

Rapid and Flexible Few Shot Learning-Based Classification of Scanning Transmission Electron Microscopy Data

Sarah Reehl¹, Elizabeth Kautz², Matthew Olszta³, Derek Hopkins¹, Bethany Matthews⁴, Le Wang¹, Yingge Du¹ and Steven Spurgeon¹

¹Pacific Northwest National Laboratory, United States, ²Pacific Northwest National Laboratory - PNNL, Richland, Washington, United States, ³Pacific Northwest National Laboratory, Washington, United States, ⁴Pacific Northwest National Laboratory, Richland, Washington, United States

Control of property-defining materials defects for quantum computing and energy storage depends on the ability to precisely probe structure and chemistry at the highest spatial and temporal resolutions. Modern scanning transmission electron microscopy (STEM) is well-suited to this task, having yielded rich insights into defect populations in many systems. However, the dilute nature and complexity of materials defects, coupled with their varied representations in STEM data, makes reliable, accurate, high-throughput statistical defect analysis a significant challenge. Possible analysis approaches include low-level pixel processing, or even the application of machine learning methods for classification and image segmentation. However, the latter requires large sets of labeled training data that are difficult to obtain for many practical materials science studies. Here, we describe the use of an emerging few shot learning capability for rapid and flexible STEM data classification. This approach requires minimal information at the start of the analysis and uses a generally pre-trained encoder network to make inferences on experimental data. Our results show drastic improvements in data annotation costs, reproducibility, and scalability in comparison to neural network training from scratch. We demonstrate how few shot techniques can quickly extract feature maps and global statistics from a variety of STEM data, enabling a new quantitative understanding of defect populations.

This research was supported by the Chemical Dynamics Initiative (CDi) Laboratory Directed Research and Development (LDRD) program at Pacific Northwest National Laboratory (PNNL). PNNL is a multi-program 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 I3T Commercialization and Nuclear Process Science Initiative (NPSI) LDRD projects at PNNL. A portion of the work was performed at the W. R. Wiley Environmental Molecular Sciences Laboratory, a DOE User Facility sponsored by the Office of Biological and Environmental Research and located at PNNL.

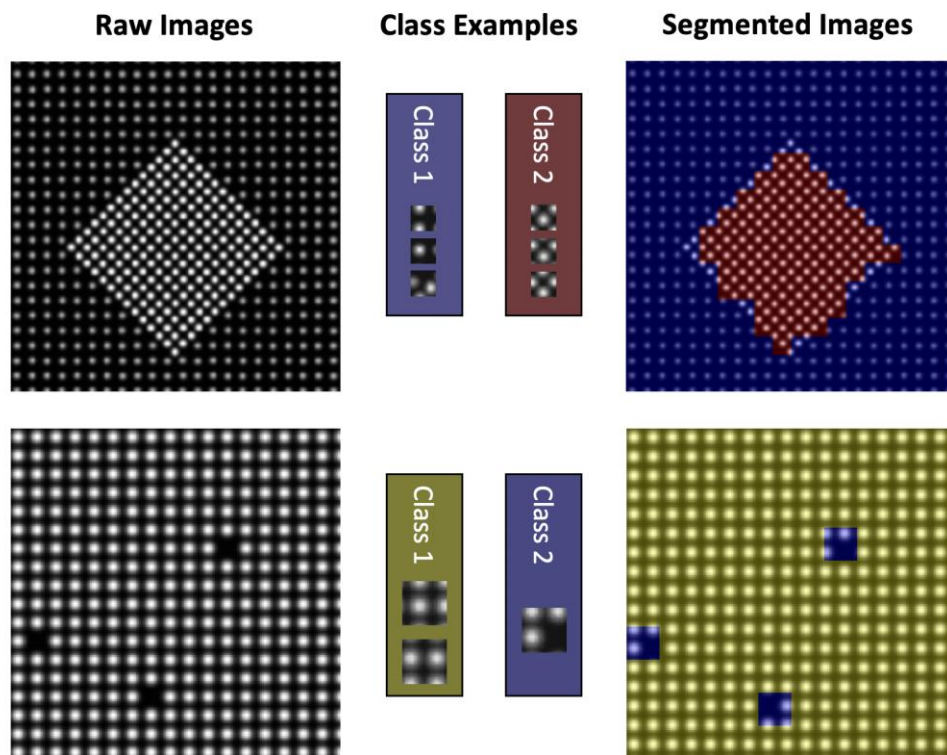


Figure 1. Output of few-shot based segmentation using a limited number of class examples for two synthetic atomic-resolution STEM datasets.