



Deep learning modeling in microscopy imaging: A review of materials science applications

Marco Ragone ^a, Reza Shahabazian-Yassar ^a, Farzad Mashayek ^b, Vitaliy Yurkiv ^b

^a Department of Mechanical and Industrial Engineering, University of Illinois at Chicago, Chicago, IL 60607, USA

^b Department of Aerospace and Mechanical Engineering, University of Arizona, Tucson, AZ 85721, USA



ARTICLE INFO

Keywords:

Deep learning
Microscopy images
Synthetic image generation
Structure identification
3D structural reconstruction
Physical properties evaluation

ABSTRACT

The accurate analysis of microscopy images representing various materials obtained in scanning probe microscopy, scanning tunneling microscopy, and transmission electron microscopy, is in general time consuming as it requires the inspection of multiple data bases for the correct interpretation of the observed crystal structures. This task is especially demanding in microscopy video analysis involving a vast amount of image data. The recent development of deep learning (DL) algorithms has paved the way for cutting-edge microscopy studies in materials science, often outperforming conventional image analysis methods. This paper reviews the state-of-the-art in DL-based synthetic data generation, materials structure identification, three-dimensional structural reconstruction, and physical properties evaluation for different types of microscopy images. First, the fundamental concepts of DL relevant to materials science applications are reviewed. Subsequently, the combined experimental measurements and numerical simulations for preparing dedicated microscopy image for DL analysis are discussed. Then, the review concentrates on the core topic of the paper, that is the critical assessment of DL advances in materials' structural and physical properties evaluation. We believe that the future development and deployment of DL for practical microscopy data analysis will rely on the progress and improvement of advanced algorithms and innovative methods for training data generation.

1. Introduction

Revealing the structure–property relationship of materials is a fundamental objective in materials science. The knowledge of features such as atomic structure, morphology, phase, chemical compositions, and presence of defects allows one to control the material's physical properties for improving engineering applications of energy storage and conversion devices, catalysis, nanoelectronics, biomedicine, etc. Microscopy imaging is an established methodology of materials visualization at different length scales, allowing for an in-depth structure investigation and physical properties evaluation. In recent years, the analysis of microscopy images has been boosted by deep learning (DL), a class of machine learning (ML) algorithms including advanced techniques dedicated to image processing applications. This review focuses on the usage of DL techniques in specific areas of microscopy and materials science with the aim of guiding a researcher on the choice of a method for a targeted application.

Different microscopy techniques are appropriate depending on the specific application of interest. For example, optical microscopy is used to study materials in images at the micron-scale resolution [1]. Scanning probe microscopy (SPM) techniques, including atomic force microscopy (AFM), and scanning tunneling microscopy (STM), are instead considered for the inspection of atomic structure and electronic states of complex surfaces in the scale range from nanometers to millimeters [2]. In applications requiring a characterization at the atomic scale, transmission electron microscopy (TEM), and in particular high-resolution transmission electron microscopy

<https://doi.org/10.1016/j.pmatsci.2023.101165>

Received 25 May 2022; Received in revised form 7 February 2023; Accepted 19 July 2023

Available online 20 July 2023

0079-6425/© 2023 Elsevier Ltd. All rights reserved.

(HRTEM), allow observing features such as the positions of atoms and atomic columns, and also the presence and local transformation of defects [3,4]. Scanning transmission electron microscopy (STEM) is another high-resolution microscopy technique, typically adopted in the high-angle annular dark-field (HAADF) mode, to investigate the morphology of materials and their chemical composition from the resulting Z-contrast [5].

The analysis of microscopy images for predicting the material's structural properties involves non-trivial challenges in practical applications. The main effort is the acquisition of "high-quality" microscopy images to ease the interpretation of the signal contrast and encoding information of the material's structural properties. This is a complex task, because of practical challenges such as the sensitive calibration of the microscope parameters used for tuning the lenses aberration and adjusting the eventual presence of noise due to the sample's contamination. A further challenge is the utilization of the signal contrast to identify the material's structural properties of interest, a task performed experimentally or algorithmically. Then another cumbersome task is the estimation of the physical properties from the structural features extracted from microscopy images. Finally, with the continued development of *in situ* and *operando* microscopy techniques producing tera bytes of image data, it is paramount to integrate the microscopes with advanced tools to evaluate materials properties, allowing for fast and real-time analysis.

In recent years, computer vision algorithms built upon DL have been established as state-of-the-art techniques for image processing. Convolutional neural network (CNN), a class of DL models with superior capability of feature extraction from images, has been adopted in image recognition problems such as classification [6], object detection [7–10], and semantic segmentation [11,12]. The progress in these methods has grown the attraction for their application to microscopy images representing 2D and 3D materials, to solve problems such as the identification of atoms [13], atomic column heights estimation [14], localization of defects, and chemical composition evaluation [15]. The main advantage of learning algorithms is the ability to unpin in real time the correlations between the intensity of the image signal contrast generated by the microscope, and the structural and physical properties of materials, by exploring hidden patterns among data [16].

In materials science, data-driven algorithms have been considered not only for DL applications to microscopy images, but also for high-throughput screening of materials. In these studies, DL models applied to materials' metadata rather than microscopy images, and other general ML techniques such as support vector machine (SVM), random forest (RF), and k-nearest neighbor (kNN), have been applied to experimentally acquired and simulated datasets of materials. These datasets have been collected from publicly available resources, for example, the Materials Project Database [17], Automatic Flow for Materials Discovery (AFLOWLIB) [18], Polymer Genome [19], Citrination [20], and Materials Platform for Data Science (MPDS) [21]. For instance, data-driven algorithms have been successfully used in Li-ion battery (LIB) research for screening solid-state electrolyte candidates [22–24], for predicting the crystal structure of cathode materials [25], and for estimating the redox potential of molecular electrode materials [26,27]. Another application of interest is high entropy alloys (HEAs) exploration, where ML and DL models have shown exceptional performance in the prediction of the configuration energy [28], phase formation [29–31], crystal structure [32], and mechanical and thermal properties [33].

The research field of data-driven algorithms applied to materials science is also known as *materials informatics* [34–36], with various reviews published in the prior literature about the high-throughput screening of materials [37–40], as well as applications of DL to microscopy images representing materials [41–45]. Kalinin et al. summarized the progress of DL in STEM [41], while Azuri and co-workers discussed the role of CNNs in SPM [43]. Ge et al. prepared a more comprehensive review, describing the applications to different microscopy techniques such as scanning electron microscopy (SEM), STEM, and SPM [42]. Furthermore, Xu et al. [44] and Zheng et al. [45] focused on the applications of DL to microscopy images of Li-ion batteries materials and materials for energy storage technologies, respectively. Despite the useful literature's categorization on the class of microscopy images and targeted materials, the previously published reviews do not provide a clear distinction on the utilization of different DL methods for addressing specific materials science problems. However, describing how DL techniques can be employed for resolving a particular materials science challenge is of primary importance for microscopists interested in advancing their experimental findings using methods belonging to this innovative class of image processing algorithms.

With the continued advancements in *materials informatics*, a wide area of microscopy imaging problems has been explored using DL models in recent years and expanding these methods to new cutting-edge applications is attractive research on a structure–property relationship. For this reason, it is essential to track the development of DL in microscopy images with respect to the investigated materials science applications. Herein, we present a review of DL in microscopy imaging, classified on the main materials science problems targeted in the literature, i.e., synthetic image generation, structure identification, 3D reconstruction, and physical properties estimation. In this review, we discuss each topic with a characterization of their corresponding microscopy imaging studies and related challenges and most importantly, the description of how they have been addressed with DL. The main objective of this paper is not only to provide a more consistent approach to categorize the literature compared to the previous reviews, but most importantly, to point the reader to the most appropriate choice of DL method to be combined with microscopy analysis in the investigation of a particular materials science problem of interest. The review begins with a description of the fundamental concepts of DL algorithms as related to microscopy image analysis, then it discusses the technical aspects of different DL techniques from the viewpoint of the summarized material science applications, which is an important characterization not reported in previous reviews. In addition, we also provide a detailed explanation of essential steps for preparing appropriate datasets of microscopy images. The in-depth description of the preparation of microscopy images datasets is another contribution of our review, whose goal is to guide a researcher to understand how images should be appropriately collected and pre-processed for their successive application in DL workflows. After reviewing the literature, we conclude with our perspective on the future development and opportunities of DL modeling in microscopy imaging for innovative solutions in materials science applications.

2. Overview of deep learning techniques

Herein, the fundamental concepts of DL techniques suitable for microscopy applications are presented. Deep learning is defined by the pioneering work of LeCun *et al.* [46], as a class of ML algorithms with the multi-level representation of input data, transformed through the composition of non-linear layered modules with increasing levels of abstraction and complexity. At subsequent levels of representation, a DL model is capable of learning features of the input data, with the objective of inferring a prediction of interest. In general, the prediction of a DL model is a continuous value or a probability score in regression and classification problems, respectively. The desired prediction is called label or ground truth, and if it is available in the dataset, the learning problem is supervised, while if it is not provided, the learning problem is unsupervised.

The combination of layers is built in a complex structure called a neural network, a term describing the capability of DL algorithms to model the human brain's behavior. In a fully connected (FC) neural network, or artificial neural network (ANN), a layer represents a collection of neurons connected through a network of synapses responsible for the transferring of data information, modeled with algebraic operations $z_{n \times 1}^{(kk)} = \sigma(w_{m \times n}^T z_{m \times 1}^{(k)})$. In such an operation, $z_{m \times 1}^{(k)}$ and $z_{n \times 1}^{(kk)}$ are one-dimensional arrays representing the layers k and kk , with m and n neurons, respectively, while $w_{m \times n}$ is a matrix representing the network of synapses connecting the two layers, whose components are called weights. In Fig. 1, such an operation is performed in the ANN part of the neural network. The first, intermediate, and final layers of a neural network are called the input, hidden, and output or score layer, respectively, and σ is the activation function, modulating the neurons with non-linearities. The most widely used activation function for the hidden layers is the rectified linear unit (ReLU) [47]. On the other hand, the choice of activation functions for the output layer depends on the specific task to perform: in regression problems with continuous targets, ReLU or Linear activations are typically used, while Sigmoid or Softmax activations are adopted to predict probability scores in binary and multi-class classification schemes, respectively [48]. The ANNs are typically employed with numerical metadata, fed to the model as one-dimensional arrays of features. In materials science applications, a set of features could be the structural properties, with a target physical property or material's class as output. The term "deep learning" refers to the concept of building a deep neural network (DNN) with many stacked layers for learning complex and non-linear data patterns.

More advanced architectures are the CNNs for image, video, audio, and volumetric data processing [49–51], and the recurrent neural networks (RNNs) for time series analysis [52]. In the case of CNNs for images applications, the input data are image arrays with shape $W \times H \times N_c$, where W and H are the width and height of the images, and N_c is the image's number of channels. In RGB images, $N_c = 3$ corresponds to the red, green, and blue primary colors, while in grayscale representations such as microscopy images, $N_c = 1$

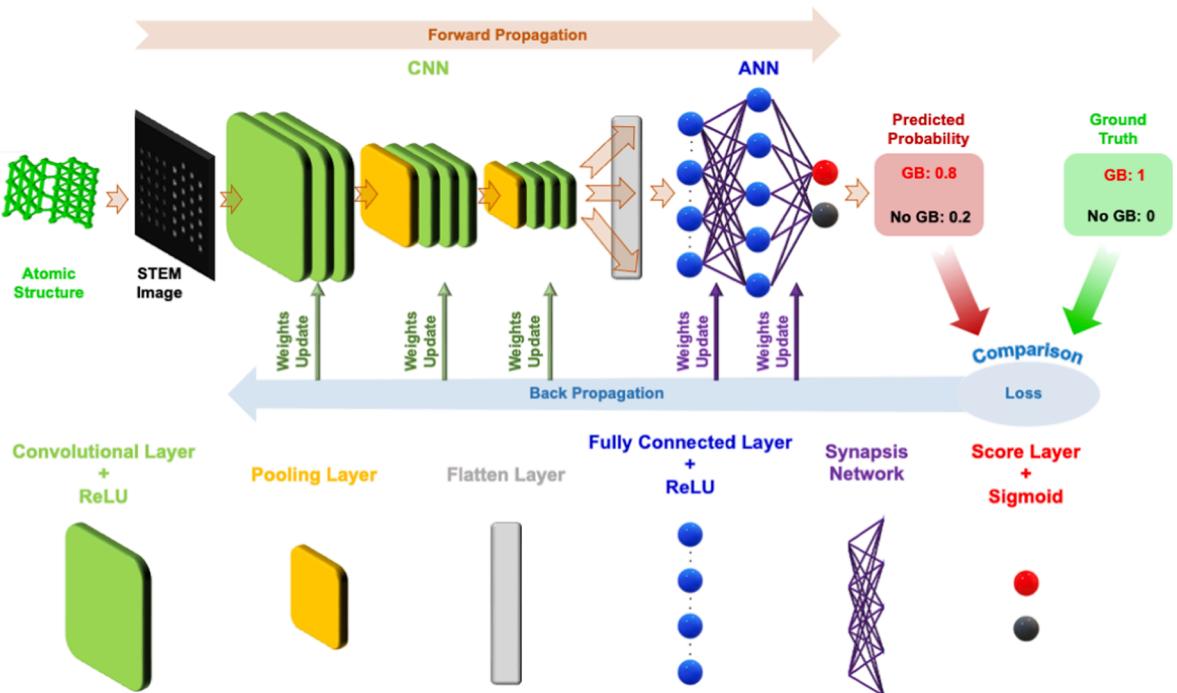


Fig. 1. Schematic of classification neural network comprising CNN for feature extraction and ANN for final inference. The input image representing an atomic structure with GB is forwardly propagated in the network, and probabilities of GB and no GB structural classes are predicted by the score layer. The GB class has the highest predicted probability of 0.8, meaning that the network can estimate the class correctly, as the STEM image represents a GB structure. The loss calculated between the predicted probability and the ground truth is back propagated in the network to update the weights of the convolutional and FC layers to further improve the predicted probability of GB class in the next training epoch.

corresponds to the gray channel. The relationship describing a CNN's layer can be formulated as $z_{W \times H \times N_c(kk)}^{(kk)} = f(z_{W \times H \times N_c(kk)}^{(k)})$, where f is a convolution correlation, defined by filters called kernels, sliding throughout the image while performing algebraic operations depending on the specific layer. There are four main types of layers in CNNs, i.e., convolutional layers, pooling layers, upsampling layers, and batch normalization layers. Convolutional layers are responsible for the extraction of relevant features of the images for inferring a prediction. The correlation f has the following mathematical expression: $z_{W \times H \times N_c(kk)}^{(kk)}(i, j) = \sigma(K_{m \times n} * z_{W \times H \times N_c(kk)}^{(k)}(i : i + m, j : j + n))$, with $0 \leq i \leq W$ and $0 \leq j \leq H$. The value of a pixel in position (i, j) in the output image $z_{W \times H \times N_c(kk)}^{(kk)}$, is equal to the inner product between the kernel $K_{m \times n}$ with width m and height n , and the portion of the image $z_{W \times H \times N_c(kk)}^{(k)}$ defined by the pixel's coordinates (i, j) , $(i + m, j)$, $(i + m, j + n)$, and $(i, j + n)$. The kernels $K_{m \times n}$ represent the CNN's layers parameters, and they have the analogous role of the matrix of weights $w_{m \times n}$ in ANNs. Within the convolution, the image size $W \times H$ is usually preserved between the input and output image, while N_c depends on the number of filters carried by the convolutional layers. The number of filters is progressively doubled while flowing deeply within the network, as the image size $W \times H$ is halved by the pooling layers. The objective of the pooling layers is to extract smaller portions of the image representing highly detailed objects [53]. On the other hand, upsampling layers reverse this operation by progressively restoring the original image size if required. Upsampling layers are used in fully convolutional networks (FCNs) for semantic segmentation tasks [12]. Finally, batch normalization layers are introduced in the network to address the distribution's variation of the layer's inputs, a problem known as covariate shift [54]. As the images are being processed through these layers, the CNN learns to extract the necessary features to infer a prediction. In Fig. 1, such operations are performed in the CNN part of the neural network. In typical applications, a CNN and an ANN are combined in a single network: the CNN is responsible for feature extraction, while the ANN is responsible for the final inference. The connection between a CNN and an ANN is performed with a Flatten layer reshaping 2D image arrays from the last convolutional layer of the CNN, into a 1D array used as input to the first FC layer of the ANN. In Fig. 1, such operation is performed between the ANN and CNN parts of the neural network, where the flatten layer is represented in gray.

Learning is performed through an iterative optimization process defined as training. At each training iteration (i.e., epoch), batches of data are forwardly propagated through the network, and a loss function calculates the divergence between the estimate and the ground truth. In regression problems, the mean squared error (MSE) loss function is typically used, while binary cross-entropy and categorical cross-entropy are considered for binary and multiclass classification, respectively [55]. The computed loss is back-propagated through the network to update the layer's parameters, with the objective of minimizing the loss in the next training epoch. The backpropagation is implemented with the gradient descent algorithm, involving the computation of the partial derivatives of the loss with respect to the layer's parameters in the update's correlation. The most efficient gradient descent algorithm is the stochastic gradient descent [56], where the updates are performed on batches of data randomly sampled from the dataset. Widely used stochastic gradient descent algorithms are Adam, RMSProp, and Momentum [56,57]. A representation of a neural network for the classification of an STEM image into grain boundary (GB) and no GB structural classes, together with forward and backpropagation workflow, is illustrated in Fig. 1.

A popular training technique is “transfer learning” with fine-tuning. Transfer learning involves applying a pre-trained DL model to a target dataset of interest [58]. VGG16 [59], Inception Net [60], ResNet [61], AlexNet [6], and DenseNet [62], are state-of-the-art DL classifiers, pre-trained on the ImageNet dataset. Similarly, RCNN [63], Fast-RCNN [64], Faster RCNN [65], Mask-RCNN [7], SSD [8], YOLO [9], and RetinaNet [10], are widely used object detection networks pre-trained on the COCO and PASCAL VOC datasets. The advantage of transfer learning is that a pre-trained model benefits in general of satisfactory performance, as it has been already trained on a large dataset comprising images representing a variety of object categories. However, since a pre-trained model must be applied on the dataset of interest, fine-tuning is typically considered for re-training the inference layers of the neural network on such a dataset, with fewer iterations required to achieve the desired performance compared to a “from scratch” training.

While a DL model learns to predict a target on a set of training data, it is paramount to verify its generalization capability of correct predictions on a diverse set of data. Such a task is called a test, and it is performed on test data not seen by the network during training. The model should be trained and tested on the same data distribution, for this reason, the training set is usually randomly sampled from 80% of the entire dataset, while the remaining 20% is used for the test set. Overfitting is the problem of an accurate prediction on the training set, with a less accurate performance on the test set. Overfitting could be prevented with methods such as early stopping, network reduction, expansion of the number of training data, and regularization [66]. On the contrary, underfitting is the issue of non-accurate training and validation performance, and it could be addressed by training for more epochs, increasing network complexity, or the amount of data. In general, the calibration of the modeling settings to prevent overfitting and underfitting is a necessary parametric study in ML and DL analysis called the bias-variance trade-off.

The performance of a DL model is evaluated with metrics dedicated to the problem of interest. In regression tasks, the R^2 score is generally used to estimate the similarity between the predictions and the continuous ground truth values. In materials science applications, the R^2 score could be used for tasks such as physical properties estimation or 3D reconstruction, where the model infers continuous values such as the Young's modulus or material's thickness. On the other hand, accuracy, precision, and recall are considered in classification problems. For a specific class, the accuracy is the proportion of the correctly predicted samples, the true positive TP (i.e., class samples correctly predicted in the class), and the true negative TN (i.e., non-class samples correctly predicted not in the class). Precision is instead the proportion of the TP with respect to the sum of the TP and the false positive FP (i.e., non-class samples incorrectly predicted in the class). Finally, recall is the fraction of the TP and the sum of the TP and the false negative FN (i.e., class samples incorrectly predicted not in the class). In materials science applications, accuracy, precision, and recall could be suitable metrics in the evaluation of the model's performance in structure classification problems, where phases, structural forms, type

of defects, rotational states, and crystal structures represent classes of interest. In object detection and semantic segmentation problems, where object bounding boxes and image masks are respectively predicted, the intersection over union (IoU) metric is adopted. The IoU is calculated as the proportion of the number of overlapping pixels between the predicted and the ground truth box or mask and the total number of pixels between the predicted and ground truth box or mask. IoU is analogous to accuracy, where the overlapping pixels represent the TPs, while the pixels in the predicted box or mask not belonging to the ground truth are the FPs, and the missed pixels in the ground truth box or masks are the FNs. In materials science applications, IoU could be used in structure detection and structure segmentation problems, where bounding boxes or masks are used to identify atoms, defects, nanoparticles, crystals, or phases.

DL models can be built with a variety of libraries, including TensorFlow [67], PyTorch [68], MXNet [69], and Caffe [70], usually implemented in the Python programming language. Due to the significant computational cost and high memory requirements, especially in image processing applications, GPUs are more suitable hardware compared to CPUs for running DL models. NVIDIA's GPUs are widely used because of their standard cuDNN library, designed to efficiently process DL frameworks in the CUDA environment [71]. For instance, Tesla V100 and Tesla A100 are efficient NVIDIA GPUs for accelerating DL training due to the Tensor core-based data center and high-performance computing (HPC) capabilities. GPU servers are usually accessed by creating dedicated Docker containers, typically in cloud computing platforms like EC2 Amazon Web Service (AWS) [72], and Google Cloud Platform (GCP) [73]. GPUs offer parallelization capability through dedicated implementations within the DL libraries [74]. The data parallelization technique consists in spreading batches of the data to different GPUs, on which training is performed simultaneously [75]. On the other hand, the model parallelization method implemented with the Horovod module [76], allows to train a neural network's components on different GPUs. Finally, GPUs with advanced computing capabilities can benefit from techniques like mixed precision [77], and accelerated linear algebra (XLA) [78], to speed up the algebraic operations within a neural network.

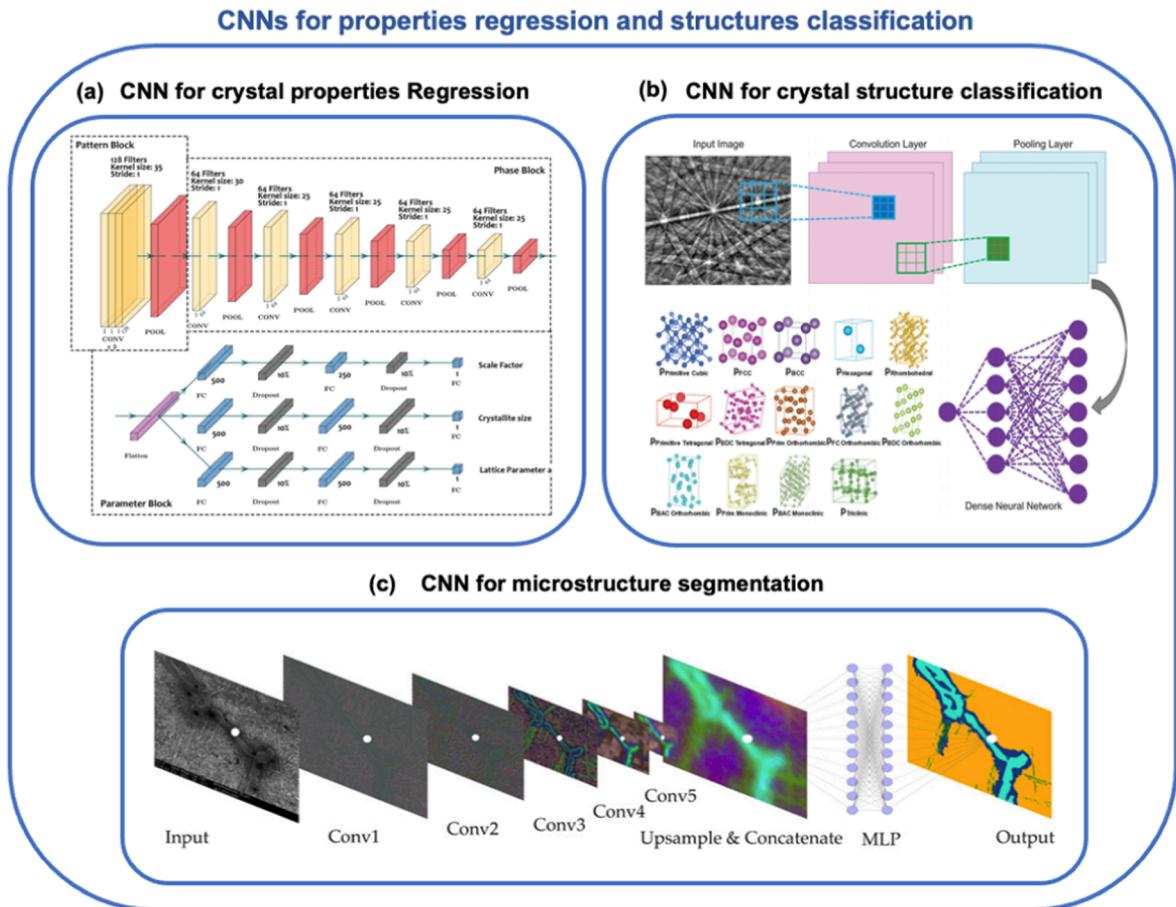


Fig. 2. (a) Schematics of CNNs architectures for crystal properties prediction (regression) proposed by Dong et al. [79], (b) crystal structure identification (classification) developed by Kaufmann et al. [80], and (c) carbon steel microstructure segmentation built by De Cost et al. [81]. Reproduced from References [80,79], and [81].

3. Deep learning techniques for microscopy imaging in materials science applications

The present section discusses different DL techniques adopted in the reviewed microscopy imaging and materials science applications. First, DL techniques related to different categories of neural networks are reported. Then, the section is concluded with a description of the transfer learning technique and its applications in this review. The reviewed neural network's architectures have been developed for more general tasks such as image classification, object detection, semantic segmentation, and generative modeling. In materials science applications, they have been adapted to address tasks as structure identification, 3D reconstruction, physical properties estimation, and synthetic image generation. It should be noted that each of these problems is not related to a specific category of neural network architecture. On the contrary, the same type of DL model could be suitable for different applications, from properties estimation to structure classification. Although a general architecture category can be flexible to a different scenario, each problem involves an original implementation and network design.

3.1. Convolutional neural networks for properties estimation and structures classification

A widely used architecture is the state-of-the-art CNN comprising convolutional layers for feature extraction in the model's initial stages, and FC layers for the final predictions, which could be continuous values for physical and structural properties estimation or probability scores for materials or microstructure classification. For instance, Dong and co-workers developed a regression CNN called Parameter Quantification Network (PQ-Net) for predicting crystal properties such as scale factor, crystallite size, and lattice parameter of multi-phase Ni-Pd/CeO₂-ZrO₂/Al₂O₃ catalytic materials detected with X-ray diffraction images [79]. On the other hand, Kaufmann et al. built a CNN applied to electron backscatter diffraction (EBSD) patterns for the classification of crystal structures [80]. A pixel-wise classifier built on the PixelNet architecture has been also proposed by De Cost et al. for segmenting carbon steel microstructures, where a phase class was predicted for each pixel in the images [81]. The schematics of the CNNs incorporating convolutional and FC layers are represented in Fig. 2.

3.2. Convolutional neural networks for structures detection

Another class of DL architectures adopted in materials science is the object detection neural network for performing tasks such as the detection of nanoparticles and crystals. In these types of DL models, the detection relies on the prediction of bounding boxes around the objects of interest, together with a probability score of the object's class. In architectures such as Mask-RCNN, object masks are also predicted in addition to the bounding boxes. For instance, Masubuchi et al. used a Mask-RCNN for the detection of 2D graphene, hBN, MoS₂, and WTe₂ crystals represented in optical microscopy images [82]. A representation of the Mask-RCNN architecture and the inferred detections from the input images is shown in Fig. 3.

3.3. Fully convolutional neural networks for structures segmentation

The FCN is instead designed for semantic segmentation tasks [11]. Segmentation refers to the FCN's pixel-wise predictions of objects class masks, which are generally more precise than bounding boxes methods. The pioneering FCN architecture is the U-Net, introduced by Ronneberger et al., for the segmentation of biomedical electron microscopy images, winning the International

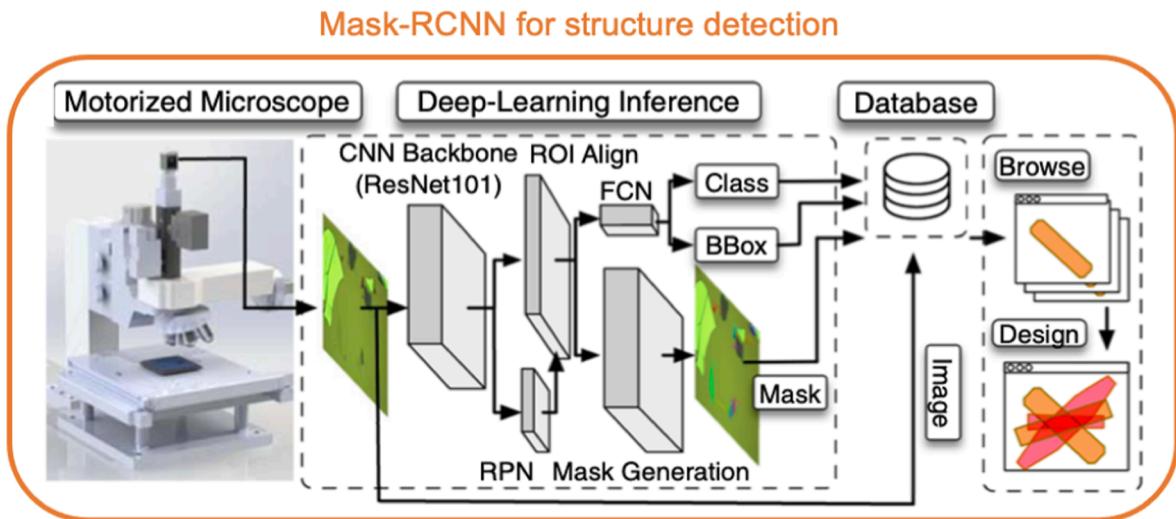
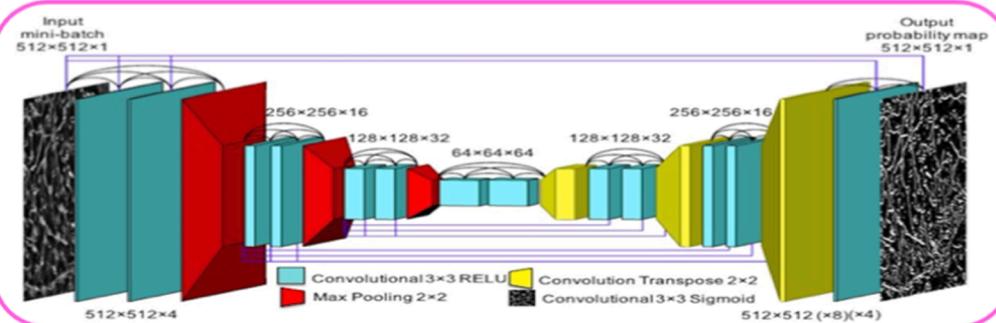


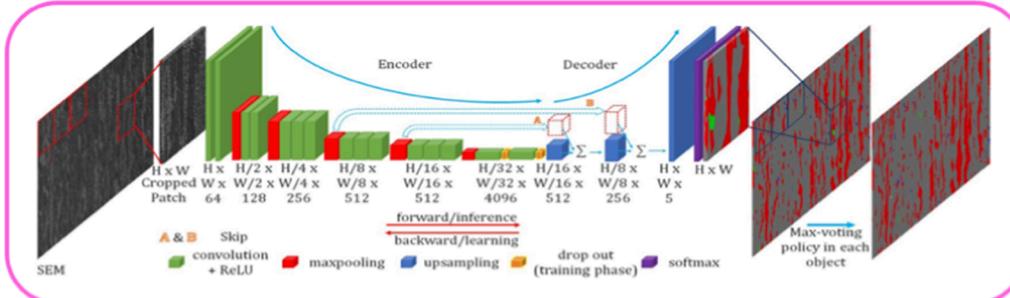
Fig. 3. Schematic of Mask-RCNN architecture adopted by Masubuchi et al. [82] for the detection of 2D graphene, hBN, WTe₂, and MoS₂ crystals. Reproduced from Reference [82].

FCNs for structures segmentation

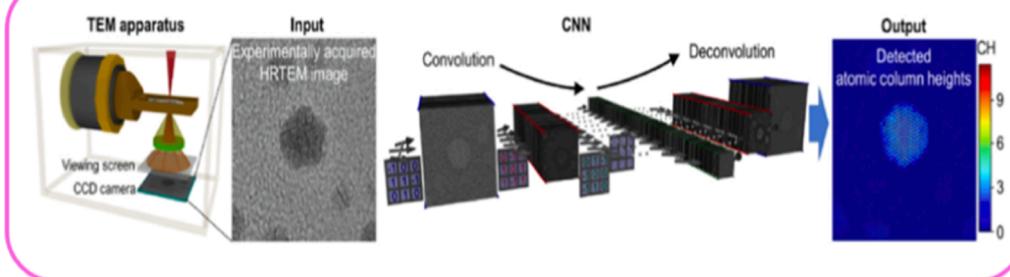
(a) FCN for defects segmentation



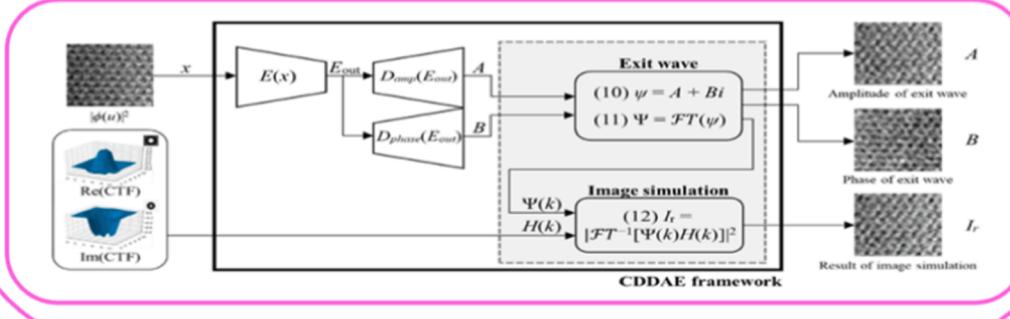
(b) FCN for microstructure segmentation



(c) FCN for atomic column heights estimation



(d) FCN for synthetic images generation

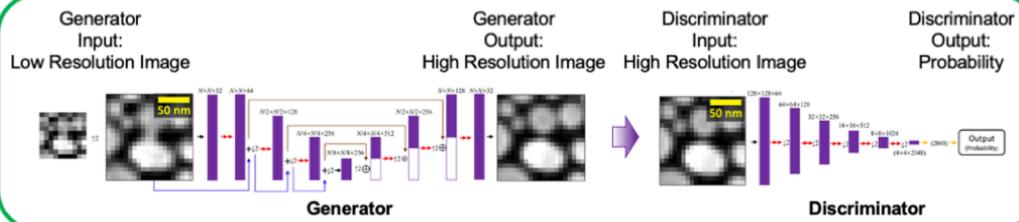


(caption on next page)

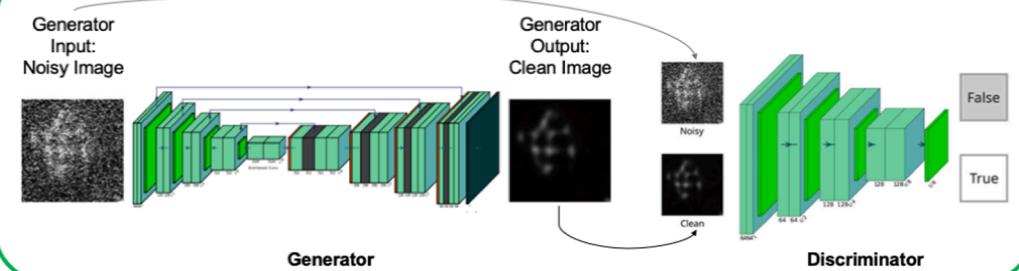
Fig. 4. Examples of FCNs for structure segmentation. (a) FCN adopted for segmentation of defects by Roberts et al. [84]: the input of the FCN is an DCI STEM image of structural steel alloy, and the output is a defects segmentation map. (b) Segmentation of phases by Azimi et al. [85]: the input of the FCN is an SEM image of low carbon steel matrix, and the output is a microstructure segmentation map. (c) Segmentation of atomic CHs by Ragone et al. [14]: the input of the FCN is an HRTEM image of gold nanoparticles, and the output is a segmentation map reporting the values of the atomic column heights. (d) Exit wave's CTF's amplitude and phase reconstruction by Lee et al. [86]: the input of the FCN is a TEM image of 2D graphene and the output is a segmentation map representing the amplitude and the phase of exit wave's CTF. Reproduced from References [84,85,14] and [86].

GANs for synthetic images and structures generation

(a) GANs for high resolution images generation



(b) GANs for denoised images generation



(c) GANs for synthetic structures generation

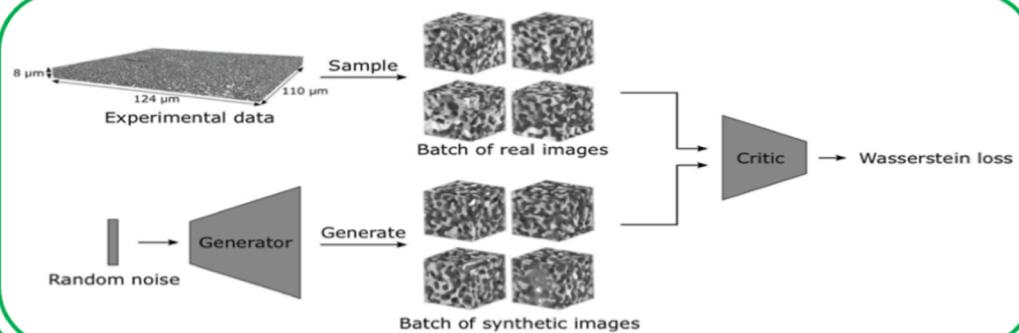


Fig. 5. Examples of GANs adopted for images and structure generations. (a) GAN adopted for enhancing the resolution of SEM images of gold nanoparticles on a carbon substrate [96]: the input of the generator is a low resolution SEM image, and the output is a corresponding high resolution image. (b) Denoising HAADF STEM image of Pt atomic structures [97]: the input of the generator is noisy STEM image, and the output is a corresponding noise-free image. (c) Generation of the 3D microstructure of solid oxide fuel cell electrodes [98]: the input of the generator is a set of random numbers sampled from a Gaussian distribution and the output is a random but realistic 3D synthetic structure. Reproduced from References [96,97] and [98].

Symposiums on Biomedical Imaging (ISBI) cell tracking challenge in 2015 [12]. The U-Net architecture is characterized by a symmetrical structure of convolutional layers, where the image size, halved by pooling layers in the encoder side of the network, is progressively restored in the decoder side by upsampling layers, and segmentation masks with the same size of the input image are predicted. The original U-Net architecture has received modifications through the years, for example, with the addition of residual blocks in the fully residual encoder-decoder network (FRED-Net) proposed by Arsalan et al. for the Iris dataset segmentation [83].

In materials science, the FCN architecture has been the most popular architecture in different applications, from structure segmentation to 3D reconstruction and synthetic image generation problems. For instance, Roberts et al. proposed an encoder-decoder model called DefectSegNet, for the segmentation of dislocation lines, precipitates, and voids crystallographic defects of structural steel alloys represented in diffraction contrast imaging (DCI) STEM images [84]. A model built on the state-of-the-art FCN architecture developed by Long et al. [11] was also applied to SEM images by Azimi et al., for the segmentation of microstructural phases in low-carbon steel matrix [85]. In the 3D reconstruction application, Ragone et al. adopted a regression FCN for the estimation of the atomic CHs in metallic nanoparticles [14]. In their application, the pixel-wise predictions are continuous values representing the atomic CHs, rather than probability scores representing objects class probabilities as in standard semantic segmentation methods. In the context of synthetic image generation, Lee et al. have proposed a convolutional dual-decoder autoencoder (CDDAE) for the reconstruction of the amplitude and phase of the exit wave's contrast transfer function (CTF) for the generation of denoised TEM images of 2D graphene [86]. The model's architectures employed in these applications are illustrated in Fig. 4.

3.4. Generative adversarial networks for images and structures generation

The generative adversarial network (GAN) is another class of neural network architectures adopted in materials science for synthetic image generation applications. GAN is a generative DL model introduced by Goodfellow et al., built as a combination of a generator neural network for producing artifact synthetic images, and a discriminator (i.e., critic) neural network whose task is to discriminate between real-life images (data distribution) and the synthetic images computed by the generator (model distribution), in an adversarial feedback loop [87]. Goodfellow et al. demonstrated that GAN could be used to generate new plausible images for the MNIST digit dataset, the CIFAR-10 small object photograph dataset, and the Toronto Face Database. The original GAN has been improved with the introduction of more advanced models such as Conditional GAN [88], Stack GAN [89], and Cycle GAN [90]. In particular, deep convolutional GAN (DC-GAN) is the class of GANs employing convolutional layers for image processing applications [91].

The GANs have been used for various applications, among whose super-resolution [92], image denoising [93], and 3D structures generation [94,95] have received the greatest attention in materials science research. For instance, Haan et al. proposed a GAN for enhancing the resolution of SEM images representing gold nanoparticles on a carbon substrate [96]. Similarly, Wang and co-workers

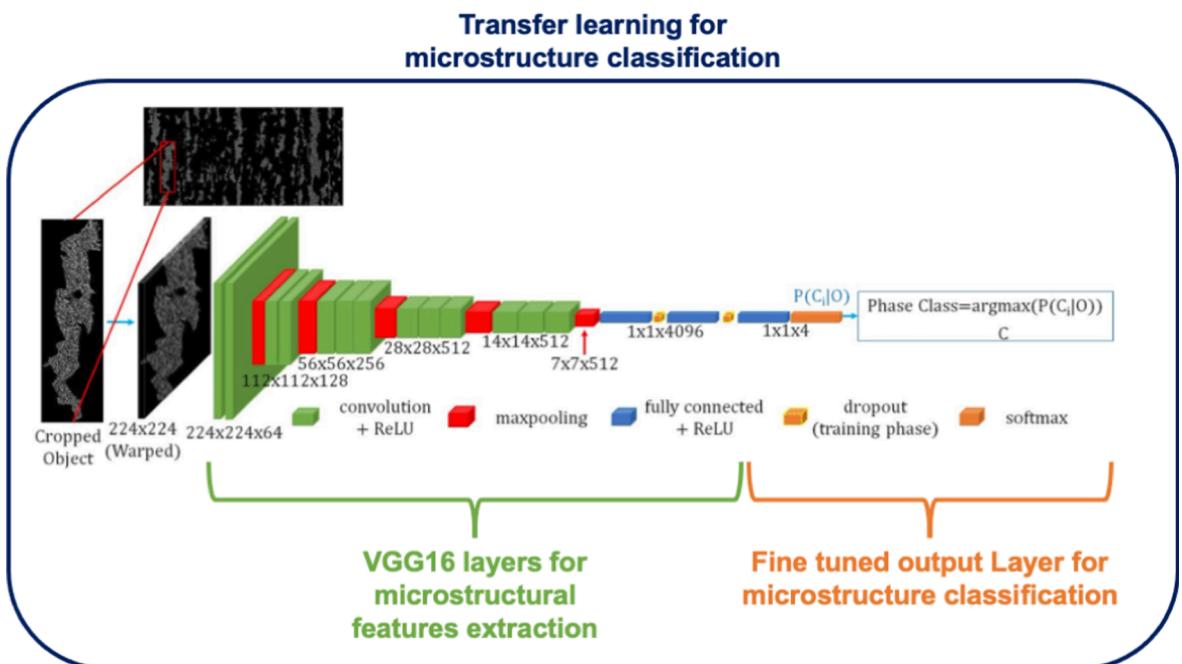


Fig. 6. Transfer learning approach proposed by Azimi et al. [85] for the classification of microstructures of advanced steel represented in SEM images. The VGG16 architecture has been used as base model for extracting the microstructural features in the SEM images, and a fine-tuned classification layer has been added for predicting the four class probabilities of martensite, tempered martensite, bainite and perlite phases. Reproduced from Reference [85].

developed a GAN for denoising HAADF STEM images of Pt atomic structures [97]. In both cases, the generator is an FCN built on the U-Net architecture [12], with the input a low-resolution or noisy image, and the output a synthetic high-resolution or noise-free image, respectively. Within the field of 3D structure generation, Hsu et al. developed a GAN for generating random but realistic 3D microstructures of solid oxide fuel cell electrodes [98], where the input of the generator is a set of random values sampled from a Gaussian distribution called a latent vector. The schematics of the GANs frameworks are reported in Fig. 5.

3.5. Transfer learning in microscopy imaging for materials science applications

This section presents some examples of transfer learning methods in materials microscopy imaging analysis. Transfer learning is a DL technique that is applicable to all the categories of neural networks described in the previous sections. Transfer learning has been demonstrated as a powerful method in various DL problems involving different types of architecture, and relevant applications have also been reported in materials science, where pre-trained CNNs, RCNNs, FCNs, and GANs have been used.

The most widely used pre-trained model is the VGG16, one of the most accurate architectures for classification tasks developed in the literature. Azimi et al. adopted transfer learning with fine tuning of VGG16 architecture for the classification of advanced steel's microstructures in the martensite, tempered martensite, bainite, and perlite phases represented in SEM images [85]. For this purpose, the last layer of VGG16 architecture comprising 1000 neurons for classifying 1000 categories in the ImageNet dataset, was replaced with a four-neuron layer for predicting class probabilities of the microstructure phases. Fig. 6 illustrates the fine-tuned VGG16 model for the prediction of the phase class probabilities proposed by Azimi et al. [85].

A pre-trained VGG16 has been also adopted by De Cost and co-workers for building the feature extraction part of their model for phase classification and segmentation of carbon steel microstructures [81,99]. Transfer learning with fine tuning using VGG16 architecture was also considered by Luo and co-workers for the classification of carbon nanotubes/nanofibers (CNTs/CNFs) [100]. In their proposed method, the last layer of VGG16 was replaced with five neurons layers for predicting structural classes such as a cluster of loosely packed fibers, single fibers, condensed matrix structures with embedded CNT/CNF, oversized mixed structures with CNT or CNFs on surfaces and non-CNT structures.

Zhang et al. also applied a pre-trained VGG16 for the estimation of the local thickness from convergent beam diffraction (CBED) patterns of SrTiO₃ samples represented in STEM images at atomic resolutions [101]. In their 3D reconstruction application, Zhang and co-workers proposed both classification CNN and regression CNN for thickness estimation. In the classification scheme, the authors built the output layer with 50 neurons to predict class thickness incrementing from 2 nm to 100 nm, while in the regression scheme, one neuron was considered to predict the continuous thickness value from 0.5 nm to 35 nm. Modarres et al. compared the performance of Inception-v3, Inception-v4 [60], and ResNet [61] pre-trained on the ImageNet dataset, for the classification of materials represented in SEM images into ten structural classes, i.e., 0D particles, 1D nanowires and fibers, 2D films and coated surfaces, and 3D patterned surfaces such as pillars [102]. Furthermore, pre-trained object detection networks based on RCNN were applied respectively by Okunev et al. [103,104] and Zhang et al. [105] for the detection of metallic nanoparticles in STM and TEM images. In general, in all the aforementioned methods and all the reports where transfer learning was applied, the desired performance was achieved by training the pre-trained model for a relatively small number of epochs on the microscopy dataset for the materials science application of interest.

4. Microscopy images dataset preparation

The preparation of an appropriate dataset of microscopy images is a fundamental step for training and testing a neural network in materials science applications. As microscopy images have unique features compared to natural scene images, the construction of a dedicated dataset is not straightforward, and it requires special preliminary considerations. In some applications, datasets of real-life experimental images can be used, while in other cases, microscopy image simulations are required. The following sections describe the acquisition of microscopy images datasets following these two approaches.

4.1. Experimental microscopy images datasets

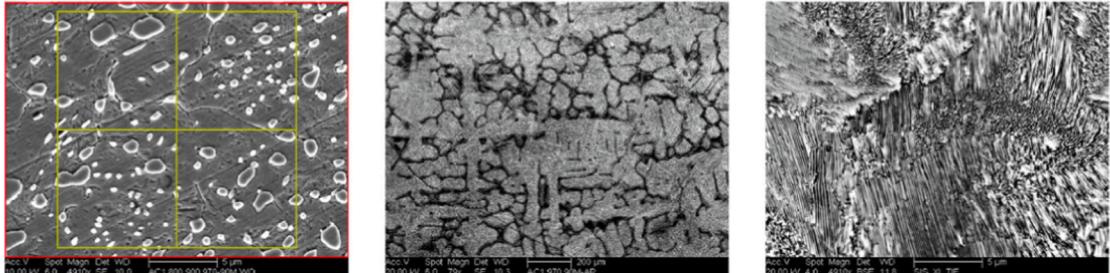
Training DL models require a vast number of images, usually in the thousands. The in-house acquisition of a sufficiently large database of microscopy images with desired conditions for training DL models is generally time-consuming and challenging in many applications. Typical issues are the cumbersome calibration of the microscope parameters to acquire noise-free images and the minimization of signal drift and distortion typically present in microscopy analysis [106]. Then, the quality of experimental microscopy images is also influenced by the sample's stability and contamination during the microscope's acquisition. In addition, it is important to ensure a sufficient pixel signal variation in a DL dataset. In microscopy images, the signal is dictated by the imaging and microscope conditions, as well as by the structural content and the physical properties of the represented materials. While the variation of the former could be relatively straightforward, the latter requires experimentalists to analyze an infeasible number of samples and structures. For these reasons, it is often prohibitive to collect a large database of microscopy images from in-house experiments. Exceptions could be video analysis, where large datasets could be built from sequences of image frames, but these cases are limited to specific DL time series applications. On the contrary, experimental microscopy images are more suited for on-the-fly inference stages to demonstrate the applicability of a trained neural network to real-life applications.

Nevertheless, some publicly available datasets of experimental microscopy images have been collected by different research groups, and they are suitable for training DL models. For instance, the Ultrahigh Carbon Steel DataBase (UHCSDB) [107] comprises 961 SEM images of Ultrahigh Carbon Steel (UHCS), acquired under varying heat treatments and magnifications, and labeled with the

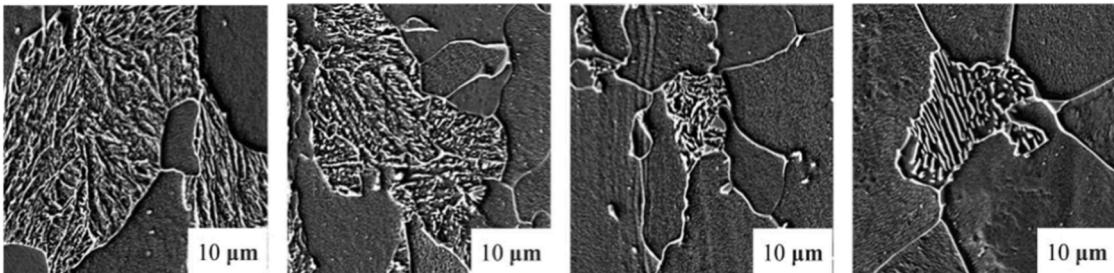
microconstituents of UHCS [108,109]. The UHCSDB was adopted by DeCost et al. for the classification and segmentation of UHCS phases such as spheroidite, proeutectoid cementite network, and perlite [81,99]. Examples of SEM images from the UHCSDB belonging to the three classes are illustrated in Fig. 7.

A further example of a microscopy images database is the steel image dataset developed by the Material Engineering Center Saarland (MECS) in Saarbrücken, Germany [110]. The dataset contains 21 light optical microscopy (LOM) and SEM image of steel structures in the martensitic, tempered martensitic, bainitic, and pearlitic microconstituents. As the average image size is 7000×8000 pixels, the dataset can be enlarged by cropping sub-images of smaller size. The MECS dataset was used by Azimi and co-workers for the classification and segmentation of such steel microconstituents [85]. Another publicly available database is the SEM images dataset from CNR Istituto di Officina dei Materiali (CNR-IOM) in Trieste. The dataset comprises more than 150,000 SEM images of material

(a) SEM images of carbon steel microstructures from UHCS database



(b) SEM images of carbon steel microstructures from MECS steel database



(c) SEM images of nano structures from CNR-IOM database

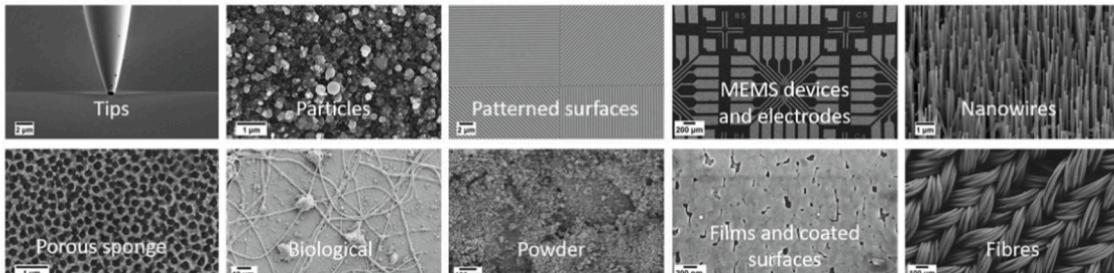


Fig. 7. Databases of experimentally acquired SEM images used for training DL models. (a) Examples of SEM images from the UHCS database representing the spheroidized cementite, proeutectoid cementite network, and perlite classes considered by De Cost et al. [99]. The yellow frames in the first image represent four cropped sub-images for expanding the dataset. (b) Example of SEM images from the MECS steel database used by Azimi and co-workers for the classification and segmentation of steel microconstituents [85]. (c) Examples of SEM images from the CNR-IOM databases adopted by Modarres et al. for structural forms classification [102]. Reproduced from References [99,85] and [102]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

structures such as particles, 1D nanowires and fibers, 2D films and coated etc., resulting from more than five years of experiments performed by many researchers [111]. The CNR-IOM dataset was adopted by Modarres et al. for structural forms classification [102].

In general, the pre-processing of datasets of real-life images involves a variety of challenges in many DL applications. A typical problem is the so-called class imbalance, meaning that certain classes count a larger number of images compared to other classes. This problem could be addressed by collecting more images until a balanced image distribution is obtained. However, the acquisition of more images of certain materials or structures could be time-consuming and infeasible, especially if materials categories are rare to observe. For instance, the UHCSDB has a strong class imbalance of the spheroidite and proeutectoid cementite network with 374 and 212 images, respectively, compared to the uncommon pearlite + widmanstätten and martensite/bainite with 27 and 36 images, respectively [99]. Data augmentation is an effective technique to address problems related to a shortage of images in a dataset [112]. Data augmentation is a method for generating new synthetic images from an initial distribution, using geometric operations such as random rotation, mirroring, cropping, and zooming and with the random variation of brightness, contrast, blur, etc. For instance, Badmos et al. adopted data augmentation to artificially increase the number of images representing Lithium-ion battery materials with defects to balance the more populated non-defective distribution [113]. Similarly, data augmentation was used by Azimi et al. to address the class imbalance of the MECS database, with 11 martensitic, 4 pearlitic, 4 bainitic, and 2 tempered martensitic samples, and to further increase the number of images in the dataset [85]. Generative DL models such as GANs are also useful methods for data augmentation. This review includes a description of the utilization of GANs as data augmentation technique in summarized microscopy imaging and materials science applications. In general, data augmentation is an essential step of any DL study, and most of the works reported in this review adopted data augmentation to enlarge the datasets of microscopy images.

Another non-trivial task is the images annotation for generating the necessary ground truth for supervised DL studies. As it has been discussed in sections 2 and 3, the images ground truth depends on the specific DL task. In the examples provided in Fig. 2, the ground truth is represented by crystal properties (regression), crystal structures (classification), and phases masks (segmentation), while in the application reported in Fig. 3 and Fig. 4, the ground truth consists in object masks with bounding box coordinates (object detection with segmentation) and object masks (segmentation), respectively. In the case of experimentally acquired microscopy image datasets, the annotation is usually performed by expert materials scientists, whose task is to manually assign the ground truth with high

Images of structural steel and phase masks

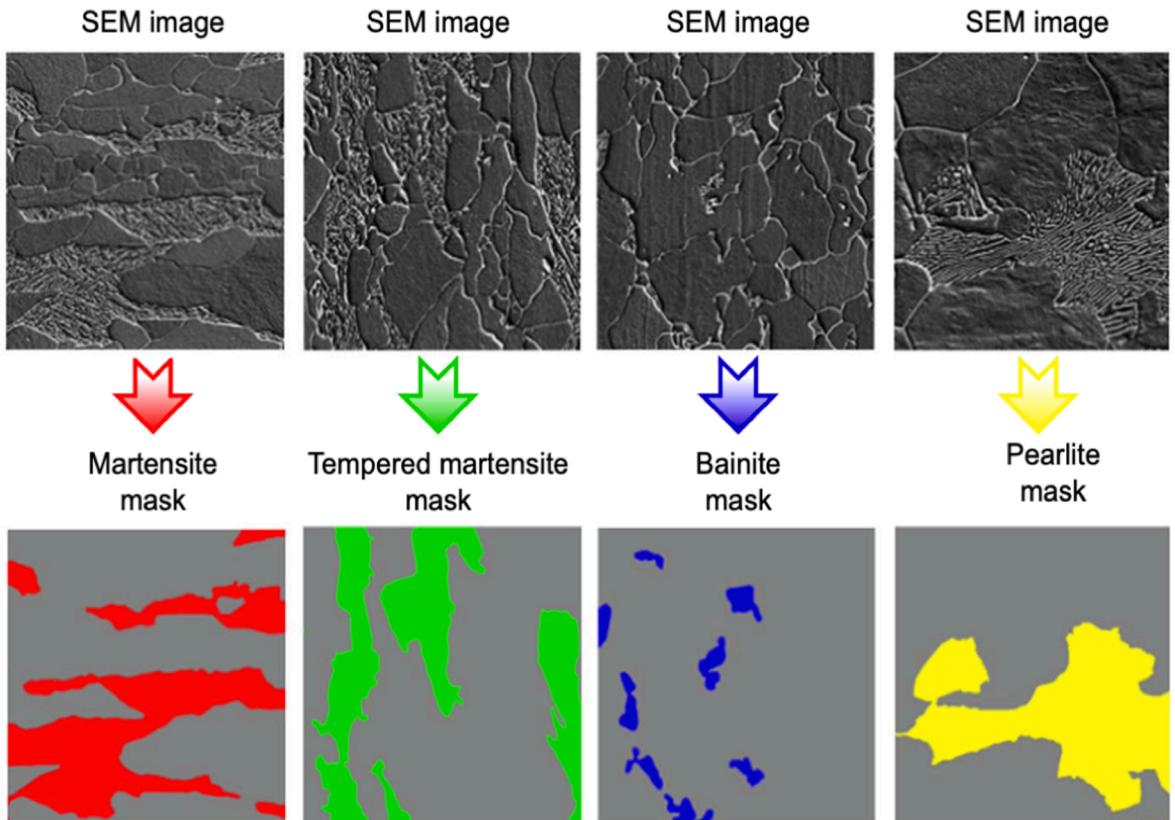


Fig. 8. Mask annotation of martensite, tempered martensite, bainite and pearlite phases for the segmentation-based DL model proposed by Azimi et al. [85], applied to SEM and LOM images of structural steel. The annotation has been manually performed by materials science experts and metallographers. Reproduced from Reference [85].

precision to not bias the DL model to learn incorrect labels. The correct ground truth assignment is a posed problem of data annotation in many DL analyses, and a variety of methods exist in the literature to search for “mislabeled” data originating from involuntary human mistakes in the labeling process [114,115]. The likelihood of wrong annotation is generally higher in images of difficult interpretation, as it could be the case of microscopy images representing complex materials. In addition, the manual annotations of materials ground truth could be a laborious and time-consuming process, and in some applications, it is not available directly from the dataset, and it must be performed in-house according to the target of interest. Fig. 8 shows the annotations of pixel-wise masks for segmenting martensite, tempered martensite, bainite, and pearlite phases of structural steel in SEM images adopted by Azimi et al., executed by materials science experts and metallographers [85].

Due to the complications involved in the collection and preparation of experimental microscopy images, it is often preferable, and sometimes mandatory, to perform images simulations to construct an appropriate dataset for DL analysis. In the following section, the process of microscopy images simulations and its advantages in DL modeling are discussed.

4.2. Simulated microscopy images datasets

Microscopy simulation is an efficient tool for the interpretation of images acquired in experimental measurements, where the extraction of quantitative sample information requires to reproduce the process of the signal contrast formation. Some of these applications are the investigation of atoms and defects arrangement [116,117], the evaluation of samples thickness and tilt [118,119], and the determination of the chemical composition of atomic columns [120]. Dedicated algorithms have been developed in the literature for the simulation of different categories of microscopy images [121,122]. For instance, the multislice algorithm, which involves modeling the traveling of an electron beam through the matter, is a wide-spread technique for the simulation of electron microscopy images [123].

With the advent of materials data mining, microscopy images simulations have received growing interest for building datasets for training DL models. Compared to experimental microscopy images, simulated images have several advantages. First, there is no bound on the dataset's size, assuming a reasonable computational cost of the images' simulations. This is a crucial aspect for DL, where datasets comprising thousands of images are required. Then, within the simulations, the microscope parameters can be controlled to obtain images with varying imaging conditions, being beneficial for the generalization capabilities of DL models. In addition, since the microscope imaging conditions leverage the performance of learning algorithms, as demonstrated by Zhong et al. [124], the precise knowledge of the microscope parameters allows to monitor the reliability of predictions under varying signal distributions. For instance, Madsen et al. demonstrated that the defocus and electron dose values in HRTEM images simulations are the most sensitive parameters influencing the accuracy of their DL model in the estimation of the atomic column heights of gold nanoparticles [13]. Another relevant advantage of image simulations is the algorithmic annotations of images' ground truth. The process of microscopy simulations is always associated with a prior model's construction of the materials represented in the simulated images. In materials models, features such as physical and structural properties, microstructure classes, and atoms and defects' locations can be algorithmically evaluated and then used to generate labels for training a neural network. A further benefit of this approach is the precise inspection of the correlation between the pixels' signal of simulated images and the corresponding ground truth. This is an important stage for DL analysis since it allows one to inspect the data patterns to be learned by the neural network. Finally, the generation of materials models with varying features such as size, zone axis, exposure, chemical composition etc., allows to construct a dataset with a

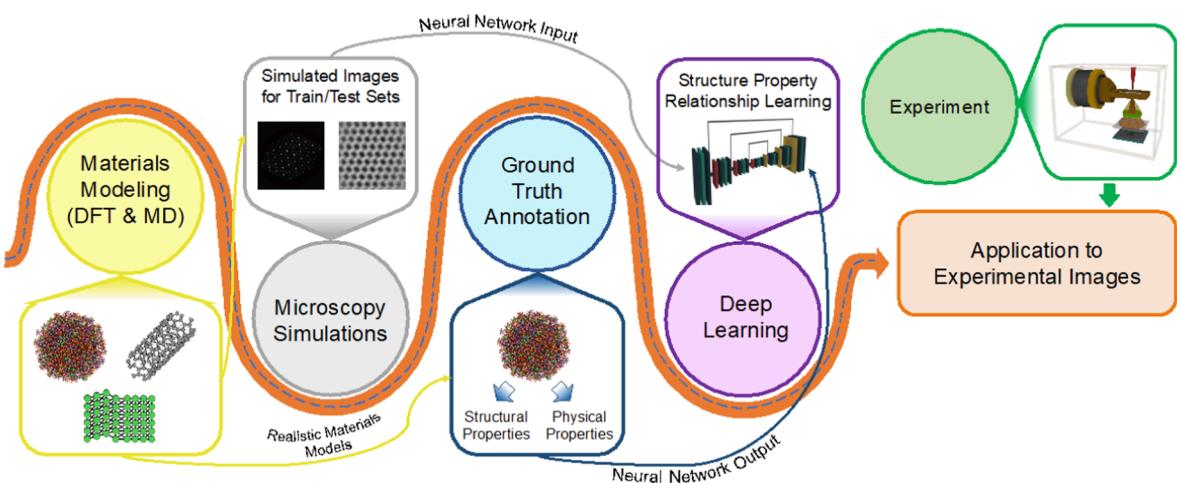


Fig. 9. Schematic of the workflow for the generation of simulated microscopy images dataset for training DL model to be successively applied to experimentally acquired microscopy images. DFT and MD are used to generate realistic materials models, adopted for the generation of training and test datasets with microscopy images simulations and ground truth annotation (i.e., structural properties and/or physical properties). The DL model trained on the simulated dataset to learn the structure–property relationship is then applied to predict microscopy images acquired in experimental measurements.

broad distribution of materials-induced signal in the simulated images.

Nevertheless, the simulation of materials models is a delicate step of algorithmic dataset generation. First, it is paramount to generate realistic materials models in terms of thermodynamic stability and structure optimization. Density functional theory (DFT) [125] and *ab initio* molecular dynamics (MD) [126] are widely used computational methods for the simulation and screening of realistic materials models. Then, it is imperative to build materials models with properties in agreement with the samples observed in experimental measurements for a successful application of a DL model to real-life microscopy images. This task is not straightforward to address in the cases of complicated interpretation of the properties of the materials analyzed in experiments. Fig. 9 shows the schematic of the workflow for the construction of a modeling dataset. Materials models could be computed with DFT and MD, and then adopted for generating the corresponding simulated microscopy images. Once a neural network is trained on simulated images to learn a desired target, then it could be applied to experimental microscopy images representing real-life materials.

The algorithmic generation of datasets, comprising materials modeling and microscopy images simulation, is a widely used approach in many DL applications in materials science. Herein, some examples of the simulated dataset's utilization are reported. Ziatdinov et al. trained an FCN model with theoretical STEM images of the atomic lattice of graphene and MoSe₂ doped with tungsten

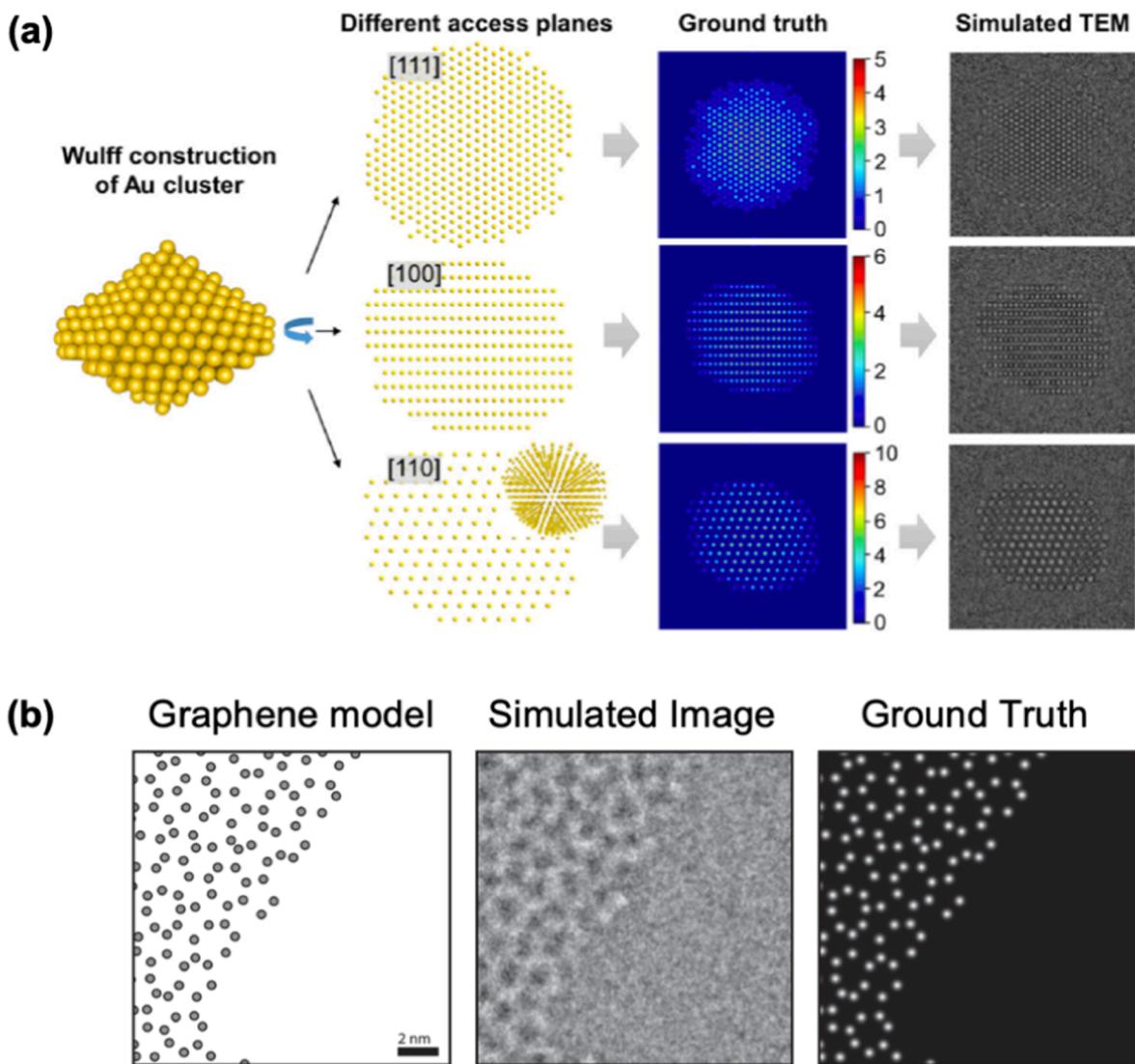


Fig. 10. Examples of modeling workflows for the generation of datasets of simulated images. (a) Ragone et al. [14] presented a method for the construction of a simulated dataset for training a DL model to predict atomic column heights of gold (Au) nanoparticles. The Wulff construction method has been adopted for generating realistic Au clusters. The modeling's nanoparticles have been exposed to different access planes to generate column heights with a varying arrangements using a semantic segmentation approach. (b) Madsen et al. [13] proposed a framework to build a simulated dataset for training a DL model to locate C atoms in graphene structures. Graphene structures computed using Voronoi tessellation have been used to obtain the corresponding simulated HRTEM images. In both cases, image simulations have been performed using the multislice algorithm. Reproduced from References [13,14].

($\text{Mo}_{1-x}\text{WSe}_2$) for the identification of defects such as vacancies and dopants [15]. Similarly, Lee and co-workers adopted annular dark field (ADF) STEM image simulations for generating a dataset for training an FCN to detect defects in a monolayer 2D transition metal dichalcogenide $\text{WSe}_{2-2x}\text{Te}_{2x}$ [127]. Madsen et al. [13] and Ragone et al. [14] developed atomic models of gold (Au) nanoparticles to simulate HRTEM images for building a dataset for atomic column height estimation. In their modeling approach, Ragone et al. adopted the Wulff construction method to generate realistic gold (Au) cluster, rotated to different zone axes, e.g. [111,100] and [110] to increase the variability of the structure-induced signal in the simulated TEM images. The column heights ground truth was generated algorithmically via a semantic segmentation approach [14]. In addition, Madsen et al. [13] designed a framework including modeling of graphene structures using Voronoi tessellation and HRTEM images simulation for training a DL model to locate the position of carbon atoms in a graphene lattice. The modeling workflows for the generation of a simulated dataset of TEM images representing gold nanoparticles and graphene structure proposed by Madsen et al. [13] and Ragone et al. [14] are represented in Fig. 10(a) and (b), respectively.

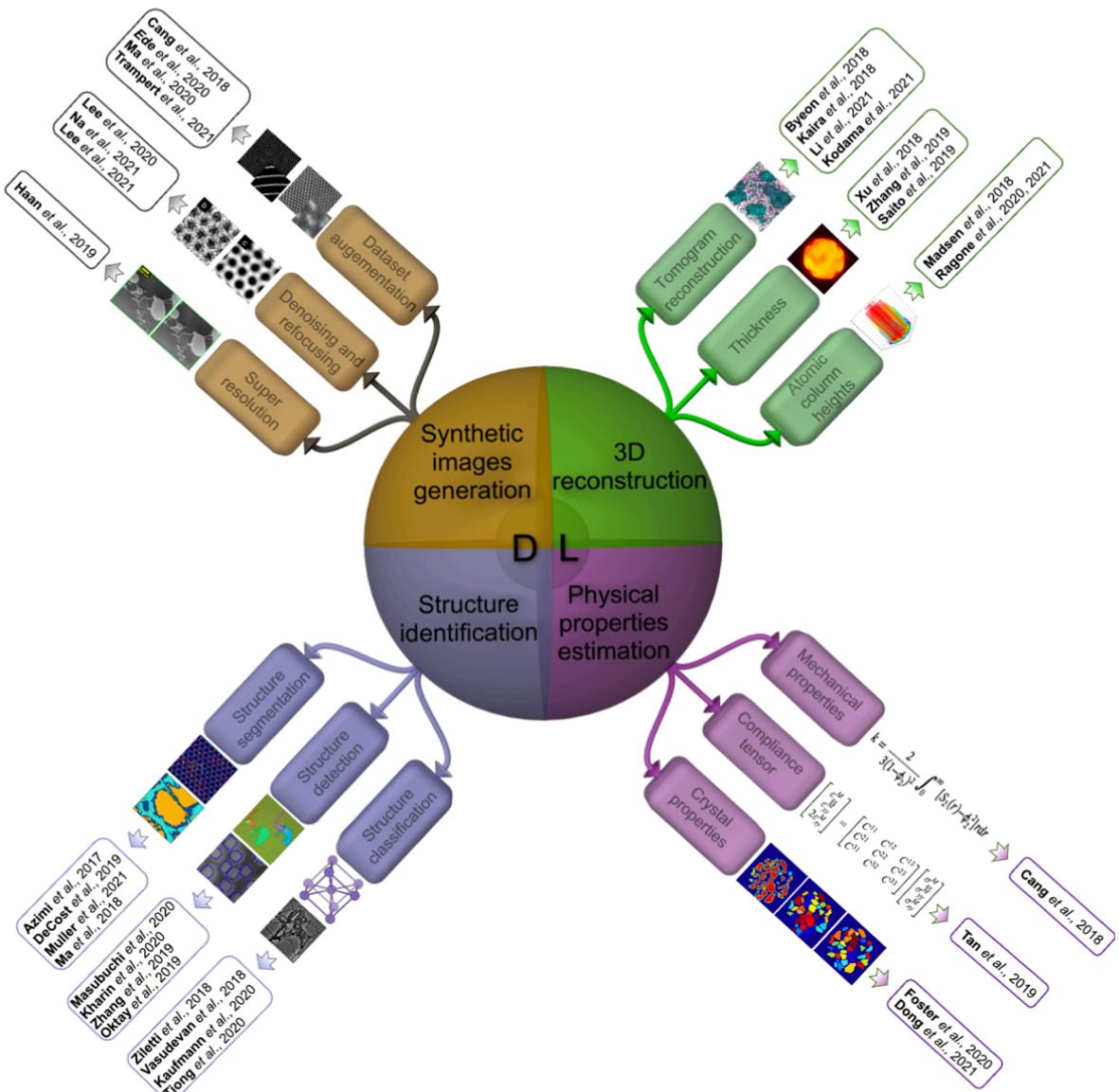


Fig. 11. Literature review of the most important DL applications in microscopy imaging problems in materials science research: synthetic data generation, structure identification, 3D reconstruction, and physical properties estimation. Each of these applications comprises different types of tasks. Synthetic data generation has been used for generating super-resolution, denoised and refocused, and augmented images. Likewise, structure identification has been addressed with a classification of crystal structure, rotational states, defects, structural forms, and phases, with the detection of nanoparticles and crystals, and with the segmentation of atoms, atomic columns, defects, nanoparticles, and phases. Similarly, 3D reconstruction has been performed with CHs and thickness estimation or tomograph and hologram reconstruction. Finally, crystal properties, compliance tensor, and mechanical properties are the material's physical properties estimated via DL.

Nanoparticle models and simulated images generated using Beer's law, Gaussian noise, and transfer function modulation were used by Yao and co-workers for training a DL model to identify nanoparticles' dynamics in liquid phase TEM videos [128]. Förster et al. employed MD to compute carbon nanotube (CNT) structures, serving as input to multislice HRTEM images simulation to develop a dataset for training a neural network for chirality estimation [129]. Ziatdinov et al. trained a CNN on STM images simulated via Markov chain Monte Carlo sampler to classify rotational states of complex molecular assemblies on surfaces, modeled with DFT calculations [130]. Yeom and co-workers adopted phase field (PF) simulations to generate images of Al-Zn for training a SegNet model for the alloys segmentation task [131]. Simulated diffraction patterns [79,80,101,132,133], and simulated tomographs [134] in alternative to microscopy images were also used for DL analysis of 3D reconstruction, structure classification, and physical properties estimation.

In summary, the algorithmic generation of microscopy images datasets is a convenient and popular approach in many DL applications in materials science. The workflow of this method is characterized by an initial simulation of materials models, typically with DFT and MD calculations, followed by simulations of microscopy images of the corresponding materials models, as illustrated in Fig. 11. Within this modeling workflow, DFT and MD calculations are typically implemented with software/codes such as the Vienna Ab Initio Simulation Package (VASP) [135], Atomistic Tool Kit (ATK) [136], LAMMPS, etc. [137]. Alternately, the atomic simulation environment (ASE) [138] is an open-source library for atomistic simulations implemented in the Python programming language. Similarly, a variety of computational packages are available for microscopy image simulations, such as Computem [139], QSTEM [140], Dr. Probe [141], PRISM [142], etc. PyQSTEM is a Python version of QSTEM with a dedicated interface to ASE, providing a convenient platform for dataset simulations in DL analysis [143].

5. Review of deep learning application in microscopy imaging in materials science

In this section, a general overview of the identified applications of DL in microscopy imaging and materials science, including synthetic image generation, structure identification, 3D reconstruction, and physical properties estimation, is provided. Each of these applications plays a central role in the study of the structure–property relationship of materials. Synthetic image generation is crucial to expand the number of images for materials inspection, allowing for the analysis of the broad distribution of structures and properties, and improving imaging quality by refocusing and denoising. Then, structure identification is essential for extracting the relevant structural information necessary for physical properties predictions. In this context, 3D reconstruction is fundamental in applications where the physical properties depend mainly on 3D structural arrangements. Finally, properties estimation is important in linking the structural properties observed in microscopy images to the physical behavior of materials. In all these materials science fields, the development of innovative and cutting-edge DL methods has advanced the state-of-the-art, opening a new frontier in the investigation of structure-properties relationships. This review presents a summary of the methods followed by a description of their advantages and limitations. A quantitative comparison between the performance of DL methods with a conventional microscopy imaging techniques is also provided. A schematic of the reviewed applications and related tasks is illustrated in Fig. 11.

5.1. Deep learning applications for synthetic image generation

In this section, the methods for the generation of synthetic images via DL are presented. Synthetic image generation refers to the ability of DL models to produce artificial images with desired characteristics. There are three distinct tasks in this area of application, i.e., super-resolution image generation, image denoising and refocusing, and data augmentation. The aim of these applications is the enhancement of the dataset's quality by improving imaging conditions (i.e., super-resolution, image denoising, and refocusing) and increasing the amount of data with artifact materials structures or synthetic microscopy images (i.e., data augmentation).

5.1.1. Super-resolution

A typical shortcoming of the traditional electron microscopy imaging analysis is the inherently destructive nature of the electron beam-inducing charging effects and sample deformation, especially for dielectric and soft materials. For this reason, shorter exposure times or low electron doses are typically considered to minimize the beam damage during image acquisition, resulting in the lack of high-resolution representations. Computational techniques such as multi-resolution data fusion [144], and image deconvolution for spatial blur reduction [145], and cumbersome experimental methods as electron ptychography [146], are often used for enhancing the resolution of electron microscopy images, however, they have their own limitations in practical applications.

In recent years, Dong et al. demonstrated that CNNs could be used for efficiently predicting super-resolution images from the corresponding low-resolution images as input [147]. In materials science, Haan et al. adopted a similar method and developed a convolutional GAN for generating high-resolution SEM images of gold nanoparticles and hydrogels [96]. Their approach outperforms the former conventional methods, as it applies to a wider range of samples and requires a single SEM image as an input exclusively. Their GAN comprises a U-Net generator responsible for high-resolution image estimation and a VGG16 discriminator distinguishing between low-resolution and high-resolution images. The GAN was trained on pairs of low-resolution images acquired with 14.2 nm pixel size ($10000 \times$ magnification) and high-resolution images with 7.1 nm pixel size ($20000 \times$ magnification). The performance of the model was evaluated by comparing the percentage of detected gaps between nanoparticles, since more gaps could be detected with a refined image representation. For instance, in a low-resolution image, 13.9% of gaps between nanoparticles were not detectable, while in the GAN-enhanced resolution image, only 3.7% of gaps could not be identified. Although the method developed by Haan and co-workers was demonstrated to be successful for SEM images, it could be applied to enhance the image resolution in other areas of microscopy such as TEM, STEM, etc. Such flexibility is due to the versatile nature of learning algorithms, as they are not bounded to a

specific physical problem, but they can be adapted to different scenarios by simply changing the dataset of interest. Despite the method proposed by Haan et al. has the advantage of generating on-the-fly super-resolution microscopy images with a versatile nature, a shortcoming is the lower performance of the GAN on images representing samples susceptible to electron beam damage, or cases affected by the movement of the samples during the images acquisition.

5.1.2. Refocusing and denoising

A common problem affecting the quality of microscopy images is a noised, out-of-focus, or blurred acquisition, resulting from an incorrect manual adjustment of microscope parameters, improper hardware settings, and imaging automation errors. Deblurring, autofocusing, and denoising techniques have been developed in the literature, but they have restrictions in practical applications. For instance, deconvolution techniques based on local linear and non-linear filters [148,149], non-local self-similarity filters [150], manifold representation [151], and Bayesian restoration [152], have been developed to restore microscopy images by removing blur and recovering sharpness. However, these approaches are limited to “non-blind” restoration, where the deblurring kernel must be known a priori, while in most real-world scenarios, such deblurring kernels are usually unknown. Similarly, autofocus methods based on image quality metrics as image variance, image entropy, and image gradient [153], have been proposed to improve microscopy image’s sharpness [154,155]. Although these metrics are widely used to assess image quality in general, their high dependence on the image’s contrast and brightness makes them inappropriate for focusing evaluations on microscopy analysis. In TEM, image denoising can be achieved by reconstructing the electron beam exit-wave function, as the latter incorporates the sample’s information before being distorted by the lens’ aberrations. Nevertheless, standard exit-wave function reconstruction methods require a cumbersome acquisition of images at varying defocus values (i.e., focal series). These challenges motivate the development of novel approaches for refocusing and denoising microscopy images, and DL based techniques have demonstrated efficient capability to achieve these goals.

In recent years, CNNs have been established as innovative refocusing and de-blurring methods in different applications of natural scene images [156]. The success of these methods has paved the way for their application also in microscopy imaging studies. Lee et al. proposed an autofocus method for SEM images built on a dual DL network, including an autofocusing-evaluation network (AENet) and an autofocusing-control network (ACnet), working together in an iterative process [157]. The AENet was used to evaluate the focus quality of an input image, while the ACnet was employed for online control of the working distance (WD) microscope parameter to optimize the acquisition of in-focus SEM images to be given as the input to the AENet. Lee et al. concluded that their developed DL method is efficient in finding WD values allowing for the experimental acquisition in-focus SEM images.

Similarly, Na et al. developed a multi-scale refocusing network (MRN) to restore defocused SEM images representing martensitic steel and precipitation-hardened alloys using the filters of their trained CNN [158]. In their proposed framework, data augmentation was also used to build a non-uniformly defocused dataset to allow the MRN to restore images affected by varying defocusing distribution. The most important contribution of this work was the successful refocusing in complicated “blind” settings, with no a-priori knowledge of deblurring filters. However, Na et al. have not demonstrated the applicability of the method on out-of-distributions cases, such as images with varying focus blur levels or representing other classes of microstructures. On the other hand, Lee and co-workers developed a convolutional dual-decoder autoencoder (CDDAE) for reconstructing the phase and the amplitude of the exit-wave function of input microscopy images, applied subsequently to the corresponding contrast transfer function (CTF) for the restoration of denoised images [86]. The authors applied their innovative method to a dataset of TEM images representing graphene structures. By evaluating the denoising performance of the CDDAE with metrics such as signal-to-noise ratio (SNR), peak-signal-to-noise ratio (PSNR), and structural similarity index map (SSIM), Lee et al. demonstrated that their framework outperformed conventional denoising methods such as Wiener filtering [159]. For instance, the SNR, PSNR, and SSIM of the CDDAE compared to the conventional Wiener filtering were respectively 18.3390 and 16.7880, 26.7727 and 25.2226, and 0.9440 and 0.9229, indicating the higher quality of the image reconstruction provided by the CDDAE. Nonetheless, the proposed method struggles in cases where impurities like Cu and Si are present in the graphene structure, whose identification requires the comparison of the phase and amplitude derived with the CDDAE with the ones obtained with the conventional microscopy methods.

In all the reported applications, DL represents an efficient and accurate method for improving microscopy images quality. Also, in this application, the aforementioned refocusing and denoising techniques could be applied to broader areas of microscopy images and analyzed materials.

5.1.3. Dataset augmentation

The acquisition of a broad dataset of structures or microscopy images could be prohibitive in applications where costly and time-consuming experimental measurements or intensive computation are required. For this reason, generative DL models such as GANs and variational autoencoders (VAEs) have been proposed to generate both artifact materials samples as well as synthetic microscopy images. The main advantage of these methods is the negligible effort and on-the-fly generation of datasets with desired characteristics. The success of these DL-based generative models relies on the ability of reproducing the realistic morphological features of materials, as well as authentic imaging microscopic conditions. Here, we first review the application of DL for artifact microstructure generation, then we discuss the synthetic microscopy image’s formation.

The availability of a broad ensemble of materials is essential for structure–property relationship studies based on materials data mining. Standard methods to generate simulated but realistic microstructures include statistical function-based [160] and physical descriptor-based numerical models [161]. Statistical methods have disadvantages such as information loss in the dimensionality reduction process and computational intractability of high-order correlation functions, while physical models require a time-consuming comparison of the probability density functions (PDFs) of the real and artifact microstructure characteristics. To address these shortcomings, DL-based approaches have been developed to generate synthetic microstructures through a volumetric

representation using generative convolutional neural networks, e.g., VAEs and GANs. The main advantage of generative DL models is that they do not need an a priori explicit knowledge of the PDF of the distribution of the synthetic microstructures, as it is implicitly learned by a neural network during training.

Cang and co-workers proposed a generative VAE model with morphological constraints to create a dataset of artificial sandstone microstructures from a limited set of realistic samples [162]. The generated microstructures were further used to train a ResNet for predicting mechanical properties such as Young's modulus, diffusion coefficient, and permeability coefficient. With this analysis, Cang et al. [162] demonstrated that the VAE's generation of artifact samples leads to a more accurate performance in physical properties prediction compared to the generation via the state-of-the-art generative Markov Random Field (MRF) model. In fact, for all the predicted physical properties, the R^2 score close to 0.95 was achieved using input microstructures generated by the proposed VAE, while a lower value of 0.85 was obtained using MRF. Similarly, Li et al. designed a transfer learning encoder-decoder deep convolutional network based on VGG19 [59] to generate microstructures for a wide range of materials with varying geometrical features and complexity, requiring a single target microstructure as input [163]. By comparing the difference between the real and reconstructed microstructure, Li and co-workers demonstrated that their approach outperforms conventional statistical and physical modeling methods. For example, the error rates of the proposed DL method in comparison to the two-point correlation approach in reconstructing structures like carbonate, ceramics, sandstone, and polymer composites were respectively 7.63% and 45.59%, 1.31% and 37.92%, 3.61% and 24.19%, 3.58%, and 30.65%. Thus, the DL approach results in a more precise microstructure reconstruction compared to the two-point correlation. However, a limitation of their method is the applicability solely to stochastic microstructures, while the effectiveness on deterministic microstructures like period crystallography structures could require further model refinement. In addition, the demonstration of the proposed approach is limited only to 2D microstructures. Another application was demonstrated by Gayon-Lombardo et al., who developed a DC-GAN for the reconstruction of arbitrarily large three-phase microstructures for LIB cathodes and solid oxide fuel cell (SOFC) anodes applications [164]. An innovative contribution of their work is the capability to generate microstructures with periodic boundaries, allowing for the utilization of periodic boundary conditions for accelerating multiphysics simulations required for microstructure optimization and physical properties estimation. A similar GAN-based method for the generation of a synthetic ensemble of heterogeneous energetic (HE) materials was presented by Chun and co-workers [165]. The performance of the designed GAN was tested by comparing the distribution of morphological metrics of voids (i.e., void diameter, aspect ratio, and concentration), crystal phases, and shock response of the real and generated microstructures. A benefit introduced by Chun et al. is the potential generation of novel morphologies, simply controlling the spatial distribution of the generated microstructures. Lastly, Yang et al. advanced the aforementioned methods by proposing a Stack GAN for the generation of high resolution, highly variable, periodic, and highly controllable microstructures [166]. Compared to the state-of-the-art architectures, the Stack GAN allows to capture more information from conditional latent variables adopted in GANs, leading to create microstructures with a higher level of detail in their structural information. Yang and co-workers applied their framework for the generation of barium titanate oxygen (BTO) piezoceramic microstructure, demonstrating the capability of their Stack GAN to successfully reproduce with high fidelity morphological features such as micropores and grain boundaries. For instance, the proposed GAN was capable to generate synthetic structures with an average pores volume fraction of 5.52%, close to the 5.89% of the real microstructures. In addition, the authors showed that by manipulating the GAN's input noise vectors they can control the grain's size, orientation, volume fraction, and permittivity and thus creating a broad distribution of microstructures for inspecting the correlation with physical properties such as piezoelectricity.

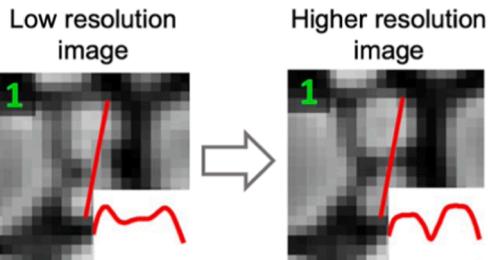
In general, in all of the reviewed works, it was demonstrated that generative DL models are useful techniques for creating an ensemble of materials microstructure with morphological features as grain boundaries and voids or physical properties such as Young's modulus and diffusivity, in agreement with the structures observed in experimental measurements. These capabilities of generative DL models resolve the issue of analyzing an infeasible number of materials samples for collecting a sufficiently large dataset. In the following sections, it is discussed how similar approaches could be used to generate synthetic microscopy images, lifting the requirement of acquiring a large dataset of images through time-consuming experimental measurements and laborious human annotation.

Image segmentation for microstructure investigation is a cumbersome labeling process, as it requires a highly precise and expert pixel-wise annotation, which hinders the collection of sufficiently large datasets for materials exploration. In section 4.2, it was discussed how the challenges of images annotation could be addressed by constructing simulated datasets, where the ground truth is obtained from materials models computed with DFT or MD. However, in some applications, certain theoretical approximations and simplifications prevent the simulation of realistic materials models, resulting in unrealistic images and ground truth. To address this problem, Ma et al. developed an image-to-image conversion model based on a GAN architecture to predict realistic images by corresponding simulated and theoretical images [167]. For this purpose, the authors adopted a dataset optical image of polycrystalline iron labeled, while the corresponding simulated images were computed by slicing 3D models of polycrystalline materials simulated with the Monte Carlo Potts algorithm and labeled with grain's semantic masks. The authors demonstrated that training a U-net segmentation model on a dataset constructed by mixing real images and the synthetic images generated by the GAN could lead to a value of 0.504 of the mean average precision (MAP) metric, whereas lower values between 0.2 and 0.4 were obtained using conventional segmentation methods.

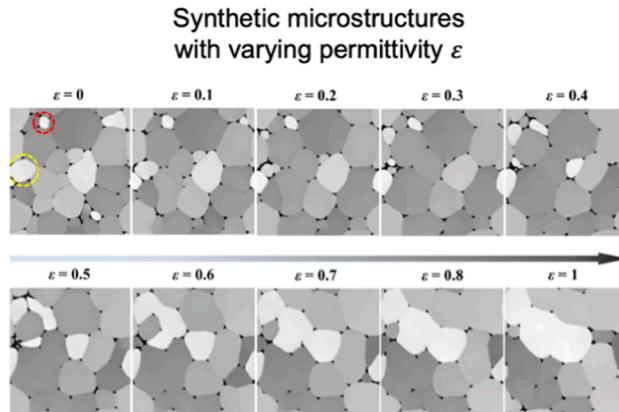
As previously discussed, the materials radiation damage induced by the electron beam is a limiting factor of electron microscopy, and consequently, reduced exposure acquisitions result in lower-resolution images. An alternative method to prevent the sample's damage is the sampling of a sparse random subset of pixels (i.e., partial scans), followed by the image's reconstruction using an inpainting [168] or compressed sensing [169] algorithm. Spiral [170] and jittered grid-like [171] partial scans are typically adopted in microscopy applications; however, their completion is not straightforward using conventional methods. In recent years, DL algorithms

have been successfully applied to image infilling problems, including image completion [172] and irregular gap infilling [173]. Thus, these models could be adapted for the completion of partial scans in electron microscopy. For this purpose, Ede et al. developed a GAN to complete realistic STEM images from partial scans, such as spirals and jittered grids, by training on a dataset of pairs of STEM raster scans and partial scans extracted by the authors [174]. By comparing the image reconstruction obtained with the proposed GAN and with the conventional method, Ede and co-workers showed that their approach reconstructs more realistic images. For example, the GAN-reconstructed images had realistic noise and structural characteristics, whereas the images obtained with standard methods exhibit more blurred representation. In addition, the authors demonstrated that GAN's performance is robust when varying the

(a) Resolution enhancement



(b) Synthetic microstructure generation



(c) Image completions from partial scans

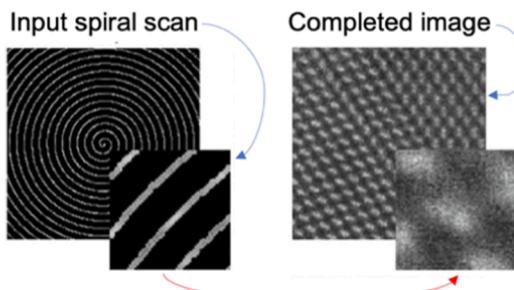


Fig. 12. Examples of synthetic images and structures generation using GANs. In (a), the resolution enhancement application demonstrated by Haan et al. [96]. A low-resolution SEM image cannot show gaps between nanoparticles, resulting in weak signal variation between nanoparticles. On the other hand, the high-resolution image generated by the GAN can capture the gaps, and a signal drop emerges in correspondence to the background. In (b), synthetic microstructures generated by the Stack GAN developed by Yang and co-workers [166]. Linear interpolations between the GAN's input noise vectors allow to create microstructure with varying values of permittivity ε . In (c), image completion from partial scans proposed by Ede et al. [174]. Spiral partial scans from the upper left corner of the global scan giving a smaller portion of the structure represented in the image, are given in input to the GAN to reconstruct the image representing the structure corresponding to the global scan. The blue arrows point to the partial scan in input and the reconstructed image representing the global structure in output, while the red arrow suggests the correspondence between the partial spiral scan and the image representing a portion of the structure. Reproduced from References [96,166], and [174]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

coverage of the partial scans. Thus, the model developed by Ede et al. allows for a high-fidelity reconstruction of partial scans in microscopy images, which are necessary to investigate beam-sensitive materials such as organic crystals, nanotubes, or nanoparticle dispersions. However, the authors did not demonstrate the applicability of the proposed GAN to a different scenario like SEM images or other scan systems. A successful application of synthetic images of cracked surfaces and grain boundaries, and x-ray tomographic synthetic images of Li-ion battery electrode structures, in enhancing neural network performance was also demonstrated by Trampert et al. [176] and Müller et al. [177]. In the application of Müller et al., the synthetic images generated via a style-transfer Cycle GAN were combined with realistic images to construct the training dataset. In general, the aforementioned methods not only offer a strategy to augment the datasets of microscopy images for training DL models, but they also demonstrate that synthetic images allow for accurate model performance. For example, Trampert et al. showed that combining synthetic images with real images for training a surface crack detection model provided an MSE of 0.205, whereas a higher value of 0.273 was obtained by training exclusively on real images. Examples of GAN's applications in synthetic images and structure generation are provided in Fig. 12.

5.2. Deep learning applications to structure identification

In this section, the applications of DL modeling for structure identification are described. With structure identification, we refer to the capability of DL models to identify atoms, atomic columns, defects, chemical species, nanoparticles, crystals, and phases within microscopy images. Structure identification plays a central role in the microscopy analysis of materials, as it allows for an automated and accelerated extraction or classification of structural features, preventing a time-consuming human inspection. The demonstrated success of DL models in various image processing problems broke the ground for their application to multiscale structural visual recognition in microscopy images. In this review, the description of structure identification is categorized in structure segmentation, structure detection and structure classification. Each of these approaches involves different categories of DL techniques, as well as different analyzed structures at varying physical scales.

5.2.1. Structure segmentation

Structure segmentation is a crucial analysis in materials science, allowing for the identification of different categories of structures within microscopy images through a pixel-wise prediction of structure masks. The structure masks could represent atoms and atomic columns [13,15], defects and chemical species at the atomic scales [15,84,127,175], nanoparticles at the nano scale [128,131,176,177], and microstructural phases at the micro scale [81,85,178]. Structure segmentation is the most common DL application in materials science, and it involves the training of FCNs, usually built on the U-Net architecture.

The groundbreaking demonstrations of structure segmentation at the atomic scale were introduced by Ziatdinov et al. [15] and

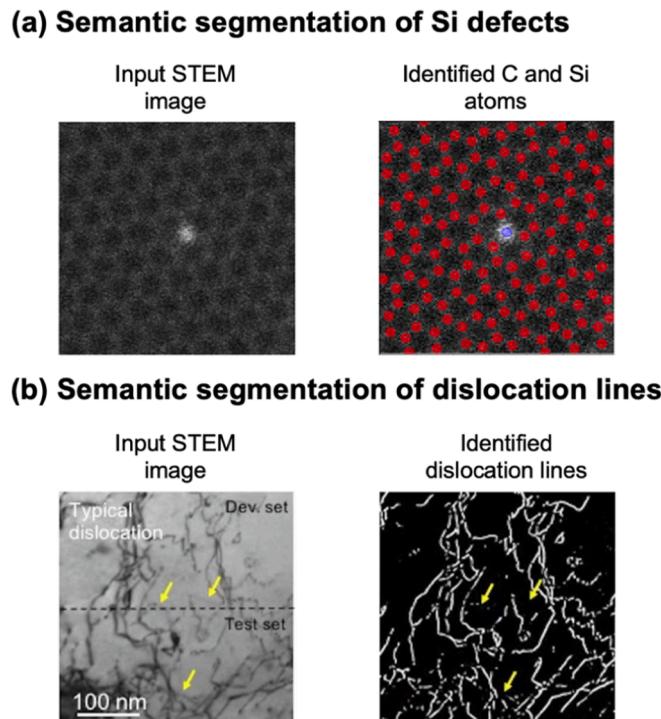


Fig. 13. Examples of defects segmentation using FCNs. (a) Segmentation of C atoms and Si point defects atoms in a graphene structure performed by Ziatdinov et al. [15], and (b) segmentation of dislocation lines defects in steel with DefectSegNet proposed by Roberts et al. [84]. Reproduced from References [15] and [84].

Madsen et al. [13], with application to STEM and HRTEM images, respectively. In both works, the authors introduced the innovative idea of labeling the pixels of atomically resolved images with segmentation masks of atoms, atomic columns, defects like dopants and vacancies, and atoms of different chemical species. Ziatdinov and co-workers employed blobs for the atomic defects and chemical species labeling, while Madsen et al. adopted atomic masks computed with a spatial Gaussian distribution, peaked in the pixels corresponding to the centers of atoms or atomic column representations. Ziatdinov and co-workers applied their framework for the identification of silicon atoms implanted into graphene vacancies, including tracking of the local transformation of such defects during a chemical reaction. Most importantly, their method can detect the structural changes of the chemical bonding of a molecular rotor formed by the Si dopant and the graphene atoms. The authors proved that the Si-C bond length derived by their atom's identification method was ~ 172 pm, in agreement with the 177–179 pm value obtained with DFT calculations for planar SiC structures. On the other hand, Madsen et al. demonstrated that their proposed method is successful in identifying atomic columns of metallic nanoparticles, and under desired imaging conditions, it could also estimate the number of atoms in such columns, a structural feature known as column height. Furthermore, the authors showed that their DL model is capable to capture the positions of atoms diffusing on the surface of the nanoparticles in a gaseous atmosphere.

The success of the aforementioned methods led the way for further in-depth development of DL algorithms for defect detection in microscopy images. Building on the framework proposed by Ziatdinov et al. [15], Maksov and co-workers developed an FCN to identify lattice defect and phase evolution in WS₂, subjected to structural transformations induced by the electron beam in STEM [175]. Their method also incorporates an unsupervised clustering method to classify the identified defects and quantification of dynamics and reactions of sulfur vacancies and Mo dopants. However, a limitation of such a technique is the physics-agnostic nature of the model, as the structural transformations are identified with a pure visual method without incorporating information related to the involved physical processes. In a more recent report, Lee et al. demonstrated a similar approach to locate and classify point defects with sub-picometer precision in a monolayer 2D transition metal dichalcogenide (WSe_{2-2x}Te_{2x}). The analysis proposed by Lee et al. revealed that Se vacancies determine oscillating and complex strain fields in the lattice, causing the alternation of expansion and contraction of the lattice's rings. Roberts et al. introduced DefectSegNet, an FCN for the segmentation of crystallographic defects in structural steel alloys such as dislocation lines, precipitates, and voids [84]. Comparing the detection performance of the neural network with human experts, Roberts et al. showed that their DL model is more efficient, as well as more accurate, in the high-throughput quantifications of the investigated defects. For instance, the FP error in identifying defects like dislocation was 4% for DefectSegNet, whereas the higher 20% was obtained by human analysis. Furthermore, in the determination of precipitates size, an error of 2% was reported for DefectSegNet, while a worst 13% by a human expert. In addition, the computational time required by the neural network to identify defects was reported as ~ 2 s, while approximately 30 min were required for the human inspection. Examples of defects segmentations of Si impurities in graphene matrix (Ziatdinov et al. [15]) and dislocation lines in steel (Roberts et al. [84]) are reported in Fig. 13.

The identification of metallic nanoparticles and the inspection of their dynamics is another important field of study in materials science. In fact, metallic nanoparticles are promising candidates for industrial catalysis and energy storage applications, as they benefit from exceptional catalytic activity due to their high surface energy. For this reason, it is of great interest to precisely quantify the nanoparticles' size and tracking their location during the dynamic evolution process. Semantic segmentation of microscopy images via DL has been demonstrated as an efficient technique to achieve these goals. Compared to the previously reviewed applications where atomic scale segmentation masks are assigned, in this case, the pixel labeling is performed for nanoscale objects. Oktay et al. developed a multi-output CNN (MO-CNN) for simultaneously estimating nanoparticles locations, shape, and size via segmentation of TEM images [179]. Horwath et al. trained a U-Net-type model to segment metallic nanoparticles represented in high-resolution environmental transmission electron microscopy (ETEM), performing a parametric study of the network's hyperparameters to achieve optimal performance [177]. Similarly, Yao and co-workers implemented a segmentation model built on the U-Net architecture to investigate the diffusion, interaction, reaction kinetics, and assembly dynamics in the liquid phase of colloidal nanoparticles from TEM videos [128]. Furthermore, Yeom et al. applied a SegNet-based CNN to x-ray computed tomography (XCT) images to segment Al-Zn alloys synthesized from in-situ solidification experiments. The authors demonstrated that the most accurate nanoparticle detection performance was achieved by training on images representing microstructural and imaging features such as diffusive particle-background interfaces, presence of noise, and absence of nanoparticles at the images' edges. Ma et al. proposed a similar approach for the segmentation of Al-La alloys, demonstrating that their DL model outperformed conventional segmentation methods in the identification of dendrites along the alloys' boundaries [176]. Indeed, Ma and co-workers showed that their proposed technique could achieve an IoU around 93%, while lower values of 87%, 65%, and 56% were achieved by other segmentation methods like Watershed, Otsu, and Adaptive Gaussian, respectively. Within the framework, a method for the rectification of the predicted segmentation masks was also developed to potentially eliminate from the images the noise resulting from the contamination formed during the sample preparation process. However, a limitation of this method is the lower performance in segmenting close dendrite structures, a task that could require incorporating 3D information within the proposed framework.

Semantic segmentation has been also adopted for structure identification tasks at the microscale. In sections 3.1 and 3.3, we have illustrated the pioneering microstructure segmentation methods proposed by DeCost et al. [81] and Azimi et al. [85]. In both works, the authors demonstrated that neural networks are robust and effective tools for determining the distribution, size, and grain boundaries of microstructures, e.g., martensite, bainite, pearlite, and spheroidite. Azimi and co-workers [85] demonstrated that their proposed FCN was capable of detecting different microstructures with a precision of 99.08%, 91.78%, 84.19%, and 97.23% for the martensite, tempered martensite, bainite, and pearlite respectively. DeCost et al. [81], adopting a similar method but with a different neural network architecture, reported a precision of 81.8% on ferritic matrix, 94.9% on spheroidite, and 72.2% on Widmanstätten. The different values of precision on distinct microstructures are due to the class imbalance in the adopted dataset, where the most populated microstructure class is more favorable to be correctly predicted. DeCost et al. [81] also reported that their Pixelnet exhibited

IoU values of 90.0% and 54.8% in predicting ferritic matrix and spheroidite, while the Otsu method was limited to 86.2% and 53.7% on the same dataset. In a recent publication, Müller et al. applied microstructure segmentation to graphite-silicon composite electrodes of Li-ion batteries [178]. The authors trained a 3D U-Net model to segment x-ray tomographic microscopy images in pore space, graphite, silicon particles and carbon black-binder domain classes. In particular, it was illustrated how the performed segmentation provides insights on the evolution of the microstructure of composite anodes during the battery's cycling, by revealing that morphological changes are more likely to occur around the silicon particles rather than the graphite particles. However, a drawback of their method is the segmentation of pore space and graphite, due to the low contrast between these two phases within the microscopy images. Examples of the structure segmentation at the micro, nano and atomic scales are shown in Fig. 14.

5.2.2. Structure detection

Structure detection is a similar application to structure segmentation; however, it relies on object detection, a widespread technique in computer vision for the identification of object categories using bounding boxes. Object detection has been applied in materials science to identify nanoparticles [103–105], and crystals [82] within microscopy images, using state-of-the-art models such as Cascade RCNN and Mask RCNN. It should be noted that Mask RCNN predicts a segmentation mask in addition to the bounding boxes, but since it is built on the RCNN object detection model, we have decided to include its applications in the structure detection section rather than structure segmentation, where pure segmentation models based on the U-Net are adopted. The main objective of these applications is the identification of nanoparticles and crystals to estimate their shape and size, which are morphological features of vital importance in the structure–property relationship analysis of these materials.

Okunev et al. proposed a pre-trained Cascade RCNN applied to STM images for an automated localization of platinum nanoparticles deposited on highly oriented pyrolytic graphite (HOPG), with the further objective of determining their size and height [103]. Within their framework, the detection was performed with the estimation of contour boundaries around the nanoparticles, labeled by a material science expert in the ground truth formation stage. Although the network was not capable of detecting nanoparticles in images with poor imaging quality, it could automatically identify 175 out of 254 nanoparticles marked by an operator, resulting in a detection accuracy of 68%, however, with a significant increase in the detection process speedup. In a subsequent publication, Okunev et al. adopted Cascade Mask-RCNN and introduced a further boundary contour refinement with a 2D Gaussian fit algorithm approximating the nanoparticles shapes, with the objective of providing a more precise nanoparticles detection [104]. The method was demonstrated on a “rough” dataset, where greater attention was given in labeling a high number of nanoparticles rather than defining contour boundaries with precise size and shape. Then, the model was applied on a “precise” dataset, where a lower number of nanoparticles were labeled compared to the previous study, but with more careful attention to the definition of the nanoparticle’s borders. While in the first analysis, the network showed high accuracy in nanoparticles detection but poor performance in size evaluation, in the second study, the nanoparticles’ shape and size were predicted with enhanced precision. A similar method of

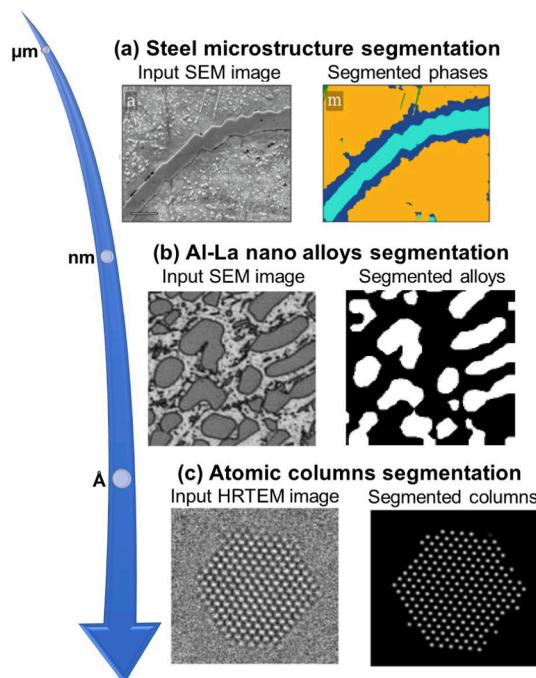


Fig. 14. Examples of multiscale structure segmentation. (a) Segmentation at the micro scale of cementite, ferritic, and spheroidite phases of carbon steel performed by De Cost et al. [81]. (b) Segmentation at the nano scale of Al-La alloys demonstrated by Ma et al. [176]. (c) Segmentation at the atomic scale of atomic columns of gold nanoparticles developed by Madsen et al. [13]. Reproduced from References [81,176] and [13].

nanoparticle shape and size evaluation based on Mask-RCNN detection and edge fitting was proposed by Zhang and co-workers [105]. The authors showed that the model is accurate even in the complex scenario of agglomerated and adhered nanoparticles. In addition, the proposed edge fitting algorithm allowed for a contour boundary reconstruction for incomplete representation of nanoparticles at the edge of the image or joint nanoparticles after adhesion. The authors reported an IoU of 0.92 computed on 630 analyzed nanoparticles.

The Mask RCNN was also used by Masubuchi et al. for searching exfoliated 2D crystals sparsely distributed over SiO₂/Si substrates represented in optical microscopy images and classifying them into graphene, hBN, MoS₂, and WTe₂ materials classes [82]. The authors demonstrated that their detection method is robust under variation of microscopy conditions such as color balance and illumination and has satisfactory generalization capability in predicting exfoliated 2D crystals not included in the training set. For instance, the model trained on WTe₂ could accurately detect flakes like WSe₂ and MoTe₂ with different signal contrast in the images. Satisfactory precision and recall values of 0.95 and 0.97 were reported by the authors. Furthermore, their main contribution was the integration of the DL algorithm in a motorized optical microscope through a dedicated software pipeline implemented by the authors in the C++ and Python programming languages. Thus, Masubuchi et al. developed one the first practical demonstration of the integration of DL models with a microscope for automated and real-time *in situ* searching of materials structures. Fig. 15 illustrates examples of the reviewed crystals and nanoparticles detection studies.

5.2.3. Structure classification

Structure classification is another important objective of structure identification in materials science. A comprehensive variety of materials' structures exists in nature, and each structural class has unique physical properties. Intelligible examples are the classification of crystal into structural classes such as cubic, trigonal, hexagonal, etc., or the classification of microstructures into pearlite, spheroidite, martensite, etc. As it is often a demanding task to categorize materials based on their structural features, DL methods have been extensively adopted in the literature as an alternative to a more laborious and time-consuming manual classification. In the applications reported in this review, materials have been classified according to crystal structure [80,132,180–182], rotational classes [130], defects [113,183,184], at the atomic scale, structural forms at the nanoscale [100,102], and microstructural phases at the micro scale [85,99]. Comparing to structure segmentation and structure detection, where structures are identified and classified within the images, in structure classification, the class is assigned at the image level, meaning that each image represents a different category of structures.

Crystal's classification is the most targeted application of structure classification using DL models, and prominent methods have been published in the literature. In most of the applications, diffraction patterns have been used since they are more directly correlated to crystal symmetry information than microscopy images. Ziletti et al. developed a pioneering framework based on CNN to efficiently classify crystal structures, including structures with defects, impurities, or experimental noise, outperforming conventional methods

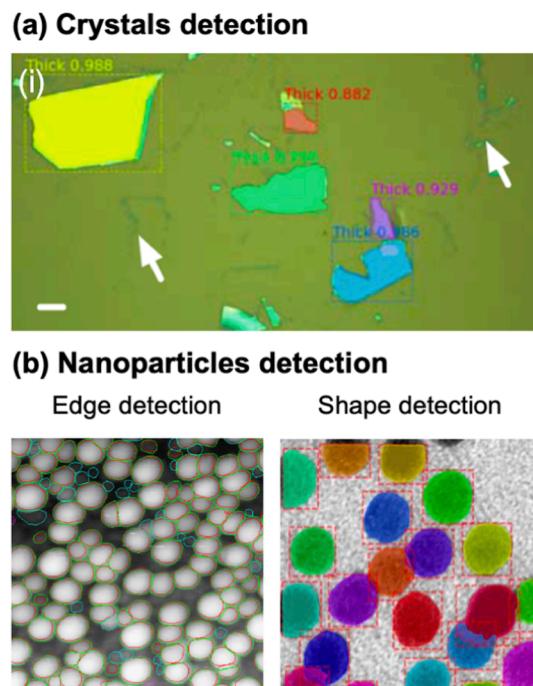


Fig. 15. Structure detection using Mask-RCNN. (a) Detection of 2D crystals sparsely distributed over SiO₂/Si substrates represented in optical microscopy images reported by Masubuchi et al. [82]. (b) Edge and shape detection of nanoparticles presented by Okunev et al. and Zhang et al. [104,105]. Reproduced from References [82] and [104,105].

based on space groups, relying on the symmetry of idealized crystal structures [181]. The proposed method was demonstrated on algorithmically computed diffraction fingerprints images of simulated body-centered tetragonal 139 and 141, rhombohedral 166, hexagonal 194, diamond, simple cubic, face-centered cubic, and body-centered cubic crystal structures, including cases with a concentration of defects up to 60%. The authors demonstrated that their model could achieve an accuracy of 100%, providing a perfect prediction even in the case of highly defective crystal structures. Kaufmann and co-workers proposed a similar method for Bravais lattice classification but applied to simulated electron backscatter diffraction images [80]. In the developed framework, the crystal structures were successfully classified with an accuracy higher than 90% using ResNet and Xception architectures. A Bravais lattice classification approach applied to a 2D fast Fourier transform of simulated and experimental STM and STEM images was developed by Vasudevan et al. [180]. The neural network was also applied to Mo-doped WS₂ structures subjected to electron beam-induced transformations to determine its phase transformation kinetics. However, a limitation of the proposed method is that it is not applicable to non-centrometric crystals. Crystal structure classification applied to simulated diffraction patterns images representing crystals belonging to seven lattice families and 32-point symmetry groups used as class labels was developed by Aguiar et al. [182]. Similar to previous reports, an exceptional accuracy over 90% was achieved for most of the classes, with a peak of 99.24% for hexagonal-trapezohedral and the lowest value of 65.80% for tetragonal-pyramidal crystals. A related work was published by Tiong et al., where 72 space groups (SGs) (out of the 230 SGs formed by the combination of 32-point groups and 14 Bravais lattices) were considered for crystal classification. Comparing to previous methods where few high-level categories were considered by classifying crystals into symmetry classes, Bravais lattices, or point groups, the SGs-based model introduced by Tiong et al. allowed for a classification of crystals with a superior level of complexity. The authors implemented an advanced DL model named multistream DenseNet (MSDN), achieving a satisfactory accuracy around 80% in classifying crystals into the considered 72 SGs.

The knowledge of structural and rotational configurations of individual molecules in a lattice substrate is fundamental for investigating the kinetics of such molecules and related short-range orders phenomena leading to conformational changes reactions. For this purpose, Ziatdinov et al. introduced a machine vision approach to classify structural and rotational states of individual building blocks of sumanene molecules assembled on gold lattice surfaces represented in STM images [130]. In this method, blow-up and blown-down structural states and three rotational states on a [2,2,2] gold surface plus a rotation class corresponding to eventual imperfections of the molecular film or underlying substrate were considered. The authors demonstrated that the proposed technique could be employed for a precise reconstruction of the pair density function to analyze the disorder-property relationship of the system and retrieving reaction pathways involving molecular conformation transformations.

Defect identification is a crucial analysis in materials science, and defect classification approaches have been proposed in the literature in addition to the reported defect segmentation methods. For instance, Cheon et al. applied a CNN to SEM images for classifying wafer surface defects into the spot, rock-shaped particles, ring-shaped particles, misalignment, and scratch defect classes [183]. The authors illustrated that the developed CNN could outperform other learning algorithms such as multi-layer perceptron (MLP), support vector machine (SVM), and stacked autoencoder (SAE). For instance, their CNN achieved a defects classification of 96.2%, while MLP, SVM, and SAE were limited to 92.8%, 92.5%, and 91.8%. Similarly, Imoto and co-workers adopted transfer learning with fine-tuning of the Inception CNN [60], for an automatic defect classification of semiconductor manufacturing materials represented in SEM images [184]. Comparing with conventional automated defect classification methods achieving an accuracy of 77.2%, their method provides superior accuracy of 87.3%, as well as significantly reduced time in manual image inspection. Similarly, Badmos et al. used a transfer learning with a finetuning-based approach for the automated detection of microstructural defects in Lithium-ion batteries electrodes' materials reported in light microscopy images of sectioned battery cells [113]. The model was implemented to perform a binary classification into the "presence of defects" such as layer contamination, metal particle contaminant, non-uniform coating, and "without defects" classes. The authors demonstrated first that transfer learning benefits of superior accuracy compared to a CNN trained purely on battery electrodes images; then, they showed that fine-tuned VGG19 [59] leads to more accurate estimation compared to fine-tuned Inception [60] and Xception CNNs [185]. Although, a limitation of their approach is the difficulty in the acquisition of sufficiently high-resolution experimental images required for accurately training the model. A further downside of the proposed method is the bias of the preparation process and imaging conditions on the pixel's intensities of the collected experimental images.

At the nano-scale, structure classification has been applied to categorize structural forms of objects such as particles, nanowires, films, and pillars [102], carbon nanotubes (CNTs)/nanofibers (CNFs) [100]. Modarres et al. proposed an original transfer learning of Inception and ResNet models to classify nanomaterials represented in SEM image into the following ten classes: tips, particles, patterned surfaces, MEMS devices and electrodes, nanowires, porous sponge, biological, powder, films and coated surfaces, and fibers [102]. With an in-depth analysis of the performance achieved in each class, it was concluded that classes with the most representative features as patterned surfaces, MEMS, particles, nanowires, and tips could achieve accuracies higher than 90%, while the other classes were predicted with lower but still satisfactory accuracies. In a later report, Luo et al. performed a classification of TEM images covering all the major categories of airborne carbon nanostructures employed in the manufacturing of CNT and CNF materials [100]. The authors initially partitioned the images in five classes corresponding to the most observed structural forms of CNTs and CNFs, namely cluster, fiber, matrix, matrix-surface, and non-CNT, originated for example, by the stacking arrangements of graphene sheets, inter-fiber interaction, and aggregation and structural mixing. In a further step, the classification was refined with the introduction of five additional classes to categorize the non-CNT structures in graphene sheets, soot particles, high-density particles, polymer residuals, and others. By applying transfer learning of VGG16 and data augmentation, the authors showed that their approach leads to a classification accuracy of 90.9% and 84.5% in the two cases. However, a drawback of the presented methodology is the lower performance in classifying images from structural classes with high similarity, like matrix and matrix-surface.

The classification of microstructural phases has particular importance since it represents one of the pioneering demonstrations of

the application of DL to microscopy images for materials classification. Azimi et al. were the first in the field to present a framework based on a pre-trained VGG16 model applied to SEM images to classify the microstructure of low-carbon steel into martensitic, tempered martensitic, bainitic, and pearlitic phases [85]. The innovative method proposed by Azimi et al. could achieve a high accuracy of 93.94%, outperforming an earlier non-DL technique limited to 48.89%. Additionally, the authors reported individual structural classes accuracies of 99.08%, 91.78%, 84.19%, and 92.23% for martensitic, tempered martensitic, bainitic, and pearlitic phases, respectively. As mentioned in section 4.1, the MECS [110] publicly available database of experimental SEM images labeled by materials science experts was used, and image augmentation was adopted to address the problem of imbalance among the microstructural classes. A similar and cutting-edge method for low-carbon steels' microstructure classification using DL applied to the SEM images in the UHCS database [108] was introduced by DeCost et al. [99]. Similar as in the work developed by Azimi et al. [85], the transfer-learning of VGG16 was adopted; however, additional microstructural classes were considered, including spheroidite, proeutectoid cementite network, pearlite, pearlite + spheroidite, Widmanstätten cementite, pearlite + Widmanstätten and Martensite/Bainite. A classification accuracy higher than 95% was reported. One of the most important contributions of their work was the exploration of the images' features using a vector of locally aggregated descriptors (VLAD) [186] applied to the hidden features maps of the VGG16, fed to the t-distributed stochastic neighbor embedding (t-SNE) algorithm for projecting in a 2D space the obtained images' high dimension features vectors. The method was employed to visualize the distributions and relationships of the images' extracted high-dimensional microstructural features by clustering their projections in the 2D t-SNE microstructure maps. For instance, the authors showed that spheroidite, pearlite, and network images tend to form distinct and extended clusters, while martensitic images are agglomerated in a cluster close to pearlite due to their similar morphologies. The proposed approach is a useful technique to visualize the structural features learned by a DL model in materials classification and to analyze microscopy images of materials in

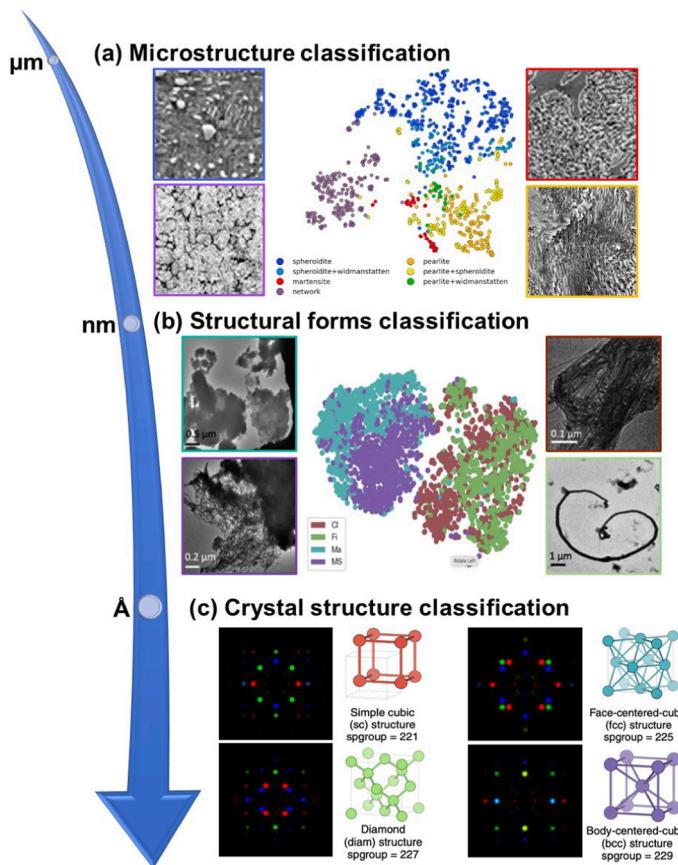


Fig. 16. Examples of multiscale structure classification using CNNs. In (a), DeCost et al. adopted transfer learning of the VGG16 model to classify SEM images of steel into microstructural classes such as spheroidite, (blue) proeutectoid cementite network, (purple) martensite, (red) and pearlite + spheroidite. (yellow) For illustration purposes, the classified images have been projected in to the t-SNE scatter plot map [99]. In (b), Luo et al. used a pre-trained VGG16 to classify the structural forms of the major categories of airborne carbon nanostructures employed in the manufacturing of CNT and CNF material [100]. As in (a), the t-SNE diagram has been used to project the images into the condensed matrix structures, (Ma, light blue) oversized mixed structures, (MS, purple) clusters of packed fibers, (Cl, brown), and single fibers (Fi, green). In (c), Ziletti et al. trained a CNN on computed diffraction fingerprints images to classify crystal structures into classes, e.g., simple cubic, diamond, face-centered-cubic, and body-centered-cubic [181]. Reproduced from References [99,100], and [181]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

general. In Fig. 16, examples of microstructures, structural forms and crystal structures classification, are illustrated.

5.3. Deep learning applications to 3D reconstruction

The 3D reconstruction of materials' structure is another field of study of high interest for the determination of structure–property relationships of materials with complex morphology. In this section, the applications of DL for the 3D reconstruction of materials are categorized into nanoparticle's atomic column heights estimation, thickness evaluation, and tomography reconstruction.

5.3.1. Atomic column heights estimation

As mentioned in earlier sections, the estimation of structural characteristics of nanoparticles plays a central role in the evaluation of their physical properties, such as surface energy in catalysis applications, and DL methods have proven effective capabilities in determining size and shape by predicting segmentation masks or bounding boxes. However, these approaches offer a coarse characterization of nanoparticles' structure limited to 2D representations, while a more in-depth analysis of their atomic arrangements could reveal more precise morphological features in the 3D space. Within nanoparticles, atoms are aligned in ordinate columns along different crystallographic directions depending on their crystal structure. The number of atoms in the atomic columns is referred in materials science as column height, and its evaluation is the typical objective for the 3D reconstruction of nanoparticles. Techniques based on linear interpolation, parameter estimation theory, and least square estimator applied to STEM images have demonstrated the capability to accurately predict the column heights of single chemical element nanoparticles of different sizes [187–189]. In general, these methods assume that in STEM images, the pixel intensity is linearly correlated with the number of atoms distributed in the atomic columns. On the other hand, the relationship between the pixel intensities in the HRTEM image and the column heights is highly non-linear and sensitive to many microscope parameters as defocus, electron dose, etc. The state-of-the-art techniques for extracting the features of the crystal structures from the experimentally acquired HRTEM images include the cumbersome installment of a dedicated image corrector and the extraction of the exit wave function from a focal series [190–192]. The determination of the 3D configuration of metallic nanoparticles from HRTEM images has been performed in several attempts presented in the literature; however, these studies are limited to restricted areas of application, such as simulated images, single surface orientation, and liquid type of substrate [193–195].

As previously reviewed, Madsen et al. introduced a novel DL framework based on semantic segmentation for local atomic structure identification in HRTEM images, with the potential of estimating the atomic column heights of metallic nanoparticles [13]. The proposed neural network segments the atomic columns and classifies them according to their height by training on simulated HRTEM images computed with varying values of the microscope parameters to ensure a sufficient model's robustness with respect to the microscopy imaging conditions. For each nanoparticle, a focal series with defocus values randomly picked within a defined range was given as input to the network to mimic the prediction of the column heights from a focal series. The authors demonstrated that with their technique, the column heights could be predicted with an error up to two atoms in simulated images. Although a successful model's performance was demonstrated on simulated images representing samples perfectly aligned with the electron beam, a limitation of the proposed method is that its application was not demonstrated on experimental images, and larger column height errors were obtained for samples with a slightly tilted zone axis. This innovative DL approach was further extended by Ragone and co-workers, who developed a regression-based neural network version of the original model designed by Madsen et al. [14]. The most significant contribution was the application of the trained DL model to experimental HRTEM images to predict the column heights in real-life nanoparticles with different sizes and zone axes. Furthermore, Ragone et al. demonstrated that the same technique could be successfully applied in the prediction of the atomic column heights of chemically complex nanoparticles, e.g., HEAs [196]. The neural network trained on simulated STEM images was then adopted to predict the atomic column heights in experimental images of a $\text{Pt}_{0.2}\text{Ni}_{0.2}\text{Pd}_{0.2}\text{Co}_{0.2}\text{Fe}_{0.2}$ HEA for each of the five chemical elements with an error up to one atom in most of the columns. However, a disadvantage of the proposed method is that the trained network is restricted to only such class of HEA and composition, whereas the application to other HEA involving other chemical elements in a different ratio requires additional model training and evaluation.

5.3.2. Thickness estimation

A further method for retrieving 3D structural information of materials is the estimation of the sample's thickness. Position-averaged convergent beam electron diffraction (PACBED) patterns obtained with STEM can be used to evaluate the local thickness and tilt of analyzed materials [197], usually by comparing experimental PACBED patterns with the numerical simulation through least square fitting (LSF) [198] or human perspective. A faster method of thickness and tilt estimation using DL applied to PACBED patterns was introduced by Xu and LeBeau [133]. A transfer learning CNN based on AlexNet architecture [6] was trained on SrTiO_3 and $\text{PbMg}_{0.33}\text{Nb}_{0.66}\text{O}_3$ crystals to predict thickness ranging from 6 to 120 nm and from 8 nm to 70 nm in PACBED acquired with a 13.6 mrad and 19.1 mrad probes respectively. By comparing the predictions of the CNN and LSF, the authors demonstrated the optimal performance of the DL model for thickness below 60 nm with an error up to 2 nm, while for higher thickness, special treatments in PACBED patterns preprocessing were required to minimize discrepancy and blurring effects. For instance, 96.5% of the CNN's predictions matched with LSF within an error just up to 3 nm for 13.6 mrad patterns, while a 94.2% match within 3 nm was reported for 19.1 mrad patterns. A limitation of the method developed by Xu and LeBeau is the reduction of thickness agreement between CNN and LSF for samples with a size larger than 60 nm, with the lowest performance reported for sizes between 100 and 120 nm. The technique proposed by Xu and LeBeau was advanced in a later publication by Zhang and co-workers, with the development of classification and regression CNNs built on the VGG16 architecture for thickness estimation of atomic columns of SrTiO_3 crystals from PACBED patterns of HAADF STEM images [101]. While the classification CNN categorizes crystals with thickness classes ranging from 2 nm to 100 nm

with 2 nm increment, the regression CNN provides a refined estimation by predicting continuous values between 0.5 nm and 35 nm. Comparing the CNNs' predictions with sample thickness extracted from HAADF STEM images corresponding to the PACBED patterns, Zhang et al. showed that the regression CNN yields a more accurate prediction for thickness below 35 nm, with an error up to 1 nm for 70% of the cases and up to 2 nm for 94% of the cases. On the other hand, the classification CNN estimates the thickness with an error up to 2 nm only for 66% of the samples.

5.3.3. Tomograph reconstruction

Tomography is another state-of-the-art technique widely adopted in microscopy to investigate the volumetric distribution of materials' structure, consisting of the acquisition of a tilt series of image projections by rotating a sample along a crystallographic axis to varying exposed orientations [199]. Such a series of stacked 2D projection images representing a material's volume is referred to as a 3D tomographic image dataset. For instance, Chen et al. [200] and Miao et al. [201] presented pioneering studies based on atomic electron tomography (AET) to reconstruct the 3D structure of metallic nanoparticles, allowing for precise localization of atomic displacement, grain boundaries, dislocation, and evaluation of physical properties such as the strain tensor. Although AET is an advanced method for the 3D reconstruction of materials, it involves a cumbersome and challenging reconstruction of the materials volume from the corresponding series of 2D projections, usually performed with complicated algorithms such as the simultaneous iterative reconstruction technique (SIRT) [202], and equal slope tomography (EST) [203]. In addition, the quantification of the material's structure from tomography imaging datasets requires an appropriate segmentation of features in the scanned material's

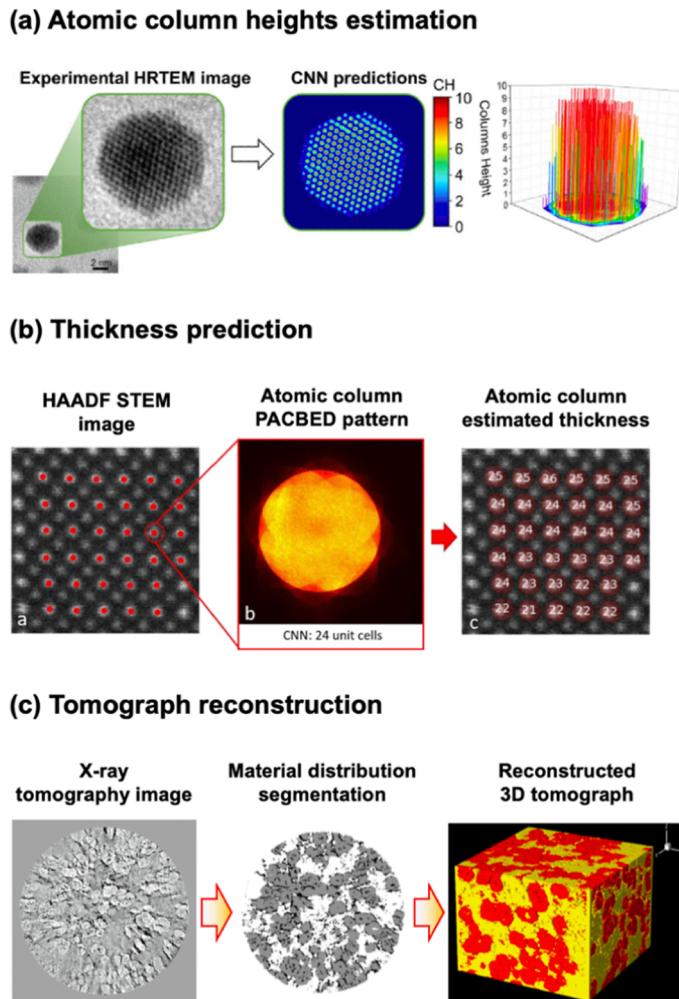


Fig. 17. Examples of 3D reconstruction of materials structure. In (a), Ragone et al. trained a DL model to segment HRTEM images representing gold nanoparticles to estimate the number of atoms in their atomic columns [14]. In (b), Zhang and co-workers trained a CNN on PACBED patterns to predict the thickness of atomic columns SrTiO₃ crystals represented in corresponding HAADF STEM images (in the figure, the predicted thickness is expressed in a unit cell, but the CNN estimates nm values) [101]. In (c), Kodama et al. trained an FCN on X-ray tomography images to segment 2D slices of NMC, SE, and void distribution of cathode materials of Li-ion batteries. The 3D tomograph is then reconstructed by combining the 2D segmented slices predicted by the DL model [205]. Reproduced from References [14,101], and [205].

volume.

Recently, DL algorithms have been demonstrated to be alternative and promising approaches for post-processing tomography image data for the reconstruction of 3D morphological features of materials [134,204,205]. Li and co-workers used CNN to reveal the 3D distribution of L1₂-type δ'-Al₃(LiMg) nanoparticles with face-centered cubic crystal structure from spatial distribution maps (SDMs) obtained with atom probe tomography (APT) [134]. The model was first trained to classify simulated APT-SDMs reporting Al-Li-Mg FCC bulk or L1₂ precipitate's structure categories, achieving a classification accuracy of 100%. Then, the authors demonstrated that the L1₂ precipitates could be successfully identified also in experimentally acquired APT-SDMs, considering crystal structure information such as occupancy sites and atoms positions and its effect on SDM's signal contrast, outperforming traditional methods based solely on composition difference evaluation. Alternatively, Kaira et al. adopted a CNN for the segmentation of the volumetric stack of nanotomography images of an AlCu alloy acquired with transmission X-ray microscopy (TXM), coupled with post-scanning SEM imaging to retrieve 3D morphological parameters [204]. With this method, the 3D tomograph is reconstructed by combining the predicted segmented images of 2D slices. The authors illustrated that the automated segmentation based on DL could accurately quantify the 3D morphology of nanoscale precipitates, coarse θ precipitates nucleating heterogeneously in high energy sites of grain boundaries, and plate-like θ' precipitates. A similar method was proposed by Kodama et al., for the segmentation of 3D images from X-ray nanotomography of cathode materials such as LiCoO₂ or LiNi_{0.5}Mn_{0.3}Co_{0.2}O₂ (NMC), with sulfide solid electrolyte (SE) for all-solid-state Li-ion battery applications [205]. A U-Net was trained to automatically segment the NMC, SE and void components of the cathode assembly, obtaining volumetric fractions of 37%, 43% and 19% respectively, matching the values obtained with experimental measurements. Examples of 3D reconstruction applications are shown in Fig. 17.

5.4. Deep learning applications to physical properties estimation

Physical properties estimation is the ultimate and most important task in the workflow of structure-properties relationship analysis. In previous sections, we have reviewed how DL algorithms have contributed to improve the state-of-knowledge on different steps of a such workflow, from the pre-processing of microscopy images to the identification and reconstruction of structural features of materials. While the summarized reports were more focused on the “structure” side, the actual prediction of the physical properties of the investigated structures was not included. Herein, we complete the review with a description of the methods related to the “properties” side, where DL models were employed for the estimation of physical properties, e.g., crystal [79,129] and mechanical properties [162,206].

5.4.1. Crystal properties evaluation

Crystal properties such as scale factors, lattice parameters, and crystallite sizes are the most important to characterize the operating conditions of materials, and tomographic diffraction imaging are usually used to reconstruct 3D mapping of such properties [207]. In the previous section, we described how DL models have been successfully applied to tomographic image data for 3D structure reconstruction purposes. Dong et al. proposed a similar approach and developed a regression CNN applied to simulated and experimental X-ray diffraction (XRD) computed tomography patterns representing single phase Ni FCC structure and multi-phase Ni-Pd/CeO₂-ZrO₂/Al₂O₃ catalytic materials, for the prediction of scale factor, lattice parameters, and crystallite size volumetric mapping [79,129]. Comparing the model's predictions with the ground truth in the simulated XRD tomography dataset, the authors showed that the crystal properties mapping could be estimated accurately for both the simpler single phase Ni structure as well as the more complex multiphase Ni-Pd/CeO₂-ZrO₂/Al₂O₃ system. Then, the predictions for the experimental XRD images were validated by measuring the difference with respect to the properties mapping obtained with conventional Rietveld analysis [208], revealing an optimal agreement achieved by the DL model. While a slight difference of 2% was measured between the two methods the advantage provided by DL was the much faster processing time of 10 s compared to the 4.4 h required by the Rietveld method. Another application of crystal properties estimation was presented by Förster and co-workers, with the development of CNN for the prediction of chiral indices of carbon nanotubes (CNTs) represented in HRTEM images [129]. The importance of chiral indices lies in their determination of CNT's structure and influence on metallic and semiconducting properties; however, their evaluation using conventional methods, e.g., analysis of electron diffraction patterns from TEM [209], or comparison of experimental and simulated HRTEM images [210], is usually tedious and time-consuming. On the other hand, Förster et al. showed that their network, trained first on a simulated HRTEM dataset, is capable to predict on the fly the chirality of CNTs, with a 71% agreement with a manual and longer evaluation based on the aforementioned conventional methods. In addition, the authors claimed that in the case of disagreement, the error committed by their automated DL method is minimal, and an agreement of 100% could be achieved, when removing from the dataset the most challenging nanotubes.

5.4.2. Mechanical properties evaluation

Among different physical properties, mechanical properties are the most sensitive to the structural features of materials. For instance, the Young's modulus is strongly correlated with the atomic bond strength, while the diffusivity and fluid permeability of porous materials is mainly dictated by the local pores' size distribution and degree of connectivity. Cang et al. demonstrated that mechanical properties, such as the Young's modulus, diffusivity, and fluid permeability of geological porous materials such as sandstones, could be predicted using a regression ResNet trained on corresponding microstructure mapping [162]. The main objective of the proposed method is to use DL to link the sandstone's microstructure represented in the input mapping, to the mechanical properties of interest, without developing an explicit physical model. In their innovative framework, the model's inputs consist of

synthetic microstructural mapping derived with a generative VAE and DCGAN, while the ground truth of the mechanical properties values is computed with effective medium theory [211]. The effectiveness of the method was outlined by the superior R^2 in the mechanical properties' prediction obtained with the developed DL method, compared to a conventional approach-based Markov Random Field algorithm, especially for a large number of analyzed samples. For example, an R^2 of 0.95 was achieved by the proposed DL methods in evaluating the Young's modulus on 1000 samples, while the Markov Random Field algorithm was limited to 0.85, and similar metrics were reported for the diffusion and permeability coefficients. A similar framework combining synthetic microstructural mapping generation using DCGAN and CNN for the prediction of mechanical properties was developed in a later report by Tan and co-workers [206]. The DCGAN was employed to generate images of a porous microstructure with geometrical constraints, such as elliptical holes with random size and location, and the CNN was trained on these images to predict the corresponding compliance tensor of the microscopic stress field, computed with FEM simulations in the algorithmic ground truth generation step. The authors showed that with the proposed method, the compliance tensor could be predicted with a small error up to 5%, with a four-order of magnitude reduction in the computational time compared to standard FEM-based calculations. In addition, the model exhibited satisfactory performance on a dataset of microstructure with more complex topology compared to the ones used for training, proving the robustness of the developed DL method in the prediction of structure–property relationships. Fig. 18 shows examples of the reviewed applications related to physical properties estimation.

6. Advantages and limitations of deep learning in microscopy imaging and materials science

The applications of reviewed DL techniques to various microscopy imaging problems have demonstrated several advantages,

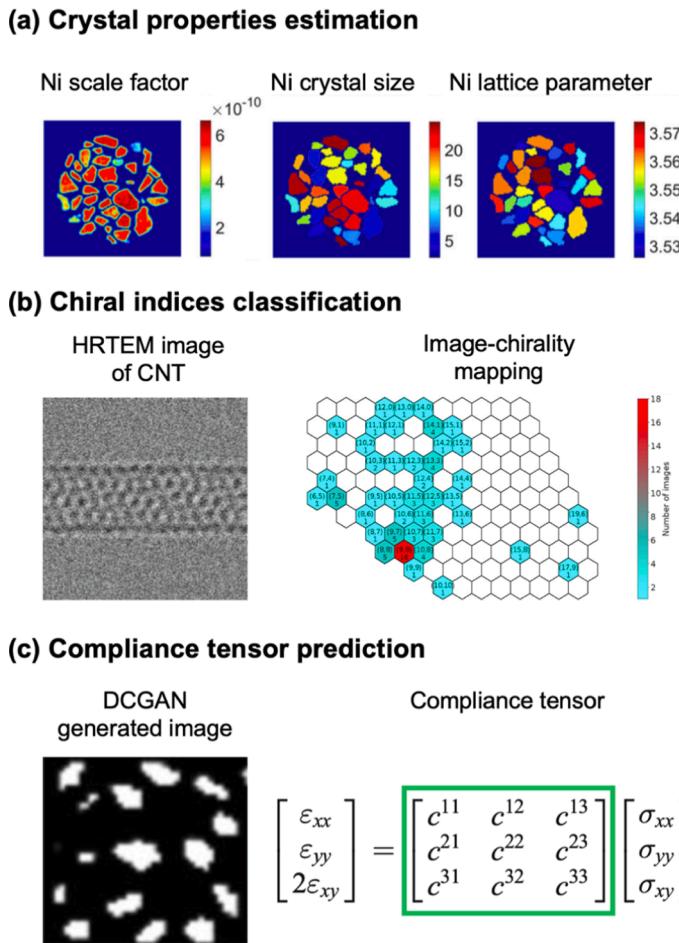


Fig. 18. Examples of material's physical properties estimation. In (a), prediction maps of crystal properties such as scale factor, crystal size, and lattice parameters of multiphase Ni-Pd catalytic materials are represented in X-ray diffraction computed tomography patterns presented by Dong et al. [79,129]. In (b), the classification of chiral indices of CNTs represented in HRTEM images performed by Förster and co-workers[129]. The image-chirality mapping illustrated the number of images predicted in a certain chirality class. In (c), Tan et al. developed a DCGAN for the generation of synthetic images of microstructures, used to train a CNN for the prediction of the corresponding compliance tensor [206]. Reproduced from References [207,129], and [206].

significantly improving materials science research. At the same time, some drawbacks and limitations were revealed depending on the application area. This section summarizes the advantages and disadvantages of the reported DL methods.

The main advantage of the reported DL techniques is the fast and automated processing of large databases of microscopy images, where the material's features can be detected on-the-fly, compared to more time-consuming conventional image processing techniques or human analysis. Another significant DL advantage is the ability to provide quantitative measures via superior image processing metrics, such as accuracy, precision, IoU, MSE, SNR, PSNR, and SSIM, in predicting materials' features of interest, compared to more standard methods used in electron microscopy. The DL-boosted performance in predicting various material features such as defect types and morphology, atoms and nanoparticle location, or 3D structural configurations significantly surpasses conventional methods. Furthermore, DL algorithms can provide remarkable improvement to the quality of microscopy images by enhancing their resolution, removing noise, and restoring focus. In the reviewed papers, improving image quality has allowed, for example, to identify more precise gaps between nanoparticles or to provide insights into a graphene structure by reconstructing the phase and amplitude of refocused images. Additionally, DL algorithms can be used to augment limited databases of materials' microstructures or microscopy images acquired in the experiments; this tool is beneficial to not only a more precise model training but also a broader area of applications in materials analysis. Finally, being data-driven models, DL methods can be used to link microscopy images directly to materials' properties. This capability is imperative for studying structure–property relationships of materials, where DL methods significantly accelerate the process of relating structure to an appropriate property.

However, DL algorithms also have their own limitations and challenges in the analysis of microscopy images and materials. The main drawback of these methods is their limited applicability only to a data domain similar to the one adopted in the training process. A weaker model's performance is commonly encountered when microscopy images or material structures are different from the ones learned by a model. For example, a model trained on STEM images is expected to perform poorly when applied to HRTEM images. This is due to the different types of correlation between the signal encoded in the pixels of a microscopy image and the material's features for distinct categories of microscopy techniques. In such a case, new training on images acquired with the target microscopy technique is required. In some cases, even if the category of microscopy images is the same, the value of microscope parameters adopted in the images acquisition covers different ranges. In such cases, the error of the model's prediction is expected to be higher compared to the error of the prediction based on the same or similar parameters range. A similar diminishing performance arises when a model is applied to images representing materials with structural features diverging from those used in the training domain. For example, if a model is trained to identify platinum atoms within specific crystal structure(s), it is expected to have a similar performance in identifying gold atoms but poor performance in detecting iron atoms due to atom size and property difference. Similarly, a neural network trained to predict the 3D structures of nanoparticles with sizes up to 5 nm will not accurately predict nanoparticles with a larger size of 20 nm. It should be noted that the restricted applicability to data distribution close to the training set is a general problem of learning algorithms in almost all scientific and industrial applications. For this reason, it is extremely important to focus the research efforts on the collection of broad datasets and to perform parametric studies to gain insights into the performance of data-driven models on different data distributions and to expand their generalization ability. In this survey, summaries of such studies described in the reviewed articles have been reported.

In this regard, a further challenge of DL applications in microscopy and materials science is the acquisition and preparation of datasets suitable for training and evaluating a neural network. In Section 4, the demanding process of assembling a dataset of microscopy images through experimental measurements or simulation methods has been described. While different approaches have been adopted in the literature, as described in this review, this step still requires careful and time-consuming analyses. While DL models can perform fast and on-the-fly predictions, another issue concerning these methods is the computational time required to properly train a model. Being computationally time-dependent on the size of the input images, such a problem is particularly critical in materials science, where high-resolution microscopy images are adopted to analyze materials' features. Finally, a further limitation of DL is that it does not provide any insights on the physics characterizing a material represented in a microscopy image. Indeed, DL relies exclusively on data patterns between materials' features and pixel intensities of the images as a "black box," thus, they cannot provide, for example, a physical law describing a specific material and its behavior. In a more general sense, the reviewed DL techniques do not guarantee satisfaction of laws of physics.

In conclusion, this novel class of DL techniques has demonstrated the potential to be superior for image analysis for various applications in materials science. Although less mature than traditional methods, such as the geometric phase analysis (GPA) [212,213], finding local intensity extrema [214,215], direct comparison with a template [216], the paraboloid method (PAM) or the maximum likelihood (MAL) method [217], there is undoubtedly a promising future for DL techniques. Broadly speaking, the challenges include finding ways to improve prediction on untrained data set within the same class of problems, reduce training time, simplifying the training data preparation and relating prediction to underlying physics.

7. Future directions

The reviewed reports represent the dawn of a new pathway in the research of materials' structure–properties relationships, paved by the rising artificial intelligence and materials informatics methods combined with advanced microscopy techniques. Within this innovative field of materials research, new opportunities and challenges are emerging. In general, there are two main areas that necessitate being explored and where novel contributions could be proposed. First, it is important to pose future attention to the dataset's development stage. As it was reviewed, experimental measurements and numerical simulations are used to collect microscopy images. From the experimental measurements' perspective, the effort of microscopists should be focused on the accurate identification of structural features of materials and microscopy imaging conditions, allowing for appropriate input parameters for

simulations in agreement with real-life cases. Such work requires particular attention in in-situ TEM analysis, where terabytes of image data are usually collected. An opportunity in the data development stage is represented by the 4D-STEM field [218], an emerging new direction in TEM research. On the other hand, computational scientists should deeply investigate the data patterns linking the signal contrast of microscopy images with the structure or properties of materials, using, for example, the computed materials models. Although the algorithmic generation of microscopy images' datasets is a well-established approach, exploring such data patterns requires more investigation in future reports to elucidate the physics-based correlations learned by DL models.

Even though most of the DL methods have been applied to a structure-related component of the structure-properties relationships analysis, there are few reports associated with the actual properties' estimation constituent. In other words, synthetic images generation, structure identification, and 3D structure reconstruction were mainly focused on extracting images and structural content. In only a few applications reviewed in section 5.4, the retrieved information has been used to estimate the structure's physical properties. Coupled with this, these analyses have been limited mainly to the predictions of crystal and mechanical properties. Future contributions could be proposed to fill the gap with the actual estimation of materials' properties. For instance, a potential development of synthetic image generation could demonstrate how high-quality synthetic images lead to an improved interpretation of the observed structures and the prediction of their physical properties. Similarly, a promising direction of defects identification is represented by the prediction of the change in the physical properties of materials in the presence of the identified defects. On the other hand, the 3D reconstruction application could be extended to innovative studies for the evaluation of material properties based on the complexity of their estimated 3D geometry. Finally, physical properties like thermal, magnetic, and electric properties could be the target of DL-based predictions in forthcoming analysis. These prospective research contributions could be performed by advancing the current state-of-the-art reviewed in this paper, with findings from combined experimental measurements and numerical simulations, as well as further DL investigations. The future directions of the DL applications in microscopy imaging and materials science for the four applications reviewed in this paper are summarized in Fig. 19.

8. Conclusions

Using advanced DL techniques for direct analysis of experimentally acquired microscopy images of materials has the potential to be

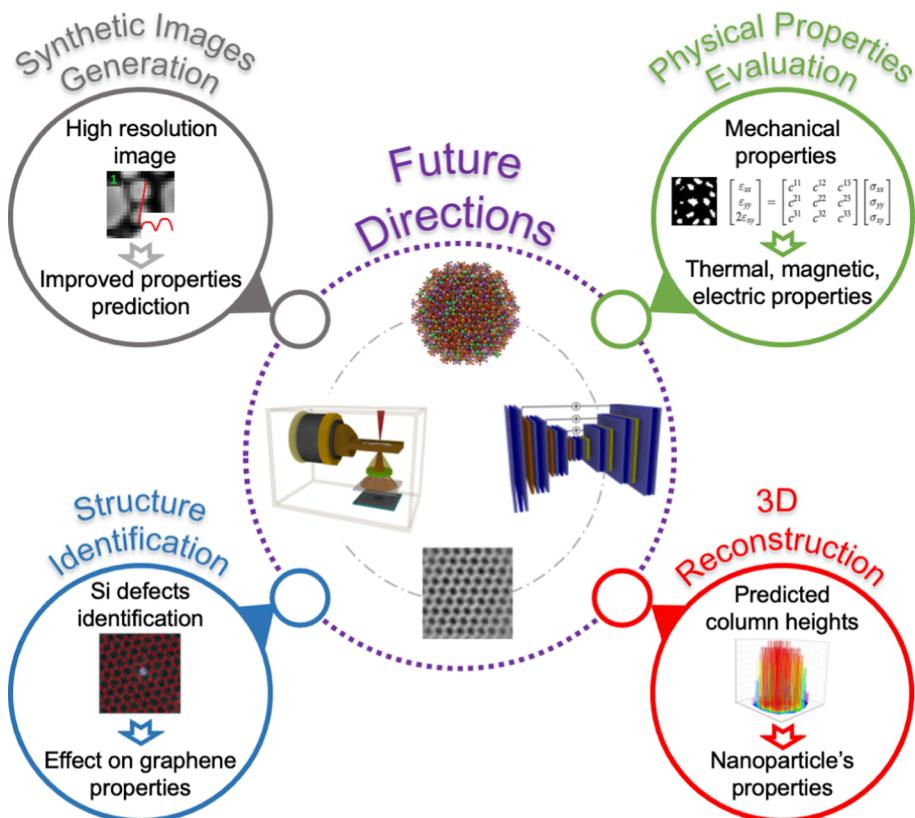


Fig. 19. Summary of future directions of DL in microscopy imaging and materials science for each of the reviewed applications. Synthetic images generation could be advanced by improving the predictions of physical properties from synthetic high-resolution images. In structure identification, a potential study is represented by the analysis of the effect of the identified Si dopants on the graphene structure. Then, the predicted NPs' CHs in the 3D reconstruction study could be used to evaluate NPs' physical properties of interest. Finally, thermal, magnetic, and electric properties are other physical properties of interest that could be predicted using DL models in future contributions.

substantially more efficient than conventional visual imaging methods by using advanced DL techniques. There is an inherent inefficiency associated with conventional visual imaging methods, such as biased by prior knowledge or analogy-driven assumptions. Direct analysis of experimental images using DL algorithms is yet less mature than the conventional visual methods. However, they certainly carry a great promise to excel traditional methods for a fast or on-the-fly analysis, interpretation, and feature extraction from experimentally acquired microscopy images.

In this review, we have summarized the DL methodologies reported in the literature to address four categories of applications in microscopy imaging and materials science: synthetic dataset generation, structure identification, 3D structures reconstruction, and physical properties evaluation. First, the description of the main DL concepts has been introduced. Then, we have presented a description of the categories of neural network architectures adopted for various tasks like materials classification, detection, and segmentation. Furthermore, we have reviewed the workflow and related challenges of the preparation of microscopy image datasets through experimental measurements or numerical simulations and how these two approaches have been followed in prior published reports. Then, the paper concentrates on the most relevant sections about the DL applications to the four identified microscopy imaging and materials science problems.

Synthetic image generation mainly focused on generative DL algorithms for resolution enhancement, denoising/refocusing, and dataset augmentation. Structure identification methods such as structure segmentation, structure detection, and structure classification were the most explored applications in the prior literature. Different types of structures and elements at different physical scales, such as atoms and defects at the atomic scale, nanoparticles at the nanoscale, and microconstituents at the microscale, were identified. On the other hand, in 3D structural reconstruction problems, DL was used for atomic column heights estimation of metallic nanoparticles, materials' thickness prediction, and analysis of 3D tomographic data. Finally, in physical properties evaluation, DL models were used to predict crystal properties, for example, scale factor, lattice parameters, chiral indices, and microstructures' mechanical properties (e.g., Young's modulus and microscopic stress field). One of the objectives of this paper is to review how DL algorithms outperformed the existing traditional methods in addressing these complicated tasks, opening innovative solutions for images generation, structure classification and detection, 3D reconstruction, and physical properties estimation.

As evidenced by the over two hundred papers cited in this review, DL algorithms have proven exceptional potential in various microscopy imaging applications in materials science, creating new paradigms in structure-properties relationship studies. Through a quantitative comparison of the performance of the developed DL frameworks with respect to conventional image processing techniques, the review elucidates how these novel approaches, based on properly trained neural networks, can offer superior capabilities in extracting the physical properties of materials from microscopy images. Despite the promising horizon depicted by DL, learning-based algorithm present their own limitations. The most common challenge in all the articles summarized in this review is the difficulty in correctly predicting out-of-domain cases, where materials or microscopy images are characterized by features not included in a training set distribution. Examples reported in this review could be the identification of atoms of different chemical species from the ones used to train a model, or the application of a neural network trained to a certain class of microscopy image, to images acquired with a different microscopy technique.

At the same time, it is also evident that there is not yet any direct automatic procedure for a fast or on-the-fly analysis of experimental microscopy images. Thus, building on the revolution of artificial intelligence and the continued advancements of microscopy techniques, experimental and computational scientists have the opportunity to work together to improve the state-of-knowledge of materials at different physical scales by exploring microscopy images with innovative data-driven methods. This growing scientific field will impact not only the way we study materials nowadays but also their design and integration into current and new technological applications.

Declaration of Competing Interest

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

Data availability

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

Acknowledgments

The authors acknowledge the financial support from the National Science Foundation, award DMR-2311104, as well as CNS-1828265.

References

- [1] Vander Voort GF. Optical Microscopy. *Encycl Condens Matter Phys* 2005;175–82. Doi:10.1016/B0-12-369401-9/00579-9.
- [2] Wiesendanger R, Roland W. *Scanning Probe Microscopy and Spectroscopy: Methods and Applications*. Cambridge University Press; 1994. p. 637.
- [3] Williams DB, Barry Carte C. *Transmission Electron Microscopy*. 2nd Ed. Springer; 2009. p. 282.
- [4] Li J, Johnson G, Zhang S, Su D. In Situ Transmission Electron Microscopy for Energy Applications. *Joule* 2019;3:4–8. <https://doi.org/10.1016/j.joule.2018.12.007>.
- [5] Pennycook S, Nellist PD. *Scanning Transmission Electron Microscopy*. 1st Ed. Springer; 2011. p. 774.

- [6] Krizhevsky A, Sutskever I, Hinton GE. ImageNet Classification with Deep Convolutional Neural Networks. *Commun ACM* 2017;60. doi:10.1145/3065386.
- [7] He K, Gkioxari G, Dollar P, Girshick R. Mask R-CNN, arXiv:1703.06870 2017.
- [8] Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu CY, et al. SSD: Single shot multibox detector. *Lect Notes Comput Sci (Including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)* 2016;9905 LNCS:21–37. doi:10.1007/978-3-319-46448-0_2.
- [9] Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: Unified, real-time object detection. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit* 2016;2016-Decem:779–88. doi:10.1109/CVPR.2016.91.
- [10] Lin T, Ai F, Doll P. Focal Loss for Dense Object Detection, arXiv:1708.02002 2017.
- [11] Long J, Shelhamer E, Darrel T. Fully Convolutional Networks for Semantic Segmentation. 2015 IEEE Conf Comput Vis. Pattern Recognit 2015;10:227–8. <https://doi.org/10.1109/CVPR.2015.7298965>.
- [12] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. ArXiv150504597 [Cs] 2015.
- [13] Madsen J, Liu P, Kling J, Wagner JB, Hansen TW, Winther O, et al. A Deep Learning Approach to Identify Local Structures in Atomic-Resolution Transmission Electron Microscopy Images. *Adv Theory Simulations* 2018;1800037. <https://doi.org/10.1002/adts.201800037>.
- [14] Ragone M, Yurkiv V, Song B, Ramsubramanian A, Shahbazian-Yassar R, Mashayek F. Atomic column heights detection in metallic nanoparticles using deep convolutional learning. *Comput Mater Sci* 2020;180:109722. <https://doi.org/10.1016/j.commatsci.2020.109722>.
- [15] Ziatdinov M, Dyck O, Maksov A, Li X, Sang X, Xiao K, et al. Deep Learning of Atomically Resolved Scanning Transmission Electron Microscopy Images: Chemical Identification and Tracking Local Transformations. *ACS Nano* 2017;11:12742–52. <https://doi.org/10.1021/acsnano.7b07504>.
- [16] Vasudevan RK, Choudhary K, Mehta A, Smith R, Kusne G, Tavazza F, et al. Materials science in the artificial intelligence age: High-throughput library generation, machine learning, and a pathway from correlations to the underpinning physics. *MRS Commun* 2019;9:821–38. <https://doi.org/10.1557/mrc.2019.95>.
- [17] Jain A, Ong SP, Hautier G, Chen W, Richards WD, Dacek S, et al. Commentary: The materials project: A materials genome approach to accelerating materials innovation. *APL Mater* 2013;1. <https://doi.org/10.1063/1.4812323>.
- [18] Curtarolo S, Setyawan W, Wang S, Xue J, Yang K, Taylor RH, et al. AFLOWLIB.ORG: A distributed materials properties repository from high-throughput ab initio calculations. *Comput Mater Sci* 2012;58:227–35. <https://doi.org/10.1016/j.commatsci.2012.02.002>.
- [19] Kim C, Chandrasekaran A, Huan TD, Das D, Ramprasad R. Polymer Genome: A Data-Powered Polymer Informatics Platform for Property Predictions. *J Phys Chem C* 2018;122:17575–85. <https://doi.org/10.1021/acs.jpcc.8b02913>.
- [20] Informatics C. Open Citrination Platform.
- [21] Materials Platform for Data Science.
- [22] Sendek AD, Cubuk ED, Antoniuk ER, Cheon G, Cui Y, Reed EJ. Machine Learning-Assisted Discovery of Solid Li-Ion Conducting Materials. *Chem Mater* 2019;31:342–52. <https://doi.org/10.1021/acs.chemmater.8b03272>.
- [23] Ahmad Z, Xie T, Maheshwari C, Grossman JC, Viswanathan V. Machine Learning Enabled Computational Screening of Inorganic Solid Electrolytes for Suppression of Dendrite Formation in Lithium Metal Anodes. *ACS Cent Sci* 2018;4:996–1006. <https://doi.org/10.1021/acscentsci.8b00229>.
- [24] Cubuk ED, Sendek AD, Reed EJ. Screening billions of candidates for solid lithium-ion conductors: A transfer learning approach for small data. *J Chem Phys* 2019;150. doi:10.1063/1.5093220.
- [25] Attarian Shandiz M, Gauvin R. Application of machine learning methods for the prediction of crystal system of cathode materials in lithium-ion batteries. *Comput Mater Sci* 2016;117:270–8. <https://doi.org/10.1016/j.commatsci.2016.02.021>.
- [26] Allam O, Cho BW, Kim KC, Jiang SS. Application of DFT-based machine learning for developing molecular electrode materials in Li-ion batteries. *RSC Adv* 2018;8:39414–20. <https://doi.org/10.1039/c8ra0712h>.
- [27] Okamoto Y, Kubo Y. Ab Initio Calculations of the Redox Potentials of Additives for Lithium-Ion Batteries and Their Prediction through Machine Learning. *ACS Omega* 2018;3:7868–74. <https://doi.org/10.1021/acsomega.8b00576>.
- [28] Zhang J, Liu X, Bi S, Yin J, Zhang G, Eisenbach M. Robust data-driven approach for predicting the configurational energy of high entropy alloys. *Mater Des* 2020;185:108247. <https://doi.org/10.1016/j.matdes.2019.108247>.
- [29] Lee SY, Byeon S, Kim HS, Jin H, Lee S. Deep learning-based phase prediction of high-entropy alloys: Optimization, generation, and explanation. *Mater Des* 2021;197:109260. <https://doi.org/10.1016/j.matdes.2020.109260>.
- [30] Pei Z, Yin J, Hawk JA, Alman DE, Gao MC. Machine-learning informed prediction of high-entropy solid solution formation: Beyond the Hume-Rothery rules. *Npj Comput Mater* 2020;6. <https://doi.org/10.1038/s41524-020-0308-7>.
- [31] Zhou Z, Zhou Y, He Q, Ding Z, Li F, Yang Y. Machine learning guided appraisal and exploration of phase design for high entropy alloys. *Npj Comput Mater* 2019;5:1–9. <https://doi.org/10.1038/s41524-019-0265-1>.
- [32] Kaufmann K, Vecchio KS. Searching for high entropy alloys: A machine learning approach. *Acta Mater* 2020;198:178–222. <https://doi.org/10.1016/j.actamat.2020.07.065>.
- [33] Dai FZ, Wen B, Sun Y, Xiang H, Zhou Y. Theoretical prediction on thermal and mechanical properties of high entropy (Zr0.2Hf0.2Ti0.2Nb0.2Ta0.2)C by deep learning potential. *J Mater Sci Technol* 2020;43:168–74. <https://doi.org/10.1016/j.jmst.2020.01.005>.
- [34] Ramprasad R, Batra R, Pilania G, Manodi-Kanakkithodi A, Kim C. Machine learning in materials informatics: Recent applications and prospects. *Npj Comput Mater* 2017;3. doi:10.1038/s41524-017-0056-5.
- [35] Agrawal A, Choudhary A. Deep materials informatics: Applications of deep learning in materials science. *MRS Commun* 2019;9:779–92. <https://doi.org/10.1557/mrc.2019.73>.
- [36] Rickman JM, Lookman T, Kalinin SV. Materials informatics: From the atomic-level to the continuum. *Acta Mater* 2019;168:473–510. <https://doi.org/10.1016/j.actamat.2019.01.051>.
- [37] Butler KT, Davies DW, Cartwright H, Isayev O, Walsh A. Machine learning for molecular and materials science. *Nature* 2018;559:547–55. <https://doi.org/10.1038/s41586-018-0337-2>.
- [38] Schmidt J, Marques MRG, Botti S, Marques MAL. Recent advances and applications of machine learning in solid-state materials science. *Npj Comput Mater* 2019;5. doi:10.1038/s41524-019-0221-0.
- [39] Wei J, Chu X, Sun X, Xu K, Deng H, Chen J, et al. Machine learning in materials science InfoMat 2019;1:338–58. <https://doi.org/10.1002/inf2.12028>.
- [40] Pilania G. Machine learning in materials science: From explainable predictions to autonomous design. *Comput Mater Sci* 2021;193:110360. <https://doi.org/10.1016/j.commatsci.2021.110360>.
- [41] Kalinin SV, Lupini AR, Dyck O, Jesse S, Ziatdinov M, Vasudevan RK. Lab on a beam-Big data and artificial intelligence in scanning transmission electron microscopy. *MRS Bull* 2019;44:565–75. <https://doi.org/10.1557/mrs.2019.159>.
- [42] Ge M, Su F, Zhao Z, Su D. Deep learning analysis on microscopic imaging in materials science. *Mater Today Nano* 2020;11. doi:10.1016/j.mtnano.2020.100087.
- [43] Azuri I, Rosenhek-Goldman I, Regev-Rudzki N, Fantner G, Cohen SR. The role of convolutionsl neural networks in scanning probe microscopy: a review. *Beilstein J Nanotechnol* 2021;12:878–901. <https://doi.org/10.3762/bjnano.12.66>.
- [44] Xu H, Zhu J, Finegan DP, Zhao H, Lu X, Li W, et al. Guiding the Design of Heterogeneous Electrode Microstructures for Li-Ion Batteries: Microscopic Imaging, Predictive Modeling, and Machine Learning. *Adv Energy Mater* 2021;11:1–34. <https://doi.org/10.1002/aenm.202003908>.
- [45] Zheng H, Lu X, He K. In situ transmission electron microscopy and artificial intelligence enabled data analytics for energy materials. *J Energy Chem* 2021. <https://doi.org/10.1016/j.jecchem.2021.12.001>.
- [46] Lecun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;521:436–44. <https://doi.org/10.1038/nature14539>.
- [47] Agarap AF. Deep Learning using Rectified Linear Units (ReLU). *ArXiv180308375 [Cs, Stat]* 2018.
- [48] Nwankpa C, Ijomah W, Gachagan A, Marshall S. Activation Functions: Comparison of trends in Practice and Research for Deep Learning, arXiv:1811.03378 2018;1–20.
- [49] Albawi S, Mohammed TAM, Alzawi S. Understanding of a Convolutional Neural Network. IEEE 2017.

- [50] Gu J, Wang Z, Kuen J, Ma L, Shahroudy A, Shuai B, et al. Recent advances in convolutional neural network. ArXiv 2017. <https://doi.org/10.1101/j.neucom.2018.09.038>.
- [51] Yamashita R, Nishio M, Kinh Gian Do R, Togashi K. Convolutional neural networks: an overview and application in radiology. Insights. Imaging 2018;9: 611–29. https://doi.org/10.1007/978-981-15-7078-0_3.
- [52] Sherstinsky A. Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network. Phys D Nonlinear Phenom 2020;404:1–43. <https://doi.org/10.1016/j.physd.2019.132306>.
- [53] Scherer D, Müller A, Behnke S. Evaluation of pooling operations in convolutional architectures for object recognition. Lect Notes Comput Sci (Including Subser Lect Notes Artif Intell Lect Notes Bioinformatics) 2010;6354 LNCS:92–101. Doi:10.1007/978-3-642-15825-4_10.
- [54] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. 32nd Int Conf Mach Learn ICML 2015;1: 448–56.
- [55] Janocha K, Czarnecki WM. On loss functions for deep neural networks in classification. Schedae Informaticae 2016;25:49–59. <https://doi.org/10.4467/20838476SL16.004.6185>.
- [56] Ruder S. An overview of gradient descent optimization algorithms. ArXiv160904747 [Cs] 2016.
- [57] Kingma DP, Ba JL. Adam: A method for stochastic optimization. 3rd Int Conf Learn Represent ICLR 2015 - Conf Track Proc 2015:1–15.
- [58] Ribani R, Marengoni M, Mackenzie UP. A Survey of Transfer Learning for Convolutional Neural Networks. 2019 32nd SIBGRAPI Conf Graph Patterns Images Tutorials 2019:47–57. Doi:10.1109/SIBGRAPI-T.2019.00010.
- [59] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. 3rd Int Conf Learn Represent ICLR 2015 - Conf Track Proc 2015:1–14.
- [60] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, et al. Going deeper with convolutions. Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit 2015;07–12-June:1–9. Doi:10.1109/CVPR.2015.7298594.
- [61] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit 2016;2016-Decem:770–8. Doi:10.1109/CVPR.2016.90.
- [62] Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. Proc - 30th IEEE Conf Comput Vis Pattern Recognition, CVPR 2017 2017;2017-Janua:2261–9. Doi:10.1109/CVPR.2017.243.
- [63] Girshick R, Donahue J, Darrell T, Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation. Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit 2014;2015 Inter:580–7. Doi:10.1109/CVPR.2014.81.
- [64] Girshick R. Fast R-CNN. Proc IEEE Int Conf Comput Vis 2015;2015 Inter:1440–8. Doi:10.1109/ICCV.2015.169.
- [65] Ren S, He K, Girshick R, Sun J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. IEEE Trans Pattern Anal Mach Intell 2017;39:1137–49. <https://doi.org/10.1109/TPAMI.2016.2577031>.
- [66] Ying X. An Overview of Overfitting and its Solutions. J Phys Conf Ser 2019;1168. <https://doi.org/10.1088/1742-6596/1168/2/022022>.
- [67] Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, Corrado G, Davis A, Dean J, Devin M, Ghemawat S, Goodfellow I, Harp A, Irving G, Isard M, Jia Y, Jozefowicz R, Kaiser L, Kudlur M, Levenberg J, Mane D, Monga R, Moore S, Murray D, Olah C, Schuster M, Shlens J, Steiner B, Sutskever I, Talwar K, Tucker P, Vanhoucke V, Vasudevan V, Viegas F, Vinyals O, Warden P, Wattenberg M, Wicke M, Yu Y, Zheng X. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. ArXiv:1603.04467 2016.
- [68] Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. PyTorch: An imperative style, high-performance deep learning library. Adv Neural Inf Process Syst 2019;32.
- [69] Chen T, Li M, Li Y, Lin M, Wang N, Wang M, et al. MXNet: A Flexible and Efficient Machine Learning Library for Heterogeneous Distributed Systems 2015:1–6.
- [70] Jia Y, Shelhamer E, Donahue J, Karayev S, Long J, Girshick R, Guadarrama S, Darrel T. Caffe : Convolutional Architecture for Fast Feature Embedding, arXiv: 1408.5093 2014.
- [71] Chetlur S, Woolley C, Vandermersch P, Cohen J, Tran J, Catanzaro B, Shelhamer E. cuDNN : Efficient Primitives for Deep Learning, arXiv:1410.0759 2014.
- [72] Amazon Web Service.
- [73] Google Cloud Platform.
- [74] Jia Z, Zaharia M, Aiken A. Beyond data and model parallelism for deep neural networks, arXiv:180705358 2018.
- [75] Shallue CJ, Dahl GE. Measuring the Effects of Data Parallelism on Neural Network Training, arXiv1811.03600 2019;20:1–49.
- [76] Sergeev A, Balso M Del. Horovod : fast and easy distributed deep learning in TensorFlow, arXiv:1802.05799 2018.
- [77] Narang S, Diamos G, Elsen E, Micikevicius P, Alben J, Garcia D, Ginsburg B, Houston M, Kuchaiev O, Venkatesh G, Wu H, Mixed Precision Training, arXiv: 1710.03740 2018:1–12.
- [78] Tensorflow: XLA.
- [79] Dong H, Butler KT, Matras D, Price SWT, Odarchenko Y, Khatry R, et al. A deep convolutional neural network for real-time full profile analysis of big powder diffraction data. Npj Comput Mater 2021;7:1–9. <https://doi.org/10.1038/s41524-021-00542-4>.
- [80] Kaufmann K, Zhu C, Rosengarten AS, Maryanovsky D, Harrington TJ, Marin E, et al. Crystal symmetry determination in electron diffraction using machine learning 2020;367:564–8.
- [81] DeCost BL, Lei B, Francis T, Holm EA. High throughput quantitative metallography for complex microstructures using deep learning: A case study in ultrahigh carbon steel. Microsc Microanal 2019;25:21–9. <https://doi.org/10.1017/S1431927618015635>.
- [82] Masubuchi S, Watanabe E, Seo Y, Okazaki S, Sasagawa T, Watanabe K, et al. Deep-learning-based image segmentation integrated with optical microscopy for automatically searching for two-dimensional materials. Npj 2D Mater Appl 2020;4:4–6. <https://doi.org/10.1038/s41699-020-0137-z>.
- [83] Arsalan M, Kim DS, Lee MB, Owais M, Park KR. FRED-Net: Fully residual encoder–decoder network for accurate iris segmentation. Expert Syst Appl 2019;122: 217–41. <https://doi.org/10.1016/j.eswa.2019.01.010>.
- [84] Roberts G, Haile SY, Sainju R, Edwards DJ, Hutchinson B, Zhu Y. Deep Learning for Semantic Segmentation of Defects in Advanced STEM Images of Steels. Sci Rep 2019;9:1–12. <https://doi.org/10.1038/s41598-019-49105-0>.
- [85] Azimi SM, Britz D, Engstler M, Fritz M, Mücklich F. Advanced steel microstructural classification by deep learning methods. Sci Rep 2018;8:1–14. <https://doi.org/10.1038/s41598-018-20037-5>.
- [86] Lee J, Lee Y, Kim J, Lee Z. Contrast transfer function-based exit-wave reconstruction and denoising of atomic-resolution transmission electron microscopy images of graphene and cu single atom substitutions by deep learning framework. Nanomaterials 2020;10:1–11. <https://doi.org/10.3390/nano10101977>.
- [87] Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative Adversarial Nets. Adv Neural Inf Process Syst 2014;3:2672–80. https://doi.org/10.3156/jsoft.29.5_177_2.
- [88] Mirza M, Osindero S. Conditional Generative Adversarial Nets. ArXiv:14111784 2014:1–7.
- [89] Zhang H, Xu T, Li H, Zhang S, Wang X, Huang X, et al. StackGAN++: Realistic Image Synthesis with Stacked Generative Adversarial Networks. IEEE Trans Pattern Anal Mach Intell 2017;41:1947–62. <https://doi.org/10.1109/TPAMI.2018.2856256>.
- [90] Zhu J, Park T, Isola P, Efros A. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. ArXiv:170310593 2020.
- [91] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. 4th Int Conf Learn Represent ICLR 2016 - Conf Track Proc 2016:1–16.
- [92] Allen-Zhu Z, Li Y. Forward Super-Resolution: How Can GANs Learn Hierarchical Generative Models for Real-World Distributions, arXiv:2106.02619 2021.
- [93] Zhong Y, Liu L, Zhao D, Li H. A generative adversarial network for image denoising. Multimed Tools Appl 2020;79:16517–29. <https://doi.org/10.1007/s11042-019-7556-x>.
- [94] Wu J, Zhang C, Xue T, Freeman WT, Tenenbaum JB. Learning a probabilistic latent space of object shapes via 3D generative-adversarial modeling. Adv Neural Inf Process Syst 2016:82–90.
- [95] Gadelha M, Maji S, Wang R. 3D shape induction from 2D views of multiple objects. Proc - 2017 Int Conf 3D Vision, 3DV 2017 2018:402–11. Doi:10.1109/3DV.2017.00053.

- [96] de Haan K, Ballard ZS, Rivenson Y, Wu Y, Ozcan A. Resolution enhancement in scanning electron microscopy using deep learning. *Sci Rep* 2019;9:1–7. <https://doi.org/10.1038/s41598-019-48444-2>.
- [97] Wang F, Henninen TR, Keller D, Erni R. Noise2Atom: unsupervised denoising for scanning transmission electron microscopy images. *Appl Microsc* 2020;50. <https://doi.org/10.1186/s42649-020-00041-8>.
- [98] Hsu T, Epting WK, Kim H, Abernathy HW, Hackett GA, Rollett AD, et al. Microstructure Generation via Generative Adversarial Network for Heterogeneous, Topologically Complex 3D Materials. *JOM* 2021;73:90–102. <https://doi.org/10.1007/s11837-020-04484-y>.
- [99] DeCost BL, Francis T, Holm EA. Exploring the microstructure manifold: Image texture representations applied to ultrahigh carbon steel microstructures. *Acta Mater* 2017;133:30–40. <https://doi.org/10.1016/j.actamat.2017.05.014>.
- [100] Luo Q, Holm EA, Wang C. A transfer learning approach for improved classification of carbon nanomaterials from TEM images. *Nanoscale Adv* 2020. <https://doi.org/10.1039/d0na00634c>.
- [101] Zhang C, Feng J, DaCosta LR, Voyles PM. Atomic resolution convergent beam electron diffraction analysis using convolutional neural networks. *Ultramicroscopy* 2020;210:112921. <https://doi.org/10.1016/j.ultramic.2019.112921>.
- [102] Aversa R, Modarres MH, Brandino GP, Cozzini S, Leto A, Ciancio R. Neural Network for Nanoscience Scanning Electron Microscope Image Recognition. *Sci Rep* 2017;7:1–12. <https://doi.org/10.1038/s41598-017-13565-z>.
- [103] Okunev AG, Nartova AV, Matveev AV. Recognition of nanoparticles on scanning probe microscopy images using computer vision and deep machine learning. *Sib* 2019 - Int Multi-Conference Eng Comput Inf Sci Proc 2019:940–3. Doi:10.1109/SIBIRCON48586.2019.8958363.
- [104] Okunev AG, Mashukov MY, Nartova AV, Matveev AV. Nanoparticle recognition on scanning probe microscopy images using computer vision and deep learning. *Nanomaterials* 2020;10:1–16. <https://doi.org/10.3390/nano10071285>.
- [105] Zhang F, Zhang Q, Xiao Z, Wu J, Liu J, Liu Y. Spherical nanoparticle parameter measurement method based on mask r-cnn segmentation and edge fitting. *Assoc Comput Mach* 2019;219:205–202. <https://doi.org/10.1145/3373509.3373590>.
- [106] Jin P, Li X. Correction of image drift and distortion in a scanning electron microscope. *J Microsc* 2015;260:268–80. <https://doi.org/10.1111/jmi.12293>.
- [107] DeCost BL, Hecht MD, Francis T, Webler BA, Picard YN, Holm EA. UHCSDB: UltraHigh Carbon Steel Micrograph DataBase: Tools for Exploring Large Heterogeneous Microstructure Datasets. *Integr Mater Manuf Innov* 2017;6:197–205. <https://doi.org/10.1007/s40192-017-0097-0>.
- [108] Hecht MD, Webler BA, Picard YN. Digital image analysis to quantify carbide networks in ultrahigh carbon steels. *Mater Charact* 2016;117:134–43. <https://doi.org/10.1016/j.matchar.2016.04.012>.
- [109] Hecht MD, Picard YN, Webler BA. Coarsening of Inter- and Intra-granular Proeutectoid Cementite in an Initially Pearlitic 2C–4Cr Ultrahigh Carbon Steel. *Metall Mater Trans A Phys Metall Mater Sci* 2017;48:2320–35. <https://doi.org/10.1007/s11661-017-4012-2>.
- [110] Gola J, Britz D, Staudt T, Winter M, Schneider AS, Ludovic M, et al. Advanced microstructure classification by data mining methods. *Comput Mater Sci* 2018;148:324–35. <https://doi.org/10.1016/j.commatsci.2018.03.004>.
- [111] CNR-IOM Dataset.
- [112] Shorten C, Khoshgoftaar TM. A survey on Image Data Augmentation for Deep Learning. *J Big Data* 2019;6. Doi:10.1186/s40537-019-0197-0.
- [113] Badmos O, Kopp A, Bernthaler T, Schneider G. Image-based defect detection in lithium-ion battery electrode using convolutional neural networks. *J Intell Manuf* 2020;31:885–97. <https://doi.org/10.1007/s10845-019-01484-x>.
- [114] Yi K, Wu J. Probabilistic End-to-end Noise Correction for Learning with Noisy Labels, arXiv:1903.07788 2019.
- [115] Pleiss G, Elenberg E. Identifying Mislabeled Data using the Area Under the Margin Ranking, arXiv:2001.10528 2020.
- [116] Van Aert S, Verbeeck J, Erni R, Bals S, Luysberg M, Van DD, et al. Quantitative atomic resolution mapping using high-angle annular dark field scanning transmission electron microscopy. *Ultramicroscopy* 2009;109:1236–44. <https://doi.org/10.1016/j.ultramic.2009.05.010>.
- [117] Du H, Jia CL, Koehl A, Barthel J, Dittmann R, Waser R, et al. Nanosized Conducting Filaments Formed by Atomic-Scale Defects in Redox-Based Resistive Switching Memories. *Chem Mater* 2017;29:3164–73. <https://doi.org/10.1021/acs.chemmater.7b00220>.
- [118] Rosenauer A, Gries K, Müller K, Pretorius A, Schowalter M, Avramescu A, et al. Measurement of specimen thickness and composition in Al_x Ga_{1-x}N/GaN using high-angle annular dark field images. *Ultramicroscopy* 2009;109:1171–82. <https://doi.org/10.1016/j.ultramic.2009.05.003>.
- [119] Fitting L, Thiel S, Schmehl A, Mannhart J, Muller DA. Subtleties in ADF imaging and spatially resolved EELS: A case study of low-angle twist boundaries in SrTiO₃. *Ultramicroscopy* 2006;106:1053–61. <https://doi.org/10.1016/j.ultramic.2006.04.019>.
- [120] Erni R, Heinrich H, Kostorz G. Quantitative characterisation of chemical inhomogeneities in Al-Ag using high-resolution Z-contrast STEM. *Ultramicroscopy* 2003;94:125–33. [https://doi.org/10.1016/S0304-3991\(02\)00249-8](https://doi.org/10.1016/S0304-3991(02)00249-8).
- [121] Kirkland EJ. *Advanced Computing in Electron Microscopy*. 2nd Ed. 2010.
- [122] Germer TA, Marx E. Simulations of optical microscope images. *Metrol Insp Process Control Microlithogr XX* 2006;6152:615201. <https://doi.org/10.1117/12.656370>.
- [123] Kirkland EJ. Computation in electron microscopy. *Acta Crystallogr Sect A Found Adv* 2016;72:1–27. <https://doi.org/10.1107/S205327331501757X>.
- [124] Zhong X, Gallagher B, Eves K, Robertson E, Mundhenk TN, Han TYJ. A study of real-world micrograph data quality and machine learning model robustness. *Npj Comput Mater* 2021;7:1–11. <https://doi.org/10.1038/s41524-021-00616-3>.
- [125] Kohn W, Becke AD, Parr RG. Density functional theory of electronic structure. *J Phys Chem* 1996;100:12974–80. <https://doi.org/10.1021/jp9606691>.
- [126] Hospital A, Goñi JR, Orozco M, Gelpí JL. Molecular dynamics simulations: Advances and applications. *Adv Appl Bioinforma Chem* 2015;8:37–47. <https://doi.org/10.2147/AABC.S70333>.
- [127] Lee CH, Khan A, Luo D, Santos TP, Shi C, Janicek BE, et al. Deep learning enabled strain mapping of single-atom defects in two-dimensional transition metal dichalcogenides with sub-picometer precision. *Nano Lett* 2020;20:3369–77. <https://doi.org/10.1021/acs.nanolett.0c00269>.
- [128] Yao L, Ou Z, Luo B, Xu C, Chen Q. Machine Learning to Reveal Nanoparticle Dynamics from Liquid-Phase TEM Videos. *ACS Cent Sci* 2020;6:1421–30. <https://doi.org/10.1021/acscentsci.0c00430>.
- [129] Förster GD, Castan A, Loiseau A, Nelayah J, Alloyeau D, Fossard F, et al. A deep learning approach for determining the chiral indices of carbon nanotubes from high-resolution transmission electron microscopy images. *Carbon N Y* 2020;169:465–74. <https://doi.org/10.1016/j.carbon.2020.06.086>.
- [130] Ziatdinov M, Maksov A, Kalinin SV. Learning surface molecular structures via machine vision. *Npj Comput Mater* 2017;3:1–9. <https://doi.org/10.1038/s41524-017-0038-7>.
- [131] Yeom J, Stan T, Hong S, Voorhees PW. Segmentation of experimental datasets via convolutional neural networks trained on phase field simulations. *Acta Mater* 2021;214:116990. <https://doi.org/10.1016/j.actamat.2021.116990>.
- [132] Tiang LCO, Kim J, Han SS, Kim D. Identification of crystal symmetry from noisy diffraction patterns by a shape analysis and deep learning. *Npj Comput Mater* 2020;6:1–11. <https://doi.org/10.1038/s41524-020-00466-5>.
- [133] Xu W, LeBeau JM. A deep convolutional neural network to analyze position averaged convergent beam electron diffraction patterns. *Ultramicroscopy* 2018;188:59–69. <https://doi.org/10.1016/j.ultramic.2018.03.004>.
- [134] Li Y, Zhou X, Colnaghi T, Wei Y, Marek A, Li H, et al. Convolutional neural network-assisted recognition of nanoscale L12 ordered structures in face-centred cubic alloys. *Npj Comput Mater* 2021;7:1–9. <https://doi.org/10.1038/s41524-020-00472-7>.
- [135] Kresse G, Furthmüller J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys Rev B - Condens Matter Mater Phys* 1996;54:11169–86. <https://doi.org/10.1103/PhysRevB.54.11169>.
- [136] Atomistix Toolkit version 2017.2., Synopsys QuantumWise A/S.
- [137] Thompson AP, Aktulga HM, Berger R, Bolintineanu DS, Brown WM, Crozier PS, et al. LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Comput Phys Commun* 2022;271:108171. <https://doi.org/10.1016/j.cpc.2021.108171>.
- [138] Larsen AH, Jens Jørgen Mortensen JB, Castelli IE, Christensen R, Marcin Dulak JF, Groves MN, et al. *J Phys Condens Matter* 2017;Vol.;29 273002.
- [139] Kirkland EJ. *Advanced Computing in Electron Microscopy*. Springer; 2010.
- [140] QSTEM: Quantitative TEM/STEM Simulations 2017.
- [141] Barthel JD. Probe: A software for high-resolution STEM image simulation. *Ultramicroscopy* 2018;193:1–11. <https://doi.org/10.1016/j.ultramic.2018.06.003>.

- [142] Ophus C. A fast image simulation algorithm for scanning transmission electron microscopy. *Adv Struct Chem Imaging* 2017;3:1–11. <https://doi.org/10.1186/s40679-017-0046-1>.
- [143] Koch C. PyQSTEM. PhD Thesis ASU 2002.
- [144] Sreehari S, Venkatakrishnan S V, Bouman KL, Simmons JP, Drummy LF, Bouman CA. Multi-Resolution Data Fusion for Super-Resolution Electron Microscopy, 2017 IEEE Proceed 2160-7516/17.
- [145] Roels J, Aelterman J, De Vylder J, Luong H, Saeyns Y, Philips W. Bayesian deconvolution of scanning electron microscopy images using point-spread function estimation and non-local regularization. *Proc Annu Int Conf IEEE Eng Med Biol Soc EMBS* 2016;2016-Octob:443–7. Doi:10.1109/EMBC.2016.7590735.
- [146] Jiang Y, Chen Z, Han Y, Deb P, Gao H, Xie S, et al. Electron ptychography of 2D materials to deep sub-ångström resolution. *Nature* 2018;559:343–9. <https://doi.org/10.1038/s41586-018-0298-5>.
- [147] Dong C, Loy CC, He K, Tang X. Image Super-Resolution Using Deep Convolutional Networks. *IEEE Trans Pattern Anal Mach Intell* 2016;38:295–307. <https://doi.org/10.1109/TPAMI.2015.2439281>.
- [148] Kushwhala HS, Tanwar S, Rathore KS, Srivastava S. De-noising filters for TEM (Transmission Electron Microscopy) image of nanomaterials. *Proc - 2012 2nd Int Conf Adv Comput Commun Technol ACCT* 2012 2012:276–81. Doi:10.1109/ACCT.2012.41.
- [149] Pantelis RS, Rothnagel R, Huang CY, Muller D, Woolford D, Landsberg MJ, et al. The discriminative bilateral filter: An enhanced denoising filter for electron microscopy data. *J Struct Biol* 2006;155:395–408. <https://doi.org/10.1016/j.jsb.2006.03.030>.
- [150] Wei DY, Yin CC. An optimized locally adaptive non-local means denoising filter for cryo-electron microscopy data. *J Struct Biol* 2010;172:211–8. <https://doi.org/10.1016/j.jsb.2010.06.021>.
- [151] Aharon M, Elad M, Bruckstein A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans Signal Process* 2006;54:4311–22. <https://doi.org/10.1109/TSP.2006.881199>.
- [152] Ahmed SS, Messali Z, Ouahabi A, Trepout S, Messaoudi C, Marco S. Nonparametric denoising methods based on contourlet transform with sharp frequency localization: Application to low exposure time electron microscopy images. *Entropy* 2015;17:3461–78. <https://doi.org/10.3390/e17053461>.
- [153] Ahumada AJ. Computational image quality metrics: A review. *SID Dig* 1993;24:305–8.
- [154] Sun Y, Duthaler S, Nelson BJ. Autofocusing in computer microscopy: Selecting the optimal focus algorithm. *Microsc Res Tech* 2004;65:139–49. <https://doi.org/10.1002/jemt.20118>.
- [155] Kudryavtsev AV, Dembélé S, Piat N. Autofocus on moving object in scanning electron microscope. *Ultramicroscopy* 2017;182:216–25. <https://doi.org/10.1016/j.ultramic.2017.07.008>.
- [156] Sun J, Cao W, Xu Z, Ponce J. Learning a convolutional neural network for non-uniform motion blur removal. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit* 2015;07-12-June:769–77. Doi:10.1109/CVPR.2015.7298677.
- [157] Lee W, Nam HS, Kim YG, Kim YJ, Lee JH, Yoo H. Robust autofocusing for scanning electron microscopy based on a dual deep learning network. *Sci Rep* 2021;11:1–12. <https://doi.org/10.1038/s41598-021-00412-5>.
- [158] Na J, Kim G, Kang SH, Kim SJ, Lee S. Deep learning-based discriminative refocusing of scanning electron microscopy images for materials science. *Acta Mater* 2021;214. <https://doi.org/10.1016/j.actamat.2021.116987>.
- [159] Chen J, Benesty J, Huang Y, Doclo S. New insights into the noise reduction Wiener filter. *IEEE Trans Audio, Speech Lang Process* 2006;14:1218–33. <https://doi.org/10.1109/TSA.2005.860851>.
- [160] Bostanabad R, Bui AT, Xie W, Apley DW, Chen W. Stochastic microstructure characterization and reconstruction via supervised learning. *Acta Mater* 2016;103:89–102. <https://doi.org/10.1016/j.actamat.2015.09.044>.
- [161] Song X, He L, Yang W, Wang Z, Chen Z, Guo J, et al. Additive manufacturing of bi-continuous piezocomposites with triply periodic phase interfaces for combined flexibility and piezoelectricity. *J Manuf Sci Eng Trans ASME* 2019;141. <https://doi.org/10.1115/1.4044708>.
- [162] Cang R, Li H, Yao H, Jiao Y, Ren Y. Improving direct physical properties prediction of heterogeneous materials from imaging data via convolutional neural network and a morphology-aware generative model. *Comput Mater Sci* 2018;150:212–21. <https://doi.org/10.1016/j.commatsci.2018.03.074>.
- [163] Li X, Zhang Y, Zhao H, Burkhardt C, Brinson LC, Chen W. A Transfer Learning Approach for Microstructure Reconstruction and Structure-property Predictions. *Sci Rep* 2018;8:1–13. <https://doi.org/10.1038/s41598-018-31571-7>.
- [164] Gayon-Lombardo A, Mosser L, Brandon NP, Cooper SJ. Pores for thought: generative adversarial networks for stochastic reconstruction of 3D multi-phase electrode microstructures with periodic boundaries. *Npj Comput Mater* 2020;6:1–11. <https://doi.org/10.1038/s41524-020-0340-7>.
- [165] Chun S, Roy S, Nguyen YT, Choi JB, Udaykumar HS, Baek SS. Deep learning for synthetic microstructure generation in a materials-by-design framework for heterogeneous energetic materials. *Sci Rep* 2020;10:1–15. <https://doi.org/10.1038/s41598-020-70149-0>.
- [166] Yang W, Wang Z, Yang T, He L, Song X, Liu Y, et al. Exploration of the Underlying Space in Microscopic Images via Deep Learning for Additively Manufactured Piezoceramics. *ACS Appl Mater Interfaces* 2021. <https://doi.org/10.1021/acsami.1c12945>.
- [167] Ma B, Wei X, Liu C, Ban X, Huang H, Wang H, et al. Data augmentation in microscopic images for material data mining. *Npj Comput Mater* 2020;6. <https://doi.org/10.1038/s41524-020-00392-6>.
- [168] Hujasak K, Myers BD, Roth E, Li Y, David VP. Suppressing electron exposure artifacts: An electron scanning paradigm with Bayesian machine learning. *Microsc Microanal* 2016;22:778–88. <https://doi.org/10.1017/S1431927616011417>.
- [169] Candès E, Romberg J. Sparsity and incoherence in compressive sampling. *Inverse Probl* 2007;23:969–85. <https://doi.org/10.1088/0266-5611/23/3/008>.
- [170] Sang X, Lupini AR, Unocic RR, Chi M, Borisevich AY, Kalinin SV, et al. Dynamic scan control in STEM: spiral scans. *Adv Struct Chem. Imaging* 2016;2. <https://doi.org/10.1186/s40679-016-0020-3>.
- [171] Stevens A, Luzzi L, Yang H, Kovarik L, Mehdi BL, Liyu A, et al. A sub-sampled approach to extremely low-dose STEM. *Appl Phys Lett* 2018;112. <https://doi.org/10.1063/1.5016192>.
- [172] Wu X, Li RL, Zhang FL, Liu JC, Wang J, Shamir A, et al. Deep Portrait Image Completion and Extrapolation. *IEEE Trans Image Process* 2020;29:2344–55. <https://doi.org/10.1109/TIP.2019.2945866>.
- [173] Liu G, Reda FA, Shih KJ, Wang TC, Tao A, Catanzaro B. Image Inpainting for Irregular Holes Using Partial Convolutions. *Lect Notes Comput Sci (Including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)* 2018;11215 LNCS:89–105. Doi:10.1007/978-3-030-01252-6_6.
- [174] Ede JM, Beanland R. Partial Scanning Transmission Electron Microscopy with Deep Learning. *Sci Rep* 2020;10:1–10. <https://doi.org/10.1038/s41598-020-65261-0>.
- [175] Maksov A, Dyck O, Wang K, Xiao K, Geohegan DB, Sumpter BG, et al. Deep learning analysis of defect and phase evolution during electron beam-induced transformations in WS₂. *Npj Comput Mater* 2019;5:1–8. <https://doi.org/10.1038/s41524-019-0152-9>.
- [176] Ma B, Ban X, Huang H, Chen Y, Liu W, Zhi Y. Deep learning-based image segmentation for Al-La alloy microscopic images. *Symmetry (Basel)* 2018;10:1–13. <https://doi.org/10.3390/sym10040107>.
- [177] Horwath JP, Zakharov DN, Mégré R, Stach EA. Understanding important features of deep learning models for segmentation of high-resolution transmission electron microscopy images. *Npj Comput Mater* 2020;6:1–9. <https://doi.org/10.1038/s41524-020-00363-x>.
- [178] Müller S, Sauter C, Shunmugasundaram R, Wenzler N, De Andrade V, De Carlo F, et al. Deep learning-based segmentation of lithium-ion battery microstructures enhanced by artificially generated electrodes. *Nat Commun* 2021;12:1–12. <https://doi.org/10.1038/s41467-021-26480-9>.
- [179] Oktay AB, Gurses A. Automatic detection, localization and segmentation of nano-particles with deep learning in microscopy images. *Micron* 2019;120:113–9. <https://doi.org/10.1016/j.micron.2019.02.009>.
- [180] Vasudevan RK, Laanait N, Ferragut EM, Wang K, Geohegan DB, Xiao K, et al. Mapping mesoscopic phase evolution during E-beam induced transformations via deep learning of atomically resolved images. *Npj Comput Mater* 2018;4. <https://doi.org/10.1038/s41524-018-0086-7>.
- [181] Ziletti A, Kumar D, Scheffler M, Ghiringhelli LM. Insightful classification of crystal structures using deep learning. *Nat Commun* 2018;9:1–10. <https://doi.org/10.1038/s41467-018-05169-6>.
- [182] Aguiar JA, Gong ML, Unocic RR, Tasdizen T, Miller BD. Decoding crystallography from high-resolution electron imaging and diffraction datasets with deep learning. *Sci Adv* 2019;5:1–10. <https://doi.org/10.1126/sciadv.aaw1949>.

- [183] Cheon S, Lee H, Kim CO, Lee SH. Convolutional Neural Network for Wafer Surface Defect Classification and the Detection of Unknown Defect Class. *IEEE Trans Semicond Manuf* 2019;32:163–70. <https://doi.org/10.1109/TS.M.2019.2902657>.
- [184] Imoto K, Nakai T, Ike T, Haruki K, Sato Y. A CNN-Based transfer learning method for defect classification in semiconductor manufacturing. *IEEE Trans Semicond Manuf* 2019;32:455–9. <https://doi.org/10.1109/TS.M.2019.2941752>.
- [185] Chollet F. Xception: Deep learning with depthwise separable convolutions. *Proc - 30th IEEE Conf Comput Vis Pattern Recognition, CVPR 2017 2017;2017-Janua:1800–7. Doi:10.1109/CVPR.2017.195.*
- [186] Jegou H, Douze M, Schmid C, Perez P. Aggregating local descriptors into a compact image representation. *Comput Vis Pattern Recognit (CVPR), 2010 IEEE Conf* 2010:3304–11.
- [187] Lebeau JM, Findlay SD, Allen LJ, Stemmer S. Standardless atom counting in scanning transmission electron microscopy. *Nano Lett* 2010;10:4405–8. <https://doi.org/10.1021/nl102025s>.
- [188] Van Aert S, De Backer A, Martinez GT, Goris B, Bals S, Van Tendeloo G, et al. Procedure to count atoms with trustworthy single-atom sensitivity. *Phys Rev B - Condens Matter Phys* 2013;87:1–6. <https://doi.org/10.1103/PhysRevB.87.064107>.
- [189] De Backer A, van den Bos KHW, Van den Broek W, Sijbers J, Van Aert S. StatSTEM: An efficient approach for accurate and precise model-based quantification of atomic resolution electron microscopy images. *Ultramicroscopy* 2016;171:104–16. <https://doi.org/10.1016/j.ultramic.2016.08.018>.
- [190] Thust A, Coene WMJ, Op de Beeck M, Van Dyck D. Focal-series reconstruction in HRTEM: Simulation studies on non-periodic objects. *Ultramicroscopy* 1996;64:211–30. [https://doi.org/10.1016/0304-3991\(96\)00011-3](https://doi.org/10.1016/0304-3991(96)00011-3).
- [191] Coene WMJ, Thust A, Op de Beeck M, Van Dyck D. Maximum-likelihood method for focus-variation image reconstruction in high resolution transmission electron microscopy. *Ultramicroscopy* 1996;64:109–35. [https://doi.org/10.1016/0304-3991\(96\)00010-1](https://doi.org/10.1016/0304-3991(96)00010-1).
- [192] Hsieh WK, Chen FR, Kai JJ, Kirkland AI. Resolution extension and exit wave reconstruction in complex HREM. *Ultramicroscopy* 2004;98:99–114. <https://doi.org/10.1016/j.ultramic.2003.08.004>.
- [193] Chen FR, Kisielowski C, Van Dyck D. 3D reconstruction of nanocrystalline particles from a single projection. *Micron* 2015;68:59–65. <https://doi.org/10.1016/j.micron.2014.08.009>.
- [194] Jia CL, Mi SB, Barthel J, Wang DW, Dunin-Borkowski RE, Urban KW, et al. Determination of the 3D shape of a nanoscale crystal with atomic resolution from a single image. *Nat Mater* 2014;68:1044–9. <https://doi.org/10.1038/nmat4087>.
- [195] Park J, Elmlund H, Ercius P, Yuk JM, Limmer DT, Chen Q, et al. 3D structure of individual nanocrystals in solution by electron microscopy. *Science* 2015;349:290–5. <https://doi.org/10.1126/science.aab1343>.
- [196] Ragone M, Saray MT, Long L, Shahbazian-Yassar R, Mashayek F, Yurkiv V. Deep learning for mapping element distribution of high-entropy alloys in scanning transmission electron microscopy images. *Comput Mater Sci* 2022;201:110905. <https://doi.org/10.1016/j.commatsci.2021.110905>.
- [197] LeBeau JM, Findlay SD, Allen LJ, Stemmer S. Position averaged convergent beam electron diffraction: Theory and applications. *Ultramicroscopy* 2010;110:118–25. <https://doi.org/10.1016/j.ultramic.2009.10.001>.
- [198] Ophus C, Ercius P, Huijben M, Ciston J. Non-spectroscopic composition measurements of SrTiO₃-La_{0.7}Si_{0.3}MnO₃ multilayers using scanning convergent beam electron diffraction. *Appl Phys Lett* 2017;110. <https://doi.org/10.1063/1.4975932>.
- [199] Chen C-C, Marks LD, Scott MC, Ercius P, Yang Y, Ramezani-Dakhel H, et al. Three-dimensional coordinates of individual atoms in materials revealed by electron tomography. *Nat Mater* 2015;14:1099–103. <https://doi.org/10.1038/nmat4426>.
- [200] Chen C-C, Zhu C, White ER, Chiu C-Y, Scott MC, Regan BC, et al. Three-dimensional imaging of dislocations in a nanoparticle at atomic resolution. *Nature* 2013;496:74–7. <https://doi.org/10.1038/nature12009>.
- [201] Miao J, Ercius P, Billinge SJL. Atomic electron tomography: 3D structures without crystals. *Science* 2016;353. Doi:10.1126/science.aaf2157.
- [202] Gilbert P. Iterative methods for the three-dimensional reconstruction of an object from projections. *J Theor Biol* 1972;36:105–17. [https://doi.org/10.1016/0022-5193\(72\)90180-4](https://doi.org/10.1016/0022-5193(72)90180-4).
- [203] Miao J, Förster F, Levi O. Equally sloped tomography with oversampling reconstruction. *Phys Rev B - Condens Matter Mater Phys* 2005;72:3–6. <https://doi.org/10.1103/PhysRevB.72.052103>.
- [204] Shashank Kaira C, Yang X, De Andrade V, De Carlo F, Scullin W, Gursoy D, et al. Automated correlative segmentation of large Transmission X-ray Microscopy (TXM) tomograms using deep learning. *Mater Charact* 2018;142:203–10. <https://doi.org/10.1016/j.matchar.2018.05.053>.
- [205] Kodama M, Ohashi A, Adachi H, Miyuki T, Takeuchi A, Yasutake M, et al. Three-dimensional structural measurement and material identification of an all-solid-state lithium-ion battery by X-Ray nanotomography and deep learning. *J Power Sources Adv* 2021;8:100048. <https://doi.org/10.1016/j.powera.2021.100048>.
- [206] Tan RK, Zhang NL, Ye W. A deep learning-based method for the design of microstructural materials. *Struct Multidiscip Optim* 2020;61:1417–38. <https://doi.org/10.1007/s00158-019-02424-2>.
- [207] Vamvakarios A, Jacques SDM, Di Michiel M, Matras D, Middelkoop V, Ismagilov IZ, et al. 5D operando tomographic diffraction imaging of a catalyst bed. *Nat Commun* 2018;9:1–11. <https://doi.org/10.1038/s41467-018-07046-8>.
- [208] Coelho AA. TOPAS and TOPAS-Academic: an optimization program integrating computer algebra and crystallographic objects written in C++. *J Appl Crystallogr* 2018;51:210–8. <https://doi.org/10.1107/S1600576718000183>.
- [209] He M, Yang T, Shang D, Xin B, Chernov AI, Obraztsova ED, et al. High temperature growth of single-walled carbon nanotubes with a narrow chirality distribution by tip-growth mode. *Chem Eng J* 2018;341:344–50. <https://doi.org/10.1016/j.cej.2018.02.051>.
- [210] Warner JH, Young NP, Kirkland AI, Briggs GAD. Resolving strain in carbon nanotubes at the atomic level. *Nat Mater* 2011;10:958–62. <https://doi.org/10.1038/nmat3125>.
- [211] Torquato S. Random Heterogeneous Materials: Microstructure and Macroscopic Properties. New York, US: Springer Science & Business Media; 2013.
- [212] Hjeltnes MJ, Snoeck E, Kilaas R. Quantitative measurement of displacement and strain fields from HREM micrographs. *Ultramicroscopy* 1998;74:131–46.
- [213] Madsen J, Liu P, Wagner JB, Hansen TW, Schiøtz J. Accuracy of surface strain measurements from transmission electron microscopy images of nanoparticles. *Adv Struct Chem Imaging* 2017;3. <https://doi.org/10.1186/s40679-017-0047-0>.
- [214] Galindo PL, Kret S, Sanchez AM, Laval JY, Yáñez A, Pizarro J, et al. The Peak Pairs algorithm for strain mapping from HRTEM images. *Ultramicroscopy* 2007;107:1186–93. <https://doi.org/10.1016/j.ultramic.2007.01.019>.
- [215] Bierwolf R, Hohenstein M, Phillip F, Brandt O, Crook GE, Ploog K. Direct measurement of local lattice distortions in strained layer structures by HREM. *Ultramicroscopy* 1993;49:273–85. [https://doi.org/10.1016/0304-3991\(93\)90234-O](https://doi.org/10.1016/0304-3991(93)90234-O).
- [216] Zuo JM, Shah AB, Kim H, Meng Y, Gao W, Rouvière JL. Lattice and strain analysis of atomic resolution Z-contrast images based on template matching. *Ultramicroscopy* 2014;136:50–60. <https://doi.org/10.1016/j.ultramic.2013.07.018>.
- [217] Kübel C, Thust A. True Image. In: Weirich TE, Lábár JL, Zou X, editors. Electron Crystallography. NATO Sci. Ser. II Math. Phys. Chem., Dordrecht: Springer; 2006, p. vol. 211.
- [218] Ophus C. Four-Dimensional Scanning Transmission Electron Microscopy (4D-STEM): From Scanning Nanodiffraction to Ptychography and Beyond. *Microsc Microanal* 2019;563–82. <https://doi.org/10.1017/S1431927619000497>.