

Miti Shah
EE367
2/20/2026

Strain reconstruction of nanobeam 4d-STEM datasets using a machine learning approach

Background

Transmission electron microscopy (TEM) is a technique that allows for high resolution imaging of materials at atomic length scales. It uses electrons as the imaging mode (instead of light) and measures the scattered electrons as they pass through a thin sample. 4D scanning transmission electron microscopy is a detector development to the TEM with which a diffraction pattern (specifically a converged beam electron diffraction or CBED pattern) can be recorded at each real space pixel position. The name 4d is a description of the dataset accumulated since you get a 2d diffraction pattern at each position in a 2d real space grid giving a 4d dataset. This means that you can get local material structural information since you can analyze the diffraction pattern at each pixel position for a certain material property. Then you can repeat it for the whole grid of patterns and obtain a real space map of the material property in the size of the grid detector. I plan to focus on calculating strain of materials using these 4d-stem datasets. Understanding strain characteristics of materials has various applications since it can determine semiconductor device failure, material defects and interfacial properties. Therefore being able to obtain accurate and high precision strain mapping is proving more important. [1]

Related Work

There are many algorithms that use 4D-STEM to obtain a strain of a material. One such algorithm is called template matching. For each CBED pattern, the disk positions are found using cross correlation. The positions are then used to get the basis u and v vectors of the pattern. These basis vectors change based on the compressive or tensile strain at each pixel position. Therefore, once the basis vector is calculated at each position, the change from a reference is used to generate a strain map. This technique can fail due to false detections and errors in disk position due to noise in the pattern. [2]

Another technique uses a cepstral transform to determine the basis vectors. The cepstral transform is the fourier transform of the log of the image. The basis vectors are determined through finding the sub pixel peak positions of the bragg peaks. This method varies in accuracy of strain maps based on its application. For example, if only the two first-order peaks are used to refine the basis vectors, the resulting strain map is very noisy due to imprecision in sub pixel peak detection, especially given noise. However, if all the peaks of the averaged pattern are used to refine the basis vectors then a more robust strain map is obtained since the small peak position errors are reduced through using more than one peak. [3]

Due to the variability with current methods, I plan on using machine learning as a way to denoise and analyze CBED patterns for better strain calculation. There has been a paper which

used a FCU-Net to denoise CBED patterns. This focused on denoising the entire CBED pattern and decomposing it into the dynamical diffraction pattern for easier strain analysis [4].

Project Goals

I plan on developing a supervised machine learning algorithm that improves strain mapping precision. Current methods are unstable especially when the images are noisy. Therefore, I think given the right training dataset and loss function, a CNN could learn to essentially denoise the datasets and allow for more accurate peak position localization. The training and validation datasets would need to be simulated to model electron diffraction patterns. Current packages exist which enable this for a variety of different crystal systems such as abTEM [5]. Validation using experimental datasets would prove to be difficult since there is no ground truth. Therefore mainly simulated data would be used for training. Experimental dataset can be used for final algorithm tests where seeing how accurate the strain maps are can give an idea of the algorithm success.

Timeline and Milestones

Week 1:

- Generate training datasets using multislice simulation of various crystals
- Test different CNN structures and loss functions to determine best combination for given problem

Week 2:

- Tune training data and generate more inputs and labels either by adding more crystal structures or augmenting current data
- Finalize CNN structure and test different parameters for better convergence and results

Week 3:

- Run training and validation on various datasets for a large number of iterations
- Prepare final report and poster presentation

References and related works

1. Colin Ophus. *Microscopy and Microanalysis* (2019) Volume 25, Issue 3 563–582.
<https://doi.org/10.1017/S1431927619000497>
2. Pekin T C et al. *Ultramicroscopy* (2017) 176 170-176.
<https://doi.org/10.1016/j.ultramic.2016.12.021>
3. Padgett E et al. *Ultramicroscopy* (2020) 214 112994.
<https://doi.org/10.1016/j.ultramic.2020.112994>
4. Munshi J et al. *NPJ Computational Materials* (2022) 8 254.
<https://doi.org/10.1038/s41524-022-00939-9>
5. Madsen J and Susi T. *Microscopy and Microanalysis* (2020) Volume 26, Issue S2 448-450.
<https://doi.org/10.1093/mam/ozaf103>