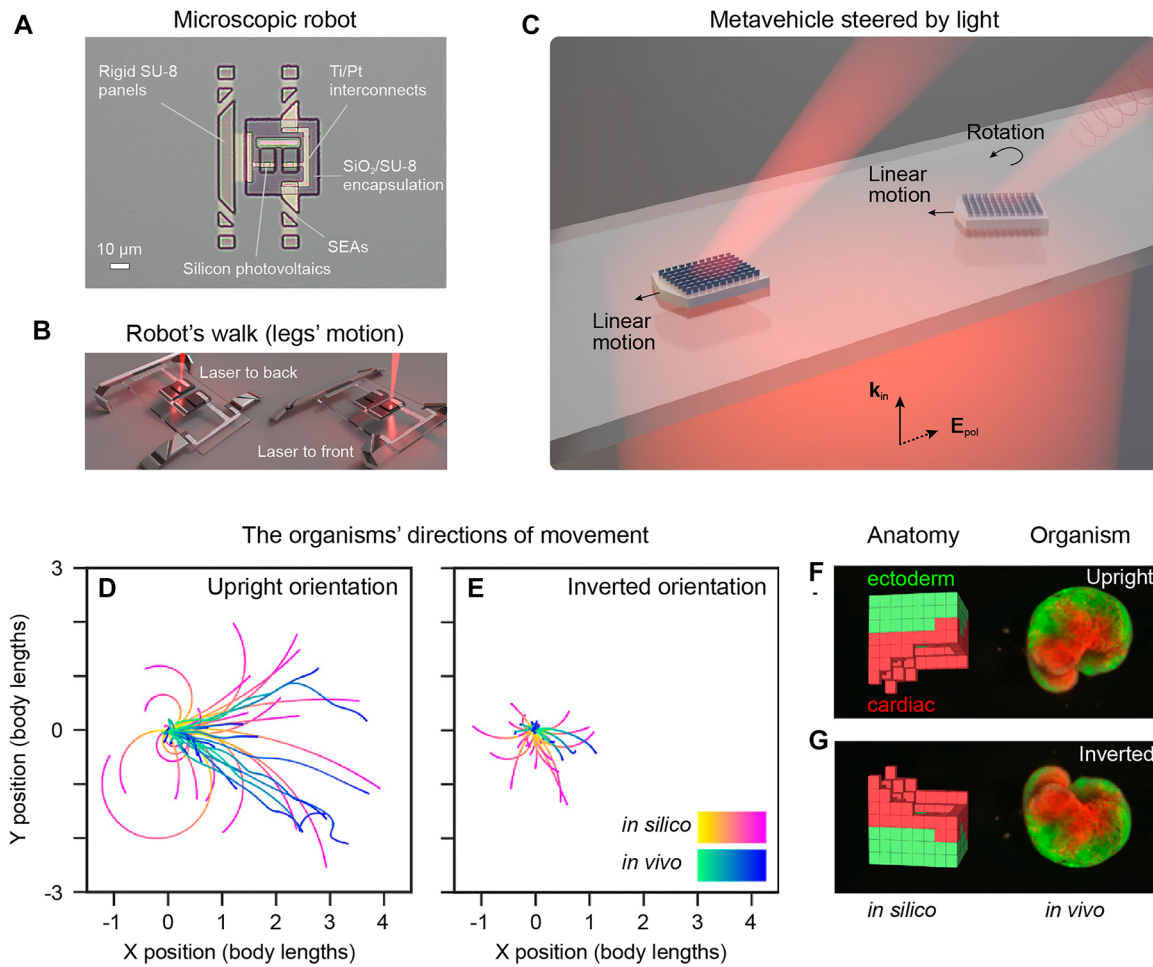
Figure 9: Toward embodied intelligence in microscopic particles.
A Optical image of a microscopic robot. It has two parts: a body with internal electronics and legs that actuate. The electronics are simple circuits made from silicon p−n junctions and metal interconnects, encapsulated between a layer of silicon dioxide and a layer of SU-8 photoresist. The legs are made from a new class of voltage-controlled surface electrochemical actuators (SEAs) and rigid photoresist panels. The panels control the folded shape of the leg while the SEAs produce motion. **B** By directing laser light to photovoltaics that alternately biases the front and back legs, the robot walks along patterned surfaces. **C** By incorporating an engineered metasurface into a microparticle, the changes in linear and angular momentum of the incident light can be employed to propel and steer the microparticle across a surface, realizing a metavehicle. **D−G** Advanced in silico design using artificial intelligence is employed to generate multicellular microorganisms with specific behaviors. Panels A−B are reproduced with permission from Ref. [185] (Copyright 2020 Springer-Nature), panel C from Ref. [186] (Copyright 2021 Springer-Nature), and panels D−G from Ref. [187] (Copyright 2020 NAS).

providing a fast, automatized, and noise-resilient route to optimize the output of optics and photonics experiments. However, in designing, executing, and validating the performance of deep-learning-based methods, one is faced with several considerations and challenges distinct from those experienced when employing conventional methods. This section will review the essential challenges and provide simple guidelines for the effective deployment of deep-learning-based techniques in the study of light–matter interactions.

5.1 Training data augmentation and simulation

The first challenge is to obtain high-quality training data. This is a challenge common to all supervised machine-learning methods and involves the generation of matching input/output pairs. For the network to generalize to unseen data, the network must learn to recognize relevant features of the input data. This requires that the input data represent the full range of expected cases where

the trained network will be applied. Determining whether the input data is sufficiently general can be non-trivial, particularly in cases where the physical rules connecting input and output are unknown. In addition to acquiring a representative set of network inputs for the training, the corresponding target outputs need to be constructed. For example, for image classification and segmentation, the target outputs are often built through manual labeling, which is a time-consuming process, limiting the amount of training data that can be constructed within a reasonable amount of time and effort. Further, the use of user-labeled data for network training makes the network output subjective, as it reflects the biases of the user constructing the labels.

Data normalization and augmentation are often employed to reduce the amount of data required for network training. Data normalization aims at making the data easily interpretable by a neural network. While data normalization does speed up the learning process of neural networks, it also implies a loss of data, which might be of relevance, particularly for regression tasks. Therefore, choosing the right type of data normalization, which quantitatively retains the information of interest, is crucial for successful deep learning deployment in science. A trivial but sometimes overlooked example is that the sample-to-sample variability within a dataset is lost when individual samples from that dataset are independently normalized. In cases where such variability is expected to be important, e.g., for quantitative applications, a global normalization across the entire dataset is to be preferred.

The purpose of data augmentation is to perform multiple transformations to the input, which have a predictable effect on the expected network output. For example, it is possible to generate multiple training images from a single input image, thereby extending the available training set by image rotations, scaling, and cropping. However, one should be wary of transformations that do not necessarily conserve the relation between the input and output, particularly in cases where the underlying theory connecting input and output is unknown. For example, in most cases, the analysis of microscopy images is invariant under translations and rotations. However, depending on the context, transformations that alter the pixels themselves, such as scaling or elastic transformations, may introduce unpredictable artifacts in the analysis. As a general rule, one should only employ augmentations that have a predictable effect on the analysis.

One way to partially overcome these challenges is to use simulated data for training neural networks. This approach can generate data on the fly during training, enabling essentially unlimited training data. Nonetheless, verifying that the training data represents the experiment realistically is still non-trivial. Essentially, the simulation must be sufficiently exact to capture the relevant features in the experimental data accurately. As an illustrative example, in Ref. [59], simulated microscopy images of single particles were used to train a network to estimate the position of particles within an image. Since the experimental data consisted of particles whose intensity profiles could be well described by Bessel functions, training the network on combinations of two-dimensional Bessel functions with varying width, position, and intensity was sufficient to provide good localization accuracy. In contrast, in Ref. [128], simulated data were used to train a network to quantify the size and refractive index of individual nanoparticles from their scattering patterns. In this case, the relevant information is encoded in the Fourier spectrum of the scattered field, and in order to capture the relevant features of the scattering patterns, the simulated data needed to consist of simulated Mie scattering patterns of particles passed through a synthetic replica of the experimental optical system including its aberrations.

Sufficiently complex deep-learning models can learn sophisticated correlations in the input data. The models can also be sensitive to out-of-distribution shifts. Slight changes in the parameters of the microscope in the last example would invalidate the trained model and require retraining. Detecting and making deep learning solutions robust to such perturbations is an active field of research, and multiple strategies have been proposed to partially address this issue [190, 191].

5.2 Architecture and hyperparameter optimization

The choice of network architecture is currently more of a form of art than an exact science. In general, if there exist a transformation connecting the input data to the desired output, any network architecture with sufficiently many adjustable parameters will be able to learn an approximation to this transformation. What can vary between different choices of architectures is the accuracy of the approximation, the training time, the training data amount required to reach this approximation, and the execution speed of the deployed network. What matters the most for the choice of network architecture and hyperparameters is the amount and type of available data and the desired output type. More specifically, inherent symmetries in the input data can often be used to guide the choice of

network architecture. For example, when detecting and/or classifying objects in an image, the spatial location of the object within the image rarely matters for classification. Convolution kernels are intrinsically translationally invariant and are thus routinely used for such tasks. Beyond this, physically informed neural networks provide a strategy for imposing physical symmetries and constraints on the network prediction, typically by penalizing predictions that do not obey the specified symmetries [192].

In this way, the architecture of a neural network carries an inductive bias, which determines what relations and features of the data are most easily learnt and prioritized by the network [193]. Being aware of these biases can aid the development of efficient network architectures for solving a specific problem.

Regarding the choice of hyperparameters, such as the depth of the network and the number of adjustable parameters, this will typically depend on the amount of available data and the type of transformation the network is to learn. The deeper the network, the more complex changes the network can learn. On the other hand, if the available data is limited, there is a chance of overfitting the training data. Choosing the right balance requires iterative training with varying hyperparameters while carefully validating the network results on a large test set. It is often a good practice to start by training a relatively small network and keep increasing its size as long as its performance increases without overfitting the training data.

An open challenge is automatically determining mechanisms to define the best architecture and hyperparameters. For example, that has been done using some evolutionary architecture such as the neuroevolution of augmenting topologies (NEAT) genetic algorithm developed in 2002 [194] and its subsequent derivations [195]. However, such approaches have turned out to be quite slow in convergence and require many computational resources.

5.3 Performance benchmarking, validation, and reproducibility

Any use of deep learning in science needs to be motivated by a superior performance compared to standard methods by some predefined metric. Depending on the application, such metrics may include analysis speed, accuracy, or robustness to noise.

In some cases, such benchmarking is straightforward. For instance, for the inverse problem solvers, the viability of the obtained solution can be checked by explicitly solving the forward problem with the obtained solution as an input. However, in many cases, deep learning solutions are applied to situations beyond the capabilities of existing techniques and where no theory exists that maps the network prediction to the targeted output. Validating the output of the network in such cases poses a considerable challenge. To our knowledge, the best practice is to generate an experimental data set that can be analyzed using some traditional method. Once the deep learning solution has been validated against the traditional method on this data set, the next step is to generate a synthetic data set that simulates the experimental setup and where the ground truth is known perfectly. The quality of the synthetic data set can be evaluated by comparing the output of the traditional method and deep learning solution to the simulated ground truth for a range of parameters where the conventional method is known to perform well. Once the simulation quality has been validated, the final step is to tune the parameters in the synthetic data set to the instances where the traditional method fails and again validate the deep learning solution against the simulated ground truth for the updated data set. In this way, the deep learning solution can be shown to perform comparably to a traditional method for the cases where the traditional method is expected to work. Furthermore, the deep learning solution should outperform the traditional method on the synthetic data sets as well.

Such careful benchmarking and validation become particularly important when adversarial approaches are used to manipulate the data [196]. As we have seen in the previous sections, adversarial networks are trained to generate synthetic data based on some input and are designed to *look* realistic rather than to represent the ground truth fairly. Therefore, adversarial approaches are at risk of extrapolating the input data, effectively making up non-existing data [196]. There is an increasing awareness of these issues in the machine learning community, and several recent studies demonstrate that advances in image reconstruction and image analysis, provided by the use of GANs, can also be achieved without adversarial learning [118, 124, 197].

Finally, a related issue is the result reproducibility by the deep learning-powered analysis. In particular, the published deep learning solutions are typically made for specific data sets; therefore, they are unlikely to provide valid results when directly employed by a different research group on a similar, but not statistically identical, data set. Thus, there has lately been a surge of interest toward the development of publicly accessible software for customizing deep learning solutions without the steep learning curve commonly associated with machine learning [51, 198]. A set of guidelines were recently proposed

for enhancing the reproducibility of the deep learning techniques [199]. As a minimal requirement, the data, model, and analysis code should be made publicly available (bronze standard). To meet the silver standard, authors should make the dependencies of the analysis installable with a single command, with the code properly documented. Further, all random components of the analysis should be made deterministic for reproducibility. Finally, the authors should make the full analysis reproducible with a single command to meet the gold standard.

6 Conclusions and guidelines

In this review, we have presented the current state and future perspectives for the application of deep learning in the fields of optics and photonics. Owing to the challenges related to validating and reproducing deep learning results, a set of community-wide recommendations for deep learning reporting and validation in biology was recently published [200], with the acronym DOME (**D**ata, **O**ptimization, **M**odel, **E**valuation). The recommendations are summarized as a checklist of questions that should be addressed when reporting deep learning results. These recommendations largely apply to the fields of optics and photonics as well; namely, every deep learning application should be able to provide answers to these questions:

- **Data:** How large is the data set used for training the model? How large is the validation set? Are validation and training set independent? Is the distribution of data different in the training and validation sets? Has the data set been used previously? Are the data publicly available?
- **Optimization:** What type of deep learning algorithm was used? Is the algorithm new? If so, why was it chosen over existing algorithms? Does the model use output from other deep learning algorithms as input? How were the data encoded and preprocessed prior to prediction? How many parameters does the model consist of? Was the feature selection performed? If so, how? Are the number of parameters much larger than the number of data points in the training set? If so, how was the overfitting ruled out? If not, how was underfitting ruled out? Are the hyperparameter configurations, optimization schedule, model files, and optimization parameters reported?
- **Model:** Is the model black box or interpretable? If the model is interpretable, can you give clear examples of this? Is the model classification or regression? What

is the typical execution time? Is the software publicly available?
- **Evaluation:** How was the method evaluated? Which metrics were used for the evaluation? Was a comparison to standard algorithms made on benchmark data sets? Was a comparison to simpler baselines performed? Are the raw evaluation files available?

Finally, we remark that, beyond a large range of applications that have already been successful, there are still many fields where deep learning can have a large impact and, therefore, need to be explored. First, the execution speed of trained deep learning models provides new possibilities for experiments through the automatization of current setups. This can free these experiments from the need for continuous human intervention and supervision, permitting the acquisition of large-scale statistics that currently would be prohibitively work-intensive. In turn, this will allow scientists to study also relatively rare events that human operators on small sample sizes might disregard as outliers (e.g., in biomolecule pulling experiments using optical tweezers). Second, there is a great drive towards understanding not only *what* deep learning can do but also *how* it achieves it. This means gaining insights into the understanding and interpreting the black box that deep learning often represents. Once this work advances, it will open new possibilities to discover the theory underlying various phenomena by observing what the network learns. Third, nanophotonics can provide essential tools to make physical implementations of neural networks, which have remarkable advantages in terms of increasing computational speed and minimizing power consumption. In fact, as we have seen in the previous sections, there are already several proposals along these lines. One of the critical issues in this field is integrating these new neuromorphic computing technologies with current computational technologies based on Boolean electronic circuits. Fourth, arguably the holy grail of the field would be to employ machine learning concepts and techniques to realize microscopic particles capable of real intelligent behavior by autonomously processing and responding to information from their environment.

References

[1] Y. S. Kivshar and G. P. Agrawal, *Optical Solitons: From Fibers to Photonic Crystals*, 2003.

[2] M. Gu, Q. Zhang, and S. Lamon, "Nanomaterials for optical data storage," *Nat. Rev. Mater.*, vol. 1, 2016, Art no. 3660.

[3] J. A. Jackman, A. R. Ferhan, and N. J. Cho, "Nanoplasmonic sensors for biointerfacial science," *Chem. Soc. Rev.*, vol. 46, pp. 3615–3660, 2017.

[4] S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vucković, and A. W. Rodriguez, "Inverse design in nanophotonics," *Nat. Photonics*, vol. 12, pp. 659–670, 2018.

[5] F. Cichos, K. Gustavsson, B. Mehlig, and G. Volpe, "Machine learning for active matter," *Nat. Mach. Intell.*, vol. 2, pp. 94–103, 2020.

[6] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[7] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, pp. 533–536, 1986.

[8] J. Jiang, M. Chen, and J. A. Fan, "Deep neural networks for the evaluation and design of photonic devices," *Nat. Rev. Mater.*, vol. 6, pp. 679–700, 2020.

[9] M. Vai and S. Prasad, "Automatic impedance matching with a neural network," *IEEE Microw. Guid. Wave Lett.*, vol. 3, no. 10, pp. 353–354, 1993.

[10] M. M. Vai, W. Shuichi, L. Bin, and S. Prasad, "Reverse modeling of microwave circuits with bidirectional neural network models," *IEEE Trans. Microw. Theor. Tech.*, vol. 46, no. 10, pp. 1492–1494, 1998.

[11] M. F. O. Hameed, S. S. A. Obayya, K. Al-Begain, A. M. Nasr, and M. I. Abo el Maaty, "Accurate radial basis function based neural network approach for analysis of photonic crystal fibers," *Opt. Quant. Electron.*, vol. 40, no. 11, pp. 891–905, 2009.

[12] L. Mescia, G. Fornarelli, D. Magarielli, F. Prudenzano, M. De Sario, and F. Vacca, "Refinement and design of rare earth doped photonic crystal fibre amplifier using an ann approach," *Opt. Laser. Technol.*, vol. 43, no. 7, pp. 1096–1103, 2011.

[13] T. Asano and S. Noda, "Optimization of photonic crystal nanocavities based on deep learning," *Opt. Express*, vol. 26, no. 25, pp. 32704–32717, 2018.

[14] T. Asano and S. Noda, "Iterative optimization of photonic crystal nanocavity designs by using deep neural networks," *Nanophotonics*, vol. 8, no. 12, pp. 2243–2256, 2019.

[15] T. Christensen, C. Loh, S. Picek, et al., "Predictive and generative machine learning models for photonic crystals," *Nanophotonics*, vol. 9, no. 13, pp. 4183–4192, 2020.

[16] L. Yang, J. Ren, Y. Li, and H. Chen, "Inverse design of photonic topological state via machine learning," *Appl. Phys. Lett.*, vol. 114, no. 18, p. 181105, 2019.

[17] C. Qiu, X. Wu, Z. Luo, et al., "Simultaneous inverse design continuous and discrete parameters of nanophotonic structures via back-propagation inverse neural network," *Opt. Commun.*, vol. 483, p. 126641, 2021.

[18] Y. Chen, J. Zhu, Y. Xie, N. Feng, and Q. H. Liu, "Smart inverse design of graphene-based photonic metamaterials by an adaptive artificial neural network," *Nanoscale*, vol. 11, no. 19, pp. 9749–9755, 2019.

[19] D. Liu, Y. Tan, E. Khoram, and Z. Yu, "Training deep neural networks for the inverse design of nanophotonic structures," *ACS Photonics*, vol. 5, no. 4, pp. 1365–1369, 2018.

[20] R. Unni, K. Yao, X. Han, M. Zhou, and Y. Zheng, "A mixture-density-based tandem optimization network for on-demand inverse design of thin-film high reflectors," *Nanophotonics*, vol. 10, pp. 4057–4065, 2021.

[21] R. Unni, K. Yao, and Y. Zheng, "Deep convolutional mixture density network for inverse design of layered photonic structures," *ACS Photonics*, vol. 7, no. 10, pp. 2703–2712, 2020.

[22] D. Zhang, Q. Bao, W. Chen, Z. Liu, G. Wei, and J. J. Xiao, "Inverse design of an optical film filter by a recurrent neural adjoint method: an example for a solar simulator," *J. Opt. Soc. Am. B*, vol. 38, no. 6, pp. 1814–1821, 2021.

[23] A.-P. Blanchard-Dionne and O. J. F. Martin, "Successive training of a generative adversarial network for the design of an optical cloak," *OSA Continuum*, vol. 4, no. 1, pp. 87–95, 2021.

[24] C. Qian, B. Zheng, Y. Shen, et al., "Deep-learning-enabled self-adaptive microwave cloak without human intervention," *Nat. Photonics*, vol. 14, no. 6, pp. 383–390, 2020.

[25] Z. Zhen, C. Qian, Y. Jia, et al., "Realizing transmitted metasurface cloak by a tandem neural network," *Photon. Res.*, vol. 9, no. 5, pp. B229–B235, 2021.

[26] P. Komorowski, P. Czerwińska, M. Surma, P. Zagrajek, R. Piramidowicz, and A. Siemion, "Three-focal-spot terahertz diffractive optical element-iterative design and neural network approach," *Opt. Express*, vol. 29, no. 7, pp. 11243–11253, 2021.

[27] Z. Liu, Z. Zhu, and W. Cai, "Topological encoding method for data-driven photonics inverse design," *Opt. Express*, vol. 28, no. 4, pp. 4825–4835, 2020.

[28] S. An, C. Fowler, B. Zheng, et al., "A deep learning approach for objective-driven all-dielectric metasurface design," *ACS Photonics*, vol. 6, no. 12, pp. 3196–3207, 2019.

[29] S. An, B. Zheng, H. Tang, et al., "Multifunctional metasurface design with a generative adversarial network," *Adv. Opt. Mater.*, vol. 9, no. 5, p. 2001433, 2021.

[30] H. N. Bui, J. S. Kim, and J. W. Lee, "Design of tunable metasurface using deep neural networks for field localized wireless power transfer," *IEEE Access*, vol. 8, pp. 194868–194878, 2020.

[31] W.-Q. Chen, Da-S. Zhang, S.-Yu. Long, Z.-Z. Liu, and J.-J. Xiao, "Nearly dispersionless multicolor metasurface beam deflector for near eye display designed by a physics-driven deep neural network," *Appl. Opt.*, vol. 60, no. 13, pp. 3947–3953, 2021.

[32] J. A. Fan, "Freeform metasurface design based on topology optimization," *MRS Bull.*, vol. 45, no. 3, pp. 196–201, 2020.

[33] F. Ghorbani, S. Beyraghi, J. Shabanpour, H. Oraizi, H. Soleimani, and M. Soleimani, "Deep neural network-based automatic metasurface design with a wide frequency range," *Sci. Rep.*, vol. 11, no. 1, p. 7102, 2021.

[34] F. Ghorbani, J. Shabanpour, S. Beyraghi, H. Soleimani, H. Oraizi, and M. Soleimani, "A deep learning approach for

inverse design of the metasurface for dual-polarized waves," *Appl. Phys. A*, vol. 127, no. 11, p. 869, 2021.

[35] X. Han, Z. Fan, Z. Liu, C. Li, and L. Jay Guo, "Inverse design of metasurface optical filters using deep neural network with high degrees of freedom," *InfoMat*, vol. 3, no. 4, pp. 432–442, 2021.

[36] E. S. Harper, E. J. Coyle, J. P. Vernon, and M. S. Mills, "Inverse design of broadband highly reflective metasurfaces using neural networks," *Phys. Rev. B*, vol. 101, no. 19, p. 195104, 2020.

[37] Z. Liu, L. Raju, D. Zhu, and W. Cai, "A hybrid strategy for the discovery and design of photonic structures," *IEEE J. Emerg. Sel. Top. Circ. Syst.*, vol. 10, no. 1, pp. 126–135, 2020.

[38] C. C. Nadell, B. Huang, J. M. Malof, and W. J. Padilla, "Deep learning for accelerated all-dielectric metasurface design," *Opt. Express*, vol. 27, no. 20, pp. 27523–27535, 2019.

[39] J. Noh, Y.-H. Nam, S. So, et al., "Design of a transmissive metasurface antenna using deep neural networks," *Opt. Mater. Express*, vol. 11, no. 7, pp. 2310–2317, 2021.

[40] T. Qiu, X. Shi, J. Wang, et al., "Deep learning: a rapid and efficient route to automatic metasurface design," *Adv. Sci.*, vol. 6, no. 12, p. 1900128, 2019.

[41] N. B. Roberts and M. K. Hedayati, "A deep learning approach to the forward prediction and inverse design of plasmonic metasurface structural color," *Appl. Phys. Lett.*, vol. 119, no. 6, p. 061101, 2021.

[42] M. Ali Shameli, A. Fallah, and L. Yousefi, "Developing an optimized metasurface for light trapping in thin-film solar cells using a deep neural network and a genetic algorithm," *J. Opt. Soc. Am. B*, vol. 38, no. 9, pp. 2728–2735, 2021.

[43] X. Shi, T. Qiu, J. Wang, X. Zhao, and S. Qu, "Metasurface inverse design using machine learning approaches," *J. Phys. Appl. Phys.*, vol. 53, no. 27, p. 275105, 2020.

[44] J. R. Thompson, J. A. Burrow, P. J. Shah, et al., "Artificial neural network discovery of a switchable metasurface reflector," *Opt. Express*, vol. 28, no. 17, pp. 24629–24656, 2020.

[45] D. Xu, Yu. Luo, J. Luo, et al., "Efficient design of a dielectric metasurface with transfer learning and genetic algorithm," *Opt. Mater. Express*, vol. 11, no. 7, pp. 1852–1862, 2021.

[46] C. Yeung, J.-M. Tsai, B. King, et al., "Multiplexed supercell metasurface design and optimization with tandem residual networks," *Nanophotonics*, vol. 10, no. 3, pp. 1133–1143, 2021.

[47] L. Yuan, L. Wang, X. S. Yang, H. Huang, and B. Z. Wang, "An efficient artificial neural network model for inverse design of metasurfaces," *IEEE Antenn. Wireless Propag. Lett.*, vol. 20, no. 6, pp. 1013–1017, 2021.

[48] Q. Du, Q. Zhang, and G. Liu, "Deep learning: an efficient method for plasmonic design of geometric nanoparticles," *Nanotechnology*, vol. 32, no. 50, p. 505607, 2021.

[49] J. He, C. He, C. Zheng, Q. Wang, and J. Ye, "Plasmonic nanoparticle simulations and inverse design using machine learning," *Nanoscale*, vol. 11, no. 37, pp. 17444–17459, 2019.

[50] J. Peurifoy, Y. Shen, Li. Jing, et al., "Nanophotonic particle simulation and inverse design using artificial neural networks," *Sci. Adv.*, vol. 4, no. 6, p. eaar4206, 2018.

[51] B. Midtvedt, S. Helgadottir, A. Argun, J. Pineda, D. Midtvedt, and G. Volpe, "Quantitative digital microscopy with deep learning," *Appl. Phys. Rev.*, vol. 8, no. 1, p. 011310, 2021.

[52] P. R. Wiecha, A. Lecestre, N. Mallet, and G. Larrieu, "Pushing the limits of optical information storage using deep learning," *Nat. Nanotechnol.*, vol. 14, pp. 237–244, 2019.

[53] D. Marković, A. Mizrahi, D. Querlioz, and J. Grollier, "Physics for neuromorphic computing," *Nature Reviews Physics*, vol. 2, no. 9, pp. 499–510, 2020.

[54] B. Mehlig, *Machine Learning with Neural Networks*, Cambridge University Press, 2021.

[55] *Long Short Term Memory: Deep Dive*, 2021. https://faun .pub/long-short-term-memory-deep-dive-4830e22b28ac [accessed: Feb. 07, 2022].

[56] A. Vaswani, N. Shazeer, N. Parmar, et al., "Attention is all you need," *Adv. Neural Inf. Process. Syst.*, vol. 2017, pp. 5999–6009, 2017.

[57] A. Géron, *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*, Sebastopol, CA, O'Reilly Media, Inc., 2019.

[58] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Hinton. Imagenet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, 2017. https://doi.org/10 .1145/3065386.

[59] S. Helgadottir, A. Argun, and G. Volpe, "Digital video microscopy enhanced by deep learning," *Optica*, vol. 6, no. 4, pp. 506–513, 2019.

[60] A. Radford and L. Metz, *Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks*, New York, ICLR, 2015.

[61] J. Zhou, G. Cui, S. Hu, et al., "Graph neural networks: a review of methods and applications," *AI Open*, vol. 1, 2020. https://doi.org/10.1016/j.aiopen.2021.01.001.

[62] H. Zhang, M. Gu, X. D. Jiang, et al., "An optical neural chip for implementing complex-valued neural network," *Nat. Commun.*, vol. 12, no. 1, p. 2021.

[63] S. Ira Ktena, S. Parisot, E. Ferrante, et al., "Metric learning with spectral graph convolutions on brain connectivity networks," *Neuroimage*, vol. 169, pp. 431–442, 2018.

[64] X. Li, Y. Zhou, N. Dvornek, et al., "Braingnn: interpretable brain graph neural network for fmri analysis," *Med. Image Anal.*, vol. 74, p. 102233, 2021.

[65] J. Pineda, B. Midtvedt, H. Bachimanchi, et al., Geometric deep learning reveals the spatiotemporal fingerprint of microscopic motion. arXiv preprint arXiv:2202.06355, 2022.

[66] K. Koshelev, S. Kruk, E. Melik-Gaykazyan, et al., "Subwavelength dielectric resonators for nonlinear nanophotonics," *Science*, vol. 367, no. 6475, pp. 288–292, 2020.

[67] V. Mylnikov, S. T. Ha, Z. Pan, et al., "Lasing action in single subwavelength particles supporting supercavity modes," *ACS Nano*, vol. 14, no. 6, pp. 7338–7346, 2020.

[68] A. Tittl, A. Leitis, M. Liu, et al., "Imaging-based molecular barcoding with pixelated dielectric metasurfaces," *Science*, vol. 360, no. 6393, pp. 1105–1109, 2018.

[69] M. M. R. Elsawy, S. Lanteri, R. Duvigneau, J. A. Fan, and P. Genevet, "Numerical optimization methods for metasurfaces," *Laser Photon. Rev.*, vol. 14, no. 10, p. 1900445, 2020.

[70] W. Ma, Z. Liu, Z. A. Kudyshev, A. Boltasseva, W. Cai, and Y. Liu, "Deep learning for the design of photonic structures," *Nat. Photonics*, vol. 15, no. 2, pp. 77—90, 2021.

[71] K. Yao, R. Unni, and Y. Zheng, "Intelligent nanophotonics: merging photonics and artificial intelligence at the nanoscale," *Nanophotonics*, vol. 8, no. 3, pp. 339—366, 2019.

[72] B. Shen, P. Wang, R. Polson, and R. Menon, "An integrated-nanophotonics polarization beamsplitter with $2.4 \times 2.4 \, \mu m^2$ footprint," *Nat. Photonics*, vol. 9, no. 6, pp. 378—382, 2015.

[73] Z. A. Kudyshev, A. V. Kildishev, V. M. Shalaev, and A. Boltasseva, "Machine-learning-assisted metasurface design for high-efficiency thermal emitter optimization," *Appl. Phys. Rev.*, vol. 7, no. 2, p. 021407, 2020.

[74] E. Vahidzadeh and K. Shankar, "Artificial neural network-based prediction of the optical properties of spherical core—shell plasmonic metastructures," *Nanomaterials*, vol. 11, no. 3, p. 633, 2021.

[75] Y.-W. Chang, L. Natali, O. Jamialahmadi, S. Romeo, J. B. Pereira, and G. Volpe. Neural network training with highly incomplete datasets. arXiv preprint arXiv:2107.00429, p. 2107.00429, 2021.

[76] P. R. Wiecha, A. Arbouet, C. Girard, and O. L. Muskens, "Deep learning in nano-photonics: inverse design and beyond," *Photon. Res.*, vol. 9, no. 5, pp. B182—B200, 2021.

[77] R. Singh, A. Agarwal, and B. W. Anthony, "Mapping the design space of photonic topological states via deep learning," *Opt. Express*, vol. 28, no. 19, pp. 27893—27902, 2020.

[78] M. H. Tahersima, K. Kojima, T. Koike-Akino, et al., "Deep neural network inverse design of integrated photonic power splitters," *Sci. Rep.*, vol. 9, no. 1, p. 1368, 2019.

[79] J. Jiang and J. A. Fan, "Global optimization of dielectric metasurfaces using a physics-driven neural network," *Nano Lett.*, vol. 19, no. 8, pp. 5366—5372, 2019.

[80] J. Jiang, D. Sell, S. Hoyer, et al., "Free-form diffractive metagrating design based on generative adversarial networks," *ACS Nano*, vol. 13, no. 8, pp. 8872—8878, 2019.

[81] G. Genty, L. Salmela, J. M. Dudley, et al., "Machine learning and applications in ultrafast photonics," *Nat. Photonics*, vol. 15, no. 2, pp. 91—101, 2021.

[82] R. S. Hegde, "Deep learning: a new tool for photonic nanostructure design," *Nanoscale Adv.*, vol. 2, no. 3, pp. 1007—1023, 2020.

[83] O. Khatib, S. Ren, M. Jordan, and W. J. Padilla, "Deep learning the electromagnetic properties of metamaterials—a comprehensive review," *Adv. Funct. Mater.*, p. 2101748, 2021. https://doi.org/10.1002/adfm.202101748.

[84] Z. Liu, D. Zhu, L. Raju, and W. Cai, "Tackling photonic inverse design with machine learning," *Adv. Sci.*, vol. 8, no. 5, p. 2002923, 2021.

[85] S. Mao, L. Cheng, C. Zhao, F. N. Khan, Q. Li, and H. Y. Fu, "Inverse design for silicon photonics: from iterative

optimization algorithms to deep neural networks," *Appl. Sci.*, vol. 11, no. 9, p. 3822, 2021.

[86] Q. Wu, X. Li, Li. Jiang, et al., "Deep neural network for designing near- and far-field properties in plasmonic antennas," *Opt. Mater. Express*, vol. 11, no. 7, pp. 1907—1917, 2021.

[87] A. Sheverdin, F. Monticone, and C. Valagiannopoulos, "Photonic inverse design with neural networks: the case of invisibility in the visible," *Phys. Rev. Appl.*, vol. 14, no. 2, p. 024054, 2020.

[88] Z. Huang, X. Liu, and J. Zang, "The inverse design of structural color using machine learning," *Nanoscale*, vol. 11, no. 45, pp. 21748—21758, 2019.

[89] L. Xu, M. Rahmani, Y. Ma, et al., "Enhanced light—matter interactions in dielectric nanostructures via machine-learning approach," *Adv. Photon.*, vol. 2, no. 2, p. 026003, 2020.

[90] M. V. Zhelyeznyakov, S. Brunton, and A. Majumdar, "Deep learning to accelerate scatterer-to-field mapping for inverse design of dielectric metasurfaces," *ACS Photonics*, vol. 8, no. 2, pp. 481—488, 2021.

[91] X. Tu, W. Xie, Z. Chen, et al., "Analysis of deep neural network models for inverse design of silicon photonic grating coupler," *J. Lightwave Technol.*, vol. 39, no. 9, pp. 2790—2799, 2021.

[92] X. Xu, C. Sun, Li. Yu, Z. Jia, J. Han, and W. Huang, "An improved tandem neural network for the inverse design of nanophotonics devices," *Opt. Commun.*, vol. 481, p. 126513, 2021.

[93] P. Dai, Y. Wang, Y. Hu, et al., "Accurate inverse design of fabry—perot-cavity-based color filters far beyond srgb via a bidirectional artificial neural network," *Photonics Res.*, vol. 9, no. 5, pp. B236—B246, 2021.

[94] C.-Xu. Liu, G.-L. Yu, and G.-Y. Zhao, "Neural networks for inverse design of phononic crystals," *AIP Adv.*, vol. 9, no. 8, p. 085223, 2019.

[95] T. Coen, H. Greener, M. Mrejen, L. Wolf, and H. Suchowski, "Deep learning based reconstruction of directional coupler geometry from electromagnetic near-field distribution," *OSA Continuum*, vol. 3, no. 8, pp. 2222—2231, 2020.

[96] D. Yuan, D. Li, C. Zhang, et al., "Inverse design of two-dimensional graphene/h-bn hybrids by a regressional and conditional gan," *Carbon*, vol. 169, pp. 9—16, 2020.

[97] Z. Liu, D. Zhu, S. P. Rodrigues, K.-T. Lee, and W. Cai, "Generative model for the inverse design of metasurfaces," *Nano Lett.*, vol. 18, no. 10, pp. 6570—6576, 2018.

[98] Y. Mao, H. Qi, and X. Zhao, "Designing complex architectured materials with generative adversarial networks," *Sci. Adv.*, vol. 6, no. 17, p. eaaz4169, 2020.

[99] D. Zhu, Z. Liu, L. Raju, A. S. Kim, and W. Cai, "Building multifunctional metasystems via algorithmic construction," *ACS Nano*, vol. 15, no. 2, pp. 2318—2326, 2021.

[100] R. Lin, Z. Alnakhli, and X. Li, "Engineering of multiple bound states in the continuum by latent representation of freeform structures," *Photonics Res.*, vol. 9, no. 4, pp. B96—B103, 2021.

[101] W. Ma, F. Cheng, Y. Xu, Q. Wen, and Y. Liu, "Probabilistic representation and inverse design of metamaterials based on a deep generative model with semi-supervised

learning strategy," *Adv. Mater.*, vol. 31, no. 35, p. 1901111, 2019.

[102] 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.*, vol. 63, no. 8, p. 284212, 2020.

[103] P. Naseri and S. V. Hum, "A generative machine learning-based approach for inverse design of multilayer metasurfaces," *IEEE Trans. Antenn. Propag.*, vol. 69, pp. 5725−5739, 2021.

[104] J. Jiang and J. A. Fan, "Simulator-based training of generative neural networks for the inverse design of metasurfaces," *Nanophotonics*, vol. 9, no. 5, pp. 1059−1069, 2020.

[105] A. Makhzani, J. Shlens, N. Jaitly, I. Goodfellow, and B. Frey. Adversarial autoencoders. arXiv preprint arXiv:1511.05644, 2015.

[106] N. Coudray, P. S. Ocampo, T. Sakellaropoulos, et al., "Classification and mutation prediction from non−small cell lung cancer histopathology images using deep learning," *Nat. Med.*, vol. 24, no. 10, pp. 1559−1567, 2018.

[107] S. Cheng, S. Liu, J. Yu, et al., "Robust whole slide image analysis for cervical cancer screening using deep learning," *Nat. Commun.*, vol. 12, no. 1, pp. 1−10, 2021.

[108] O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9351, Springer, 2015, pp. 234−241.

[109] T. Falk, D. Mai, R. Bensch, et al., "U-Net: deep learning for cell counting, detection, and morphometry," *Nat. Methods*, vol. 16, no. 1, pp. 67−70, 2019.

[110] A. Gomariz, T. Portenier, P. M. Helbling, et al., "Modality attention and sampling enables deep learning with heterogeneous marker combinations in fluorescence microscopy," *Nat. Mach. Intell.*, vol. 3, no. 9, pp. 799−811, 2021.

[111] A. Speiser, L.-R. Müller, P. Hoess, et al., "Deep learning enables fast and dense single-molecule localization with high accuracy," *Nat. Methods*, vol. 18, no. 9, pp. 1082−1090, 2021.

[112] J. M. Newby, A. M. Schaefer, P. T. Lee, M. Gregory Forest, and S. K. Lai, "Convolutional neural networks automate detection for tracking of submicron-scale particles in 2D and 3D," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 115, no. 36, pp. 9026−9031, 2018.

[113] L. E. Altman and D. G. Grier, "CATCH: characterizing and tracking colloids holographically using deep neural networks," *J. Phys. Chem. B*, vol. 124, no. 9, pp. 1602−1610, 2020.

[114] N. Granik, L. E. Weiss, E. Nehme, et al., "Single-particle diffusion characterization by deep learning," *Biophys. J.*, vol. 117, no. 2, pp. 185−192, 2019.

[115] A. Argun, T. Thalheim, S. Bo, F. Cichos, and G. Volpe, "Enhanced force-field calibration via machine learning," *Appl. Phys. Rev.*, vol. 7, no. 4, p. 41404, 2020.

[116] H. Wang, Y. Rivenson, Y. Jin, et al., "Deep learning enables cross-modality super-resolution in fluorescence

microscopy," *Nat. Methods*, vol. 16, no. 1, pp. 103−110, 2019.

[117] W. Ouyang, A. Aristov, M. Lelek, X. Hao, and C. Zimmer, "Deep learning massively accelerates super-resolution localization microscopy," *Nat. Biotechnol.*, vol. 36, 2018. https://doi.org/10.1038/nbt.4106.

[118] N. Wagner, F. Beuttenmueller, N. Norlin, et al., "Deep learning-enhanced light-field imaging with continuous validation," *Nat. Methods*, vol. 18, no. 5, pp. 557−563, 2021.

[119] Z. Wang, L. Zhu, H. Zhang, et al., "Real-time volumetric reconstruction of biological dynamics with light-field microscopy and deep learning," *Nat. Methods*, vol. 18, no. 5, pp. 551−556, 2021.

[120] Y. Wu, Y. Luo, G. Chaudhari, et al., "Bright-field holography: cross-modality deep learning enables snapshot 3D imaging with bright-field contrast using a single hologram," *Light Sci. Appl.*, vol. 8, no. 1, 2019, Art no. 25.

[121] Y. Rivenson, T. Liu, Z. Wei, Y. Zhang, K. de Haan, and A. Ozcan, "PhaseStain: the digital staining of label-free quantitative phase microscopy images using deep learning," *Light Sci. Appl.*, vol. 8, no. 1, pp. 1−11, 2019.

[122] S. Helgadottir, B. Midtvedt, J. Pineda, et al., "Extracting quantitative biological information from bright-field cell images using deep learning," *Biophys. Rev.*, vol. 2, no. 3, p. 031401, 2021.

[123] Y. Li, Y. Xue, and L. Tian, "Deep speckle correlation: a deep learning approach towards scalable imaging through scattering media," vol. 5, no. 10, pp. 1181−1190, 2018.

[124] C. Qiao, Di. Li, Y. Guo, et al., "Evaluation and development of deep neural networks for image super-resolution in optical microscopy," *Nat. Methods*, vol. 18, no. 2, pp. 194−202, 2021.

[125] Y. Rivenson, Y. Zhang, H. Günaydın, Da. Teng, and A. Ozcan, "Phase recovery and holographic image reconstruction using deep learning in neural networks," *Light Sci. Appl.*, vol. 7, no. 2, p. 17141, 2018.

[126] Y. Rivenson, H. Wang, Z. Wei, et al., "Virtual histological staining of unlabelled tissue-autofluorescence images via deep learning," *Nat. Biomed. Eng.*, vol. 3, no. 6, pp. 466−477, 2019.

[127] H. Lin, H. J. Lee, N. Tague, et al., "Microsecond fingerprint stimulated Raman spectroscopic imaging by ultrafast tuning and spatial-spectral learning," *Nat. Commun.*, vol. 12, no. 1, pp. 1−12, 2021.

[128] B. Midtvedt, E. Olsén, F. Eklund, et al., "Fast and accurate nanoparticle characterization using deep-learning-enhanced off-Axis holography," *ACS Nano*, vol. 15, pp. 2240−2250, 2021.

[129] B. Manifold, S. Men, R. Hu, and D. Fu, "A versatile deep learning architecture for classification and label-free prediction of hyperspectral images," *Nat. Mach. Intell.*, vol. 3, no. 4, pp. 306−315, 2021.

[130] K. Thadson, S. Visitsattapongse, and S. Pechprasarn, "Deep learning-based single-shot phase retrieval algorithm for surface plasmon resonance microscope based refractive index sensing application," *Sci. Rep.*, vol. 11, no. 1, pp. 1−14, 2021.

[131] R. Hamerly, L. Bernstein, A. Sludds, M. Soljačić, and D. Englund, "Large-scale optical neural networks based on photoelectric multiplication," *Phys. Rev. X*, vol. 9, no. 2, p. 021032, 2019.

[132] M. A. Nahmias, T. F. de Lima, A. N. Tait, H.-T. Peng, B. J. Shastri, and P. R. Prucnal, "Photonic multiply-accumulate operations for neural networks," *IEEE J. Sel. Top. Quant. El.*, vol. 26, no. 1, pp. 1−18, 2020.

[133] D. A. B. Miller, "Self-configuring universal linear optical component [invited]," *Photon. Res.*, vol. 1, no. 1, p. 1, 2013.

[134] L. Xing, Y. Rivenson, N. T. Yardimci, et al., "All-optical machine learning using diffractive deep neural networks," *Science*, vol. 361, no. 6406, pp. 1004−1008, 2018.

[135] J. Moughames, X. Porte, M. Thiel, et al., "Three-dimensional waveguide interconnects for scalable integration of photonic neural networks," *Optica*, vol. 7, no. 6, p. 640, 2020.

[136] Y. Shen, N. C. Harris, S. Skirlo, et al., "Deep learning with coherent nanophotonic circuits," *Nat. Photonics*, vol. 11, no. 7, pp. 441−446, 2017.

[137] J. Feldmann, N. Youngblood, C. D. Wright, H. Bhaskaran, and W. H. P. Pernice, "All-optical spiking neurosynaptic networks with self-learning capabilities," *Nature*, vol. 569, no. 7755, pp. 208−214, 2019.

[138] Q. Zhang, H. Yu, M. Barbiero, B. Wang, and M. Gu, "Artificial neural networks enabled by nanophotonics," *Light Sci. Appl.*, vol. 8, no. 1, 2019. https://doi.org/10.1038/s41377-019-0151-0.

[139] X. Xu, M. Tan, B. Corcoran, et al., "Photonic perceptron based on a kerr microcomb for high-speed, scalable, optical neural networks," *Laser Photonics Rev.*, vol. 14, no. 10, p. 2000070, 2020.

[140] X. Xu, M. Tan, B. Corcoran, et al., "11 TOPS photonic convolutional accelerator for optical neural networks," *Nature*, vol. 589, no. 7840, pp. 44−51, 2021.

[141] B. J. Shastri, A. N. Tait, T. F. de Lima, et al., "Photonics for artificial intelligence and neuromorphic computing," *Nat. Photonics*, vol. 15, no. 2, pp. 102−114, 2021.

[142] G. Mourgias-Alexandris, A. Tsakyridis, N. Passalis, A. Tefas, K. Vyrsokinos, and N. Pleros, "An all-optical neuron with sigmoid activation function," *Opt. Express*, vol. 27, no. 7, p. 9620, 2019.

[143] A. L. Hodgkin and A. F. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve," *J. Physiol.*, vol. 117, no. 4, pp. 500−544, 1952.

[144] X. Ji, F. A. S. Barbosa, S. P. Roberts, et al., "Ultra-low-loss on-chip resonators with sub-milliwatt parametric oscillation threshold," *Optica*, vol. 4, no. 6, p. 619, 2017.

[145] A. N. Tait, T. F. de Lima, E. Zhou, et al., "Neuromorphic photonic networks using silicon photonic weight banks," *Sci. Rep.*, vol. 7, no. 1, 2017. https://doi.org/10.1038/s41598-017-07754-z.

[146] W. Gerstner, R. Kempter, J. L. van Hemmen, and H. Wagner, "A neuronal learning rule for sub-millisecond temporal coding," *Nature*, vol. 383, no. 6595, pp. 76−78, 1996.

[147] S. Song, K. D. Miller, and L. F. Abbott, "Competitive hebbian learning through spike-timing-dependent synaptic plasticity," *Nat. Neurosci.*, vol. 3, no. 9, pp. 919−926, 2000.

[148] R. Mirek, A. Opala, P. Comaron, et al., "Neuromorphic binarized polariton networks," *Nano Lett.*, vol. 21, no. 9, pp. 3715−3720, 2021.

[149] G. Mourgias-Alexandris, N. Passalis, G. Dabos, A. Totović, A. Tefas, and N. Pleros, "A photonic recurrent neuron for time-series classification," *J. Lightwave Technol.*, vol. 39, no. 5, pp. 1340−1347, 2021.

[150] B. Shi, N. Calabretta, and R. Stabile, "Inp photonic integrated multi-layer neural networks: architecture and performance analysis," *APL Photonics*, vol. 7, no. 1, p. 010801, 2022.

[151] Y. Zuo, Y. Zhao, Y.-C. Chen, S. Du, and J. Liu, "Scalability of all-optical neural networks based on spatial light modulators," *Phys. Rev. Appl.*, vol. 15, no. 5, p. 054034, 2021.

[152] A. N. Tait, T. F. de Lima, M. A. Nahmias, et al., "Silicon photonic modulator neuron," *Phys. Rev. Applied*, vol. 11, no. 6, p. 064043, 2019.

[153] R. Amin, J. K. George, S. Sun, et al., "ITO-based electro-absorption modulator for photonic neural activation function," *APL Materials*, vol. 7, no. 8, p. 081112, 2019.

[154] J. K. George, A. Mehrabian, R. Amin, et al., "Neuromorphic photonics with electro-absorption modulators," *Opt. Express*, vol. 27, no. 4, p. 5181, 2019.

[155] M. M. P. Fard, I. A. D. Williamson, M. Edwards, et al., "Experimental realization of arbitrary activation functions for optical neural networks," *Opt. Express*, vol. 28, no. 8, p. 12138, 2020.

[156] C. Liu, Q. Ma, Z. J. Luo, et al., "A programmable diffractive deep neural network based on a digital-coding metasurface array," *Nat. Electron.*, vol. 5, no. 2, pp. 113−122, 2022.

[157] T. Wan, S. Ma, F. Liao, L. Fan, and Y. Chai, "Neuromorphic sensory computing," *Sci. China Inf. Sci.*, vol. 65, no. 4, p. 141401, 2021.

[158] X. Guo, J. Xiang, Y. Zhang, and Y. Su, "Integrated neuromorphic photonics: synapses, neurons, and neural networks," *Adv. Photonics Res.*, vol. 2, no. 6, p. 2000212, 2021.

[159] A. Argyris, "Photonic neuromorphic technologies in optical communications," *Nanophotonics*, vol. 11, no. 5, pp. 897−916, 2022.

[160] S. Abdollahramezani, O. Hemmatyar, H. Taghinejad, et al., "Tunable nanophotonics enabled by chalcogenide phase-change materials," *Nanophotonics*, vol. 9, no. 5, pp. 1189−1241, 2020.

[161] Z. Cheng, C. Ríos, W. H. P. Pernice, C. D. Wright, and H. Bhaskaran, "On-chip photonic synapse," *Sci. Adv.*, vol. 3, no. 9, 2017. https://doi.org/10.1126/sciadv.1700160.

[162] M. Miscuglio, A. Mehrabian, Z. Hu, et al., "All-optical nonlinear activation function for photonic neural networks [invited]," *Opt. Mater. Express*, vol. 8, no. 12, p. 3851, 2018.

[163] J. Robertson, M. Hejda, J. Bueno, and A. Hurtado, "Ultrafast optical integration and pattern classification for neuromorphic photonics based on spiking VCSEL neurons," *Sci. Rep.*, vol. 10, no. 1, 2020. https://doi.org/10.1038/s41598-020-62945-5.

[164] M. Skontranis, G. Sarantoglou, S. Deligiannidis, A. Bogris, and C. Mesaritakis, "Time-multiplexed spiking convolutional neural network based on VCSELs for unsupervised image classification," *Apple. Sci.*, vol. 11, no. 4, p. 1383, 2021.

[165] C. Mesaritakis, A. Kapsalis, A. Bogris, and D. Syvridis, "Artificial neuron based on integrated semiconductor quantum dot mode-locked lasers," *Sci. Rep.*, vol. 6, no. 1, 2016, Art no. 39317.

[166] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[167] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[168] N. C. Harris, J. Carolan, D. Bunandar, et al., "Linear programmable nanophotonic processors," *Optica*, vol. 5, no. 12, p. 1623, 2018.

[169] J. Feldmann, N. Youngblood, M. Karpov, et al., "Parallel convolutional processing using an integrated photonic tensor core," *Nature*, vol. 589, no. 7840, pp. 52–58, 2021.

[170] L. G. Wright, T. Onodera, M. M. Stein, et al., "Deep physical neural networks trained with backpropagation," *Nature*, vol. 601, no. 7894, pp. 549–555, 2022.

[171] C. Wu, H. Yu, S. Lee, R. Peng, I. Takeuchi, and M. Li, "Programmable phase-change metasurfaces on waveguides for multimode photonic convolutional neural network," *Nat. Commun.*, vol. 12, no. 1, p. 2021. https://doi.org/10.1038/s41467-020-20365-z.

[172] H. J. Charlesworth and M. S. Turner, "Intrinsically motivated collective motion," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 116, pp. 15362–15367, 2019.

[173] A. Strandburg-Peshkin, C. R. Twomey, N. W. F. Bode, et al., "Visual sensory networks and effective information transfer in animal groups," *Curr. Biol.*, vol. 23, pp. R709–R711, 2013.

[174] A. Attanasi, A. Cavagna, L. Del Castello, et al., "Information transfer and behavioural inertia in starling flocks," *Nat. Phys.*, vol. 10, pp. 691–696, 2014.

[175] H. C. Berg and R. M. Berry, "E.coli in motion," *Phys. Today*, vol. 58, pp. 64–64, 2005.

[176] A. Sengupta, F. Carrara, and R. Stocker, "Phytoplankton can actively diversify their migration strategy in response to turbulent cues," *Nature*, vol. 543, pp. 555–558, 2017.

[177] G. M. Viswanathan, M. G. E. Da Luz, E. P. Raposo, and H. E. Stanley, *The physics of foraging: an introduction to random searches and biological encounters*, New York, Cambridge University Press, 2011.

[178] O. Bénichou, C. Loverdo, M. Moreau, and R. Voituriez, "Intermittent search strategies," *Rev. Mod. Phys.*, vol. 83, 2011, Art no. 81.

[179] T. Vicsek and A. Zafeiris, "Collective motion," *Phys. Rep.*, vol. 517, nos. 3–4, pp. 71–140, 2012.

[180] J. M. Yeomans, "Nature's engines: active matter," *EuroPhys News*, vol. 48, pp. 21–25, 2017.

[181] T. Sanchez, D. T.N. Chen, S. J. DeCamp, M. Heymann, and Z. Dogic, "Spontaneous motion in hierarchically assembled active matter," *Nature*, vol. 491, no. 7424, pp. 431–434, 2012.

[182] J. Urzay, A. Doostmohammadi, and J. M. Yeomans, "Multi-scale statistics of turbulence motorized by active matter," *J. Fluid Mech.*, vol. 822, pp. 762–773, 2017.

[183] J. Palacci, S. Sacanna, A. P. Steinberg, D. J. Pine, and P. M. Chaikin, "Living crystals of light-activated colloidal surfers," *Science*, vol. 339, pp. 936–940, 2013.

[184] I. Buttinoni, J. Bialké, F. Kümmel, H. Löwen, C. Bechinger, and T. Speck, "Dynamical clustering and phase separation in suspensions of self-propelled colloidal particles," *Phys. Rev. Lett.*, vol. 110, no. 23, p. 238301, 2013.

[185] M. Z. Miskin, A. J. Cortese, K. Dorsey, et al., "Electronically integrated, mass-manufactured, microscopic robots," *Nature*, vol. 584, p. 2020. https://doi.org/10.1038/s41586-020-2626-9.

[186] D. Andrén, D. G. Baranov, S. Jones, G. Volpe, R. Verre, and M. Käll, "Microscopic metavehicles powered and steered by embedded optical metasurfaces," *Nat. Nanotechnol.*, vol. 16, p. 2021.

[187] S. Kriegman, D. Blackiston, M. Levin, and J. Bongard, "A scalable pipeline for designing reconfigurable organisms," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 117, no. 4, pp. 1853–1859, 2020.

[188] M. C. Gather and S. H. Yun, "Single-cell biological lasers," *Nat. Photonics*, vol. 5, no. 7, pp. 406–410, 2011.

[189] K. Franze, J. Grosche, S. N. Skatchkov, et al., "Müller cells are living optical fibers in the vertebrate retina," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 104, no. 20, pp. 8287–8292, 2007.

[190] M. Arjovsky, L. Bottou, I. Gulrajani, et al., *Invariant Risk Minimization*, 2019, arXiv e-prints.

[191] X. Dong, Z. Yu, W. Cao, Y. Shi, and Q. Ma, "A survey on ensemble learning," *Front. Comput. Sci.*, vol. 14, pp. 241–258, 2020.

[192] Y. Chen and L. Dal Negro, "Physics-informed neural networks for imaging and parameter retrieval of photonic nanostructures from near-field data," *APL Photonics*, vol. 7, 2022, Art no. 010802.

[193] P. W. Battaglia, J. B. Hamrick, V. Bapst, et al., *Relational inductive biases, deep learning, and graph networks*, arxiv preprint 1806.01261, 2018.

[194] K. O. Stanley and R. Miikkulainen, "Evolving neural networks through augmenting topologies," *Evol. Comput.*, vol. 10, pp. 99–127, 2002.

[195] E. Papavasileiou, J. Cornelis, and B. Jansen, "A systematic literature review of the successors of 'neuroevolution of augmenting topologies'," *Evol. Comput.*, vol. 29, pp. 1–73, 2020.

[196] R. F. Laine, I. Arganda-Carreras, R. Henriques, and G. Jacquemet, "Avoiding a replication crisis in deep-learning-based bioimage analysis," *Nat. Methods*, vol. 18, no. 10, pp. 1136–1144, 2021.

[197] L. Jin, B. Liu, F. Zhao, et al., "Deep learning enables structured illumination microscopy with low light levels and enhanced speed," *Nat. Commun.*, vol. 11, no. 1, pp. 1–7, 2020.

[198] E. Gómez-de Mariscal, C. García-López-de Haro, W. Ouyang, et al., "DeepImageJ: a user-friendly environment to run deep learning models in ImageJ," *Nat. Methods*, vol. 18, no. 10, pp. 1192–1195, 2021.

[199] B. J. Heil, M. M. Hoffman, F. Markowetz, S.-I. Lee, C. S. Greene, and S. C. Hicks, "Reproducibility standards for machine learning in the life sciences," *Nat. Methods*, vol. 18, no. 10, pp. 1132–1135, 2021.

[200] I. Walsh, D. Fishman, D. Garcia-Gasulla, et al., "DOME: recommendations for supervised machine learning validation in biology," *Nat. Methods*, vol. 18, no. 10, pp. 1122–1127, 2021.