

## Review

# Optical multilayer thin film structure inverse design: From optimization to deep learning

Taigao Ma,<sup>1,3</sup> Mingqian Ma,<sup>2</sup> and L. Jay Guo<sup>2,\*</sup><sup>1</sup>Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA<sup>2</sup>Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109, USA<sup>3</sup>Present address: Visa Research, Austin, TX 78759, USA\*Correspondence: [guo@umich.edu](mailto:guo@umich.edu)<https://doi.org/10.1016/j.isci.2025.112222>

## SUMMARY

Optical multilayer thin film structures have been widely used in numerous photonic domains and applications. The key component to enable these applications is the inverse design. Different from other photonic structures such as metasurface or waveguide, multilayer thin film is a one-dimensional structure, which deserves its own treatment for the design process. Optimization has always been the standard design algorithm for decades. Recent years have witnessed a rapid development of integrating different deep learning algorithms to tackle the inverse design problems. A natural question to ask is: how do these algorithms differ from each other? Why do we need to develop so many algorithms and what type of challenges do they solve? What is the state-of-the-art algorithm in this domain? Here, we review recent progress and provide a guide tour through this research area, starting from traditional optimization to recent deep learning approaches. Challenges and future perspectives are also discussed.

## INTRODUCTION

Optical multilayer thin film structure<sup>1–3</sup> (shortened as multilayer structure) is a type of photonic platform that consists of multiple layers of material stacked together, with thickness typically in the range of tens of nanometers to several hundred nanometers. Because of their ease of fabrication, they have been widely used in many different photonic applications. For example, most photovoltaics<sup>4,5</sup> devices are based on thin film structures to achieve high power conversion efficiency in large-scale manufacturing. Multilayer thin film-based transmissive/reflective filters<sup>6,7</sup> are widely used as functional coatings and protective layers for optical lenses and glass pieces. The light interference phenomenon inside multilayer structures also provides an ideal platform to make structural color coatings<sup>8,9</sup> as an eco-friendly and advanced decoration with high aesthetic expression. Other important applications include displays,<sup>10,11</sup> touch screens,<sup>12</sup> radiative cooling devices,<sup>13</sup> to name a few.

For a given multilayer structure, obtaining its optical responses is easy and straightforward, through well-developed electromagnetic simulation algorithms, such as Transfer Matrix Methods<sup>14</sup> (TMM). However, the inverse problem, to design a multilayer structure with desired optical responses is non-trivial and usually does not have a closed-form. The inverse design requires determining the total number of layers and identifying the best combination of materials at each layer as well as their corresponding thickness, which is a combinatorial optimization problem. At the early stage, many designs are based on researchers' physics intuition or optical knowledge. However, this method is only suitable for some simple and conventional

structures. As design complexity increases, optimization-based methods have been widely adapted to automatically identify optimal structures, including needle optimization,<sup>15,16</sup> particle swarm optimization,<sup>17</sup> genetic algorithms,<sup>18</sup> and so forth. To speed up the design process, in recent several years, researchers have started to use deep learning as an alternative approach for inverse design.<sup>19–23</sup> After learning from the training dataset, these deep learning models can establish a direct mapping from optical responses to the multilayer structure and, therefore, can finish the design process within a fraction of a second, with a speed much faster than the optimization approach. Many different deep learning models have already emerged and demonstrated excellent performance in diverse applications, including tandem networks,<sup>19,24,25</sup> Variational Auto-Encoders (VAE),<sup>26</sup> Generative Adversarial Networks (GAN),<sup>27</sup> Mixture Density Networks (MDN),<sup>20,21</sup> and so forth. There are also many nice benchmarking works that compare different inverse design methods.<sup>28–30</sup>

With such significant progress in multilayer structure inverse design, a comprehensive review of these inverse design algorithms could benefit the relevant research community by providing a high-level perspective on the development and future directions. Though there are many excellent reviews that summarize deep learning-based inverse design for photonic and nano-photonic,<sup>31–38</sup> there is essentially no review that focuses on the multilayer structure inverse design. Different from other photonic structures such as metasurface or waveguide, multilayer thin film is a one-dimensional structure, which requires special treatment and considerations during the design process. For example, recently many researchers started to treat



multilayer structures as a sequence,<sup>23,39,40</sup> which is a special type of data that has not been considered in other types of photonic structures.

In this review, we will fill this gap by providing a perspective of multilayer structure inverse design and looking through the research roadmap. In the first part, we will provide an overview of multilayer structure inverse design by dissecting the representations of both multilayer structures and optical responses and present the overall challenges. Then we will go through different mainstream inverse design methods and summarize their achievements from a historical perspective: starting from the traditional optimization-based methods over decades ago, then the deep learning using vectorized representation starting around several years ago, the recent combined methods using both optimization and deep learning, and finally the deep learning using sequential representation, which was just proposed in recent two years. We hope this review could provide insights into current research development and promote the exploration of new research frontiers.

## MULTILAYER STRUCTURE INVERSE DESIGN

In this section, we will provide a general introduction to multilayer inverse design fundamentals. Inverse design usually deals with finding the suitable structures that can satisfy the desired optical response. Therefore, there are two aspects that researchers are interested in: multilayer structures and optical responses. We will first describe core concepts that we need to figure out when designing a multilayer structure, as well as three different methods when representing a multilayer structure. This is the part where readers can tell the difference between multilayer design and other nanophotonic inverse designs. Following this, several common types of optical responses with general interest and wide applications will be discussed. In the last part of this section, we will summarize several challenges during inverse design and provide an illustrative table to show how existing inverse design methods can tackle these challenges in different aspects. Details of these methods will be discussed in the following sections.

### Design parameters for multilayer structure

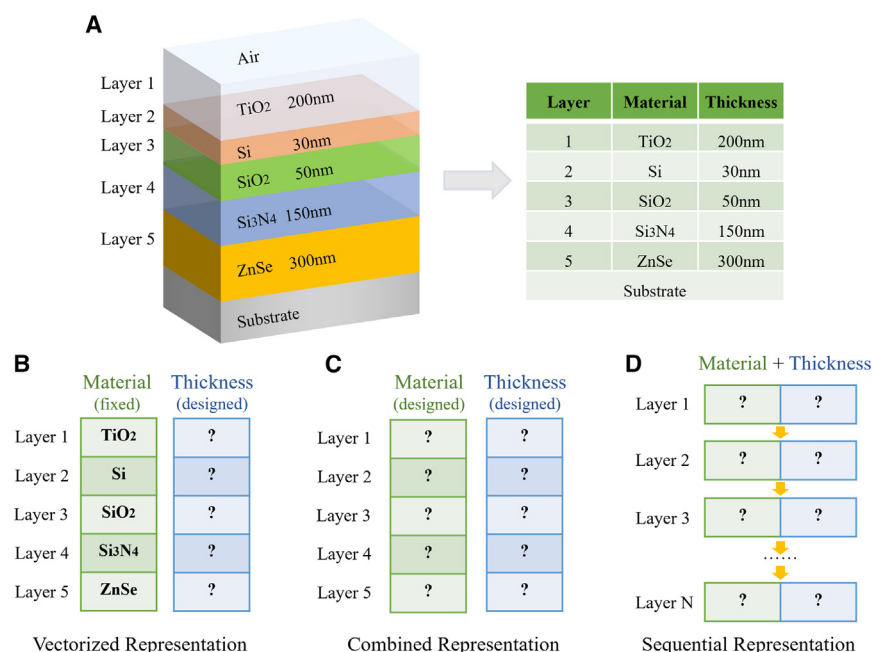
As the name suggests, a multilayer structure consists of multiple layers of materials stacking on top of each other. Usually, each layer can be made from a different material with different thickness, and these different layers contribute together to control light propagation and interaction inside the medium, and collectively determine the multilayer structure's optical responses. For a multilayer structure, there are three sets of parameters to be determined during the design process: the total number of layers, material combinations at each layer, and their corresponding thickness.

The total number of layers usually depends on the type of optical responses. For example, only 3–4 layers of suitable materials can already make a nice structural color filter.<sup>8,41,42</sup> However, in some other cases, such as designing a perfect reflective or transmissive filter, the total number of layers can easily exceed 100. Generally, there is no specific criterion for how many layers should be used for a given application; however, it is always

desirable to use structures with fewer layers considering the fabrication complexity and cost. Material combinations are another important design parameter. In optical frequency, different materials usually have different refractive indices, which uniquely determine how light will pass through and interact with the layers. Based on their absorption ability in optical frequency, most materials can be divided into two types: dielectric, or metallic. Dielectric materials such as silicon dioxide (SiO<sub>2</sub>) or titanium dioxide (TiO<sub>2</sub>) typically have low optical absorption and allow a significant portion of light to pass through in the visible range, while metallic materials such as silver (Ag) or gold (Au) have high absorption and are usually opaque in optical frequency. Note that we also include semiconductor materials in the dielectric category. Identifying the correct material selections at each layer can effectively determine how light is propagating inside the layers as well as light refraction at each interface, and thus, the optical responses in the layered system. However, this is not a trivial task as materials are usually dispersive, i.e., the refractive index is dependent on the incident light wavelength. The last important design parameter is the thickness at each layer. The major usage of thickness is to introduce a desirable propagation phase while light passes through each layered medium. In some special cases, this can create constructive or destructive interference phenomenon. One example is the Distributed Bragg Reflector (DBR)<sup>43–45</sup> that consists of alternating materials with high/low refractive index profiles. When each layer's thickness is around one-quarter of the effective wavelength, strong interference in the reflection side can lead to a perfect reflector that is close to 100%.

As an example, in Figure 1A, we illustrate a simple five-layer structure (left side) and use a table to illustrate these three design parameters (right side). This structure is made on a glass substrate which is not treated as a design parameter. As a notation, we count the number of layers from top to bottom, i.e., the first layer is closer to the air and the fifth layer is closer to the substrate. We will use this notation for the rest of this review. Ideally, we hope that an inverse design algorithm can identify all three design parameters based on the provided target optical responses. However, there are significant difficulties when trying to incorporate them into a computer-based inverse design algorithm. For example, how can we let an algorithm to understand the number of layers, which is supposed to be a varying number? In addition, how to represent different materials in the algorithm? Remember that materials usually have dispersion, which hinders us from using a single refractive index to distinguish them.

Perhaps the easiest design parameter to deal with is the thickness, which is a continuous variable that aligns well with computational considerations. Currently, there have been three different ways of representing a multilayer structure inside the inverse design algorithm and the first type is the vectorized representation that only relates to the thickness parameter. We illustrate this idea in Figure 1B. Within this representation, one needs to first determine how many layers to use as well as material choices at each layer based on their prior knowledge or experience, and only use the algorithm to find out the thickness. Common algorithms include some optimization-based methods such as Particle Swarm Optimization<sup>17</sup> (PSO), and many machine learning based methods such as tandem networks.<sup>19</sup>



**Figure 1. Visualization of multilayer structures as well as three common types of representation**

(A) An example of a five-layer thin film structure (left) as well as the three design parameters presented inside a table (right), including the total number of layers, the material combinations, and their thickness at each layer.

(B–D) Shows three widely used representation for a multilayer structure in inverse design algorithms. (B) is the vectorized representation, where material and total number of layers are fixed and only continuous thickness will be designed. (C) is the combined representation, where total number of layers are fixed and we need to design both material and thickness. Usually, material is a discrete variable that is selected from a pre-defined material database. (D) is the sequential representation, where all three parameters will be designed in a layer-by-layer way.

### Optical responses

In the previous section we mainly talk about the output of the inverse design,

namely, the design parameters. Here, we would like to briefly summarize the input of inverse design: the targeted optical responses for the designed structure. In optical frequency, multilayer structures have been widely used as transmissive/reflective filters, structural color coatings, absorbers, radiative cooling devices, and so forth. These applications can be attributed to the following types of optical responses.

- (1) *Transmission or Reflection Spectrum*: They are mostly used as design targets for many special filters. For example, a band-pass/band-notch filter requires a transmission window/block region at the desired frequency. Designing DBR is looking for a near 100% reflection spectrum within the expected wavelength region. On the other hand, a near-zero reflection spectrum is needed for designing anti-reflection coatings.
- (2) *Absorption Spectrum*: Light absorbance is very important for many energy-harvesting and detection applications. In solar-thermal harvesting devices, a high absorption over a broadband wavelength that covers the whole solar spectrum is always desirable. Many detection devices such as spectrometers require a perfect absorption spectrum to completely remove potential crosstalk and improve the measurement accuracy.
- (3) *Structural Color*: Compared to traditional color, structural color has unique advantages of higher spatial resolution, better stability, more environmental friendliness, and so forth. They can be well described using a color space such as RGB, xyY, or Lab values.

There are also many other important optical responses not listed here, e.g., combined spectrum and angular responses, which present higher level difficulties. Different from multilayer structures that require special representations in the design

The major limitation of this vectorized representation comes from the fact that its design performance is heavily determined by the pre-selected materials. Considering the large variety of materials in nature, it is possible that these selected materials may not lead to a good performance for any thickness choices,<sup>9</sup> which can leave the design algorithm running in vain. An inverse design algorithm that can automatically determine the material at each layer would be preferable. This leads to the second type of combined representation (see Figure 1C). Here, the material is represented by a discrete variable and thickness is a continuous variable. Usually, there are several candidate materials to be selected and they are denoted by different integers. For example, when there are two materials, we can use 0 to represent one material and 1 for the other material. In some special cases when all candidate materials have low dispersion, a single refractive index value can be used to represent the material. Many global optimization methods use this type of representation, such as genetic algorithm<sup>18</sup> and memetic algorithm.<sup>46</sup>

This combined representation has been popular for a long time. Only very recently, with the development of sequential machine learning models such as Recurrent Neural Networks<sup>47,48</sup> (RNN) and transformers,<sup>49</sup> researchers started to establish the third idea of sequential representation, describing a multilayer structure with a sequence (see Figure 1D). Instead of splitting material and thickness separately, this sequential representation treats both material and thickness on an equal footing, and assembles a multilayer structure in a layer-by-layer manner, just like in actual fabrication. Two recent models, the sequential decision network called OMLPPO,<sup>23,40</sup> and the conditional sequence generation network called OptoGPT,<sup>39,50</sup> are adapting this representation. Both methods have been demonstrated to automatically determine the optimal number of layers not known beforehand and terminate the design process when a target is reached.

**Table 1. A comparison of different inverse design algorithm for multilayer structure**

Methods	Reference	Year	Multilayer Representation	Global Design?	Efficient Design?	Diverse Design?	Flexible Design?
Needle Optimization	Tikhonravov et al. <sup>16</sup>	2007	Combined	Yes	No	Yes	No
Genetic Algorithm	Schubert et al. <sup>18</sup>	2008	Combined	Yes	No	No	No
PSO	Rabady and Ababneh <sup>17</sup>	2014	Vectorized	No	No	No	No
Memetic Algorithm	Shi et al. <sup>46</sup>	2018	Combined	Yes	No	No	No
Tandem Networks	Liu et al. <sup>19</sup>	2018	Vectorized	No	Yes	No	No
MDN	Unni et al. <sup>20</sup>	2020	Vectorized	No	Yes	Yes	No
VAE	Zandehshahvar et al. <sup>26</sup>	2021	Vectorized	No	Yes	No	No
GAN	Dai et al. <sup>51</sup>	2022	Vectorized	No	Yes	Yes	No
GLONet	Jiang and Fan <sup>52</sup>	2021	Combined	Yes	No	Yes	No
NEUTRON	Wang and Guo <sup>53</sup>	2022	Combined	Yes	No	Yes	No
OMLPPO	Wang et al. <sup>23</sup>	2021	Sequential	Yes	No	Yes	No
OptoGPT	Ma et al. <sup>39</sup>	2024	Sequential	Yes	Yes	Yes	Yes

algorithm, these optical responses (i.e., spectra) are usually continuous variables and can be easily represented as a high-dimension vector when integrated into the design algorithm. Therefore, we will not provide more description on this.

### Challenges

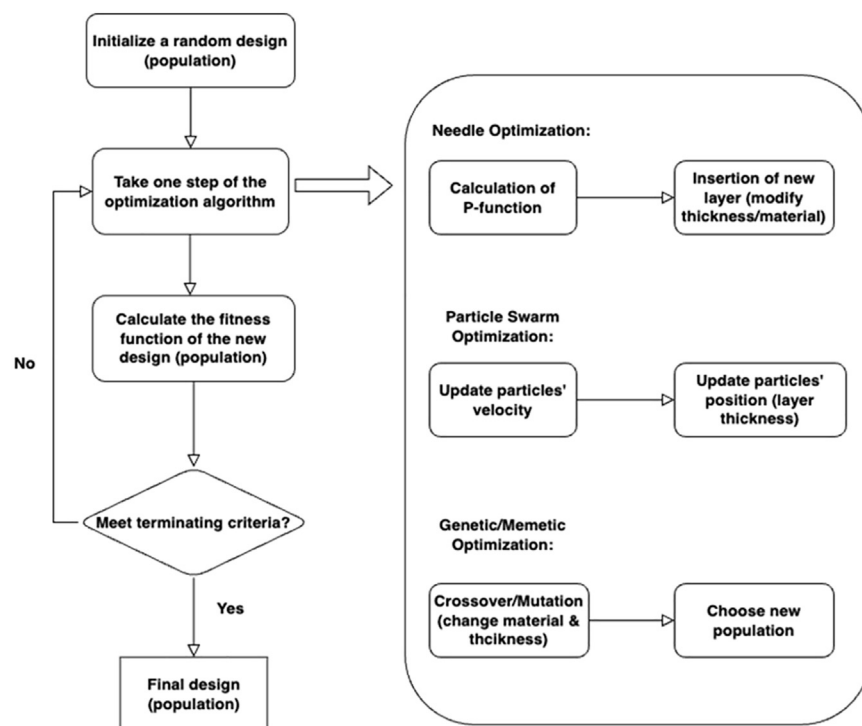
An ideal inverse design algorithm should be able to quickly output multilayer structures given any type of response as the input. However, this can lead to significant difficulties and challenges. The first challenge is how to design the multilayer structure globally (Global Design). By global we mean the algorithm can automatically design for all three sets of design parameters, including the total number of layers, the material sequence, and their thickness at each layer. As a comparison, local design only deals with thickness design, where the performance is heavily constrained by the fixed material selections. Global design is non-trivial as the design space is extremely large. For example, considering a five-layer structure where each layer has ten different material choices and the thickness also only has ten different discrete values, the design space can expand to  $10^{10}$  different combinations of structures. It is impossible to search all potential designs exhaustively, therefore, advanced algorithms are required to navigate through this large design space and identify the optimal structure.

The second challenge is how to finish a design quickly (Efficient Design). Traditional optimization algorithms are based on iterative evaluations and use fitness functions to update structures to reach user-defined goals. Although many numerical methods such as TMM can finish simulating a multilayer structure and obtaining its optical characteristics within 1 s, the iterative evaluation process often requires thousands or even more simulations, resulting in a significantly long processing time. The situation can be worse when multiple design targets are required. Speeding up the design process and making it more efficient would benefit everyone in this field.

In addition to these two major challenges related to the algorithm itself, researchers and engineers are also interested in leveraging inverse design algorithms to assist the practical fabrication process. Based on this, the third challenge is how to obtain

multiple different designed structures with minimal effort so that researchers can select the best one for their fabrication needs based on the availability of materials and deposition methods (Diverse Design). In principle, this is feasible because of the one-to-many mapping feature in the inverse design, meaning multiple structures can lead to similar optical properties. However, most existing algorithms cannot tackle this challenge because they can only output one definite design. An algorithm with probabilistic output can potentially solve this difficulty. The fourth challenge is how to integrate practical constraints into the design process (Flexible Design). With this challenge solved, researchers may restrict the material selections and thickness range at any desired layer based on their fabrication or design needs. For example, some may require the top layer to be dielectric material to protect the next metallic layers from oxidation. The latter two challenges have not yet been extensively explored.

In this review, we will focus on these four challenges and summarize how researchers came up with innovative ideas and advanced algorithms to tackle each of them. Decades ago, some global optimizations, including genetic algorithms<sup>18</sup> and memetic algorithms,<sup>46</sup> have been proposed to simultaneously design materials and thickness from a global perspective. Later around 2017, with the great success of machine learning, many different neural network-based methods were utilized to speed up the inverse design process. However, both methods have their own disadvantages: optimization-based algorithms are not efficient, while deep learning based algorithms cannot design structures globally. Therefore, later around 2020, the combination of optimization and deep learning provides a solution to alleviate these problems to design in the global space while still making the algorithms fast. However, the iterative process still takes quite a long time. Finally, in recent two or three years, with the development of sequential learning, researchers started to treat the multilayer structure as a sequence and eventually developed OptoGPT that can solve all four challenges, declaring the victory of the long-lasting design issues in multilayer structure. In Table 1, we summarize all these impactful algorithms and list their accomplishments.



**Figure 2. A general pipeline of optimization-based methods for inverse design**

tor of layered thickness that can adjust its trajectories based on its own experience and the experience of neighboring particles, balancing the exploration and exploitation. Rabady et al.<sup>17</sup> used PSO to design band-pass filters and demonstrated good performance (see [Figure 3A](#)). PSO has also been applied in various real-case multilayer thin film applications, including transmissive spectrum design,<sup>54</sup> inverse extraction of thickness and optical constants,<sup>55</sup> transmissive filters for image differentiation,<sup>56</sup> and so forth.

To further explore the global design of both material and thickness, Tikhonravov et al.<sup>15</sup> introduced needle optimization for optical coating design. Needle Optimization iteratively adds or adjusts layers in a thin film design to improve its optical properties toward the design target by modifying the thickness or refractive index

## OPTIMIZATION-BASED METHODS

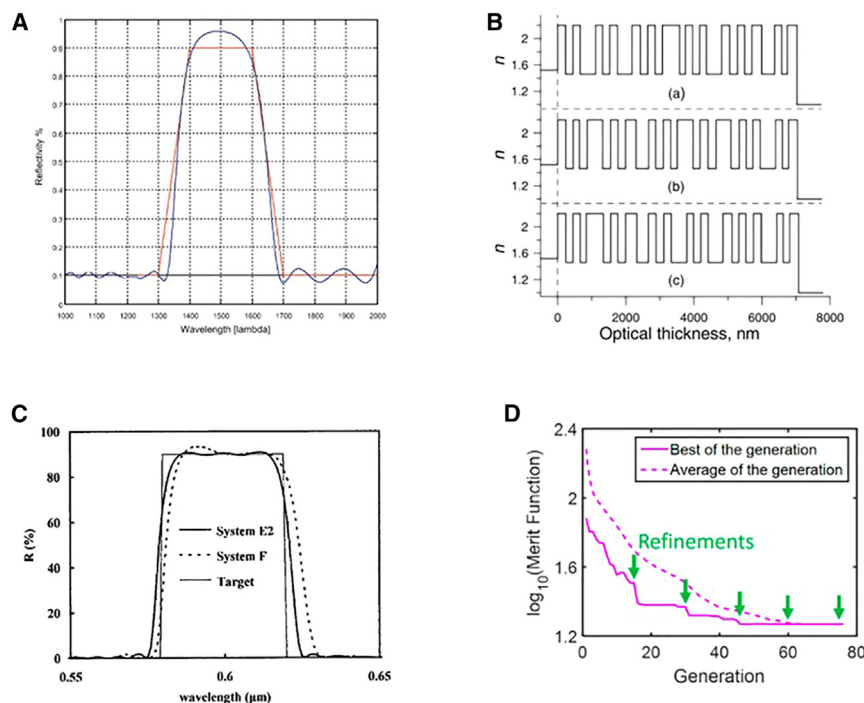
Since decades ago, heuristic optimization-based methods have become essential tools in tackling inverse design challenges, offering robust and flexible approaches to explore the vast solution space and find near-optimal designs. Although dealing with inverse design, these methods are mainly based on iterative forward simulation and follow a similar pipeline shown in [Figure 2](#). When given a design target, the algorithm will first start from a random structure and obtain its simulated optical responses. Then, the difference between the simulated responses and target responses will be evaluated and used to calculate the fitness function. If the fitness function does not meet the termination criteria, the algorithm will update the initial structure and keep running this iterative process until reaching satisfaction. After sufficient iterations, a structure with desired optical responses can be found. Different optimization methods, including needle optimization, particle swarm optimization, genetic algorithms, differ from each other in the way how they update the structure in each iteration, with illustrations can be found in [Figure 2](#). A summary and benchmark of optimization-based methods on photonic applications can be found in Schneider et al.<sup>30</sup> and Campbell et al.<sup>38</sup>

Optimization-based methods mainly deal with thickness or material design. In the beginning, researchers use the vectorized representation to only optimize the thickness. One such example is Particle swarm optimization (PSO), an evolutionary algorithm inspired by the social behavior of birds. PSO maintains a set of particles that move through the solution space and are influenced by their own best-known position and the best-known positions of the entire particles. In this case, each particle is a vec-

tor of the material configuration (see [Figure 3A](#)). The needle optimization evolves gradually, allowing for the generation of multiple solutions with nearly the same performance metrics, providing the designers with a range of options to select the most practical and manufacturable design. Another example is the application of genetic algorithms (GAs) to the design of optical multilayer structures by Martin et al.<sup>18</sup> Genetic algorithms are stochastic optimization methods inspired by the principles of natural selection and genetics. They operate by generating an initial population of potential solutions and iteratively evolving these solutions through selection, crossover, and mutation operations to maximize a fitness function, which is usually taken as the merit function commonly used in traditional approaches. In each iteration, the individuals with the highest fitness score are more likely to be selected for reproduction, ensuring that their genetic material is passed on to the next generation. Crossover combines parts of two parent solutions to create offspring with potentially better performance, while mutation introduces random changes to individual solutions to maintain diversity within the population and prevent premature convergence to local optima. The authors demonstrated the effectiveness of this algorithm by designing the refractive index and thickness of each structure layer for three different optical filters: antireflection coating, beam splitter, and rejection filter (see [Figure 3C](#)). This idea has also been applied to the design of broadband reflectors,<sup>57</sup> edge filters,<sup>58</sup> and so forth.

However, using the refractive index to represent materials restricts the design target to be a narrow band because of materials' dispersion. To design for a broader wavelength, Shi et al.<sup>46</sup> introduced a mixed-integer programming (MIP) with a genetic algorithm solution to optimize the design of multilayer thin





**Figure 3. Several design examples for optimization-based methods**

(A) Design performance of Particle swarm optimization (PSO) algorithm for a band-pass optical interference filter in the range of 1000–2000 nm wavelengths. Reprinted from Rabady and Ababneh<sup>17</sup> Copyright @ 2014, Elsevier.

(B) A demonstration of versatile design using Needle Optimization. Reprinted from Tikhonravov et al.<sup>16</sup> Copyright @ 2007, Optica Publishing Group.

(C) Genetic Algorithm designing a rejection filter of 90% over the 0.55–0.65  $\mu\text{m}$  region. Reprinted from Schubert et al.<sup>18</sup> Copyright @ 2008, Optica Publishing Group.

(D) Iterative refinement of memetic algorithm with mixed-integer programming (MIP) to design better generations. Reprinted from Shi et al.<sup>46</sup> Copyright @ 2018, ACS Publications.

films. MIP can handle both discrete and continuous variables in the optimization process. In this context, the discrete variables represent different dielectric materials, while continuous variables correspond to the thickness of each material layer. This algorithm improves upon traditional genetic algorithms by introducing discrete materials into the crossover and mutation steps (see Figure 3C). The authors demonstrated state-of-the-art performance using the algorithm on two design tasks with a broad wavelength range: radiative cooling devices and incandescent light bulb filters. The GA with the MIP algorithm's ability to design a variety of solutions with distinct material and thickness permutations discovered innovative designs for multilayer thin films at a global solution space for wide wavelengths, materials, and thicknesses.

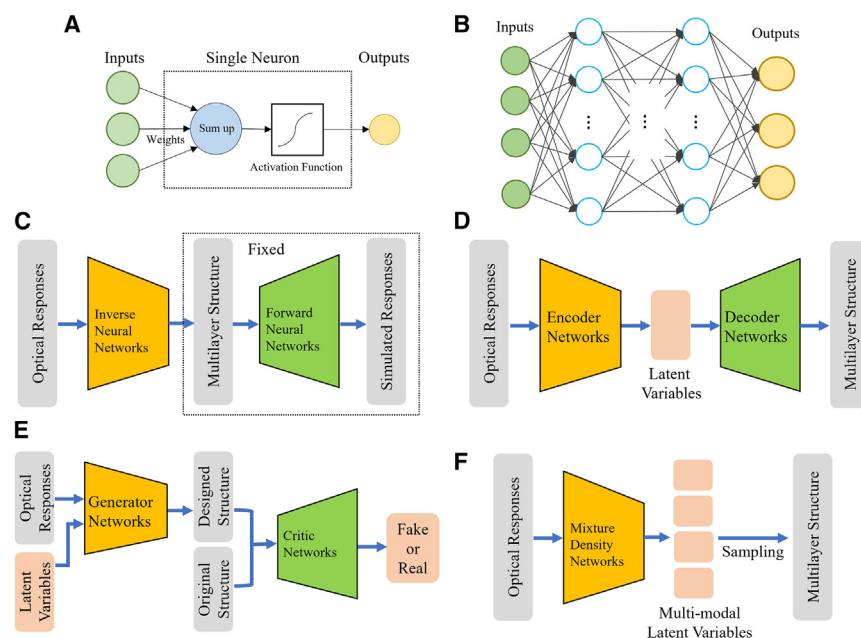
## DEEP LEARNING METHODS WITH VECTORIZED REPRESENTATION

The fundamental unit inside neural networks is the artificial neuron. As shown in Figure 4A, a single neuron first takes one or multiple inputs, then sums up all these inputs multiplied by learnable weights and outputs results through a nonlinear activation function. In this way, information can be passed in, get processed, and propagated to the next level. Usually, a single neural network (NN) can contain thousands or millions of neurons, leading to millions or billions of learnable weights. After training on a large dataset, these learnable weights are updated to capture the general mapping between the inputs and outputs and process extremely complicated information such as image classification,<sup>59–61</sup> speech recognition,<sup>62–64</sup> sequence decision,<sup>65–67</sup> language understanding,<sup>49,68–70</sup> and so forth. Different types of neural networks differ in the way how the neurons are connected,

and therefore, are suitable for processing different types of data. For example, Convolutional Neural Networks (CNN)<sup>59</sup> use the convolutional layers to capture spatial correlations, thus widely used for processing and analyzing image-type data. Recurrent Neural Networks (RNN)<sup>47</sup> employ recurrent connections to obtain temporal dependencies, making them well-suited for sequential data.

In the previous discussion, we concluded that most of these optical responses (inputs) can be denoted as vectors while multilayer structures (outputs) can be denoted using either vectorized, combined, or sequential representations. However, combined representations require a discretized output for material design, which is non-trivial for NN since their outputs are continuous variables. Also, it was not until recent two or three years ago that researchers figured out the sequential representation. Therefore, most of the early works that used deep learning for inverse design seek to solve the vector-to-vector mapping. A general working process is to first generate a training dataset containing multilayer structures and corresponding optical responses using simulation (e.g., TMM), then train the NN model to fit the inverse mapping function from optical responses to multilayer structures. Once trained, the model can be used to design multilayer structures for different optical responses instantaneously, providing a much faster alternative than optimizations, which usually take minutes or hours.

In terms of the deep learning models for vector-to-vector mapping, perhaps the most straightforward way is the Multilayer-Perceptron (MLP), where neurons are stacked layer-by-layer and all neurons are fully connected to other neurons in nearby layers (see Figure 4B). However, directly using MLP to inverse design will give inaccurate results. Because of the one-to-many mapping issue, i.e., there are multiple possible structures for a given optical response, it is difficult to minimize the loss function (the difference between model outputs and target outputs) and have the NN model converged during training. Eventually, the model will learn to output an averaged structure, which



**Figure 4. Illustration of different types of neural networks**

(A–F) Visualizations of (A) a single neuron, (B) Multilayer-Perceptron (MLP), (C) tandem networks, (D) Variational Auto-Encoders (VAE), (E) Generative Adversarial Networks (GAN), and (F) Mixture Density Networks (MDN).

usually does not have the desired optical responses.<sup>34</sup> Therefore, special buildup of neural networks is required to handle this one-to-many mapping issue. Many different types of neural networks have been proposed and demonstrated effectiveness in inverse design. Based on the type of outputs, they can be classified into two different categories: (1) deterministic NN, mainly including tandem networks, and (2) stochastic NN (or generative NN), including Generative Adversarial Networks (GAN), Variational Auto-Encoders (VAE), and Mixture Density Networks (MDN). In the remaining section, we are going to discuss each model and summarize its variants and applications.

### Deterministic NN

In deterministic NN, the model always maps the input into a single output. The most popular one is the tandem network. As the name “tandem” suggests, tandem networks connect two NN in series (see Figure 4C): the Forward Neural Networks (FNN) and the Inverse Neural Networks (INN). The FNN takes in the structural parameters and outputs the predictions of their corresponding optical responses. Once trained, we can use FNN to quickly predict the optical responses for a given structure with high accuracy. On the other hand, INN takes in the target optical responses and outputs predicted structures. The training of tandem networks involves a two-step training procedure for each individual network. In the first step, we need to train FNN separately to obtain a fast surrogate predictor for optical responses. The second step is to connect the output of INN to this pre-trained FNN and use FNN to supervise the learning of INN. In this way, we can resolve the one-to-many mapping issue by only predicting one structure suggested by FNN.

Tandem networks is first introduced in Liu et al.<sup>19</sup> to design transmission spectra for multilayer structures. The multilayer system in consideration is a twenty-layer structure with SiO<sub>2</sub>/Si<sub>3</sub>N<sub>4</sub> alternating material combinations. Using this method, the

authors have demonstrated excellent design performance for arbitrary spectra, as well as the special Gaussian shaped spectra (see Figure 5A). In addition, it only takes a fraction of a second to finish a design, which is much faster than optimization-based methods. Later, extensive works have been dedicated to utilizing tandem networks to design different multilayer structures and tackle various design problems efficiently, including designing a Fabry-Perot-cavity-based color filter,<sup>71</sup> organic light-emitting diode,<sup>72</sup> and so forth. In addition to multilayer structures, tandem networks are

also widely used in other types of photonic structures and target responses.<sup>73–76</sup> For example, Gao et al.<sup>73</sup> used tandem networks to design reflective structural color for silicon nanorod metasurface; Ma et al.<sup>74</sup> combined tandem networks to design a reflection spectrum for three-dimensional chiral metamaterials; He et al.<sup>75</sup> utilized tandem networks to design the far-field radiation profile for plasmonic nanoparticles.

There are also many variants to improve the design performance of tandem networks. For example, Xu et al.<sup>24</sup> modified and benchmarked different activation functions and loss functions used in tandem networks to find the optimal schema. They found that using ReLU activation function<sup>77</sup> together with intermediate structural loss can significantly increase the design accuracy (see Figure 5B). On the other hand, many researchers seek to find alternate neural networks to replace the MLP architecture used in INN and FNN.<sup>22,78</sup> For example, Chen et al.<sup>22</sup> proposed to use transformer architecture for both INN and FNN by leveraging the self-attention mechanism to effectively extract and process information. In this work, they proposed a meta-material spectrum transformer (MST) to design a high-performance broadband solar metamaterial absorber that exhibits a high average absorbance of 94% in a broad solar spectrum (see Figure 5C). In addition, considering that the original tandem network can only output one designed structure, Yuan et al.<sup>25</sup> came up with the idea of connecting the INN to multiple different FNNs, which they called multi-head tandem networks, to expand its capability of outputting multiple structures.

### Stochastic NN

Different from deterministic NN, stochastic NN (also called generative NN) can output multiple different predictions given the same input. This is because these models take in one latent variable randomly sampled from a normal distribution as an extra input, which enables them to generate different outputs

conditioned on this extra input. Therefore, they can intrinsically solve the one-to-many mapping issue. There are three major types of generative NN that are widely used in inverse design: VAE,<sup>79</sup> GAN,<sup>80</sup> and MDN.<sup>81</sup> We illustrate their model architectures in Figure 4D–4F, respectively. There are already many reviews<sup>32,33,35</sup> and benchmark<sup>28,29,82,83</sup> that discuss these models, but toward a more general audience on the photonic inverse designs. Here, we mainly summarize the work done for multilayer structure designs.

Figure 4D shows the model architecture for VAE, which originates from information theory: it first uses an NN (called encoders) to encode the high-dimensional input information into a low-dimension hidden representation, usually with a normal distribution, then uses another NN (decoders) to decode hidden representations back into the corresponding output. During inverse design, each input (target optical responses) will be encoded to different normal distributions with different means and variances. Sampling from this distribution and sending it through the decoder will give the designed structure. Since this sampling process is random, VAE can output multiple structures when decoding from different latent variables, thus solving the one-to-many mapping issue naturally. For example, Zandehshahvar et al.<sup>26</sup> used the VAE to design multilayer structures composed of consecutive layers of SiO<sub>2</sub> and TiO<sub>2</sub>. In their work, they demonstrated that when given a single transmission spectrum as input, their model can output three different structures (Figure 6A left side). The simulation shows that their simulated spectra are close to the target spectrum (Figure 6A right side). VAE has also demonstrated successful applications in other types of photonic structures, including Fano-type resonant nanopillars,<sup>84</sup> hybrid plasmonic/phase-change material metasurface,<sup>85</sup> a periodic array of Au nanoribbons,<sup>86</sup> and so forth.

Another type of stochastic NN is GAN (Figure 4E). It also contains two neural networks: the generator NN that generates designed structures based on the random latent variables and the input optical responses, and the critic NN that learns to distinguish if the input structure comes from a model or not. Different from VAE, GAN originated from game theory and these two NN learn to generate by combating with each other: the generator tries to generate outputs that are distributed as close to the training dataset as possible to fool the critic, while the critic tries to distinguish the generated outputs from the ground truth in the training dataset. A successful application is demonstrated in Dai's work,<sup>51</sup> where they used GAN to design a transmissive Fabry-Perot-cavity-based color filter using the three-layer Ag-SiO<sub>2</sub>-Ag structure and identified an average number of 3.58 solutions for each color. Experimental results in Figure 6B also validated the effectiveness of these designs.

MDN is the third type of stochastic NN (see Figure 4F). Different from VAE and GAN which utilize single-modal normal distribution, MDN models the design parameters as a multimodal probability distribution. Here, each distribution in this multi-modality is parameterized by a different normal distribution with unique mean, variance, and weight. The total number of distributions is a hyperparameter that is usually pre-determined and can be tens or hundreds. After training, the designed structure will be obtained after sampling from this predicted multimodal distribution, which inherently solves the one-to-many mapping issue. As an

example, Unni et al.<sup>20</sup> utilized MDNs with 16 mixtures to design 10 layers of alternating SiO<sub>2</sub> and TiO<sub>2</sub> structure. Their results in Figure 6C demonstrated that the MDN can give multiple structures that satisfy the desired transmission spectrum.

All these four different neural networks, including tandem networks, GAN, VAE, and MDN have been exploited for multilayer inverse design with vectorized representation. However, each model still has its own limitations. For example, tandem networks are simple to understand and easy to train, but they can only output one designed structure for one input target, which can lead to difficulty when multiple structures are desired. On the other hand, all the other three models are stochastic models and can output diverse structures. Both GAN and VAE samples from a single-modal distribution, but GAN usually has higher diversity than VAE, while VAE is more robust than GAN. With a multi-modality distribution, MDN can learn a more diverse solution than both GAN and VAE. However, MDN is harder to train because of the increased number of modalities and parameters. A delicate adjustment of a number of modalities based on the specific dataset is required to obtain better performance. For an exhaustive discussion, we recommend some benchmark works that compare these models in a fair setting.<sup>28,29,83</sup>

## OPTIMIZATION-COMBINED DEEP LEARNING

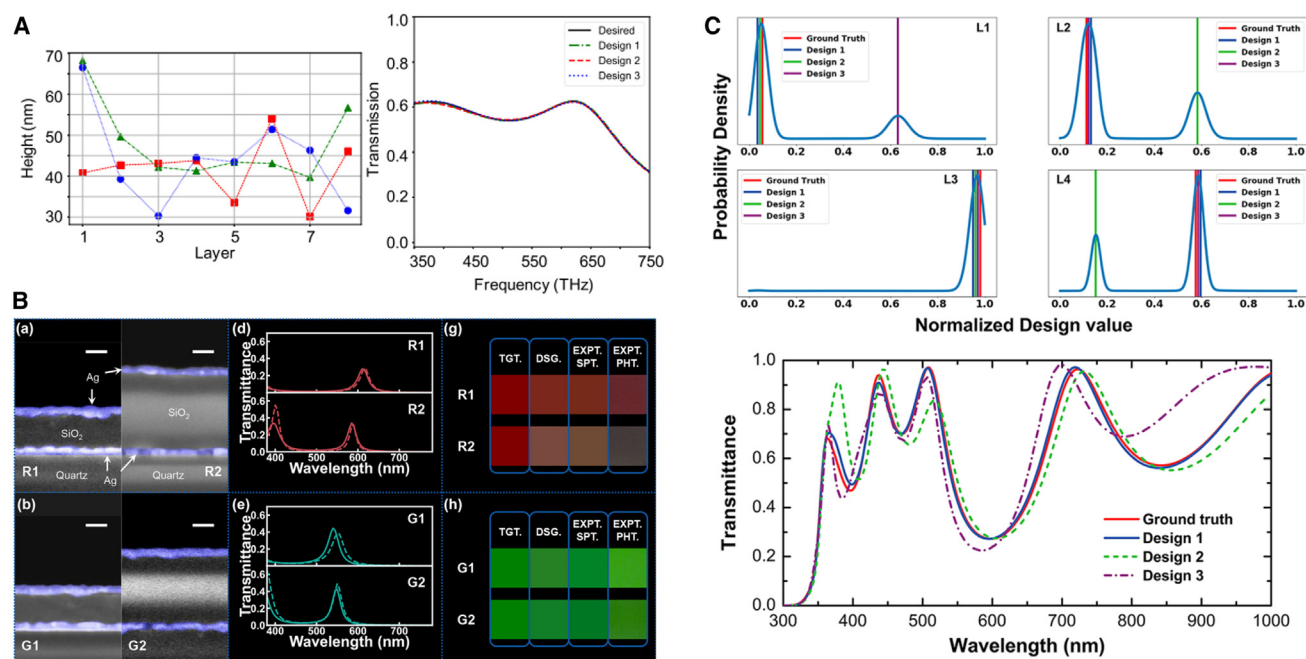
Although optimization and deep learning based methods are widely used for inverse design, both have their own strengths and limitations. Optimization has much better flexibility to incorporate the material into design consideration. Designing in this global space with both materials and thicknesses usually leads to better performance than deep learning-based methods. However, optimization requires iterative evaluations to minimize the fitness function, which can be time-consuming. In addition, one needs to restart the optimization process from scratch whenever encountering a new design target, making it task-specific. On the other hand, because deep learning based methods usually use neural networks to learn a general mapping from the space of optical responses to the space of multilayer structure, they can finish a design task much faster than optimization and can quickly adapt to a different design target. The probability output for stochastic NN also enables multiple designs to solve one-to-many mapping issues. However, these neural networks usually have vectorized outputs, making it difficult to incorporate the discrete materials into outputs, thus, resulting in limited design performance. A natural question to ask is if it is possible to combine both methods together and obtain greater benefits. It is encouraging to see that many researchers have already started to do so. Based on how these two methods are integrated together, we can classify them into three categories and we will summarize each of them later in discussion.

### Neural networks as a fast surrogate simulator

Neural networks are not only fast for inverse design, but also a fast surrogate simulator to predict the optical responses given the structure.<sup>19,50,87</sup> Considering that optimization requires an iterative evaluation process, it is straightforward to use NN as the surrogate simulator to speed up the optimization process. For example, Unni et al.<sup>21</sup> proposed a mixture-density-based







**Figure 6. Stochastic NN for inverse design with multiple outputs**

(A) VAE can output multiple designs that satisfy the desired transmission spectrum. Reprinted from Zandehshahvar et al.<sup>26</sup> Copyright © 2021, Optica Publishing Group.

(B) A three-layer structural color designed by GAN showing different structures. Experimental results also verified the design performance. Reprinted from Dai et al.<sup>51</sup> Copyright © 2022, Elsevier.

(C) The output probability from MDN showing a multi-modality distribution which corresponds to different designs. All designs show desired performance in the transmission spectrum. Reprinted from Unni et al.<sup>20</sup> Copyright © 2020, ACS Publications.

coating with 1.81% averaged reflection, which is lower than traditional optimization methods (See Figure 7D). In a later work, Zhang et al.<sup>92</sup> used GLONets to design the optical transfer function of a multilayer film at wavelengths of 532 nm and 633 nm and successfully achieved incoherent differentiation with a high resolution of 6.2  $\mu\text{m}$ , opening the door for high-speed imaging processing, object tracking, and disease diagnosis.

### Two-step design process

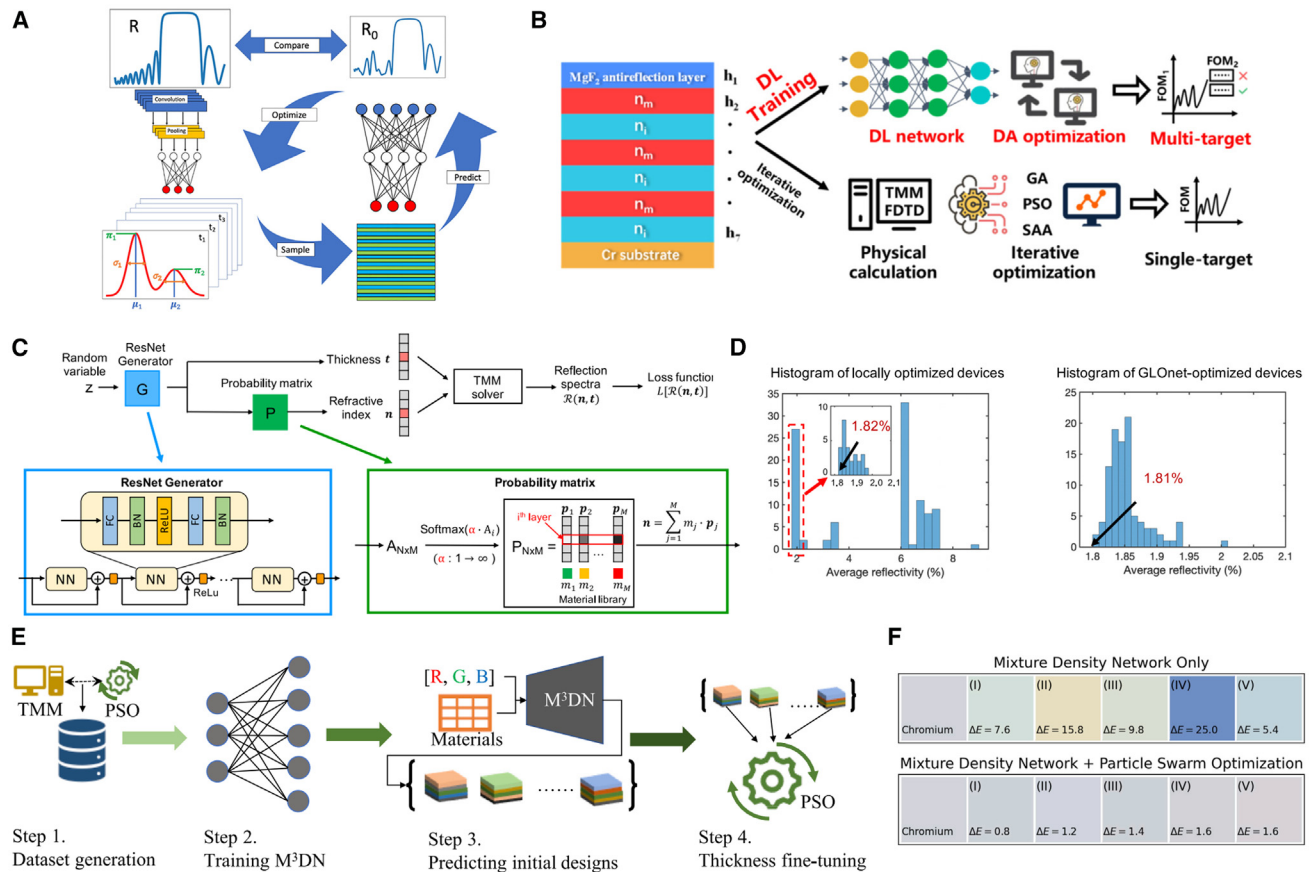
Different from previous two methods that use NN as a surrogate model, the third type splits the design process into two steps: first use NN to design material combinations, then use optimization to identify the suitable thickness. An example is the algorithm of Neural Particle Swarm Optimization (NEUTRON), which was introduced by Wang et al.<sup>53</sup> to inverse design structural colors. The idea was illustrated in Figure 7E. Here, instead of design material and thickness at the same time, NEUTRON first uses a classifier to predict the most suitable materials at each layer for a given layered structure, then uses MDN to give the initial thickness guess in a probability distribution form and combine PSO to refine and optimize the thickness and achieve optimal performance. Using this two-step process, they demonstrated a much more accurate structural color design compared to single MDN, and used it to output structures for a large number of colors in a painting. We give one design example in Figure 7F.

### DEEP LEARNING METHODS WITH SEQUENTIAL REPRESENTATION

Sequence modeling and understanding is an important research field in machine learning as most of the data in our world can be treated as sequences, including natural language,<sup>49</sup> human voice,<sup>93</sup> time-series,<sup>94</sup> video stream,<sup>95</sup> and so forth. Different from vectorized data with pre-determined shapes, sequential data usually have variable length and require special architecture in NN, e.g., the recurrent connection inside RNN,<sup>47,48</sup> to handle the serialized information. Recently, researchers have observed that the varying numbers of layers in a multilayer structure is analogous to the varying lengths in sequential data. Therefore, they have started to treat multilayer structures as a sequence and use sequential NN to deal with inverse design. In this way, researchers can completely automate the inverse design process for the first time: NN can design the material and thickness simultaneously at each layer, and can also design the total number of layers by smartly terminating the design process whenever NN thinks the current design is good enough. This points to a bright future for multilayer inverse design. Currently, there are two different sequential learning methods used for inverse design: sequential decision process and conditional sequence generation.

### Sequential decision process

The first type of sequential learning model is based on the sequential decision process,<sup>66</sup> where the model needs to decide



**Figure 7. Multiple ways to combine optimization and deep learning together**

(A) A mixture-density-based tandem optimization network. Reprinted from Unni et al.<sup>21</sup> Copyright © 2021, Elsevier.

(B) Deep learning combined with multi-objective double annealing algorithms. Reprinted from Qu et al.<sup>87</sup> Copyright © 2023, Elsevier.

(C) Global optimization networks (GLOnets) and (D) the design performance comparison using GLOnets and traditional optimization. Reprinted from Jiang and Fan.<sup>52</sup> Copyright © 2021, Elsevier.

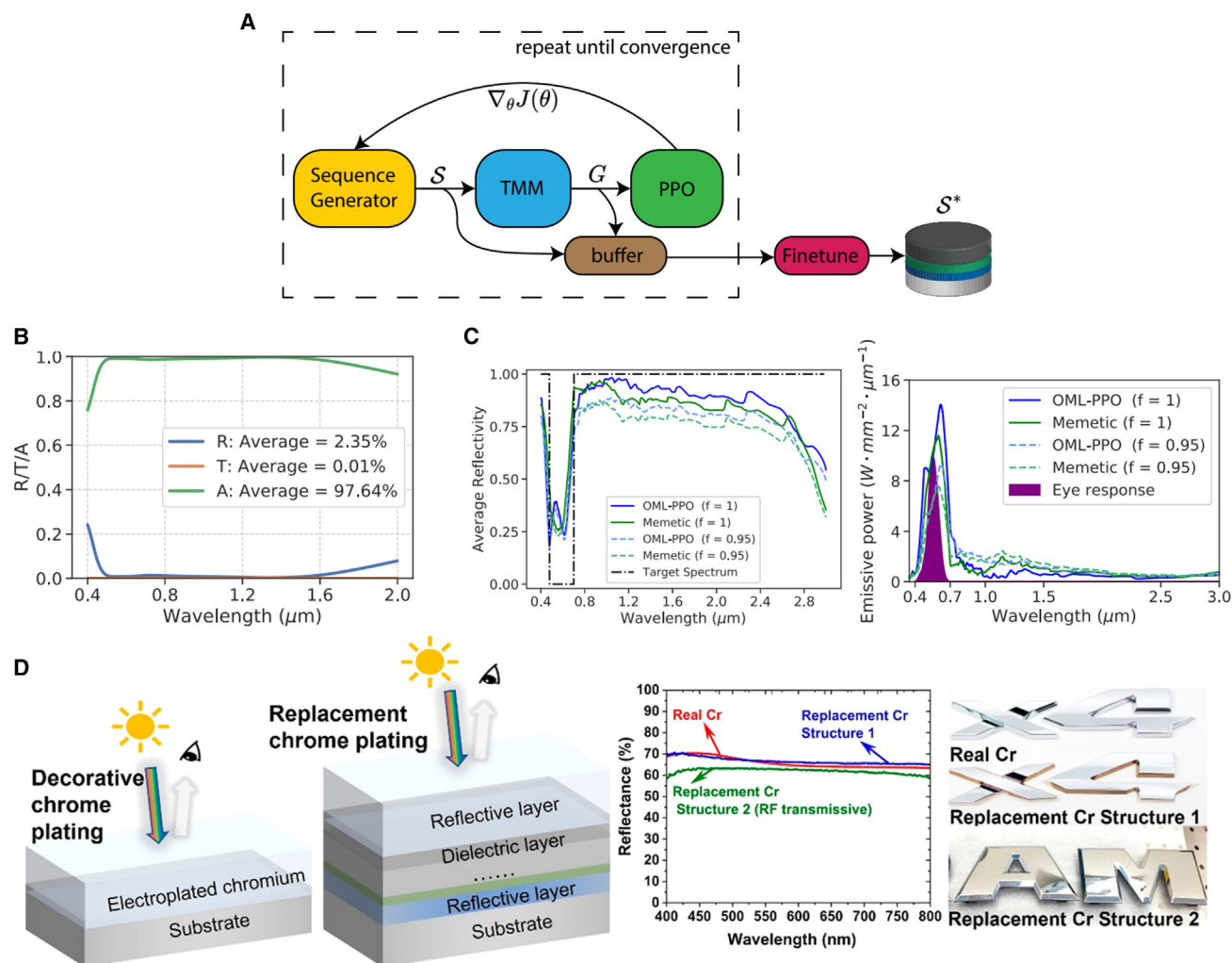
(E) NEUTRON: Neural particle swarm optimization for material-aware inverse design and its design performance combined to NN only (F). Reprinted from Wang and Guo.<sup>53</sup> Copyright © 2022, Cell Press.

what to do next based on the current situation. It has been widely used in robot manipulations,<sup>96</sup> game playing,<sup>65,67</sup> self-driving cars,<sup>97,98</sup> and so forth. In 2021, Wang et al.<sup>23</sup> proposed to use a sequential decision process for multilayer design and introduced the algorithm of OMLPPO based on the reinforcement learning algorithms.<sup>66,99</sup> In this work, OMLPPO designs and finalizes the multilayer design in a layer-by-layer manner: after figuring out the material and thickness for the first few layers, it will decide what type of material to use for the next layer as well as its corresponding thickness. This process will keep going until the designed structure reaches the desired performance, or the sequence length (namely, the number of layers) equals the maximum allowable number of layers. The proposed OMLPPO algorithm learns to make sequential decisions in a trial-and-error way and uses iterative evaluation to update designs and approach the ideal structure. The complete pipeline is shown in Figure 8A. Although this iterative process looks similar to heuristic optimizations, OMLPPO can learn multilayer structure design domain knowledge during training and leverage it to improve

the design performance. For example, the algorithm automatically designed a five-layer ultra-wideband absorber with 97.64% average absorption, which is ~2% more than the one designed by human experts (Figure 8B). They also used the algorithm to design incandescent light bulb filters and achieved an 8.5% higher visible light enhancement factor than the structure designed by optimization (Figure 8C). Because of its high performance, OMLPPO has also been adapted to solve other non-trivial design tasks. For example, Saha et al.<sup>40</sup> identified two multilayer structures that have a similar visible reflection spectrum as decorative Cr, serving as an environmentally friendly replacement for the harmful and toxic chemical finishing process (Figure 8D).

### Conditional sequence generation

Another type of sequential learning model is conditional sequence generation, where the model needs to generate serialized output based on the input. One important example is the language modeling<sup>100</sup>: to predict the probability of a sequence of words conditioned on previous words and construct coherent



**Figure 8. Sequential decision process for multilayer design**

(A) The working pipeline of OMLPPO algorithm.

(B) The perfect absorption spectrum designed by OMLPPO, which is 2% higher than demonstrated human designs.

(C) The incandescent light bulb filters designed by OMLPPO, which exhibits 8% performance improvement than optimization methods.

(A–C) are reprinted from Wang et al.<sup>23</sup> Copyright © 2021, IOP Publishing.

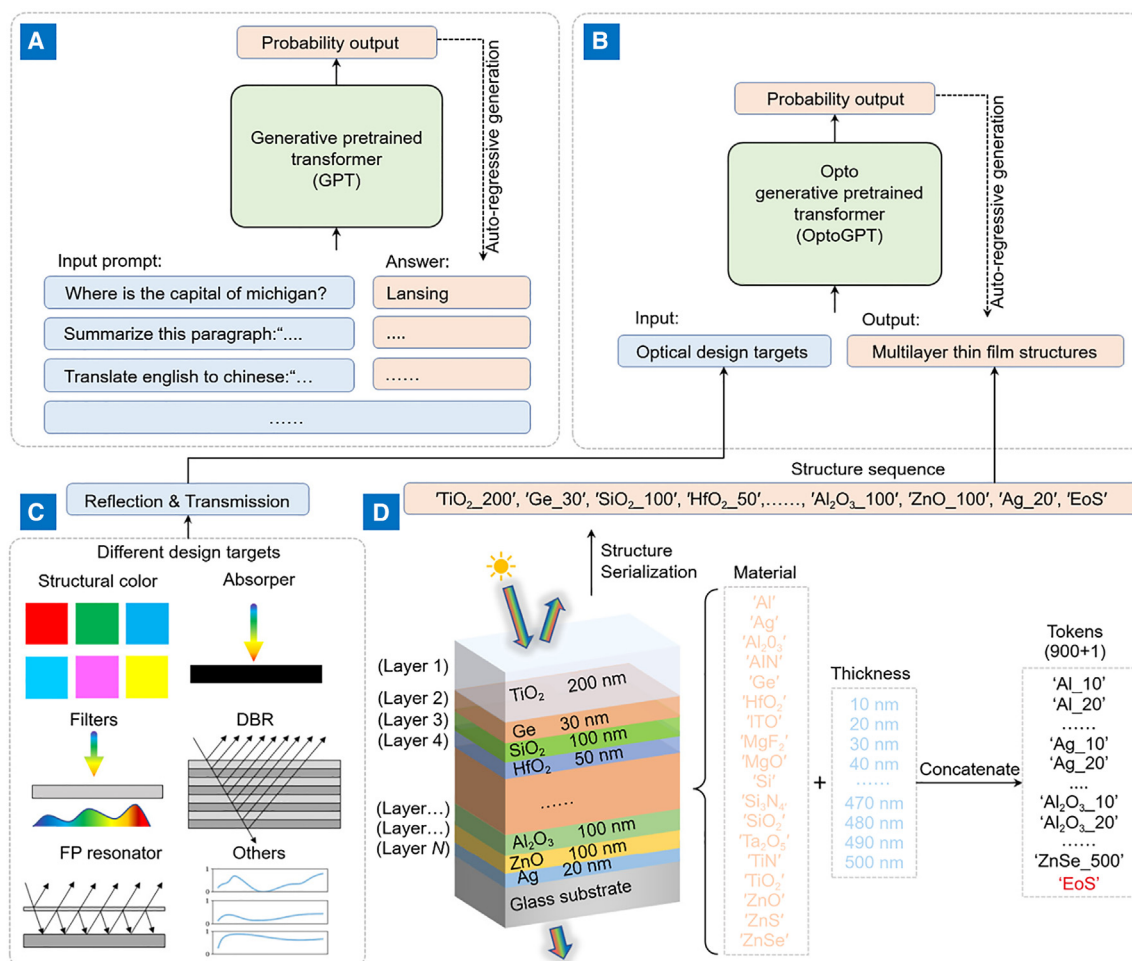
(D) The decorative Chrome-looking reflective coating designed by OMLPPO. Reprinted from Saha et al.<sup>40</sup> Copyright © 2023, ACS Publications.

and contextually relevant texts. It is the backbone for a wide range of natural language processing (NLP) applications, including machine translation,<sup>101</sup> question answering,<sup>102</sup> text generation,<sup>103</sup> and so forth. The recent introduction of GPT-type models<sup>69,103–105</sup> which use the decoder part of the Transformer<sup>49</sup> model, not only improves the learning efficiency of language modeling but also enhances the model's ability to generate more coherent texts by scaling up the model size and training dataset, leading to the development of powerful language models, including ChatGPT,<sup>106</sup> Llama,<sup>107</sup> BART,<sup>70</sup> and so forth. The success of the GPT model in language modeling also intrigues researchers to solve other sequential generation problems in scientific and engineering applications, including DNA sequence generation,<sup>108</sup> protein prediction,<sup>109</sup> molecule modeling,<sup>110,111</sup> and so forth.

Inspired by these successful GPT models, in 2024, Ma et al.<sup>39</sup> proposed the OptoGPT model and completely tackled these four challenges for multilayer design. Figure 9A and 9B show the analogy between GPT models and the proposed OptoGPT model. For GPT models, given some language texts as input, e.g., a question or task description, they can generate a sequence of words that are related to the inputs. Similarly, given the input of optical responses (in their case, is reflection and transmission spectrum), the OptoGPT model can output the serialized representation of multilayer structures directly, without any iterative process.

To make OptoGPT work, we first introduce the “structure serialization” technique to convert the material and thickness information at each layer into a single token by simply concatenating them together. As shown in Figure 9D, a token is a human-readable notation similar to a word or a phrase in language. Then,





**Figure 9. Idea of conditional sequence generation using OptoGPT model**

(A and B) Shows the diagram of general GPT models in NLP and the OptoGPT model in multilayer inverse design, respectively.

(C) Different types of input to the OptoGPT model as the design target.

(D) One example of the "structure tokenization" and "structure serialization" for an N-layer structure on the glass substrate. All figures are reprinted from Ma et al.<sup>39</sup> Copyright © 2024, OE Journals.

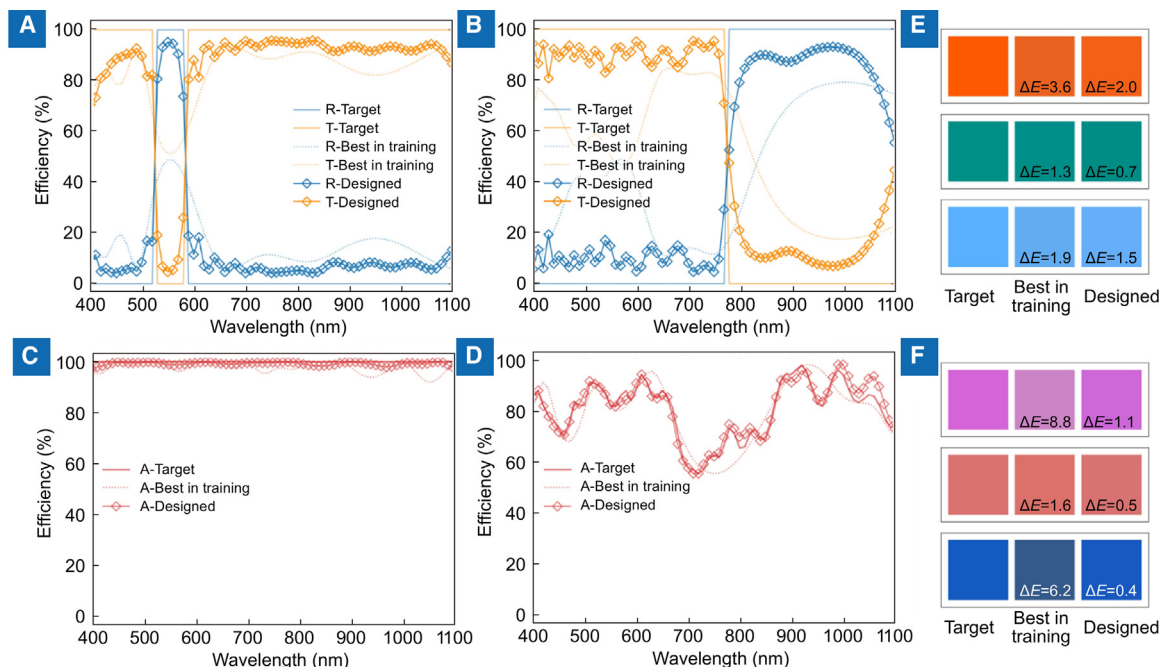
we proposed the "structure serialization" technique and added these tokens one by one to form the sequence representation of a multilayer structure. In this way, the OptoGPT model can effectively design for both material and thickness information, and automatically determine the optimal number of layers. The input to the OptoGPT model is optical responses. To make OptoGPT well suited for different types of optical responses, we propose to design for both transmission and reflection spectrum and use multiple techniques to extend the input to diverse responses, including absorption spectrum, reflective and transmissive structural color, and so forth. Examples can be found in Figure 9C.

OptoGPT is a foundation model<sup>112</sup> that once trained it can be used to solve extensive downstream tasks and applications to design multilayer thin film structures. OptoGPT demonstrated a unified inverse design toward a range of applications, including transmissive/reflective filter, high reflection filter, perfect absorber, arbitrary absorber, reflective/transmissive structural color, and so forth. Figure 10 gives some inverse design ex-

amples. Notably, each design task can be completed within 0.1 s. In addition, because the output is based on probability sampling, therefore, OptoGPT can give different designs by running inference multiple times, which solves the diversity challenges. To solve the flexible design challenge, we further proposed "probability resampling" to design structures that can satisfy arbitrary constraints during fabrication or other practical requirements. Using finetuning and "mixed sampling," OptoGPT can even design structures for different incident angles and polarization states, as well as simultaneous design for multiple angles requirements. In this way, OptoGPT solves all previously mentioned challenges in multilayer inverse design, significantly pushing the research boundaries forward.

## DISCUSSION AND OUTLOOK

In conclusion, this review summarized the approaches of different types of inverse design algorithms specifically for



**Figure 10. Examples of different types of optical designs that OptoGPT can design with**

(A) Design for a band-notch filter at 550 nm.

(B) Design for high reflection in near-infrared.

(C) Design for perfect absorber.

(D) Design for arbitrary absorber.

(E–F) Design for reflective/transmissive type structural color. All figures are reprinted from Ma et al.<sup>39</sup> Copyright © 2024, OE Journals.

multilayer thin film structures, starting from traditional optimization-based methods, to the recent popular deep learning based algorithms. Optimization-based methods require iterative simulations of multilayer structures as well as evaluations of the fitness function, which can be time-consuming for complicated structures and multiple designs. On the other hand, deep learning based methods learn a general mapping from some established datasets, therefore, they are much faster than optimization. However, considering that collecting training datasets and training modes could take much time, it would be wise to first estimate the time consumption as compared with the task to be accomplished, and then decide which method to use. In addition, optimization-based methods are more flexible when dealing with new materials from the lab, or some custom multilayer structures with different surrounding mediums and substrates. In both cases, one would need to collect a new dataset and retrain deep learning models.

Compared to other types of photonic structures, multilayer structures are special as they can be represented using three different methods. Each of them has its unique advantages and drawbacks, and corresponds to a specific type of inverse design algorithms. It was only until recently that researchers have identified the sequential representation, a unique data type that is specific to layered structure. This new method has shown promising potential to reshape the inverse design.

Despite the great success deep learning-based methods have achieved, there are still remaining questions to be solved.

Perhaps the biggest question too deep learning is the “black-box problem,” i.e., the predictions from NN cannot be properly explained due to the high-level complexity and nonlinearity inside NN. In inverse design, this explainability is very important to understand why certain structures are preferred during the design process, e.g., if the designed structure is coming from some existing physical behaviors or potentially includes a new physical mechanism that has not been discovered before. However, uncovering the black-box problem has not been extensively explored yet. New ideas and approaches are needed to tackle this issue and reveal hidden physical principles.

Another issue is that many deep learning based methods require a large dataset. Although after training, NN is extremely fast for inverse design, generating the training dataset itself can take a long time. This leads to the dilemma of whether the speed-up in the inverse design process is worthwhile. In addition to the data-based deep learning, currently another type of neural operator-based learning has been popular in solving complicated partial differential equations, such as PINN<sup>113,114</sup> and DeepONet.<sup>115,116</sup> By integrating the physics rules and mathematical equations into NN, these models can directly learn the complicated mapping from input to output without any data. More efforts are needed in order to incorporate these methods into inverse design.

Finally, there are still several issues that require special attention for both methods. For example, it is important to consider the physical conditions for the size of the incoming light beams

during the design process, especially for the applications where the input light spot size is very small, such as functional filters for laser beams. If these conditions are not properly incorporated into design considerations, the multilayer structure may not perform as expected at some higher incidence angle. Currently, these issues have not been extensively explored for both optimization and deep learning-based methods. New methods and techniques are required to solve these challenges. One possible direction to take is to treat the specialized cases as “out-of-distribution” datasets, where we have recently proposed a method to handle some of these situations, based on the pre-trained transformer and incorporating perturbations and genetic optimization.<sup>117</sup>

## ACKNOWLEDGMENTS

We thank the National Science Foundation FET-2309403 for the support of this work.

## AUTHOR CONTRIBUTIONS

T. M. and L. G conceived the review framework. All authors contribute to the writing.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

## REFERENCES

1. T. Babeva, ed. (2021). *Optical Thin Films and Structures* (MDPI - Multidisciplinary Digital Publishing Institute). <https://doi.org/10.3390/books978-3-0365-0893-1>.
2. (2013). Index. In *Optical Thin Films and Coatings* Woodhead Publishing Series in Electronic and Optical Materials, A. Piegari and F. Flory, eds. (Woodhead Publishing), pp. 811–835. <https://doi.org/10.1533/9780857097316.backmatter>.
3. Macleod, H.A. (2018). *Thin-film Optical Filters*, Fifth edition (CRC Press, Taylor & Francis Group).
4. Morales-Masis, M., De Wolf, S., Woods-Robinson, R., Ager, J.W., and Balif, C. (2017). Transparent Electrodes for Efficient Optoelectronics. *Adv. Electron. Mater.* 3, 1600529. <https://doi.org/10.1002/aeml.201600529>.
5. He, Z., Zhong, C., Su, S., Xu, M., Wu, H., and Cao, Y. (2012). Enhanced power-conversion efficiency in polymer solar cells using an inverted device structure. *Nat. Photonics* 6, 591–595. <https://doi.org/10.1038/nphoton.2012.190>.
6. Yang, C., Ji, C., Shen, W., Lee, K.-T., Zhang, Y., Liu, X., and Guo, L.J. (2016). Compact multilayer film structures for ultrabroadband, omnidirectional, and efficient absorption. *Acs Photonics* 3, 590–596.
7. Fink, Y., Winn, J.N., Fan, S., Chen, C., Michel, J., Joannopoulos, J.D., and Thomas, E.L. (1998). A Dielectric Omnidirectional Reflector. *Science* 282, 1679–1682. <https://doi.org/10.1126/science.282.5394.1679>.
8. Rorem, B.A., Cho, T.H., Farjam, N., Lenef, J.D., Barton, K., Dasgupta, N.P., and Guo, L.J. (2022). Integrating Structural Colors with Additive Manufacturing Using Atomic Layer Deposition. *ACS Appl. Mater. Interfaces* 14, 31099–31108. <https://doi.org/10.1021/acsami.2c05940>.
9. Yang, Z., Ji, C., Liu, D., and Guo, L.J. (2019). Enhancing the Purity of Reflective Structural Colors with Ultrathin Bilayer Media as Effective Ideal Absorbers. *Adv. Opt. Mater.* 7, 1900739. <https://doi.org/10.1002/adom.201900739>.
10. Zhang, Q., Li, B., Huang, S., Nomura, H., Tanaka, H., and Adachi, C. (2014). Efficient blue organic light-emitting diodes employing thermally activated delayed fluorescence. *Nat. Photonics* 8, 326–332. <https://doi.org/10.1038/nphoton.2014.12>.
11. Jeong, C., Park, Y.-B., and Guo, L.J. (2021). Tackling light trapping in organic light-emitting diodes by complete elimination of waveguide modes. *Sci. Adv.* 7, eabg0355. <https://doi.org/10.1126/sciadv.abg0355>.
12. Park, Y.-B., Lee, S., Tobah, M., Ma, T., and Guo, L.J. (2023). Optimized optical/electrical/mechanical properties of ultrathin metal films for flexible transparent conductor applications: review [Invited]. *Opt. Mater. Express* 13, 304. <https://doi.org/10.1364/OME.473277>.
13. Raman, A.P., Anoma, M.A., Zhu, L., Rephaeli, E., and Fan, S. (2014). Passive radiative cooling below ambient air temperature under direct sunlight. *Nature* 515, 540–544. <https://doi.org/10.1038/nature13883>.
14. Byrnes, S.J. (2016). Multilayer optical calculations. Preprint at arXiv. <https://doi.org/10.48550/arXiv.1603.02720>.
15. Tikhonravov, A.V., Trubetskov, M.K., and DeBell, G.W. (1996). Application of the needle optimization technique to the design of optical coatings. *Appl. Opt.* 35, 5493–5508.
16. Tikhonravov, A.V., Trubetskov, M.K., and DeBell, G.W. (2007). Optical coating design approaches based on the needle optimization technique. *Appl. Opt.* 46, 704–710. <https://doi.org/10.1364/AO.46.000704>.
17. Rabady, R.I., and Ababneh, A. (2014). Global optimal design of optical multilayer thin-film filters using particle swarm optimization. *Optik* 125, 548–553.
18. Schubert, M.F., Mont, F.W., Chhajed, S., Poxson, D.J., Kim, J.K., and Schubert, E.F. (2008). Design of multilayer antireflection coatings made from co-sputtered and low-refractive-index materials by genetic algorithm. *Opt. Express* 16, 5290–5298.
19. Liu, D., Tan, Y., Khoram, E., and Yu, Z. (2018). Training deep neural networks for the inverse design of nanophotonic structures. *Acs Photonics* 5, 1365–1369.
20. Unni, R., Yao, K., and Zheng, Y. (2020). Deep convolutional mixture density network for inverse design of layered photonic structures. *ACS Photonics* 7, 2703–2712.
21. Unni, R., Yao, K., Han, X., Zhou, M., and Zheng, Y. (2021). A mixture-density-based tandem optimization network for on-demand inverse design of thin-film high reflectors. *Nanophotonics* 10, 4057–4065. <https://doi.org/10.1515/nanoph-2021-0392>.
22. Chen, W., Gao, Y., Li, Y., Yan, Y., Ou, J.-Y., Ma, W., and Zhu, J. (2023). Broadband Solar Metamaterial Absorbers Empowered by Transformer-Based Deep Learning. *Adv. Sci.* 10, 2206718. <https://doi.org/10.1002/advs.202206718>.
23. Wang, H., Zheng, Z., Ji, C., and Jay Guo, L. (2021). Automated multi-layer optical design via deep reinforcement learning. *Mach. Learn. Sci. Technol.* 2, 025013.
24. Xu, X., Sun, C., Li, Y., Zhao, J., Han, J., and Huang, W. (2021). An improved tandem neural network for the inverse design of nanophotonics devices. *Opt. Commun.* 481, 126513. <https://doi.org/10.1016/j.optcom.2020.126513>.
25. Yuan, X., Wang, S., Gu, L., Xie, S., Ma, Q., and Guo, J. (2024). Multi-headed tandem neural network approach for non-uniqueness in inverse design of layered photonic structures. *Opt. Laser. Technol.* 176, 110997. <https://doi.org/10.1016/j.optlastec.2024.110997>.
26. Zandehshahvar, M., Kiarashi, Y., Chen, M., Barton, R., and Adibi, A. (2021). Inverse design of photonic nanostructures using dimensionality reduction: reducing the computational complexity. *Opt. Lett.* 46, 2634–2637. <https://doi.org/10.1364/OL.425627>.
27. Baucour, A., Kim, M., and Shin, J. (2022). Data-driven concurrent nanostructure optimization based on conditional generative adversarial networks. *Nanophotonics* 11, 2865–2873. <https://doi.org/10.1515/nanoph-2022-0005>.
28. Ma, T., Tobah, M., Wang, H., and Guo, L.J. (2022). Benchmarking deep learning-based models on nanophotonic inverse design problems. *OES* 1, 210012. <https://doi.org/10.29026/oes.2022.210012>.

29. Ren, S., Mahendra, A., Khatib, O., Deng, Y., Padilla, W.J., and Malof, J.M. (2022). Inverse deep learning methods and benchmarks for artificial electromagnetic material design. *Nanoscale* **14**, 3958–3969. <https://doi.org/10.1039/D1NR08346E>.
30. Schneider, P.-I., Garcia Santiago, X., Soltwisch, V., Hammerschmidt, M., Burger, S., and Rockstuhl, C. (2019). Benchmarking Five Global Optimization Approaches for Nano-optical Shape Optimization and Parameter Reconstruction. *ACS Photonics* **6**, 2726–2733. <https://doi.org/10.1021/acsp Photonics.9b00706>.
31. Liu, Z., Zhu, D., Raju, L., and Cai, W. (2021). Tackling Photonic Inverse Design with Machine Learning. *Adv. Sci.* **8**, 2002923. <https://doi.org/10.1002/advs.2002923>.
32. Ma, W., Liu, Z., Kudyshev, Z.A., Boltasseva, A., Cai, W., and Liu, Y. (2021). Deep learning for the design of photonic structures. *Nat. Photonics* **15**, 77–90.
33. Wiecha, P.R., Arbouet, A., Girard, C., and Muskens, O.L. (2021). Deep learning in nano-photonics: inverse design and beyond. *Photonics Res.* **9**, B182–B200.
34. Jiang, J., Chen, M., and Fan, J.A. (2020). Deep neural networks for the evaluation and design of photonic devices. *Nat. Rev. Mater.* **6**, 679–700.
35. Hegde, R.S. (2020). Deep learning: a new tool for photonic nanostructure design. *Nanoscale Adv.* **2**, 1007–1023. <https://doi.org/10.1039/C9NA00656G>.
36. Li, Z., Pestourie, R., Lin, Z., Johnson, S.G., and Capasso, F. (2022). Empowering Metasurfaces with Inverse Design: Principles and Applications. *ACS Photonics* **9**, 2178–2192. <https://doi.org/10.1021/acsp Photonics.1c01850>.
37. Park, J., Kim, S., Nam, D.W., Chung, H., Park, C.Y., and Jang, M.S. (2022). Free-form optimization of nanophotonic devices: from classical methods to deep learning. *Nanophotonics* **11**, 1809–1845.
38. Campbell, S.D., Sell, D., Jenkins, R.P., Whiting, E.B., Fan, J.A., and Werner, D.H. (2019). Review of numerical optimization techniques for meta-device design [Invited]. *Opt. Mater. Express* **9**, 1842. <https://doi.org/10.1364/OME.9.001842>.
39. Taigao, M., Wang, H., and Guo, L.J. (2024). OptoGPT: A foundation model for inverse design in optical multilayer thin film structures. *Opto Electron. Adv.* **7**. <https://www.oejournal.org/article/doi/10.29026/oea.2024.240062>.
40. Saha, A., Ma, T., Wang, H., and Guo, L.J. (2023). Environmentally Sustainable and Multifunctional Chrome-like Coatings Having No Chromium Designed with Reinforcement Learning. *ACS Appl. Mater. Interfaces* **15**, 28772–28780. <https://doi.org/10.1021/acsaami.3c02993>.
41. Wang, D., Liu, Z., Wang, H., Li, M., Guo, L.J., and Zhang, C. (2023). Structural color generation: from layered thin films to optical metasurfaces. *Nanophotonics* **12**, 1019–1081. <https://doi.org/10.1515/nanoph-2022-0063>.
42. Feng, W.-J., Ji, C., and Guo, L.J. (2023). Primary and Secondary Reflective Color Realized by Full-Solution-Processed Multi-Layer Structures. *Adv. Opt. Mater.* **11**, 2300456. <https://doi.org/10.1002/adom.202300456>.
43. Wang, S. (1974). Principles of distributed feedback and distributed Bragg-reflector lasers. *IEEE J. Quantum Electron.* **10**, 413–427. <https://doi.org/10.1109/JQE.1974.1068152>.
44. Zhang, C., ElAfandy, R., and Han, J. (2019). Distributed Bragg Reflectors for GaN-Based Vertical-Cavity Surface-Emitting Lasers. *Appl. Sci.* **9**, 1593. <https://doi.org/10.3390/app9081593>.
45. Kimura, M., Okahara, K., and Miyamoto, T. (1979). Tunable multilayer-film distributed-Bragg-reflector filter. *J. Appl. Phys.* **50**, 1222–1225. <https://doi.org/10.1063/1.326154>.
46. Shi, Y., Li, W., Raman, A., and Fan, S. (2018). Optimization of Multilayer Optical Films with a Memetic Algorithm and Mixed Integer Programming. *ACS Photonics* **5**, 684–691. <https://doi.org/10.1021/acsp Photonics.7b01136>.
47. Schmidt, R.M. (2019). Recurrent Neural Networks (RNNs): A gentle Introduction and Overview. Preprint at arXiv. <https://doi.org/10.48550/arXiv.1912.05911>.
48. Staudemeyer, R.C., and Morris, E.R. (2019). Understanding LSTM – a tutorial into Long Short-Term Memory Recurrent Neural Networks. Preprint at arXiv. <https://doi.org/10.48550/arXiv.1909.09586>.
49. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., and Polosukhin, I. (2017). Attention is all you need. In *Advances in neural information processing systems*, p. 30.
50. Ma, T., Wang, H., and Guo, L.J. (2023). OL-transformer: A Fast and Universal Surrogate Simulator for Optical Multilayer Thin Film Structures (ICML SynS & ML ICML Workshop).
51. Dai, P., Sun, K., Yan, X., Muskens, O.L., de Groot, C.H.K., Zhu, X., Hu, Y., Duan, H., and Huang, R. (2022). Inverse design of structural color: finding multiple solutions via conditional generative adversarial networks. *Nanophotonics* **11**, 3057–3069. <https://doi.org/10.1515/nanoph-2022-0095>.
52. Jiang, J., and Fan, J.A. (2020). Multiobjective and categorical global optimization of photonic structures based on ResNet generative neural networks. *Nanophotonics* **10**, 361–369.
53. Wang, H., and Guo, L.J. (2022). NEUTRON: Neural Particle Swarm Optimization for Material-Aware Inverse Design of Structural Color. *iScience* **25**, 104339. <https://doi.org/10.1016/j.isci.2022.104339>.
54. Siabi-Garjan, A., and Hassanzadeh, R. (2018). A computational approach for engineering optical properties of multilayer thin films: Particle swarm optimization applied to Bruggeman homogenization formalism. *Eur. Phys. J. Plus* **133**, 419. <https://doi.org/10.1140/epjp/i2018-12203-y>.
55. Ruan, Z.-H., Yuan, Y., Zhang, X.-X., Shuai, Y., and Tan, H.-P. (2016). Determination of optical properties and thickness of optical thin film using stochastic particle swarm optimization. *Sol. Energy* **127**, 147–158. <https://doi.org/10.1016/j.solener.2016.01.027>.
56. Jin, C., and Yang, Y. (2021). Transmissive nonlocal multilayer thin film optical filter for image differentiation. *Nanophotonics* **10**, 3519–3525. <https://doi.org/10.1515/nanoph-2021-0313>.
57. Abdullah, T.A., Md Imran, H., Chowdhury, S.I., and Haque, M.A. (2024). A Genetic Algorithm Approach for A Broadband Reflector Design. In *2024 IEEE 14th Annual Computing and Communication Workshop and Conference (CCWC)*, pp. 597–602. <https://doi.org/10.1109/CCWC60891.2024.10427648>.
58. Li, D.G., and Watson, A.C. (1999). Genetic algorithms in optical thin film optimisation design. In *Proceedings Third International Conference on Computational Intelligence and Multimedia Applications (ICCIMA'99)*, pp. 86–90. <https://doi.org/10.1109/ICCIMA.1999.798507>.
59. Krizhevsky, A., Sutskever, I., and Hinton, G.E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In *Advances in Neural Information Processing Systems* (Curran Associates, Inc). [https://proceedings.neurips.cc/paper\\_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf).
60. Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255. <https://doi.org/10.1109/CVPR.2009.5206848>.
61. He, K., Zhang, X., Ren, S., and Sun, J. (2015). Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. In *2015 IEEE International Conference on Computer Vision (ICCV)* (Santiago, Chile), pp. 1026–1034. <https://doi.org/10.1109/ICCV.2015.123>.
62. Li, J. (2022). Recent Advances in End-to-End Automatic Speech Recognition. Preprint at arXiv. <https://doi.org/10.48550/arXiv.2111.01690>.
63. Gaikwad, K., Gawali, W., and Yannawar, P. (2010). Article: A Review on Speech Recognition Technique. *Int. J. Comp. Appl.* **10**, 16–24.
64. Gulati, A., Qin, J., Chiu, C.-C., Parmar, N., Zhang, Y., Yu, J., Han, W., Wang, S., Zhang, Z., Wu, Y., et al. (2020). Conformer: Convolution-augmented Transformer for Speech Recognition. Preprint at arXiv. <https://doi.org/10.48550/arXiv.2005.08100>.



65. Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., et al. (2015). Human-level control through deep reinforcement learning. *Nature* 518, 529–533.
66. Sutton, R.S., and Barto, A.G. (2018). *Reinforcement Learning: An Introduction* (MIT Press).
67. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature* 529, 484–489. <https://www.nature.com/articles/nature16961>.
68. Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Preprint at arXiv. <https://doi.org/10.18653/V1/N19-1423>.
69. Brown, T.B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., et al. (2020). Language Models are Few-Shot Learners. Preprint at arXiv. <https://doi.org/10.48550/arXiv.2005.14165>.
70. Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., Stoyanov, V., and Zettlemoyer, L. (2020). BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (Association for Computational Linguistics)*, pp. 7871–7880. <https://doi.org/10.18653/v1/2020.acl-main.703>.
71. Dai, P., Wang, Y., Hu, Y., de Groot, C.H., Muskens, O., Duan, H., and Huang, R. (2021). Accurate inverse design of Fabry–Perot-cavity-based color filters far beyond sRGB via a bidirectional artificial neural network. *Photon. Res.* 9, B236. <https://doi.org/10.1364/PRJ.415141>.
72. Kim, S., Shin, J.M., Lee, J., Park, C., Lee, S., Park, J., Seo, D., Park, S., Park, C.Y., and Jang, M.S. (2021). Inverse design of organic light-emitting diode structure based on deep neural networks. *Nanophotonics* 10, 4533–4541. <https://doi.org/10.1515/nanoph-2021-0434>.
73. Gao, L., Li, X., Liu, D., Wang, L., and Yu, Z. (2019). A bidirectional deep neural network for accurate silicon color design. *Adv. Mater.* 31, 1905467.
74. Ma, W., Cheng, F., and Liu, Y. (2018). Deep-learning-enabled on-demand design of chiral metamaterials. *ACS Nano* 12, 6326–6334.
75. He, J., He, C., Zheng, C., Wang, Q., and Ye, J. (2019). Plasmonic nanoparticle simulations and inverse design using machine learning. *Nanoscale* 11, 17444–17459. <https://doi.org/10.1039/C9NR03450A>.
76. Xiong, J., Shen, J., Gao, Y., Chen, Y., Ou, J.-Y., Liu, Q.H., and Zhu, J. (2023). Real-Time On-Demand Design of Circuit-Analog Plasmonic Stack Metamaterials by Divide-and-Conquer Deep Learning. *Laser Photon. Rev.* 17, 2100738. <https://doi.org/10.1002/lpor.202100738>.
77. Eckle, K., and Schmidt-Hieber, J. (2019). A comparison of deep networks with ReLU activation function and linear spline-type methods. *Neural Netw.* 110, 232–242. <https://doi.org/10.1016/j.neunet.2018.11.005>.
78. Gao, Y., Chen, W., Li, F., Zhuang, M., Yan, Y., Wang, J., Wang, X., Dong, Z., Ma, W., and Zhu, J. (2024). Meta-Attention Deep Learning for Smart Development of Metasurface Sensors. *Adv. Sci.* 11, 2405750. <https://doi.org/10.1002/advs.202405750>.
79. Kingma, D.P., and Welling, M. (2022). Auto-Encoding Variational Bayes. Preprint at arXiv. <https://doi.org/10.48550/arXiv.1312.6114>.
80. Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2014). Generative Adversarial Networks. Preprint at arXiv. <https://doi.org/10.48550/arXiv.1406.2661>.
81. Bishop, C.M. (1994). Mixture density networks. <https://publications.aston.ac.uk/id/eprint/373/>.
82. Deng, Y., Dong, J., Ren, S., Khatib, O., Soltani, M., Tarokh, V., Padilla, W.J., and Malof, J.M. Benchmarking Data-Driven Surrogate Simulators for Artificial Electromagnetic Materials. 12.
83. Ren, S., Padilla, W., and Malof, J. (2021). Benchmarking deep inverse models over time, and the neural-adjoint method. Preprint at arXiv. <https://doi.org/10.48550/arXiv.2009.12919>.
84. Hemmatyar, O., Abdollahramezani, S., Kiarashinejad, Y., Zandehshahvar, M., and Adibi, A. (2019). Full color generation with Fano-type resonant HfO<sub>2</sub> nanopillars designed by a deep-learning approach. *Nanoscale* 11, 21266–21274. <https://doi.org/10.1039/C9NR07408B>.
85. Kiarashinejad, Y., Abdollahramezani, S., Zandehshahvar, M., Hemmatyar, O., and Adibi, A. (2019). Deep Learning Reveals Underlying Physics of Light–Matter Interactions in Nanophotonic Devices. *Adv. Theory Simul.* 2, 1900088. <https://doi.org/10.1002/adts.201900088>.
86. Kiarashinejad, Y., Abdollahramezani, S., and Adibi, A. (2020). Deep learning approach based on dimensionality reduction for designing electromagnetic nanostructures. *NPJ Comput. Mater.* 6, 12. <https://doi.org/10.1038/s41524-020-0276-y>.
87. Qu, Y., Jing, L., Shen, Y., Qiu, M., and Soljačić, M. (2019). Migrating Knowledge between Physical Scenarios Based on Artificial Neural Networks. *ACS Photonics* 6, 1168–1174. <https://doi.org/10.1021/acsp Photonics.8b01526>.
88. Ma, W., Chen, W., Li, D., Liu, Y., Yin, J., Tu, C., Xia, Y., Shen, G., Zhou, P., Deng, L., and Zhang, L. (2023). Deep learning empowering design for selective solar absorber. *Nanophotonics* 12, 3589–3601. <https://doi.org/10.1515/nanoph-2023-0291>.
89. Design of multilayer optical thin-films based on light scattering properties and using deep neural networks <https://opg.optica.org/oe/fulltext.cfm?uri=oe-29-20-32627&id=459831>.
90. Jiang, J., and Fan, J.A. (2019). Global optimization of dielectric metasurfaces using a physics-driven neural network. *Nano Lett.* 19, 5366–5372.
91. Jiang, J., and Fan, J.A. (2019). Simulator-based training of generative models for the inverse design of metasurfaces. *Nanophotonics* 9, 1059–1069. <https://doi.org/10.1515/nanoph-2019-0330>.
92. Zhang, X., Bai, B., Sun, H., Jin, G., and Valentine, J. (2022). Incoherent Optoelectronic Differentiation Based on Optimized Multilayer Films. *Laser Photon. Rev.* 16, 2200038. <https://doi.org/10.1002/lpor.202200038>.
93. Latif, S., Zaidi, A., Cuayahuitl, H., Shamshad, F., Shoukat, M., and Qadir, J. (2023). Transformers in Speech Processing: A Survey. Preprint at arXiv. <https://doi.org/10.48550/arXiv.2303.11607>.
94. Wen, Q., Zhou, T., Zhang, C., Chen, W., Ma, Z., Yan, J., and Sun, L. (2023). Transformers in Time Series: A Survey. Preprint at arXiv. <https://doi.org/10.48550/arXiv.2202.07125>.
95. Yaqoob, A., Bi, T., and Muntean, G.-M. (2020). A Survey on Adaptive 360° Video Streaming: Solutions, Challenges and Opportunities. *IEEE Commun. Surv. Tutorials* 22, 2801–2838. <https://doi.org/10.1109/COMST.2020.3006999>.
96. Brunke, L., Greeff, M., Hall, A.W., Yuan, Z., Zhou, S., Panerati, J., and Schoellig, A.P. (2022). Safe Learning in Robotics: From Learning-Based Control to Safe Reinforcement Learning. *Annu. Rev. Control Robot. Auton. Syst.* 5, 411–444. <https://doi.org/10.1146/annurev-control-042920-020211>.
97. Spielberg, S., Tulsyan, A., Lawrence, N.P., Loewen, P.D., and Bhushan Gopaluni, R. (2019). Toward self-driving processes: A deep reinforcement learning approach to control. *AIChE J.* 65, e16689. <https://doi.org/10.1002/aic.16689>.
98. Liang, X., Wang, T., Yang, L., and Xing, E. (2018). CIRL: Controllable Imitative Reinforcement Learning for Vision-Based Self-Driving. In *Computer Vision – ECCV 2018. Lecture Notes in Computer Science*, vol. 11211, V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, eds., (Springer, Cham). [https://doi.org/10.1007/978-3-030-01234-2\\_36](https://doi.org/10.1007/978-3-030-01234-2_36).
99. Schulman, J., Wolski, F., Dhariwal, P., Radford, A., and Klimov, O. (2017). Proximal Policy Optimization Algorithms. Preprint at arXiv. <https://doi.org/10.48550/arXiv.1707.06347>.
100. Rosenfeld, R. (2000). Two decades of statistical language modeling: where do we go from here? *Proc. IEEE* 88, 1270–1278. <https://doi.org/10.1109/5.880083>.

101. Bahdanau, D., Cho, K., and Bengio, Y. (2016). Neural Machine Translation by Jointly Learning to Align and Translate. Preprint at arXiv. <https://doi.org/10.48550/arXiv.1409.0473>.
102. Izacard, G., and Grave, E. (2021). Leveraging Passage Retrieval with Generative Models for Open Domain Question Answering. Preprint at arXiv. <https://doi.org/10.48550/arXiv.2007.01282>.
103. Radford, A., Narasimhan, K., Salimans, T., and Sutskever, I. (2018). Improving Language Understanding by Generative Pre-training, p. 12.
104. Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., and Sutskever, I. (2019). Language Models Are Unsupervised Multitask Learners, p. 24.
105. Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C.L., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., et al. (2022). Training language models to follow instructions with human feedback. Preprint at arXiv. <https://doi.org/10.48550/arXiv.2203.02155>.
106. Ray, P.P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. Internet Things Cyber-Phys. Syst. 3, 121–154. <https://doi.org/10.1016/j.iotcps.2023.04.003>.
107. Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., et al. (2023). LLaMA: Open and Efficient Foundation Language Models. Preprint at arXiv. <https://doi.org/10.48550/arXiv.2302.13971>.
108. Benegas, G., Batra, S.S., and Song, Y.S. (2023). DNA language models are powerful predictors of genome-wide variant effects. Proc. Natl. Acad. Sci. USA 120, e2311219120. <https://doi.org/10.1073/pnas.2311219120>.
109. Madani, A., Krause, B., Greene, E.R., Subramanian, S., Mohr, B.P., Holton, J.M., Olmos, J.L., Xiong, C., Sun, Z.Z., Socher, R., et al. (2023). Large language models generate functional protein sequences across diverse families. Nat. Biotechnol. 41, 1099–1106. <https://doi.org/10.1038/s41587-022-01618-2>.
110. Liu, P., Ren, Y., Tao, J., and Ren, Z. (2024). GIT-Mol: A Multi-modal Large Language Model for Molecular Science with Graph, Image, and Text. Comput. Biol. Med. 171, 108073. <https://doi.org/10.1016/j.compbiomed.2024.108073>.
111. (2022). Language models can learn complex molecular distributions. Nat. Commun. 13, 3293. <https://www.nature.com/articles/s41467-022-30839-x>.
112. Bommasani, R., Hudson, D.A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M.S., Bohg, J., Bosselut, A., Brunskill, E., et al. (2022). On the Opportunities and Risks of Foundation Models. Preprint at arXiv. <https://doi.org/10.48550/arXiv.2108.07258>.
113. Hao, Z., Liu, S., Zhang, Y., Ying, C., Feng, Y., Su, H., and Zhu, J. (2023). Physics-Informed Machine Learning: A Survey on Problems, Methods and Applications. Preprint at arXiv. <https://doi.org/10.48550/arXiv.2211.08064>.
114. Mao, Z., Jagtap, A.D., and Karniadakis, G.E. (2020). Physics-informed neural networks for high-speed flows. Comput. Methods Appl. Mech. Eng. 360, 112789.
115. Lu, L., Jin, P., Pang, G., Zhang, Z., and Karniadakis, G.E. (2021). Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. Nat. Mach. Intell. 3, 218–229. <https://doi.org/10.1038/s42256-021-00302-5>.
116. Wang, S., Wang, H., and Perdikaris, P. (2021). Learning the solution operator of parametric partial differential equations with physics-informed DeepONets. Sci. Adv. 7, eabi8605. <https://doi.org/10.1126/sciadv.abi8605>.
117. Ma, M., Ma, T., and Guo, L.J. (2024). Solving Out-Of-Distribution Challenges in Optical Foundation Models Using Self-Improving Data Augmentation.