

<https://doi.org/10.1038/s41524-024-01226-5>

Deep learning for three-dimensional segmentation of electron microscopy images of complex ceramic materials

Check for updates

Yu Hirabayashi¹, Haruka Iga¹, Hiroki Ogawa¹, Shinnosuke Tokuta¹, Yusuke Shimada^{2,3} & Akiyasu Yamamoto ^{1,3}

The microstructure is a critical factor governing the functionality of ceramic materials. Meanwhile, microstructural analysis of electron microscopy images of polycrystalline ceramics, which are geometrically complex and composed of countless crystal grains with porosity and secondary phases, has generally been performed manually by human experts. Objective pixel-based analysis (semantic segmentation) with high accuracy is a simple but critical step for quantifying microstructures. In this study, we apply neural network-based semantic segmentation to secondary electron images of polycrystalline ceramics obtained by three-dimensional (3D) imaging. The deep-learning-based models (e.g., fully convolutional network and U-Net) by employing a dataset based on a 3D scanning electron microscopy with a focused ion beam is found to be able to recognize defect structures characteristic of polycrystalline materials in some cases due to artifacts in electron microscopy imaging. Owing to the training images with improved depth accuracy, the accuracy evaluation function, intersection over union (IoU) values, reaches 94.6% for U-Net. These IoU values are among the highest for complex ceramics, where the 3D spatial distribution of phases is difficult to locate from a 2D image. Moreover, we employ the learned model to successfully reconstruct a 3D microstructure consisting of giga-scale voxel data in a few minutes. The resolution of a single voxel is 20 nm, which is higher than that obtained using a typical X-ray computed tomography. These results suggest that deep learning with datasets that learn depth information is essential in 3D microstructural quantifying polycrystalline ceramic materials. Additionally, developing improved segmentation models and datasets will pave the way for data assimilation into operando analysis and numerical simulations of in situ microstructures obtained experimentally and for application to process informatics.

In functional materials, the controlling of the microstructures that dominate the material performance is essential. Recently, research on the multi-dimensionality of the obtained structural information has been developing. One example is the acquisition of three-dimensional (3D) spatial observations. Typical methods for obtaining 3D structures include optical microscopy and X-ray computed tomography^{1–7}, serial sectioning^{8–12} using a scanning electron microscope with a focused ion beam (FIB-SEM), and transmission electron microscope computed tomography^{13–15}. These 3D images can be characterized as 3D voxel data rather than the conventional

2D pixel data, which allows information, such as phase connectivity, shape, and surface topography, to be obtained with high accuracy¹⁶.

Meanwhile, as these microstructural images have become increasingly multi-dimensional, the data obtained have also become enormous. Consequently, there have been recent projects for objective and automated large-scale data analysis using computer vision approaches^{1,3–5,17–24}. Semantic segmentation, which can accurately extract the material phase related to functional appearance pixel by pixel, is notable^{24–27}. Alternatively, machine learning approaches of semantic segmentation, especially 3D

¹Department of Applied Physics, Tokyo University of Agriculture and Technology, Koganei, Tokyo 184-8588, Japan. ²Department of Advanced Materials Science and Engineering, Kyushu University, Kasuga, Fukuoka 816-8580, Japan. ³JST-CREST, Kawaguchi, Saitama 332-0012, Japan. e-mail: akiyasu@cc.tuat.ac.jp

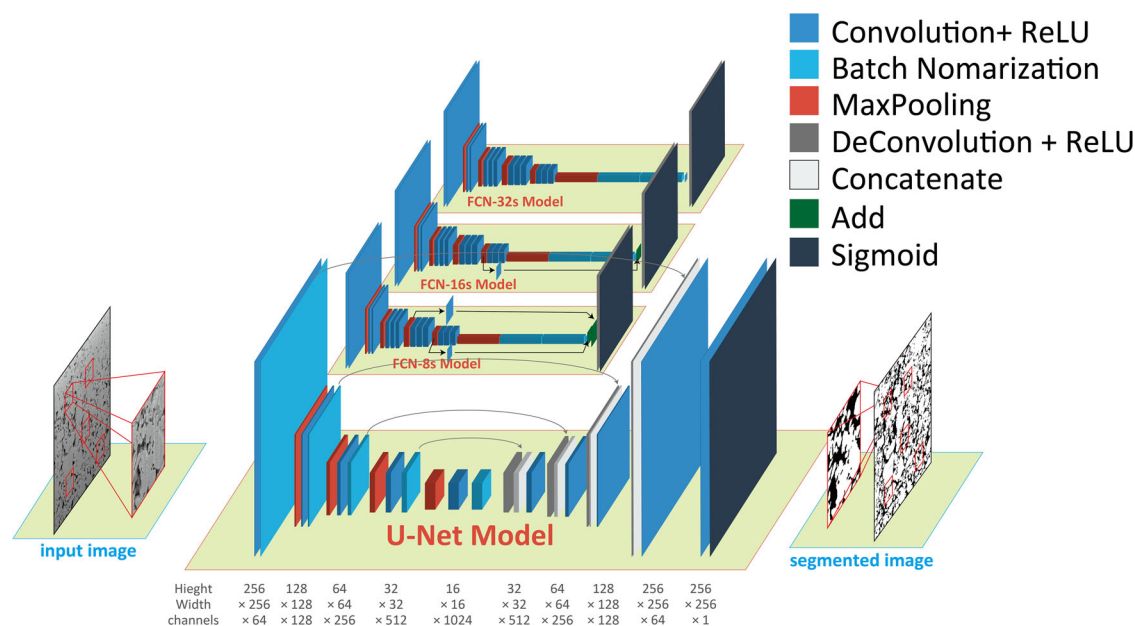


Fig. 1 | Conceptual diagram of the neural network structure. The deep-learning-based semantic segmentation models (FCN-32s, 16s and 8s, and U-Net). The rectangles represent layers: light blue corresponds to the convolution layer + ReLU layer, red to the MaxPooling layer, and green to the upsampling layer in the FCN series; for U-Net, light blue represents the convolution layer + ReLU + Batch-Normalization layer, red represents the MaxPooling layer, gray represents the layer of upsampling as well as copying encoder features, and white represents the concatenate layer. FCN-8s, 16s and U-Net have skip connections (indicated by arrows),

so that these models do not lose the feature details when they are transmitted at each layer of the network, and can obtain high-resolution output while preserving the details of the features. Although FCN integrates feature maps from different layers by adding values for each channel (add), U-Net concatenates the outputs of feature maps obtained from the Encoder by adding them in parallel to the Decoder's feature maps (concatenate). Therefore, U-Net can learn while distinguishing between Encoder and Decoder feature maps.

segmentation^{2,4,5,25}, are focused on X-ray computed tomography, which is nondestructive and easier to acquire 3D microstructure owing to its higher transparency than electrons. The semantic segmentation of 3D microstructures based on electron microscopes, which in principle have higher resolution and can be applied to materials comprising light elements, is expected; however, currently, it is limited to 2D images^{21,26,27}.

In quantitative analysis of microstructural images by electron microscopy, functional polycrystalline materials containing considerable microstructural information, such as porosity, grain boundaries, structural defects, and secondary phases, necessitate accurate segmentation. However, some factors were difficult to recognize due to weak contrast or pseudo-luminance changes caused by experimental artifacts. Besides, machine-based batch segmentation methods, such as the thresholding method, lack accuracy. Therefore, conventional segmentation, which is performed subjectively by experts, has drawbacks, such as inaccurate surface reconstruction due to slight discrepancies in the decision between images and enormous time consumption.

There are two major categories of semantic segmentation methods: classical computer vision and machine learning-based approaches. The thresholding method is one of the classical semantic segmentation methods. Different phases in a microstructural image of a material often appear as regions of different contrast values. When there are two or more phases with different contrasts, the thresholding method that uses the peaks and valleys corresponding to the contrast histogram for segmentation is simple and effective. Software, such as ImageJ²⁸, is well-known. Studies that use the thresholding method for microstructural segmentation by electron microscopy include functional materials, such as superconducting materials^{29,30}, lithium-ion batteries³¹, thermoelectric materials³², nanoporous materials³³, geomaterials³⁴, and superalloys³⁵. On the other hand, the field of computer vision has made remarkable progress with advances in machine learning, like neural networks, and the range of applications is expanding. Typical problems in image recognition include object classification, identification, and detection. Classification involves separating an input image into predefined categories, for which highly accurate models, such as VGG-16 by

Simonyan et al.³⁶ and AlexNet by Krizhevsky et al.³⁷, are known. Detection requires taking an input image and determining where the target is located. For instance, it is applied in pedestrian detection and fingerprint recognition. Categorizing this detection pixel-based with high accuracy is defined as machine learning-based semantic segmentation^{38,39}. The basic model for semantic segmentation is the fully convolutional network (FCN) presented by Long et al.⁴⁰. The FCN model significantly improved segmentation accuracy by transferring pretrained classifier weights, fusing different layer representations, and performing end-to-end learning on whole images⁴⁰. Based on these models, the U-Net and DeepLab models have been improved for medical images¹⁹ and automatic driving identification³⁹, respectively.

As an example of functional polycrystalline ceramic materials, this study performed neural network-based semantic segmentation on microstructural images of iron-based high-temperature superconductors^{41,42} obtained by serial sectioning using a scanning electron microscope. The accuracy evaluation results were compared with the conventional automatic thresholding methods. A giga-scale 3D microstructure reconstruction with a single voxel size of 20 nm based on the learned models was demonstrated.

Results

Models and Datasets

The four semantic segmentation methods used in this study are the classical thresholding method (Otsu method⁴³), the local adaptive thresholding method (Sauvola method⁴⁴), FCN models, and U-Net model¹⁹, which perform deep-learning based on a network structure (Fig. 1). Deep learning of semantic segmentation models with neural networks requires training data: a secondary electron image of a cross-sectional 3D microstructure is cut into 896×896 pixels (Fig. 2a). A training image was created by manually segmenting each of the $\sim 800,000$ pixels into two phases: the positive phase for the superconducting phase and the negative phase for structural defects like voids and impurities (Fig. 2b). A group of supervised graduate students with experience in material synthesis or electron microscopy performed the manual segmentation. First, a draft of roughly segmented images was produced by manually bucket-filling regions in the image, which were

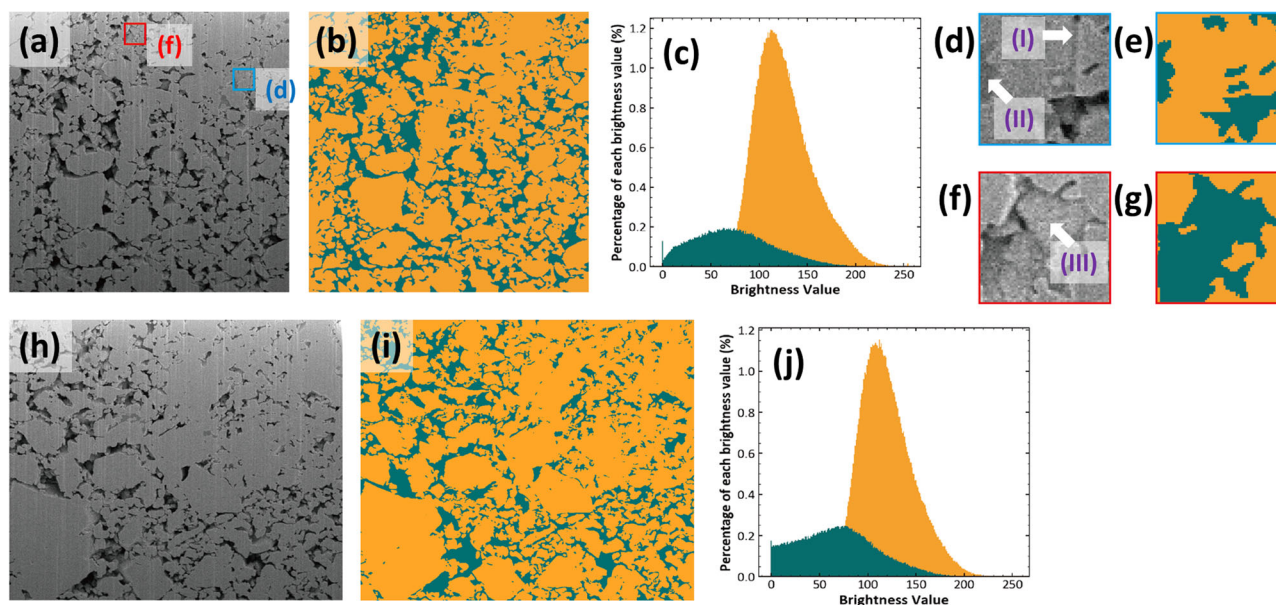


Fig. 2 | Image datasets. a–g Original secondary electron images (a, d, f) paired with their manually segmented images (b, e, g) from the training image and the contrast histogram (c). d–g correspond to the magnified images of the areas enclosed by the blue and red squares in a, respectively. Arrow (I) in (d) indicates the ion polishing

trace artifact. h–j Original secondary electron image of the test image (h) paired with its manually segmented image (i) and contrast histogram (j). The yellow and blue regions correspond to the superconducting phase (positive phase) and microstructural defects, such as voids and impurities (negative phase).

preclassified in eight tones, using a painting software (Adobe Photoshop or Clip Studio Paint Pro). Then, the process of searching for missegmented pixels through visual inspection and improving the image draft was repeated thrice. Artifacts due to ion polishing and impurity phases (indicated by arrows (I) and (II) in Fig. 2d, respectively) were classified as positive and negative phases, respectively. Polycrystalline materials generally contain voids. Noteworthy is that the boundaries of positive phases on smooth slopes adjacent to voids (the extent boundary to which they exist on the corresponding xy cross-section for positive phases that extend continuously in the z (depth) direction) are in many cases difficult to distinguish even by human (for example, arrow (iii) in Fig. 2f). Exploiting the 3D-SEM observation feature using FIB¹⁰, we improved the accuracy of the training image by determining the positive phase based on the difference in brightness between the cross-section of the target and the cross-sections above and below and by considering the continuity of each microstructure and artifact in the z (depth) direction. Using the same method as for the training images, we created a test image used for accuracy evaluation with a size of 1100×924 pixels (Fig. 2h, i). To avoid the manifestation of the overfitting effects and overestimation of segmentation accuracy of deep-learning models, the xy slice cross-sectional position on the z axis of this test image was significantly different from that of the training images.

Quantitative comparison

Table 1 shows the results of accuracy evaluation based on the confusion matrix for Otsu's and Sauvola's thresholding methods, FCN models, and

U-Net model. For the evaluation functions of precision, recall, and intersection over union (IoU), the neural network-based model U-Net model was the best. The confusion matrix, ROC curve, and Precision-Recall curve are shown in Supplementary Table 1, Supplementary Fig. 1a, b, respectively, with corresponding text in Supplementary Note 1.

The Otsu's classic thresholding method provided the smallest IoU values, especially the recall, which corresponds to the percentage of positive phases in the correct image accurately identified as positive phases, which was about 65%, which is significantly lower than the other models. This outcome is primarily due to the misrecognition of the salt pepper-like noise within the positive phase as a defect (i.e., corresponding to a false negative). On the other hand, the best recall value was obtained by Sauvola's local adaptive thresholding method. Precision is an evaluation function that decreases as negative phases are misrecognized as positive phases. It shows differences among the neural network models, with a tendency for precision to increase with the resolution of features concatenated during upsampling. U-Net, which concatenates features at all resolutions, has the highest value compared to the other models. The IoU, which evaluates the overall segmentation accuracy of these models, was highest for U-Net, reaching 94.6%. Note that the IoU value is surprisingly high for polycrystalline ceramics, which contain voids, have continuous contrast variation, and are relatively difficult to segment. It is one of the highest values compared to steel materials, ex. steel (93.9%, Azimi)²¹ and complex-phase steel (>90%, Durmaz)²⁶, which contain few voids and have marked contrast among phases.

Table 1 | Performance metrics for the classic and deep-learning-based segmentation approaches

Model name	Precision	Recall	IoU
Otsu	0.9417 ± 0.0005	0.6478 ± 0.0049	0.6136 ± 0.0047
Sauvola	0.9316 ± 0.0004	0.9936 ± 0.0001	0.9246 ± 0.0004
FCN-32s	0.8456 ± 0.0003	0.9524 ± 0.0005	0.8095 ± 0.0004
FCN-16s	0.9147 ± 0.0002	0.9425 ± 0.0003	0.8642 ± 0.0002
FCN-8s	0.9597 ± 0.0001	0.9574 ± 0.0001	0.9188 ± 0.0002
U-Net	0.9751 ± 0.0002	0.9712 ± 0.0001	0.9464 ± 0.0002

Qualitative comparison (successful cases)

The characteristics of each segmentation method are discussed qualitatively, taking specific microstructural structures as examples. Figure 3 shows, from left to right, the original secondary electron image, the segmented images by the Otsu and Sauvola thresholding methods, FCN-32s, FCN-16s, FCN-8s, and U-Net, and the correct image. Figure 3a is the macroscopic view image, and Fig. 3b–e show the local microstructure, which is a partially enlarged version of (a).

First, focusing on the macroscopic view (Fig. 3a), it can be observed that in the Otsu method, the upper part of the image is often misidentified as the positive phase and the lower part as the negative phase, unlike the other methods. This is because the cross-sectional SEM images in this experiment

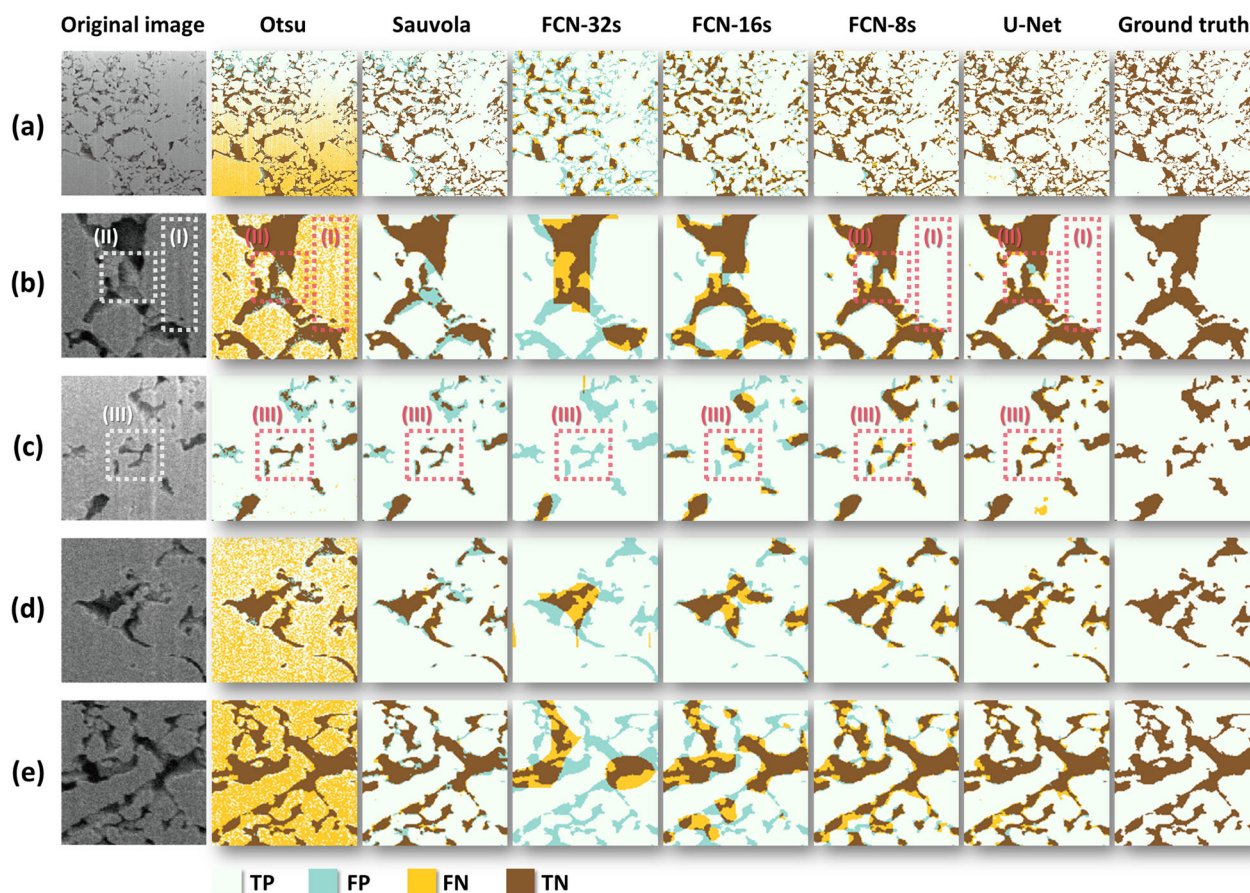


Fig. 3 | Segmentation results for successful cases. Shown from left to right are the original secondary electron images, segmentation images by the Otsu and Sauvola thresholding methods, FCN-32s, FCN-16s, FCN-8s, and U-Net, and the correct image. **a** shows the macroscopic field-of-view image (768×768), and **b–e** show a partially enlarged image of **a**. These are examples of areas where semantic segmentation models based on neural networks have been relatively successful. The dotted squares represent regions (I), (II), and (III). Region (I) in **b** shows ion

polishing traces in the vertical direction. Region (II) in the same image (**b**) contains the ‘valley’ where a void (negative phase) exists between the left and right positive phases. Region (III) in **c** contains the small, independent voids in the positive phase matrix. The pale blue, aqua, yellow and brown pixels in the segmented images correspond to TP (True Positive), FP (False Positive), FN (False Negative), and TN (True Negative), respectively.

were acquired from a 38° direction by cutting the center of the sample using the FIB for its experimental ease and versatility. This acquisition method decreases the background intensity at the bottom of the image due to geometric artifacts where the surrounding cross-section absorbs the generated secondary electrons. In contrast, the Sauvola method, the FCN and U-Net models show that the same original image can be segmented with little effect of changes in background intensity.

Region (I) in Fig. 3b has ion polishing traces in the vertical direction, which appear as dark contrast stripes in the original secondary electron image. Therefore, the Otsu method shows a stripe pattern extending vertically in the corresponding region. However, the FCN and U-Net models did not show any misidentification derived from these ion polishing traces. This result suggests that these neural network models successfully learned the features of ion polishing traces. The Sauvola method also succeeded in segmenting stripes with dark contrast (region (I)), but for bright contrast, missegmentation was observed in the surrounding areas (Fig. 3c). Region (II) in the same image (b) is the ‘valley’ where a void exists between the left and right positive phases and where the positive phase is deeper than the corresponding cross-section. The thresholding methods incorrectly identify part of the areas with bright contrast, especially on the lower side, as the positive phase. Meanwhile, FCN-8s and U-Net correctly identified these areas as a negative phase, which was not affected by the depth reflection.

Figure 3c is a magnified image of the upper part of Fig. 3a, where the contrast is relatively bright. Because the voids are sparsely distributed, and the entire image’s contrast is bright, the thresholding methods do not

segment the relatively shallow part of the voids very well. Next, we consider the differences between the FCN models, focusing on the small, independent voids in the region (III): FCN-32s ignore the voids, and FCN-16s roughly identify them, but their shapes are very different. In contrast, FCN-8s identify the voids, including their rough shape. This is consistent with the quantitative trend in the confusion matrix (Supplemental Table 1), where the False Positive (FP) values were 13.1%, 6.7%, and 3.3% for FCN-32s, FCN-16s, and FCN-8s, respectively. Table 1 shows that although no significant changes were observed among the recall values of the FCN models, the lower precision for FCN-32s than that of the other FCN models and U-Net is mainly because of the large FP, which may reflect the characteristics of high-resolution electron microscopy images of ceramics containing fine voids. This depends on the number of times the upsampling layer is expanded; the higher the value, the less specific the identification. This is thought to have reduced the loss of positional information.

Figure 3d is a close-up of the upper part of Fig. 3a, where the contrast is relatively dark. In the Otsu method, salt pepper-like misidentifications are scattered within the positive phase. Figure 3e shows one of the darkest areas in Fig. 3a. In this region, the accuracy of the Otsu method is significantly degraded, and only the edges are correctly identified as positive phases. However, in the FCN and U-Net models, the contrast brightness or darkness does not seem to have much effect on the segmentation accuracy. The local adaptive thresholding (Sauvola) method no longer missegments the superconducting phase as nonsuperconducting, however, missegments the nonsuperconducting phase with relatively bright contrast (e.g., the edges

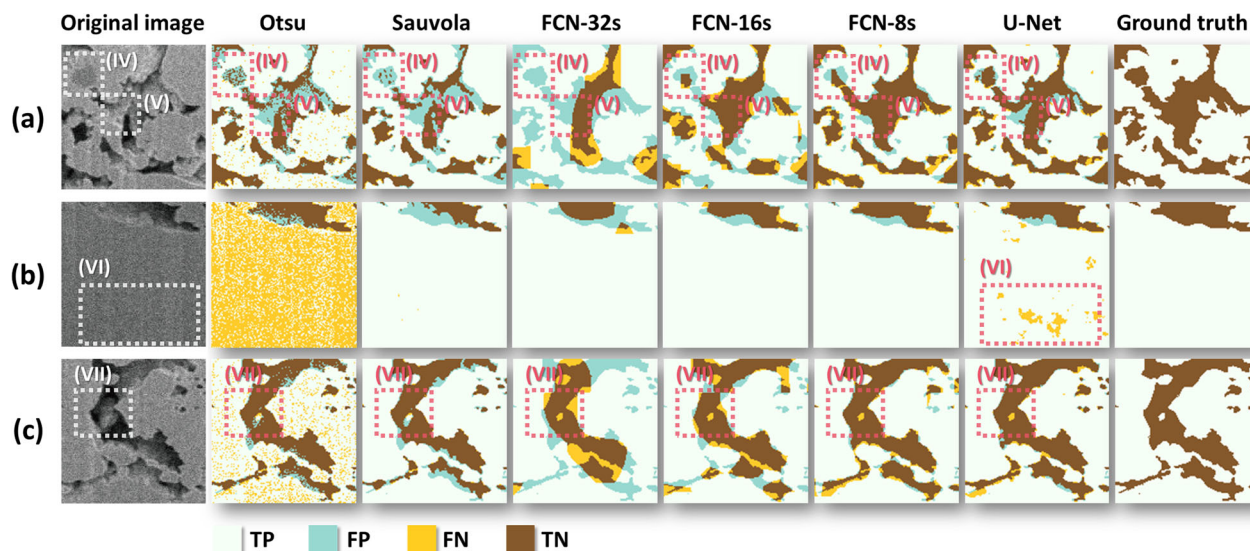


Fig. 4 | Segmentation results for failure cases. a–c From left to right, the original secondary electron image, the segmentation images by the Otsu and Sauvola thresholding methods, FCN-32s, FCN-16s, FCN-8s, and U-Net, and the correct image. These are examples of where the neural network-based semantic segmentation failed to identify. The dotted squares represent regions (IV), (V), (VI), and (VII). Region (IV) in a shows an impurity phase with dark contrast. Region (V) in

a shows a submarine ridge-like superconducting phase in the void deeper than the image cross-section, with relatively high brightness. Region (VI) in b comprises the superconducting phase with negligible defects. Region (VII) in c shows an island-like superconducting phase surrounded by voids. The pale blue, aqua, yellow and brown pixels in the segmented images correspond to TP, FP, FN, and TN, respectively.

and 'valleys'). These images are examples of semantic segmentation based on neural networks successfully performed without being affected by artifacts from electron microscopy observations.

Qualitative comparison (failure cases)

Figure 4a–c shows the original secondary electron image, segmentation images by the Otsu and Sauvola thresholding methods, FCN-32s, FCN-16s, FCN-8s, and U-Net, and correct image, from left to right, as in Fig. 3. These are examples of regions where semantic segmentation did not work well.

In the original secondary electron image in Fig. 4a, there is an impurity phase (IV) and a shallow void (V) where the superconducting phase is reflected in the depth (z axis) direction. Focusing on the impurity phase (IV), the thresholding methods show noise due to its relatively low brightness. Accurate segmentation is difficult even with neural network models. Although U-Net can identify most of them, the accuracy is lower than the segmentation of other voids. This may be because the number of impurity phases in the training images was only six, and the training was insufficient.

Region (V) in Fig. 4a has a peaked superconducting phase in the void deeper than the image cross-section ('submarine ridge'); the brightness is relatively high due to the secondary electron image feature. Consequently, the ridge is misidentified by the thresholding methods and U-Net model. However, among the FCN models, FCN-8s segmented properly. This is because FCN-8s incorporate more global features than U-Net. Thus, it is less affected by the local increase in contrast to the peaked superconducting phase.

The original secondary electron image in Fig. 4b shows that most of the superconducting phase is composed of few defects, whereas U-Net misidentified the superconducting phase as defects mainly in the region (VI). This is because the narrow receptive field of the U-Net discriminated the superconducting phase in a narrow range, and the filter for void recognition was dominant even if the contrast difference was small, resulting in the misidentification.

In Fig. 4c, there is an island-like superconducting phase surrounded by voids (i.e., the region indicated by (VII)), which any of the neural network models, including U-Net, did not identify. In contrast, the thresholding methods succeeded to some extent in the segmentation. The island-like area was determined by considering the secondary electron images of the upper and lower layers of the image. It will be an interesting future challenge to see

if a neural network model can accurately segment points that are difficult to determine even with the human eye.

Accurate training images are indispensable for developing a better semantic segmentation model. It is considered an effective method for acquiring 3D microstructures and using the data of the upper and lower layers of the target cross-section for creating training images.

3D reconstruction

Figure 5 shows the 3D reconstructed images of the 620 stacked original secondary electron images and the 620 stacked images of the positive phases segmented by each deep-learning model. Figure 5a–g shows macroscopic regions ($768 \times 768 \times 620$ voxels), and Fig. 5h–n shows relatively localized microregions ($256 \times 256 \times 206$ voxels) cut from the center of a–g. Focusing on the continuity along the z axis, discontinuous background artifacts are observed for the Otsu thresholding method (Fig. 5b). In contrast, the Sauvola thresholding method (Fig. 5c), the neural network-based models appear to reconstruct the microstructure with continuity in the z axis relatively smoothly (Fig. 5d–g). This suggests that the segmentation is well reproduced and accurate between adjacent images in the z axis. The superconducting phase is identified with the same high accuracy as obtained in the test images throughout. The magnified images in Fig. 5k–m show that the FCN-32s captured relatively globally rough defect features. In contrast, in that order, the U-Net, FCN-8s, and FCN-16s identified more detailed defect objects, as seen in the region (III) in Fig. 3c.

As a quantitative evaluation, the filling ratio of the superconducting phase in each z section image is plotted for each semantic segmentation method in Fig. 6. The mean and standard deviation are shown in Table 2. The z dispersion of the positive phase ratio is because the polycrystalline material's microstructure can be locally coarse and dense. Compared to the Otsu thresholding method, the smooth variation of the positive phase ratio between successive layers in the z direction in the Sauvola thresholding method and the neural network-based methods agrees considerably with the qualitative results observed in Fig. 5. The percentage of positive phases in the training and test images, which the experts manually segmented, were 74.2% and 79.7%, respectively. The difference from the percentage of positive phases predicted by the U-Net and FCN-8s models was small, within 2%.

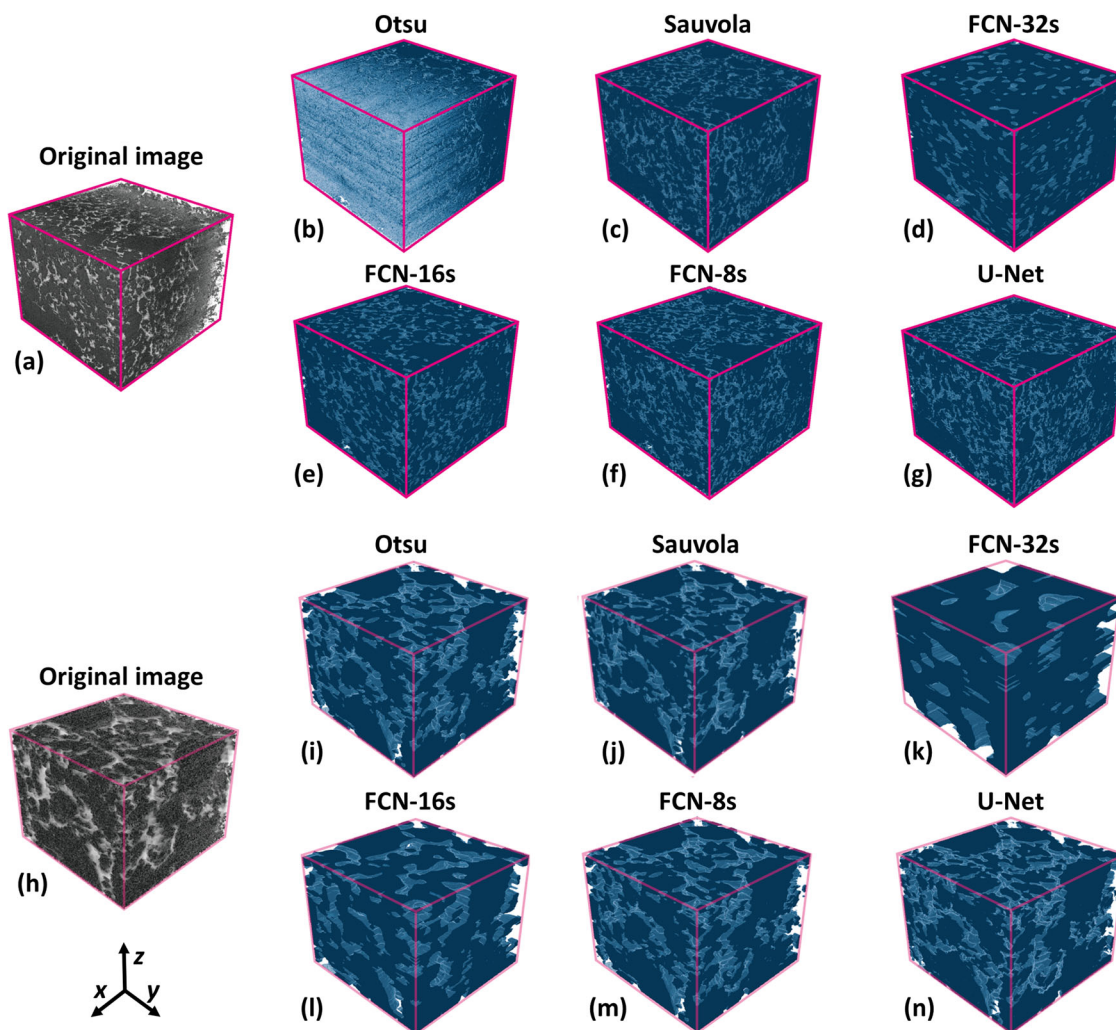


Fig. 5 | 3D reconstructed images from each segmentation model. (Upper a–g) represents a wide area of $768 \times 768 \times 620$ voxels, (lower h–n) a narrower area of $256 \times 256 \times 206$ voxels cutout of the central part of the upper (a–g), enlarged.

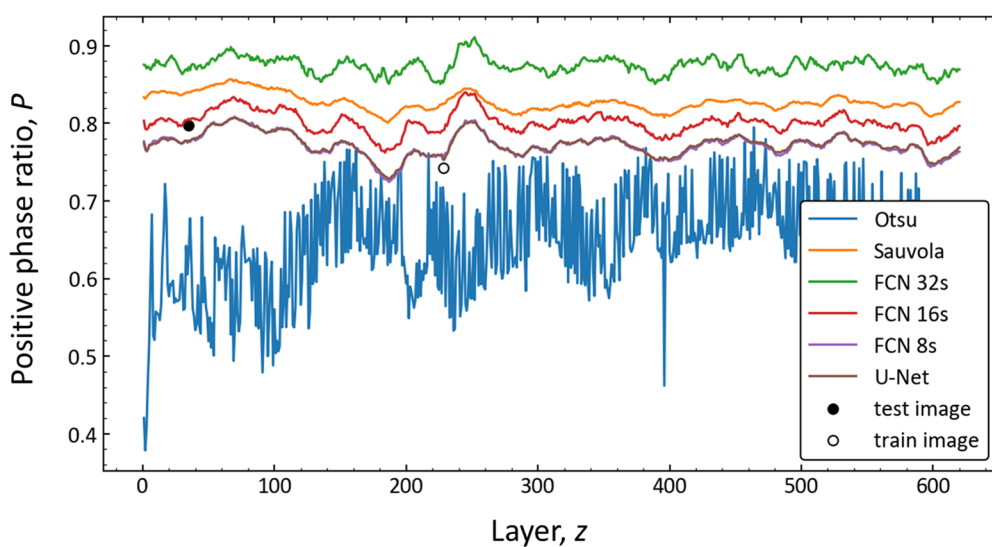


Fig. 6 | Comparison of the variance of the positive phase (superconducting phase) ratio for each z-cross-section by each segmentation method. FCN-8s and U-Net models show almost the same trend (overlapped). The positive phase ratios in the

test and training images, which the expert manually segmented, are also shown for reference.

Table 2 | Mean and standard deviation of the percentage of positive phases for 620 cross-sections for each model

Otsu	Sauvola	FCN-32s	FCN-16s	FCN-8s	U-Net
0.6466	0.8272	0.8743	0.8009	0.7720	0.7728
± 0.0649	± 0.0111	± 0.0104	± 0.0137	± 0.0151	± 0.0143

Compared to the U-Net and FCN-8s models, the FCN-16s, Sauvola, and FCN-32s models tended to overestimate the filling ratio in that order, which was proportional to IoU value among the deep-learning models. On the other hand, it is interesting to note that the Sauvola thresholding method overestimated the filling ratio compared to the FCN-16s model, which has a lower IoU value. This is due to the fact that FP and FN are nearly identical or FN is higher than FP in the neural network models, thus balancing each other out, whereas the Sauvola thresholding method shows a very small FN (0.5%) and the impact of FP is significant.

Discussion

This study demonstrated a method for the automatic and rapid reconstruction of electron microscopy-based 3D microstructures of polycrystalline functional materials with high accuracy using semantic segmentation based on neural network-based models. Compared with the conventional automatic thresholding method, this method significantly increased the tolerance to artifacts associated with electron microscopy, such as polishing marks added during sample preparation and edge brightness derived from the electron microscopic observation. Additionally, by learning patterns that incorporate surrounding information through convolution, they are less susceptible to changes in the brightness of single pixels, which have the advantage of being more noise resistant, such as salt pepper. The segmentation accuracy of the present model is 94.6% for IoU, which is among the highest for an automatic segmentation method^{21,26}, but inferior to that of an expert. However, by improving the model and dataset, AI could successfully identify boundary regions in the depth direction, which even experts cannot distinguish.

The ability to reconstruct electron microscopy-based 3D microstructures of polycrystalline functional materials on a voxel basis with higher precision is expected to make it possible to quantitatively analyze microstructural factors in three dimensions, which has been done mainly for 2D microstructural images. Specifically, 2D images may not produce a reliable quantification of the hidden 3D network structure of voids, secondary phases, and grain boundary phases, as well as the internal surface area, and curvature, especially for materials with highly anisotropic structural feature. Through contrasting experimental electrical or magnetic property mapping, we can elucidate the mechanism of functional manifestation based on the 3D microstructure of the bulk material, including the depth direction. Moreover, machine learning of the 3D voxel big data would result in new microstructural features related to material function, which are not immediately visible in SEM images. Alternatively, in systems where transport phenomena are related to the functionality, such as critical current^{45,46} and phase transition⁴⁷ in superconductors, thermal and electrical conduction in thermal-interface/thermoelectric materials^{48,49}, and ionic conduction in batteries⁵⁰, the percolation theory states that the conduction mechanism varies greatly depending on the system dimension^{51,52}. In the case of 3D bulk materials, the 3D connectivity of the target phase significantly impacts the macroscopic transport properties. In the case of superconductors, the degree of texture in grain orientation⁴⁵ and network of voids and grain boundary phases⁴⁶ are known to significantly affect macroscopic critical current. In the case of thermal-interface materials, high thermal properties of the epoxy-based hybrid composites with binary fillers were reported, where a combination of graphene fillers with (high-aspect ratio) and Cu-nanoparticles fillers (small aspect ratio and nm-scale dimensions) contribute to thermal and electronic percolations⁴⁸. The ability to directly use 3D microstructural information from 3D-SEM, which has become increasingly popular recently⁵³, is expected to provide insights into microstructural factors and

feedback on process design while understanding transport mechanisms previously discussed based on inferences from 2D microstructural images.

Alternatively, the ability to handle a huge amount of data (i.e., more than a billion) on a real voxel basis will pave the way for a 'digital twin' of material microstructures that connects experimental data and computational simulations as the dataset infrastructure for microstructures of various functional materials is developed in the future. For example, it will be possible to integrate experimental data from large-area 3D microstructure observation⁵⁴, in situ observation methods⁵⁵, and operando analysis⁵⁶ with high spatial/temporal resolution, which have been difficult due to the large data size, into multi-scale and multi-dimensional simulations of the microstructural formation and physical properties. This can lead to the development of more accurate prediction models and the application of microstructure data to process informatics.

Methods

Sample preparation

The sample used in this study is a polycrystalline bulk Ba122, which is one of the iron-based superconductors^{41,42}. Mechanically alloyed Ba122 powder was prepared by high-energy planetary ball-milling of elemental metals weighed so that the composition was BaFe_{1.84}Co_{0.16}As₂. The 8% Co-doped Ba122 polycrystalline bulk was prepared by sintering the alloyed Ba122 powder in a vacuum at 600 °C for 48 hours. All powder processes were performed in a glove box in a high-purity Ar atmosphere to minimize oxygen contamination that could cause impurity phases^{57,58}.

3D-SEM imaging

The three-dimensional structural observation was performed by serial sectioning using FIB-SEM (Thermo Scientific Helios 600i)⁵⁹. The secondary electron images were acquired with an acceleration voltage of 5 kV and an Everhart–Thornley (ET) detector. The angle between the Ga ion and electron guns was 52°. The number of pixels in each image is $(x, y) = (1536, 1024)$, and the 3D microstructure was acquired by stacking 620 images with a pitch of 20 nm in the z direction; the equivalent size of one voxel in real space is $(x, y, z) = (20.8 \text{ nm}, 26.4 \text{ nm}, 20 \text{ nm})$. As the images contain areas without the sample, an area measuring 1100×924 pixels was selected from the central part to be used for segmentation.

Models

This study uses four semantic segmentation methods: the classical thresholding method (Otsu method), the local adaptive thresholding method (Sauvola method), the FCN models based on machine learning, and the U-Net model. The Implementation Details section describes the thresholding method; the FCN models are FCN-32s, FCN-16s, and FCN-8s. Their accuracy varies depending on the original convolutional neural network (CNN) models. Figure 1 shows the typical network architecture of the FCN and U-Net models. These models perform segmentation by extracting features using the existing CNN model, performing deconvolution based on these features, and restoring the original image size. FCN-16s performs the same deconvolution as FCN-32s to restore the original image size. It combines the features at one higher resolution layer with the tensor in the 3D (*concatenate*). FCN-16s is a model that concatenates the features in a higher resolution layer and achieves better accuracy than FCN-32s, whereas FCN-8s performs a similar concatenation in two higher resolution layers than FCN-32s to restore the original image size, resulting in even better accuracy than the FCN-16s. U-Net is an improved version that can concatenate the features in all resolution layers, allowing it to focus on even more detailed objects than the FCN models.

Automated training and testing dataset generation

The training dataset for the neural network models was prepared using data expansion from training images. First, for a pair of original secondary electron images of a certain z-section obtained by the 3D imaging mentioned above and its manually segmented image, a training dataset of 1000 images was created by cutting them from random positions to 256×256 -

pixel size and adding rotation and flipping operations randomly. Next, a test dataset of 1000 images for evaluating the accuracy of the classical thresholding method and neural network models were created using data expansion in the same way: for a pair of original secondary electron images of 1100×924 -pixel size and their manually segmented image, a 256×256 -pixel size cropping was performed from random positions, and rotation and flipping operations were also applied. Consequently, 10 training datasets and 10 test datasets were created. These datasets will be published elsewhere.

Implementation details

As described in the Model section, we used four semantic segmentation methods: Otsu's thresholding method, Sauvola's thresholding method, FCN models, and the U-Net model. For the classic thresholding method, the automatic thresholding method (Otsu method) was performed using OpenCV. As shown in Fig. 2c, (i), there is a difference in the brightness trend between pixels corresponding to positive and negative phases. The automatic thresholding method uses this brightness difference to segment the two phases using a specific brightness value as the threshold boundary. For the local adaptive thresholding method, the Sauvola method was performed using scikit-image after applying Gaussian filter. For the deep-learning models, the learning rate, lr , was calculated using the following Eq. (1) with $initial_lr = 0.001$, $\gamma = 0.5$, and $step_size = 20$, where γ is the decay rate, a measure of how much the learning rate decreases with each step size relative to the epochs.

$$lr = initial_lr \times \gamma^{\left(\frac{epoch}{step_size}\right)} \quad (1)$$

The number of training epochs is 120, and the time required for the training is 2 h. In addition, the segmentation of 620 images of 768×768 pixels for 3D reconstruction takes only a few minutes. The time required for performing automated semantic segmentation is significantly faster than that for manual segmentation of each pixel [several days for a training image of 896×896 (802,816) pixels; Fig. 2a].

The training was performed using Python 3.8.8 and TensorFlow 2.4.1 on Nvidia Quadro RTX5000 16 GB GPU.

Loss function

BCE Dice Loss, commonly used in semantic segmentation, was the loss function. X_i is the input image (original secondary electron image); y_i is the correct image (manually identified image); p_i is the predicted image output when input to the model. A per-pixel loss function $\mathcal{L}(x_i)$ calculated from the following Eq. (2) is averaged over all pixels (N : number of pixels) to obtain an image loss function.

$$\mathcal{L}(x_i) = \frac{1}{N} \sum_{i=1}^N -[y_i \log p_i + (1 - y_i) \log(1 - p_i)] + \frac{1}{N} \sum_{i=1}^N \left(1 - \frac{2p_i y_i + \gamma}{p_i + y_i + \gamma}\right) \quad (2)$$

Evaluation function

In this study, we used the confusion matrix as the evaluation function. This method compares the predicted, correct images, assigns one pixel in the predicted image to one of TP, FN, FP, or TN, and counts the number of pixels by applying this process to all pixels. TP is called true positive, where the pixel in the correct image is a positive phase, and the pixel in the predicted image is also the positive phase. FN is called false negative, where the correct one is a positive phase, and the prediction is in the negative phase. FP is called false positive, where the correct one is the negative phase, and the prediction is a positive phase. TN is called true negative, where the correct one is the negative phase, and the prediction is also the negative phase. In other words, TP and TN correspond to the correct cases. The models' evaluation index can be calculated by the functions calculated using the confusion matrix values.

Recall: Percentage of positive phase in the correct image that is correctly identified as positive: Recall = TP/(TP + FN)

Precision: Percentage of positive phase correctly identified among the predicted positive phase: Precision = TP/(TP + FP)

IoU: Rigorous accuracy evaluation is known as the Jaccard index: IoU = TP/(TP + FP + FN)

Data availability

Image datasets will be publicly available at <https://github.com/YamamotoLaboratory/3D-SEM-Segmentation>.

Code availability

Codes will be publicly available at <https://github.com/YamamotoLaboratory/3D-SEM-Segmentation>.

Received: 12 April 2023; Accepted: 9 February 2024;

Published online: 05 March 2024

References

- Yang, T. et al. High strength and damage-tolerance in echinoderm stereom as a natural bicontinuous ceramic cellular solid. *Nat. Commun.* **13**, 6083 (2022).
- Müller, S. et al. Deep learning-based segmentation of lithium-ion battery microstructures enhanced by artificially generated electrodes. *Nat. Commun.* **12**, 6205 (2021).
- Trageser, J. E. et al. The effect of differential mineral shrinkage on crack formation and network geometry. *Sci. Rep.* **12**, 22264 (2022).
- Badran, A. Automated segmentation of computed tomography images of fiber-reinforced composites by deep learning. *J. Sci.* **55**, 16273–16289 (2020).
- Shashank Kaira, C. S. et al. Automated correlative segmentation of large transmission X-ray microscopy (TXM) tomograms using deep learning. *Mater. Charact.* **142**, 203–210 (2018).
- Kobayashi, M., Matsuyama, T., Kouno, A., Toda, H. & Miura, H. Construction of finite element meshes for polycrystal grains model from X-ray CT image. *Mater. Trans.* **57**, 2089–2096 (2016).
- Kim, J. H. et al. Microscopic role of carbon on MgB₂ wire for critical current density comparable to NbTi. *NPG Asia Mater.* **4**, e3 (2012).
- Jangid, D. K. et al. Adaptable physics-based super-resolution for electron backscatter diffraction maps. *npj Comput. Mater.* **8**, 255 (2022).
- Hagita, K., Higuchi, T. & Jinnai, H. Super-resolution for asymmetric resolution of FIB-SEM 3D imaging using AI with deep learning. *Sci. Rep.* **8**, 5877 (2018).
- Alkemper, J. & Voorhees, P. W. Quantitative serial sectioning analysis. *J. Microsc.* **201**, 388–394 (2001).
- Groeber, M. A., Haley, B. K., Uchic, M. D., Dimiduk, D. M. & Ghosh, S. 3D reconstruction and characterization of polycrystalline microstructures using a FIB-SEM system. *Mater. Charact.* **57**, 259–273 (2006).
- Adachi, Y., Morooka, S., Nakajima, K. & Sugimoto, Y. Computer-aided three-dimensional visualization of twisted cementite lamellae in eutectoid steel. *Acta Mater.* **56**, 5995–6002 (2008).
- Gondrom, S. et al. X-ray computed laminography: an approach of computed tomography for applications with limited access. *Nucl. Eng. Des.* **190**, 141–147 (1999).
- Hata, S. et al. Electron tomography imaging methods with diffraction contrast for materials research. *Microscopy* **69**, 141–155 (2020).
- Midgley, P. A. & Weyland, M. 3D electron microscopy in the physical sciences: the development of Z-contrast and EFTEM tomography. *Ultramicroscopy* **96**, 413–431 (2003).
- Kaneko, K. et al. Structural and morphological characterization of cerium oxide nanocrystals prepared by hydrothermal synthesis. *Nano Lett.* **7**, 421–425 (2007).
- Chan, H., Cherukara, M., Loeffler, T. D., Narayanan, B. & Sankaranarayanan, S. K. R. S. Machine learning enabled autonomous microstructural characterization in 3D samples. *npj Comput. Mater.* **6**, 1 (2020).

18. Davydenko, T., Sinclair, D., Chawla, N. & Tahmasebi, P. Deep-layers-assisted machine learning for accurate image segmentation of complex materials. *Mater. Charact.* **192**, 112175 (2022).
19. Ronneberger, O., Fischer, P. & Brox, T. U-net: convolutional networks for biomedical image segmentation. *Med. Image Comput. Comput. Assist. Interv. MICCAI 2015. Lecture Notes in Computer Science* (eds. Navab, N., Hornegger, J., Wells, W. & Frangi, A.) **9351**, 234–241 (Berlin: Springer, 2015).
20. Wang, J., Lv, P., Wang, H. & Shi, C. SAR-U-Net: squeeze-and-excitation block and atrous spatial pyramid pooling based residual U-Net for automatic liver segmentation in computed tomography. *Comput. Methods Programs Biomed.* **208**, 106268 (2021).
21. Azimi, S. M., Britz, D., Engstler, M., Fritz, M. & Mücklich, F. Advanced steel microstructural classification by deep learning methods. *Sci. Rep.* **8**, 2128 (2018).
22. Ajioka, F., Wang, Z. L., Ogawa, T. & Adachi, Y. Development of high accuracy segmentation model for microstructure of steel by deep learning. *ISIJ Int.* **60**, 954–959 (2020).
23. Bellens, S., Vandewalle, P. & Dewulf, W. Deep learning based porosity segmentation in X-ray CT measurements of polymer additive manufacturing parts. *Procedia CIRP* **96**, 336–341 (2021).
24. Yeom, J., Stan, T., Hong, S. & Voorhees, P. W. Segmentation of experimental datasets via convolutional neural networks trained on phase field simulations. *Acta Mater.* **214**, 116990 (2021).
25. James, J. et al. Segmentation of tomography datasets using 3D convolutional neural networks. *Comp. Mater. Sci.* **216**, 111847 (2023).
26. Durmaz, A. R. et al. A deep learning approach for complex microstructure inference. *Nat. Commun.* **12**, 6272 (2021).
27. Stuckner, J., Harder, B. & Smith, T. M. Microstructure segmentation with deep learning encoders pre-trained on a large microscopy dataset. *npj Comput. Mater.* **8**, 200 (2022).
28. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).
29. Bagni, T. et al. Machine learning applied to X-ray tomography as a new tool to analyze the voids in RRP Nb₃Sn wires. *Sci. Rep.* **11**, 7767 (2021).
30. Badica, P. et al. Compressive properties of pristine and SiC-Te-added MgB₂ powders, green compacts and spark-plasma-sintered bulks. *Ceram. Int.* **44**, 10181–10191 (2018).
31. Almar, L., Joos, J., Weber, A. & Ivers-Tiffée, E. Microstructural feature analysis of commercial Li-ion battery cathodes by focused ion beam tomography. *J. Power Sources* **427**, 1–14 (2019).
32. Byrnes, J., Mitchell, D. R. G. & Aminorroaya Yamini, S. Thermoelectric performance of thermally aged nanostructured bulk materials—a case study of lead chalcogenides. *Mater. Today Phys.* **13**, 100190 (2020).
33. Prill, T., Schladitz, K., Jeulin, D., Faessel, M. & Wieser, C. Morphological segmentation of FIB-SEM data of highly porous media. *J. Microsc.* **250**, 77–87 (2013).
34. Hashemi, M. A., Khaddour, G., François, B., Massart, T. J. & Salager, S. A tomographic imagery segmentation methodology for three-phase geomaterials based on simultaneous region growing. *Acta Geotech.* **9**, 831–846 (2014).
35. Smith, T. M. et al. Characterization of nanoscale precipitates in superalloy 718 using high resolution SEM. imaging. *Mater. Charact.* **148**, 178–187 (2019).
36. Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **1409**, 1556 (2014).
37. Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet classification with deep convolutional neural networks. *Commun. ACM* **60**, 84–90 (2017).
38. Jianbo, S. & Malik, J. Normalized cuts and image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **22**, 888–905 (2000).
39. Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K. & Yuille, A. L. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Trans. Pattern Anal. Mach. Intell.* **40**, 834–848 (2018).
40. Shelhamer, E., Long, J. & Darrell, T. Fully convolutional networks for semantic segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**, 640–651 (2017).
41. Kamihara, Y., Watanabe, T., Hirano, M. & Hosono, H. Iron-based layered superconductor La[O_{1-x}F_x]FeAs (x=0.05–0.12) with T_c=26 K. *J. Am. Chem. Soc.* **130**, 3296–3297 (2008).
42. Hosono, H., Yamamoto, A., Hiramatsu, H. & Ma, Y. Recent advances in iron-based superconductors toward applications. *Mater. Today* **21**, 278–302 (2018).
43. Otsu, N. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* **9**, 62–66 (1979).
44. Sauvola, J. & Pietikainen, M. Adaptive document image binarization. *Pattern Recognit.* **33**, 225–236 (2000).
45. Eisterer, M. Predicting critical currents in grain-boundary limited superconductors. *Phys. Rev. B* **99**, 094501 (2019).
46. Yamamoto, A., Shimoyama, J., Kishio, K. & Matsushita, T. Limiting factors of normal-state conductivity in superconducting MgB₂: an application of mean-field theory for a site percolation problem. *Supercond. Sci. Technol.* **20**, 658–666 (2007).
47. Hanzawa, K. et al. Insulator-like behavior coexisting with metallic electronic structure in strained FeSe thin films grown by molecular beam epitaxy. *Phys. Rev. B* **99**, 035148 (2019).
48. Barani, Z. et al. Thermal properties of the binary-filler hybrid composites with graphene and copper nanoparticles. *Adv. Funct. Mater.* **30**, 1904008 (2020).
49. Huang, Y. et al. Tailoring the electrical and thermal conductivity of multi-component and multi-phase polymer composites. *Int. Mater. Rev.* **65**, 129–163 (2020).
50. Saroha, R. et al. Self-supported hierarchically porous 3D carbon nanofiber network comprising Ni/Co/NiCo₂O₄ nanocrystals and hollow N-doped C nanocages as sulfur host for highly reversible Li-S batteries. *Chem. Eng. J.* **446**, 137141 (2022).
51. Kirkpatrick, S. Percolation and conduction. *Rev. Mod. Phys.* **45**, 574–588 (1973).
52. Obara, T. & Yamamoto, A. Quantitative analysis of meandering and dimensional crossover of conduction path in 3D disordered media by percolation modeling. *Supercond. Sci. Technol.* **33**, 074004 1–074004 7 (2020).
53. Yang, Y. et al. One dimensional wormhole corrosion in metals. *Nat. Commun.* **14**, 988 (2023).
54. Burnett, T. L. et al. Large volume serial section tomography by Xe plasma FIB dual beam microscopy. *Ultramicroscopy* **161**, 119–129 (2016).
55. Wang, X. et al. Atomic-scale friction between single-asperity contacts unveiled through in situ transmission electron microscopy. *Nat. Nanotechnol.* **17**, 737–745 (2022).
56. Wang, H., Kline, D. J. & Zachariah, M. R. In-operando high-speed microscopy and thermometry of reaction propagation and sintering in a nanocomposite. *Nat. Commun.* **10**, 3032 (2019).
57. Tokuta, S., Shimada, Y. & Yamamoto, A. Evolution of intergranular microstructure and critical current properties of polycrystalline Co-doped BaFe₂As₂ through high-energy milling. *Supercond. Sci. Technol.* **33**, 0940101–0940108 (2020).
58. Tokuta, S., Hasegawa, Y., Shimada, Y. & Yamamoto, A. Enhanced critical current density in K-doped Ba122 polycrystalline bulk superconductors via fast densification. *iScience* **25**, 103992 (2022).
59. Shimada, Y., Tokuta, S., Yamanaka, A., Yamamoto, A. & Konno, T. J. Three-dimensional microstructure and critical current properties of ultrafine grain Ba(Fe,Co)₂As₂ bulk superconductors. *J. Alloys Compd.* **923**, 166358 (2022).

Acknowledgements

This work was supported by JST CREST (JPMJCR18J4), JSPS KAKENHI (JP21H01615 and JP18H01699), and Nanotechnology Platform (A-18-TU-0037) of the MEXT, Japan.

Author contributions

The work was conceptualized by A.Y. The 3D imaging was performed by Y.S. The semantic segmentations were performed by Y.H., and H.O. Datasets were prepared by H.I. The sample was synthesized by S.T. The paper was written by Y.H., H.I., and A.Y. with input from Y.S. All authors contributed to the discussions and analyses of the data, and approved the final version.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41524-024-01226-5>.

Correspondence and requests for materials should be addressed to Akiyasu Yamamoto.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024