



Phong C. H. Nguyen
School of Data Science,
University of Virginia,
Charlottesville, VA 22904
e-mail: phongnguyen@virginia.edu

Joseph B. Choi
School of Data Science,
University of Virginia,
Charlottesville, VA 22904
e-mail: nmf8dm@virginia.edu

H. S. Udaykumar¹
Department of Mechanical Engineering,
University of Iowa,
Iowa City, IA 52242
e-mail: hs-kumar@uiowa.edu

Stephen Baek¹
School of Data Science,
University of Virginia,
Charlottesville, VA 22904;
Department of Mechanical and Aerospace
Engineering,
University of Virginia,
Charlottesville, VA 22903
e-mail: baek@virginia.edu

Challenges and Opportunities for Machine Learning in Multiscale Computational Modeling

Many mechanical engineering applications call for multiscale computational modeling and simulation. However, solving for complex multiscale systems remains computationally onerous due to the high dimensionality of the solution space. Recently, machine learning (ML) has emerged as a promising solution that can either serve as a surrogate for, accelerate or augment traditional numerical methods. Pioneering work has demonstrated that ML provides solutions to governing systems of equations with comparable accuracy to those obtained using direct numerical methods, but with significantly faster computational speed. These high-speed, high-fidelity estimations can facilitate the solving of complex multiscale systems by providing a better initial solution to traditional solvers. This paper provides a perspective on the opportunities and challenges of using ML for complex multiscale modeling and simulation. We first outline the current state-of-the-art ML approaches for simulating multiscale systems and highlight some of the landmark developments. Next, we discuss current challenges for ML in multiscale computational modeling, such as the data and discretization dependence, interpretability, and data sharing and collaborative platform development. Finally, we suggest several potential research directions for the future. [DOI: 10.1115/1.4062495]

Keywords: machine learning, artificial intelligence, computational modeling, multiscale modeling, data-driven engineering, machine learning for engineering applications, multiscale modeling and simulation

1 Introduction

Multiscale computational modeling has emerged as a central part of many mechanical engineering applications in recent years. The need for multiscale models arises from the limitations of single-scale models in describing physical laws and phenomena at different spatial and/or temporal scales. Larger-scale models alone cannot capture the fine-scale features or resolve key details of complex mechanical systems, while smaller-scale models are computationally inefficient for simulating physical phenomena that manifest at a larger system scale [1]. Multiscale modeling addresses this trade-off between physical fidelity and efficiency by using different numerical models to represent physical laws and phenomena at different scales and bridging scales using closure models. Over the past decade, a growing number of studies have adopted multiscale modeling and successfully demonstrated its ability to predict complex phenomena in areas such as fluid mechanics [2,3], materials modeling and design [4,5], and manufacturing and process modeling [6,7].

Broadly speaking, there are two main strategies for linking models at different scales. *Hierarchical* multiscale modeling approaches [6,8–10], also known as *sequential* or *decoupled* multiscale modeling, precompute constitutive relations (e.g., elasticity constants, energy release rate) using fine-scale models. These

fine-scale quantities of interest (QoIs) are typically computed “offline” and then utilized during system-scale computations that seek to model observable and measurable larger-scale phenomena. On the contrary, *concurrent* multiscale modeling approaches compute the required QoIs for coarse-scale models “on-the-fly,” as needed, by coupling with the fine-scale models [11–14]. While hierarchical approaches require large amounts of memory and storage space to store pre-computed QoIs, concurrent approaches face formidable computation costs as the problem size increases. These challenges limit the scalability and practicality of multiscale computational modeling in real-world applications.

Recent works suggest that machine learning (ML) has the potential to overcome the limitations of traditional multiscale modeling methods (Fig. 1). For instance, to address the issue of memory and storage requirements in hierarchical modeling, ML techniques can be applied to learn a coarse, low fidelity, and low-cost representation, referred to as *representation learning*, of pre-computed QoIs [15]. This allows for fast retrieval of information without the need for large memory and storage spaces. Concurrent multiscale modeling can also benefit from modern deep learning (DL) [16] models, which can replace traditional numerical solvers for “on-the-fly” computation of QoIs. Previous studies have demonstrated that ML-based simulation can achieve comparable fidelity to direct numerical simulation while being considerably faster in computation time [17]. Moreover, ML can accelerate the solution of inverse problems, providing direct design solutions that meet design constraints without requiring the traditional iterative optimization process [18].

¹Corresponding authors.

Manuscript received January 13, 2023; final manuscript received April 11, 2023; published online May 25, 2023. Assoc. Editor: Satyandra K. Gupta.

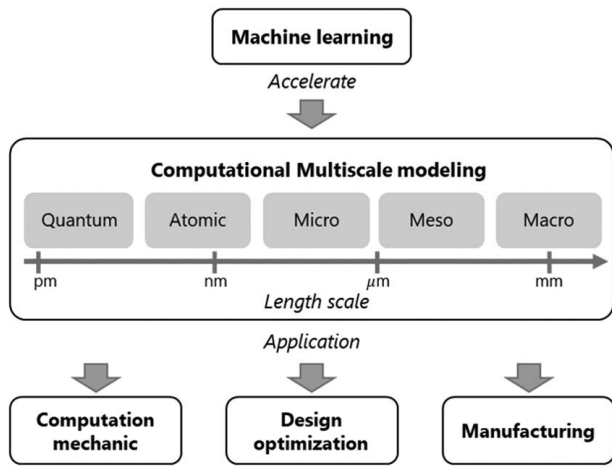


Fig. 1 Machine learning can be applied to accelerate multiscale computational modeling. Examples of applications include computational mechanics, design optimization, and manufacturing modeling.

Despite the promise of using ML in multiscale modeling, several critical challenges have emerged from pioneering works. One of the most significant obstacles is the data dependency of ML, particularly with modern DL models. ML/DL has achieved remarkable success in computer vision by leveraging massive datasets with high-quality annotations, such as image labels and captions. However, this is not typically the case for multiscale modeling problems, which require intensive calculations and laborious data collection. Therefore, alongside the development of fast and efficient data generation methods, the research community may benefit from “small data learning” techniques, such as transfer learning and meta-learning (Secs. 3.1.1 and 3.1.2), and collaborative/shared learning methods (Sec. 3.4).

Additionally, the majority of ML models are dependent on the discretization used during training, which can limit their scalability and practicality for multiscale problems where different discretization settings are used. This issue can be addressed by developing discretization-independent ML models such as those that employ continuous convolution or neural operators (Sec. 3.2), which are promising and warrant further investigation.

Lastly, the “black-box” nature of ML models poses a challenge to their trustworthiness and interpretability in practical applications. While ML predictions can be accurate, the assimilated information during training that drive predictions may not always reflect the underlying physical phenomena. Hence, it is important to explain how predictions were derived and examine the physics-awareness of a model, in addition to the evaluation of their statistical performance. To achieve this, the development of techniques for interpreting ML predictions, such as neural network visualization (Sec. 3.3), is crucial and demands additional efforts.

This paper examines the promise of ML in accelerating the development of multiscale computational modeling and offers potential solutions to address the challenges. The paper is organized as follows. In Sec. 2, we present several examples of successful applications of ML in multiscale modeling. Section 3 discusses the challenges that previous works have revealed, as well as our perspective on potential solutions. Finally, we provide conclusions and remarks in Sec. 4.

2 Examples of Machine Learning in Multiscale Computational Modeling

2.1 Computational Mechanics

2.1.1 Fluid Modeling. The intersection of ML and fluid mechanics presents a natural synergy, owing to their similarity in

dealing with large and complex datasets. In traditional fluid mechanics, one has to deal with a large amount of data coming from various sources, including experiments, field measurements or large-scale simulations. The analysis of fluid mechanics data still relies largely on domain expertise and traditional statistical analysis, which involve many simplifying assumptions [19]. Recently, the rise of ML, especially DL, has been a significant driving force in fluid mechanics research, rendering opportunities for tackling more difficult fluid mechanics problems such as turbulent flows and its modeling. By leveraging the computational efficiency of ML, advancements in fluid flow modeling can be achieved, which can have a profound impact on a wide range of applications, such as biomedical engineering, aerospace engineering, and aerial/hydro-robotics and autonomous systems, among others.

In fact, DL techniques such as artificial neural networks have been employed to simulate fluid dynamics very early, since the 1990s [20,21] and are still being developed [19]. Along with the breakthrough in DL architecture development, the progress in fluid flow modeling has been significantly accelerated. ML approaches such as recurrent neural networks (RNNs) have been utilized to model flows in extreme environments [22,23]. Unlike traditional feedforward neural networks, RNNs have loops that allow information to be passed from one step of the network to the next, enabling them to maintain information about previous inputs, and influence the current output. This characteristic makes RNNs capable of considering the inherent order of data and therefore well-suited to fluid mechanics modeling, where the data on the evolving flow fields are supplied to the machine learning algorithm in the form of sequences of images, videos, or time series.

Another DL technique that has also demonstrated success in fluid modeling is generative adversarial networks (GAN). The GAN framework is composed of two neural networks, namely the generator and the discriminator, that are trained in adversarial manner to generate high-quality synthetic data. Motivated by the capability of GAN to capture and generate realistic complex geometric patterns, Kim et al. [24] have applied GAN to model various complex fluid behaviors. These authors demonstrated that GAN is capable of accurately capturing the velocity field enabling many opportunities for further application including fast reconstruction of fluid flow, time-resampling, or fluid data compression.

Although current DL-based approaches have achieved several successes, they have a limitation that trained models are typically only effective for interpolating data points and may not be as reliable when extrapolating beyond the known data range [19]. This reduces the generalizability and applicability of trained models and hinders their use in real-world applications. To address this limitation, it is crucial not only to expand the training dataset to ensure that most future prediction tasks fall within the interpolation of the training data but also to develop new modeling methods with better extrapolation capabilities.

2.1.2 Materials Microstructure Modeling. Materials microstructure modeling is another domain of application that has benefited from the advancement of ML. ML can greatly facilitate the materials microstructure modeling process by providing efficient and accurate methods for predicting materials properties and behavior of given material compositions [25] or microstructure [26]. ML can achieve this modeling capability due to its ability to identify and learn from data complex patterns that may be difficult to capture using traditional statistical methods. This modeling capability allows ML algorithms to learn complex relationships between a wide range of microstructure morphology parameters (e.g., grain size, grain orientation, or distribution of defects) and the output material properties or mechanical response. These rapid and accurate predictions then can be used to design new materials with specific properties or to optimize the currently available materials with improved performance [27,28].

Recent advances in DL, with the increase in approximation power of neural networks, have further accelerated progress in

ML-based materials microstructure modeling. Modern DL methods such as convolutional neural networks (CNNs) can produce high-fidelity estimation of materials microstructure response to loads with an accuracy comparable to traditional numerical methods but with multiple order faster computation time [17,29]. Physics-informed ML (PIML) [30–32] is another branch of the DL method that has been widely adopted to simulate complex governing equations in multiscale materials modeling with increasing success. In PIML, prior physics knowledge is embedded in the design of DL algorithms, either via learning objectives [30], or architecture design [17,31]. As reported in pioneering works [17], the embedded prior physics knowledge helps reduce the need for large datasets while maintaining the prediction accuracy and generalizability of the DL model. The strong predictive capabilities and computational efficiency of DL-based methods make them ideally suited as surrogates for design optimization [33], multiscale structure-properties-performance (SPP) linkage modeling [34,35], and uncertainty quantification tasks [36].

2.1.3 Reducing the Nonlinearity of Dynamical Systems. Many complex mechanical systems are highly nonlinear, posing a significant challenge for obtaining stable and reliable numerical solutions, which often requires an extremely fine grid resolution to resolve physics at features like shocks, interfaces, and reactive fronts. The presence of nonlinearities and the potential for emergence of features at spatial and temporal hierarchies of scales renders the computation of such problems time-consuming. To address these challenges, the Koopman operator theory [37,38] has been drawing attention recently. This theory is based on the idea that any highly nonlinear dynamical system can be represented in terms of an infinite-dimensional linear operator, called the Koopman operator. This linear representation of nonlinear dynamical systems facilitates the analysis, design, and control of complex system via well-established methods for linear system. Here, the challenge is in approximating the Koopman operator from data. This has led to improved numerical methods such as dynamic mode decomposition (DMD) [39], which utilized linear observables, and its extended version (eDMD) [40], which augments the original DMD with nonlinear observables. These methods have demonstrated successes in modeling nonlinear flows [39], the Duffing oscillator, and the Kuramoto–Sivashinsky partial differential equation [40], among many others. DMD extracts rich and physically meaningful basis functions, which can notably improve the prediction accuracy of numerous classification and forecasting tasks involving time-series data [40,41]. In addition to DMD and its variants, DL with the representation power of neural networks has been applied exclusively to model the Koopman-invariant subspaces of observables [38,42,43]. The advantage of the Koopman-based methods is that they do not require prior knowledge or an explicitly defined model, and therefore can be scalable to various applications with unknown physics.

2.2 Topology Optimization. Topology optimization (TO) is another a group of computational design problems that benefit greatly from ML. TO determines the optimal distribution of material in a given design domain to achieve design objectives (e.g., high strength) while satisfying design constraints (e.g., low weight) [44]. With such capabilities, TO plays crucial role in many materials and structural design applications [4,37–39]. However, many TO methods face the challenge of repeatedly computing an expensive objective function, which involves finite element (FE) solvers at each iteration [18]. These computational processes can be time-consuming and demand significant computational resources, especially for problems that involve nonlinear analysis, such as plasticity, viscoelasticity, and dynamics. To this end, ML models can act as surrogates to replace computationally expensive FE calculations in the objective function, thereby accelerating design optimization processes [45–47]. Moreover, ML-based models are computationally inexpensive and differentiable, making them suitable for

gradient-based TO methods, which can further facilitate the optimization processes [48].

In addition to being used as surrogate models to replace expensive numerical models, ML can also be used to directly model the topology optimization process. In this approach, ML models take the design objective and constraints as input and give the final topology design as output. Various topology optimization problems, ranging from structures under static load [49] to conductive heat transfer [50], have been effectively tackled using CNN-based approaches. Deep generative models, such as GAN, can also be used to predict the optimal topology design, given design objectives and constraints. As reported in the literature, GAN-based methods have demonstrated great success in the optimization of material microstructures for desired properties [18,51,52]. Finally, reinforcement learning has also shown several successes for recurrent optimization problems, as documented in prior research [28,53].

2.3 Design for Manufacturing. ML-based methods are also advantageous in design for manufacturing (DFM) due to their strong modeling capability and computational efficiency. Assimilating process-structure-property linkages is a crucial task in computational modeling for DFM but remains challenging. Advanced manufacturing techniques, such as metallic additive manufacturing, involve various physical and chemical processes that occur across multiple scales during the transformation of raw materials. As these processes are complex, traditional numerical approaches such as finite element analysis or cellular automata can be expensive and inadequate for process optimization, monitoring, and control. In such cases, ML can be used to simplify the computational models by serving as surrogates or even replace the expensive numerical models entirely [54].

For instance, Tapia et al. [55] employed the Gaussian process (GP) to predict the depth of the melt pool in laser powder bed fusion of 316L stainless steel. The GP model was trained using data from high-fidelity physics-based numerical simulations and was able to accurately predict the melt pool depth of single-track prints based on inputs including laser power, velocity, and spot size. Furthermore, the use of GP allows for the assessment of uncertainty in predictions, which facilitates the making of robust and reliable design and control decisions [56]. In the same vein, Wang et al. [57] proposed a physics-informed data-driven modeling framework that can connect multiple data-driven surrogate models at different levels, allowing for the effective estimation of manufacturing uncertainty.

3 Challenges in Machine Learning for Multiscale Computational Modeling and Potential Solutions

As we have seen so far, there have been notable successes reported in the literature regarding the application of ML to multiscale computational modeling. However, there are still several critical obstacles that need to be surmounted before ML can be used in real-world applications. In this section, we discuss these challenges and offer potential solutions.

3.1 Data Dependency. The use of ML for multiscale computational modeling presents several challenges that require attention. Although modern ML techniques like DL can achieve high predictive performance, they necessitate a considerable amount of training data. However, collecting such training data is often an expensive and tedious process, especially when dealing with experimental data. Synthetic data generated from numerical simulation may offer an alternative, but producing them can also be computationally demanding due to the complexity of multiscale problems [17,58]. Insufficient training data can result in erroneous predictions or overfitting, reducing the generalizability, and trustworthiness of ML algorithms.

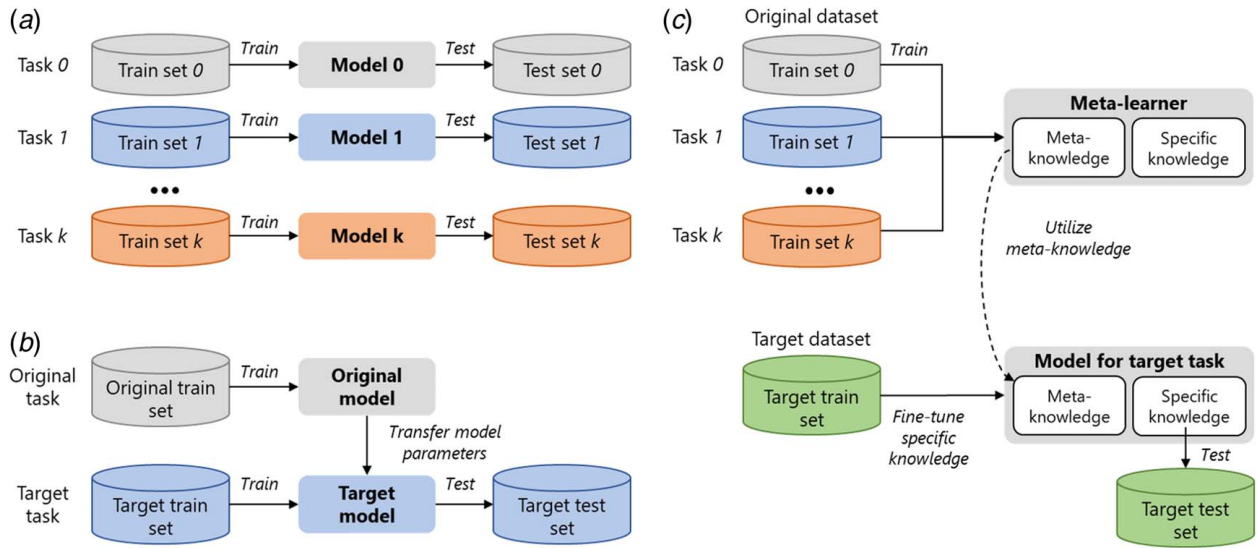


Fig. 2 Three types of training ML models: (a) training from the scratch for each task, (b) transfer learning, and (c) meta-learning

Besides the amount of training data, the quality of data is also crucial for the success of ML algorithms. Training data with missing values or many outliers/noises can cause ML models to learn incorrect patterns, resulting in poor prediction performance, and preventing the application of ML for real-world practice.

In the following subsections, we discuss several potential directions in this regard, including transfer learning and meta-learning.

3.1.1 Transfer Learning: Utilizing Prior Knowledge From Previous Training. Training ML models for computational modeling requires a large amount of data, which can take a lot of time and effort to collect. Unfortunately, when a new problem arises, the same data collection process must be repeated, and the ML model needs to be retrained with the newly collected dataset. However, repeating this long and tedious data collection process is not desirable because collecting data for ML models used in multiscale computational modeling is often expensive. Moreover, retraining an ML model from scratch (Fig. 2(a)) may not always yield similar statistical performance compared to previously trained results. In such cases, one may consider utilizing prior knowledge gained from previous training sessions to reduce the need for obtaining a large training dataset for a new problem.

In this vein, transfer learning [59] (Fig. 2(b)) has proven to be successful in many generic ML applications, such as image recognition [60] and language modeling [61]. In transfer learning, model parameters of a pre-trained ML model, trained on a similar but different dataset, are used as the starting point for the current training session. In the new training session, relatively minor changes are applied to a subset of model parameters, while a big portion of model parameters can be “frozen” and not updated. This allows the ML model to fine-tune its model parameters to learn the new problem with only a small number of newly added data. Transfer learning can reduce the need for a large dataset while maintaining the prediction performance of the ML models. Previous studies have shown that transfer learning is effective in training ML models to predict the deformation of structures under static loads in new loading scenarios [62]. Transfer learning has also been found to be successful in predicting the response of energetic material microstructures when subjected to new shock loading conditions [35].

3.1.2 Meta-Learning: Learning Generalized Knowledge for Quick Adaptations. A step further from the idea of transfer learning is the more generalizable concept of “meta-learning” (Fig. 2(c)). Meta-learning allows ML algorithms to quickly learn and adapt to new tasks by leveraging experience gained from solving similar

tasks. In this context, transfer learning can be considered as a specific case of meta-learning.

Conceptually, in meta-learning, a meta-learner is constructed to learn generalized knowledge (meta-knowledge) across different learning tasks and utilize it as a prior to learn specific knowledge for each task, thereby maximizing the likelihood

$$P(\theta_i|D_i, D_0) = \int_{\phi} P(\theta_i|\phi, D_i) P(\phi|D_0) d\phi \approx P(\theta_i|\phi^*, D_i) P(\phi^*|D_0) \quad (1)$$

In Eq. (1), the meta-learner is represented by meta-learning parameters ϕ and the knowledge on a specific task σ_i can be represented as a combination of the meta-knowledge ϕ and the specific knowledge $\theta_i|\phi$. The meta-knowledge is acquired via training on the reference dataset D_0 which maximizes the posteriori $\phi^* = \arg \max_{\phi} P(\phi|D_0)$. In more detail, given the reference dataset D_0 , the algorithm will first learn generalizable knowledge by leveraging the richness of the reference dataset D_0 , and then use it as a basis to learn the specific knowledge θ_i^* for the specific task σ_i using only a small data set D_i .

Previous works showed that meta-learning approaches could provide good prediction performance on unseen tasks despite being trained with relatively small datasets. For instance, Liu et al. [63] initialized physics-informed neural networks (PINN) with the Reptile algorithm, a meta-learning initialization method. The authors demonstrated that PINN can be trained more quickly with initialization guided by meta-learning algorithms than with the conventional randomized initialization methods. This new PINN training paradigm was successfully validated with a variety of forward and inverse test problems. In another study, Li et al. [64] proposed a PIML model for tool wear prediction problems in milling process modeling. The authors incorporated a physics-informed loss term that ensured the meta-learner learned the robust relationship between tool wear-rate and cutting forces, thereby facilitating parameter estimation and enhancing the interpretability of the PIML model.

Previous examples have demonstrated that meta-learning is well-suited for the small-data context of ML in the computational modeling of multiscale problems. Meta-learning allows for rapid adaptation to new operating conditions, such as new materials, geometric domains, or boundary conditions, without requiring a large training dataset. In addition, the “learning-to-learn” procedure can also aid to bridge the knowledge gap between simulations and real-world experiments. Currently, ML models are commonly

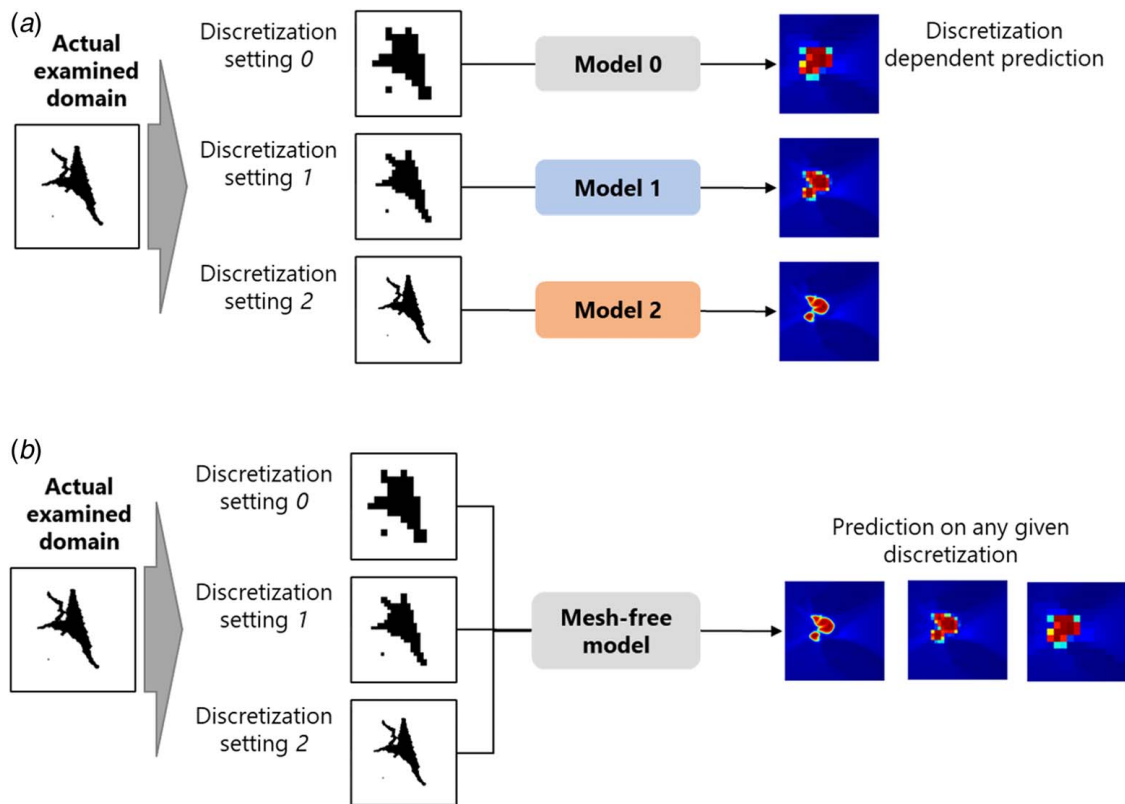


Fig. 3 The difference between discretization (a) variant and (b) invariant models. The discretization variant models are not generalized beyond the discretization setting of training data. Meanwhile, discretization invariant model can be trained on multiple discretization settings and its prediction is not dependent on the discretization of input fields.

trained with simulation data, as observational data from physical experiments are typically limited and expensive. However, there is a substantial difference between the idealized environment of numerical simulations and the real-world environment of physical experiments, which may cause ML models to perform poorly in real-world settings. In such situations, meta-learning can be used to rapidly adapt pre-trained ML models from numerical simulations to real-world environments. Specifically, a meta-learner can be trained to learn both meta-knowledge representing underlying physics laws and specific knowledge representing the operating environment of datasets, i.e., ideal assumptions of numerical simulation or real-world environments. At the deployment stage, meta-learning can then be utilized to quickly adapt ML models on ideal environment assumptions to real-world environments, using just a small amount of real-world experiment data. With this approach, ML models exclusively trained with numerical simulations can be quickly adapted to operate in real-world environments.

3.2 Discretization Matters: The Need for Discretization-Independent Models. Applying ML for multiscale computational modeling also presents another challenge that is the dependence of ML algorithms on the discretization of the analyzed domain. Once trained, most of the common ML techniques, such as CNNs, are limited to a specific discretization setting (Fig. 3(a)) and do not generalize well for settings that differ from the discretization of the training data. This issue of discretization dependence poses a significant challenge for training ML models for multiscale computational modeling.

The dependency of ML algorithms on the discretization requires efforts to determine the optimal discretization setting that ensures convergence properties for all data samples in the dataset. However, the convergence property of a domain under examination

heavily depends on its geometric characteristics, which notably vary across all data samples in the training dataset. Therefore, finding the optimal discretization setting is extremely costly and time-consuming. For this reason, developing discretization-independent (Fig. 3(b)) ML models is preferable as these models can make predictions regardless of the discretization of the examined domain.

Initial efforts have been devoted to developing discretization-independent ML models for computational modeling of physical problems. For instance, Wandel et al. [65] transformed the discretized representation of state variable fields into continuous representations by interpolation with Hermite spline kernels. The Hermite spline constants were stored for each discrete grid point and could be handled by a CNN. This method successfully modeled several complex flow phenomena, such as Karman vortex streets, the Magnus effect, the Doppler effect, interference patterns, and wave reflections. Quantitative assessment also showed that using CNN on a continuous representation of state variables can improve the accuracy of ML models while still maintaining their computational efficiency advantage. In another work, Sun et al. [66] proposed the use of a continuous convolution kernel based on the Zernike polynomial for scalar field regression. This method exhibited an increase in prediction performance compared to the conventional approach with discrete graph convolutions, as well as the capability of making predictions on arbitrary geometric domains with different discretization settings. The utilization of neural operators is another approach that can feature discretization-independent characteristics [67–71]. In this approach, neural networks are trained to learn operators, namely *neural operators*, that map between infinite-dimensional input and output function spaces. Since neural operators are trained to operate on function spaces, their predictions are not affected by discretization settings of examined domain while the accuracy of prediction is still assured [71].

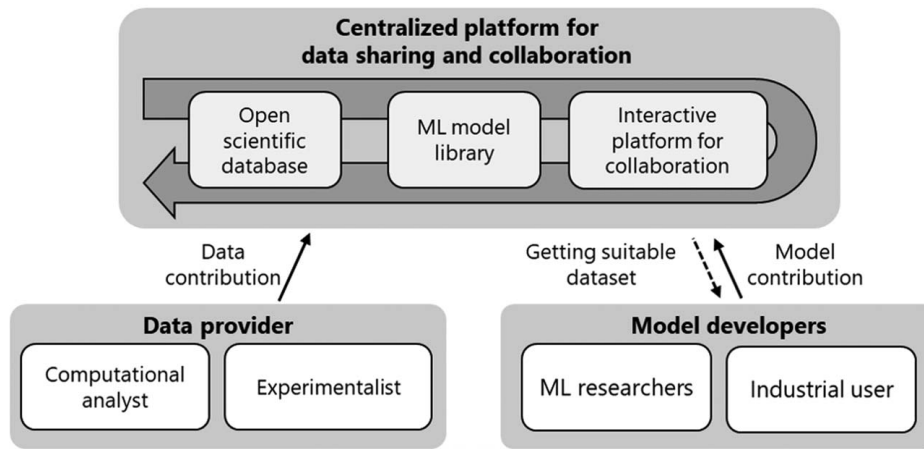


Fig. 4 A centralized platform is necessary for data and model sharing as well as supporting collaboration between data providers and model developers

3.3 Interpretability of Machine Learning Models. ML has demonstrated great potential in multiscale computational modeling, as discussed in previous sections. However, there is still a significant concern about the trustworthiness of these predictions, which prevents their use in real-world applications. This concern arises from the “black-box” nature of ML predictions. ML models are often trained to optimize predefined objective functions that may not fully describe the real-world physics, leading to the possibility of ML models merely “memorizing” observed data but not being aware of or representing the underlying physical phenomena. As reported by Nguyen et al. [17], the physics-naïve architecture of ML models developed for generic purposes, despite having relatively high prediction performance, often fails to capture the underlying physical laws. Due to these reasons, it is crucial to ensure the trustworthiness and physics-awareness of ML predictions before utilizing these models for practical computational modeling.

The need to justify the trustworthiness of ML predictions necessitates the development of methods to interpret ML model predictions, i.e., understanding of the rationale behind ML predictions. Several initial efforts have been made to develop method for interpreting ML predictions in physical science. For instance, Nguyen et al. [17] used saliency map visualization to highlight “critical” microstructure locations that have a high impact on the creation of energy localization in energetic materials modeling. The interpretation results were compared to previous findings, and a high level of agreement between the saliency map identification and the literature on the characteristics of “critical” energetic materials microstructure was demonstrated. In another work, Lellep et al. [72] utilized Shapley additive explanation algorithms to investigate the main factors that affect the ML prediction in fluid dynamics. The authors demonstrated the applicability of the method in various fluid dynamic datasets, including the evaluation of plane Couette flow.

3.4 Centralized Platform for Data Sharing and Collaborations. Large databases, such as ImageNet [73], MNIST [74], and others, play a crucial role in the success of ML in computer vision and language modeling. Similarly, centralized databases offer a potential solution to the challenge of a shortage of high-quality training data for ML algorithms in computational modeling. Moreover, with the advancement of cloud computing technology, these platforms can be expanded for shared and collaborative learning, bringing together computational analysts, experimentalists, and ML researchers to facilitate effective collaboration. Unfortunately, such a platform is currently lacking.

Figure 4 depicts our proposed centralized platform for data sharing and collaboration in ML for computational modeling, which is inspired by OpenML [75], an example of a successful

sharing platform in the ML community. The proposed platform consists of three main components. Similar to several available platforms for scientific data sharing [76,77], our proposed platform includes an open scientific database for sharing and retrieving data. Additionally, ML model libraries are included, allowing for the sharing of previously successful ML models. Finally, the proposed platform includes an interactive environment that allows scientists to collaborate on a particular project. With the above features, the proposed centralized data sharing platform can create a common format and protocol to facilitate data and model sharing. Because engineering data are generally heterogeneous and can come from multiple sources, having such a sharing platform will enrich the dataset that can be used to train ML models, thus helping to overcome the shortage of training data.

Another important feature of the proposed data sharing framework is its ability to facilitate collaborations via an interactive platform. In this collaborative platform, several most suitable datasets and ML models can be suggested to researchers for specific problems. Researchers can then decide whether to utilize the suggested datasets and models or work with their own. Additionally, this platform allows ML developers to request or recommend additional data samples for simulation or experimentation based on the performance of the model on the current dataset. Consequently, experimentalists or computational analysts can confirm the validity or suggest alternate data samples based on their expertise.

Compared to the conventional method of data sharing, this collaborative mechanism offers several advantages. First, this mechanism can reduce the time to collect and preprocess training data since the quality of training datasets suggested by the collaborative platform has been justified via previous training sessions. Additionally, through this mechanism, ML developers can gain access to meaningful and valuable data, thereby improving the resulting ML model’s predictive performance. Finally, this mechanism aids to justify the robustness of ML algorithms. In computational modeling, ML models are typically trained on a single dataset, which can cause ML predictions to be biased toward the training dataset. With this collaborative environment, an ML model can be trained and validated on multiple different datasets, enhancing its robustness and generalizability.

4 Conclusion

Multiscale computational modeling plays a critical role in many mechanical engineering applications, such as computational mechanics, mechanical design and optimization, and manufacturing/process modeling. Despite the promise of capturing complex physics phenomena, the progress in multiscale computational

modeling research is hampered by the high computation cost. To this end, ML can be a viable option that can aid in the analysis of multiscale physical systems by either augmenting or serving as a surrogate for multiscale computational models.

However, several obstacles must be surmounted to successfully apply ML to the computational modeling of multiscale physical systems. The first pressing issue is big data dependency, which requires immediate attention. Unlike traditional ML problems, such as image or text classifications, collecting scientific data is difficult and expensive, particularly for multiscale physical systems. The lack of large, high-quality training datasets can significantly reduce the generalizability and robustness of ML models. Investigating the use of prior knowledge through physics-informed ML, transfer learning, and meta-learning are all viable options to overcome this challenge. Another issue is dealing with the dependency of ML algorithms on the discretization setting of the examined domain. To overcome this issue, developing ML on continuous representation or applying neural operators are potential future directions. In addition to improving the statistical performance of ML-based computational models, it is equally important to validate the physics-awareness and trustworthiness of ML predictions. Finally, the development of a centralized platform for data and model sharing is necessary to facilitate the development of algorithms and foster collaboration in the field of multiscale computational modeling.

Aside from a few pioneering works, the above-mentioned issues are still understudied. Therefore, there is an opportunity for future research to explore better solutions to these challenges. Once these obstacles are overcome, ML has the potential to significantly accelerate the development of multiscale computational modeling and facilitate the discovery of novel and more effective mechanical systems.

Acknowledgment

This paper is based upon work supported by the U.S. Air Force Office of Scientific Research (AFOSR) Multidisciplinary University Research Initiative (MURI) program under Grant No. FA9550-19-1-0318 and by the National Science Foundation (NSF) Designing Materials to Revolutionize and Engineer our Future (DMREF) program under Grant No. 2203580.

Conflict of Interest

There are no conflicts of interest.

Data Availability Statement

No data, models, or code were generated or used for this paper.

References

- [1] Michopoulos, J. G., Farhat, C., and Fish, J., 2005, "Modeling and Simulation of Multiphysics Systems," *ASME J. Comput. Inf. Sci. Eng.*, **5**(3), pp. 198–213.
- [2] Martys, N. S., and Hagedorn, J. G., 2002, "Multiscale Modeling of Fluid Transport in Heterogeneous Materials Using Discrete Boltzmann Methods," *Mater. Struct.*, **35**(10), pp. 650–658.
- [3] Chen, S., Wang, M., and Xia, Z., 2014, "Multiscale Fluid Mechanics and Modeling," *Proc. IUTAM*, **10**, pp. 100–114.
- [4] Fish, J., Wagner, G. J., and Keten, S., 2021, "Mesoscopic and Multiscale Modelling in Materials," *Nat. Mater.*, **20**(6), pp. 774–786.
- [5] Sivapuram, R., Dunning, P. D., and Kim, H. A., 2016, "Simultaneous Material and Structural Optimization by Multiscale Topology Optimization," *Struct. Multidiscipl. Optim.*, **54**(5), pp. 1267–1281.
- [6] Gawad, J., Van Bael, A., Eyckens, P., Samaey, G., Van Houtte, P., and Roose, D., 2013, "Hierarchical Multi-Scale Modeling of Texture Induced Plastic Anisotropy in Sheet Forming," *Comput. Mater. Sci.*, **66**, pp. 65–83.
- [7] Markl, M., and Körner, C., 2016, "Multiscale Modeling of Powder Bed-Based Additive Manufacturing," *Annu. Rev. Mater. Res.*, **46**(1), pp. 93–123.
- [8] Sridhar, A., Kouznetsova, V. G., and Geers, M. G. D., 2016, "Homogenization of Locally Resonant Acoustic Metamaterials Towards an Emergent Enriched Continuum," *Comput. Mech.*, **57**(3), pp. 423–435.
- [9] Oskay, C., and Fish, J., 2007, "Eigen Deformation-Based Reduced Order Homogenization for Failure Analysis of Heterogeneous Materials," *Comput. Meth. Appl. Mech. Eng.*, **196**(7), pp. 1216–1243.
- [10] Yu, Q., and Fish, J., 2002, "Multiscale Asymptotic Homogenization for Multiphysics Problems With Multiple Spatial and Temporal Scales: A Coupled Thermo-Viscoelastic Example Problem," *Int. J. Solids Struct.*, **39**(26), pp. 6429–6452.
- [11] Muralidharan, K., Deymier, P. A., and Simmons, J. H., 2003, "A Concurrent Multiscale Finite Difference Time Domain/Molecular Dynamics Method for Bridging an Elastic Continuum to an Atomic System," *Model. Simul. Mat. Sci. Eng.*, **11**(4), pp. 487–501.
- [12] Fish, J., and Chen, W., 2004, "Discrete-to-Continuum Bridging Based on Multigrid Principles," *Comput. Meth. Appl. Mech. Eng.*, **193**(17), pp. 1693–1711.
- [13] Smith, G. S., Tadmor, E. B., Bernstein, N., and Kaxiras, E., 2001, "Multiscale Simulations of Silicon Nanoindentation," *Acta Mater.*, **49**(19), pp. 4089–4101.
- [14] Nair, A. K., Warner, D. H., Hennig, R. G., and Curtin, W. A., 2010, "Coupling Quantum and Continuum Scales to Predict Crack Tip Dislocation Nucleation," *Scr. Mater.*, **63**(12), pp. 1212–1215.
- [15] Chan, Y.-C., Ahmed, F., Wang, L., and Chen, W., 2021, "METASet: Exploring Shape and Property Spaces for Data-Driven Metamaterials Design," *ASME J. Mech. Des.*, **143**(3), p. 031707.
- [16] Goodfellow, I., Bengio, Y., and Courville, A., 2016, *Deep Learning*, The MIT Press, Cambridge, MA.
- [17] Nguyen, P. C. H., Nguyen, Y.-T., Choi, J. B., Seshadri, P. K., Udaykumar, H. S., and Baek, S., 2022, "PARC: Physics-Aware Recurrent Convolutional Neural Networks to Assimilate Meso-Scale Reactive Mechanics of Energetic Materials," *Sci. Adv.*, **9**(17), p. eadd6868.
- [18] Kollmann, H. T., Abueidda, D. W., Koric, S., Guleryuz, E., and Sobh, N. A., 2020, "Deep Learning for Topology Optimization of 2D Metamaterials," *Mater. Des.*, **196**, p. 109098.
- [19] Brunton, S. L., Noack, B. R., and Koumoutsakos, P., 2020, "Machine Learning for Fluid Mechanics," *Annu. Rev. Fluid Mech.*, **52**(1), pp. 477–508.
- [20] Grant, I., and Pan, X., 1995, "An Investigation of the Performance of Multi-layer, Neural Networks Applied to the Analysis of PIV Images," *Exp. Fluids*, **19**(3), pp. 159–166.
- [21] Teo, C. L., Lim, K. B., Hong, G. S., and Yeo, M. H. T., 1991, "A Neural Net Approach in Analyzing Photograph in PIV," Conference Proceedings 1991 IEEE International Conference on Systems, Man, and Cybernetics, Vol. 3, pp. 1535–1538.
- [22] Wan, Z. Y., Vlachas, P., Koumoutsakos, P., and Sapsis, T., 2018, "Data-Assisted Reduced-Order Modeling of Extreme Events in Complex Dynamical Systems," *PLoS One*, **13**(5), p. e0197704.
- [23] Vlachas, P. R., Byeon, W., Wan, Z. Y., Sapsis, T. P., and Koumoutsakos, P., 2018, "Data-Driven Forecasting of High-Dimensional Chaotic Systems With Long Short-Term Memory Networks," *Proc. R. Soc. A: Math. Phys. Eng. Sci.*, **474**(2213), p. 20170844.
- [24] Kim, B., Azevedo, V. C., Thuerey, N., Kim, T., Gross, M., and Solenthaler, B., 2019, "Deep Fluids: A Generative Network for Parameterized Fluid Simulations," *Comput. Graphics Forum*, **38**(2), pp. 59–70.
- [25] Meredig, B., Agrawal, A., Kirklin, S., Saal, J. E., Doak, J. W., Thompson, A., Zhang, K., Choudhary, A., and Wolverton, C., 2014, "Combinatorial Screening for New Materials in Unconstrained Composition Space With Machine Learning," *Phys. Rev. B*, **89**(9), p. 094104.
- [26] Bhaduri, A., Gupta, A., and Graham-Brady, L., 2022, "Stress Field Prediction in Fiber-Reinforced Composite Materials Using a Deep Learning Approach," *Compos. B: Eng.*, **238**, p. 109879.
- [27] Rao, Z., Tung, P.-Y., Xie, R., Wei, Y., Zhang, H., Ferrari, A., Klaver, T. P. C., et al., 2022, "Machine Learning-Enabled High-Entropy Alloy Discovery," *Science*, **378**(6615), pp. 78–85.
- [28] Nguyen, P. C. H., Vlassis, N. N., Bahmani, B., Sun, W., Udaykumar, H. S., and Baek, S. S., 2022, "Synthesizing Controlled Microstructures of Porous Media Using Generative Adversarial Networks and Reinforcement Learning," *Sci. Rep.*, **12**(1), p. 9034.
- [29] Gupta, A., Bhaduri, A., and Graham-Brady, L., 2022, "Accelerated Multiscale Mechanics Modeling in a Deep Learning Framework," arXiv:2212.14601.
- [30] Raissi, M., Perdikaris, P., and 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**, pp. 686–707.
- [31] Pokharel, R., Pandey, A., and Scheinker, A., 2021, "Physics-Informed Data-Driven Surrogate Modeling for Full-Field 3D Microstructure and Micromechanical Field Evolution of Polycrystalline Materials," *JOM*, **73**(11), pp. 3371–3382.
- [32] Haghighat, E., Raissi, M., Moure, A., Gomez, H., and Juanes, R., 2021, "A Physics-Informed Deep Learning Framework for Inversion and Surrogate Modeling in Solid Mechanics," *Comput. Meth. Appl. Mech. Eng.*, **379**, p. 113741.
- [33] Zhang, X., Xie, F., Ji, T., Zhu, Z., and Zheng, Y., 2021, "Multi-Fidelity Deep Neural Network Surrogate Model for Aerodynamic Shape Optimization," *Comput. Meth. Appl. Mech. Eng.*, **373**, p. 113485.
- [34] Yan, W., Lin, S., Kafka, O. L., Lian, Y., Yu, C., Liu, Z., Yan, J., et al., 2018, "Data-Driven Multi-Scale Multi-Physics Models to Derive Process-Structure-Property Relationships for Additive Manufacturing," *Comput. Mech.*, **61**(5), pp. 521–541.

- [35] Nguyen, P. C. H., Nguyen, Y.-T., Seshadri, P. K., Choi, J. B., Udaykumar, H. S., and Baek, S., 2023, "A Physics-Aware Deep Learning Model for Energy Localization in Multiscale Shock-to-Detonation Simulations of Heterogeneous Energetic Materials," *Propell. Explos. Pyrotech.*, **48**(4), p. e202200268.
- [36] Olivier, A., Shields, M. D., and Graham-Brady, L., 2021, "Bayesian Neural Networks for Uncertainty Quantification in Data-Driven Materials Modeling," *Comput. Meth. Appl. Mech. Eng.*, **386**, p. 114079.
- [37] Koopman, B. O., 1931, "Hamiltonian Systems and Transformation in Hilbert Space," *Proc. Natl. Acad. Sci. U. S. A.*, **17**(5), pp. 315–318.
- [38] Lusch, B., Kutz, J. N., and Brunton, S. L., 2018, "Deep Learning for Universal Linear Embeddings of Nonlinear Dynamics," *Nat. Commun.*, **9**(1), p. 4950.
- [39] Rowley, C. W., Mezic, I., Bagheri, S., Schlatter, P., and Henningson, D. S., 2009, "Spectral Analysis of Nonlinear Flows," *J. Fluid. Mech.*, **641**, pp. 115–127.
- [40] Li, Q., Dietrich, F., Bollt, E. M., and Kevrekidis, I. G., 2017, "Extended Dynamic Mode Decomposition With Dictionary Learning: A Data-Driven Adaptive Spectral Decomposition of the Koopman Operator," *Chaos: Interdiscipl. J. Nonlinear Sci.*, **27**(10), p. 103111.
- [41] Surana, A., 2020, "Koopman Operator Framework for Time Series Modeling and Analysis," *J. Nonlinear Sci.*, **30**(5), pp. 1973–2006.
- [42] Yeung, E., Kundu, S., and Hodas, N. O., 2019, "Learning Deep Neural Network Representations for Koopman Operators of Nonlinear Dynamical Systems," 2019 American Control Conference (ACC), Philadelphia, PA, July 10–12, pp. 4832–4839.
- [43] Takeishi, N., Kawahara, Y., and Yairi, T., 2017, "Learning Koopman Invariant Subspaces for Dynamic Mode Decomposition," Proceedings of the 31st International Conference on Neural Information Processing Systems, Curran Associates Inc., Red Hook, NY, pp. 1130–1140.
- [44] Bendsøe, M. P., and Sigmund, O., 2004, *Topology Optimization*.
- [45] Woldseth, R. V., Aage, N., Bærentzen, J. A., and Sigmund, O., 2022, "On the Use of Artificial Neural Networks in Topology Optimisation," *Struct. Multidiscipl. Optim.*, **65**(10), p. 294.
- [46] White, D. A., Arrighi, W. J., Kudo, J., and Watts, S. E., 2019, "Multiscale Topology Optimization Using Neural Network Surrogate Models," *Comput. Meth. Appl. Mech. Eng.*, **346**, pp. 1118–1135.
- [47] Ghavamian, F., and Simone, A., 2019, "Accelerating Multiscale Finite Element Simulations of History-Dependent Materials Using a Recurrent Neural Network," *Comput. Meth. Appl. Mech. Eng.*, **357**, p. 112594.
- [48] Wang, L., Chan, Y.-C., Ahmed, F., Liu, Z., Zhu, P., and Chen, W., 2020, "Deep Generative Modeling for Mechanistic-Based Learning and Design of Metamaterial Systems," *Comput. Meth. Appl. Mech. Eng.*, **372**, p. 113377.
- [49] Abueidda, D. W., Koric, S., and Sobh, N. A., 2020, "Topology Optimization of 2D Structures With Nonlinearities Using Deep Learning," *Comput. Struct.*, **237**, p. 106283.
- [50] Lin, Q., Hong, J., Liu, Z., Li, B., and Wang, J., 2018, "Investigation Into the Topology Optimization for Conductive Heat Transfer Based on Deep Learning Approach," *Int. Commun. Heat Mass Transfer*, **97**, pp. 103–109.
- [51] Behzadi, M. M., and Ilies, H. T., 2022, "GANTL: Toward Practical and Real-Time Topology Optimization With Conditional Generative Adversarial Networks and Transfer Learning," *ASME J. Mech. Des.*, **144**(2), p. 021711.
- [52] Yang, Z., Li, X., Catherine Brinson, L., Choudhary, A. N., Chen, W., and Agrawal, A., 2018, "Microstructural Materials Design Via Deep Adversarial Learning Methodology," *ASME J. Mech. Des.*, **140**(11), p. 111416.
- [53] Sui, F., Guo, R., Zhang, Z., Gu, G. X., and Lin, L., 2021, "Deep Reinforcement Learning for Digital Materials Design," *ACS Mater. Lett.*, **3**(10), pp. 1433–1439.
- [54] Johnson, N. S., Vulimiri, P. S., To, A. C., Zhang, X., Brice, C. A., Kappes, B. B., and Stebner, A. P., 2020, "Invited Review: Machine Learning for Materials Developments in Metals Additive Manufacturing," *Addit. Manuf.*, **36**, p. 101641.
- [55] Tapia, G., Khairallah, S., Matthews, M., King, W. E., and Elwany, A., 2018, "Gaussian Process-Based Surrogate Modeling Framework for Process Planning in Laser Powder-Bed Fusion Additive Manufacturing of 316L Stainless Steel," *Int. J. Adv. Manuf. Technol.*, **94**(9), pp. 3591–3603.
- [56] Bessa, M. A., Bostanabad, R., Liu, Z., Hu, A., Apley, D. W., Brinson, C., Chen, W., and Liu, W. K., 2017, "A Framework for Data-Driven Analysis of Materials Under Uncertainty: Countering the Curse of Dimensionality," *Comput. Meth. Appl. Mech. Eng.*, **320**, pp. 633–667.
- [57] Wang, Z., Liu, P., Ji, Y., Mahadevan, S., Horstemeyer, M. F., Hu, Z., Chen, L., and Chen, L.-Q., 2019, "Uncertainty Quantification in Metallic Additive Manufacturing Through Physics-Informed Data-Driven Modeling," *JOM*, **71**(8), pp. 2625–2634.
- [58] Roy, S., Rai, N. K., Sen, O., and Udaykumar, H. S., 2020, "Structure-Property Linkage in Shocked Multi-material Flows Using a Level-Set-Based Eulerian Image-to-Computation Framework," *Shock Waves*, **30**(5), pp. 443–472.
- [59] Weiss, K., Khoshgoftaar, T. M., and Wang, D., 2016, "A Survey of Transfer Learning," *J. Big Data*, **3**(1), p. 9.
- [60] Plested, J., and Gedeon, T., 2022, "Deep Transfer Learning for Image Classification: A Survey," arXiv:2205.09904.
- [61] Do, C. B., and Ng, A. Y., 2005, "Transfer Learning for Text Classification," *Advances in Neural Information Processing Systems*, Y. Weiss, B. Schölkopf, and J. Platt, eds., MIT Press, Cambridge, MA.
- [62] Xu, C., Cao, B. T., Yuan, Y., and Meschke, G., 2023, "Transfer Learning Based Physics-Informed Neural Networks for Solving Inverse Problems in Engineering Structures Under Different Loading Scenarios," *Comput. Meth. Appl. Mech. Eng.*, **405**, p. 115852.
- [63] Liu, X., Zhang, X., Peng, W., Zhou, W., and Yao, W., 2021, "A Novel Meta-Learning Initialization Method for Physics-Informed Neural Networks," *Neural Comput. Appl.*, **34**(17), pp. 14511–14534.
- [64] Li, Y., Wang, J., Huang, Z., and Gao, R. X., 2022, "Physics-Informed Meta Learning for Machining Tool Wear Prediction," *J. Manuf. Syst.*, **62**, pp. 17–27.
- [65] Wandel, N., Weinmann, M., Neidlin, M., and Klein, R., 2021, "Spline-PINN: Approaching PDEs Without Data Using Fast, Physics-Informed Hermite-Spline CNNs," arXiv:2109.07143.
- [66] Sun, Z., Rooke, E., Charton, J., He, Y., Lu, J., and Baek, S., 2020, "ZerNet: Convolutional Neural Networks on Arbitrary Surfaces Via Zernike Local Tangent Space Estimation," *Comput. Graphics Forum*, **39**, pp. 204–216.
- [67] Kovachki, N., Li, Z., Liu, B., Azizzadenesheli, K., Bhattacharya, K., Stuart, A., and Anandkumar, A., 2021, "Neural Operator: Learning Maps Between Function Spaces," arXiv:2108.08481.
- [68] Rahman, M. A., Florez, M. A., Anandkumar, A., Ross, Z. E., and Azizzadenesheli, K., 2022, "Generative Adversarial Neural Operators," arXiv:2205.03017.
- [69] Rahman, M. A., Ross, Z. E., and Azizzadenesheli, K., 2022, "U-NO: U-Shaped Neural Operators," arXiv:2204.11127.
- [70] Li, Z., Zheng, H., Kovachki, N., Jin, D., Chen, H., Liu, B., Azizzadenesheli, K., and Anandkumar, A., 2021, "Physics-Informed Neural Operator for Learning Partial Differential Equations," arXiv:2111.03794.
- [71] Li, Z., Kovachki, N., Azizzadenesheli, K., Liu, B., Bhattacharya, K., Stuart, A., and Anandkumar, A., 2020, "Fourier Neural Operator for Parametric Partial Differential Equations," arXiv:2010.08895.
- [72] Lellep, M., Prexl, J., Eckhardt, B., and Linkmann, M., 2022, "Interpreted Machine Learning in Fluid Dynamics: Explaining Relaminarisation Events in Wall-Bounded Shear Flows," *J. Fluid. Mech.*, **942**, p. A2.
- [73] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L., 2009, "ImageNet: A Large-Scale Hierarchical Image Database," 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, June 20–25, pp. 248–255.
- [74] Deng, L., 2012, "The MNIST Database of Handwritten Digit Images for Machine Learning Research [Best of the Web]," *IEEE Signal Process. Mag.*, **29**(6), pp. 141–142.
- [75] Vanschoren, J., van Rijn, J. N., Bischl, B., and Torgo, L., 2014, "OpenML: Networked Science in Machine Learning," *SIGKDD Explor. Newsl.*, **15**(2), pp. 49–60.
- [76] Zhao, H., Li, X., Zhang, Y., Schadler, L. S., Chen, W., and Brinson, L. C., 2016, "Perspective: NanoMine: A Material Genome Approach for Polymer Nanocomposites Analysis and Design," *APL Mater.*, **4**(5), p. 53204.
- [77] Hu, J., Stefanov, S., Song, Y., Ome, S. S., Louis, S.-Y., Siriwardane, E. M. D., Zhao, Y., and Wei, L., 2022, "MaterialsAtlas.Org: A Materials Informatics Web App Platform for Materials Discovery and Survey of State-of-the-Art," *NPJ Comput. Mater.*, **8**(1), p. 65.