Micromechanics-based deep-learning for composites
Challenges and future perspectives

Mirkhalaf, Mohsen; Rocha, Iuri

DOI
10.1016/j.euromechsol.2024.105242

Publication date
2024

Document Version
Final published version

Published in
European Journal of Mechanics, A/Solids

European Journal of Mechanics / A Solids 105 (2024) 105242

Contents lists available at ScienceDirect

European Journal of Mechanics / A Solids

journal homepage: www.elsevier.com/locate/ejmsol

Full length article

Micromechanics-based deep-learning for composites: Challenges and future perspectives

Mohsen Mirkhalaf [a,*], Iuri Rocha [b]

[a] Department of Physics, University of Gothenburg, 41296 Gothenburg, Sweden
[b] Delft University of Technology, Faculty of Civil Engineering and Geosciences, P.O. Box 5048, 2600GA Delft, The Netherlands

ARTICLE INFO

Keywords:
Micromechanics
Artificial neural networks
Composite materials

ABSTRACT

During the last few decades, industries such as aerospace and wind energy (among others) have been remarkably influenced by the introduction of high-performance composites. One challenge, however, for modeling and designing composites is the lack of computational efficiency of accurate high-fidelity models. For design purposes, using conventional optimization approaches typically results in cumbersome procedures due to huge dimensions of the design space and high computational expense of full-field simulations. In recent years, deep learning techniques have been found to be promising methods to increase the efficiency and robustness of a variety of algorithms in multi-scale modeling and design of composites. In this perspective paper, a short overview of the recent developments in micromechanics-based machine learning for composites is given. More importantly, existing challenges for further model enhancements and future perspectives of the field development are elaborated.

1. Introduction

Following successful applications of Machine Learning (ML) techniques in image processing, language processing etc., these techniques are currently finding a large number of applications in materials science as well (see e.g. Le et al. (2015) for mechanical response, Lu et al. (2019) for electrical response, and Wei et al. (2018) for thermal conductivity). The movement of different science and engineering fields towards usage of ML techniques is facilitated by (i) availability of unprecedented amount of data from experiments and simulations, (ii) quick growth of computer power, and (iii) availability of advanced open-source libraries such as TensorFlow and PyTorch.

Having accurate constitutive models is crucial for modeling and analysis of structural components with complex microscopic behavior. A possibility is to consider the material homogeneous and describe its behavior by a phenomenological constitutive model. These constitutive models typically have a number of parameters which are calibrated using experimental results (see *e.g.*, Mirkhalaf et al. (2016a, 2017), Ferreira et al. (2023) for polymers, and Cózar et al. (2022), Rodrigues Lopes et al. (2022), Oddy et al. (2022) for composites). Accurate predictions can be obtained using this approach. However, to (i) have physics-based models, (ii) establish structure–property relationships, and (iii) have a deep understanding of deformation, damage and failure mechanisms, the use of multi-scale modeling approaches, where the heterogeneities of the material sub-scale(s) are taken into account, becomes a necessity. Hence, many micro-mechanical and coupled multi-scale models have been developed for different classes of materials (see *e.g.*, Mirkhalaf et al. (2016b, 2019a) for polymers, and Melro et al. (2013), Arteiro et al. (2014), Mirkhalaf et al. (2022), Castricum et al. (2022) for composites).

For composite materials, the micro-structural properties of interest depend heavily on the type of composite under study. As an example, Short Fiber Reinforced Composites (SFRCs) have a large number of micro-structural features including fiber volume fraction, fiber orientation distribution, fiber geometrical aspect, constitutive properties of the matrix and reinforcements, among others. For woven composites, this becomes even more complex due to the fact that (i) yarns in woven composites are typically interlaced with each other, and thus complex stress states are developed when these materials are subjected to loads, (ii) the mesoscopic length scale has by itself a microscopic sub-scale: yarns in woven composites are themselves unidirectional composites. Hence, three different length scales are distinguished for a structure made from a woven composite, namely macroscopic, mesoscopic, and microscopic length scales. Fig. 1 shows numerical Representative Volume Elements (RVEs) for sub-scales of a woven composite and an SFRC.

Thus, accurate prediction of the mechanical response of composites, and establishing structure–property relationships for these materials through micromechanics-based approaches, requires high-fidelity

* Corresponding author.
E-mail addresses: mohsen.mirkhalaf@physics.gu.se (M. Mirkhalaf), i.rocha@tudelft.nl (I. Rocha).

https://doi.org/10.1016/j.euromechsol.2024.105242
Received 22 August 2023; Received in revised form 8 December 2023; Accepted 11 January 2024
Available online 13 January 2024

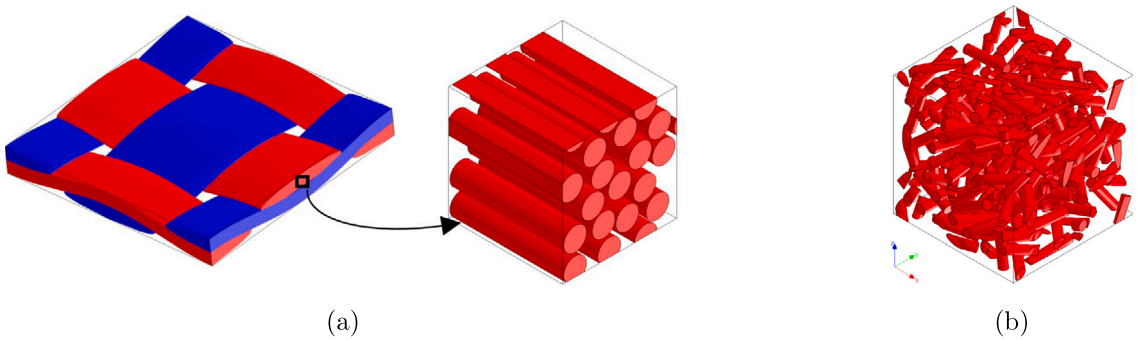

Fig. 1. Numerical RVEs of (a): meso-scale and micro-scale of a woven composites, (b): micro-scale of a short fiber reinforced composite.

three-dimensional full-field simulations. These simulations are computationally demanding. The computational cost becomes even higher when coupled multi-scale (macro–micro) simulations are needed. As a result, using these methods for modeling and analysis and real-life structures is still challenging despite the fast growth of the computational power.

In recent years, a number of Artificial Neural Network (ANN) models were developed to act as a surrogate for constitutive models (see *e.g.*, Wang and Sun (2018), Avery et al. (2021), Masi and Stefanou (2022), Benaimeche et al. (2022), Aldakheel et al. (2023a) and Liu et al. (2021) for a review). An ANN model is typically comprised of multiple layers of artificial neurons whose individual response is simple, but their joint properties can yield to very complex phenomena. These ANN models are developed based on two main categories of data, namely experimental results, and physics-based simulations. ANN models developed based on experiments typically describe phenomena for which a thorough physical understanding is not available yet. For simulation-based ANN models, two different approaches can be followed: (i) the material is considered homogeneous and described by a phenomenological constitutive model (see *e.g.*, Liu et al. (2020a,b)), (ii) the material heterogeneities at different sub-scales are considered in the simulations (see *e.g.*, Mozaffar et al. (2019), Wu et al. (2020), Mentges et al. (2021), Friemann et al. (2023), Ghane et al. (2023a)). This is a very efficient approach to circumvent the considerable computational cost of high-fidelity micromechanics-based and coupled multi-scale models, and hence, making these methods applicable to actual real-life structures. This is particularly important when there is a need to predict the non-linear behavior of these materials where energy dissipation mechanisms are activated, and deformation path plays an important role in their mechanical behavior.

ANN-enhanced multi-scale models can be used not only for modeling the material response, but also for material design purposes (see *e.g.*, Lew and Buehler (2023)). In the design process of composite materials, the micro-structural morphology is tailored together with a specific combination of constituent materials in order to fulfill a specific design requirement. However, such a design process is very challenging due to (i) huge dimensions of the design space with a large number of properties (morphological characteristics and material parameters) to be obtained, and (ii) computationally expensive full-field simulations. Using surrogate ANN models at the material sub-scales remarkably affects this process, with potential for significantly accelerating material design and certification.

In this perspective paper, an overview of recent developments in micromechanics-based ANN models for composite materials is given in Section 2. The goal of this paper is by no means to present a comprehensive review of the developments of the field. The main focus is to (i) illustrate major challenges for further developments of the field, and (ii) elaborate on potential approaches to address the existing challenges. These are given in Section 3. Some concluding remarks are given in Section 4.

2. Overview of recent developments

Initial applications of artificial neural networks in computational materials science and engineering date back to the 1990s (see *e.g.*, Wu et al. (1992), Oishi et al. (1995)). However, only during the last decade, a considerable attention from the communities got directed towards deep learning methods (see Ramprasad et al. (2017), Bock et al. (2019) for a review). In the following, some of the recent developments in this field and related to composite materials, are discussed. The goal here is to merely set the stage for discussing important trends and challenges in Section 3. For a more comprehensive review of the topic, an interested reader is referred to Liu et al. (2021).

2.1. Constitutive surrogate modeling

One of the subjects which has benefited greatly from usage of ANNs is multi-scale modeling of materials. ANNs have shown great capabilities to mitigate the computational cost of high-fidelity multi-scale models, see *e.g.*, Mozaffar et al. (2019), Mentges et al. (2021), Buehler (2022), Maia et al. (2023). This is highly relevant for composite materials due to their heterogeneous sub-scale(s) and multiple micro-structural properties with a direct impact on their macroscopic behavior. Fig. 2 shows a schematic representation of ANN surrogate model development using composite micro-mechanical simulations.

Elastic properties of different composite materials have been successfully predicted using different micro-mechanical models, see *e.g.*, Heidari-Rarani et al. (2018) for an overview on unidirectional (UD) composites, Mirkhalaf et al. (2020) for short fiber reinforced Composites (SFRCs), and Melro et al. (2012) for woven composites. These models can be categorized into two main approaches, namely mean-field and full-field models. Using mean-field (MF) models results in quick and rather accurate prediction of elastic properties of composites. However, to have high-fidelity predictions using full-field models (such as FE models), the associated computational effort increases substantially. Moreover, due to a large number of micro-structural properties, covering the design space is a time consuming process. Hence, ANN models are developed to surrogate elastic micro-mechanical models of composite materials. Some of these models are discussed below.

Le et al. (2015) proposed a computational homogenization approach for non-linear elastic heterogeneous materials using ANNs. Random discrete values of the effective potential were calculated using Finite Element Method (FEM) computations for an RVE, and an ANN was trained using the obtained values in the simulations. An effective constitutive law is defined using the effective potential, which is numerically determined using the trained ANN. Compared to nested multilevel FE^2 method, the proposed approach proved to be much more efficient as it eliminates the need for local FEM computations in structure calculations. Xu et al. (2015) used micro-mechanical modeling and ANNs to predict the elastic modulus of 3D multi-phase and multi-layer (MPML) braided composites. Two-scale micro-mechanical simulations were conducted to generate the required data. The trained ANN model

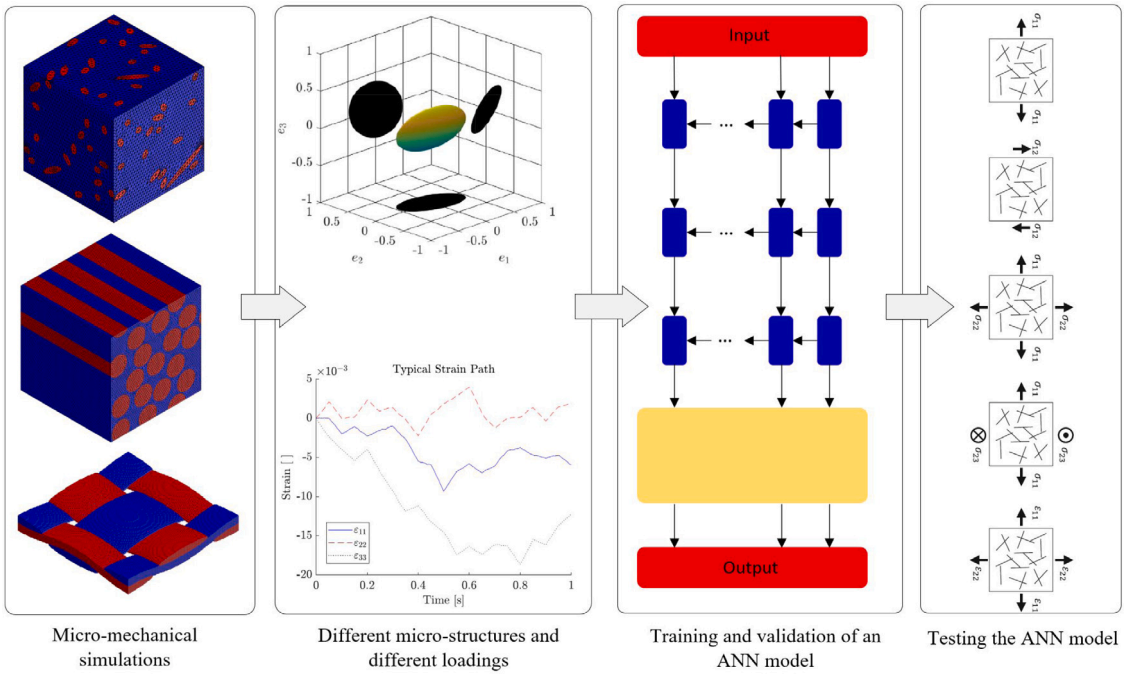

Fig. 2. Schematic representation of micromechanics-based ANN model development for composite materials.

maps micro-structural parameters to the composite's elastic modulus. Comparisons were also conducted with experimental results. Mentges et al. (2021) developed an ANN model for elastic properties of SFRCs. A two-step homogenization method was used for generating the required dataset. Finite Element Analysis (FEA) was used in the first homogenization step and Orientation Averaging (OA) in the second step (Mirkhalaf et al., 2020). As a result, an extensive database, accommodating a wide range for each of the micro-structural parameters, was generated. An optimal number of hidden layers and neurons in each layer was obtained using a grid-search hyper-parameter tuning technique. Ghane et al. (2023a) developed a multi-scale deep learning model for elastic properties of woven composites. The required data for training and validation of the models were generated using high fidelity FE simulations. Once the data was generated, first, an ANN model was trained and validated for the microscopic FE simulations. The micro-ANN model is then used at the meso-scale level to compute the elastic properties of a plain-weave fabric composite in a high-fidelity meso-scale model (see numerical RVEs for two hierarchical sub-scales of a woven composite in Fig. 1(a)). Based on meso-scale predictions, a second ANN is trained to estimate the resulting elastic behavior of woven composite based on the different combinations of fiber and matrix materials. A remarkable speed-up is obtained for both ANN models.

In addition to feed-forward and recurrent ANN models, other deep learning tools have also been employed in learning structure–property mappings and surrogate constitutive models for composite materials. Convolutional Neural Networks (CNNs) provide a structured way to incorporate full-field micro-structural information as input features without a need to handle high-dimensional feature spaces (*e.g.*, pixels or voxels) by assuming neighboring features are highly correlated (see Section 3.2.2 for a related discussion).

Aldakheel et al. (2023a) used convolutional neural networks (CNNs) to develop an accurate and computationally efficient model for homogenized elastic stresses of three-phase composite materials. The training dataset is formed by running standard finite element simulations (FEM) on a cubical micro-structure sample, with variations in the numbers of inclusions/phases. Also, transfer learning was used for new realistic micro-structures. The results achieved using the CNN model and its transfer learning exhibit a high precision in forecasting homogenized stresses, while substantially reducing the computational time.

In addition to elastic properties, yield, post-yield and viscous behavior of materials are of paramount importance when it comes to material design for different applications. As opposed to elastic response of materials, plasticity is typically a path-dependent phenomenon (De Souza Neto et al., 2008). This means that in addition to instantaneous strain state of a material point, the strain-path (evolution of deformation to reach the current strain state) is very important to determine the material response. As a result, plasticity models are typically challenging to implement and require return-mapping algorithms. Hence, a material plasticity simulation is computationally much more demanding than an elastic analysis. Hence, developments of accurate and computationally efficient ANN models for a material inelastic response is a very relevant and favorable approach to follow.

Lefik et al. (2009) presented one of the first studies on micro mechanics-based ANNs for elasto-plastic composites. Numerical simulations were conducted on a unit cell to identify the homogenized properties and generate the required input–output dataset for training a network. Different ANNs were trained and used in multi-scale analyses to facilitate stress–strain recovery at the sub-scale and to estimate material yielding states.

The path-dependent nature of composites make Recurrent Neural Networks (RNNs) an advantageous tool for learning complex constitutive behavior in a comprehensive way. This is due to the ability of RNNs to simultaneously learn a mapping between inputs and outputs together with complex latent space dynamics operators that explain time dependencies in the observed data. Some of the earliest results in employing RNNs for path-dependent material behavior were presented by Ghavamian et al. (Ghavamian and Simone, 2019) and (Mozaffar et al., 2019). In Mozaffar et al. (2019), RVE simulations of a composite were used for training and validating the RNN model. The models also included micro-structural features as extra inputs to the RNN, and the authors experimented with three different architectures with different approaches for incorporating these extra features. In Ghavamian and Simone (2019), RNNs are trained to reproduce Perzyna-type viscoplasticity with data coming from highly-specialized strain paths extracted from FE^2 simulations.

Friemann et al. (2023) developed a micromechanics-based ANN model for path-dependent elasto-plastic response of SFRCs (Friemann et al., 2023). To create the required database for training and validating the ANN model, mean-field simulations were conducted, and 40,000 data samples were generated. Arbitrary loading paths were generated by sampling 6-dimensional vectors via a random walk approach. Due to the time-series nature of the data for the model and path-dependency of plastic behavior, recurrent neural networks (RNN) such as Long-Short-Term-Memory (LSTM) and Gated Recurrent Unit (GRU) architectures were used to develop the model. A very good performance and computational efficiency of the RNN model was obtained using the GRU architecture. Although limited to conducting two-dimensional analyses, Wu et al. (2020) successfully developed a competent ANN model to serve as a surrogate for micro-mechanical modeling of an elasto-plastic composite. To achieve this, they performed two-dimensional full-field finite element simulations, which provided the necessary dataset for training and validating an RNN. The results showed promising performance. Expanding their research further, Wu and Noels (2022) went on to enhance the RNN surrogate model to capture the evolution of local micro-structure state variables when subjected to complex loading paths.

Aldakheel et al. (2023b) employed CNNs to map full two-dimen sional heterogeneous micro-structures to effective material properties for magneto-static simulations. The authors also extended the study to three-dimensional RVEs. A related approach can also be found in Rao and Liu (2020) but for elastic constants for mechanical equilibrium problems. Gupta et al. (2023) used a CNN to predict full fields for the complete stress tensor of linear-elastic composite RVEs and compute stiffness constants from the resulting fields, with the goal of accelerating FE^2 simulations. Krokos et al. (2022b) proposes an innovative multi-fidelity approach with a CNN that takes a coarse-scale stress field (assuming homogeneity) and a gray-scale representation of the actual heterogeneous micro-structure to compute corrected fine-scale stress fields. Robust predictions are ensured by Bayesian regularization through stochastic variational inference (see Section 3.2.1 for a related discussion). The same idea of performing convolutions in space can be extended to handle time series, resulting in one-dimensional Temporal Convolutional Networks (TCNs). In Wang et al. (2022), Wang et al. propose employing TCNs as an alternative to RNNs for modeling strain path-dependent materials with very long time dependencies. The convolution operator maps complete strain paths into a set of latent features that can then be mapped to stresses with a conventional feed-forward decoder. The approach is also used in Vlassis and Sun (2021) in a similar context. In both cases, temporal convolutions seem to lead to models with similar performance as RNNs, although to the best of our knowledge, a thorough and consistent comparison between methods including accuracy, dataset size requirements and training time has not been performed yet.

Another related class of models that has been gaining traction after the advent of large language models are attention-based architectures (Vaswani et al., 2023). An early application of the approach to solid mechanics is the work by Wang et al. (2020), in which the authors opt for an encoder–decoder architecture with an unrolled attention mechanism to reproduce path-dependent plasticity. The authors demonstrate the attention mechanism leads to more robust predictions when compared to an RNN-based encoder–decoder model, although comparisons with more straightforward and popular RNN architectures are not explored. Attention-based architectures can also be found in other interesting works dealing with mechanics tasks other than surrogate modeling (Buehler, 2022; Buehler and Buehler, 2022).

Finally, a current trend is to train machine learning models that combine data-driven and physics-based components. This ranges from special network architectures and loss function formulations that enforce thermodynamical and other physical constraints (Masi et al., 2021; Linka et al., 2021; Danoun et al., 2022) to hybrid architectures including intact material models embedded in broader network architectures (Liu et al., 2019; Maia et al., 2023; Rocha et al., 2023), including using physics-based constitutive models as feature extractors in a role akin to temporal convolution operators (Rocha et al., 2023). Further discussion on these models is left for Sections 3.2.3 and 3.2.4.

2.2. Material design and discovery

Aside from constructing surrogate models for material behavior, machine learning also sees extensive use in material design and physics discovery applications. Furtado et al. (2021) explored a range of machine learning techniques to approximate the notched strength of a range of laminated composites with different layups and used them to efficiently explore the design space and obtain probabilistic measures of strength and design allowables taking into account uncertainties in material properties and geometrical parameters.

Rather than densely exploring the design space, machine learning can also be used to obtain optimum material designs in a goal-oriented way through the use of surrogate models for computationally expensive numerical simulations. In Chuaqui et al. (2021), Chuaqui et al. used Bayesian Optimization to efficiently obtain fiber orientations for 10-ply laminates that maximize open-hole tension strength. Densely exploring the 10-dimensional design space would require an inordinate amount of expensive finite element simulations. Bayesian optimization offers an elegant solution by treating the FE predictions as a black-box model and employing probabilistic machine learning to model the epistemic uncertainty incurred by not computing the FEM model for every possible design. By balancing *exploration* (probing regions with high uncertainty) and *exploitation* of the design space (probing regions next to current global optima), the space is traversed in a goal-oriented way and the number of FE simulations is minimized. A similar approach was also employed by Ribeiro et al. (2020) in the context of functionally-graded plates.

Instead of using machine learning to directly fit a nonlinear surrogate model for constitutive material behavior, it can also be used in a *physics discovery* context by automatically selecting the best physics-based model to fit a dataset or by discovering altogether new thermo dynamically-consistent models. In Vlassis and Sun (2021), Vlassis et al. combined a set of deep neural networks with gradient-informed training and a level set approach for automatically discovering new hardening laws from data. The approach results in new elasto-plastic material formulations that can also serve as efficient surrogates for microscopic material behavior. Fuhg et al. (2022) employed a specialized tensor basis neural network architecture to automatically discover the type and orientation of anisotropy in hyper-elastic materials based on stress–strain snapshots and construct robust surrogates with polyconvex energy potentials. On the automated model selection front, Flaschel et al. (2022, 2023) developed an elegant learning framework based on sparse regression to automatically select and fit material models coming from a broad dictionary of constitutive models. Apart from leading to parsimonious and robust models for material behavior, the framework is also able to do the discovery directly from force–displacement snapshots coming from Digital Image Correlation (DIC) experiments, from which stress–strain behavior is indirectly inferred.

Other recent trends in design exploration include using deep learning models to obtain optimum designs without relying on optimization algorithms by instead inverting structure–property relationships (Kumar et al., 2020), optimization and material behavior discovery through Reinforcement Learning (Wang et al., 2019; Yu et al., 2022), and relying on generative machine learning for efficiently sampling new material and structural designs (Fuhr and Sumpter, 2022). More details on some of these trends are left for Sections 3.4.1 and 3.4.2.

3. Challenges and future perspective

In this section, the main challenges for development of micro mechanics-based ANN models are discussed. Also, future perspectives of the research developments to address the existing challenges are elaborated. Fig. 3 gives the structure and a summary of the content of this section.

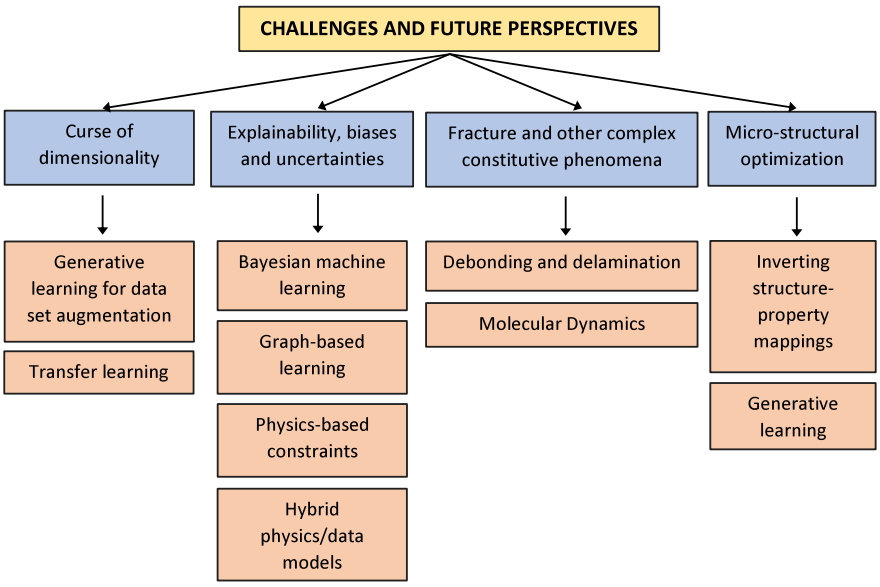

Fig. 3. A summary of challenges and future perspectives of micromechanics-based ANN models for composite materials.

3.1. Curse of dimensionality

Deep learning models need an excessive amount of data for training and validation purposes. For tasks with large feature spaces (*e.g.*, mapping a combination of strains, material properties and geometry to stresses), dataset sizes are exponentially larger than for simpler tasks (*e.g.*, simply mapping strains to stresses). To avoid this so-called *curse of dimensionality*, data should be either generated efficiently or the need for large datasets should be mitigated.

Further complicating this issue are difficulties related to building accurate models of material behavior in the first place. RVEs are fundamental units used in computational models to represent the heterogeneous nature of composites accurately. However, generating RVEs that faithfully capture the intricate arrangement of reinforcing phases within the matrix material is a very challenging task (Bargmann et al., 2018). The micro-structural features may be irregular, highly variable, or exhibit multiple length scales, making it difficult to define a unique RVE that encompasses all relevant characteristics. This is particularly true for SFRCs (Mirkhalaf et al., 2019b, 2022) and woven composites (Pierreux et al., 2019; Oddy et al., 2022). Additionally, the desired RVE size should be large enough to capture statistical variability but small enough to be computationally feasible (Mirkhalaf et al., 2016b). Overcoming these challenges is critical to achieving reliable simulations of composite materials and obtaining meaningful predictions of their mechanical behavior and performance.

The data requirement issue is particularly profound for non-linear path-dependent full-field simulations of composites. Depending on the micro-structure type, size and refinement of its spatial discretization, any single simulation could take days to finish. Hence, generating enough data in such cases would take an enormous amount of time.

3.1.1. Generative learning for dataset augmentation

A potential solution to the data requirement challenge could be using generative algorithms. Generative Adversarial Networks (GANs) is one of the generative models which could potentially mitigate this data requirement issue. GANs was initially proposed by Goodfellow et al. (2014) and can be described as two connected networks which are competing with each other. These networks are referred to as generator and discriminator. With the generator network, the goal is to create high quality "fake" data which can be used as training data. The role of the discriminator network is to recognize the "fake" data from the real data. The two networks are trained simultaneously and are in competition with each other. During this process, very high-quality fake data is generated which can be used in the training process of other machine learning models.

GANs can be effectively used to generate micro-structural samples. Henkes and Wessels (2022) used GANs to generate 3-dimensional micro-structures by learning the underlying material properties from a single μCT-scan. Using GANs could potentially provide a solution to the RVE generation challenge for different composite materials and different morphological properties.

Also, other data augmentation approaches, such as mathematical methods, could potentially be used to address the data scarcity issue. Very recently, Cheung et al. (2023) proposed a data augmentation approach for full-field micro-mechanical simulations of SFRCs. The proposed method involves considering multiple configurations, and thus, transforming the existing data from its original coordinate system into multiple coordinate systems via rotations of input and output data. The obtained results from different RNNs trained on the original dataset and augmented datasets showed the effectiveness of the data augmentation approach.

3.1.2. Transfer learning

To circumvent the challenge of data requirement, a potential solution is to use representational inductive transfer method (Torrey and Shavlik, 2009; Tan et al., 2018)

Using transfer learning, it is possible to transfer the knowledge of a trained network to another network. In a recent study by Ghane et al. (2023b), the application of transfer learning was introduced to tackle initialization hurdles in RNN models when studying the cyclic elastoplastic behavior of woven composites. Initially, a network underwent training with 6-dimensional, dynamically changing strain patterns, followed by fine-tuning using a distinct dataset comprising particular cyclic loadings. Transfer learning could also be used for cases where obtaining a huge amount of data is very difficult and time-consuming such as path-dependent high-fidelity micro-mechanical full-field simulations. In such cases, a source of low fidelity data (such as mean-field simulations) is required. Since it is not expensive to generate this kind of data, it is possible to develop big datasets. Then, a network is trained and validated which is essentially the pre-trained network for high-fidelity (such as full-field FE simulations) data. The pre-trained network is then enhanced using the limited amount of expensive high-fidelity data. This approach helps to remarkably decrease the amount of time and computational power needed to

generate the required datasets. Recently, Cheung and Mirkhalaf (2024) employed this approach and developed a very accurate RNN model for elasto-plastic SFRCs using a pre-trained network with mean-field simulations (Friemann et al., 2023) together with a limited number of full-field simulations. We believe that this approach will gain much more attention from the community to develop ANN models for different composites, at different length scales, and addressing different complex phenomena such as plasticity, damage and fracture.

It is worth mentioning that using transfer learning can perhaps be considered also as a potential remedy to the difficulty of RVE generation process: High-fidelity data can be obtained from morphologies for which generating an RVE is a straightforward process, and low-fidelity data can be obtained for the entire modeling/design space.

3.2. Explainability, biases and uncertainties

Predictions made with machine learning models are often not interpretable and with no indication on the confidence levels associated with them. From the discussion in the previous section, even if measures are taken to mitigate the curse of dimensionality, the fact that input spaces are often high-dimensional means that trained models will regardless often operate in extrapolation (see Balestriero et al. (2021) for an insightful discussion on the subject). It is therefore crucial to promote robustness and explainability while handling any remaining prediction uncertainties in a consistent way.

A straightforward approach for promoting robustness is to introduce regularization through traditional machine learning tools such as dropout layers, weight decay and L_1 sparsification. These effectively reduce the complexity of the model being trained and tend to reduce prediction variability away from the training space. In machine learning nomenclature, this is referred to as increasing model *bias* and reducing *variance* (Bishop, 2006). However, not all biases are created equal, and the aforementioned data-driven sources of bias come with a set of caveats: (i) they introduce new hyperparameters which can only be tuned by sacrificing a portion of the available data (*e.g.*, through cross validation); (ii) they are not *a priori* related to any domain knowledge associated with the learning task or the collected data; and (iii) they do not improve interpretability, the inner workings of the models still remain largely opaque.

While data-driven regularization remains nevertheless crucial in most applications, it can be complemented by other, less opaque approaches. This section presents a number of alternative sources of model bias that can promote parsimony and sparseness in ways that can be more directly interpretable by composite material modeling experts and that make the resulting predictions more transparent for interpretation.

3.2.1. Bayesian machine learning

The Bayesian formalism leverages probability theory to translate *a priori* assumptions and beliefs into model bias. A key departure of Bayesian machine learning with respect to conventional approaches is letting go of the assumption that model parameters and predictions are deterministic but can rather be described by probability densities. By observing data, these densities move from their *prior distributions* to *posterior distributions* according to Bayes' Theorem (Bishop, 2006). Instead of letting the observed data fully determine model behavior, it is instead a compromise between the data and assumptions made in the prior distributions which can come directly from domain knowledge and engineering experience.

Models incorporating Bayesian aspects tend to be more parsimonious and robust against overfitting while being able to use their complete datasets for training (Bishop, 2006) *i.e.* separating some data as validation dataset is not necessary. Moreover, model predictions become probabilistic, providing an intuitive measure of prediction uncertainty. Finally, by further committing to a Bayesian workflow and adopting a prior distribution over models, comparing different models (*e.g.*, through Bayes' factors) and combining predictions coming from different models can be done in a principled and elegant way (*e.g.*, through mixture models).

Bayesian learning frameworks have promising applications for a wide range of tasks. Fuhg et al. (Fuhg and Bouklas, 2022) used Gaussian Process models for constructing surrogate models of composite material behavior, Gaynutdinova et al. (2023) employ Markov Chain Monte Carlo simulations for inferring micro-structural parameters in composite materials from observed DIC data, Girolami et al. (2021) propose a fully-probabilistic version of FEM in which model uncertainty is propagated in a principled way, Pförtner et al. (2023) develop a broad range of numerical solvers with extra smoothness inherited from Bayesian priors, and Chuaqui et al. (2021) leverages Gaussian Processes for efficient optimization. Moreover, techniques such as Stochastic Variational Inference make the Bayesian formalism readily accessible in deep learning applications (Wilson and Izmailov, 2020), including not only models for regression but also for manifold learning and generative discovery — *e.g.*, variational autoencoders, Gaussian diffusions (Prince, 2023).

Bayesian machine learning also facilitates seamless online data assimilation, model updating and enforcement of physical constraints. In Rocha et al. (2021), a Gaussian Process material model surrogate is updated on the fly with data coming from a small number of microscale computations of a two-phase composite, with the prediction uncertainty provided by the Bayesian formalism being used to guide the active learning process. Krokos et al. (2022b) also exploit Bayesian models in predicting stresses in porous materials, employing uncertainty both for selective learning and online enforcement of Neumann boundary conditions through an Ensemble Kalman approach (Krokos et al., 2022a). All these developments show a great deal of promise, and we therefore fully expect Bayesian learning approaches to become more widespread in the near future.

3.2.2. Graph-based learning

Bias can also follow directly from the choice of machine learning model being used. When a CNN is picked to extract features from images, it is implicitly assumed that neighboring pixels are likely to have similar colors, and this in turn introduces bias to the model. For recurrent networks and time convolutions, a bias is introduced by assuming consecutive time steps are highly correlated.

Graph Neural Networks (GNNs) are generalizations of CNNs that allow for incorporating geometric bias without the need of a regular grid of input features (*e.g.*, pixels, voxels). From an *a priori* assumed graph structure with nodes linked together by edges, GNNs learn with a message passing mechanism through which information is shared between neighboring nodes and gradually propagates through the graph. The beneficial bias of the model comes from adopting a graph structure and connectivity that reflect actual geometrical features of the material being modeled.

Although GNNs are a relatively new machine learning development, they are already garnering a substantial interest from the solid mechanics community. For instance, Yang and Buehler (2022) use GNNs to build surrogate models for molecular dynamics simulations of crystalline solids and predict homogenized material properties. In this case, graph nodes and connectivity follow naturally from the atomic structure being modeled. At the continuum scale, Maurizi et al. (2022) directly translate the structure of an FE mesh into a graph and trains GNNs to predict displacements, strains and stresses in composite micromodels. Bias from FE meshes is also exploited by Vlassis and Sun (2023b), and a GNN-based variational autoencoder is used to learn latent state variables for plasticity models.

3.2.3. Physics-based constraints

Composites have been traditionally modeled with traditional physics-based models augmented with robust phenomenological relations that fill knowledge gaps with experimental observations (*e.g.*, S-N curves, time–temperature superposition curves). When switching to machine learning approaches, an early trend was to opt for fully data-driven models, and therefore, giving up on decades of advances on physics-based modeling. The community is however gradually correcting its course, inspired in part by recent advances in ANN-based PDE solvers such as so-called physics-informed neural networks (PINNs) (Raissi et al., 2019) and operator-type networks (Lu et al., 2021). In PINNs, physical bias is introduced by augmenting the loss function with extra terms usually computed from derivatives of network outputs (Haghighat et al., 2020). In operator networks, bias is introduced through changes in network architecture that give latent features extra constraints and interpretability (Yao et al., 2020; Yin et al., 2022; Patel et al., 2022).

Several authors have been working on physics-constrained models for heterogeneous materials. Masi et al. (2021) employ a mix of custom network architectures and augmented loss functions in order to promote thermodynamic consistency in ANN-based plasticity models. Linka et al. (2021) follow a related approach for hyper-elasticity but also include extra non-kinematic inputs that can for instance be used to accommodate micro-scale structural features. In an elegant framework combining a set of feed-forward, recurrent and graph neural networks, Vlassis and Sun (2023b) rely on splitting well-defined constitutive tasks between different machine learning models and combining them into a physics-based return mapping algorithm for handling hyperelasto-plasticity.

We expect this research direction to maintain its rapid expansion moving forward. Most of the proposed frameworks focus on a single class of material models (*e.g.*, hyper-elasticity, elasto-plasticity), and hence, extensions to broader classes of models are to be expected. The community can also take further inspiration from other recent advances in fundamental machine learning – *e.g.*, time-continuous recurrent neural networks (Hasani et al., 2020) – and exploiting potential links with physics-based principles.

3.2.4. Hybrid physics/data models

An alternative to the highly-specialized architectures of the previous section is to embed classical constitutive models directly into machine learning architectures in a form that keeps their robust physics-based algorithms essentially intact during training and prediction. In order to make a distinction from the aforementioned strategies, we refer to these as *hybrid models*. By focusing the machine learning effort on the task of how intact physics-based models can be combined or augmented, hybrid approaches can lead to architectures that (i) are directly applicable for any class of material model, (ii) are easier to implement in FE software and (iii) directly leverage existing constitutive model computer code with little to no changes. Moreover, in the context of surrogate modeling for FE^2 simulations, the choice of which physics-based models to embed becomes trivial, as models describing microscopic material phases are readily available (*e.g.*, elasticity for fibers, elasto-plasticity for matrix). There are different potential approaches to develop hybrid models, see Fig. 4 for a schematic representation of three examples of hybrid architectures.

In Liu et al. (2019), Liu (2021), Liu et al. propose a heavily-modified ANN architecture based on trainable homogenization-based building blocks. From a bottom layer of intact constitutive models for different material phases, hierarchical homogenization is performed until a single equivalent homogeneous material block remains which is then trained to reproduce the stiffness of an RVE. Remarkably, the model is trained exclusively with elastic-regime snapshots but can nevertheless predict path-dependent nonlinear behavior accurately. On the other hand, making predictions is not a straightforward operation, requiring a local Newton–Raphson solver to iteratively compute the strain contributions at the bottom layer.

The models by Maia et al. (2023) and Rocha et al. (2023) rely on architectures much closer to those of conventional surrogate models (*e.g.*, RNNs) and are trained with the same type of snapshots. In Maia et al. (2023), an encoder–decoder architecture learns how to distribute local strains between a number of *fictitious material points* and combine their stress predictions into a homogenized RVE stress. In Rocha et al. (2023), a single material model is made more flexible by an ANN-based encoder that learns an evolution in time for its material properties. Both architectures rely on simple forward passes for prediction and can capture strain path dependency without being trained for it, since time dependencies are fully handled by the embedded classical constitutive models (through the thermodynamic state variables $\boldsymbol{\alpha}$ of Fig. 4).

3.3. Fracture and other complex constitutive phenomena

There is a large body of literature on ANN models for elasticity, hyper-elasticity and plasticity. In contrast, there is very little data-driven models on different damage mechanism, such as fiber–matrix debonding, and fracture. Generating datasets for these more complex phenomena is considerably more difficult, due to the inherent challenging and computationally expensive nature of these simulations.

3.3.1. Debonding and delamination

Many studies on composite materials consider a perfect bonding between matrix and fibers. This is a rather reasonable assumption when the purpose is obtaining only the elastic properties. However, it is well-understood that matrix–fiber interfacial properties have a considerable impact on the overall macroscopic behavior of these materials, particularly at large deformations and beyond the elastic domain (Rabinovitch, 2014; Rocha et al., 2020; De Leon and Sweat, 2023). In order to consider imperfect fiber–matrix bonding, an approach is to use a cohesive zone between matrix and fibers is FE simulations (Meyer et al., 2022; Poggenpohl et al., 2022). This will however cause a dramatic increase of the required computational power and calculation time. To address this issue, transfer learning (see Section 3.1.2) is probably a good choice. Different datasets, excluding and including debonding and with different fidelity levels, could be a solution to address this in an ANN model. Another potential solution could be using generative models (see Section 3.1.1) to generate synthetic data using a rather small original dataset.

3.3.2. Molecular dynamics

Perfect bonding between matrix and reinforcements is not always an accurate and relevant assumption. One potential method to characterize a matrix–fiber interface and obtain the interfacial shear strength is to use Molecular Dynamics (MD) (Yan et al., 2021; Cho et al., 2020). MD simulations provide a very powerful and accurate tool to understand dynamics and structure of matter. However, in each MD step, many iterative computations are performed which makes this method computationally expensive. It is possible to use ANNs to surrogate MD simulations. We believe that GNNs (see Section 3.2.2) seems to be a proper choice for surrogate MD simulations for measuring interfacial properties (see *e.g.*, Li et al. (2022)). GNNs enables encoding information not only related to particles properties, but also their interactions to their neighboring particles.

3.4. Micro-structural optimization

Data-driven computational analysis is also beneficial for exploring multi-scale and high-dimensional design spaces and take advantage of data to find best designs for different objective macroscopic properties. Traditionally, efforts in composite material optimization have been focused at the meso-scale, with micro-scale structure being largely regarded as fixed. However, recent advances in manufacturing such as 3D

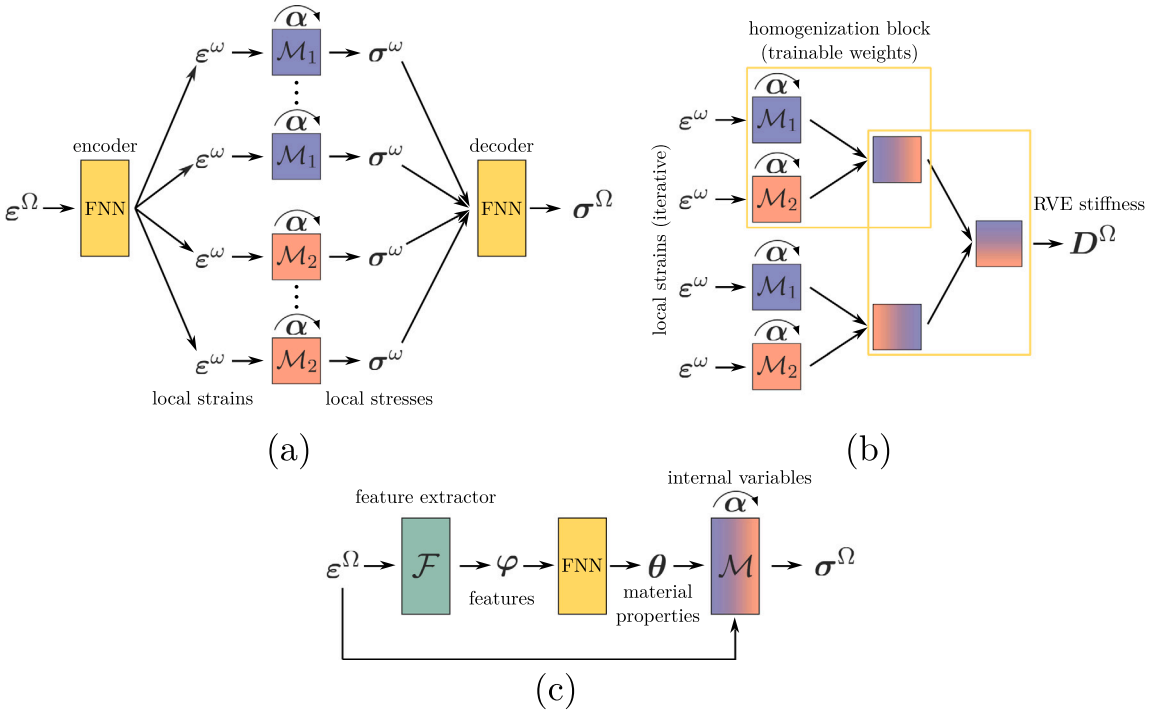

Fig. 4. Architectures of the hybrid surrogate models of (a) PRNNs (Maia et al., 2023), (b) DMNs (Liu et al., 2019), and (c) the evolving material properties approach from Rocha et al. (2023) for two-phase composites.

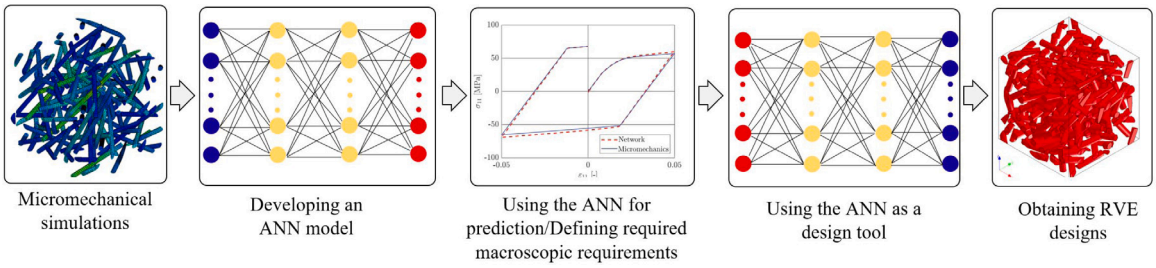

Fig. 5. Schematic representation of using an ANN for designing composites.

printing and automatic fiber placement are starting to open doors for further design optimization at very small scales. We therefore expect an increased focus on research in micro-structural material optimization moving forward, and machine learning offers a set of efficient tools for handling the resulting extremely high-dimensional design spaces. Recent trends are focused on leveraging machine learning to (i) avoid expensive iterative optimization procedures by directly mapping target properties to suitable optima, and (ii) learning lower-dimensional latent descriptors of microscopic geometry in order to avoid the huge dimensionalities involved in topology optimization.

3.4.1. *Inverting structure–property mappings*

Classical optimization techniques involve iterating over several feasible designs until a (local) optimum is found. In the context of optimizing micro-scale topology, the computational effort involved quickly becomes prohibitive, especially in an FE^2 context (*e.g.*, micro-structure grading throughout the meso-scale (Telgen et al., 2022)).

Instead of resorting to traditional optimization algorithms, promising new developments are being made in using machine learning to invert structure–property mappings and train models that can directly give optimized micro-structures for a given desired property with no iterations. A schematic representation of a composite design process using a micromechanics-based ANN model is shown in Fig. 5.

Kumar et al. (2020) proposed an architecture with a pair of feed-forward ANNs with inverse and forward structure–property mappings for stiffness of spinodoid-type micro-structures generated from solving the Cahn–Hilliard equation. By chaining the networks together, a regularized inverse mapping is obtained from which optimized micro-structures can be quickly obtained. The authors have since successfully extended the approach to truss- (Bastek et al., 2022) and growth-based (Van 't Sant et al., 2023) micro-structures.

3.4.2. *Generative learning*

In the aforementioned works, a low-dimensional parametrization of micro-scale geometry is readily available — *e.g.*, the coefficients of the Cahn–Hilliard equation. In general, an interpretable representation of microscopic geometry with limited dimensionality is not available, and resorting to pixel- or voxel-based topology optimization quickly becomes unacceptably high-dimensional.

A current trend is to employ generative machine learning models to learn low-dimensional representations for micro-structure that can be efficiently sampled from, in order to find candidate optima. These models, particularly in the form of large language models (LLMs), hold immense potential in revolutionizing materials science, especially in the realms of modeling and material design (see *e.g.*, Buehler (2023b)). These models, equipped with sophisticated algorithms and trained on vast datasets, can effectively capture intricate patterns and

relationships within materials datasets. In the field of material design, generative models offer the capability to predict novel material properties and even propose entirely new compositions by extrapolating from existing datasets (see *e.g.*, Ni et al. (2023), Buehler (2023a)). Vlassis and Sun (2023a) use a conditional denoising U-net to generate micro-structures based on the Mechanical MNIST dataset (Lejeune, 2020) targeting specific hyper-elastic energy potentials. The authors also condition on image-based latents for extra customization (*e.g.*, a rough hand-drawn sketch of the desired micro-structure shape), demonstrating how design constraints could be taken into account. Qian et al. (2022) combine a GAN for micro-structure generation, a CNN for structure–property mapping and a Genetic Algorithm for optimization in order to find composite micro-structures with optimized stiffness.

We expect research in this direction to receive even more attention in the near future. In particular, we expect the integration of design tools with Large Language Models (LLMs) to become a trend (Hu and Buehler, 2023), inspired by the recent successes of models such as ChatGPT (OpenAI, 2023) and DALL-E (Ramesh et al., 2022). Moreover, the approach is not limited to the micro-scale, and similar ideas can also be directly extended for instance to generations at the molecular scale (Guo et al., 2022; Fuhg et al., 2022). Recently, Buehler (2023b) proposed MeLM which is a flexible multi-modal mechanics language model. MeLM tackles nonlinear forward and inverse problems across various materials and scales. Demonstrating proficiency in tasks like bio-inspired honeycomb design, carbon nanotube mechanics, and protein unfolding, MeLM utilizes autoregressive attention mechanisms to represent large multi-particle systems. It excels in solving complex mechanics design problems, suggesting novel material architectures and enabling materials discovery. Also, Buehler (2023a) proposed a Large Language Model (LLM), called MechGPT to connect knowledge across diverse fields and explore relationships. Through a fine-tuned LLM approach, MechGPT is employed for various computational tasks such as knowledge retrieval, language tasks, hypothesis generation, and linking information across different domains.

4. Concluding remarks

Machine learning techniques, particularly ANNs and other deep learning models, are becoming important tools for modeling and analysis of different materials including composites. These methods enable the development of highly accurate and remarkably efficient models. In this perspective paper, applications of artificial neural networks together with composites micromechanics for surrogate modeling, design and discovery of these materials were discussed. A brief overview of recent developments in the field was given. However, despite a deluge of developments, there are still challenges to be solved and further enhancements to be developed in the quest for more general and accurate ANN models for different classes of composites. An effort was given to elaborate on the current limitations and challenges together with potential solutions to mitigate the issues. Further promising developments are expected, and we therefore believe that academic and industrial communities dealing with modeling and design of composites will develop/have access to ever more advanced and specifically tailored machine learning-based tools in the years to come.

CRediT authorship contribution statement

Mohsen Mirkhalaf: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Visualization, Writing – original draft, Writing – review & editing. **Iuri Rocha:** Investigation, Methodology, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgments

Mohsen Mirkhalaf gratefully acknowledges financial support from the Swedish Research Council (VR grant number: 2019-04715). Iuri Rocha gratefully acknowledges the TU Delft AI Labs Programme for their support through the SLIMM AI Lab.

References

Aldakheel, Fadi, Elsayed, Elsayed S., Zohdi, Tarek I., Wriggers, Peter, 2023a. Efficient multiscale modeling of heterogeneous materials using deep neural networks. Comput. Mech. 72 (1), 155–171, Cited by: 5; All Open Access, Green Open Access, Hybrid Gold Open Access.

Aldakheel, Fadi, Soyarslan, Celal, Palanisamy, Hari Subramani, Elsayed, Elsayed Saber, 2023b. Machine learning aided multiscale magnetostatics. Mech. Mater. 184, 104726.

Arteiro, A., Catalanotti, G., Melro, A.R., Linde, P., Camanho, P.P., 2014. Micro-mechanical analysis of the in situ effect in polymer composite laminates. Compos. Struct. 116, 827–840.

Avery, Philip, Huang, Daniel Z., He, Wanli, Ehlers, Johanna, Derkevorkian, Armen, Farhat, Charbel, 2021. A computationally tractable framework for nonlinear dynamic multiscale modeling of membrane woven fabrics. Internat. J. Numer. Methods Engrg. 122 (10), 2598–2625, Cited by: 14; All Open Access, Green Open Access.

Balestriero, Randall, Pesenti, Jerome, LeCun, Yann, 2021. Learning in high dimension always amounts to extrapolation. arXiv:2110.09485.

Bargmann, Swantje, Klusemann, Benjamin, Markmann, Jürgen, Schnabel, Jan Eike, Schneider, Konrad, Soyarslan, Celal, Wilmers, Jana, 2018. Generation of 3D representative volume elements for heterogeneous materials: A review. Prog. Mater. Sci. 96, 322–384.

Bastek, Jan-Hendrik, Kumar, Siddhant, Telgen, Bastian, Glaesener, Raphaël N., Kochmann, Dennis M., 2022. Inverting the structure– property map of truss metamaterials by deep learning. Proc. Natl. Acad. Sci. 119 (1), e2111505119.

Benaimeche, Mohamed Amine, Yvonnet, Julien, Bary, Benoit, He, Qi-Chang, 2022. A k-means clustering machine learning-based multiscale method for anelastic heterogeneous structures with internal variables. Internat. J. Numer. Methods Engrg. 123 (9), 2012–2041, Cited by: 8; All Open Access, Green Open Access.

Bishop, Christopher M., 2006. Pattern Recognition and Machine Learning. In: Information Science and Statistics, Springer, New York.

Bock, Frederic E., Aydin, Roland C., Cyron, Christian J., Huber, Norbert, Kalidindi, Surya R., Klusemann, Benjamin, 2019. A review of the application of machine learning and data mining approaches in continuum materials mechanics. Front. Mater. 6.

Buehler, Markus J., 2022. Modeling atomistic dynamic fracture mechanisms using a progressive transformer diffusion model. J. Appl. Mech. 89 (12), 121009.

Buehler, Markus J., 2023a. MechGPT, a language-based strategy for mechanics and materials modeling that connects knowledge across scales, disciplines and modalities. Appl. Mech. Rev. 1–82.

Buehler, Markus J., 2023b. Melm, a generative pretrained language modeling framework that solves forward and inverse mechanics problems. J. Mech. Phys. Solids 181, 105454.

Buehler, Eric L., Buehler, Markus J., 2022. End-to-end prediction of multimaterial stress fields and fracture patterns using cycle-consistent adversarial and transformer neural networks. Biomed. Eng. Adv. 4, 100038.

Castricum, B.A., Fagerström, M., Ekh, M., Larsson, F., Mirkhalaf, S.M., 2022. A computationally efficient coupled multi-scale model for short fiber reinforced composites. Composites A 163, 107233.

Cheung, Hon Lam, Mirkhalaf, Mohsen, 2024. A multi-fidelity data-driven model for highly accurate and computationally efficient modeling of short fiber composites. Compos. Sci. Technol. 246, 110359, https://www.sciencedirect.com/science/article/pii/S0266353823004530.

Cheung, Hon Lam, Uvdal, Petter, Mirkhalaf, Mohsen, 2023. Augmentation of scarce data – a new approach for deep-learning modeling of composites. Pre-print at: http://dx.doi.org/10.48550/arXiv.2311.14557.

Cho, Kiho, Sul, Jung-Hoon, Stenzel, Martina H., Farrar, Paul, Prusty, B. Gangadhara, 2020. Experimental cum computational investigation on interfacial and mechanical behavior of short glass fiber reinforced dental composites. Composites B 200, 108294.

Chuaqui, T.R.C., Rhead, A.T., Butler, R., Scarth, C., 2021. A data-driven Bayesian optimisation framework for the design and stacking sequence selection of increased notched strength laminates. Composites B 226, 109347.

Cózar, I.R., Otero, F., González, P. Maimíand E.V., Miot, S., Turon, A., Camanho, P.P., 2022. A three-dimensional plastic-damage model for polymer composite materials. Composites A 163, 107198.

Danoun, Aymen, Prulière, Etienne, Chemisky, Yves, 2022. Thermodynamically consistent Recurrent Neural Networks to predict non linear behaviors of dissipative materials subjected to non-proportional loading paths. Mech. Mater. 173, 104436.
De Leon, A., Sweat, R.D., 2023. Interfacial engineering of CFRP composites and temperature effects: A review. Mech. Compos. Mater. 59 (3), 419–440.
De Souza Neto, E.A., Perić, D., Owen, D.R.J., 2008. Computational Methods for Plasticity: Theory and Applications. pp. 1–791.
Ferreira, Bernardo P., Carvalho Alves, A. Francisca, Andrade Pires, F.M., 2023. An efficient finite strain constitutive model for amorphous thermoplastics: Fully implicit computational implementation and optimization-based parameter calibration. Comput. Struct. 281, 107007.
Flaschel, Moritz, Kumar, Siddhant, De Lorenzis, Laura, 2022. Discovering plasticity models without stress data. npj Comput. Mater. 8 (1), 1–10.
Flaschel, Moritz, Kumar, Siddhant, De Lorenzis, Laura, 2023. Automated discovery of generalized standard material models with EUCLID. Comput. Methods Appl. Mech. Engrg. 405, 115867.
Friemann, J., Dashtbozorg, B., Fagerström, M., Mirkhalaf, S.M., 2023. A micromechanics-based recurrent neural networks model for path-dependent cyclic deformation of short fiber composites. Internat. J. Numer. Methods Engrg. 124 (10), 2292–2314.
Fuhg, Jan N., Bouklas, Nikolaos, 2022. On physics-informed data-driven isotropic and anisotropic constitutive models through probabilistic machine learning and space-filling sampling. Comput. Methods Appl. Mech. Engrg. 394, 114915.
Fuhg, J.N., Bouklas, N., Jones, R.E., 2022. Learning hyperelastic anisotropy from data via a tensor basis neural network. J. Mech. Phys. Solids 168, 105022.
Fuhr, Addis S., Sumpter, Bobby G., 2022. Deep generative models for materials discovery and machine learning-accelerated innovation. Front. Mater. 9.
Furtado, C., Pereira, L.F., Tavares, R.P., Salgado, M., Otero, F., Catalanotti, G., Arteiro, A., Bessa, M.A., Camanho, P.P., 2021. A methodology to generate design allowables of composite laminates using machine learning. Int. J. Solids Struct. 233, 111095.
Gaynutdinova, L., Rokoš, O., Havelka, J., Pultarová, I., Zeman, J., 2023. Bayesian approach to micromechanical parameter identification using integrated digital image correlation. Int. J. Solids Struct. 280, 112388.
Ghane, E., Fagerström, M., Mirkhalaf, S.M., 2023a. A multiscale deep learning model for elastic properties of woven composites. Int. J. Solids Struct. 282, 112452.
Ghane, Ehsan, Fagerström, Martin, Mirkhalaf, Mohsen, 2023b. Recurrent neural networks and transfer learning for elasto-plasticity in woven composites. Pre-print at: http://dx.doi.org/10.48550/arXiv.2311.13434.
Ghavamian, F., Simone, A., 2019. Accelerating multiscale finite element simulations of history-dependent materials using a recurrent neural network. Comput. Methods Appl. Mech. Engrg. 357, 112594.
Girolami, Mark, Febrianto, Eky, Yin, Ge, Cirak, Fehmi, 2021. The statistical finite element method (statFEM) for coherent synthesis of observation data and model predictions. Comput. Methods Appl. Mech. Engrg. 375, 113533.
Goodfellow, Ian J., Pouget-Abadie, Jean, Mirza, Mehdi, Xu, Bing, Warde-Farley, David, Ozair, Sherjil, Courville, Aaron, Bengio, Yoshua, 2014. Generative adversarial nets. 3, (January), pp. 2672–2680,
Guo, Minghao, Thost, Veronika, Li, Beichen, Das, Payel, Chen, Jie, Matusik, Wojciech, 2022. Data-efficient graph grammar learning for molecular generation. arXiv:2203.08031.
Gupta, Ashwini, Bhaduri, Anindya, Graham-Brady, Lori, 2023. Accelerated multiscale mechanics modeling in a deep learning framework. Mech. Mater. 184, 104709.
Haghighat, Ehsan, Raissi, Maziar, Moure, Adrian, Gomez, Hector, Juanes, Ruben, 2020. A deep learning framework for solution and discovery in solid mechanics. arXiv:2003.02751.
Hasani, Ramin, Lechner, Mathias, Amini, Alexander, Rus, Daniela, Grosu, Radu, 2020. Liquid time-constant networks. arXiv:2006.04439.
Heidari-Rarani, M., Bashandeh-Khodaei-Naeini, K., Mirkhalaf, S.M., 2018. Micromechanical modeling of the mechanical behavior of unidirectional composites – A comparative study. J. Reinf. Plast. Compos. 37 (16), 1051–1071.
Henkes, Alexander, Wessels, Henning, 2022. Three-dimensional microstructure generation using generative adversarial neural networks in the context of continuum micromechanics. Comput. Methods Appl. Mech. Engrg. 400, 115497.
Hu, Yiwen, Buehler, Markus J., 2023. Deep language models for interpretative and predictive materials science. APL Mach. Learn. 1 (1), 010901.
Krokos, Vasilis, Bordas, Stéphane P.A., Kerfriden, Pierre, 2022a. A Graph-based probabilistic geometric deep learning framework with online physics-based corrections to predict the criticality of defects in porous materials. arXiv:2205.06562.
Krokos, Vasilis, Bui Xuan, Viet, Bordas, Stéphane P.A., Young, Philippe, Kerfriden, Pierre, 2022b. A Bayesian multiscale CNN framework to predict local stress fields in structures with microscale features. Comput. Mech. 69 (3), 733–766.
Kumar, Siddhant, Tan, Stephanie, Zheng, Li, Kochmann, Dennis M., 2020. Inverse-designed spinodoid metamaterials. npj Comput. Mater. 6 (1), 1–10.
Le, B.A., Yvonnet, J., He, Q.-C., 2015. Computational homogenization of nonlinear elastic materials using neural networks. Internat. J. Numer. Methods Engrg. 104 (12), 1061–1084.
Lefik, M., Boso, D.P., Schrefler, B.A., 2009. Artificial neural networks in numerical modelling of composites. Comput. Methods Appl. Mech. Engrg. 198 (21–26), 1785–1804.
Lejeune, Emma, 2020. Mechanical MNIST: A benchmark dataset for mechanical metamodels. Extreme Mech. Lett. 36, 100659.
Lew, Andrew J., Buehler, Markus J., 2023. Single-shot forward and inverse hierarchical architected materials design for nonlinear mechanical properties using an attention-diffusion model. Mater. Today 64, 10–20.
Li, Zijie, Meidani, Kazem, Yadav, Prakarsh, Barati Farimani, Amir, 2022. Graph neural networks accelerated molecular dynamics. J. Chem. Phys. 156 (14).
Linka, Kevin, Hillgärtner, Markus, Abdolazizi, Kian P., Aydin, Roland C., Itskov, Mikhail, Cyron, Christian J., 2021. Constitutive artificial neural networks: A fast and general approach to predictive data-driven constitutive modeling by deep learning. J. Comput. Phys. 429, 110010.
Liu, Zeliang, 2021. Cell division in deep material networks applied to multiscale strain localization modeling. Comput. Methods Appl. Mech. Engrg. 384, 113914.
Liu, Xin, Tao, Fei, Du, Haodong, Yu, Wenbin, Xu, Kailai, 2020a. Learning nonlinear constitutive laws using neural network models based on indirectly measurable data. J. Appl. Mech. 87 (8), 081003.
Liu, Xin, Tao, Fei, Yu, Wenbin, 2020b. A neural network enhanced system for learning nonlinear constitutive law and failure initiation criterion of composites using indirectly measurable data. Compos. Struct. 252, 112658.
Liu, Xin, Tian, Su, Tao, Fei, Yu, Wenbin, 2021. A review of artificial neural networks in the constitutive modeling of composite materials. Composites B 224, 109152.
Liu, Zeliang, Wu, C.T., Koishi, M., 2019. A deep material network for multiscale topology learning and accelerated nonlinear modeling of heterogeneous materials. Comput. Methods Appl. Mech. Engrg. 345, 1138–1168.
Lu, Xiaoxin, Giovanis, Dimitris G., Yvonnet, Julien, Papadopoulos, Vissarion, Detrez, Fabrice, Bai, Jinbo, 2019. A data-driven computational homogenization method based on neural networks for the nonlinear anisotropic electrical response of graphene/polymer nanocomposites. Comput. Mech. 64 (2), 307–321.
Lu, Lu, Jin, Pengzhan, Karniadakis, George Em, 2021. DeepONet: Learning nonlinear operators for identifying differential equations based on the universal approximation theorem of operators. Nat. Mach. Intell. 3 (3), 218–229.
Maia, M.A., Rocha, I.B.C.M., Kerfriden, P., van der Meer, F.P., 2023. Physically recurrent neural networks for path-dependent heterogeneous materials: Embedding constitutive models in a data-driven surrogate. Comput. Methods Appl. Mech. Engrg. 407, 115934.
Masi, Filippo, Stefanou, Ioannis, 2022. Multiscale modeling of inelastic materials with Thermodynamics-based Artificial Neural Networks (TANN). Comput. Methods Appl. Mech. Engrg. 398, 115190.
Masi, Filippo, Stefanou, Ioannis, Vannucci, Paolo, Maffi-Berthier, Victor, 2021. Thermodynamics-based artificial neural networks for constitutive modeling. J. Mech. Phys. Solids 147.
Maurizi, Marco, Gao, Chao, Berto, Filippo, 2022. Predicting stress, strain and deformation fields in materials and structures with graph neural networks. Sci. Rep. 12 (1), 21834.
Melro, A.R., Camanho, P.P., Andrade Pires, F.M., Pinho, S.T., 2013. Micromechanical analysis of polymer composites reinforced by unidirectional fibres: Part I – Constitutive modelling. Int. J. Solids Struct. 50 (11), 1897–1905.
Melro, A.R., Camanho, P.P., Pinho, S.T., 2012. Influence of geometrical parameters on the elastic response of unidirectional composite materials. Compos. Struct. 94 (11), 3223–3231.
Mentges, N., Dashtbozorg, B., Mirkhalaf, S.M., 2021. A micromechanics-based artificial neural networks model for elastic properties of short fiber composites. Composites B 213, 108736.
Meyer, Christopher S., Haque, Bazle Z. (Gama), Gillespie Jr., John W., 2022. Bridging length scales from micro to mesoscale through rate-dependent traction-separation law predictions. Composites B 231.
Mirkhalaf, S.M., Andrade Pires, F.M., Simoes, R., 2016a. An elasto-viscoplastic constitutive model for polymers at finite strains: Formulation and computational aspects. Comput. Struct. 166, 60–74.
Mirkhalaf, S.M., Andrade Pires, F.M., Simoes, Ricardo, 2017. Modelling of the post yield response of amorphous polymers under different stress states. Int. J. Plast. 88, 159–187.
Mirkhalaf, M., van Dommelen, J.A.W., Govaert, L.E., Furmanski, J., Geers, M.G.D., 2019a. Micromechanical modeling of anisotropic behavior of oriented semicrystalline polymers. J. Polym. Sci. B 57 (7), 378–391.
Mirkhalaf, SM, Eggels, E, Anantharanga, AT, Larsson, F, Fagerstrom, M, et al., 2019b. Short fiber composites: Computational homogenization vs orientation averaging. In: ICCM22 2019. Engineers Australia, p. 3000.
Mirkhalaf, S.M., Eggels, E.H., van Beurden, T.J.H., Larsson, F., Fagerström, M., 2020. A finite element based orientation averaging method for predicting elastic properties of short fiber reinforced composites. Composites B 202, 108388.
Mirkhalaf, S.M., Pires, F.M. Andrade, Simoes, R., 2016b. Determination of the size of the Representative Volume Element (RVE) for the simulation of heterogeneous polymers at finite strains. Finite Elem. Anal. Des. 119, 30–44.
Mirkhalaf, S.M., van Beurden, T.J.H., Ekh, M., Larsson, F., Fagerström, M., 2022. An FE-based orientation averaging model for elasto-plastic behavior of short fiber composites. Int. J. Mech. Sci. 219, 107097.
Mozaffar, M., Bostanabad, R., Chen, W., Ehmann, K., Cao, J., Bessa, M.A., 2019. Deep learning predicts path-dependent plasticity. Proc. Natl. Acad. Sci. 116 (52), 26414–26420.

Ni, Bo, Kaplan, David L., Buehler, Markus J., 2023. Generative design of de novo proteins based on secondary-structure constraints using an attention-based diffusion model. Chem 9 (7), 1828–1849.

Oddy, Carolyn, Ekh, Magnus, Fagerström, Martin, 2022. Macroscale modelling of 3D-woven composites: Elasto-plasticity and progressive damage. Int. J. Solids Struct. 250, 111696.

Oishi, A., Yamada, K., Yoshimura, S., Yagawa, G., 1995. Quantitative nondestructive evaluation with ultrasonic method using neural networks and computational mechanics. Comput. Mech. 15 (6), 521–533.

OpenAI, 2023. GPT-4 technical report. arXiv:2303.08774.

Patel, Dhruv, Ray, Deep, Abdelmalik, Michael R.A., Hughes, Thomas J.R., Oberai, Assad A., 2022. Variationally mimetic operator networks. arXiv:2209.12871.

Pförtner, Marvin, Steinwart, Ingo, Hennig, Philipp, Wenger, Jonathan, 2023. Physics-informed Gaussian process regression generalizes linear pde solvers. arXiv:2212.12474.

Pierreux, Gerrit, Van Hemelrijck, Danny, Massart, Thierry J., 2019. Automated generation of 3D orthogonal woven composites RVEs including yarn cross-section variations. Compos. Sci. Technol. 176, 90–102.

Poggenpohl, Lukas, Holthusen, Hagen, Simon, Jaan-Willem, 2022. Failure zone homogenization at mode II and mixed mode loading including gradient-extended damage and interface debonding at finite strains. Compos. Struct. 298.

Prince, Simon J.D., 2023. Understanding Deep Learning. MIT Press.

Qian, Chao, Tan, Ren Kai, Ye, Wenjing, 2022. Design of architectured composite materials with an efficient, adaptive artificial neural network-based generative design method. Acta Mater. 225, 117548.

Rabinovitch, Oded, 2014. An extended high order cohesive interface approach to the debonding analysis of FRP strengthened beams. Int. J. Mech. Sci. 81, 1–16.

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. J. Comput. Phys. 378, 686–707.

Ramesh, Aditya, Dhariwal, Prafulla, Nichol, Alex, Chu, Casey, Chen, Mark, 2022. Hierarchical text-conditional image generation with CLIP latents. arXiv:2204.06125.

Ramprasad, Rampi, Batra, Rohit, Pilania, Ghanshyam, Mannodi-Kanakkithodi, Arun, Kim, Chiho, 2017. Machine learning in materials informatics: Recent applications and prospects. npj Comput. Mater. 3 (1).

Rao, Chengping, Liu, Yang, 2020. Three-dimensional convolutional neural network (3D-CNN) for heterogeneous material homogenization. Comput. Mater. Sci. 184, 109850.

Ribeiro, Leonardo Gonçalves, Maia, Marina Alves, Parente Jr., Evandro, de Melo, Antônio Macário Cartaxo, 2020. Surrogate based optimization of functionally graded plates using radial basis functions. Compos. Struct. 252, 112677.

Rocha, I.B.C.M., Kerfriden, P., van der Meer, F.P., 2021. On-the-fly construction of surrogate constitutive models for concurrent multiscale mechanical analysis through probabilistic machine learning. J. Comput. Phys.: X 9, 100083.

Rocha, I.B. C.M., Kerfriden, P., van der Meer, F.P., 2023. Machine learning of evolving physics-based material models for multiscale solid mechanics. Mech. Mater. 184, 104707.

Rocha, Daniel Belchior, de Souza, Alana Gabrieli, Szostak, Marek, Rosa, Derval dos Santos, 2020. Polylactic acid/Lignocellulosic residue composites compatibilized through a starch coating. Polym. Compos. 41 (8), 3250–3259.

Rodrigues Lopes, Igor A., Camanho, Pedro P., Andrade Pires, Francisco M., Arteiro, Albertino, 2022. An invariant-based elasto-visco-plastic model for unidirectional polymer composites at finite strains. Int. J. Solids Struct. 236–237, 111292.

Tan, Chuanqi, Sun, Fuchun, Kong, Tao, Zhang, Wenchang, Yang, Chao, Liu, Chunfang, 2018. A survey on deep transfer learning. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 11141 LNCS, pp. 270–279.

Telgen, Bastian, Sigmund, Ole, Kochmann, Dennis M., 2022. Topology optimization of graded truss lattices based on on-the-fly homogenization. J. Appl. Mech. 89 (6).

Torrey, Lisa, Shavlik, Jude, 2009. Transfer learning. In: Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques. pp. 242–264.

Van 't Sant, Sikko, Thakolkaran, Prakash, Martínez, Jonàs, Kumar, Siddhant, 2023. Inverse-designed growth-based cellular metamaterials. Mech. Mater. 182, 104668.

Vaswani, Ashish, Shazeer, Noam, Parmar, Niki, Uszkoreit, Jakob, Jones, Llion, Gomez, Aidan N., Kaiser, Lukasz, Polosukhin, Illia, 2023. Attention is all you need. arXiv:1706.03762.

Vlassis, Nikolaos N., Sun, WaiChing, 2021. Sobolev training of thermodynamic-informed neural networks for interpretable elasto-plasticity models with level set hardening. Comput. Methods Appl. Mech. Engrg. 377, 113695.

Vlassis, Nikolaos N., Sun, WaiChing, 2023a. Denoising diffusion algorithm for inverse design of microstructures with fine-tuned nonlinear material properties. Comput. Methods Appl. Mech. Engrg. 413, 116126.

Vlassis, Nikolaos N., Sun, WaiChing, 2023b. Geometric learning for computational mechanics Part II: Graph embedding for interpretable multiscale plasticity. Comput. Methods Appl. Mech. Engrg. 404, 115768.

Wang, Kun, Sun, WaiChing, 2018. A multiscale multi-permeability poroplasticity model linked by recursive homogenizations and deep learning. Comput. Methods Appl. Mech. Engrg. 334, 337–380, Cited by: 202; All Open Access, Bronze Open Access, Green Open Access.

Wang, Kun, Sun, WaiChing, Du, Qiang, 2019. A cooperative game for automated learning of elasto-plasticity knowledge graphs and models with AI-guided experimentation. Comput. Mech. 64 (2), 467–499.

Wang, Jia-Ji, Wang, Chen, Fan, Jian-Sheng, Mo, Y.L., 2022. A deep learning framework for constitutive modeling based on temporal convolutional network. J. Comput. Phys. 449, 110784.

Wang, Chen, Xu, Li-yan, Fan, Jian-sheng, 2020. A general deep learning framework for history-dependent response prediction based on UA-Seq2Seq model. Comput. Methods Appl. Mech. Engrg. 372, 113357.

Wei, Han, Zhao, Shuaishuai, Rong, Qingyuan, Bao, Hua, 2018. Predicting the effective thermal conductivities of composite materials and porous media by machine learning methods. Int. J. Heat Mass Transfer 127, 908–916.

Wilson, Andrew Gordon, Izmailov, Pavel, 2020. Bayesian deep learning and a probabilistic perspective of generalization. 2020-December,

Wu, X., Ghaboussi, J., Garrett, J.H., 1992. Use of neural networks in detection of structural damage. Comput. Struct. 42 (4), 649–659.

Wu, Ling, Nguyen, Van Dung, Kilingar, Nanda Gopala, Noels, Ludovic, 2020. A recurrent neural network-accelerated multi-scale model for elasto-plastic heterogeneous materials subjected to random cyclic and non-proportional loading paths. Comput. Methods Appl. Mech. Engrg. 369, 113234.

Wu, Ling, Noels, Ludovic, 2022. Recurrent Neural Networks (RNNs) with dimensionality reduction and break down in computational mechanics; Application to multi-scale localization step. Comput. Methods Appl. Mech. Engrg. 390, 114476.

Xu, Yingjie, You, Tao, Du, Chenglie, 2015. An integrated micromechanical model and BP neural network for predicting elastic modulus of 3-D multi-phase and multi-layer braided composite. Compos. Struct. 122, 308–315.

Yan, Yishu, Xu, Junbo, Zhu, Huajian, Xu, Yinxiang, Wang, Min, Wang, Bingyin, Yang, Chao, 2021. Molecular dynamics simulation of the interface properties of continuous carbon fiber/ polyimide composites. Appl. Surf. Sci. 563, 150370.

Yang, Zhenze, Buehler, Markus J., 2022. Linking atomic structural defects to mesoscale properties in crystalline solids using graph neural networks. npj Comput. Mater. 8 (1), 1–13.

Yao, Houpu, Gao, Yi, Liu, Yongming, 2020. FEA-Net: A physics-guided data-driven model for efficient mechanical response prediction. Comput. Methods Appl. Mech. Engrg. 363.

Yin, Minglang, Zhang, Enrui, Yu, Yue, Karniadakis, George Em, 2022. Interfacing finite elements with deep neural operators for fast multiscale modeling of mechanics problems. Comput. Methods Appl. Mech. Engrg. 115027.

Yu, Chi-Hua, Tseng, Bor-Yann, Yang, Zhenze, Tung, Cheng-Che, Zhao, Elena, Ren, Zhi-Fan, Yu, Sheng-Sheng, Chen, Po-Yu, Chen, Chuin-Shan, Buehler, Markus J., 2022. Hierarchical multiresolution design of bioinspired structural composites using progressive reinforcement learning. Adv. Theory Simul. 5 (11), 2200459.