



Accelerating the discovery of new materials with deep learning

Melanie Vollmar*

Diamond Light Source Ltd, Harwell Science and Innovation Campus, Didcot, OX11 0DE, United Kingdom.

*Correspondence e-mail: melanie.vollmar@diamond.ac.uk

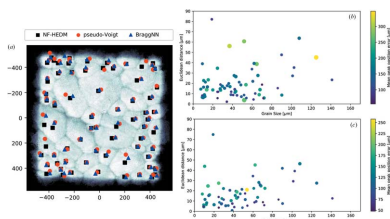
Keywords: high-energy diffraction microscopy; new materials in energy research; deep learning.

Design and development of new materials are crucial to progress and innovation, especially in areas concerned with generation, transmission and use of energy. In recent years, new materials have become even more important, considering the global shift towards renewable energies and sustainability in all areas of the energy sector and the insatiable hunger for energy consumption (Ma *et al.*, 2021; Kordas, 2017).

Commonly, new materials are proposed based on intuition, past experience and studying empirical results. However, pushing the boundaries to explore as yet understudied areas of potential materials is not routinely done. This is due to a combination of factors. On one hand, there is the difficulty of breaking out of established design strategies to identify novel materials, and on the other hand, there is the challenge of accelerating existing imaging techniques and the software required to analyze the acquired experimental data for these new materials. Taking a lead from the developments in machine learning and artificial intelligence, the first issue has been addressed by using a number of different algorithms to design novel materials as is reviewed in Liu *et al.* (2021). To investigate newly designed materials and to verify that the produced sample is consistent with theoretical assumptions, high-energy X-ray characterization methods such as high-energy diffraction microscopy (HEDM) (Park *et al.*, 2017) have become routinely available over recent years. HEDM is based on three-dimensional X-ray diffraction which provides a non-destructive way to analyze the internal grain structure of materials (Poulsen *et al.*, 2001; Poulsen, 2012). Use of diffraction and tomographic imaging elucidates the inner granular structure of materials down to the micrometre level. Nevertheless, the analysis of such data remains computationally expensive. Liu *et al.* (2022) use deep learning to address the second shortcoming of new materials development, the lack of fast data analysis.

While rotation imaging of a sample can be completed within minutes, for example at a synchrotron facility, the analysis of the data is substantially slower. In particular, finding Bragg peaks and determining their precise peak location, as well as reconstructing the material grain structure based on this information, are very time consuming. A Voigt profile in 2D or 3D needs to be fitted to determine the peak location, which is computationally expensive, hence, a pseudo-Voigt is usually used for approximation. After identifying the center-of-mass of the Bragg peaks, the individual grains in the material are reconstructed (Sharma *et al.*, 2012*a,b*). Even so, depending on the nature of the material under investigation and computational resources available, it can take up to several weeks to determine the precise peak positions and reconstruct the grain structure. As such, the slow processing and analysis of HEDM data in material sciences hampers a routine trial-and-error assessment of newly designed materials and dramatically reduces the speed of new discoveries.

In the current issue of **IUCrJ**, Liu *et al.* (2022) have developed a deep neural network based system, *BraggNN*, to precisely locate Bragg peaks in HEDM images. In their machine learning-based model (Fig. 1), Liu *et al.* (2022) build on the established and well documented work around convolutional neural networks (Fukushima, 1980) in image recognition. The aim for this system is to provide a near real-time data analysis and feedback application – greatly reducing the time needed for data analysis compared with standard methods (Abeykoon *et al.*, 2019). Its sole purpose is to rapidly identify the center-of-mass of a Bragg peak to sub-pixel precision. Fig. 2 compares three methods, near-field HEDM, pseudo-Voigt far-field HEDM and *BraggNN*, and how well each



OPEN ACCESS

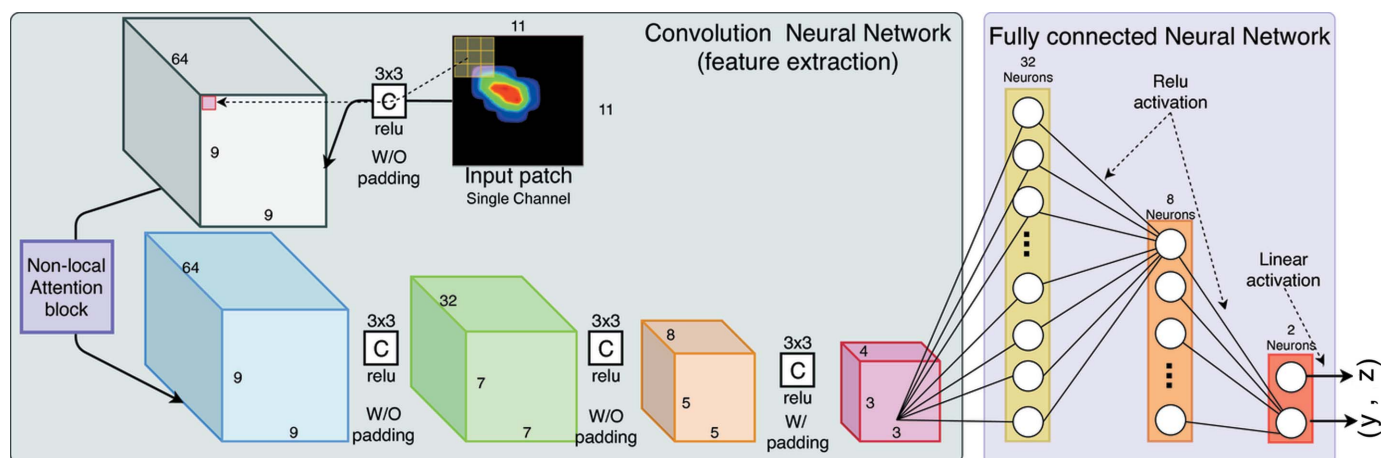


Figure 1
Neural network architecture as published by Liu *et al.* (2022). The image shows the details for the convolutional part of the network used for feature extraction and the fully connected part used for inference.

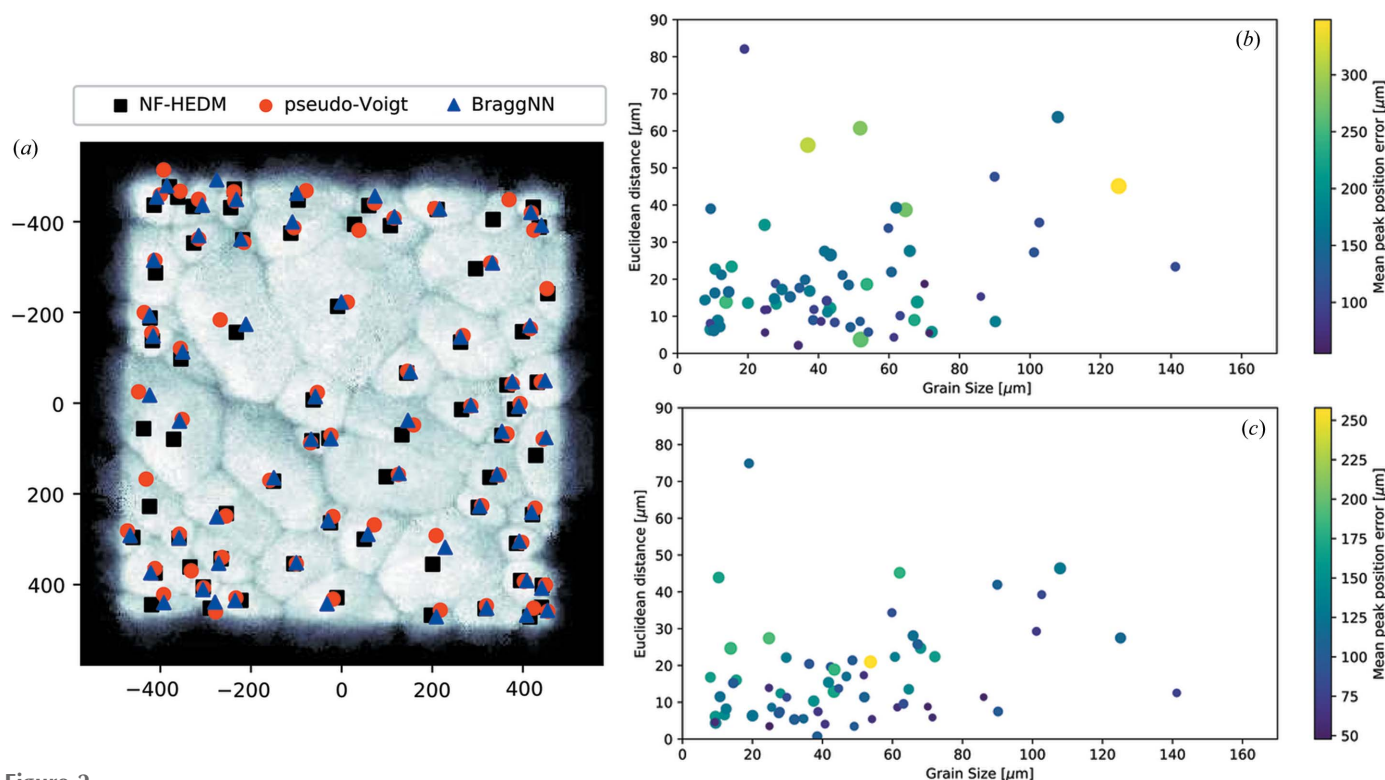


Figure 2
Grain identification by three different algorithms. Comparison between near-field HEDM, pseudo-Voigt far-field HEDM and *BraggNN* for the test material used by Liu *et al.* (2022). (a) Gives the grain position in the material for near-field HEDM (black squares), pseudo-Voigt far-field HEDM (red circles) and *BraggNN* (blue triangles). (b)–(c) Represent the differences in the grain positions between near-field HEDM (b), pseudo-Voigt far-field HEDM (c) and *BraggNN* as determined Liu *et al.* (2022).

performed in identifying the precise location of the Bragg peaks, allowing for the reconstruction of the grains in the sample material. As can be seen, applying the deep learning algorithm *BraggNN* enables the correct reconstruction of all but a few grains, usually at the edge of the sample, while at the same time greatly reducing the analysis time. Compared with optimized pseudo-Voigt fitting (Sharma *et al.*, 2012a) a 57× reduction in computation time was achieved on a CPU system

and a 350× reduction when GPUs were accessed. Even a good graphics card GPU allowed for a 250× reduction in computational time for center-of-mass detection for the Bragg peaks compared with standard pseudo-Voigt fitting. So, even groups with limited computational resources can benefit greatly from this new algorithm.

BraggNN is one step towards on-the-fly data analysis for the field of material sciences. If the application described here

proves as efficient and robust as it was for the sample used for development, then this new approach to HEDM data analysis offers a significant step forward to assess new materials more quickly and increase the turn-over rate from theoretical conception to practical application.

References

- Abeykoon, V., Liu, Z., Kettimuthu, R., Fox, G. & Foster, I. (2019). *IEEE/ACM 1st Annual Workshop on Large-scale Experiment-in-the-Loop Computing (XLOOP)*, 18 November, 2019, Denver, CO, USA, pp. 8–13. IEEE.
- Fukushima, K. (1980). *Biol. Cybern.* **36**, 193–202.
- Kordas, K. (2017). *Front. Mater.* **4**, 38.
- Liu, Y., Esan, O. C., Pan, Z. & An, L. (2021). *Energy AI*, 100049.
- Liu, Z., Sharma, H., Park, J.-S., Kenesei, P., Miceli, A., Almer, J., Kettimuthu, R. & Foster, I. (2022). *IUCrJ*, **9**, 104–113.
- Ma, Y., Ma, Y., Wang, Q., Schweidler, S., Botros, M., Fu, T., Hahn, H., Brezesinski, T. & Breitung, B. (2021). *Energy Environ. Sci.* **14**, 2883–2905.
- Park, J.-S., Zhang, X., Kenesei, P., Wong, S. L., Li, M. & Almer, J. (2017). *Microsc. Today*, **25**, 36–45.
- Poulsen, H. F. (2012). *J. Appl. Cryst.* **45**, 1084–1097.
- Poulsen, H. F., Nielsen, S. F., Lauridsen, E. M., Schmidt, S., Suter, R. M., Lienert, U., Margulies, L., Lorentzen, T. & Juul Jensen, D. (2001). *J. Appl. Cryst.* **34**, 751–756.
- Sharma, H., Huizenga, R. M. & Offerman, S. E. (2012a). *J. Appl. Cryst.* **45**, 693–704.
- Sharma, H., Huizenga, R. M. & Offerman, S. E. (2012b). *J. Appl. Cryst.* **45**, 705–718.