doi:10.1017/S1431927621003482
Microsc. Microanal. 27 (Suppl 1), 2021
© Microscopy Society of America 2021

Deep Learning Enabled Atom-by-Atom Analysis of 2D materials on the Million-Atom Scale

Chia-Hao Lee[1], Abid Khan[1], Di Luo[1], Chuqiao Shi[2], Yue Zhang[1], M. Abir Hossain[1], Arend van der Zande[1], Bryan Clark[1] and Pinshane Huang[1]

[1]University of Illinois at Urbana-Champaign, United States, [2]Rice University, United States

Two-dimensional (2D) materials present an unprecedented opportunity to probe structure-property relationships; in projection we can characterize almost every individual atom, making 2D materials well-suited for high precision defect analysis. In particular, understanding the types and densities of point defects in 2D materials is important for producing 2D materials that can achieve theoretical limits of charge mobility and quantum yield for high quality optoelectronics [1]. Recently, machine learning techniques have been widely applied to electron microscopy data for defect identification, image segmentation, and image denoising [2-6]. Machine learning also opens up new possibilities to analyze large datasets of atomic resolution information. For example, we have previously shown that machine learning methods can be used to circumvent fundamental limitations from electron beam damage and achieve sub-pm precision measurements of defect structures in 2D transition metal dichalcogenides (TMDC) [6]. Here, we use a combination of automated acquisition, electron scattering simulations, and machine learning to acquire, process, and analyze aberration-corrected scanning transmission electron microscopy (STEM) images of 2D materials. We analyzed 56,814 point defects in an atomic resolution, ~million atom dataset to determine the precise atomic structures, spatial distributions, and concentrations of 7 different types of point defects in a 2D TMDC ($WSe_{2-2x}Te_{2x}$).

To achieve robust defect identification, a key step is generating high quality, realistic training data to train the deep learning models (fully convolutional networks, FCN). First, we simulate semi-quantitative annular dark-field (ADF) STEM images with a wide range of microscope conditions including aberrations and probe current variations. Next, we include common imperfections presented in real experimental data, such as linear and non-linear drift distortion, scan jittering, Poisson and Gaussian noise. Moreover, we added low frequency variation to emulate surface contamination. These steps generate realistic STEM images (Figure 1) for a robust FCN training, which achieves a 98% true positive rate. The FCN trained using simulated data performs comparably to an FCN trained on human-labeled experimental data, but with much less manual labor and human intervention required. The FCNs identified 56,814 defect sites, or roughly 10% of the atomic columns measured. The identified defects are mostly Te substitutions (86.6%) and Se vacancies (12.5%). Class averaged [6] images of each defect type are shown in Figure 2. Besides the common Te substitutions and Se vacancies, the FCNs also discover rare metal-site defects such as Se_W antisite defects (0.8%) and W vacancies (0.08%) that are 2 orders of magnitude less frequent than the Se vacancies. Our technique provides detailed atomic structures, spatial distributions, and concentrations of each defect type simultaneously, which are often inaccessible by optical techniques [7]. Although we demonstrate the technique with 2D TMDC, this workflow can in principle be applied to other 2D materials and scanning probe techniques, paving the way towards large-scale, all-atom analysis of 2D materials [8].

https://doi.org/10.1017/S1431927621003482 Published online by Cambridge University Press

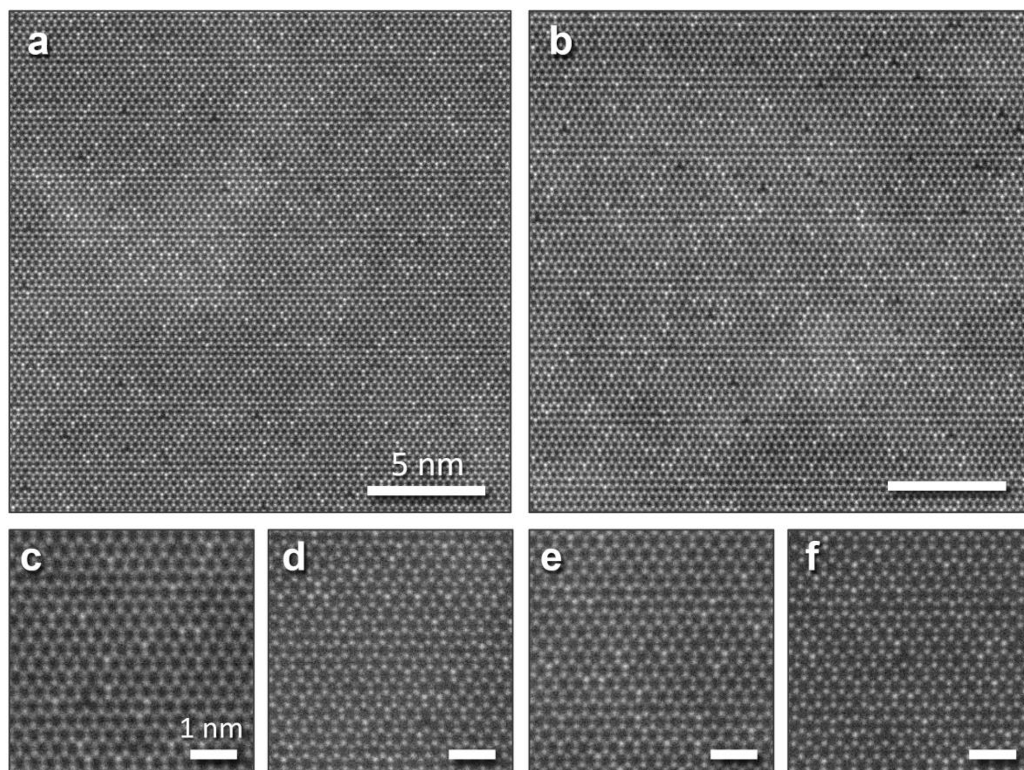
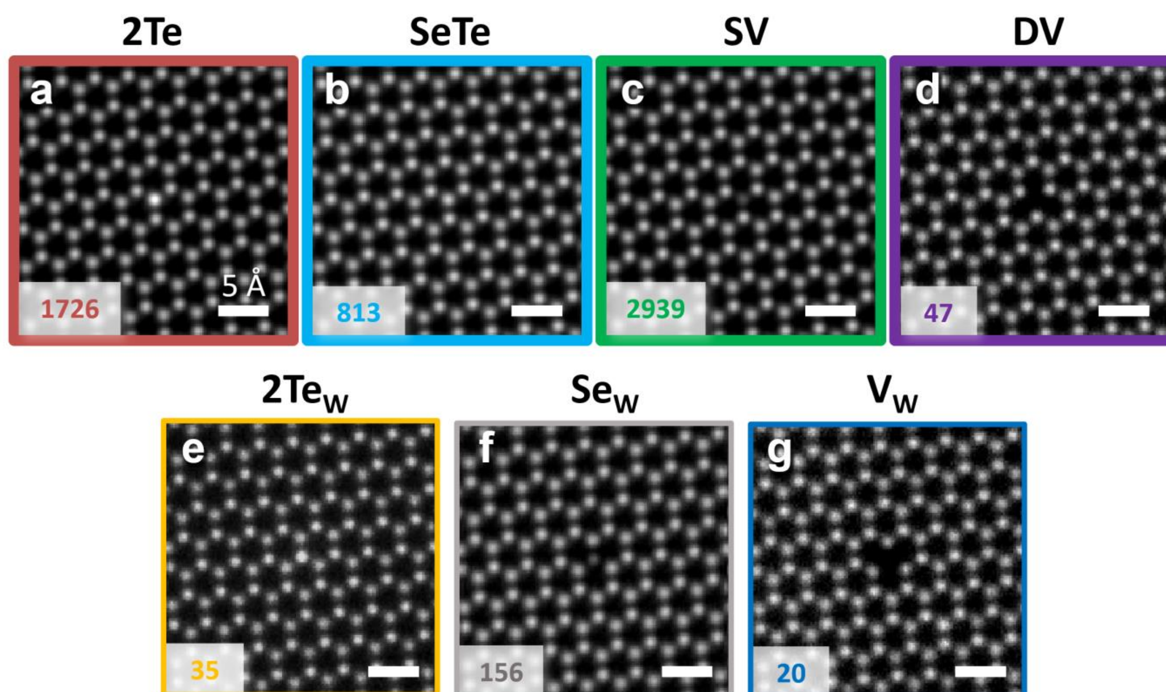


Figure 1. Figure 1: Simulated ADF-STEM images generated by our workflow for FCN training, including electron counting noise, aberrations, detector noise, image distortions, and surface contamination. Representative images with (a-b) low magnification and (c-f) high magnification.



https://doi.org/10.1017/S1431927621003482 Published online by Cambridge University Press

Figure 2. Figure 2: Class-averaged images of FCN-identified atomic defects segmented from experimental STEM datasets of WSe2-2xTe2x. Other than the class averaging, we did not apply any sort of smoothing, filtering, or probe deconvolution to our STEM images. (a-d) Chalcogen-site defects. From left to right: Double Te substitution (2Te), single Te substitution (SeTe), single Se vacancy (SV), double Se vacancy (DV). (e-g) Metal-site defects. From left to right: Double Te antisite (2TeW, two overlapping Te atoms residing at W sublattice), single Se antisite (SeW), W vacancy (VW). The number of images summed is labeled at the bottom left corner of each image.

References

1. D. Edelberg et al., Nano Lett. 19.7 (2019), p. 4371-4379
2. M. Ziatdinov et al., ACS Nano 11 (2017), p. 12742-12752
3. W. Li et al., npj Comp. Matl. 4.1 (2018), p. 1-9
4. J. Madsen et al., Adv. Theory Simul. 1 (2018), p. 1800037
5. R. Lin et al., arXiv: 2012.09093 (2020)
6. C.-H. Lee et al., Nano Lett. 20.5 (2020), p. 3369-3377
7. Q. Qian et al., Nanoscale 12.3 (2020), p. 2047-2056
8. This work was supported by the DOE award number DE-SC0020190 and was carried out in part in the Materials Research Laboratory at UIUC.

https://doi.org/10.1017/S1431927621003482 Published online by Cambridge University Press