✴ MICROSCOPY

Applications of deep learning in electron microscopy

Kevin P. Treder[1], Chen Huang[2,*], Judy S. Kim[1,2] and Angus I. Kirkland[1,2]

[1]Department of Materials, University of Oxford, Oxford, Oxfordshire OX1 3PH, UK
[2]Rosalind Franklin Institute, Harwell Research Campus, Didcot, Oxfordshire OX11 0FA, UK
*To whom correspondence should be addressed. E-mail: chen.huang@rfi.ac.uk

Abstract

We review the growing use of machine learning in electron microscopy (EM) driven in part by the availability of fast detectors operating at kiloHertz frame rates leading to large data sets that cannot be processed using manually implemented algorithms. We summarize the various network architectures and error metrics that have been applied to a range of EM-related problems including denoising and inpainting. We then provide a review of the application of these in both physical and life sciences, highlighting how conventional networks and training data have been specifically modified for EM.

Key words: artificial intelligence, machine learning, deep learning, electron microscopy, neural networks, cryo-EM

Introduction

Machine learning is a relatively new field, which has recently seen extensive application in various aspects of electron microscopy (EM).

Direct electron detectors have had a significant effect on EM, which has led to an increase in the amount of data generated [1]. In addition, complex specimens and new imaging modes are increasingly used, which often results in data sets that are difficult to analyse using manual approaches [2,3]. Thus, new methods that are less reliant on manual intervention and expertise are required.

Deep learning was originally developed within machine learning and computer vision and has led to significant progress in a variety of computer vision problems [4]. Recently, there have been a number of reports that use deep learning to improve automated image analysis in EM. In this review, we survey research that uses artificial intelligence (AI) and machine learning (ML) in EM. We focus on deep learning techniques across a variety of neural network architectures and their applications in exemplars taken from both physical and life sciences. Traditional computational techniques, such as principal component analysis (PCA), sparse coding, k-means clustering, random forest and others, are not covered in this review except in a comparative or supplementary context. We further restrict ourselves to the field of EM and exclude research that uses other radiations.

Many EM-related applications, including denoising and image restoration, are essentially direct extensions of their counterparts in general image processing. However, EM-specific applications typically require specific curation or generation of network training data.

The broad scope of this review is hence 2-fold; firstly, we summarize the basic principles of deep learning as applied to computer vision together with common error metrics described in the literature. Subsequently, we provide an overview of EM applications in both *physical* and *life sciences*. Depending on instrument outputs and the analysis task under consideration, specific network architectures are encountered more frequently than others. The most prominent ones are highlighted, and key EM-related applications are discussed.

It should be noted that other reviews and articles detailing other microscopy techniques, available resources for AI, further ML techniques and big data are available to provide a complementary review of the field as a whole [1,5–9].

Deep learning in computer vision
Principles

Deep learning is a form of ML that learns a high-level representation of data through hierarchical architectures [10], often referred to as neural networks. The network passes raw data input through non-linear functional modules and transforms them into increasingly abstract representations. This allows the network to automatically learn complex functions and determine the representation needed for a specific task such as classification and detection [11]. The essence of their success lies in how the representations are generated, and in many cases, these are entirely learned from data using a general-purpose learning procedure [11].

Krizhevsky *et al.* [12] achieved ground-breaking results in the ImageNet Large Scale Visual Recognition Challenge, using convolutional neural networks (CNNs), and these have been adopted by the computer vision community and more broadly applied in diverse applications including consumer electronics, pedestrian detection, autonomous driving and computer-aided medical diagnosis [13,14].

Different outputs are obtained depending on the assigned computer vision task, thus requiring a clear distinction: given a set of mutually exclusive classes $c = [c_1, c_2, c_3, \ldots, c_n]$, *classification* is defined as a function f_{class} that assigns one or

K.P. Treder *et al. Applications of deep learning in electron microscopy*

i101

multiple instances of the real-valued vector c to a real-valued input image I_{in}:

$$f_{class} : I_{in}^{j \times k \times l} \rightarrow c^{n \times 1} \tag{1}$$

where $j, k, l \epsilon \mathbb{N}$ denote the image dimensions, e.g. the image size in pixels and the number of channels.

This forms the foundation for more complex operations such as *object detection* (f_{od}) and *semantic segmentation* (f_{seg}) [15]. For the former, the aim is to detect the presence of a class instance c_n and to determine its coordinates and measures, for which a bounding box can be applied [13]:

$$f_{od} : I_{in}^{j \times k \times l} \rightarrow D^{m \times 5} \tag{2}$$

In Eq. (2), m refers to the number of instances detected while each entry i of the array $D^{m \times 5}$ is represented by a vector $d_i = [c_n, x, y, w, h]$, with x, y the image coordinates and w, h the width and height of the bounding box, respectively. *Semantic segmentation* is subsequently characterized by a classification of individual pixels, which partitions an image into meaningful regions [14]:

$$f_{seg} : I_{in}^{j \times k \times l} \rightarrow I_{out}^{j \times k \times l} \tag{3}$$

for which an input image I_{in} is mapped into an output image I_{out}.

Computer vision in EM

Although the principles of deep learning as described above were originally developed using photographs of everyday life scenes, they equally hold true when applied to electron micrographs. Figure 1 illustrates a number of exemplary micrographs and data that are typically encountered in EM. The 'aimed output' can be achieved through different paths, and the estimated processing time is the fastest for CNN-based solutions.

Other important computer vision applications in EM include image *inpainting* and *denoising* (Fig. 1). Inpainting describes the interpolation of neighbouring pixels in defective parts of an image without noticeable changes to the rest of the image [16] for which CNNs and generative adversarial networks (GANs) [17,18] are the state-of-the art methods. For comparison to Eq. (3), this can be expressed as an image restoring function:

$$f_{restore} : I_{in}^{j \times k \times l} \rightarrow I_{out}^{j \times k \times l} \tag{4}$$

In contrast to a conventional CNN, a GAN architecture sets a generative model G against a discriminative model D where both are typically neural networks on their own [19]. While D learns to determine whether a sample stems from the model or data distribution, G tries to generate false outputs such that D makes mistakes [17]. Hence, the training procedure follows a mini–max optimization problem where the saddle point is reached at minimum G and a maximum for D. At this point, G can be thought to have captured the real distribution of true examples [19]. Denoising aims to remove additive noise from an image that inevitably arises during the acquisition, compression and transmission of images [20].

We summarize the most commonly used network architectures and their applications in EM in Table 1. There are intrinsic links between some networks; many segmentation networks are the encoder–decoder versions of their respective classification counterparts, for example, SegNet/VGG, ResUNet/ResNet and Fully Convolutional DenseNet/DenseNet. Additional references providing in-depth explanations of the fundamentals of neural networks and deep learning are given in refs. [11,21–23].

Error measurement

Training error and test error are well-defined concepts in ML [131], and it is standard practice to include network training information, such as training curves that illustrate the change in loss or accuracy as the training progresses. These provide valuable information for determining if a neural network is trained correctly and sufficiently without overfitting [21].

Error measurement for classification and object detection

In the cases of classification and object detection, the performance can be evaluated by a confusion matrix [132], where the prediction and the ground truth are compared (Table 2).

For EM applications, the confusion matrix is widely used to evaluate particle picking performance in cryo-EM. A discussion and clarification of the divergence in terminology has been given by Langlois and Frank [133], and a set of generally agreed metrics (sensitivity = TP/P, specificity = TN/N, accuracy = (TP + TN)/(P + N), error = (FP + FN)/(P + N)) and the receiver operating characteristic curve [134] have been established.

Error measurement for image segmentation

For EM applications, the output from segmentation networks often needs to be abstracted into quantities such as length, width, area [26,86], atomic coordinates [79,84,85] or labelled regions [37].

For semantic segmentation where the output masks directly correspond to a measurement, for example, in the case of dimension measurement, the segmentation error can be simply calculated as an extended case of classification averaged across all pixels. If the physical properties of the pixels are calibrated, the error estimate will have the same unit as the measured quantity.

Where atom columns are resolved, segmentation network outputs are often abstracted into atomic coordinates, and simple adjustments to error metrics, such as comparing the centroids of each segmented atom instead of the pixel arrays, can assist the interpretation of performance. For example, Lee *et al.* [90] have introduced a tracing error defined as the error between two sets of atomic coordinates, the ground truth and the prediction, to assess the performance of their network for surface atom detection.

A more thorough consideration of error metrics in semantic segmentation of EM images comes from competitions. Arganda-Carreras *et al.* [135] have documented the rationale and progression in the choice of error metrics for the SBEM segmentation competition (International Symposium on Biomedical Imaging 2012 (ISBI'12)) and have proposed two separate sets of metrics—Rand scoring and information theoretic scoring (Table 3).

Rand scoring defines p_{ij} as the probability that a randomly chosen pixel belongs to a segment i in prediction S and segment j in ground truth T. Consequently $\sum_{ij} p_{ij} = 1$ and the

probability of randomly selecting two pixels that belong to
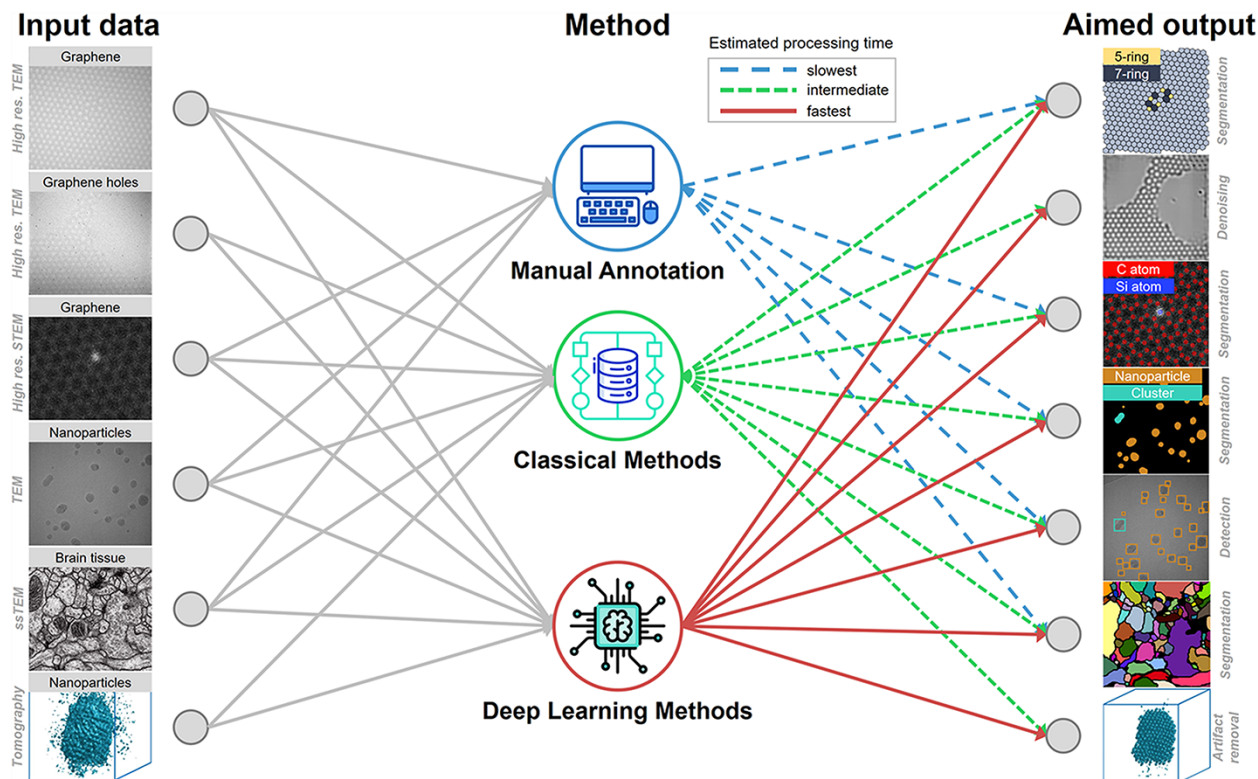
Fig. 1. Examples of different processing routes for an aimed output. Classical methods refer to (semi-) automated algorithms and ML methods which differ depending on the given analysis task. The output of a given network is mainly dependent on the training images used. Input data images of 'High res. STEM Graphene', 'ssTEM Brain Tissue' and 'Tomography Nanoparticles' as well as respective outputs reproduced from [79], [135] and [90] respectively. Icons used in the 'Methods' column were obtained from flaticon.com. Icon authors: 'Icongeek26', 'Freepik', 'xnimrodx'. High res. STEM input and 'Aimed output' images reprinted with permission from [79], Copyright 2021, American Chemical Society.

the same i and j is $\sum_{ij} p_{ij}^2$. Marginal distributions are given by $s_i = \sum_j p_{ij}$ and $t_j = \sum_i p_{ij}$.

An alternative metric, information theoretic scoring, is based on the mutual information: $I(S;T) = \sum_{ij} p_{ij} \log p_{ij} - \sum_i s_i \log s_i - \sum_j t_j \log t_j$ and entropy: $H(S) - \sum_i s_i \log s_i$.

F-scores for both metrics were defined as the weighted harmonic mean of the split score and merge score and were used as the evaluation standard for the ISBI'12 competition [135].

Evaluation of image denoising

The performance of denoising networks can be evaluated in a similar way to established denoising methods. Perfect (infinite-dose) simulated images can be used as the signal while different types of noise are added to generate noisy images. The denoised version of such data can then be compared to the original input signal to calculate the signal-to-noise ratio (SNR). Subsequently, metrics such as peak signal-to-noise ratio (PSNR) and a structural similarity index measure (SSIM) are often used [136]. As an example, if the reference image f and the degraded image g are grayscale images of size $M \times N$ and bit depth n, then the PSNR is defined as

$$PSNR(f,g) = 10 \log\left((2^n - 1)^2 / MSE(f,g)\right) \quad (5)$$

where the mean squared error (MSE) has the form

$$MSE(f,g) = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} (f_{ij} - g_{ij})^2 \quad (6)$$

The SSIM is then defined as

$$SSIM(f,g) = l(f,g)\, c(f,g)\, s(f,g) \quad (7)$$

where

$$\begin{cases} l(f,g) = \frac{2\mu_f \mu_g + C_1}{\mu_f^2 + \mu_g^2 + C_1} \\ c(f,g) = \frac{2\sigma_f \sigma_g + C_2}{\sigma_f^2 + \sigma_g^2 + C_2} \\ s(f,g) = \frac{\sigma_{fg} + C_3}{\sigma_f \sigma_g + C_3} \end{cases} \quad (8)$$

As can be seen from Eq. (7), the SSIM is the product of three comparison measurements that correspond to the closeness of luminance (*l*), contrast (*c*) and structure (*s*) between two images. The individual comparison functions (Eq. (8)) are calculated using the mean luminance (μ_f and μ_g) and standard deviation (σ_f and σ_g) of the respective images and their covariance (σ_{fg}). C_1, C_2 and C_3 are positive constants that serve to stabilize the denominators. If f and g are identical, then the SSIM equals 1, while SSIM = 0 would suggest the two images are completely uncorrelated.

As discussed by Horé and Ziou [136], these two metrics are analytically linked and in many cases give a similar assessment of the image quality, albeit with a slight difference in sensitivity towards different types of degradation. As a consequence, Wang *et al.* [106] have proposed a new metric called 'Con-

Table 1. Neural network architectures and respective applications in EM. The superscript numbers I–IV refer to the task I, classification; II, object detection; III, semantic segmentation and IV, image inpainting and denoising (image restoration) introduced in Eqs. (1)—(4) where applicable

Underlying network	Applications
Multi-layer perceptron (MLP) [24]	*Materials analysis* Carbon nanotubes (CNTs): quality assessment[I] [25] and image segmentation[III] [26] Nanoparticles: size analysis in agglomerates[I] [27] Semiconductors: composition and thickness analysis[I] [28] *Phase retrieval* CBED analysis [29–31] Image wave reconstruction in holography[I] [32,33] *Techniques* EELS low-loss region analysis[I] [34] *Connectomics* Brain tissue segmentation[III] [35–37] *Cryo-EM* Particle picking for single-particle analysis (SPA)[II] [38,39]
Autoencoder (AE) [40]	*Material analysis* Transformations in graphene [41] *Cryo-EM* Particle picking for SPA[II] [42] Denoising of cryo-EM images[IV] [43]
LeNet [44]	*Material analysis* CNT chiral index determination[I] [45] Semiconductors: phase boundary analysis [46] *Cryo-EM* Particle picking for SPA[II] [47]
AlexNet [12]	*Material analysis* Metal alloys: Microstructure recognition[I] [48] *Diffraction* Position averaged convergent beam electron diffraction (PACBED) analysis[I] [49] *Cryo-EM* Particle picking for SPA[II, III] [50–52] *General applications* Micrograph labelling[I] [53]
VGG [54]	*Material analysis* Metal alloys: microstructure classification and segmentation[I, III] [55,56] Semiconductors: defect recognition in III–V/Si crystalline materials[I, III] [57] Classification of carbon nanomaterials[I] [58] *Cryo-EM* Pruning of false positives in SPA[I] [59]
DenseNet [60]	*Material analysis* Metal alloys: defect analysis[III] [61] *Techniques* Super-resolution reconstruction of TEM images [62] *Virus analysis* Virus image classification[I] [63]
ResNet [64]	*Material analysis* Nanoparticles: detection in SEM images[II] [65] *Cryo-EM* Part of cryo-EM analysis pipeline[II, IV] [66]
InceptionNet [67]	*Cryo-EM* Particle picking for SPA[III] [68] *Cryo-ET*

(continued)

Table 1. (Continued)

Underlying network	Applications
	Subdivision of tomograms into structurally homogeneous subsets[I] [69] *General applications* Categorization of EM images [70]
YOLO [71]	*Cryo-EM* Particle picking for SPA[II] [72,73] Ice-thickness determination[I, II] [74] Part of cryo-EM analysis pipeline[I, II] [75]
FCN [76]	*Cryo-EM* Particle picking for SPA [77]
SegNet [78]	*Material analysis* 2D materials: segmentation of graphene and MoWSe$_2$ STEM images[III] [79], phase evolution and dynamics in Mo-doped WS$_2$[III] [80], generation of graphene defect libraries[III] [81] *Phase retrieval* Prediction of structure and phase from diffraction patterns [82]
U-Net [83]	*Material analysis* 2D materials: analysis of beam-induced reactions of Si atoms on graphene[III, IV] [84], atomic structure analysis in graphene and nanoparticles[III] [85] Nanoparticles: segmentation for *in situ* experiments[III] [86,87], segmentation and defect identification[III] [88], column height determination in metallic nanoparticles[III] [89] *Tomography* Atomic-resolution tomography [90] Reconstruction of noisy 3D EDX tomograms [91] *Cryo-EM* Data pruning approaches for SPA [92] Image denoising[IV] [66,93–95] Resolution estimation and validation of cryo-EM maps[I] [96] Atom structure determination from cryo-EM maps [97] *Virus analysis* Optimising network architectures for virus analysis[III] [98] *General applications* ADF-STEM image training library and image analysis app [99] Image-to-image inverse problems [100]
ResUNet [101]	*Material analysis* 2D materials: Analysis of WeSe$_{2-2x}$Te$_{2x}$ point defects[III] [102] *Connectomics* Segmentation of brain tissue images[III] [103]
GAN [17]	*Tomography* Information recovery and artefact removal[IV] [104,105] *Techniques* Denoising of STEM images[IV] [106] STEM micrograph inpainting[IV] [107] SEM image-resolution enhancement [108]
Other	*Material analysis* Nanoparticles: analysis of nanoparticle diffusion behaviour in liquid-cell TEM[I] [109], detection and segmentation of nanoparticles[II, III] [110] Metal alloys: defect analysis[I, II] [111]

(continued)

Table 1. (Continued)

Underlying network	Applications
	Diffraction
	Diffraction pattern classification[I] [112]
	Phase retrieval
	Solving the ptychographic inversion problem [113]
	Techniques
	Analysis of manganese EELS spectra[I] [114]
	Denoising of EM micrographs[IV] [115]
	Cryo-EM
	Particle picking for SPA[II] [116]
	Structure modelling[I, II, III] [117–119]
	Cryo-ET
	Subcellular annotation and feature extraction[III] [120]
	Open-source platform for data analysis [121]
	Connectomics
	Brain tissue segmentation[III] [122–124]
	Virus analysis
	Segmentation of feline calicivirus images[III] [125], classification and detection of herpesvirus[I, II] [126,127], detection of SARS, MERS and HIV viruses[II] [128], image classification [129]
	General applications
	Entry-level platform for deep learning tasks[II, III, IV] [130]

Table 2. The confusion matrix

TP (true positive)	FP (false positive)	P′ = TP + FP (predicted positive)
FN (false negative)	TN (true negative)	N′ = FN + TN (predicted negative)
P = TP + FN (actual positive)	N = FP + TN (actual negative)	

Table 3. Segmentation error metrics: Rand scoring and information theoretic scoring

	Rand scoring	Information theoretic scoring
Split score	$V^{Rand}_{split} = \frac{\sum_{ij} p_{ij}^2}{\sum_k t_k^2}$	$V^{info}_{split} = \frac{I(S;T)}{H(S)}$
Merge score	$V^{Rand}_{merge} = \frac{\sum_{ij} p_{ij}^2}{\sum_k s_k^2}$	$V^{info}_{merge} = \frac{I(S;T)}{H(T)}$
F-score ($\alpha = 0.5$):	$V^{Rand}_{\alpha} = \frac{\sum_{ij} p_{ij}^2}{\alpha \sum_k s_k^2 + (1-\alpha) \sum_k t_k^2}$	$V^{info}_{\alpha} = \frac{I(S;T)}{(1-\alpha)H(S)+\alpha H(T)}$

secutive Similarity (CSS)' based on SSIM to circumvent the situation where no ground truth data are available to calculate the SSIM.

Applications in physical sciences

Atomic-resolution scanning transmission electron microscopy

Early examples of the use of deep neural networks for semantic segmentation of EM data are based on atomic-resolution scanning transmission electron microscopy (STEM) with sophisticated post-processing workflows [79–81,84] using

fully convolutional networks (FCNs) [76,78,83]. For example, a training data set of exclusively simulated images has been used in a supervised training approach to construct a neural network for automated defect segmentation and identification in graphene and $Mo_{1-x}W_xSe_2$ monolayers [79]. The central idea in this study was to form a training base with limited *a priori* information, which could then be applied to infer a wider variety of defect types in experimental images.

Further studies have applied a two-step deep learning approach to identify Si impurities and vacancy defects in graphene [81] and the segmentation of Mo-doped WS_2 during electron irradiation to study the dynamics of phase evolution [80]. In both cases, identified defect structures were categorized either manually or by using an unsupervised Gaussian mixture model. Subsequent steps involved modelling of defect structures using density function theory and a comparison with experimental STM images of the same sample. Kalinin *et al.* [41] also suggest a bottom-up description of systems undergoing chemical transformation using minimal physical descriptors and general rotational invariance implemented in a variational autoencoder [137].

Similarly, Lee *et al.* [102] have analysed defect structures in $WSe_{2-2x}Te_{2x}$ monolayers using a ResUNet to classify specific defects recorded using low-dose Annular Dark Field (ADF) images that subsequently formed the basis for class-averaged images of isolated defect types. This allowed accurate measurements of atomic spacings, which revealed strain oscillations around vacancies.

A substantially different approach to analyse atomic-resolution STEM images has been reported by Dennler and co-workers [57]. Instead of using a direct segmentation approach, the authors employed a VGG-16 classification network in conjunction with graph-based heuristics and automatically generated convolutional filter banks to analyse crystal structure defects in III–V/Si samples. This pipeline yielded a multi-class segmentation map, through which an image was separated into non-defective regions, regions with symmetry faults and blurred regions. Importantly, this approach outperformed classical Fourier space Bragg filtering.

Atomic-resolution transmission electron microscopy

An early use of neural networks applied to atomic-resolution transmission electron microscopy (TEM) is given in Kirschner and Hillebrand [28] who trained a neural network to determine the composition and thickness of a compound semiconductor in 2000.

Madsen *et al.* [85] reported the first application of a deep neural network to high-resolution TEM images of graphene monolayers, and the use of the same network for atom counting in nanoparticles has been reported by Ragone and colleagues [89]. Although these authors employed a network with a U-Net architecture, improvements were achieved by explicit use of realistic Wulff constructed nanoparticles in the training data set as well as regression-based output for column height determination.

More recently, Groschner *et al.* [88] demonstrated the use of a modified U-Net for segmenting Au and CdSe nanoparticles in TEM images with a subsequent application of a random forest classifier [138] to categorize nanoparticles that contain a stacking fault.

Other (S)TEM applications

In addition to atomic-resolution (S)TEM, several research groups have used deep learning methods to help characterize materials at lower resolution. Examples include the characterization of nanoparticles [86,87,109,110] and agglomerates [27], dislocations in metals and alloys [61,111], phase-boundary regions [46] and chiral indices in carbon nanotubes (CNTs) [45]. Studies by Jamali *et al.* [109], Frei *et al.* [27] and Förster *et al.* [45] are notable in that they use pure image classification instead of a segmentation approach in which a modified version of a generic temporal convolution network [139], a MLP and an architecture based on the LeNet-5 were exploited. This provided new insights into nanoparticle diffusion in liquid-cell TEM [109] as well as advances towards a time-efficient determination of the chiral index of CNTs with improved statistical significance [45].

Scanning electron microscopy

Neural network–based characterization has also been performed using scanning electron microscopy (SEM) data. This is illustrated in the study of Al-Khedher *et al.* [25] that reports the use of a MLP to characterize the curvature and alignment values of CNTs from SEM images. Trujillo *et al.* [26] also chose a similar MLP network for segmenting CNTs from SEM/TEM images, and Modarres *et al.* [70] utilized more advanced neural network models and performed transfer learning training on a diverse SEM data set to achieve accurate classification of nanostructures. A similar classification study using TEM images of carbon nanomaterials was reported by Luo *et al.* [58] by applying transfer learning with a VGG-16 network.

Azimi *et al.* [56] showed that segmentation analysis of steel microstructures that was traditionally carried out by human experts can be automated without sacrifices in accuracy by using a VGG-16 network and an FCN. A combination of traditional ML methods, such as Visual Bag of Words, and deep learning methods has been shown by Chowdhury *et al.* [48] to be a useful approach when multiple features appear in the same field of view. DeCost *et al.* [55] used the features in intermediate layers within the VGG network architecture combined with unsupervised learning techniques, to classify the local structure of ultrahigh carbon steel, which also demonstrated superior performance compared to non-deep learning methods. Finally, it has been shown [57,65] that synthetic data sets can be used for nanoparticle object detection in SEM images using RetinaNet [140] with a pre-trained ResNet-50 backbone.

Diffraction

Several studies have contributed to the implementation of deep learning methods in reciprocal space. Aguiar *et al.* [112] provide an example of how public data set and classification networks can be combined in a diffraction pattern characterization tool. Pennington *et al.* [31] together with Xu and LeBeau [49] have similarly described the analysis of convergent beam electron diffraction (CBED) patterns in which automated measurements of disk size, pattern centre, rotation, thickness and tilt were demonstrated using a SiTiO₃ sample with high accuracy and processing rates of 0.1 s/pattern [49]. Oxley *et al.* [141] have also shown that by training a deep CNN with simulations of a variety of LaAlO₃

and $SrTiO_3$ interfaces, the neural network can distinguish physically rough interfaces from chemically diffuse interfaces and even locate buried steps within these interfaces.

Spectroscopy

An important capability of electron microscopes is the measurement of chemical properties using energy-dispersive X-ray (EDX) analysis [91] or electron energy-loss spectroscopy (EELS). However, these methods have specific limitations. The calibration of the spectrometers used for EELS can be subject to drift of the zero-loss peak (ZLP) energy, which complicates the comparison of absolute peak positions [114]. To resolve this, Chatzidakis and Botton [114] implemented a deep learning method for calibration-invariant EELS spectrum analysis. Similar to other methods that harness publicly available data, the authors accessed EELS spectra of Mn^{2+}, Mn^{3+} and Mn^{4+} acquired under varying experimental conditions from in-house stored data that served as a training data set. The network was able to identify spectral peaks merely by their shape while retaining translation invariance. This method's success was further underlined by generating a correct output for spectra that were digitized from other publications.

Roest *et al.* [34] have studied the analysis of the ultra-low-loss region of EELS spectra, which exhibits overlapped contributions from the ZLP as well as those from low-loss inelastic interactions. Specifically, the authors aimed to parametrize the ZLP for subsequent subtraction from the spectrum. A Monte Carlo replica method was used to construct a probability distribution in the experimental data space, which was propagated by a neural network to a probability density in the space of ZLP models. Subsequently, the Neural Network Parton Distribution Function approach [142] was used, which was originally developed for the studies of the quark and gluon substructure in the proton [143]. It was shown that this procedure can disentangle the ZLP contribution from low-loss inelastic scattering and was applied to WS_2 nanoflowers characterized by a 2 H/3 R polytypic crystalline structure. Consistent with previous studies, a bandgap distribution $E_{BG} \cong 1.6^{+0.3}_{-0.2}$ eV with an indirect bandgap preference was measured.

Tomography

An important contribution to the sub-field of electron tomography is described in a study published in 2019 [104], which proposes a two-step joint model to inpaint missing-wedge information in sinograms and to reduce residual artefacts in reconstructed tomograms based on GANs [17]. These authors demonstrated superior performance with respect to established methods such as weighted back projection (WBP), simultaneous algebraic reconstruction technique (SART) or total variation minimization (TVM) on a wide range of data sets. In a follow-up study published a year later [105], missing wedge artefacts were successfully removed from an atomic-resolution tomogram, which improved the resolution from ~1.42 Å to 0.71 Å. In contrast to earlier work, this study used a trained single GAN for artefact removal only, whereas a U-Net++ architecture [144] was used for the generator. Remarkably, a very generic training data set containing images from open data sources such as the Massachusetts General Hospital (MGH), the National Biomedical Imaging Archive (NBIA), the ImageNet library, as well as

simulated images resembling cross-sectional images of faceted and rounded nanocrystals were used in both cases for training, potentially suggesting applicability in related fields such as computer tomography (CT), optical diffraction imaging (ODT) and magnetic resonance imaging (MRI) [105].

Lee *et al.* [90] reported an atomic structure measurement of 15 pm precision in 2021 in which three-dimensional (3D) U-Net [145] was trained on image simulations for tomogram artefact removal and augmentation. This approach resulted in a reduction of averaged tracing errors to below 1% measured on 1000 simulated tomograms, and the resolution in an experimental tomogram of a 4-nm Pt nanoparticle was significantly improved. The largest improvements were achieved at the surface of the nanoparticle, which the authors attribute to otherwise unidentifiable surface atoms related to the missing wedge. The accurate determination of 3D atomic coordinates allowed a detailed strain analysis, which confirmed theoretical predictions of a compressive stress along <100> and tensile strain along <111>.

Finally, Han *et al.* [91] have employed a multi-neural network procedure for STEM-EDX tomography where InP/ZnSe/ZnS quantum dots were successfully reconstructed in 3D from noisy data.

Phase retrieval

Some of the earliest applications of neural networks in EM are found for phase retrieval methods. The earliest study [32] dates back to 1996, and subsequently, Meyer and Heindl [33] successfully trained a MLP using simulated images for image wave calculations. Importantly, an application to experimental images was demonstrated, giving an early indication of the general capabilities of neural networks in EM.

By reformulating the existing stacked Bloch wave algorithm, Pennington *et al.* [31] and van den Broek *et al.* [29,30] have shown that it is possible to make use of the optimization tools originally developed for training artificial neural networks and have applied them to accelerate the iterative reconstruction of a 3D object potential from CBED data for which the solution is otherwise computationally expensive.

Cherukara *et al.* [82] have used two SegNet-like networks to retrieve phase information from coherent diffraction imaging (CDI) experiments, although the networks were only trained with simulated images and applied to simplified phase objects with convex shapes and uniform amplitude. A follow-up study from the same group showed that a combined version of these networks, PtychoNN, can reconstruct experimental X-ray CDI data, giving similar results as the ePIE algorithm [146] but significantly faster [113].

Wang *et al.* have extended the application of a modified U-Net to phase retrieval and denoising from STEM images [100,106] by varying the training data set. The denoising aspect of this contribution will be discussed subsequently.

Denoising and inpainting

By analogy to semantic segmentation, image pairs of noisy and clean image pairs can be used to train a neural network for denoising [20]. However, Lehtinen *et al.* [147] have presented a solution that shows that clean images are not required. In their approach, the authors highlight a benefit of the L_2 minimization, which is not immediately obvious in that its estimate is unchanged if the targets in a training set are replaced with random numbers with a matching expectation. Consequently, training targets can be in principle corrupted with zero-mean noise without changing what the network learns, and hence, both the training input *and* targets can be drawn from corrupted images. This suggests that a network can, for instance, learn to remove photon noise given only pairs of noisy images. Based on this background, unsupervised approaches [148–152] are of particular interest where image pairs cannot be acquired or simulated.

Wang *et al.* [106] reconsidered aspects of the studies mentioned above and implemented a denoising network solution using a GAN [17]. Noise2Atom [106] applies the discriminator D to differentiate between images composed of pure Gaussian peaks and images that were translated from a noisy experimental input into a clean one. The performance of the method was benchmarked against established denoising algorithms including the multiresolution deep convolutional neural network (MCNN) [100]. Contrary to Noise2Atom, MCNN is a CNN trained on simulated images, and the authors note that it is difficult to include all potential noise levels encountered in STEM imaging.

Ede and Beanland have studied both image denoising [115] and inpainting [107] for which (S)TEM images from their published database [153] were corrupted with Possion noise or binary masks resembling partial scans, respectively. These authors employed a network structure similar to DeepLab3 [154,155] and outperformed eight other established denoising methods [115].

In the context of image inpainting, the low SNR limitation encountered when imaging beam-sensitive materials can also be inversely interpreted. Sparse sampling can reduce the electron dose significantly with missing information inferred through algorithms such as natural neighbour interpolation or exemplar-based inpainting [156] aiming for minimal information loss. This is equivalent to an approach already demonstrated using compressive sensing [157,158] but requires a fast deflection system for application to STEM imaging. Published data assumed patterns that were most likely to be realistic for standard deflection systems and used STEM images with diverse characteristics to obtain a training data set. Results from the multiscale GAN allowed for dose reductions with a factor of 17.9 with a 3.8% test error and 87.0 with a test error of 6.2%, respectively. One difficulty in using this method is that the generative part cannot reliably complete micrographs with unpredictable structure in regions that are not covered by the electron beam [107].

More general EM applications

This sub-section briefly summarizes the use of neural networks for more general-purpose applications, which do not directly fall into the above categories but are equally important.

Metadata annotation of individual micrographs can be as time-consuming as segmentation. Weber *et al.* [53] have used a transfer learning approach applied on an AlexNet-based network architecture to transfer manually labelled images to the entire datastore of the National Center for Electron Microscopy (NCEM). This study achieved a classification accuracy of $\sim 80\%$, suggesting a potential strategy to increase efficiency and simplicity in data searching particularly for big-data projects.

As suitable training data often remain difficult to generate independently of a particular computer vision task, there is significant interest in using publicly available resources for successful application of deep learning in EM. One such example is described by Ede [153] who posted data sets used in various studies, as well as models and interactive visualizations for public use. Similarly, Xin and co-workers recently published a desktop app with a graphical user interface for the analysis of ADF-STEM images [99]. This package includes a training library that consists of a diverse set of simulated images as well as the U-Net-based FCN called AtomSegNet that can be trained for various tasks such as super-resolution, localization and denoising.

Van Chamier and co-workers [130] provided a solution for making deep learning approaches accessible to the wider microscopy community. ZeroCostDL4Mic is an entry-level deep learning deployment platform harnessing Google Colaboratory (Colab). It allows users to perform important computer vision tasks such as image segmentation, object detection, image restoration and denoising without requiring in-depth technical or coding knowledge. As the authors note, this environment can be used for prototyping workflows or execution of small to medium-sized projects using state-of-the-art network architectures. Although the study used biomedical images, application into the physical sciences is an obvious extension.

Development of automatic alignment using deep learning is also likely to become an important application area in EM. As an example, Rotunno *et al.* [159] recently reported aligning a TEM orbital angular momentum sorter using a deep learning based misalignment measurement.

Applications in life sciences
Cryo-EM

In the last decade, the 'resolution revolution' [160] driven mainly by the introduction of direct-electron detectors marked the beginning of a new era in the field of structural biology. Single-particle cryo-EM (SPA cryo-EM) is no longer merely complementary to X-ray crystallography but has emerged as the favoured method for many structural determinations [161,162]. Despite this impact, specimen preparation for cryo-EM and data analysis remain complicated and time-consuming. In SPA, units of the biological specimen under investigation are dispersed in random orientations in thin vitreous ice. A high-resolution 3D reconstruction is subsequently obtained by averaging classes [163] with defined projection angles. However, typically thousands of individual images are required for successful reconstruction, not only to cover all different projections but also to obtain a sufficiently high SNR. Due to its complexity, the selection of suitable molecules, 'particle picking' is often performed in a manual or semi-automated fashion [162,164] and is regarded as one of the bottlenecks towards an efficient and fully automated cryo-EM processing pipeline [38,39,50,165]. However, from the perspective of computer science, 'particle picking' can be considered understood as object detection.

Pioneering work by Ogura and Sato in the early 2000s [38,39,165] using ML before the availability of direct electron detection used a three-layer 'pyramidal-type' neural network that converted a two-dimensional image input into a binary, one-dimensional output at the centre of identified particles [38]. This work included a thorough hidden layer analysis for a better understanding of the learning process along with extended data augmentation and PCA Eigenimage initialization [39]. In this relatively early implementation of a neural network–based approach, the authors recognized that the preparation of a training data set can be difficult due to low contrast and high background noise levels in cryo-EM images [165].

The first modern, deep neural network–based solution to 'particle picking' was implemented by Wang *et al.* [50] in 2016. Five publicly available data sets and a sliding window approach were exploited to train the network and infer 'unseen' images with a performance close to that of a human expert. Subsequent studies by Xiao and Yang [116] and Zhu *et al.* [47] significantly improved the inference time and high precision/recall performance despite a reduction of the training data set size. Bepler *et al.* [42] and Wagner and colleagues [72,73], used only positive training examples to train respective CNNs. The training procedure used in the particle-picking pipeline Topaz [42] has been reformulated as a constrained optimization problem in which a novel generalized expectation criterion was applied to account for potential overfitting. Remarkable performance was achieved for an asymmetric particle using only a sparsely labelled training data set. This differs from crYOLO, which employs an object detection architecture that defines its computer vision task in terms of a regression problem [71], resulting in high computational efficiency. A generalized version of this network that was trained on 45 data sets of manually picked, simulated and particle-free images has been published, allowing applications to unseen experimental data. Generalized neural networks like crYOLO can reduce manual involvement time in the particle picking procedure to zero. This has advantages for cryo-EM where it is often difficult to determine an actual ground truth from low-contrast, low-SNR images and has been exploited by Yao *et al.* [77] in the training of a generalized network on simulated cryo-EM images. In conjunction with a FCN that allows for fast inference for arbitrary image sizes, high correlation coefficients were achieved when reconstruction was compared to six published cryo-EM data sets [166–171]. A well-conceived data processing method allows for automatic training data generation from experimental images, and this has been demonstrated in various publications [51,52,172,173], in which particular attention was paid to account for low-SNR images. George *et al.* [68] implemented a solution based on full-resolution residual network [174] with semantic segmentation and Inception-v4 [175] feature extraction rather than object detection. In this case, a graphical user interface (GUI) was also provided, which allows for contrast enhancement and generation of a training data set in a semi-automated manner.

In the light of these developments, deep learning approaches have also been applied to other critical steps commonly used in the 3D reconstruction process [164]. These include improvements in data collection [74], data pruning to reduce the number of false positives [59,92], 3D alignment of data [176] as well as studies on 3D resolution determination [96,177].

A recent study has been published [178] that addresses solving heterogeneous cryo-EM structures using cryoDRGN (deep reconstructing generative networks), which employs

a specialized image-encoder–volume-decoder architecture. In this approach, particle images are embedded into a low-dimensional manifold latent space, which describes the heterogeneity of the data set. Subsequently, representative parts of the particle can be reconstructed by feeding in specified regions of the latent space into the decoder. This has been used, for example, to discover heterogeneous states of the RAG complex and the Pf80S ribosome, which were averaged by earlier used reconstruction methods, underlining the potential of deep learning-based methods.

Ultimately, automation spanning from raw data input to the generation of a 3D reconstruction within short time frames will be required to further transform cryo-EM into a data-driven technique.

Examples of such frameworks have been presented by Tegunov and Cramer [66] and Li and colleagues [75], in which a modular design consisting of tailored image processing routines and dedicated neural networks are combined. However, the application philosophies differ while Warp [66] provides a user-friendly user interface (UI) that aims to conduct pre-processing steps during image acquisition, the high-throughput pipeline presented by Li *et al.* [75] does not require any user interaction. Both methods contain highly generalized deep learning modules for particle picking and have achieved similar [66,75] or higher-resolution reconstructions from published data.

Structure modelling

The final step required to achieve an atomic structure of the molecule which is essential to understand the underlying mechanisms, molecular dynamics and structure comparison that, for example, enable structure-based drug design is structural modelling. While this article aims to address advances achieved in EM using deep learning, the interested reader is directed to detailed reviews of Malhotra *et al.* [179] and Kim *et al.* [180], which review this step. In contrast to classification methods, deep-learning-based structure modelling deals with 3D data, from which secondary structure elements, the backbone structure and even the Cα atoms play a central role in amino acid characterization [97,117–119]. Several reported results have outperformed state-of-the art prediction methods in terms of both Cα atoms identified and resolution in successful modelling of the SARS-CoV-2 spike glycoprotein as well as the human receptor ACE2 and extraction of hidden dynamic information from static 3D cryo-EM density maps. Figure 2 shows examples of deep neural networks applied in the field of cryo-EM and structure modelling.

Cryo-electron tomography

An important task in cryo-electron tomography data processing is to correctly classify subtomograms. Promising performance when compared to SVM-based methods was demonstrated by Xu *et al.* [69], who trained deep neural networks with simulated subtomogram voxels of 22 macromolecular complexes. This has later been extended into an AI platform for analysing cryo-tomography data [121]. Progress on neural network–assisted cellular cryo-electron tomograms annotation has also been reported by Chen *et al.* [120].

Volume segmentation and connectomics

In neuroscience, it is critically important to visualize and understand the internal neuronal connections. Combining advanced sample preparation, serial block-face scanning electron microscopy (SBEM), big data, automated image segmentation and 3D reconstruction, an exciting branch of biotechnology, 'connectomics' has emerged. At the core of this technology, volumes of brain tissues are sectioned into thousands of thin slices for SBEM imaging. The images are segmented before being stacked to form a 3D model to study the internal wiring of the neural circuits. Neural networks are a natural fit for the segmentation task and were introduced to the connectomics workflows as early as 2007 [35,36]. However, it was not until the ISBI'12 conference that competition with large volume of hand-labelled data and carefully designed evaluation metrics (see the section on 'error measurement') that a deep learning approach became recognized as the state-of-the-art method [135]. While still not fully automated, the connectomic reconstruction of mouse retina [37] and fly brain [181] has attracted significant interest accompanied by further advancement in ML methods [103,122–124] and innovative data annotation through cloud-based environments [123] and crowdsourcing [37,135]. Using developments in multibeam SEM, Shapson-Coe *et al.* [182] have reported the reconstruction of a 1-mm^3 volume of brain tissue from 1.4 PB of SEM image data. A recent review on CNN-based segmentation methods in connectomics has been published by Ishii *et al.* [183] where some of the key publications [184–187] arising from the SNEMI3D open challenge (based on published data by Kasthuri *et al.* [188]) were compared.

Virus recognition

Finally, publications have emerged in the field of virus recognition and analysis [63,98,125–129], including the implementation of a more shallow FCN architecture [125], weight reduction of a U-Net without performance loss [98], training data enrichment by using a GAN for the generation of synthetic images [127] and evaluations of different object detection networks in terms of their performance in identifying the viral species severe acute respiratory syndrome coronavirus (SARS-CoV), middle east respiratory syndrome coronavirus (MERS-CoV), human immunodeficiency virus (HIV) and SARS-CoV-2 [128].

Denoising

As already introduced in the context of particle picking, Topaz [42] includes a denoising network trained using the Noise2Noise [147] scheme on thousands of electron micrographs acquired using direct electron detectors [94]. The generated network successfully improves the SNR by >2 dB on average over simple low-pass-filtered data and by ~ 20 dB over the raw images, while also performing well for micrographs acquired on indirect scintillator coupled detectors. Improved particle picking performance was manifest by an increase in the number of real picks by a factor of 2.15 for a particle that was particularly difficult to identify. Similarly, Buchholz *et al.* [93] and Tegunov *et al.* [66] have employed the Noise2Noise [147] training scheme.

Palovcak *et al.* [95] have further studied the quantitative evaluation of SNR enhancement, bias introduction and denoising effects on image processing when applying Noise2Noise [147] training to cryo-EM data sets. It was shown that CNNs trained by this training scheme provide robust results and that the SNR was significantly increased
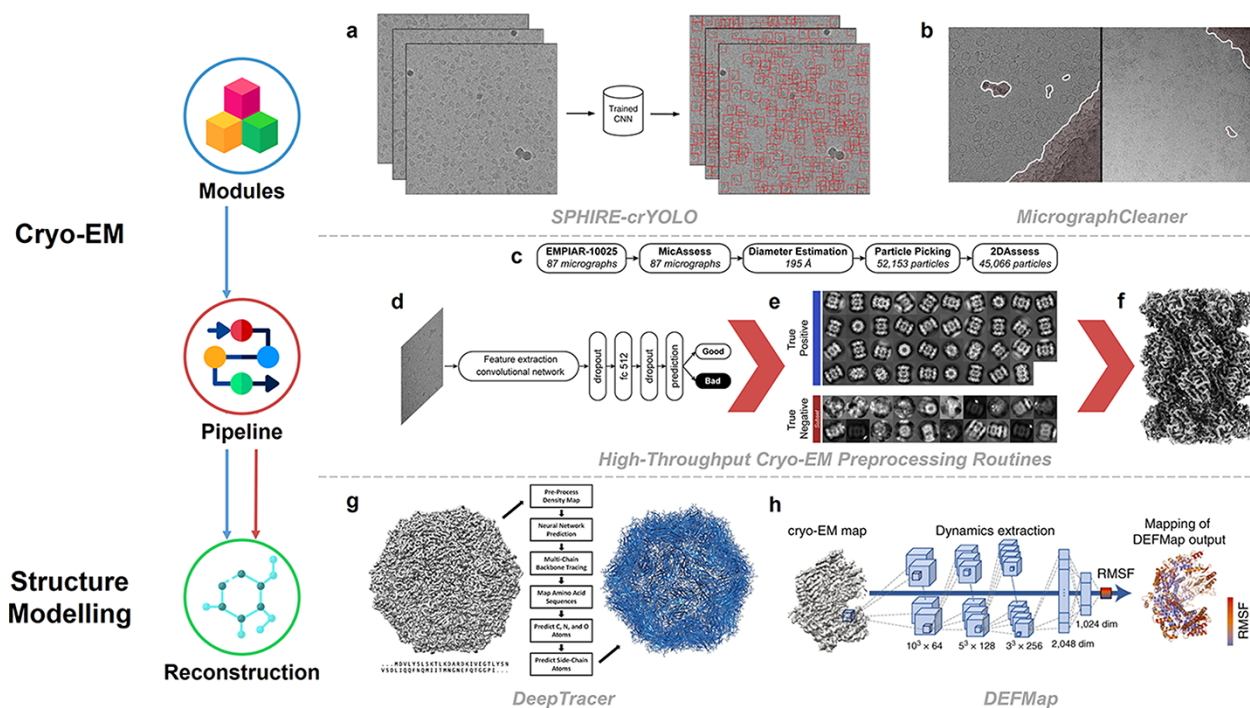
Fig. 2. Examples of deep neural networks applied in the field of cryo-EM and structure modelling. Modules describe neural networks that are trained and applied for a specific task, e.g. (a) SPHIRE-crYOLO (modified from Fig. 1, [72]) for particle picking and (b) Micrograph Cleaner (modified from Fig. 1, [92]) for reducing false positives in SPA. Pipeline refers to applications that employ multiple neural networks for automated multi-stage processing of datasets. (c) Workflow for the T20S proteasome (EMPIAR-10025) of the pipeline introduced by Li *et al.* [75]. Note that the particle picking step in this pipeline is conducted by the 'general' crYOLO network (d) Architecture of MicAssess, a neural network that assesses motion corrected cryo-EM micrographs for their quality. (e) 'good' and 'bad' 2D averages obtained by 2D Assess, a network employed for 2D classification of identified particles. (f) 3D electron density volume obtained from the particle stack outputted by the pipeline. Reconstruction refers to applications employing neural networks to reconstruct structures of dynamics associated with atomic fluctuations from cryo-EM electron density maps. (g) Model determination pipeline of DeepTracer (modified from Fig. 1, [97]). (h) Workflow of the prediction stage in DEFMap [119]. Icons used for 'Modules', 'Pipeline', and 'Reconstruction' were obtained from flaticon.com. Icon author: 'Freepik' for all three icons respectively. (c – f) reprinted from Structure, 28 (7), Li *et al.* [75], with permission from Elsevier. (h) reprinted by permission from Matsumoto *et al.* [119].

at all spatial frequencies. However, portions of uncorrelated noise are transformed into a statistically correlated bias at high spatial frequencies, which is possibly due to high spatial frequency features not being distinguished from noise at low SNR. However, this bias was yet shown to be removable by averaging large numbers of particle images as the bias is effectively random.

In contrast to the above, Lei *et al.* [43] approach denoising for cryo-EM by using a supervised Cascade of Denoising Autoencoders (CDAE), in which AEs of individual blocks are firstly trained with simulated data and subsequently fine-tuned with experimental data.

Super-resolution

Suveer *et al.* [62] have trained a modified DenseNet for super-resolution reconstruction with the intention of shortening acquisition times for biological specimens. The study showcased a successful proof of concept, using low and high image pairs obtained by rigid registration and compared with downsampled ones. A related study was published by de Haan *et al.* [108], focusing on super-resolution of SEM images of Au nanoparticles and a hydrogel.

Conclusions and future perspectives

We have reviewed the use of deep learning as currently applied to EM. We have highlighted the various network architectures

commonly used and provided metrics that are commonly used to assess their performance. We have also summarized and referenced the application of the use of neural networks across a range of EM-specific applications in the life and physical sciences.

In the future, it can be foreseen that deep learning techniques will become more prolific and powerful within the EM research community. Faster and better network architectures will undoubtedly extend capabilities in tackling more complex EM-related tasks in real-time data processing, enhanced data handling and instrument control. In combination, this will assist the development of EM research towards more statistically meaningful observations of many instances of heterogeneous samples as opposed to capturing isolated phenomena under restrictive imaging conditions. There are also still largely unexplored application areas for deep learning in EM, including the use of reinforcement learning and sophisticated AI-based instrumental control of complex optics. As a consequence, many routine EM tasks will become automated and no longer require expert human intervention.

Funding

Conflict of Interest

The authors declare that they have no conflict of interest.

References

1. Spurgeon S R, Ophus C, Jones L, Petford-Long A, Kalinin S V, Olszta M J, Dunin-Borkowski R E, Salmon N, Hattar K, Yang W-C D, Sharma R, Du Y, Chiaramonti A, Zheng H, Buck E C, Kovarik L, Penn R L, Li D, Zhang X, Murayama M, and Taheri M L (2021) Towards data-driven next-generation transmission electron microscopy. *Nat. Mater.* 20: 274–279.
2. Zhang D, Zhu Y, Liu L, Ying X, Hsiung C-E, Sougrat R, Li K, and Han Y (2018) Atomic-resolution transmission electron microscopy of electron beam–sensitive crystalline materials. *Science* 359: 675–679.
3. Ilett M, S'ari M, Freeman H, Aslam Z, Koniuch N, Afzali M, Cattle J, Hooley R, Roncal-Herrero T, Collins S M, Hondow N, Brown A, and Brydson R (2020) Analysis of complex, beam-sensitive materials by transmission electron microscopy and associated techniques. *Phil. Trans. R. Soc. A* 378: 20190601.
4. Voulodimos A, Doulamis N, Doulamis A, and Protopapadakis E (2018) Deep learning for computer vision: a brief review. *Comput. Intell. Neurosc.* 2018: 1–13.
5. Kalinin S V, Sumpter B G, and Archibald R K (2015) Big–deep–smart data in imaging for guiding materials design. *Nat. Mater.* 14: 973–980.
6. Kalinin S V, Lupini A R, Dyck O, Jesse S, Ziatdinov M, and Vasudevan R K (2019) Lab on a beam—big data and artificial intelligence in scanning transmission electron microscopy. *MRS Bulletin* 44: 565–575.
7. Dan J, Zhao X, and Pennycook S J (2019) A machine perspective of atomic defects in scanning transmission electron microscopy. *InfoMat.* 1: 359–375.
8. Ge M, Su F, Zhao Z, and Su D (2020) Deep learning analysis on microscopic imaging in materials science. *Mater. Today Nano* 11: 100087.
9. Ede J M (2021) Deep learning in electron microscopy. *Mach. Learn* 2: 011004.
10. Guo Y, Liu Y, Oerlemans A, Lao S, Wu S, and Lew M S (2016) Deep learning for visual understanding: a review. *Neurocomputing* 187: 27–48.
11. LeCun Y, Bengio Y, and Hinton G (2015) Deep learning. *Nature* 521: 436–444.
12. Krizhevsky A, Sutskever I, and Hinton G E (2012) ImageNet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* 25: 1097–1105.
13. Liu L, Ouyang W, Wang X, Fieguth P, Chen J, Liu X, and Pietikäinen M (2020) Deep learning for generic object detection: a survey. *Int. J. Comput. Vis.* 128: 261–318.
14. Hao S, Zhou Y, and Guo Y (2020) A brief survey on semantic segmentation with deep learning. *Neurocomputing* 406: 302–321.
15. Rawat W and Wang Z (2017) Deep convolutional neural networks for image classification: a comprehensive review. *Neural Comput.* 29: 2352–2449.
16. Jam J, Kendrick C, Walker K, Drouard V, Hsu J G-S, and Yap M H (2021) A comprehensive review of past and present image inpainting methods. *Comput. Vis. Image Underst.* 203: 103147.
17. Goodfellow I J, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, and Bengio Y (2014) Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* 27: 2672–2680.
18. Creswell A, White T, Dumoulin V, Arulkumaran K, Sengupta B, and Bharath A A (2018) Generative adversarial networks: an overview. *IEEE Signal Process. Mag.* 35: 53–65.
19. Gui J, Sun Z, Wen Y, Tao D, and Ye J (2020) A review on generative adversarial networks: algorithms, theory, and applications. *arXiv*, 2001.06937. https://arxiv.org/abs/2001.06937. 14 June 2021.
20. Fan L, Zhang F, Fan H, and Zhang C (2019) Brief review of image denoising techniques. *Vis. Comput. Ind. Biomed. Art* 2: 7.
21. Goodfellow I, Bengio Y, and Courville A (2016) *Deep Learning* (MIT Press, Cambridge, Massachusetts; London, England).
22. Géron A (2019) *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems* (O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472).
23. Zhang A, Lipton Z C, Li M, and Smola A J (2021) Dive into deep learning. *arXiv*, 2106.11342. https://d2l.ai/. 24 August 2021.
24. Rosenblatt F (1958) The perceptron: a probabilistic model for information storage and organization in the brain. *Psychol. Rev.* 65: 386–408.
25. Al-Khedher M A, Pezeshki C, McHale J L, and Knorr F J (2007) Quality classification via Raman identification and SEM analysis of carbon nanotube bundles using artificial neural networks. *Nanotechnology* 18: 355703.
26. Trujillo M C R, Alarcón T E, Dalmau O S, and Ojeda A Z (2017) Segmentation of carbon nanotube images through an artificial neural network. *Soft. Comput.* 21: 611–625.
27. Frei M and Kruis F E (2018) Fully automated primary particle size analysis of agglomerates on transmission electron microscopy images via artificial neural networks. *Powder Technol.* 332: 120–130.
28. Kirschner H and Hillebrand R (2000) Neural networks for HREM image analysis. *Inf Sci (NY)* 129: 31–44.
29. van den Broek W and Koch C T (2012) Method for retrieval of the three-dimensional object potential by inversion of dynamical electron scattering. *Phys. Rev. Lett.* 109: 245502.
30. van den Broek W and Koch C T (2013) General framework for quantitative three-dimensional reconstruction from arbitrary detection geometries in TEM. *Phys. Rev. B* 87: 184108.
31. Pennington R S, van den Broek W, and Koch C T (2014) Third-dimension information retrieval from a single convergent-beam transmission electron diffraction pattern using an artificial neural network. *Phys. Rev. B* 89: 205409.
32. Heindl E, Rau W D, and Lichte H (1996) The phase-shift method in electron-off-axis holography: using neural network techniques. *Ultramicroscopy* 64: 87–97.
33. Meyer R and Heindl E (1998) Reconstruction of off-axis electron holograms using a neural net. *J. Microsc.* 191: 52–59.
34. Roest L I, van Heijst S E, Maduro L, Rojo J, and Conesa-Boj S (2021) Charting the low-loss region in electron energy loss spectroscopy with machine learning. *Ultramicroscopy* 222: 113202.
35. Jain V, Murray J F, Roth F, Turaga S, Zhigulin V, Briggman K L, Helmstaedter M N, Denk W, and Seung H S (2007) Supervised learning of image restoration with convolutional networks. In: *2007 IEEE 11th International Conference on Computer Vision*, pp 1–8 (Rio de Janeiro, Brazil).
36. Turaga S C, Murray J F, Jain V, Roth F, Helmstaedter M, Briggman K, Denk W, and Seung H S (2010) Convolutional networks can learn to generate affinity graphs for image segmentation. *Neural Comput.* 22: 511–538.
37. Helmstaedter M, Briggman K L, Turaga S C, Jain V, Seung H S, and Denk W (2013) Connectomic reconstruction of the inner plexiform layer in the mouse retina. *Nature* 500: 168–174.
38. Ogura T and Sato C (2001) An automatic particle pickup method using a neural network applicable to low-contrast electron micrographs. *J. Struct. Biol.* 136: 227–238.
39. Ogura T and Sato C (2004) Automatic particle pickup method using a neural network has high accuracy by applying an initial weight derived from eigenimages: a new reference free method for single-particle analysis. *J. Struct. Biol.* 145: 63–75.
40. LeCun Y (1987) Modèles connexionistes de l'apprentissage. PhD Thesis.

41. Kalinin S V, Dyck O, Jesse S, and Ziatdinov M (2021) Exploring order parameters and dynamic processes in disordered systems via variational autoencoders. *Sci. Adv.* 7: eabd5084.

42. Bepler T, Morin A, Rapp M, Brasch J, Shapiro L, Noble A J, and Berger B (2019) Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat. Methods* 16: 1153–1160.

43. Lei H and Yang Y (2021) CDAE: a cascade of denoising autoencoders for noise reduction in the clustering of single-particle cryo-EM images. *Front. Genet.* 11: 627746.

44. Lecun Y, Bottou L, Bengio Y, and Haffner P (1998) Gradient-based learning applied to document recognition. *Proc. IEEE* 86: 2278–2324.

45. Förster G D, Castan A, Loiseau A, Nelayah J, Alloyeau D, Fossard F, Bichara C, and Amara H (2020) A deep learning approach for determining the chiral indices of carbon nanotubes from high-resolution transmission electron microscopy images. *Carbon* 169: 465–474.

46. Yoshioka H and Honda T (2021) Determination of the interface between amorphous insulator and crystalline 4H–SiC in transmission electron microscope image by using convolutional neural network. *AIP Adv.* 11: 015101.

47. Zhu Y, Ouyang Q, and Mao Y (2017) A deep convolutional neural network approach to single-particle recognition in cryo-electron microscopy. *BMC Bioinform.* 18: 348.

48. Chowdhury A, Kautz E, Yener B, and Lewis D (2016) Image driven machine learning methods for microstructure recognition. *Comput. Mater. Sci.* 123: 176–187.

49. Xu W and LeBeau J M (2018) A deep convolutional neural network to analyze position averaged convergent beam electron diffraction patterns. *Ultramicroscopy* 188: 59–69.

50. Wang F, Gong H, Liu G, Li M, Yan C, Xia T, Li X, and Zeng J (2016) DeepPicker: a deep learning approach for fully automated particle picking in cryo-EM. *J. Struct. Biol.* 195: 325–336.

51. Zhang J, Wang Z, Chen Y, Han R, Liu Z, Sun F, and Zhang F (2019) PIXER: an automated particle-selection method based on segmentation using a deep neural network. *BMC Bioinform.* 20: 41.

52. Al-Azzawi A, Ouadou A, Max H, Duan Y, Tanner J J, and Cheng J (2020) DeepCryoPicker: fully automated deep neural network for single protein particle picking in cryo-EM. *BMC Bioinform.* 21: 509.

53. Weber G H, Ophus C, and Ramakrishnan L (2018) Automated labeling of electron microscopy images using deep learning. In: *2018 IEEE/ACM Machine Learning in HPC Environments (MLHPC)*, pp. 26–36 (Dallas, TX, USA).

54. Simonyan K and Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. *arXiv*, 1409.1556. https://arxiv.org/abs/1409.1556. 10 May 2021.

55. DeCost B L, Francis T, and Holm E A (2017) Exploring the microstructure manifold: image texture representations applied to ultrahigh carbon steel microstructures. *Acta. Mater.* 133: 30–40.

56. Azimi S M, Britz D, Engstler M, Fritz M, and Mücklich F (2018) Advanced steel microstructural classification by deep learning methods. *Sci. Rep.* 8: 2128.

57. Dennler N, Foncubierta-Rodriguez A, Neupert T, and Sousa M (2021) Learning-based defect recognition for quasi-periodic HRSTEM images. *Micron* 146: 103069.

58. Luo Q, Holm E A, and Wang C (2020) A transfer learning approach for improved classification of carbon nanomaterials from TEM images. *Nanoscale Adv.* 3: 206–213.

59. Sanchez-Garcia R, Segura J, Maluenda D, Carazo J M, and Sorzano C O S (2018) Deep Consensus, a deep learning-based approach for particle pruning in cryo-electron microscopy. *IUCrJ* 5: 854–865.

60. Huang G, Liu Z, van der Maaten L, and Weinberger K Q (2017) Densely connected convolutional networks. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp 2261–2269 (Honolulu, HI, USA).

61. Roberts G, Haile S Y, Sainju R, Edwards D J, Hutchinson B, and Zhu Y (2019) Deep learning for semantic segmentation of defects in advanced STEM images of steels. *Sci. Rep.* 9: 12744.

62. Suveer A, Gupta A, Kylberg G, and Sintorn I-M (2019) Super-resolution reconstruction of transmission electron microscopy images using deep learning. In: *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pp 548–551 (Venice, Italy).

63. Nanni L, de Luca E, Facin M L, and Maguolo G (2020) Deep learning and handcrafted features for virus image classification. *J. Imaging* 6: 143.

64. He K, Zhang X, Ren S, and Sun J (2016) Deep residual learning for image recognition. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp 770–778 (Las Vegas, NV, USA).

65. Kharin A Y (2020) Deep learning for scanning electron microscopy: synthetic data for the nanoparticles detection. *Ultramicroscopy* 219: 113125.

66. Tegunov D and Cramer P (2019) Real-time cryo-electron microscopy data preprocessing with Warp. *Nat. Methods* 16: 1146–1152.

67. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, and Rabinovich A (2014) Going deeper with convolutions. *arXiv*, 1409.4842. https://arxiv.org/abs/1409.4842. 06 July 2021.

68. George B, Assaiya A, Roy R J, Kembhavi A, Chauhan R, Paul G, Kumar J, and Philip N S (2021) CASSPER is a semantic segmentation-based particle picking algorithm for single-particle cryo-electron microscopy. *Commun. Biol.* 4: 200.

69. Xu M, Chai X, Muthakana H, Liang X, Yang G, Zeev-Ben-Mordehai T, and Xing E P (2017) Deep learning-based subdivision approach for large scale macromolecules structure recovery from electron cryo tomograms. *Bioinformatics* 33: i13–i22.

70. Modarres M H, Aversa R, Cozzini S, Ciancio R, Leto A, and Brandino G P (2017) Neural network for nanoscience scanning electron microscope image recognition. *Sci. Rep.* 7: 13282.

71. Redmon J, Divvala S, Girshick R, and Farhadi A (2016) You only look once: unified, real-time object detection. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp 779–788 (Las Vegas, NV, USA).

72. Wagner T, Merino F, Stabrin M, Moriya T, Antoni C, Apelbaum A, Hagel P, Sitsel O, Raisch T, Prumbaum D, Quentin D, Roderer D, Tacke S, Siebolds B, Schubert E, Shaikh T R, Lill P, Gatsogiannis C, and Raunser S (2019) SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. *Commun. Biol.* 2: 218.

73. Wagner T and Raunser S (2020) The evolution of SPHIRE-crYOLO particle picking and its application in automated cryo-EM processing workflows. *Commun. Biol.* 3: 61.

74. Yokoyama Y, Terada T, Shimizu K, Nishikawa K, Kozai D, Shimada A, Mizoguchi A, Fujiyoshi Y, and Tani K (2020) Development of a deep learning-based method to identify 'good' regions of a cryo-electron microscopy grid. *Biophys. Rev.* 12: 349–354.

75. Li Y, Cash J N, Tesmer J J G, and Cianfrocco M A (2020) High-throughput cryo-EM enabled by user-free preprocessing routines. *Structure* 28: 858–869.

76. Long J and Darrell T (2015) Fully convolutional networks for semantic segmentation. In: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* pp. 3431–3440 (Boston, MA, USA).

77. Yao R, Qian J, and Huang Q (2019) Deep-learning with synthetic data enables automated picking of cryo-EM particle images of biological macromolecules. *Bioinformatics* 36: 1252–1259.

78. Badrinarayanan V, Handa A, and Cipolla R (2015) SegNet: a deep convolutional encoder-decoder architecture for robust semantic pixel-wise labelling. *arXiv*, 1505.07293. https://arxiv.org/abs/1505.07293. 8 June 2021.

79. Ziatdinov M, Dyck O, Maksov A, Li X, Sang X, Xiao K, Unocic R R, Vasudevan R, Jesse S, and Kalinin S V (2017) Deep learning of atomically resolved scanning transmission electron microscopy images: chemical identification and tracking local transformations. *ACS Nano* 11: 12742–12752.

80. Maksov A, Dyck O, Wang K, Xiao K, Geohegan D B, Sumpter B G, Vasudevan R K, Jesse S, Kalinin S V, and Ziatdinov M (2019) Deep learning analysis of defect and phase evolution during electron beam-induced transformations in WS2. *Npj Comput. Mater.* 5: 12.

81. Ziatdinov M, Dyck O, Li X, Sumpter B G, Jesse S, Vasudevan R K, and Kalinin S V (2019) Building and exploring libraries of atomic defects in graphene: scanning transmission electron and scanning tunneling microscopy study. *Sci. Adv.* 5: eaaw8989.

82. Cherukara M J, Nashed Y S G, and Harder R J (2018) Real-time coherent diffraction inversion using deep generative networks. *Sci. Rep.* 8: 16520.

83. Ronneberger O, Fischer P, and Brox T (2015) U-Net: convolutional networks for biomedical image segmentation. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, *Vol. 9351*, pp 234–241 (Munich, Germany).

84. Ziatdinov M, Dyck O, Jesse S, and Kalinin S V (2019) Atomic mechanisms for the Si atom dynamics in graphene: chemical transformations at the edge and in the bulk. *Adv. Funct. Mater.* 29: 1904480.

85. Madsen J, Liu P, Kling J, Wagner J B, Hansen T W, Winther O, and Schiøtz J (2018) A deep learning approach to identify local structures in atomic-resolution transmission electron microscopy images. *Adv. Theory Simul.* 1: 1800037.

86. Horwath J P, Zakharov D N, Mégret R, and Stach E A (2020) Understanding important features of deep learning models for segmentation of high-resolution transmission electron microscopy images. *Npj Comput. Mater.* 6: 108.

87. Yao L, Ou Z, Luo B, Xu C, and Chen Q (2020) Machine learning to reveal nanoparticle dynamics from liquid-phase TEM videos. *ACS Cent. Sci.* 6: 1421–1430.

88. Groschner C K, Choi C, and Scott M C (2021) Machine learning pipeline for segmentation and defect identification from high-resolution transmission electron microscopy data. *Microsc. Microanal.* 27: 1–8.

89. Ragone M, Yurkiv V, Song B, Ramsubramanian A, Shahbazian-Yassar R, and Mashayek F (2020) Atomic column heights detection in metallic nanoparticles using deep convolutional learning. *Comput. Mater. Sci.* 180: 109722.

90. Lee J, Jeong C, and Yang Y (2021) Single-atom level determination of 3-dimensional surface atomic structure via neural network-assisted atomic electron tomography. *Nat. Commun.* 12: 1962.

91. Han Y, Jang J, Cha E, Lee J, Chung H, Jeong M, Kim T-G, Chae B G, Kim H G, Jun S, Hwang S, Lee E, and Ye J C (2021) Deep learning STEM-EDX tomography of nanocrystals. *Nat. Mach. Intell.* 3: 267–274.

92. Sanchez-Garcia R, Segura J, Maluenda D, Sorzano C O S, and Carazo J M (2020) MicrographCleaner: a python package for cryo-EM micrograph cleaning using deep learning. *J. Struct. Biol.* 210: 107498.

93. Buchholz T-O, Jordan M, Pigino G, and Jug F (2019) CryoCARE: content-aware image restoration for cryo-transmission electron microscopy data. In: *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pp 502–506 (Venice, Italy).

94. Bepler T, Kelley K, Noble A J, and Berger B (2020) TopazDenoise: general deep denoising models for cryoEM and cryoET. *Nat Commun.* 11: 5208.

95. Palovcak E, Asarnow D, Campbell M G, Yu Z, and Cheng Y (2020) Enhancing the signal-to-noise ratio and generating contrast for cryo-EM images with convolutional neural networks. *IUCrJ* 7: 1142–1150.

96. Avramov T K, Vyenielo D, Gomez-Blanco J, Adinarayanan S, Vargas J, and Si D (2019) Deep learning for validating and estimating resolution of cryo-electron microscopy density maps. *Molecules* 24: 1181.

97. Pfab J, Phan N M, and Si D (2021) DeepTracer for fast de novo cryo-EM protein structure modeling and special studies on CoV-related complexes. *Proc. Natl. Acad. Sci. USA* 118: e2017525118.

98. Matuszewski D J and Sintorn I-M (2019) Reducing the U-Net size for practical scenarios: virus recognition in electron microscopy images. *Comput. Methods Programs Biomed.* 178: 31–39.

99. Lin R, Zhang R, Wang C, Yang X-Q, and Xin H L (2021) TEMImageNet training library and AtomSegNet deep-learning models for high-precision atom segmentation, localization, denoising, and deblurring of atomic-resolution images. *Sci. Rep.* 11: 5386.

100. Wang F, Eljarrat A, Müller J, Henninen T R, Erni R, and Koch C T (2020) Multi-resolution convolutional neural networks for inverse problems. *Sci. Rep.* 10: 5730.

101. Diakogiannis F I, Waldner F, Caccetta P, and Wu C (2020) ResUNet-a: a deep learning framework for semantic segmentation of remotely sensed data. *ISPRS J. Photogramm. Remote Sens.* 162: 94–114.

102. Lee C-H, Khan A, Luo D, Santos T P, Shi C, Janicek B E, Kang S, Zhu W, Sobh N A, Schleife A, Clark B K, and Huang P Y (2020) Deep learning enabled strain mapping of single-atom defects in two-dimensional transition metal dichalcogenides with sub-picometer precision. *Nano Lett.* 20: 3369–3377.

103. Fakhry A, Zeng T, and Ji S (2017) Residual deconvolutional networks for brain electron microscopy image segmentation. *IEEE Trans. Med. Imaging* 36: 447–456.

104. Ding G, Liu Y, Zhang R, and Xin H L (2019) A joint deep learning model to recover information and reduce artifacts in missing-wedge sinograms for electron tomography and beyond. *Sci. Rep.* 9: 12803.

105. Wang C, Ding G, Liu Y, and Xin H L (2020) 0.7 Å resolution electron tomography enabled by deep-learning-aided information recovery. *Adv. Intell. Syst.* 2: 2000152.

106. Wang F, Henninen T R, Keller D, and Erni R (2020) Noise2Atom: unsupervised denoising for scanning transmission electron microscopy images. *Appl. Microsc.* 50: 23.

107. Ede J M and Beanland R (2020) Partial scanning transmission electron microscopy with deep learning. *Sci. Rep.* 10: 8332.

108. de Haan K, Ballard Z S, Rivenson Y, Wu Y, and Ozcan A (2019) Resolution enhancement in scanning electron microscopy using deep learning. *Sci. Rep.* 9: 12050.

109. Jamali V, Hargus C, Ben-Moshe A, Aghazadeh A, Ha H D, Mandadapu K K, and Alivisatos A P (2021) Anomalous nanoparticle surface diffusion in LCTEM is revealed by deep learning-assisted analysis. *Proc. Natl. Acad. Sci. USA* 118: e2017616118.

110. Oktay A B and Gurses A (2019) Automatic detection, localization and segmentation of nano-particles with deep learning in microscopy images. *Micron* 120: 113–119.

111. Li W, Field K G, and Morgan D (2018) Automated defect analysis in electron microscopic images. *Npj Comput. Mater.* 4: 36.

112. Aguiar J A, Gong M L, Unocic R R, Tasdizen T, and Miller B D (2019) Decoding crystallography from high-resolution electron imaging and diffraction datasets with deep learning. *Sci. Adv.* 5: eaaw1949.

113. Cherukara M J, Zhou T, Nashed Y, Enfedaque P, Hexemer A, Harder R J, and Holt M V (2020) AI-enabled high-resolution

scanning coherent diffraction imaging. *Appl. Phys. Lett.* 117: 044103.

114. Chatzidakis M and Botton G A (2019) Towards calibration-invariant spectroscopy using deep learning. *Sci. Rep.* 9: 2126.

115. Ede J M and Beanland R (2019) Improving electron micrograph signal-to-noise with an atrous convolutional encoder-decoder. *Ultramicroscopy* 202: 18–25.

116. Xiao Y and Yang G (2017) A fast method for particle picking in cryo-electron micrographs based on fast R-CNN. *AIP Conf. Proc.* 1836: 020080.

117. Xu K, Wang Z, Shi J, Li H, and Zhang Q C (2019) A2-Net: molecular structure estimation from cryo-EM density volumes. *Proc. AAAI Conf Artif. Intell.* 33: 1230–1237.

118. Si D, Moritz S A, Pfab J, Hou J, Cao R, Wang L, Wu T, and Cheng J (2020) Deep learning to predict protein backbone structure from high-resolution cryo-EM density maps. *Sci. Rep.* 10: 4282.

119. Matsumoto S, Ishida S, Araki M, Kato T, Terayama K, and Okuno Y (2021) Extraction of protein dynamics information from cryo-EM maps using deep learning. *Nat. Mach. Intell.* 3: 153–160.

120. Chen M, Dai W, Sun S Y, Jonasch D, He C Y, Schmid M F, Chiu W, and Ludtke S J (2017) Convolutional neural networks for automated annotation of cellular cryo-electron tomograms. *Nat. Methods* 14: 983–985.

121. Zeng X and Xu M (2020) AITom: open-source AI platform for cryo-electron tomography data analysis. *arXiv,* 1911.03044. https://arxiv.org/abs/1911.03044. 21 June 2021.

122. Zeng T, Wu B, and Ji S (2017) DeepEM3D: approaching human-level performance on 3D anisotropic EM image segmentation. *Bioinformatics* 33: 2555–2562.

123. Haberl M G, Churas C, Tindall L, Boassa D, Phan S, Bushong E A, Madany M, Akay R, Deerinck T J, Peltier S T, and Ellisman M H (2018) CDeep3M—Plug-and-Play cloud-based deep learning for image segmentation. *Nat. Methods* 15: 677–680.

124. Quan T M, Hildebrand D G C, and Jeong W-K (2021) Fusion-Net: a deep fully residual convolutional neural network for image segmentation in connectomics. *Front. Comput. Sci.* 3: 613981.

125. Ito E, Sato T, Sano D, Utagawa E, and Kato T (2018) Virus particle detection by convolutional neural network in transmission electron microscopy images. *Food Environ. Virol.* 10: 201–208.

126. Devan K S, Walther P, von Einem J, Ropinski T, Kestler H A, and Read C (2019) Detection of herpesvirus capsids in transmission electron microscopy images using transfer learning. *Histochem. Cell Biol.* 151: 101–114.

127. Devan K S, Walther P, von Einem J, Ropinski T, Kestler H, and Read C (2021) Improved automatic detection of herpesvirus secondary envelopment stages in electron microscopy by augmenting training data with synthetic labelled images generated by a generative adversarial network. *Cell. Microbiol.* 23: e13280.

128. Zhang L and Yan W Q (2020) Deep learning methods for virus identification from digital images. In: *2020 35th International Conference on Image and Vision Computing New Zealand (IVCNZ)*, pp 1–6 (Wellington, New Zealand).

129. Xiao C, Chen X, Xie Q, Li G, Xiao H, Song J, and Han H (2021) Virus identification in electron microscopy images by residual mixed attention network. *Comput. Methods Programs Biomed.* 198: 105766.

130. von Chamier L, Laine R F, Jukkala J, Spahn C, Krentzel D, Nehme E, Lerche M, Hernández-Pérez S, Mattila P K, Karinou E, Holden S, Solak A C, Krull A, Buchholz T-O, Jones M L, Royer L A, Leterrier C, Shechtman Y, Jug F, Heilemann M, Jacquemet G, and Henriques R (2021) Democratising deep learning for microscopy with ZeroCostDL4Mic. *Nat. Commun.* 12: 2276.

131. Bishop C (2006) *Pattern Recognition and Machine Learning*, (Springer-Verlag, New York).

132. Murphy K P (2012) *Machine Learning: A Probabilistic Perspective*, (MIT press, Cambridge, Massachusetts; London, England).

133. Langlois R and Frank J (2011) A clarification of the terms used in comparing semi-automated particle selection algorithms in Cryo-EM. *J. Struct. Biol.* 175: 348–352.

134. Fawcett T (2006) An introduction to ROC analysis. *Pattern Recogn. Lett.* 27: 861–874.

135. Arganda-Carreras I, Turaga S C, Berger D R, Cireşan D, Giusti A, Gambardella L M, Schmidhuber J, Laptev D, Dwivedi S, Buhmann J M, Liu T, Seyedhosseini M, Tasdizen T, Kamentsky L, Burget R, Uher V, Tan X, Sun C, Pham T D, Bas E, Uzunbas M G, Cardona A, Schindelin J, and Seung H S (2015) Crowdsourcing the creation of image segmentation algorithms for connectomics. *Front. Neuroanat.* 9: 142.

136. Horé A and Ziou D (2010) Image quality metrics: PSNR vs. SSIM. In: *2010 20th International Conference on Pattern Recognition*, pp 2366–2369 (Istanbul, Turkey).

137. Kingma D P and Welling M (2014) Auto-encoding variational Bayes. *arXiv,* 1312.6114. https://arxiv.org/abs/1312.6114. 1 June 2021.

138. Breiman L (2001) Random forests. *Mach. Learn.* 45: 5–32. https://arxiv.org/abs/1312.6114. 30 May 2021.

139. Bai S, Kolter J Z, and Koltun V (2018) An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv,* 1803.01271. https://arxiv.org/abs/1803.01271. 30 May 2021.

140. Lin T-Y, Goyal P, Girshick R, He K, and Dollár P (2017) Focal loss for dense object detection. *arXiv,* 1708.02002. https://arxiv.org/abs/1708.02002. 30 May 2021.

141. Oxley M P, Yin J, Borodinov N, Somnath S, Ziatdinov M, Lupini A, Jesse S, Vasudevan R K, and Kalinin S V (2020) Deep learning of interface structures from simulated 4D STEM data: cation intermixing vs. roughening. *Mach. Learn Sci. Technol* 1: 04LT01.

142. Rojo J (2018) Machine learning tools for global PDF fits. *arXiv*, 1809.04392. https://arxiv.org/abs/1809.04392. 31 May 2021.

143. Gao J, Harland-Lang L, and Rojo J (2017) The structure of the proton in the LHC precision era. *arXiv*, 1709.04922. https://arxiv.org/abs/1709.04922. 31 May 2021.

144. Zhou Z, Siddiquee M M R, Tajbakhsh N, and Liang J (2018) UNet++: a nested U-Net architecture for medical image segmentation. *DLMIA 2018, ML-CDS 2018* 11045: 3–11.

145. Çiçek Ö, Abdulkadir A, Lienkamp S S, Brox T, and Ronneberger O (2016) 3D U-Net: learning dense volumetric segmentation from sparse annotation. *MICCAI* 2016: 424–432.

146. Maiden A M and Rodenburg J M (2009) An improved ptychographical phase retrieval algorithm for diffractive imaging. *Ultramicroscopy* 109: 1256–1262.

147. Lehtinen J, Munkberg J, Hasselgren J, Laine S, Karras T, Aittala M, and Aila T (2015) Noise2Noise: learning image restoration without clean data. In: *Proceedings of the 35th International Conference on Machine Learning, PMLR, Vol. 80*, pp 2965–2974 (Stockholm, Sweden).

148. Batson J and Royer L (2019) Noise2Self: blind denoising by self-supervision. In: *Proceedings of the 36th International Conference on Machine Learning, PMLR, Vol. 97*, pp 524–533 (Long Beach, CA, USA).

149. Krull A, Buchholz T-O, and Jug F (2019) Noise2Void - learning denoising from single noisy images. In: *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp 2124–2132 (Long Beach, CA, USA).

150. Laine S, Karras T, Lehtinen J, and Aila T (2019) High-quality self-supervised deep image denoising. *Adv. Neural Inf. Process. Syst.* 32: 6970–6980.

151. Krull A, Vičar T, Prakash M, Lalit M, and Jug F (2020) Probabilistic Noise2Void: unsupervised content-aware denoising. *Front. Comput. Sci.* 2: 5.

152. Prakash M, Lalit M, Tomancak P, Krull A, and Jug F (2020) Fully unsupervised probabilistic Noise2Void. In: *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, pp 154–158 (Iowa City, IA, USA).

153. Ede J M (2020) Warwick electron microscopy datasets. *Mach. Learn.* 1: 045003.

154. Chen L-C, Papandreou G, Schroff F, and Adam H (2017) Rethinking atrous convolution for semantic image segmentation. *arXiv*, 1706.05587. https://arxiv.org/abs/1706.05587. 14 June 2021.

155. Chen L-C, Zhu Y, Papandreou G, Schroff F, and Adam H (2018) Encoder-decoder with atrous separable convolution for semantic image segmentation. *Comput Vision – ECCV 2018* 11211: 833–851.

156. Trampert P, Bourghorbel F, Potocek P, Peemen M, Schlinkmann C, Dahmen T, and Slusallek P (2018) How should a fixed budget of dwell time be spent in scanning electron microscopy to optimize image quality? *Ultramicroscopy* 191: 11–17.

157. Stevens A, Yang H, Carin L, Arslan I, and Browning N D (2014) The potential for Bayesian compressive sensing to significantly reduce electron dose in high-resolution STEM images. *Microscopy* 63: 41–51.

158. Stevens A, Luzi L, Yang H, Kovarik L, Mehdi B L, Liyu A, Gehm M E, and Browning N D (2018) A sub-sampled approach to extremely low-dose STEM. *Appl. Phys. Lett.* 112: 043104.

159. Rotunno E, Tavabi A H, Rosi P, Frabboni S, Tiemeijer P, Dunin-Borkowski R E, and Grillo V (2021) Alignment of electron optical beam shaping elements using a convolutional neural network. *Ultramicroscopy* 228: 113338.

160. Kühlbrandt W (2014) The resolution revolution. *Science* 343: 1443–1444.

161. Cheng Y (2015) Single-particle cryo-EM at crystallographic resolution. *Cell* 161: 450–457.

162. Cheng Y (2018) Single-particle cryo-EM - how did it get here and where will it go. *Science* 361: 876–880.

163. Frank J (1989) Image analysis of single macromolecules. *Electron. Microsc. Rev.* 2: 53–74.

164. Sigworth F J (2016) Principles of cryo-EM single-particle image processing. *Microscopy* 65: 57–67.

165. Ogura T and Sato C (2004) Auto-accumulation method using simulated annealing enables fully automatic particle pickup completely free from a matching template or learning data. *J. Struct. Biol.* 146: 344–358.

166. Liao M, Cao E, Julius D, and Cheng Y (2013) Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature* 504: 107–112.

167. Bartesaghi A, Matthies D, Banerjee S, Merk A, and Subramaniam S (2014) Structure of β-galactosidase at 3.2-Å resolution obtained by cryo-electron microscopy. *Proc. Natl. Acad. Sci. USA* 111: 11709–11714.

168. Wong W, Bai X, Brown A, Fernandez I S, Hanssen E, Condron M, Tan Y H, Baum J, and Scheres S H (2014) Cryo-EM structure of the Plasmodium falciparum 80S ribosome bound to the anti-protozoan drug emetine. *Elife* 3: e03080.

169. Ru H, Chambers M G, Fu T-M, Tong A B, Liao M, and Wu H (2015) Molecular mechanism of V(D)J recombination from synaptic RAG1-RAG2 complex structures. *Cell* 163: 1138–1152.

170. Danev R and Baumeister W (2016) Cryo-EM single particle analysis with the Volta phase plate. *Elife* 5: e13046.

171. Lee C-H and MacKinnon R (2017) Structures of the human HCN1 hyperpolarization-activated channel. *Cell* 168: 111–120.

172. Al-Azzawi A, Ouadou A, Tanner J J, and Cheng J (2019) A super-clustering approach for fully automated single particle picking in cryo-EM. *Genes* 10: 666.

173. Al-Azzawi A, Ouadou A, Tanner J J, and Cheng J (2019) AutoCryoPicker: an unsupervised learning approach for fully automated single particle picking in Cryo-EM images. *BMC Bioinform.* 20: 326.

174. Pohlen T, Hermans A, Mathias M, and Leibe B (2017) Full-resolution residual networks for semantic segmentation in street scenes. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp 3309–3318 (Honolulu, HI, USA).

175. Szegedy C, Ioffe S, Vanhoucke V, and Alemi A (2017) Inception-v4, Inception-Resnet and the impact of residual connections on learning. In: *AAAI'17: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, pp 4278–4284 (San Francisco, CA, USA).

176. Jiménez-Moreno A, Střelák D, Filipovič J, Carazo J M, and Sorzano C O S (2021) DeepAlign, a 3D alignment method based on regionalized deep learning for cryo-EM. *J. Struct. Biol.* 213: 107712.

177. Ramírez-Aportela E, Mota J, Conesa P, Carazo J M, and Sorzano C O S (2019) DeepRes: a new deep-learning- and aspect-based local resolution method for electron-microscopy maps. *IUCrJ* 6: 1054–1063.

178. Zhong E D, Bepler T, Berger B, and Davis J H (2021) CryoDRGN: reconstruction of heterogeneous cryo-EM structures using neural networks. *Nat. Methods* 18: 176–185.

179. Malhotra S, Träger S, Peraro M D, and Topf M (2019) Modelling structures in cryo-EM maps. *Curr. Opin. Struc. Biol.* 58: 105–114.

180. Kim D N, Gront D, and Sanbonmatsu K Y (2020) Practical considerations for atomistic structure modeling with cryo-EM maps. *J. Chem. Inf. Model.* 60: 2436–2442.

181. Dorkenwald S, McKellar C, Macrina T, Kemnitz N, Lee K, Lu R, Wu J, Popovych S, Mitchell E, Nehoran B, Jia Z, Bae J A, Mu S, Ih D, Castro M, Ogedengbe O, Halageri A, Ashwood Z, Zung J, Brittain D, Collman F, Schneider-Mizell C, Jordan C, Silversmith W, Baker C, Deutsch D, Encarnacion-Rivera L, Kumar S, Burke A, Gager J, Hebditch J, Koolman S, Moore M, Morejohn S, Silverman B, Willie K, Willie R, Yu S, Murthy M, and Seung H S (2020) FlyWire: online community for whole-brain connectomics. *bioRxiv* 2020: 30.274225.

182. Shapson-Coe A, Januszewski M, Berger D R, Pope A, Wu Y, Blakely T, Schalek R L, Li P, Wang S, Maitin-Shepard J, Karlupia N, Dorkenwald S, Sjostedt E, Leavitt L, Lee D, Bailey L, Fitzmaurice A, Kar R, Field B, Wu H, Wagner-Carena J, Aley D, Lau J, Lin Z, Wei D, Pfister H, Peleg A, Jain V, and Lichtman J W (2021) A connectomic study of a petascale fragment of human cerebral cortex. *bioRxiv* 2021.29.446289.

183. Ishii S, Lee S, Urakubo H, Kume H, and Kasai H (2020) Generative and discriminative model-based approaches to microscopic image restoration and segmentation. *Microscopy* 69: 79–91.

184. Lee K, Zung J, Li P, Jain V, and Seung H S (2017) Superhuman accuracy on the SNEMI3D connectomics challenge. *arXiv*, 1706.00120. https://arxiv.org/abs/1706.00120. 20 August 2021.

185. Januszewski M, Kornfeld J, Li P H, Pope A, Blakely T, Lindsey L, Maitin-Shepard J, Tyka M, Denk W, and Jain V (2018) High-precision automated reconstruction of neurons with flood-filling networks. *Nat. Methods* 8: 323.

186. Haehn D, Hoffer J, Matejek B, Suissa-Peleg A, Al-Awami A, Kamentsky L, Gonda F, Meng E, Zhang W, Schalek R, Wilson A, Parag T, Beyer J, Kaynig V, Jones T, Tompkin J, Hadwiger M, Lichtman J, and Pfister H (2017) Scalable interactive visualization for connectomics. *Informatics* 4: 29.

187. Meirovitch Y, Mi L, Saribekyan H, Matveev A, Rolnick D, and Shavit N (2019) Cross-classification clustering: an efficient multi-object tracking technique for 3-D instance segmentation in connectomics. In: *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp 8417–8427 (Long Beach, CA, USA).

188. Kasthuri N, Hayworth K J, Berger D R, Schalek R L, Conchello J A, Knowles-Barley S, Lee D, Vázquez-Reina A, Kaynig V, Jones T R, Roberts M, Morgan J L, Tapia J C, Seung H S, Roncal W G, Vogelstein J T, Burns R, Sussman D L, Priebe C E, Pfister H, and Lichtman J W (2015) Saturated reconstruction of a volume of neocortex. *Cell* 162: 648–661.