Anomaly Detection and Clustering of Atomic-Resolution STEM Images Using Semantic Segmentation

Surabhi Sathish,[a] Claudia Sosa Espada[b] and Thomas Karagiannis[c]

[a] Department of Mechanical Engineering, University of Bath, Bath, BA2 7AY, United Kingdom

[b] School of Physics, University College Dublin, Dublin 4, Ireland

[c] School of Chemical and Bioprocess Engineering, University College Dublin, Belfield, Dublin 4, Ireland

Abstract

The rapid growth of atomic-resolution STEM datasets necessitates automated and unbiased approaches for defect identification beyond manual analysis.[1–4] Here, we present a supervised machine-learning framework for defect discovery and clustering in atomic-resolution HAADF-STEM images. The approach employs transfer learning using a convolutional neural network (CNN) trained on bulk-like regions of CdTe and $SrTiO_3$ datasets to classify images according to the defects present. High classification performance is achieved, with accuracies exceeding 94% for $SrTiO_3$ and 99% for CdTe. To further generate defect-sensitive reconstruction error maps, image segmentation was carried out through a U-net architecture. We reformulate defect classification as weakly supervised multi-label semantic segmentation. Because pixel-level annotations are unavailable, we generate pseudo-labels using local intensity deviations, allowing the network to learn spatial localisation of defects rather than global classification alone. Quantitative descriptors derived from these maps are subsequently clustered using unsupervised k-means, with the number of clusters optimised via silhouette metrics and visualized using t-SNE. This workflow enables autonomous separation of bulk regions and multiple defect classes without labelled data. The proposed framework advances anomaly detection toward scalable, weakly supervised defect taxonomy and provides a reproducible pathway for high-throughput, data-driven electron microscopy analysis.

Methodology and Results

Initial defect identification was performed using a supervised image classification model to establish a baseline performance. A convolutional neural network based on the MobileNet architecture was implemented using transfer learning (fine-tuning in the last 40 layers with data set (batch size = 32, epochs = 25) to enhance feature extraction and improve convergence on limited microscopy datasets. The data were divided into training, validation, and test sets using a 70:20:10 split. Model parameters were optimised via backpropagation, and performance was evaluated on unseen test images, yielding validation and test accuracies of 94% for $SrTiO_3$ and 99.3% for CdTe.. Representative classification results are shown in *Figure 1*.

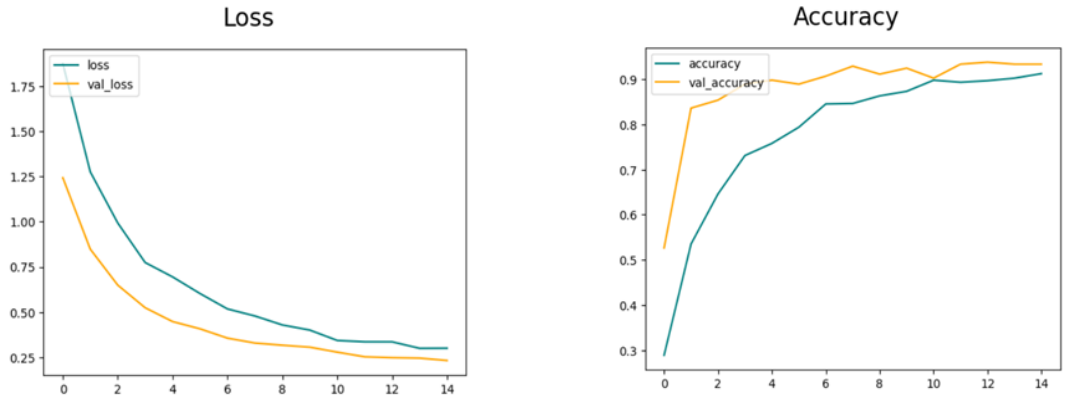

Figure 1: Training loss and accuracy for the image classification models against the number of epochs.

While effective for single-label classification, this approach is insufficient for microscopy images containing multiple, spatially overlapping defects. To address this limitation, a semantic segmentation model based on the U-Net architecture was developed. The encoder–decoder structure, combined with skip connections, enabled both contextual feature learning and precise spatial localisation. Synthetic multi-defect images were generated to emulate realistic experimental conditions. In the absence of pixel-level labels, weak supervision was implemented through pseudo-ground truth derived from local intensity

deviations relative to bulk lattice regions. The segmentation network produced saliency heat maps representing pixel-wise defect likelihoods, as shown in *Figure 2*, where brighter regions indicate stronger model confidence.

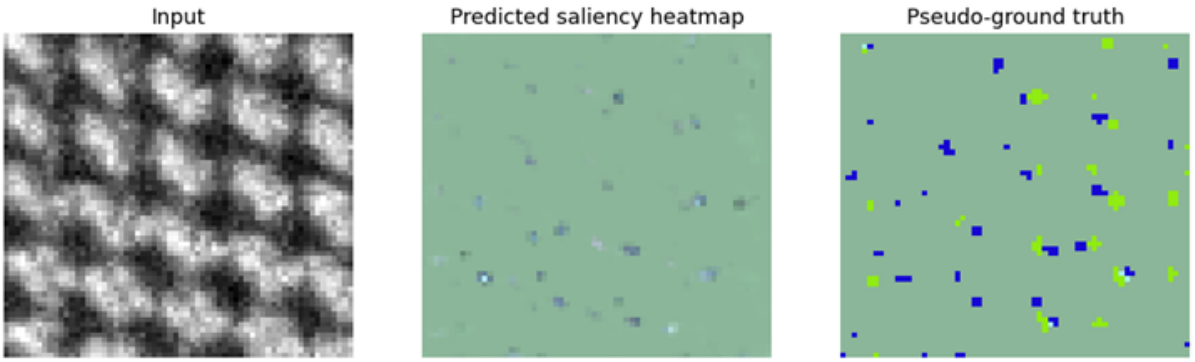

Figure 2: Output from the segmentation model. The input image is a reconstruction including three types of defect characteristics.

Class-specific probability maps were subsequently extracted to quantify the presence of individual defect types within the same image. *Figure 3* illustrates these probability distributions, with colour intensity reflecting prediction confidence. This approach enables simultaneous localisation and discrimination of multiple defect classes within a single field of view.

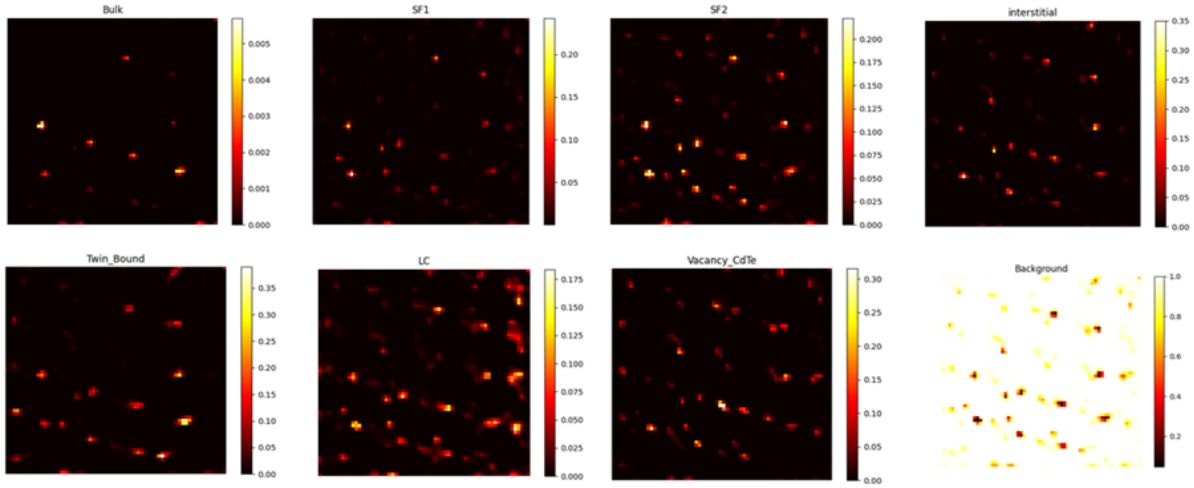

Figure 3: Heat maps for each defect showing the confidence level of the model.

Conclusions

We present a machine-learning framework for automated defect classification, localisation, and clustering in atomic-resolution HAADF-STEM images. High classification accuracies were achieved for both materials studied, exceeding 99% for CdTe and 94% for $SrTiO_3$. Differences in accuracy reflect intrinsic variations in crystal structure and defect complexity, with CdTe exhibiting stronger defect contrast than $SrTiO_3$. The segmentation model enables reliable localisation of multiple, coexisting defect types within a single image, while the class-specific probability heatmaps provide quantitative defect identification, making the framework well prepared for realistic laboratory conditions.

References

(1) Ayyubi, R. A. W.; Buban, J. P.; Klie, R. F. Automated Defect Detection in Atomic Resolution STEM Images: A Machine Learning Approach with Variational Convolutional Autoencoders. *Microscopy and Microanalysis* **2024**, *30* (Supplement_1), ozae044.180. https://doi.org/10.1093/mam/ozae044.180.

(2) Prifti, E.; Buban, J. P.; Thind, A. S.; Klie, R. F. Variational Convolutional Autoencoders for Anomaly Detection in Scanning Transmission Electron Microscopy. *Small* **2023**, *19* (16), 2205977. https://doi.org/10.1002/smll.202205977.

(3) Gui, C.; Zhang, Z.; Li, Z.; Luo, C.; Xia, J.; Wu, X.; Chu, J. Deep Learning Analysis on Transmission Electron Microscope Imaging of Atomic Defects in Two-Dimensional Materials. *iScience* **2023**, *26* (10), 107982. https://doi.org/10.1016/j.isci.2023.107982.

(4) Barakati, K.; Zhang, X.; Wang, X.; Kalinin, S. V. Machine Learning-Enhanced TEM Image Analysis Techniques. *Microscopy and Microanalysis* **2024**, *30* (Supplement_1), ozae044.1027. https://doi.org/10.1093/mam/ozae044.1027