Automated Defect Detection in STEM Images Using Machine Learning

Joshua Marvin[1], Noah Holt[2,*], Ramji Subedi[1], Kamal Khanal[3], Mohd Tauhid Khan[4]

1. Department of Mechanical and Aerospace Engineering, Oklahoma State University, USA

2. Department of Chemistry, Oklahoma State University, USA

3. Department of Physics, Oklahoma State University, USA

4. Department of Chemical Engineering, Oklahoma State University, USA

*Email address: noah.holt@okstate.edu

Abstract

We developed an ensemble deep learning framework for automated defect detection in high-resolution scanning transmission electron microscopy (STEM) images of CdTe and $SrTiO_3$ (STO). The approach follows patch-level classification using custom convolutional neural networks (CNNs) trained on frequency-enhanced image representations to identify crystalline and defect regions. Training is performed on labeled 64×64 image patches using a tailored preprocessing pipeline that enhances structural contrast through normalization, Gaussian filtering, and Fourier-domain feature amplification. Models converge rapidly within a small number of epochs and achieve robust binary classification performance. The trained networks are then deployed for large-area image analysis, enabling efficient localization of defect-rich regions and showing the importance of machine learning for accelerating microscopy-based materials characterization.

Methodology

Training data consist of labeled 64×64 pixel STEM image patches extracted from CdTe and STO datasets [1] and processed through a consistent preprocessing pipeline which follows max-intensity normalization, Gaussian smoothing, logarithmic Fourier transform magnitude enhancement and applying threshold to suppress background contributions as shown in fig (2). A custom convolutional neural network was constructed using multiple convolutional layers initialized with He normal initialization, each followed by batch normalization and SELU activation, with max pooling and fully connected layers culminating in a sigmoid-activated output neuron for binary classification. Binary classification was performed to distinguish bulk crystalline regions from defect-containing regions based on dataset-provided annotations. Data augmentation is done by rotational transformations to 90°, 180°, 270° improving generalization. Class imbalance is reduced by dataset mixing and random shuffling prior to training, ensuring balanced exposure to bulk and defect patches and improves the reliability and stability of defect detection. Models are trained

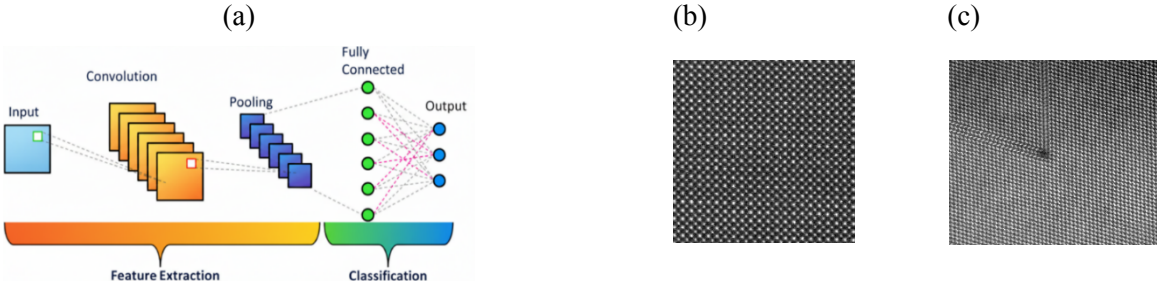


Fig 1: (a) Schematic of FFT-enhanced preprocessing and CNN-based feature detection workflow, (b) and (c) representative STEM image of STO and CdTe used for testing.

using binary cross-entropy loss and the Adam optimizer for a small number of epochs. 3 for the bulk and 9 for the SF-1 models respectively, for example. Trained networks are serialized and then loaded later for inference. During deployment, large microscopy images are scanned using an overlapping sliding window, classified patch by patch using an ensemble of pretrained models and reconstructed into spatial masks that highlight regions of interest across the original image. Here, Bulk denotes defect-free crystalline regions. LC corresponds to Lomer–Cottrell defects (junctions) formed by dislocation reactions, producing localized strain contrast. Vacancies represent missing atomic columns, while interstitials indicate excess atomic columns. Stacking faults (SF1/SF2) arise from disruptions in the stacking sequence of atomic planes, and twin boundaries are mirror-symmetric crystalline regions separated by coherent planar defects.

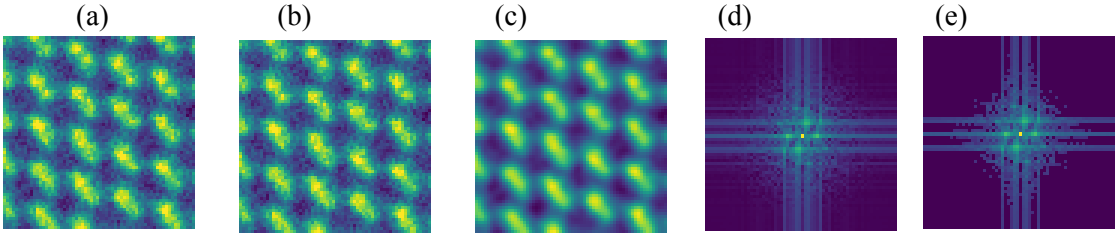
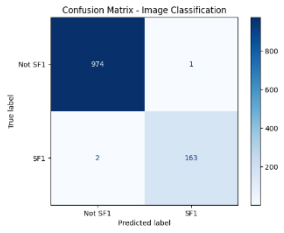


Fig 2: Image-processing pipeline showing: (a) original image, (b) intensity normalization, (c) background removal via Gaussian blurring, (d) FFT of the normalized image, and (e) FFT after threshold

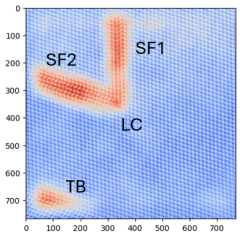
Results and discussion

Quantitative evaluation of the trained model shows performance of binary classification as shown in figure 3. The confusion matrix shows a high true positive and true negative rate, with 974 correctly classified non-defect patches and 163 correctly classified defect patches, and only three total misclassifications across 1140 test samples. The corresponding classification report indicates precision, recall, and F1-scores of approximately 0.99–1.00 for both classes, resulting in an overall accuracy of ~99%. These results confirm that the network effectively distinguishes bulk crystalline regions from defect-containing regions, while maintaining low false-positive and false-negative rates. The high classification consistency further validates the effectiveness of the FFT-enhanced preprocessing pipeline and the patch-based CNN architecture for automated defect detection in STEM images and result after testing the model is shown in figure 3. (b) where we see proper detections of defects. Models were also constructed for STO images; however, the provided testing image contains no defects, severely limiting evaluation. When tested on images outside the given dataset, the model successfully identified defect regions, demonstrating its generalizability.

(a)



(b)



(c)

	precision	recall	f1-score	support
0	1.00	1.00	1.00	975
1	0.99	0.99	0.99	165
accuracy			1.00	1140
macro avg	1.00	0.99	0.99	1140
weighted avg	1.00	1.00	1.00	1140

Fig 3. (a) Confusion matrix for binary defect classification, (b) spatial map of predicted defect regions highlighting SF, LC, and TB features, and (c) corresponding classification metrics.

References:

1. https://github.com/RAW-Ayyubi/defect-classification-in-stem-images