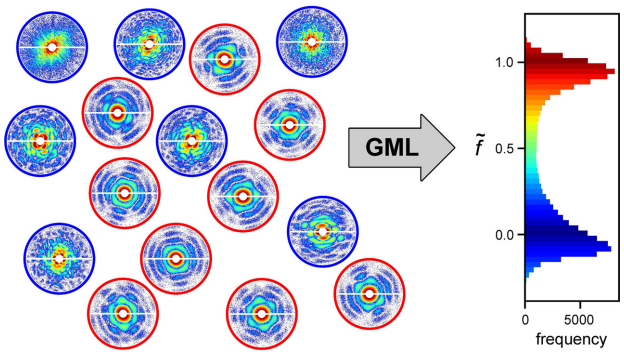
Scilight

18 February 2021

Geometric machine learning algorithm helps sieve through structural data of tiny biological objects

Aili McConnon

By fusing manifold embedding with machine learning, the researchers create an algorithm that can pick out X-ray diffraction patterns of single particles from an ocean of data.



Recent progress in single-particle-imaging with X-rays has made it possible to obtain 3D structural movies for studying tiny biological objects such as viruses. The quality of these movies depends on the number of diffraction pattern snapshots collected during the experiment. Cruz-Chú et al. developed a Geometric Machine Learning (GML) algorithm for identifying snapshots containing only a single-particle, out of a dataset with tens of millions of diffraction patterns collected by X-ray free-electron lasers (XFELs).

“Our method constitutes a critical step in identifying the large number of XFEL single-particle snapshots needed to extract structural variability, energy landscapes and 3D movies of the function of biomolecules at high resolution,” said author Peter Schwander.

XFELs allow researchers to capture snapshots of tiny biological objects in action. However, the process also produces gigantic datasets that are often difficult to navigate, with extraneous data that are diffraction patterns of multiple particles, no particles, or unwanted debris. The researchers’ GML algorithm allowed them to filter the data sets efficiently and retain the relevant snapshots that only contained a single particle.

“The goal of our work was to extract the single-particle snapshots efficiently with high precision and yield,” said author Eduardo R. Cruz-Chú.

The same type of GML approach can be applied to a wide range of different scientific disciplines that require accurate predictions from heterogeneous data. “Just recently, we successfully applied GML to predict the fetal gestational age from ultrasound measurements with unprecedented precision,” said author Russell Fung.

Source: “Selecting XFEL single-particle snapshots by geometric machine learning,” by Eduardo R. Cruz-Chu, Ahmad Hosseinizadeh, Ghoncheh Mashayekhi, Russell Fung, Abbas Ourmazd, and Peter Schwander, *Structural Dynamics* (2021). The article can be accessed at https://doi.org/10.1063/4.0000060.