[BIOINFO 2024]

Deep Learning Strategy for Predicting Interaction between Protein and Ligand

Sanghwa Yoon[1], Yang Jae Kang[1,2,*]

[1] *Division of Bio & Medical Bigdata Department (BK4 Program), Gyeongsang National University*
[2] *Division of Life Science Department, Gyeongsang National University*
**Corresponding author: kangyangjae@gnu.ac.kr*

Protein-ligand interactions play a crucial role in biological processes, including signal transduction, enzyme catalysis, and immune responses. Predicting these interactions is vital in drug discovery, but current machine learning models that use one-dimensional protein sequences and SMILES for ligands have limitations in capturing structural features. To overcome these challenges, we propose a deep learning-based approach that utilises structural data to enhance the prediction of protein-ligand interactions. Using a dataset of 8,345 proteins and around 840,000 ligands from the ChEMBL database, proteins are represented as two-dimensional distance maps, and ligands are encoded using Coulomb matrices derived from their structures. A convolutional neural network with ResNet architecture is employed to effectively learn and predict interactions from this structural information. Our model shows superior performance against state-of-the-art baselines when evaluated on benchmark datasets, including BindingDB, DUD-E, and PDBbinds. Moreover, molecular modeling studies reveal that the ligands identified by our model possess significant interaction potential with target proteins, emphasizing the model's utility for screening drug candidates. By incorporating two-dimensional structural representations into the deep learning framework, our approach enhances the prediction accuracy of protein-ligand interactions. This strategy has the potential to accelerate the drug discovery process by enabling more efficient screening of potential therapeutics and reducing the need for costly experimental validation.