doi:10.1017/S1431927622011321
Microsc. Microanal. 28 (Suppl 1), 2022

A Semi-Supervised Machine Learning Workflow to Extract Quantitative Insights From Ultrafast Electron Microscopy Datasets

Arun Baskaran[1*], Faran Zhou[2], Thomas E. Gage[1], Haihua Liu[1], Ilke Arslan[1], Haidan Wen[2,3], Maria K.Y. Chan[1*]

[1.] Center for Nanoscale Materials, Argonne National Laboratory, Lemont IL, USA.
[2.] X-ray Science Division, Argonne National Laboratory, Lemont IL, USA.
[3.] Material Science Division, Argonne National Laboratory, Lemont IL, USA.
* Corresponding author: abaskaran@anl.gov, mchan@anl.gov

Ultrafast Electron Microscopy (UEM) has emerged as a novel characterization technique to study the time-resolved dynamics of nanoscale phenomena [1]. The temporal resolution obtained in this technique, ~600 fs, enables the real-space imaging of phenomena such as coherent acoustic phonons [2] and plasmons [1]. A typical image dataset generated during a UEM experiment includes a series of time-indexed images, along with metadata regarding the pixel resolution and the time of impact of the stimulating beam. This series of images together describe the dynamic behavior of the nanoscale feature(s) relevant to the study. During post-processing, one of the predominant motivations for analysis pertains to the quantification of this dynamic behavior. In previous studies, such analyses have been conducted using manual methods, such as the space-time contour plots in [3]. The emergence of machine learning (ML) techniques as a supplementary toolbox presents an opportunity to automate the post-processing of UEM images, reduce human-induced bias in the processing steps, and decrease the turnaround time for analysis.

(a) (b)

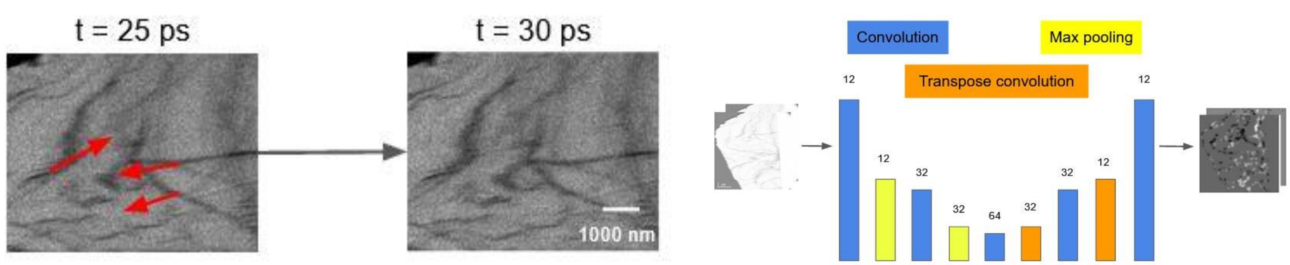

Figure 1. (a) A major motivation for post-processing analysis pertains to the quantification of the dynamics of nanoscale features. The direction of motion of 3 selected phonon wavefronts during this time-period is shown using red arrows. (b) An overview of the neural network architecture used in this work. The input and the output are of the same dimensions. The number of trainable filters in each layer of the network is denoted on top of the layer.

In this study, we have implemented a ML based workflow to extract quantitative information regarding the motion of coherent acoustic phonons that propagating through a characteristic two-dimensional material. The specific details regarding the material and the experiment setup will be discussed elsewhere, but this workflow is generalizable to other typical UEM datasets. Given an image dataset, the dynamic behavior of features in the dataset is quantified in this workflow by computing pixel-level optical flow vectors [4] between image pairs, which maps a pixel in the first image to a displacement vector corresponding to its position in the second image. Two types of image pairs are constructed: using successive images and using a base image (taken before stimulation) and an image taken after the time of stimulation. The use of machine learning for optical flow computation [5] has gained popularity in the last few years owing to the development of advanced machine learning models and access to better computing resources. In the current work, a neural network modeled after U-net [6] is trained under a semi-supervised learning algorithm to compute pixel-level optical flow between image pairs. The model architecture is presented in Figure 1. The model takes as input a pair of grayscale images. The input images are fed forward through an encoder block, during which a lower-dimensional feature representation of the image pair is learnt by the model. In the decoder part of the model, the encoded representation is progressively upsampled back to the input dimensions. The target output of the model is the pixel-level optical flow vectors corresponding to the input image pair. A major challenge to the workflow is the lack of manually annotated ground truth data. To overcome this, pseudo-ground truth labels are generated using the optical flow computation provided as part of the OpenCV library [7]. The 'relevant' pixels are isolated in a pre-processing step by analyzing the distribution of peaks in the Fast Fourier Transform (FFT) of the time-profiles of pixel intensities. This unique pre-processing step improves the accuracy of optical flow computation. The model is pre-trained in a supervised manner using this pseudo ground truth. Following this, the model is fine-tuned in an unsupervised manner. The loss function that is implemented during the fine-tuning step measures the ability of the model to reconstruct the target image using the source image and the computed optical flow vector.

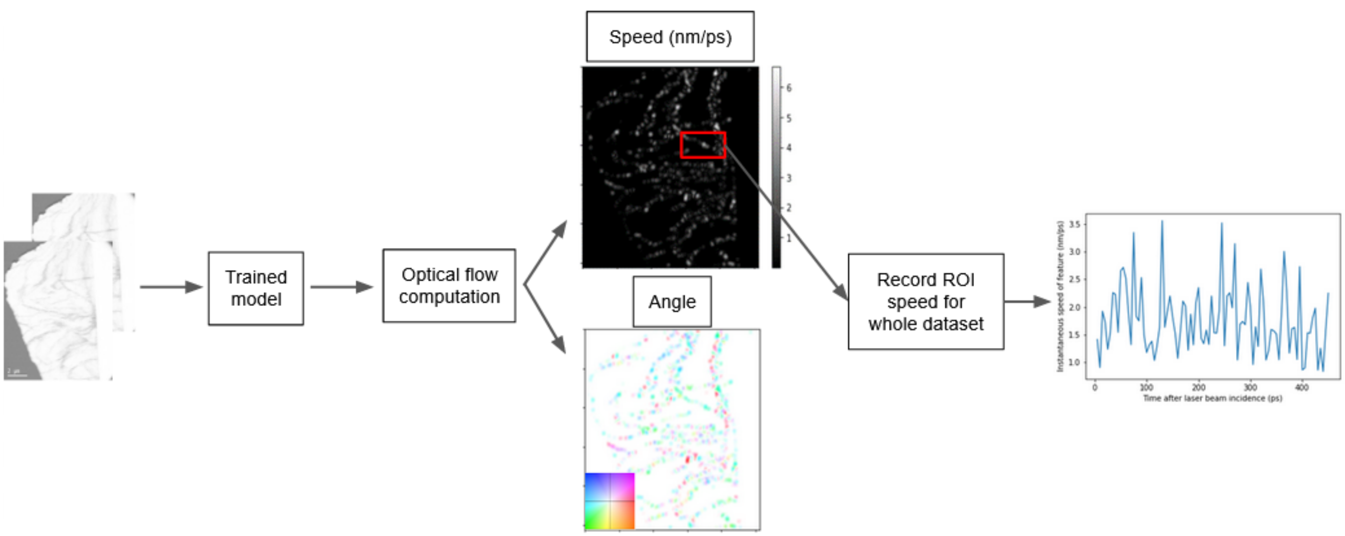

Figure 2. A characteristic example of how the workflow is used to extract meaningful inference from any given image pair in the dataset. The optical flow vector is transformed to pixel-level speed and angle of motion. Furthermore, a feature vector can be constructed for any given region of interest (ROI) using the raw output and the transformed output. Work is in progress currently to explore the utility of such feature vectors for downstream classification tasks.

A characteristic way to interpret an output from the trained neural network is shown in figure 2. The optical flow vector that is obtained at every pixel is transformed to more intuitive measurements such as speed and direction. This body of work presents an opportunity for deeper analysis by leveraging the optical flow computations performed for individual image pairs. For example, as shown in the right-half of figure 2, a feature vector for a selected region of interest (ROI) can be constructed by concatenating the average speed within the ROI from successive image pairs. This feature vector can be used for further downstream analysis. Currently, work is in progress to explore the utility of such a feature vector towards training a classification model to identify the different modes of dynamic behaviors exhibited by the acoustic phonons. This workflow is generalizable to a real-space image dataset obtained from UEM, and is a step towards automated analysis of motion quantification of nanoscale features [8].

References:

[1] H Liu et al., Nano Letters **21** (2021). https://doi.org/10.1021/acs.nanolett.1c01824
[2] D Cremons, DA Plemmons and DJ Flannigan, Structural Dynamics **4** (2017). http://dx.doi.org/10.1063/1.4982817
[3] SA Reisbick, Y Zhang and DJ Flannigan, Journal of Physical Chemistry A **124** (2020) https://dx.doi.org/10.1021/acs.jpca.9b12026
[4] M Zhai et al., Pattern Recognition **114** (2021). https://doi.org/10.1016/j.patcog.2021.107861
[5] J Hur and S Roth, arXiv:2004.02853 (2020).
[6] O Ronneberger, P Fischer and T Brox, arXiv:1505.04597
[7] opencv/opencv: Open Source Computer Vision Library, https://github.com/opencv/opencv (accessed 02/07/2022)
[8] This work was performed at the Center for Nanoscale Materials, a U.S. Department of Energy Office of Science User Facility, and supported by the U.S. Department of Energy, Office of Science, under Contract No. DE-AC02-06CH11357. In addition, this research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory.

https://doi.org/10.1017/S1431927622011321 Published online by Cambridge University Press