Explore the symmetry encoding in atomic pair distribution function (PDF) with convolutional neural network (CNN)

[a]Chia-Hao Liu, [b]Yunzhe Tao, [c]Daniel J. Hsu, [d]Qiang Du, [e]Simon J. L. Billinge

[a]Department of Applied Physics and Applied Mathematics, Columbia University, New York, New York, 10027, USA, cl3077@columbia.edu.
[b]Department of Applied Physics and Applied Mathematics, Columbia University, New York, New York, 10027, USA, vinzertao@gmail.com.
[c]Department of Computer Science, Columbia University, New York, New York, 10027, USA, djhsu@cs.columbia.edu.
[d]Department of Applied Physics and Applied Mathematics, Columbia University, New York, New York, 10027, USA, qd2125@columbia.edu.
[e]Department of Applied Physics and Applied Mathematics, Columbia University, New York, New York, 10027, USA, sb2896@columbia.edu.

Discovery of new materials depends critically on knowing the material structure which is generally done using x-ray diffraction. The first step in structure solution is to "index" the diffraction patterns to determine the space-group symmetry. However, the indexing approach fails when the material is nanostructured. Atomic pair distribution function (PDF) is powerful in this case of probing local structure, but the space group information is encoded in a yet discovered way. In this work, we utilize a convolutional neural network (CNN) model to approximate the unknown encoding of the space group information in PDF. The CNN model is trained on more than 100,000 data and it yields a 91% of accuracy with the top-6 estimates against testing set and it demonstrates robust results on predicting space group of experimental PDFs, from both crystalline and nanocrystalline samples.