

Rapid and Flexible Semantic Segmentation of Electron Microscopy Data Using Few-Shot Machine Learning

Sarah Akers

Pacific Northwest National Laboratory

Elizabeth Kautz

Pacific Northwest National Laboratory

Andrea Trevino-Gavito

Pacific Northwest National Laboratory

Matthew Olszta

Pacific Northwest National Laboratory

Bethany Matthews

Pacific Northwest National Laboratory

Le Wang

Pacific Northwest National Laboratory

Yingge Du

Pacific Northwest National Laboratory

Steven Spurgeon (✉ steven.spurgeon@pnl.gov)

Pacific Northwest National Laboratory <https://orcid.org/0000-0003-1218-839X>

Article

Keywords: Atomic-scale Electron Microscope Images, Structure-property Relationships, Rapid Image Classification, Microstructural Feature Mapping

Posted Date: March 19th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-346102/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at npj Computational Materials on November 17th, 2021. See the published version at <https://doi.org/10.1038/s41524-021-00652-z>.

Rapid and Flexible Semantic Segmentation of Electron Microscopy Data Using Few-Shot Machine Learning

Sarah Akers,¹ Elizabeth Kautz,² Andrea Trevino-Gavito,^{1,3} Matthew Olszta,² Bethany Matthews,² Le Wang,⁴ Yingge Du,⁴ and Steven R. Spurgeon^{2, a)}

¹⁾*National Security Directorate, Pacific Northwest National Laboratory, Richland, WA, USA 99352*

²⁾*Energy and Environment Directorate, Pacific Northwest National Laboratory, Richland, WA, USA 99352*

³⁾*Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL, USA 60208*

⁴⁾*Physical and Computational Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA, USA 99352*

(Dated: 19 March 2021)

Semantic segmentation of key microstructural features in atomic-scale electron microscope images is critical to improved understanding of structure-property relationships in many important materials and chemical systems. However, the present paradigm involves time-intensive manual analysis that is inherently biased, error-prone, and unable to accommodate the large volumes of data produced by modern instrumentation. While more automated approaches have been proposed, many are not robust to a high variety of data, and do not generalize well to diverse microstructural features and material systems. Here, we present a flexible, semi-supervised few-shot machine learning approach for semantic segmentation of scanning transmission electron microscopy images of three oxide material systems: (1) epitaxial heterostructures of SrTiO₃ / Ge, (2) La_{0.8}Sr_{0.2}FeO₃ thin films, and (3) MoO₃ nanoparticles. We demonstrate that the few-shot learning method is more robust against noise, more reconfigurable, and requires less data than conventional image analysis methods. This approach can enable rapid image classification and microstructural feature mapping needed for emerging high-throughput and autonomous microscope platforms.

^{a)}Electronic mail: steven.spurgeon@pnl.gov

I. INTRODUCTION

Material microstructures govern the functionality of many important technologies, including catalysts, energy storage devices, and emerging quantum computing architectures. Scanning transmission electron microscopy (STEM) has long served as a foundational tool to study microstructures because of its ability to simultaneously resolve structure, chemistry, and defects with atomic-scale resolution for a range of materials classes.¹⁻³ STEM has helped elucidate the nature of microstructural features ranging from complex dislocation networks to secondary phases and point defects, leading to refined structure-property models.^{2,4,5} Traditionally, STEM images have been analyzed by a domain expert manually or semi-automatically, utilizing *a priori* knowledge of the system to identify known and unknown features. While this approach is suitable for measuring a limited number of features for small data volumes, it is impractical for samples possessing high density, rare, or noisy features.^{6,7} Moreover, manual approaches are difficult to scale to include multiple data modalities and cannot be performed at high speed, hindering our ability to perform *in situ* and complementary or correlative studies harnessing the full potential of modern instruments.⁸ At a more fundamental level, variability in how such measurements are conducted and a lack of standardized approaches contributes to the broader issue of reproducibility in experimentation.⁹ Though these limitations apply to all materials classes, they are particularly pronounced for complex oxides, whose properties are heavily influenced by even trace amounts of unwanted defects.¹⁰⁻¹² Hence, there is an urgent need to develop new approaches to characterize microstructural features with greater accuracy, speed, and statistical rigor than is possible with existing methodologies.

A central challenge in quantitatively describing microscopy image data (i.e. micrographs) is the wide variety of possible microstructural features and data modalities. The same instrument that is used to examine interfaces at atomic-resolution one day may be used to examine particle morphology or grain boundary distributions at lower magnification the next. In every study, the goal is to extract quantitative and semantically-meaningful microstructural descriptors to link measurements to underlying physical models.^{13,14} For example, estimating the area fraction of a specific phase or abundance of a feature through image segmentation is an important part of understanding synthesis products and phase transformation kinetics.¹⁵⁻¹⁹ Although several image segmentation methods exist (e.g. Otsu,²⁰ the water-

shed algorithm,²¹ k-means clustering²²), these are often not easily generalizable to different material systems, image types, and may require significant tailored image preprocessing.

Machine learning (ML) methods, specifically convolutional neural networks (CNNs), have recently been adopted for the recognition and characterization of microstructural data across length scales.^{23–26} Classification tasks have been performed to either assign a label to an entire image that represents a material or microstructure class (e.g. “dendritic,” “equiaxed,” etc.),^{26–29} or to assign a label to each pixel in the image so that they are classified into discrete categories.^{25,30–32} The latter classification type is segmentation of an image to identify local features (e.g. line defects, phases, crystal structures), referred to as semantic segmentation. However, many challenges remain in the practical application of semantic segmentation methods, such as the large data set size required for training and the difficulty of developing methods that are generalizable to a wide variety of data. Typically, data analysis via deep learning methods requires large amounts of labeled training data (such as the large image data set available through the ImageNet database).^{33,34} The ability to analyze data sets on the basis of limited training data, as often encountered in microscopy,^{35,36} is an important frontier in materials and data science. Recent advances have led to developments that allow human-level performance in one-shot, or few-shot learning problems,^{37,38} but there are limited studies on such methods in the materials science domain. While many characterization tools may provide just a few data points, a single electron micrograph (and potentially additional imaging / spectral channels) may encompass many microstructural features of interest. The one-shot or few-shot learning concept also has significant implications for the study of transient or unstable materials, as well as those where limited samples are available for analysis due to long lead-time experimentation (such as corrosion or neutron irradiation studies). In other cases, there exists data from previous studies that may be very limited or poorly understood, for which advanced data analysis methods could be applied.³⁹

In this work, we present a rapid and flexible approach to recognition and segmentation of STEM images using few-shot machine learning. Three oxide materials systems were selected for model development (epitaxial heterostructures of SrTiO₃ (STO) / Ge, La_{0.8}Sr_{0.2}FeO₃ (LSFO) thin films, and MoO₃ nanoparticles) due to the range of microstructural features they possess, and their importance in semiconductor, spintronic, and catalysis applications.^{40,41} We demonstrate that with only 5–8 sub-images (termed chips) that represent examples of a specific microstructural feature (e.g. a crystal motif or particular

particle morphology), our model yields segmentation results comparable to those produced by a domain expert for all oxide systems studied here. The successful image mapping can be attributed to the low noise sensitivity and high learning capability of few-shot machine learning in comparison to other segmentation methods (e.g. Otsu thresholding, watershed, k-means clustering, etc.). The few-shot approach rapidly identifies varying microstructural features across STEM data streams, which can inform real-time image data collection and analysis. More broadly, our findings underscore the power of image-driven machine learning to enable improved microstructural characterization for materials discovery and design.

II. RESULTS AND DISCUSSION

A deep learning approach known as few-shot learning was developed for semantic segmentation of STEM images. The premise of this few-shot learning model is to use very few labeled examples (< 10) per class for the model to identify regions of an image that correspond to each class. The general approach to image segmentation using few-shot learning is schematically described in Figure 1 and involves breaking an input image into a grid of sub-images (referred herein as chips), model initialization, inference, and output of a segmented micrograph. The process of chipping relies on domain-specific knowledge of the materials microstructure, as indicated in the annotations in Figure 1.A.

A. Preprocessing

To separate and measure distinct phases which have varying contrast in the STEM images, preprocessing of original image data was required. A histogram equalization (HE) technique designed to enhance local image qualities without introducing global artifacts termed contrast limited adaptive HE (CLAHE)^{42,43} was selected for use in this work. The details of the CLAHE implementation are described in Table I. CLAHE was first performed on original images and then the processed image was sectioned into a set of smaller sub-images, as shown in Figure 1.B. The chip size varied between 95×95 pixels and 32×32 pixels, however all chips are resized to 256×256 in the ResNet101 embedding module. The variable size allowed for each chip to be large enough to capture a microstructural motif and small enough to provide granularity between adjoining spatial regions, as shown in Figure 1.

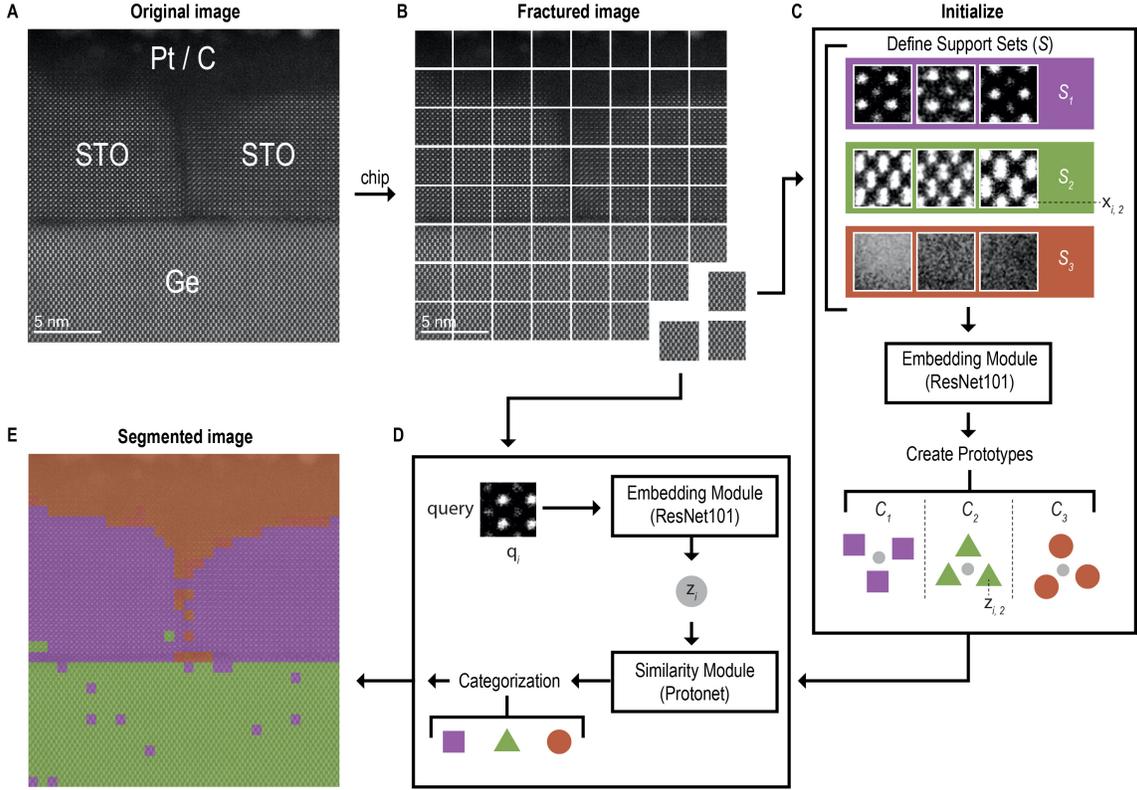


FIG. 1. A schematic of the few-shot approach to segmentation. The raw STO / Ge image (A) is broken into several smaller chips (B) and a few user defined chips are used to represent desired segmentation classes in the support set (C). Each chip then acts as a query and is compared against a prototype (D), defined by the support set, and categorized according to the minimum Euclidean distance between the query and each prototype, yielding the segmented image (E).

The final preprocessing step is an enhancement technique⁴⁴ that marks the position and size of atomic columns using a Laplacian of Gaussians (LoG) blob detection routine.⁴⁵ This step was used on the LSFO system to enhance the extremely subtle differences between classes.

B. Model Architecture

The few-shot model inputs the preprocessed STEM image, typically with high resolution on the order of 3000×3000 pixels, that has been broken down into a series of smaller chips, \mathbf{x}_{ik} , typically not larger than 100 by 100 pixels. A handful of these chips are used

as examples, or a support set, to define each of one or several classes. While most image applications for few-shot learning use disparate \mathbf{x}_{ik} to define a support set for each class (S_k),^{46–50} here S_k was created by breaking the original image into a grid of smaller sub-images. A subset of chips were labeled for each class. The set of N labeled examples for $k = 1, \dots, K$ classes makes up the support set defined by: $S = \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_N, y_N)\}$, where \mathbf{x}_i represents an image i and y_i is the corresponding true class label.

A Prototypical Network⁵¹ was selected in this work, given its lightweight design and simplicity. This model is based on the premise that each S_k may be represented by a single *prototype*, \mathbf{c}_k . To compute \mathbf{c}_k , each \mathbf{x}_{ik} first goes through an embedding function f_ϕ which maps a D -dimensional image into an M -dimensional representation through learnable parameters ϕ . The transformed chips, or $f_\phi(\mathbf{x}_{ik}) = \mathbf{z}_{ik}$, then creates the prototype for class k as the mean vector of the embedded support points \mathbf{c}_k , as follows:

$$\mathbf{c}_k = \frac{1}{N_{S_k}} \sum_{(\mathbf{z}_i, y_i) \in S_k} \mathbf{z}_i \quad (1)$$

After class prototypes are created, an untrained Prototypical Network classifies a new data point, or *query* q_i , by first transforming the query through the embedding function and then calculating a distance, e.g. Euclidean distance, between the embedded query vector and each of the class prototype vectors. After the distances are computed, a softmax normalizes the distance into class probabilities, where the class with the highest probability becomes the label for the query.⁵¹ The final output of the model, for each q_i , is the respective class label.

C. Model Inference

In order to quantify phase fractions in a STEM image (which can range from nm to μm in spatial dimension) each chip is used as a query point, q_i , so that the entire set of query points, Q , makes up the full image. The size of Q is directly proportional to the size of each chip and the size of the full image, as shown in Table II. All q_i first go through the embedding function and distances to each prototype are computed using the selected distance function. The network then produces a distribution over each of the K classes by computing a softmax over the distances and assigns a class label according to the highest normalized value.⁵¹

The model-specific implementation and parameters are given in Table II. While the selection of model parameters is often tedious, specific model parameters in the few-shot context are generally straightforward, since we often leverage pretrained models for the embedding architecture. Here, a residual network with 101 layers, ResNet101,⁵² was used as the embedding architecture. ResNet was specifically selected owing to its success in several related image recognition tasks.⁵² Model weights for ResNet101 are available from PyTorch⁵³ pytorch/vision v0.6.0, as trained on the image database ImageNet.⁵⁴ Additionally, the Euclidean distance metric was used, since this metric generally performs well across a wide variety of benchmark datasets and classification tasks.⁵¹ These pretrained models come with specified parameters and trained model weights. However, any embedding architecture may be used, especially those well-suited for segmentation tasks.⁵⁵

The similarity module can be any few-shot or meta-learning architecture as well; however, Protonets are generally simple and easy to implement. Parameters not necessarily specific to the models—namely chip size and batch size—should take the size of each distinct micrograph into consideration in addition to computational memory capacity. A chip should generally encompass a single micrograph and may take trial and error depending on the size of the full image and magnification. The batch size is simply the number of chips to evaluate at once. Generally, a machine with at least 16 GB of RAM and 2.7 GHz of processing power can reasonably compute model predictions at a rate of about 1 chip per 0.5 seconds, with a batch size of 100 chips measuring 64×64 pixels. The compute time naturally depends on processing power in addition to the chip size and the number of parameters in the embedding module. In the case of training a few-shot model, rather than simple inference as shown here, at least one GPU is necessary and may take several days to reach convergence given a sizeable database, e.g. a typical image database like ImageNet⁵⁴ contains 14 million images. The scope of this manuscript will only discuss the former, using an untrained few-shot model and pure inference to make judgments about an image.

D. Classification

The segmentation output of few-shot classification using the Prototypical architecture for three oxide systems is shown in Figure 2. The model output is a superpixel classification, i.e. every pixel that belongs to a chip receives the same label and corresponding color,

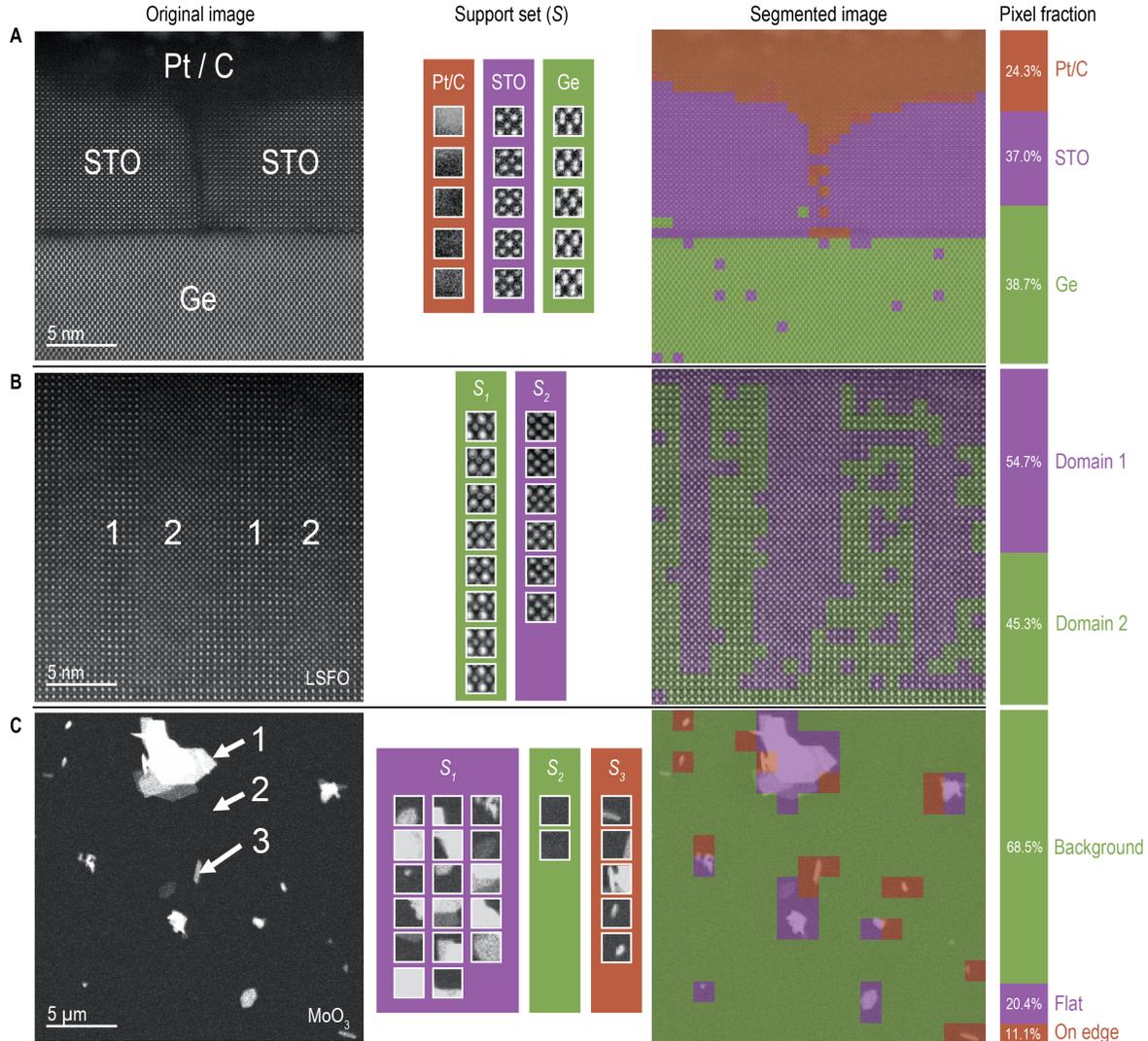


FIG. 2. (Left) Raw images from three oxide material systems: STO / Ge, LSFO, and MoO₃ (top to bottom row). (Center) respective support set and (right) output of few-shot segmentation. The estimated proportion of each support set class is shown in the inset bar chart for each system.

much in the same way other computer vision applications approach segmentation.⁵⁶ Here, the support set classes define the set of possible output labels. The percentage of chips belonging to each class, shown in Figure 2 (right), can be scaled from percentages to area using pixel scale conversions for a total area estimate for each distinct micrograph.

The STO / Ge system presents a particular challenge for most image analysis techniques in that the contrast varies irregularly across the whole image. The sample also contains

multiple interfaces and is representative of typical thin film-substrate imaging data. The selected LSFO image shows a secondary phase in the perovskite-structured matrix, and the secondary phase appears to have a gradient from top to bottom, which drastically diminishes the very subtle differences between the two micrographs. Separation of the two interpenetrating microstructural domains is necessary to understand the synthesis process and resulting properties, such as electrical conductivity. While preprocessing can adjust for some of these irregularities, traditional threshold-based segmentation techniques such as Otsu’s Method²⁰ and watershedding²¹ are not robust enough for a consistent solution and even adaptive methods can fail against a gradient and certainly when applied generally across multiple images. The few-shot technique is not completely immune to these irregularities as seen in Figure 2 (row A center) where the secondary micrograph region contains a handful of misclassified chips. In the LSFO system, few-shot is slightly more inconsistent in identifying the (green) microstructural features. However, these issues may be corrected with a *post hoc* spatial smoothing. For example, chips completely surrounded by one label within some radius would weight the class probability, or adjustments in the chips that define the support set. Despite some irregularities, the few-shot method is much more robust to noise within the full input image when compared against other segmentation techniques. Additionally, this few-shot method is easily generalizable to several different material systems, since a single support set defined by one image can be applied without adjustment to multiple images of the same type for an unmatched time savings in the analysis of image series. This feature is particularly important in the case of large area mapping, as shown for the Mo₃ nanoparticles, where it is necessary to collect image montages to survey the wide variety of possible particle morphologies. Here again the few-shot method successfully distinguishes several nanoparticle orientations from the carbon support background, with minimal instances of inaccurate labeling. Note the ability of the few-shot approach to accommodate the visual complexity of S_1 seen in 3 (row C middle), with a range of shapes, contrast, and sizes defining this ‘flat’ category. While S_1 here is defined with several more chips than the others, the model is able to reasonably perform a segmentation task impossible to contrast based methods alone. The ability of this model to generalize well to different material systems is demonstrated in Figure 2, which shows that illustrate varying microstructural features were successfully mapped for STO, LSFO, and MoO₃.

E. Comparison to Other Methods

Initially, several image analysis techniques were explored in an effort to quantify microstructural features of interest in specific micrographs, i.e. segmentation. It was immediately obvious that no single segmentation method would perform well in the absence of preprocessing steps, such as contrast adjustments, smoothing, and sharpening. Ideally, the aim of preprocessing in these analyses is to globally minimize artificial contrast textures and locally emphasize object edges, a critical noise reduction step for most segmentation routines.⁵⁷ Given that preprocessing and segmentation are often inseparable,⁵⁸ we examine comparable segmentation methods in the context of both segmentation and preprocessing together. In an effort to compare the few-shot approach with more widely-used segmentation methods, an example image from the STO / Ge system was analyzed using techniques with varying noise sensitivity and segmentation capabilities, with results presented in Figure 3.

The simplest approach to segmentation falls under a family of thresholding techniques shown in the first row of Figure 3. The three methods shown in the top row are designed to separate pixels in an image into two or more classes, based on an intensity threshold. The threshold in these methods is determined using information about the distribution of pixel intensities either globally (top row left) or locally using a neighborhood of pixels (top row center and right). The neighborhood methods are commonly more sensitive to noise, while Otsu’s more global technique appears to separate foreground pixels (light) from background (dark) relatively well.

Moving beyond simple thresholding, we begin to look towards separating pixels into classes other than background and foreground. The segmentation methods shown in Figure 3 typically have the ability to separate intensities into multiple classes again defined by the distribution of pixel intensities in the image. Two classes are specified for these routines in order to demonstrate the premise that, ideally, the image could be segmented according to the two distinct micrographs. These approaches also typically involve blurring filters and/or morphological operations⁵⁹ in order to remove pixels that are not a part of a larger group or shape. While shape edges are more defined in the middle row of Figure 3 than in the top row, we note that the resulting segmentation still appears to be background/foreground and misses the distinction between micrograph structures. One obvious limitation of a direct implementation of these methods is that the resulting classes will always be based on

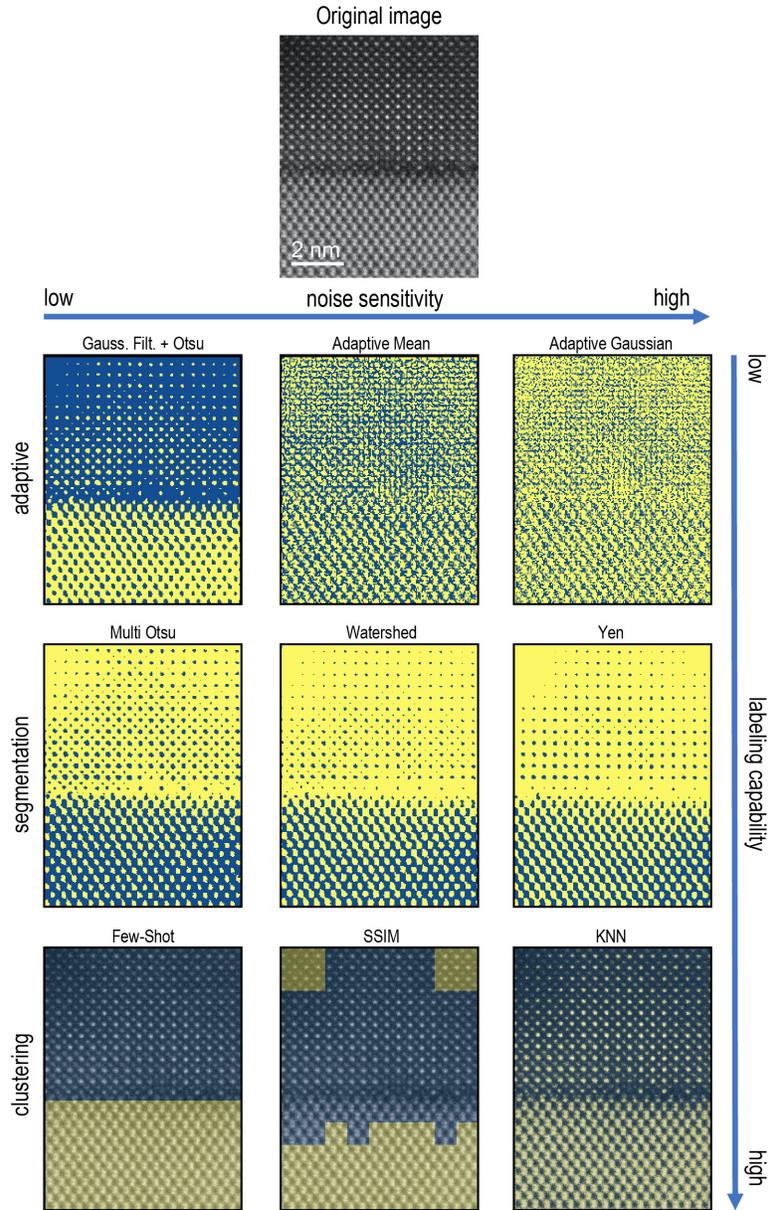


FIG. 3. An image from the STO / Ge system (top) is analyzed with a suite of image processing techniques with varying noise sensitivity and labelling capabilities. The thresholding techniques (top) typically separate background from foreground in an image. The strict segmentation techniques (middle) have the ability to separate the image further into multiple classes, though the classes are defined solely on pixel intensity. The clustering techniques (bottom) separate an image into classes based calculated image properties such as a centroid (KNN), structural similarity threshold (SSIM), or a prototype (few-shot).

intensity and not on the size or shape of the underlying micrographs. It may be possible to layer these methods with a shape detection routine where shapes of approximately the same size may be clustered into the same class. However, we found that clustering shapes post foreground/background segmentation was not able to distinctly separate microstructural features in an unsupervised manner, i.e. without tedious and manual intervention.

Rather than adding a shape clustering routine to an already segmented image, we implemented cluster based methods on either the raw image, or neighborhoods of the raw image in the bottom row of Figure 3. A common unsupervised K-Nearest Neighbors (KNN) clustering method is shown in the bottom row right, where we again see clustering results based on pixel intensity or background/foreground separation. Bottom row middle shows the first non-intensity based approach. An average structural similarity index measure (SSIM) is computed pairwise for 100×100 pixel non-overlapping neighborhoods as a measure of similarity between regions. The average SSIM for each neighborhood is a bimodal distribution that can be grouped into two classes as shown in bottom row center of Figure 3. However, the cutoff in SSIM must be manually determined. Lastly, the few-shot segmentation technique described in this manuscript is shown in bottom row left, where we see perfect segmentation between the two regions of distinct micrographs. We have included few-shot as a clustering technique, since a neighborhood is compared to a prototype analogous to the way clustering techniques compare to a centroid.

III. CONCLUSION AND FUTURE WORK

Here we developed a flexible few-shot learning approach to STEM image segmentation that can significantly accelerate mapping and identification of phases, defects, and other microstructural features of interest in comparison to more traditional image processing methods. We studied three different materials systems (STO / Ge, LSFO, and MoO₃), with varying atomic-scale features and hence diversity in image data for model development. Segmented images using the few-shot learning approach show good qualitative agreement with original micrographs.

When compared to other techniques, we find that noise sensitivity and/or labeling capability remain challenges for adaptive segmentation and clustering algorithms. The few-shot techniques explored in this manuscript provide powerful resources to combat these issues

and remain flexible enough to accommodate a suite of materials. While few-shot machine learning has been increasingly successful in rapidly generalizing to new classification tasks containing only a few samples with supervised information, it is a known problem that the empirical risk minimizer can be slightly unreliable,⁶⁰ leading to uncertainty in the reliability of the model for any given support set. We can mitigate some of this uncertainty with careful selection of the support set in order to avoid driving the model toward a non-optimal solution with mistakes in the support set, for instance. Another way we can alleviate some of the uncertainty in solutions is to actually train the few-shot model with large volumes of labeled data, when available, so that new classification tasks already have the benefit of optimization, even in a completely new and different material systems. Spatial statistics may also provide a means for smoothing spurious segmentation predictions from an otherwise uniform material region.

Aside from uncertainties in performance, the benefits in model generalizability also limit the amount of physically meaningful information we can extract from the model itself. One route for future exploration is the use of multiple data streams, such as spectral (electron energy loss spectroscopy (EELS) and energy-dispersive X-ray spectroscopy (EDS)), as well as diffraction (4D-STEM) in a multi-modal few-shot model. Insights in the multi-modal context would help to extract more physically meaningful information and likely enhance the segmentation performance.

In theory, this same model should generalize to a wide variety of STEM images and our preliminary results on other systems are indicative of this. Yet still unknown are the effects of the size and variety of the examples in a support set for a given class. Simulation studies are being designed to help answer these questions and more, including the possibility of using other models for embedding and the benefit of model training. Recent work in few-shot for texture segmentation⁶¹ also indicates that pixel level annotation may be possible, pushing the accuracy of this approach beyond chip-level resolution. Ideally, few-shot segmentation can also help curate and annotate these large training sets, something that is otherwise extremely costly to do manually and at scale. In summary, this approach offers a potentially powerful means to standardize and automate the analysis of materials microstructures for a single image, paving the way to new high-throughput characterization architectures.

IV. METHODS

A. Experimental Materials and Methods

The three experimental systems were prepared as follows. SrTiO_3 films were deposited onto Ge substrates using molecular beam epitaxy (MBE), as described elsewhere.⁴⁰ $\text{La}_{0.8}\text{Sr}_{0.2}\text{FeO}_3$ films were deposited onto SrTiO_3 (001) substrates using MBE, according to a procedure described elsewhere.⁴¹ Cross-sectional STEM samples of the thin films were prepared using a FEI Helios NanoLab DualBeam Focused Ion Beam (FIB) microscope and a standard lift out procedure. Bulk MoO_3 particles were drop cast onto a lacey carbon grid from suspension in ethanol. High-angle annular dark field (STEM-HAADF) images of the STO / Ge were collected on a probe-corrected JEOL ARM-200CF microscope operating at 200 kV, with a convergence semi-angle of 20.6 mrad and a collection angle of 90–370 mrad. STEM-HAADF images of the LSFO and MoO_3 were collected on a probe-corrected JEOL GrandARM-300F microscope operating at 300 kV, with a convergence semi-angle of 29.7 mrad and a collection angle of 75–515 mrad. The original image data analyzed in this work varied between 3042×3044 pixels (for STO / Ge), 2048×2048 (for LSFO), and 512×512 for MoO_3 . Because of its beam sensitivity, the STO / Ge images shown were collected using a frame-averaging approach; a series of 10 frames were acquired with a 1024×1024 px sampling and $2 \mu\text{s px}^{-1}$, then non-rigid aligned and upsampled $2\times$ using the SmartAlign plugin.⁶² Tens of images were collected from each material system and a range of selected defect features were used in this study.

B. Computational Methods

The specific implementation of the preprocessing techniques and parameters for the few-shot model are described in Tables II and II, respectively. All methods were implemented using the Python programming language v.3.6 available at <http://www.python.org>. Each image was processed using a 16 GB RAM 2.7 GHz Intel Core i7 MacBook Pro.

TABLE I. Image preprocessing parameters listed by material system and task with respective libraries/methods for implementation.

Task	Method / Library	Material System		
		LSFO	STO / Ge	MoO ₃
Image Reading	cv2.imread	flag=0 (grayscale)	flag=0 (grayscale)	flag=0 (grayscale)
Equalization	CLAHE	clipLimit=2.0, tileGridSize=(8, 8)	clipLimit=2.0, tileGridSize=(8, 8)	clipLimit=2.0, tileGridSize=(8, 8)
Chipping	image_slicer	tiles=500	tiles=1000	tiles=256
Resizing (pixels)	ResNet	256 × 256	256 × 256	256 × 256
LoG	skimage	max_sigma=30, num_sigma=8, threshold=0.1	N/A	N/A

TABLE II. Few-shot specific implementation information including model parameters, image information, and computing device used for the images shown in Figure 2.

Parameter	Material System		
	LSFO	STO / Ge	MoO ₃
Support Classes	2	3	3
Batch Size	8	8	16
Image Size (pixels)	2048 × 2048	3042 × 3044	512 × 512
Total Chips	500	1000	256
Chip Size (pixels)	34 × 35	95 × 95	32 × 32
Encoding	ResNet101 pretrained=True	ResNet101 pretrained=True	ResNet101 pretrained=True
Distance Metric	Euclidean	Euclidean	Euclidean
Similarity Module	Protonet	Protonet	Protonet
Computing	device=CPU	device=CPU	device=CPU

V. ACKNOWLEDGMENTS

The authors would like to thank Dr. Bryan Stanfill for useful discussions. This research was supported by a Chemical Dynamics Initiative (CDi) Laboratory Directed Research and Development (LDRD) project at Pacific Northwest National Laboratory (PNNL). PNNL is a multiprogram national laboratory operated for the U.S. Department of Energy (DOE) by Battelle Memorial Institute under Contract No. DE-AC05-76RL0-1830. Initial code development was performed on Nuclear Processing Science Initiative (NPSI) and I3T Commercialization Program LDRD projects. The growth and STEM data collection of the STO / Ge was supported by the U.S. Department of Energy (DOE), Office of Basic Energy Sciences, Division of Materials Science and Engineering under award no. 10122. A portion of the STEM imaging shown was performed in the Radiological Microscopy Suite (RMS), located in the Radiochemical Processing Laboratory (RPL) at PNNL. Thin film synthesis and additional characterization was performed using the Environmental Molecular Sciences Laboratory (EMSL), a national scientific user facility sponsored by the Department of Energy's Office of Biological and Environmental Research and located at PNNL.

VI. COMPETING INTERESTS STATEMENT

The authors declare no competing interests.

REFERENCES

- ¹I. MacLaren and Q. M. Ramasse, *Int. Mater. Rev.* **59**, 115 (2014).
- ²S. Pennycook, M. Varela, C. Hetherington, and A. Kirkland, *MRS Bull.* **31**, 36 (2006).
- ³M. Varela, A. Lupini, K. V. Benthem, A. Borisevich, M. Chisholm, N. Shibata, E. Abe, and S. Pennycook, *Annu. Rev. Mater. Res.* **35**, 539 (2005).
- ⁴S. J. Pennycook, *Ultramicroscopy* **180**, 22 (2017).
- ⁵M. P. Oxley, A. R. Lupini, and S. J. Pennycook, *Reports Prog. Phys.* **80**, 026101 (2017).
- ⁶J. A. Aguiar, M. L. Gong, R. R. Unocic, T. Tasdizen, and B. D. Miller, *Sci. Adv.* **5**, eaaw1949 (2019).
- ⁷P. M. Voyles, *Curr. Opin. Solid State Mater. Sci.* **21**, 141 (2017).
- ⁸S. R. Spurgeon, C. Ophus, L. Jones, A. Petford-Long, S. V. Kalinin, M. J. Olszta, R. E. Dunin-Borkowski, N. Salmon, K. Hattar, W.-c. D. Yang, R. Sharma, Y. Du, A. Chiaramonti, H. Zheng, E. C. Buck, L. Kovarik, R. L. Penn, D. Li, X. Zhang, M. Murayama, and M. L. Taheri, *Nat. Mater.* (2020), 10.1038/s41563-020-00833-z.
- ⁹M. Baker, *Nature* **533**, 452 (2016).
- ¹⁰F. Gunkel, D. V. Christensen, Y. Z. Chen, and N. Pryds, *Appl. Phys. Lett.* **116**, 120505 (2020).
- ¹¹Z. Huang, Ariando, X. Renshaw Wang, A. Rusydi, J. Chen, H. Yang, and T. Venkatesan, *Adv. Mater.* **30**, 1802439 (2018).
- ¹²H. L. Tuller and S. R. Bishop, *Annu. Rev. Mater. Res.* **41**, 369 (2011).
- ¹³A. Belianinov, R. Vasudevan, E. Strelcov, C. Steed, S. M. Yang, A. Tselev, S. Jesse, M. Biegalski, G. Shipman, C. Symons, A. Borisevich, R. Archibald, and S. Kalinin, *Adv. Struct. Chem. Imaging* **1**, 6 (2015).
- ¹⁴L. Vlcek, A. Maksov, M. Pan, R. K. Vasudevan, and S. V. Kalinin, *ACS Nano* **11**, 10313 (2017).
- ¹⁵J. P. Horwath, D. N. Zakharov, R. M egret, and E. A. Stach, *npj Comput. Mater.* **6**, 108 (2020).
- ¹⁶O. S. Ovchinnikov, A. O’Hara, S. Jesse, B. M. Hudak, S.-z. Yang, A. R. Lupini, M. F. Chisholm, W. Zhou, S. V. Kalinin, A. Y. Borisevich, and S. T. Pantelides, *Adv. Struct. Chem. Imaging* **6**, 3 (2020).

- ¹⁷A. Maksov, O. Dyck, K. Wang, K. Xiao, D. B. Geohegan, B. G. Sumpter, R. K. Vasudevan, S. Jesse, S. V. Kalinin, and M. Ziatdinov, *npj Comput. Mater.* **5**, 12 (2019).
- ¹⁸R. K. Vasudevan, M. Ziatdinov, S. Jesse, and S. V. Kalinin, *Nano Letters* **16**, 5574 (2016).
- ¹⁹E. Kautz, W. Ma, S. Jana, A. Devaraj, V. Joshi, B. Yener, and D. Lewis, *Materials Characterization*, 110379 (2020).
- ²⁰N. Otsu, *IEEE transactions on systems, man, and cybernetics* **9**, 62 (1979).
- ²¹H. Digabel and C. Lantuéjoul, in *Proc. 2nd European Symp. Quantitative Analysis of Microstructures in Material Science, Biology and Medicine*, Vol. 19 (Riederer Verlag, 1978) p. 8.
- ²²J. MacQueen *et al.*, in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, Vol. 1 (Oakland, CA, USA, 1967) pp. 281–297.
- ²³J. M. Ede, **149**, 130 (2020), arXiv:2009.08328.
- ²⁴S. R. Kalidindi and M. De Graef, *Annu. Rev. Mater. Res.* **45**, 171 (2015).
- ²⁵B. L. DeCost, T. Francis, and E. A. Holm, arXiv:1805.08693 [cs] (2018).
- ²⁶A. Chowdhury, E. Kautz, B. Yener, and D. Lewis, *Computational Materials Science* **123**, 176 – 187 (2016).
- ²⁷S. M. Azimi, D. Britz, M. Engstler, M. Fritz, and F. Mücklich, *Sci. Rep.* **8:2128** (2018), 10.1038/s41598-018-20037-5.
- ²⁸B. L. DeCost and E. A. Holm, *Comput. Mater. Sci.* **110**, 126–133 (2015).
- ²⁹A. Baskaran, G. Kane, K. Biggs, R. Hull, and D. Lewis, *Comput. Mater. Sci.* **177**, 109593 (2020).
- ³⁰M. Ziatdinov, O. Dyck, A. Maksov, X. Li, X. Sang, K. Xiao, R. R. Unocic, R. Vasudevan, S. Jesse, and S. V. Kalinin, *ACS Nano* **11**, 12742 (2017).
- ³¹G. Roberts, S. Y. Haile, R. Sainju, D. J. Edwards, B. Hutchinson, and Y. Zhu, *Sci. Rep.* **9**, 12744 (2019).
- ³²D. Chen, D. Guo, S. Liu, and F. Liu, *Symmetry* **12**, 639 (2020).
- ³³A. Krizhevsky, I. Sutskever, and G. E. Hinton, in *Advances in Neural Information Processing Systems*, Vol. 25, edited by F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger (Curran Associates, Inc., 2012) pp. 1097–1105.
- ³⁴J. Deng, W. Dong, R. Socher, L.-j. Li, K. Li, and L. Fei-fei, in *In CVPR* (2009).
- ³⁵M. Larmuseau, M. Sluydts, K. Theuwissen, L. Duprez, T. Dhaene, and S. Cottenier, *Scr. Mater.* **193**, 33 (2021).

- ³⁶J.-F. Díez-Pastor, P. Latorre-Carmona, Á. Arnaiz-González, J. Ruiz-Pérez, and D. Zurro, *Microsc. Microanal.* **26**, 1158 (2020).
- ³⁷H. Altae-Tran, B. Ramsundar, A. S. Pappu, and V. Pande, *ACS central science* **3**, 283 (2017).
- ³⁸E. M. Rutter, J. H. Lagergren, and K. B. Flores, *Domain Adaptation and Representation Transfer and Medical Image Learning with Less Labels and Imperfect Data*, , 190 (2019).
- ³⁹E. J. Kautz, A. R. Hagen, J. M. Johns, and D. E. Burkes, *Computational Materials Science* **161**, 107 (2019).
- ⁴⁰Y. Du, P. V. Sushko, S. R. Spurgeon, M. E. Bowden, J. M. Ablett, T.-L. Lee, N. F. Quackenbush, J. C. Woicik, and S. A. Chambers, *Phys. Rev. Mater.* **2**, 094602 (2018).
- ⁴¹L. Wang, Y. Du, P. V. Sushko, M. E. Bowden, K. A. Stoerzinger, S. M. Heald, M. D. Scafetta, T. C. Kaspar, and S. A. Chambers, *Phys. Rev. Mater.* **3**, 025401 (2019).
- ⁴²S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman, and K. Zuiderveld, *Computer vision, graphics, and image processing* **39**, 355 (1987).
- ⁴³K. Zuiderveld, *Graphics gems* , 474 (1994).
- ⁴⁴B. P. Marsh, N. Chada, R. R. S. Gari, K. P. Sigdel, and G. M. King, *Scientific reports* **8**, 1 (2018).
- ⁴⁵S. Van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, and T. Yu, *PeerJ* **2**, e453 (2014).
- ⁴⁶O. Vinyals, C. Blundell, T. Lillicrap, K. Kavukcuoglu, and D. Wierstra, *arXiv preprint arXiv:1606.04080* (2016).
- ⁴⁷S. Rahman, S. Khan, and F. Porikli, *IEEE Transactions on Image Processing* **27**, 5652 (2018).
- ⁴⁸H. Edwards and A. Storkey, *arXiv preprint arXiv:1606.02185* (2016).
- ⁴⁹M. Ren, E. Triantafillou, S. Ravi, J. Snell, K. Swersky, J. B. Tenenbaum, H. Larochelle, and R. S. Zemel, *arXiv preprint arXiv:1803.00676* (2018).
- ⁵⁰N. Hilliard, L. Phillips, S. Howland, A. Yankov, C. D. Corley, and N. O. Hodas, *arXiv preprint arXiv:1802.04376* (2018).
- ⁵¹J. Snell, K. Swersky, and R. Zemel, in *Advances in neural information processing systems* (2017) pp. 4077–4087.

- ⁵²K. He, X. Zhang, S. Ren, and J. Sun, in *Proceedings of the IEEE conference on computer vision and pattern recognition* (2016) pp. 770–778.
- ⁵³A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimeshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, in *Advances in Neural Information Processing Systems 32*, edited by H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett (Curran Associates, Inc., 2019) pp. 8024–8035.
- ⁵⁴J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, in *2009 IEEE conference on computer vision and pattern recognition* (IEEE, 2009) pp. 248–255.
- ⁵⁵J. P. Horwath, D. N. Zakharov, R. Megret, and E. A. Stach, *npj Computational Materials* **6**, 1 (2020).
- ⁵⁶R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, *IEEE transactions on pattern analysis and machine intelligence* **34**, 2274 (2012).
- ⁵⁷A. Hughes, Z. Liu, M. Raftari, and M. E. Reeves, *A workflow for characterizing nanoparticle monolayers for biosensors: Machine learning on real and artificial SEM images*, Tech. Rep. (PeerJ PrePrints, 2014).
- ⁵⁸R. C. Gonzales and R. E. Woods, “Digital image processing,” (2002).
- ⁵⁹A. Buades, B. Coll, and J. M. Morel, *CMLA Preprint* **5** (2004).
- ⁶⁰Y. Wang, Q. Yao, J. T. Kwok, and L. M. Ni, *ACM Computing Surveys (CSUR)* **53**, 1 (2020).
- ⁶¹I. Ustyuzhaninov, C. Michaelis, W. Brendel, and M. Bethge, *arXiv preprint arXiv:1807.02654* (2018).
- ⁶²L. Jones, H. Yang, T. J. Pennycook, M. S. J. Marshall, S. Van Aert, N. D. Browning, M. R. Castell, and P. D. Nellist, *Adv. Struct. Chem. Imaging* **1**, 8 (2015).

Figures

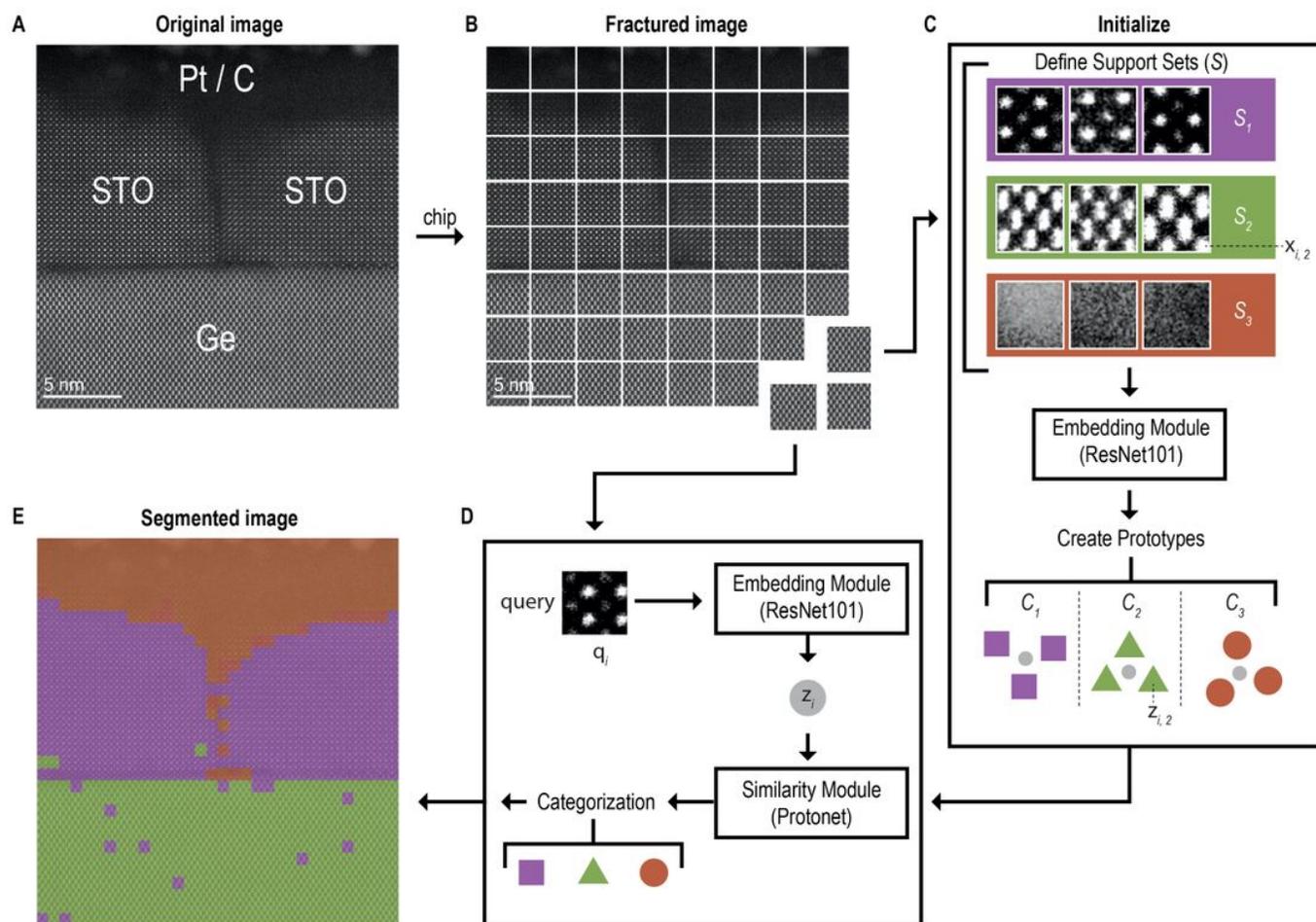


Figure 1

A schematic of the few-shot approach to segmentation. The raw STO / Ge image (A) is broken into several smaller chips (B) and a few user defined chips are used to represent desired segmentation classes in the support set (C). Each chip then acts as a query and is compared against a prototype (D), defined by the support set, and categorized according to the minimum Euclidean distance between the query and each prototype, yielding the segmented image (E).

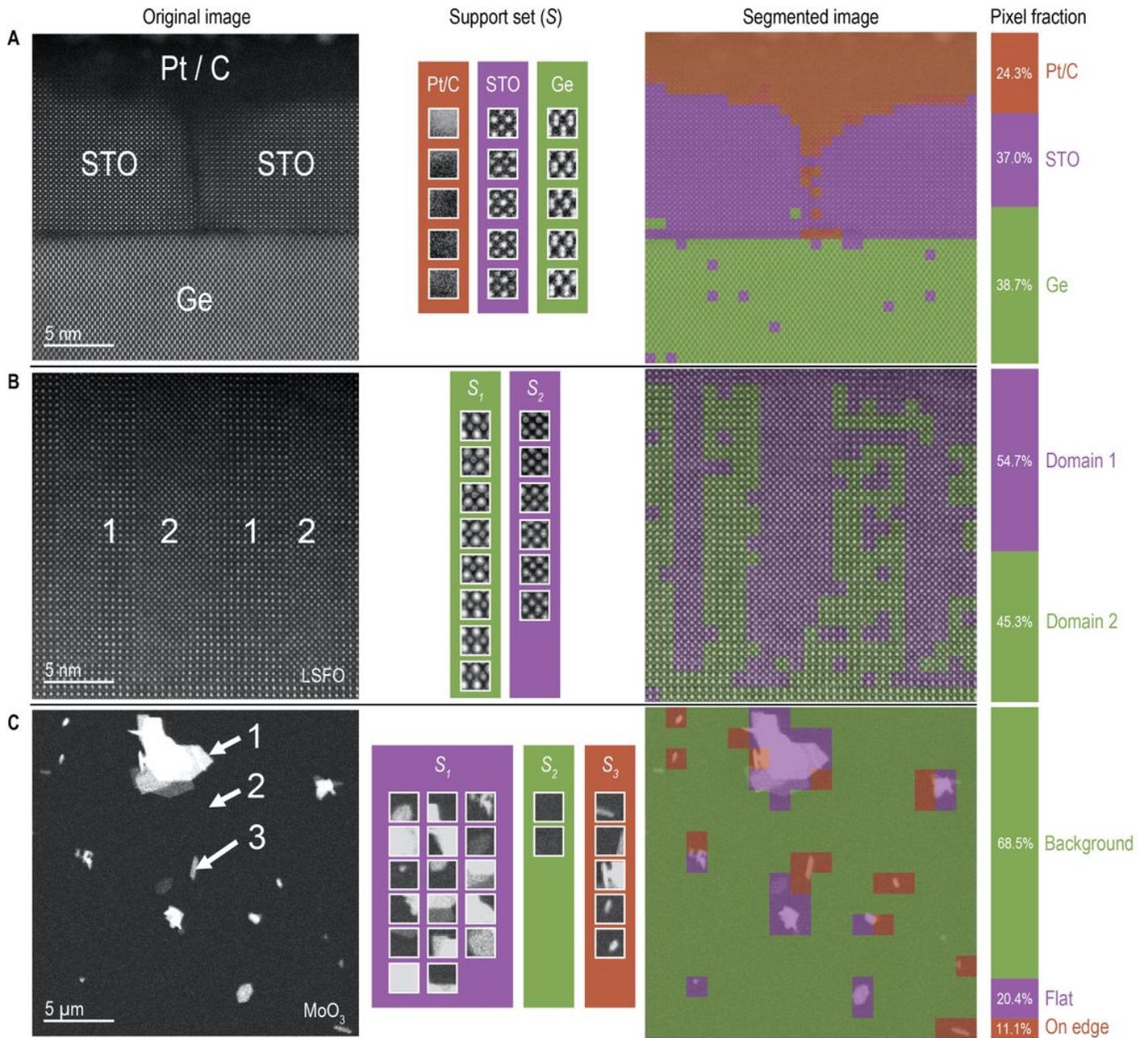


Figure 2

(Left) Raw images from three oxide material systems: STO / Ge, LSF0, and MoO₃ (top to bottom row). (Center) respective support set and (right) output of few-shot segmentation. The estimated proportion of each support set class is shown in the inset bar chart for each system.

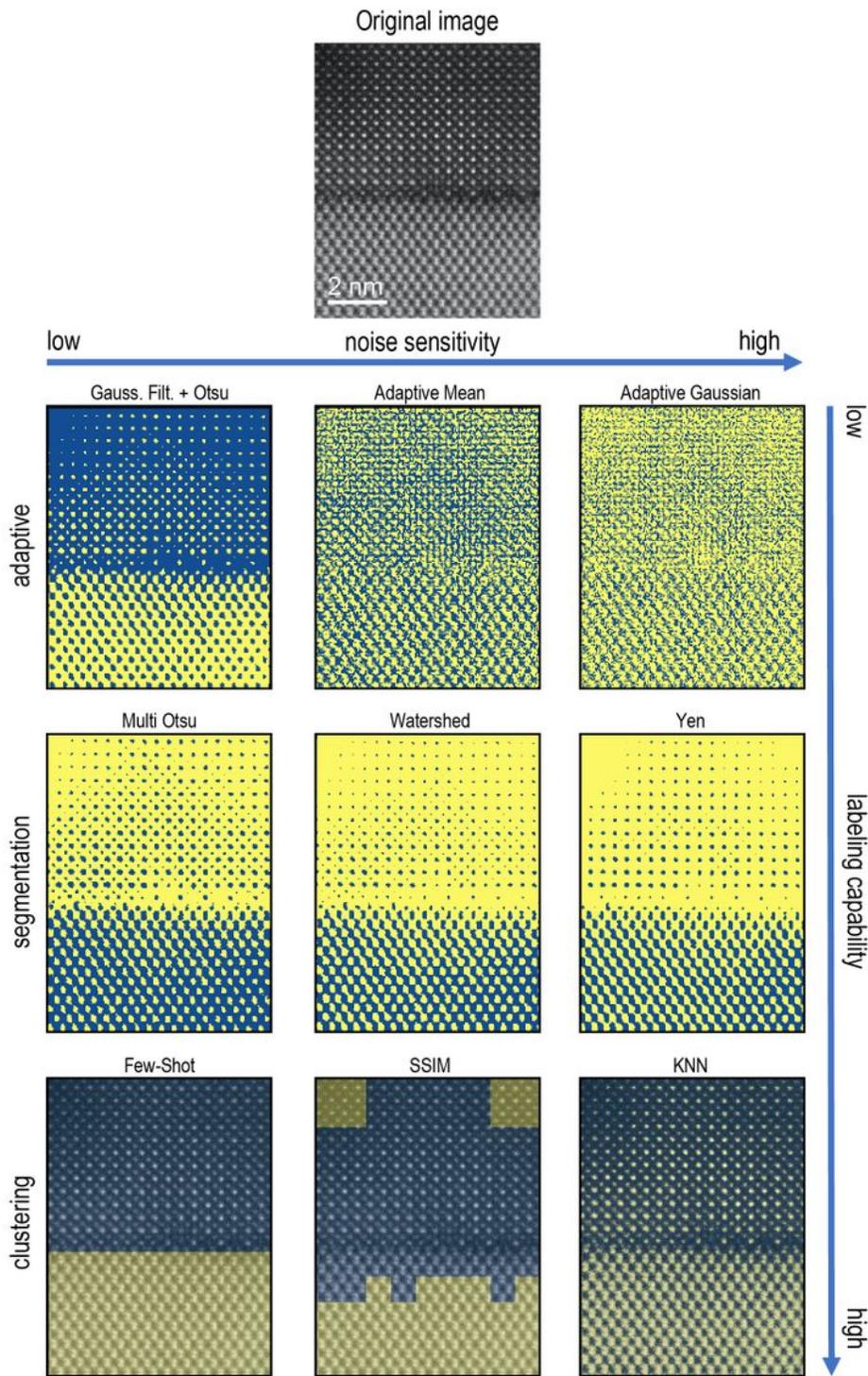


Figure 3

An image from the STO / Ge system (top) is analyzed with a suite of image processing techniques with varying noise sensitivity and labelling capabilities. The thresholding techniques (top) typically separate background from foreground in an image. The strict segmentation techniques (middle) have the ability to separate the image further into multiple classes, though the classes are defined solely on pixel intensity.

The clustering techniques (bottom) separate an image into classes based calculated image properties such as a centroid (KNN), structural similarity threshold (SSIM), or a prototype (few-shot).