IJMMM
International Journal of
Minerals, Metallurgy and Materials
矿物冶金与材料学报（英文版）

From microstructure to performance optimization: Innovative applications of computer vision in materials science

Chunyu Guo, Xiangyu Tang, Yu’e Chen, Changyou Gao, Qinglin Shan, Heyi Wei, Xusheng Liu, Chuncheng Lu, Meixia Fu, Enhui Wang, Xinhong Liu, Xinmei Hou, and Yanglong Hou

Cite this article as:
Chunyu Guo, Xiangyu Tang, Yu’e Chen, Changyou Gao, Qinglin Shan, Heyi Wei, Xusheng Liu, Chuncheng Lu, Meixia Fu, Enhui Wang, Xinhong Liu, Xinmei Hou, and Yanglong Hou, From microstructure to performance optimization: Innovative applications of computer vision in materials science, *Int. J. Miner. Metall. Mater.*, 33(2026), No. 1, pp. 94-115. https://doi.org/10.1007/s12613-025-3217-4

INTERNATIONAL JOURNAL OF
Minerals, Metallurgy and Materials

Volume 33, Number 1, January 2026
Page 94

Review https://doi.org/10.1007/s12613-025-3217-4

From microstructure to performance optimization: Innovative applications of computer vision in materials science

Chunyu Guo[1,2),*], *Xiangyu Tang*[2),*], *Yu'e Chen*[2,3)], *Changyou Gao*[3)], *Qinglin Shan*[4)], *Heyi Wei*[4)], *Xusheng Liu*[5)], *Chuncheng Lu*[2)], *Meixia Fu*[6)], *Enhui Wang*[2),✉], *Xinhong Liu*[5)], *Xinmei Hou*[2),✉], *and Yanglong Hou*[1)]

1) School of Materials, Sun Yat-sen University, Shenzhen 518107, China
2) Institute for Carbon Neutrality, University of Science and Technology Beijing, Beijing 100083, China
3) Wuhu Xinxing Cast Pipe Co., Ltd., Wuhu 214000, China
4) Tangshan Iron and Steel Group Co., Ltd., Tangshan 063000, China
5) Henan Key Laboratory of High Temperature Functional Ceramics, School of Material Science and Engineering, Zhengzhou University, Zhengzhou 450052, China
6) School of Automation, University of Science and Technology Beijing, Beijing 100083, China
(Received: 10 January 2025; revised: 16 June 2025; accepted: 24 June 2025)

Abstract: The rapid advancements in computer vision (CV) technology have transformed the traditional approaches to material microstructure analysis. This review outlines the history of CV and explores the applications of deep-learning (DL)-driven CV in four key areas of materials science: microstructure-based performance prediction, microstructure information generation, microstructure defect detection, and crystal structure-based property prediction. The CV has significantly reduced the cost of traditional experimental methods used in material performance prediction. Moreover, recent progress made in generating microstructure images and detecting microstructural defects using CV has led to increased efficiency and reliability in material performance assessments. The DL-driven CV models can accelerate the design of new materials with optimized performance by integrating predictions based on both crystal and microstructural data, thereby allowing for the discovery and innovation of next-generation materials. Finally, the review provides insights into the rapid interdisciplinary developments in the field of materials science and future prospects.

Keywords: microstructure; deep learning; computer vision; performance prediction; image generation

1. Introduction

The microstructure of a material directly influences its mechanical, thermal, and electrical properties [1–2]. Understanding the relationship between the microstructure of a material and its performance has long been a challenge in materials science [3–5]. Advancements made in imaging techniques, such as scanning electron microscopy (SEM) and transmission electron microscopy (TEM) over the past few decades have enabled the visualization of microstructures with high resolutions. However, effectively analyzing and interpreting the vast amount of image data generated using these technologies remains challenging. Traditional methods often rely on manual annotations and qualitative judgments, which are not only labor-intensive but also prone to human error and subjectivity.

Integrating computer vision (CV) with materials science helps overcome this challenge. It is a rapidly developing subfield of artificial intelligence (AI) that applies sophisticated algorithms to extract meaningful information from images automatically [6–7]. With the advent of deep learning (DL), particularly convolutional neural networks (CNNs), CV has been successful in tasks such as object detection, semantic segmentation, and feature classification [8–9]. These tasks have revolutionized microstructure analysis, enabling high-throughput, precise, and unbiased evaluation of complex material systems. Moreover, its ability to quickly process large datasets with minimal human intervention makes CV particularly valuable for accelerating material design and optimizing workflows.

Microstructure-based performance prediction, microstructure information generation, and microstructure defect detection are key tasks that have seen significant improvements with the application of DL-driven CV [10–12]. These tasks are particularly powerful in analyzing the complex relationships between the structure of a material and its performance and behavior under different conditions [13]. For example, performance optimization frameworks based on mi-

*These authors contributed equally to this work.
✉ Corresponding authors: Enhui Wang E-mail: wangenhui@ustb.edu.cn; Xinmei Hou E-mail: houxinmeiustb@ustb.edu.cn

crostructural analysis can predict material behavior under various loading conditions, facilitating the selection of materials with properties optimal for specific applications. By automating the analysis of large datasets, these frameworks offer a systematic approach that can identify performance bottlenecks and suggest microstructural modifications to enhance material performance [14]. Crystal structure-based prediction frameworks, such as crystal graph convolutional neural networks (CGCNNs) [15], have become crucial tools for understanding the fundamental relationship between the atomic structure of a material and its properties. These frameworks offer valuable insights into the stability, formation energy, and other key properties of a material that directly influence its performance. By learning from the large datasets of crystal structures, these frameworks enable the design of materials with tailored properties, thereby accelerating the discovery of high-performance materials [16]. Despite these advancements, the application of CV in materials science is still in its early stages, and many challenges are yet to be overcome. For example, optimizing model architectures for accurate and scalable predictions, overcoming data limitations, and ensuring interpretability are key hurdles that must be addressed to fully unlock the potential of CV in material performance optimization.

This review aims to explore the advancements in the application of CV for microstructure analysis and performance optimization of materials. It begins with a brief overview of the development of CV. Subsequently, the applications of DL-driven CV in four key areas (Fig. 1), namely microstructure-based performance prediction, microstructure information generation, microstructure defect detection, and crystal structure-based property prediction, are introduced. The review highlights the typical frameworks developed in the four areas, along with their contributions to materials science and their potential to shape the future of material discovery and design. Finally, the prospects of this interdisciplinary approach in transforming the study and design of advanced materials are discussed.

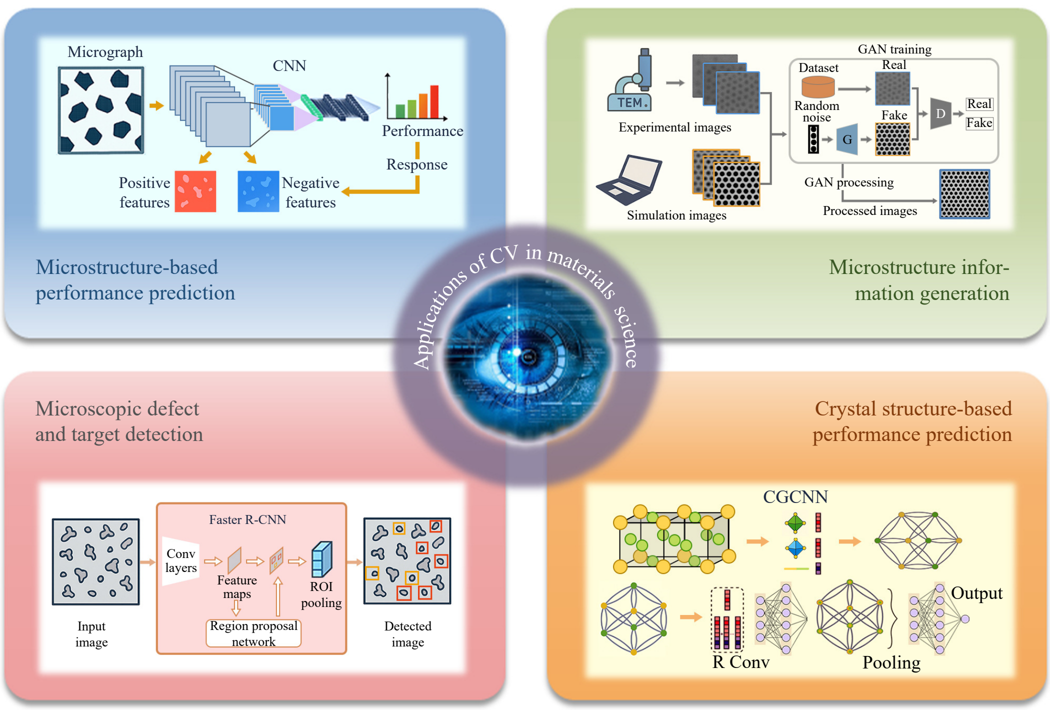

Fig. 1. Applications of DL-driven CV in microstructure-based performance prediction, microstructure information generation, microstructure defect detection, and crystal structure-based property prediction.

2. Overview of CV development

As a pivotal field within AI, CV enables machines to derive meaningful insights from visual data. Early progress in CV was constrained by the high computational and storage demands of image processing, but the field has since advanced rapidly with the development of AI algorithms and modern computing technologies.

Today, CV serves as a crucial link between visual data and actionable insights, playing a transformative role in fields such as materials science. By enabling high-throughput analysis of microstructure images, defect detection, and predictions of material performance and structural evolution, CV revolutionizes the design, analysis, and optimization of materials. The application of CV underscores its profound and growing impact on materials science by uncovering intricate relationships between microstructures and material properties that accelerate material discovery and process optimization. This section outlines the evolution of CV in four stages: the early stage (1950s—1960s), which introduced digital imaging foundations; the conceptual foundations stage (1970s—1980s), which established theoretical underpinnings; the handcrafted features stage (1990s—2000s), characterized by algorithmic innovations; the DL stage (2010s—

present), which redefined CV using neural networks.

2.1. Early stage (1950s—1960s)

In 1957, Kirsch, an engineer at the U.S. National Bureau of Standards, used a standard electronic automatic computer to scan the world's first digital image [17], marking the first instance of converting images into a digital format suitable for analysis and processing by computers. In 1963, Roberts [18] succeeded in extracting three-dimensional (3D) contour information from images, which is now considered pioneering work in the field of CV.

In 1966, Papert [19] introduced the concept of CV to the broad field of AI. Minsky and Papert [20] explored the foundations of pattern recognition and its applications in computational systems. These efforts laid the groundwork for the core subfields of CV, including object detection, image analysis, and 3D image reconstruction, establishing CV as a distinct domain.

2.2. Conceptual foundations stage (1970s—1980s)

The 1970s witnessed a shift in CV from abstract theoretical concepts to solving specific practical problems. During this stage, the core challenges faced by CV, including low-level vision, object recognition, character recognition, 3D modeling, and motion estimation, were formally identified [21]. In 1977, Marr [22] proposed a series of fundamental principles for CV, including hierarchical processing, primal sketching, and two-and-a-half dimensional sketching. These concepts laid the foundation for many core ideas that subsequently influenced DL development.

A pivotal development in CNNs occurred in 1980, when Fukushima [23] proposed Neocognitron, a model that introduced convolutional operations for feature extraction using local connectivity and weight sharing. This innovation not only reduced computational complexity but also preserved spatial information, influencing CNN architecture. In 1989, LeCun refined these concepts by introducing LeNet-5 [24], a 5-layer CNN designed for handwritten digit recognition. LeNet-5 demonstrated its real-world applicability by processing handwritten checks, marking a milestone in the integration of theory and practice.

2.3. Handcrafted features stage (1990s—2000s)

During the 1990s and 2000s, the challenges faced in training deep neural networks, including vanishing gradients and inefficient backpropagation [25], drove researchers toward handcrafted feature engineering. The handcrafted features stage emphasized the design of algorithms for manually extracting meaningful image information.

Notable advances made in CV during this stage included the scale-invariant feature transform [26] and speeded-up robust features [27], which addressed robust object recognition and image matching even in the presence of scale, rotation, and lighting variations. Algorithms such as the Canny edge detector [28] and histograms of oriented gradients [29] became essential tools for edge detection and object classification. In the late 1990s, active appearance models [30] emerged as a key tool for facial analysis, combining shape and texture information to accurately model and track facial features. Optical flow techniques [31] have also gained prominence in estimating motion in video sequences, benefitting applications such as gesture recognition, video compression, and robot navigation.

The cascade structure and Haar-like features of the Viola–Jones framework [32] introduced in 2001 revolutionized face detection [33]. These handcrafted approaches excelled in structured, well-defined tasks but required extensive domain expertise while struggling with high-dimensional datasets. This limitation underscored the need for automated and scalable methods and established the stage for DL.

2.4. Deep learning stage (2010s—present)

By the 2010s, handcrafted feature engineering in CV had reached its limits, particularly when addressing complex and large-scale challenges. The rise of DL led to a paradigm shift in the field of CV. A defining moment in this phase occurred in 2012, when the deep CNN AlexNet (Fig. 2) [34] achieved a groundbreaking victory in the ImageNet image classification competition. For the first time, DL surpassed traditional handcrafted methods, with its top-5 classification accuracy improvements exceeding 10%. This milestone highlighted the remarkable advantages of hierarchical feature learning enabled by DL, driving rapid innovation and widespread adoption of DL in CV [35–36].

Over the next decade, DL continued to expand the capabilities of CV across diverse applications. Beginning in 2014, models such as region-based CNNs (R-CNNs) [37] for object detection, fully convolutional networks (FCNs) [38] for semantic segmentation, and holistically nested edge detection (HED) [39] for edge detection set new standards in their

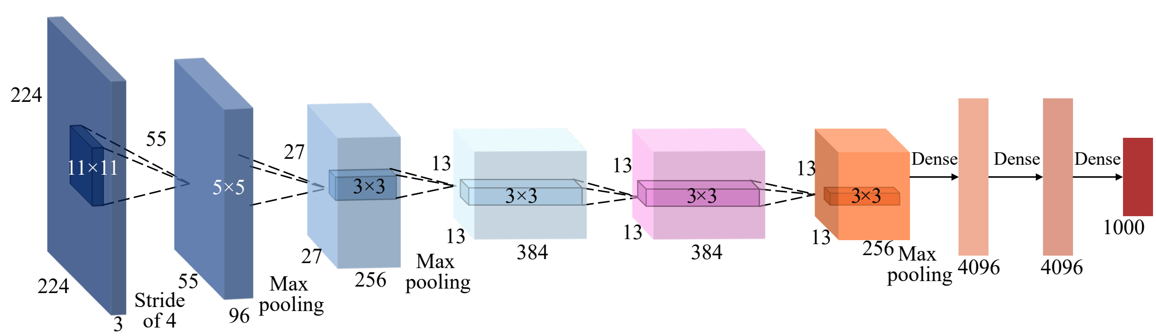

Fig. 2. AlexNet architecture [34].

respective areas. U-Net [40], initially developed for biomedical image segmentation, has also been widely adopted because of its symmetric architecture, which excels in learning spatial hierarchies while retaining fine-grained details in segmentation tasks. Residual networks (ResNets) [41] address the vanishing gradient problem by incorporating residual connections, thereby enabling the training of ultradeep neural networks. The introduction of generative adversarial networks (GANs) [42] has opened up new avenues for image synthesis, super-resolution, and style transfer, significantly broadening the scope of research and applications within CV. In 2017, unified frameworks, such as Mask R-CNN [43], enabled simultaneous object detection and instance segmentation within a single architecture, thereby enhancing the efficiency and comprehensiveness of CV solutions. Vision transformers (ViTs) [44], introduced in 2020, utilize self-attention mechanisms to effectively model global contextual information, surpassing CNNs in tasks such as image classification and object detection on large-scale datasets. The ViTs have significantly influenced the evolution of CV, inspiring the development of hybrid architectures that integrate transformers with convolutional layers to achieve enhanced accuracy and flexibility.

DL-based CV technologies are increasingly applied for analyzing material microstructures, predicting material performance, generating microstructure images, and detecting defects. Researchers have also developed specialized DL models for materials science, such as the crystal graph CNN (CGCNN) [15], which can predict the properties of materials directly from their atomic configurations. Overall, CV technologies have enabled high-throughput analysis, accelerating the discovery and optimization of materials. These advancements are crucial for tasks such as alloy design, composite material analysis, and manufacturing process optimization. They have significantly improved the precision and scalability of material characterizations, which will be discussed in detail throughout the following sections.

3. Application of CV in microstructure analysis

Building upon the historical development of CV in materials science, the following sections systematically examine four critical application domains where DL-driven CV has made transformative contributions. Section 3.1 focuses on microstructure-based performance prediction of materials and the establishment of structure–property relationships through advanced image analysis. Section 3.2 explores microstructure information generation through generative models that address data limitations and enable predictive simulations. Section 3.3 focuses on microstructure defect detection, demonstrating the automated identification and characterization of material imperfections. Finally, section 3.4 extends the discussion to atomic/molecular-scale crystal structure analysis, revealing the fundamental mechanisms that determine material properties. Together, these CV technologies connect microstructural and atomic/molecular-level analyses, delivering unprecedented capabilities in materials property prediction and development efficiency.

3.1. Microstructure-based performance prediction of materials

The quest to decode the composition, microstructure, and property relationships of materials lies at the heart of materials science. Quantifying and predicting these relationships are essential for refining material preparation processes, adjusting material compositions, and tailoring material microstructures to achieve desired properties. Recent advances in DL-driven CV have revolutionized this pursuit through two complementary paradigms: direct performance prediction from microstructure images and segmentation-based quantification of critical features. Highly flexible DL-driven CV technologies have enabled the extraction of critical information about materials from their microstructure images and have established strong correlations between microstructures and material properties [45–46]. These approaches collectively address longstanding challenges associated with manual feature engineering and its limited scalability.

Predicting the performance of a material directly from its microstructure images has become a widely adopted CV workflow. The DL models, such as CNNs and advanced hybrid architectures, support these workflows by processing high-dimensional image data to automatically extract relevant features. A detailed study of these models enables researchers to pinpoint critical microstructural features and identify the key factors that influence material performance. These insights deepen our understanding of structure–performance relationships in materials and help determine strategies for performance optimization. For example, Kondo *et al.* [47] developed a CNN-based framework to establish structure–property relationships in ceramics using SEM images, specifically linking microstructures to ionic conductivities (Fig. 3(a)). Utilizing data augmentation and a customized CNN architecture, they demonstrated that accurate predictions could be made with as few as seven micrographs, significantly reducing the data requirements compared to conventional methods. The CNN framework exhibited enhanced predictive accuracy for ionic conductivity directly from raw images, attaining a coefficient of determination (R^2 = 0.64), which surpassed that of kernel ridge regression (KRR, R^2 = 0.55) under identical experimental conditions. Moreover, the framework eliminated the need for manual feature engineering, thereby avoiding researcher bias and enabling the automatic extraction of physically interpretable features such as void distribution and void-free areas through intermediate feature visualization. An optimal representative volume element size essential for accurate predictions was also identified. Pei *et al.* [48] developed a variational autoencoder (VAE)-based framework for predicting alloy compositions. By analyzing the microstructure of Cr ferritic-martensitic steels, the framework effectively distinguished the microstructures of materials with highly similar morphologies. Leveraging principal component analysis, it identified critical elemental contributions to martensitic microstructure features and introduced an innovative inverse alloy design method capable of handling complex multicomponent systems.

The CV workflows for directly predicting material performance from microstructure images have drastically accelerated the analysis of complex material systems, delivering rapid predictions and uncovering structure–performance relationships in materials. They have streamlined the processes by efficiently processing high-dimensional data and automatically extracting critical features, significantly surpassing traditional methods in terms of speed and efficiency. Du *et al.* [49] developed a topological data analysis framework using persistent homology to extract essential topological features from the polarization data of polar oxide superlattices (Fig. 3(b)). By employing persistent images as descriptors, they constructed support vector regression-based CNN models for the automated and precise classification and regression of topological states. Trained on high-dimensional phase-field simulation data for PTO/STO superlattices, the framework demonstrated exceptional efficiency, generating strain and electric field phase diagrams within seconds. Additionally, they explored the dynamic evolution of topological features during electric field switching, revealing correlations between homology group lifetimes and skyrmion dynamics. Xu *et al.* [50] introduced a CNN-based DL protocol to analyze polaritonic wave images captured using scattering-type scanning near-field optical microscopy. Their method achieved over 1000-fold faster processing speeds than traditional techniques, enabling the rapid extraction of wavelengths and quality factors within 150 ms. Trained on simulated datasets, the CNN accurately analyzed experimental charge-transfer plasmon polariton images at graphene/α-$RuCl_3$ interfaces and simultaneously identified multiple polaritonic modes.

In addition to its efficiency, CV surpasses traditional methods in accurately predicting material performance by leveraging microstructural details. Ferdousi *et al.* [51] developed lightweight hybrid composites and investigated their process–structure–property relationships using statistical methods, theoretical modeling, and a CNN model (Fig. 3(c)). Their results revealed that the CNN model based on microstructure images provided significantly more accurate predictions of material strength, with a root mean square error 48.6% lower than that of the hybrid theoretical model. DeMille *et al.* [52] developed a framework for designing carbon nanotube (CNT) bundle microstructures using a genetic algorithm (GA) informed by a CNN. The GA selects the microstructural features of CNT bundles, while the CNN rapidly predicts mechanical properties—including elastic moduli, shear moduli, and Poisson's ratios—based on micromechanical finite element simulations. The CNN achieved high predictive accuracy ($R^2 > 0.96$ for elastic and shear moduli; $R^2 > 0.83$ for Poisson's ratios). The CNN-informed GA outperformed 79% to 100% of the solutions identified by brute-force search, making it a highly efficient tool for optimizing CNT bundle microstructures.

Building on this foundation, recent studies have introduced multimodal frameworks designed for specific material systems and properties, further enhancing the accuracy and interpretability of material property prediction [53–55]. Ren *et al.* [53] proposed a multimodal CNN-based framework for predicting the tensile properties of dual-phase steels by integrating compositional data with multisource microstructure images (Fig. 3(d)). This multimodal approach enabled accurate prediction of ultimate tensile strength and uniform elongation across a broad stress–strain range. Compared with traditional mean-field constitutive models, the multimodal approach improved prediction accuracy by 20%–36% and increased the prediction efficiency by more than 10 times. Moreover, reverse visualization techniques were used to determine strain distribution within various phase morphologies, thereby improving model interpretability.

These case studies demonstrate that DL-driven CV technologies for predicting material performance directly from the microstructure images of the materials have emerged as powerful and versatile tools. These workflows efficiently handle high-dimensional data, extract critical features, and uncover structure–performance relationships in materials. Another representative CV framework focuses on the automatic and accurate segmentation, classification, and quantification of phase structural features of materials from their microstructures, further establishing their relationships with the material properties through complementary approaches. The DL models used in frameworks such as R-CNN, FCN, and U-Net generate labeled images matching the dimensions of the input micrographs, where the phases or structures of interest are highlighted using specific red-green-blue (RGB) markers [56].

The DL-driven CV frameworks offer significant advantages over traditional segmentation methods in terms of speed, accuracy, and scalability, enhancing their reliability when extracting key information from microscopic images. Jiang *et al.* [57] developed a DL-assisted framework to analyze the detachment behavior of Ni-rich $LiNi_{0.8}Mn_{0.1}Co_{0.1}O_2$ (NMC) particles from the carbon binder domain in lithium-ion battery cathodes. A Mask R-CNN model was applied to automatically identify and segment over 650 NMC particles from 3D tomographic reconstructions, overcoming the inefficiency of conventional labor-intensive manual methods. This approach exhibited outstanding accuracy and efficiency, particularly in handling heavily damaged particles and identifying multiple fragments originating from the same particle. A comparative analysis showed that the Mask R-CNN model significantly outperformed conventional watershed and separation algorithms. Na *et al.* [58] introduced an FCN-based framework for unified microstructure segmentation, integrating weakly supervised learning with scribble annotations and active learning to minimize annotation costs while enhancing segmentation reliability. Comprehensive experiments revealed the versatility of the framework across various material classes, structural constituents, and imaging modalities, demonstrating performance comparable to or exceeding that

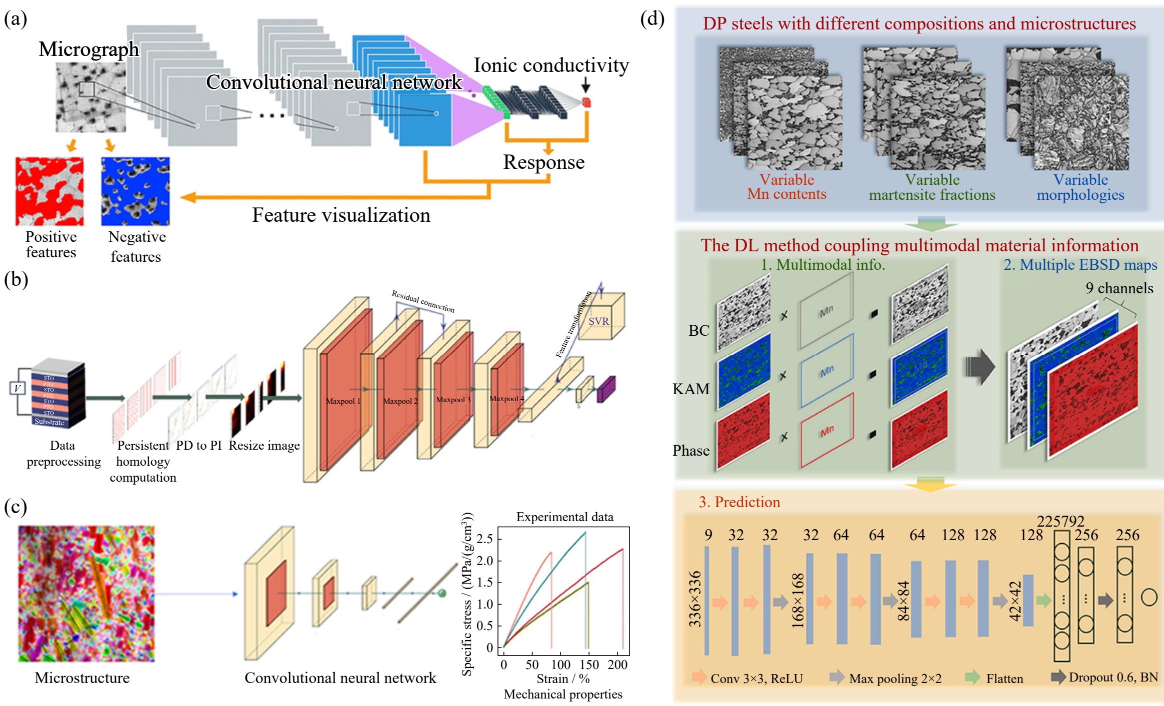

Fig. 3. Typical CV workflows for predicting the performance of a material from its microstructure: (a) workflow of the CNN-based framework for establishing structure–property relationships from SEM images [47]; (b) workflow of the topological data analysis framework using persistent homology for extracting essential topological features from its polarization data in polar oxide superlattices [49]; (c) workflow of the CNN-based framework for investigating the mechanical behaviors of 3D-printed hybrid composites [51]; (d) workflow of the CNN-based framework for predicting tensile properties of a material using the multimodal DP steel database [53]. (a) Reprinted from *Acta Mater.*, 141, Kondo *et al.*, Microstructure recognition using convolutional neural networks for prediction of ionic conductivity in ceramics, 29-38, Copyright 2017, with permission from Elsevier. (b) Reprinted from *Acta Mater.*, 282, Du *et al.*, Topological data analysis assisted machine learning for polar topological structures in oxide superlattices, 120467, Copyright 2024, with permission from Elsevier. (c) Reprinted from *Composptes Part. B*, 265, Ferdousi *et al.*, Investigation of 3D printed lightweight hybrid composites via theoretical modeling and machine learning, 110958, Copyright 2023, with permission from Elsevier. (d) Reprinted from *Acta Mater.*, 252, Ren *et al.*, Building a quantitative composition–microstructure–property relationship of dual-phase steels via multimodal data mining, 118954, Copyright 2023, with permission from Elsevier.

of fully supervised methods. The framework showed strong generalization to unseen micrographs from different sample sections, underscoring its robustness. This innovative approach transformed microstructure segmentation into an automated, high-throughput process, facilitating the accelerated design and optimization of materials. Dong *et al.* [59] developed a 3D DL model based on the U-Net architecture, which merged hyperspectral reflectance and RGB images to achieve accurate layer mapping of two-dimensional (2D) materials. Using this DL model, MoS_2 flakes with varying thicknesses (monolayer, bilayers, trilayers, and multilayers) were successfully identified and segmented, demonstrating the model's capability for high-dimensional data analysis.

With precise segmentation results, researchers can efficiently extract valuable information from microstructure images, including quantity, distribution, size, and morphology. This significantly streamlines data extraction and enables a clearer understanding of how preparation processes and microstructures of materials influence material properties. Chan *et al.* [60] developed a tool based on a Mask R-CNN model for segmenting and classifying polyhedral crystals in SEM images. Using this tool, the microcrystal size and product distribution of colloidal crystal products assembled from deoxyribonucleic acid (DNA) were extracted from over 13000 SEM images. By analyzing the crystal growth pathways, Chan et *al.* achieved targeted control over the size range of colloidal crystals by tuning DNA bond strength and thermal treatment conditions.

Building on this foundation, researchers have introduced more comprehensive analysis workflows by integrating post-processing methods, such as batch statistical analysis and machine learning, to uncover correlations between segmentation results and material performance. These CV-based approaches improve interpretability and support accurate prediction of material properties, as well as optimized material design and performance evaluation. Gorynski *et al.* [61] presented an R-CNN-based framework for the precise quantification of microstructural features to enhance understanding of structure–property relationships in materials (Fig. 4(a)). The R-CNN approach combined object detection with instance segmentation, enabling accurate identification and analysis of individual grains within a microstructure. Furthermore, they introduced Voronoi++, a postprocessing tool that simplified grain boundaries into polygons, allowing for measurement of angles and neighbor counts. Compared with conventional American Society for Testing and Materials (ASTM) methods such as the linear intercept and planimetric approaches [62], this framework improved relative precision from 10% to 3% and reduced manual workload by over 90%, thereby facilitates robust analysis of the structure–property relationships in materials. Kulesh *et al.* [63] developed a DL-driven framework for optimizing FePt heat-assisted magnetic recording (HAMR) media using TEM image segmentation to accelerate the analysis of structure–property relationships in materials. Their U-Net 3+ model accurately segmented FePt grains, identifying grain boundaries and sizes. By integrating segmented data with a gradient boosting regressor, the study identified optimal processing conditions to achieve desired grain size and distribution, which is key to improving HAMR media performance. Liu *et al.* [64] developed a U-Net-based framework to extract accurate microstructural information from microstructure images to analyze γ' precipitate coarsening in Co-based superalloys (Fig. 4(b)). The framework efficiently segmented and quantified γ/γ' microstructure images, capturing features such as γ' number, volume fraction, size, and morphology fraction. A random forest model was then used to identify key descriptors influencing γ' coarsening, including the Young's modulus difference between the γ and γ' phases and the valence electron number of constituent elements. This robust method enabled the design of advanced Co-based superalloys with enhanced γ' coarsening resistance and improved mechanical performance.

The DL-driven CV frameworks have revolutionized the analysis of microstructure images of materials, enabling the precise prediction of material performance with outstanding efficiency and accuracy. Compared with traditional manual image analysis, these frameworks dramatically reduce the time required for microstructure image analysis while offering deep insights into the relationships between material microstructures and performance. Beyond simplifying material characterization, the frameworks establish a solid foundation for optimizing material performance and accelerating the design of advanced materials, underscoring the potential of the frameworks as indispensable tools for advancing materials science and engineering.

3.2. Microstructure information generation

Another emerging research focus is the application of DL-driven CV technologies for image generation. Unlike classification or regression DL models, which primarily rely on labeled data for supervised learning, image generation DL models often employ unsupervised or semi-supervised learning approaches. The image generation models, such as the GAN, VAE, and diffusion models, do not require extensively labeled datasets, making them particularly valuable in scenarios where data labeling is time consuming or challenging. These models extract latent patterns and reconstruct information from existing microstructural data, enabling the synthesis of high-quality microstructure images, simulation of material states under hypothetical conditions, and enhancement of low-quality images [65].

Generative CV frameworks can synthesize realistic microstructures and predict their evolution. They address critical challenges, such as data scarcity, high costs of traditional experiments or simulations, and the need to explore uncharted design spaces. Recent studies exemplify the application of generative CV technologies across diverse material systems, highlighting their ability to reveal complex relationships between material processing and microstructures [66–67]. Cao *et al.* [66] developed a framework based on a conditional generative adversarial network (cGAN) to predict the microstructural features of Ti–6Al–4V alloys fabricated using laser powder bed fusion (LPBF) (Fig. 5(a)). The developed framework accurately reconstructed the microstructure images of materials and predicted their martensite morphologies and grain sizes based on laser-processing parameters, such as laser power and scan speed, with an accuracy of approximately 80%. Additionally, the framework could reveal the correlations between laser-processing parameters and grain morphology of materials through grain-size distribution maps, showcasing the scalability of the framework and its efficiency for laser-processing optimization toward tailored microstructure control. Azqadan *et al.* [67] developed a framework based on a denoising diffusion probabilistic model (DDPM) to generate high-resolution SEM images of AZ80 magnesium alloy microstructures. Trained on a dataset encompassing various cast-forging parameters, the framework generated realistic microstructures by accurately capturing their critical features, such as the $Mg_{17}Al_{12}$ phase morphology, dynamically recrystallized regions, and grain sizes. By bridging experimental and synthetic data, the study demonstrated the ability of DDPMs to save time and costs while enabling the exploration of the relationships between material processing and microstructures.

In addition to generating microstructure images of materials based on specific parameters, generative CV frameworks can be used to predict the time-dependent evolution of material microstructures. By modeling dynamic processes in materials, such as phase transformations, grain growth, or morphological changes under varying conditions, these frameworks provide a deep understanding of how microstructures evolve over time. Im *et al.* [68] introduced a generative cellular automata approach combining the principles of cellular automata and the U-Net architecture to simulate the growth of chiral gold nanoparticles. The approach effectively replicated key experimental observations, providing insights into the mechanisms driving the chiral morphology evolution and showcasing the potential of generative frameworks to model dynamic microstructural transformations. Sciazko *et al.* [69] proposed a physically constrained unsupervised image-to-image translation (UNIT) framework to predict microstructural changes in solid oxide fuel cell electrodes, addressing the challenge of unpaired data (Fig. 5(b)). By incorporating physical constraints, the framework modeled the reduction process of NiO and predicted its time-dependent microstructural evolutions, such as the changes in triple-phase boundary density and pore tortuosity. The UNIT framework has

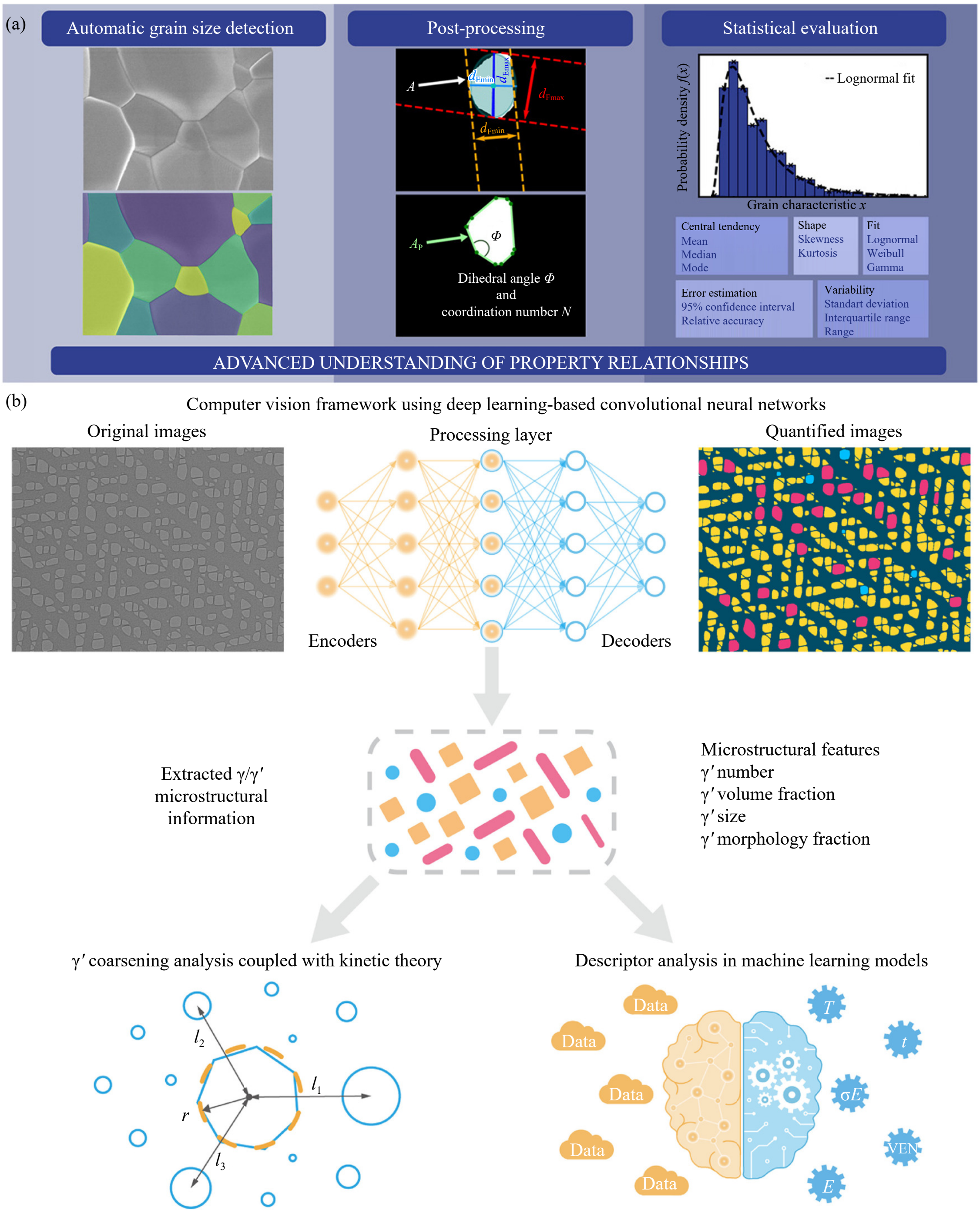

Fig. 4. Typical CV workflows for segmenting microstructures and quantifying their phases and structures: (a) workflow of the R-CNN-based framework for precisely quantifying microstructural features of a material to enhance the understanding of its structure–property relationship (A and A_p represent the marked area and polygon area, respectively; d_{Emax} and d_{Emin} represent the major and minor axis lengths of the ellipse, respectively; d_{Fmax} and d_{Fmin} represent the longest and shortest projected Feret diameters, respectively) [61]; (b) workflow of the U-Net-based framework for extracting accurate microstructural information from its γ/γ′ microstructure images required for the γ′ coarsening analysis (r represents the radius of γ′ particles; l_1, l_2, and l_3 represent the distance between the centroids of adjacent γ′ particles; T and t represent the aging temperature and aging time, respectively; E and σE represent Young's modulus and Young's modulus difference, respectively; VEN represents valence electron number) [64]. (a) Reprinted from *Acta Mater.*, 256, Gorynski *et al.*, Machine learning based quantitative characterization of microstructures, 119106, Copyright 2023, with permission from Elsevier. (b) Reprinted from *Acta Mater.*, 235, Liu *et al.*, Evolution analysis of γ′ precipitate coarsening in Co-based superalloys using kinetic theory and machine learning, 118101, Copyright 2022, with permission from Elsevier.

proven its versatility in analyzing diverse material systems and offers a critical tool for understanding and optimizing electrode performance.

In contrast to traditional thermodynamic methods, the DL-driven CV approaches still exhibit limited prediction accuracy and reliability. Computational thermodynamic tools, such as Thermo-Calc and DICTRA, based on the CALPHAD methodology and first-principles thermodynam-

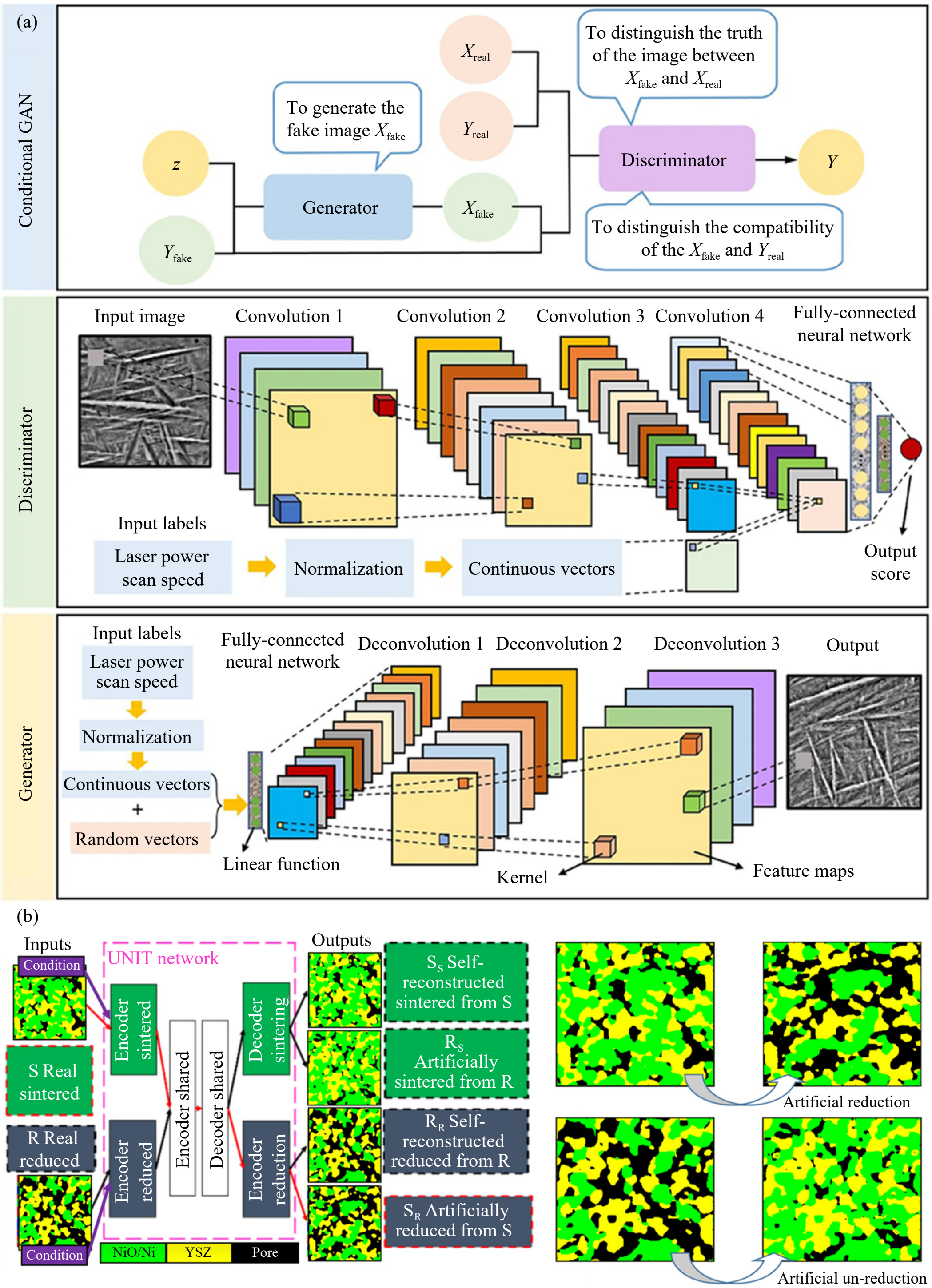

Fig. 5. Typical CV workflows for microstructure generation and evolution prediction in materials: (a) workflow of the cGAN-based framework for reconstructing and predicting the microstructures of Ti–6Al–4V fabricated using LPBF [66] and (b) workflow of the UNIT framework for predicting microstructural evolutions in solid oxide fuel cell electrodes [69]. (a) Reprinted from Ref. [66]; (b) Reprinted from Ref. [69].

ic data [70], have been extensively validated through experimental research and industrial applications. These tools can physically and rigorously predict phase equilibria, diffusion kinetics, and microstructural evolutions in materials. Their reliability results from the use of comprehensive thermodynamic databases and their strict adherence to physical laws, rendering them essential for both academic research and industrial process optimization. DL-driven CV approaches can derive empirical structure–property correlations in materials directly from their microstructure images, without depending on thermodynamic modeling. Although these emerging techniques show considerable potential, they are still in their developmental phase, and their widespread adoption is hampered by several critical challenges. These challenges include inherent dataset biases, interpretability limitations, and insufficient validation across diverse applications. Thus, the predictions generated by DL-driven CV approaches require rigorous verification of experimental results and established theoretical frameworks. Synergistic integration of these approaches is promising. For example, CALPHAD-generated thermodynamic constraints can be incorporated into DL frameworks to ensure physically consistent predictions, whereas DL-driven CV approaches can facilitate rapid high-throughput screening of materials. A combination of computational screening and thermodynamic validation will allow traditional thermodynamic methods to validate the most promising candidates identified through computational screening, thereby establishing an efficient closed-loop system that significantly reduces computational costs while maintaining thermodynamic fidelity.

In the prediction of material stress–strain fields, generative CV frameworks have effectively addressed key challenges such as low computational efficiency, data scarcity, and the complexity of traditional simulation methods. The frameworks allow for the rapid and accurate prediction of the stress–strain behavior of materials based on their microstructures. Yang *et al.* [71] proposed a CNN-based framework to predict microscale elastic strain fields in 3D high-contrast two-phase composites (Fig. 6(a)). Their approach, which overlooked traditional feature designs, significantly outperformed benchmark methods in terms of accuracy, computational efficiency, and generalizability, facilitating the exploration of process–structure–property relationships in materials in detail. Based on the results of previous studies, Gupta *et al.* [72] enhanced the capabilities of CV by introducing a framework specifically designed for multiscale mechanics modeling (Fig. 6(b)). They targeted fiber-reinforced composite microstructures and leveraged U-Net to predict local stress tensor fields under various loading conditions. Their study demonstrated the practicality of combining multiscale modeling with highly efficient computational workflows by seamlessly integrating upscaling and downscaling to compute macroscale material properties while capturing subelement-level stress distributions within the materials.

To facilitate predictive modeling, Yang *et al.* [73] utilized a cGAN-based framework to directly translate complex microstructural geometries into physical stress and strain fields (Fig. 6(c)). The approach displayed exceptional generalization capabilities by accommodating diverse material geometries, boundary conditions, and hierarchical structures. By predicting the stress fields around cracks, the framework also advanced material designs efforts associated with crack-resistant structures. The significant computational efficiency gains achieved by Yang *et al.* aligned with those achieved by Gupta *et al.*, emphasizing the versatility of generative DL models across different material systems. Building on these developments, Rashid *et al.* [74] introduced a neural operator-based framework using a Fourier neural operator to predict the stress–strain fields in 2D composite microstructures of materials. Unlike previous approaches, the approach of Rashid *et al.* could strongly generalize unseen object geometries with minimal training data. In addition to improving efficiency, the framework also processes low-resolution input data to produce detailed high-resolution stress–strain outputs, demonstrating the ability of the DL methods to handle data of varying quality and resolutions. For more complex systems, Wu *et al.* [75] developed a hybrid framework that combined a cGAN with an enhanced electro-chemo-mechanical (E-C-M) model to predict stress corrosion cracking in austenitic stainless steel placed in high-temperature water. By combining the strengths of the cGAN for stress–strain predictions and the E-C-M model for localized corrosion simulations, the framework exemplifies how hybrid approaches can address complex multiphysics problems. The integration of electrochemical and mechanical predictions enhances the design and optimization of corrosion-resistant materials, complementing earlier efforts, which focused only on structural and mechanical predictions.

Several CV frameworks have been developed to enhance both data quality and data acquisition efficiency in material microstructure characterization workflows. The frameworks leverage advanced DL architectures and algorithms to address key challenges, such as noise reduction, resolution enhancement, and high-quality dataset generation. By incorporating domain-specific knowledge, such as physics-informed loss functions or reciprocal space representations, the CV frameworks enable the reconstruction of accurate microstructural features, even under challenging experimental conditions [76]. Moreover, they significantly expedite the data acquisition processes, reduce reliance on manual intervention, and provide scalable solutions for handling large-scale datasets. Feng *et al.* [77] developed a cGAN-based framework to accelerate the multipoint statistics (MPS) reconstruction of 3D porous media. Through layer-by-layer reconstruction, the framework eliminates the need for sequential point-by-point simulations and captures and reconstructs topologically complex porous structures with excellent fidelity. The reconstructed microstructures and target systems, evaluated using metrics such as two-point correlation, lineal-path, and cluster functions, show excellent agreement. The model significantly reduces computational time and memory requirements, thereby addressing the limitations of traditional MPS-based approaches. Na *et al.* [78] introduced a multiscale refocusing network (MRN) to restore defocused SEM images (Fig. 7(a)). Using multiscale CNN architecture, the MRN effectively restores the defocused regions in the SEM images of martensitic steel and precipitation-hardened alloys, demonstrating exceptional qualitative and quantitative performance. The MRN not only enhances the SEM image quality of large datasets but also accelerates image acquisition, allowing for efficient data-driven material informatics. Khan *et al.* [79] developed a framework based on a cycle GAN to generate realistic scanning transmission electron microscopy (STEM) images (Fig. 7(b)). By integrating a reciprocal space discriminator, the framework minimizes the discrepancies between simulated and experimental data while preserving the ground truth. The resulting realistic and adaptable datasets enable high-throughput processing and flexible

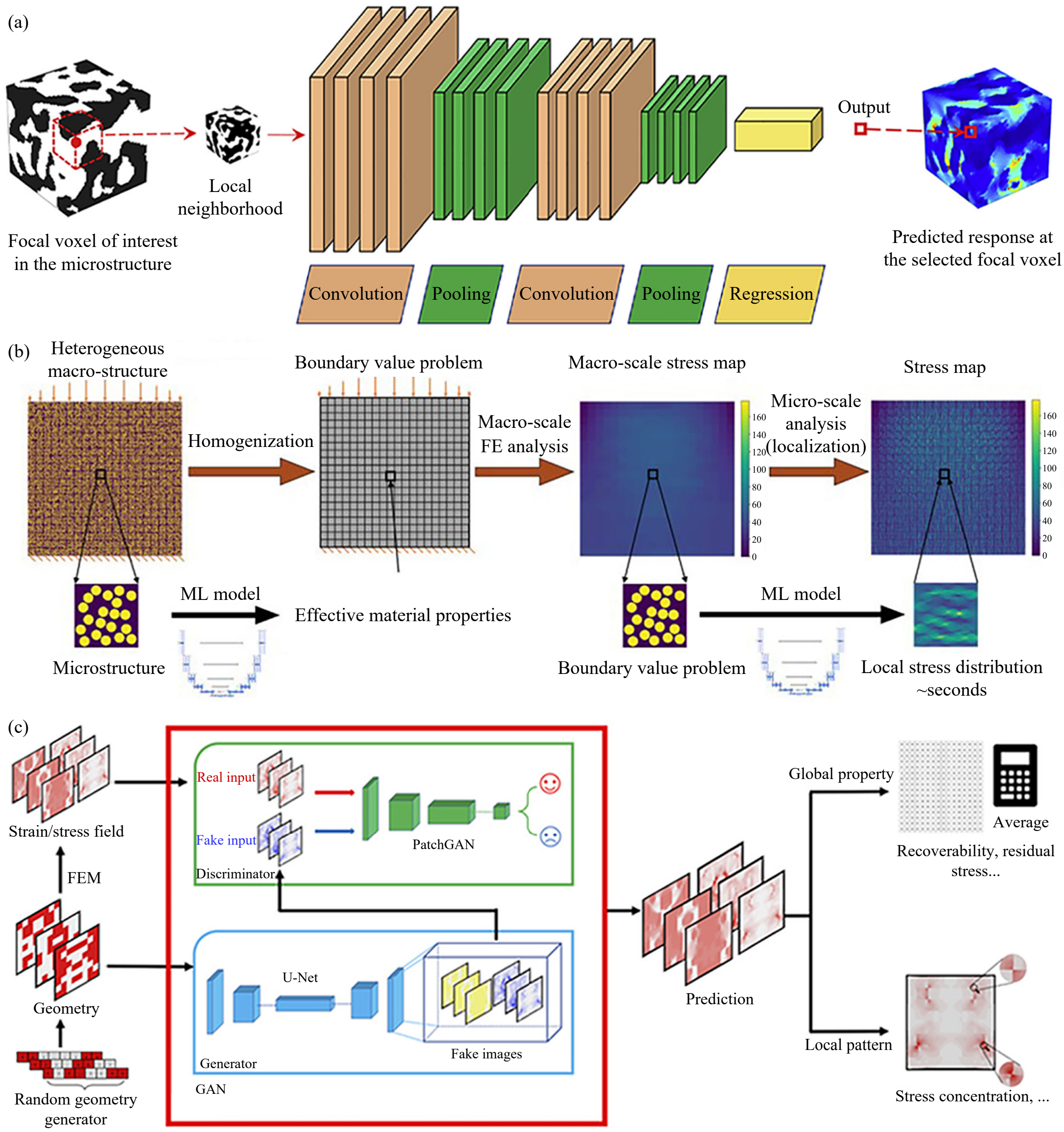

Fig. 6. Typical CV workflows for predicting material stress–strain fields based on material microstructures: (a) workflow of CNN-based framework for predicting microscale elastic strain fields in 3D high-contrast two-phase composites [71]; (b) workflow of the U-Net-based framework for multiscale mechanics analysis of fiber-reinforced composite microstructures [72]; (c) workflow of cGAN-based framework for predicting stress and strain fields directly from the material microstructure geometries [73]. (a) Reprinted from *Acta Mater.*, 166, Yang *et al.*, Establishing structure–property localization linkages for elastic deformation of three-dimensional high contrast composites using deep learning approaches, 335-345, Copyright 2019, with permission from Elsevier. (b) Reprinted from *Mech. Mater.*, 184, Gupta *et al.*, Accelerated multiscale mechanics modeling in a deep learning framework, 104709, Copyright 2023, with permission from Elsevier. (c) From Z.Z. Yang, C.H. Yu, and M.J. Buehler, *Sci. Adv.*, 7, eabd7416 (2021) [73]. Reprinted with permission from AAAS.

machine learning applications in microscopy, offering a transformative step toward fully autonomous atomic-resolution material research. Khan *et al.* trained an FCN to identify single-atom defects in STEM images. The defect-detection technique of the FCN is discussed in the next section. Jangid *et al.* [80] designed a physics-guided super-resolution framework for electron backscatter diffraction (EBSD) orientation maps used in crystallographic analysis (Fig. 7(c)). Their approach accounted for rotational symmetry and misorientation by leveraging quaternion-based orientation representations and physics-informed loss functions. Attention-based models, such as holistic attention networks, significantly enhance the resolution, accuracy, and noise reduction in grain morphology analysis of materials. By accelerating the high-throughput EBSD mapping for both 2D and 3D datasets, this adaptable framework underscores the potential of integrating domain-specific physics into CV frameworks.

The advancements in CV technologies involving DL-

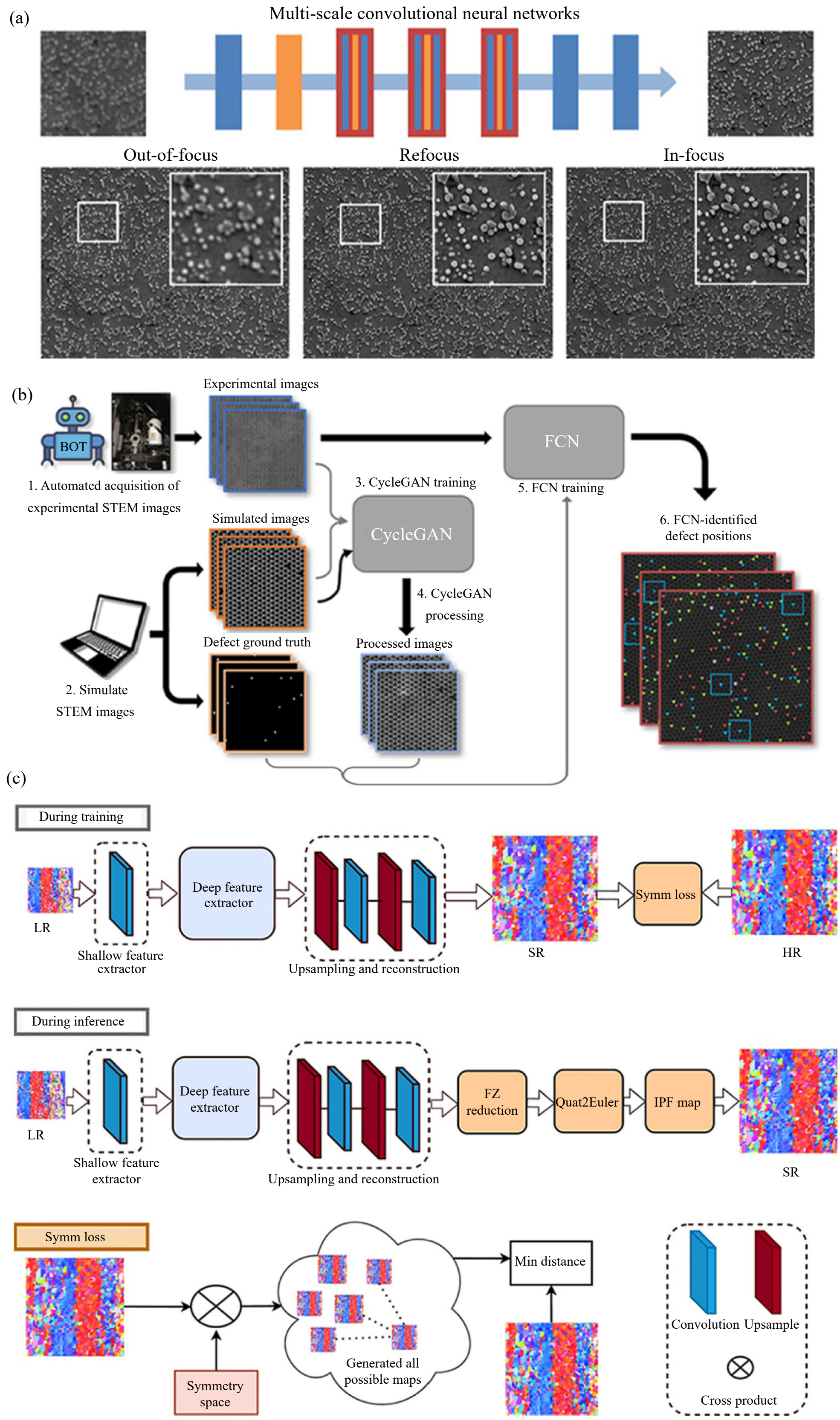

Fig. 7. Typical CV workflows for enhancing the data quality and data acquisition efficiency in material microstructure characterization: (a) workflow of the MRN for refocusing defocused SEM images [78]; (b) workflow of CycleGAN-based framework for generating realistic STEM images [79]; (c) workflow of the super-resolution framework for electron backscatter diffraction orientation maps [80]. (a) Reprinted from *Acta Mater.*, 214, Na *et al.*, Deep learning-based discriminative refocusing of scanning electron microscopy images for materials science, 116987, Copyright 2021, with permission from Elsevier. (b) Reprinted from Ref. [79]. (c) Reprinted from Ref. [80].

driven image generation and evolution prediction offer a transformative approach for addressing the key challenges faced in materials science. The generative CV frameworks not only bridge the gap between experimental and synthetic data but also reveal the complex relationships between material microstructures and processing conditions or time dependencies, thereby enhancing the efficiency of material design and process optimization. The ability of CV frameworks to simulate dynamic microstructural changes in materials and predict their performance will be crucial for accelerating the development of materials with tailored performance.

3.3. Microstructure defect detection

The study of microscopic defects, such as cracks, voids, dislocations, and impurities, in materials is crucial for optimizing their properties and ensuring the quality of components manufactured using those materials. With the advancement of DL-driven CV technologies, significant strides have been made in automating the detection, classification, and characterization of material defects at the microscopic level [81–83]. Unlike traditional methods that rely on manual inspection or conventional image-processing techniques, DL-based CV technologies offer robust, scalable, and efficient tools for analyzing large microscopy image datasets.

DL-driven defect detection frameworks can be categorized based on their primary objectives: static defect characterization, dynamic evolution tracking, and defect–property correlation. The ability of CV to distinguish microstructural defects has been demonstrated. Madsen *et al.* [84] introduced a CNN-based framework to analyze high-resolution TEM images, offering a robust classification of atomic structures and detection of local defects under different experimental conditions. Ziatdinov *et al.* [85] developed an FCN-based framework to analyze atomically resolved STEM images, using a weakly supervised approach to extract chemical and structural information. The framework could identify atomic positions, classify defects, and track material transformations, even for unseen defect types. It revealed reversible transitions between 3-fold and 4-fold Si dopants in graphene, mixed-coordination Si dimers, and dynamic transformations, such as molecular "rotor" motion. The scalable automated approach mimicked human reasoning, enabling real-time defect analysis, and laid the foundation for autonomous "self-driving" microscopes. Li *et al.* [86] developed a framework that used a cascade object detector, a CNN, and localized image analysis methods for automated defect detection in electron microscopy images (Fig. 8(a)). The framework had a recall rate of 0.842 and an accuracy rate of 0.837 for detecting dislocations and voids, exceeding the manually labeled values of 0.804 and 0.790, respectively, while maintaining robust performance across different image contrast, brightness, and magnification values. With continuing advancements, the accuracy and efficiency of defect detection using CV have progressively outpaced human capabilities, thereby widening the performance gap. Badmos *et al.* [87] developed a CNN-based framework for the automated detection of microstructural defects in lithium-ion battery electrodes and focused on identifying defects that affected the performance and lifespan of the batteries. A fine-tuned VGG19 model, using transfer learning, exhibited excellent F1 scores, the harmonic means of precision and recall, of 0.99 and 1.00 for defect-containing images and defect-free images, respectively. The framework automatically identified critical features, such as foreign particles and deformed layers, in materials while using activation map visualizations for precise defect localization. It enabled the efficient analysis of thousands of micrographs and offered a scalable and practical solution for quality control in battery manufacturing. Shen *et al.* [88] used a Faster R-CNN-based framework for advanced defect detection in electron microscopy images, specifically targeting dislocation loops in irradiated ferritic alloys. The framework accurately detected a range of defect morphologies, including elliptical loops and solid dots, using a small training dataset. It provided quantitative information on defect sizes and areal densities, demonstrating an F1 score of 0.78 at a speed of 0.1 s per image, which is hundred times faster than manual analysis. This innovative approach highlights the potential of DL for scalable and efficient defect detection in high-resolution electron microscopy images with the ability to accommodate new defect types as they appear. Dey *et al.* [89] developed a DL-based system for classifying and localizing defects in SEM images by focusing on grain boundary misorientations and voids. Using an ensemble model with ResNet and VGGNet architectures, along with a preference-based strategy, they demonstrated significant improvements in defect classification and detection. The system also removed SEM image noise using an unsupervised denoising technique, reducing false positives and enhancing accuracy.

Researchers have now gained a deep understanding of defect evolution in synthetic and service environments and its impact on material properties. Their studies have offered robust data support for investigating defect mechanisms in materials, laying a solid foundation for optimizing material design and improving material performance. Maksov *et al.* [90] developed a deep CNN-based framework to analyze the defects and phase evolution in Mo-doped WS_2 subjected to electron-beam-induced transformations. The framework, trained on a single STEM image, rapidly detected thousands of lattice defects, classified them via unsupervised clustering, and extracted diffusion parameters for sulfur vacancies. It also reconstructed spatiotemporal diagrams and showed the transition probabilities of defect complexes, allowing for physics-guided AI tools for dynamic atomic-scale analysis. Maxim *et al.* [91] combined electron beam (e-beam) manipulation with FCN-based analysis to monitor the deterministic motion of Si dopants in graphene. They captured structural changes in a material during atom manipulation, revealing symmetry-breaking phenomena and providing statistical insights into the defect configurations in the material. The integration of DL and e-beam control highlights the potential for atom-by-atom defect engineering. Shen *et al.* [92] developed a YOLOv3-based framework to analyze *in situ* TEM ion irradiation videos. Their system achieved human-level accuracy (an F1 score of 0.89) in detecting dislocation loops in FeCrAl alloys, while enabling scalable and consistent defect tracking. By combining detection with dynamic analysis, the system provided insights into defect growth rates and enabled real-time experimental adjustments. To further enhance the defect tracking efficiency, Sainju *et al.* [93] introduced DefectTrack, a multiple object tracking (MOT) framework based on a high-resolution network designed for *in situ* TEM videos. DefectTrack achieved an MOT accuracy of 66.43% and an F1 score of 79.38%. It processed full datasets in 57.14 s, unlike human experts who required 5.25 h to analyze just one-tenth of the defect clusters. Moreover, its accuracy in predicting defective lifetime distributions significantly surpassed that of humans, demonstrating both efficiency and precision gains over traditional manual analyses. It underscored the potential of automated tracking tools for promoting the understanding of irradiation-induced defect dynamics.

Researchers have now gained a comprehensive and in-depth understanding of defect evolution in materials within both synthetic and service environments and the influence of the defects on material properties. Previous studies have provided the data required for investigating defect mechanisms in materials, thereby establishing a solid scientific foundation for optimizing material design and enhancing material performance. Lee *et al.* [94] developed an FCN-based framework to achieve sub picometer precision in mapping strain fields induced by single-atom defects in 2D transition metal dichalcogenides. Their method used aberration-corrected STEM images to classify point defects in materials and by generating high signal-to-noise class averages, enabling precise 2D atomic spacing measurements. The method revealed complex oscillating strain fields around the Se vacancies in $WSe_{2-2x}Te_{2x}$, which aligned with density functional theory (DFT) predictions and offered insights into defect-induced strain phenomena in the material. Jacobs *et al.* [95] developed a Mask R-CNN-based framework to analyze defects in the TEM images of irradiated FeCrAl alloys. The framework accurately predicted defect shapes, sizes, and densities, which were crucial for understanding the effects of irradiation on materials. It demonstrated robust performance with limited training data and predicted irradiation-induced hardening with an accuracy of 10–20 MPa. Chen *et al.* [96] developed a U-Net++-based framework to analyze *in situ* TEM videos, enabling a detailed investigation of the stability and evolution of irradiation-induced voids in nickel. In contrast to the traditional manual measurement method requiring approximately 30 min, the framework completed data analysis in 0.6 s and achieved an intersection over union of 0.94 for large pore detection, demonstrating its strong ability to detect medium-to-large voids. Frame-by-frame segmentation of TEM videos revealed the temperature at which the voids transitioned between growth and shrinkage and quantified the void shrinkage rate as a function of temperature and void size. The study offered new insights into void stability mechanisms, emphasizing the influence of vacancy mobility and vacancy cluster retention at low temperatures. Komninos *et al.* [97] proposed a transformer-based framework for predicting the remaining useful life (RUL) of composite structures under fatigue conditions. Their proposed framework effectively correlated crack propagation characteristics with RUL predictions, enhancing the interpretability and accuracy. Bilyk *et al.* [98] developed a Mask R-CNN-based framework to analyze dislocation loops in ion-irradiated CrFeMnNi alloys (Fig. 8(b)). Using data augmentation and image normalization, they segmented and characterized the dislocation loops in materials, offering insights into their mobility and supporting mechanisms, such as Ostwald ripening driven by vacancy diffusion.

These developments collectively underscore the transformative role played by DL-driven CV in the detection and analysis of microstructural defects. The CV not only significantly enhances the efficiency and accuracy of identifying material defects but also provides essential tools for predicting material performance degradation and guiding material optimization. By bridging the gap between experimental observations and theoretical insights, these approaches have enabled researchers to gain a deep understanding of defect behavior in materials, enabling them to address complex challenges in materials science with high precision and scalability.

3.4. Crystal structure-based material property prediction

The crystal structure of a material, the 3D spatial arrangement of atoms in the material, exhibits unique periodic geometric configurations directly governing the interatomic interaction strengths, bond length/angle distributions, and electron cloud arrangements in the material. The atomic-scale structure–property relationship of a material represents a fundamental research paradigm in materials science, offering a mechanistic basis for understanding the various physical and chemical properties of the material. Conventional investigations primarily use deterministic computational methods, such as DFT and molecular dynamics simulations to predict material properties from crystal structures using quantum mechanical equations. However, these approaches face significant challenges in terms of computational costs and scale limitations. The large datasets generated using computational methods contain material properties, such as band structures, formation energies, and elastic tensors, and offer physically interpretable training sets for AI, thereby effectively mitigating the impact of experimental data scarcity. Recently, researchers have developed topological representations of crystal structures [99] that can be fed to graph neural networks (GNNs). The representations preserve both the symmetry and periodicity of the crystals while introducing physical constraints into DL models.

Thus, DL-driven CV frameworks have achieved end to end prediction ability by predicting material properties from

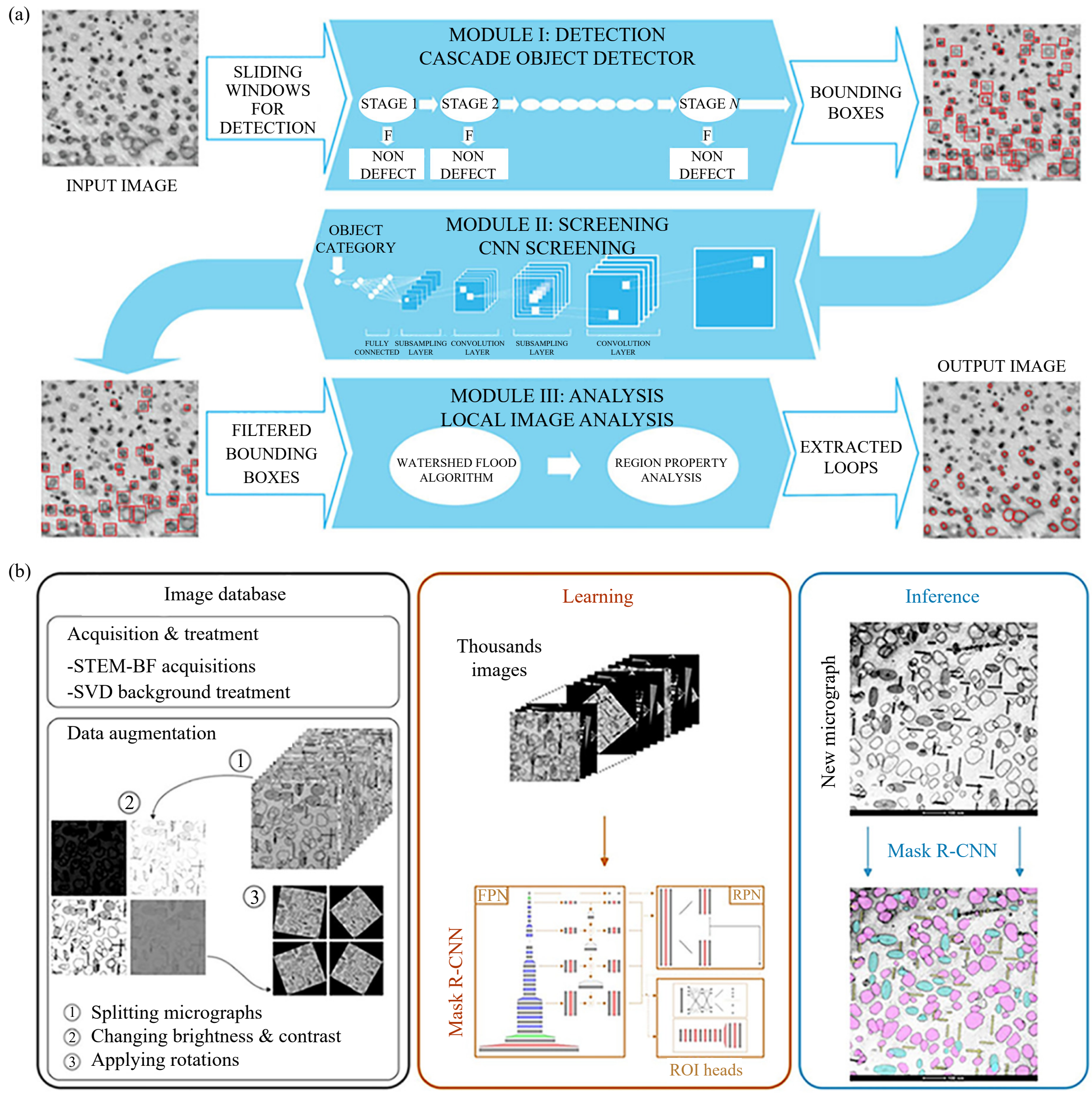

Fig. 8. Typical CV workflows for defect detection: (a) workflow of a cascade object detector, a CNN, and localized image analysis methods framework combined for automated defect detection in electron microscopy images [86] and (b) workflow of the Mask R-CNN-based framework for detecting dislocation loops [98]. (a) Reprinted from Ref. [86]. (b) Reprinted from Ref. [98].

their crystal structures using large material databases generated through DFT. The frameworks reflect the adaptability of CV to different length scales, from the microscopic to atomic and molecular scales. Xie and Grossman [15] introduced the CGCNN, a framework that represents materials as graphs with atoms and chemical bonds represented as nodes and edges, respectively (Fig. 9(a)). The CGCNN directly learns the material properties from the atomic arrangement in the material crystal structure, offering a universal and interpretable representation of crystalline materials. The CGCNN, trained on the data obtained from the "Materials Project" database, demonstrated high accuracy in predicting eight different material properties, such as formation energy and band gap, for crystals with different structure types and compositions. Moreover, the interpretability of the model allows for the extraction of local chemical contributions to global material properties. For example, the model not only accurately predicts perovskite material properties but also reveals empirical rules that guide perovskite material design and significantly reduces the search space required for high-throughput screening of perovskite materials. The CGCNN exhibits higher out-of-distribution (OOD) prediction errors than the baseline model for three benchmark datasets of MatBench, the standardized task platform for machine learning performance testing and benchmarking in the field of materials science, including dielectric, elasticity, and perovskite datasets. However, it outperformed the other models when working on OOD tasks within the perovskite dataset, highlighting its strong adaptability to specific material systems [100]. This performance discrepancy arises because the CGCNN primarily relies on local atomic environments, which constrain its ability to handle global structural variations, such as changes in crystal symmetry, across different material systems.

Park and Wolverton [101] further refined the CGCNN developed by Xie and Grossman and introduced the improved CGCNN, known as iCGCNN, (Fig. 9(b)). This enhanced framework incorporated additional features, such as Voronoi tessellation, three-body correlations of neighboring atoms, and optimized chemical representations of interatomic bonds, which significantly improved predictive accuracy for material properties. After undergoing training on a large dataset and incorporating the mentioned refinements, the iCGCNN outperformed the original model in two key areas. First, in predicting the thermodynamic stability, the iCGCNN showed a 20% improvement in accuracy over the original CGCNN. Secondly, in high-throughput searches for stable compounds in the $ThCr_2Si_2$ structure type, the iCGCNN demonstrated a success rate that was 2.4 times higher than that of the original CGCNN, significantly accelerating the discovery of stable materials. This improved model proved to be more efficient in material design and discovery than the original model, offering high accuracy and applicability across a broad range of materials.

Karamad *et al.* [102] further improved the CGCNN by introducing an orbital graph CNN (OGCNN) that incorporated atomic orbital interaction features into graph representations of materials (Fig. 9(c)). This enhancement improved the ability of the model to predict material properties, such as magnetic and electronic characteristics, by encoding orbital-orbital interactions within the local chemical environment of atoms. The OGCNN also contained an encoder-decoder network, which allowed it to extract important features from atomic, orbital, and topological data. When benchmarked against the CGCNN and other state-of-the-art material representation methods, including many-body tensor representations and smooth overlap of atomic positions, the OGCNN demonstrated a significantly higher predictive accuracy. This improvement makes the OGCNN a powerful tool for material discovery, enabling accurate predictions across a wide range of crystalline materials.

Several studies have further explored the versatility of GNN in crystal graph analyses. Chen *et al.* [103] developed the MatErials Graph Network (MEGNet), a GNN-based framework designed to predict material properties of both molecules and crystals (Fig. 10(a)). The MEGNet uses global state features such as size, shape, and spatial arrangement, which enhance its ability to understand molecular structures and improve prediction accuracy for material properties. Its architecture included sequential operations that updated bond, atom, and global state attributes. Atomic and bond features were updated based on the connectivity within the atomic graph, whereas global-state features encoded molecu-

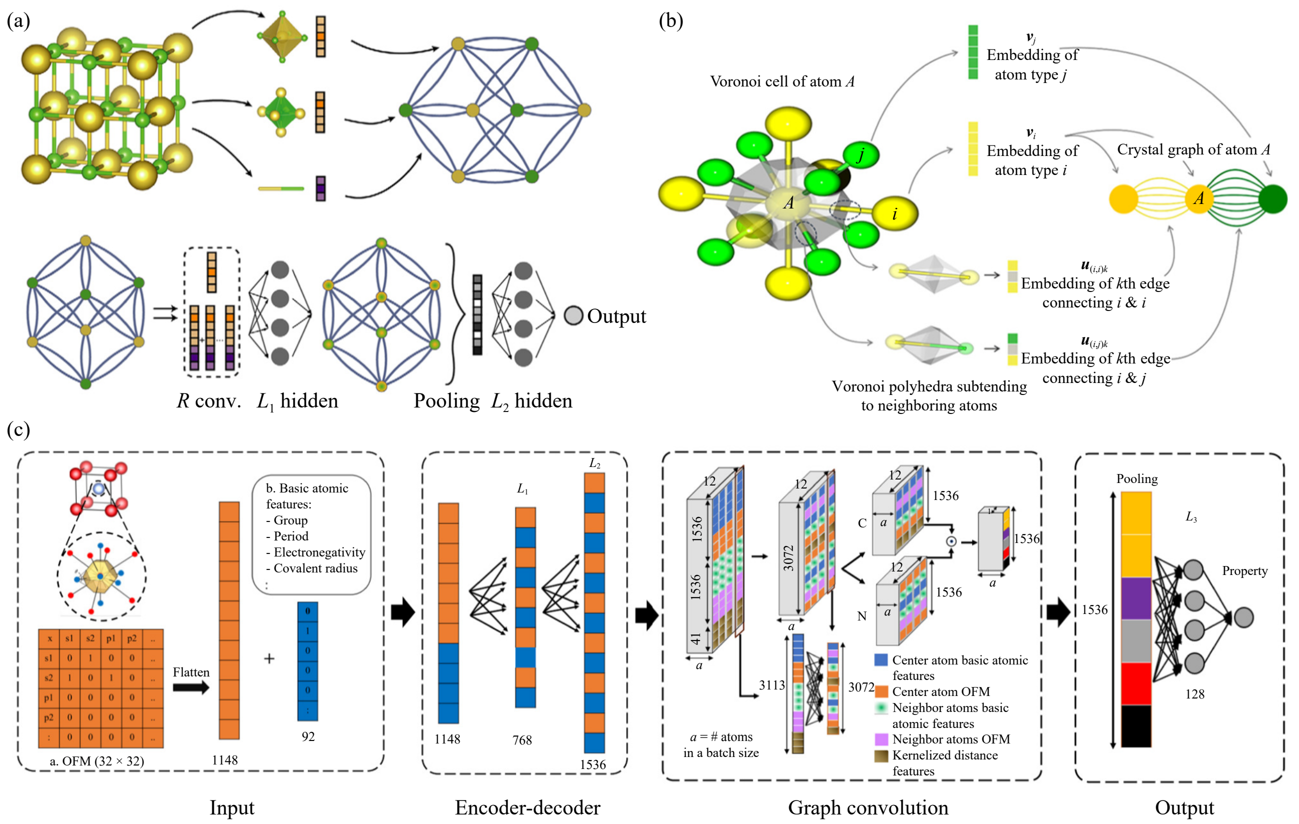

Fig. 9. Workflow of the CGCNN and its improved version: (a) workflow of the CGCNN [15]; (b) workflow of improved iCGCNN (v_i and v_j represent the atomic properties embedding vectors of atoms *i* and *j*, respectively; $u_{(i,i)k}$ and $u_{(i,j)k}$ represent the geometric-correlation edge-embedding vectors for the neighboring atom pairs *i–i* and *i–j*, respectively) [101]; (c) workflow of OGCNN framework [102]. (a) Reprinted with permission from Xie and Grossman, *Phys. Rev. Lett.*, 120, 145301 (2018) [15]. Copyright 2018 by the American Physical Society. (b) Reprinted with permission from Park and Wolverton, *Phys. Rev. Mater.*, 4, 063801 (2020) [101]. Copyright 2020 by the American Physical Society. (c) Reprinted with permission from Karamad *et al.*, *Phys. Rev. Mater.*, 4, 093801 (2020) [102]. Copyright 2020 by the American Physical Society.

lar information, enabling long-range interactions. The MEGNet used Gaussian distributions to model atomic relationships and addressed the challenges posed by nonuniformly distributed data. Its accuracy surpassed that of DFT for properties such as formation energies, bandgaps, and elastic moduli when trained on approximately 60000 crystals from the Materials Project. Louis *et al.* [104] introduced a global attention mechanism into GNN with their framework, i.e. global attention mechanism with graph neural network (GATGNN), which integrates augmented graph-attention layers (AGAT) and a global attention layer (Fig. 10(b)). The GATGNN architecture enables the framework to capture local atomic relationships through the AGAT layers, whereas the global attention layer computes the importance of atomic interactions and adapts to the irregular structures of the crystal graphs. By assigning weights to connections between nodes and aggregating information from neighboring nodes based on the assigned weights, the framework effectively captures long-range dependencies and varying node contributions. The multilayer graph convolution and multi-head attention mechanisms further enhance the capacity of the framework to learn atomic and bond features, significantly improving both the predictive performance and generalization ability of the framework. The approach allows the GATGNN to make material property predictions with increased accuracy and explains how individual atoms contribute to the overall behavior of the material. Based on the architectural strengths of GATGNN, the research team proposed its deep extension, deeperGATGNN [105], which addressed the oversmoothing issue in traditional GNNs by integrating differentiable group normalization (DGN) and residual skip connections. The DGN dynamically clusters node features into groups and normalizes them separately, effectively preserving the distinct characteristics of atomic communities and preventing the homogenization of node representations in the deep layers. Meanwhile, residual skip connections enable cross-layer feature fusion, retain critical shallow-level atomic interaction information, and stabilize network training with over 30 graph convolution layers. According to the experimental results, the DeeperGATGNN demonstrated state-of-the-art performance on five out of the six benchmark datasets for material property prediction. The DeeperGATGNN outperformed OOD prediction for complex crystal structures, such as the perovskite dataset in MATBench through deep attention mechanisms and normalization techniques, particularly in global feature modeling and capturing long-range dependencies [100]. However, it has large-scale data computational resource requirements, leading to poor performance in low-attribute-density tasks.

Wen *et al.* [106] developed BonDNet, a GNN-based framework designed for the rapid and accurate prediction of bond dissociation energies (BDEs) for charged molecules, highlighting the ability of graph-based models to predict molecular properties extending beyond bulk material properties (Fig. 10(c)). BonDNet uses chemically inspired features, including global features, such as molecular charge, to effectively represent bond-dissociation reactions. It uses the differences between the atom, bond, and global features of reactants and products to understand bond-breaking reactions. The framework demonstrated an impressive mean absolute error (MAE) of 0.022 eV for unseen test data, surpassing previous models in terms of accuracy. The ability of BonDNet to handle both homolytic and heterolytic BDEs for molecules of any charge, along with its ability to gain chemical insights through feature analysis, makes it a powerful tool that can be used to understand complex chemical processes, including drug metabolism, biofuel combustion, and photochemical decontamination. Choudhary and DeCost [107] developed an atomistic line graph neural network (ALIGNN), a novel GNN-based framework designed to improve material property prediction by incorporating both interatomic bond graphs and their corresponding line graphs that capture bond-angle information (Fig. 10(d)). The traditional GNNs used for crystal graphs focus primarily on atomic distances, overlooking the influence of bond angles, crucial for distinguishing atomic structures. The ALIGNNs incorporate the influence of bond angles by passing messages on both atomistic and line graphs, allowing them to propagate bond angle information alongside atomic distance data. This integrated approach enhances the ability of the framework to capture detailed atomic structural information, leading to improved framework performance in various atomistic prediction tasks. The ALIGNN was tested on solid-state and molecular properties from databases such as JARVIS-DFT, Materials Project, and QM9, and it outperformed some of the previous GNN frameworks with regard to prediction accuracy and training speed. Meanwhile, ALIGNN shows excellent adaptability in OOD tasks, significantly outperforming the baseline in SparseY single tests on dielectric, elasticity, and perovskite datasets, with improvements of 7.3%, 32.8%, and 9.7%, respectively [100]. However, because of its reliance on line graph encoding and two-level convolution operations, the training and inference costs of ALIGNN are high, particularly for large-scale datasets. Gao *et al.* [108] proposed a generic crystal pattern graph neural network (GCPNet), a novel framework that integrates complete geometric features, such as bond angles and lattice vectors, into crystal pattern graphs and applies a graph convolutional attention operator (GCAO) and a two-level update mechanism to address the limitations of existing GNNs used in material property prediction. The GCPNet dynamically aggregates local atomic interactions via graph convolution while adaptively weighting long-range dependencies through attention mechanisms, demonstrating state-of-the-art performance on five benchmark datasets with MAE reductions in the range of 14.69%–49.61% while reducing bulk crystal formation energy errors to 0.0264 eV/atom. The interpretability of the GCPNet enabled the extraction of local site energies in perovskites, guiding high-throughput screening with an efficiency higher than that of CGCNN by 32%; its robustness was validated by its stable performance, even in the presence of 128-dimensional GCAO layers without any overfitting.

In summary, crystal structure-based material–property prediction using advanced DL-driven CV has demonstrated

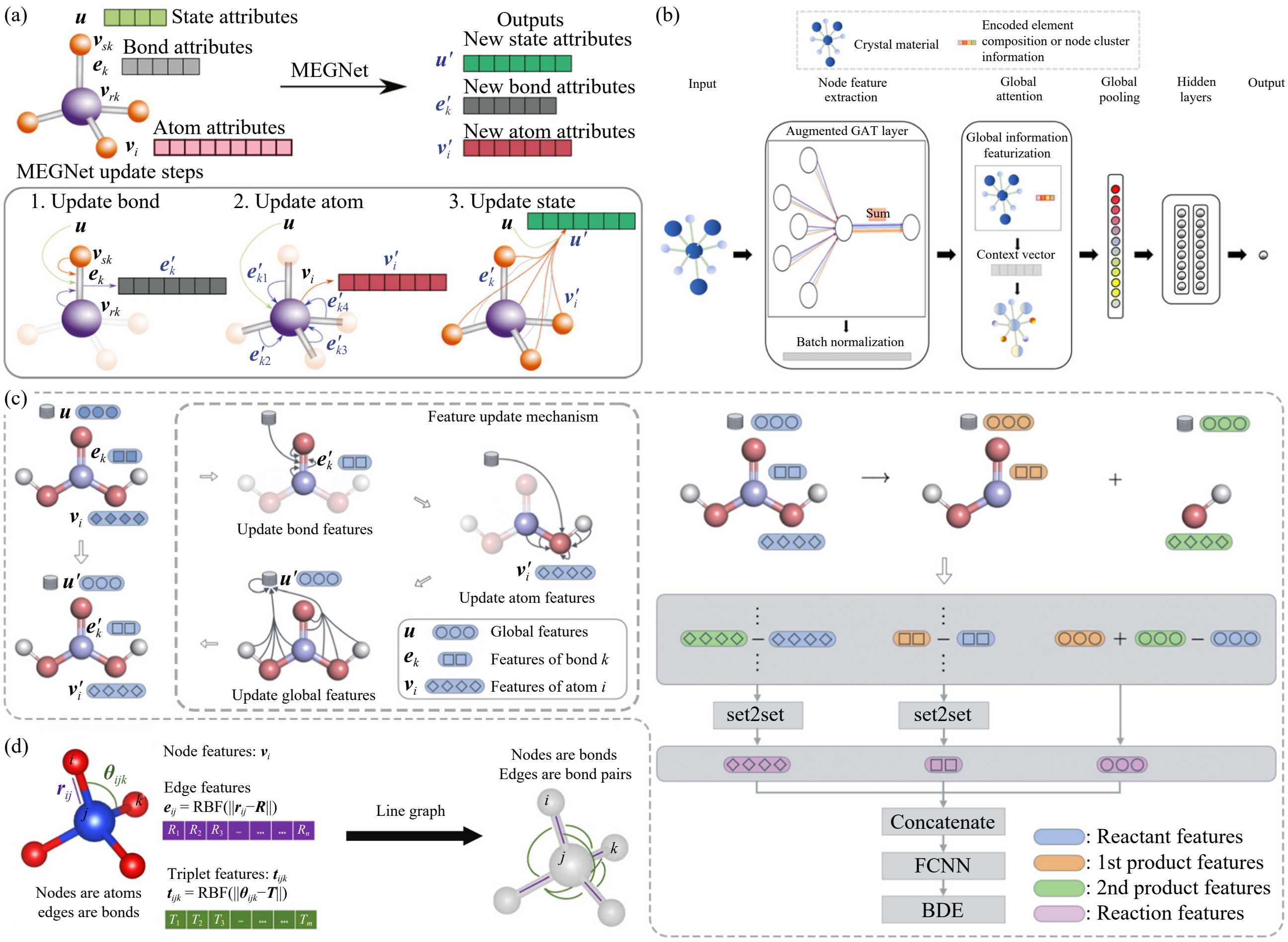

Fig. 10. Typical GNN-based workflows for predicting material properties from crystal structures: (a) workflow of the MEGNet framework (e_k represents bond attribute vector for bond k; u represents the global state vector storing the molecule/crystal level or state attributes; e'_k represents the updated bond vector incorporating the original bond attributes, the features of the bonded atoms, and the global state; v'_i represents the updated atomic vector combining the intrinsic features of atom i, the features of its neighboring atoms, the associated bond features, and the global state; u' represents the updated global state vector aggregating all atomic features, bond features, and the global state) [103]; (b) workflow of the GATGNN framework [104]; (c) workflow of the BonDNET framework [106]; (d) workflow of the ALIGNN framework (r_{ij} represents the atomic displacement vector from atom i to atom j; θ_{ijk} represents the bond angle formed by the triplet (i, j, k); e_{ij} represents the edge features constructed by applying a radial basis function (RBF) expansion to the interatomic distance; t_{ijk} represents the triplet features generated through an RBF expansion of the bond-angle cosine; R is the set of distance expansion centers; T is the set of angle expansion centers) [107]. (a) Reprinted with permission from Chen *et al.*, *Chem. Mater.*, 31, 3564-3572 (2019) [103]. Copyright 2019 American Chemical Society. (b) Reproduced from Ref. [104] with permission from the PCCP Owner Societies. (c) Reprinted from Ref. [106]. (d) Reprinted from Ref. [107].

significant potential for delivering accurate, scalable, and interpretable predictions. The development of frameworks, such as CGCNN, OGCNN, and ALIGNN, has significantly progressed material discovery, enabling researchers to use crystal structure data in the design of new materials with optimized properties. These advancements allow for a promising future in which machine learning will play a central role in the design and discovery of next-generation materials.

4. Summary and outlook

The application of CV has led to transformative advancements in different fields of materials science. Through the development of innovative frameworks, DL-driven CV frameworks have made significant progress in microstructure-based performance prediction, microstructure information generation, microstructure defect detection, and crystal structure-based material-property prediction. These advances have accelerated the microstructural characterization, performance evaluation, and optimization of materials.

The CV frameworks for microstructure-based performance prediction of materials have proven to be effective in accurately forecasting material performance, shortening analysis times, and offering deep insights into the structure–performance relationship of a material. Image generation and evolution prediction CV frameworks bridge the gap between experimental and synthetic data, thereby improving material design and process optimization. The frameworks can also simulate dynamic microstructural changes and stress fields in materials during their use, which is crucial for assessing the material behavior under various real-world conditions. Similarly, microstructure defect detection has seen significant improvements, with CV frameworks now being capable of real-time defect detection and tracking in materials, providing valuable insights into material defect evolution and its impact on material performance under various conditions. The CV frameworks based on crystal graphs exhibit a significantly higher computational efficiency than DFT calculations when utilizing crystal structure data, offering a scalable solution for material property prediction and high-throughput screening. The DFT remains the gold standard for accuracy because of its rigorous quantum mechanical foundations. In practice, these approaches can be synergistically integrated into a complementary workflow. The CV frameworks can rapidly prescreen large material spaces following which DFT is applied to validate high-priority candidates. This collaborative strategy not only substantially reduces the computational cost of DFT through DL but also preserves the physical rigor of first-principles calculations.

Despite the remarkable progress made in the application of CV in materials science, some limitations and ample opportunities for improvement still exist. Three key cross-cutting challenges emerged from this review: ensuring dataset reliability and quality, improving model interpretability, and developing integrated multinetwork frameworks. Firstly, many studies rely on synthetic or limited experimental datasets that often lack real-world validation or proper labeling, underscoring the urgent need for standardized benchmarking protocols. Secondly, although attention mechanisms and visualization tools have enhanced model transparency, physics-informed DL approaches remain underexplored. Thirdly, an important transition from single-network models to sophisticated, nested multinetwork architectures that continuously refine material performance evaluations is taking place. Unsupervised image generation models, such as GANs, offer a promising solution to the data scarcity problem by synthesizing high-fidelity microstructure images for training. Recent advances in physics-informed transfer learning frameworks, such as those combining CALPHAD-generated thermodynamic constraints with active learning strategies, have further enhanced cross-domain adaptability. Moreover, the adoption of transfer learning expands the scope and applicability of the frameworks, enabling broad, fast, and efficient screening of materials.

To fully realize these advances, cross-disciplinary collaboration between materials scientists, CV experts, and industry is essential. As these models become increasingly refined and integrated into the materials design pipeline, they will allow for sustainable and efficient development of high-performance materials that meet the growing demands of modern technology, which is crucial for creating next-generation materials and driving innovation across multiple industries.

Acknowledgements

This work was financially supported by the National Science Fund for Distinguished Young Scholars, China (No. 52025041), the National Natural Science Foundation of China (Nos. 52450003, U2341267, and 52174294), the National Postdoctoral Program for Innovative Talents, China (No. BX20240437), and the Fundamental Research Funds for the Central Universities, China (Nos. FRF-IDRY-23-037 and FRF-TP-20-02C2).

Conflict of Interest

Xinmei Hou is an editorial board member for this journal and was not involved in the editorial review or the decision to publish this article. The authors declare that there are no competing interests

References

[1] W.S. Xia, X.B. Zhao, J.W. Wang, *et al*., New strategy to improve the overall performance of single-crystal superalloys by designing a bimodal γ′ precipitation microstructure, *Acta Mater*., 257(2023), art. No. 119200.

[2] Z.X. Zhang, Y.Q. Guan, Q. Huang, *et al*., Insight of the microstructure evolution and performance enhancement of spinodal decomposition in (Ti, Zr)C composite carbide ceramics: Multiscale simulation and experimental investigation, *Acta Mater*., 282(2025), art. No. 120487.

[3] X.Y. Zhou, H.H. Wu, Y. Wu, *et al*., Formation and strengthening mechanism of ordered interstitial complexes in multi-principle element alloys, *Acta Mater*., 281(2024), art. No. 120364.

[4] C.Y. Guo, E.H. Wang, Y.S. Liu, Y.P. Zheng, T. Yang, and X.M. Hou, Effect of Sn doping concentration on the oxidation of Al-containing MAX phase (Ti_3AlC_2) combining simulation with experiment, *Fundam. Res*., 2(2022), No. 1, p. 114.

[5] Y.B. Huang, N. Xu, H.L. Lu, Y. Ren, S.L. Li, and Y.D. Wang, Microstructures and micromechanical behaviors of high-entropy alloys investigated by synchrotron X-ray and neutron diffraction techniques: A review, *Int. J. Miner. Metall. Mater*., 31(2024), No. 6, p. 1333.

[6] J. Wei, X. Chu, X.Y. Sun, *et al*., Machine learning in materials science, *InfoMat*, 1(2019), No. 3, p. 338.

[7] C.Y. Guo, X.J. Duan, Z. Fang, *et al*., A new strategy for long-term complex oxidation of MAX phases: Database generation and oxidation kinetic model establishment with aid of machine learning, *Acta Mater*., 241(2022), art. No. 118378.

[8] D. Ouyang, B. He, A. Ghorbani, *et al*., Video-based AI for beat-to-beat assessment of cardiac function, *Nature*, 580(2020), No. 7802, p. 252.

[9] S. Ravuri, K. Lenc, M. Willson, *et al*., Skilful precipitation nowcasting using deep generative models of radar, *Nature*, 597(2021), No. 7878, p. 672.

[10] A. Agrawal and A. Choudhary, Perspective: Materials informatics and big data: Realization of the “fourth paradigm” of science in materials science, *APL Mater*., 4(2016), No. 5, art. No. 053208.

[11] G.R. Schleder, A.C.M. Padilha, C.M. Acosta, M. Costa, and A. Fazzio, From DFT to machine learning: Recent approaches to materials science–A review, *J. Phys. Mater*., 2(2019), No. 3, art. No. 032001.

[12] Y.Z. Fu, A.R.J. Downey, L. Yuan, T.Y. Zhang, A. Pratt, and Y. Balogun, Machine learning algorithms for defect detection in metal laser-based additive manufacturing: A review, *J. Manuf. Process*., 75(2022), p. 693.

[13] Z.H. Tian, C.L. Shang, C.L. Zhang, *et al*., Review of precipitation strengthening in ultrahigh-strength martensitic steel, *Int. J. Miner. Metall. Mater*., 32(2025), No. 2, p. 256.

[14] K. Choudhary, B. DeCost, C. Chen, *et al*., Recent advances

and applications of deep learning methods in materials science, *npj Comput. Mater.*, 8(2022), art. No. 59.
[15] T. Xie and J.C. Grossman, Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties, *Phys. Rev. Lett.*, 120(2018), No. 14, art. No. 145301.
[16] X.Y. Shi, L.M. Zhou, Y.H. Huang, Y.J. Wu, and Z.J. Hong, A review on the applications of graph neural networks in materials science at the atomic scale, *Mater. Genome Eng. Adv.*, 2(2024), No. 2, art. No. e50.
[17] R.A. Kirsch, SEAC and the start of image processing at the national bureau of standards, *IEEE Ann. Hist. Comput.*, 20(1998), No. 2, p. 7.
[18] L.G. Roberts, *Machine Perception of Three Dimensional Solids* [Dissertation], Massachusetts Institute of Technology, Cambridge, 1963.
[19] S.A. Papert, *The Summer Vision Project*, MIT Artificial Intelligence Laboratory, Cambridge, MA, 1966 [2025-01-07]. https://dspace.mit.edu/handle/1721.1/6125
[20] M. Minsky and S. Papert, An introduction to computational geometry, *Cambridge tiass.*, *HIT*, 479(1969), No. 480, p. 104.
[21] R. Szeliski, *Computer Vision*: *Algorithms and Applications*, Springer, London, 2011.
[22] D. Marr, *Vision*: *A Computational Investigation into the Human Representation and Processing of Visual Information*, The MIT Press, Cambridge, MA, 2010.
[23] K. Fukushima, Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position, *Biol. Cybern.*, 36(1980), No. 4, p. 193.
[24] Y. LeCun, B. Boser, J.S. Denker, *et al*., Backpropagation applied to handwritten zip code recognition, *Neural Comput.*, 1(1989), No. 4, p. 541.
[25] D.E. Rumelhart, G.E. Hinton, and R.J. Williams, Learning internal representations by error propagation, [In] D.E. Rumelhart and J.L. McClelland, eds., *Parallel Distributed Processing*: *Explorations in the Microstructure of Cognition*, *vol. 1*: *Foundations*, MIT Press, Cambridge, 1986, p. 318.
[26] D.G. Lowe, Distinctive image features from scale-invariant keypoints, *Int. J. Comput. Vis.*, 60(2004), No. 2, p. 91.
[27] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, Speeded-up robust features (SURF), *Comput. Vis. Image Underst.*, 110(2008), No. 3, p. 346.
[28] J. Canny, A computational approach to edge detection, *IEEE Trans. Pattern Anal. Mach. Intell.*, PAMI-8(1986), No. 6, p. 679.
[29] N. Dalal and B. Triggs, Histograms of oriented gradients for human detection, [in] *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (*CVPR'05*), San Diego, 2005.
[30] T.F. Cootes, C.J. Taylor, D.H. Cooper, and J. Graham, Active shape models-their training and application, *Comput. Vis. Image Underst.*, 61(1995), No. 1, p. 38.
[31] B.K.P. Horn and B.G. Schunck, Determining optical flow, *Artif. Intell.*, 17(1981), No. 1-3, p. 185.
[32] P. Viola and M. Jones, Rapid object detection using a boosted cascade of simple features [in] *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, Kauai, HI, 2001.
[33] R. Lienhart and J. Maydt, An extended set of haar-like features for rapid object detection, [in] *Proceedings of International Conference on Image Processing*, Rochester, 2002.
[34] A. Krizhevsky, I. Sutskever, and G.E. Hinton, ImageNet classification with deep convolutional neural networks, *Commun. ACM*, 60(2017), No. 6, p. 84.
[35] O. Russakovsky, J. Deng, H. Su, *et al*., ImageNet large scale visual recognition challenge, *Int. J. Comput. Vis.*, 115(2015), No. 3, p. 211.
[36] Y. LeCun, Y. Bengio, and G. Hinton, Deep learning, *Nature*, 521(2015), No. 7553, p. 436.
[37] R. Girshick, J. Donahue, T. Darrell, and J. Malik, Rich feature hierarchies for accurate object detection and semantic segmentation, [in] *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, 2014.
[38] J. Long, E. Shelhamer, and T. Darrell, Fully convolutional networks for semantic segmentation, [in] *2015 IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*). Boston, 2015, p. 3431.
[39] S.N. Xie and Z.W. Tu, Holistically-nested edge detection, [in] *2015 IEEE International Conference on Computer Vision* (*ICCV*), Santiago, 2015, p. 1395.
[40] O. Ronneberger, P. Fischer, and T. Brox, U-net: Convolutional networks for biomedical image segmentation, [in] *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015*: *18th International Conference*, Munich, 2015, p. 18.
[41] K.M. He, X.Y. Zhang, S.Q. Ren, and J. Sun, Deep residual learning for image recognition, [in] *2016 IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*), Las Vegas, 2016, p. 770.
[42] I. Goodfellow, J.P. Abadie, M. Mirza, *et al*., Generative adversarial nets, [in] *Neural Information Processing Systems*, Montreal, 2014, p. 2672.
[43] K.M. He, G. Gkioxari, P. Dollár, and R. Girshick, Mask R-CNN, [in] *2017 IEEE International Conference on Computer Vision* (*ICCV*), Venice, 2017, p. 2980.
[44] A. Dosovitskiy, An image is worth 16x16 words: Transformers for image recognition at scale, [in] *International Conference on Learning Representations*, Vienna, 2021.
[45] C.G. Shen, C.C. Wang, M.H. Huang, N. Xu, S.V.D. Zwaag, and W. Xu, A generic high-throughput microstructure classification and quantification method for regular SEM images of complex steel microstructures combining EBSD labeling and deep learning, *J. Mater. Sci. Technol.*, 93(2021), p. 191.
[46] A. Leitherer, B.C. Yeo, C.H. Liebscher, and L.M. Ghiringhelli, Automatic identification of crystal structures and interfaces via artificial-intelligence-based electron microscopy, *npj Comput. Mater.*, 9(2023), art. No. 179.
[47] R. Kondo, S. Yamakawa, Y. Masuoka, S. Tajima, and R. Asahi, Microstructure recognition using convolutional neural networks for prediction of ionic conductivity in ceramics, *Acta Mater.*, 141(2017), p. 29.
[48] Z.R. Pei, K.A. Rozman, Ö.N. Doğan, *et al*., Machine-learning microstructure for inverse material design, *Adv. Sci.*, 8(2021), No. 23, art. No. 2101207.
[49] G. Du, L.M. Zhou, Y.H. Huang, Y.J. Wu, and Z.J. Hong, Topological data analysis assisted machine learning for polar topological structures in oxide superlattices, *Acta Mater.*, 282(2025), art. No. 120467.
[50] S.H. Xu, A.S. McLeod, X.Z. Chen, *et al*., Deep learning analysis of polaritonic wave images, *ACS Nano*, 15(2021), No. 11, p. 18182.
[51] S. Ferdousi, R. Advincula, A.P. Sokolov, W. Choi, and Y.J. Jiang, Investigation of 3D printed lightweight hybrid composites via theoretical modeling and machine learning, *Composites Part B*, 265(2023), art. No. 110958.
[52] K.J. DeMille, R. Hall, J.R. Leigh, I. Guven, and A.D. Spear, Materials design using genetic algorithms informed by convolutional neural networks: Application to carbon nanotube bundles, *Composites Part B*, 286(2024), art. No. 111751.
[53] D. Ren, C.C. Wang, X.L. Wei, Q.Q. Lai, and W. Xu, Building a quantitative composition-microstructure-property relationship of dual-phase steels via multimodal data mining, *Acta Mater.*, 252(2023), art. No. 118954.
[54] J.L. Wu, X.L. Yin, and H. Xiao, Seeing permeability from im-

ages: Fast prediction with convolutional neural networks, *Sci. Bull.*, 63(2018), No. 18, p. 1215.

[55] X.Q. Jia, W. Li, Q. Lu, K. Zhang, H. Du, Y.T. Xu, and X.J. Jin, Towards accurate prediction for ultra-low carbon tempered martensite property through the cross-correlated substructures, *Mater. Des.*, 211(2021), art. No. 110126.

[56] G. Külekçi, K. Hacıefendioğlu, and H.B. Başağa, Enhancing mineral processing with deep learning: Automated quartz identification using thin section images, *Int. J. Miner. Metall. Mater.*, 32(2025), No. 4, p. 802.

[57] Z.S. Jiang, J.Z. Li, Y. Yang, *et al.*, Machine-learning-revealed statistics of the particle-carbon/binder detachment in lithium-ion battery cathodes, *Nat. Commun.*, 11(2020), No. 1, art. No. 2310.

[58] J. Na, S.J. Kim, H. Kim, S.H. Kang, and S. Lee, A unified microstructure segmentation approach via human-in-the-loop machine learning, *Acta Mater.*, 255(2023), art. No. 119086.

[59] X.C. Dong, H.W. Li, Z.T. Jiang, *et al.*, 3D deep learning enables accurate layer mapping of 2D materials, *ACS Nano*, 15(2021), No. 2, p. 3139.

[60] R.R. Chan, J. Pietryga, K.M. Landy, K.J. Gibson, and C.A. Mirkin, Microcrystal growth pathways investigated with machine learning segmentation and classification in scanning electron microscopy, *ACS Nano*, 18(2024), No. 48, p. 33073.

[61] C. Gorynski, M. Frei, F.E. Kruis, and M. Winterer, Machine learning based quantitative characterization of microstructures, *Acta Mater.*, 256(2023), art. No. 119106.

[62] ASTM International, *ASTM Standard E112–13*: *Standard Test Methods for Determining Average Grain Size*, ASTM International, West Conshohocken, 2013.

[63] N. Kulesh, A. Bolyachkin, I. Suzuki, Y.K. Takahashi, H.S. Amin, and K. Hono, Data-driven optimization of FePt heat-assisted magnetic recording media accelerated by deep learning TEM image segmentation, *Acta Mater.*, 255(2023), art. No. 119039.

[64] P. Liu, H.Y. Huang, X. Jiang, *et al.*, Evolution analysis of γ′ precipitate coarsening in co-based superalloys using kinetic theory and machine learning, *Acta Mater.*, 235(2022), art. No. 118101.

[65] I. Lobato, T. Friedrich, and S. Van Aert, Deep convolutional neural networks to restore single-shot electron microscopy images, *npj Comput. Mater.*, 10(2024), art. No. 10.

[66] Z.H. Cao, Q. Liu, Q.C. Liu, X.B. Yu, J.J. Kruzic, and X.P. Li, A machine learning method to quantitatively predict alpha phase morphology in additively manufactured Ti–6Al–4V, *npj Comput. Mater.*, 9(2023), art. No. 195.

[67] E. Azqadan, H. Jahed, and A. Arami, Predictive microstructure image generation using denoising diffusion probabilistic models, *Acta Mater.*, 261(2023), art. No. 119406.

[68] S.W. Im, D. Zhang, J.H. Han, *et al.*, Investigating chiral morphogenesis of gold using generative cellular automata, *Nat. Mater.*, 23(2024), No. 7, p. 977.

[69] A. Sciazko, Y. Komatsu, T. Shimura, and N. Shikazono, Prediction of electrode microstructure evolutions with physically constrained unsupervised image-to-image translation networks, *npj Comput. Mater.*, 10(2024), art. No. 49.

[70] S.Y. Zhu, D. Sarıtürk, and R. Arróyave, Accelerating CALPHAD-based phase diagram predictions in complex alloys using universal machine learning potentials: Opportunities and challenges, *Acta Mater.*, 286(2025), art. No. 120747.

[71] Z.J. Yang, Y.C. Yabansu, D. Jha, *et al.*, Establishing structure-property localization linkages for elastic deformation of three-dimensional high contrast composites using deep learning approaches, *Acta Mater.*, 166(2019), p. 335.

[72] A. Gupta, A. Bhaduri, and L.G. Brady, Accelerated multiscale mechanics modeling in a deep learning framework, *Mech. Mater.*, 184(2023), art. No. 104709.

[73] Z.Z. Yang, C.H. Yu, and M.J. Buehler, Deep learning model to predict complex stress and strain fields in hierarchical composites, *Sci. Adv.*, 7(2021), No. 15, art. No. eabd7416.

[74] M.M. Rashid, T. Pittie, S. Chakraborty, and N.M.A. Krishnan, Learning the stress-strain fields in digital composites using Fourier neural operator, *iScience*, 25(2022), No. 11, art. No. 105452.

[75] Y.L. Wu, J.M. Wang, X.L. Guo, and L.F. Zhang, Conditional generative adversarial network-based predictive method for crack initiation in a dual-phase austenite stainless weld, *Corros. Sci.*, 240(2024), art. No. 112494.

[76] Y.F. Shen, R. Pokharel, T.J. Nizolek, A. Kumar, and T. Lookman, Convolutional neural network-based method for real-time orientation indexing of measured electron backscatter diffraction patterns, *Acta Mater.*, 170(2019), p. 118.

[77] J.X. Feng, Q.Z. Teng, X.H. He, and X.H. Wu, Accelerating multi-point statistics reconstruction method for porous media *via* deep learning, *Acta Mater.*, 159(2018), p. 296.

[78] J. Na, G. Kim, S.H. Kang, S.J. Kim, and S. Lee, Deep learning-based discriminative refocusing of scanning electron microscopy images for materials science, *Acta Mater.*, 214(2021), art. No. 116987.

[79] A. Khan, C.H. Lee, P.Y. Huang, and B.K. Clark, Leveraging generative adversarial networks to create realistic scanning transmission electron microscopy images, *npj Comput. Mater.*, 9(2023), art. No. 85.

[80] D.K. Jangid, N.R. Brodnik, M.G. Goebel, *et al.*, Adaptable physics-based super-resolution for electron backscatter diffraction maps, *npj Comput. Mater.*, 8(2022), art. No. 255.

[81] T. Niu, M. Nasim, R.G.S. Annadanam, *et al.*, Recent studies on void shrinkage in metallic materials subjected to in situ heavy ion irradiations, *JOM*, 72(2020), No. 11, p. 4008.

[82] A. Visheratina, A. Visheratin, P. Kumar, M. Veksler, and N.A. Kotov, Chirality analysis of complex microparticles using deep learning on realistic sets of microscopy images, *ACS Nano*, 17(2023), No. 8, p. 7431.

[83] Z.D. Liu, D.Y. Li, Q.Q. Zhu, C.X. Zhang, J.Y. Ma, and J.J. Zhao, Intelligent method to experimentally identify the fracture mechanism of red sandstone, *Int. J. Miner. Metall. Mater.*, 30(2023), No. 11, p. 2134.

[84] J. Madsen, P. Liu, J. Kling, *et al.*, A deep learning approach to identify local structures in atomic-resolution transmission electron microscopy images, *Adv. Theory Simul.*, 1(2018), No. 8, art. No. 1800037.

[85] M. Ziatdinov, O. Dyck, A. Maksov, *et al.*, Deep learning of atomically resolved scanning transmission electron microscopy images: Chemical identification and tracking local transformations, *ACS Nano*, 11(2017), No. 12, p. 12742.

[86] W. Li, K.G. Field, and D. Morgan, Automated defect analysis in electron microscopic images, *npj Comput. Mater.*, 4(2018), art. No. 36.

[87] O. Badmos, A. Kopp, T. Bernthaler, and G. Schneider, Image-based defect detection in lithium-ion battery electrode using convolutional neural networks, *J. Intell. Manuf.*, 31(2020), No. 4, p. 885.

[88] M.R. Shen, G.Z. Li, D.X. Wu, *et al.*, Multi defect detection and analysis of electron microscopy images with deep learning, *Comput. Mater. Sci.*, 199(2021), art. No. 110576.

[89] B. Dey, D. Goswami, S. Halder, K. Khalil, P. Leray, and M.A. Bayoumi, Deep learning-based defect classification and detection in SEM images, [in] *Metrology*, *Inspection*, *and Process Control XXXVI*, San Jose, 2022, p. 83.

[90] A. Maksov, O. Dyck, K. Wang, *et al.*, Deep learning analysis of defect and phase evolution during electron beam-induced transformations in WS_2, *npj Comput. Mater.*, 5(2019), art. No. 12.

[91] Z. Maxim, S. Jesse, B.G. Sumpter, S.V. Kalinin, and O. Dyck,

Tracking atomic structure evolution during directed electron beam induced Si-atom motion in graphene via deep machine learning, *Nanotechnology*, 32(2021), No. 3, art. No. 035703.

[92] M.R. Shen, G.Z. Li, D.X. Wu, *et al*., A deep learning based automatic defect analysis framework for in situ TEM ion irradiations, *Comput. Mater. Sci.*, 197(2021), art. No. 110560.

[93] R. Sainju, W.Y. Chen, S. Schaefer, *et al*., DefectTrack: A deep learning-based multi-object tracking algorithm for quantitative defect analysis of *in situ* TEM videos in real-time, *Sci. Rep.*, 12(2022), art. No. 15705.

[94] C.H. Lee, A. Khan, D. Luo, *et al*., Deep learning enabled strain mapping of single-atom defects in two-dimensional transition metal dichalcogenides with sub-picometer precision, *Nano Lett.*, 20(2020), No. 5, p. 3369.

[95] R. Jacobs, M.R. Shen, Y.H. Liu, *et al*., Performance and limitations of deep learning semantic segmentation of multiple defects in transmission electron micrographs, *Cell Rep. Phys. Sci.*, 3(2022), No. 5, art. No. 100876.

[96] W.Y. Chen, Z.G. Mei, L. Ward, *et al*., *In-situ* TEM investigation of void swelling in nickel under irradiation with analysis aided by computer vision, *Acta Mater.*, 254(2023), art. No. 119013.

[97] P. Komninos, A.E.C. Verraest, N. Eleftheroglou, and D. Zarouchas, Intelligent fatigue damage tracking and prognostics of composite structures utilizing raw images via interpretable deep learning, *Composites Part B*, 287(2024), art. No. 111863.

[98] T. Bilyk, A.M. Goryaeva, M.C. Marinica, *et al*., Accurate quantification of dislocation loops in complex functional alloys enabled by deep learning image analysis, *Sci. Rep.*, 14(2024), art. No. 25168.

[99] X.J. Duan, Z. Fang, T. Yang, *et al*., Maximizing the mechanical performance of Ti_3AlC_2-based MAX phases with aid of machine learning, *J. Adv. Ceram.*, 11(2022), No. 8, p. 1307.

[100] S.S. Omee, N.H. Fu, R.Z. Dong, M. Hu, and J.J. Hu, Structure-based out-of-distribution (OOD) materials property prediction: A benchmark study, *npj Comput. Mater.*, 10(2024), art. No. 144.

[101] C.W. Park and C. Wolverton, Developing an improved crystal graph convolutional neural network framework for accelerated materials discovery, *Phys. Rev. Mater.*, 4(2020), No. 6, art. No. 063801.

[102] M. Karamad, R. Magar, Y.T. Shi, S. Siahrostami, I.D. Gates, and A.B. Farimani, Orbital graph convolutional neural network for material property prediction, *Phys. Rev. Mater.*, 4(2020), No. 9, art. No. 093801.

[103] C. Chen, W.K. Ye, Y.X. Zuo, C. Zheng, and S.P. Ong, Graph networks as a universal machine learning framework for molecules and crystals, *Chem. Mater.*, 31(2019), No. 9, p. 3564.

[104] S.Y. Louis, Y. Zhao, A. Nasiri, *et al*., Graph convolutional neural networks with global attention for improved materials property prediction, *Phys. Chem. Chem. Phys.*, 22(2020), No. 32, p. 18141.

[105] S.S. Omee, S.Y. Louis, N.H. Fu, *et al*., Scalable deeper graph neural networks for high-performance materials property prediction, *Patterns*, 3(2022), No. 5, art. No. 100491.

[106] M.J. Wen, S.M. Blau, E.W.C.S. nSmith, S. Dwaraknath, and K.A. Persson, BonDNet: A graph neural network for the prediction of bond dissociation energies for charged molecules, *Chem. Sci.*, 12(2021), No. 5, p. 1858.

[107] K. Choudhary and B. DeCost, Atomistic line graph neural network for improved materials property predictions, *npj Comput. Mater.*, 7(2021), art. No. 185.

[108] H.D. Gao, X.W. Guo, G.L. Li, C. Li, and C.Q. Yang, GCPNet: An interpretable generic crystal pattern graph neural network for predicting material properties, *Neural Networks*, 188(2025), art. No. 107466.