Microscopy

DOI: 10.1111/jmi.13413

ORIGINAL ARTICLE

Deep learning assisted high-resolution microscopy image processing for phase segmentation in functional composite materials

Ganesh Raghavendran1Bing Han1,2Fortune Adekogbe4Shuang Bai2,3Bingyu Lu1William Wu5Minghao Zhang3Ying Shirley Meng1,3

¹Aiiso Yufeng Li Family Department of Chemical and Nano Engineering, University of California San Diego, La Jolla, California, USA

²Department of Materials Science and Engineering, University of California San Diego, La Jolla, California, USA

³Pritzker School of Molecular Engineering, University of Chicago, Chicago, Illinois, USA

⁴Department of Chemical and Petroleum Engineering, University of Lagos, Lagos, Nigeria

⁵Department of Computer Science, University of California San Diego, La Jolla, California, USA

Correspondence

Minghao Zhang, Pritzker School of Molecular Engineering, University of Chicago, Chicago, IL 60637, USA. Email: miz016@uchicago.edu

Ying Shirley Meng, Aiiso Yufeng Li Family Department of Chemical and Nano Engineering, University of California San Diego, La Jolla, CA 92093, USA. Email: shirleymeng@uchicago.edu

Funding information

National Science Foundation, Grant/Award Number: ECCS-1542148; Thermo Fisher Scientific

Abstract

In the domain of battery research, the processing of high-resolution microscopy images is a challenging task, as it involves dealing with complex images and requires a prior understanding of the components involved. The utilisation of deep learning methodologies for image analysis has attracted considerable interest in recent years, with multiple investigations employing such techniques for image segmentation and analysis within the realm of battery research. However, the automated analysis of high-resolution microscopy images for detecting phases and components in composite materials is still an underexplored area. This work proposes a novel workflow for FFT-based segmentation, periodic component detection and phase segmentation from raw high-resolution Transmission Electron Microscopy (TEM) images using a trained U-Net segmentation model. The developed model can expedite the detection of components and their phase segmentation, diminishing the temporal and cognitive demands associated with scrutinising an extensive array of TEM images, thereby mitigating the potential for human errors. This approach presents a novel and efficient image analysis approach with broad applicability beyond the battery field and holds potential for application in other related domains characterised by phase and composition distribution, such as alloy production.

KEYWORDS

deep learning, FFT, GUI tool, solid electrolyte interphase, transmission electron microscopy, U-Net

© 2025 The Author(s). Journal of Microscopy published by John Wiley & Sons Ltd on behalf of Royal Microscopical Society.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

1 | INTRODUCTION

Microsco

In the era of information, the emphasis has transitioned from data collection to data processing. Currently, an abundance of data is accessible across various research domains, and the primary hurdle lies in extracting meaningful information through its processing. This principle holds true within the realm of materials science, where researchers endeavour to derive valuable insights from their experimental samples, using high-resolution microscopy imaging.¹ The post-processing of these images has always been a challenge, owing to the complexity of the high-resolution image and the need for prior knowledge of components involved. Additionally, a comprehensive examination necessitates the acquisition of multiple images of the same sample, thereby augmenting the burden of post-processing.

In recent years, high-resolution imaging has become an indispensable tool in the field of battery material research, playing a vital role in the development of strategies to address the increasing energy demand. High-Resolution Transmission Electron Microscopy (HRTEM) stands out as a powerful technique used for investigating the microstructure of battery materials, encompassing cathodes, anodes, and electrolytes, with atomic resolution.² Among the various battery systems under investigation, Li metal batteries have received significant attention due to their propensity to store at least 33% more power per pound than traditional Li-ion batteries, rendering them suitable for a wide range of applications, including electric vehicles, renewable energy integration, and grid-scale energy storage.³ It is important to note that cycled Li metal, due to presence of oxides, carbonates, and sulfides, exhibits sensitivity to electron beam. Consequently, cryogenic Electron Microscopy (cryo-EM) technology has emerged, expanding the feasibility and necessity of high-resolution imaging in battery research, particularly for studying electron beam-sensitive anode materials like electrochemically cycled Li metal.^{4,5} Solid Electrolyte Interphase (SEI) formed in the anode materials after cycling is regarded as one of the least understood systems in the battery community.^{6,7} Studying the SEI is of scientific interest due to its crucial role in enabling longterm cycling in battery systems, as well as its complex composition of both organic and inorganic compounds. HRTEM, particularly when used in conjunction with cryogenic techniques, enables the elucidation of the components of the thin SEI layer, which typically has a thickness in nanometre scale.⁸ Such characterisation methods facilitate SEI engineering and the development of improved electrolyte systems.

In addition to the arduous task of sample preparation for HRTEM, processing the resulting images can also

present significant challenges due to the high resolution and large amount of data produced.⁹ Furthermore, in situ TEM studies targeting the dynamic interplay of properties, structures, and compositions within nanostructures yield substantial datasets acquired at elevated frame rates, posing an exceedingly formidable task for comprehensive data analysis.¹⁰ Recently, novel image analysis techniques have emerged within the framework of deep learning, a data processing approach that has gained tremendous popularity over the past decade. The proliferation of deep learning techniques within the field of image analysis has been spurred by several factors, such as the expanded availability of labelled datasets of significant size, notable advances in the realm of deep learning research, and the emergence of powerful high-performance frameworks like PyTorch and TensorFlow.^{11,12}

In the field of battery research, several studies have employed machine learning and deep learning techniques for image analysis.¹³⁻¹⁵ The image analysis investigations have largely focused on segmenting microscopy and tomography images of electrode materials.^{16,17} Automated analysis of high-resolution microscopy images for detection of phases and components in composite battery materials is an area that remains relatively unexplored. In contrast, there have been several notable works in materials science that have focused on detecting phases and components from HRTEM images. For instance, Liu et al. employed an unsupervised clustering algorithm combined with a scanning window technique to detect and group different phases from a TEM image of an AM Inconel 718 alloy.¹⁸ The technique can detect and map multiple phases, but there is a possibility of the same phase getting identified more than once and the chemical composition of the phases is not identified. While the method exhibits efficiency in phase mapping, its suitability for analysing battery materials is more limited compared to alloys. Zhang et al. proposed an improved Local Contrast Attention-UNet (LCA-UNet) for detecting Fast Fourier Transform (FFT) features from TEM images of Zirconium oxides.¹⁹ The authors utilised a scanning window technique to generate the FFT, then identified the features and labelled the corresponding phase window. The study used a 64×64 pixels region of the 256×256 -pixels window for detection, potentially losing information outside the chosen region. The employed method disregarded the inclusion of asymmetrical characteristics, potentially leading to the loss of pertinent information, and the designated window was a square region rather than accurately representing the intended area. Furthermore, the intricate composition of the cycled Li metal anode samples may lead to low-contrast images that are challenging to process. In fact, using smaller-scale 2-D FFTs may exacerbate this issue, even with advanced deep learning models such as LCA-UNet.



FIGURE 1 Schematic of the program flow: High volume of TEM data is acquired for analysis \rightarrow FFT features are detected using deep learning model \rightarrow Components identified from detected features using a material database developed with ICDD (International Centre for Diffraction Data) database \rightarrow Each feature is isolated and mapped using IFFT \rightarrow Component detection program is used for high volume TEM processing \rightarrow Intensity profiling and mapping of components obtained.

The workflow in this study for FFT-based segmentation, periodic component detection and phase segmentation from HRTEM images of battery materials is presented in Figure 1. First, a 2-D FFT is generated from the HRTEM image, and a trained U-Net segmentation model is used to detect features from the generated 2-D FFT cropped half (1024 \times 512 pixels). Exploiting the innate symmetry present in the 2-D FFTs, half cropped 2-D FFTs (1024×512 pixels) were used for model training. The half cropped 2-D FFTs performed better than the full 2-D FFTs (1024×1024 pixels). The d-spacing of the identified features from model generated 2-D FFT is calculated and compared with a material database to detect the components present in the TEM image. Finally, the inverse FFT (IFFT) image is generated for each component by isolating the features based on detected components, and watershed segmentation is used to map the periodic components present in the TEM image. For high-volume TEM processing, this process is repeated for all the TEM images to be analysed, and the intensity profile for each component is generated to observe their evolution over the imaging period. The developed segmentation model can expedite component detection and phase segmentation, not only in the battery field but also in alloy manufacturing. The qualitative and quantitative analysis of TEM images in this work decreases the time and effort necessary for evaluating substantial

quantities of TEM images while mitigating the potential for human error caused by cognitive task-induced state fatigue²⁰ of analysing a large set of TEM images (e.g. \sim 100 images).

2 | MATERIALS AND METHODS

2.1 | TEM sample preparation

LiF powder was ground in an Ar gas glove box and the ground LiF powder was dispersed on the TEM Cu grid. The TEM Cu grid is loaded to Melbuild vacuum transfer holder under Ar atmosphere and the sample is transferred into the TEM column without any air exposure. The experiment is performed at room temperature for beam damage analysis. For cryo temperature measurement, liquid nitrogen was added to the dewar when the holder is fully inserted into the column. The system stabilised at \sim -165°C for cryo temperature beam damage analysis.

2.2 | TEM imaging

We used a Thermo Fisher Talos F200X TEM electron microscope system with super-low-dose TEM techniques

to characterise samples. The low dose HRTEM images were acquired with controllable electron dose rate (50–1000 e A^{-2} s⁻¹) at FEI Ceta 16 M camera and low dose system. The pixel size of the TEM image corresponds to 0.037 nm/pixel.

2.3 | FFT generation from TEM images

TEM images can contain complex structures with multiple crystallographic and amorphous phases. Therefore, extracting and interpreting statistical information and uncovering the underlying physical mechanisms can pose significant challenges. FFT is useful for identifying periodic patterns in TEM images, such as lattice fringes, and for extracting information about the crystallographic orientation and symmetry of the sample.²¹ For this study, the Cryo-EM images of cycled Li metal anode in different electrolytes are used to introduce variability and to reduce redundancy. Images of magnification >300 kx are chosen for getting SEI layer of the anode in better resolution. The Cryo-EM images of format .dm3/.dm4 generated by GATAN micrograph software are used specifically. The TEM image is generated from the .dm3/.dm4 files using the DM3lib python library²² and the 2-D FFT is generated from the TEM image using the FFT function from the OpenCV2 library in Python. The mathematical representation of the 2D Discrete Fourier Transform (DFT) for an image signal f(x,y), across an x-y plane is:

$$F(u,v) = \frac{1}{M \times N} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) e^{-2\pi i \left(\frac{ux}{M} + \frac{vy}{N}\right)}, \quad (1)$$

where F(u,v) is the Discrete Fourier Transform of f(x,v), u and v are the spatial frequencies in the x and v directions, respectively, and M, N are their associated sizes. The sum is taken over the entire x-y plane. In the field of image processing, the Fourier transform is commonly employed to analyse and manipulate the frequency components of an image. To achieve a more visually meaningful representation, the zero-frequency component is shifted to the centre of the transformed array, and the resulting complex number is converted into an absolute value. A logarithmic operation is then applied to the absolute value for perceptual scaling.²³ This transformed matrix is subsequently converted into an intensity image containing only the normalised logarithm of the absolute values of the complex numbers, wherein the pixel values range from 0 (representing black) to 1 (representing white). The resulting image is then cropped from the centre to generate a reduced 2-D FFT with a size of 2048×2048 pixels.

2.4 | FFT preprocessing

GATAN²⁴ software-generated FFTs are used as a reference for pre-processing training 2-D FFTs, as their manual display adjustments allow for effective noise reduction and feature enhancement. However, manual adjustment is not feasible for large datasets. Therefore, an automated Python-based denoising method was developed.

Analysis of GATAN's display controls (Figure S2a) reveals that suppressing pixel intensities below the histogram's mean value (marked by dashed red line) significantly improves feature visibility. This observation inspired the Python denoising process. First, the mean pixel value of the noisy FFT generated using Python (Figure S2b) is calculated. This mean is then subtracted from every pixel in the image and the minimum value is clipped to 0, effectively removing negative values. This process, similar to GATAN's manual adjustment, reduces the background noise but has poor feature contrast (Figure S2c).

To further improve visual similarity with the GATAN FFTs, the brightness and contrast of the denoised image are adjusted. This step involves linear scaling using the OpenCV library's convertScaleAbs function. Scaling parameters (alpha = 3 for contrast, beta = 11 for brightness) that maximise similarity with the GATAN output were empirically determined (Figure S2d).

To quantify the similarity between the processed FFTs and the reference GATAN FFT, the Structural Similarity Index Measure (SSIM) metric is used. SSIM,²⁵ implemented using the scikit-image library in Python, is chosen for its ability to assess perceived similarity, which aligns with the goal of generating FFTs that are visually similar to the human-adjusted GATAN outputs. SSIM considers luminance, contrast, and structure, making it more suitable than pixel-based metrics like Mean Squared Error (MSE) which are less sensitive to structural changes²⁶ (Figure S2).

A positive correlation between SSIM scores and the performance metrics of models trained on different FFT datasets (noisy, denoised, enhanced) was observed. This correlation, detailed in the ablation study section (Table S5), highlights the importance of effective FFT denoising for effective model training.

2.5 | Component detection from FFT—radial profiling

Radial profiling is a widely used and straightforward approach for analysing 2-D FFTs.²⁷ To perform radial profiling, the distance of each pixel from the centre point is first calculated using the Euclidean distance formula:

$$r = \sqrt{(x - x_0)^2 + (y - y_0)^2}.$$
 (2)

Here, *r* is the radial distance, (x, y) are coordinates of the pixel and (x_0, y_0) are the coordinates of the centre of the 2-D FFT. Then, the pixels that are equidistant from the centre are binned together, and the average intensity within each bin is computed. This results in a one-dimensional profile representing the average intensity as a function of radial distance.

The resulting integrated values are plotted against the real distance of each circular band, which is determined by calculating the reciprocal distance from the centre of the 2-D FFT (Figure S3g–i). To convert the reciprocal distance to real distance, the pixel size information of the TEM image file (.dm3/.dm4) is extracted using DM3lib and applied to the following formula:

$$real\ distance = \frac{pixelsize \times N}{r}.$$
 (3)

Here, *pixelsize* is the calibration information obtained from the metadata of the TEM file that indicates the real size of a pixel in the TEM image, N is the corresponding size of the TEM image, r is the radial distance of the pixel in the FFT calculated from Equation [2]. The periodic components present in the TEM image can be identified by the peaks obtained from the diffraction-like graph generated using the circular integration method. The position of the peak corresponds to the d-spacing of the periodic components, aiding in the identification of the components present in the image. However, the analysis of the noisy 2-D FFTs using the circular integration method results in a diffraction-like graph that is also noisy, making it difficult to detect the peaks accurately (Figure S3g–j).

2.6 | Automated FFT training set generation

The preprocessing of 2-D FFTs generated from HRTEM images (detailed in previous section) for the purpose of training a segmentation model involves the use of Gaussian filters, which is a standard image processing technique available in the widely used OpenCV python package.²⁸ The Gaussian filter was selected thanks to its isotropic properties which, when applied uniformly across the image, maintain the relative intensities and positions of symmetrically arranged diffraction spots while effectively suppressing high-frequency noise—characteristics particularly important for 2-D FFTs. 2-D FFTs were chosen for filter application rather than raw TEM images

Microscopy

because the feature detection algorithm operates in the frequency domain, where periodic features can be isolated and background noise can be reduced through Gaussian filtering. Additionally, direct application of Gaussian blur on TEM images could potentially smudge or blur the sharp lattice fringes, which would result in significant loss of critical crystallographic information. By applying the filtering in the frequency domain instead, the integrity of the original lattice structures can be preserved while still improving signal quality for subsequent analysis.

A comprehensive parametric study was conducted comparing different Gaussian kernel sizes ((1,1), (3,3), and (5,5)) against unfiltered images ('No Blur') across various feature sizes (7–40 pixels). Each filter configuration was quantitatively evaluated using Signal-to-Noise Ratio (SNR) and Contrast-to-Noise Ratio (CNR) metrics (Figure S4). SNR and CNR were calculated using binary masks derived from labelled images to identify signal (feature—labelled white) and noise (background—labelled black) regions. The SNR was computed as 20 times the logarithm (base 10) of the ratio between the mean signal intensity (μ_{signal}) and the standard deviation of the noise (σ_{noise}), expressed in decibels (dB).

$$SNR = 20 \times \log_{10} \left(\frac{\mu_{signal}}{\sigma_{noise}} \right). \tag{4}$$

The CNR was determined by calculating the difference between the mean signal intensity (μ_{signal}) and mean noise intensity (μ_{noise}), divided by the standard deviation of the noise (σ_{noise}).

$$CNR = \frac{\mu_{signal} - \mu_{noise}}{\sigma_{noise}}.$$
 (5)

These metrics quantitatively assess the quality of feature detection, where higher SNR values indicate better signal clarity and higher CNR values indicate better feature distinguishability from the background noise.

Superior performance was demonstrated by the Gaussian filter with kernel size (3,3), with the highest average SNR (21.03 \pm 1.92 dB) and CNR (7.83 \pm 3.35) values being yielded across all object sizes. Best enhancement for smaller features (7–22 px) was provided by the (3,3) kernel, while slightly better results for larger features (31–40 px) were shown with (5,5). However, since features across all size ranges are contained in the dataset, (3,3) was selected as the best-suited kernel size (Figure S5).

In contrast, the production of labelled data involves manual input from experts in the field who utilise Adobe Photoshop software to identify and mark features in the 2-D FFTs.

To generate a training set for the proposed deep learningbased framework for FFT-based segmentation and analysis of TEM images, a total of 80 reduced 2-D FFTs $(2048 \times 2048 \text{ pixels})$, collected during multiple experiments, were utilised. To standardise the dataset, the 2-D FFTs were resized through interpolation (resized through interpolation cv2.INTER_AREA²⁹) to a uniform size of 1024×1024 pixels. Resizing was performed in the FFT domain, rather than on the original TEM images, to preserve the full range of spatial frequencies captured in the original high-resolution images. Resizing the TEM images before FFT computation could have led to the loss of sharp features crucial for detecting fine crystallographic details. By resizing in the FFT domain, we retain the original spatial information while controlling the frequency range used for training. Furthermore, the ground truth labels for the 2-D FFTs were generated at the full resolution of 2048×2048 pixels and then resized to 1024×1024 pixels to match the input images. This ensures that no information is lost in the labels during the resizing process, maintaining the accuracy and precision of the training data. This approach also reduces computational demands and memory requirements, which were essential given the limitations of our hardware.

RMS

To increase the diversity of the dataset, the 80 2-D FFTs were augmented by applying affine transformations and image rotations, resulting in a total of 1986 images. The affine transformations used on the images include: Rotation—Randomly rotated slightly by setting a range of $\pm 0.2^{\circ}$, Horizontal Translation—Randomly shifted by up to 5% of the image's width in either the left or right direction, Shearing-Slanting the shape of an image, towards the left or right direction by 0.05°, and Zooming-Randomly zoomed in or out by up to 5%. Data augmentation was applied to the 1024×1024 pixel images, which were then cropped to 1024×512 pixels to create our primary training dataset (Figure S1). To investigate the impact of both dataset size and image sizes on model accuracy, a similar data augmentation technique was applied to generate four distinct datasets. Two datasets consisted of 1986 images each, one with a size of 256×256 pixels (resized through interpolation cv2.INTER_AREA) and the other with 512×512 pixels (resized through interpolation cv2.INTER_AREA). Additionally, two smaller datasets of 500 images each were generated with the same respective image sizes using data augmentation. While all these datasets were used for model evaluation as discussed in the Results section, our final optimised model was trained using the 1024×512 pixel dataset. This approach ensures that the deep learning model is trained on a diverse range of images and FFT-based segmentation is more accurate.

2.7 | Programming and training machine learning models

The study employed Python for all programming activities, with the Keras framework of Tensorflow library³⁰ utilised for machine learning. The final dataset consisted of 1986 images with dimensions of 1024×512 pixels, randomly divided into training (90%, or 1771 images) and validation (10%, or 197 images) sets. While we utilised the reduced size images (256×256 and 512×512 pixels) for comparative analysis as shown in Figure S7, these were not used for our final proposed model as they resulted in lower performance metrics (Table S1). The UNet-type architecture used in our study comprised four levels in both contracting and expansive paths (with the fifth level serving as the bottleneck), with each level containing two convolutional layers followed by either max-pooling (in the contracting path) or up-sampling (in the expansive path) (Figure S6). The models were trained for 100 epochs on Dell workstations with Nvidia RTX A4000, 16GB, 4DP GPUs, with each model taking approximately 10-11 h to train. The training period was deemed sufficient for experimentation with network architecture, data preprocessing, and hyperparameter tuning. The stability of the models was ensured by tracking loss as a function of epoch number, indicating general convergence. Finally, the binary segmentation map, which classified individual pixels as particle or background, was obtained by thresholding predicted softmax output for each pixel.

2.8 | Hyperparameter tuning for U-Net model

Hyper parameter tuning was carried out using Bayesian Optimisation via the Keras Tuner Python Library. Bayesian Optimisation uses probabilistic models to guide the hyperparameter search process. It is particularly useful for black box functions and problems where the objective function takes a long time to evaluate.

Binary Cross-Entropy (BCE) function was employed as the loss function which is effective in handling imbalanced class distribution as is our case, where the region of FFT features are very small compared to the background; The formula for BCE is given by:

$$BCE \ Loss = -[y \times \log(p) + (1 - y) \times \log(1 - p)], \quad (6)$$

where *y* is the true label (0 or 1), *p* is the predicted probability of the positive class and log is natural logarithm.

Metrics used for the evaluation of neural network's segmentation performance are: the Dice coefficient,

Intersection over Union (IoU), Pixel Accuracy, and Recall. To define the metrics, we use several key terms based on the ground truth set G and the predicted set P. True Positives (TP) represent the intersection of G and P, signifying pixels correctly identified as belonging to the target object. Conversely, False Positives (FP) denote pixels incorrectly predicted as positive, meaning they are present in P but absent in G. True Negatives (TN.) indicate pixels correctly identified as not belonging to the object, while False Negatives (FN) represent pixels that were actually part of the object in G but were missed by the prediction and therefore not included in P.

The metrics are then formulated as follows:

$$Dice = \frac{2 \times |TP|}{(|TP| + |FP| + |FN|)},\tag{7}$$

where, || denotes the cardinality of the set.

$$IoU = \frac{|TP|}{(|TP| + |FP| + |FN|)},$$
 (8)

$$Pixel Accuracy = \frac{|TP| + |TN|}{(|TP| + |TN| + |FP| + |FN|)}, \quad (9)$$

$$Recall = \frac{|TP|}{(|TP| + |FN|)}.$$
 (10)

In this work, the objective for the optimisation was minimising the validation loss and 15 trials were allowed with 100 epochs per trial. A callback was added to the Tuner to ensure that if the model's validation loss did not continue to increase after 3 epochs, the training is terminated to save experimentation time.

In each trial, the optimisation occurred over 6 parameters. These include the learning rate, convolutional kernel size, number of convolutional filters, convolutional transpose kernel size, activation function and 3 drop out values treated separately. The learning rate was set to vary between three values which were 1e-4, 1e-5 and 1e-6. The convolutional kernel size was set to vary between 3 and 5. The convolutional transpose kernel size was set to vary between 2 and 3. The number of convolutional filters in the first layer was set to vary between 32 and 128, with subsequent layers doubling this number as per the standard U-Net architecture. In our final implementation, we used 128 filters in the first level of the encoder, doubling at each level to reach 2048 filters at the bottleneck level. The activation function was set to vary between ReLu, eLu and GeLu. All 3 dropout values were set to vary between 0 and 0.5 with steps of 0.1.

3 | RESULTS AND DISCUSSIONS

Micro

3.1 | FFT feature detection using deep learning—high-resolution image segmentation

Deep learning technique was utilised to extract periodic features from the 2-D FFTs, effectively reducing the impact of non-periodic noise contributions in the diffraction-like graph obtained by radial profiling. Semantic segmentation is an essential technique employed in machine learning for the identification and classification of individual pixels in an image into distinct classes or entities.³¹ By mimick-ing the way in which humans perceive and analyse visual data, semantic segmentation is a fundamental step in enabling machines to comprehend and interpret micrography images accurately. In the context of micrography, semantic segmentation holds significant potential in distinguishing between different cell types or structures at the pixel level, thereby aiding in various fields, including medical diagnosis, biological research, and particle sizing.³²

U-Net architecture is a widely used deep learning technique for semantic segmentation. U-Net is composed of two main pathways, namely, the contracting and expansive paths. The contracting path focuses on capturing the context of the input image, whereas the expansive path enables precise localisation by up sampling the feature maps.³³ The U-Net segmentation model is employed to detect features in the 2-D FFT (Figure S6) where very few annotated images are used. A large set of training data and corresponding ground truth are generated using data augmentation as detailed in previous sections. To optimise the training set conditions and model parameters, four factors are taken into consideration, namely, the size of the training set, size of the training image, the number of filters, and the threshold of the segmented image.

To assess the influence of training dataset size and image size on segmentation accuracy, a series of experiments were conducted. The model was trained using varying quantities of images and at multiple sizes, specifically comparing performance across two distinct dataset sizes and three image sizes.

A clear correlation between the number of training images and the model's performance was revealed. As depicted in Figure S7, a noticeable improvement in the precision and clarity of feature detection was observed when the dataset size was increased.³⁴ These findings are further supported by the quantitative metrics presented in Table S1.

Furthermore, the impact of image size was examined by maintaining a fixed dataset size and varying the pixel dimensions of the training images. The results, again

Microscopy Internal of Inte

illustrated in Figure S7 and quantified in Table S1, demonstrated that finer details were facilitated and the issue of missing features observed at smaller image sizes was mitigated when larger image sizes were used. Notably, a consistent enhancement in feature identification was shown with the progression from the smallest to the largest image size. While improved segmentation accuracy is offered by larger

In semantic segmentation, convolutional layers with a higher number of filters can extract more complex patterns from image data effectively.³⁵ Therefore, this study evaluates the effect of increasing filters on the semantic segmentation task. Based on our assessment, it was determined that 128 filters in the first layer yielded better overall results for segmentation (Figure S8; Table S2).

Since a sigmoid activation function is used in the output layer of the U-Net model for binary segmentation, probability maps with values between 0 and 1 for each pixel are produced. A threshold must be applied to convert these probabilities into binary segmentation masks (where pixel values > threshold are classified as foreground, and values \leq threshold as background). After the U-Net model trained on 1024 × 512 pixels images was finalised, the best-performing threshold for converting the model's sigmoid output probabilities into binary segmentation masks was determined. Threshold values ranging from 0.0 to 1.0 were systematically evaluated using multiple segmentation quality metrics and 0.90 was selected as the best-performing threshold (Figures S9 and S10).

The U-Net model, as described previously, was utilised to detect features in the 2-D FFT (Figure 2A) and generate a corresponding labelled image with white markings indicating the detected features. The labelled 2-D FFT was subsequently masked onto the original 2-D FFT (Figure 2B), and circular integration (Figure 2C) was applied to generate the diffraction-like graph (Figure 2D). An example of this process is illustrated in Figure 2, where the 2-D FFT obtained from a TEM image of a Li metal anode was processed to obtain the diffraction-like graph. The Li (011) (2.42 Å) and Li_2O (111) (2.65 Å) components were detected using the peak positions and the user-defined database created using ICDD database.³⁶ The segmentation model, U-Net, was trained using 1024×1024 pixels images. However, the model suffered from poor symmetry in some cases of the resultant FFT (Figure 2E), as demonstrated in Figure 2F with regions marked in red dots, and imprecise segmentation due to the detected features being merged, as shown in regions marked in blue dots. Moreover, the peak detection of crystalline components using circular integration (Figure 2G) sometimes failed at very low intensity, as shown in Figure 2H, where the $Li_2O(022)$ component was not detected due to the low intensity of the feature. Figure 2H shows a case where peak detection

failed for low-intensity features, particularly for the Li_2O (022) component. This failure is primarily due to the missing detected features on the symmetrical side of the Li_2O (022) component (marked in red). Since the radial profiling is done for set of pixels at a fixed radial distance from the centre, the missing features affect the total intensity of the peak obtained.

3.2 | FFT half image training technique

To address the imprecise segmentation, the symmetry of the 2-D FFT was utilised by dividing the image into two halves, using only half of the image for training, and reducing the training image size by half to 1024×512 pixels. The mathematical basis for this lies in the properties of the DFT for real-valued signals. From Equation [1], if f(x,y) is a real-valued signal, then its DFT F(u,v) satisfies the condition $F(-u,-v) = F^*(u,v)$, where $F^*(u,v)$ is the complex conjugate of F(u,v).

The DFT of real-valued signals exhibits Hermitian symmetry, meaning that the negative frequency components are the complex conjugate of the positive frequency components. Therefore, one half of the FFT domain contains all the unique information present in the full FFT.³⁷ By using only half the domain, we effectively eliminate redundant information without losing any essential data. Therefore, U-Net's performance would not be significantly affected by choosing the opposite half. Using half of the FFT reduces the input size to the U-Net by 50% which not only reduces the computational load but increases the accuracy of the model (Table S3).

To obtain a complete 2-D FFT, half of the 2-D FFT is first duplicated and then horizontally mirrored. The mirrored image is then mirrored vertically (Figures 3A and S11). Flipping operation is carried out using cv2.flip function.³⁸ The challenge in processing the 2-D FFT (Figure 3B) using model trained with full 2-D FFTs (1024 \times 1024 pixels) was overcome by employing a model trained on half 2-D FFTs (1024 \times 512 pixels) for feature detection, as shown in Figure 3C. Following the initial segmentation process, individual regions were identified and characterised using connected component labelling. The cv2.connectedComponents³⁹ function (OpenCV library) was applied to the resulting binary image to assign unique labels to each contiguous region (Figures S14b and 3D). Subsequently, region properties, including centroid, major axis length, and minor axis length, were extracted using the skimage.measure.regionprops_table⁴⁰ function (scikit-image library), which were then used to obtain the intensity profile after masking (Figure 3E). The centroid was used as the centre of each extracted feature,



FIGURE 2 Image analysis of (A) computer generated FFT of TEM image of the SEI in a cycled Li metal anode: (B) detected FFT features using U-Net model trained with full (1024×1024 pixels) images and (C) circular integration technique on segmented image to obtain the (D) diffraction-like graph of the 2-D FFT. Image analysis of (E) another computer generated FFT of TEM image of the SEI in a cycled Li metal anode: (F) model trained with full (1024×1024 pixels) images suffer from improper symmetry in segmentation (marked in dashed red) and merged features (marked in dashed purple). (G) Circular integration technique on segmented image (H) failed peak detection in diffraction-like graph for low intensity features.



FIGURE 3 (A) Schematic of the usage of half 2-D FFTs for feature detection. Image analysis of (B) computer generated FFT of TEM image of the SEI in a cycled Li metal anode. (C) Model trained with half (1024 × 512 pixels) images preserves symmetry and produces clear distinct features. (D) Individual features detected from the segmented image with mask for IFFT. (E) Feature intensity graph generated from extracted feature properties.

Components	d-spacing (Å) (calculated)	d-spacing (Å) (from GATAN)	% Match of program calculated d-spacing with GATAN calculated	Feature size (pixels) (IFFT mask)	Pixel value count (a.u.)
Li (011)	2.41641	2.416	100%	20	28078
Li ₂ O (111)	2.62746	2.6528	99.04%	17	287173
Li ₂ O (022)	1.59336	1.593	100%	11	4601

TABLE 1 Tabulated results from the instance segmented image.

and the size of each feature was calculated as:

Size of the feature =

$$\sqrt{\left(\left[Major\ axis\ length\right]^{2} + \left[Minor\ axis\ length\right]^{2}\right)}.$$
(11)

Although the region properties are initially calculated as floating-point values, they are converted to integers for subsequent processing.

The components detected from the instance segmentation provide not only the exact position of the components detected, which agrees closely with the values generated from the GATAN Micrograph software. This size information can be used to design masks of the appropriate size to prevent noise from the FFT from being included in the mask, and a more accurate and discrete intensity profile was obtained after masking. Table 1 presents a comparison between our automated measurements and those obtained manually using GATAN software, a standard in the field. This comparison validated our method's accuracy, as demonstrated by a >99% d-spacing match, and confirmed its reproducibility across varied FFT generation methods. Through this validation, the reliability of our automated workflow was established, positioning it as an alternative to time-consuming manual analysis.

The features detected by the model trained with full size (1024×1024 pixels) images tend to produce features at least 30% bigger than the model trained with half-size images (Figure 4A–C). The intensity profile shows the robustness of the new model trained with half images (Figure 4D–F). The size of the segmented feature is important as a larger sized feature makes the distinct features more prone to merging, resulting in inaccurate IFFT generation from the masking. To quantify the closeness in shape and size between the segmented feature and its ground truth, we utilise the Hausdorff distance.⁴¹ This measure determines the maximum mismatch between two sets of points, representing the boundaries of the segmented feature (S) and its corresponding ground truth (G). The Hausdorff distance (H(S, G)) is defined as:

$$H(S,G) = max\{h(S,G), h(G,S)\},$$
(12)

where

$$h(S,G) = \max_{s \in S} \min_{g \in G} ||s - g||,$$
 (13)

$$h(G,S) = \max_{g \in G} \min_{s \in S} ||g - s||, \qquad (14)$$

where *s* is a point in the set *S*, *g* is a point in the set *G* and ||s - g|| is the Euclidean distance between the points *s* and *g*. h(S, G) calculates the maximum distance from a point in the segmentation to its closest point in the ground truth. h(G, S) calculates the maximum distance from a point in the ground truth to its closest point in the segmentation. The Hausdorff distance H(S, G) takes the maximum of these two values, representing the worst-case mismatch between the two sets. A lower Hausdorff distance value signifies a greater conformity between the segmented feature and it's corresponding ground truth. In the ideal scenario of perfect segmentation, where the segmented features and ground truth are indistinguishable, the Hausdorff distance would be 0.

U-Net model trained with 1024×512 pixels 2-D FFTs was benchmarked against: (1) U-Net model trained with 1024×1024 pixels 2-D FFTs, (2) CNN model trained with 1024×512 pixels images, and (3) K-Means clustering algorithm optimised for 1024×512 pixels images. Evaluation metrics employed were the Dice coefficient, IoU, pixel accuracy, precision, recall, and Hausdorff distance (Figure S12). Descriptions of the CNN architecture and the K-Means clustering method are provided in Section S1 of supplementary information. Superior performance across all metrics, including the Hausdorff distance, was observed for the U-Net model trained with 1024×512 pixels 2-D FFTs (Table S4).

An ablation study was also conducted to investigate the impact of 2-D FFT pre-processing on U-Net model performance (Figure S13). The performance of U-Net models trained with 2-D FFTs at various stages of preprocessing was compared. A positive correlation was observed between model performance and the similarity score of the pre-processed 2-D FFT to the GATANgenerated 2-D FFT (Table S5; Figure S3).

11



FIGURE 4 Intensity profiles for selected features in (A) computer generated FFT of TEM image of the SEI in a cycled Li metal anode. Line scans of three features compared with segmented images generated by (B) model trained with full (1024 × 1024 pixels) images and (C) model trained with half images (1024 × 512 pixels). (D-F) Intensity variation for each feature in the raw image (solid black line), model trained with full (1024×1024 pixels) images (dashed), and model trained with half images (1024×512 pixels) (starred-solid).

3.3 **Component detection and mapping**

Features detected in the 2-D FFT were matched to corresponding components in the database based on their d-spacing values. The component with the nearest dspacing value was assigned to each feature. Subsequently, features were grouped by component, and the d-spacing values and particle sizes within each group were averaged to yield a single representative value for each component (Figure S14c). The markers corresponding to the grouped components are used to reconstruct 2-D FFTs with isolated features (Figure S14d,e).

IFFT (Inverse FFT) was done on the masked 2-D FFTs of isolated components to map the component distribution on the TEM image. The IFFT generated had regions of varied intensity and indistinct boundaries between regions of interest (Figure S15). Watershed segmentation⁴² was utilized to extract bright regions of interest from IFFT maps. The segmentation was done in three steps: (1) initial binary separation was performed using Otsu's automated thresholding with binary inversion, (2) noise was removed via morphological opening operations using a square kernel,

and (3) background regions were identified through dilation to ensure clear structure separation. Parameters of watershed segmentation was optimized through a systematic grid search across multiple parameter combinations, wherein kernel size $(3 \times 3, 5 \times 5, 7 \times 7)$, opening iteration count (0, 1, 2), and dilation iteration count (0, 1, 2)2) were varied. Each combination was evaluated against 15 manually annotated ground truth images using five quantitative metrics: Dice coefficient, IoU, pixel accuracy, precision, and recall. Parameters were selected based on the highest average rank across all metrics (Figure S16). Through experimental comprehensive evaluation, the best-performing configuration was found to be a 7×7 kernel size with 2 opening iterations and 0 dilation iterations. It was observed that configurations with low opening iterations (0) combined with high dilation iterations (2) produced poor segmentation results, as insufficient noise removal prior to dilation caused boundary expansion of noise artefacts, inappropriate region merging, and resulting in a completely dark image (Figure S17).

Given the robust model, this tool can be used for high throughput TEM image analysis. The detection of the periodic components and the mapping of the components can be achieved with the tool (Figure S18).

3.4 | LiF beam damage analysis

The investigation of the SEI in lithium-ion batteries requires the use of TEM for atomic resolution imaging. Electron beam and environmental factors are known to cause irradiation damage in the SEI. To avoid such damage, the lithium sample must be maintained at low temperatures, but even then, the dose rate and exposure time can impact the sample products and its rate.⁷ The present study aims to analyse the beam damage of lithium fluoride (LiF), which is a commonly occurring SEI component, using a high-volume TEM image processing tool. To this end, we obtained data on LiF beam damage from a study that investigated the effects of beam damage on the imaging of SEI components in a cycled lithium metal anode, while considering the dose rate and exposure time. In our study, we analyse the images of the LiF particles at cryogenic and room temperatures, and with various dose rates ranging from 50 e A^{-2} s⁻¹ to 1000 e A^{-2} s⁻¹ using the developed TEM processing methodology. The TEM images are obtained by the Mel-Build holder using an optimised workflow.

A video was recorded during TEM imaging to monitor the beam damage of LiF particles. The Velox⁴³ software was used to save the resulting video file in the mrc format, which contained 100 slices corresponding to 2.46 s of electron beam exposure each. The mrc python library⁴⁴ was utilised to extract image data from the file, but since it was unable to retrieve the pixel size information, this information was manually entered along with the mrc file for subsequent processing. Batch file processing was employed for the automatic analysis of the components present in the sample, which were mapped using feature extraction from the 2-D FFT. The intensity of each component was recorded and compared to its corresponding intensity in the subsequent slice to track the evolution of the components during beam damage (Figure 5A). This allows for effective assessment of the beam damage of LiF particles during TEM imaging. The analysis of the intensity distribution of the detected components over 100 frames of beam exposure of LiF (Figure 5A) reveals that the Li (011) phase diffuses out of the LiF first at 6th frame or after 14.76 s of beam exposure (Figure 5B). The Li (011) facet is the most commonly observed one during nucleation, despite the slightly lower surface energy of Li (002) facet which is observed at 22nd frame or after 54.12 s of beam exposure (Figure 5C). This is due to the lower thermodynamic overpotential required to obtain Li (011) facet, which overpowers the surface energy effect.⁴⁵ The inten-

sity of the Li (011) phase decreases and starts to increase midframe due to the appearance of underlying LiF particles. Li₂O (111) and Li₂CO₃ (002) phases are also observed, which could be attributed to the surrounding environment of the TEM chamber.⁴⁶ LiF (002) is dominant over LiF (111) phase because it has a lower surface energy and a higher binding energy.⁴⁷ LiF (111) is observed at the initial frames of the beam exposure (Figure S19) and fades away as Li (011) phase grows. There is no information provided in literature that directly answers whether using high dose rates during TEM can convert one phase with lower surface energy to one with higher surface energy. However, it is known that high dose rates can cause radiation damage to the sample, which can affect the crystal structure and properties of the material being studied. Therefore, it is possible that high dose rates during TEM could affect the surface energy of a material, but further research is needed to determine the exact mechanism and conditions under which this could occur. The consistent intensity of the (002) peak in LiF is observed to be a result of the presence of newly formed LiF particles on the bottom surface after the degradation of the topmost LiF layer due to beam damage. This observation is corroborated by the increase in the intensity of the Li (011) phase, which subsequently decreases but then increases again. The increase in the intensity of the Li (011) phase is indicative of the formation of Li (011) due to the degradation and resurfacing of LiF particles. The degradation of LiF into Li is observed to be very limited at low dose rate and at cryo temperature (Figure S19). A Python-based GUI tool has been designed and developed for future use and broader distribution. This intuitive interface offers a potential advantage over current state-of-the-art analysis programs, particularly when dealing with large datasets. The GUI aims to streamline the analysis workflow, making the advanced capabilities of the program widely accessible to a wider range of users and facilitating efficient processing of complex TEM data (Figure S20).

4 | CONCLUSIONS

This paper introduces a novel deep learning-based framework for identifying and evaluating the phases present in TEM images of the SEI of Li metal anodes. The framework utilises the FFT diffractograms of the TEM images to analyse the feature positions and determine the SEI components. The proposed method also incorporates techniques to handle high-resolution images and exploit the symmetry of the FFT diffractograms for better model performance. The effectiveness of the framework is further improved by introducing additional training data. Additionally, a comprehensive analysis of the TEM images



FIGURE 5 (A) Intensity distribution of detected components over 100 frames of LiF at electron dosage of 1000 e $A^{-2} s^{-1}$ and cryo temperature. TEM image slice and program generated IFFT mapping at the initial detection of (D) LiF (002)—after 2.46 s beam exposure, (B) Li (011)—after 14.76 s beam exposure, and (C) Li (002)—after 54.12 s of beam exposure.

through intensity profiling and component mapping provides valuable insights into the SEI component evolution during imaging.

In future studies, we propose utilising the diffraction pattern generated from high frame rate TEM image patches for more detailed and accurate analysis. This alternative approach is superior to the sliding window technique for generating FFT diffractograms from TEM images as the latter may lack clarity and clear features. Furthermore, our proposed deep learning model can be applied to analyse not only the SEI components of Li metal anodes, but also other composite systems involving periodic components (Na metal anodes, Si anodes, etc.) provided an appropriate database with the respective compounds is used. The developed workflow can be further improved by integrating it for operando detection and mapping of the phases, thereby enabling the analysis of component evolution during TEM imaging.

AUTHOR CONTRIBUTIONS

G.R., M.Z. and Y.S.M. conceived the ideas. B.H. and S.B. performed TEM imaging. G.R. and F.A developed the model for TEM processing. G.R, F.A. and W.W. worked on the development of GUI application. All authors discussed the results and commented on the manuscript.

ACKNOWLEDGEMENTS

This work is supported by the funding and collaboration agreement between UCSD and Thermo Fisher Scientific on Advanced Characterization of Energy Materials. TEM was performed at the San Diego Nanotechnology Infrastructure (SDNI), a member of the National Nanotechnology Coordinated Infrastructure, which is supported by the National Science Foundation (grant ECCS-1542148). The workstation to build the model was provided by the University of Chicago under YSM's startup package. F.A. acknowledges support from the Research Experience for Nigerian Engineering Undergraduates (RENEU) program at the University of Chicago.

CONFLICT OF INTEREST STATEMENT The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The datasets generated and/or analysed during the current study are available in the *TEMGUILESC* Github repository (https://github.com/ganyguru/TEMGUILESC).

CODE AVAILABILITY

The underlying code and training/validation datasets for this study is available in *TEMGUILESC* Github and can be accessed via this link (https://github.com/ganyguru/TEMGUILESC).

REFERENCES

- Hill, J., Mulholland, G., Persson, K., Seshadri, R., Wolverton, C., & Meredig, B. (2016). Materials science with large-scale data and informatics: Unlocking new opportunities. *MRS Bulletin*, *41*, 399–409. https://doi.org/10.1557/mrs.2016.93
- 2. Wang, H., Jang, Y., Huang, B., Sadoway, D. R., & Chiang, Y. (1999). TEM study of electrochemical cycling-induced damage and disorder in LiCoO2 cathodes for rechargeable lithium batteries. *Journal of The Electrochemical Society*, *146*, 473–480. https://doi.org/10.1149/1.1391631
- 3. Cheng, X.-B., Zhang, R., Zhao, C.-Z., & Zhang, Q. (2017). Toward safe lithium metal anode in rechargeable batteries: A review. *Chemical Reviews*, *117*, 10403–10473. https://doi.org/10.1021/acs. chemrev.7b00115

- Xu, K. (2014). Electrolytes and interphases in Li-ion batteries and beyond. *Chemical Reviews*, 114, 11503–11618. https://doi.org/ 10.1021/cr500003w
- Winter, M., Barnett, B., & Xu, K. (2018). Before Li ion batteries. *Chemical Reviews*, 118, 11433–11456. https://doi.org/10.1021/ acs.chemrev.8b00422
- Li, Y., Li, Y., Pei, A., Yan, K., Sun, Y., Wu, C.-L., Joubert, L.-M., Chin, R., Koh, A. L., Yu, Y., Perrino, J., Butz, B., Chu, S., & Cui, Y. (2017). Atomic structure of sensitive battery materials and interfaces revealed by cryo-electron microscopy. *Science*, 358, 506–510. https://doi.org/10.1126/science.aam6014
- Cheng, D., Lu, B., Raghavendran, G., Zhang, M., & Meng, Y. S. (2022). Leveraging cryogenic electron microscopy for advancing battery design. *Matter*, *5*, 26–42. https://doi.org/10.1016/j.matt. 2021.11.019
- Wang, X., Zhang, M., Alvarado, J., Wang, S., Sina, M., Lu, B., Bouwer, J., Xu, W., Xiao, J., Zhang, J.-G., Liu, J., & Meng, Y. S. (2017). New insights on the structure of electrochemically deposited lithium metal and its solid electrolyte interphases via cryogenic TEM. *Nano Letters*, *17*, 7606–7612. https://doi.org/10. 1021/acs.nanolett.7b03606
- Taheri, M. L., Stach, E. A., Arslan, I., Crozier, P. A., Kabius, B. C., LaGrange, T., Minor, A. M., Takeda, S., Tanase, M., Wagner, J. B., & Sharma, R. (2016). Current status and future directions for in situ transmission electron microscopy. *Ultramicroscopy*, *170*, 86–95. https://doi.org/10.1016/j.ultramic.2016.08. 007
- Tao, F. (Feng), & Salmeron, M. (2011). In situ studies of chemistry and structure of materials in reactive environments. *Science*, *331*, 171–174. https://doi.org/10.1126/science. 1197461
- Chen, X.-W., & Lin, X. (2014). Big data deep learning: Challenges and perspectives. *IEEE Access*, 2, 514–525. https://doi.org/ 10.1109/ACCESS.2014.2325029
- Badea, M. S., Felea, I. I., Florea, L. M., & Vertan, C. (2016). The use of deep learning in image segmentation, classification and detection. arXiv:1605.09612. https://doi.org/10.48550/arXiv.1605. 09612
- Ling, C. (2022). A review of the recent progress in battery informatics. *npj Computational Materials*, *8*, 33. https://doi.org/10. 1038/s41524-022-00713-x
- Lombardo, T., Duquesnoy, M., El-Bouysidy, H., Årén, F., Gallo-Bueno, A., Jørgensen, P. B., Bhowmik, A., Demortière, A., Ayerbe, E., Alcaide, F., Reynaud, M., Carrasco, J., Grimaud, A., Zhang, C., Vegge, T., Johansson, P., & Franco, A. A. (2022). Artificial intelligence applied to battery research: Hype or reality? *Chemical Reviews*, *122*, 10899–10969. https://doi.org/10.1021/acs. chemrev.1c00108
- Furat, O., Wang, M., Neumann, M., Petrich, L., Weber, M., Krill, C. E., & Schmidt, V. (2019). Machine learning techniques for the segmentation of tomographic image data of functional materials. *Frontiers in Materials*, 6, https://doi.org/10.3389/fmats.2019. 00145
- Yao, W., Chouchane, M., Li, W., Bai, S., Liu, Z., Li, L., Chen, A. X., Sayahpour, B., Shimizu, R., Raghavendran, G., Schroeder, M. A., Chen, Y.-T., Tan, D. H. S., Sreenarayanan, B., Waters, C. K., Sichler, A., Gould, B., Kountz, D. J., ... Meng, Y. S. (2023). A 5 V-class cobalt-free battery cathode with high loading enabled

RMS

15

by dry coating. *Energy & Environmental Science*, *16*, 1620–1630. https://doi.org/10.1039/D2EE03840D

- Zhang, M., Chouchane, M., Shojaee, S. A., Winiarski, B., Liu, Z., Li, L., Pelapur, R., Shodiev, A., Yao, W., Doux, J.-M., Wang, S., Li, Y., Liu, C., Lemmens, H., Franco, A. A., & Meng, Y. S. (2022). Coupling of multiscale imaging analysis and computational modeling for understanding thick cathode degradation mechanisms. *Joule*, *7*, 201–220. https://doi.org/10.2139/ ssrn.4241175
- Liu, S., Amin-Ahmadi, B., Liu, R., Zheng, Q., & Zhang, X. (2023). Automated phase segmentation and quantification of high-resolution TEM image for alloy design. *Materials Characterization*, *199*, 112779. https://doi.org/10.1016/j.matchar.2023. 112779
- Zhang, Q., Bai, R., Peng, B., Wang, Z., & Liu, Y. (2023). FFT pattern recognition of crystal HRTEM image with deep learning. *Micron*, *166*, 103402. https://doi.org/10.1016/j.micron.2022. 103402
- Behrens, M., Gube, M., Chaabene, H., Prieske, O., Zenon, A., Broscheid, K.-C., Schega, L., Husmann, F., & Weippert, M. (2023). Fatigue and human performance: An updated framework. *Sports Medicine*, *53*, 7–31. https://doi.org/10.1007/s40279-022-01748-2
- Jany, B. R., Janas, A., & Krok, F. (2020). Automatic microscopic image analysis by moving window local Fourier Transform and Machine Learning. *Micron*, *130*, 102800. https://doi.org/10.1016/ j.micron.2019.102800
- Pierre-Ivan Raynal. DM3 Python Reader. https://bitbucket.org/ piraynal/pydm3reader
- Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., & Garcia-Rodriguez, J. (2017). A review on deep learning techniques applied to semantic segmentation. arXiv:1704.06857. https://doi.org/10.48550/arXiv.1704.06 857
- 24. GATAN Digital Micrograph Software. https://www.gatan.com/ products/tem-analysis/digitalmicrograph-software
- 25. Structural similarity index (scikit-image). https://scikitimage.org/docs/0.24.x/auto_examples/transform/plot_ssim. html
- Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, *13*, 600–612. https://doi.org/10.1109/TIP.2003.819861
- Stępniowski, W. J., Michalska-Domańska, M., Norek, M., & Czujko, T. (2014). Fast Fourier transform based arrangement analysis of poorly organized alumina nanopores formed via self-organized anodization in chromic acid. *Materials Letters*, 117, 69–73. https://doi.org/10.1016/j.matlet.2013.11. 099
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Müller, A., Nothman, J., Louppe, G., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2012). Scikit-learn: Machine Learning in Python. *The Journal* of Machine Learning Research, 12, 2825–2830. http://jmlr.org/ papers/v12/pedregosal1a.html
- OpenCV—Geometric Image Transformations. https://docs. opencv.org/3.4/da/d54/group_imgproc_transform.html
- 30. Keras-Tensorflow. https://www.tensorflow.org/guide/keras

- Kim, J.-G., Yoo, S. J., Kim, C.-Y., & Jou, H.-T. (2014). Circular fast Fourier transform application: A useful script for fast Fourier transform data analysis of high-resolution transmission electron microscopy image. *Applied Microscopy*, 44, 138–143. https://doi. org/10.9729/AM.2014.44.4.138
- 32. Guo, Y., Liu, Y., Georgiou, T., & Lew, M. S. (2018). A review of semantic segmentation using deep neural networks. *International Journal of Multimedia Information Retrieval*, 7, 87–93. https://doi.org/10.1007/s13735-017-0141-z
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. arXiv:1505.04597. https://doi.org/10.48550/arXiv.1505. 04597
- 34. Caicedo, J. C., Roth, J., Goodman, A., Becker, T., Karhohs, K. W., Broisin, M., Molnar, C., McQuin, C., Singh, S., Theis, F. J., & Carpenter, A. E. (2019). Evaluation of deep learning strategies for nucleus segmentation in fluorescence images. *Cytometry Part A*, 95, 952–965. https://doi.org/10.1002/cyto.a.23863
- Ahmed, W. S., & Karim, A. A. A. (2020). The impact of filter size and number of filters on classification accuracy in CNN. In 2020 International Conference on Computer Science and Software Engineering (CSASE) (pp. 88–93). IEEE. https://doi.org/10.1109/ CSASE48920.2020.9142089
- American Society for Testing and Materials (ASTM). International Centre for Diffraction Data. https://www.icdd.com/
- Salehi, S. A., Amirfattahi, R., & Parhi, K. K. (2013). Pipelined architectures for real-valued FFT and Hermitian-Symmetric IFFT with real datapaths. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 60, 507–511. https://doi.org/10.1109/TCSII. 2013.2268411
- OpenCV—Operations on arrays. https://docs.opencv.org/4.x/ d2/de8/group_core_array.html
- OpenCV—Structural Analysis and Shape Descriptors. https:// docs.opencv.org/3.4/d3/dc0/group_imgproc_shape.html
- skimage.measure—scikit-image. https://scikit-image.org/docs/ stable/api/skimage.measure.html
- Karimi, D., & Salcudean, S. E. (2019). Reducing the Hausdorff distance in medical image segmentation with convolutional neural networks. arXiv:1904.10030. https://doi.org/10. 48550/arXiv.1904.10030
- OpenCV—Image Segmentation with Watershed Algorithm. https://docs.opencv.org/4.x/d3/db4/tutorial_py_watershed. html
- Thermo ScientificTM VeloxTM. https://www.thermofisher.com/ order/catalog/product/VELOX
- PyPi.org. mrcfile 1.4.3 Library. https://pypi.org/project/mrcfile/, https://doi.org/10.1038/s41467-020-19206-w
- Nagy, K. S., Kazemiabnavi, S., Thornton, K., & Siegel, D. J. (2019). Thermodynamic overpotentials and nucleation rates for electrodeposition on metal anodes. *ACS Applied Materials & Interfaces*, *11*, 7954–7964. https://doi.org/10.1021/acsami. 8b19787
- 46. Liang, C., Zhang, X., Xia, S., Wang, Z., Wu, J., Yuan, B., Luo, X., Liu, W., Liu, W., & Yu, Y. (2020). Unravelling the roomtemperature atomic structure and growth kinetics of lithium metal. *Nature Communications*, *11*, 5367. https://doi.org/10.1038/ s41467-020-19206-w
- 47. Chattopadhyay, S., Lipson, A. L., Karmel, H. J., Emery, J. D., Fister, T. T., Fenter, P. A., Hersam, M. C., & Bedzyk, M. J. (2012).

Microscopy

In situ X-ray study of the solid electrolyte interphase (SEI) formation on graphene as a model li-ion battery anode. *Chemistry of Materials*, *24*, 3038–3043. https://doi.org/10.1021/cm301584r

RMS

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article. How to cite this article: Raghavendran, G., Han, B., Adekogbe, F., Bai, S., Lu, B., Wu, W., Zhang, M., & Meng, Y. S. (2025). Deep learning assisted high-resolution microscopy image processing for phase segmentation in functional composite materials. *Journal of Microscopy*, 1–16. https://doi.org/10.1111/jmi.13413