[BIOINFO 2025]

Improving Protein-Ligand Interaction Prediction using Electron-Aware Embeddings

Donghyeon Kim[1], and Dongsup Kim[1,*]

[1]*Department of Bio and Brain Engineering, Korea Advanced Institute of Science and Technology*
**Corresponding author: kds@kaist.ac.kr*

In the fields of chemistry and life sciences, effective molecular representation is a critical component for the success of deep learning models that handle molecules. For decades, established methods have included the use of physicochemical properties and structural fingerprints. Among these, Extended-Connectivity Fingerprints (ECFP) have been particularly successful, offering high computational efficiency while preserving structure-based information. However, ECFPs suffer from significant drawbacks, such as information loss due to bit collisions and a lack of interpretability. To overcome these limitations, recent research has shifted towards Graph Neural Network (GNN) and Transformer-based embeddings to better incorporate three-dimensional molecular information.

This study proposes a novel approach, using molecular electron density as a new embedding, inspired by the fundamental principle that molecular interactions are governed by electron interactions. We aim to validate the expressive power of this electron-aware representation and demonstrate its superior ability to capture non-covalent interactions in protein-ligand binding compared to conventional docking programs.