doi:10.1017/S1431927620013100
Microsc. Microanal. 26 (Suppl 2), 2020

High Throughput Crystal Structure Classification

Jess Tate[1], Jeffery Aguiar[2], Matthew Gong[1] and Tolga Tasdizen[1]

[1]University of Utah, Salt Lake City, Utah, United States, [2]Idaho National Laboratory, Idaho Falls, Idaho, United States

Advances in imaging and experimental techniques have made materials science abundant, increasing the potential for new discoveries but requiring new techniques to trans- form the vast amount of data into information that can be readily interpreted by researchers. Data-driven models, specifically deep and machine learning, are one set of tools that have been applied with noticeable success to some burdensome materials and microscopy problems. However, the ultimate goal with these models has yet to be realized: to automatically process, characterized, and compressed the increasingly common high- resolution, multi-modal imaging data on a single system. One crucial component of many workflows is classifying the crystal structure of a material, yet is mostly performed by trained expert who divine structural information from minute diffraction pattern variations. [1, 2, 3] The process of determining a crystal's space group often involves a lengthy process requiring fitting to a series of non-linear equations and intimate knowledge of a sample to be performed properly, including standardized approaches such as Rietveld refinement. [4, 5] The heavy dependence in complex matching, time intensive processes, potential for dimensionality reduction, and the vast amount of data generated makes crystal classification an ideal case for automation.

We have recently developed a convolutional neural network (CNN) to classify the crystal structure from material diffraction data. [6] More recent developments of our model focused on improving accuracy with additional material information, such as the chemical composition, and improving the robustness of the model by training the model to perturbations in the location and number of peaks in the diffraction pattern [7]. These development efforts are crucial to the ultimate goal of distributing our model as a community resource to facilitate the processing of large volume diffraction data. To that end, we present the application of our crystal structure classification tool in two types of high-throughput applications: identifying potential structures within a sample of mixed composition and data recorded at many time steps (Figure 1).

To showcase how our data-driven model can help identify structures within a mixed composition, we applied our classification tool to a series of selected area electron diffraction patterns (SAED) obtained during $^{177}Lu \rightarrow {}^{177}Hf$, a β-decay process. Our model can predict multiple structures within material from a signal diffraction pattern by generating an exhaustive list of possible combinations of its peaks, predicting the structure based on each subset of peaks, and accumulating the predictions to identify likely structures. Space-groups that receive any prediction may exist in the sample, yet multiple predictions, especially of rare space-groups, may provide evidence of unexpected crystal structures that exist in the material. Some of the structures predicted by this approach include those which correspond to known phases of the known starting, transition, and end phased of the reaction (Figure 2). One space-group that the model predicted is 216 (f43m), which is notable because it points to the presence of a theoretical transition compound: $^{177}Lu_xHf_{1-x}O_3$ (Figure 2). This example also highlights the use of our model in a time-dependent data set, as the combinatorial prediction was applied throughout the experiment [8].

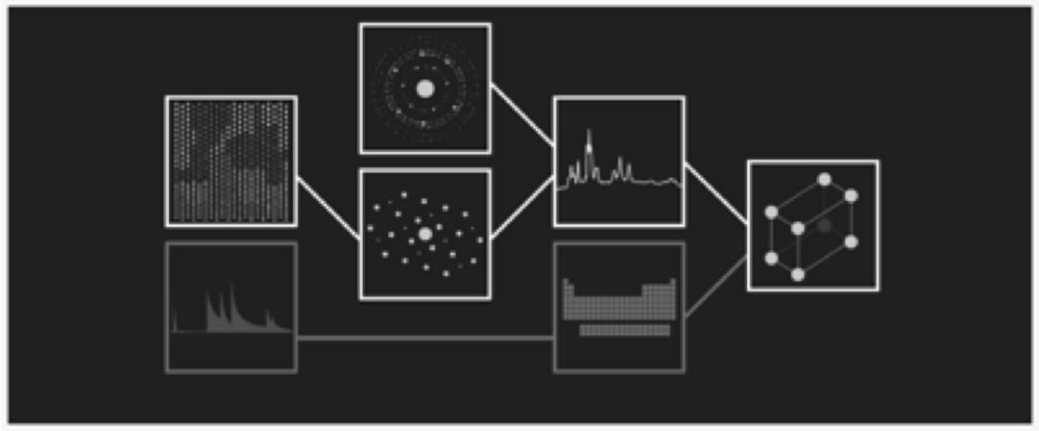

Figure 1. Robust data workflow for high throughput analysis. Illustrated above is a workflow for merging diffraction and chemistry data gathered from various modalities. Leveraging neural networks and machine learning, we have developed toolsets, and workflow models to benefit the community aiding their ability to perform high throughput analysis for several imaging modalities.

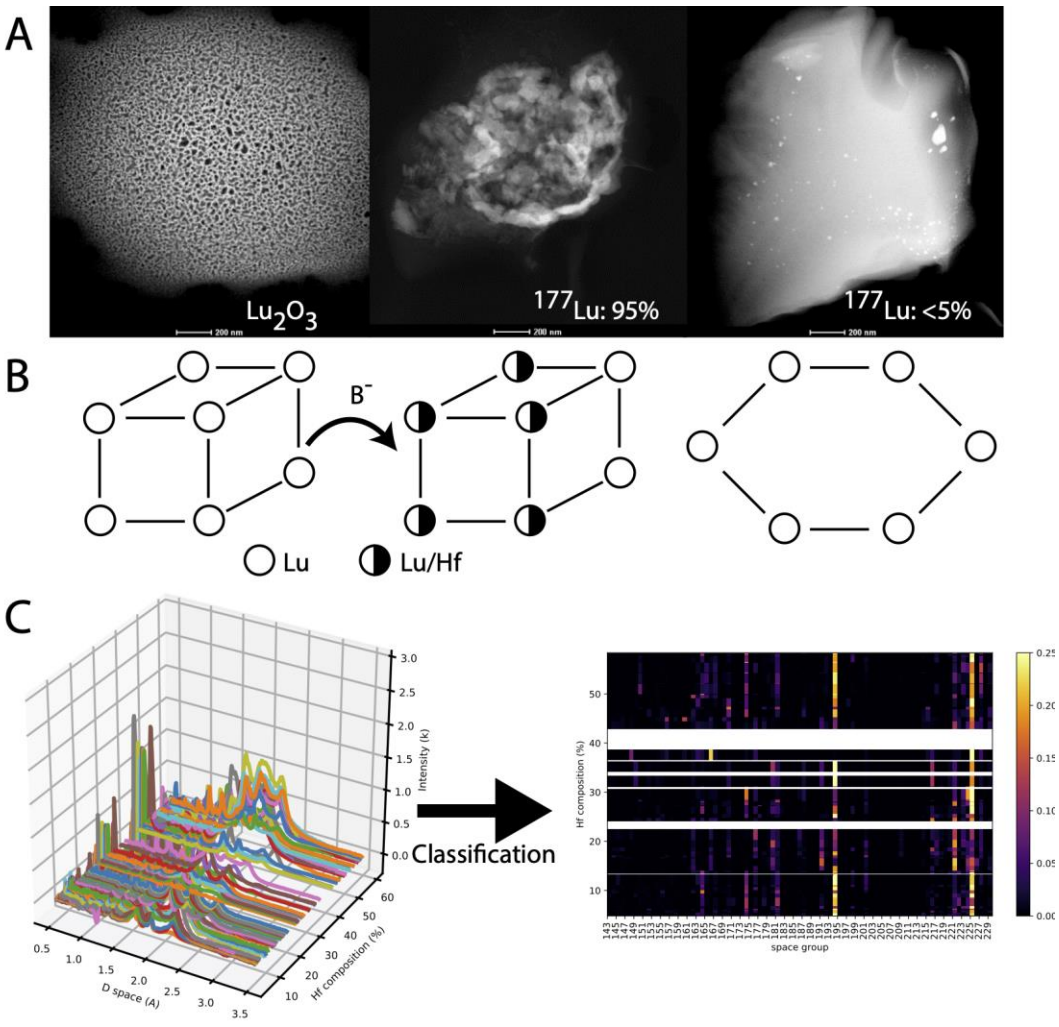

Figure 2. Application of the crystal structure classification tool to 177Lu as it beta decays into 177Hf. A) shows the TEM imaging of the sample at various stages of decay, B) shows probable crystal structures

https://doi.org/10.1017/S1431927620013100 Published online by Cambridge University Press

during the reaction, and C) shows the diffraction patterns and the classifications of possible structures within the sample.

References

1. R. K. Vasudevan, A. Tselev, A. P. Baddorf, and S. V. Kalinin. Big-data reflection high energy electron diffraction analysis for understanding epitaxial film growth processes. ACS Nano, 8(10):10899–10908, 10 2014.
2. R. K. Vasudevan, A. Belianinov, A. G. Gianfrancesco, A. P. Baddorf, A. Tselev, S. V. Kalinin, and S. Jesse. Big data in reciprocal space: Sliding fast fourier transforms for determining periodicity. Applied Physics Letters, 106(9):091601, 2015.
3. J. Dongarra, P. Beckman, T. Moore, and P. Aerts etal. The international exascale software project roadmap. The International Journal of High Performance Computing Applications, 25(1):3–60, 2011.
4. A. A. Coelho. Indexing of powder diffraction patterns by iterative use of singular value decomposition. Journal of Applied Crystallography, 36(1):86–95, Feb 2003.
5. A. A. Coelho. An indexing algorithm independent of peak position extraction for X- ray powder diffraction patterns. Journal of Applied Crystallography, 50(5):1323–1330, Oct 2017.
6. J. A. Aguiar, M. L. Gong, R. R. Unocic, T. Tasdizen, and B. D. Miller. Decoding crys- tallography from high-resolution electron imaging and diffraction datasets with deep learning. Science Advances, 5(10), 2019.
7. J. A. Aguiar, M. L. Gong, and T. Tasdizen. Crystallographic prediction from diffraction and chemistry data for higher throughput classification using machine learning. Computational Materials Science, 173:109409, 2020.
8. Work supported through the INL Laboratory Directed Research & Development (LDRD) Program under DOE Idaho Operations Office Contract DE-AC07-05ID14517. This work was performed, in part, at the Center for Integrated Nanotechnologies, an Office of Science User Facility operated for the U.S. Department of Energy (DOE) Office of Science. Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC. a wholly owned subsidiary of Honeywell International, Inc., for the U.S. DOE's National Nuclear Security Administration under contract DE-NA-0003525. The views expressed in the article do not necessarily represent the views of the U.S. DOE or the United States Government. In part, this research was conducted at the Center for Nanophase Materials Sciences, which is a DOE Office of Science User Facility.