doi:10.1017/S1431927620021169
Microsc. Microanal. 26 (Suppl 2), 2020

A Self-Supervised Workflow for Particle Picking in Cryo-EM

Qun Liu, Donal McSweeney and Sean McSweeney

Brookhaven National Laboratory, Upton, New York, United States

High-resolution single-particle cryo-EM data analysis relies on accurate particle picking. However, finding suitable parameters for particle picking remains a time-consuming and laborious process. Convolutional neural networks (CNN) have been increasingly used for automated particle picking in cryo-EM single-particle analysis. These CNN-based methods may differ in the formation of network architecture. Nevertheless, they all require particle data for training and the training quality determines the picking results and subsequent single-particle analysis. The training data can be composed of either manually picked particles or ab initio picking by various feature detection methods.

Though the CNN-based methods are promising for automated particle picking, improving the quality of training data and subsequent picking is so far limited. Therefore, an effective strategy is needed such that CNNs can be trained in a self-supervised manner for improved particle picking. Considering the established utility of 2D class averages in selecting particles and CNNs in pattern recognition, we propose that the combination of the two could improve the quality of training data via iterative training, particle picking, and 2D class averaging. To test this hypothesis, we devised a self-supervised particle picking workflow for automated particle picking.

Our workflow includes an iterative strategy to use 2D class average to improve training particles and a progressively improved CNN model for particle picking. To automate the selection of particles, we define a threshold (%/Res) using the ratio of percentage class distribution and resolution as a cutoff. Our workflow has been tested using six publicly available data sets of different conditions ranging from 1.3 MDa ribosome to 53 kDa streptavidin. With minimal input parameters, particles picked by the workflow support high-resolution reconstructions at 3.0 Å or better. Our workflow paves the way toward automated single-particle Cryo-EM data analysis at the stage of particle picking.

References

McSweeney, DM, McSweeney, SM, and Liu, Q. A Self-Supervised Workflow for Particle Picking in Cryo-EM. (2020). Submitted.