doi:10.1017/S1431927622011424
Microsc. Microanal. 28 (Suppl 1), 2022

Automated Acquisition and Deep Learning of 2D Materials on the Million-Atom Scale

Chia-Hao Lee[1], Abid Khan[2], Yue Zhang[3], M. Abir Hossain[3], Arend van der Zande[3], Bryan Clark[2] and Pinshane Huang[1]*

[1.] Department of Materials Science and Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA.
[2.] Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL, USA.
[3.] Department of Mechanical Science and Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA.
* Corresponding author: pyhuang@illinois.edu

New approaches for data acquisition and information extraction are critical to fully exploit large datasets and unveil the hidden information that are normally overlooked and discarded. In our work, we develop a series of techniques including automated acquisitions of scanning transmission electron microscopy (STEM) images, realistic training data enhanced by generative adversarial network (GAN), defect identification with deep learning, and class averaging for high signal-to-noise ratio (SNR) images to explore the realm of all-atom analysis of 2D materials [1].

First, we build a scripting framework using *pywinauto* to integrate different functionalities from multiple control software. For example, electromagnetic lens settings, stage control, aberration correction, and acquisition parameters. This allows us to automatically acquire atomic-resolution datasets on the million-atom scale, or ~μm^2 size within a day. These large datasets contain valuable information but also significant variations in experimental parameters that are difficult to model. Next, we construct a CycleGAN-enhanced workflow to generate realistic training data from both simulated and experimental images. The CycleGAN extract subtle features from the experimental images and apply it onto the simulated images, outputting high-quality, experiment-like simulated data with the benefits of being automatically labeled. The resulting high-quality training data enables accurate defect identification using a fully convolutional network (FCN) [2]. The performance is comparable to or even exceeding human labeling with much less human intervention. We examine hundreds of atomic resolution images with FCN and identify tens of hundreds point defects, including vacancies, substitution, and antisite defects from 2D materials. The defect distribution and statistics are used to correlate with its photoluminescence yield, showing an inversely proportional relation across the µm scale. We further perform class averaging to generate high SNR images from nominally identical defects, which allows us to circumvent the fundamental limitations set by electron beam damage and achieve 0.3 pm precision of the 2D coordinates of these defects. This technique also enables direct observation of pm-scale strain field oscillations around the defect core and how nearby defects interact with each other through coupled strain fields. Moving forward, these techniques may pave the way towards self-driving microscopy, unsupervised feature identification, and high-precision analysis of beam sensitive samples [3].

References:

[1] CH Lee et al., Nano Letters **20** (2020), p. 3369.
[2] Code available at: https://github.com/ClarkResearchGroup/stem-learning
[3] This work is supported primarily through DOE BES under DE-SC0020190. Additional support from

AFOSR FA9550-7-1-0213. This work was carried out in part in the Materials Research Laboratory at UIUC.