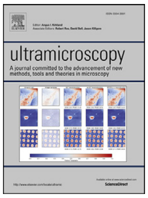
Ultramicroscopy 188 (2018) 59–69

Contents lists available at ScienceDirect

Ultramicroscopy

journal homepage: www.elsevier.com/locate/ultramic



A deep convolutional neural network to analyze position averaged convergent beam electron diffraction patterns

W. Xu, J.M. LeBeau[*]

Department of Materials Science and Engineering, North Carolina State University, Raleigh, NC 27695, USA

ARTICLE INFO

Article history:
Received 18 July 2017
Revised 26 February 2018
Accepted 2 March 2018
Available online 5 March 2018

Keywords:
Machine learning
Convolutional neural networks
Position averaged convergent beam electron diffraction (PACBED)
Automation

ABSTRACT

We establish a series of deep convolutional neural networks to automatically analyze position averaged convergent beam electron diffraction patterns. The networks first calibrate the zero-order disk size, center position, and rotation without the need for pretreating the data. With the aligned data, additional networks then measure the sample thickness and tilt. The performance of the network is explored as a function of a variety of variables including thickness, tilt, and dose. A methodology to explore the response of the neural network to various pattern features is also presented. Processing patterns at a rate of ∼0.1 s/pattern, the network is shown to be orders of magnitude faster than a brute force method while maintaining accuracy. The approach is thus suitable for automatically processing big, 4D STEM data. We also discuss the generality of the method to other materials/orientations as well as a hybrid approach that combines the features of the neural network with least squares fitting for even more robust analysis. The source code is available at https://github.com/subangstrom/DeepDiffraction.

1. Introduction

For a highly convergent and coherent, ångström-sized electron probe, the corresponding convergent beam electron diffraction (CBED) disks strongly overlap to form a complex interference pattern. These patterns depend sensitively upon the position of the probe within the unit cell, but by averaging these patterns together, a position averaged CBED (PACBED) pattern is created [1]. The patterns then depend strongly on sample thickness and tilt, and also reveal crystal polarity, changes in composition, octahedral distortions, and strain [2–8].

While PACBED patterns have been shown to be very sensitive to nanometer-level sample thickness differences and sub-milliradian tilt [1], the patterns change with sample thickness in a non-intuitive way due to dynamical diffraction. Even so, visual inspection is often sufficient to match experimental PACBED to a library of simulated ones. This process, however, is inherently subjective and time-consuming. To reduce human error and enhance the repeatability of the measurements, a semi-automated approach is usually employed. To this end, least square fitting (LSF) has been the primary tool [3,4,8–10]. The parameter of interest, e.g. thickness, tilt, polarity, etc., is found by searching for the best fit amongst a library of patterns. While LSF can be precise and accurate [10], it is often time-consuming to fit a broad range of parameters and avoid local minima. Beyond processing speed, patterns alignment is an additional limiting factor. Generally, some pretreatment of the data is required by the user to locate the precise pattern center and calibrate the pixel size/scale. Specimen tilt complicates this analysis by displacing the center of intensity mass from its true position. Furthermore, alignment is obfuscated by significant CBED disk overlap, which precludes the use of a Hough transform to locate the disk centers [10–13].

Beyond brute force methods, convolutional neural networks (CNNs) have enabled breakthrough image recognition performance, even within very complex scenes [14–21]. For example, CNNs have become the standard for applications ranging from face recognition to self-driving cars. By combining multiple, deep convolutional layers with an appropriate training set, a CNN can automatically “learn” high-level representations needed for robust image classification. While neural networks have shown promise for electron microscopy analysis, these powerful tools have only recently begun to be applied [22,23]. This is particularly relevant to automated PACBED analysis, as these networks have the potential to overcome many of the limitations of other methods.

In this work, we develop a set of deep CNNs to automatically analyze PACBED patterns, extracting pattern size, center, rotation, specimen thickness, and specimen tilt. The training and processing speeds are accelerated by the implementation of GPU calculations. Further, we show that the network architecture enables fully automatic PACBED analysis without the need for human

[*] Corresponding author.
E-mail address: jmlebeau@ncsu.edu (J.M. LeBeau).

https://doi.org/10.1016/j.ultramic.2018.03.004

supervision. The approach is compared to LSF using the same datasets and is found to be orders of magnitude faster after training. Finally, we report various observations including application to 4D STEM datasets, generalizability of the trained networks to other materials, and a hybrid approach to combine neural networks with LSF for fast, robust analysis of additional parameters.

2. Materials & methods

2.1. Experiment

$SrTiO_3$ single crystals oriented along [100] and $PbMg_{1/3}Nb_{2/3}O_3$, oriented along [110], were used throughout this study. The crystals were thinned to electron transparency using wedge-polishing and low energy ion-milling using a Fischione 1050 Ar ion mill. The PMN sample was carbon coated to reduce sample charging. A probe-corrected FEI Titan G2 STEM microscope was operated at 200 kV with probe convergence semi-angle of either 13.6 mrad and 19.1 mrad. The former convergence semi-angle is consistent with an uncorrected STEM, and also has many advantages in atomic resolution EDS elemental quantification [9]. The latter is consistent with that used for aberration-corrected, sub-angstrom STEM imaging. PACBED patterns were recorded using a Gatan UltraScan 1000XP CCD camera.

$SrTiO_3$ PACBED patterns from experiment were used to create a database for performance testing. Using the 13.6 mrad probe, a total number of 231 PACBED patterns were recorded from regions 6–120 nm thick. For the 19.1 mrad probe, a total of 156 PACBEDs were captured at thicknesses ranging from 8 to 70 nm. In both cases, the patterns exhibited random tilts up to ~ 4 mrad. The CCD acquisition time was varied from 0.1 to 5 s, with a probe current of about 80 pA to incorporate varying levels of noise into the database. A 10×10 4D STEM dataset was collected over a $60 \times 60\,\mathrm{nm}^2$ region of the sample with an acquisition time of 1 s/pattern.

2.2. Simulation

To establish a library for neural network training, PACBED patterns were simulated using the Bloch wave method. The Many-Beam dynamical-simulations and least-squares FITting (MBFIT) software was used for this purpose [24]. Note that the original MBFIT source code was modified to generate the PACBED output with overlapped diffraction disks [6]. Patterns were calculated in 1 nm increments with thicknesses ranging from 1 to 120 nm at 13.6 mrad, and 1–80 nm for 19.1 mrad. At each thickness, a tilt series was also simulated with up to 4 mrad tilt along [100] and [010]. The tilts were separated by 0.25 mrad when tilt was <1 mrad, and 0.5 mrad otherwise. In total, 4560 and 3040 PACBED patterns were simulated for 13.6 mrad and 19.1 mrad, respectively.

2.3. Convolutional neural network

The convolutional neural networks applied here are based on the AlexNet architecture, a description of which can be found in Ref. [25]. As seen in Fig. 1, the CNN applied consists of multiple convolutional, nonlinear activation and pooling layers in its network architecture, followed by fully connected layers and softmax to conduct tasks for image classification. Each convolutional layer takes an input image (or feature map) (X^l) that is convolved with a bank of convolution kernels to produce new feature maps (X^{l+1}). In this way, pattern features are found through this convolutional process. The value at each pixel of the k^{th} feature map, $X^{l+1}_{i^{l+1},j^{l+1},k}$, is given by Eq. (1), with unity stride and zero padding.

$$X^{l+1}_{\mathrm{i}^{l+1},\mathrm{j}^{l+1},k} = \sum_{i=0}^{m-1}\sum_{j=0}^{m-1}\sum_{d^l=0}^{D^l} w_{i,j,d,k} X^{l}_{i^{l+1}+i,j^{l+1}+j,d^l} + b^l_k \tag{1}$$

where $w_{i,\ j,\ d,\ k}$ and b^l_k are weights and bias of the kth kernel. The kernels are $m \times m \times D^l \times D^{l+1}$ in size. D^l and D^{l+1} represent the channel size of the input and number of kernels in the convolutional layer, respectively. The output feature image X^{l+1} has the same number of channels as the number as convolutional kernels in D^{l+1}, representing a collection of convolved feature maps that aim to capture different image features from the input. Moreover, the convolutional kernel weights are shared among the feature maps, which provides feature location invariance and leads to reduced network complexity.

The convolved map, X^{l+1}, is then passed through a quasi-linear activation function. In this study, a rectified linear unit (ReLU) function, $\max\{0, X^{l+1}_{\mathrm{i,j,k}}\}$, is implemented after each convolutional layer [26,27]. After ReLU, a channel-wise local response normalization is applied to the feature maps, aiming to improve network generalization [25]. The normalized feature maps are down-sampled via max-pooling, where the map is sub-divided and only the maximum within each sub-region is retained. The pooling process not only controls the number of parameters in the neural network, but also makes the feature detection less dependent on scale and/or orientation [28].

After max-pooling, the feature map is treated as input for the next convolutional layer, where the maps become sensitive to higher level features. In the complete (deep) CNN, multiple levels of convolutional–ReLU–normalization–max-pooling layers are sequentially stacked. This process abstracts features of different ‘length-scales’ across the image and across the field of view. These abstracted units are then connected to neurons within the fully connected layers that reorganize the image features for identification, which is analogous to biological neurons. As a final step, the softmax activation layer converts the classified outputs to probabilities. For error backpropagation during the network training, a cross-entropy loss function is used [29].

2.4. PACBED measurement algorithm

The overall neural network configuration to measure PACBED parameters can also be seen in Fig. 1, which contains five separate CNNs trained. In the first stage, the zero order disk size, disk center, and pattern rotation angle are measured. The flow of the procedure is illustrated in Fig. 2. There are a number of automated procedures that are applied before passing to the CNNs. Rough estimates of the of the pattern center and size are provided by fitting the integrated intensity of the PACBED pattern along both horizontal and vertical direction to Gaussian functions. Precision here is not essential as these variables are iterated.

The roughly centered and cropped patterns are then passed to subsequent CNNs to refine the center and shift measurements. To measure shift along both horizontal and vertical directions, the same CNN is used, but with the pattern rotated 90°. Updated center and size variables are used to realign the original PACBED dataset until convergence. The rotation angle is then determined via another trained CNN, but without the need for iteration. Prior to thickness and tilt measurements, uniform background is subtracted from the pattern to account for the contribution from inelastic scattering. This improves the network performance, particularly when determining sample thicknesses above ~ 70 nm. For more details, see Section 4.1. As part of this process, the background subtraction parameter is converged while determining thickness and tilt. It is also important to note that only positive tilt values along [100] and [010] are measured using the CNN. This is justified by the four-fold symmetry for the zone considered. We first align the patterns to the same rotation angle and then measure the sign of the tilt by finding the pattern quadrant that has the greatest intensity. This approach greatly reduces the