Title
Machine learning-based characterization of surface defects in REBCO tapes

Journal
Superconductor Science and Technology

ISSN
0953-2048

Authors
Menon, Nandana
Croteau, Jean-Francois
Abraimov, Dmytro
et al.

Publication Date
2026-09-04

DOI
10.1088/1361-6668/aea2d9

Superconductor
Science and
Technology

PAPER

Machine learning-based characterization of surface defects in REBCO tapes

To cite this article: Nandana Menon *et al* 2026 *Supercond. Sci. Technol.* **39** 095015

You may also like

Superconductor Science and Technology

PAPER

RECEIVED
31 May 2026

REVISED
23 July 2026

ACCEPTED FOR PUBLICATION
4 September 2026

PUBLISHED
24 September 2026

Machine learning-based characterization of surface defects in REBCO tapes

Nandana Menon[1,*], Jean-Francois Croteau[1], Dmytro Abraimov[2], Hannah Matos-Pimentel[3], Yavuz Oz[2,4], Ian Pong[1], Carina Zha[3], Nicole Bishop[3,5], Jozef Kvitkovic[3], Jun Lu[2] and Jeremy Levitan[2]

[1] Lawrence Berkeley National Laboratory, Berkeley, CA, United States of America
[2] National High Magnetic Field Laboratory, Florida State University, Tallahassee, FL, United States of America
[3] Applied Superconductivity Center, National High Magnetic Field Laboratory, Florida State University, Tallahassee, FL, United States of America
[4] Y. Oz is currently at GE Healthcare, Florence, SC, United States of America
[5] N. Bishop is currently at FSU College of Medicine, Tallahassee, FL, United States of America
[*] Author to whom any correspondence should be addressed.

E-mail: nmenon@lbl.gov

Keywords: REBCO, microstructural defects, scanning electron microscopy, critical current, machine learning, image segmentation

Abstract

REBCO coated conductors are of significant interest for high-field superconducting applications owing to their exceptional critical current retention under high magnetic fields. However, depending on the fabrication route, microstructural inhomogeneities such as porosity, Cu_xO precipitates, and *a*-axis oriented grains can emerge at various length scales, disrupting current transport and limiting the critical current density. This study investigates and characterizes such defects in commercial REBCO conductors of varying specifications. Top-view scanning electron microscope (SEM) images of the REBCO layer were acquired following chemical etching of Cu and Ag layers to expose the microstructure for analysis. Conventional image analysis and segmentation techniques prove insufficient for reliably quantifying these defects, while manual identification remains prohibitively labor-intensive. To overcome these limitations, a machine learning approach is explored to enable rapid, automated, and accurate defect detection. Specifically, an open-source computer vision model, Mask region-based convolutional neural network, is fine-tuned on domain-specific SEM image data. The fine-tuned model achieved a validation mean average precision of 45.23% and enabled quantitative defect analysis across 63 tapes. Partial correlation analysis, controlling for confounding between defect types, revealed independent associations with critical current density that varied in strength and sign across defect types and measurement conditions. These findings motivate further targeted characterization to establish the microstructural origins of these relationships and inform conductor optimization.

1. Introduction

The growing demand for REBCO coated conductors, driven by next-generation accelerator and fusion magnet programs, has spurred a rapid increase in the number of industrial manufacturers worldwide and accelerated the optimization of fabrication processes [1]. At present, REBCO conductors are produced through a variety of routes, the most common being metal organic chemical vapor deposition (MOCVD), metal organic deposition, pulsed laser deposition, and co-evaporation. Significant efforts have been directed toward developing high-throughput methodologies to meet this growing demand, including advances such as modified MOCVD processes and optimized deposition parameters that improve yield and conductor performance [2]. Nevertheless, the fabrication of REBCO conductors remains an inherently complex and sensitive process. Even within a single fabrication route, manufacturers may employ different stoichiometries and doping strategies tailored to specific performance targets. For instance, SuperPower Inc. produces both Advanced Pinning (AP) and the High Magnetic (HM) tape variants with varying Zr % content, optimized for different operating conditions. Beyond compositional choices, processing parameters such as REBCO layer thickness, deposition temperature, and oxygen partial pressure among others can each independently influence the final tape characteristics. Even when all parameters are nominally fixed, deviations from ideal growth conditions can arise due to limitations in process control. Small variations in these tightly interrelated parameters can therefore result in measurable differences in tape characteristics and superconducting performance, often manifesting as spatial inhomogeneities across the width and length of a tape. These inhomogeneities are frequently reflected in the formation of non-superconducting secondary phases and structural defects. While some secondary phases, such as Y_2O_3 [3], $BaZrO_3$ [4], and Y_2BaCuO_5 [5], can be intentionally engineered at the nanoscale to serve as flux pinning centers, most others are undesirable consequences of fluctuating fabrication variables. Misoriented grains, voids, cracks, and precipitates such as $RECu_xO_2$ and Cu_xO are known current-limiting defects. The detrimental role of *a*-axis grains in particular has been well demonstrated: by patterning a section of tape free from *a*-axis grain growth, the peak pinning force density was shown to increase by as much as 40% [6]. Despite such insights, a significant performance gap persists between long-length coated conductors produced in industrial settings and short scientific film samples fabricated under controlled, static conditions. Bridging this gap requires a deeper understanding of how local processing variations translate into microstructural defects and, ultimately, into performance non-uniformity. As production scales up to meet industrial demand, ensuring consistent conductor quality becomes increasingly critical. Robust quality assurance methodologies must therefore be developed in parallel with fabrication advances, underpinned by a thorough understanding of the relationship between processing conditions, microstructure, and superconducting performance.

Characterizing these micron and sub-micron features is a necessary step toward understanding the relationship between fabrication conditions and superconducting performance, yet it remains severely limited by the time and effort required for image acquisition and processing. As a result, manufacturers often resort to quick qualitative assessments based on the overall appearance of the REBCO surface in scanning electron microscope (SEM) images, which are insufficient for capturing the statistical distribution and severity of defects across a tape. This bottleneck hinders a systematic understanding of which production parameters truly need to be controlled and what the cumulative effect of small process variations is on the final conductor quality. Previously, researchers have attempted to quantify defects by manually counting, labeling and measuring them [7]. Traditional image analysis, such as pixel-intensity based thresholding, have been used to analyze cracks from slitting in REBCO [8]. However, this method is not robust to varying imaging conditions requiring the user to iteratively refine image processing conditions for every image. Moreover, the weak boundary contrast of the secondary-phase defects against the surrounding REBCO matrix in SEM images prevents reliable segmentation using thresholding or clustering techniques.

Advancements in computer vision and machine learning (ML), combined with decreasing computational costs, are enabling novel defect detection and image-based analysis pipelines across a wide range of scientific and industrial domains. Within the superconductor community, early applications of ML to microstructural analysis are beginning to emerge. A recent study demonstrated the use of generative adversarial networks to infer the internal defect morphology of REBCO tapes from surface magnetic field maps, though its efficacy in practical systems remains to be validated [10]. Bagni *et al* [11] implemented U-Net, a convolution neural network (CNN) architecture, for the semantic segmentation of crack formation and propagation in Nb_3Sn under transverse compressive stress from tomographic images. However, semantic segmentation assigns a class label to each pixel without distinguishing between individual object instances, and therefore lacks the capacity for instance segmentation, wherein objects belonging to distinct categories are delineated separately. More directly related to the present work, an earlier effort by Matos-Pimentel *et al* [9, 12] attempted to segment surface defects in SEM images of REBCO coated conductors using Trainable Weka Segmentation [13], a primitive ML plugin within the Fiji image processing platform [14]. While this approach produced some degree of semantic segmentation (individual defects were not classified), it struggled with the reliable separation of closely located features (as seen in figure 2 of [12]). This limitation likely stems from the fact that the plugin's core feature extraction methods were originally developed and tuned for biological structures, such as cellular membranes, rather than the distinct morphological characteristics of inorganic surface defects.

The present work builds upon this foundation by adopting a more capable model architecture. Image segmentation and classification are increasingly pursued using ML, and open-source models developed by industry offer researchers access to state-of-the-art capabilities without the burden of training from scratch. Transfer learning further enables these pretrained models to be adapted to specialized domains such as SEM-based microstructural analysis with relatively modest fine-tuning effort, making it particularly well-suited to scientific applications where labeled data is scarce. Leveraging this, Mask region-based

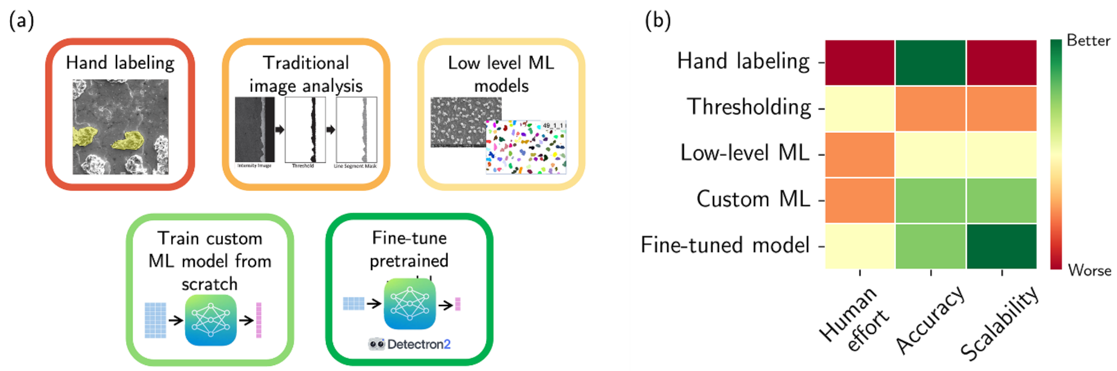


Figure 1. Defect segmentation approaches, ranging from manual hand labeling to fine-tuned open-source computer vision models. (a) Schematic overview of the five approaches discussed, and (b) qualitative comparison of each approach across three metrics: human effort, segmentation accuracy, and scalability. As the methodology advances from hand labeling to fine-tuning a pretrained model, human effort decreases while segmentation accuracy and scalability increase. Reproduced from [8]. © The Author(s). Published by IOP Publishing Ltd CC BY 4.0. The example for 'Low-level ML models' shows Weka segmentation applied to one of the tapes considered in this study, with the Weka interface figure adapted from [9], with permission from the authors.

convolutional neural network (R-CNN) was previously demonstrated by Croteau *et al* [15] to be highly effective in quantifying crack distributions from digital microscopy images of Nb_3Sn stacks, significantly accelerating image analysis and enabling the extraction of statistically meaningful data. The same architecture is adapted here for the detection and classification of surface defects in SEM images of REBCO coated conductors. A scalable Python-based framework is developed around this model to automate defect segmentation and quantification. The resulting defect statistics are subsequently used to investigate potential correlations between the microstructure of the tape, processing conditions, and measured superconducting transport performance. Figure 1 summarizes the common defect analysis strategies discussed thus far, comparing each approach across three key metrics: human effort, accuracy, and scalability. Hand labeling, while serving as the ground truth reference, demands considerable human effort and is inherently limited in scalability across large numbers of images and samples. Traditional image analysis is comparatively straightforward and offers improved scalability, but requires parameter tuning when imaging conditions vary and is not easily generalized to defects that lack clear contrast with their background. Low-level ML models [9, 12] and models trained from scratch [11] both require substantial annotated training data, with the latter achieving higher segmentation accuracy at the cost of greater resource investment. The transfer learning via fine-tuning approach adopted in this work also requires annotated training data, albeit significantly less than training from scratch, and offers the greatest scalability as the model can be readily adapted to related segmentation tasks with minimal additional input. While hand labeling, custom ML, and fine-tuned models are comparable in accuracy, a well-trained model can surpass human labeling in consistency by applying learned segmentation criteria uniformly across all images, free from annotator fatigue or subjectivity. However, insufficient training data or inter-annotator inconsistencies in the training labels can confound model performance. Given these two competing factors, the two ML approaches are considered marginally lower in accuracy than hand labeling.

2. Methodology

2.1. Dataset

The dataset comprises 63 graded SuperPower Inc. AP tapes (7.5 % Zr) that were purchased for the 40 T magnet project at NHMFL [16]. The tapes include 4 mm wide sections that were mechanically slit from the original 12 mm wide tapes. As a result, the data set was a mix of double slit (4 mm wide section from the middle) and single-slit (4 mm wide sections from either sides) tapes. The tapes constitute variations in the production parameters as well as in the MOCVD system used. These span a range of processing conditions, encompassing 2 different machines, 6 grades, and 10 distinct sets of process parameters. The specific proprietary details of these variables were not disclosed by the manufacturer, which limits the extent to which the influence of any of these factors could be isolated. Tapes were grown with intentionally different production parameters, including REBCO thicknesses, to produce graded conductors necessary to manufacture a high-field all superconducting magnet [16]. For each sample examined,

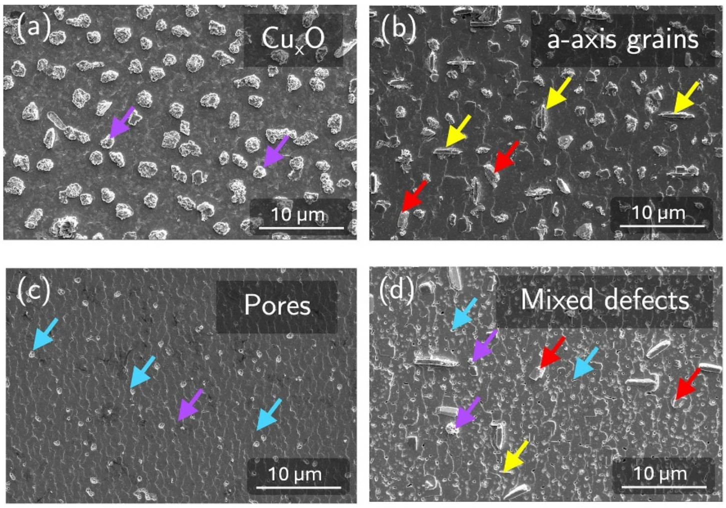


Figure 2. Surface defects observed on REBCO layer following chemical etching of the Cu and Ag layers. Tapes exhibiting predominantly (a) Cu_xO particles (select particles indicated by purple markers), (b) needle-like *a*-axis oriented grains (yellow arrows), along with other misoriented grains (red-arrows), collectively referred to as *a*-axis grains throughout this work, (c) pores (cyan arrows) accompanied by smaller Cu_xO particles, and (d) a heterogeneous mixture of defects characterised by a highly discontinuous REBCO layer with large misoriented grains, pores, and Cu_xO particles.

10 mm long sections were cut. The Cu and Ag layers were chemically etched to expose the REBCO layer. The Cu layer was etched using ammonium persulfate (APS-100), and the Ag layer was etched using a mixture of ammonium hydroxide and 30% hydrogen peroxide solution in water (5:2 ratio) [17]. Both etchants are selective toward their respective metallic layers and act rapidly, minimizing potential attack on the underlying REBCO surface [18]. Top-surface imaging was then performed using a Thermo Fisher Helios G4 UC focused ion beam SEM (FIB-SEM). To protect the REBCO edge during cross-sectional analysis, the original silver layer was left intact over approximately one-third of the sample. FIB milling was then employed to prepare cross-sections through this protective silver layer, allowing for precise measurements of the REBCO film thickness. Imaging conditions, including resolution, dwell time, and pixel size, were kept consistent across the dataset, with only minor variations in brightness and contrast. Energy-dispersive x-ray spectroscopy was routinely performed to validate the nature of identified defects. This was especially critical for confirming the higher Cu content associated with the Cu_xO precipitates observed on the REBCO surface. Note that Cu_xO here refers collectively to all Cu–O compounds that sometimes contain trace amounts of Y and Gd. Figure 2 shows representative SEM images highlighting the major defects observed. The dataset also consisted of samples with a mixture of the defects, as seen in figure 2(d).

Transport measurements were performed at both ends of each spool, and the critical current density (J_c) reported here represents the average of the two measurements, hereinafter denoted $J_{c,\mathrm{avg}}$. Measurements at 77 K under self-field and at 4.2 K under an applied background field of 14 T at two orientations relative to the tape surface, 90° and 18°, are considered in this study. The reader is directed to [17] for more details on the imaging and transport measurements.

2.2. Data pre-processing

Pre-processing the data for consistency is key to training a reliable ML model. Figure 3 outlines the pre-processing workflow followed prior to training the ML model for detection and segmentation. Before any processing, the pixel-to-physical length information was extracted from the metadata of each raw image in its *.tiff* format. All images used in this study share a consistent pixel size (i.e. scale). The images were then cropped to remove the information bar at the bottom since this was unnecessary data adding up the computation size. Note, all images were of the same resolution (1536×1094) so cropping to remove the information bar yielded all images of the same size (1536×1024). In cases where the dataset comprises mixed sized images, padding (addition of pixels at borders) would be necessary to have a consistent input size for the model. A rough visual inspection was performed over larger areas of each tape to confirm the general defect distribution. However, for training, only one high-fidelity image, acquired from roughly the mid-section of the tape, was used per sample. This image was taken to be representative of the entire 10 mm segment analyzed, with certain exceptions discussed in section 3.2.

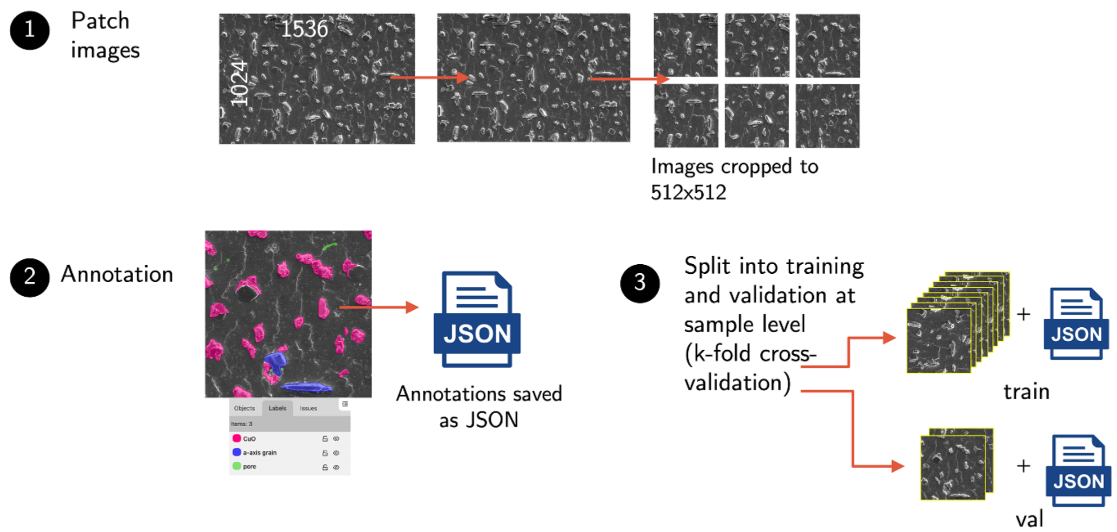


Figure 3. Data preparation workflow for training and validating Mask-RCNN model on SEM images of REBCO. This involves: (1) Patching images into smaller subsets (here 512 pixel × 512 pixel); (2) Annotating the images by marking the boundaries of the defects and assigning their respective labels. Annotations in COCO format are saved as JSON files; (3) Splitting the dataset (images + annotations) at the image/sample level into training and validation dataset for k-fold cross-validation.

Thereafter, the images were divided into patches of size 512 pixels × 512 pixels, resulting in six patches per image. Patching is beneficial in (i) increasing the overall computational efficiency in handling large-scale images, (ii) improving the localized learning per image, and (iii) reducing images into more manageable subsets for annotation. A total of 110 patches, were labeled using CVAT.ai.cloud [19], an open data annotation platform. To prevent information from a single source image being shared across training and evaluation sets, splitting was performed at the source-image level rather than the patch level, with all patches derived from a given image assigned exclusively to one partition. This was necessary because patches from the same image share acquisition-specific characteristics, such as background noise texture and imaging artifacts, which could otherwise leak information between sets. The annotation process was simplified by using the hosted ML model which can approximate a polygon around a defect by separating the foreground from the background. Adjustments were made to the polygon, if necessary, to accurately outline the defect and then labeled into the three defect categories considered in this study: (1) pores, (2) Cu_xO particles, and (3) *a*-axis grains (including *a*-axis oriented grains and other misoriented outgrowths). The annotations were saved in the common objects in context (COCO) format, a widely used JSON-based standard for annotating computer vision datasets [20].

2.3. Transfer learning using mask R-CNN

Instance segmentation of the SEM images was performed using Mask R-CNN [21], implemented using Detectron2 [22], a PyTorch-based platform developed by Facebook AI Research (FAIR) for object detection and segmentation. It is a two-stage model that first generates region proposals representing potential object locations, then refines and classifies each proposal. For every detected instance, the model outputs a bounding box, a pixel-level segmentation mask, a class label, and an associated confidence score. This state-of-the-art vision architecture was originally trained on a large dataset of generic images. Rather than building a model from scratch, which demands substantial training data and computational resources, the predictive capabilities of such models can be leveraged for domain-specific tasks through transfer learning. Transfer learning refers to the reuse of knowledge gained from solving a *source* problem to improve performance on a related *target* problem. While multiple strategies exist for employing transfer learning, for vision-based problems such as the one addressed in this study where only limited set of labeled images are available, a common approach is to *fine-tune* the source model's pre-trained weights on the target dataset. The performance of the *fine-tuned* model was assessed using standard COCO evaluation metrics [20]. These metrics include:

Table 1. Distribution of instances by class and object scale in the dataset.

Class	Scale						Total (%)
	Small		Medium		Large		
Pores	1186	(22.53%)	122	(2.32%)	2	(0.04%)	24.89%
CuxO	2458	(46.69%)	834	(15.84%)	37	(0.70%)	63.24%
a-axis grains	330	(6.27%)	289	(5.49%)	6	(0.11%)	11.87%
Total:	5264						100%

- Intersection over Union (IoU): The ratio of the intersection to the union of the predicted and ground truth bounding box (or mask), quantifying spatial overlap. A predefined IoU threshold (0–1) determines whether a prediction is counted as a true positive (above threshold) or false positive (below threshold).
- Average precision (AP, 0%–100%): Computed from the precision–recall curve, where precision is the ratio of true positives to predicted positives and recall is the ratio of true positives to actual positives. The COCO-style AP is the mean over IoU thresholds from 0.50 to 0.95 in steps of 0.05. AP50 and AP75 denote AP at IoU thresholds of 0.50 and 0.75, respectively. AP by object scale is denoted as APs, APm, and APl, for small ($< 32^2$ pixels), medium (32^2–96^2 pixels), and large ($> 96^2$ pixels) instances, corresponding to $< 745.5\,\mathrm{nm}^2$, 745.5–$6709.7\,\mathrm{nm}^2$, and $> 6709.7\,\mathrm{nm}^2$, respectively, given the fixed pixel size of 26.97 nm used in this study.
- mAP (mean average precision): The mean of the AP scores across all classes in the dataset.

A deeper insight into what the data contains is provided in table 1. The dataset was highly imbalanced both in terms of class distribution and object scale. Note that this is specific to the images acquired in this study and is not a generalization of feature distributions in REBCO tapes. Cu_xO dominates the instances, constituting for $\sim 63\%$ of the total training instances, while *a*-axis grains were the most underrepresented. In terms of scale imbalance, the majority of instances were small objects ($\sim 75\%$), with very few large instances. This imbalance had direct implications on the trends in model performance, which will be discussed in the section 3.1.

3. Results and discussion

3.1. Model training

Standard Detectron2 defaults were used for domain adaptation while fine-tuning. Training used stochastic gradient descent with base learning rate 1×10^{-4}, batch size 2, and 5000 iterations. Validation was performed every 100 iterations on the held-out validation dataset. At each training iteration, the total loss, defined as the sum of the classification, bounding box, and mask losses described in [21], was calculated for the training and validation datasets to monitor learning progress and detect potential overfitting (i.e. when validation loss $\gg$ training loss). All training campaigns were performed on a single node of the Einsteinium GPU cluster at Lawrence Berkeley National Laboratory, equipped with one NVIDIA A40 GPU (48 GB VRAM) and 4 CPU cores. Upon completion, the trained model configuration and weights were exported as *.yaml* and *.pth* files, respectively. These files were then used for inference. An overview of the fine-tuning process is outlined in figure 4.

Detectron2's *Model Zoo* provides a collection of ready-to-use, pre-trained models trained on standard benchmarks that is maintained by FAIR. To identify the most suitable backbone for the SEM images, a systematic comparison of Mask R-CNN configurations available through the Model Zoo was conducted. These configurations vary in network depth, feature extraction strategy, and multi-scale feature aggregation. Since these models were originally optimized for natural image benchmarks, their relative performance does not necessarily transfer to SEM images and hence warrants an ablation study. This study considers two ResNet backbones (ResNet-50 and ResNet-101), each paired with one of three feature extraction strategies. C4 uses only the final backbone layer as a single-scale feature map. DC5 is a variant of C4 that adds dilated convolutions for higher spatial resolution. Feature pyramid network (FPN) combines multi-scale features across backbone layers. All resulting Mask R-CNN variants were evaluated under both 1× and 3× training schedules (where available), yielding nine configurations in total. All

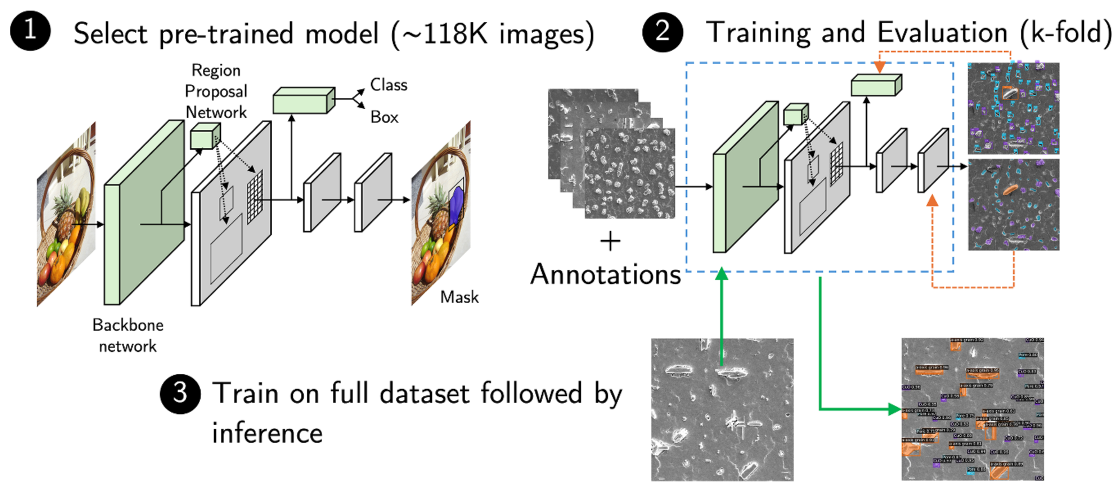


Figure 4. Schematic of the model training, validation and testing pipeline used in this work. Broadly, the steps include (1) selecting the right pre-trained model, originally trained on COCO dataset [20]; (2) Fine-tuning it by training and evaluating on the domain-specific, SEM images coupled with their annotations for several iterations; (3) Training validated model on full dataset followed by inference.

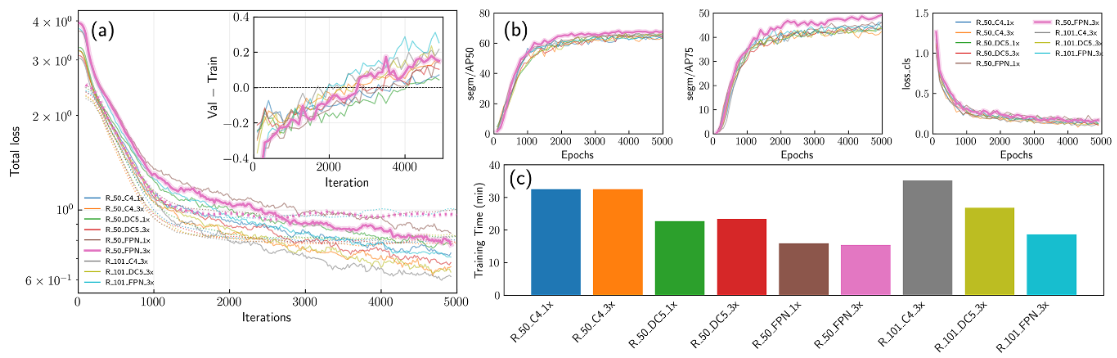


Figure 5. Model backbone ablation results. (a) Total loss curves for each backbone where the solid line indicates the total loss on the training dataset and the dotted line indicates the total loss on the validation dataset across 5000 iterations. The inset shows the difference between validation total loss and training total loss at the iterations where the model was validated. The difference indicates the degree of overfit. (b) Segmentation metrics (AP50 and AP75) and classification loss observed during validation. The selected backbone, `R_50_FPN_3x` is highlighted in plots (a) and (b). (c) Training time for each model backbone for the same number of iterations with the same training configurations.

models were fine-tuned using identical hyperparameters on a fixed training set, with performance evaluated on a corresponding held-out validation set using an 80:20 train–validation split.

Figure 5(a) presents the training and validation loss curves. While all models exhibit broadly similar convergence behavior, a separation emerges toward later training stages. The FPN-based models (`R_50_FPN_1x`, `R_50_FPN_3x`, and `R_101_FPN_3x`) show slightly higher final total loss compared to C4 and DC5 variants. The inset illustrates the generalization gap (validation minus training loss), which remains close to zero for all models, indicating comparable and modest levels of overfitting. No single model stands out as dramatically worse in terms of overfitting. However, loss alone does not fully capture model performance. Figure 5(b) compares detection and segmentation quality using AP metrics (AP50 and AP75), alongside classification loss. Again, no single model consistently dominates across all metrics, highlighting inherent trade-offs. While classification loss is comparable across the backbones tested, `R_50_FPN_3x` stands out as the best-performing backbone for both AP50 and AP75. The FPN variants are also consistently the lightest among those evaluated, with `R_50_FPN_3x` representing the shallowest choice, making it the most suitable option for this segmentation problem with minimal risk of overfitting. This selection is further supported by its computational efficiency, as shown in figure 5(c), where this configuration exhibits the shortest training time among all evaluated models.

Once the model backbone was selected, the model's generalizability was evaluated using k-fold cross-validation. For each fold, splitting of training and validation patches was performed at the sample level, ensuring that no patches originating from the same sample (SEM image) appeared in both the training and validation sets. This prevents data leakage between folds. Specifically, five-fold cross-validation was performed, wherein the 110 patches were partitioned into five approximately equal, non-overlapping folds at the sample level. In each iteration, four folds were used for training and the remaining fold was held out for validation. This procedure yielded five independent train–validation splits, each approximating an 80:20 division of the data, and provided a robust estimate of model generalizability. Although the model was trained for the full 5000 iterations, validation performance was tracked throughout training to identify the optimal stopping point. Table 2 reports the mean and standard deviation of each metric across folds at this best checkpoint, offering an estimate of both the model's expected performance and its sensitivity to the specific composition of the training set. The median number of iterations corresponding to the best checkpoint across folds was 4300.

The bounding box mAP reached 59.69 $\pm$ 1.98% on training data, closely matched by a segmentation mAP of 58.82 $\pm$ 1.74%, indicating balanced performance between localization and mask prediction. Validation mAP followed a similar pattern (44.68 $\pm$ 6.58% bounding box, 45.23 $\pm$ 6.79% segmentation), with substantially higher variance reflecting the smaller size of each validation fold. At lower IoU thresholds, performance was substantially higher (AP50 $\sim$81% training, $\sim$65% validation), reflecting reliable object detection, while stricter overlap criteria (AP75 $\sim$67% training, $\sim$50% validation) confirmed strong boundary precision.

Class-wise performance reflected the underlying class distribution described in table 1. Generally, the model struggled more with smaller objects, as evidenced by the lower APs values, due to the inherent difficulty of detecting fine-scale features in dense microscopy images. Note that large-object AP (APl) could not be reliably computed for the validation folds (indicated by '–'). This is due to the limited number of large-object instances per fold, combined with the sample-level fold splitting resulting in an insufficient number of large-object instances in some validation folds for a stable AP estimate. Cu_xO was consistently the best-performing class across both training and validation (bounding box AP of $\sim$67% for training and $\sim$56% for validation, and segmentation AP of $\sim$68% for training and $\sim$58% for validation), consistent with its dominant representation in the dataset (63% of instances). The *a*-axis grains achieved comparatively high training AP ($\sim$63% for both bounding box and segmentation) despite comprising only $\sim$12% of instances, but showed a drop in validation performance ($\sim$39%–41%), accompanied by elevated variance (std up to $\pm$10.4%), suggesting greater sensitivity to fold composition given its limited sample size. Notably, however, Cu_xO exhibited higher variance than *a*-axis grains despite its larger representation due to differences in object-scale distribution, as the majority of Cu_xO instances are small, which are inherently more difficult to localize and segment precisely, whereas 47% of *a*-axis grain occurrences fall in the medium-to-large range, where detection tends to be more stable. Pores exhibited the lowest AP values in both training ($\sim$45%–49%) and validation ($\sim$36%–39%) despite its moderate representation, likely reflecting greater morphological variability and boundary ambiguity relative to the other classes. Overall, validation metrics closely tracked training metrics across all categories, with the expected increase in variance due to smaller fold sizes, supporting that the model generalizes reasonably well to unseen samples rather than overfitting to the training set.

Although these values may appear modest relative to typical classification or regression metrics, they are consistent with those for instance segmentation tasks. AP is computed at the pixel level across multiple IoU thresholds, making it highly sensitive to even minor discrepancies between predicted and ground truth masks. Such sensitivity is particularly pronounced in low-contrast, multi-class SEM images, where precise delineation is inherently challenging. Comparable AP ranges are commonly reported for complex segmentation tasks on benchmark datasets such as COCO2017 [21]. Furthermore, prior work by Russakovsky *et al* [23] highlights the difficulty of visually distinguishing IoU differences, underscoring the stringency of this evaluation metric.

After confirming validation performance, a production model was trained on the entire set of annotated patches with the same hyperparameters. This model was then used to generate predictions on the SEM images of the 63 samples with confidence threshold at 0.7 (i.e. only detections where the model assigned $\geqslant$70% certainty to a predicted class were retained). Inference on an Apple M3 Pro CPU, 6 cores took approximately 2 s per patch. The predicted bounding boxes and masks for each patch were transformed from local patch coordinates to the full image coordinate system and stitched together. Representative cases are shown in figures 6(a) and (b), where panel (a) corresponds to the same image as figure 2(a) in [12]. These examples clearly demonstrate that the fine-tuned Mask R-CNN model is adept at distinguishing boundaries between closely occurring instances, a task the Weka-based model struggled with. All quantitative analyses were performed in Python. Developing both the modeling framework and the statistical analysis within a single Python environment makes the entire image

Table 2. Performance metrics (in %) of the fine-tuned model at the best checkpoint on the training and validation sets, reported as mean ± standard deviation across 5-fold cross-validation.

Training performance						
	mAP	AP50	AP75	APs	APm	APl
Bounding box	59.69 ± 1.98	81.11 ± 2.18	67.79 ± 2.61	54.18 ± 1.80	81.18 ± 1.62	92.67 ± 2.34
Segmentation	58.82 ± 1.74	80.96 ± 2.05	66.33 ± 2.08	52.85 ± 1.53	80.57 ± 1.29	96.97 ± 1.55
	Per-class AP					
	Pores		Cu_xO		*a*-axis grains	
Bounding box	48.66 ± 2.09		67.36 ± 2.02		63.04 ± 3.43	
Segmentation	45.16 ± 1.90		67.85 ± 1.73		63.45 ± 3.45	
Validation performance						
	mAP	AP50	AP75	APs	APm	APl
Bounding box	44.68 ± 6.58	64.59 ± 7.52	50.39 ± 7.61	40.89 ± 6.78	59.21 ± 13.23	—
Segmentation	45.23 ± 6.79	64.77 ± 7.82	50.00 ± 8.16	40.01 ± 6.80	65.88 ± 12.79	—
	Per-class AP					
	Pores		Cu_xO		*a*-axis grains	
Bounding box	38.61 ± 7.54		56.49 ± 14.20		38.94 ± 9.43	
Segmentation	36.38 ± 7.40		57.93 ± 14.64		41.37 ± 10.44	

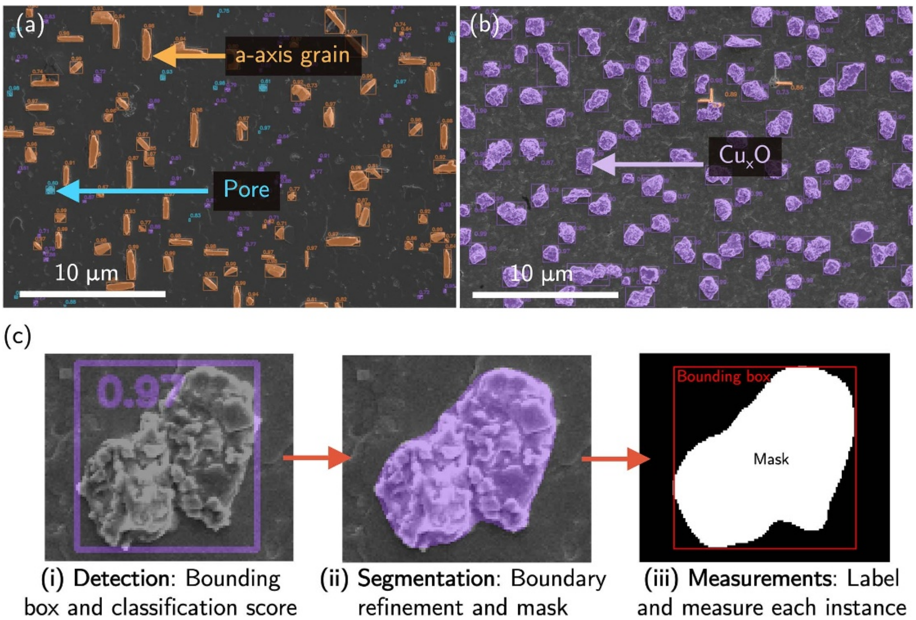


Figure 6. Instance segmentations results on validation set from the final model showing two samples with higher (a) *a*-axis grain and pore density, and (b) Cu_xO particles. (c) Illustration of the key steps in the inference workflow - (i) detect and identify, (ii) segment, and (iii) measure relevant geometric descriptors. Steps (i) and (ii) are part of the Mask-RCNN inference process and step (iii) is implemented using `scikit-image` package in Python. Note that the images shown are representative examples selected to illustrate model performance on each class, and are not necessarily representative of typical feature densities across the dataset.

analysis pipeline easily scalable and reproducible. The scikit-image [24] library was used to post-process and extract measurements of each segmented instance as shown in figure 6(c).

3.2. Microstructural variation across the tape

Quantifying spatial organization of defects is essential for understanding whether they exhibit non-random patterns such as clustering or regular dispersion. To understand this, within the analyzed region of the samples, nearest neighbor analyses were performed on all the detected defects. By calculating the mean of the Euclidean distances from each feature to its nearest neighbors ($\overline{D}_{\mathrm{obs}}$), a nearest neighbor index (NNI) was determined. For a given area of the sample, $NNI = \frac{\overline{D}_{\mathrm{obs}}}{\overline{D}_{\mathrm{exp}}}$, where $\overline{D}_{\mathrm{exp}}$ is the expected average distance for a random distribution of the same number of points in the same area ($\overline{D}_{\mathrm{exp}} = 0.5\sqrt{\text{Area enclosing all defects}/\text{Total number of defects}}$). An NNI less than 1 indicates clustering,

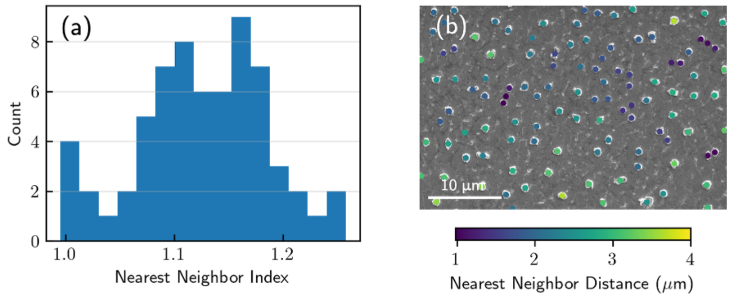


Figure 7. Spatial distribution analysis of segmented defects across all samples using nearest neighbor analysis. (a) Histogram of the nearest neighbor index (NNI) across all analyzed samples, with values greater than 1 indicating a tendency toward random to dispersed spatial arrangements of defects. (b) Representative SEM image marked with the centroids of segmented defects, colored according to their nearest neighbor distance (in μm), illustrating the local spatial distribution of defects across the REBCO layer surface.

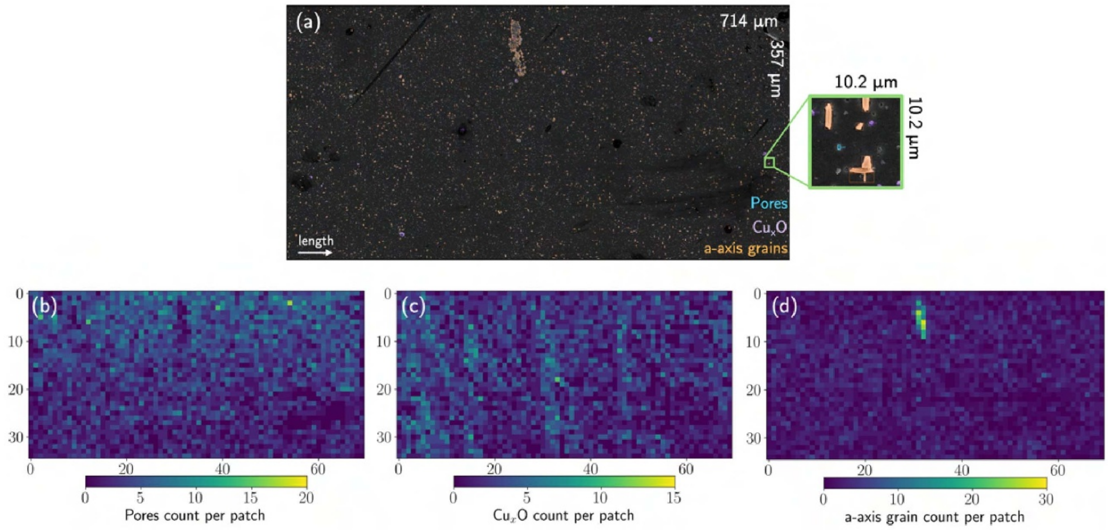


Figure 8. (a) Model predictions on a low-fidelity, automated large-area scan acquired using Thermo Scientific Maps Software, with a magnified view of the marked region shown alongside. Although the individual frames comprising the composite large-area scan have comparable pixel sizes to the training images, the dwell time was approximately 80$\times$ shorter, resulting in reduced signal-to-noise ratio and grainier images. (b)–(d) Spatial frequency maps of segmented defects across the large-area scan. Each pixel in the map corresponds to a single patch in the composite scan, with color intensity representing the number of detected defects per patch. Maps are shown for (b) pores, (c) Cu_xO precipitates, and (d) *a*-axis grains. The spatial distribution reveals clustering and regional variations in defect density across the REBCO surface.

while a value greater than 1 indicates dispersion. A value of exactly 1 represents a random distribution, and values above 2.15 indicate maximum dispersion, typically a perfectly uniform hexagonal spacing [25]. To mitigate edge effects, defects were excluded if their nearest neighbor distance exceeded their distance to the image boundary. This ensures that only defects whose true nearest neighbor lies within the image are retained in the analysis.

Across all samples, NNI values ranged from approximately 0.99 to 1.25, consistent with random to weak spatial dispersion (figure 7(a)). While most values are clearly greater > 1, they remain well below those associated with highly regular spatial arrangements (NNI $\sim$ 2.15 for a perfectly uniform lattice), indicating that the observed defect distributions are only weakly ordered. Statistical analysis confirmed that the weak dispersion is statistically significant ($p\leqslant 0.05$) in 48 of 63 samples. Overall, within individual micrographs, the defect distributions are best described as weakly dispersed rather than strongly ordered. Visual inspection of centroid distributions (a representative case is shown in figure 7(b)) corroborated these findings, showing irregular but slightly spaced defect arrangements.

It should be noted that the analyzed single SEM image per sample only covers an area of 41 μm $\times$ 27 μm which is relatively small for NNI analysis to be representative of the full tape width and length. As a result, localized clusters of defects, if any, may not be fully captured at this scale and would require larger-area scans for proper characterization. Figure 8 shows model predictions on one such automated large-area scan that was taken using Thermo Scientific Maps Software. The scan covers an area of

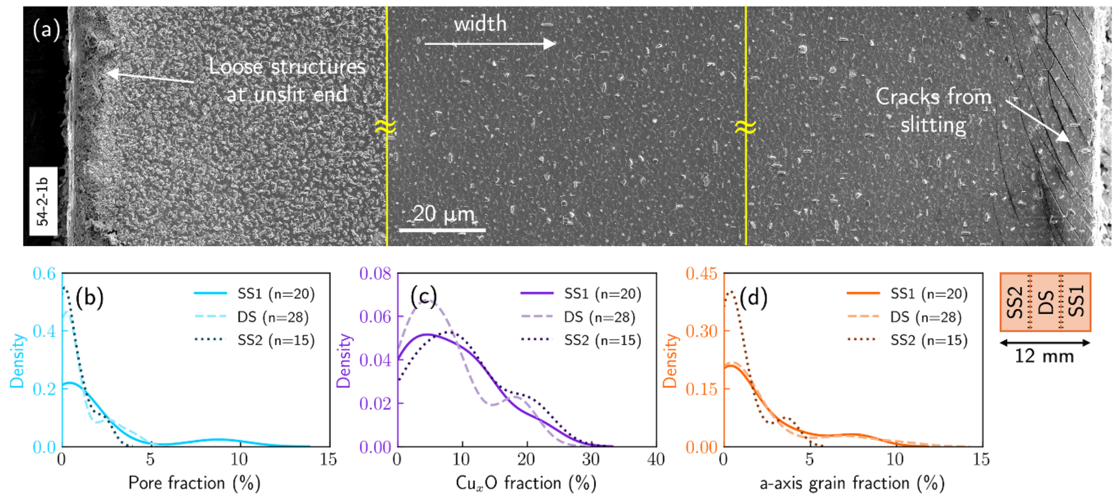


Figure 9. (a) Representative top-surface SEM images of a single-slit sample 54-2-1b across the tape width, showing three regions of interest: high concentration of mixed defects at the unslit edge with some loose structures at the edge (left), the central region (middle) showing typical expected distribution of defects, and mechanical slitting-induced cracks at the slit edge (right). All three images have the same scale of 20 μm. Kernel density estimates of the defect fraction distributions for (b) pore area fraction (%) (c) Cu_xO area fraction (%), and (d) *a*-axis grain area fraction (%), segmented across three slit types: Single-slit tapes SS1 ($n = 20$) and SS2 ($n = 15$), double-slit tapes , DS ($n = 28$), as illustrated schematically in the diagram next to (d). The numbers in the parentheses indicate the number of samples for each slit type.

714 μm $\times$ 357 μm and was divided into 2450 patches of size 10.2 μm $\times$ 10.2 μm. The acquisition time for this scan was approximately 17 min and represents only ∼0.6 % of the 10 mm $\times$ 4 mm section extracted for SEM visualization. The resulting image is of relatively low fidelity, as individual frames within the mapped area were acquired at a lower dwell time as well as at a single focal plane without stacking along the depth, leading to suboptimal focus across the composite. Despite these limitations, the model was able to reasonably segment the defects. Figures 8(b)–(d) show the spatial distribution of defects, obtained by counting the number of defects of each class within each patch. A local cluster of *a*-axis grains, located roughly in the upper-middle region of the image, is captured by the model and highlighted in figure 8(d). Beyond this cluster, the defects are distributed fairly evenly throughout the imaged area.

The variation across the full 12 mm tape width was studied in two parts: (1) differences between 4 mm slit sections, and (2) variations within each slit section across its width. Multiple MOCVD tapes with a width of 12 mm were fabricated and subsequently slit into three 4 mm sections. The 63 samples analyzed in this study were randomly selected 4 mm wide REBCO tapes obtained from this set of slit tapes. These samples were categorized as single-slit (SS1 and SS2) or double-slit (DS). SS1 and SS2 correspond to the outer sections of the original 12 mm tape (each having one slit edge and one original edge), while DS refers to the middle section (bounded by two slit edges). Figure 9(a) shows representative images from an SS2 tape. Visual inspection indicates that, in single-slit samples, the unslit (original) edge exhibits a higher concentration of defects, whereas closer to the slit edge shows a defect distribution similar to that of the tape interior. The slit edges had cracks characteristic of the mechanical slitting process.

Figure 9(b) shows kernel density estimation (KDE) plots depicting the distribution of defect area fraction, defined as the percentage of the total REBCO tape area within an SEM image occupied by a given defect type, for SS1, DS, and SS2 tapes. Since these defect area fractions were analyzed from SEM images acquired approximately in the middle of each tape, edge effects were not captured, and the defect distributions appear similar regardless of the type of slit. The distributions indicate that most tapes have low pore and *a*-axis grain area fractions. SS2 showed the greatest variability in total defect fraction and in-field critical current. Pore fraction remained similarly distributed across all slit orientations, while *a*-axis grain fraction was least variable in SS2 and DS samples exhibited the greatest variability in Cu_xO fraction.

To better understand how the defect concentration stabilizes in single-slit tapes with increasing distance from the edge, multiple images were acquired along the width of three SS2 tapes. Figure 10 shows the variation of each defect category as a function of distance from the unslit edge for three SS2 samples. The concentration of defects, primarily Cu_xO and *a*-axis grains, is significantly elevated within

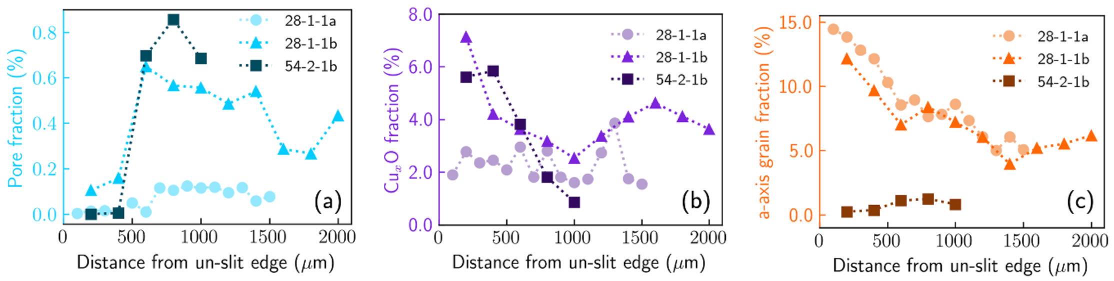


Figure 10. Spatial variation of defect fractions as a function of distance from the unslit edge for three SS2 samples (`28-1-1a`, `28-1-1b`, and `54-2-1b`): (a) pore area fraction (%), (b) Cu_xO area fraction (%), and (c) *a*-axis grain area fraction (%). Each data point represents the average defect fraction extracted from a single image at the corresponding position across the tape width. Sample `54-2-1b` corresponds to the same sample shown in figure 9(a).

approximately 800 μm of the unslit edge, beyond which it stabilizes. This is attributed to poor temperature control at the tape edges during the MOCVD process. This edge-localized defect concentration also has important implications for magnet design, as reported in [26], where tapes of the SS1/SS2 type exhibited a counterintuitive finding via magneto-optical imaging. Despite the presence of edge microcracks in single-slit tapes, the critical current density was lower at the unslit edge, reflecting the dominant role of microstructural defects over mechanical cracking in that region.

Since this elevated defect concentration was observed only within approximately 1 mm of the tape unslit edge and since the mid-section is representative of the bulk microstructural characteristics as shown by the KDE plots in figures 9(b)–(d), all subsequent analyses considered only images acquired away from the edges.

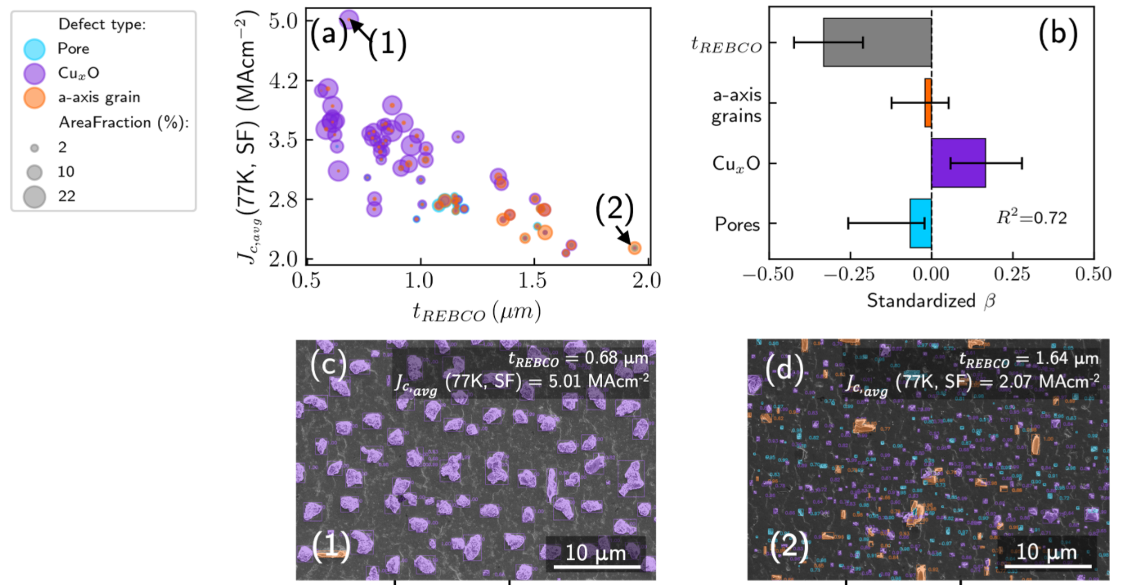
3.3. Effect of microstructure on critical current

Using the methodology described in figure 4, geometrical descriptors were extracted from all segmented defects and subjected to statistical analysis to explore their potential correlations with the average critical current density ($J_{c,avg}$). The discussion is divided into two parts, covering performance at 77 K under self-field conditions and at 4.2 K under a 14 T background field, at two orientations relative to the sample plane (90° and 18°). The primary descriptor used throughout this discussion is the area fraction (or defect fraction). Additional descriptors, such as the total count of each defect type and the major axis length of the best-fit ellipse around each defect as a proxy for defect size, can also be readily extracted from the segmented images.

3.3.1. Transport $J_{c_{avg}}$ at 77 K under self-field

Figure 11(a) shows the average critical current density ($J_{c,avg}$) measured at 77 K under self-field as a function of REBCO layer thickness (t_{REBCO}) for all analyzed samples. Several trends emerge from the data. The marker size and color encode defect information derived from the corresponding SEM image, where the size of the marker is proportional to the area fraction. The dominant defect in most samples is Cu_xO (purple markers), which is present throughout the full thickness range. At larger thicknesses, a notable increase of *a*-axis grain area fraction (orange markers) is observed, suggesting that crystallographic misalignment becomes increasingly prevalent in thicker films which also explains the degradation of $J_{c,avg}$. Pores (cyan markers) appear relatively minor across most samples but are more prevalent in the mid-range of REBCO thickness. Thinner films (around 0.5–0.7 μm) achieved the highest critical current densities of up to ∼5 $MAcm^{-2}$, while thicker films approaching 2.0 μm show significantly reduced values, often falling below 2.5 $MAcm^{-2}$. This thickness dependence is a well-known phenomenon in coated conductors, commonly attributed to increased defect accumulation and microstructural degradation in thicker films [27].

The collective influence of the defect fractions and t_{REBCO} (X) on the critical current density (Y) was quantified using multiple linear regression ($Y = \beta X + \varepsilon$). The standardized regression coefficient (β) and corresponding 95 % confidence intervals are shown in figure 11(b). The β values for each predictor indicates their relative influence. A positive correlation is observed with Cu_xO, while pores and *a*-axis grains show marginal inverse correlations. Among the predictors considered, REBCO thickness exhibits the strongest, and notably negative, influence. The coefficient of determination, R^2 indicates that these four predictors collectively explain 72 % of $J_{c,avg}$ (77 K, SF). Two samples are highlighted in figure 11(a): point (1), a sample with exceptionally high $J_{c,avg}$ (∼5 $MAcm^{-2}$) at low thickness, and point (2), which represents a notably thick film among the lowest $J_{c,avg}$ values (∼2 $MAcm^{-2}$). The corresponding area



Pore	Cu$_x$O	a-axis grain
0%	17.8%	0.1%

Pore	Cu$_x$O	a-axis grain
0.8%	2.7%	1.9%

Figure 11. (a) Scatter plot of the average critical current density, $J_{c,avg}$ (77 K, SF), as a function of REBCO layer thickness, t_{REBCO}, for all samples. Each data point represents an individual sample, with the size of the overlaid circles scaled proportionally to the area fraction of pores (cyan), Cu$_x$O (purple), and *a*-axis grains (orange). A general trend of decreasing $J_{c,avg}$ (77 K, SF) with increasing t_{REBCO} is observed across the dataset. (b) Standardized regression coefficients (β) from multiple linear regression of $J_{c_{avg}}$ (77 K, SF) on defect fractions and t_{REBCO} along with error bars represent 95 % confidence intervals and the coefficient of determination R^2. Two samples are identified in (a), labeled (1) and (2), corresponding to thinner, higher-performing samples and thicker, lower-performing samples, respectively. (c) Sample (1) marked in (a) with the highest $J_{c,avg}$ (77 K, SF) and (d) Sample (2) marked in (a) with the lowest $J_{c,avg}$ (77 K, SF), along with their defect area fractions (%).

that was analyzed within the SEM image is shown for each sample along with the calculated defect fraction in figures 11(c) and (d). Sample (1) shows only Cu$_x$O while Sample (2) has little to no Cu$_x$O but more *a*-axis grains than the former. These extremes bracket the performance envelope of the dataset and may warrant further microstructural investigation given the positive correlation observed between Cu$_x$O and $J_{c,avg}$.

To further break down these observations and identify possible groupings, *K*-means clustering was applied to samples based on the five key features plotted in figures 11(a) and (b): t_{REBCO}, $J_{c,avg}$ (77 K, SF), and the area fractions of Cu$_x$O, pores, and *a*-axis grains. Similar to figure 11(b), this approach considers all features simultaneously, allowing samples with similar combinations of microstructural and critical current to be naturally grouped together. This is particularly useful for analyzing samples where multiple defect types coexist and may interact in their influence on $J_{c,avg}$, making it difficult to isolate the effect of any single variable. Prior to clustering, the features were standardized so that each variable contributed equally. *K*-means was then performed in this standardized feature space, partitioning the samples into three clusters by minimizing within-cluster variance. To visualize the results, principal component analysis (PCA) was used to project the five-dimensional feature space onto two dimensions, with the cluster decision boundaries overlaid. It should be noted that the clustering itself was performed in the full feature space, and PCA was used solely as a visualization tool. By examining how the distributions of each feature differ across clusters, one can identify which combinations of microstructural features are associated with high or low $J_{c,avg}$, and whether certain defect populations tend to coexist. As seen in figure 12, all the samples sort themselves into three distinct regions. The black cluster corresponds to low thickness ($\sim$0.5 μm – 1 μm), high $J_{c,avg}$ (77 K, SF), low pore and *a*-axis grain area fraction but high Cu$_x$O area fraction. The orange cluster corresponds to high thickness ($>$ 1.5 μm) , low $J_{c,avg}$ (77 K, SF), few pores and Cu$_x$O, high occurrence of *a*-axis grains. Finally, the yellow cluster corresponds to moderate thickness ($\sim$1.2 μm – 1.5 μm), moderate $J_{c,avg}$ (77 K, SF), and some mixed defects. This grouping is also explained by the arrows (called variable loadings) overlaid on the plot. The length of each arrow reflects the strength of a variable's contribution to the principal components, while its direction indicates its association with the observed clusters. The correlation between any two variables can be inferred from the angle between their respective arrows. Smaller angles indicate positive correlation, right angle indicates no correlation, and opposing directions indicate negative correlation. The arrows

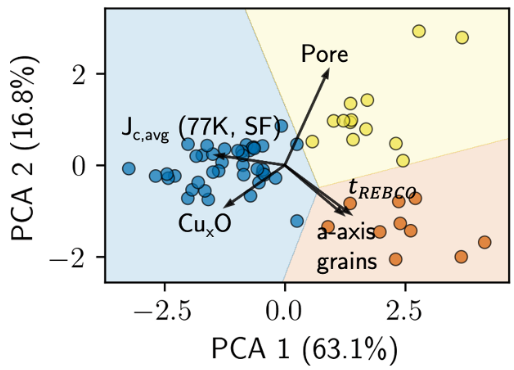


Figure 12. Visualization of sample grouping based on five variables using PCA and clustering. Clustering reveals patterns among samples based on $J_{c,avg}$ (77 K, SF), REBCO thickness, and the three defect area fractions. Arrows indicate the influence of each variable and the angle between two indicates the correlation between them. Values in parentheses indicate the percentage of variance explained by each principal component.

for *a*-axis grain fraction and REBCO thickness (t_{REBCO}) are closely aligned, confirming that these two quantities are strongly positively correlated. Both arrows point away from the $J_{c,avg}$ arrow, suggesting that an increase in *a*-axis grain fraction as a result of greater t_{REBCO} layer thickness is associated with suppressed $J_{c,avg}$ (77 K, SF). The pore fraction arrow is approximately orthogonal to $J_{c,avg}$ (77 K, SF), indicating that these two variables are largely uncorrelated. Finally, the acute angle between the $J_{c,avg}$ (77 K, SF) and Cu_xO arrows indicates a positive correlation between these variables, suggesting that samples with higher Cu_xO fraction tend to exhibit higher critical current density under these conditions. While it is well established that *a*-axis grains obstruct current flow and thereby degrade superconducting performance, the observed superior performance of thinner samples with a significant Cu_xO area fraction remains less straightforward to interpret. Whether the contribution to higher $J_{c,avg}$ is merely a result of thickness and Cu_xO has no impact cannot be determined simply from the surface analyzed in SEM images. This phenomenon has also been discussed in prior studies that reported Cu_xO precipitates up to 1 μm in diameter commonly distributed throughout YBCO thin films without any apparent impact on electrical performance [28]. Notably, among the sputtered YBCO thin films studied by Westerheim *et al* [28], the best-performing films tended to have the largest precipitates. However, the resulting surface roughness can pose challenges for multilayer deposition.

3.3.2. Transport $J_{c_{avg}}$ at 4.2 K in 14 T background field

Similar statistical analyses were performed to investigate the potential impact of defects on $J_{c,avg}$ at 4.2 K under a background field of 14 T, at field orientations of 90° and 18° relative to the sample plane. Interestingly, looking at figures 13(a) and (b), unlike the $J_{c,avg}$ at 77 K, self-field, the thickness of the REBCO layer appears to be uncorrelated to the tape's electrical performance at 4.2 K with background field. The performance was negatively affected by increasing number of defects. No distinct clustering was observed via *K*-means, and the PCA plots are therefore not shown. The results of the multiple linear regression are shown in panels (c) and (d). Here, the four predictors, three defect fractions and t_{REBCO}, explain only 43 % and 36 % of in-field $J_{c,avg}$ at the two respective orientations. Comparatively, the higher explanatory power at 77 K SF suggests that the measured defect types and layer thickness are the primary governing factors of self-field performance, while the lower R^2 values at 14 T indicate that additional microstructural factors not captured by the current SEM characterization play a more important role under high applied fields. The individual β coefficients carry considerable uncertainty and should not be over-interpreted here due to possible multicollinearity between the predictors which will be discussed later.

Examining the samples with the highest and lowest in-field $J_{c,avg}$ in figures 13(e) and (f), the microstructures of samples (3) and (4) present strikingly contrasting characteristics. Sample (3), exhibiting the highest in-field $J_{c,avg}$, contains a significant number of *a*-axis grains, whereas sample (4), with the lowest $J_{c,avg}$, contains almost none but instead has a considerably higher proportion of Cu_xO. Whether the *a*-axis grains in sample (3) extend through the full thickness of the REBCO layer to impede current pathways can only be conclusively assessed through cross-sectional imaging which is beyond the scope of this work. Additionally, surface SEM imaging also cannot resolve the depth distribution of Cu_xO precipitates. The observed inverse relationship in sample (4) may in part reflect a detection limitation rather than a true reduction in Cu_xO formation with increasing thickness, since precipitates formed during

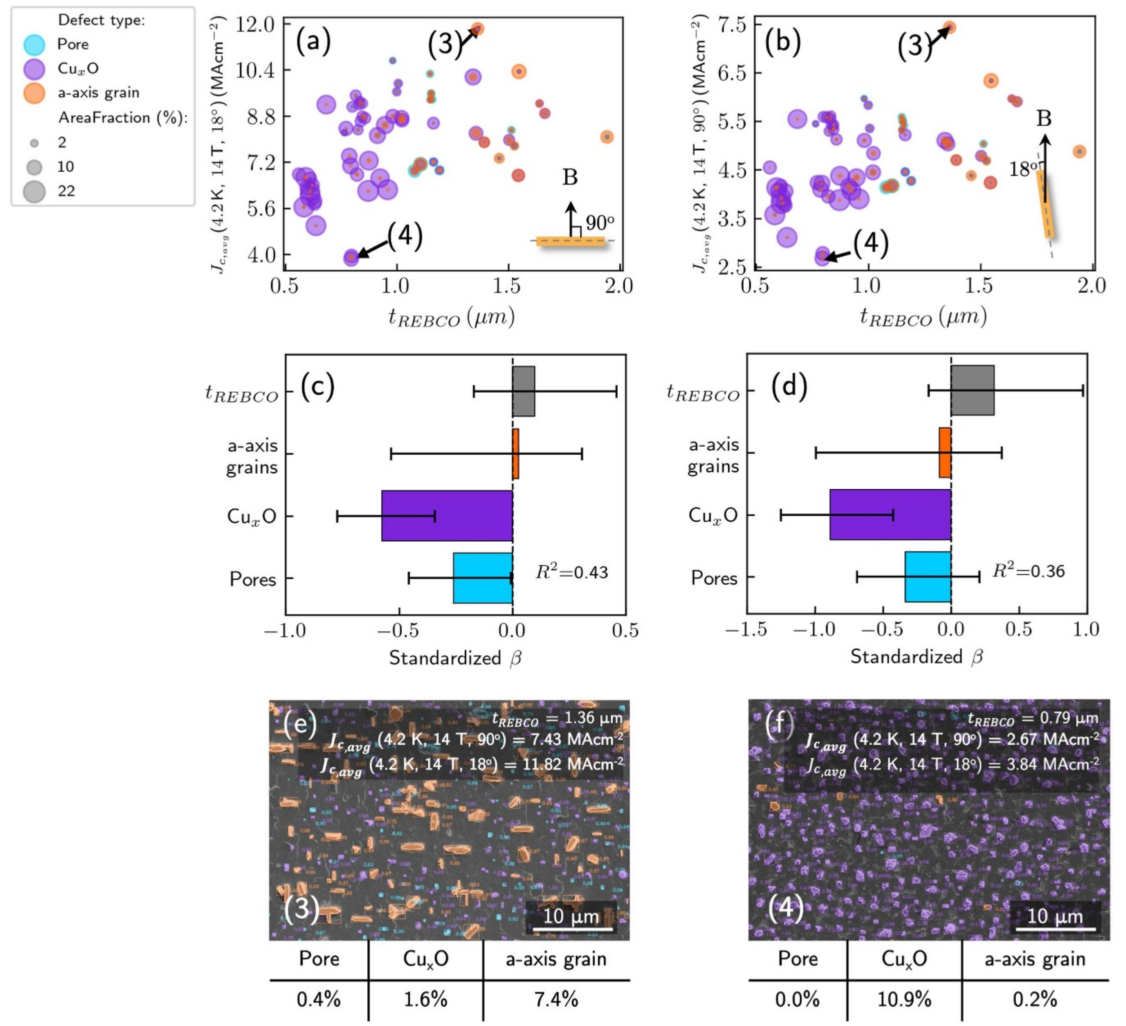


Figure 13. Scatter plot of the average critical current density, $J_{c,avg}$, measured at 4.2 K with a background field of 14 T at (a) 90°, and (b) 18° to the sample plane, as a function of REBCO layer thickness t_{REBCO}, for all samples. Each data point represents an individual sample, with the size of the overlaid circles scaled proportionally to the area fraction of pores (cyan), Cu$_x$O (purple), and *a*-axis grains (orange). No clear trend between $J_{c,avg}$ and t_{REBCO} is observed across the dataset. Notably, samples clustering in the lower-left quadrant with larger scatter point sizes suggest a higher proportion of defects that are detrimental to $J_{c,avg}$. Standardized regression coefficients (β) from multiple linear regression of (c) $J_{c,avg}$ (4.2 K, 14 T, 90°) and (d) $J_{c,avg}$ (4.2 K, 14 T, 18°) on defect fractions and t_{REBCO}. Two samples are highlighted in (a): Sample (3) shown in (e) with the highest in-field $J_{c,avg}$ and Sample (4) shown in (f) with the lowest in-field $J_{c,avg}$, along with their defect area fractions (%).

early growth stages could become buried beneath subsequently deposited REBCO material in thicker films and thus escape detection in top-view imaging. Distinguishing between a genuine reduction in Cu$_x$O density and this geometric burial effect would require cross-sectional characterization, such as FIB-SEM or transmission electron microscopy (TEM).

3.4. Pairwise and partial correlation analysis

Combining all model predictions, statistical correlations between defect fractions and $J_{c,avg}$ were computed. Since defect fraction data is skewed and not a normal distribution, as shown in figures 14(a)–(c), the Spearman rank correlation coefficient (ρ) was used. Spearman's ρ measures the strength and direction of monotonic associations between ranked variables, ranging from –1 to 1, where the magnitude corresponds to correlation strength: very weak for $|\rho| < 0.2$, weak for $0.2 \leqslant |\rho| < 0.4$, moderate for $0.4 \leqslant |\rho| < 0.6$, strong for $0.6 \leqslant |\rho| < 0.8$, and very strong for $|\rho| \geqslant 0.8$. It is worth noting that $\rho = 0$ does not necessarily imply the absence of a relationship, only the absence of a monotonic one. Figure 14(d) shows the pairwise Spearman's ρ values across all variables, with *p*-values indicated by asterisks denoting statistical significance.

Looking at each category of defect, pore fraction shows a strong negative correlation with $J_{c,avg}$(77 K, SF) ($\rho = -0.69$), with a sign reversal and moderate correlation observed for in-field measurements ($\rho = 0.49$ at both 90° and 18°, respectively). A moderate positive correlation with t_{REBCO} ($\rho = 0.57$)

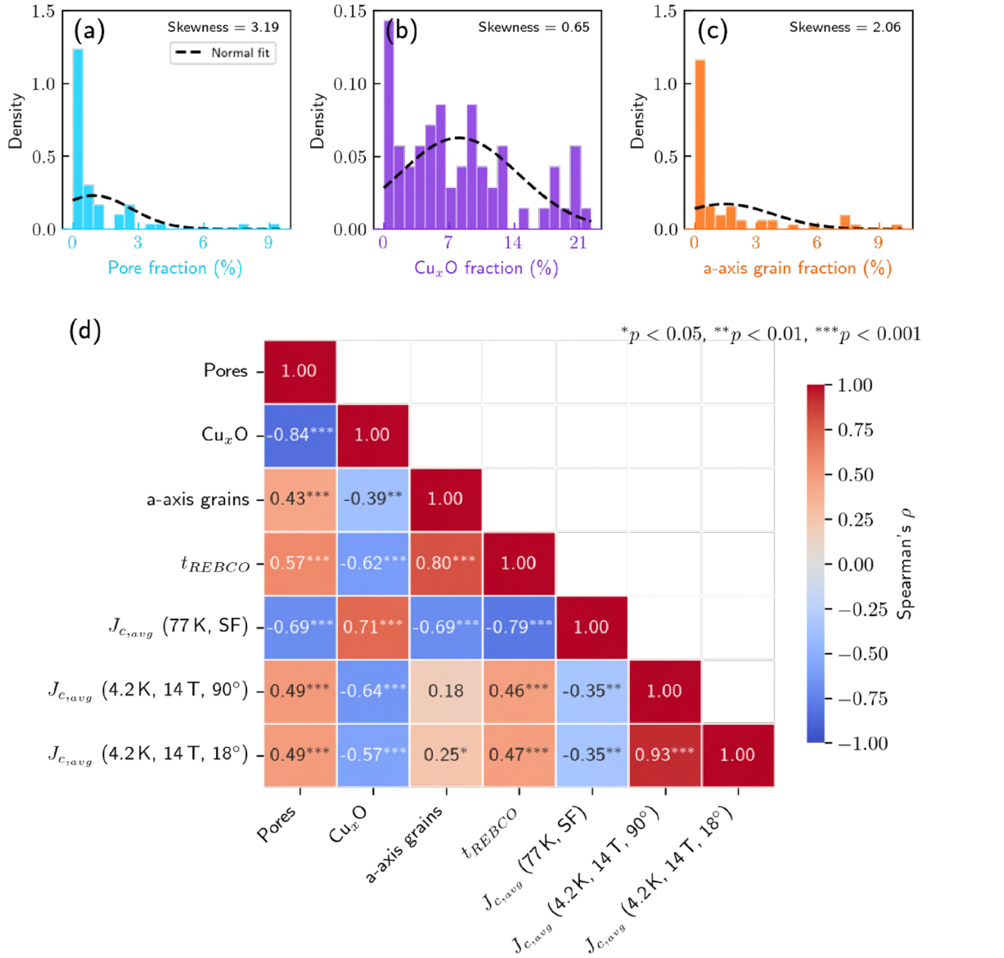


Figure 14. Statistical analysis of defect fractions and their correlations with critical current density. Probability density distributions of (a) pore fraction (%), (b) Cu_xO fraction (%), and (c) *a*-axis grain fraction (%) across the sample set, with skewness values annotated for each distribution. A normal distribution fit (dashed line) is overlaid for reference. Note that *y*-axis values represent probability density and can exceed 1, as bar heights are scaled such that the total area under each histogram integrates to 1. All three distributions exhibit positive skewness, indicating non-normal distributions with the fraction of pores and *a*-axis grains more skewed toward lower values. (d) Spearman correlation matrix illustrating pairwise correlations between the defect fractions, REBCO thickness, and average critical current density. Asterisks denote varying degree of statistical significance (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

is also observed. The *a*-axis grain fraction is strongly correlated with layer thickness ($\rho = 0.80$), consistent with the trends observed in figure 12, and shows a strong negative correlation with $J_{c,avg}$(77 K, SF) ($\rho = -0.69$). At 14 T, only weak positive correlations are observed ($\rho = 0.18$ and $\rho = 0.25$) at 90° and 18°, with the former not being statistically significant. For Cu_xO fraction, a strong positive correlation is observed with $J_{c,avg}$(77 K, SF) ($\rho = 0.71$), while strong and moderate negative correlations are found with $J_{c,avg}$(4.2 K, 14 T, 90°) and $J_{c,avg}$(4.2 K, 14 T, 18°) ($\rho = -0.64$ and $\rho = -0.57$, respectively), consistent with the trends discussed previously. A negative correlation with t_{REBCO} ($\rho = -0.62$) is also noted.

However, interpreting these pairwise correlations as calculated is misleading due to significant confounding with other variables. In particular, the pores fraction and the Cu_xO fraction exhibit a very strong negative correlation ($\rho = -0.84$), and all defect fractions are correlated to various degrees with t_{REBCO}. These inter-dependencies make it impossible to attribute observed $J_{c_{avg}}$ trends to any single defect type based on pairwise analysis alone. To isolate the independent contribution of each defect type, partial correlations were computed (using `Pingouin` [29] in Python), controlling simultaneously for t_{REBCO} and all other defect fractions. This addresses key questions such as: among samples with similar thickness and similar pore and *a*-axis grain fractions, do samples with higher Cu_xO fraction tend to exhibit higher or lower $J_{c_{avg}}$? The results are summarized in table 3 along with the corresponding *p*-values.

After controlling for confounding variables, pore fraction shows no statistically significant independent correlation with $J_{c,avg}$ under any measurement condition ($p > 0.05$ throughout). This suggests that its pairwise correlations were driven primarily by its strong negative correlation with Cu_xO fraction and its correlation with thickness, rather than by a direct effect on current-carrying performance. For *a*-axis

Table 3. Partial correlation of defects with J_c after controlling for t_{REBCO} and other defects.

Defect	J_c (77 K, SF)	J_c (4.2 K, 14 T, 90°)	J_c (4.2 K, 14 T, 18°)
Pores	$\rho = -0.17$ ($p = 0.186$)	$\rho = -0.08$ ($p = 0.518$)	$\rho = -0.02$ ($p = 0.856$)
Cu_xO	$\rho = 0.25^*$ ($p = 0.047$)	$\rho = -0.38^{**}$ ($p = 0.001$)	$\rho = -0.23$ ($p = 0.073$)
a-axis grains	$\rho = -0.26^*$ ($p = 0.037$)	$\rho = -0.28^*$ ($p = 0.027$)	$\rho = -0.18$ ($p = 0.163$)

grain fraction, weak negative partial correlations remain statistically significant at 77 K, SF ($\rho = -0.26$, $p = 0.037$) and at 4.2 K, 14 T, 90° ($\rho = -0.28$, $p = 0.027$), while no significant correlation is observed at 4.2 K, 14 T, 18° ($\rho = -0.18$, $p = 0.163$).

Cu_xO fraction continues to show an independent effect on $J_{c,avg}$, though this effect is not uniformly significant across both in-field conditions. At 77 K, SF, Cu_xO fraction shows a weak positive partial correlation ($\rho = 0.25$, $p = 0.047$). At 4.2 K, 14 T, 90°, Cu_xO fraction is independently detrimental, with a moderate negative partial correlation ($\rho = -0.38$, $p = 0.001$). However, at 4.2 K, 14 T, 18°, the partial correlation is weaker and no longer statistically significant ($\rho = -0.23$, $p = 0.073$). This pattern, weakly beneficial at self-field but detrimental at high perpendicular field, represents a measurable independent effect at two of the three tested conditions, though it should be interpreted cautiously given the loss of significance at 18° and the possibility of unmeasured variables co-varying with Cu_xO fraction.

4. Conclusions

An end-to-end instance segmentation framework was developed to automate defect detection in REBCO coated conductors from SEM images spanning multiple production batches. Built on a fine-tuned Mask R-CNN architecture, the segmentation model achieved an average mAP of 45.23 % on the validation folds. Analysis across 63 tapes revealed that multiple defect types coexist within each sample, with Cu_xO particles constituting the dominant defect population, followed by *a*-axis grains and minimal porosity, though the dominant defect type shifts with REBCO layer thickness. The ML-based approach demonstrated clear advantages over conventional methods, increasing detection capabilities, accelerating qualification, and enabling scalable, quantitative defect analysis.

Correlation analysis between defect populations and critical current density revealed negative correlations with pore fraction and *a*-axis grain fraction, and a positive correlation between Cu_xO content and $J_{c,avg}$ measured at 77 K in self-field. Since Cu_xO is non-superconducting, this correlation is unlikely to be intrinsic and may instead arise from co-variation with other microstructural features. Tapes with higher Cu_xO content tended to exhibit lower fractions of *a*-axis grains and reduced pore density, both of which correlated negatively with $J_{c,avg}$ (77 K, SF) and are known to degrade performance by disrupting connectivity. This combination of factors likely contributes to the observed positive correlation. To isolate the individual contribution of each defect type, partial correlation analysis was performed. A weakly positive correlation between Cu_xO particle density and $J_{c,avg}$ (77 K, SF) persisted even after accounting for these confounding factors. This trend, which was also qualitatively visible in the SEM images, lacks a clear mechanistic explanation. Additionally, the possibility of unconsidered variables co-varying with Cu_xO cannot be excluded. This suggests that Cu_xO density may be acting as a proxy for an underlying microstructural feature that influences J_c, rather than contributing directly. In contrast, at 4.2 K, 14 T, Cu_xO fraction shows a statistically significant independent negative effect on $J_{c,avg}$ at 90°, though this effect weakens and is no longer significant at 18°, suggesting the detrimental in-field effect may be angle-dependent. The *a*-axis grain fraction similarly shows a weak but significant negative correlation at self-field, and in-field measurements at 90°, with no significant effect at 18°. Pore fraction shows no independent effect under any condition tested.

The observed correlations and their statistical significance may reflect genuine physical effects or statistical artifacts arising from the specific samples considered here, as well as noise in the predictions. Distinguishing between these possibilities would require further studies with larger sample sets. Moreover, this study characterizes defect populations using a single SEM image per tape, acquired at the mid-length, mid-width location . Although each image contains a substantial number of defect instances, yielding reasonably well-sampled local statistics, it does not capture potential longitudinal variation in defect density that could arise from MOCVD process fluctuations along longer tape lengths. Extending this framework to multiple locations per tape, or to reel-to-reel imaging, represents a natural next step for establishing the extent of within-tape defect variability, and is now considerably more tractable given that manual annotation has been replaced by the automated detection pipeline demonstrated here.

Additionally, on the imaging side, top-surface SEM imaging is inherently limited in its ability to capture the complex three-dimensional morphology of defects extending through the REBCO layer thickness, as well as their subsequent effect on superconducting properties. A complete characterization of these features therefore requires cross-sectional imaging using FIB and TEM. Even so, surface-level analysis can provide valuable guidance on which samples and regions warrant further investigation via these more intensive methods, rather than requiring an exhaustive search across the full sample set. Such cross-sectional analysis is currently being pursued as a future extension of this work. The segmentation model thus serves as an efficient initial screening tool.

Looking ahead, the broader vision of this work extends toward a more comprehensive and community-driven framework for ML-based microstructural characterization of REBCO conductors. A key step in this direction is the establishment of an open, anonymized dataset aggregating SEM images acquired across different instruments and imaging conditions, which would significantly improve the generalizability and robustness of trained models across fabrication routes. The longer-term objective is to develop a user-friendly interface integrating the ML model into a unified workflow that streamlines image labeling, model retraining, dataset evaluation, and statistical analysis. Such a platform would substantially reduce the time and expertise currently required to extract quantitative microstructural data, making comprehensive defect characterization more accessible to the broader superconductor research community and enabling richer, more systematic use of existing image acquisition infrastructure.

Acknowledgments

The work performed at Lawrence Berkeley National Laboratory was supported by the Office of High Energy, through the US Magnet Development Program, and Nuclear Physics, U. S. Department of Energy, under Contract No. DE-AC02-05CH11231. The work performed at National High Field Magnet Laboratory was supported by National Science Foundation Cooperative Agreement No. DMR-1644779, DMR-1938789, DMR-2131790, and the State of Florida. This manuscript has been authored by authors at Lawrence Berkeley National Laboratory under Contract No. DE-AC02-05CH11231 with the U.S. Department of Energy. The U.S. Government retains, and the publisher, by accepting the article for publication, acknowledges, that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for U.S. Government purposes.

Data availability statement

A demo of the model can be accessed at https://huggingface.co/spaces/nmenon97/REBCO-defect-detection.

The data cannot be made publicly available upon publication because they contain commercially sensitive information. The data that support the findings of this study are available upon reasonable request from the authors.

Author contributions

Nandana Menon 0000-0001-8701-0436
Formal analysis (lead), Investigation (lead), Methodology (lead), Software (lead), Validation (lead), Visualization (lead), Writing – original draft (lead), Writing – review & editing (equal)

Jean-Francois Croteau 0000-0002-3486-8401
Conceptualization (lead), Formal analysis (supporting), Investigation (supporting), Methodology (equal), Project administration (equal), Resources (equal), Supervision (equal), Writing – review & editing (equal)

Dmytro Abraimov 0000-0002-0587-4917
Data curation (lead), Formal analysis (supporting), Funding acquisition (supporting), Investigation (supporting), Methodology (supporting), Project administration (supporting), Resources (supporting), Supervision (supporting), Writing – review & editing (supporting)

Hannah Matos-Pimentel 0000-0002-5053-3003
Data curation (supporting), Formal analysis (supporting), Investigation (supporting), Methodology (supporting)

Yavuz Oz
Data curation (supporting)

Ian Pong 0000-0002-8996-4249
Conceptualization (supporting), Project administration (supporting), Resources (supporting), Supervision (supporting), Writing – review & editing (supporting)

Carina Zha
Data curation (supporting)

Nicole Bishop
Data curation (supporting)

Jozef Kvitkovic 0000-0001-5369-821X
Data curation (supporting), Writing – review & editing (supporting)

Jun Lu 0000-0001-8521-489X
Data curation (supporting)

Jeremy Levitan 0000-0002-4149-5199
Data curation (supporting)

References

[1] Zhao Y, Wu Y, Goyal A, Huhtinen H, Paturi P and Tsuchiya Y 2025 Commercial compact fusion triggered REBCO tape industry: pulsed laser deposition technology opportunities and challenges *Superconductivity* **15** 100188
[2] Majkic G, Galstyan E and Selvamanickam V 2015 High performance 2G-HTS wire using a novel MOCVD system *IEEE Trans. Appl. Supercond.* **25** 1–4
[3] Chen Z, Kametani F, Chen Y, Xie Y, Selvamanickam V and Larbalestier D C 2009 A high critical current density MOCVD coated conductor with strong vortex pinning centers suitable for very high field use *Supercond. Sci. Technol.* **22** 055013
[4] Galstyan E, Pratap R, Majkic G, Kochat M, Mohan V and Selvamanickam V 2019 Correlation between microstructure and in-field performance of Zr-added REBCO coated conductors made by advanced MOCVD *IEEE Trans. Appl. Supercond.* **29** 1–6
[5] Murakami M, Yamaguchi K, Fujimoto H, Nakamura N, Taguchi T and Koshizuka N 1992 Flux pinning by non-superconducting inclusions in melt-processed YBaCuO superconductors *Cryogenics* **32** 930–5
[6] Li P, Abraimov D, Xu A and Larbalestier D 2011 Observation of important current-limiting defects in a recent high pinning force MOCVD IBAD-MgO coated conductor *Supercond. Sci. Technol.* **25** 025002
[7] Yang Z, Song P, Feng F and Qu T 2025 Critical current degradation of REBCO coated conductor tapes caused by multiple cracks under tensile loading *Superconductivity* **15** 100185
[8] Hartnett W N, Ramirez J, Olson T E R, Hopp C T, Jewell M C and Knoll A R 2021 Characterization of edge damage induced on REBCO superconducting tape by mechanical slitting *Eng. Res. Express* **3** 035007
[9] Abraimov D, Kvitkovic J, Oz Y, Lu J, Levitan J and Bradford G 2023 Origins of low-temperature in-field Jc variability in modern MOCVD ReBCO tapes *16th European Conf. on Applied Superconductivity (EUCAS) (Bologna, Italy)* (available at: https://nationalmaglab.org/library/presentations/NHMFL_Presentation-10885.pdf)
[10] Huang D, Zheng S, Gao P and Wang X 2026 Deep learning approach for defect identification and current distribution reconstruction in REBCO-coated conductors based on magnetic field mapping *Superconductivity* **17** 100220
[11] Bagni T, Mauro D, Majkut M, Rack A and Senatore C 2022 Formation and propagation of cracks in RRP Nb3Sn wires studied by deep learning applied to x-ray tomography *Supercond. Sci. Technol.* **35** 104003
[12] Matos-Pimentel H, Zha C, Bishop N, Mao K, Kvitkovic J and Lu J 2024 Use of SEM/FIB and Machine Learning to Characterize REBCO Conductors *Microsc. Microanal.* **30** ozae044217
[13] Arganda-Carreras I, Kaynig V, Rueden C, Eliceiri K W, Schindelin J and Cardona A 2017 Trainable Weka Segmentation: a machine learning tool for microscopy pixel classification *Bioinformatics* **33** 2424–6
[14] Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M and Pietzsch T 2012 Fiji: an open-source platform for biological-image analysis *Nat. Methods* **9** 676–82
[15] Croteau J F, Vallone G, Menon N, D'Addazio M, Niccoli F and Pong I 2025 Crack identification and characterization in deformed Nb3Sn Rutherford cable stacks using machine learning *IEEE Trans. Appl. Supercond.* **35** 1–7
[16] Bai H, Bird M D, Cooley L D, Dixon I R, Kim K L and Larbalestier D C 2020 The 40 T superconducting magnet project at the national high magnetic field laboratory *IEEE Trans. Appl. Supercond.* **30** 1–5
[17] Abraimov D, Gillman J, Zha C, Oz Y, Pimentel H and Mao S 2025 Optimization of transport critical currents at 4.2 K – 20 K at magnetic fields up to 31 T for MOCVD REBCO conductors with variable Zr and growth conditions *IEEE Trans. Appl. Supercond.* **35** 1–7
[18] Song H, Hunte F and Schwartz J 2012 On the role of pre-existing defects and magnetic flux avalanches in the degradation of YBa2Cu3O7–*x* coated conductors by quenching *Acta Mater.* **60** 6991–7000
[19] Computer Vision Annotation Tool (CVAT) 2023 CVAT.ai Corporation (available at: https://cvat.ai/)
[20] Lin T Y, Maire M, Belongie S, Bourdev L, Girshick R and Hays J Microsoft COCO: Common Objects in Context (arXiv:1405.0312)
[21] He K, Gkioxari G, Dollár P and Girshick R M 2017 R-CNN (available at: https://arxiv.org/abs/1703.06870v3)

[22] Wu Y, Kirillov A, Massa F, Lo W Y and Girshick R 2019 Detectron2 (available at: https://github.com/facebookresearch/detectron2)
[23] Russakovsky O, Li L J and Fei-Fei L 2015 Best of both worlds: Human-machine collaboration for object annotation *2015 IEEE Conf. on Computer Vision and Pattern Recognition (CVPRBoston, MA, USA)* (IEEE) pp 2121–31
[24] Svd W, Schönberger J L, Nunez-Iglesias J, Boulogne F, Warner J D and Yager N 2014 scikit-image: image processing in Python. *PeerJ* **2** e453 (arXiv:1407.6245)
[25] Philo C and Philo P 2022 2.15 or not 2.15? an historical-analytical inquiry into the nearest-neighbor statistic *Geogr. Anal.* **54** 333–56
[26] Bang J, Lee J, Bradford G, Kim K, Hu X and Abraimov D 2024 Evidence that transverse variability of critical current density can greatly mitigate screening current stress in high field REBCO magnets *Sci. Rep.* **14** 31703
[27] Mogro-Campero A, Turner L G, Hall E L, Lewis N, Peluso L A and Balz W E 1990 Film thickness dependence of critical current density and microstructure for epitaxial YBa2Cu3O7-x films *Supercond. Sci. Technol.* **3** 62
[28] Westerheim A C, Anderson A C, Oates D E, Basu S N, Bhatt D and Cima M J 1994 Relation between electrical properties and microstructure of YBa2Cu3O7–x thin films deposited by single target off axis sputtering *J. Appl. Phys.* **75** 393–403
[29] Vallat R 2018 Pingouin: statistics in Python *J. Open Source Softw.* **3** 1026