

Deep learning analysis on microscopic imaging in materials science

M. Ge ^{a, b}, F. Su ^{a, b, **}, Z. Zhao ^{a, b, ***}, D. Su ^{c, *}

^a School of Artificial Intelligence, Beijing University of Posts and Telecommunications, Beijing, 100876, China

^b Beijing Key Laboratory of Network System and Network Culture, Beijing University of Posts and Telecommunications, Beijing, 100876, China

^c Beijing National Laboratory for Condensed Matter Physics, Institute of Physics, Chinese Academy of Sciences, Beijing, 100190, China



ARTICLE INFO

Article history:

Received 1 May 2020

Received in revised form

18 May 2020

Accepted 27 May 2020

Available online 4 June 2020

Keywords:

Image analysis

Materials informatics

TEM

SEM

SPM

ABSTRACT

Microscopic imaging providing the real-space information of matter, plays an important role for understanding the correlations between structure and properties in the field of materials science. For the microscopic images of different kinds of objects at different scales, it is a time-consuming task to retrieve useful information on morphology, size, distribution, intensity etc. Alternatively, deep learning has shown great potential in the applications on complicated systems for its ability of extracting useful information automatically. Recently, researchers have utilized deep learning methods on imaging analysis to identify structures and retrieve the linkage between microstructure and performance. In this review, we summarize the recent progresses of the applications of deep learning analysis on microscopic imaging, including scanning electron microscopy (SEM), transmission electron microscopy (TEM), and scanning probe microscopy (SPM). We present sequentially the basic concepts of deep learning methods, the review of the applications on imaging analysis, and our perspective on the future development. Based on the published results, a general workflow of deep learning analysis is put forward.

© 2020 Published by Elsevier Ltd.

1. Introduction

A primary goal of materials science is to unpin the correlations between materials' structure and their physical or chemical properties. Here the structure means morphology, phase, atomic structure, surface facet, interfacial structure, etc. With the knowledge of so-called structure–property correlations, it is possible to tune the parameters of synthesis to control structure of materials and therefore optimize new materials for applications. This is a fundamental paradigm for many types of materials research, including metallic materials, semiconductors, and functional materials for catalysis, biomedicine, electronics, batteries and so on. Proper structure and property characterization methods are vital for understanding the underlying correlation mechanisms. Among the techniques of structure characterization, microscopic imaging is particularly important because it can provide real space information on the structure of materials in a large range of scale.

The typical imaging tools include optical microscopy, transmission electron microscopy (TEM), scanning electron microscopy (SEM), and scanning probe microscopy (SPM). Optical microscopy is the most common tool, which allows us to see the details of object at a spatial resolution of micron. TEM with an aberration corrector provides the project information of a thin specimen at a spatial resolution of sub-Armstrong [1–4]; SEM gives information on surface morphology (secondary electron) and composition (backscattering electron); and SPM, including atomic force microscopy (AFM), scanning tunneling microscopy (STM), etc., have been widely used to study atomic structure and electronic states on surface. These techniques can help to recognize details from millimeter to tens of picometer and give unique information on materials features of morphology, phase, crystallography, magnetic structure, and molecular as well as atomic structure.

Although the images from these tools have different physical meanings, the processes for image analysis have some similarities. Taking an example of TEM imaging shown in Fig. 1, there are mainly four challenges for the analysis [5]. The first challenge is how to achieve materials-specific information through experimental or simulated data, or in other words, how to utilize images to infer structural information. The second challenge is how to use the information to predict or generate physical and chemical properties via generative models. The third is how to reconstruct materials

* Corresponding author.

** Corresponding author.

*** Corresponding author.

E-mail addresses: sufeil@bupt.edu.cn (F. Su), zhaozc@bupt.edu.cn (Z. Zhao), dongsu@iphy.ac.cn (D. Su).

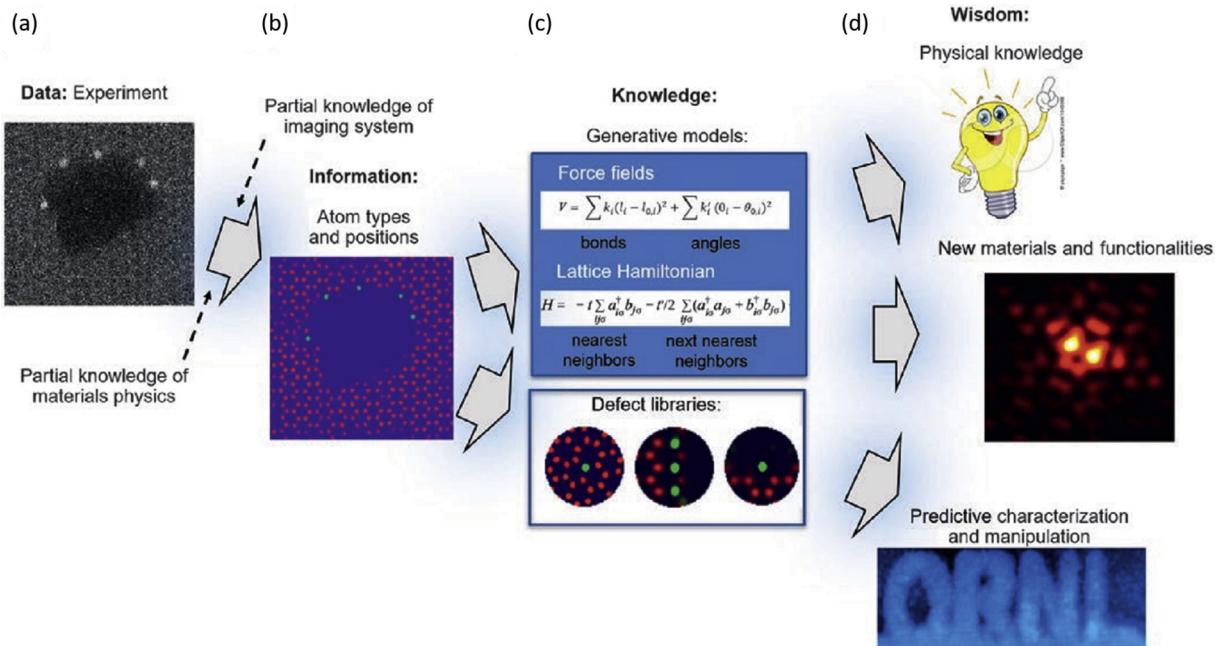


Fig. 1. Technical challenges for microscopic imaging analysis [5]. From (a) experimental data, the first step is whether we can obtain (b) the partial knowledge of materials physics. The second step is whether the obtained materials-specific information with uncertainties can be used to infer physical and chemical properties through (c) generative models. The third step is whether the materials information can be used to reconstruct materials behaviors. The last one is whether we can get real-time analysis results of the data stream from the microscope (Reproduced from Ref. [5]).

behaviors or how to predict the performance. The last one is whether we can realize real-time analysis of the data stream from the microscopes. We believe that these challenges can be generalized to any imaging techniques where physical knowledge is abstracted from experimental results. Normally, this workflow is successfully built based on repeated and multimodal experimental studies. In a complicated system, the link is often not clear and takes time to be revealed. Ideally, if this knowledge-reveal process of *in-situ* microscopic techniques can be finished in real time or a short time, it may greatly help investigators to study the dynamics of the system. For example, according to the analyzed results, the investigators may know how to tune experimental parameters or what to do next during the experiment. Utilization of data-driven technologies, such as machine learning, can help to fulfill this task, such as on the discovery of functional materials [6], on imaging characterization [7], on physically relevant properties prediction [8], on materials design [9], etc. Furthermore, data-driven technologies can help to improve the performance of prediction and decision-making processes too [10]. These demands have driven the rapid growth of the utilization of machine learning in recent years.

As a part of artificial intelligence (AI) techniques, machine learning has been developed and researched on for several decades [11]. It is a research direction of computational models and algorithms aiming to make a computer system work on specific tasks based on data and experience. Traditional machine learning methods usually need the hard-coded features as input, which are basically designed with human's assistance and can hardly be changed. Inspired by the enormous scale of data, the development of advanced computer algorithms, and the breakthroughs in computing hardware, especially the rise of graphics processing units (GPUs), a new type of representation-learning methods named 'deep learning' [12] has come up. As shown in Fig. 2, deep learning belongs to the family of machine learning and of course the broad field of artificial intelligence [13,14]. The ability of

learning representations automatically via composing simple but non-linear modules, makes deep learning more powerful than most of the traditional machine learning methods, which usually need manually chosen feature-extraction algorithms as pre-processing. Moreover, with enough modules (also known as layers) stack-up, deep learning can realize complex functions or systems.

Very recently, deep learning has speeded up its applications in many areas such as social network analysis [15], information retrieval [16], speech and audio processing [17], visual data processing [18], natural language processing [19], and so on, as listed in Fig. 3. There have been attempts of using deep learning approaches in the field of materials science as well [20–25], including structure prediction and design [26–28], learning of chemistry [29], structure–property linkage analysis [30–35] and structure characterization [36–39]. Among all the research areas, microscopic imaging analysis, which objects to retrieve the real-space information of materials, is of particular importance. As a devotional, non-periodic record of different morphologic subjects, images normally contain massive information and huge complexity which are suitable for AI analysis. An emerging need for materials scientists is to know what deep learning analysis can help and how to utilize this tool on materials data analysis. In this review, we focus on research progresses of deep learning on microscopic imaging. In the following text, we firstly introduce the concepts and different approaching models of deep learning, and then we review the applications of deep learning on SEM, TEM, SPM etc. Summaries of the progresses in the microscopic imaging analysis are presented. Finally, we elucidate our perspective on the challenges and opportunities for the utilization of deep learning on the image analysis in the part of summary and outlook.

2. Concepts and workflow of deep learning

First, we briefly introduce the basic concepts of deep learning for the convenience of the following content. Then we introduce

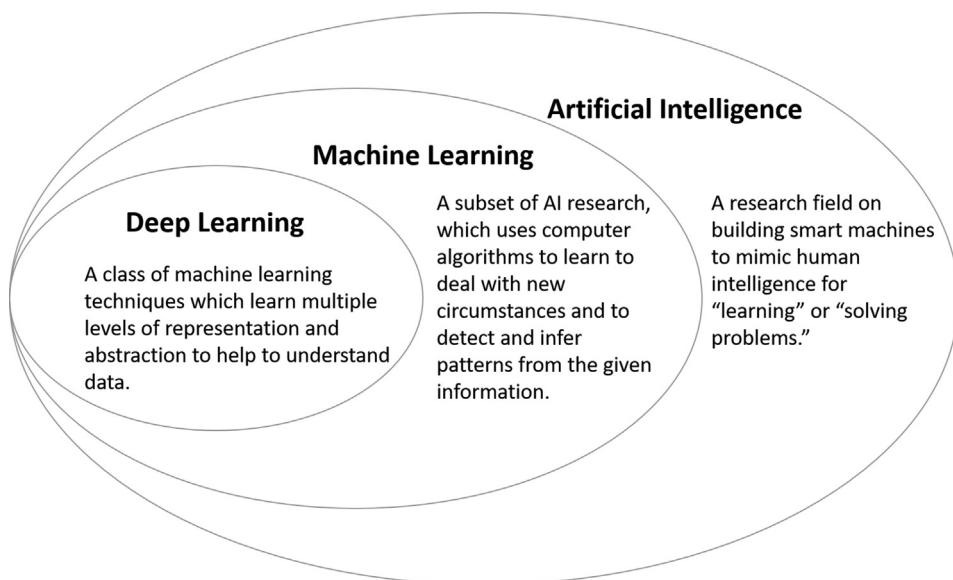


Fig. 2. The relationship of deep learning, machine learning and artificial intelligence [13,14].



Fig. 3. Applications of deep learning in various fields [70].

the difficulties of microscopic imaging analysis tasks by summarizing the differences between microscopy images and natural scene images. As a conclusion, we propose the workflow for the utilization of deep learning on microscopic imaging which mainly contains five stages, according to the way that deep learning works and the demands of analyzing microscopy images.

2.1. Key concepts of deep learning

An overall schematic of a neural network for object identification [40] is shown in Fig. 4. To link the input image with the correct output label, the model is designed to implement the sophisticated identification process by breaking it up into simple but non-linear mathematic operations, such as convolutions, normalization, activations etc. The layers are connected by propagating the calculated information (or features) through the whole neural network. The information flow can be straight forward through all the stacked layers, or recurrent by feeding the present information and the previous information into the current layer. The network can achieve more and more abstract and structural information through stacking-up those combinations of operations. For example, the first layer of the network may only extract edges in the images, and the third layer enables to learn more structural information like object parts. By designing and training the network, each simple operation can adjust their parameters automatically based on the constraints of the optimization rules, which are designed to make the whole system work efficiently. In the following part, several key concepts are briefly described.

For datasets, the expected results of a designed model are named as 'ground truth' (or label), and ground truth data vary for different tasks [41]. For example, ground truth could be the labeled features used for classification, and the ground truth for microscopic imaging can be regarded as the analyzed image with information of geometric positions, radius and areas. Usually, the

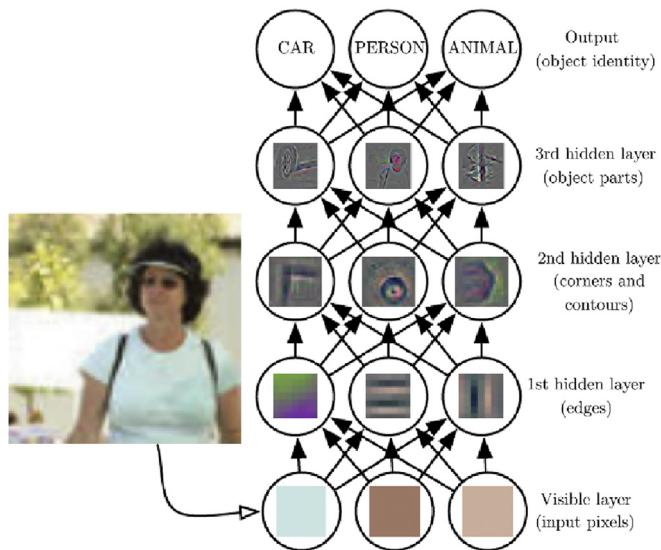


Fig. 4. Schematic of a multilayer convolutional neural network used for object identification. Usually each "layer" contains a combination of operations, including convolutions, ReLU, max pooling and normalization. The combination differs in different models. Input pixels are called as visible layer due to the visibility by human and the hidden layers are named because the variables cannot be observed directly. From bottom to top, each circle denotes the features extracted by the layer. The higher layer achieves more abstract information than the lower ones, which means the neural network can learn to extract useful structural information automatically. The final output is the probability for object identity and the probability of being a person should be the highest in this example (Reproduced from Ref. [40]).

datasets are split as training set, test set and validation set [42]. A training set is used to make the model learn and a validation set is usually used to tune the hyperparameters during training. A test set is the sample of data that the model has not seen to evaluate the model's performance after the training.

For a deep learning model, if there is ground truth provided, then the model is called 'supervised' model, otherwise it is named 'unsupervised' model. Each neural network could be composed of multiple types of layers, including convolutional layers, deconvolutional layers, activation layers, pooling layers, fully connected layers, normalization layers and so on. 'Features' denote the hidden representations or outputs of each layer. Each layer presents a mathematical operation and stacking up those layers could model a very complex function. For example, a two-dimensional convolution operation [40] can be written as

$$S(i,j) = (I * K)(i,j) = \sum_m \sum_n I(i+m, j+n)K(m,n) \quad (1)$$

where the asterisk denotes cross-convolution operation, I denotes the input image, K denotes a two-dimensional kernel and S denotes the output feature map. And ReLU, also named rectified linear unit [43], is a widely used activation function

$$g(z) = \max(0, z) \quad (2)$$

and the function is applied element-wise [40]. Pooling can be seemed as a sampling operation and the most common choice is max pooling, which returns the maximum value of the sampling kernel [40].

For the optimization of a deep learning network, the minibatch stochastic methods are usually applied, which are algorithms that use a batch of fixed number of training examples to estimate the gradients. And back propagation is the main mechanism for the training of the network, which is a process of the calculation of the gradients of each layer or operation. When the batch size is one, the optimization algorithm is usually called stochastic method or online method. Different optimization methods acquire information from minibatches in different ways [40]. There are many efficient optimization methods, such as momentum [44], Adam [45], RMSprop [46] etc. All the configurations, including network architecture, optimization method and rules, training settings, should be carefully designed to meet the requirements of the task.

During training process, potential problems may occur, such as overfitting, underfitting and gradient explosion. Overfitting means the situation when a model learns the noise of the training data and underfitting occurs when the model cannot learn the underlying trend of the data. Overfitting and underfitting are bias-variance trade-off and the objective of deep learning model is to balance the bias and variance error. Usually, overfitting can be prevented by many solutions, including enlarging the training set, increasing regularization, stopping the training process earlier, randomly dropping out part of the features and so on. Underfitting can be solved by training more iterations, decreasing regularization and using more complex models. The overfitting problem can also be tackled as a representation learning problem when there are only a few training samples [47–49]. Gradient explosion often comes with Recurrent Neural Networks [50], which use the same module several times via a recurrent connection. Gradient explosion can be dealt with by gradient clipping. There are many other potential problems and researchers have also proposed the respective training skills to fix them.

A pretrained model usually means a model that has already been trained using other datasets, which are usually Imagenet [51] data for the task of classification, or COCO dataset for the tasks of

object detection and segmentation [52]. Well pretrained models, such as ResNet [53], VGG [54], Inception Net [55–57], which can achieve great performance on their training dataset, are thought to have the ability of extracting useful features. Thus, pretrained models are widely used as robust and efficient feature extractor in other tasks, including semantic segmentation (or pixel-wise segmentation) [58], object detection [59] etc., by utilizing their architectures and the parameters in a new neural network.

For the performance evaluation of the model, accuracy, precision and recall are generally used in most of tasks. For a classification problem, a confusion matrix is usually utilized to display the results, which is a summary of the prediction. It shows the count values of the correct or incorrect predicted results. The confusion matrix consists of four elements for a binary classification task: true positive (TP, correctly predicted positive values), true negative (TN, correctly predicted negative values), false positive (FP, incorrectly predicted positive values), and false negative (FN, incorrectly predicted negative values). Given a confusion matrix, the accuracy, precision and recall can be calculated as following:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (4)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (5)$$

Accuracy represents the fraction that the model predicts right, precision denotes the fraction of all the positive predictions which are correct, and recall denotes the fraction of actual positives which are correctly predicted. Usually a trained deep learning model can be assessed by its performance on test dataset. However, there may exist mistakes in the outputs of the trained model for an unseen instance (or an out-of-distribution instance), even when its accuracy is high enough on test set. One way to improve the reliability of deep learning model is to enlarge the training dataset and human experts' assistance is also a solution. Another method is to estimate the uncertainty and make use of it to help with decision making [60].

Deep learning usually involves with numerous scales of parameters and data, including convolution values, activation values, gradient values, etc., which need to be changed during each step of training process [40]. Thus, suitable hardware and frameworks are critical to implement efficient deep learning methods. The criteria of evaluating hardware and framework are not only the computing performances but also energy efficiency, chip price, compatibility and so on. Compared to CPUs, GPUs are more widely utilized due to the high parallelism and high memory bandwidth. What's more, the development of general-purpose GPUs (GP-GPUs) also offers wide availability and generality, convenient programming environment and well-supported software stacks for implementing neural networks [61]. For example, NVIDIA's GPUs achieve popular usage, due to their proposed standard libraries which make deep learning methods easy to establish [62]. Open-source frameworks enable researchers to design, implement a deep learning model and accelerate the training process based on the available hardware. There are many developed open-source frameworks with interpreted language frontends. Among them, mature and popular frameworks such as Theano, Torch, Caffe, TensorFlow, PyTorch etc., which are user-friendly for people who are familiar with Python or MATLAB, can support different kinds of hardware (either CPU or GPU) and achieve high performance [63]. To illustrate the configuration of software, the image classification task on ImageNet [51]

dataset is introduced as an example. A network named Fix-EfficientNet proposed by Touvron et al. achieved 88.5% top-1 score (the percentage of the one with the highest output probability to be the target label, a score to evaluate the performance of accuracy) on the task with the number of parameters as 480 M, which was the best reported network with external data until now [64]. The Fix-EfficientNet was implemented and trained using PyTorch framework on a cluster with several GPUs. Compared to another one of BiT-L [65], with 87.54% top-1 score and 928 M parameters, the FixEfficientNet achieved a great performance on both the number of parameters and accuracy. Besides the accuracy and number of parameters mentioned above, the performance of hardware is also vital for the efficiency of training and inference. Companies of NVIDIA and ATI have introduced many widely used products. Studies had shown that both of them have their own advantages and favor different applications and energy efficiency. For example, the NVIDIA Fermi GPU performs great on double-precision problems and the ATI Radeon GPU is more energy-efficient [66]. The configuration of model architecture and hardware can directly affect the performance and efficiency of the model. The number of operations, which is determined by model structure, decides the inference time. There are studies showed that the energy constraint of hardware was related to the upper bound on the accuracy and model complexity [67]. Bianco et al. showed that on the task of ImageNet, the deep learning model with a very low complexity still needed a minimum GPU memory of about 0.6 GB. The Inception-V4 model, with batch-size as 1 and memory consumption as 0.8 GB, could achieve a classification accuracy of about 80% and the inference time of 18.96 ms per image. As an example, the previous experiment was implemented based on PyTorch, with a workstation equipped with an Intel Core i7-7700 CPU @ 3.60 GHZ, 16 GB DDR4 RAM 2400 MHz, NVIDIA Titan X Pascal GPU with 12 GB memory [68]. To be summarized, the hardware part is highly dependent on the deep learning models and frameworks. It is well-known that GPU can accelerate the machine learning computing and the choice of GPUs should consider the storage needed for the model, as well as the computation and energy efficiency. Hard drives with high capacity are necessary for storage of large scale of data. If the data size exceeds the available memory of processing units, the data can be processed part by part. For the storage of massive scale of data, cloud storage and computing techniques can be utilized.

2.2. Difference between microscopy images and natural images

According to the FAIR guiding principles [69], scientific data should be findable, accessible, interoperable and reusable. It is important to develop data analyzing tools to help the management of the enormous number of images and moreover to unpin the correlation between materials' structures and properties based on these images. Although deep learning has made huge improvement on natural scene image analysis tasks [70], such as object detection, classification, prediction and so on, there are many differences between tasks of natural scene analysis and the ones of microscopic imaging analysis. Compared to natural scene images, microscopy images in materials science have featured characters, which make the tasks of microscopic imaging analysis very different and difficult.

First, raw experimental microscopy images usually associate with high-level noise and distortions [71,72]. There are some traditional unified image preprocessing algorithms to reduce the noise like Gaussian smooth filter, order filter, etc. Traditional algorithms usually have difficulties on tackling with high density noise, due to their heavily dependence on local information.

However, deep learning algorithms can overcome the limitation of noise via well-defined ground truth and model architecture [39].

Second, the datasets of microscopy images are usually gray-scaled, different from natural scene images. Natural scene images are mostly colorscaled (Red, Green, Blue (RGB) or RGBD which has another depth channel) and have rich natural textures. To increase the quantity and diversity of datasets, augmentation is essential for the situation of limited data, including geometrical transformations, color space augmentations, kernel filters etc. [73]. Several recent works on using deep learning in microscopic imaging analysis augmented the training dataset via simple transformation such as rotation [74,75], in order to enrich the training dataset and avoid overfitting. Moreover, most of the atomic resolved images have unique periodicity of lattices and the characterization methods should follow the distribution of atoms.

Third, information of a microscopy image has strong associations with physical or chemical constraints and experimental environments, which makes the structures of microscopy image datasets much more sophisticated. There are specific physical and chemical parameters for each of the microscopy image, including the type of materials, imaging technologies, equipment parameters, environment factors and so on. What is more, microscopic images usually have very high resolution. Thus, the scale of the microscopy data is larger, and the structure of each data-node is more complex than natural scene image data. The complexity and diversity of microscopy datasets make the discovery of materials-specific fingerprint with human effort very difficult. Whereas, deep learning methods are good at processing and mining large-scale datasets, which decides their adaptability to large scale microscopic imaging data.

Fourth, analysis tasks are usually required to be done on dynamic datasets or 'movies', in order to observe the process of materials transformation. Deep learning models like Long Short Term Memory (LSTM) networks [76] can 'remember' information for long periods of time and have been widely applied on processing time sequences in several tasks, including Natural Language Processing (NLP) [17], traffic forecast [77], human action recognition

[78] etc. Utilization of spatial-temporal neural networks can help with the localization and analysis of the materials' dynamic process.

In addition, the greatest challenge may be the hard-achievable 'ground truth' label. For a supervised task, labeled data (ground truth) is necessary. Usually, the ground truth can be achieved by human investigators and computer simulation, which is time-consuming and inefficient. Moreover, the training of supervised deep learning model and the validation or analysis of experiments need large scale of accurately labeled data. The above-mentioned factors make the generation of ground truth very important and challenging. There are several works using data labeled by human assistance [75,79]. Many algorithms utilize simulated data [80–83], data processed beforehand [39,84] instead. One way to generate reliable simulated dataset using software or theoretical models, is to consider the variance of multiple physical and chemical factors which can affect the generated results massively. Another way is to consider using the few-shot learning methods [85], which are developed to tackle with feeding a deep learning model with a very small training set.

Finally, many tasks of microscopic imaging are not fully determinative and need further analysis or explanation. Unlike most of tasks of natural scenes or language processing, which have clear and definite knowledge as basis, many phenomena in materials science are unexpected or unforeseen and need scientists to make further study on the causes and knowledge of them.

Considering all the characters mentioned above, research on microscopic imaging in materials science with deep learning should be carefully designed and implemented, under the supervision of physical and chemical constraints and knowledge.

2.3. Workflow of using deep learning on microscopy images analysis

The workflow can be summarized to the process shown in Fig. 5, which contains five stages, including task analysis, data preparation, model design, feature analysis, and validation.

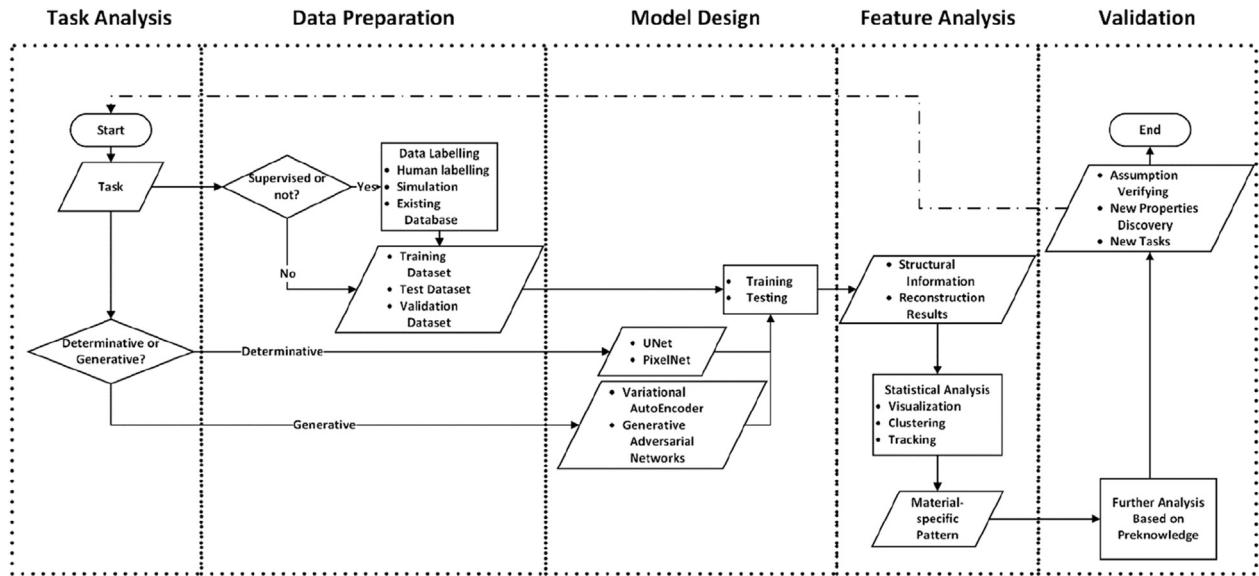


Fig. 5. Summarized workflows for deep learning analysis on images, which contain five stages. The first step is task analysis, for example, cataloging the task as a classification problem or a regression problem. The second is the preparation for training data, which can be achieved by human labeled [38] or simulation [80]. Appropriate model design is crucial in the analysis process, such as DefectSegNet for semantic segmentation [74], Inception V3 [56] for feature extracting, Variational AutoEncoder [102] for learning connections and so on. Then comes to the feature analysis to extract physical and chemical information through the outputs of the trained model. The last step is validation and further analysis based on prior knowledge. Here, the rectangles denote operations, the diamonds denote decisions to make and parallelograms denote the intermediate data or results.

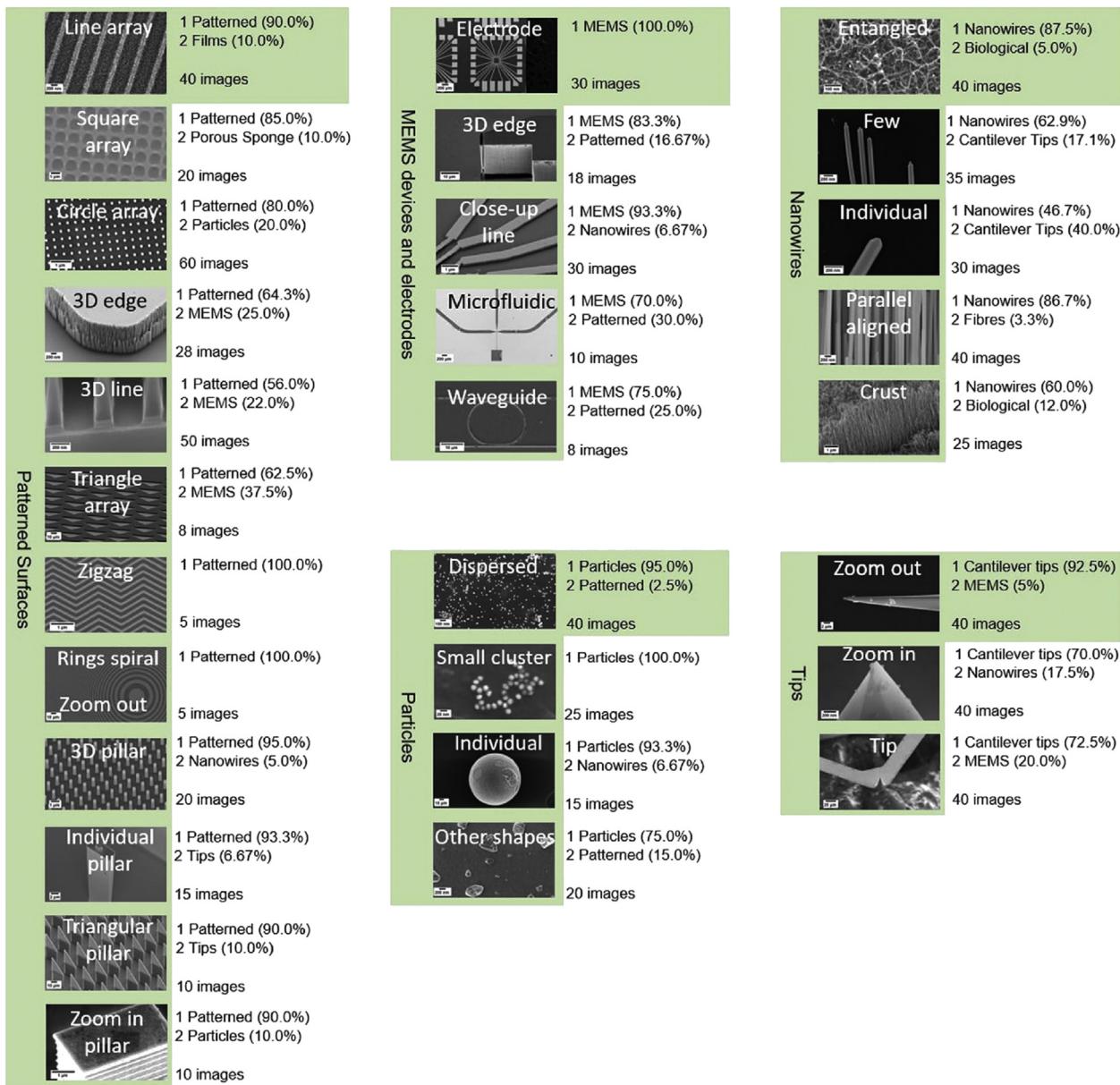


Fig. 6. Deep learning analysis on SEM images: cataloging image features [79]. This figure illustrates the classification outputs of the categories with the most scattered results in Table 1. In each subcategory, the top two ranked values are the percentage of the classification results of images belonged to that category.

Table 1

The classification results of ref. [79]. The tag Ln (n = [0,9]) denotes the category. The notation is the following: L0 = Porous_Sponge, L1 = Patterned_surface, L2 = Particles, L3 = Films_Coated_Surface, L4 = Powder, L5 = Tips, L6 = Nanowires, L7 = Biological, L8 = MEMS_devices_and_electrodes, L9 = Fibres. The last column shows the number of images of each category in the test set.

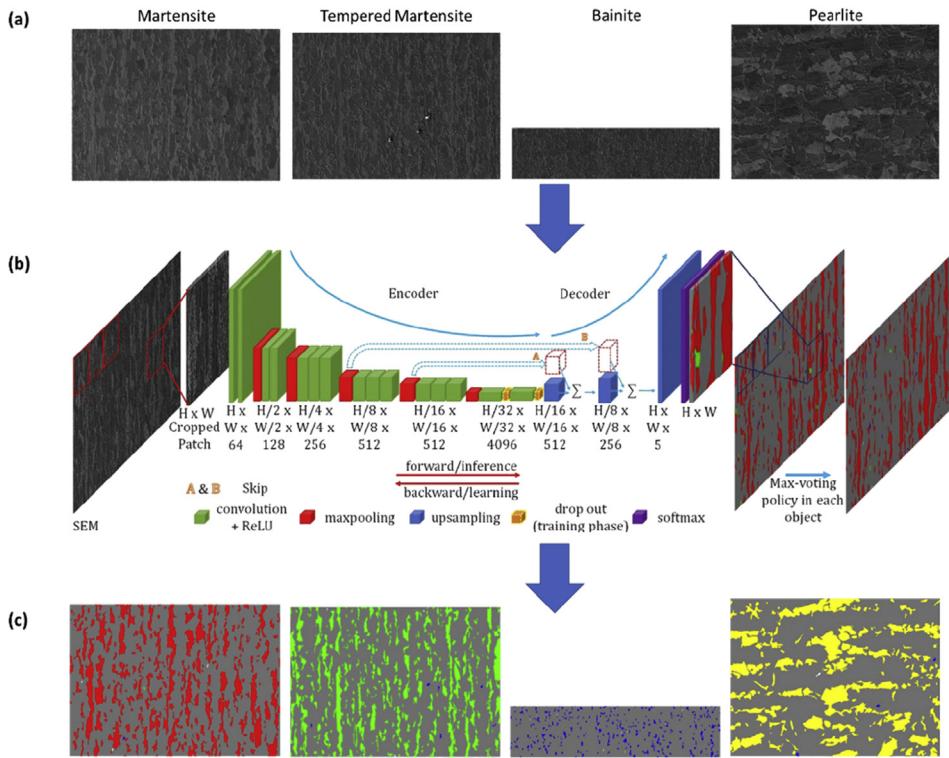


Fig. 7. Analysis for low-carbon steel based on a network which uses the Fully Convolutional Neural Networks as a backbone [86]. The authors propose a max-voted fully convolutional neural networks (FCNN) [MVFCNN] to assign labels to each pixel, including martensite, tempered martensite, bainite, pearlite and background. (a) are input examples of the four categories and (b) presents the network architecture used for this study. (c) shows the output results of the images in (a). The colors of martensite, tempered martensite, bainite and pearlite are red, green, blue and yellow respectively.

task specification and analysis, including determination of the task category, physical constraints, and other requirements. For example, a characterization or quantification task could be described as a determinative problem, and on the other side, a linkage learning task could be presented as a regression problem. The second step is dataset preparation via human assistance or simulation tools. If the learning task is supervised, then data sets should contain the labels, or in other words, the expected results for the input data. The third is to design the model according to the task and the existing data. The next step is feature analysis, which helps to extract the pattern from the outputs of the trained model in the perspective of materials science. The last is validation and further analysis for making prediction on properties or performance, verifying assumption or observing specific phenomenon. The overall objective of the utilization of deep learning is to model the connection between experimental data and the respective properties, and then using the connection to reconstruct or predict the performance.

3. Application of deep learning on microscopic imaging

Herein, we review recently published reports on the analysis of scanning electron microscopy (SEM), transmission electron microscopy (TEM)/scanning transmission electron microscopy (STEM), and scanning probe microscopy (SPM) techniques.

3.1. Analysis of SEM images

Deep learning has been used to solve two problems for SEM images analysis: One is the classification of SEM images based on

their morphological features [38,79,86], and the other is resolution enhancement to improve the image quality [87].

Modarres et al. utilized a pretrained deep learning model for SEM images classification as shown in Fig. 6 and Table 1 [79]. The authors compared the results of four different deep learning models (Inception-slim, Inception-v3, Inception-v4 and ResNet) and proved that Inception-v3 [56], which was composed of symmetric and asymmetric blocks including convolutions, pooling (or subsampling), concatenations etc., performed better than others on both accuracy and computational efficiency. The authors manually labeled 18,577 images into 10 categories as the training set. Table 1 shows the classification accuracy of the algorithm on a test set with 1,853 images in total. The results of test set indicated that, for categories like patterned surface and microelectromechanical system (MEMS) devices, which had divergent image characters, the neural network could not achieve a good performance. Further analyses were performed on dataset with a more detailed category setting and most of the subcategories were correctly labeled.

Another case is a pixel-wise classification on carbon steel images [38,86,88,89]. Pixel-wise classification is a kind of classification by assigning a specific label to each pixel of the image, which have the advantages of providing details of object shape and area compared to the global image classification in the above case. Azimi et al. proposed an object-based convolutional neural network to make classification on constituents and phases as shown in Fig. 7 [86]. A network called max-voted fully convolutional neural networks (FCNN) [MVFCNN] was developed to make pixel-wise segmentation on low carbon-steel SEM or Light Optical Microscopy (LOM) images. By comparing the object-based and pixel-wise classification networks, the authors concluded that the later one produced more accurate and detailed segmentation results, in

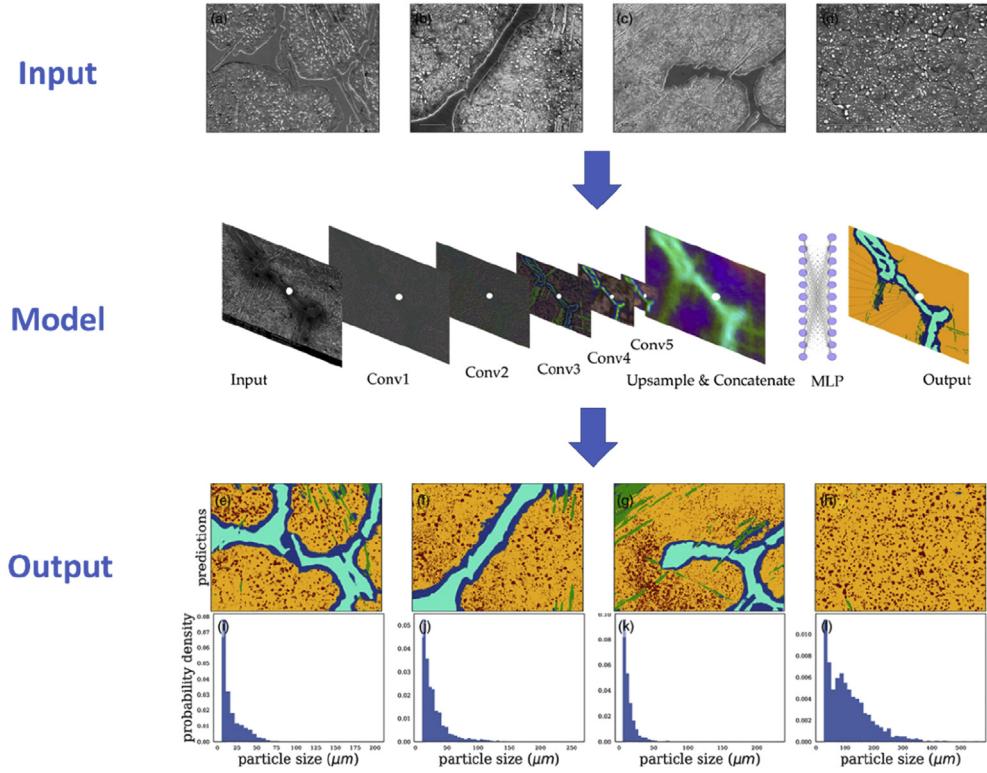


Fig. 8. Pixel-wise segmentation of the ultrahigh carbon steel microstructure dataset using PixelNet [38]. The overall image can be segmented to four categories, including pro-eutectoid grain boundary cementite (light blue), ferritic matrix (dark blue), spheroidite particles (yellow) and Widmanstätten cementite (green). The region of spheroidite particles are further segmented into small spheroidite particles (red).

which the features of martensite, tempered martensite, bainite, pearlite and background were clearly distinguished. There were reports of pixel-wise classification on ultrahigh carbon-steel microstructures [38,88,89] as well. DeCost et al. developed an ultrahigh carbon steel dataset named UltraHigh Carbon Steel DataBase (UHCSDB) from SEM imaging results and the processing and visualizing tools for this dataset [88,89]. Based on the dataset and a network named PixelNet [90], which used a combination of the convolutional results from multiple scales as the final feature map, DeCost et al. segmented and analyzed small microscopic constituents, including pro-eutectoid grain boundary cementite, ferritic matrix, spheroidite particles, Widmanstätten cementite, as shown in Fig. 8 [38]. These cases show that a properly designed and trained network can efficiently catalog the physical features over massive data. The final classification results can differ from global category of the image to small particles characterization, by the utilization of the selected network, the prepared ground truth and priori knowledge on the target.

Deep learning methods were applied to enhance the resolution of SEM images, which helped to retrieve the useful features of the samples that were fragile to the focused electron beam, as shown in Fig. 9 [87]. Generative adversarial networks (GANs) [91] were used for this task, which were developed based on the idea of game theory and usually included two branches of networks, the generator and the discriminator. In the case of image enhancement, the generator aimed to reconstruct high-resolution images from low-resolution ones and the discriminator was trained to distinguish which ones were generated images, or fake ones. When the discriminator could not tell the difference, then the generator succeeded to learn how to make predictions as real as possible. The training dataset used in this study was constructed on a gold-on-

carbon sample, with the low- and high-resolution images taken at different magnifications. The analyses showed the capability of the trained model on improving the quality of SEM images, with a two-fold increase in resolution. The success on image enhancement proved that deep learning has the ability of predicting microscopic data from limited information.

3.2. Analysis of STEM images

Scanning TEM (STEM) imaging technique has unique capabilities for materials science: (1) high angle annular dark field (HAADF)-STEM can provide an atomic number-contrast (Z-contrast) imaging where one can distinguish the elements directly from the contrast. (2) STEM imaging has a much better spatial resolution, which is determined by the electron probe size and can be down to sub-Armstrong. Thus, STEM has been intensively used to collect real-space information of crystallographic structure and defects [74,80–84,92–95]. Neural network of deep learning enables the exploring of the crystallographic phases, atomic configurations and dynamic transformations more efficient. The objectives for STEM imaging analysis are to characterize the atomic structure, nature of defects, as well as the morphology of samples, and then correlate the structural information with material-specific properties and performance, as indicated in Fig. 1.

The crystallographic defects could be segmented using deep learning framework [74], which is shown in Fig. 10. Roberts et al. proposed a network called DefectSegNet to detect and classify crystallographic defects pixel by pixel in advanced diffraction contrast imaging (DCI) STEM images into three categories, including dislocation lines, precipitates and voids. The DefectSegNet was designed with the inspiration of UNet [96] and DenseNet

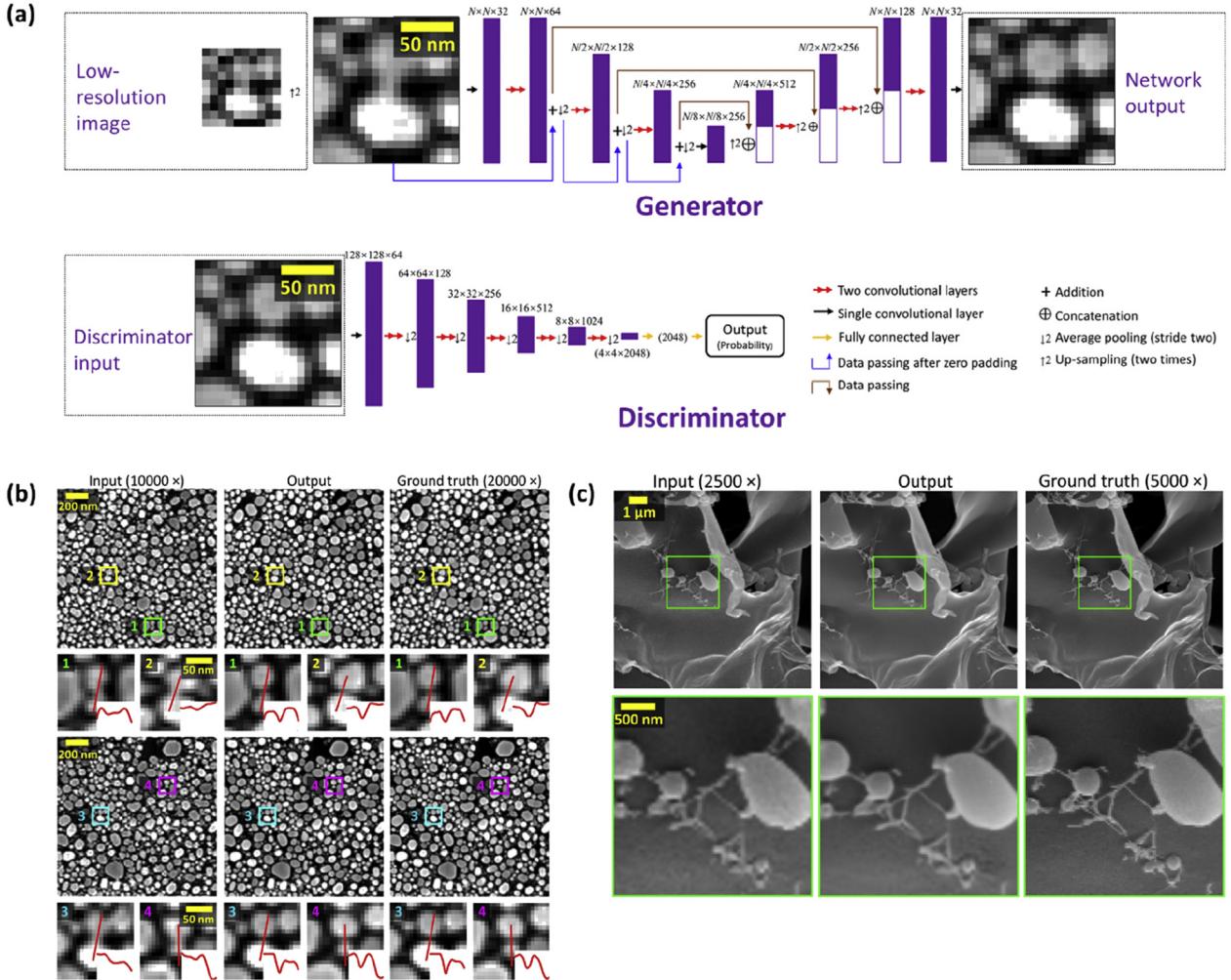


Fig. 9. Super-resolution reconstruction of SEM imaging using deep learning [87]. (a) is the model of Generative Adversarial Network, which contains two main branches of Generator and Discriminator. (b) and (c) are examples of gold nanoparticles and hydrogel samples respectively. The network can produce clear gap between nanoparticles as shown in (b). However, images with the beam-induced damage in (c) do not achieve good performance using the network.

[97], where the former was well known for the great performance in biomedical segmentation and the latter one enabled full usage of all the convolutional layers by creating short paths from early layers to later layers. The network was trained using high-quality DCI STEM defect images of HT-9 martensitic steels, which were obtained using a JEOL ARM200CF microscope operated at 200 kV and a convergence semi-angle of 6.2 mrad and bright-field collection angle of 9 mrad. Comparisons proved that the deep learning based semantic segmentation outperformed the defects quantification by human experts in both pixel-wise accuracy and time efficiency.

The analysis of atomic resolved STEM images normally follows the workflows of Fig. 5. After a task analysis, one may prepare the dataset generated by manual labelling or simulation based on the atomic structure of sample. The simulated STEM images can provide an accurate ground truth containing our interested feature, which can be done using some imaging simulation software, such as JEMS or Crystal Maker etc. Usage of simulated images is very efficient but have their limits because these images do not contain the features from real situations such as noises, scanning distortion, straining effect and other factors. Whereas, experimental images labeled by researchers can be more realistic and reliable, but time-consuming. As shown in Fig. 11, given a proper trained network,

structural information can be extracted, such as atomic positions [82], dumbbells [93]), column heights of atoms [80], lattice types [81] and defects types [84]. For example, Ziatdinov et al. utilized an efficient ‘atom finder’ tool to identify atomic positions by combining deep learning network and image processing algorithms [82,93]. The trained networks outputted the probability of each pixel to distinguish whether it was on atom column or background. Then Laplacian of Gaussian blob detection was used to extract the center of each circular atom [82] and ellipse fitting method to identify dumbbells [93]. Moreover, based on the detected atoms, threefold and fourfold Si atoms could be characterized using graph representation [82]. With the structural information extracted by deep learning models, patterns and configurations were further derived according to the achieved material-specific information and the priori knowledge of certain materials. What is more, a library of defects could also be constructed [95], in order to help with the research on materials functionality.

Maksov et al. performed unsupervised classification and analysis of the defect structures and phase evolution in dynamic STEM imaging of layered WS₂ [84]. As shown in Fig. 12, the authors trained a neural network with the first frame of a movie on the motion of Mo-doped WS₂ under electron beam to detect defects,

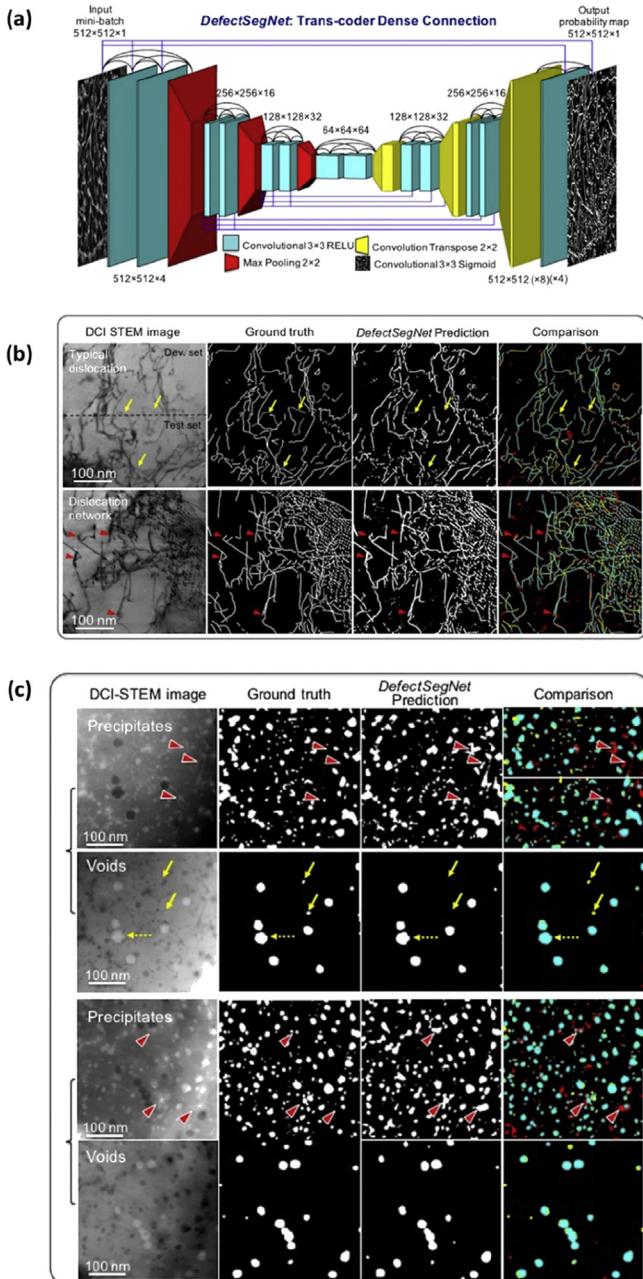


Fig. 10. Segmentation and classification of crystallographic defects in the STEM images of structural alloy [74]. DefectSegNet was proposed for the task, as shown in (a). (b) shows the segmentation results of line dislocations and (c) are the results of precipitates and voids on structural alloys. The color turquoise denotes TP, black denotes TN, red denotes FP and yellow denotes FN. The yellow arrows in (b) present uncommon dislocation lines and red ones present the overestimation of FP. The yellow arrows in (c) denote the overlapped voids with precipitates and red arrows point to the source of FP for precipitate predictions. Note that TP, TN, FP, FN denote true positive, true negative, false positive and false negative respectively.

and generalize the model for the remain frames. Then, the same group used a Gaussian mixture model (GMM) along the spatial and temporal dimensions, to cluster all defects into five groups, which distinguished and provided visualization of the trajectories of these different defects. Among them, two types of defects associated with Mo dopants show peculiar switching behaviors and their symmetry distortions were extracted by analyzing inner structures using

atom finder and principal component analysis (PCA). The two defects types were then classified into four subclasses including undistorted Mo_w defect and three ($\text{Mo}_\text{w} + \text{V}_\text{s}$) complexes. The transition matrix of these four subclasses suggested that Mo_w defects might associate with an S vacancy and Mo dopants might capture S vacancy which were created in the neighborhood during the e-beam irradiation. In this work, the authors showed that they were able to identify dominant point defects, analyze diffusion for the selected defect species, and also study transformation pathways.

Ziatdinov et al. used similar method to study the motion of individual Si atoms in graphene and observed symmetry breaking in the Si–C defect configuration under e-beam manipulations [92]. As shown in Fig. 13, the authors utilized the focused STEM electron beam to activate Si dopant atoms along circular and linear trajectories respectively that had been defined beforehand. A deep convolutional network was developed to reconstruct lattice or impurity atoms with the training dataset which contained 3,000 simulated images. From the extracted atom positions, distribution of lattice was reconstructed, and bond lengths were clustered according to the statistical histogram. From the analyses of Si atom position and bond length, the authors found that the variation of bond length was below 10% along with a dopant atom, which indicated that there was no significant bond distortion associated with the motion of dopant atoms. However, by further analyzing of the distributions of threefold Si–C bond lengths along the time sequence, a symmetry breaking was observed, which might be related to a larger tilt in the sample. As shown in Fig. 14, deep learning methods could be also used to describe thermodynamics and kinetics of beam induced reactions of Si atoms on the edge and in the bulk of graphene [94]. In this case, a network was trained to characterize all the atoms, especially the ones on the edge of nanohole. Then local configurations around Si impurities were cropped and clustered as state descriptors, which further provided a view of transition probabilities between different states. The authors finally identified the stability of 1D ordered Si structure on the edge of a graphene nanohole and the coupling of Si impurities to topological lattice reconstructions in the bulk of graphene.

Another application of deep learning was to infer 3D rotation distortions of atoms through 2D annular bright field (ABF) images [83] as shown in Fig. 15. Previous methods to extract 3D local information were implemented by analyzing column shapes in ABF micrographs with human experts' inspection and priori knowledge on quantitative 3D octahedral rotation information [98], and their performance was limited by the column shapes which were manually identified. To deal with the limitations of traditional methods, Laanait et al. used three angles to represent all the rotations and trained a 12-layer convolutional network with simulated dataset of the prototype perovskite SrTiO_3 oriented in [110] projection, to detect atomic distortions of oxide perovskite materials. The trained neural network showed great performance on extracting symmetry and magnitudes of octahedral rotations on experimental data and well generalized to different chemical compositions from trained data, including CaTiO_3 , $\text{La}_{0.7}\text{Sr}_{0.3}\text{MnO}_3$, $(\text{LaAlO}_3)_{0.3}(\text{Sr}_2\text{AlTaO}_6)_{0.7}$, and $\text{Eu}_{0.7}\text{Sr}_{0.3}\text{MnO}_3$. In above cases, deep learning shows great power on underlying information retrieval which is hard for traditional computer-guided analyzing approaches.

3.3. Analysis of SPM images

Scanning probe microscopy (SPM) is a class of tools for imaging surfaces from a region of nanometers to millimeters [4]. To form images, a scanning probe microscope scan its tip over the surface and detects electronic feedback signals with either constant

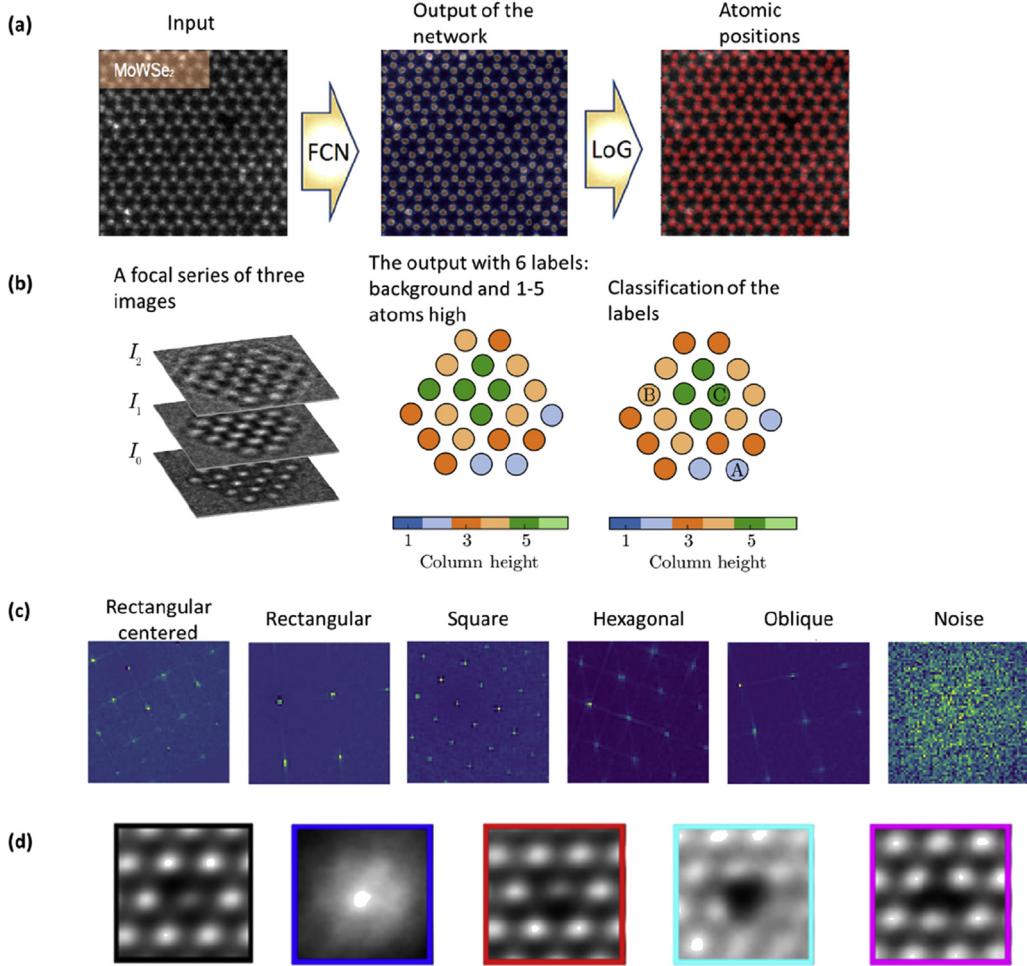


Fig. 11. In the task of atomic resolution STEM imaging analysis, deep learning network can be used to retrieve the structural information, including (a) atomic positions [82], (b) column heights [80], (c) lattice types [81] and (d) defect types [84].

interaction mode or constant height mode. The common SPM approaches include scanning tunneling microscopy (STM), atomic force microscopy (AFM) etc. In fact, many deep learning-based methods, which are derived based on STEM datasets, can also be generalized to the tasks on SPM data [81,82,93], which is probably another evidence of the power of deep learning on image analyses.

It is well known that the tip size of SPM is crucial for the quality and spatial resolution of SPM images. Misused tip may generate artifacts on the final image [99]. One application is automatic tip sharpening in SPM images, mainly on hydrogen terminated (100) Si substrate [75], as shown in Fig. 16. Techniques of *in-situ* tip conditioning is important for atomic-scale characterization and manipulation. Thus Rashidi et al. proposed a neural network, which was trained with about 3,500 STM images of silicon dangling bonds, and the autonomous tips reconditioning could achieve an accuracy of 99%.

Another utilization is the identification and classification of surface molecule structural and rotational states [100], as shown in Fig. 17. By combining the Markov network with convolutional neural network, the framework could identify each molecule to be bowl-up or bowl-down state and classify the rotation into four classes. The neural network was trained on synthetic STM images, which were generated using Markov chain Monte Carlo sampler with density functional theory calculations of electronic charge density distribution as inputs. The trained model was also tested on

experimental STM data of buckybowls on gold (111). By analyzing all the pair distributions of the molecular states, an association between the switching of bowl-up/down states and formation of rotational disorder in the inverted molecules could be observed. A 'two-stage' reaction path for bowl inversion and its effect on the neighboring molecules was also analyzed by exploring the output local correlations. The study shows that neural networks made the discovery of hidden properties more efficient.

Neural networks can also help with modeling the connection between the thermodynamic conditions with STM images. Vlcek et al. developed a framework to analyze elemental segregation in a $\text{La}_{1-x}\text{Ca}_x\text{MnO}_3$ (LCMO) thin film [101], as shown in Fig. 18. The authors first proposed a statistical model of an equilibrium system, which contained only three adjustable variables to predict 3D structures given STM images and electron spectroscopy results. Based on previous studies of LCMO, they developed an equation of the energy u_i of configuration i of the thin-layer system:

$$u_i = w_{CL} \sum_{\{C,L\}} \delta_i^{b,CL} + w_{LL} \sum_{\{L,L\}} \delta_i^{s,LL} + w_{LO} \sum_{\{L,O\}} \delta_i^{s,LO} \quad (6)$$

where $C(\text{Ca}^{2+})$ and $L(\text{La}^{3+})$ denoted metal cations and O denoted surface oxygens. The indicator b and s denoted whether there was at least one of the atoms belonged to the bulk (b) or both belonged to the surface layer (s). δ^{XY} was 1 when the particle was X or Y ,

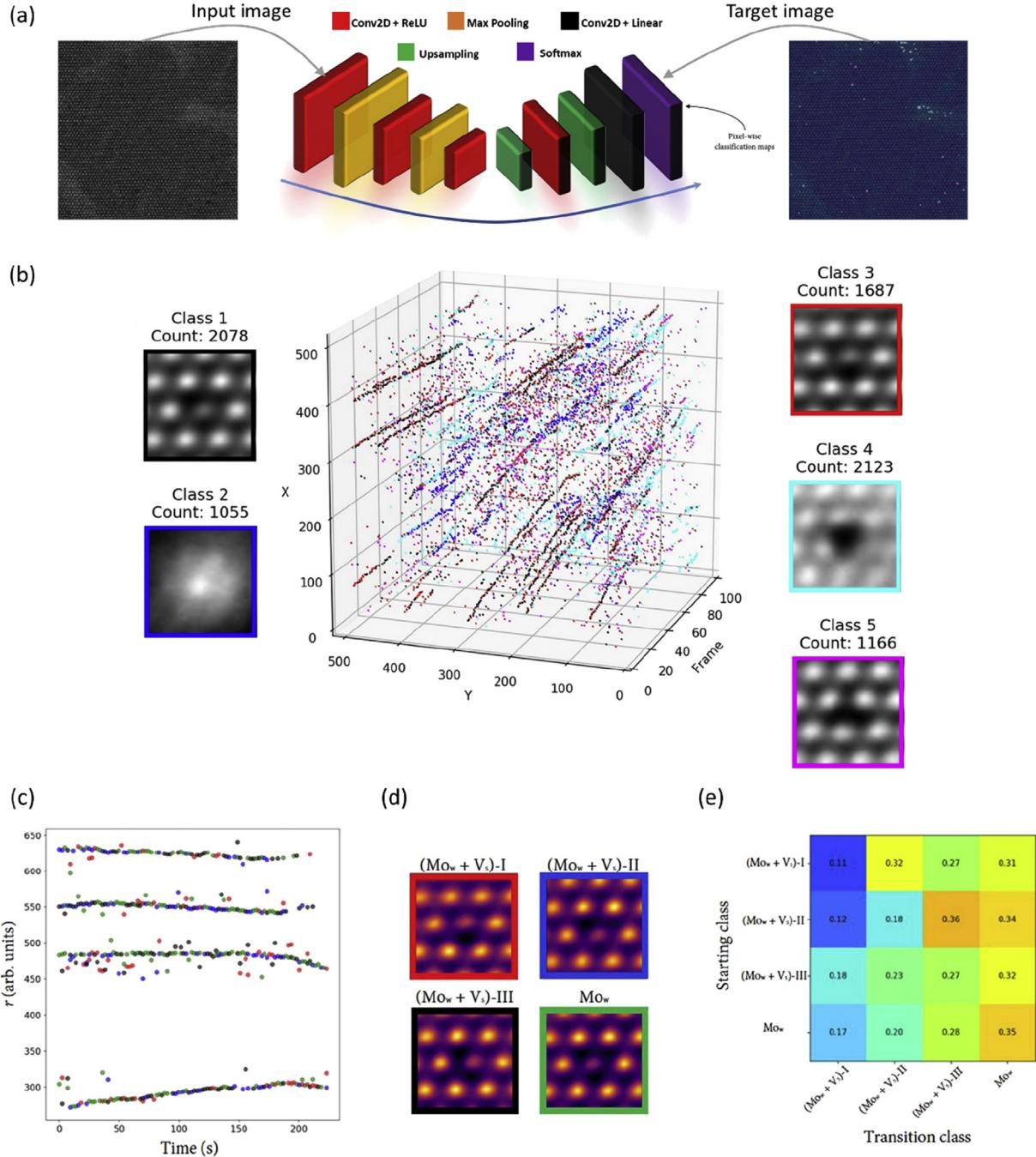


Fig. 12. By the utilization of deep learning methods and statistical algorithms, lattice defects and phase evolution can be efficiently extracted and analyzed from dynamic atomic-resolution STEM imaging on layered WS₂ [84]. (a) shows the architecture of the neural network. The input is the first frame of the movie and the target image is generated by FFT and filtering. The trained network can output a pixel-wise classification of defects. (b) denotes the Gaussian mixture model (GMM) unmixing results of all the detected defects and their trajectories along the time axis. There are mainly three types of the trajectories: weakly moving trajectories, stronger diffusion and flickering. Among them, class 1 and class 3 have shown relatively continuous behavior. The two classes associated with Mo dopant are classified into four subclasses, including undistorted Mo_w defect and three (Mo_w + V_s) complexes, shown in (d). (c) shows the transition flows and transition matrix of the four subclasses.

otherwise it was 0. The three variables w_{CL} , w_{LL} , w_{LO} presented the energy of interactions between the nearest-neighbor pairs of Ca²⁺ and La³⁺ cations in bulk, between the nearest and the next nearest neighbor pairs of La³⁺ cations in the surface unshielded by surface oxygens, and between La³⁺ and the nearest surface. The learned parameters were given in reduced units: ($w' = w/k_B T$) $w'_{CL} = 0.194$, $w'_{LL} = 0.739$, $w'_{LO} = -0.890$, which could be interpreted in terms of elastic and electrostatic effective interactions. With the

parameters, 3D microstructures were generated by canonical Monte Carlo methods. Based on the optimized statistical model, connection between structural configurations and thermodynamic conditions, such as temperature or chemical potential, could be explored with a generative model called Variational Auto-Encoder (VAE) [102]. VAE enabled to encode atomic configurations into a few numbers of latent parameters and the relation between the encoded latent parameters with temperature and compositions

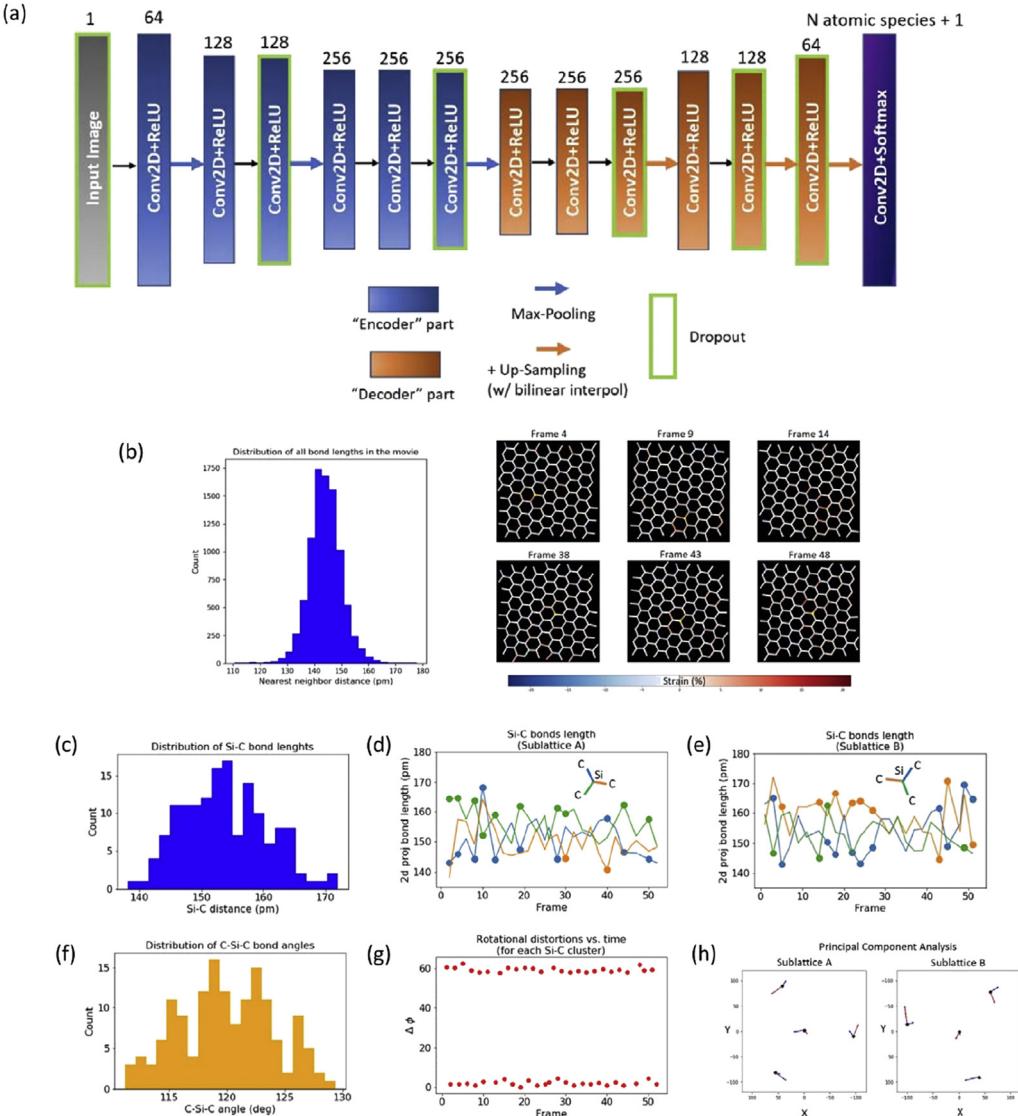


Fig. 13. Illustration of a learning methods based framework to analyze structural evolution and possible defects configuration during e-beam induced Si atom motion in graphene [92]. (a) shows the deep learning model used in the work, which outputs the probability map of different types of atoms, including lattice atoms, impurities etc. Based on the pixel-wise classification of atoms, the lattices are reconstructed by nearest atoms searching algorithms and all the bond lengths are computed. (b) denotes the clustering of bond lengths in a histogram and the generated real-space strain map. Strain is defined as $s = (a - a_0)/a_0$, where a_0 is the mean value of the histogram. The Si atoms are colored as yellow. (c) is the histogram of all the Si-C bond lengths, where the Si-C cluster is defined as the Si impurity atom with its lattice atoms in the first coordinate sphere. (d) and (e) show the variation of the bond lengths of two Si-C sublattice along the time. The mean values for the three bonds of the sublattice in (d) are 151 pm, 150 pm and 157 pm. The values for the sublattice in (e) are 155 pm, 159 pm and 153 pm. (f) is the distribution of C-Si-C bond angles for all the frames. A rotation distortion of Si-C cluster is plotted in (g). From (d) and (e), a symmetry breaking in Si-C can be observed. In order to discover the distortion mode, PCA is performed on the two sublattice and the first two components are shown in (h). However, no presence of a dominant symmetry breaking mode is observed from the analysis.

could be observed. The experimental results showed clear trends between latent parameters, temperature and compositions when the Ca concentration value was 0.5, indicating that there was entropic stabilization at a 0.5 surface stoichiometry. By modeling and exploring the linkage between materials' properties and experimental conditions, neural networks may help with the process of designing materials with special needs.

4. Summary and outlook

In this review, we summarized the works for applications of deep learning on different kinds of microscopic imaging analysis. Based on the published reports, we propose a general workflow for imaging analysis with deep learning, which includes specific task

analysis, preparation of work, model design, feature analysis and validation. So far, the applications of deep learning have mainly focused on structural information extraction, such as recognition and tracking of morphology, phase, defect and so on. Moreover, generative models are also used by a few works, including super-resolution reconstruction and linkage construction between experimental conditions and microstructures. Generative models [91,102] are powerful tools to model the connections between inputs and outputs and can also help with dimension reduction which may be useful in theoretical calculation.

Looking into the future, the rising of utilizing deep learning methods for microscopic imaging analysis will present both challenges and opportunities. As noted above, most of the existed works utilized similar models and architectures, which are

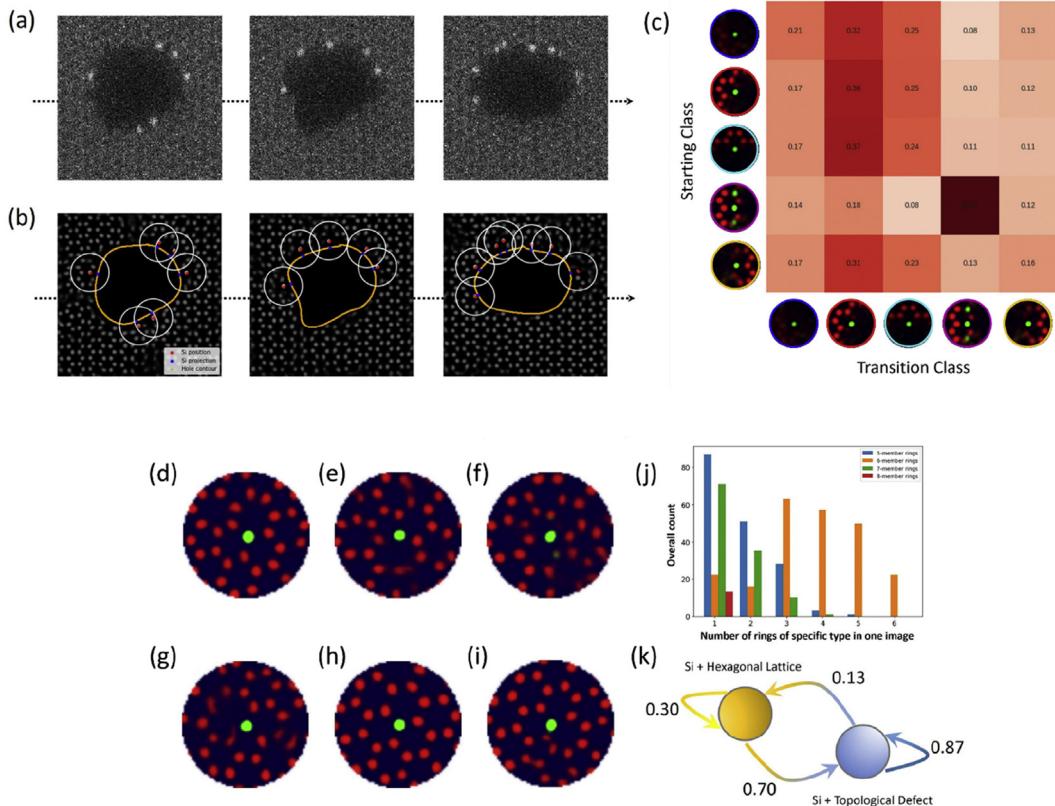


Fig. 14. Neural network can help with the study of the dynamic behavior of e-beam irradiated Si atoms in the bulk and on the edges of nanohole [94]. First, a neural network is designed and trained to localize and identify the types of atoms. The training data is simulated with three types of graphene edges, including zigzag, armchair and bearded edge. Along the edges of the simulated data, vacancies and substitutional Si impurities are introduced randomly. (a) illustrates the evolution of Si dopants, driven by 60 keV electron beam irradiation. (b) shows the respective outputs of the network, where blue dots denote the Si atoms on the edge of the nanohole and white circles denote the neighborhood of the Si atoms. The configurations of Si atoms and the neighborhood can be denoted as state descriptors. All the states around the edges can be clustered into 16 groups using GMM unmixing methods and a further classification is performed with structural similarity search to categorize the 16 groups into 5, shown in (c). The green dots denote Si atoms and the red ones denote C atoms. (c) shows the transition matrix between the 5 states. The analysis method for Si atoms in the bulk is using ring network. (d–i) denote the 5-, 6-, 7- and 8-member rings around a Si impurity. Topological defects can be observed from all the cropped images, except for (h). Statistical analysis is plotted in (j) and (k) shows the transition probabilities of switching between states with and without topological defects.

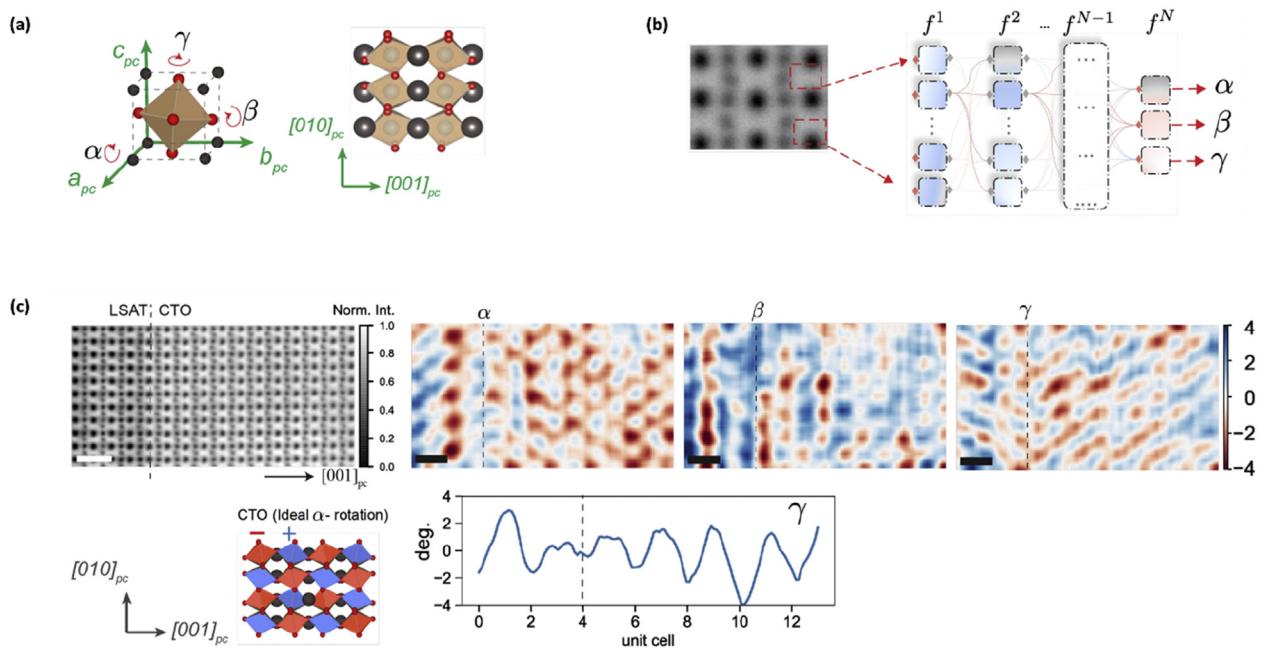


Fig. 15. Rotation reconstructions of oxygen octahedra from atomic ABF-STEM [83]. All the rotations are characterized as three angles indicated in the atomic model of (a). (b) represents the 12-layer network used for analysis, by generating parameters of three angles from an ABF-STEM image. (c) shows the results on an ABF-STEM image of a CaTiO_3 (CTO) epitaxial thin film on a single crystal substrate ($\text{LaAlO}_3_{0.3}(\text{Sr}_2\text{AlTaO}_6)_{0.7}$ (LSAT)). The output shows that the network can correctly output the octahedral rotation symmetry.

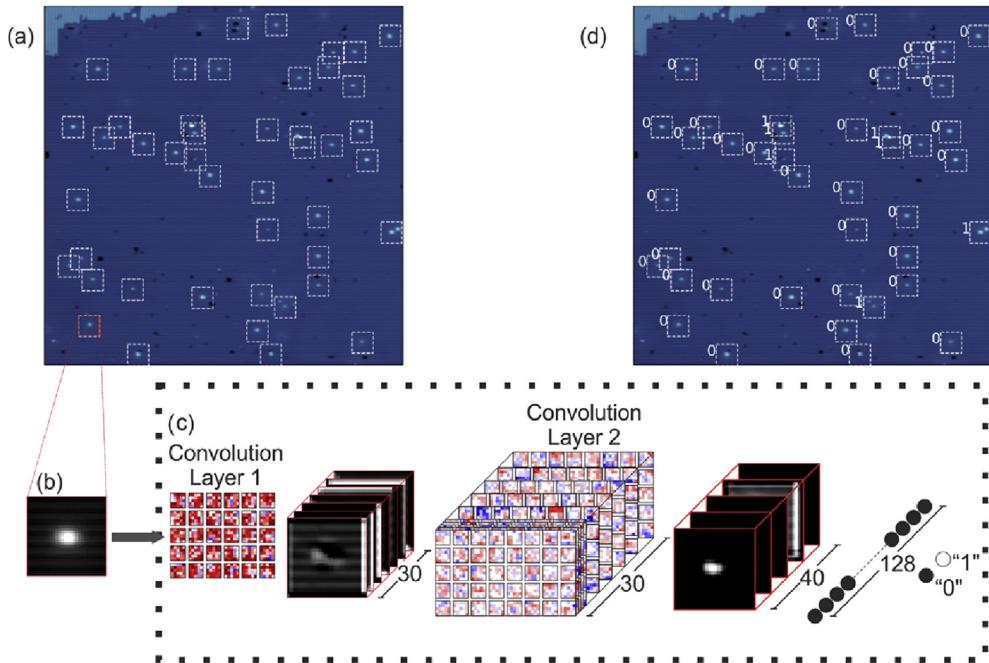


Fig. 16. Automatic tip sharpening in STM images with deep learning [75]. (a) is the overall raw image and (b) is the cropped image for analysis. (c) is the workflow of tip sharpening. Labels with 0 denote for sharp tips and labels with 1 are for double tips. (d) is the labeled image of (a).

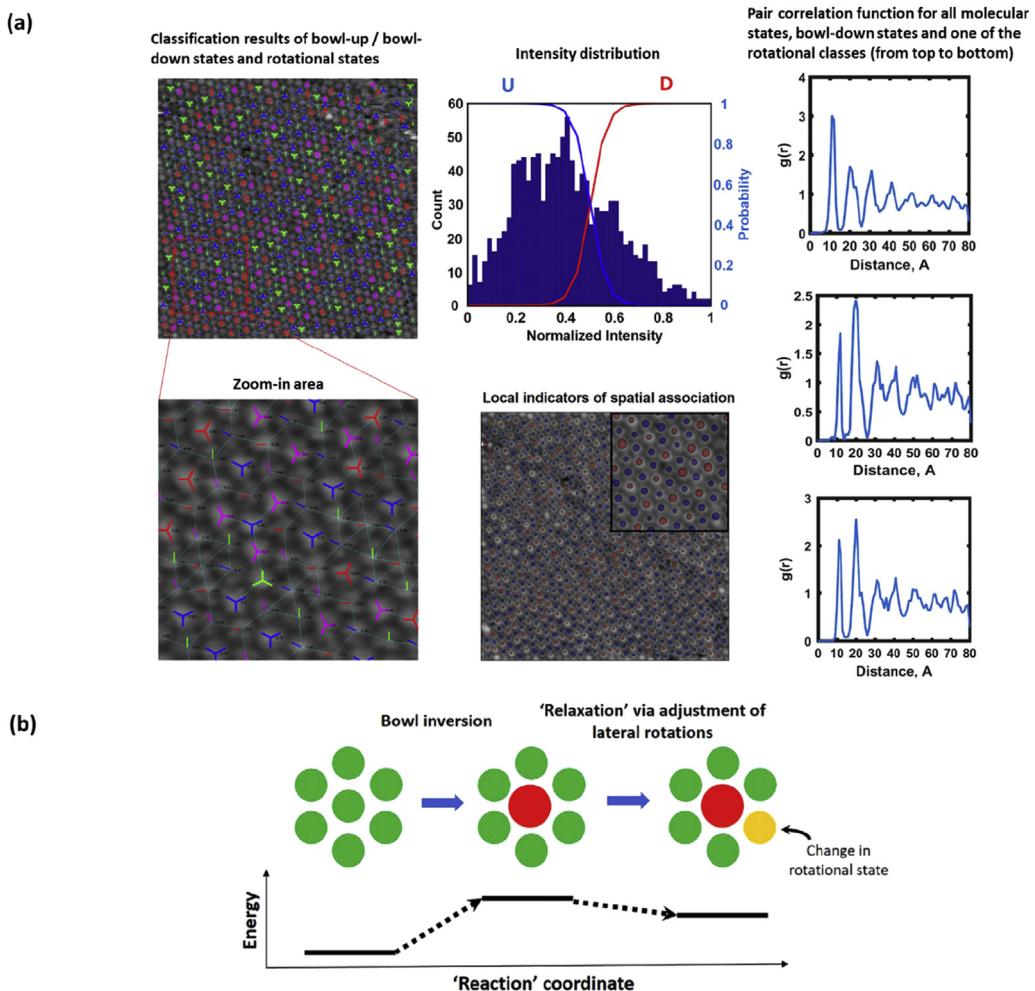


Fig. 17. Analysis of STM images of buckybowls on gold (111), with surface molecular structural and rotational states identification and classification [100]. (a): Experimental results on buckybowls on gold (111) are shown. The left column denotes the classification results of all the structural and rotational states and the zoom-in area of the red box. Based on the results, SPM intensity of all the identified molecules can be analyzed and the spatial correlations between each nearby molecule is calculated using Moran's I method. As shown in the middle bottom, the red and blue circles denote bowl-down and bowl-up states. The right column shows the pair distribution functions for all the states. By exploring the local correlations between bowl states and the rotational states disorder of the neighboring molecules, a two-stage' reaction path was proposed, as shown in (b).

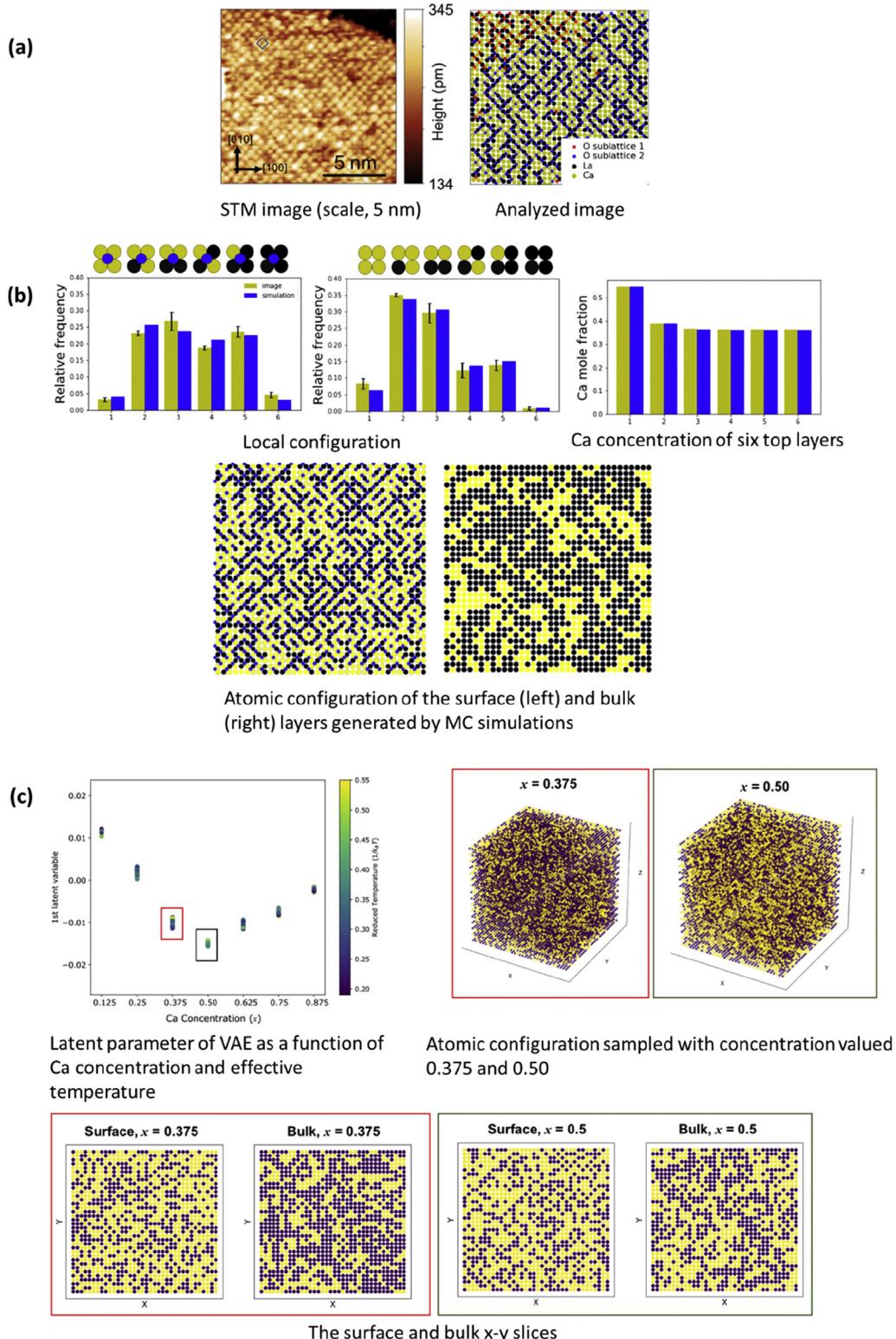


Fig. 18. A framework that uses a statistical model of an equilibrium system to predict 3D microstructural with STM images and X-ray photoelectron spectroscopy (XPS) measurements and a deep generative model called VAE to analyze the conditions region of entropic stability on $\text{La}_{5/8}\text{Ca}_{3/8}\text{MnO}_3$ thin film [101]. The authors first inferred atomic structures through STM images using principal component analysis (PCA) filtering and a motif matching approach, with the additional constraints including total stoichiometry and XPS, as shown in (a), where black, yellow and blue denote La^{3+} , Ca^{2+} and surface O respectively. Then the local configurations are presented using histograms as signatures or fingerprints, to illustrate interactions between atoms. The histograms of the chosen local configurations and the respective snapshots of atomic configurations in the top layer and bulk are shown in (b). The atomic configurations are generated using MC simulations and the fingerprints of the simulated results are shown in blue in the histograms. (c) shows the analysis results of VAE, with top left presenting the latent parameters as a function of concentration and temperature.

generally auto-encoders composed of stacked blocks of convolutional layers, normalization layers and activation layers. As the application of deep learning is blooming, there are other powerful algorithms, like Long Short Term Memory (LSTM) [76] or Graph Neural Networks (GNN) [103], for tackling complex data structure of materials research data. It is commonsense that data scientists and materials scientists or microscopists have to work together to apply deep learning tool to solve materials questions, which means that both parties have to learn the knowledge or code, paradigm or architecture and other fields' languages. It would be a painful, challenging but exciting intellectual journey. We hope our previous review could give some hints on the background of deep learning. For a microscopist, to grasp some popular software and use it for some problems can be a good start. On the other hand, one important task for the expert of deep learning is to develop user-friendly software/platform and provide needed trainings for microscopists. These software/platform should be both enough specializing into certain materials subjects and enough generalizing into images with different experimental conditions.

Deep learning holds the promises of efficiently retrieving hidden information from massive data. *In-situ* TEM enables to record the structural changes as a function of time, and generate 3–4 orders of magnitude data than the *ex-situ* TEM works [2,104–108]. Owing to this urgent demand, AI techniques including deep learning would be rapidly developed and applied for *in-situ* TEM in near future. Deep learning approaches have also been utilized in crystallography analysis in reciprocal space, for example, automatic classification of crystal structures from electron diffractions [109], auto analysis of convergent beam electron diffraction patterns [110,111] and specific structure-performance correlations in lithium ion conductors [6]. Because of the physical correlation between electron crystallography and microscopic analysis, the deep learning analyses of these two different kinds of subjects can also be correlated to reveal a full structure-property scenario of materials. Due to limited space, this review does not include the application of deep learning on X-ray based microscopic imaging techniques [112]. The deep learning tools have been intensively utilized on the X-ray techniques for medical sciences which have been intensively reviewed.

Although deep learning is particularly powerful in this big data era for its ability of learning to identify patterns and extract structures from complex data, we need to pay attention on the operation in real cases. First, the design of model and loss function should reflect the basic physical and chemical principles, as the way that several frontier researches work on the relationship between computer science and physical sciences [113]. The *a priori* knowledge is necessary for meaningful analysis. The achieved configurations from a trained model should always be compared with the results of other confirmed physical constraints. Second, the model should be run with testing cases which helped us to judge the validity. In other words, it is necessary to generate meaningful ground truth for network training. As mentioned in section 2.2, labelling data by human is time consuming and most of the training datasets mentioned in this review were generated by simulation. However, simulated data is usually man-made cases and have certain discrepancy from real subjects. Particular carefulness is necessary to make the neural network well adapted to real experimental data. With more and more data being collected by different labs in the field, it is plausible to acquire the imaging data into a database and hopefully to retrieve further information. The successful utilization of deep learning on biology and medical image processing encourages the development of deep leaning on the microscopic imaging processing, which may truly advance the field of materials science. In addition, controlled experiments can be designed and implemented according to the information achieved

from deep learning models to verify trained model or discover specific correlations.

As an emerging field, materials informatics has been proposed aiming to go beyond the traditional paradigm of design and discovery of materials [114,115]. Similar to the concept of informatics, which was established in fields such as biology, drug discovery, astronomy, etc., materials informatics is a combination of information science and materials science and holds the promises of speeding up the materials research. For example, several *a b initio* calculation databases have been built in the field of batteries for providing theoretical guidance for materials explorations [116,117]. As an essential technique in the research loop of materials informatics, deep learning analysis has unique capabilities in dealing with massive experimental results and will be more and more important in materials science.

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.

Acknowledgments

This work is supported by Chinese National Natural Science Foundation (U1931202, 61532018), and Science and Technology Foundation of Beijing Municipal Science & Technology Commission, China.

References

- [1] D.B. Williams, C.B. Carter, The transmission electron microscope, in: *Transm. Electron Microsc. Textb. Mater. Sci.*, Springer US, Boston, MA, 1996, pp. 3–17, https://doi.org/10.1007/978-1-4757-2519-3_1.
- [2] J. Li, G. Johnson, S. Zhang, D. Su, In situ transmission electron microscopy for energy applications, *Joule* 3 (2019) 4–8, <https://doi.org/10.1016/j.joule.2018.12.007>.
- [3] J.I. Goldstein, D.E. Newbury, J.R. Michael, N.W. Ritchie, J.H.J. Scott, D.C. Joy, *Scanning Electron Microscopy and X-Ray Microanalysis*, Springer, 2017.
- [4] R. Wiesendanger, W. Roland, *Scanning Probe Microscopy and Spectroscopy: Methods and Applications*, Cambridge University Press, 1994.
- [5] S.V. Kalinin, A.R. Lupini, O. Dyck, S. Jesse, M. Ziatdinov, R.K. Vasudevan, Lab on a beam—big data and artificial intelligence in scanning transmission electron microscopy, *MRS Bull.* 44 (2019) 565–575, <https://doi.org/10.1557/mrs.2019.159>.
- [6] Y. Zhang, X. He, Z. Chen, Q. Bai, A.M. Nolan, C.A. Roberts, D. Banerjee, T. Matsunaga, Y. Mo, C. Ling, Unsupervised discovery of solid-state lithium ion conductors, *Nat. Commun.* 10 (2019) 1–7, <https://doi.org/10.1038/s41467-019-13214-1>.
- [7] S. Somnath, C.R. Smith, S.V. Kalinin, M. Chi, A. Borisevich, N. Cross, G. Duscher, S. Jesse, Feature extraction via similarity search: application to atom finding and denoising in electron and scanning probe microscopy imaging, *Adv. Struct. Chem. Imaging* 4 (2018) 3, <https://doi.org/10.1186/s40679-018-0052-y>.
- [8] L.M. Ghiringhelli, J. Vybiral, E. Ahmetcik, R. Ouyang, S.V. Levchenko, C. Draxl, M. Scheffler, Learning physical descriptors for materials science by compressed sensing, *New J. Phys.* 19 (2017), 023017, <https://doi.org/10.1088/1367-2630/aa57bf>.
- [9] S.V. Kalinin, B.G. Sumpter, R.K. Archibald, Big-deep-smart data in imaging for guiding materials design, *Nat. Mater.* 14 (2015) 973–980, <https://doi.org/10.1038/nmat4395>.
- [10] T.M. Dieb, K. Tsuda, *Machine Learning-Based Experimental Design in Materials Science*, in: *Nanoinformatics*, Springer, Singapore, 2018, pp. 65–74.
- [11] E. Alpaydin, *Introduction to Machine Learning*, MIT Press, 2020.
- [12] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, *Nature* 521 (2015) 436–444, <https://doi.org/10.1038/nature14539>.
- [13] L. Deng, D. Yu, others, Deep learning: methods and applications, *Found. Trends® Signal Process.* 7 (2014) 197–387, <https://doi.org/10.1561/2000000039>.
- [14] S. Russell, P. Norvig, *Artificial Intelligence: A Modern Approach*, third ed., Prentice Hall Press, USA, 2009.
- [15] S. Vosoughi, P. Vijayaraghavan, D. Roy, Tweet2Vec: learning tweet embeddings using character-level CNN-LSTM encoder-decoder, in: *Proc. 39th Int. ACM SIGIR Conf. Res. Dev. Inf. Retr.*, Association for Computing Machinery,

- New York, NY, USA, 2016, pp. 1041–1044, <https://doi.org/10.1145/2911451.2914762>.
- [16] P.-S. Huang, X. He, J. Gao, L. Deng, A. Acero, L. Heck, Learning deep structured semantic models for web search using clickthrough data, in: Proc. 22nd ACM Int. Conf. Inf. Knowl. Manag. Association for Computing Machinery, New York, NY, USA, 2013, pp. 2333–2338, <https://doi.org/10.1145/2505515.2505665>.
- [17] A. Graves, A. Mohamed, G. Hinton, Speech recognition with deep recurrent neural networks, in: 2013 IEEE Int. Conf. Acoust. Speech Signal Process., IEEE, 2013, pp. 6645–6649, <https://doi.org/10.1109/ICASSP.2013.6638947>.
- [18] Z. Zhao, B. Zhao, F. Su, Person re-identification via integrating patch-based metric learning and local salience learning, *Pattern Recogn.* 75 (2018) 90–98, <https://doi.org/10.1016/j.patcog.2017.03.023>.
- [19] Y. Wu, M. Schuster, Z. Chen, Q.V. Le, M. Norouzi, W. Macherey, M. Krikun, Y. Cao, Q. Gao, K. Macherey, others, Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation, *ArXiv Prepr.*, 2016, ArXiv160908144.
- [20] A. Agrawal, A. Choudhary, Deep materials informatics: applications of deep learning in materials science, *MRS Commun.* (2019) 1–14, <https://doi.org/10.1557/mrc.2019.73>.
- [21] J. Schmidt, M.R. Marques, S. Botti, M.A. Marques, Recent advances and applications of machine learning in solid-state materials science, *Npj Comput. Mater.* 5 (2019) 1–36, <https://doi.org/10.1038/s41524-019-0221-0>.
- [22] J.M. Rickman, T. Lookman, S.V. Kalinin, Materials informatics: from the atomic-level to the continuum, *Acta Mater.* (2019), <https://doi.org/10.1016/j.actamat.2019.01.051>.
- [23] R.K. Vasudevan, K. Choudhary, A. Mehta, R. Smith, G. Kusne, F. Tavazza, L. Vlcek, M. Ziatdinov, S.V. Kalinin, J. Hattrick-Simpers, Materials science in the artificial intelligence age: high-throughput library generation, machine learning, and a pathway from correlations to the underpinning physics, *MRS Commun.* 9 (2019) 821–838, <https://doi.org/10.1557/mrc.2019.95>.
- [24] J. Wei, X. Chu, X.-Y. Sun, K. Xu, H.-X. Deng, J. Chen, Z. Wei, M. Lei, Machine learning in materials science, *InfoMat* 1 (2019) 338–358, <https://doi.org/10.1002/inf2.12028>.
- [25] S.V. Kalinin, O. Dyck, N. Balke, S. Neumayer, W.-Y. Tsai, R. Vasudevan, D. Lingerfelt, M. Ahmadi, M. Ziatdinov, M.T. McDowell, others, Toward electrochemical studies on the nanometer and atomic scales: progress, challenges, and opportunities, *ACS Nano* 13 (2019) 9735–9780, <https://doi.org/10.1021/acsnano.9b02687>.
- [26] K. Ryan, J. Lengyel, M. Shatruk, Crystal structure prediction via deep learning, *J. Am. Chem. Soc.* 140 (2018) 10158–10168, <https://doi.org/10.1021/jacs.8b03913>.
- [27] X. Li, Y. Zhang, H. Zhao, C. Burkhardt, L.C. Brinson, W. Chen, A transfer learning approach for microstructure reconstruction and structure-property predictions, *Sci. Rep.* 8 (2018), <https://doi.org/10.1038/s41598-018-31571-7>.
- [28] Z. Yang, X. Li, L.C. Brinson, A.N. Choudhary, W. Chen, A. Agrawal, Microstructural materials design via deep adversarial learning methodology, *J. Mech. Des.* 140 (2018) 111416, <https://doi.org/10.1115/1.4041371>.
- [29] D. Jha, L. Ward, A. Paul, W. Liao, A. Choudhary, C. Wolverton, A. Agrawal, Elemnet: deep learning the chemistry of materials from only elemental composition, *Sci. Rep.* 8 (2018) 17593, <https://doi.org/10.1038/s41598-018-35934-y>.
- [30] W. Ye, C. Chen, Z. Wang, I.-H. Chu, S.P. Ong, Deep neural networks for accurate predictions of crystal stability, *Nat. Commun.* 9 (2018) 3800, <https://doi.org/10.1038/s41467-018-06322-x>.
- [31] T. Xie, J.C. Grossman, Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties, *Phys. Rev. Lett.* 120 (2018) 145301, <https://doi.org/10.1103/PhysRevLett.120.145301>.
- [32] Z. Yang, Y.C. Yabansu, R. Al-Bahrani, W. Liao, A.N. Choudhary, S.R. Kalidindi, A. Agrawal, Deep learning approaches for mining structure-property linkages in high contrast composites from simulation datasets, *Comput. Mater. Sci.* 151 (2018) 278–287, <https://doi.org/10.1016/j.commatsci.2018.05.014>.
- [33] A. Cecen, H. Dai, Y.C. Yabansu, S.R. Kalidindi, L. Song, Material structure-property linkages using three-dimensional convolutional neural networks, *Acta Mater.* 146 (2018) 76–84, <https://doi.org/10.1016/j.actamat.2017.11.053>.
- [34] Z. Yang, Y.C. Yabansu, D. Jha, W. Liao, A.N. Choudhary, S.R. Kalidindi, A. Agrawal, Establishing structure-property localization linkages for elastic deformation of three-dimensional high contrast composites using deep learning approaches, *Acta Mater.* 166 (2019) 335–345, <https://doi.org/10.1016/j.actamat.2018.12.045>.
- [35] C. Hu, Y. Zuo, C. Chen, S. Ping Ong, J. Luo, Genetic algorithm-guided deep learning of grain boundary diagrams: addressing the challenge of five degrees of freedom, *Mater. Today* (2020), <https://doi.org/10.1016/j.mattod.2020.03.004>.
- [36] R. Liu, A. Agrawal, W. Liao, A. Choudhary, M. De Graef, Materials discovery: understanding polycrystals from large-scale electron patterns, in: 2016 IEEE Int. Conf. Big Data Big Data, IEEE, 2016, pp. 2261–2269, <https://doi.org/10.1109/BigData.2016.7840857>.
- [37] D. Jha, S. Singh, R. Al-Bahrani, W. Liao, A. Choudhary, M. De Graef, A. Agrawal, Extracting grain orientations from ebsd patterns of polycrystalline materials using convolutional neural networks, *Microsc. Microanal.* 24 (2018) 497–502, <https://doi.org/10.1017/S1431927618015131>.
- [38] B.L. DeCost, B. Lei, T. Francis, E.A. Holm, High throughput quantitative metallography for complex microstructures using deep learning: a case study in ultrahigh carbon steel, *Microsc. Microanal.* 25 (2019) 21–29, <https://doi.org/10.1017/S1431927618015635>.
- [39] R.M. Patton, J.T. Johnston, S.R. Young, C.D. Schuman, D.D. March, T.E. Potok, D.C. Rose, S.-H. Lim, T.P. Karnowski, M.A. Ziatdinov, others, 167-PFlops deep learning for electron microscopy: from learning physics to atomic manipulation, in: Proc. Int. Conf. High Perform. Comput. Netw. Storage Anal., IEEE Press, 2018, p. 50, <https://doi.org/10.1109/SC.2018.00053>.
- [40] I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning*, MIT Press, 2016.
- [41] S. Krige, Ground truth data, content, metrics, and analysis, in: *Comput. Vis. Metr. Surv. Taxon. Anal.*, Apress, Berkeley, CA, 2014, pp. 283–311, https://doi.org/10.1007/978-1-4302-5930-5_7.
- [42] B.D. Ripley, *Pattern Recognition and Neural Networks*, Cambridge University Press, 2007.
- [43] R.H. Hahnloser, R. Sarpeshkar, M.A. Mahowald, R.J. Douglas, H.S. Seung, Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit, *Nature* 405 (2000) 947, <https://doi.org/10.1038/35016072>.
- [44] I. Sutskever, J. Martens, G. Dahl, G. Hinton, On the importance of initialization and momentum in deep learning, *Int. Conf. Mach. Learn.* (2013) 1139–1147.
- [45] D.P. Kingma, J. Ba, Adam: A Method for Stochastic Optimization, *ArXiv Prepr.*, 2014, ArXiv14126980.
- [46] A. Graves, Generating Sequences with Recurrent Neural Networks, *ArXiv Prepr.*, 2013, ArXiv1308050.
- [47] S. Belharbi, Neural Networks Regularization through Representation Learning, 2018, CoRR, abs/1807.05292, <http://arxiv.org/abs/1807.05292>.
- [48] Y. Bengio, A. Courville, P. Vincent, Representation learning: a review and new perspectives, *IEEE Trans. Pattern Anal. Mach. Intell.* 35 (2013) 1798–1828, <https://doi.org/10.1109/TPAMI.2013.50>.
- [49] A. Mahendran, A. Vedaldi, Visualizing deep convolutional neural networks using natural pre-images, *Int. J. Comput. Vis.* 120 (2016) 233–255, <https://doi.org/10.1007/s11263-016-0911-8>.
- [50] J.L. Elman, Finding structure in time, *Cognit. Sci.* 14 (1990) 179–211, https://doi.org/10.1207/s15516709cog1402_1.
- [51] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, others, Imagenet large scale visual recognition challenge, *Int. J. Comput. Vis.* 115 (2015) 211–252, <https://doi.org/10.1007/s11263-015-0816-y>.
- [52] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, C.L. Zitnick, Microsoft coco: common objects in context, in: *Eur. Conf. Comput. Vis.*, Springer, 2014, pp. 740–755, https://doi.org/10.1007/978-3-319-10602-1_48.
- [53] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 770–778, <https://doi.org/10.1109/CVPR.2016.90>.
- [54] K. Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, *ArXiv Prepr.*, 2014, ArXiv14091566.
- [55] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, in: *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 1–9, <https://doi.org/10.1109/CVPR.2015.7298594>.
- [56] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Rethinking the inception architecture for computer vision, in: *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 2818–2826, <https://doi.org/10.1109/CVPR.2016.308>.
- [57] C. Szegedy, S. Ioffe, V. Vanhoucke, A.A. Alemi, Inception-v4, inception-resnet and the impact of residual connections on learning, in: *Thirty-First AAAI Conf. Artif. Intell.*, 2017.
- [58] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 3431–3440, <https://doi.org/10.1109/TPAMI.2016.2572683>.
- [59] K. He, G. Gkioxari, P. Dollár, R. Girshick, Mask r-cnn, in: *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 2961–2969, <https://doi.org/10.1109/ICCV.2017.322>.
- [60] Y. Gal, Uncertainty in deep learning, *Univ. Camb.* 1 (2016) 3.
- [61] Y. LeCun, 1.1 deep learning hardware: past, present, and future, in: *2019 IEEE Int. Solid-State Circuits Conf.-ISSCC*, IEEE, 2019, pp. 12–19, <https://doi.org/10.1109/ISSCC.2019.8662396>.
- [62] S. Chetlur, C. Woolley, P. Vandermersch, J. Cohen, J. Tran, B. Catanzaro, E. Shelhamer, cudnn: Efficient Primitives for Deep Learning, *ArXiv Prepr.*, 2014, ArXiv14100759.
- [63] W. Dai, D. Berleant, Benchmarking contemporary deep learning hardware and frameworks: a survey of qualitative metrics, in: *2019 IEEE First Int. Conf. Cogn. Mach. Intell. CogMI*, IEEE, 2019, pp. 148–155, <https://doi.org/10.1109/CogMI48466.2019.00029>.
- [64] H. Touvron, A. Vedaldi, M. Douze, H. Jégou, Fixing the Train-Test Resolution Discrepancy: FixEfficientNet, *ArXiv Prepr.*, 2020, ArXiv200308237.
- [65] A. Kolesnikov, L. Beyer, X. Zhai, J. Puigcerver, J. Yung, S. Gelly, N. Houlsby, Big Transfer (BiT), General Visual Representation Learning, 2019.
- [66] Y. Zhang, L. Peng, B. Li, J.-K. Peir, J. Chen, Performance and power comparisons between Nvidia and ATI GPUs, *Int. J. Comput. Sci. Inf. Technol.* 6 (2014) 1, <https://doi.org/10.5121/ijcsit.2014.6601>.
- [67] A. Canziani, A. Paszke, E. Culurciello, An Analysis of Deep Neural Network Models for Practical Applications, *arXiv* 2016, ArXiv Prepr., 2016, ArXiv160507678.
- [68] S. Bianco, R. Cadene, L. Celona, P. Napoletano, Benchmark analysis of representative deep neural network architectures, *IEEE Access* 6 (2018) 64270–64277, <https://doi.org/10.1109/ACCESS.2018.2877890>.

- [69] M.D. Wilkinson, M. Dumontier, I.J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L.B. da Silva Santos, P.E. Bourne, others, The FAIR Guiding Principles for scientific data management and stewardship, *Sci. Data* 3 (2016), <https://doi.org/10.1038/sdata.2016.18>.
- [70] S. Pouyanfar, S. Sadiq, Y. Yan, H. Tian, Y. Tao, M.P. Reyes, M.-L. Shyu, S.-C. Chen, S. Iyengar, A survey on deep learning: algorithms, techniques, and applications, *ACM Comput. Surv. CSUR* 51 (2019) 92, <https://doi.org/10.1145/3234150>.
- [71] N. Marturi, S. Dembélé, N. Piat, Scanning electron microscope image signal-to-noise ratio monitoring for micro-nanomanipulation, *Scanning: J. Scanning Microsc.* 36 (2014) 419–429, <https://doi.org/10.1002/sca.21137>.
- [72] L. Jones, P.D. Nellist, Identifying and correcting scan noise and drift in the scanning transmission electron microscope, *Microsc. Microanal.* 19 (2013) 1050–1060, <https://doi.org/10.1017/S1431927613001402>.
- [73] C. Shorten, T.M. Khoshgoftaar, A survey on image data augmentation for deep learning, *J. Big Data* 6 (2019) 60, <https://doi.org/10.1186/s40537-019-0197-0>.
- [74] G. Roberts, S.Y. Haile, R. Sainju, D.J. Edwards, B. Hutchinson, Y. Zhu, Deep learning for semantic segmentation of defects in advanced STEM images of steels, *Sci. Rep.* 9 (2019) 1–12, <https://doi.org/10.1038/s41598-019-49105-0>.
- [75] M. Rashidi, R.A. Wolkow, Autonomous scanning probe microscopy in situ tip conditioning through machine learning, *ACS Nano* 12 (2018) 5185–5189, <https://doi.org/10.1021/acsnano.8b02208>.
- [76] S. Hochreiter, J. Schmidhuber, Long short-term memory, *Neural Comput.* 9 (1997) 1735–1780, <https://doi.org/10.1162/neco.1997.9.8.1735>.
- [77] Z. Zhao, W. Chen, X. Wu, P.C. Chen, J. Liu, LSTM network: a deep learning approach for short-term traffic forecast, *IET Intell. Transp. Syst.* 11 (2017) 68–75, <https://doi.org/10.1049/iet-its.2016.0208>.
- [78] X. Duan, L. Wang, C. Zhai, N. Zheng, Q. Zhang, Z. Niu, G. Hua, Joint spatio-temporal action localization in untrimmed videos with per-frame segmentation, in: 2018 25th IEEE Int. Conf. Image Process. ICIP, IEEE, 2018, pp. 918–922, <https://doi.org/10.1109/ICIP.2018.8451692>.
- [79] M.H. Modarres, R. Aversa, S. Cozzini, R. Ciancio, A. Leto, G.P. Brandino, Neural network for nanoscience scanning electron microscope image recognition, *Sci. Rep.* 7 (2017) 13282, <https://doi.org/10.1038/s41598-017-13565-z>.
- [80] J. Madsen, P. Liu, J. Kling, J.B. Wagner, T.W. Hansen, O. Winther, J. Schiøtz, A deep learning approach to identify local structures in atomic-resolution transmission electron microscopy images, *Adv. Theory Simul.* 1 (2018), <https://doi.org/10.1002/adts.201800037>.
- [81] R.K. Vasudevan, N. Laanait, E.M. Ferragut, K. Wang, D.B. Geohegan, K. Xiao, M. Ziatdinov, S. Jesse, O. Dyck, S.V. Kalinin, Mapping mesoscopic phase evolution during E-beam induced transformations via deep learning of atomically resolved images, *Npj Comput. Mater.* 4 (2018) 30, <https://doi.org/10.1038/s41524-018-0086-7>.
- [82] M. Ziatdinov, O. Dyck, A. Maksov, X. Li, X. Sang, K. Xiao, R.R. Unocic, R. Vasudevan, S. Jesse, S.V. Kalinin, Deep learning of atomically resolved scanning transmission electron microscopy images: chemical identification and tracking local transformations, *ACS Nano* 11 (2017) 12742–12752, <https://doi.org/10.1021/acsnano.7b07504>.
- [83] N. Laanait, Q. He, A.Y. Borisevich, Reconstruction of 3-D Atomic Distortions from Electron Microscopy with Deep Learning, *ArXiv Prepr.*, 2019. [ArXiv1902.06876](https://arxiv.org/abs/1902.06876).
- [84] A. Maksov, O. Dyck, K. Wang, K. Xiao, D.B. Geohegan, B.G. Sumpter, R.K. Vasudevan, S. Jesse, S.V. Kalinin, M. Ziatdinov, Deep learning analysis of defect and phase evolution during electron beam-induced transformations in WS₂, *Npj Comput. Mater.* 5 (2019) 12, <https://doi.org/10.1038/s41524-019-0152-9>.
- [85] J. Snell, K. Swersky, R. Zemel, Prototypical networks for few-shot learning, in: *Adv. Neural Inf. Process. Syst.* 2017, pp. 4077–4087.
- [86] S.M. Azimi, D. Britz, M. Engstler, M. Fritz, F. Mücklich, Advanced steel microstructural classification by deep learning methods, *Sci. Rep.* 8 (2018) 2128, <https://doi.org/10.1038/s41598-018-20037-5>.
- [87] K. de Haan, Z.S. Ballard, Y. Rivenson, Y. Wu, A. Ozcan, Resolution Enhancement in Scanning Electron Microscopy Using Deep Learning, *ArXiv Prepr.*, 2019. [ArXiv1901.11094](https://arxiv.org/abs/1901.11094).
- [88] B.L. DeCost, T. Francis, E.A. Holm, Exploring the microstructure manifold: image texture representations applied to ultrahigh carbon steel microstructures, *Acta Mater.* 133 (2017) 30–40, <https://doi.org/10.1016/j.actamat.2017.05.014>.
- [89] B.L. DeCost, M.D. Hecht, T. Francis, B.A. Webler, Y.N. Picard, E.A. Holm, UHCSDB: UltraHigh carbon steel micrograph database, *Integr. Mater. Manuf. Innov.* 6 (2017) 197–205, <https://doi.org/10.1007/s40192-017-0097-0>.
- [90] A. Bansal, X. Chen, B. Russell, A. Gupta, D. Ramanan, Pixelnet: Representation of the Pixels, by the Pixels, and for the Pixels, *ArXiv Prepr.*, 2017. [ArXiv1702.06506](https://arxiv.org/abs/1702.06506).
- [91] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets, in: *Adv. Neural Inf. Process. Syst.* 2014, pp. 2672–2680.
- [92] M. Ziatdinov, S. Jesse, R.K. Vasudevan, B.G. Sumpter, S.V. Kalinin, O. Dyck, Tracking Atomic Structure Evolution during Directed Electron Beam Induced Si-Atom Motion in Graphene via Deep Machine Learning, *ArXiv Prepr.*, 2018. [ArXiv1809.04785](https://arxiv.org/abs/1809.04785).
- [93] M. Ziatdinov, O. Dyck, A. Maksov, B.M. Hudak, A.R. Lupini, J. Song, P.C. Snijders, R.K. Vasudevan, S. Jesse, S.V. Kalinin, Deep Analytics of Atomically-Resolved Images: Manifest and Latent Features, *ArXiv Prepr.*, 2018. [ArXiv180105133](https://arxiv.org/abs/180105133).
- [94] M. Ziatdinov, O. Dyck, S. Jesse, S.V. Kalinin, Atomic Mechanisms for the Si Atom Dynamics in Graphene: Chemical Transformations at the Edge and in the Bulk, *ArXiv Prepr.*, 2019. [ArXiv190109322](https://arxiv.org/abs/190109322).
- [95] M. Ziatdinov, O. Dyck, X. Li, B.G. Sumpter, S. Jesse, R.K. Vasudevan, S.V. Kalinin, Building and exploring libraries of atomic defects in graphene: scanning transmission electron and scanning tunneling microscopy study, *Sci. Adv.* 5 (2019), eaaw8989, <https://doi.org/10.1126/sciadv.aaw8989>.
- [96] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: *Int. Conf. Med. Image Comput. Comput.-Assist. Interv.*, Springer, 2015, pp. 234–241, https://doi.org/10.1007/978-3-319-24574-4_28.
- [97] G. Huang, Z. Liu, L. Van Der Maaten, K.Q. Weinberger, Densely connected convolutional networks, in: *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 4700–4708, <https://doi.org/10.1109/CVPR.2017.243>.
- [98] Q. He, R. Ishikawa, A.R. Lupini, L. Qiao, E.J. Moon, O. Ovchinnikov, S.J. May, M.D. Biegalski, A.Y. Borisevich, Towards 3D mapping of BO₆ octahedron rotations at perovskite heterointerfaces, unit cell by unit cell, *ACS Nano* 9 (2015) 8412–8419, <https://doi.org/10.1021/acsnano.5b03232>.
- [99] K.L. Westra, A.W. Mitchell, D.J. Thomson, Tip artifacts in atomic force microscope imaging of thin film surfaces, *J. Appl. Phys.* 74 (1993) 3608–3610, <https://doi.org/10.1063/1.354498>.
- [100] M. Ziatdinov, A. Maksov, S.V. Kalinin, Learning surface molecular structures via machine vision, *Npj Comput. Mater.* 3 (2017) 31, <https://doi.org/10.1038/s41524-017-0038-7>.
- [101] L. Vlcek, M. Ziatdinov, A. Maksov, A. Tselev, A.P. Baddorf, S.V. Kalinin, R.K. Vasudevan, Learning from imperfections: predicting structure and thermodynamics from atomic imaging of fluctuations, *ACS Nano* 13 (2019) 718–727, <https://doi.org/10.1021/acsnano.8b07980>.
- [102] D.P. Kingma, M. Welling, Auto-encoding Variational Bayes, *ArXiv Prepr.*, 2013. [ArXiv1312.6114](https://arxiv.org/abs/1312.6114).
- [103] F. Scarselli, M. Gori, A.C. Tsoi, M. Hagenbuchner, G. Monfardini, The graph neural network model, *IEEE Trans. Neural Network* 20 (2008) 61–80, <https://doi.org/10.1109/TNN.2008.2005605>.
- [104] W. Yuan, B. Zhu, X.-Y. Li, T.W. Hansen, Y. Ou, K. Fang, H. Yang, Z. Zhang, J.B. Wagner, Y. Gao, others, Visualizing H₂O molecules reacting at TiO₂ active sites with transmission electron microscopy, *Science* 367 (2020) 428–430, <https://doi.org/10.1126/science.aae2474>.
- [105] N. Li, D. Su, In-situ structural characterizations of electrochemical intercalation of graphite compounds, *Carbon Energy* 1 (2019) 200–218, <https://doi.org/10.1002/cey.2.21>.
- [106] Y. Jiang, Z. Zhang, W. Yuan, X. Zhang, Y. Wang, Z. Zhang, Recent advances in gas-involved in situ studies via transmission electron microscopy, *Nano Res.* 11 (2018) 42–67, <https://doi.org/10.1007/s12274-017-1645-9>.
- [107] S. Hwang, X. Chen, G. Zhou, D. Su, In situ transmission electron microscopy on energy-related catalysis, *Adv. Energy Mater.* 10 (2020) 1902105, <https://doi.org/10.1002/aenm.201902105>.
- [108] Y. Jiang, H. Li, Z. Wu, W. Ye, H. Zhang, Y. Wang, C. Sun, Z. Zhang, In situ observation of hydrogen-induced surface faceting for palladium–copper nanocrystals at atmospheric pressure, *Angew. Chem. Int. Ed.* 55 (2016) 12427–12430, <https://doi.org/10.1002/anie.201605956>.
- [109] A. Ziletti, D. Kumar, M. Scheffler, L.M. Ghiringhelli, Insightful classification of crystal structures using deep learning, *Nat. Commun.* 9 (2018) 2775, <https://doi.org/10.1038/s41467-018-05169-6>.
- [110] W. Xu, J.M. LeBeau, A deep convolutional neural network to analyze position averaged convergent beam electron diffraction patterns, *Ultramicroscopy* 188 (2018) 59–69, <https://doi.org/10.1016/j.ultramic.2018.03.004>.
- [111] J. Aguiar, M. Gong, R. Unocic, T. Tasdizen, B. Miller, Decoding crystallography from high-resolution electron imaging and diffraction datasets with deep learning, *Sci. Adv.* 5 (2019), eaaw1949, <https://doi.org/10.1126/sciadv.aaw1949>.
- [112] Z. Jiang, J. Li, Y. Yang, L. Mu, C. Wei, X. Yu, P. Pianetta, K. Zhao, P. Cloetens, F. Lin, Y. Liu, Machine-learning-revealed statistics of the particle-carbon/binder detachment in lithium-ion battery cathodes, *Nat. Commun.* 11 (2020) 2310, <https://doi.org/10.1038/s41467-020-16233-5>.
- [113] D.M. Dimiduk, E.A. Holm, S.R. Niezgoda, Perspectives on the impact of machine learning, deep learning, and artificial intelligence on materials, processes, and structures engineering, *Integr. Mater. Manuf. Innov.* 7 (2018) 157–172, <https://doi.org/10.1007/s40192-018-0117-8>.
- [114] K. Rajan, Materials informatics, *Mater. Today* 8 (2005) 38–45, [https://doi.org/10.1016/S1369-7021\(05\)71123-8](https://doi.org/10.1016/S1369-7021(05)71123-8).
- [115] P.-N. Tan, M. Steinbach, V. Kumar, *Introduction to Data Mining*, first ed., Addison-Wesley Longman Publishing Co., Inc., USA, 2005.
- [116] A. Jain, S.P. Ong, G. Hautier, W. Chen, W.D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, K.A. Persson, The Materials Project: a materials genome approach to accelerating materials innovation, *APL Mater.* 1 (2013), 011002, <https://doi.org/10.1063/1.4812323>.
- [117] Materials DFT Database. www.atomly.net (accessed: May 2020).