

## DefectNet – A Deep Convolutional Neural Network for Semantic Segmentation of Crystallographic Defects in Advanced Microscopy Images

Graham Roberts<sup>1</sup>, Rajat Sainju<sup>2</sup>, Brian Hutchinson<sup>3,4</sup>, Mychailo B. Toloczko<sup>1</sup>, Danny J. Edwards<sup>1</sup>, Yuanyuan Zhu<sup>1,2\*</sup>

<sup>1</sup> Nuclear Sciences Division, <sup>4</sup> Computing & Analytics Division, Pacific Northwest National Laboratory, Richland, WA 99352, USA

<sup>2</sup> Department of Materials Science and Engineering, Institute of Materials Science, University of Connecticut, Storrs, CT 06269, USA

<sup>3</sup> Computer Science Department, Western Washington University, Bellingham, WA 98225, USA

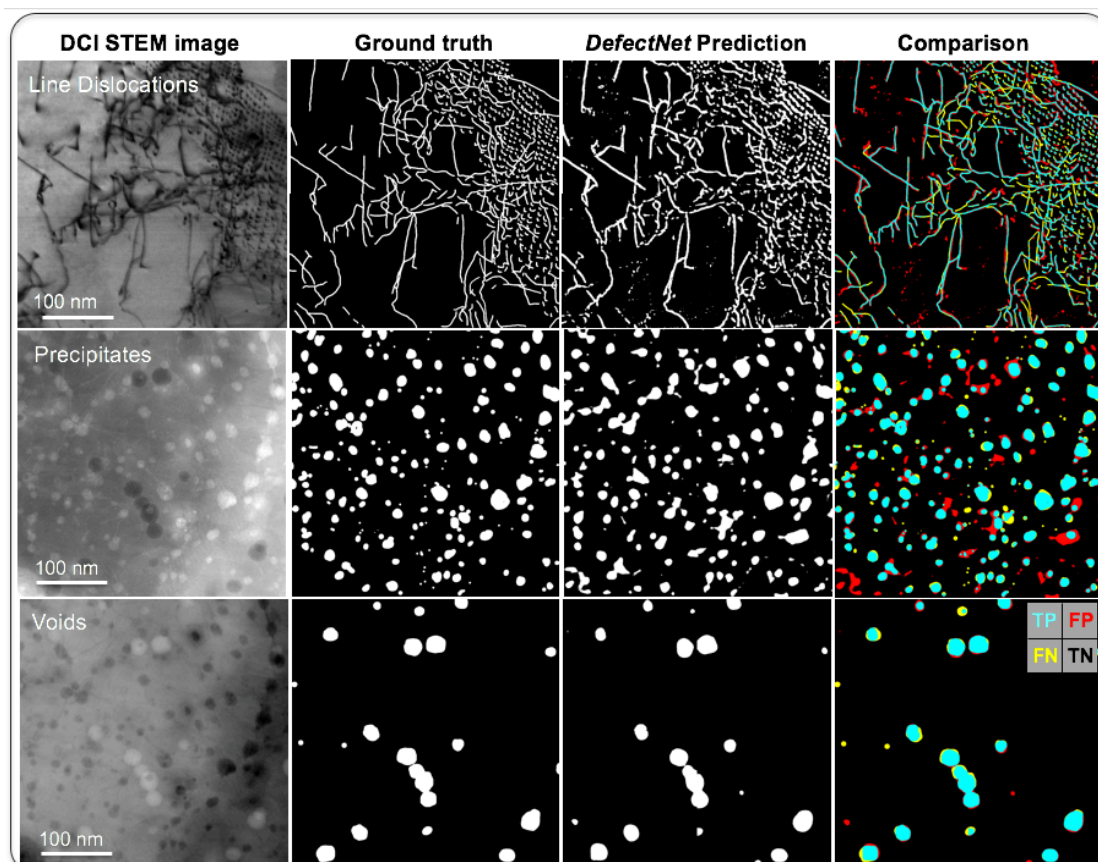
\* Corresponding author: [yuanyuan.2.zhu@uconn.edu](mailto:yuanyuan.2.zhu@uconn.edu)

The current practice of identifying defects in microscopy images and deriving metrics such as dislocation density and precipitates/voids diameter remains largely in the purview of human analysis. The lack of automated defect analysis for statistically meaningful quantification of a variety types of crystallographic defects is causing an increasingly large bottleneck for rational alloy design. The first and most important step of automating defect analysis is perceptual defect identification. In terms of digital image processing, semantic segmentation best emulates human recognition of defect features – it tells what defects are in an image and where they are located. In this work, we developed a novel deep convolutional neural network (CNN) model, called *DefectNet*, for robust and automated semantic segmentation of three crystallographic defects including line dislocations, precipitates, and voids commonly observed in structural metals and alloys [1]. Defect semantic segmentation in TEM micrographs is a challenging deep learning task due to the nature of the image itself. Unlike everyday photographs, the interpretation of image contrast in TEM micrographs is usually not straightforward; multiple contrast mechanisms often contribute to the observation of defect features. Here, we aim at resolving this image-induced challenge by optimizing the image quality. In previous work, we established an experimental protocol for a diffraction contrast imaging scanning transmission electron microscopy (DCI STEM) technique and tailored it specifically for imaging defects in popular iron-based structural alloys [2]. Thus, the *DefectNet* was trained on a small set of high-quality DCI STEM defect images obtained from HT-9 martensitic steels. The performance of the resulting model for each defect was assessed quantitatively by standard semantic segmentation evaluation metrics, and the resulting defect density and size measurements compared to that from a group of human experts.

Figure 1 presents the *DefectNet* semantic segmentation predictions for the development and test sets of the three crystallographic defects. Compared to the ground truth label, the deep learning predicted defect maps show excellent resemblance. In the comparison maps color-coded by the confusion matrix at the pixel level, we can see that the majority of pixels in the prediction map are in black and turquoise and thus correctly classified as the background and the corresponding defect. Table 1 summarizes the semantic segmentation performance of the *DefectNet* on the test sets. The current *DefectNet* was trained over a limited number of labeled DCI STEM images, but it has achieved an excellent prediction performance on the test sets, with an overall averaged pixel accuracy of  $94.61 \pm 1.13\%$ , precision of  $72.12 \pm 2.73\%$ , recall of  $79.22 \pm 3.27\%$ , and region intersection over union (IU) of  $61.79 \pm 2.13\%$ , comparable to state-of-the-art deep learning semantic segmentation algorithm. The success and source of error in DCNN prediction was analyzed for each defect features in terms of the feature representation and quality, and the design of deep CNN architecture. Compared with the defect quantification results produced by human experts, computer-based defect analysis is overall more accurate, reproducible, and at least two orders of magnitude more efficient.

## References:

- [1] G Roberts, SY Haile, R Sainju, D Patel, DJ Edwards, B Hutchinson and Y Zhu. *in preparation*  
 [2] Y Zhu, C Ophus, MB Toloczko and DJ Edwards, *Ultramicroscopy*, **193**(2018) 12-23.  
 [3] Authors thank Dr. Colin Ophus, Mr. Deep Patel and Mr. Simon Y. Haile for supporting image labeling and CNN training. We acknowledge funding from the U.S. DOE Office of Fusion Energy Sciences under contract DE-AC05-76RL01830, Office of Nuclear Energy's Nuclear Energy Enabling Technologies program project CFA 16-10570, PNNL Research Computing program and the Nvidia Corporation for the GPUs used in this research.



**Figure 1.** Pixelwise *DefectNet* semantic segmentation prediction of line dislocations, precipitates and voids using DCI STEM images. The comparison maps are color-coded based on the confusion matrix including True Positive (turquoise), True Negative (black), False Positive (red) and False Negative (yellow) at each pixel for both development and test sets.

<i>DefectNet</i> Performance	Pixel accuracy	Precision	Recall	IU
<b>Dislocations</b>	91.60±1.77%	55.37±2.22%	69.10±1.93%	44.34±0.63%
<b>Precipitates</b>	93.39±1.00%	72.06±4.44%	78.38±2.05%	59.85±2.07%
<b>Voids</b>	98.85±0.56%	89.17±1.28%	90.17±5.84%	81.19±3.68%
<b>Defect Overall</b>	<b>94.61±1.13%</b>	<b>72.12±2.73%</b>	<b>79.22±3.27%</b>	<b>61.79±2.13%</b>

**Table 1.** Semantic segmentation performance of the *DefectNet* for the crystallographic defects in steel. The number reported here are averaged metrics with standard deviation over test sets.