NRC·CNRC

NRC Publications Archive
Archives des publications du CNRC

Beyond order: perspectives on leveraging machine learning for disordered materials

Yazdani Sarvestani, Hamidreza; Nadigotti, Surabhi; Fatehi, Erfan; Aranguren van Egmond, Derek; Ashrafi, Behnam

This publication could be one of several versions: author's original, accepted manuscript or the publisher's version. / La version de cette publication peut être l'une des suivantes : la version prépublication de l'auteur, la version acceptée du manuscrit ou la version de l'éditeur.

For the publisher's version, please access the DOI link below./ Pour consulter la version de l'éditeur, utilisez le lien DOI ci-dessous.

Publisher's version / Version de l'éditeur:

https://doi.org/10.1002/adem.202402486

Advanced Engineering Materials, 2025-01-18

NRC Publications Archive Record / Notice des Archives des publications du CNRC :

https://nrc-publications.canada.ca/eng/view/object/?id=674480c6-2874-4673-b19c-8289d002a497

https://publications-cnrc.canada.ca/fra/voir/objet/?id=674480c6-2874-4673-b19c-8289d002a497

Access and use of this website and the material on it are subject to the Terms and Conditions set forth at https://nrc-publications.canada.ca/eng/copyright

READ THESE TERMS AND CONDITIONS CAREFULLY BEFORE USING THIS WEBSITE.

L'accès à ce site Web et l'utilisation de son contenu sont assujettis aux conditions présentées dans le site https://publications-cnrc.canada.ca/fra/droits

LISEZ CES CONDITIONS ATTENTIVEMENT AVANT D'UTILISER CE SITE WEB.

Questions? Contact the NRC Publications Archive team at PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca. If you wish to email the authors directly, please see the first page of the publication for their contact information.

Vous avez des questions? Nous pouvons vous aider. Pour communiquer directement avec un auteur, consultez la première page de la revue dans laquelle son article a été publié afin de trouver ses coordonnées. Si vous n'arrivez pas à les repérer, communiquez avec nous à PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca.

National Research Council Canada | Conseil national de recherches Canada

Canada

Check for updates

ADVANCED ENGINEERING MATERIALS
www.aem-journal.com

Beyond Order: Perspectives on Leveraging Machine Learning for Disordered Materials

Hamidreza Yazdani Sarvestani, Surabhi Nadigotti, Erfan Fatehi, Derek Aranguren van Egmond, and Behnam Ashrafi*

Disordered structures, characterized by their lack of periodicity, present significant challenges in fields such as materials science and biology. Conventional methods often fall short of capturing the intricate properties and behaviors of these complex systems. For example, the prediction of material properties in amorphous polymers and high-entropy alloys has historically been inaccurate due to their inherent disorder, which arises from the probabilistic nature of structural defects and nonuniform atomic arrangements. However, the rise of machine learning (ML) offers a revolutionary approach to understanding and predicting the behavior of disordered materials. This perspective article explores how ML techniques, including neural networks and generative models, provide unprecedented insights into materials with inherent disorder, driving advances in industries such as energy storage, drug discovery, and structural engineering. By leveraging powerful algorithms, researchers can now predict structural properties, identify hidden patterns, and accelerate the discovery of novel materials. Case studies illustrate the ability of ML to overcome data scarcity, enhance model reliability, and enable real-time analysis of disordered structures. While challenges such as data quality and computational costs remain, the integration of ML with traditional methods marks a transformative leap in our ability to navigate the disordered landscape, setting the stage for ground-breaking discoveries.

1. Introduction

1.1. Overview of Disordered Materials and Structures

Disordered materials and structures are ubiquitous in nature, emerging as the result of biological growth phenomena and energy distribution. In engineering, synthetic disordered materials have recently become a unique subset of architected materials. Here, they are characterized by their nonperiodic representative volume elements (RVEs), distinguishing them from more conventional periodic materials. Their lack of a clear pattern grants them unique mechanical and physical properties, making them advantageous in fields such as bioinspired systems and advanced metamaterials.[1–4] For instance, stochastic metamaterials in engineering are defined by their nonrepeating connectivity, constructed from lattices without a defining unit cell, where no two nodes share identical orientations or beam connections.[5] This randomness results in materials with enhanced robustness, adaptability, and other tailored properties not achievable through conventional designs. However, characterizing and understanding these systems is challenging due to their complex geometry, which arises from irregular and nonperiodic structures that lack symmetry, requiring extensive sampling across multiple scales to accurately define their RVEs. This challenge is akin to analyzing the mechanical properties of ceramic systems, where the statistical impact of a random flaw distribution within the material must be accounted for, necessitating detailed analysis to capture the variability and ensure accurate predictions[6,7] (see **Figure 1**).

1.2. Formation and Design of Disordered Metamaterials

Disordered metamaterials are formed by introducing randomness into their structural arrangements, often through advanced techniques such as additive manufacturing (AM) or stochastic design algorithms. These materials lack a defining unit cell, instead featuring nonrepeating connectivity and geometric irregularities that contribute to their unique mechanical and functional properties. For instance, randomizing the connectivity between nodes and beams creates isotropic behaviors and

H. Yazdani Sarvestani, S. Nadigotti, B. Ashrafi
Aerospace Manufacturing Technology Center
National Research Council Canada
5145 Decelles Avenue, Montreal, QC H3T 2B2, Canada
E-mail: hamidreza.yazdani@nrc-cnrc.gc.ca

E. Fatehi
Digital Technologies Research Center
National Research Council Canada
5145 Decelles Avenue, Montreal, QC H3T 2B2, Canada

D. Aranguren van Egmond
Advanced Materials Research Facility
Clean Energy Innovation Research Center
National Research Council Canada
2620 Speakman Drive, Mississauga, ON L5K 1B4, Canada

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/adem.202402486.

© 2025 The Author(s). Advanced Engineering Materials published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/adem.202402486

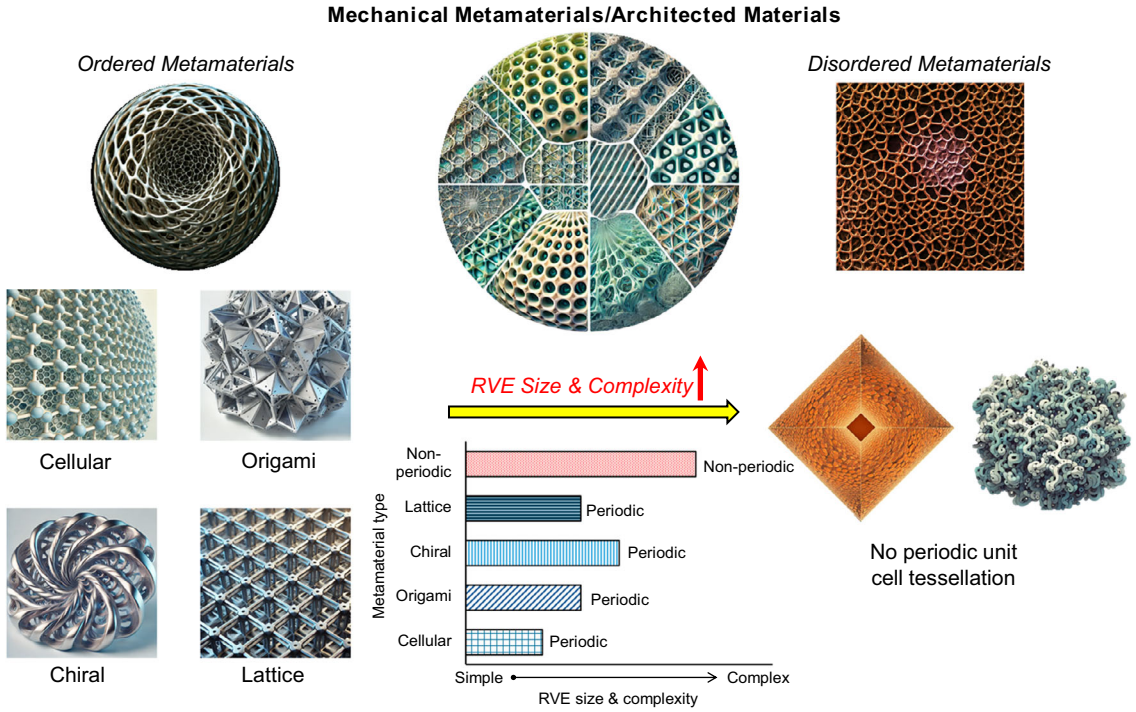

Figure 1. Categorizing mechanical metamaterials/architected materials into ordered and disordered types. Ordered metamaterials, such as cellular, origami, chiral, and lattice structures, feature periodic unit cell tessellation. Disordered metamaterials, lacking this periodicity, often exhibit larger RVEs and increased complexity. This classification highlights the diverse structural arrangements within the field of metamaterials.

robustness under complex loading conditions, while incorporating hierarchical structures blends macrolevel randomness with microlevel variations, enabling tunable multifunctional performance. AM further facilitates the precise fabrication of highly complex and irregular designs, making it a key technology for producing disordered metamaterials. Structural design rules for these materials typically include introducing stochastic variations in connectivity to enhance energy absorption, optimizing porosity and randomness to control mechanical strength and thermal conductivity, and layering hierarchical configurations to achieve enhanced performance across scales.[8]

Machine learning (ML) plays a transformative role in this process by exploring vast configuration spaces and predicting optimal designs. Generative models and reinforcement learning algorithms can train on datasets of existing designs to identify configurations that maximize desired attributes, such as strength, energy dissipation, or thermal resistance. By iteratively refining designs through simulation or experimental feedback, ML accelerates the discovery and optimization of disordered metamaterials, unlocking new possibilities for advanced material design.[9]

1.3. Importance of Disordered Structures in Various Fields

1.3.1. Biological Systems

In biology,[10,11] materials such as bone,[12,13] skin,[14] and turtle shells[15] exemplify how natural systems leverage disorder for enhanced mechanical performance, such as toughness and adaptability. These biological systems often inspire the design of disordered metamaterials that mimic their capacity to distribute stress and deformation unpredictably, which enhances durability under mechanical loads.[16–18]

1.3.2. Medicine

In medicine, intrinsically disordered proteins (IDPs) play a critical role in drug discovery.[19] Their flexibility and adaptable binding sites enable interactions with a wide range of molecular targets, making them crucial for therapies in cancer, neurological diseases, and beyond.[20] The ability to manipulate and model such disordered systems offers a pathway to developing more effective, targeted treatments, showcasing the broader implications of understanding disorder at both the structural and molecular levels[21,22] (see **Figure 2**).

1.3.3. Engineering Materials

In engineering, disordered structures such as foams represent an important category of materials with diverse applications.[23] Structural foams, for instance, are lightweight materials characterized by irregular cellular structures, which grant them excellent mechanical properties such as high energy absorption, thermal insulation, and low density.[24] These properties make them invaluable in applications ranging from aerospace to construction and packaging. The disorder in these materials allows for tunable properties, where pore size distribution, connectivity, and density can be manipulated to meet specific requirements.

© 2025 The Author(s). Advanced Engineering Materials published by Wiley-VCH GmbH

15272648, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adem.202402486 by National Research Council, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

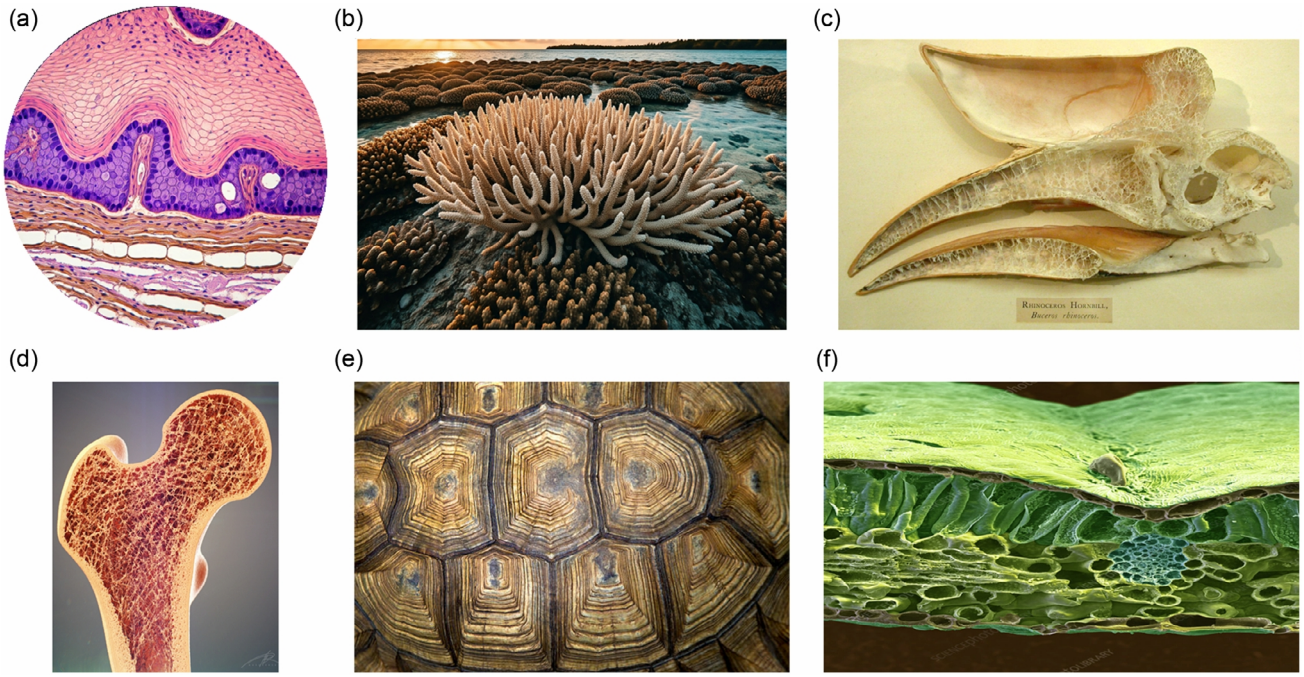

Figure 2. Diverse structural architectures found in natural materials and biological structures: a) human skin, b) coral, c) rhinoceros hornbill beak,[133] d) human bone,[134] e) tortoise shell,[135] and f) plant leaf[136] each demonstrate unique, intricate designs that contribute to their function. These examples highlight the inspiration that nature can provide for the development of innovative materials and structures.

The stochastic nature of foam structures also poses challenges in modeling and simulation, requiring advanced techniques such as finite element analysis (FEA) and machine learning to accurately predict their behavior under varying conditions.[25] Expanding the understanding and control of disordered engineering materials like foams has the potential to further improve their functionality and broaden their applications.[26]

1.4. Challenges in Studying Disordered Materials & Structures

The complexity of disordered structures poses significant challenges to traditional analytical methods.[27] For example, in materials science, amorphous polymers and metallic glasses cannot be accurately characterized using techniques like X-ray diffraction (XRD), which is better suited for periodic crystalline materials. In such cases, machine learning models have demonstrated the ability to identify key features, such as pore size distribution and atomic arrangement, from complex datasets. It is important to note that these features are not equivalent to data obtained through analytical methods like XRD. Instead, they serve as "proxy" features, offering alternative pathways to extract insights that may complement or parallel those derived from traditional methods. A comparison of these techniques is shown in **Table 1**, which outlines the limitations of conventional methods versus the strengths of ML approaches for studying disordered materials.[28] This is particularly evident in amorphous systems, where the differentiation between crystalline and amorphous phases can be subjective and prone to error. Variations in measurement, such as peak intensity distortions or sample preparation inconsistencies, further complicate analysis.[29] To overcome these limitations, the application of ML techniques has emerged as a powerful tool. ML models, especially deep learning algorithms like convolutional neural networks (CNNs), are capable of identifying hidden patterns in large, complex datasets where

Table 1. Comparison of traditional versus ML approaches in the study of disordered structures. A summary of key attributes where traditional methods and ML differ in effectiveness.

Attribute	Traditional approach	ML approach
Time efficiency	Labor-intensive; slower due to manual processes	High speed; time-efficient with automated processes
Data requirements	High-quality, large datasets required for analysis	Adaptable to varied data quality and smaller datasets
Scalability	Limited scalability with large datasets	Highly scalable; efficient with large and complex datasets
Accuracy	Varies with methods; often lower due to data handling limits	Higher accuracy due to advanced algorithms
Interpretability	Direct and interpretable; simpler models	Challenging to interpret; ongoing improvements in XAI

© 2025 The Author(s). Advanced Engineering Materials published by Wiley-VCH GmbH

15272648, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adem.202402486 by National Research Council, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

traditional methods fall short. These algorithms excel at uncovering the intricate correlations within disordered systems, enabling researchers to make sense of data that previously defied interpretation. As a result, ML offers a fresh perspective on how we study and understand the behavior of disordered materials.

1.5. Introduction to ML in Scientific Research

The role of ML in scientific discovery is rapidly expanding, driven by its ability to handle complex, high-dimensional data across a wide range of fields.[22,30,31] Supervised learning is particularly valuable for tasks such as predicting material properties from labeled datasets, while unsupervised learning excels at discovering patterns and structures within unclassified data. Deep learning, with techniques like CNNs, is especially effective in domains involving image data, such as X-ray and electron microscopy used in material science. In addition, generative models, such as variational autoencoders (VAEs) and generative adversarial networks (GANs), have opened up new possibilities in generating candidates of both molecular-scale materials and macroscale porous and disordered materials. By generating plausible structural models and simulating the behavior of complex systems, these techniques enable the discovery of novel materials with desired properties, even when real-world experimental data is limited. In effect, the use of ML for these purposes helps scientists and engineers to zoom in and examine the role of "local defects" on bulk properties in either molecular or macroscopic structural scales. Thus, allowing a new type of "machine intuition" to augment our understanding of these large design spaces where scientific intuition may fall short.[32] ML is thus revolutionizing the way researchers approach the study of disordered structures. It not only enhances our understanding of existing materials but also accelerates the discovery of new, optimized materials for applications in fields as diverse as energy storage, drug design, and advanced manufacturing.[33]

Despite the significant advancements in ML and its applications in various fields, there remains a substantial gap in understanding and leveraging ML for disordered structures. This perspective article has aimed to bridge this gap by providing a comprehensive overview of the current state of research, highlighting the challenges and opportunities, and outlining potential future directions. By exploring the applications of ML techniques, discussing data handling and model development, and addressing the limitations and challenges, this perspective article offers a valuable resource for researchers and practitioners interested in advancing the field of disordered materials.[30]

Traditional methods for analyzing disordered structures face several limitations due to their dependency on extensive data, manual processes, and limited scalability with complex datasets. ML offers distinct advantages, including speed, adaptability to diverse data qualities, and higher predictive accuracy. Table 1 summarizes key differences between traditional and ML approaches, highlighting how ML addresses specific limitations inherent in conventional methods.

2. Applications of ML in Disordered Structures

2.1. Functionality Advantages of Disordered Structures

Disordered structures exhibit unique functional advantages over their ordered counterparts due to their intrinsic randomness and irregularity. These features enable superior material properties that are often unattainable with traditional periodic designs, making disordered structures highly valuable across a range of applications

2.1.1. Mechanical Strength and Robustness

The stochastic nature of disordered materials such as metallic glasses and randomized lattices provides enhanced tolerance to localized defects. Unlike ordered structures that propagate stress along regular paths, the irregular configurations in disordered structures diffuse stress, improving toughness and durability.[34,35]

2.1.2. Energy Absorption and Dissipation

Cellular foams and disordered metamaterials excel in energy absorption due to their irregular pore distribution and heterogeneous densities. These characteristics make them highly effective in applications such as impact protection, vibration damping, and acoustic insulation.[23,24]

2.1.3. Thermal and Electrical Properties

Disordered materials, such as amorphous semiconductors, exhibit unique thermal and electronic behavior due to the absence of periodicity. For example, amorphous silicon offers reduced thermal conductivity, making it a preferred choice for thermal insulation applications.[36,37]

2.1.4. Environmental Adaptability

Inspired by biological systems, disordered metamaterials can adapt to dynamic environmental conditions. Their nonuniform configurations allow for multifunctionality, including resistance to deformation and dynamic load redistribution.[38]

ML further enhances these functional advantages by optimizing disordered designs for specific properties. Techniques like GANs and reinforcement learning identify configurations that maximize mechanical, thermal, or energy dissipation properties. These advancements are paving the way for next-generation materials tailored to meet increasingly demanding performance criteria.

2.2. ML Techniques Relevant to Disordered Structures

Disordered structures, with their inherent complexity, require advanced analytical methods capable of deciphering their unpredictable behaviors. Deep learning models like CNNs, generative models, and VAEs have shown remarkable promise in understanding these systems. For example, in a recent study on

© 2025 The Author(s). Advanced Engineering Materials published by Wiley-VCH GmbH

15272648, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adem.202402486 by National Research Council, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

amorphous metallic alloys, CNNs were able to achieve a 10% improvement in predicting mechanical properties compared to traditional molecular dynamics (MD) simulations. Such improvements highlight the potential of ML in enhancing material property predictions, significantly reducing computation times while maintaining accuracy. These techniques provide several advantages for studying nonperiodic structures.[39–41]

For example, CNNs excel in image recognition and pattern detection, making them ideal for analyzing the geometries of porous and cellular materials. When applied to disordered structures, CNNs can identify key structural features that humans would likely miss, which influence properties such as strength, stiffness, and durability. Generative models, like deep convolutional generative adversarial networks (DCGANs), further expand the analysis by simulating realistic 3D structures from 2D images. These models not only reduce the computational load but also allow researchers to generate new material structures that meet specific design criteria, optimizing their properties for targeted applications.

In the context of porous materials, generative models combined with Navier–Stokes (NS)[42] or lattice Boltzmann methods (LBM)[43] have been successfully used to predict permeability and porosity, crucial properties for energy storage and filtration applications. By leveraging ML to process large datasets, researchers can gain deeper insights into the disordered microstructures of materials, allowing for more accurate predictions and efficient design processes.[44]

2.3. Predicting Structural Properties of Disordered Materials

One of the key contributions of ML to the study of disordered structures is its ability to predict material properties from complex, disordered data. Traditional methods often struggle to provide accurate predictions due to the lack of periodicity in disordered materials, as this irregularity makes it difficult to apply conventional models that rely on symmetry, repetitive patterns, or defined unit cells for calculations. Disordered systems often exhibit a high degree of variability at both atomic and macroscales, leading to challenges in accurately modeling their properties using deterministic approaches. However, ML models have demonstrated the ability to infer critical properties such as mechanical strength, thermal stability, and glass transition temperatures (T_g) in amorphous materials and complex alloys by uncovering hidden patterns and correlations in large, complex datasets.[45–51]

2.3.1. ML Model Prediction Process

The ML model prediction process for disordered materials involves several critical steps, each contributing to the final predictive accuracy and reliability

Data Preprocessing: Raw data from experiments or simulations often contains noise, missing values, or inconsistencies. Data preprocessing ensures that the input data is clean and standardized. For disordered materials, this includes normalizing property values, filling gaps in experimental data, and removing outliers that may skew predictions. In cases of limited datasets, techniques such as data augmentation or synthetic data generation using GANs can be employed to expand the training dataset.[52–54]

Feature Selection and Engineering: The structural complexity of disordered materials requires identifying the most relevant features that correlate with the target properties. Features such as atomic coordination, pore size distribution, and local stress fields are extracted using advanced analytical methods or simulation outputs. Dimensionality reduction techniques, such as principal component analysis (PCA), are often used to retain only the most informative features, improving both model efficiency and interpretability.[55]

Model Training: Training ML models involves optimizing parameters to minimize prediction errors. For disordered materials, commonly used models include random forests, neural networks, and graph neural networks (GNNs). The training process often includes hyperparameter tuning, such as adjusting the learning rate or regularization coefficients, to balance model complexity and generalization. Physics-informed constraints are sometimes incorporated into the training phase to align predictions with established physical principles.[52,56]

Model Validation and Testing: Cross-validation techniques, such as *k*-fold validation, are used to assess the generalizability of the ML model. For disordered materials, validation often includes comparison to independent experimental results or hybrid simulations (e.g., combining MD with density functional theory (DFT)). This ensures that the model accurately captures the variability inherent in disordered systems and performs well on unseen data.[57]

2.3.2. Applications and Impact

For instance, ML techniques have been employed to predict T_g of amorphous molecular materials, which are commonly used in electronic applications due to their flexibility and tunable properties. By combining MD simulations with custom force field parameters, researchers can accurately predict T_g, helping to determine the long-term stability and performance of these materials.[58–60] The ML-driven prediction process involves force field parameterization, MD simulations,[61,62] and bilinear regression fitting to identify optimal temperature intervals for T_g estimation. This allows for a balance between computational efficiency and predictive accuracy. Such predictive capabilities have also extended to the mechanical properties of amorphous metals and high-entropy alloys (HEAs). ML models, trained on experimental data, have successfully predicted how these disordered systems will behave under stress, allowing for the rapid design of alloys with enhanced strength and durability, without the need for exhaustive experimental trials.[63–67]

2.4. Generalization Ability of ML Models in Disordered Materials

The ability of ML models to generalize across different types of disordered structural materials is critical for their applicability in predicting material properties under unseen conditions. Generalization is influenced by the diversity of the training dataset, the choice of model architecture, and the robustness of the validation process.

© 2025 The Author(s). Advanced Engineering Materials published by Wiley-VCH GmbH

15272648, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adem.202402486 by National Research Council, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2.4.1. Diversity in Training Data

For disordered materials, ensuring a diverse and representative training dataset is essential. Disordered systems, such as amorphous polymers, metallic glasses, and HEAs, exhibit significant variability in structural configurations and properties. Training datasets that include a broad range of these variations improve the model's ability to make accurate predictions for new materials. For instance, including datasets from different synthesis methods or environmental conditions enhances the model's adaptability to unseen scenarios.[52,68]

2.4.2. Model Architecture and Features

Advanced ML architectures, such as GNNs and ensemble methods, are particularly effective for disordered materials because they can capture complex relationships between features like atomic coordination, pore size distribution, or local stress fields. Models that incorporate domain-specific features, such as bond angles or local density, show improved performance when generalizing across materials with different degrees of disorder.[69]

2.4.3. Validation Techniques for Generalization

Cross-validation methods, such as leave-one-out validation or *k*-fold cross-validation, are commonly used to evaluate model performance. For disordered materials, additional techniques like transfer learning and multifidelity modeling can assess how well a model predicts material properties in conditions or configurations not represented in the training set. These methods involve fine tuning models trained on ordered systems to adapt to disordered counterparts, or combining high-accuracy simulation data with lower-fidelity experimental data to enhance predictions.[57,70]

2.4.4. Challenges in Generalization

The inherent randomness and lack of periodicity in disordered materials pose unique challenges for generalization. Unlike ordered materials, where properties are often dictated by unit cells or periodic features, disordered systems require models to account for stochastic and hierarchical structures. For instance, predicting the mechanical properties of porous disordered materials may fail in configurations with extreme porosity or irregular defect distributions unless these cases are included in the training.[36]

2.4.5. Recent Advances Enhancing Generalization

Recent advancements, such as uncertainty quantification techniques (e.g., Bayesian neural networks), provide confidence intervals for predictions, helping researchers assess model reliability under unseen conditions. Additionally, active learning frameworks, where the model iteratively queries new data points from unexplored regions of the design space, improve its adaptability to new materials or scenarios.[71,72]

2.5. Applications and Case Studies

2.5.1. Amorphous Polymers

ML models trained on polymers with varying cross-link densities have shown excellent generalization in predicting mechanical properties for new compositions under different temperatures and loading conditions.[73]

2.5.2. Metallic Glasses

Transfer learning has been used to adapt models trained on crystalline alloys to predict phase stability and toughness in metallic glasses, achieving accurate results despite the structural differences.[74]

2.5.3. High-Entropy Alloys

Ensemble models have demonstrated robust performance in predicting phase diagrams and mechanical properties across diverse HEAs, even for compositions outside the training set.[75]

By addressing the challenges of diversity, architectural choice, and validation techniques, ML models can significantly improve their generalization ability, enabling reliable predictions for unseen disordered materials and conditions.

2.6. Evaluating ML Models for Revealing Hidden Patterns in Disordered Systems

ML models are increasingly recognized for their ability to reveal hidden patterns in disordered systems, providing insights that enhance our understanding of these complex materials.[76] The effectiveness of ML models in achieving this goal can be evaluated using both practical examples and quantitative metrics, ensuring reliability and interpretability.[77]

2.6.1. Practical Examples

Microstructural Analysis of Metallic Glasses: CNNs have been applied to high-resolution 3D imaging data to identify microstructural defects in metallic glasses. These defects, which are challenging to detect using conventional microscopy, play a critical role in determining mechanical properties such as yield strength and fracture toughness. Insights from CNNs have enabled the optimization of metallic glass compositions for improved performance in biomedical implants and structural components.[78]

Atomic Correlations in Amorphous Polymers: GNNs have been used to analyze atomic-scale data in amorphous polymers, uncovering key bonding patterns that influence thermal stability and flexibility. By correlating local atomic arrangements with macroscopic properties, GNNs have guided the design of advanced polymers for flexible electronics and high-temperature insulation.[73]

2.6.2. Quantitative Metrics to Assess Effectiveness

To ensure the reliability of ML-driven insights, a robust evaluation framework is necessary. Commonly used metrics include

© 2025 The Author(s). Advanced Engineering Materials published by Wiley-VCH GmbH

15272648, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adem.202402486 by National Research Council, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

the following:[79] accuracy: measures the proportion of correct predictions out of all predictions made by the model; precision: reflects the model's ability to avoid false positives by measuring the proportion of true positive predictions among all positive predictions; recall (sensitivity): indicates the model's ability to detect all relevant patterns, minimizing false negatives; F1 score: combines precision and recall into a single metric, providing a balanced evaluation of the model's performance; and area under the receiver operating characteristic curve: captures the trade-off between true positive and false positive rates, offering insights into classification accuracy across different thresholds.

2.6.3. Advancing Understanding of Disordered Materials

The combination of powerful ML models and rigorous evaluation metrics not only improves predictive accuracy but also deepens our understanding of disordered systems. By quantifying patterns and correlations in datasets, ML models can inform material design and optimization, accelerating the development of disordered materials with tailored properties for applications ranging from energy storage to structural engineering.[80]

2.7. Accelerating Material Discovery and Inverse Design

ML has revolutionized material discovery, particularly for disordered systems where high-throughput screening is required.[80] Traditional experimental and computational approaches are often constrained by the vast compositional space and structural variability in disordered materials. ML overcomes these limitations by learning from existing datasets and rapidly predicting potential candidates for new materials, significantly accelerating the discovery process. For example, ML has been successfully applied in the discovery of new catalysts and energy storage materials.[81] In this process, ML models first learn the relationships between material composition and performance, then predict the behavior of novel compositions that have yet to be tested. These predictions are often validated using techniques like DFT[82] or MD simulations, streamlining the screening process. By integrating ML with traditional computational methods, researchers can dramatically reduce the time and resources required to identify new high-performance materials. The ability to rapidly screen vast compositional spaces has made ML an invaluable tool in designing disordered materials with specific properties, whether for mechanical robustness, energy efficiency, or environmental resilience.[80]

One of the key reasons for the accuracy and reliability of ML models in disordered materials discovery is their ability to process complex, nonlinear relationships in large datasets. Techniques like neural networks, random forests, and graph-based models excel at uncovering hidden patterns that are often missed by conventional methods. Regularization techniques, such as dropout and weight decay, further enhance the generalizability of these models, preventing overfitting and ensuring consistent performance on unseen data. Additionally, integrating domain-specific knowledge into the feature engineering process helps reduce noise and improve prediction quality, particularly for disordered systems where structural variability is high.

To overcome the challenge of limited datasets—a common issue for disordered materials—ML models often leverage data augmentation and synthetic data generation techniques. GANs and VAEs are widely used to expand training datasets while maintaining fidelity to real-world properties, thereby improving model robustness. Moreover, hybrid approaches combining ML with physics-based methods, such as MD and DFT, ensure that ML predictions are grounded in established physical principles, further enhancing reliability.

Beyond discovery, ML has enabled inverse design, where models are trained to achieve specific target properties by working backward from desired outcomes to identify the optimal material compositions and structures.[83,84] Unlike traditional trial-and-error approaches, inverse design leverages generative models, reinforcement learning, and optimization techniques to explore the vast design space efficiently. For disordered materials, this means the ability to specify mechanical, thermal, or electrical properties and allow ML algorithms to propose candidate configurations that meet these criteria. Such capabilities significantly accelerate innovation and open new possibilities for tailoring disordered materials for specialized applications.[85–89]

As these techniques continue to advance, the potential for ML to drive innovation in material science is bound to expand, opening new frontiers in the study of disordered systems (see **Figure 3**).

2.8. Differences in Applying ML to Disordered versus Ordered Structural Design

The application of ML to disordered structural design presents distinct challenges and opportunities compared to its use in ordered systems. While both involve optimizing material properties and structural configurations, disordered structures require specialized approaches due to their inherent randomness and lack of periodicity.

2.8.1. Data Characteristics

Ordered structures often benefit from well-defined and relatively smaller datasets due to their periodicity and symmetry. For example, crystal structures can be described using unit cells, simplifying ML input data. In contrast, disordered structures demand larger, more diverse datasets to capture the variability and stochastic nature of their configurations, such as random atomic arrangements or heterogeneous microstructures.[27]

2.8.2. Feature Representation

ML models for ordered structures rely on features such as symmetry descriptors, lattice constants, or bandgaps, which are straightforward to compute from their periodicity. For disordered materials, the absence of symmetry requires more complex representations, such as graph-based models (e.g., GNNS) to capture irregular atomic connectivity, or voxel-based representations for microstructural heterogeneity.[27]

© 2025 The Author(s). Advanced Engineering Materials published by Wiley-VCH GmbH

15272648, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adem.202402486 by National Research Council, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

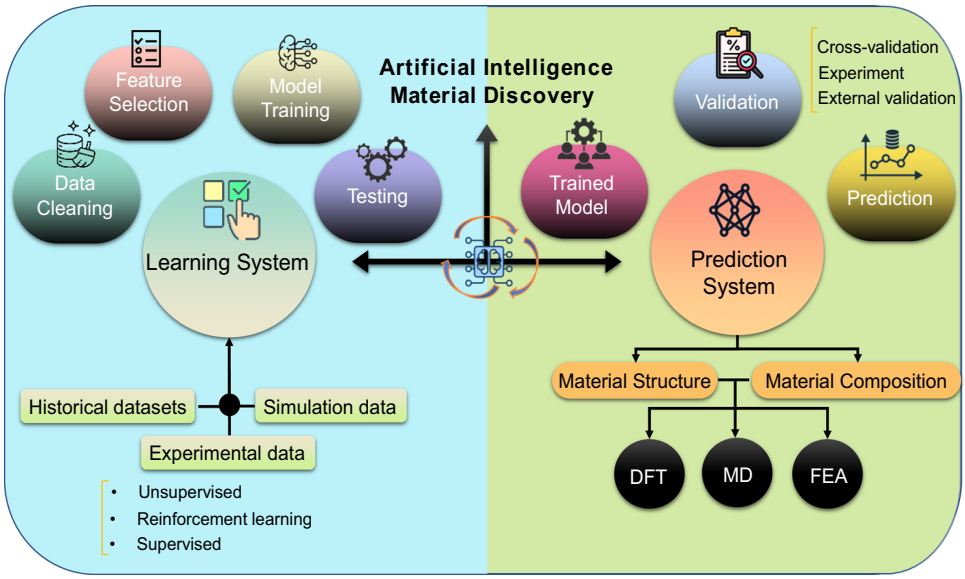

Figure 3. Outlines the workflow of ML in material discovery: Starting with data cleaning and feature selection, ML models are trained and tested to predict material properties based on composition and structure. DFT provides theoretical calculations to enhance accuracy. Feedback loops allow for continuous model improvement, ultimately accelerating the discovery of novel materials.

2.8.3. Optimization Challenges

Ordered designs often involve constrained optimization in a well-defined design space, leveraging known relationships between structure and properties. Disordered designs, on the other hand, require exploration of a vast and less predictable configuration space. Generative models, such as VAEs or GANs, are particularly effective for proposing viable disordered configurations, making ML indispensable in navigating this complexity.

2.8.4. Validation and Interpretability

In ordered systems, ML predictions can often be validated using traditional analytical methods such as XRD or computational techniques like DFT. For disordered materials, validation is more challenging, requiring multiscale simulations (e.g., combining MD with continuum models) or experimental techniques like high-resolution microscopy to ensure accuracy.

2.8.5. Performance Advantages

Disordered materials often exhibit emergent properties, such as enhanced toughness, energy absorption, or multifunctional adaptability, which are difficult to achieve in ordered systems. ML enables the design of such materials by identifying nonintuitive relationships and leveraging the randomness inherent in disordered configurations to maximize performance.[5]

The application of ML to disordered structures demands tailored approaches that embrace their randomness and complexity. By leveraging advanced models and diverse datasets, ML offers unique advantages for exploring and optimizing disordered designs, unlocking their full potential for advanced applications.

3. Data Handling and Preprocessing

Effective data handling and preprocessing are critical for developing reliable ML models, particularly for disordered structures. The limited availability of high-quality data for disordered materials arises from several factors. Unlike periodic materials, disordered structures lack uniformity and symmetry, which makes them difficult to characterize consistently using traditional analytical methods. Experimental data collection often requires advanced techniques, such as 3D imaging or high-resolution microscopy, which are both labor intensive and expensive. Additionally, the inherent variability of disordered systems necessitates large datasets to capture the statistical distribution of their properties, further compounding the challenge of obtaining comprehensive and high-quality data.

To address these limitations, advanced techniques such as GANs and transfer learning have emerged as practical solutions. GANs generate synthetic datasets that mimic the distribution of real data, allowing for the expansion of training datasets without requiring costly experimental efforts. For example, GANs have been successfully applied to reconstruct 3D microstructures of porous materials, capturing intricate details such as pore size distributions and defect patterns. By augmenting training datasets with these synthetic structures, ML models achieve improved generalization and predictive accuracy, even in data-scarce scenarios.

Transfer learning further mitigates data scarcity by enabling ML models trained on well-established datasets, such as those of crystalline systems, to adapt to disordered materials. This technique reduces the dependency on large, labeled datasets specific to disordered systems and has been particularly effective in fields where experimental data is difficult or costly to obtain. By leveraging knowledge from pre-trained models, transfer learning accelerates the adaptation of ML models for amorphous or disordered systems.[90]

© 2025 The Author(s). Advanced Engineering Materials published by Wiley-VCH GmbH

15272648, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adem.202402486 by National Research Council, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

15272648, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adem.202402486 by National Research Council, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Additional data augmentation methods, such as transformations (e.g., rotations, noise injection), further enhance model robustness by introducing variability into existing datasets. Combined with feature selection and dimensionality reduction techniques like PCA, these preprocessing steps ensure that ML models focus on the most significant structural characteristics, eliminating noise and irrelevant data. This not only improves the efficiency of model training but also enhances prediction accuracy.[90]

Preprocessing techniques such as normalization and scaling are equally critical to ensure that datasets are properly prepared for training, minimizing biases and improving model performance. Together, these data handling and preprocessing strategies ensure that ML models for disordered structures are robust, scalable, and capable of addressing the unique challenges posed by these complex systems.

4. Model Development and Optimization

The development of ML models for disordered structures involves selecting the right algorithms, tuning hyperparameters, and ensuring robust performance across diverse datasets. The choice of model depends largely on the complexity of the data and the target properties. For instance, CNNs are often used for image-based analyses, while support vector machines and decision trees are useful for regression and classification tasks.[91]

Hyperparameter tuning is crucial for optimizing model performance. Techniques like grid search and random search allow researchers to identify the most effective model configurations. Regularization methods such as L_1 and L_2 are employed to prevent overfitting, ensuring that the model generalizes well beyond the training data.[92,93]

Cross validation is a commonly used technique to assess model performance, particularly in data-limited scenarios. By splitting the data into multiple subsets, researchers can evaluate the model's robustness and minimize the risk of overfitting. This method ensures that the final model performs consistently across a variety of conditions.[94]

Ensemble Learning methods, such as random forests and gradient boosting, combine multiple models to improve predictive accuracy.[95] These approaches help balance model complexity and performance, making them ideal for complex disordered systems where capturing diverse structural features is essential.

Overall, careful model selection, tuning, and validation are critical to developing ML models that are both accurate and computationally efficient, enabling faster discovery of novel disordered materials with desired properties.[5,96,97]

5. Challenges and Limitations

5.1. Data Scarcity and Model Generalization

Data scarcity and inconsistent quality remain major limitations in applying ML to disordered structures. Unlike crystalline materials, which exhibit periodicity and uniformity, disordered systems are inherently complex and variable, posing significant challenges for both data collection and characterization. Traditional analytical methods, such as XRD and electron microscopy, are often ill suited to disordered structures due to their reliance on regular patterns and symmetry. Furthermore, advanced experimental setups—such as high-resolution imaging or multiscale testing—are required to capture the unique features of disordered materials. These methods are not only time intensive and expensive but also prone to variability, further complicating the generation of reliable datasets.

The application-specific nature of disordered systems further compounds the issue. For example, material properties such as mechanical strength or thermal conductivity can vary widely depending on environmental conditions like temperature, pressure, or humidity. As a result, data collection often requires large-scale experiments under diverse scenarios, which are not always feasible. Together, these factors contribute to the limited availability of high-quality data, hindering the development of robust ML models for disordered materials.[98]

To address these challenges, advanced ML techniques such as transfer learning and synthetic data augmentation have emerged as practical solutions. Transfer learning allows models trained on large datasets from related domains—such as ordered crystalline systems—to be fine tuned for disordered materials, reducing the reliance on extensive labeled datasets. Synthetic data generation methods, particularly GANs, have proven especially effective in expanding the training pool for ML models. By training a generator to create realistic data samples and a discriminator to evaluate their authenticity, GANs produce synthetic datasets that closely resemble real-world examples.

For disordered structures, GANs can replicate intricate features such as pore size distributions, atomic arrangements, and random defect patterns. A notable case study involved using GANs to generate synthetic 3D microstructures of porous materials. By augmenting the training dataset with these synthetic structures, researchers achieved significant improvements in predicting mechanical properties such as permeability and elastic modulus. This approach reduced the dependency on extensive experimental data and enabled the exploration of material configurations that had not yet been physically fabricated.

Despite their potential, challenges remain in ensuring that synthetic data generalizes well to real-world systems. Validating the representativeness of synthetic data and incorporating domain-specific knowledge into the generation process are essential steps toward improving the reliability of augmented datasets. Additionally, integrating synthetic data with traditional computational methods, such as molecular dynamics or FEA, can provide physics-based validation and further enhance model performance.[99]

The use of GANs and other synthetic data augmentation techniques represents a transformative step in overcoming data scarcity for disordered materials. By enabling researchers to explore uncharted material spaces and expand the capabilities of ML models, these methods are paving the way for new discoveries and innovations in the study of disordered structures.[100]

5.2. Limitations of ML Models in Disordered Structures

While ML has shown remarkable promise in analyzing and designing disordered structures, several limitations remain,

 © 2025 The Author(s). Advanced Engineering Materials published by Wiley-VCH GmbH

particularly due to the inherent complexity and variability of these systems.

5.2.1. Data Scarcity and Quality

Disordered structures often lack extensive, high-quality datasets. Experimental characterization of such materials is expensive, time consuming, and frequently limited to specific configurations, making it difficult to create comprehensive training datasets. The variability in disordered structures further exacerbates this issue, as capturing the full range of possible configurations requires extensive sampling across multiple scales.[101]

5.2.2. Feature Representation Challenges

The irregularity of disordered structures makes feature extraction and representation nontrivial. Unlike ordered systems, where periodicity and symmetry can simplify descriptors, disordered systems require complex representations, such as graphs, voxels, or high-dimensional embeddings. These representations are computationally intensive and may not always capture the critical properties governing material behavior.[102]

5.2.3. Overfitting in Complex Models

ML models trained on small or biased datasets risk overfitting, leading to poor generalization when applied to unseen configurations. This limitation is particularly critical for disordered structures, where the high-dimensional parameter space and randomness make it difficult for models to generalize effectively across all configurations.[103]

5.2.4. Lack of Interpretability

Many ML models, particularly deep learning approaches, act as "black boxes," providing predictions without clear insights into the underlying decision-making process. This lack of interpretability hinders the ability to validate results and understand the key factors influencing disordered structure behavior, limiting their applicability in scientific discovery.

5.2.5. Integration with Physics-Based Models

While hybrid ML-physics models are effective, integrating ML predictions with physics-based simulations (e.g., MD or FEA) can be computationally expensive and challenging to implement. Inaccuracies in either component may propagate, reducing the reliability of the combined approach.[104]

5.2.6. Uncertainty Quantification

Disordered structures often exhibit a wide range of possible behaviors due to their stochastic nature. ML models typically struggle to quantify prediction uncertainty, making it difficult to evaluate the reliability of results. Recent advancements in Bayesian neural networks and Monte Carlo dropout partially address this issue, but these methods add computational overhead and require careful implementation.[105]

5.2.7. Scalability Across Length Scales

While ML excels at specific length scales (e.g., atomic or macroscopic), bridging these scales remains a challenge. Disordered structures often exhibit hierarchical properties that span multiple scales, requiring seamless integration of ML models tailored to different levels of analysis.[106]

5.3. Enhancing Interpretability of Key Structural Features

One of the challenges in applying ML to disordered structures is the limited interpretability of the key structural features that govern material properties. Unlike ordered systems, where periodicity simplifies feature identification, disordered systems require advanced approaches to understand the complex relationships between structure and functionality.

5.3.1. Techniques for Enhancing Interpretability

Feature Importance Analysis: Techniques such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) provide insights into the contributions of individual features to ML predictions. For disordered materials, these methods can identify critical parameters, such as pore size distribution, atomic coordination numbers, or local strain fields, that influence mechanical or thermal properties.[107]

Graph-Based Representations: For atomic or microstructural data, GNNs provide interpretable outputs by associating predictions with specific nodes or edges. For example, GNNs can highlight atomic clusters or bonding patterns that contribute to enhanced strength or stability in metallic glasses.[108]

Visualization Techniques: Visualizing the learned features from CNNs or autoencoders helps uncover patterns in microstructural or morphological data. For disordered systems, these visualizations can reveal how randomness in geometry or connectivity affects material performance.[109]

Domain Knowledge Integration: Embedding domain-specific knowledge into ML models increases interpretability by aligning the learning process with established physical principles. For example, using physics-informed neural networks (PINNs) ensures that predictions adhere to fundamental conservation laws, making the results more reliable and interpretable.[109]

Sensitivity Analysis: Conducting sensitivity analysis on input features quantifies the impact of structural variations on material properties. This technique is particularly useful for hierarchical disordered systems, where multiple scales interact in complex ways.[109]

5.3.2. Case Studies and Applications

Recent studies have demonstrated the utility of these techniques in improving interpretability for disordered materials. For example, SHAP analysis applied to ML models predicting thermal conductivity in amorphous silicon identified local structural motifs that impede phonon transport. Similarly, GNNs have

© 2025 The Author(s). Advanced Engineering Materials published by Wiley-VCH GmbH

15272648, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adem.202402486 by National Research Council, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

been used to interpret the effects of atomic coordination variability on phase stability in high-entropy alloys. By employing these strategies, researchers can enhance the interpretability of ML models, providing not only accurate predictions but also meaningful insights into the structural features driving material behavior. This increased understanding bridges the gap between data-driven approaches and traditional scientific inquiry, fostering innovation in the study of disordered systems.

5.4. Overfitting and Model Generalization

Overfitting is another challenge, particularly in cases where ML models are trained on small, highly specific datasets. Overfitting occurs when the model becomes too closely aligned with the training data, losing its ability to generalize to new, unseen data. Regularization techniques, such as cross validation, dropout, and early stopping, help mitigate this by ensuring the model remains robust across a wider range of input conditions.[110]

5.5. Computational Costs and Efficiency

As ML models become more complex, the computational costs associated with training and deploying them increase significantly. Techniques like ensemble learning and hyperparameter tuning, while improving model performance, often require extensive computational resources. This is particularly challenging when working with high-dimensional data from disordered systems. Balancing the need for model accuracy with computational efficiency is crucial, especially for real-time applications like material design and drug discovery.[111]

5.6. Interpretability and Transparency

ML models, particularly deep learning models, are often seen as black boxes, making it difficult to interpret how they arrive at specific predictions. In fields where model outcomes directly influence material design or critical applications, transparency is essential. Explainable AI (XAI) methods like SHAP and LIME are increasingly being adopted to improve interpretability, allowing researchers to trace predictions back to specific features. However, striking a balance between model complexity and interpretability remains a key challenge for broader ML adoption in material science.

6. Manufacturing of ML-Optimized Disordered Structures

The successful implementation of ML-optimized disordered structures relies on efficient and scalable manufacturing techniques, with AM, or 3D printing, emerging as a highly promising approach.[112,113] AM offers several key advantages, including customization through precise control of geometric features and material composition, flexibility in producing a diverse range of disordered structures with intricate geometries and heterogeneous properties, and efficiency in reducing lead times and material waste compared to traditional methods. Additionally, AM enables very quick iteration between dissimilar candidate designs and allows for the simultaneous fabrication of many candidate designs, greatly accelerating the testing and validation process.

Materials selection is a critical first step in this process, as the choice of materials directly impacts the performance and manufacturability of disordered structures. Identifying materials with the desired properties and compatibility with AM processes is essential but can be challenging due to the unique requirements of disordered systems, such as their mechanical robustness and thermal stability.[114]

However, manufacturing these complex materials presents significant challenges. Ensuring repeatability of material properties between identical designs is a major hurdle, as the random nature of disordered structures can amplify the influence of internal manufacturing defects. These defects, such as voids or incomplete fusions, can significantly affect the mechanical and functional properties of the fabricated structures, complicating the comparison of theoretical (simulated) material properties to their real-world counterparts. Consistent quality and reproducibility remain difficult to achieve, especially for large-scale production, and additional postprocessing steps, such as heat treatment or surface finishing, are often required to meet property specifications.

Addressing these challenges involves several strategies. Developing tailored materials specifically for AM can improve compatibility and performance. Process optimization techniques, including advanced monitoring and control systems, can enhance accuracy and repeatability during fabrication. Simulation tools play a critical role in predicting manufacturing outcomes, optimizing design parameters, and reducing trial-and-error efforts. Implementing robust quality control measures, such as in situ monitoring and nondestructive evaluation methods, ensures that fabricated disordered structures meet the required specifications.

By overcoming these obstacles and harnessing advancements in AM technology, the full potential of ML-optimized disordered structures can be realized across various applications, unlocking new opportunities in material science and engineering (see **Figure 4**a).

7. Future Directions and Opportunities

7.1. Integration of ML with Traditional Computational Methods

The integration of ML with traditional computational methods, such as MD and DFT, has transformed the study of disordered materials at the atomic and microstructural scales. These methods provide detailed insights into atomic interactions, bonding, and electronic structures, which are critical for understanding the behavior of disordered systems at the molecular level.[115] For instance, combining ML with DFT has significantly reduced the computational cost and time required for discovering new high-performance alloys, cutting discovery time by up to 40%. This integration enables researchers to validate ML-predicted candidates by evaluating their thermodynamic stability and electronic properties.[60]

At the atomic scale, ML models such as GNNs enhance MD and DFT simulations by predicting phase stability, bond

© 2025 The Author(s). Advanced Engineering Materials published by Wiley-VCH GmbH

15272648, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adem.202402486 by National Research Council, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

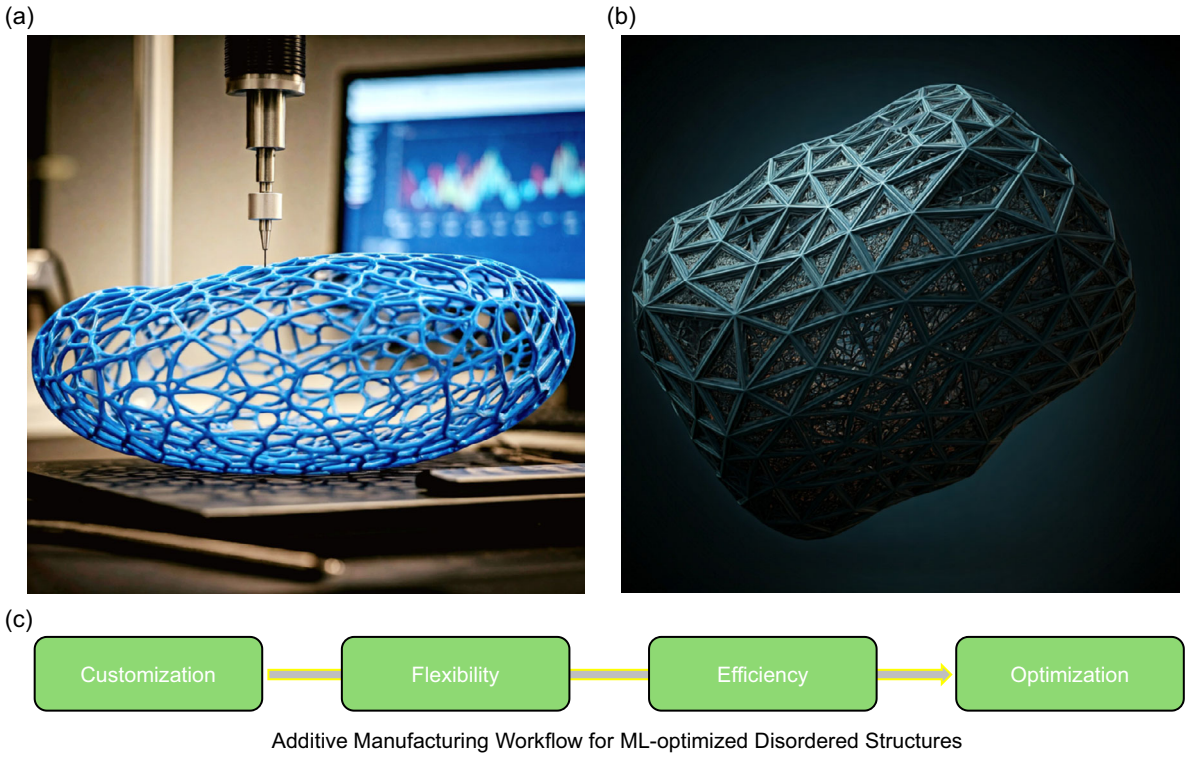

Figure 4. a) The intersection of technology and materials science: a ML-optimized disordered structure, fabricated using AM for enhanced performance and functionality. b) The future of materials: A complex lattice, born from advanced manufacturing and ML. c) This flowchart shows the main stages of the AM process enhanced by ML: Customization: precise control over structure and material properties. Flexibility: enables complex geometries and heterogeneous properties. Efficiency: reduces lead times and material waste. Optimization: ML refines parameters, improving performance and meeting targeted specifications. This ML-driven AM approach enables advanced, tailored disordered structures for applications in energy storage, aerospace, and biomedical fields.

strengths, and defect distributions. For example, ML models trained on MD-generated datasets have been used to predict glass transition temperatures in amorphous polymers, providing rapid evaluations of thermal stability without exhaustive simulations.[44] In catalysis, ML combined with DFT descriptors has facilitated the discovery of ternary platinum alloys for catalytic applications, optimizing activity and stability for oxygen reduction reactions in fuel cells.[116]

At the microstructural scale, ML models enhance the analysis of grain boundaries, pore distributions, and void configurations. ML-MD coupling has been applied to predict the permeability and porosity of porous media, accurately replicating fluid flow and diffusion properties in disordered structures. These advancements allow researchers to explore and optimize microstructures for improved mechanical and thermal performance.[117,118]

The primary strength of this integration lies in its ability to provide atomic-level precision while leveraging ML to accelerate traditionally computationally intensive simulations. By combining ML's data-driven predictions with physics-based insights from MD and DFT, researchers can gain a more comprehensive understanding of disordered systems at the molecular and microstructural levels.

Moving forward, the integration of ML with atomistic simulations will likely unlock further advancements, such as real-time monitoring of solvent interactions with disordered surfaces, essential for energy storage and catalytic processes.

7.2. ML-Driven Multiscale Design of Disordered Structures

The design of disordered structures spans multiple length scales, from atomic to macroscopic, each presenting unique technical requirements and challenges. ML facilitates this multiscale design by tailoring disorder to optimize material performance for specific applications across atomic, microstructural, and macroscopic scales.

At the atomic scale, disorder arises from the randomization of atomic arrangements, as seen in high-entropy alloys and amorphous solids. ML models, such as GNNs, are particularly effective at this scale due to their ability to model non-Euclidean relationships between atoms. These models enable predictions of phase stability, electronic properties, and mechanical performance by analyzing atomic configurations, accelerating the discovery of materials with enhanced strength and thermal stability.[119]

At the microstructural scale, disorder is introduced through variations in grain sizes, orientations, and void distributions, as observed in porous media and composite materials. GANs and CNNs play a significant role at this scale, simulating and

© 2025 The Author(s). Advanced Engineering Materials published by Wiley-VCH GmbH

15272648, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adem.202402486 by National Research Council, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

optimizing microstructures for properties such as fracture resistance, energy absorption, and manufacturability. For example, GANs have been used to generate synthetic microstructures that replicate experimental observations, enabling the exploration of novel configurations without extensive physical testing.[120]

At the macroscopic scale, disorder is incorporated into architected structures, such as foams, lattices, and stochastic metamaterials. Here, ML aids in optimizing geometries to achieve specific performance metrics, such as load-bearing capacity, acoustic damping, or thermal conductivity. Techniques like reinforcement learning are particularly effective for navigating the vast design space, while FEA is often integrated with ML to validate and refine macroscale designs. For instance, ML models combined with FEA have been used to design hierarchical lattice structures that maximize load capacity while minimizing weight.[121]

The multiscale design framework leverages the synergy between ML and traditional computational methods, addressing the distinct challenges at each scale. Atomic-scale design focuses on precision and physics-based simulations, while macroscale design emphasizes structural optimization and manufacturability. By seamlessly integrating ML across these scales, researchers can develop hierarchical disordered materials with tailored properties for diverse applications, including energy storage, aerospace, and biomedical devices.

7.3. Advancements in ML Algorithms and Their Potential Impact

As ML algorithms evolve, their applications in material science, particularly for disordered structures, will grow exponentially. The development of more efficient deep learning architectures, such as GNNs and PINNs, offers new ways to model disordered materials with unprecedented accuracy. These networks can incorporate physical constraints and symmetries directly into their learning processes, making them ideal for modeling materials where traditional assumptions break down.[122]

For disordered systems, GNNs are particularly promising because they can capture relationships in non-Euclidean spaces, such as the irregular atomic arrangements found in metallic glasses and high-entropy alloys. By representing these materials as graphs, where nodes represent atoms and edges represent interatomic bonds or forces, GNNs can effectively model the complex interactions between atoms in disordered systems. This approach allows for the identification of local structural motifs, such as clusters or defects, that significantly influence material properties.

Recent studies have demonstrated the power of GNNs in predicting the mechanical and thermal properties of disordered systems. For instance, GNNs have been applied to metallic glasses to predict yield strength and elastic moduli based on atomic configurations, outperforming traditional empirical models that rely on simplified structural descriptors.[123] Similarly, in high-entropy alloys, GNNs have been used to predict phase stability and deformation mechanisms, enabling the design of alloys with enhanced toughness and thermal stability.[87]

Beyond predicting material properties, GNNs also excel in uncovering the underlying principles governing the behavior of disordered materials. For example, researchers have used GNNs to identify key atomic arrangements that contribute to the unique thermal conductivity of amorphous silicon, providing insights into heat transport mechanisms in noncrystalline materials.[124,125] Moreover, GNNs can be integrated with generative models to propose novel atomic configurations for disordered materials with targeted properties, such as lightweight, high-strength alloys or thermally conductive glasses.

The adaptability of GNNs also extends to multiscale modeling, where they can bridge atomistic and mesoscopic scales. By combining GNNs with MD simulations, researchers can develop hybrid models capable of predicting the behavior of disordered systems under real-world conditions, such as high temperatures or mechanical loads. As GNN architectures continue to evolve, their ability to handle increasingly complex datasets and provide interpretable predictions will further enhance their role in the study and design of disordered materials.

PINNs offer another avenue of innovation by embedding physical laws (e.g., conservation of mass, energy, and momentum) directly into the learning process. This allows for the creation of models that are not only data-driven but also physically consistent, enabling more reliable predictions for systems where data is sparse or noisy—an issue that is common in the study of disordered structures.

7.4. Machine Learning-Driven Multiscale Modeling of Disordered Metamaterials

To address the connection between ML, metamaterials, and multiscale modeling, we recognize that disordered metamaterials often exhibit structural complexity that spans multiple scales. These systems demand hybrid approaches combining finite element methods (FEM) and MD simulations to effectively model their continuum and atomic-level behaviors. Metamaterials, particularly those characterized by stochastic or hierarchical configurations, serve as a prime example of how ML can bridge these scales by enhancing the predictive accuracy of multiscale frameworks.

ML algorithms, such as neural networks and generative models, can significantly improve the integration of FEM and MD. By predicting key material behaviors at different scales, ML can reduce the computational costs associated with multiscale simulations. For instance, it can interpolate results between MD simulations and FEM, enabling real-time insights into the behavior of disordered metamaterials under varying conditions. Additionally, hierarchical metamaterials can benefit from ML's ability to identify hidden patterns, optimize structural configurations, and simulate emergent behaviors, which are often challenging to capture with conventional methods alone.

The inclusion of ML in multiscale frameworks not only improves efficiency but also facilitates the design of metamaterials with tailored properties for applications ranging from energy absorption to biomedical devices. Future work should emphasize the development of data-driven multiscale models where ML augments traditional computational techniques, enabling a deeper understanding of disordered structures and their hierarchical properties.

© 2025 The Author(s). Advanced Engineering Materials published by Wiley-VCH GmbH

15272648, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adem.202402486 by National Research Council, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

7.5. Enhancing Interpretability and Transparency in ML Models

As ML models become increasingly complex, ensuring their interpretability and transparency remains a critical challenge. In the context of disordered materials, understanding how an ML model arrives at a particular prediction is crucial for gaining trust in its outcomes, especially when the results influence the design of materials for critical applications like energy storage, aerospace, and biomedical devices.

XAI techniques are becoming a focus of research to address this challenge. Methods such as SHAP and LIME are already being applied to interpret ML models in material science, providing insight into which features are most influential in predicting material properties. As these techniques improve, they will enable researchers to better understand the relationships between structural disorder and material performance, fostering more confident and reliable design processes.

Moreover, as transparency becomes a requirement in high-stakes industries, ML models for material science must be designed with auditability in mind. This involves creating robust pipelines for tracking model decisions, ensuring the reproducibility of results, and addressing potential biases in training datasets. Future research must focus on developing ML frameworks that balance accuracy with transparency, particularly when designing disordered materials where experimental validation is challenging.

7.6. Overcoming Data Scarcity and Quality Challenges

Data scarcity and quality issues remain significant obstacles in applying ML to disordered structures. While ML models excel when provided with large datasets, gathering comprehensive and high-quality data for disordered systems is often challenging and expensive. This difficulty is exacerbated in extreme environments, such as high-temperature or high-pressure conditions, where experimental data is inherently limited. Additionally, the unique and nonrepetitive designs of disordered materials require that they be evaluated on a case-by-case basis, complicating efforts to generalize across families of geometries and applications.

To address these challenges, advanced techniques such as transfer learning, synthetic data generation, and collaborative data-sharing platforms are emerging as promising solutions. Transfer learning enables ML models trained on large, related datasets—such as those of ordered crystalline systems—to be fine-tuned for disordered materials, reducing the reliance on extensive labeled datasets. This approach has already shown success in fields like medical imaging and is beginning to demonstrate potential in material science for adapting models to amorphous or disordered systems.

GANs and other synthetic data augmentation methods have also emerged as effective tools for tackling data scarcity. These techniques generate realistic datasets by mimicking the distribution of real data, thereby expanding the training pool and improving ML model generalization. For example, GANs have been applied to create synthetic microstructures for porous materials, enabling the exploration of material spaces that are otherwise inaccessible due to data limitations. As these synthetic data generation methods continue to evolve, they promise to facilitate robust ML training and open up new possibilities in material discovery.

Looking ahead, interdisciplinary collaboration and the development of robust data-sharing platforms will be critical for overcoming these challenges. Such platforms can aggregate experimental and simulated datasets from multiple research groups, fostering broader accessibility to high-quality data and accelerating innovation in disordered materials research. By addressing data scarcity through these advanced techniques and collaborative efforts, the field can unlock the full potential of ML for the design and optimization of disordered structures.

7.7. Automation and Data Generation

Data scarcity poses a significant challenge in studies requiring iterative, costly experiments, especially when extensive data generation is needed for comprehensive analysis. Although advancements in 3D printing have enhanced the fabrication of macrodisordered structures, the process remains resource intensive, demanding considerable time and human involvement. Furthermore, selecting an appropriate starting design or disordered structure for analysis is often unclear. Automated, robotic data-driven campaigns present a viable solution by enabling faster, more efficient data generation. Such approaches have shown considerable success in fields like chemistry and materials science.[126–128]

Integrating optimization algorithms, especially those powered by artificial intelligence, can greatly enhance iterative material production processes. Through robotic systems, samples can be physically fabricated and assessed, guiding decisions on the optimal next material and design for testing. In this context, active learning and Bayesian optimization are especially promising. Active learning selectively queries data points expected to maximize learning efficiency, while Bayesian optimization identifies the best candidates for experimentation, iterating toward a target design with minimal trials. Bayesian optimization, in particular, utilizes probabilistic models to predict outcomes, balancing exploration (testing new designs) with exploitation (refining known promising designs). This capability makes it a valuable tool for efficiently navigating the experimental space and approaching the global minimum with fewer resources.[129–131] By enabling automated, data-driven decision-making, these methods significantly enhance the speed and precision of iterative experimental loops, accelerating advancements of discovery new materials and designs.

7.8. Collaboration Across Disciplines

The future of ML in disordered materials research will require increased interdisciplinary collaboration. The complexity of disordered systems means that expertise in ML, material science, computational physics, and even biology must converge to develop effective solutions. Collaborative efforts between computational researchers and experimentalists are already yielding new insights, particularly in fields like protein folding, drug design, and multifunctional material discovery.

© 2025 The Author(s). Advanced Engineering Materials published by Wiley-VCH GmbH

15272648, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adem.202402486 by National Research Council, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Increased collaboration will also foster the development of shared data platforms, where experimental and computational results can be systematically archived and made accessible to the broader scientific community. These platforms will accelerate the pace of discovery by providing a wealth of data for training and validating ML models, facilitating cross-disciplinary innovations, and ensuring that discoveries in one field are rapidly translated into others.

7.9. Exploring New Domains of Application

Finally, the application of ML to disordered structures is just beginning to unlock its full potential. New frontiers include the study of disordered quantum systems, magnetism, and superconductivity—areas where the behavior of materials is highly sensitive to structural disorder. ML, with its capacity to model complex, high-dimensional relationships, is uniquely positioned to drive breakthroughs in these fields[122,132] (see Figure 4b).

In biological systems, ML could advance our understanding of IDPs, shedding light on their role in disease mechanisms and offering new avenues for therapeutic interventions. Similarly, ML applications in energy materials, particularly for battery technologies and fuel cells, could lead to the discovery of new disordered materials with improved performance, durability, and environmental resilience. Such disordered energy materials would benefit from optimizing for high surface and active chemical reaction sites.

As ML continues to evolve, its ability to tackle the complexity of disordered structures will undoubtedly expand, pushing the boundaries of what is possible in material design and discovery.

8. Conclusion

ML is revolutionizing the study and optimization of disordered structures, offering ground-breaking opportunities across disciplines such as materials science, biology, and engineering. By employing advanced techniques like deep learning, generative models, and graph neural networks, researchers can now tackle the inherent complexity of disordered systems with unprecedented accuracy. These advancements are driving innovations in energy storage, drug design, and advanced manufacturing, unlocking new possibilities that were previously unattainable.

A key strength of ML lies in its ability to process complex, nonperiodic data and uncover hidden patterns and correlations. This capability has led to significant breakthroughs in identifying material properties, accelerating material discovery, and enabling the design of novel structures. Examples include using GANs to generate synthetic datasets for porous materials and leveraging reinforcement learning for inverse design. Such advances highlight ML's unique capacity to address the variability and irregularity inherent in disordered systems.

However, the application of ML to disordered materials is not without its challenges. Data scarcity remains a critical barrier, as the generation of high-quality experimental datasets is often costly and time intensive. Additionally, ensuring the interpretability of complex ML models and overcoming computational demands are essential for broader adoption. Advances in generative models, explainable AI, and hybrid methods that integrate ML with traditional computational approaches—such as molecular dynamics and DFT—are helping address these issues. These integrations not only enhance predictive accuracy but also provide physics-based insights into the behavior of disordered systems.

Looking to the future, ML's role in disordered systems research will depend on fostering interdisciplinary collaboration and building robust data-sharing infrastructures. Collaborative platforms and open-source initiatives can help bridge gaps in data availability while ensuring equitable access to high-quality datasets. Advances in computational power and algorithmic efficiency will further expand ML's applicability, enabling real-time analysis and multiscale modeling of disordered materials.

The potential for ML to transform areas such as energy technologies, catalysis, and biomedical engineering is immense. By addressing current limitations and embracing interdisciplinary approaches, the field can unlock a new era of innovation. ML is poised to drive faster, more accurate discoveries, enabling transformative applications that impact both fundamental research and real-world industries. As this technology evolves, its contributions will undoubtedly reshape the study and design of disordered materials, paving the way for a future where the challenges of complexity are met with precision and innovation.

Acknowledgements

Financial support from the National Research Council Canada (NRC) (HYBA24) is gratefully acknowledged.

Open Access funding provided by the National Research Council Canada library.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Hamidreza Yazdani Sarvestani: data curation (lead); investigation (lead); project administration (lead); resources (lead); supervision (lead); visualization (lead); writing—original draft (lead). **Surabhi Nadigotti**: data curation (supporting); investigation (supporting); visualization (equal); writing—original draft (supporting). **Erfan Fatehi**: investigation (supporting); writing—original draft (supporting). **Derek Aranguren van Egmond**: formal analysis (supporting); investigation (supporting); writing—original draft (supporting). **Behnam Ashrafi**: conceptualization (equal); funding acquisition (equal); resources (equal); writing—original draft (equal).

Keywords

amorphous materials, disordered structures, machine learning, materials science, neural networks

Received: October 25, 2024
Revised: December 29, 2024
Published online:

[1] H. Liu, L. Li, Z. Wei, M. M. Smedskjaer, X. R. Zheng, M. Bauchy, *Adv. Sci.* **2024**, *11*, 2304834.

© 2025 The Author(s). Advanced Engineering Materials published by Wiley-VCH GmbH

15272648, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adem.202402486 by National Research Council, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

[2] X. Guo, M. Guzmán, D. Carpentier, D. Bartolo, C. Coulais, *Nature* **2023**, *618*, 506.
[3] C. Lu, M. Hsieh, Z. Huang, C. Zhang, Y. Lin, Q. Shen, F. Chen, L. Zhang, *Engineering* **2022**, *17*, 44.
[4] X. Yu, J. Zhou, H. Liang, Z. Jiang, L. Wu, *Prog. Mater. Sci.* **2018**, *94*, 114.
[5] S. Yu, X. Piao, N. Park, *Nat. Commun.* **2020**, *11*, 4842.
[6] M. Zaiser, S. Zapperi, *Nat. Rev. Phys.* **2023**, *5*, 679.
[7] S. Bonfanti, R. Guerra, M. Zaiser, S. Zapperi, *APL Mater.* **2021**, *9*, 020904.
[8] J. Song, J. Yan, B. Yi, *Front. Mech. Eng.* **2024**, *19*, 3.
[9] H. Jin, E. Zhang, B. Zhang, S. Krishnaswamy, G. E. Karniadakis, H. D. Espinosa (Preprint), arXiv:2311.13812, v1, Submitted: May **2023**.
[10] Z. Wang, Z. Du, L. Wang, G. He, I. P. Parkin, Y. Zhang, Y. Yue, *Nano Energy* **2024**, *121*, 109250.
[11] J. H. Lee, N. D. Loh, Z. Y. Yeo, Y. K. Ong, D. Balakrishnan, C. M. A. Limpo, A. Datta, C. Cetin, S. Ning, C. Wong, J. Shi, F. Hou, J. Lin, T. Minamikawa, T. Ito, H. Kamisuki, S. Pennycook, P. Matsudaira, B. Özyilmaz, *Adv. Mater.* **2024**, *36*, 2402628.
[12] M. E. Launey, R. O. Ritchie, *Adv. Mater.* **2009**, *21*, 2103.
[13] I. Giorgio, M. Spagnuolo, U. Andreaus, D. Scerrato, A. M. Bersani, *Math. Mech. Solids* **2021**, *26*, 1074.
[14] D. W. Green, K. K.-H. Lee, J. A. Watson, H.-Y. Kim, K.-S. Yoon, E.-J. Kim, J.-M. Lee, G. S. Watson, H.-S. Jung, *Sci. Rep.* **2017**, *7*, 41023.
[15] S. Krauss, E. Monsonego-Ornan, E. Zelzer, P. Fratzl, R. Shahar, *Adv. Mater.* **2009**, *21*, 407.
[16] C. A. La Porta, M. C. Lionetti, S. Bonfanti, S. Milan, C. Ferrario, D. Rayneau-Kirkhope, M. Beretta, M. Hanifpour, U. Fascio, M. Ascagni, L. De Paola, Z. Budrikis, M. Schiavoni, E. Falletta, A. Caselli, O. Chepizhko, A. Tuissi, A. Vailati, S. Zapperi, *Proc. Natl. Acad. Sci.* **2019**, *116*, 18777.
[17] L. J. Gibson, *J. R. Soc. Interfaces* **2012**, *9*, 2749.
[18] S. Hamza, N. Slimane, Z. Azari, G. Pluvinage, *Appl. Surf. Sci.* **2013**, *264*, 485.
[19] D. E. Koshland Jr, G. Némethy, D. Filmer, *Biochemistry* **1966**, *5*, 365.
[20] M. Arkin, *Curr. Opin. Chem. Biol.* **2005**, *9*, 317.
[21] Y. Cheng, T. Legall, C. J. Oldfield, J. P. Mueller, Y.-Y. J. Van, P. Romero, M. S. Cortese, V. N. Uversky, A. K. Dunker, *Trends Biotechnol.* **2006**, *24*, 435.
[22] S. Yu, C.-W. Qiu, Y. Chong, S. Torquato, N. Park, *Nat. Rev. Mater.* **2021**, *6*, 226.
[23] L. J. Gibson, *MRS Bull.* **2003**, *28*, 270.
[24] J. Banhart, *Prog. Mater. Sci.* **2001**, *46*, 559.
[25] S. Gaitanaros, S. Kyriakides, A. M. Kraynik, *Int. J. Solids Struct.* **2012**, *49*, 2733.
[26] W. S. Sanders, L. J. Gibson, *Mater. Sci. Eng., A* **2003**, *347*, 70.
[27] C. Chen, Y. Zuo, W. Ye, X. Li, S. P. Ong, *Nat. Comput. Sci.* **2021**, *1*, 46.
[28] C. Nunes, A. Mahendrasingam, R. Suryanarayanan, *Pharm. Res.* **2005**, *22*, 1942.
[29] A. Radford (Preprint), arXiv:1511.06434, v1, Submitted: Nov. **2015**.
[30] S. D. Griesemer, Y. Xia, C. Wolverton, *Nat. Comput. Sci.* **2023**, *3*, 934.
[31] M. Mortazavi, R. Ehsani, S. Carpin, A. Toudeshki, *Predicting Tree Water Status in Pistachio and Almond Orchards using Supervised Machine Learning*, Available at SSRN 4511076, **2023**.
[32] J. Spitaler, S. K. Estreicher, *Front. Mater.* **2018**, *5*, 70.
[33] J. Schmidt, M. R. G. Marques, S. Botti, M. A. L. Marques, *npj Comput. Mater.* **2019**, *5*, 83.
[34] F. Luborsky, in *Amorphous Metallic Alloys, Butterworths Monographs in Materials* (Ed: F. E. Luborsky), Butterworths, USA **1983**, pp. 1–7.
[35] A. L. Greer, *Science* **1995**, *267*, 1947.
[36] S. R. Elliott, *Physics of Amorphous Materials*, Longman Scientific & Technical, USA **1990**.
[37] W. Pontuschka, L. Barbosa, *Non-Cryst. Mater. Optoelectron.* **2004**, *13*, 363.
[38] J. Vincent, *Structural Biomaterials*, Princeton University Press, USA **2012**.
[39] A. Graves, A. Graves, *Supervised Sequence Labelling with Recurrent Neural Networks*, Springer, Germany **2012**, pp. 37–45.
[40] A. Graves (Preprint), arXiv:1308.0850, v1, Submitted: Aug. **2013**.
[41] D. P. Kingma, P. Dhariwal, *Advances in Neural Information Processing Systems*, Vol. *31*, MIT Press, USA **2018**.
[42] H. G. Weller, G. Tabor, H. Jasak, C. Fureby, *Comput. Phys.* **1998**, *12*, 620.
[43] B. Crouse, D. M. Freed, N. Koliha, G. Balasubramanian, R. Satti, D. Bale, S. Zuklic, *Paper SCA2016-058 presented at the Int. Symp. of the Society of Core Analysts held in Snow Mass*, Colorado, USA **2016**.
[44] K. M. Guan, T. I. Anderson, P. Creux, A. R. Kovscek, *Comput. Geosci.* **2021**, *156*, 104905.
[45] N. B. Kotadiya, A. Mondal, P. W. M. Blom, D. Andrienko, G.-J. A. H. Wetzelaer, *Nat. Mater.* **2019**, *18*, 1182.
[46] N. B. Kotadiya, H. Lu, A. Mondal, Y. Ie, D. Andrienko, P. W. M. Blom, G.-J. A. H. Wetzelaer, *Nat. Mater.* **2018**, *17*, 329.
[47] K.-H. Lin, A. Prlj, L. Yao, N. Drigo, H.-H. Cho, M. K. Nazeeruddin, K. Sivula, C. Corminboeuf, *Chem. Mater.* **2019**, *31*, 6605.
[48] X. Yin, Z. Song, Z. Li, W. Tang, *Energy Environ. Sci.* **2020**, *13*, 4057.
[49] T. Chatterjee, K. Wong, *Adv. Opt. Mater.* **2019**, *7*, 1800565.
[50] J. Jeong, M. Kim, J. Seo, H. Lu, P. Ahlawat, A. Mishra, Y. Yang, M. A. Hope, F. T. Eickemeyer, M. Kim, Y. J. Yoon, I. W. Choi, B. P. Darwich, S. J. Choi, Y. Jo, J. H. Lee, B. Walker, S. M. Zakeeruddin, L. Emsley, U. Rothlisberger, A. Hagfeldt, D. S. Kim, M. Grätzel, J. Y. Kim, *Nature* **2021**, *592*, 381.
[51] K.-H. Lin, A. Prlj, C. Corminboeuf, *J. Mater. Chem. C* **2018**, *6*, 960.
[52] T. Hastie, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer, Germany **2009**.
[53] R. J. Little, D. B. Rubin, *Statistical Analysis with Missing Data*, Vol. *793*, John Wiley & Sons, USA **2019**.
[54] C. Shorten, T. M. Khoshgoftaar, *J. Big Data* **2019**, *6*, 1.
[55] I. Guyon, A. Elisseeff, *J. Mach. Learn. Res.* **2003**, *3*, 1157.
[56] L. Breiman, *Mach. Learn.* **2001**, *45*, 5.
[57] R. Kohavi, *A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection*, Morgan Kaufman Publishing, USA **1995**.
[58] C. Kim, A. Chandrasekaran, T. D. Huan, D. Das, R. Ramprasad, *J. Phys. Chem. C* **2018**, *122*, 17575.
[59] Y. Zhang, X. Xu, *Polym. Chem.* **2021**, *12*, 843.
[60] G. Pilania, C. N. Iverson, T. Lookman, B. L. Marrone, *J. Chem. Inf. Model.* **2019**, *59*, 5013.
[61] M. A. F. Afzal, A. R. Browning, A. Goldberg, M. D. Halls, J. L. Gavartin, T. Morisato, T. F. Hughes, D. J. Giesen, J. E. Goose, *ACS Appl. Polym. Mater.* **2021**, *3*, 620.
[62] A. Odinokov, A. Freidzon, A. Bagaturyants, *Chem. Phys. Lett.* **2015**, *633*, 41.
[63] K.-H. Lin, L. Paterson, F. May, D. Andrienko, *npj Comput. Mater.* **2021**, *7*, 179.
[64] R. Xie, A. R. Weisen, Y. Lee, M. A. Aplan, A. M. Fenton, A. E. Masucci, F. Kempe, M. Sommer, C. W. Pester, R. H. Colby, E. D. Gomez, *Nat. Commun.* **2020**, *11*, 893.
[65] S. Yin, Z. Shuai, Y. Wang, *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 970.
[66] R. Barbosa-Da-Silva, R. Stefani, *Mol. Simul.* **2013**, *39*, 234.
[67] J. Xu, B. Chen, *J. Mol. Model.* **2005**, *12*, 24.
[68] P. Domingos, *Commun. ACM* **2012**, *55*, 78.
[69] J. Zhou, G. Cui, S. Hu, Z. Zhang, C. Yang, Z. Liu, L. Wang, C. Li, M. Sun, *AI Open* **2020**, *1*, 57.
[70] S. J. Pan, Q. Yang, *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345.
[71] Y. Gal, *Uncertainty in Deep Learning*, University of Cambridge, UK **2016**.

© 2025 The Author(s). Advanced Engineering Materials published by Wiley-VCH GmbH

15272648, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adem.202402486 by National Research Council, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

[72] B. Settles, *Active Learning Literature Survey*, University of Wisconsin, USA **2009**.
[73] W. Lee, S. Chong, J. Kim (Preprint), arXiv:2401.06152, v1, Submitted: Dec. **2023**.
[74] J. Zhang, M. Zhao, C. Zhong, J. Liu, K. Hu, X. Lin, *Nanoscale* **2023**, *15*, 18511.
[75] B. O. Mukhamedov, K. V. Karavaev, I. A. Abrikosov, *Phys. Rev. Mater.* **2021**, *5*, 104407.
[76] V. Sundararaghavan, N. Zabaras, *Comput. Mater. Sci.* **2005**, *32*, 223.
[77] M. Fernandez, P. G. Boyd, T. D. Daff, M. Z. Aghaji, T. K. Woo, *J. Phys. Chem. Lett.* **2014**, *5*, 3056.
[78] F. Corvacho, V. Muñoz, M. Sepúlveda-Macías, G. Gutiérrez, *Mater. Today Commun.* **2024**, *38*, 108298.
[79] G. Naidu, T. Zuva, E. M. Sibanda, in *Computer Science On-line Conf.*, Springer **2023**.
[80] Y. Liu, T. Zhao, W. Ju, S. Shi, *J. Materiomics* **2017**, *3*, 159.
[81] Y. Liu, X. Zou, S. Ma, M. Avdeev, S. Shi, *Acta Materialia* **2022**, *238*, 118195.
[82] X. Liu, J. Zhang, Z. Pei, *Prog. Mater. Sci.* **2023**, *131*, 101018.
[83] J. Noh, G. H. Gu, S. Kim, Y. Jung, *Chem. Sci.* **2020**, *11*, 4871.
[84] J. Wang, Y. Wang, Y. Chen, *Materials* **2022**, *15*, 1811.
[85] H. Ronellenfitsch, N. Stoop, J. Yu, A. Forrow, J. Dunkel, *Phys. Rev. Mater.* **2019**, *3*, 095201.
[86] Z. Wei, K. Wei, X. Yang, *Compos. Struct.* **2024**, *340*, 118210.
[87] D. Dold, D. Aranguren Van Egmond, *Cell Rep. Phys. Sci.* **2023**, *4*, 1.
[88] S. Bonfanti, S. Hiemer, R. Zulkarnain, R. Guerra, M. Zaiser, S. Zapperi, *Nat. Comput. Sci.* **2024**, *4*, 574.
[89] H. Pahlavani, K. Tsifoutis-Kazolis, M. C. Saldivar, P. Mody, J. Zhou, M. J. Mirzaali, A. A. Zadpoor, *Adv. Mater.* **2024**, *36*, 2303481.
[90] S. Candemir, X. V. Nguyen, L. R. Folio, L. M. Prevedello, *Radiol.: Artif. Intell.* **2021**, *3*, e210014.
[91] H.-T. Thai, *Structures*, Elsevier, Netherlands **2022**.
[92] L. Yang, A. Shami, *Neurocomputing* **2020**, *415*, 295.
[93] M. Feurer, F. Hutter, *Automated Machine Learning: Methods, Systems, Challenges*, Springer Nature, Germany **2019**, pp. 3–33.
[94] D. Berrar, *Cross-validation*, Elsevier, Netherlands **2019**.
[95] A. Botchkarev (Preprint), arXiv:1809.03006, v1, Submitted: Sep. **2018**.
[96] P. Schratz, J. Muenchow, E. Iturritxa, J. Richter, A. Brenning, *Ecol. Modell.* **2019**, *406*, 109.
[97] M. Bücker, G. Szepannek, A. Gosiewska, P. Biecek, *J. Oper. Res. Soc.* **2022**, *73*, 70.
[98] L. Alzubaidi, J. Bai, A. Al-Sabaawi, J. Santamaría, A. S. Albahri, B. S. N. Al-Dabbagh, M. A. Fadhel, M. Manoufali, J. Zhang, A. H. Al-Timemy, Y. Duan, A. Abdullah, L. Farhan, Y. Lu, A. Gupta, F. Albu, A. Abbosh, Y. Gu, *J. Big Data* **2023**, *10*, 46.
[99] Y. Mao, Q. He, X. Zhao, *Sci. Adv.* **2020**, *6*, eaaz4169.
[100] S. C. Shen, M. J. Buehler, *Commun. Eng.* **2022**. *1*, 37.
[101] A. Nandy, C. Duan, H. J. Kulik, *Curr. Opin. Chem. Eng.* **2022**, *36*, 100778.
[102] S. B. Torrisi, M. Z. Bazant, A. E. Cohen, M. G. Cho, J. S. Hummelshøj, L. Hung, G. Kamat, A. Khajeh, A. Kolluru, X. Lei, H. Ling, J. H. Montoya, T. Mueller, A. Palizhati, B. A. Paren, B. Phan, J. Pietryga, E. Sandraz, D. Schweigert, Y. Shao-Horn, A. Trewartha, R. Zhu, D. Zhuang, S. Sun, *APL Mach. Learn.* **2023**, *1*, 1.
[103] R. Chang, Y.-X. Wang, E. Ertekin, *npj Comput. Mater.* **2022**, *8*, 242.
[104] J. Willard, X. Jia, S. Xu, M. Steinbach, V. Kumar, *ACM Comput. Surv.* **2022**, *55*, 1.
[105] Y. Gal, P. Koumoutsakos, F. Lanusse, G. Louppe, C. Papadimitriou, *Nat. Rev. Phys.* **2022**, *4*, 573.
[106] O. Sedehi, A. M. Kosikova, C. Papadimitriou, L. S. Katafygiotis, *Mech. Syst. Signal Process.* **2024**, *208*, 111021.
[107] D. Mane, A. Magar, O. Khode, S. Koli, K. Bhat, P. Korade, *J. Electr. Syst.* **2024**, *20*, 1252.
[108] H. Xuanyuan, T. Zhao, D. Luo, *Proc. of the AAAI Conf. on Artificial Intelligence*, Vancouver, Canada **2023**.
[109] A. Khan, E. B. Mobaraki, in *2023 IEEE 10th Int. Conf. on Data Science and Advanced Analytics (DSAA)*, IEEE, LA, USA **2023**.
[110] K. Kouvaris, J. Clune, L. Kounios, M. Brede, R. A. Watson, (Preprint), arXiv:1508.06854, v1, Submitted: Aug. **2015**.
[111] O. Gómez-Carmona, D. Casado-Mansilla, F. A. Kraemer, D. López-De-Ipiña, J. García-Zubia, *Future Gener. Comput. Syst.* **2020**, *112*, 670.
[112] H. Yazdani Sarvestani, A. H. Akbarzadeh, H. Niknam, K. Hermenean, *Compos. Struct.* **2018**, *200*, 886.
[113] T. Lacelle, K. L. Sampson, H. Yazdani Sarvestani, A. Rahimizadeh, J. Barroeta Robles, M. Mirkhalaf, M. Rafiee, M. B. Jakubinek, C. Paquet, B. Ashrafi, *APL Mater.* **2023**, *11*, 1.
[114] H. Yazdani Sarvestani, T. Lacelle, A. Sohrabi-Kashani, A. Shashoua, V. Karamzadeh, H. Ravanbakhsh, A. Robitaille, H. Lavoie, C. Paquet, M. B. Jakubinek, B. Ashrafi, *Appl. Mater. Today* **2024**, *39*, 102327.
[115] L. Je, G. W. Huber, R. C. Van Lehn, V. M. Zavala, *Curr. Opin. Chem. Eng.* **2022**, *36*, 100796.
[116] M. Melchionna, P. Fornasiero, *ACS Catal.* **2020**, *10*, 5493.
[117] T. Xie, J. C. Grossman, *Phys. Rev. Lett.* **2018**, *120*, 145301.
[118] G. Huang, F. Huang, W. Dong, *Chem. Eng. J.* **2024**, *492*, 152294.
[119] H. Zhang, R. Huang, J. Chen, J. M. Rondinelli, W. Chen (Preprint), arXiv:2408.16337, v1, Submitted: Aug. **2024**.
[120] A. Madanchi, M. Kilgour, F. Zysk, T. D. Kühne, L. Simine, *J. Chem. Phys.* **2024**, *160*, 1.
[121] F. Aldakheel, C. Soyarslan, H. S. Palanisamy, E. S. Elsayed, *Mech. Mater.* **2023**, *184*, 104726.
[122] G. R. Schleder, A. C. Padilha, C. M. Acosta, M. Costa, A. Fazzio, *J. Phys.: Mater.* **2019**, *2*, 032001.
[123] E. Ross, D. Hambleton, in *Proc. of AAG2020-Advances in Architectural Geometry*, Vol. *24*, Paris, France **2021**, pp. 466–485.
[124] Z. Fang, Q. Yan, *npj Comput. Mater.* **2024**, *10*, 91.
[125] V. Fung, J. Zhang, E. Juarez, B. G. Sumpter, *npj Comput. Mater.* **2021**, *7*, 84.
[126] C. W. Coley, N. S. Eyke, K. F. Jensen, *Angew. Chem., Int. Ed.* **2020**, *59*, 22858.
[127] B. Burger, P. M. Maffettone, V. V. Gusev, C. M. Aitchison, Y. Bai, X. Wang, X. Li, B. M. Alston, B. Li, R. Clowes, N. Rankin, B. Harris, R. S. Sprick, A. I. Cooper, *Nature* **2020**, *583*, 237.
[128] P. Nikolaev, D. Hooper, F. Webber, R. Rao, K. Decker, M. Krein, J. Poleski, R. Barto, B. Maruyama, *npj Comput. Mater.* **2016**, *2*, 1.
[129] K. L. Snapp, B. Verdier, A. E. Gongora, S. Silverman, A. D. Adesiji, E. F. Morgan, T. J. Lawton, E. Whiting, K. A. Brown, *Nat. Commun.* **2024**, *15*, 4290.
[130] G. Tom, S. P. Schmid, S. G. Baird, Y. Cao, K. Darvish, H. Hao, S. Lo, S. Pablo-García, E. M. Rajaonson, M. Skreta, N. Yoshikawa, S. Corapi, G. D. Akkoc, F. Strieth-Kalthoff, M. Seifrid, A. Aspuru-Guzik, *Chem. Rev.* **2024**, *124*, 9633.
[131] E. Fatehi, M. Thadani, G. Birsan, R. W. Black, (Preprint), arXiv:2305.12541, v1, Submitted: May **2023**.
[132] V. L. Deringer, M. A. Caro, G. Csányi, *Adv. Mater.* **2019**, *31*, 1902765.
[133] M. F. Kinnaird, T. G. O'Brien, *The Ecology and Conservation of Asian Hornbills: Farmers of the Forest*, University of Chicago Press, USA **2007**.
[134] S. Weiner, W. Traub, *FASEB J.* **1992**, *6*, 879.
[135] T. Lyson, G. Bever, T. Scheyer, A. Hsiang, J. Gauthier, *Curr. Biol.* **2013**, *23*, 1113.
[136] S. Torre, T. Fjeld, H. R. Gislerød, R. Moe, *J. Am. Soc. Hortic. Sci.* **2003**, *128*, 598.

© 2025 The Author(s). Advanced Engineering Materials published by Wiley-VCH GmbH

15272648, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adem.202402486 by National Research Council, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License