CHAPTER ONE

# Artificial neural networks and their applications in computational materials science: A review and a case study

**Shaoping Xiao[a,*], John Li[b], Stéphane Pierre Alain Bordas[c], and Tae-Yeon Kim[d]**

[a]Department of Mechanical Engineering, Iowa Technology Institute, University of Iowa, Iowa City, IA, United States
[b]Department of Computer Science, Northwestern University, Evanston, IL, United States
[c]Department of Engineering, University of Luxembourg, Luxembourg, France
[d]Civil Infrastructure and Environmental Engineering, Khalifa University of Science and Technology, Abu Dhabi, United Arab Emirates
*Corresponding author. e-mail address: e-mail address: shaoping-xiao@uiowa.edu

## Contents



*Advances in Applied Mechanics*, Volume 57
ISSN 0065-2156, https://doi.org/10.1016/bs.aams.2023.09.001
Copyright © 2023 Elsevier Inc.
All rights reserved.

## Abstract

Current advances in artificial intelligence (AI), especially machine learning and deep learning, provide an alternative approach to problem-solving for engineers and scientists in various disciplines, including materials science. Artificial neural networks (ANNs), including their variations as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), have become one of the most effective machine learning approaches. This paper comprehensively reviews ANNs and their applications in different computational materials science research topics, such as multiscale modeling, microstructure-dependent material properties, and model-free constitutive relationships. In addition, we intend to share AI insights in the materials science community and promote the applications of ANNs in our research.

## 1. Introduction

Considerable effort has been dedicated to developing next-generation materials (Peng et al., 2017; Ray & Cooney, 2018) and structures for use in multiscale and multi-physical problems. In particular, advanced materials promote such platforms by coupling the predominant material properties to create multifunctional composites with enhanced mechanical, thermal, and other material capabilities. However, understanding such complex phenomena is highly dependent on systematic and accurate estimations of the effective physical properties, if possible. Therefore, rapid advancement requires numerical modeling and simulations capable of quickly and accurately determining such properties. Furthermore, computation has assisted the materials science community in various achievements as an important discovery tool, including rapid process development, quick microstructural analysis, fast property evaluation, and significant performance improvement.

Traditional computational methods have been extensively used to study physical phenomena at different length and time scales independently (Attarian & Xiao, 2022). Those methods include finite element methods (FEMs) (Belytschko, Liu, & Moran, 2000), meshfree particle methods (MPM) (Li & Liu, 2002; Rabczuk, Belytschko, & Xiao, 2004), phase-field methods (PFM) (Boettinger et al., 2003), molecular dynamics (MD) (Ghaffari, Zhang, & Xiao, 2017; Samanta et al., 2019), and quantum mechanics (QM) (Griffiths, 1995). In addition, many other advanced numerical methods have been developed recently. Peridynamics, introduced by Silling (Silling & Lehoucq, 2010; Silling, 2000), is a nonlocal integral-type numerical method for continuum mechanics. Notably, the internal forces in the governing equations of peridynamics are

calculated via integrations instead of derivatives. As a result, this method can directly handle spatial discontinuities. It has been successfully applied to fracture mechanics (Bobaru & Zhang, 2015; Silling & Askari, 2014), as well as the studies of plastic deformation (Madenci & Oterkus, 2016), composite materials (Yaghoobi & Chorzepa, 2017; Tuhami & Xiao, 2022), and heterogeneous materials (Jung & Seok, 2016). Besides bond-based peridynamics (Ghaffari et al., 2019) mentioned above, state-based peridynamics (Silling et al., 2007; Silling, 2010) has also been developed. Another recently developed method is the lattice element method (Rizvi, Nikolić, & Wuttke, 2019), a numerical method that investigates rock materials' fracture without predefining a crack path.

Multiscale modeling (Tadmor & Miller, 2011) is an efficient approach to studying the physical phenomena of materials when considering the interactive effects between multiple spatial and temporal scales. Early development focused on the architecture of either hierarchical (i.e., sequential) or concurrent multiscale methods. Concurrent multiscale methods (Wagner & Liu, 2003; Xiao & Belytschko, 2004; Xiao & Hou, 2007a, 2007b; Xiao et al., 2008; Miller & Tadmor, 2009; Rahman et al., 2017; Tadmor & Miller, 2017) employ an appropriate model to couple multiple length/time scales so that simulations at different scales are conducted simultaneously. Most of the developed concurrent multiscale methods are atomistic/continuum coupling methods, in which the molecular model is overlapped with the continuum model. However, the scale-coupling or scale-overlapping challenge in concurrent multiscale methods doesn't exist in hierarchical approaches (Tadmor, Phillips, & Ortiz, 2000). Indeed, researchers pay more attention to passing information between scales. Homogenization (Arroyo & Belytschko, 2003; Ericksen, 1984; Xiao & Yang, 2005, 2006; Xiao, Andersen, & Yang, 2008; Yang & Xiao, 2008), including the RVE techniques (Ghaffari, Zhang, & Xiao, 2018; Grabowski et al., 2017; Subramanian, Rai, & Chattopadhyay, 2015), has been commonly employed to obtain effective material properties to bridge various scales. The current state-of-the-art multiscale methodologies can be found in several review papers (Budarapu & Rabczuk, 2017; Gooneie, Schuschnigg, & Holzer, 2017; Kanouté et al., 2009).

Riding the current wave of artificial intelligence (AI), many disciplines, especially robotics and control (Cai, Hasanbeig, et al., 2021; Cai et al., 2021; Zhu et al., 2022), have applied learning-based approaches. Notably, the data-driven approach has become another powerful tool in scientific discoveries and engineering problem-solving (Versino, Tonda, & Bronkhorst, 2017). One of its new paradigms in materials science is discovering new materials

(Wang et al., 2022) or improving material designs (Pollice et al., 2021) based on the knowledge extracted from extensive materials datasets. Himanen et al. (2019) addressed data-driven materials science's status, challenges, and perspectives. Their review focused on materials data infrastructures and discussed several critical challenges in developing a material search tool. Tripathi, Kumar, and Tripathi (2020) presented big data models for material science data management and feature preservation in another survey. They also reported several challenges in big data analysis, such as data privacy, data preprocessing, and predictive algorithms.

Machine learning (ML) (Mitchell, 1997) is an approach using statistical models to analyze data and draw inferences from its pattern. Particularly, supervised learning models learn the relationship between the input features and the output targets without explicit instructions. As a subset of ML, deep learning (DL) (Schulz & Behnke, 2012) employs artificial neural networks (ANNs) to find appropriate representations from data progressively for good performance. Zhang and Friedrich (2003) presented one of the first reviews on predicting specific material properties of polymer composites by using neural networks. According to their review, a few early works have been conducted to predict fatigue life (El Kadi & Al-Assaf, 2002), tribological properties (Rutherford et al., 1996), and some other mechanical behaviors (Zhang, Klein, & Friedrich, 2002). Neural networks were also used for composite processing optimization (Heider, Piovoso, & Gillespie, 2003) and design optimization (Ulmer II et al., 1998). In another work, Kadi (El Kadi, 2006) summarized the implementation of ANNs in the mechanical modeling of fiber-reinforced composite materials, including static deformation and failure behaviors (Olivito, 2003), creep behavior (Al-Haik, Al-Haik, Garmestani, & Savran, 2004), delamination (Valoor & Chandrashekhara, 2000), crack and damage detection (Bar, Bhat, & Murthy, 2004), impact (Chandrashekhara, Okafor, & Jiang, 1998), and vibration control (Smyser & Chandrashekhara, 1997). Kadi also reviewed the applications of fuzzy ANN in studies of damage and failure in composite materials (Jarrah, Al-Assaf, & Kadi, 2002).

Recently, the applications of ML and DL have caught more and more attention from researchers in the materials science community, and quite a few updated reviews and discussions have been reported. Rodrigues et al. (2021) proposed a roadmap for future research focusing on ML-aided discovery of new materials and analysis of chemical sensing compounds. They also elaborated on the conceptual and practical limitations when applying big data and ML to materials science research topics. Morgan and Jacobs (2020) reviewed some common types of ML models in materials

science and addressed the breadth of opportunities and the best practices for their usage. Another recent work (Choudhary et al., 2022) reviewed the applications of DL in atomistic simulation and material imaging.

This paper aims to provide an in-depth review of ANNs, including physics-informed neural networks (PINN), convolutional neural networks (CNNs), and recurrent neural networks (RNNs), and their applications in computational materials science and engineering. We will focus on several advanced research topics, such as multiscale modeling, forward and inverse problems, microstructure-dependent material property prediction, and model-free constitutive identification. In addition, we use a case study to demonstrate the applications of neural networks in studying the material failure of metal-ceramic spatially tailored materials. This paper also intends to share AI insights in the materials science community and promote the applications of ANNs in our research.

## 2. Artificial neural networks

### 2.1 Basics of artificial neural networks

A typical ANN (Dreiseitl & Ohno-Machado, 2002), shown in Fig. 1, usually consists of an input layer, an output layer, and one or more hidden layers. This kind of neural network is fully-connected because every neuron connects all the neurons on the previous and subsequent layers. For example, we consider a data set of $N$ distinct training samples $(x_I, y_I)$ where $I \in [1, N]$. Each data sample has $p$ input features ($x_I \in R^p$) and $q$ outputs ($y_I \in R^q$).

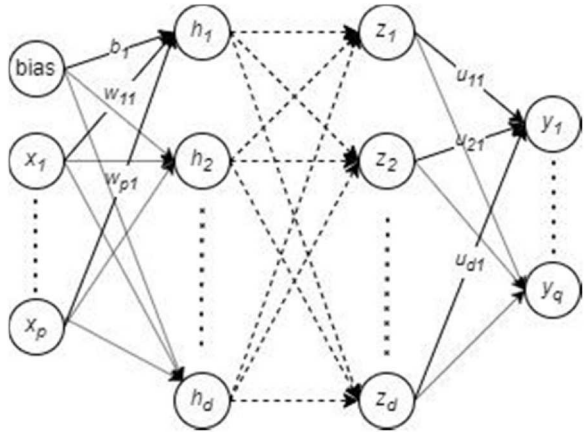


**Fig. 1** An artificial neural network (ANN).

Therefore, the corresponding neural network to approximate the relation between the input and the output of the data set has $p + 1$ neurons on the input layer and $q$ neurons on the output layer. The hidden layers can have various numbers of neurons, and it is assumed that each hidden layer has $d$ neurons in Fig. 1.

The neural network training includes a feedforward process and a backpropagation process. During the feedforward process, every neuron in the hidden layers transforms the outputs from the previous layer into a different representation, the input to the next layer. There are two steps in the transformation. For example, the input data is projected into the first hidden layer via the weights, $\boldsymbol{w}$, and biases, $\boldsymbol{b}$. Then, the projected outcome is transformed via the activation function $\varphi$, also called the transformation function. Mathematically, this transformation can be expressed as

$$h_j = \varphi_j\left(w_j^T x + b_j\right) = \varphi_j\left(\sum_{i=1}^{p} w_{ij} x_i + b_j\right) j = 1 \ldots d \tag{1}$$

It is known that the hidden layer is not limited to having only one type of activation function in neurons. There are a variety of activation functions available, and most of them are nonlinear functions, including sigmoid, hyperbolic tangent, and radial basis functions (RBFs). Particularly, the RBF neurons use distances between samples and centroids as inputs, and $L^1$, $L^2$, or $L^\infty$ norms of distances can be used.

If the last hidden layer has the output $z$, as shown in Fig. 1, the estimated $k$ th output of the $I$th training sample is calculated as

$$\tilde{y}_{Ik} = \varphi_o\left(u_k^T z\right) = \varphi_o\left(\sum_{i=1}^{d} u_{ik} z_i\right) \; k = 1 \ldots q \tag{2}$$

where $u$ are weight coefficients. $\varphi_o$ is the activation function for outputs, and it is usually an identity function for regression problems. A loss function is calculated based on the estimated output targets and the actual outputs to evaluate the neural network's performance. It is also called the data loss function as

$$L_d = L_d\left(y_1, y_2 \ldots y_N, \tilde{y}_1, \tilde{y}_2, \ldots \tilde{y}_N\right) \tag{3}$$

Indeed, training a neural network becomes an optimization problem to find appropriate weight coefficients, including $w$ and $u$, for minimizing the loss function. This is usually done using the gradient descent method or its variations in the backpropagation process.

## 2.2 Physics-informed neural networks

Training a PINN (Raissi, Perdikaris, & Karniadakis, 2019) needs a training data set and a physical-based mathematical model, i.e., partial differential equations (PDE), shown in Fig. 2. Without a loss of generality, we assume a system of PDEs with appropriate initial and boundary conditions as below.

$$\begin{aligned} &\frac{\partial \gamma(x,\ t)}{\partial t} + D[y;\ \gamma] = 0 \\ &\text{s. t. } \mathrm{y}(\mathrm{x},\ 0) = \mathrm{y}_0(x), \quad \gamma(0,\ t) = \underline{\gamma}_0(t), \quad \gamma(L,\ t) = \underline{\gamma}_L(t) \end{aligned} \tag{4}$$

where $D$ is a nonlinear operator parametrized by $\gamma$.

The training data can be collected by numerically solving PDEs in Eq. (4). It shall be noted the solutions are on the discretized spatial and temporal grids. Therefore, in addition to the input features, the neural network in Fig. 2 may also take time and spatial coordinates on the input layer. In most existing works of PINN, the fully-connected neural work was employed. Therefore, the output targets $\tilde{\gamma}_I$ can be predicted as described in Eq. (2) via the feedforward process. Such approximations to $\gamma$ result in not only the data loss function but also the residual of PDEs as

$$r = \frac{\partial \tilde{\gamma}(x,\ t)}{\partial t} + D[\tilde{\gamma};\gamma]. \tag{5}$$

where the derivatives of $\tilde{\gamma}$, i.e., $D[\tilde{\gamma};\gamma]$, can be derived via the automatic differentiation approach. Consequently, the physics loss function is defined

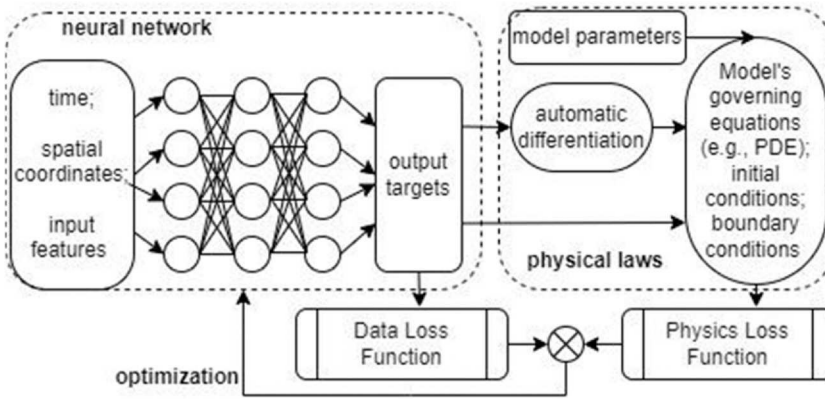


**Fig. 2** A physics-informed neural network (PINN).

below, considering residuals of governing equations and initial and boundary conditions.

$$L_p = L_p(|r|, |\tilde{y}(x, 0) - y_0(x)|, |\tilde{y}(0, t) - \underline{y}_0(t)|, |\tilde{y}(L, t) - \underline{y}_L(t)|) \tag{6}$$

Then, the total loss function, which is a combination of the data and physics loss functions, i.e., $L_d + L_p$, is implemented in the backpropagation process to optimize the neural network's weight coefficients.

It can be seen that the data loss function measures the difference between the actual outputs and their approximations predicted by the neural network. On the other hand, the physics loss function quantifies how close the input-output relationship approximated by the neural network follows the physical laws. Therefore, this neural network is physics informed. The concept of PINN has been employed in many disciplines, including computational mechanics and materials science, for both forward and inverse problems (Raissi et al., 2019).

## 2.3 Other neural networks

Other commonly used artificial neural networks include CNNs (Sainath et al., 2015) and RNNs (Schmidhuber, 2015), which mainly aim to handle image and time-series data, respectively. Convolution neural networks have been proven to be very effective and successful in image recognition and classification. Generally, an image can be represented as a matrix of pixel values, and a color image has three channels – red, green, and blue. Therefore, an image sample is indeed a three-dimensional tensor of shape, i.e., (height, width, and channels). To address a classification problem, a CNN usually has four main operations in the feedforward prediction process: convolution, nonlinearity, max pooling, and classification, as shown in Fig. 3.

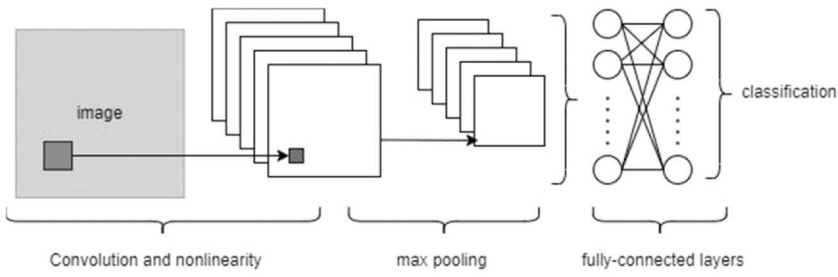


**Fig. 3** A convolutional neural network (CNN).

During the convolution and nonlinearity operations (in a convolution layer), the image is decomposed into overlapped image tiles via the sliding window search, i.e., stridden convolution. Each image tile is then fed into a small neural network with nonlinear activation functions. It shall be noted that the same neural network is applied for every single tile individually. The output feature map is a three-dimensional array with a smaller height and width than the original image. If the original image is a color one, it has a depth of three because of three channels. However, the resulting array has a depth the same as the neural network's filter number.

The next operation is max pooling, which is also called downsampling. This operation consists of extracting windows from the input feature maps (the output from the previous convolution layer) and outputting the max value of each filter. The convolution layer and the max pooling can repeat multiple times before reshaping the output feature map as a one-dimensional array to a "fully-connected" network for prediction.

Recurrent neural networks mimic the biological intelligence procedure that processes information incrementally while maintaining an internal memory (i.e., state) for past information. They are often employed for time-series data, and each data sample (i.e., time sequence) is encoded as a 2D tensor of size with time steps and input features. An RNN is a fully-connected neural network that has states. Instead of taking a data sample at one time, an RNN unit loops over time steps, as shown in Fig. 4.

At each time step $t$, the RNN considers the output from the previous time step $t$-1 as its current state, takes the $t$th input entry, and combines them to obtain the output at time $t$. After applying the activation function, the RNN sets the output as the state for the next time step. Such a recursive update can be written as

$$y^t = \varphi\left(w^T x^t + u^T y^{t-1} + b\right) \tag{7}$$

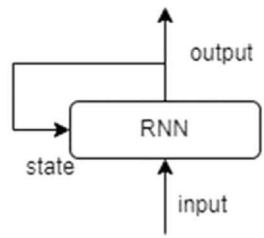


**Fig. 4** A classical recurrent neural network (RNN) unit.

where $\varphi$ is the activation function, $w$ and $u$ are weight coefficients, and $b$ is the bias. A few advanced RNN architectures have been proposed, including long-short term memory (LSTM) (Sutskever, Vinyals, & Le, 2014) and gated recurrent unit (GRU) (Cho et al., 2014).

## 3. Applications of neural networks

### 3.1 ANNs in multiscale modeling

Multiscale modeling and simulation have benefited from ML and DL methods (Alber et al., 2019), including fully connected ANNs. Le, Yvonnet, and He (2015) employed ANNs and proposed a decoupled computational homogenization method for nonlinear elastic materials. Their approach computed the training samples' effective potentials through random sampling in the parameter space. Then, ANNs were used to approximate the surface response and derive the macroscopic stress and tangent tensor components. In another work, Unger and Könke (Unger & Könke, 2009) adopted ANNs as material models in a multiscale approach to studying reinforced concrete beams. In another work, Liu, Wu, and Koishi (2019) developed a new data-driven multiscale method, i.e., a deep material network, to describe complex overall material responses of heterogeneously structured composites. They also simulated the macroscale dynamics of gas-solid mixtures by employing information collected from microscale simulations via an ANN model (Lu et al., 2012). White and co-workers (White et al., 2019) used a single-layer feedforward neural network as a surrogate model to predict the elastic response of the microscale metamaterial during the optimization of macro-scale elastic structures. Other achievements include a multiscale multi-permeability poroplasticity model (Wang & Sun, 2018), a 3D architecture of a deep material network (Liu & Wu, 2019), and ANN-assisted multiscale analysis (Balokas, Czichon, & Rolfes, 2018).

In addition, Xiao et al. (2020) proposed an alternative data-driven approach by using neural networks to pass information from the molecular model to the continuum model in a hierarchical multiscale framework. First, intensive molecular dynamics (MD) simulations were conducted to generate the dataset in which the input features included strains and temperature, while the output targets were stress components and material failure mode. Then, the generated data was used to train several DL classification and regression models. Indeed, one of the important emergent ML techniques, extreme learning machines (ELMs) (Huang, Zhu, & Siew,

2006; Guang-Bin Huang et al., 2012) were adopted. It has been shown that an ELM is a fast training method for Single-Layer Feed-forward Networks (SLFNs). An SLFN has three layers of neurons. The term "Single" stands for the only layer of linear/nonlinear neurons in the model and is the hidden layer. Finally, the well-trained learning machines were directly implemented in continuum simulations to predict material failure mode and stress components. In this approach, as shown in Fig. 5, neither constitutive relations nor effective material properties were explicitly derived as achieved in existing hierarchical multiscale methods. The learning machines served as "black boxes" to replace constitutive relations and failure mode decisions in the continuum simulations. Such "black boxes" were trained based on the dataset from molecular simulations; therefore, the proposed scheme is physical-based and data-driven.

Xiao et al. (2021) and Tuhami and Xiao (2022) have extended the above ML-based multiscale framework to study the mechanics of spatially tailored materials (STMs) via FEMs or peridynamics. Spatially tailored materials (Birman et al., 2008), also named functionally graded materials

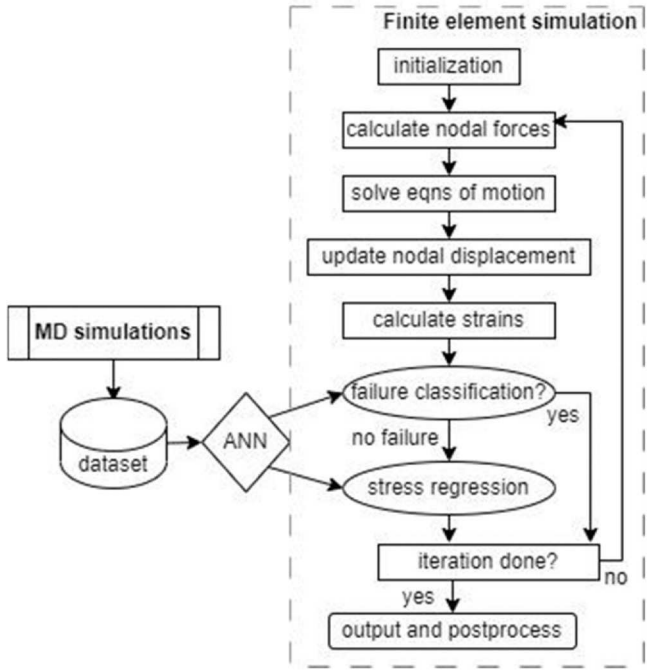


**Fig. 5** Hierarchical multiscale modeling enhanced by machine learning.

(FGMs), are one of the next-generation composites for use in multi-physical problems. They are essentially composites consisting of two or more phases of distinct materials in which the volume fractions continuously change in space. This unique class of heterogeneous composites offers advantages over traditional composites due to its ability to leverage the predominant characteristics of the constituent materials and to tailor the effective material properties according to the loading conditions and operating temperatures. A metal-ceramic STM was considered in Ref. Xiao et al. (2021) and studied via a hierarchical multiscale method from micro to macro scales. Microstructure uncertainties, including particle number, size, shape, and location, were considered during data collection via microscale simulations. After being trained via the collected dataset, the ANNs for material property and failure predictions were implemented in macroscale simulations.

## 3.2 PINNs in forward and inverse problems

Physics-informed neural networks have been adopted in various disciplines, including computational materials science. For example, Haghighat et al. (2020) demonstrated the application of PINN in solving a two-dimensional linear elasticity problem with the following governing equations:

$$\frac{\partial \sigma_{ij}}{\partial x_j} + f_i = 0 \tag{8}$$

where $\sigma$ is the Cauchy stress tensor, $f$ is the body force vector, and $i, j = 1,2$. The constitutive model is

$$\sigma_{ij} = \lambda \delta_{ij} \varepsilon_{kk} + 2\mu \varepsilon_{ij} \tag{9}$$

where $\delta_{ij}$ is the Kronecker delta, and $\lambda$ and $\mu$ are material constants. $\varepsilon_{ij}$ is the infinitesimal strain tensor and can be calculated as

$$\varepsilon_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) \tag{10}$$

where $u(x)$ is the displacement field.

The FEM solutions were used as the training dataset in their approach. Each data sample had the coordinates $(x)$ as the input features, and the output targets included $u_1(x)$, $u_2(x)$, $\sigma_{11}(x)$, $\sigma_{22}(x)$, and $\sigma_{12}(x)$. Various ANNs were employed/trained to predict each output variable individually. In addition to the data loss, the physics loss function for this specific

problem was written below based on the governing equations and the constitutive model.

$$L_p = \left|\frac{\partial\tilde{\sigma}_{11}}{\partial x_1} + \frac{\partial\tilde{\sigma}_{12}}{\partial x_2} + f_1\right| + \left|\frac{\partial\tilde{\sigma}_{12}}{\partial x_1} + \frac{\partial\tilde{\sigma}_{22}}{\partial x_2} + f_2\right| + |(\lambda+2\mu)\varepsilon_{11} + \lambda\varepsilon_{22} - \tilde{\sigma}_{11}| + |(\lambda+2\mu)\varepsilon_{22} + \lambda\varepsilon_{11} - \tilde{\sigma}_{22}| + |2\mu\varepsilon_{12} - \tilde{\sigma}_{12}| \tag{11}$$

where the output variables with tilde were predicted from ANNs. The other variables, including $\varepsilon_{11}$, $\varepsilon_{22}$, $\varepsilon_{12}$, $f_1$, and $f_2$ can be obtained from Eqs. (8) and (10) as below, through automatic graph-based differentiation (Güne et al., 2018).

$$\varepsilon_{ij} = \frac{1}{2}\left(\frac{\partial\tilde{u}_i}{\partial x_j} + \frac{\partial\tilde{u}_j}{\partial x_i}\right) \text{ and } \quad f_i = -\frac{\partial\sigma_{ij}}{\partial x_j} \tag{12}$$

Then, the PINN's weight coefficients can be updated via gradient descent approaches during the backpropagation process. Haghighat et al. (2020) also demonstrated that PINNs could be used to identify the model parameters $\lambda$ and $\mu$, as solving an inverse problem (Raissi et al., 2019).

In another work, Zhang, Yin, and Karniadakis (2020) extended PINN to solve identification problems of nonhomogeneous materials. In this inverse problem, they sought to identify soft tissue material properties based on the full-field displacement measurements under quasi-static loading. In addition to a PINN employed to approximate the unknown material parameters, another PINN was utilized to solve the corresponding forward problem. Two PINNs were trained simultaneously, as shown in Fig. 6, and the physics loss function was formulated according to the displacement boundary conditions, the incompressibility constraints, traction boundary conditions, and the governing equations (i.e., PDEs).

Recently, Zhang and Gu (2021) extended PINNs into digital material problems, in which a composite, as a 3D-printable material, was considered an assembly of material voxels. They addressed a few challenges, including discontinuous material properties, nonlinear strain due to large deformation, and neural network accuracy. In digital material design, a material configuration is generally a combination of step functions, and its derivatives are often not available. To address this challenge, they adopted the minimum energy criteria as the loss function of a PINN other than the strong governing equations (i.e., PDEs). This energy-based PINN

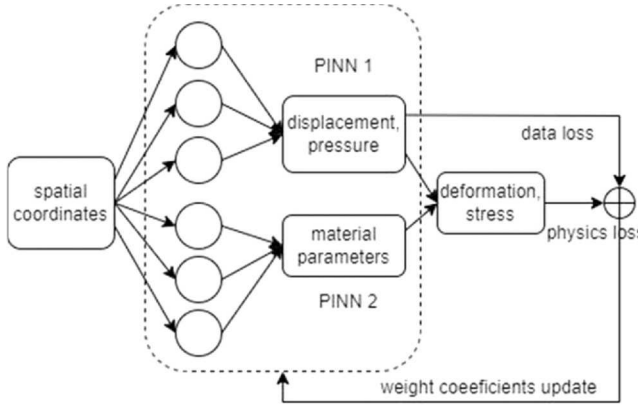


**Fig. 6** Two PINNs were used in (Zhang et al., 2020) for an inverse problem.

achieved comparable accuracy to supervised ML models. In addition, adding a hinge loss for the jacobian could enhance the proposed PINN to approximate nonlinear strain properly.

Physics-informed neural networks were also adopted in studying crack propagation of quasi-brittle materials under complex loading (Zheng et al., 2022). In this work, the researchers used PINN to reconstruct the solution of the displacement field. Without labeled data (i.e., a priori information), the PINN based on the energy minimization principle could predict crack propagation with incremental loading pattern and damage evolution model. In addition, they introduced the domain decomposition method and the finite basis algorithm to address complex boundaries and the gradient pathology problem, respectively. Therefore, their approach is robust.

Furthermore, Shukla et al. (2020) applied PINN to identify and characterize a surface-breaking crack in an aluminum alloy substrate material. PINN was supervised with realistic ultrasonic surface acoustic wave data, representing deformation on the top surface of the aluminum plate. It was physically informed by the ultrasonic surface acoustic wave equation as an inverse problem, i.e., the estimation of the unknown wave speed for given acoustic wave data. In other words, the spatially varying surface wave speeds were used as markers of characterizing/identifying the surface-breaking crack in the aluminum alloy substrate. Moreover, they used adaptive activation functions in training to accelerate convergence significantly, even with highly noisy data. Using a small portion (i.e., 10–20% of the total data) of the wave data, PINN accurately predicted the wave

speed affected by the crack in the substrate, verifying the efficiency of the PINN by reducing the cost and time of the data acquiring process.

## 3.3 CNNs in microstructure quantification

Since CNNs have been successfully employed in image processing, they were also recently used in computational materials science to quantify the microstructure of materials (especially composite materials) for predicting material properties. Previous approaches (Torquato & Haslach, 2002) utilized n-point spatial correlations (Kröner, 1977), which could effectively quantify the local neighborhoods in material microstructure as features to measure material properties. However, the number of features could be practically infinite when considering the complete set of possible local neighborhood configurations.

One of the pioneering works (Cecen et al., 2018) aimed to address this core challenge in constructing material process-structure-property linkages for new material design and improvement by using CNNs. In their approach, a 3D CNN was employed to learn the salient features of the material microstructures for material property predictions as a regression problem. Specifically, they collected 5900 microstructure images and conducted finite element simulations to calculate material properties for each microstructure. Each microstructure image consisted of $51 \times 51 \times 51$ cuboidal voxels and was convolved with 32 filters. The filter size was $10 \times 10 \times 10$, which was informed by spatial statistics. The rectifier function was used as the activation function. It shall be noted that average pooling was used instead of the max pooling as in conventional CNNs. After getting 256 features, linear regression was conducted to estimate the effective material property.

In a similar work, Rao and Liu (2020) proposed a three-dimensional CNN, shown in Fig. 7, as a homogenization surrogate model to predict the anisotropic effective material properties for microscale RVEs with random inclusions. A high-fidelity dataset was generated by using FEM simulations, and the trained CNN was capable of capturing the microstructural features of RVEs. Instead of predicting each material property with individual CNNs, a single 3D-CNN was employed to estimate all material stiffness and position components for the studied heterogeneous composites. They also discussed uncertainty quantification and the model's transferability. In another similar work (Mianroodi et al., 2022), a CNN took the nanostructured configurations as input and predicted the corresponding elasticity tensor. Other CNN-enhanced modeling and simulations include stress field

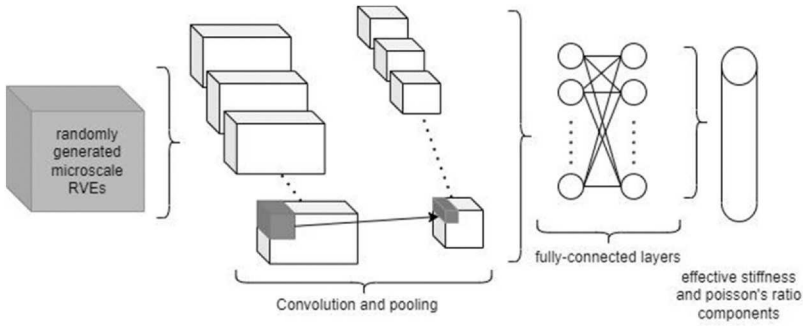


**Fig. 7** A 3D-CNN (Rao & Liu, 2020) for estimating effective anisotropic material properties.

prediction in composites (Bhaduri, Gupta, & Graham-Brady, 2022) and optimizing material structures (Yilin, Fuh Ying Hsi, & Wen Feng, 2021).

During the data collection in approaches similar to the above-mentioned works (Cecen et al., 2018), computationally expensive physics-based simulation tools (e.g., FEM) were usually employed to calculate the effective material properties for given microstructures. Such a forward model may not be practical for microstructure design problems in which the target properties are usually achieved via an iterative process. Mann and Kalidindi (2022) combined the microstructure-sensitive design (MSD) framework with the CNN-based surrogate model to reduce the computational cost. They introduced the microstructure hull concept in MSD and used 2-point spatial correlation maps as inputs in the CNN. Such a proposed strategy made exploring the complete search space of possible material properties feasible.

## 3.4 RNNs in material constitutive identification

Many previous works employed ANNs (Akbari, Mirzadeh, & Cabrera, 2015; Sabokpa et al., 2012; Singh, Rajput, & Mehta, 2016) to approximate constitutive modeling at various strain rates and high temperatures. However, most nonlinear constitutive relationships are stress/strain-history dependent. In other words, the predicted stress depends on the current deformation and the deformation history. More input features that can capture stress/strain history need to be included to address this issue. Furukawa and Hoffman (2004) used an ANN in FEM to conduct cyclic plastic analysis. The inputs in their neural networks included the current and two previous states of the back stress and inelastic strain, while the output was the increment of the back stress. Similar works include a hybrid

multilayer perceptron neural network approach to describe the statistical scatter of cyclic stress-strain curves (Janežič, Klemenc, & Fajdiga, 2010).

Since RNNs have advantages over ANNs in analyzing time-series data, they have been a better alternative approach to approximate stress-strain-time relationships. An early work done by Oeser and Freitag (2009) utilized an RNN with the input of previous load (stress) increments to estimate the new displacement (strain) increment for the studies of rheological materials. It shall be noted that the training data was collected from creep simulations. In another work, Freitag, Graf, and Kaliske (2013) employed RNNs as a complete or part of the material model in fuzzy FEM simulations. They assumed that the material parameters in numerical simulations were uncertain. Time-dependent stress and strain were modeled as fuzzy processes (Möller & Beer, 2004), which were discretized via time discretization and α-level discretization. Then, an RNN was developed to map fuzzy strain processes onto fuzzy stress processes, and it was applied to describe time-dependent material behaviors in FEM simulations.

Zopf and Kaliske (2017) coupled an RNN to the micro-sphere approach to approximate the model-free characterization of elastic and inelastic materials, including uncured natural rubber. Such coupling could address the issue that material testing cannot cover the complete stress state space. Therefore, reliable training of the proposed neural network was achieved. Recently, Stöcker et al. (2022) proposed a novel training algorithm for RNN to be more robust. The algorithm could generate adversarial examples based on the neural network prediction errors, i.e., the loss function. Consequently, the neural network could yield reliable material representations even when providing perturbed inputs. Specifically, GRU, an advanced RNN, was used in their approach.

## 4. A case study

### 4.1 Metal-ceramic spatially tailored composite materials

In this case study, we consider metal-ceramic STMs to generate the dataset and then employ various neural networks to predict material properties and mechanical behaviors. The STM materials are Ti(Ti-6Al-4V)-$TiB_2$ composites, and material properties are listed in Table 1 (Wiley, Manning, & Hunter, 1969; American Society for Metals., 1979; Munro, 2000). We only consider the room temperature (20 °C) in this case study.

**Table 1** Material properties of Ti (Ti-6Al-4V) and $TiB_2$ at 20 °C.

| | Young's modulus $E$ (GPa) | Poisson's ratio $v$ | Density $\rho$ ($kg/m^3$) | Tensile strength $\sigma_t$ (GPa) |
|---|---|---|---|---|
| Ti (Ti-6Al-4V) | 106.2 | 0.298 | 4357 | 1.17 |
| $TiB_2$ | 495.4 | 0.100 | 4505 | 3.73 |

We use the ceramic volume fraction (CVF) to represent the composition of ceramic in the studied STMs, which can be modeled as continuously variable composition materials. Particularly, the CVF can be expressed as a function of coordinates to indicate the difference between metal and ceramic at a particular spatial location in the STM. For example, in a Ti-$TiB_2$ plate, if the volume fraction varies along with the plate thickness, it can be written below as a power-law distribution.

$$v_f(z) = v_0 + (v_1 - v_0)\left(\frac{z}{h}\right)^n \tag{13}$$

where $h$ is the total thickness, and $z$ is the thickness coordinate between two surfaces $z = 0$ and $z = h$. In addition, $v_0$ and $v_1$ are the CVFs at two surfaces. It shall be noted that $n$ is the control parameter for the ceramic content distribution. If choosing $n = 1$, there is a linear distribution of ceramic volume fraction along the thickness. On the other hand, if $n = 2$, a nonlinear (i.e., quadratic) distribution exists, as shown in Fig. 8. The metal (Ti) is the matrix material when the CVF is less than 50%. After the CVF exceeds 50%, the matrix material switches to ceramic ($TiB_2$).

There may be more than one directional material variation in STM structures. Generally, Eq. (13) can be revised below for a two-dimensional graded material with a rectangular geometry.

$$v_f(x, y) = v_0 + (v_1 - v_0)\left[\eta_x\left(\frac{x}{w}\right)^{n_x} + \eta_z\left(\frac{y}{h}\right)^{n_y}\right] \tag{14}$$

where $x$ and $y$ are the coordinates of arbitrary material points taking one corner of the rectangle as the origin, $w$ and $h$ are the total width and height, and $\eta_i$ and $n_i$ are parameters controlling the ceramic content and profile in each direction. $v_0$ and $v_1$ are the minimum and maximum CVFs in the material, and they are usually 0 and 1, respectively.

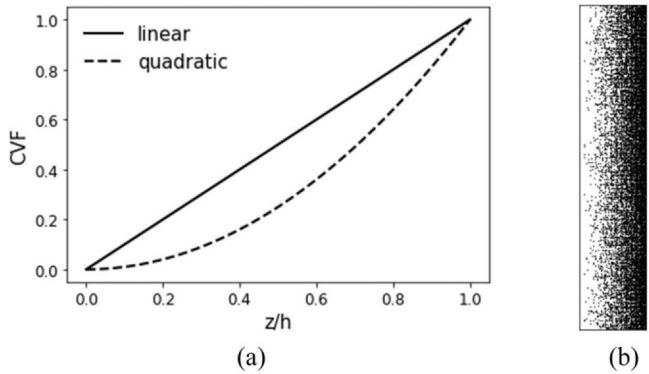


**Fig. 8** Metal-ceramics STMs: (A) CVF distributions and (B) an STM plate.

## 4.2 Peridynamics

In this case study, we will use bond-based peridynamics (Silling & Lehoucq, 2010; Silling, 2000) to model and simulate STMs at the microscale. Peridynamics is a nonlocal model of classical continuum mechanics. However, the governing equations, i.e., the partial differential equations, are reformulated below by replacing the derivative terms with volume integrals of force densities.

$$\rho \ddot{u}(x, t) = \int_{H_x} f(\eta, \xi, t)\, dV_{x'} + b(x, t) \tag{15}$$

where $\rho$ is the material density, $\ddot{u}$ is the second time derivative of the displacement vector, i.e., the acceleration vector, and $\mathbf{b}$ is the body force density vector. In bond-based peridynamics, the simulation domain is discretized into material points with finite volumes. The pairwise force density vector, $f$, corresponds to the deformation of bonds between point $x$ and material points in its horizon, i.e., $x' \in H_x$. $\xi$ and $\eta$ represent the relative position vector and the relative displacement vector, respectively, and they can be defined below, as shown in Fig. 9.

$$\xi = x' - x, \quad \eta = u(x', t) - u(x, t) \tag{16}$$

The force density can be calculated from bond strain $s$ and bond rotation $\gamma$ (Zhu & Ni, 2017) as

$$f(\eta, \xi, t) = c\, s\, n + \kappa\gamma \tag{17}$$

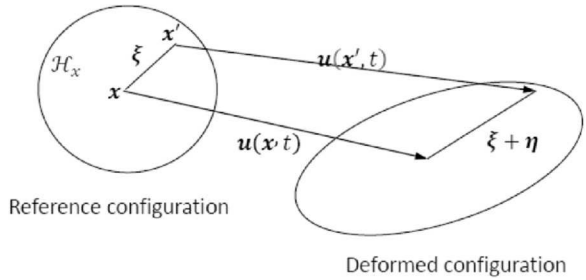


**Fig. 9** Reference and deformed configurations in bond-based peridynamics.

where

$$s(\eta, \xi) = \frac{(\|\eta + \xi\| - \|\xi\|)}{\|\xi\|} \tag{18}$$

$$n = (\eta + \xi)/\|\eta + \xi\| \tag{19}$$

$$\gamma(\eta, \xi) = \frac{1}{\|\xi\|}\eta \cdot (I - n \otimes n) \tag{20}$$

$c$ and $\kappa$ are the first and second micromoduli. Assuming an elastic material has Young's modulus $E$ and Poisson's ratio $\nu$, the micromoduli for plane strain problems can be derived as

$$c = \frac{6E}{\pi\delta^3(1 + \nu)(1 - 2\nu)}, \cdot k = \frac{6E(1 - 4\nu)}{\pi\delta^3(1 + \nu)(1 - 2\nu)} \tag{21}$$

It shall be noted that the metal-ceramic STMs have various CVFs spatially at the macroscale. However, it is assumed that the CVF is a constant at each macroscale material point. In other words, each microscale model of STM has a constant CVF. All microscale models are the two-dimension domain of 30 μm × 30 μm, assuming plane strain. The simulation domain is discretized with 2601 (51 × 51) material points for peridynamics. According to the CVF, the material points are randomly assigned as either metal or ceramic.

In addition, 1–2% porosity (Patil et al., 2019) is added to the microscale configuration by randomly selecting a small number of material points and removing them to generate vacancies. The number of vacancies follows the Poisson point distribution (Xiao et al., 2008). Fig. 10 illustrates two microscale STM configurations with 25% CVF and two with 50% CVF. In

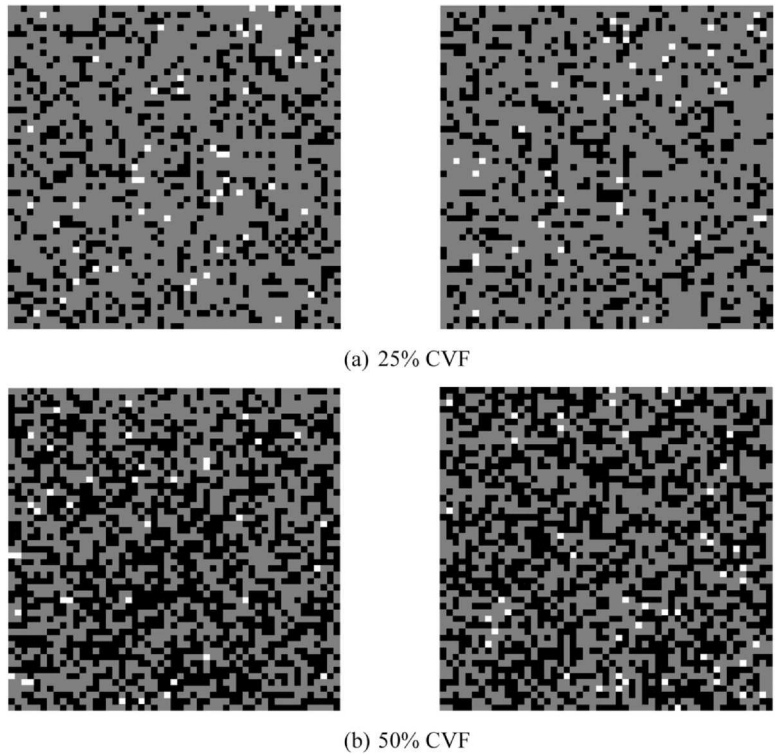


**Fig. 10** Microscale configurations of STM.

the microstructure configurations, gray represents metal, black represents ceramic, and white represents vacancy.

In the peridynamic model, we choose $\Delta x = \Delta y = 0.6\,\mu m$ and the horizon radius $\delta = 1.6\Delta x$. Since we consider metal-ceramic STMs in this study, there are three different types of bonds: metal-ceramic, metal-metal, and ceramic-ceramic. According to the material properties listed in Table 1, the first and second micromoduli can be calculated as in Table 2. They are calculated via Eq. (21) for metal-metal and ceramic-ceramic bonds. The prototype micro-elastic brittle (PMB) material model is employed, and the critical stretches, $s$, are computed based on the material tensile strength. In addition, the corresponding parameters for metal-ceramic bonds are derived via the combining rule. For example, the first micromodulus of metal-ceramic bonds is $c_{mc} = \sqrt{c_m c_c} = 622.55\frac{GPa}{\mu m^3}$.

**Table 2** First and second micromoduli in peridynamics models.

| | Metal-metal bond | Ceramic-ceramic bond | Metal-ceramic bond |
|---|---|---|---|
| $c\left(\frac{GPa}{\mu m^3}\right)$ | 326.01 | 1188.83 | 622.55 |
| $\kappa\left(\frac{GPa}{\mu m^3}\right)$ | 47.94 | 756.53 | 190.44 |
| $s$ | 0.011 | 0.008 | 0.0094 |

## 4.3 Data collection

The microscale configurations vary at the same CVF. Therefore, 25 simulations are conducted at each CVF between 1% and 99% at a 1% increment. The model is subject to the uniaxial tension in each simulation by applying the prescribed displacement on the top and fixing the bottom. Since the enforcement of boundary conditions cannot be directly applied to the boundary material points in peridynamics, fictitious walls (Ghaffari et al., 2019) are applied to eliminate such "edge softening" phenomena (Nishawala & Ostoja-Starzewski, 2017). A low strain rate is maintained, so the simulation results (e.g., strain-stress relations) can be used for quasi-static analyses. The stress is calculated by dividing the vertical component of the total bond force by the area of the middle cross-section. The Young's modulus is calculated from the slope of a strain-stress curve at 0.005% strain. The highest stress at the stress-strain curve is the tensile strength.

It is observed that the failure strength follows the Gaussian distribution at a particular CVF. Table 3 illustrates the statistical characteristics of failure strength at three different CVFs: 25%, 50%, and 75%. Obviously, the composite with a higher CVF has a higher mean failure strength and a larger standard deviation.

## 4.4 Deep learning predictive models

In the work of (Tuhami and Xiao, 2022), peridynamics was employed at the microscale to study the mechanical behaviors of metal-ceramic composites at various CVFs. The generated data were used to train multiple machine learning models to predict material properties, including Young's modulus, Poisson's ratio, and tensile strength, taking the CVF as the input. The well-trained predictive models were then implemented in continuum mechanics simulations at the macroscale. In those machine learning

**Table 3** Statistics of failure strength (GPa) at different CVFs.

| Ceramic volume fracture | 25% | 50% | 75% |
|---|---|---|---|
| Mean | 0.956 | 1.290 | 1.880 |
| Standard deviation | 0.048 | 0.053 | 0.076 |

models, fully-connected neural networks were employed. They took the CVF as the input feature and the material properties as the output target. Therefore, they were regression models. In this case study, we utilize the same way to generate the data. However, not only the CVF but also the image of the microstructure is used as the input feature. Consequently, in addition to ANN as a material failure classification model, CNN is utilized as a regression model to predict tensile strength.

### *4.4.1 Failure probability prediction via ANN*

During the data collection, 25 tensile simulations are conducted at each CVF, and various tensile strengths are obtained due to the microstructure uncertainties. In other words, although the composites have the same CVF, they fail at different strains. According to the strain-stress histories collected from microscale simulations, we can approximate the likelihood of failure occurrence at a particular strain. Consequently, we use a binary map to represent the failure probability in the strain-CVF space, as plotted in Fig. 11. The white domain indicates a 100% probability of material failure in the figure, while the black domain indicates non-failure. However, there is no clear decision boundary to distinguish failure/non-failure domains but rather a fuzzy interface, representing material failure with a probability between 0% and 100%.

Indeed, the corresponding dataset has 154,025 data samples in which the strain and CVF are in the input features while the output is material failure or non-failure, represented by "1" or "0", respectively. Instead of developing a binary classification model, we employ and train a fully connected ANN to predict the failure probability. The neural network comprises two hidden layers (128 and 64 neurons) with the "relu" activation function. In addition, the "sigmoid" activation function is employed in the output layer. Fig. 12 shows the material failure probability map that the ANN predictive model generated in the strain-CVF space. The gray area at the interface of failure/non-failure domains represents the change in the failure probability.

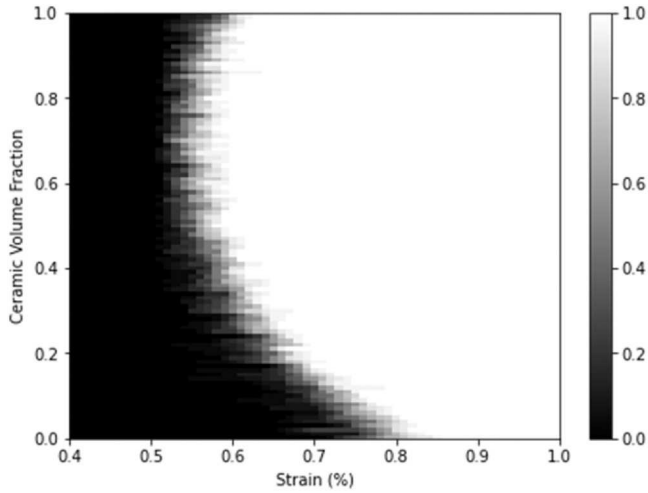


**Fig. 11** A binary map representing the failure likelihood approximated from the collected data.

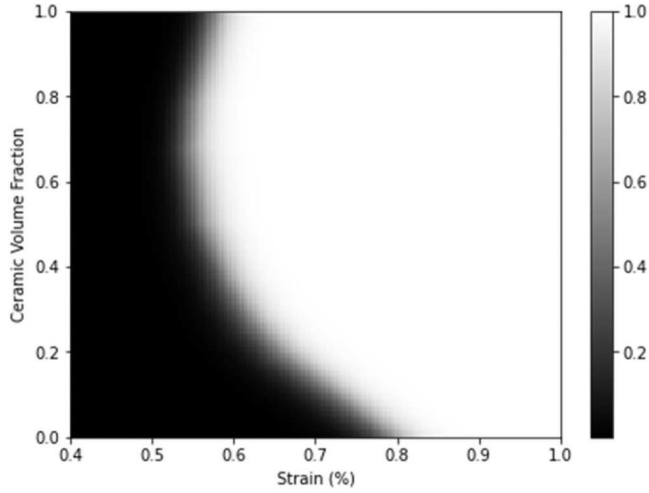


**Fig. 12** Failure probability map generated via ANN classification model.

### *4.4.2 Tensile strength prediction via CNN*

Table 3 shows that failure strength can vary at the same CVF due to microstructure uncertainties. Indeed, CNN may extract the microstructure features so that the failure strength can be deterministically predicted. There are 2525 microscale configurations generated during the data collection.

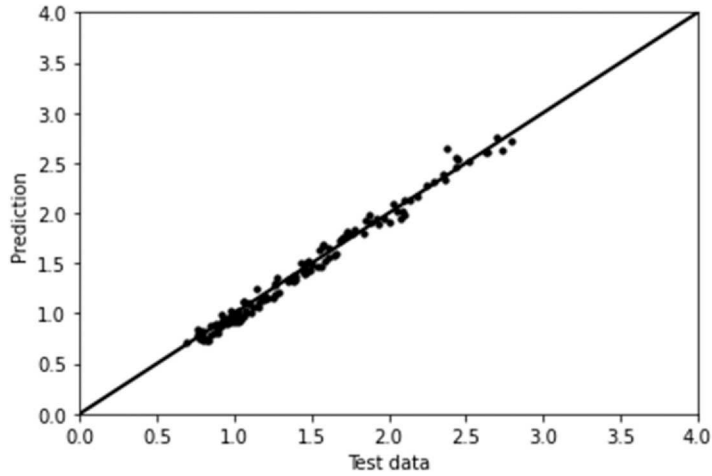


**Fig. 13** Strength (GPa) predictions of the test data.

We use those images as the inputs, and the corresponding failure strengths are the outputs. 5% of the dataset is randomly selected as the test set. The others are split into training and validation sets with a ratio of 85:15.

The CNN consists of three convolutional and max pooling layers. Each convolutional layer employs a 3 by 3 kernel. The numbers of filters are 32, 64, and 128, respectively, for convolutional layers. The max pooling uses a 2 by 2 max filter in each layer. After flattening, the fully connected neural network has two hidden layers with 64 and 32 neurons before the output layer. The "relu" activation function is utilized except for the output layer. After training, the coefficient of determination is 0.985 calculated from the test set. Compared to the actual outputs in the test set, the predictions are shown in Fig. 13, in which the diagonal line represents the ideal prediction.

## 5. Conclusions

This paper briefly introduced ANNs and some variations, such as PINNs, CNNs, and RNNs. We also reviewed the applications of those neural networks on some advanced material science research topics, including multiscale modeling, inverse problems, material image processing, and history-dependent constitutive modeling. In addition, we conducted a case study in which metal-ceramic composite materials were modeled and simulated at the microscale to generate the dataset. Then, we

developed and trained neural networks to predict material failure at the macroscale. The dataset and codes of this case study are provided on the github (https://github.com/jwli0728/ANNs-in-Material-Science). We believe that, with computer and computer technology development, the further development of AI/ML/DL algorithms will enhance scientific understanding in different disciplines, including materials discovery and design.

## Funding and acknowledgment

Xiao gratefully acknowledges the support from the National Science Foundation (#2104383) and the US Department of Education (ED#P116S210005).

## Author contributions

Xiao drafted the paper. Li did the case study. Bordas and Kim reviewed and revised the paper.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

Akbari, Z., Mirzadeh, H., & Cabrera, J.-M. (2015). A simple constitutive model for predicting flow stress of medium carbon microalloyed steel during hot deformation. *Materials & Design,* 77, 126–131. https://doi.org/10.1016/J.MATDES.2015.04.005.

Al-Haik, M. S., Garmestani, H., & Savran, A. (2004). Explicit and implicit viscoplastic models for polymeric composite. *International Journal of Plasticity*. https://doi.org/10.1016/j.ijplas.2003.11.017 [Preprint].

Alber, M., et al. (2019). Integrating machine learning and multiscale modeling—Perspectives, challenges, and opportunities in the biological, biomedical, and behavioral sciences. *npj Digital Medicine, 2*(1), 115. https://doi.org/10.1038/s41746-019-0193-y.

American Society for Metals., W. (1979). *Properties and selection—Nonferrous alloys and pure metals* (9th ed., Vol. 1). Metals Park Ohio: American Society for Metals. Available at: ⟨https://www.worldcat.org/title/metals-handbook-2-properties-and-selection-nonferrons-alloys-and-pure-metals/oclc/634910433?referer=di&ht=edition⟩ (Accessed May 29, 2019).

Arroyo, M., & Belytschko, T. (2003). A finite deformation membrane based on interatomic potentials for the transverse mechanics of nanotubes. *Mechanics of Materials, 35*(3–6), 193–215. https://doi.org/10.1016/S0167-6636(02)00270-3.

Attarian, S., & Xiao, S. (2022). Investigating the strength of Ti/TiB interfaces at multiple scales using density functional theory, molecular dynamics, and cohesive zone modeling. *Ceramics International*. https://doi.org/10.1016/J.CERAMINT.2022.07.259 [Preprint].

Balokas, G., Czichon, S., & Rolfes, R. (2018). Neural network assisted multiscale analysis for the elastic properties prediction of 3D braided composites under uncertainty. *Composite Structures, 183*(1), 550–562. https://doi.org/10.1016/j.compstruct.2017.06.037.

Bar, H. N., Bhat, M. R., & Murthy, C. R. L. (2004). Identification of failure modes in GFRP using PVDF sensors: ANN approach. *Composite Structures, 65*(2), 231–237. https://doi.org/10.1016/J.COMPSTRUCT.2003.10.019.

Belytschko, T., Liu, W. K., & Moran, B. (2000). *Nonlinear finite elements for continua and structures*. Wiley. ⟨https://books.google.com/books/about/Nonlinear_finite_elements_for_continua_a.html?id=C6goAQAAMAAJ⟩.

Bhaduri, A., Gupta, A., & Graham-Brady, L. (2022). Stress field prediction in fiber-reinforced composite materials using a deep learning approach. *Composites Part B: Engineering, 238*, 109879. https://doi.org/10.1016/J.COMPOSITESB.2022.109879.

Birman, V., et al. (2008). Response of spatially tailored structures to thermal loading. *Journal of Engineering Mathematics, 61*(2–4), 201–217. https://doi.org/10.1007/s10665-007-9151-9.

Bobaru, F., & Zhang, G. (2015). Why do cracks branch? A peridynamic investigation of dynamic brittle fracture. *International Journal of Fracture, 196*(1–2), 59–98. https://doi.org/10.1007/s10704-015-0056-8.

Boettinger, W. J., et al. (2003). *Phase-Field Simulation of Solidification, 32*, 163–194. https://doi.org/10.1146/ANNUREV.MATSCI.32.101901.155803.

Budarapu, P. R., & Rabczuk, T. (2017). Multiscale methods for fracture: A review. *Journal of the Indian Institute of Science 2017, 97*(3), 339–376. https://doi.org/10.1007/S41745-017-0041-5.

Cai, M., Hasanbeig, M., et al. (2021). Modular deep reinforcement learning for continuous motion planning with temporal logic. *IEEE Robotics and Automation Letters, 6*(4), 7973–7980 Available at: http://arxiv.org/abs/2102.12855 (Accessed April 8, 2021).

Cai, M., Xiao, S., et al. (2021). Optimal probabilistic motion planning with potential infeasible LTL constraints. *IEEE Transactions on Automatic Control*. https://doi.org/10.1109/TAC.2021.3138704 [Preprint].

Cecen, A., et al. (2018). Material structure-property linkages using three-dimensional convolutional neural networks. *Acta Materialia, 146*, 76–84. https://doi.org/10.1016/J.ACTAMAT.2017.11.053.

Chandrashekhara, K., Okafor, A. C., & Jiang, Y. P. (1998). Estimation of contact force on composite plates using impact-induced strain and neural networks. *Composites Part B: Engineering, 29*(4), 363–370. https://doi.org/10.1016/S1359-8368(98)00003-1.

Cho, K., et al. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. In *EMNLP 2014–2014 conference on empirical methods in natural language processing, proceedings of the conference* (pp. 1724–1734). Association for Computational Linguistics (ACL). Available at: ⟨https://doi.org/10.3115/v1/d14-1179⟩.

Choudhary, K., et al. (2022). Recent advances and applications of deep learning methods in materials science. *npj Computational Materials 2022, 8*(1), 1–26. https://doi.org/10.1038/s41524-022-00734-6.

Dreiseitl, S., & Ohno-Machado, L. (2002). Logistic regression and artificial neural network classification models: A methodology review. *Journal of Biomedical Informatics, 35*(5–6), 352–359. https://doi.org/10.1016/S1532-0464(03)00034-0.

Ericksen, J. L. (1984). *The cauchy and born hypotheses for crystals. Phase transformations and material instabilities in solids*. Elsevier61–77. https://doi.org/10.1016/B978-0-12-309770-5.50008-4.

Freitag, S., Graf, W., & Kaliske, M. (2013). A material description based on recurrent neural networks for fuzzy data and its application within the finite element method. *Computers & Structures, 124*, 29–37. https://doi.org/10.1016/J.COMPSTRUC.2012.11.011.

Furukawa, T., & Hoffman, M. (2004). Accurate cyclic plastic analysis using a neural network material model. *Engineering Analysis with Boundary Elements, 28*(3), 195–204. https://doi.org/10.1016/S0955-7997(03)00050-X.

Ghaffari, M. A., et al. (2019). Peridynamics with corrected boundary conditions and its implementation in multiscale modeling of rolling contact fatigue. *Journal of Multiscale Modelling, 10*(01), 1841003. https://doi.org/10.1142/s1756973718410032.

Ghaffari, M. A., Zhang, Y., & Xiao, S. (2018). Multiscale modeling and simulation of rolling contact fatigue. *International Journal of Fatigue, 108*, 9–17. https://doi.org/10.1016/J.IJFATIGUE.2017.11.005.

Ghaffari, M. A., Zhang, Y., & Xiao, S. P. (2017). Molecular dynamics modeling and simulation of lubricant between sliding solids. *Journal of Micromechanics and Molecular Physics, 2*(2), 1750009.

Gooneie, A., Schuschnigg, S., & Holzer, C. (2017). A review of multiscale computational methods in polymeric materials. *Polymers, 9*(1), https://doi.org/10.3390/POLYM9010016.

Grabowski, K., et al. (2017). Multiscale electro-mechanical modeling of carbon nanotube composites. *Computational Materials Science, 135*, 169–180. https://doi.org/10.1016/J.COMMATSCI.2017.04.019.

Griffiths, D. J. (1995). *Introduction to quantum mechanics*. Prentice Hall.

Güne, A., et al. (2018). Automatic differentiation in machine learning: A survey. *Journal of Machine Learning Research, 18*, 1–43. https://doi.org/10.5555/3122009.3242010.

Haghighat, E., et al. (2020). A deep learning framework for solution and discovery in solid mechanics. Available at: ⟨https://doi.org/10.48550/arxiv.2003.02751⟩.

Heider, D., Piovoso, M. J., & Gillespie, J. W. (2003). A neural network model-based open-loop optimization for the automated thermoplastic composite tow-placement system. *Composites Part A: Applied Science and Manufacturing, 34*(8), 791–799. https://doi.org/10.1016/S1359-835X(03)00120-9.

Himanen, L., et al. (2019). Data-driven materials science: Status, challenges, and perspectives. *Advanced Science, 6*(21), 1900808. https://doi.org/10.1002/ADVS.201900808.

Janežič, M., Klemenc, J., & Fajdiga, M. (2010). A neural-network approach to describe the scatter of cyclic stress–strain curves. *Materials & Design, 31*(1), 438–448. https://doi.org/10.1016/J.MATDES.2009.05.044.

Jarrah, M. A., Al-Assaf, Y., & Kadi, H. E. (2002). Neuro-fuzzy modeling of fatigue life prediction of unidirectional glass fiber/epoxy composite laminates. *Journal of Composite Materials, 36*(6), 685–700. https://doi.org/10.1177/0021998302036006176.

Jung, J., & Seok, J. (2016). Fatigue crack growth analysis in layered heterogeneous material systems using peridynamic approach. *Composite Structures, 152*, 403–407. https://doi.org/10.1016/J.COMPSTRUCT.2016.05.077.

El Kadi, H. (2006). Modeling the mechanical behavior of fiber-reinforced polymeric composite materials using artificial neural networks—A review. *Composite Structures, 73*(1), 1–23. https://doi.org/10.1016/J.COMPSTRUCT.2005.01.020.

El Kadi, H., & Al-Assaf, Y. (2002). Prediction of the fatigue life of unidirectional glass fiber/epoxy composite laminae using different neural network paradigms. *Composite Structures, 55*(2), 239–246. https://doi.org/10.1016/S0263-8223(01)00152-0.

Kanouté, P., et al. (2009). Multiscale methods for composites: A review. *Archives of Computational Methods in Engineering, 16*(1), 31–75. https://doi.org/10.1007/S11831-008-9028-8.

Kröner, E. (1977). Bounds for effective elastic moduli of disordered materials. *Journal of the Mechanics and Physics of Solids, 25*(2), 137–155. https://doi.org/10.1016/0022-5096(77)90009-6.

Le, B. A., Yvonnet, J., & He, Q.-C. (2015). Computational homogenization of nonlinear elastic materials using neural networks. *International Journal for Numerical Methods in Engineering, 104*(12), 1061–1084. https://doi.org/10.1002/nme.4953.

Li, S., & Liu, W. K. (2002). Meshfree and particle methods and their applications. *Applied Mechanics Reviews, 55*(1), 1–34. https://doi.org/10.1115/1.1431547.

Liu, Z., & Wu, C. T. (2019). Exploring the 3D architectures of deep material network in data-driven multiscale mechanics. *Journal of the Mechanics and Physics of Solids, 127*, 20–46. https://doi.org/10.1016/j.jmps.2019.03.004.

Liu, Z., Wu, C. T., & Koishi, M. (2019). A deep material network for multiscale topology learning and accelerated nonlinear modeling of heterogeneous materials. *Computer Methods in Applied Mechanics and Engineering, 345*, 1138–1168. https://doi.org/10.1016/J.CMA.2018.09.020.

Lu, C., et al. (2012). Multi-scale modeling of shock interaction with a cloud of particles using an artificial neural network for model representation. *Procedia IUTAM, 3*, 25–52. https://doi.org/10.1016/J.PIUTAM.2012.03.003.

Madenci, E., & Oterkus, S. (2016). Ordinary state-based peridynamics for plastic deformation according to von Mises yield criteria with isotropic hardening. *Journal of the Mechanics and Physics of Solids, 86*, 192–219. https://doi.org/10.1016/J.JMPS.2015.09.016.

Mianroodi, J. R., et al. (2022). Lossless multi-scale constitutive elastic relations with artificial intelligence. *npj Computational Materials 2022, 8*(1), 1–12. https://doi.org/10.1038/s41524-022-00753-3.

Miller, R. E., & Tadmor, E. B. (2009). A unified framework and performance benchmark of fourteen multiscale atomistic/continuum coupling methods. *Modelling and Simulation in Materials Science and Engineering, 17*(5), 053001. https://doi.org/10.1088/0965-0393/17/5/053001.

Mitchell, T. M. (1997). *Machine learning*. New York: McGraw-Hill.

Möller, B., & Beer, M. (2004). *Fuzzy randomness: Uncertainty in civil engineering and computational mechanics*. Springer.

Morgan, D., & Jacobs, R. (2020). Opportunities and challenges for machine learning in materials science. *Annual Review of Materials Research, 50*, 71–103. https://doi.org/10.1146/ANNUREV-MATSCI-070218-010015.

Munro, R. G. (2000). Material properties of titanium diboride. *Journal of research of the National Institute of Standards and Technology, 105*(5), 709–720. https://doi.org/10.6028/jres.105.057.

Nishawala, V. V., & Ostoja-Starzewski, M. (2017). Peristatic solutions for finite one- and two-dimensional systems. *Mathematics and Mechanics of Solids, 22*(8), 1639–1653. https://doi.org/10.1177/1081286516641180.

Oeser, M., & Freitag, S. (2009). Modeling of materials with fading memory using neural networks. *International Journal for Numerical Methods in Engineering, 78*(7), 843–862. https://doi.org/10.1002/nme.2518.

Olivito, R. S. (2003). A neural diagnostic system for measuring strain in FRP composite materials. *Cement and Concrete Composites, 25*(7), 703–709. https://doi.org/10.1016/S0958-9465(02)00103-8.

Patil, A. S., et al. (2019). Effect of TiB2 addition on the microstructure and wear resistance of Ti-6Al-4V alloy fabricated through direct metal laser sintering (DMLS). *Journal of Alloys and Compounds, 777*, 165–173. https://doi.org/10.1016/J.JALLCOM.2018.10.308.

Peng, H. J., et al. (2017). Review on high-loading and high-energy lithium–sulfur batteries. *Advanced Energy Materials, 7*(24), 1700260. https://doi.org/10.1002/AENM.201700260.

Pollice, R., et al. (2021). Data-driven strategies for accelerated materials design. *Accounts of Chemical Research, 54*(4), 849–860. https://doi.org/10.1021/ACS.ACCOUNTS.0C00785.

Rabczuk, T., Belytschko, T., & Xiao, S. P. (2004). Stable particle methods based on Lagrangian kernels. *Computer Methods in Applied Mechanics and Engineering, 193*(12–14), 1035–1063. https://doi.org/10.1016/J.CMA.2003.12.005.

Rahman, M. M., et al. (2017). A fully coupled space–time multiscale modeling framework for predicting tumor growth. *Computer Methods in Applied Mechanics and Engineering, 320*, 261–286. https://doi.org/10.1016/J.CMA.2017.03.021.

Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics, 378*, 686–707. https://doi.org/10.1016/J.JCP.2018.10.045.

Rao, C., & Liu, Y. (2020). Three-dimensional convolutional neural network (3D-CNN) for heterogeneous material homogenization. *Computational Materials Science, 184*, 109850. https://doi.org/10.1016/J.COMMATSCI.2020.109850.

Ray, S., & Cooney, R. P. (2018). *Thermal degradation of polymer and polymer compositesHandbook of environmental degradation of materials* (3rd ed.). William Andrew Publishing 185–206. Available at: https://doi.org/10.1016/B978-0-323-52472-8.00009-5.

Rizvi, Z. H., Nikolić, M., & Wuttke, F. (2019). Lattice element method for simulations of failure in bio-cemented sands. *Granular Matter, 21*(2), 1–14. https://doi.org/10.1007/S10035-019-0878-6/FIGURES/13.

Rodrigues, J. F., et al. (2021). Big data and machine learning for materials science. *Discover Materials, 1*(1), 1–27. https://doi.org/10.1007/S43939-021-00012-0.

Rutherford, K. L., et al. (1996). Abrasive wear resistance of TiN/NbN multi-layers: Measurement and neural network modelling. *Surface and Coatings Technology, 86–87*, 472–479. https://doi.org/10.1016/S0257-8972(96)02956-8.

Sabokpa, O., et al. (2012). Artificial neural network modeling to predict the high temperature flow behavior of an AZ81 magnesium alloy. *Materials & Design, 39*, 390–396. https://doi.org/10.1016/J.MATDES.2012.03.002.

Sainath, T. N., et al. (2015). Convolutional, long short-term memory, fully connected deep neural networks. In *ICASSP, IEEE international conference on acoustics, speech and signal processing—Proceedings* (pp. 4580–4584). Institute of Electrical and Electronics Engineers Inc. Available at: https://doi.org/10.1109/ICASSP.2015.7178838.

Samanta, A., et al. (2019). Atomistic simulation of diffusion bonding of dissimilar materials undergoing ultrasonic welding. *International Journal of Advanced Manufacturing Technology, 103*(1–4), 879–890. https://doi.org/10.1007/s00170-019-03582-9.

Schmidhuber, J. (2015). Deep Learning in neural networks: An overview. In *Neural Networks* (pp. 85–117). Elsevier Ltd. Available at: ⟨https://doi.org/10.1016/j.neunet.2014.09.003⟩.

Schulz, H., & Behnke, S. (2012). Deep learning: Layer-wise learning of feature hierarchies. *KI—Kunstliche Intelligenz, 26*(4), 357–363. https://doi.org/10.1007/S13218-012-0198-Z/FIGURES/4.

Shukla, K., et al. (2020). Physics-informed neural network for ultrasound nondestructive quantification of surface breaking cracks. *Journal of Nondestructive Evaluation, 39*(3), 1–20. https://doi.org/10.1007/S10921-020-00705-1/FIGURES/16.

Silling, S. A. (2000). Reformulation of elasticity theory for discontinuities and long-range forces. *Journal of the Mechanics and Physics of Solids, 48*(1), 175–209. https://doi.org/10.1016/S0022-5096(99)00029-0.

Silling, S. A., et al. (2007). Peridynamic states and constitutive modeling. *Journal of Elasticity, 88*(2), 151–184. https://doi.org/10.1007/s10659-007-9125-1.

Silling, S. A. (2010). Linearized theory of peridynamic states. *Journal of Elasticity, 99*(1), 85–111. https://doi.org/10.1007/s10659-009-9234-0.

Silling, S. A., & Askari, A. (2014). *Peridynamic model for fatigue cracking*. Available at: ⟨http://prod.sandia.gov/techlib/access-control.cgi/2014/1418590.pdf⟩ (Accessed June 27, 2018).

Silling, S. A., & Lehoucq, R. B. (2010). *Peridynamic theory of solid mechanics. Advances in Applied Mechanics*. Elsevier. https://doi.org/10.1016/S0065-2156(10)44002-8.

Singh, K., Rajput, S. K., & Mehta, Y. (2016). Modeling of the hot deformation behavior of a high phosphorus steel using artificial neural networks. *Materials Discovery, 6*, 1–8. https://doi.org/10.1016/J.MD.2017.03.001.

Smyser, C. P., & Chandrashekhara, K. (1997). Robust vibration control of composite beams using piezoelectric devices and neural networks. *Smart Materials and Structures, 6*(2), 178–189. https://doi.org/10.1088/0964-1726/6/2/007.

Stöcker, J., et al. (2022). A novel self-adversarial training scheme for enhanced robustness of inelastic constitutive descriptions by neural networks. *Computers & Structures, 265*, 106774. https://doi.org/10.1016/J.COMPSTRUC.2022.106774.

Subramanian, N., Rai, A., & Chattopadhyay, A. (2015). Atomistically informed stochastic multiscale model to predict the behavior of carbon nanotube-enhanced nanocomposites. *Carbon, 94*, 661–672. https://doi.org/10.1016/J.CARBON.2015.07.051.

Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. In *NIPS'14: Proceedings of the 27th* international conference on neural information processing systems (pp. 3104–3112). Montreal, CA. Available at: ⟨https://doi.org/10.5555/2969033.296917⟩.

Tadmor, E., Phillips, R., & Ortiz, M. (2000). Hierarchical modeling in the mechanics of materials. *International Journal of Solids and Structures, 37*(1–2), 379–389. https://doi.org/10.1016/S0020-7683(99)00095-5.

Tadmor, E. B., & Miller, R. E. (2011). *Modeling materials: Continuum, atomistic, and multiscale techniques*. Cambridge University Press.

Tadmor, E. B., & Miller, R. E. (2017). Benchmarking, validation and reproducibility of concurrent multiscale methods are still needed. *Modelling and Simulation in Materials Science and Engineering, 25*(7), 071001. https://doi.org/10.1088/1361-651X/aa834f.

Torquato, S., & Haslach, H. (2002). Random heterogeneous materials: Microstructure and macroscopic properties. *Applied Mechanics Reviews, 55*(4), B62–B63. https://doi.org/10.1115/1.1483342.

Tripathi, M. K., Kumar, R., & Tripathi, R. (2020). Big-data driven approaches in materials science: A survey. *Materials Today: Proceedings, 26*, 1245–1249. https://doi.org/10.1016/J.MATPR.2020.02.249.

Tuhami, A. E., & Xiao, S. (2022). Multiscale modeling of metal-ceramic spatially tailored materials via Gaussian process regression and peridynamics. *International Journal of Computational Methods*. https://doi.org/10.1142/S0219876222500256 [Preprint].

Ulmer II, C. W., et al. (1998). Computational neural networks and the rational design of polymeric materials: The next generation polycarbonates. *Computational and Theoretical Polymer Science, 8*(3–4), 311–321. https://doi.org/10.1016/S1089-3156(98)00035-X.

Unger, J. F., & Könke, C. (2009). Neural networks as material models within a multiscale approach. *Computers & Structures, 87*(19–20), 1177–1186. https://doi.org/10.1016/J.COMPSTRUC.2008.12.003.

Valoor, M. T., & Chandrashekhara, K. (2000). A thick composite-beam model for delamination prediction by the use of neural networks. *Composites Science and Technology, 60*(9), 1773–1779. https://doi.org/10.1016/S0266-3538(00)00063-4.

Versino, D., Tonda, A., & Bronkhorst, C. A. (2017). Data driven modeling of plastic deformation. *Computer Methods in Applied Mechanics and Engineering, 318*, 981–1004. https://doi.org/10.1016/J.CMA.2017.02.016.

Wagner, G. J., & Liu, W. K. (2003). Coupling of atomistic and continuum simulations using a bridging scale decomposition. *Journal of Computational Physics, 190*(1), 249–274. https://doi.org/10.1016/S0021-9991(03)00273-0.

Wang, K., & Sun, W. C. (2018). A multiscale multi-permeability poroplasticity model linked by recursive homogenizations and deep learning. *Computer Methods in Applied Mechanics and Engineering, 334*, 337–380. https://doi.org/10.1016/j.cma.2018.01.036.

Wang, Z., et al. (2022). Data-driven materials innovation and applications. *Advanced materials (Deerfield Beach, FL),* 2104113. https://doi.org/10.1002/ADMA.202104113.

White, D. A., et al. (2019). Multiscale topology optimization using neural network surrogate models. *Computer Methods in Applied Mechanics and Engineering, 346*, 1118–1135. https://doi.org/10.1016/J.CMA.2018.09.007.

Wiley, D. E., Manning, W. R., & Hunter, O. (1969). Elastic properties of polycrystalline TiB2, ZrB2 and HfB2 from room temperature to 1300 K. *Journal of the Less Common Metals, 18*(2), 149–157. https://doi.org/10.1016/0022-5088(69)90134-9.

Xiao, S., et al. (2008). Reliability analysis of carbon nanotubes using molecular dynamics with the aid of grid computing. *Journal of Computational and Theoretical Nanoscience, 5*(4), 528–534. https://doi.org/10.1166/jctn.2008.2495.

Xiao, S., et al. (2020). A machine-learning-enhanced hierarchical multiscale method for bridging from molecular dynamics to continua. *Neural Computing and Applications, 32*(18), 14359–14373. https://doi.org/10.1007/S00521-019-04480-7.

Xiao, S., et al. (2021). Machine learning in multiscale modeling of spatially tailored materials with microstructure uncertainties. *Computers and Structures, 249*, 106511. https://doi.org/10.1016/j.compstruc.2021.106511.

Xiao, S., Andersen, D. R., & Yang, W. (2008). Design and analysis of nanotube-based memory cells. *Nanoscale Research Letters, 3*, 416–420. https://doi.org/10.1007/s11671-008-9167-8.

Xiao, S., & Hou, W. (2007a). Multiscale modeling and simulation of nanotube-based torsional oscillators. *Nanoscale Research Letters, 2*, 54–59. https://doi.org/10.1007/s11671-006-9030-8.

Xiao, S., & Hou, W. (2007b). Studies of nanotube-based resonant oscillators through multiscale modeling and simulation. *Physical Review B, 75*(12), 125414. https://doi.org/10.1103/PhysRevB.75.125414.

Xiao, S., & Yang, W. (2005). A nanoscale meshfree particle method with the implementation of the quasicontinuum method. *International Journal of Computational Methods, 02*(03), 293–313. https://doi.org/10.1142/S0219876205000533.

Xiao, S., & Yang, W. (2006). Temperature-related Cauchy–Born rule for multiscale modeling of crystalline solids. *Computational Materials Science, 37*(3), 374–379. https://doi.org/10.1016/J.COMMATSCI.2005.09.007.

Xiao, S. P., & Belytschko, T. (2004). A bridging domain method for coupling continua with molecular dynamics. *Computer Methods in Applied Mechanics and Engineering, 193*(17–20), 1645–1669. https://doi.org/10.1016/J.CMA.2003.12.053.

Yaghoobi, A., & Chorzepa, M. G. (2017). Fracture analysis of fiber reinforced concrete structures in the micropolar peridynamic analysis framework. *Engineering Fracture Mechanics, 169*, 238–250. https://doi.org/10.1016/J.ENGFRACMECH.2016.11.004.

Yang, W., & Xiao, S. (2008). Extension of the temperature-related Cauchy–Born rule: Material stability analysis and thermo-mechanical coupling. *Computational Materials Science, 41*(4), 431–439. https://doi.org/10.1016/J.COMMATSCI.2007.04.023.

Yilin, G., Fuh Ying Hsi, J., & Wen Feng, L. (2021). Multiscale topology optimisation with nonparametric microstructures using three-dimensional convolutional neural network (3D-CNN) models. *Virtual and Physical Prototyping, 16*(3), 306–317. https://doi.org/10.1080/17452759.2021.1913783.

Zhang, E., Yin, M., & Karniadakis, G. E. (2020). Physics-informed neural networks for nonhomogeneous material identification in elasticity imaging. Available at: ⟨https://doi.org/10.48550/arxiv.2009.04525⟩.

Zhang, Z., & Friedrich, K. (2003). Artificial neural networks applied to polymer composites: A review. *Composites Science and Technology, 63*(14), 2029–2044. https://doi.org/10.1016/S0266-3538(03)00106-4.

Zhang, Z., & Gu, G. X. (2021). Physics-informed deep learning for digital materials. *Theoretical and Applied Mechanics Letters, 11*(1), 100220. https://doi.org/10.1016/j.taml.2021.100220.

Zhang, Z., Klein, P., & Friedrich, K. (2002). Dynamic mechanical properties of PTFE based short carbon fibre reinforced composites: Experiment and artificial neural network prediction. *Composites Science and Technology, 62*(7–8), 1001–1009. https://doi.org/10.1016/S0266-3538(02)00036-2.

Zheng, B., et al. (2022). Physics-informed machine learning model for computational fracture of quasi-brittle materials without labelled data. *International Journal of Mechanical Sciences, 223*, 107282. https://doi.org/10.1016/J.IJMECSCI.2022.107282.

Zhu, Q., & Ni, T. (2017). Peridynamic formulations enriched with bond rotation effects. *International Journal of Engineering Science, 121*, 118–129. https://doi.org/10.1016/j.ijengsci.2017.09.004.

Zhu, Y., et al. (2022). Intelligent traffic light via policy-based deep reinforcement learning. *International Journal of Intelligent Transportation Systems Research,* 1–11. https://doi.org/10.1007/S13177-022-00321-5.

Zopf, C., & Kaliske, M. (2017). Numerical characterisation of uncured elastomers by a neural network based approach. *Computers & Structures, 182*, 504–525. https://doi.org/10.1016/J.COMPSTRUC.2016.12.012.