Deep learning modeling strategy for material science: from natural materials to metamaterials

Wenwen Li¹, Pu Chen¹, Bo Xiong¹, Guandong Liu², Shuliang Dou¹, Yaohui Zhan¹, Zhiyuan Zhu³, Tao Chu¹, Yao Li⁴, * and Wei Ma¹, ³, *

¹ State Key Laboratory of Modern Optical Instrumentation, College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, People's Republic of China
² Chongqing Key Laboratory of Nonlinear Circuits and Intelligent Information Processing, Southwest University, Chongqing 400715, People's Republic of China
³ Intelligent Network Research Institute, Zhejiang Lab, Hangzhou 311100, People's Republic of China
⁴ Center for Composite Materials and Structure, Harbin Institute of Technology, Harbin 150001, People's Republic of China
⁵ School of Optoelectronic Science and Engineering & Key Laboratory of Advanced Optical Manufacturing Technologies of Jiangsu Province, Soochow University, Suzhou 215006, People's Republic of China
* Authors to whom any correspondence should be addressed.
E-mail: ma_wei@zju.edu.cn and yaoli@hit.edu.cn

Keywords: materials, deep learning, metamaterials, modeling, inverse design, optimization

Abstract
Computational modeling is a crucial approach in material-related research for discovering new materials with superior properties. However, the high design flexibility in materials, especially in the realm of metamaterials where the sub-wavelength structure provides an additional degree of freedom in design, poses a formidable computational cost in various real-world applications. With the advent of big data, deep learning (DL) brings revolutionary breakthroughs in many conventional machine learning and pattern recognition tasks such as image classification. The accompanied data-driven modeling paradigm also provides transformative methodology shift in materials science, from trial-and-error routine to intelligent material discovery and analysis. This review systematically summarize the application of DL in material science, based on a model selection perspective for both natural materials and metamaterials. The review aims to uncover the logic behind data-model relation with emphasis on suitable data structures for different scenarios in the material study and the corresponding problem-solving DL model architectures.

1. Introduction
Materials are regarded as milestones in the development and evolution of human society. Promoting the development of materials is a fundamental driving force of human society. Starting from the Stone Age, mankind discovered ore, and metallurgical technology continued to develop, which brought the age of bronze, iron, and steel. Since then, the advancement of science and technology has promoted the emergence and development of new materials. Some new materials such as artificial polymer composites, nanomaterials, and new energy materials have been applied on a large scale. The development history of the material shows the degree of freedom of material design is continuously increasing.

Traditional material research and development mainly rely on the intuition of scientists and the accumulation of experimental experience, and a large number of repetitive experiments are required to explore and verify the feasibility of the design scheme. The earliest material research and development model is mainly based on trial and error methods, similar to Edison's discovery of tungsten as filament for lamps. It is not feasible to rely solely on manpower to analyze and dig out the deep connections between large amounts of data. Although theory-driven methods were developed later, such as density functional theory (DFT), due to the complexity of material synthesis, various defects that may occur in preparation and metastability [1], computer theory and simulation are required to be closely integrated with material experiments and data [2], which makes high flux DFT in the field of materials [3]. To better systematize
materials, the United States took the lead in proposing the ‘Material Genome Project’ [4], analogous to the Human Biological Genome Project [5]. The key idea is to combine high-throughput calculation, high-throughput preparation, high-throughput detection and database systems of materials to form the basic conditions and capabilities for the innovative development of new materials.

In the 21st century, through the orderly design of the key physical dimensions of the material, people have developed a special class of man-made material that has characteristics not found in nature, namely, metamaterials. This is mainly due to their different structural arrays rather than composition, for example, by adjusting the size and arrangement of the micro-structure to achieve some unique electromagnetic properties. Metamaterials began with negative refractive index materials [6], and then the research on electromagnetic metamaterials has grown rapidly. Due to the special electromagnetic properties, such as the inverted Doppler effect [7], generalized Snell’s law [8], negative permittivity and permeability [9], it is widely used in invisible cloak [10, 11], superlens [12], and various flat optical components [13]. However, these exotic properties are achieved by introducing sub-wavelength structures in addition to material choice, resulting in an even higher degree of design freedom, which drives the generation of more advanced design algorithms.

Machine learning is the study of computer programs that mimic the learning process of humans to solve complex problems without being explicitly programmed. With the advantage of discovering the mapping relationship between experimental parameters and material properties at a low computational cost, machine learning has been applied to quickly evaluate massive material combinations, create new materials on demand, and predict material properties [14]. Several typical studies include using machine learning algorithms to improve the quality of material imaging data and extract material information from it [15], or using the predictability of machine learning to analyze the sintering density [16], atomization energies [17], band gaps [18], phase transitions [19], and crystallinity [20] of materials.

In particular, the advent of deep learning (DL) has brought revolutionary improvements in many conventional machine learning and pattern recognition tasks. In material science, it is used to predict and evaluate the fundamental properties of materials, such as wave functions [21, 22], formation energy [23, 24], and electronic ground and excited states [25], read in the chemical literature [26, 27], crystal structure prediction [28] and micro in the task of structural classification [29]. In metamaterial design tasks, DL can solve complex metamaterial design, help discover the complex and non-intuitive relationship between metamaterial structure and its optical response [30], realize on-demand inverse design [31], performance prediction [32], and improve design speed [33, 34] etc.

Compared with traditional numerical methods, such as DFT calculations for natural materials and finite difference time domain simulation for metamaterials, DL and other data-driven methods automatically learn the multilevel abstract of the underlying physical problems with the aid of the hierarchical model structure. After being trained with enough training data in advance, DL can achieve an order of magnitude faster speed than traditional numerical calculations to produce accurate results. Most contemporary DL methods adopt the form of deep neural network (DNN), so these two terminologies are often used indistinguishably in most literature. DL models are composed of serial stackings of linear processing units followed by non-linear activation layers to model highly complex data relationships. Using back-propagation, the model weights can be modified to minimize the prescribed loss function [35]. Detailed DL model structure and training process are illustrated in Box 1.

In DNNs, since the input training data and the initial value of the weight of the model are normalized, the choice of learning rate is often empirical. Generally speaking, one can choose a relatively large learning rate, such as 0.01, to accelerate the initial training speed at beginning, and gradually reduce the learning rate until the model converges [36]. In a DL network, methods such as learning rate annealing method [37], Hessian-free optimization method [38], the cyclic learning rate and the learning rate range test can be used to update or evaluate the learning rate [39].

**Box 1. DL and back-propagation algorithm.**

The basic unit of DL is an artificial neuron, whose working principle is to perform two mathematical operations, a weighted sum, and then a non-linear mapping. The weighted sum is expressed as $z = \sum w_i x_i + b$, where $w_i$ is a trainable weight vector that possesses the same dimension as the input $x_i$ and $b$ is a trainable bias term. A continuous, differentiable, non-linear function is applied to $z_i$, called the activation function $\sigma$, and the output of the neuron is expressed as $Y = \sigma(z)$. As shown in the illustration.
Neurons are connected in different ways to form a DNN, and the basic model is called a multilayer perceptron (MLP), or fully connected neural network. It consists of an input layer, an output layer, and multiple hidden layers. Each neuron in a given layer is connected to every neuron in the next layer, but there are no inter-layer connections. The neuron uses the weight $w$ and trainable bias term $b$ to process the input data $x$ layer by layer. In this way, a neural network is equivalent to a vector function characterized by a large number of parameters $w$ and $b$.

For example, when the neuron $i$ in the $(n-1)$th layer is connected to the neuron $j$ in the $n$th layer, the weight coefficient between them is denoted as $w_{ji}^n$, the bias parameter as $b_j^n$.

The $n$th layer is the output of the $j$th neuron is expressed as $a_j^n$, which can be understood as a weighted summation of the output $z_j^n$ of the previous layer and express the activation relationship between them through the nonlinear mathematical function $\sigma(\cdot)$, as follows:

$$a_j^n = \sigma \left( \sum_i w_{ji}^n a_i^{n-1} + b_j^n \right) = \sigma \left( z_j^n \right).$$  \hspace{1cm} (1)

The entire $n$th layer of the neural network will output a weight matrix $w^n$ and bias parameter $b^n$. The non-linear mathematical function $\sigma(\cdot)$ is called the activation function, including sigmoid, hyperbolic tangent, rectified linear unit, etc.

To evaluate the quality of the output from a neural network, a cost function $C(\cdot)$ can be used to measure the error between the network output and the target output. $C(\cdot)$ takes the given output of the last layer and the ground truth label as input values, and returns an output that minimizes the error. Such an error minimizing process is called training, which is accomplished by the back-propagation algorithm.

Back-propagation can be thought of as using the chain rule to forward the loss function $C(\cdot)$ of the output layer, calculating the influence of the weight on the loss function, and then combining the gradient descent method to update the weight. Here, an intermediate error vector, $\delta^n = \left( \frac{\partial C}{\partial Z_1^n}, \frac{\partial C}{\partial Z_2^n}, \frac{\partial C}{\partial Z_3^n}, \ldots \right)$, is introduced to represents the partial derivative of the cost $C$ with respect to the weighted input of the $n$th layer. In the last layer, denoted as the $n$th layer, there are:

$$\delta^n = (\nabla_{a^n} C) \odot \sigma'(z^n).$$  \hspace{1cm} (2)

Among them, $\odot$ is the Hadamard operator, the cost function $C(a^n)$ and the activation function $\sigma(z^n)$ have analytical forms. Applying the chain rule of partial derivatives and using the error $\delta^{n+1}$ of the $n+1$ layer to calculate the error vector $\delta^n$ of the $n$ layer, there are:

$$\delta^n = \left( (w^{n+1})^T \delta^{n+1} \right) \odot \sigma'(z^n).$$  \hspace{1cm} (3)

Equations (2) and (3) illustrate that the cost function representing the error between the network output value and the expected output in the back-propagation process returns to the initial layer from the last layer, and accumulates errors in this process. The cost $C(\cdot)$ is related to the weight coefficient $w_{ji}^n$ and bias parameter $b_j^n$ as follows:

$$\frac{\partial C}{\partial w_{ji}^n} = a_i^{n-1} \times \delta_j^n$$  \hspace{1cm} (4)

$$\frac{\partial C}{\partial b_j^n} = \delta_j^n.$$  \hspace{1cm} (5)
Later, the model can be trained by stochastic gradient descent. After inputting a batch of data to the network, the output and the error can be calculated by the forward propagation of the network, then the average gradient of the batch of data is calculated using the back propagation of the network to adjust the update weight and bias accordingly. A hyperparameter $\eta$ is introduced here to control the cost of the parameter being modified as part of its gradient. The following formula can be used to adjust the weight and bias [40]:

$$
    w^n \rightarrow w^n - \frac{\eta}{M} \sum_x \delta^{x,n} (a^{x,n-1})^T 
$$

$$
    b^n \rightarrow b^n - \frac{\eta}{M} \sum_x \delta^{x,n} .
$$

In the DL model, there are three types of data sets. One is the training set, which is used to iteratively adjust the weights of neurons in the process of training the neural network until the data distribution in the training set is correctly captured. Another is the validation set, which is a set of samples used to verify the performance of the model. In a conventional training process of DL model with moderate data scale, the training set usually takes up to about 80% of the total data, while the validation and test set each account for about 10% [14, 41]. When there is little data available, some advanced methods are often used, including set aside squares, K-fold cross-validation, etc [42]. The training is iterated several times until the cost value of the validation set stops decreasing, and then the training is terminated to avoid overfitting. Finally, the test set is used to evaluate the generalization ability of the model.

The key to applying DL to specific tasks in material science is the data structure and model architecture, which are visually displayed in figure 1. According to the design parameters and target properties in natural materials and metamaterials, the commonly used data structures include vector, image, graphs, and time series.

Vector data are mainly 1D discrete vectors, which are expressed in natural materials as lattice constant, the dielectric constant and permeability of the material and band gap, etc. In metamaterials, vectors are commonly used to describe nanostructure parameters, device efficiency, quality factor, optical band gap, spectral response sampled at discrete points, etc. A deep belief network can be used to analyze the potential...
parameters of the material process design process [43]. To analyze the mapping relationship between the geometric parameters of the metamaterial and the optical response, a two-way fully connected network [44] and a fully connected network model combined with an autoencoder (AE) has been used [45].

Image data is divided into 2D images and 3D images, which are not convenient to be parameterized by several discrete variables. To process image data types, a set of connected convolutional layers are often used to extract and process spatial features. This convolutional layer model is called a convolutional neural network (CNN). In the field of natural materials, the CNN model can be used to quickly analyze image data, find material defects [46], and simulate the x-ray diffraction pattern of natural materials to identify the phase and composition of the material [47]. In addition, for the use of automatic encoders to reduce the feature dimension, the combination of CNN and fully connected networks can realize the automated design of metamaterials [34] and predict the optical response of image structures [48].

Time series is generally the data of the discretization of continuous electromagnetic phenomena. The output in a given time not only depends on the input signal at the time, but also depends on the state (its internal electric field) in the previous time step. Since recurrent neural networks (RNN) return the network output to the input layer and maintain a memory that explains the past state of the system, this makes them ideal for modeling time series systems. But simply combining the current input with the previous hidden state is not suitable for sequence data with long-term dependence [49], so long-term short-term memory (LSTM) units [50] and gated recurrent units (GRU) [51] have appeared. In predicting periodic elastic-plastic materials, temporal convolutional network (TCN) is more suitable for predicting irreversible, historical, and time-related phenomena [52]. For situations with a large number of cyclic parameters, stackable cyclic cells (STAR) can maintain a stable gradient [53].

For a composite material or an electromagnetic system that is composed of physically interacting discrete objects, a graph is an ideal data structure representation. Graph neural network (GNN) model is usually used for processing. includes graph attention networks [54], graph recursive networks [55], and graph generation networks [56]. When it is necessary to quantify interacting particles to predict material properties, the GNN model can be used to achieve this ambition [57].

Since the accuracy of DL models will significantly affect their potency in material science, a natural way to improve model accuracy is to use more training data. However, it should be noted that suitable model architecture must be adopted to guarantee that the characteristics of the material data are fully exploited [58]. By combining with other traditional optimization algorithms, DL can also benefit for better accuracy [59].

The following parts of this review are deployed from the perspective of model selection, with a focus on appropriate DL model architectures in materials science for the discovery of new materials and the design of metamaterials. This paper firstly introduce several basic discriminative model architectures, including MLPs, CNNs, RNNs, and GNNs, which has been used to build accurate mappings between the design parameters of natural material or metamaterials and their physical or chemical properties, each with a unique advantage for specific data structures. On this basis, some advanced models are described, such as generative models that solve the one-to-many mapping problem in material design and hybrid models that combine DL models with traditional optimization algorithms. Following that, some challenges of DL in material applications are discussed, and corresponding solutions are proposed for model interpretability and excessive data burden. Finally, the entire review is summarized with outlooks.

2. Basic model architectures

2.1. Multilayer perceptrons

The initial DL network has a simple architecture with the basic unit called artificial neurons. The simplest structure in DL is a network composed of multiple neurons, called a perceptron. Each neuron in one layer is connected with all neurons in the next layer with its own unique weight to form a MLP. The structure is shown in Box 1. This MLP model belongs to the simplest form of feedforward neural network. In theory, it has been proved to be a universal approximator that can fit any continuous function with a finite number of neurons [60].

MLP provides acceleration and approximate simulation for traditional computational physics and inverse design, which is equivalent to an approximator. In 2018, Peurifoy et al [61] applied the MLP model to multi-layered nanoparticle light scattering, and found that the network sampling and training a small part of the data can achieve high-precision simulation, meanwhile, the simulation speed is faster than traditional numerical simulation by one order of magnitude. Combined with back-propagation, it can solve the inverse design problem of multi-parameter and gradient analysis. The MLP model can also realize the theoretical prediction and accurate simulation of the elastic properties of glass materials, and the simulation results have good correlation with the experimental results, as shown in figure 2(a) [62]. This can be used in photonic
Figure 2. MLP model architecture and application. (a) The left graph is the network architecture of the MLP model used to simulate glass materials, and the right two graphs are the linear relationship between predicted (LANN) and experimental longitudinal (Lexp) modulus values for TeO$_2$–B$_2$O$_3$–ZnO glasses doped with Eu$_2$O$_3$ NPs and Eu$_2$O$_3$ content $[62]$. (b) The picture on the left is the general schematic of the artificial neural network applied to garnet and perovskite materials, and the picture on the right is the comparison of $E_{\text{ANN}}$ and $E_{\text{DFT}}$ of the optimized ANN model on unmixed garnet, and MAEs in $E_f$ of unmixed and mixed samples in training, validation, and test sets of all garnet models $[66]$. (c) Structure of the deep-learning model for designing chiral metamaterials $[44]$. (d) The left is a series network architecture consisting of an inverse design network connected to the forward modeling network, and the right is the learning curve and test results of the series neural network $[77]$. 

Crystals to quickly evaluate the gradient of the Q factor with respect to the pore displacement $[33]$. It performs better than the spline model in evaluating hydrocarbons, with lower error $[63]$. MLP can construct the potential-energy surface of the material faster and more accurately $[64, 65]$, without the need to specify the bond or atom type. In addition, it can predict the formation energy of the material $[66]$, the band gap and melting point of the ternary compound semiconductor $[67]$, discover the topological phase transition $[68]$, describe the topological state of the material $[69]$, distinguishing the Chern insulator and the fractional Chern insulator from trivial insulators $[70]$, obtain the quantum phase boundary between the topological and trivial phases $[71]$. Ye et al $[66]$ used Pauling’s electronegativity and ionic radius as input descriptors to successfully predict the formation energy of garnet and perovskite with low average absolute errors (MAEs). The model structure and prediction results are shown in figure 2(b). This model greatly expands the space for discovering materials with potentially superior properties.

The MLP model has also been widely used in metamaterials and has become one of the most widely used DL models. A well-trained MLP model can predict and understand the effects of different structural parameters of metamaterials on the near-field radiation $[72]$ or find the optimal geometric parameters to achieve specific radiation $[73]$. When MLP is combined with an AE, it can better reduce the dimension of the design space and reduce the amount of calculation $[45]$; the model combined with the convolutional layer...
can model complex dielectric metamaterials. This method is better than traditional electromagnetic simulation software by five orders of magnitude faster [74]. Since the traditional numerical simulation methods cannot realize automatic structural design under the guidance of goals, once the design goals are changed, electromagnetic simulations are needed to be repeated, which is cumbersome and inflexible. The MLP model makes it possible to automate the design of metamaterials. For instance, the MLP can directly calculate and automatically generate the metamaterial structure under a given design goal, making the design process more efficient, faster and simpler [34].

The MLP model has developed rapidly with its model architecture and training strategies continuously improved. As shallow neural networks, evolutionary algorithms and linear regression are still limited in accuracy and practical feasibility, in order to solve the highly complex nonlinear modeling problem of physical processes, a bidirectional neural network has been developed. Typically, such as a DL model composed of a two-way neural network assembled by the partially-stacking strategies [44], as shown in figure 2(c), this model can be achieved on-demand design of 3D chiral metamaterials, and realized forward prediction more accurately and efficiently, thus having greater application prospects in the fields of nano-optics and material science. The bidirectional DNN structure including geometric prediction network and spectral prediction network designed by Malkiel et al can predict the geometric shape of complex nanostructures with high precision based on the far-field response [75].

The traditional forward modeling is to find the response for a given material, while the inverse design technique can realize the prescribed response within certain constraint in material design. Chen et al [76] proposed a DL method based on adaptive batch normalization neural network, which realized the intelligent and fast inverse design of graphene-based metamaterials with on-demand light response. However, MLP-based inverse design models often fail to converge during big data training, and one response corresponds to multiple metamaterial structure phenomena, which is called ‘one-to-many’ phenomenon or ‘non-uniqueness’. This inconsistency of data will slow down the training process. In order to solve the non-unique problem of inverse design, a training method based on forward modeling and inverse design can be used [77], as shown in figure 2(d), the network model on the left can be effectively trained on a data set containing non-unique electromagnetic scattering examples. This method is conducive to training nerves networks on large training data sets and provides help for the inverse design of complex metamaterial structures.

2.2. Convolutional neural network

As the simplest form of DL, MLP has been widely used in both natural materials and metamaterials. However, when the internal connection, target response, and design space are multi-modal or difficult to parameterize, the simple connection in the MLP model still brings some difficulties. Especially when processing image data, MLP model may lose the spatial characteristics of the image during the application process. With the significant increase of model-free parameters, CNNs began to appear to cope with high-dimensional data spaces. The basic principle of the CNN model is to extract the features of the input by performing convolution operations on the output of each layer. Through the convolution operation, CNN can capture the local correlation of the image space (such as spatial correlation and displacement invariance), thereby dealing with locally-correlated design parameters.

The CNN model was used for handwritten digit recognition as early as 1998 [78]. In recent years, the CNN model has been widely used in image processing technology since it can effectively process 2D shape data (such as image lines, edges, and structural directions). For large-scale image recognition on the order of millions, CNN requires fewer training parameters than MLP models of the same size layer, and can quickly learn thousands of objects and make accurate predictions.

In natural materials, due to the translational symmetry of the convolution operation, CNN can simulate the periodic structure, accurately and quickly predict the performance of the material according to the properties of the material [59], and learn the internal structure and chemical information of the periodic table [47], etc. In addition, CNN can also be used to automatically classify the crystal symmetry of the structure. Ziletti et al compactly encode the crystal symmetry by calculating the diffraction image and then use a CNN to predict the crystal to achieve classification [79]. The method in figure 3(a) can correctly classify crystal structures with serious defects, which paves the way for the possible noise and incomplete crystal structure identification in big data materials science. In addition to analyzing static images, there have been studies using it for dynamic STEM imaging [80] to quickly analyze data and find lattice defects, which is shown in figure 3(b). The discovery of material defects is mainly based on the fact that each defect is related to the violation of the ideal periodicity of the crystal lattice. Using a single image to train the CNN to obtain the crystal defect under the macroscopic periodicity, the extracted defect structure can be identified/classified via unsupervised clustering and decomposition techniques. Finally, local crystallography techniques can be used to further study the selected defects.
Figure 3. CNN structure diagram and applications. (a) Schematic representation of the CNN used for crystals classification [79]. (b) Training a deep CNN to recognize defects that break lattice periodicity [80]. (c) Illustration of the full inverse design problem and CNN solution. Among them, the inverse design triangle for optical metamaterials consists of reflectance/transmittance spectra, ellipsometry spectra, and metamaterial structure. On the right is a graphical representation of the CNN network structure for the reflectance/transmittance problem in the metamaterial structure [31]. (d) Structure of the DL model, where the CNN model can extract data from smaller parts of the image and extract spatial features from the structured image [48].

In metamaterials, CNN can be used to predict the optimal metamaterial design [81], and solve the inverse design problem between metamaterial structure and response [31]. As shown in figure 3(c), the developed model solves the mapping relationship between the metamaterial structure and the corresponding ellipse measure and reflectance/transmittance spectrum, highlighting the remarkable ability of neural networks in detecting large global design spaces. CNN can be applied to image labeling by combining it with other neural networks. When combined with a RNN, the absorption spectrum of the metamaterial is successfully extracted [48]. The specific model structure is shown in figure 3(d). CNN extracts spatial information such as shape and position from the image, and then RNN finds the relationship between the image and its optical features. The composite model can predict the optical response of an image structure in less than one second without requiring extensive computing power.

In recent years, the CNN model has made continuous progress in network architecture and algorithms, and its depth and breadth have been greatly improved. Advanced models such as deep CNNs [82] and residual networks [83] have emerged, and have been widely used in material science. A typical residual network Resnet-101 can be used to design anisotropic digitally coded metamaterials. CNN predicted the reflection stage and combined Resnet-101 with the binary particle swarm optimization (PSO) algorithm to achieve the design of multi-polar polarization states. This method can effectively complete the automatic design from the required reflection level performance to the target component mode, which greatly promotes the development of metamaterial intelligent design [84].

2.3. Recurrent neural network

MLP and CNN models are completely irrelevant between the previous input and the next input, and there is no connection between the layers, but the RNN model is different from them, it can process the related data before and after, and the neurons between each layer are also connected to each other and construct a weight connection relationship, whose behavior likes a memory. It maintains the state of the system and feeds back the obtained network output results to the input layer. It has been widely used in tasks such as natural language processing [85] and time series prediction [86, 87].

Although RNN can capture the historical dependence in the data, this method of combining the current input with the previous hidden state and calculating a new value to replace the activation is not suitable for long-term dependent sequential data processing. Therefore, the RNN later evolved more kinds of variants. For instance, the LSTM unit and the GRU mentioned above, as well as the time convolutional network (TCN) [88], stackable recurrent cell (STAR) [53], etc. Among them, TCN has been proved to be more...
suitable for predicting irreversible, historical and time-related phenomena than the former two [52]. For example, using multiple thermal viscoplastic constitutive models can calculate the phase transition process of the material. STAR requires fewer parameters than the first two, and can stack deeper systems [53]. These models are used in natural materials as well as metamaterials.

In the field of natural materials, RNN can predict the historically related responses of materials. For example, the RNN model based on GRU units predicts the mesoscale response of 2D elastoplastic composite RVEs under non-proportional loading conditions [89]. Furthermore, Diab W. Abueidda team applied various sequence learning models (LSTM, GRU, TCN) to quickly predict the stress and energy of cell materials [52]. The comparison found that TCN was more efficient in the calculation of training devices, which promoted the application of data-driven methods in the calculation of plastic and inelastic constitutive models. In addition to performance prediction, RNN generates atomic data in ab initio molecular dynamics (AIMD), and can also directly predict the atomic velocity and position of silicon atoms [90]. The material properties calculated by the prediction data are in good agreement with the ground real AIMD calculations. Figure 4(a) shows the accuracy of the RNN prediction can accelerate the AIMD and help the research of computational materials.

In order to improve the computational efficiency of the multiscale model of composite materials, RNN was used as an effective agent of the micro model to accelerate the multi-scale simulation of composite materials [91]. By predicting the nonlinear plastic response, RNN has proved its excellent performance in speed and accuracy at the micro level [92]. In particular, Mozaffar et al utilized RNN to transfer the complexity of plasticity theory in materials with complex microstructures and provided a substitute for DL to find constitutive models related to historical processes and microstructures [93]. The model in figure 4(b) has the ability to accurately predict the stress and most of the plastic energy region.

The LSTM-RNN model can directly predict the spectrum of the conjugated polymer from the coarse-grained representation and can explore the sensitivity level of the spectrum to the coarse-grained representation [94]. As shown in figure 4(c), this dependence is studied by using a sliding sub-window of variable length Ns along the polymer backbone, using a subset of sequential twist angles. This method provides a new tool suitable for improving the post-simulation analysis protocol.

Recognizing the relationship between materials and structures is a fundamental issue in materials science and plays a vital role in creating the next generation of materials. In the past, computational solid mechanics tried to use computer methods to predict or optimize the response of mechanical problems, such as meshless methods [95], finite element analysis [96], and isogeometric analysis [97]. However, numerical schemes for such modeling are time-consuming and computationally expensive [98, 99]. In addition, the emergence of 3D printing technology has made it possible to produce materials with unique geometric shapes. The mechanical response of this material exhibits plastic deformation. It is very challenging to predict the response of this complex microstructure. The emergence of DL models provides new solutions.

Since the standard feed-forward DNN cannot process sequential information, it is advisable to use RNN as a sequence learning model, which can be beneficial to many mechanics related problems [52]. Training RNN can accurately predict the entire historical related response of plasticity and thermoviscoplasticity from the input data of a specific domain. The sequence learning model learns the plasticity-constitutive relationship in order to predict the history of uniform stress tensor and plastic energy beyond the elastic limit based on load conditions and microstructure descriptors. Therefore, compared with other models, RNN is more suitable for predicting the mechanical properties of materials.

The application of RNN in metamaterials is less frequent. When it is combined with CNNs, image processing can be used to replace the previous numerical simulation methods to extract the absorption spectrum of the plasma metamaterial structure [48]. In figure 4(d), the CNN collects spatial information from the image, and then uses these data and the RNN to teach the model to predict the relationship between the spatial information and the absorption spectrum.

### 2.4. Graph neural network

In recent years, an increasing number of researchers have begun to be attracted by the GNN, which has been widely used in graph analysis-related research fields because of its powerful function in modeling dependencies between graph nodes. Using GNN, the physical interaction between nodes can be learned through the training process, and the essence of these interactions can be extended to different configurations of adjacent nodes, which are continuously updated through training, learning, and state information exchange between each node until the model reaches a steady state. The final output is an abstract representation of node and edge attributes and graph structure, which are further processed by the fully connected layer to obtain the required physical response.

GNN has been widely used in the field of natural materials, for modeling crystalline materials. Generally speaking, there are two ways to model crystalline materials. One is an artificial structure feature vector, and
Figure 4. RNN model architecture and applications. (a) The left picture is the radial distribution function (RDF) of ground real AIMD simulations and predictions with different gaps, averaged over the past 9000 steps; the right picture is the atomic configuration from ground real AIMD simulations and RNN predictions with different gaps ($M = 1, 2, 5$ and $10$) [90]. (b) Evaluation results for the trained model. On the left is the plastic energies, and on the right is the unidirectional loading test [93]. (c) Predicting spectra of conjugated polymers [94]. (d) Schematic of the process, from the 3D structure to the absorption curve output [48].

the other is to use the complex transformation of atomic coordinates to realize the input of crystal structure. However, the former requires case-by-case design when predicting different attributes, while the latter is difficult to interpret the model as a complex conversion result. Therefore, Tian Xie and Jeffrey C. Grossman [100] developed a crystal graph CNN (CGCNN) framework, as shown in figure 5(a). It can learn the properties of materials directly from the connections of atoms in the crystal, and the framework constructed is interpretable. It provided a flexible method for material performance prediction and design. By studying the connection of atoms in the crystal, a CGCNN model is constructed directly on the crystal diagram generated by the crystal structure, and the characteristics of the crystal structure are extracted by using this model, so as to realize the design of the material.

After that, the CGCNN model was further improved. The improved variant of the CGCNN model [101] and the geometry-information-enhanced crystal GNN [58] appeared. The former connects the clear three-body association of adjacent constituent atoms and optimizes the chemical representation of atomic bonds in the crystal diagram, which shortens the high-throughput search time and achieves superiority to CGCNN. The latter can learn complete topology and spatial geometric structure information through the
Figure 5. GNN model architecture and applications. (a) Illustration of the CGCNNs, the above is a graph in which the crystal is transformed to have nodes representing the atoms in the unit cell and edges representing the connections of the atoms. Below is the structure of the CNN at the top of the crystal diagram \[100\]. (b) Graph-based representation of an \(N\)-grain polycrystalline microstructure. The \(M\) different physical features (\(M = 5\) herein) of each grain (label \(i; i = 1, 2, \ldots, N\)) are stored using \(N\) different \(M\)-component feature vectors \(f_k^i (k = 1, 2, \ldots, M) \[57\]. (c) Architecture model GATGNN. Each model consists of five or seven AGAT layers. After extracting node features, a global attention layer is placed before the global pooling of node feature vectors. The weighted sum of the crystal feature vector is afterward fed to two hidden fully-connected layers before a final fully-connected layer outputting the predicted property \[37\]. (d) Model architecture. A graph network is composed of the encoder (ENC), the core (G) and the decoder (DEC). Its core is to first update the edges according to itself and adjacent nodes, and then update the nodes according to itself and the incoming edges. It is able to predict the tendency of each type A particle \[106\].

distance vector between adjacent nodes, which results in higher prediction accuracy than the previous CGCNN.

GNN is very effective in predicting the performance of materials. Although other neural networks are also widely used to predict the properties of materials, these models hardly take into account the interaction between particles within the material. A GNN model was developed \[57\] to obtain the embedding of polycrystalline microstructures and accurately and interpretably predict the properties of polycrystalline materials, as shown in figure 5(b). In addition, various GNN model variants are widely used to predict the properties of crystal materials. For example, the co-crystal graph network \[102\] can achieve high-precision prediction of different data in different eutectic spaces and has strong robustness and generalization. Tuples GNN (TGNN) \[103\] can accurately predict the band gap of crystal materials, and has shown good performance in the band gap prediction of four open material databases. Embedding an encoder–decoder in the orbital graph CNN \[104\] can learn the element characteristics, interactions between orbitals and topological features of crystal materials for discovering new materials.

The GNN model can also predict the crystal density of high-energy compounds \[105\] and the performance of inorganic materials \[37\]. In order to fully understand the relationship between atoms in materials and the effects of atoms on material properties, researchers \[37\] proposed a GNN model global attention GNN (GATGNN). The model architecture is shown in figure 5(c). The model is composed of a multiple graph-attention layers (GAT) and a global attention layer. The local relationship between adjacent atoms and the contribution of atoms to the material properties can be obtained, which further enhances the GNN model interpretability.

The GNN model can determine the long-term evolution of the glass system only from the initial particle position \[106\], and the model structure is shown in figure 5(d). In the shear experiment, it predicts the position of the rearranged particles. The use of some hidden structures around the particles shows that the graph network is a powerful tool for predicting the long-term dynamics of the glass system.

Table 1 visually shows the application of the aforementioned models in the field of materials. In fact, all feedforward networks can be used as approximate solvers for forward problem simulation, including MLP, CNN, GNN, etc, depending on the input data structure. For complex dielectric metamaterials, a combination of MLP models and convolutional layers can be used to achieve rapid modeling, but to solve highly complex nonlinear modeling, bidirectional neural networks will be a better choice. For modeling time series systems,
Table 1. Documents on the basic model architecture of DL. The table is classified according to the application of MLP\CNN\RNN\GNN model in the field of natural materials and metamaterials, and briefly commented.

| Model   | Target application | References        | Model functionality                                                                 |
|---------|--------------------|-------------------|-------------------------------------------------------------------------------------|
| MLP     | Natural material   | [33, 61, 62]      | Provide acceleration and approximate simulation for traditional computational physics and inverse design |
|         | Metamaterial       | [44, 71, 72, 74–76] | Predict the relationship between the structural parameters of metamaterials and the electromagnetic response; realize on-demand inverse design |
| CNN     | Natural material   | [47, 59, 78, 79]  | Simulate the periodic structure, accurately and quickly predict its performance based on the material |
|         | Metamaterial       | [31, 48, 80]      | Solve the inverse design problem between structure and response |
| GNN     | Natural material   | [52, 53, 89]      | Predict the historically related responses of materials |
|         | Metamaterial       | [93]              | Predict the relationship between spatial information and absorption spectra |
| RNN     | Natural material   | [48, 57, 58, 100–102, 104, 105] | Take into account the interaction between particles in the material, so as to flexibly predict material properties. Improve the interpretability of the model |
|         |                    | [103]             | Learn the characteristics of crystal materials and discover new materials |
|         |                    | [37]              | Determine the long-term evolution of the glass system |

a RNN model that returns the output of the network to the input layer will be an ideal choice. For graphical data, such as the modeling of crystal materials with dependencies among atoms, the GNN model is preferred. The MLP model is simple to connect and is suitable for processing discrete vector data. When faced with multi-modal or difficult to parameterize data, CNN or RNN models are better alternatives. The CNN model can be used to process image data, and the RNN model can process time series signal. It is worth noting that Deep RNN also has the problem of gradient disappearance and explosion. If there is no preprocessing step, RNN model cannot handle node-based applications [107]. Finally, GNN model is more flexible, and different models have very different properties and settings, but there is no unified learning framework.

3. Advanced deep learning models

3.1. Deep generative models

The generative model is based on the joint distribution $P(x,y)$ of input samples and output results, and optimizes the target in a probabilistic way. It contrasts with the supervised discriminant model that creates the conditional distribution $P(y|x)$. The generative model can generate data that is replicated in the same or similar way as the training data set, and is designed to deal with one-to-many mapping problems. Among the generative models, the variational AE (VAE) model and the generative adversarial network (GAN) model are two popular model categories. Since metamaterials are continuous in the design space, most of them can be generated in a variety of ways, while natural materials are often relatively discrete and cannot be combined arbitrarily (subject to the limitation of elements and charges). Therefore, generative models are more suitable in the case of metamaterial design.

In order to better introduce the VAE model, first of all, it briefly introduces the AE model, which is the basis for developing the VAE model. AE is a neural network designed to learn data features in an
unsupervised way. It consists of an encoder and a decoder. The encoder implements the dimensionality reduction processing of the training data, and the decoder reconstructs the variables mapped to the latent space into the training data. Because the AE model has good dimensionality reduction capabilities, it is used to optimize electromagnetic nanostructures. The use of AE can simplify the design and response space [108] and greatly reduce the computational complexity. At the same time, by reducing the dimensionality of the conceived design parameters, a valuable intuitive understanding can be obtained from fewer design parameters, and the role of each design parameter in the overall response of the nanostructure can be known [109].

The VAE is similar to VE, except that it increases the probability perturbation in the AE latent space. Since the trained VAE network can sample new latent vectors and become new data similar to training data after decoding, VAEs is a deep generative model.

Considering the availability of labeled data, generative models can reduce the pressure of data collection. Using the dimensionality reduction and continuity of the latent space, the generative model based on VAE can design and characterize metamaterial structures [36]. The model structure is shown in figure 6(a). The encoded 2D metamaterial shows that the model effectively uses randomly generated unlabeled data to discover the complex relationship between structure and performance.

For the multi-constrained optimization of metamaterials, VAE can naturally solve the one-to-many problem. For example, Kudyshev et al developed an optimized global framework that uses an adversarial AE (AAE) combined with a meta-heuristic optimization framework to significantly improve the optimization search efficiency of metamaterial with complex topologies, as shown in figure 6(b). The regularization technology introduced in the AAE compressed design space has led to a better multi-parameter global optimization search. Researches show that this framework can be applied to a wide range of highly constrained optimization problems [110, 111]. By encoding the metamaterial design into the latent space and sampling the latent variables as the probabilistic representation of the metamaterial, it is easy to realize the retrieval of the metamaterial structure under the required spectrum, thereby solving the one-to-many mapping in the inverse design problem [112].

The combination of VAE model and other algorithms can realize effective, on-demand and automated design of metamaterials. For instance, the developed VAE-ES model [115] which can identify target patterns with edge divergence, and can automatically and globally identify the best structure of continuous or discrete topological metamaterial in a short time. When VAE is combined with optimized support vector machines to form AMID, it can realize metamaterial inverse design and serve as a new research venue for automatic metamaterial design and efficient wave manipulation [116].

In a word, as a probabilistic generation network, the VAE model can well solve the one-to-many mapping problem compared with the deterministic model. But it also has certain limitations. The trade-off between reconstruction loss and KL divergence will force the latent space into a Gaussian space, leading to imperfect modeling. In addition, the dimensionality of the VAE latent space is arbitrary, which will cause a manual adjustment in multiple training iterations, and the resulted first solution is often not the best.

In addition to the VAE model mentioned above, there is another popular and highly successful deep generative model—the GAN. GAN is a type of deep generative model composed of generators (GN) and discriminators (DN). The GNs are trained in a confrontational manner. Learning training data can generate distributions that are indistinguishable from the training data and produce the effect of deceiving the DN. The DN is mainly used to distinguish between generated data and real data. The two conflicting goals are mutually reinforced. The ultimate training goal is to allow the GN to generate sufficiently realistic examples so that a good DN cannot accurately distinguish between the generated data and the real training data.

As an unsupervised learning model, GAN was first proposed in 2014 and is mainly used in inverse problems in image generation [117], style transmission, thermal emitter structure design [111], and super-resolution [118]. GAN has been proven to be applied for modeling continuous distributions with continuous data [119], however, more circumstances are to deal with 2D or 3D image structure. The GAN model was first applied to metamaterial design by S. Liu and his coworkers. The use of an image-based description of metamaterial can provide arbitrary patterns of the unit structure that is not easily predefined [117]. Sunae So and Junsuk Rho used conditional deep convolution GANs (cDCGAN) to design optical metamaterial antennas. The generated designs are presented in the form of images, thus basically realizing the free design of any desired optical characteristics. The designed cDCGAN structure is shown in figure 6(c), which combines the ideas of CNN and GAN. This network is not only suitable for predefined structures, but also for new structure designs that cannot be expressed by structural parameters. It provides a fast and convenient method for the design of complex metamaterial structures [113].

GAN can also be applied to photonic crystals. Figure 6(d) shows the generative confrontation network can perform a high-throughput inverse design of photonic crystals and synthesize candidate unit cells with a large number of TM band gaps [114]. Moreover, three different GAN variants have been studied: traditional
Figure 6. Generative model architecture and application. (a) The proposed DL model with a self-supervised learning mechanism for both the forward prediction and inverse design of nanophotonic structures [36]. (b) Conditional AAE-based data generation [110]. (c) On the left is a schematic diagram of data preparation for DL. On the right is the result of cDCGAN’s suggested hand-drawn spectrum [113]. (d) On the left is a GAN. Through the confrontation between generative (G) and discriminatory (D) networks, a new synthetic example (false) of a 2D unit with a TM band gap is generated from a real data set (real). The right is the fidelity of generated unit cells [114].

GAN [120], least-squares GAN [121], and deep regret analysis GAN (DRAGAN) [122]. The study found that every training each GAN variant can produce convincing unit cells and exhibit the required characteristics. Only in terms of visual quality and fidelity for the design goal of a large band gap, the three variants show different performances.

Because GANs have the problem of mode collapse, some generated results will be preferentially selected by the DN, thereby converging to a smaller subset, so the network cannot effectively explore the entire design space. This mode collapse is understood as the result of an undesirable partial balance [122], and a simple strategy to alleviate the mode collapse of GAN is to normalize the DN to constrain its gradient in the environmental data space, such as DRAGAN can achieve faster and more stable training. In addition, there is a Wasserstein GANs [123] method that does not need to balance the GN and DN capabilities, and can also solve the mode collapse problem.

In addition to the mode collapse problem, GAN is also difficult to be trained, since the two independent DNNs have to be optimized at the same time. But when combining GANs with other more traditional
optimization methods [124] can handle complex structural designs, improve training efficiency, and produce extremely impressive results.

3.2. Hybrid models with other algorithms
With the increase of design space, for either natural materials or metamaterials, traditional optimization methods often fail to solve the problem in an effective way [110]. DL can bypass the complex design process, enabling people to directly predict the target properties, or perform inverse design to retrieve the eligible designs. However, DL is a data-driven method whose performance is highly relevant to the quality and quantity of training data, so DL alone may not function well in some scenarios with extreme target responses. In such a case, combining DL with other traditional optimization algorithms may provide effective solutions. One common practice is to use DL as an alternative model of physical simulation. Compared with traditional numerical theoretical calculations, its biggest advantage is its extremely high evaluation speed. Once trained, artificial neural networks can make predictions in milliseconds, which are usually orders of magnitude faster than numerical simulations. Therefore, replacing traditional physical simulation with artificial neural networks is a natural solution for accelerating the inverse design of materials or metamaterials through global optimization heuristics. But this method still has shortcomings. When another model is used, it is only used to represent an approximation to the real model, thus introducing systematic errors. To make matters worse, the surrogate model is not guaranteed to contain singularities that the optimization algorithm may converge to in the worst case [125]. Therefore, the interactive relation between DL and traditional optimization algorithms must be carefully coordinated.

In the field of natural materials, DL can be combined with other optimization algorithms to promote material design, including genetic algorithms (GAs), topology optimization methods, Bayesian optimization, adjoint methods, etc. The hybrid model generated by an artificial neural network and adaptive group GA (SAPGA) optimizes the blast furnace smelting in the steelmaking process. RNN obtains dynamic blast furnace information to determine the current optimal solution. Compared with the ordinary steelmaking process, the hybrid model has obvious optimization capabilities and has played a role in the optimization of the ironmaking process [126]. The combination of topology optimization methods and DL algorithms (such as AAEs) promotes a wider range of optimization design and data-driven material synthesis and significantly improves computational efficiency [111]. The hybrid model of DL and global optimization algorithms (Bayesian optimization and GA) can well solve the problem that DL models require a large number of data sets, and efficiently and accurately reverse the material composition [127]. Using a hybrid finite element algorithm and feedforward neural network model, it is possible to design and design high-performance structural ceramics that experience thermal load [128]. In a hybrid model composed of DNNs and migration learning algorithms, migration learning is used to solve typical small data collection problems in materials. The DNN model can predict the complete UV–vis absorption spectrum of the material from the compound formula; through the initial migration learning training and fine-tuning of parameters, the prediction model performs well in the prediction of the spectrum of metal oxide materials containing only composition information [127].

The use of hybrid models can also solve the optimization design of multilayer coatings in perovskite solar cells [129] and evaluate the status of 3D printing structures [130]. A more complex hybrid model is used to predict the properties of the grain boundary (GB). Understanding the composition-structure-property relationship of the GB has fundamental and practical significance for the design of polycrystalline materials. Hu et al. proposed an isobaric semi-large regular integrated hybrid Monte Carlo and molecular dynamics (hybrid MC/MD) simulation and GA and DNN hybrid model [131]. The model architecture is shown in figure 7(a). The final model can not only predict the characteristics of simple symmetrical tilt and twist GBs, but also predict the characteristics of more complex and general GBs.

In the field of metamaterials, due to the complex structure design of metamaterials and the high degree of freedom element background, different hybrid models have been developed. These models mainly solve the problems in metamaterials in the following two ways. (a) Optimize the design space directly. The use of neural agents replaces all or part of the forward model. (b) By reducing the dimensionality of the output mapped to the design space [132], and reducing the amount of calculation, the optimization step can work more effectively in the reduced space. Both of these two methods have significant advantages over traditional optimization methods. These two methods are superior to traditional methods in terms of time, cost, and accuracy. Various examples are described below.

Based on the DL model and genetic model, inverse design and optimized generation of structural units in metamaterials can be realized [133]. This can also be achieved through the combination of the DL model and PSO algorithm [134]. Combining the combined pattern generation network (CPPN) and the co-evolution algorithm can solve the problem of supramolecular inverse design in metamaterials. It is used to arbitrarily
Figure 7. Hybrid model architecture and application. (a) The workflow of the GA-guided DNN prediction for GB diagrams [131]. (b) Description of modules of the hybrid framework [135]. (c) Global optimization based on a generative neural network [138].

manipulate light polarization and wavefront metal supramolecules [135]. The model architecture is shown in figure 7(b). CPPN takes the coordinates of pixels in the image as input, then predicts the corresponding pixel values and iterates all the coordinates. The predicted pixels are assembled into a pattern, and the co-evolution algorithm evaluates the performance of the framework through the inverse design of the metamaterial composed of two supramolecules. This framework model is expected to accelerate the design of large-scale metamaterials. Using transfer learning technology and GAs, it is possible to quickly and accurately design a phase modulation dielectric metamaterial scheme [136]. Combining Bayesian optimization and deep CNN (DCNN) algorithms can calculate and optimize the optical properties of metal nanostructures [137]. This model provides a wide range of applications for future nanostructure analysis and design. The combination of the topology optimization method and DL algorithm can optimize the design of thermal radiation metamaterial with efficient thermal emission shaping function. Through the control of the design space, this method can achieve high-efficiency thermal reshaping, and the search speed is 4900 times faster than the traditional direct topology optimization [111].

In traditional topology optimization, when updating the structural parameters of a metamaterial with a fixed deflection angle and wavelength (such as the refractive index of each part), changing its target or working wavelength, it is necessary to run the metamaterial simulation model again. To a certain extent, it will increase workload and reduce efficiency. So there is a global network topology optimization model (GLOnet) based on GNN and topology optimization. The model only needs to update the weight of the neural network in each iteration. Related research is shown in figure 7(c) [138]. The results show that the
production efficiency of GLOnets is equivalent to or better than the best equipment based on the accompanying topology optimization, and the required computational cost is lower.

The stability of predicting materials can be achieved by transfer learning [47], which can effectively improve model accuracy with limited training data. Another approach is to use the combined information of ICSD, QMMD and the periodic table to achieve stability prediction. The discovery of material defects in DL is mainly based on the principle that defects are related to the violation of the ideal periodicity of the crystal lattice, and the network is trained to find defects. In addition, Markov analysis is an effective method to identify the transition probability between different defect configurations. One can also use the VAE model to perform dimensionality reduction processing on the input data first, and combine the classifier to predict the combination of elements of a specific topological structure, so as to transform the prediction problem of crystal structure into a problem of predicting the possibility of a single atom’s position in the structure [28].

4. Challenges and perspectives

4.1. Interpretable deep learning models

DL models consist of a hierarchy of hidden layers to uncover the complex relationship among training data, but usually these internal representations are difficult to understand or explain, which makes it a ‘black box’ model. Material scientists have repeatedly tried to understand the internals of data-driven models to obtain some guiding information for material development [139, 140]. A sparse modeling method called least absolute shrinkage and selection operator was proposed in the early days. It uses the linear model of L1 regularization [141] to automatically select important descriptors (attributes) and reduce the dimensionality of the search space, but it does not have high prediction accuracy in the face of nonlinear material data. In order to adapt to nonlinearity as well as the trend of big material data in recent years, the interpretability research of DNNs has received more and more attention.

Interpretation of the DL model can be evaluated by how the model mechanism can be explained to humans. If the model cannot be understood by people, then it is prone to be unstable [142]. The interpretability of machine learning methods in data analysis is considered an interactive process. The model is generated based on the data used in the selected machine learning algorithm, and the explanatory tools corresponding to these algorithms are used to apply the model to the structure to interpretation. Then feedback the results of the interpretation are returned to humans, so that people understand how the model is generated according to the internal law of the data, and people can also provide suggestions to the process in a model-adapted way to achieve interaction.

An effective way to achieve interpretability is to perform dimensionality reduction. With the continuous development of materials science, the complexity of material structure lead to data of high-dimensionality. If all data attributes are retained for model training, there may be attributes that are irrelevant in the high-dimensional data [143], which will eventually make the model unexplainable and susceptible to noise. To alleviate the side effect of high-dimensional data, Yadav et al developed a model called the deformation manifold learning model to predict multi-walled carbon nanotubes (MWCNT) composed of millions of atoms, which includes unsupervised dimensionality reduction (via proposed a consequence functional principal component analysis (c-FPCA)) of the deformed manifold and supervised learning (via DNN) of deformation in the reduced dimension. Among them, the c-FPCA technology reduces the curse of dimensionality by providing a low-dimensional function representation for the deformation of MWCNT. The model extracts the dominant mode and relative contribution of structural deformation in an unsupervised manner, thereby enhancing the interpretability of the model [144]. In addition, there is also an unsupervised AE neural network to achieve dimensionality reduction. For example, Kiarashinejad et al use an unsupervised AE ANN to reduce the electromagnetic simulation value of the reflection spectrum. This unsupervised model requires training when there is less data, which helps to extend the technology to experimental data [145]. Similarly, Iten et al used the encoder–decoder ANN for interpretable physical discovery [146]. The research results show that this method of modeling neural networks after human physics deduction can assist scientific discovery without making assumptions about the system in advance.

Another way to enhance the interpretability of machine learning is to physically parameterize the training data. For example, the extinction spectrum can be preprocessed in the modal decomposition [147, 148]. Once trained, the respective neural network provides a clear interpretation of the predicted spectrum. So far, through the development of interpretation techniques [149], the black-box nature of DL models is becoming increasingly transparent. After training, DL with high predictive ability and interpretability can be applied to evaluate the importance of metamaterial structures to optical performance, identify the mechanism of
light–matter interaction [145, 150], assist in experimental design [151, 152], and research in the development of new materials [144, 153].

4.2. Burden of data collection
As a data-driven method, DL requires sufficient data to train models with high accuracy and the accuracy of a DL model is directly affected by the quantity and quality of training data. For the supervised learning model, the biggest limitation is that it relies on a large amount of data, most of which come from numerical calculation or repeated experiments. The former needs to consume a lot of computational cost and time cost, while the latter has even a longer period with considerable failure rates.

There are several ways to alleviate the burden of model data. One is to obtain higher accuracy under a given fixed amount of training data. Active learning [154, 155] is used to improve model accuracy. Active learning (also known as optimal experimental design) has been widely promoted in the field of machine learning [155, 156]. A widely used method is 'uncertainty sampling', it is assumed that the best locations for collecting training data are those locations where the model has the least confidence in its predictions. This sampling method can produce a more accurate model than the model constructed by traditional strategies.

In addition, transfer learning can increase model generalization to ease the data burden [157, 158]. Researchers can make better use of the common physical mechanism between different materials by developing new transfer learning techniques, and applying the trained network architecture to multiple similar tasks, thereby alleviating the data dependence of the model. The other method is mainly aimed at the problem of insufficient experimental data to fully train the model, imitating the computer vision community and the human genome project, building a material genome, and realizing a more open data-sharing platform. The development of big data is attributed to the joint efforts of all scientists. By publishing experimental data on the platform and achieving open source, it will be able to effectively solve the problem of insufficient data and promote the continuous improvement of algorithm models and the rapid development of materials science. So far, many large-scale material databases have been developed, such as AFLOWlib [159], OQMD [160], MatWeb [161], Materials Project [162], Cambridge Structured Database [163], etc. These open data sets can be accessed through multiple channels to realize interactive exploration and data mining.

Although the above methods can alleviate the burden of data to a certain extent, the collection of a large amount of training data itself will incur time costs, coupled with the curse of dimensionality [164] caused by high-dimensional data, requiring us to find a new way to alleviate data burden. Unsupervised and semi-supervised learning strategies can be leveraged to alleviate the burden of data collection. Since the supervised learning model requires a large amount of data to support its learning from correct labels, it causes a large data burden and is suitable for design with few parameters. Unsupervised learning and weakly supervised learning can achieve learning with fewer or no labels. Weakly supervised learning is able to uncover features in the unlabeled data set and directly output the design results through an iterative loop [36, 165], so it can be well applied to multi-parameter design. Among them, the problem of model non-convergence caused by the increase of design parameters is avoided. Several typical unsupervised models, such as principal component analysis and AEs/VAEs/GANs, can help the neural network identify the basic distribution of the data and ease the data burden.

5. Summary
From the perspective of model selection, this review surveyed the application of DL in the field of materials science, ranging from the prediction of material property to the design of metamaterials. This review primarily focus on the application of discriminative models under different data structures, which provides a powerful tool for material discovery, performance prediction and optimization of the experimental process. Then more advanced models are introduced covering generative models and hybrid models. Finally, the introduction of unsupervised models and weakly supervised models is supplemented, the difficulties and challenges of DL models are analyzed, and corresponding solutions are proposed.

Past research results have demonstrated the potency of DL methods to solve complex material problems, and these results in turn provide new guidance for DL practices in future materials research. For example, the GNN model in the DL model can be used to model nanocomposite materials. This neural network provides experimental data and ‘learns’ the relationship between stress and strain. The modeling strategy is effective for modern complex materials (such as composite materials). The use of DCNNs has been proven to be able to identify the defect parameters of 3D printing materials [130], and the use of RNN models can also predict the plastic deformation of materials with unique geometric shapes (such as 3D materials) [52].

Judging from the historical development of materials, the design flexibility of materials is constantly improving, which poses a greater challenge to computational and modeling strategies in research. It is
expected that in the future, in order to cope with multi-dimensional and highly complex design problems, the architecture and training strategy of DL models will continue to improve, leading to better accuracy and generalization capabilities. With the development of the Material Genome Project, the materials science platform led by high-throughput data will become more abundant and complete, facilitating the wider application of DL and significantly promoting the development of materials science.

**Data availability statement**

No new data were created or analysed in this study.

**Acknowledgments**

Wei Ma acknowledge the financial support from the National Key R&I Program of China (Grant No. 2021YFA1401200), the National Natural Science Foundation of China (Grant No. 62105285) and the Fundamental Research Funds for the Central Universities (No. 2021QNA5006). Shuliang Dou acknowledge the financial support from the National Natural Science Foundation of China (Grant No. 51902073). Yao Li acknowledge the financial support from the Fundamental Research Funds for the Central Universitie (Grant No. HIT.OCEF.2021004).

**ORCID iDs**

Guandong Liu  [https://orcid.org/0000-0001-8985-8924](https://orcid.org/0000-0001-8985-8924)

Zhiyuan Zhu  [https://orcid.org/0000-0003-1534-8432](https://orcid.org/0000-0003-1534-8432)

Yao Li  [https://orcid.org/0000-0003-3859-8135](https://orcid.org/0000-0003-3859-8135)

Wei Ma  [https://orcid.org/0000-0002-5665-7840](https://orcid.org/0000-0002-5665-7840)

**References**

[1] Liu Y, Niu C, Wang Z, Gan Y, Zhu Y, Sun S and Shen T 2020 Machine learning in materials genome initiative: a review J. Mater. Sci. Technol. 57 113–22
[2] Kalinin S V, Sumpter B G and Archibald R K 2015 Big–deep–smart data in imaging for guiding materials design Nat. Mater. 14 973–80
[3] Ward I, Liu R, Krishna A, Hegde V I, Agrawal A, Choudhary A and Wolverton C 2017 Including crystal structure attributes in machine learning models of formation energies via voronoi tessellations Phys. Rev. B 96 024104
[4] Nicola N 2016 Can artificial intelligence create the next wonder material Nature 533 22–25
[5] Curtarolo S, Setyawan W, Wang S, Xue J, Yang K, Taylor R H, Nelson I J, Hart G L, Sanvito S and Buongiorno-Nardelli M 2012 AFLLOWLIB. ORG: a distributed materials properties repository from high-throughput ab initio calculations Comput. Mater. Sci. 58 227–35
[6] Shelby R A, Smith D R and Schultz S 2001 Experimental verification of a negative index of refraction Science 292 77–79
[7] Lee S H, Park C M, Seo Y M and Kim C K 2010 Reversed doppler effect in double negative metamaterials Phys. Rev. B 81 241102
[8] Yu N, Genevet P, Kats M A, Aieta F, Thétou J-P, Capasso F and Gaburro Z 2011 Light propagation with phase discontinuities: generalized laws of reflection and refraction Science 334 333–7
[9] Smith D R and Kroll N 2000 Negative refractive index in left-handed materials Phys. Rev. Lett. 85 2933
[10] Valentine J, Li J, Zentgraf T, Bartal G and Zhang X 2009 An optical cloak made of dielectrics Nat. Mater. 8 568–71
[11] Ni X, Wong Z J, Mrejen M, Wang Y and Zhang X 2015 An ultrathin invisibility skin cloak for visible light Science 349 1310–4
[12] Zhang X and Liu Z 2008 Superlenses to overcome the diffraction limit Nat. Mater. 7 435–41
[13] Yu N and Capasso F 2014 Flat optics with designer metasurfaces Nat. Mater. 13 139–50
[14] Selvaratnam B and Koodali R T 2021 Machine learning aided study of sintered crystals J. Phys. Mater. 5 (2022) 014003
[15] Chmiela S, Selvaratnam B, Koodali R T and Kupferman J 2021 Machine learning in experimental materials chemistry Catal. Today 371 77–84
[16] Voiles P M 2017 Informatics and data science in materials microscopy Curr. Opin. Solid State Mater. Sci. 21 141–58
[17] Deng Z, Yin H, Jiang X, Zhang C, Zhang K, Zhang T, Xu B, Zheng Q and Qu X 2018 Machine learning aided study of sintered density in Cu-Al alloy Comput. Mater. Sci. 155 48–54
[18] Ghiringhelli I M, Vybiral J, Levchenko S V, Draxl C and Scheffler M 2015 Big data of materials science: critical role of the descriptor Phys. Rev. Lett. 114 105503
[19] Lee J, Seko A, Shitara K, Nakayama K and Tanaka I 2016 Prediction model of band gap for inorganic compounds by combination of density functional theory calculations and machine learning techniques Phys. Rev. B 93 115104
[20] Behler J, Martí-Alfaro R, Donadio D and Parrinello M 2008 Metadynamics simulations of the high-pressure phases of silicon employing a high-dimensional neural network potential Phys. Rev. Lett. 100 185501
[21] Wicker J G and Cooper R I 2015 Will it crystallise? predicting crystallinity of molecular materials CrystEngComm 17 1927–34
[22] Schütt K T, Gastegger M, Tkatchenko A, Müller K R and Maurer R J 2019 Unifying machine learning and quantum chemistry with a deep neural network for molecular wavefunctions Nat. Commun. 10 1–10
[23] Schütt K T, Arbabzadah F, Chmiela S, Müller K R and Tkatchenko A 2017 Quantum-chemical insights from deep tensor neural networks Nat. Commun. 8 1–8
[24] Chen C, Ye W, Zuo Y, Zheng C and Ong S P 2019 Graph networks as a universal machine learning framework for molecules and crystals Chem. Mater. 31 3564–72
[25] Schütt K T, Sauceda H E, Kindermans P J, Tkatchenko A and Müller K R 2018 Schnet—a deep learning architecture for molecules and materials J. Chem. Phys. 148 241122

No new data were created or analysed in this study.
Chen W-K, Liu X-Y, Fang W-H, Dral P O and Cui G 2018 Deep learning for nonadiabatic excited-state dynamics J. Phys. Chem. Lett. 9 6702–8
Kim E, Huang K, Jegelka S and Olivetti E 2017 Virtual screening of inorganic materials synthesis parameters with deep learning npj Comput. Mater. 3 1–9
Kim E, Huang K, Saunders A, McCallum A, Ceder G and Olivetti E 2017 Materials synthesis insights from scientific literature via text extraction and machine learning Chem. Mater. 29 9436–44
Ryan K, Lengyel J and Shatruk M 2018 Crystal structure prediction via deep learning J. Am. Chem. Soc. 140 10158–68
DeCost B I, Francis T and Holm E A 2017 Exploring the microstructure manifold: image texture representations applied to ultrahigh carbon steel microstructures Acta Mater. 133 30–40
Wiecha P R and Muskens O L 2019 Deep learning meets nanophotonics: a generalized accurate predictor for near fields and far fields of arbitrary 3D nanostructures Nano Lett. 20 329–38
Lininger A, Hinczewski M and Strangi G 2021 General inverse design of thin-film metamaterials with convolutional neural networks arXiv Preprint (arXiv:2104.01952)
An S, Zheng B, Shalaginov M Y, Tang H, Li H, Zhou L, Dong Y, Haerinia M, Agarwal A M and Rivero-Baleine C 2022 Deep convolutional neural networks to predict mutual coupling effects in metasurfaces Adv. Opt. Mater. 10 2102113
Asano T and Noda S 2018 Optimization of photonic crystal nanocavities based on deep learning Opt. Express 26 32704–17
Qiu T, Shi X, Wang J, Li Y, Qu S, Cheng Q, Cui T and Sui S 2019 Deep learning: a rapid and efficient route to automatic metasurface design Adv. Sci. 6 1900128
Rumelhart D E, Hinton G E and Williams R J 1986 Learning representations by back-propagating errors Nature 323 533–6
Ma W and Liu Y 2020 A data-efficient self-supervised deep learning model for design and characterization of nanophotonic structures Sci. China Phys. Mech. Astron. 63 1–8
Louis S-Y, Zhao Y, Nasiri A, Wang X, Song Y, Liu F and Hu J 2020 Graph convolutional neural networks with global attention for improved materials property prediction Phys. Chem. Chem. Phys. 22 18141–8
Martens J 2010 Deep learning via hessian-free optimization ICML 27 pp 735–42
Smith L I 2017 Cyclical learning rates for training neural networks 2017 IEEE Winter Conf. on Applications of Computer Vision (WACV) pp 164–72
Goodfellow I, Bengio Y and Courville A 2016 Deep Learning (Cambridge, MA: MIT press)
Zhang Y, Peng B, Zhou X, Xiang C and Wang D 2019 A deep convolutional neural network for topology optimization with strong generalization ability arXiv Preprint (arXiv:1901.07761)
Wang A Y-T, Murdock R J, Kawue S K, Olynyk A O, Gurlo A, Brugo J, Persson K A and Sparks T D 2020 Machine learning for materials scientists: an introductory guide toward best practices Chem. Mater. 32 4954–65
Wang P, Gao R X and Yan R 2017 A deep learning-based approach to material removal rate prediction in polishing CIRP Ann. 66 429–32
Ma W, Cheng F and Liu Y 2018 Deep-learning-enabled on-demand design of chiral metamatetals ACS Nano 12 6326–34
Hemmatyar O, Abdollahramezani S, Kiarashinejad Y, Zandehshahvar M and Adibi A 2019 Full color generation with fano-type resonant HFO 2 nanopillars designed by a deep-learning approach Nanoscale 11 21266–74
Meng M, Chua Y J, Wouterson E and Ong C P K 2017 Ultrasonic signal classification and imaging system for composite materials via deep convolutional neural networks Neurocomputing 257 128–35
Zheng X, Zheng P and Zhang R-Z 2018 Machine learning for materials scientists: an introductory guide toward best practices Chem. Rev. 118 13628–32
Sajedian I, Kim J and Rho J 2019 Finding the optical properties of plasmonic structures by image processing using a combination of convolutional neural networks and recurrent neural networks Microsyst. Nanoeng. 5 1–8
Bengio Y, Simard P and Frasconi P 1994 Learning long-term dependencies with gradient descent is difficult IEEE Trans. Neural Netw. 5 155–67
Hochreiter S and Schmidhuber J 1997 Long short-term memory Neural Comput. 9 1735–80
Chung J, Gulcehre C, Cho K and Bengio Y 2014 Empirical evaluation of gated recurrent neural networks on sequence modeling arXiv Preprint (arXiv:1412.3555)
Abueidda D W, Koric S, Sobh N A and Sehitoglu H 2021 Deep learning for plasticity and thermo-viscoplasticity Int. J. Plast. 136 102832
Turkoglu M O, D’Arongo S, Wegner J and Schindler K 2021 Gating revisited: deep multi-layer RNNs that can be trained IEEE Trans. Pattern Anal. Mach. Intell. 1
Velickovic P, Cucurull G, Casanova A, Romero A, Lio P and Bengio Y 2017 Graph attention networks arXiv Preprint (arXiv:1710.10903)
Li Y, Tarlowski J, Brockschmidt M and Zemel R 2015 Gated graph sequence neural networks arXiv Preprint (arXiv:1511.05963)
Ma T, Chen J and Xiao C 2018 Constrained generation of semantically valid graphs via regularizing variational autoencoders Adv. Neural Inf. Process. Syst. 31 7113–24
Dai M, Demirel M F, Liang Y and Hu J-M 2021 Graph neural networks for an accurate and interpretable prediction of the properties of polycrystalline materials npj Comput. Mater. 7 1–9
Cheng J, Zhang C and Dong L 2021 A geometric-information-enhanced crystal graph network for predicting properties of materials Commun. Mater. 2 1–11
Cecen A, Dai H, Yahansu Y C, Kalidindi S R and Song L 2018 Material structure–property linkages using three-dimensional convolutional neural networks Acta Mater. 146 76–84
Hornik K, Stinchcombe M and White H 1989 Multilayer feedforward networks are universal approximators Neural Netw. 2 359–66
Pearl G, Shen Y, Jing L, Yang Y, Cano-Renteria F, DeLacy B G, Ioannopoulos J D, Tagmarg M and Solajačič M 2018 Nanophotonic particle simulation and inverse design using artificial neural networks Sci. Adv. 4 eaar4206
Adam S, Halimah M, Chan K, Muhammad F, Nazrin S, Scavino E, Kamaruddin S, Az’lina A and Ghani N 2022 Structural, prediction and simulation of elastic properties for tellurite based glass systems doped with nano and micro Eu2O3 particles via artificial neural network model J. Mater. Res. Technol. 17 586–600
Li H, Collins C, Tanha M, Gordon G J and Yaron D J 2018 A density functional tight binding layer for deep learning of chemical hamiltonians J. Chem. Theory Comput. 14 5764–76
Behler J and Parrinello M 2007 Generalized neural-network representation of high-dimensional potential-energy surfaces Phys. Rev. Lett. 98 146401
[144] Yadav U, Pathrudkar S and Ghosh S 2021 Interpretable machine learning model for the deformation of multiwalled carbon nanotubes Phys. Rev. B 103 035407
[145] Kiashahnejad Y, Zandehshahvar M, Abdollahramezani S, Hemmatyar O, Pourabolghasem R and Adibi A 2020 Knowledge discovery in nanophotonics using geometric deep learning Adv. Intell. Syst. 2 1900132
[146] Iten R, Metger T, Wilming H, Del Rio L and Renner R 2020 Discovering physical concepts with neural networks Phys. Rev. Lett. 124 010508
[147] So S, Mun J and Rho J 2019 Simultaneous inverse design of materials and structures via deep learning: demonstration of dipole resonance engineering using core–shell nanoparticles ACS Appl. Mater. Interfaces 11 24264–9
[148] Blanchard-Dionne A-P and Martin O J 2020 Teaching optics to a machine learning network Opt. Lett. 45 2922–5
[149] Murdoch W J, Singh C, Kumbier K, Abbasi-Asl R and Yu B 2019 Definitions, methods, and applications in interpretable machine learning Proc. Natl Acad. Sci. 116 22071–80
[150] Selle R, Brinxer T, Bayer T, Wollenhaupt M and Baumert T 2008 Modelling of ultrafast coherent strong-field dynamics in potassium with neural networks J. Phys. B 41 074019
[151] Palmieri A M, Kovlakov E, Bianchi F, Yudin D, Straupe S, Biamonte J D and Kulik S 2020 Experimental neural network enhanced quantum tomography npj Quantum Inf. 6 1–5
[152] Cortes C L, Adhikari S, Ma X and Gray S K 2020 Accelerating quantum optics experiments with statistical learning Appl. Phys. Lett. 116 184003
[153] Iwasaki Y, Sawada R, Stanev V, Ishida M, Kirihaara A, Omori Y, Someya H, Takeuchi I, Saitoh E and Yorozu S 2019 Identification of advanced spin-driven thermoelectric materials via interpretable machine learning npj Quantum Mater. 5 1–6
[154] Cohn D A, Ghahramani Z and Jordan M I 1996 Active learning with statistical models J. Artif. Intell. Res. 4 129–45
[155] Gal Y, Islam R and Ghahramani Z 2017 Deep bayesian active learning with image data Int. Conf. on Machine Learning pp 1183–92
[156] Ducoffe M and Precioso F 2018 Adversarial active learning for deep networks: a margin based approach arXiv Preprint (arXiv:1802.09841)
[157] Pan S J and Yang Q 2009 A survey on transfer learning IEEE Trans. Knowl. Data Eng. 22 1345–59
[158] Qu Y, Jing L, Shen Y, Qiu M and Soljacic M 2019 Migrating knowledge between physical scenarios based on artificial neural networks ACS Photon. 6 1168–74
[159] Taylor R H, Rose F, Toher C, Levy O, Yang K, Nardelli M B and Curtarolo S 2014 A RESTful API for exchanging materials data in the AFLOWLIB.org consortium Comput. Mater. Sci. 93 178–92
[160] Saal J E, Kirkin S, Aykol M, Meredig B and Wolverton C 2013 Materials design and discovery with high-throughput density functional theory: the open quantum materials database (OQMD) JOM 65 1501–9
[161] MatWeb L 2014 Material property data. MatWeb ed
[162] Jain A, Ong S P, Hautier G, Chen W, Richards W D, Dacek S, Cholia S, Gunter D, Skinner D and Ceder G 2013 Commentary: the materials project: a materials genome approach to accelerating materials innovation APL Mater. 1 011002
[163] Groom C R, Bruno I J, Llightfoot M P and Ward S C 2016 The cambridge structural database Acta Crystallogr. B 72 171–9
[164] Cresci Y 2012 Pattern Recognition and Machine Learning (Amsterdam: Elsevier)
[166] Liu C-X, Yu G-L and Zhao G-Y 2019 Neural networks for inverse design of phononic crystals AIP Adv. 9 085223