Materials & Design 260 (2025) 115200

Contents lists available at ScienceDirect

Materials & Design

journal homepage: www.elsevier.com/locate/matdes

Enhanced phase segmentation of scanning electron microscopy images using U-Net with energy dispersive X-ray spectroscopy-guided labelling

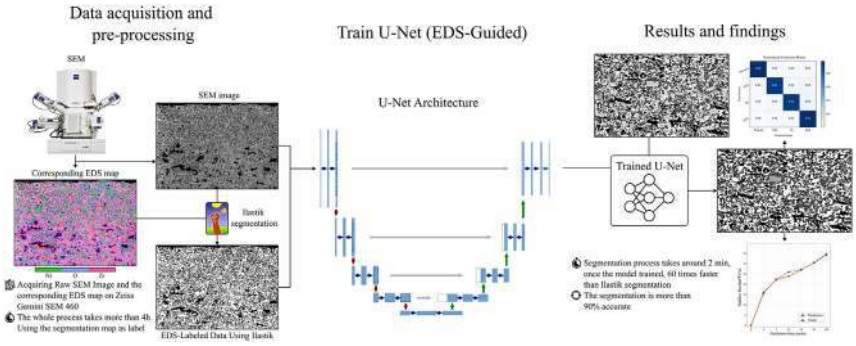
Kassem Dia [a], Thomas David [a,*], Laure Guetaz [a], Zineb Saghi [b], Geoffrey Daniel [c]

[a] Univ. Grenoble Alpes, CEA, LITEN, Grenoble, 38000, France
[b] Univ. Grenoble Alpes, CEA, LETI, Grenoble, 38000, France
[c] Univ. Paris-Saclay, CEA, Service de Genie Logiciel pour la Simulation, Gif-sur-Yvette, 91191, France

HIGHLIGHTS

- U-Net trained with EDS-guided labels segments SEM phases with high accuracy.
- Model achieves >90 % accuracy, eliminating the need for EDS during mappings.
- Deep learning analysis is 60 times faster than semi-manual segmentation.
- Model outperforms EDS-based labels, correcting segmentation misclassifications.
- U-Net demonstrates sensitivity to blur, zoom, and acquisition conditions.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Solid oxide cells
Semantic segmentation
Deep learning
Scanning electron microscopy images
Energy dispersive X-ray spectroscopy

ABSTRACT

This study introduces a deep learning-based methodology that uses U-Nets for the semantic segmentation of Scanning Electron Microscopy (SEM) images of Solid Oxide Cells (SOCs). The objective is the identification of the four key phases in hydrogen SOC electrode (porosity, yttria-stabilised zirconia (YSZ), nickel (Ni), and nickel oxide (NiO)). While the database used for training is built using Energy Dispersive X-ray Spectroscopy (EDS) maps to help label the SEM images using Random Forest classification, the model is trained to segment SEM images directly, without the need for EDS. The results show that the model achieves high segmentation accuracy on SEM images without any manual pre-labelling stepand demonstrates robustness to brightness variations, while challenges such as blur and zoom are addressed through sensitivity analysis. Additionally, a "sliding window" post-processing approach for segmenting large-scale images is used to reduce edge artifacts. The benefit of this approach is to provide an automatic workflow that does not require long EDS acquisitions or a time-consuming manual labelling procedure, offering a scalable solution for microstructural analysis of SOC electrodes.

1. Introduction

Solid oxide cells (SOCs) are high-temperature electrochemical devices that can operate either in fuel cell or electrolysis mode. They offer the advantage of producing electricity or hydrogen with high efficiency. However, durability remains the main issue preventing their widespread adoption. Therefore, it is essential to deepen the understanding of their degradation causes, which depend on their mode of operation. This involves studying how their microstructure evolves during operation. SOCs consist of two porous electrodes separated by a dense electrolyte,

* Corresponding author.
Email address: thomas.david3@cea.fr (T. David).

https://doi.org/10.1016/j.matdes.2025.115200
Received 26 August 2025; Received in revised form 6 November 2025; Accepted 21 November 2025
Available online 28 November 2025

which is typically yttrium-stabilised zirconia (YSZ), and its characterization is done through advanced imaging techniques combined with quantitative image analysis. Scanning electron microscopy (SEM) is generally used to obtain images at nanometer resolution of the morphology of the two electrodes and the distribution of the different phases. However, in order to correlate the observed microstructure degradation with the electrochemical performance of the electrode, it is necessary to extract quantitative data from the SEM images, such as porosity, phase volume fraction, and phase spatial distribution [1,2].

Dedicated image analysis approaches must be developed to automatically extract the desired information from images. However, SEM images sometimes feature very similar grayscale values, which makes it difficult to distinguish between multiple phases using direct threshold-based approaches. This is where Energy Dispersive X-ray Spectroscopy (EDS) mapping becomes valuable, as in this context it can provide significant advantage by locating the different elements, allowing distinction between phases with similar morphological features and imaging contrast.Therefore, it helps to solve the segmentation problem and to perform a quantitative study of the microstructure [3–5].

In particular, one of the challenges addressed in this study concerns the quantification of nickel oxidation (NiO) in the hydrogen electrode, which can occur during operation and may lead to critical failure. SEM images of the hydrogen electrode after nickel oxidation show four phases (porosity, NiO, Ni, and YSZ) that can be challenging to distinguish. This is especially due to the very similar grayscale intensity of YSZ and Ni. This issue and the solution using EDS mapping are illustrated in Fig. 1.

However, EDS use is limited by the time-intensive acquisition process, the requirement for specialized equipment, and insufficient spatial resolution for resolving fine microstructural details. Consequently, our primary objective is to achieve accurate segmentation of SEM images based solely on image data, without relying on additional chemical information such as EDS.

In this article, we present a novel approach in which a deep learning architecture dedicated to semantic segmentation—specifically, a U-Net—is trained to segment 2D cross-sectional SEM images of solid oxide cells (SOCs) without the need for auxiliary data. It is important to note that the proposed method could be extended to other segmentation tasks in materials science where phase discrimination is difficult due to limited contrast in imaging.

The structure of this paper is organized as follows: Section 2 provides a short review of related work. Section 3 describes the proposed methodology and its application to SEM image analysis. Section 4 presents the results and discusses their significance. Section 5 explores the robustness of the approach under varying acquisition conditions and highlights potential directions for future work. Finally, section 6 concludes the study and discusses its broader implications and perspectives.

2. Related work

Image analysis techniques are widely used in the literature for performing semantic segmentation on SEM images across various domains. They can be categorized into two major types: conventional image processing techniques using histogram-based segmentation, and more recent advanced machine learning techniques employing neural network architectures.

2.1. Conventional techniques

In the classical category, researchers have relied on thresholding methods that convert multilevel images into binary ones by dividing pixels based on their intensity levels. For example, Hu et al. [6] used an adaptive threshold segmentation algorithm (ATSA) to segment asphalt concrete images according to the image radius. Otsu thresholding has also been used for SEM image segmentation across different materials, including cement (Siddique et al. [7], Liu et al. [8]), porous polymers (Calkovsky et al. [9]), lithium-ion batteries (Almar et al. [10]), and nanoporous materials (Prill et al. [11]). This approach is ideal for simpler cases where there is a clear distinction between the foreground and the different phases in the images.

For more complex structures, others have relied on the more advanced watershed algorithm (Beucher et al. [12]), which treats the image as a topographic map, where brightness represents height, and finds the lines along the tops of the ridges. This approach is used across various research fields, including biological segmentation of neuron membranes in electron microscopy images of mouse cerebellum tissue (Liu et al. [8]), medical imaging for segmenting complex structures such as white and gray matter in brain images (Zanaty et al. [13]), geological engineering for segmenting pores, throats, and mineral grains (Zhou et al. [14]), and materials technology for segmenting SOC electrodes in X-ray computed tomography (CT) (Epting et al. [15]).

Although traditional segmentation techniques are effective and simple, they often lack accuracy in more complicated scenarios due to their sensitivity to noise, difficulty with varying illumination, inability to handle complex or gradual transitions, and inability to distinguish phases with similar SEM contrast.

2.2. Machine learning techniques

With significant advances in computer vision over the past two decades, machine learning algorithms have become widely used to

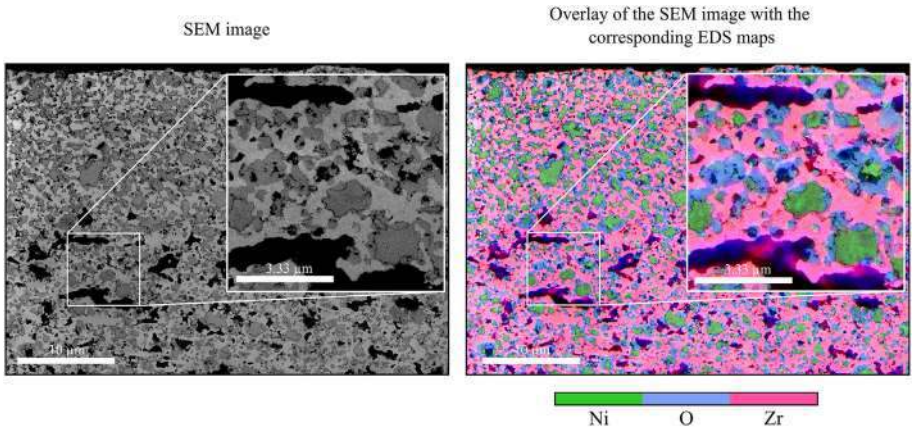


Fig. 1. Representative SEM image (left) after 20 min of oxidation time, and an overlay of the corresponding EDS maps (right), illustrating the spatial distributions of Nickel (green), Oxygen (blue), and Zirconium (pink) within the composite. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

address various image-based challenges, including object detection, image classification, and semantic segmentation. Algorithms like Support Vector Machines (SVMs), Random Forests, k-Nearest Neighbors (k-NN), and Conditional Random Fields (CRFs) have been applied to microstructure segmentation tasks (Li et al. [16], Wu et al. [17]). However, they often require a feature extraction step that tends to be time-consuming and can make it difficult to generalize the technique to new applications.

More recently, a specialized branch of machine learning, known as deep learning, has outperformed traditional algorithms, particularly in image segmentation tasks. Deep learning has found extensive applications in diverse fields such as agriculture (Zhang et al. [18], Bhagat et al. [19]), medical imaging (Rayed et al. [20], Bindhu et al. [21], Chen et al. [22]), and autonomous driving (Ghosh et al. [23]). This trend has caught the attention of the materials science community, where deep learning algorithms are also employed to solve segmentation problems and analyze microstructural components (Azimi et al. [24], Hirabayashi et al. [25]).

In particular, encoder-decoder-based architectures, such as U-Net, are frequently used for semantic segmentation. U-Nets are particularly popular for processing microscopy images, having initially been introduced for this purpose by Ronneberger et al. [26] on biomedical image segmentation. Since its introduction, U-Net has been applied in materials science to analyze for instance concrete SEM images (Bangaru et al. [27], Malik et al. [28]), 3D segmentation of electron microscopy images of ceramic materials (Hirabayashi et al. [25], Genc et al. [29]), and segmentation of graphene SEM images (Shah et al. [30]). U-Nets were also employed to study SOC SEM images, specifically for analyzing (Hwang et al. [31]), segmenting (Wang et al. [32]), and 3D segmentation of SOC Ni/YSZ anode composites (Sciazko et al. [33]).

The most widely used methods in the literature, and the focus of this study, are supervised deep learning algorithms because the type of predictions they provide and their interpretation are controlled through the training process. However, they require a labeled dataset to be trained, where most of the labelling work is manually done (Sciazko et al. [33], Hwang et al. [31]), which is time-consuming and may lack consistency because of observer variability (Zhang et al. [34]). To solve this problem, some researchers relied on EDS mapping to help in the quantitative study of the microstructure, such as correlating SEM contrast with chemical composition for accurate phase identification (Harald et al. [35]), and characterizing the chemical composition of nanoparticle hetero-aggregates (Mahr et al. [4]). EDS-guided segmentation has also been successfully applied using machine learning to multiphase materials, enabling more accurate distinction of chemically similar regions [36,37]. Juránek et al. [38] further proposed a graph-based deep learning approach that integrates SEM–BSE images with EDS data to enhance mineral phase segmentation. Such approaches provide an important complement to image-based segmentation, as they enable more robust quantification of microstructures and facilitate correlations with electrochemical performance, thereby enabling more accurate phase segmentation. Nevertheless, they still rely on the use of time-consuming EDS acquisition as input both for training and applying the machine learning model.

3. Methodology

The approach we propose here leverages high-quality cross-section SEM images and their corresponding chemical maps obtained through EDS mappings to produce reliable training datasets. The methodology is applied to microstructures containing four distinct phases: porosity, NiO, Ni, and YSZ. The training dataset, described in Section 3.1, consists of SEM images of multiple SOC samples along with their corresponding segmentation labels produced thanks to EDS data. The Deep Learning model, a U-Net architecture, and its training procedure, which we propose in this work, are described in Section 3.2. An overall schematic workflow of the proposed methodology is shown in Fig. 2.

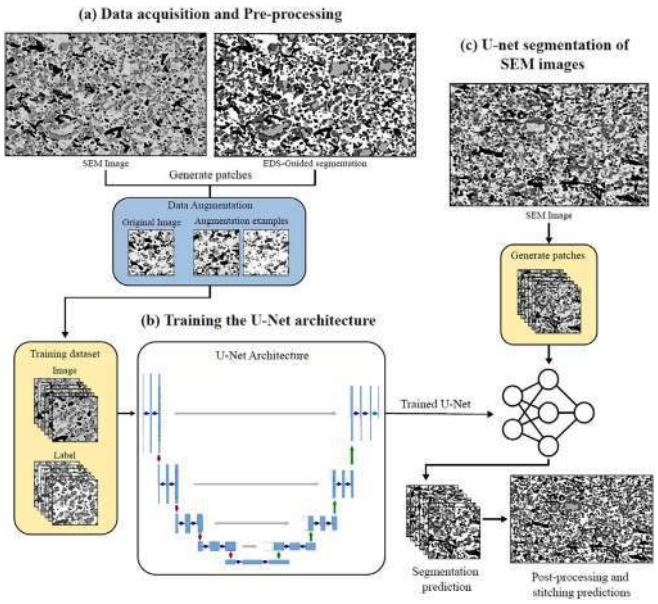


Fig. 2. Schematic outline of the main steps of the workflow. (a) Data Acquisition and pre-processing of experimental images and segmenting SEM images using EDS maps, creating patches, and applying image augmentation. The four levels of the grayscale segmentation images represent the four classes. (b) Training the U-Net model. (c) Use the trained U-Net to segment the SEM image without using its corresponding EDS map.

3.1. Data collecting

3.1.1. Data acquisition

The samples used in this study come from typical SOC hydrogen electrodes made of porous Ni/YSZ composite. The NiO phase appears in the samples after partial oxidation of Ni. The duration of this oxidation step varies for each of the samples, leading to varying NiO contents.

The samples were embedded in an epoxy resin before being polished and surface coated with a thin carbon layer to avoid charging effects. The SEM images and EDS mapping were done using a Zeiss Gemini SEM 460 equipped with an Oxford Ultim Max 100 EDS detector. Imaging was done at 1.1 kV with a pixel size of 0.04458 μm, while EDS maps were acquired at 10 kV. However, EDS was captured at a lower resolution than the SEM image. Specifically, the EDS resolution is one fourth that of the SEM image. Each SEM and EDS image takes respectively around 1.3 and 20 min to acquire. After acquisition, EDS maps were upsampled and precisely aligned on the images with a custom Python script using skimage library.

3.1.2. Pre-processing and data augmentation

Establishing a precise ground truth high-quality segmentation label is challenging. Nevertheless, a “target” label image can be generated with the help of EDS mappings. EDS provides elemental maps that facilitate the identification of different phases by highlighting the distribution of chemical components, as shown in Fig. 1. However, this labelling process is not straightforward, and performing it manually is extremely time-consuming. To address this problem, a software called Ilastik was used [39]. The SEM images and their corresponding EDS maps were provided as inputs to Ilastik to perform the target segmentation.

Ilastik’s Pixel Classification workflow uses a first manual labelling of several areas in the images and extends it to the whole image through feature extraction such as intensity, edges, and texture at multiple scales. A Random Forest classifier is trained on these features to predict the class probabilities for each pixel on the whole image, resulting in a detailed phase-segmented output. This output serves as the “target” segmentation label, providing a semi-automatic segmentation. This process

is faster than a fully manual labelling but still depends on human input for labelling a small part of the data. However, after the classifier is trained, it can still produce some errors and users may need to review the segmentation results and iterate with manual corrections to improve accuracy. This correction process is time-consuming and it does not guarantee to remove all of the errors. For a single image of 2856×4086 pixels, it takes about 4 hours to obtain a satisfying target segmentation. This value is based on an average estimation of the time required to segment all images in the dataset by microscopy specialists.

Now, the combined acquisition time for a single SEM image with its corresponding EDS map and the time required to generate its segmentation label using semi-manual segmentation approaches pose a significant limitation on the amount of data available for the training process. To address this limitation, we acquired 7 large SEM images of size 2856×4096 under the same conditions, then divided them into smaller patches of 256×256 pixels, resulting in a total of 1232 patches. The patch size was determined after a series of tests, as it effectively reduces the computational cost and memory requirements of the model. A total of 140 patches were randomly selected from various positions in the images and reserved for testing purposes and estimating the performances reported in Section 4. The remaining 1092 patches underwent data augmentation to further expand the training set, including rotations, flips, blurring, and variations in contrast and brightness. An example of augmented data is shown in Fig. 3. For each patch 10 augmented versions were generated, resulting in a total of $N = 10920$ images for training and validation. This combination of patching and augmentation exposes the model to variations in image characteristics, ensuring a better performance during the segmentation tasks.

3.2. Training the U-Net architecture

The second pillar of the approach relies on employing a deep learning model for semantic segmentation. A U-Net architecture is adopted [26], which consists of an encoder–decoder structure with skip connections. The encoder extracts hierarchical features through successive convolutional and pooling layers, while the decoder restores spatial resolution using transposed convolutions. Skip connections link corresponding layers to preserve spatial details and integrate contextual information, ensuring precise segmentation. The network ends with a 1×1 convolution followed by a softmax activation producing four output channels corresponding to the target phases.

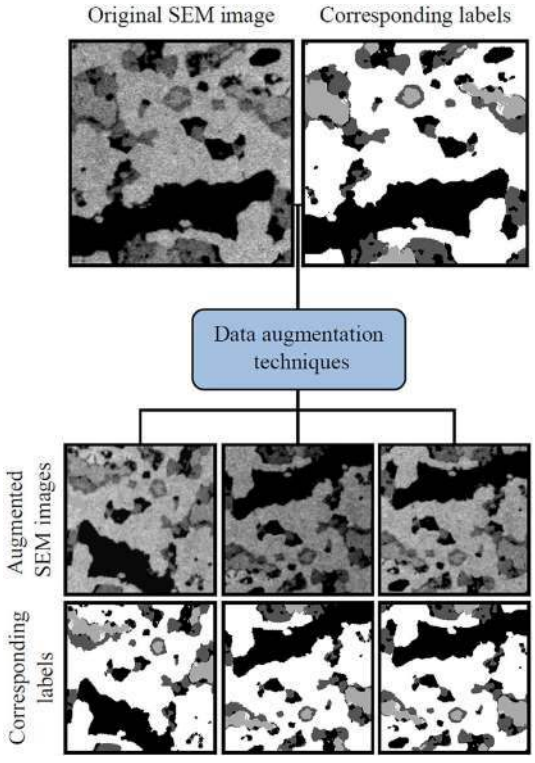


Fig. 3. Example of the workflow of data augmentation techniques applied to an SEM image patch. The SEM patch and its corresponding label undergoes augmentation processes such as flipping, rotation, and intensity adjustments.

This choice of the U-Net architecture is backed by its adaptability to microscopy and materials science images, where it consistently achieves high segmentation accuracy with limited data size. Additionally, U-Net performs effectively when trained on augmented datasets, enabling it to generalize well from a relatively small number of samples [26,40]. Compared to more recent architectures such as Attention U-Net [41]or TransUNet [42], U-Net attains comparable segmentation accuracy on the same dataset while requiring less computational time, fewer parameters to optimize, and smaller data size. In contrast, attention-based models capture long-range contextual dependencies but rely on large and diverse data distributions to generalize effectively, often necessitating extensive augmentation pipelines [42,43].

The model is trained using the categorical cross-entropy (CCE) loss function which is adapted for exclusive multi-class classification tasks:

$$\mathrm{CCE} = -\frac{1}{N} \times \frac{1}{N_p} \sum_i \sum_x \sum_k y_{i,k}(x) \log(\hat{y}_{i,k}(x)), \tag{1}$$

where N is the total number of training images, N_p is the total number of pixels in each image, and i, x, and k represent the indices of the image, the pixel within the image, and the class label, respectively. y is a binary vector that encodes the target, and $\hat{y}$ represents the predicted probability provided by the U-Net.

The Adam optimizer is used to update the network weights, with accuracy used as the evaluation metric to monitor training performance and ensure reliable segmentation results. After several optimization tests, the final training hyperparameters were set to a learning rate of 10^{-3}, 250 epochs, and a mini-batch size of 32.

Training was performed in Python 3.11.9 with TensorFlow 2.17.0 on a Dell PowerEdge R750 server equipped with an NVIDIA A100 GPU (80 GB). The complete training required approximately one hour.

4. Results and discussions

4.1. Overall performance for patches

The performance of the U-Net model was thoroughly evaluated using the confusion matrix, along with the Intersection over Union (IoU) and accuracy (Acc) metrics, defined respectively as:

$$\mathrm{IoU} = \frac{TP}{TP + FP + FN}, \qquad \mathrm{Acc} = \frac{TP + TN}{TP + TN + FP + FN}$$

where TP, TN, FP, and FN denote the numbers of true positive, true negative, false positive, and false negative pixels, respectively. An overall accuracy of 95.53 % and a mean IoU of 91.01 % were achieved across all 140 testing images, demonstrating high precision in predicting all classes. The corresponding confusion matrix is shown in Fig. 4. It highlights the model's strong performance, particularly for the accuracy on Porosity (98 %) and YSZ (97 %).

However, Ni and NiO segmentation experience more frequent misclassifications, which was anticipated due to the similarity of their grayscale values in SEM images. The model achieves 92 % accuracy on Ni phase but shows some confusion with NiO (2 %) and YSZ (5 %). It achieves 93 % accuracy on NiO phase with some misclassification as YSZ (4 %) and as Ni (3 %). Additionally, the model also demonstrates a better understanding of the image context, as shown by its ability to correct certain misclassifications visible in the target segmentation. Overall, the high diagonal values in Fig. 4 demonstrate the model's ability to effectively segment the different phases, achieving satisfactory results despite minor challenges in distinguishing visually similar classes.

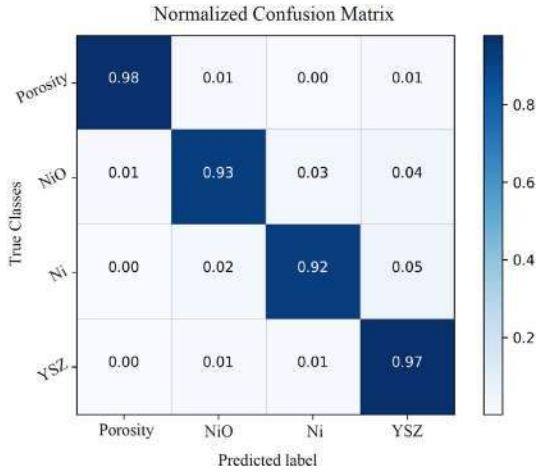


Fig. 4. Normalised confusion matrix illustrating the model's performance.

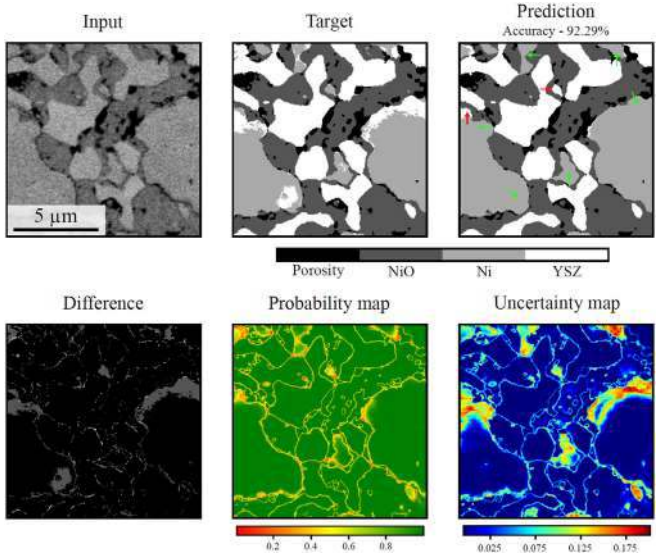


Fig. 5. Visualization of segmentation results for test sample: The first column displays (from left to right) the original SEM image, the target segmentation achieved using EDS mapping, and the model prediction with its accuracy (92.29 %). Green arrows mark regions where the model corrected mislabeled areas in the target, while red arrows indicate misclassifications. The second column includes the absolute difference between the target and prediction, the maximum predicted probability map, and the uncertainty map generated using the ensemble method [44]. In the probability map, red denotes low confidence and green high confidence; in the uncertainty map, red indicates higher uncertainty and blue lower uncertainty. Labels correspond to the four segmented phases: porosity, NiO, Ni, and YSZ. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 5 illustrates the segmentation results on a representative test image, including the input SEM image, the EDS-guided target segmentation, and the U-Net prediction. The global segmentation accuracy is 92.3 % on this image. In more detail, the U-Net model is able to correct some misclassifications in the target segmentation due to errors in the semi-manual labelling as explained in Section 3.1.2, represented by green arrows in the figure. On the contrary, red arrows highlight misclassifications of the U-Net model. Additionally, the figure shows the absolute difference, the maximum predicted probability map, and the uncertainty map. The probability map is obtained by selecting, for each pixel, the highest predicted probability across the four classes. The model exhibits strong confidence in large, homogeneous regions where higher probability values are present, while certainty decreases near phase boundaries and image edges due to grayscale transitions and the loss of global context during patch division.

The uncertainty map is computed using the ensemble method proposed by Lakshminarayanan et al. [44], a widely used technique for quantifying uncertainty in deep learning. The technique is based on training M neural network models independently, initialized with different random parameters and datasets, where each model converges to a different local minimum of the loss function. The final prediction is obtained by averaging the predictions from all models, while the variance of these predictions is interpreted as the epistemic uncertainty. Formally, if we denote by $\hat{p}_k^i$ the prediction given by model i among M models, the uncertainty map is defined as the square root of the variance, given by:

$$\mathrm{Var}(\hat{p}_k) = \frac{1}{M-1}\sum_{i=1}^{M}\left(\hat{p}_k^{(i)} - \frac{1}{M}\sum_{j=1}^{M}\hat{p}_k^{(j)}\right)^2.$$

In this study, the ensemble method was implemented by training $M = 10$ independent U-Net models on the same dataset, with training and validation data randomly shuffled for each model, resulting in 10 distinct segmentations for each test image and the uncertainty maps, as shown in Fig. 5. Elevated uncertainty mainly appears near image boundaries and phase transitions, consistent with the probability maps. However, these disagreements are not entirely uniform; in some areas, a continuous phase of high uncertainty is observed alongside the probability map. This observation suggests that certain transitions remain challenging for the model to resolve, consistent with findings in Bayesian deep learning, which highlight increased predictive uncertainty at boundaries and in transitional regions [45].

Notably, the model proved effective in distinguishing between classes with overlapping grayscale ranges, as illustrated in Fig. 6, which shows a close agreement between the predicted and target pixel value distributions where Ni, NiO, and YSZ may exhibit similar intensities. This suggests that the model leverages additional features—such as texture, size, and shape—to accurately segment the images, effectively exploiting spatial and contextual information beyond grayscale intensity alone.

To conclude on patch performance, it is important to remind that the target segmentation is not "perfect" and may contain misclassified pixels during the creation of the training set. Consequently, the reported accuracy is slightly biased but remains the most reliable estimate available to assess the model's performance. However, as shown in Fig. 5,

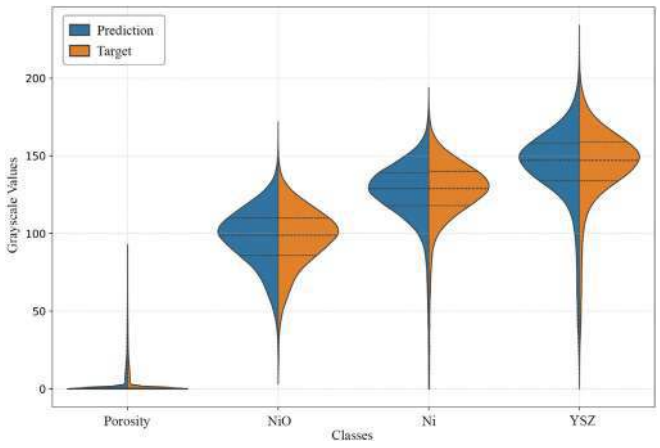


Fig. 6. Violin plot showing the distribution of grayscale values for Prediction and Target across all the test images. The left half of each violin represents the Prediction, while the right half shows the Target distribution.

some of these labelling errors are corrected by the U-Net, particularly for the Ni phases, which are more accurately segmented than in the target. This demonstrates that, despite imperfections in the labeled dataset, the network can still achieve accurate and well-generalized segmentation results.

4.2. Post-processing for large images

Once the U-Net model is trained, new SEM images are used to evaluate the model's generalization capability. Since the model's input size was fixed during training at 256×256, each image is divided into equally sized patches matching the dimensions of the training patches. These patches are then fed into the network for segmentation and subsequently stitched back together to reconstruct the full-size segmented image. However, directly stitching the segmented patches often introduces edge artifacts and discontinuities, particularly in regions with high complexity or fuzzy boundaries. To address these issues, an overlapping-patch strategy using "sliding windows" is employed. The image is divided into overlapping patches of size 256×256 pixels, with a stride of s pixels defining the overlap between adjacent patches; each patch is then independently processed by the segmentation model. Overlapping regions between patches allow each pixel to be predicted multiple times from different perspectives, which are then averaged to enhance consistency and reduce boundary artifacts. This method ensures that the final segmentation map reflects the model's confidence across predictions and provides smoother, more consistent results at patch boundaries. While this method increases the computational workload by producing more patches per image, it improves the visual consistency of the segmented outputs.

Fig. 7 compares the full-image segmentation results generated by the straightforward patch-stitching approach and the sliding window method. The sliding window method achieves better boundary consistency and reduces artifacts, demonstrating that this post-processing step is essential for large-image SEM segmentation tasks.

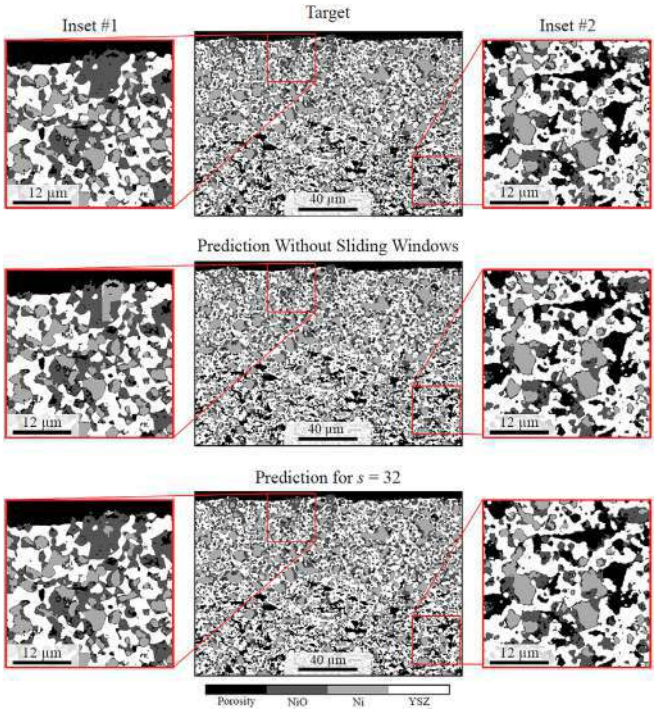


Fig. 7. Comparison of the model's segmentation predictions with and without sliding windows. To highlight key differences, magnified insets are provided and compared against the target segmentation.

Table 1
Time comparison of EDS-guided segmentation and model-based segmentation, with different stride values s.

	EDS-guided segmentation	Model segmentation	
		$s = 256$	$s = 32$
SEM acquisition time (min)	1.33	1.33	1.33
EDS acquisition time (min)	20	N/A	N/A
Segmentation time (min)	240	0.05	2.87
Total time (min)	261.33	1.38	4.20

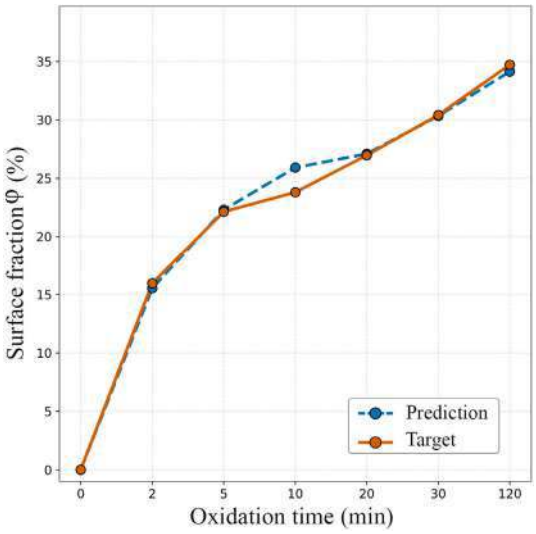


Fig. 8. Comparison of predicted versus target surface fractions φ (%) of NiO phase. The blue dashed line represents the target measured surface fraction and the orange line represents the predicted surface fraction using the proposed approach. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.3. Gain for characterization time

The proposed method offers benefits both in terms of acquisition and processing time. On the one hand, once the U-Net model is trained, the time-consuming EDS acquisition is no longer required, estimated at 20 min per image, as described in Section 3.1.1. On the other hand, our approach proposes a fully automatic prediction and removes the semi-manual labelling with Ilastik and associated post-processing adjustments, which are also time-consuming, empirically estimated at 4 hours on average, as described in Section 3.1.2. The computation time of the U-Net model depends on the stride used in the sliding window post-processing, ranging from 0.05 min for a stride $s = 256$ to 2.87 min for a stride $s = 32$. The comparison of the acquisition and processing times is summarized in Table 1.

In total, our method accelerates the characterization workflow by a factor of 65 to 190 (depending on the sliding window parameters) enabling the automatic and rapid processing of several samples in chain.

4.4. Application to material oxidation characterization

In this section, the segmentation method is applied to characterize oxidation at various exposure times: 0, 2, 5, 10, 20, 30, and 120 min. The acquisition of SEM images was performed under the same conditions. The surface fraction (φ) of the NiO phase was computed for all these exposure times for both the target segmentation, represented by the orange line in Fig. 8, and the model's predicted segmentation, represented by the blue dashed line in Fig. 8.

The prediction was performed on the full-size images by dividing them into patches, processing each patch with the model, then stitching the outputs and applying the sliding window post-processing with a stride $s = 32$. Overall, the mean absolute error (MAE) of 0.51 ± 0.68 and the root mean square error (RMSE) of 0.85 between the target and predicted fractions underscore the predictive accuracy and robustness of the proposed approach. Notably, in specific cases where only three classes are present—such as at 0 min—the model correctly identified the absence of the NiO class and did not predict it.

5. Robustness to acquisition conditions

Multiple factors can influence the quality of the acquired images, including brightness, contrast, noise level, magnification, and overall image clarity. To evaluate the model's sensitivity to these variations, additional tests were performed by artificially modifying the test images prior to prediction. The accuracy of the trained model is measured after introducing varying brightness levels and degrees of blur to the images. The results in Fig. 9(a) show that blurring had a strong effect on the segmentation performance, likely because it reduces the sharpness of phase boundaries and makes it more difficult to distinguish among different regions. In contrast, changes in brightness produced a relatively smaller impact, as the network could still adapt to moderate grayscale shifts as long as sufficient phase contrast was maintained. However, when brightness was substantially increased, accuracy declined noticeably. These artificial brightness modifications can also be observed in some of the test images shown in Figure 9(b), where natural variations in brightness and contrast during SEM imaging affected the input quality.

Additionally, Gaussian and Poisson noise were artificially added to the test images progressively to evaluate the performance of the model under varying noise conditions, where the noisy image can be

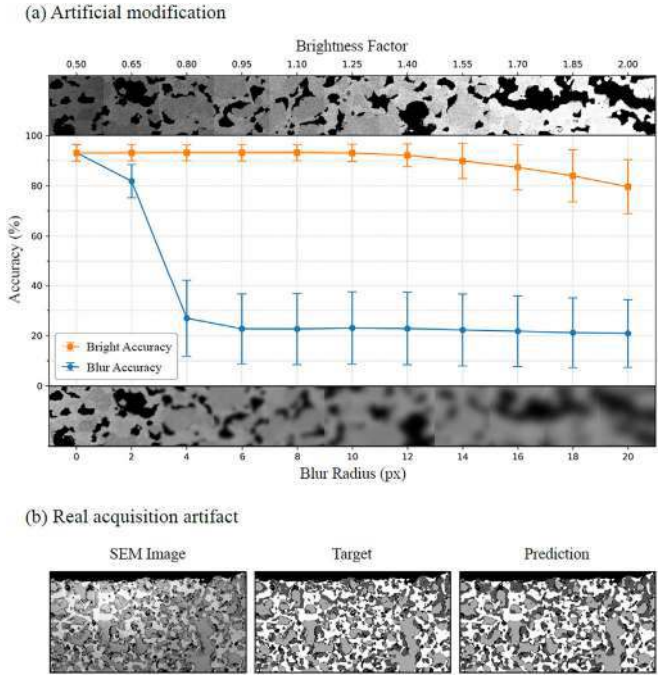


Fig. 9. (a) Effect of Blur Radius and Brightness Factor on Segmentation Accuracy. The plot illustrates the mean segmentation accuracy (%) with standard deviation intervals for varying levels of blur and brightness. The bottom x-axis represents the blur radius, while the top x-axis indicates brightness factors. Error bars correspond to the standard deviation of the accuracy measurements. (b) Model Performance on Real Acquisition Artifacts. This panel shows the model's ability to handle brightness and contrast variations in real SEM image segmentation, compared to the target segmentation. The results demonstrate improved accuracy over EDS-guided segmentation.

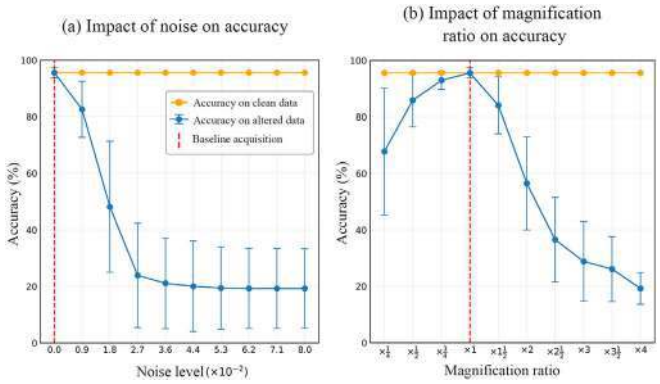


Fig. 10. (a) The impact of noise on prediction performance. Gaussian–Poisson noise was artificially added with progressively increasing noise parameter (σ) to evaluate the model's robustness. (b) The impact of magnification ratio on prediction accuracy. The magnification ratios were progressively varied to assess the model's sensitivity to changes in image scale. For both plots, the orange line represents the accuracy obtained at the original acquisition magnification, the blue line shows the accuracy on altered data, the error bars indicate the standard deviation of the results, and the red dashed line marks the acquisition magnification. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

expressed as:

$$I_{noise} = \frac{\text{Poisson}(I \times C)}{C} + \mathcal{N}(0, \sigma) \tag{2}$$

where I is the normalized clean image, I_{noise} is the same image with added noise, $C = 255$ representing the maximum intensity level of an 8-bit image, and σ is the standard deviation of the additive Gaussian noise. The results in Fig. 10(a) clearly indicate that the accuracy decreases rapidly the more the noise is added but still can give acceptable results in cases where noise level is low. Moreover, the magnification and resolution of the images were studied to assess their effect on images captured at magnification levels different from those used to train the network. To evaluate this effect, the large images were divided into patches of different sizes. Larger patches were subsequently downsampled (binned) to match the model's input size, while smaller patches were upsampled to the same target dimensions. This ensured that all images were processed at a uniform input size, allowing a consistent comparison of segmentation performance across different patch scales. The results in Fig. 10(b) suggest that the network still performed well at magnification ratios close to the original acquisition magnification, achieving an accuracy above 85 % for ratios within $\pm\frac{1}{2}$ of the original value.

These findings highlight the model's sensitivity to image clarity and emphasize the importance of maintaining high-quality acquisitions to ensure that critical features are correctly captured for the model to perform effectively.

6. Conclusion and perspectives

In this study, we introduced a new approach for phase segmentation of SEM images employing a U-Net deep learning architecture. By leveraging EDS-guided labelling during the training phase, our method provides a segmentation model that eliminates the need for time-consuming EDS mapping during prediction, providing a faster and more efficient solution for SOC microstructure analysis. The proposed segmentation model demonstrated three key functionalities:

1. **Accurately segmenting complex microstructural phases directly from SEM images:** The U-Net model achieved a high overall accuracy of 95.53 % and a mean IoU of 91.01 %, effectively distinguishing the four key phases: Porosity, Ni, NiO, and YSZ.

2. **Bypassing the time-intensive EDS mapping process to achieve similar or better segmentation outcomes:** The resulting segmentation from simply using SEM images showed the reliability of the approach without the need to go through expensive time-consuming EDS mappings.
3. **Scalability for Large Images:** Using a sliding window approach, we addressed edge artifacts and achieved smooth phase segmentation for large-scale SEM images.

In a more general sense, the findings demonstrate that deep learning-based segmentation can significantly accelerate microstructural analysis, reducing reliance on traditional manual and semi-automated techniques that often lack statistical rigor or reproducibility. Moreover, our model's ability to operate without EDS broadens its applicability in scenarios where EDS analysis might not be available or possible.

The workflow presented to train the model also exhibits two interesting features:

1. To train a DL model, a segmentation target for SEM images can easily be generated using correlated EDS maps.
2. The trained model predictions are able to surpass the target and show better generalization of features, as they avoid the misclassification errors observed in some training images (see Fig. 5).

These two key points are very general and applicable in many cases of image segmentation, where enriched data could be acquired in a first step to train a DL model, and the model could then be used to segment all other images of the same type of microstructure, without enriched data and with very high accuracy. For instance, in the case of TEM images with EDS maps, or even in X-ray imaging using fluorescence instead of EDS.

To further enhance the proposed method, several improvements can be considered. First, incorporating multi-scale models for segmentation could allow the model to handle images acquired at varying magnifications while maintaining consistent accuracy. Architectures such as multi-scale semantic segmentation U-Net [46] could be explored to achieve this goal. Second, the development of faster and more reliable post-processing techniques is essential to improve computational efficiency and ensure accurate segmentation for large-scale images. Finally, the model could be tested on different materials and microstructures to broaden its scope and confirm its generalizability across a wider range of cases. These enhancements would strengthen the model's robustness and adaptability, making it a more generalized tool for microstructural analysis.

CRediT authorship contribution statement

Kassem Dia: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Thomas David:** Writing – review & editing, Validation, Supervision, Resources, Methodology, Data curation, Conceptualization. **Laure Guetaz:** Writing – review & editing, Supervision, Resources. **Zineb Saghi:** Writing – review & editing, Supervision, Resources, Methodology. **Geoffrey Daniel:** Writing – review & editing, Validation, Supervision, Resources, Methodology, Data curation, Conceptualization.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the first author used ChatGPT o1 model in order to correct grammar errors and improve the writing flow of this article. After using this tool/service, the authors reviewed and edited the content as needed.

Declaration of competing interest

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

Acknowledgment

The authors gratefully acknowledge financial support through the MatCHMaker project which has received funding from the European Union's Horizon Europe research and innovation program under grant agreement N°101091687.

Data availability

The code used in this study is publicly available in GitHub at https://github.com/CEA-MetroCarac/SEM_Phases_Segmentation. A permanently archived version of this repository is accessible via Zenodo at https://doi.org/10.5281/zenodo.17399381. The repository includes an example of an SEM image with its corresponding label, the trained models, and the complete source code used for model training, evaluation and post-processing.

References

[1] Z. Chen, X. Wang, F. Giuliani, A. Atkinson, Analyses of microstructural and elastic properties of porous SOFC cathodes based on focused ION beam tomography, J. Power Sources 273 (January 2015) 486–494.
[2] H.-B. Lee, M.-H. Jung, Y.-H. Kim, E.-B. Park, W.-S. Jang, S.-J. Kim, K.-J. Choi, J.-Y. Park, K.-B. Hwang, J.-H. Shim, S. Yoon, Y.-M. Kim, Deep learning image segmentation for the reliable porosity measurement of high-capacity Ni-based oxide cathode secondary particles, J. Anal. Sci. Technol. 14 (November 2023) 47.
[3] H. Iwai, N. Shikazono, T. Matsui, H. Teshima, M. Kishimoto, R. Kishida, D. Hayashi, K. Matsuzaki, D. Kanno, M. Saito, H. Muroyama, K. Eguchi, N. Kasagi, H. Yoshida, Quantification of SOFC anode microstructure based on dual beam FIB-SEM technique, J. Power Sources 195 (February 2010) 955–961.
[4] C. Mahr, J. Stahl, B. Gerken, F.F. Krause, M. Schowalter, T. Grieb, L. Mädler, A. Rosenauer, Characterization of structure and mixing in nanoparticle hetero-aggregates using convolutional neural networks: 3D-reconstruction versus 2D-projection, Ultramicroscopy 265 (November 2024) 114020.
[5] M. Kishimoto, H. Iwai, M. Saito, H. Yoshida, Quantitative evaluation of transport properties of SOFC porous anode by random walk process, ECS Trans. 25 (Sept 2009) 1887–1896.
[6] J. Hu, Z. Qian, Y. Liu, Y. Xue Ed, Microstructural characteristics of asphalt concrete with different gradations by X-ray CT, J. Wuhan Univ. Technol.-Mater. Sci. 32 (June 2017) 625–632.
[7] M.A.B. Siddique, R.B. Arif, M.M.R. Khan, Digital image segmentation in MATLAB: a brief study on OTSU's image thresholding, in: 2018 International Conference on Innovation in Engineering and Technology (ICIET), December 2018, pp. 1–5.
[8] T. Liu, E. Jurrus, M. Seyedhosseini, M. Ellisman, T. Tasdizen, Watershed merge tree classification for electron microscopy image segmentation, Proc. 21st Int. Conf. Pattern Recognit (ICPR2012) (2012).
[9] M. Calkovský, E. Muller, M. Meffert, N. Firman, F. Mayer, M. Wegener, D. Gerthsen, Comparison of segmentation algorithms for FIB-SEM tomography of porous polymers: importance of image contrast for machine learning segmentation, Mater. Charact. 171 (January 2021) 110806.
[10] L. Almar, J. Joos, A. Weber, E. Ivers-Tiffée, Microstructural feature analysis of commercial li-ION battery cathodes by focused ION beam tomography, J. Power Sources 427 (July 2019) 1–14.
[11] T. Prill, K. Schladitz, D. Jeulin, M. Faessel, C. Wieser, Morphological segmentation of FIB-SEM data of highly porous media, J. Microsc. 250 (May 2013) 77–87.
[12] S. Beucher, C. Lantuéjoul, Use of watersheds in contour detection.
[13] E.A. Zanaty, A. Afifi, A watershed approach for improving medical image segmentation, Comput. Methods Biomech. Biomed. Eng. 16 (12) (2013) 1262–1272.
[14] Y. Zhou, H. Ren, Segmentation method for rock particles image based on improved watershed algorithm, Int. Conf. Comput. Sci. Serv. Syst., (August 2012) 347–349.
[15] W.K. Epting, Z. Mansley, D.B. Menasche, P. Kenesei, R.M. Suter, K. Gerdes, S. Litster, P.A. Salvador, Quantifying intermediate-frequency heterogeneities of SOFC electrodes using X-ray computed tomography, J. Am. Ceram. Soc. 100 (5) (2017) 2232–2242.
[16] C. Li, D. Wang, L. Kong, Application of machine learning techniques in mineral classification for scanning electron microscopy - energy dispersive x-ray spectroscopy (SEM-EDS) images, J. Pet. Sci. Eng. 200 (May 2021) 108178.
[17] Y. Wu, S. Misra, C. Sondergeld, M. Curtis, J. Jernigen, Machine learning for locating organic matter and pores in scanning electron microscopy images of organic-rich shales, Fuel 253 (October 2019) 662–676.
[18] S. Zhang, C. Zhang, Modified U-Net for plant diseased leaf image segmentation, Comput. Electron. Agric. 204 (January 2023) 107511.

[19] S. Bhagat, M. Kokare, V. Haswani, P. Hambarde, R. Kamble, Eff-UNet++: a novel architecture for plant leaf segmentation and counting, Ecol. Inform. 68 (May 2022) 101583.
[20] M.E. Rayed, S.M.S. Islam, S.I. Niha, J.R. Jim, M.M. Kabir, M.F. Mridha, Deep learning for medical image segmentation: state-of-the-art advancements and challenges, Inf. Med. Unlocked 47 (2024) 101504.
[21] A. Bindhu, K.K. Thanammal, Segmentation of skin cancer using fuzzy U-Network via deep learning, Meas.: Sens. 26 (April 2023) 100677.
[22] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A.L. Yuille, DeepLab: semantic image segmentation with deep convolutional Nets, atrous convolution, and fully connected CRFs, (May 2017).
[23] S. Ghosh, A. Pal, S. Jaiswal, K.C. Santosh, N. Das, M. Nasipuri, SegFast-V2: semantic image segmentation with less parameters in deep learning for autonomous driving, Int. J. Mach. Learn. Cybern. 10 (November 2019) 3145–3154.
[24] S.M. Azimi, D. Britz, M. Engstler, M. Fritz, F. Mücklich, Advanced steel microstructural classification by deep learning methods, Sci. Rep. 8 (February 2018) 2128.
[25] Y. Hirabayashi, H. Iga, H. Ogawa, S. Tokuta, Y. Shimada, A. Yamamoto, Deep learning for three-dimensional segmentation of electron microscopy images of complex ceramic materials, NPJ Comput. Mater. 10 (March 2024) 46.
[26] O. Ronneberger, P. Fischer, T. Brox, U-Net: convolutional networks for biomedical image segmentation, (May 2015).
[27] S.S. Bangaru, C. Wang, X. Zhou, M. Hassan, Scanning electron microscopy (SEM) image segmentation for microstructure analysis of concrete using U-Net convolutional neural network, Autom. Constr. 144 (December 2022) 104602.
[28] H. Malik, A.S. Idris, S.F. Toha, I.M. Idris, M.F. Daud, M.O. Tokhi, Deploying patch-based segmentation pipeline for fibroblast cell images at varying magnifications, IEEE Access 11 (2023) 98171–98181.
[29] A. Genc, L. Kovarik, H.L. Fraser, A deep learning approach for semantic segmentation of unbalanced data in electron tomography of catalytic materials, Sci. Rep. 12 (Sept 2022) 16267.
[30] A. Shah, J.A. Schiller, I. Ramos, J. Serrano, D.K. Adams, S. Tawfick, E. Ertekin, Automated image segmentation of scanning electron microscopy images of graphene using U-Net neural network, Mater. Today Commun. 35 (June 2023) 106127.
[31] H. Hwang, J. Ahn, H. Lee, J. Oh, J. Kim, J.-P. Ahn, H.-K. Kim, J.-H. Lee, Y. Yoon, J.-H. Hwang, Deep learning-assisted microstructural analysis of Ni/YSZ anode composites for solid oxide fuel cells, Mater. Charact. 172 (February 2021) 110906.
[32] Y. Wang, Y. Matsui, A. Hosomizo, M. Kishimoto, M. Han, H. Iwai, Phase segmentation of Ni/YSZ anode for solid oxide fuel cells by deep learning, ECS Trans. 111 (May 2023) 457–467.
[33] A. Sciazko, Y. Komatsu, A. Nakamura, Z. Ouyang, T. Hara, N. Shikazono, 3D microstructures of solid oxide fuel cell Ni/YSZ anodes with carbon deposition, Chem. Eng. J. 460 (March 2023) 141680.
[34] G.-X. Zhang, Y. Song, W. Zhao, H. An, J. Wang, Machine learning-facilitated multiscale imaging for energy materials, Cell Rep. Phys. Sci. 3 (Sept 2022) 101008.
[35] H. Fitzek, A. Zankel, M. Dienstleder, J. Rattenberger, H. Schröttner, F. Hofer, Correlating whole sample EDS and Raman mappings – a case study of a Chelyabinsk meteorite fragment, Micron 153 (February 2022) 103177.
[36] L. Li, J. Yang, H. Li, Quantifying the microstructure and phase assemblage of alkali-activated fly ash/slag materials by EDS mapping analysis, Mater. Des. 234 (October 2023) 112320.
[37] L. Li, S. Ding, Y. Cai, Phase-Glue: an interpretable and efficient backscattered electron-energy dispersive X-ray spectroscopy (BSE-EDS) mapping analysis approach to dissect the phase assemblage of cementitious material systems, Expert Syst. Appl. 283 (July 2025) 127923.
[38] R. Juránek, J. Výravský, M. Kolář, D. Motl, P. Zemčík, Graph-based deep learning segmentation of EDS spectral images for automated mineral phase analysis, Comput. Geosci. 165 (August 2022) 105109.
[39] S. Berg, D. Kutra, T. Kroeger, C.N. Straehle, B.X. Kausler, C. Haubold, M. Schiegg, J. Ales, T. Beier, M. Rudy, K. Eren, J.I. Cervantes, B. Xu, F. Beuttenmueller, A. Wolny, C. Zhang, U. Koethe, F.A. Hamprecht, A. Kreshuk, Ilastik: interactive machine learning for (bio)image analysis, Nat. Methods 16 (December 2019) 1226–1232.
[40] T. Falk, D. Mai, R. Bensch, O. Cicek, A. Abdulkadir, Y. Marrakchi, A. Bohm, J. Deubner, Z. Jackel, K. Seiwald, A. Dovzhenko, O. Tietz, C. Dal Bosco, S. Walsh, D. Saltukoglu, T.L. Tay, M. Prinz, K. Palme, M. Simons, I. Diester, T. Brox, O. Ronneberger, U-Net: deep learning for cell counting, detection, and morphometry, Nat. Methods 16 (January 2019) 67–70.
[41] O. Oktay, J. Schlemper, L.L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N.Y. Hammerla, B. Kainz, B. Glocker, D. Rueckert, Attention U-Net: learning where to look for the pancreas, (May 2018).
[42] J. Chen, Y. Lu, Q. Yu, X. Luo, E. Adeli, Y. Wang, L. Lu, A.L. Yuille, Y. Zhou, TransUNet: transformers make strong encoders for medical image segmentation, (February 2021).
[43] H. Cao, Y. Wang, J. Chen, D. Jiang, X. Zhang, Q. Tian, M. Wang, Swin-UNet: UNet-like pure transformer for medical image segmentation, (May 2021).
[44] B. Lakshminarayanan, A. Pritzel, C. Blundell, Simple and scalable predictive uncertainty estimation using deep ensembles, (November 2017).
[45] A. Kendall, Y. Gal, What uncertainties do we need in Bayesian deep learning for computer vision? (October 2017).
[46] Z. Alom, V.K. Asari, A. Parwani, T.M. Taha, Microscopic nuclei classification, segmentation, and detection with improved deep convolutional neural networks (DCNN), Diagn. Pathol. 17 (December 2022) 38.