Improving protein tertiary structure prediction by deep learning and distance prediction in CASP14

Jian Liu[1], Tianqi Wu[2], Zhiye Guo[1], Jie Hou[3], and Jianlin Cheng[1]

[1]University of Missouri
[2]University of Missouri - Columbia
[3]Saint Louis University

April 2, 2021

Abstract

Substantial progresses in protein structure prediction have been made by utilizing deep-learning and residue-residue distance prediction since CASP13. Inspired by the advances, we improve our CASP14 MULTICOM protein structure prediction system in three main aspects: (1) a new deep learning based protein inter-residue distance predictor (DeepDist) to improve template-free (ab initio) tertiary structure prediction, (2) an enhanced template-based tertiary structure prediction method, and (3) distance-based model quality assessment methods empowered by deep learning. In the 2020 CASP14 experiment, MULTICOM predictor was ranked 7th out of 146 predictors in protein tertiary structure prediction and ranked 3rd out of 136 predictors in inter-domain structure predic-tion. The results of MULTICOM demonstrate that the template-free modeling based on deep learning and residue-residue distance prediction can predict the correct topology for almost all template-based modeling targets and a majority of hard targets (template-free targets or targets whose templates cannot be recognized), which is a significant improvement over the CASP13 MULTICOM predictor. The performance of template-free tertiary structure prediction largely depends on the accuracy of distance pre-dictions that is closely related to the quality of multiple sequence alignments. The structural model quality assessment works reasonably well on targets for which a sufficient number of good models can be predicted, but may perform poorly when only a few good models are predicted for a hard target and the distribution of model quality scores is highly skewed.

Hosted file

`CASP14_TS_Protein_format_v5.pdf` available at https://authorea.com/users/405356/articles/516362-improving-protein-tertiary-structure-prediction-by-deep-learning-and-distance-prediction-in-casp14