Predicting Protein-Ligand Binding Affinity via Joint Global-Local Interaction Modeling

Yang Zhang\textsuperscript{a,b}\textsuperscript{*}, Gengmo Zhou\textsuperscript{a,b}, Zhewei Wei\textsuperscript{a}, Hongteng Xu\textsuperscript{a}
\textsuperscript{a} Renmin University of China, China
\textsuperscript{b} DP Technology, China
{fengyuewuya, zgm2015, zhewei, hongtengxu}@ruc.edu.cn

Abstract—The prediction of protein-ligand binding affinity is of great significance for discovering lead compounds in drug research. Facing this challenging task, most existing prediction methods rely on the topological and/or spatial structure of molecules and the local interactions while ignoring the multi-level inter-molecular interactions between proteins and ligands, which often lead to sub-optimal performance. To solve this issue, we propose a novel global-local interaction (GLI) framework to predict protein-ligand binding affinity. In particular, our GLI framework considers the inter-molecular interactions between proteins and ligands, which involve not only the high-energy short-range interactions between closed atoms but also the low-energy long-range interactions between non-bonded atoms. For each pair of protein and ligand, our GLI embeds the long-range interactions globally and aggregates local short-range interactions, respectively. Such a joint global-local interaction modeling strategy helps to improve prediction accuracy, and the whole framework is compatible with various neural network-based modules. Experiments demonstrate that our GLI framework outperforms state-of-the-art methods with simple neural network architectures and moderate computational costs.

Index Terms—Protein-ligand binding affinity, graph neural networks, long-short interactions, drug discovery

I. INTRODUCTION

In the field of drug discovery, researchers are interested in identifying the ligand (small molecular compounds) whose bindings with target proteins affect the expressions of the proteins’ functions and thus affect disease rehabilitation and our health. Therefore, the calculation of protein-ligand binding affinity is of great significance [1], [2] since the binding affinity reflects the binding strength and accordingly determines whether a ligand can be used as a candidate drug or not.

In practice, the protein-ligand binding affinity can be measured in the laboratory, which, however, is very expensive and time-consuming [2], [3]. To reduce the cost and improve the efficiency, many methods have been proposed to predict the binding affinity with the help of computer science [4], [5]. Specifically, most existing methods can be coarsely categorized into two classes: simulation-based methods and machine learning-based methods. Simulation-based methods [6] apply expert knowledge and physical principles to simulate the binding process, whose computational efficiency and scalability are often questionable. The classic machine learning-based methods [5], [7] mainly depend on hand-crafted features like ECFP (extended connectivity fingerprint) of proteins and ligands, and leverage off-the-shelf techniques like random forest [5] to predict the binding affinity. These early works are fast, but their accuracy is limited due to the information loss caused by manual feature engineering. Recently, numerous new machine learning-based methods have been developed using deep learning techniques, in which the features are learned from the molecular data rather than designed manually. In particular, various deep learning models are proposed based on convolutional neural networks (CNNs) [8]–[11] and graph neural networks (GNNs) [12]–[14], which outperform classic learning-based methods consistently.

Although the above methods have achieved encouraging progress in predicting binding affinity, they still suffer from the following drawbacks in their modeling principles. In particular, as shown in Figure 1, the binding affinity of a protein and a ligand is mainly determined by the inter-molecular interactions between them, which consist of long-range and short-range interactions [15]–[17], respectively. The long-range interactions come from non-bonded atoms that might be weak but have a large amount. In contrast, the short-range interactions correspond to the non-covalent interactions between the atoms close to each other, which have few amounts but own high energy in general. Such interactions are based on various factors, including the biochemical information of the protein and the ligand, their topological (2D) structure, their spatial (3D) structure, and etc. Facing such multi-level interactions...
and complicated structural information, existing methods, however, merely consider the short-range interactions and local non-spatial structural information, which inevitably leads to sub-optimal performance. Although some attempts have been made recently to leverage spatial [18]–[20] and pairing information [21], [22] of proteins and ligands, they do not provide systematic solutions to model both long- and short-range interactions jointly when predicting the binding affinity.

To overcome the drawbacks above, we propose a novel global-local interaction (GLI) framework to predict protein-ligand binding affinity. In particular, the proposed framework is comprised of three modules: i) A chemical info module is applied to model the intra-molecular interactions happening within proteins and ligands and embed the molecules accordingly. ii) A global interaction module is used to learn the long-range interactions between proteins and ligands, which embeds the chemical and spatial information of proteins and ligands, respectively, and then models global interaction effects based on the embeddings. iii) A local interaction module is used to learn the short-range interactions between proteins and ligands, which models the local interaction effects based on some paired atoms that own non-covalent bonds and have short distances. Finally, given a protein and a ligand, their binding affinity is predicted based on both global and local interaction effects derived from the above modules.

The main contributions of this paper are as follows: To the best of our knowledge, we are among the first to develop a general framework for protein-ligand binding affinity prediction, which captures the multi-level inter-interactions between proteins and ligands. In particular, the proposed GLI framework models the long-range and short-range interactions in the binding process, which has strong explanatory ability, generalization ability, and great potential for expansion. Our framework has high flexibility — the three modules can be designed separately and learned jointly, each of which is highly compatible with various neural network architectures. We implement our GLI framework based on some simple neural network architectures and evaluate it on the well-known PDBbind dataset [23], CSAR-HiQ [24]. Experimental results demonstrate the superiority of our GLI to the SOTA methods in both prediction accuracy and efficiency.

II. PRELIMINARIES

**Protein and Ligand Graphs.** We define a protein and a ligand as two graphs, denoted as $G^P = (V^P, E^P)$ and $G^L = (V^L, E^L)$, respectively. For each graph $G$, we denote its nodes as a set $V = \{a_i\}$, where $a_i$ is the $i$-th atom node, and its edges as a set $E = \{e_{ij}\}$, where $e_{ij}$ is the edge connecting $a_i$ to $a_j$.

The proteins and ligands are attributed, which contain significant chemical and spatial information. Accordingly, each $G^P$ (or $G^L$) can be instantiated in the following two views:

a) Chemical Info Graph $G^c = (V^c, E^c)$: $V^c$ contains the chemical information of nodes, such as their atom types, formal charges, degrees and etc. $E^c$ contains information of chemical bonds, including bond types and lengths.

b) 3D Spatial Graph $G_s = (V_s, E_s)$: $V_s$ contains the 3D coordinates of nodes, and $E_s = \{d_{ij}\}$ contains the weights of edges, where $d_{ij}$ represents the distance between $a_i$ and $a_j$. The edges in $E_s$ correspond to the node pairs whose distances are smaller than the cutoff distance $\mu$.

**Local Interaction Graph** $G^{local} = (V^{local}, E^{local})$. For each edge in $G^{local}$, the corresponding atoms are from a protein and a ligand, respectively. The distance between the atoms is smaller than the cutoff distance $\mu$, and thus, there is a non-covalent interaction occurring between the atoms.

**Problem Statement.** Given a molecule like protein or ligand, we can construct protein graph $G^P = (V^P, E^P)$, ligand graph $G^L = (V^L, E^L)$ and $G^{local} = (V^{local}, E^{local})$ with cutoff distance $\mu$. Our goal is to construct a model $f(G^P, G^L, G^{local}, \mu)$ to capture the chemical info, long-range interactions and short-range interactions in the binding process of protein and ligand, so as to predict the binding affinity accurately.

III. MODEL FRAMEWORK

Figure 2 depicts the overall framework, which is composed of three modules: the chemical info module to gather the chemical information of atoms and chemical bonds; the global interaction module to represent long-range interactions; and the local interaction module to represent short-range interactions. The predicted binding affinity comes from the long- and short-range interactions.

**A. Chemical Info Module**

Given a protein and a ligand, i.e., $\{G^P_i, G^L_i\}$ and $\{G^L_i, G^L_i\}$, we first apply a graph neural network to extract each atom’s chemical embedding:

$$C^P = f_{chem}(V^c_P, E^c_P), C^L = f_{chem}(V^c_L, E^c_L),$$

where $f_{chem}(\cdot)$ is the proposed chemical embedding model and “$\|$” represents the concatenation operation. It consists of two sub-modules. The first sub-module is an embedding layer, transforming chemical information to initial node embeddings and concatenating it to the corresponding spatial attribute; Taking the concatenated node embeddings and the topological information (i.e., $E^c_P$ and $E^c_L$) as input, the second sub-module (implemented via GAT, GCN, or GIN) extracts the chemical information embedding set of the protein and that of the ligand, denoted as $C^P$ and $C^L$, respectively. Here, each $C$ means a set of $c_i$‘s, and $c_i$ is the chemical information embedding of the atom $a_i$.

**B. Global Interaction Module**

We presented a global interaction module for learning and representing long-range interactions between protein and ligand, which are important components of inter-molecular interactions. Additionally, since it is extremely time-consuming and tedious to compute one atom by one atom, the module gathers the atomic level chemical embedding of ligand and protein as one node, respectively. Then the module estimates the interaction effect between overall protein $g^P$ and overall ligand $g^L$ as a long-range interaction $y_{global}$. 

1324
Relative Position Attention. In the calculation of long-range interactions, the position of the atom has some impact on the contribution of the interaction between atoms. For example, in a protein, atoms close to the ligand should contribute more to inter-interactions than atoms far from the ligand. By referring to attention-pooling work [25], [26], we develop a new method named "Relative Position Attention" to express the importance of atoms in long-range interaction effects. In particular, given an atom \(a_i\), we derive two \(K\)-bin histograms for the remaining atoms in \(G^P\) and those in \(G^L\), respectively, according to their distance to \(a_i\), i.e.,

\[
POS^P(a_i) = [p^P_k(a_i), ..., p^P_K(a_i)], 1 - \sum_{k=0}^K p^P_k(a_i),
\]
\[
POS^L(a_i) = [p^L_k(a_i), ..., p^L_K(a_i)], 1 - \sum_{k=0}^K p^L_k(a_i),
\]

where for \(k = 0, ..., K\),

\[
p^P_k(a_i) = \frac{1(|\{a_j|a_j \in G^P \land k \leq d_{ij} \leq k + 1\}|)}{1(|\{a_j|a_j \in G^P\}|)},
\]
\[
p^L_k(a_i) = \frac{1(|\{a_j|a_j \in G^L \land k \leq d_{ij} \leq k + 1\}|)}{1(|\{a_j|a_j \in G^L\}|)},
\]

where \(1(S)\) represents the cardinality of a set \(S\). These two histograms contain the relative position information of \(a_i\) from the protein and the ligand, respectively. Accordingly, we leverage the following relative position attention mechanism to aggregate all such relative position information, i.e.,

\[
POS(a_i) = [POS^P(a_i)||POS^L(a_i)],
\]
\[
atten(a_i) = \frac{\exp\{c_i||POS(a_i)|\}}{1 + \exp\{c_i||POS(a_i)|\}},
\]
\[
g^P = \sum_{a_i \in G^P} c_i \cdot atten(a_i),
\]

where \(atten(a_i)\) represents the position attention, \(c_i\) represents the chemical embedding of atom \(a_i\) calculated in chem info module, and \(K\) is a hyper-parameter that is set as 10 in this experiment. \(g^P\) is the comprehensive chemical information of the protein considering the position effect.

In contrast, the ligand is much smaller and more concentrated in space. Therefore, we use the add pooling method to aggregate the chemical information of the ligand \(g^L\).

\[
g^L = \text{add-pooling}(C^L),
\]

At last, we concatenate the global embeddings of protein and ligand as the global interaction vector and apply a MLP (Multi-layer Perceptron) module to calculate the long-range interaction \(\hat{y}_{global}\).

\[
\hat{y}_{global} = \text{MLP}(g^P||g^L),
\]

C. Local Interaction Module

The atoms within a range defined by the cutoff distance \(\mu\) will have strong short-range interactions during the binding process of protein and ligand. Our framework contains a local interaction module to capture such information. As aforementioned, we first construct the local interaction graph from the protein and ligand according to the cutoff distance \(\mu\). Then, to incorporate spatial information, we update the node embedding using the local spatial interaction model \(f_{spatial}\). It can be set by returning the chemical embedding directly. It is also feasible to employ spatial graph-based models (like SchNet [18], DimeNet [19] and etc.) to update chemical and spatial embedding of all nodes \(S\) by incorporating the spatial information \(V_s\).

\[
S_{local} = f_{spatial}(C_{local}, V_{local}),
\]
where $G_{\text{local}}$ means the chemical info embedding set of the nodes in the local interaction graph $G_{\text{local}}$, $V_{\text{local}}$ means the spatial info set of the nodes in local interaction graph $G_{\text{local}}$. At last, we concatenate chemical and spatial node embedding $s_i, s_j$ of each non-covalent bond $e_{ij}$ between ligand nodes and protein nodes as the short-range interaction vector. Then, we feed all the vectors into an MLP and sum the results up as the prediction of the overall short-range interactions between protein and ligand.

$$
\hat{y}_{\text{local}} = \sum_{(i,j) \in \{(i,j)|e_{ij} \in E_{\text{local}}\}} \text{MLP}(s_i, s_j).
$$

(10)

**D. Optimization Objective**

Since the binding affinity $y$ is determined by the long-range and short-range interactions between atoms of proteins and ligands, we estimate the binding affinity $\hat{y}$ by combining the results of the global interaction module $\hat{y}_{\text{global}}$ and the local interaction module $\hat{y}_{\text{local}}$.

$$
\hat{y} = \hat{y}_{\text{global}} + \hat{y}_{\text{local}}.
$$

(11)

We use L1 loss function and the mean absolute error between the predicted binding affinity $\hat{y}$ and the measured ground truth $y$ to optimize the model.

$$
L = \sum_{i=1}^{M} |y_i - \hat{y}_i|,
$$

(12)

where $L$ represents the loss function, which is implemented as the mean absolute error (MAE), and $M$ is the number of protein-ligand pairs in the dataset.

**IV. EXPERIMENTS**

In this section, we will conduct experiments on commonly used datasets to test the effects of our framework.

- **Task 1.** Compared to the existing state-of-the-art models, is it possible for our framework to produce improved or even superior prediction results?
- **Task 2.** Can different modules of our framework contribute to a substantial improvement? What contributions do the different modules make to the prediction, and are the results statistically significant?
- **Task 3.** How about the efficiency and computational consumption of our framework?

**A. Experiment Settings**

**a) Dataset:** We will train and test the baseline models and our framework on the following datasets:

- **PDBbind:** In the related research of binding affinity prediction, PDBbind [23] is a well-known and commonly used data set, which consists of three sub-data sets (general set, refined set, and core set). In this experiment, consistent with the studies [10], [21], we take the refined subset of the PDBbind v2016, PDBbind v2020 for training and the core subset of the PDBbind v2016 for testing.
- **CSAR-HiQ** [24]: This is another well-known dataset for predicting binding affinity. Consistent with the study [21], we use the refined subset of the PDBbind v2016 for training and take this data set for testing to verify the generalizability of our framework.

- **b) Introduction to the binding affinity:** In the datasets of PDBbind and CSAR-HiQ, the target binding affinity is the negative logarithm of the experimental results, such as $-logK_d$ (dissociation constant), $-logK_i$ (inhibition constant), or $-logIC_{50}$ (semi-inhibitory concentration).

- **c) Evaluation Metrics:** We use mean absolute error (MAE), root mean square error (RMSE) to evaluate the performance of different models and the significance of different modules in our GLI framework. For efficiency, we collect the average GPU RAM usage (MB) and running time (training, validating, and testing) (hours). Besides, we also performed a two-sample T-test to analyze whether different modules in our framework can bring significant performance changes.

- **d) Baseline Models and Parameters:** We compare the experiment results with the following models: **Traditional machine learning models:** RF-score [5]. **CNN models:** Pafnucy [14] and Onionnet [10]. **Topological (2D) graph-based models:** GAT [27], GCN [28]. **GIN** [13] and GCN2 [29]. **Spatial (3D) graph-based models:** SchNet [18], DimeNet [19], SphereNet [20]. **Structure-aware interaction model:** SIGN [21].

- **GLI Framework:** We implement our GLI framework based on different models in the following experiments. **GLI-0, GLI-1, GLI-2:** Taking (GAT + GCN), GIN, GCN2 as the chemical info module.

- **f) Environment and Settings:** All experiments are conducted on a machine with an NVIDIA V100 GPU (32GB RAM), Intel Xeon CPU (12 Cores, 2.5 GHz), and 92GB of RAM. We set the batch size as 32, dropout rate as 0.1 and 200 epochs. We construct the local interaction graph with cutoff distance $\mu = 5 \, \AA$. We use 10-fold cross-validation for training and validating. Related codes, processed data, and trained models for our experiments can be found at the link [30].

**B. Result Analysis**

**1) Model Comparison:** In Table I, we compare the models based on our GLI framework with the five aforementioned types of baselines. The topological graph-based methods leverage the topological information of molecules. The prediction performance of these models is inadequate, with the results of RMSE more than 1.5 in PDBbind v2016, v2020, and 1.8 in CSAR-HiQ, respectively. This could be because the topological graph-based model disregards crucial spatial information.

Pafnucy, OnionNet, SchNet, DimeNet, SphereNet, and other spatial graph models leverage the 3D spatial structures of proteins and ligands. These models have RMSE of approximately 1.454, 1.444, and 1.715 in PDBbind v2016, v2020, and CSAR-HiQ, which is better than the topological models. Due to the computational complexity of these models, however,
### Methods and Frameworks

| Methods | Type | PDBbind v2016 | PDBbind v2020 | CSAR-HiQ | GPU \(\downarrow\) | Time \(\downarrow\) |
|---------|------|--------------|--------------|----------|-------------|-------------|
|          |      | MAE \(_\downarrow\) | RMSE \(_\downarrow\) | MAE \(_\downarrow\) | RMSE \(_\downarrow\) | MAE \(_\downarrow\) | RMSE \(_\downarrow\) | (MB) | (H) |
| RF-Score | Machine Learning | 1.111(0.008) | 1.397(0.008) | 1.113(0.008) | 1.398(0.004) | 1.471(0.013) | 1.847(0.015) | - | 0.08 |
| Pfaffucy | CNN | 1.299(0.017) | 1.597(0.019) | 1.319(0.038) | 1.649(0.047) | 1.482(0.051) | 1.853(0.050) | 16,255 | 18.7 |
| OnionNet | CNN | 1.3740(0.080) | 1.8130(0.113) | 1.3960(0.184) | 1.8150(0.274) | 1.3550(0.032) | 1.7510(0.032) | 30,701 | 0.25 |
| GAT | Topological Graph | 1.3070(0.016) | 1.6290(0.018) | 1.2940(0.013) | 1.6060(0.014) | 1.6880(0.219) | 2.1120(0.238) | 1,403 | 0.16 |
| GCN | Topological Graph | 1.2900(0.017) | 1.5970(0.021) | 1.2890(0.017) | 1.5920(0.018) | 1.6200(0.316) | 2.0410(0.349) | 1,243 | 0.19 |
| GIN | Topological Graph | 1.2490(0.022) | 1.5540(0.028) | 1.2600(0.007) | 1.5680(0.007) | 1.4420(0.093) | 1.8490(0.104) | 1,557 | 0.24 |
| GCN2 | Topological Graph | 1.2920(0.028) | 1.6040(0.032) | 1.2750(0.013) | 1.5780(0.015) | 1.8720(0.128) | 2.3160(0.135) | 1,509 | 0.24 |
| GAT + GCN | Topological Graph | 1.2920(0.019) | 1.5990(0.023) | 1.2730(0.011) | 1.5930(0.014) | 1.5270(0.195) | 1.9360(0.214) | 1,480 | 0.28 |
| GAT + GIN | Topological Graph | 1.2410(0.014) | 1.5530(0.019) | 1.2460(0.019) | 1.5590(0.023) | 1.4370(0.147) | 1.8380(0.152) | 1,797 | 0.30 |
| GAT + GCN2 | Topological Graph | 1.2960(0.015) | 1.6010(0.015) | 1.2820(0.019) | 1.5960(0.023) | 1.7050(0.114) | 2.1280(0.127) | 1,835 | 0.45 |
| SchNet | Spatial Graph | 1.1570(0.035) | 1.4540(0.035) | 1.1540(0.024) | 1.4440(0.028) | 1.3300(0.064) | 1.7150(0.066) | 12,547 | 0.80 |
| DimeNet | Spatial Graph | - | - | - | - | - | - | OOM | - |
| SphereNet | Spatial Graph | - | - | - | - | - | - | OOM | - |
| SIGN \(^a\) | Structure-Aware | 1.0270(0.025) | 1.3160(0.031) | - | - | 1.3270(0.040) | 1.7350(0.031) | 20,091 | 2.73 |

\(^a\) The standard deviation of each index is indicated in brackets.

### Results and Discussion

- **Chemical Info Module**: When only the chemical info module is included, our GLI framework is equivalent to using only biochemical information of data. We use (GAT+GCN), GIN, and GCN2 as the network structures to implement the chemical info module, respectively. The experimental results indicate the chemical info module is not sufficient to make an accurate prediction.

- **Global Interaction Module**: After adding the global interaction module, we can observe that for GLI-0-cg, the predicted RMSE in PDBbind v2016, v2020 and CSAR-HiQ decreased by 0.199, 0.136, and 0.290. The p-value results of the T-test are less than 0.1%, indicating that the decrease in RMSE is statistically significant. Besides, when only the chemical info module is included, our GLI framework is equivalent to using GIN and GCN2 separately as the chemical info module and add the global interaction module. The same results were also observed in the PDBbind v2016, v2020, and CSAR-HiQ datasets. Moreover, introducing the global interaction module has a small impact on the calculation time of GLI. These results indicate that the global interaction module is a simple and effective structure with strong generalizability.

- **Local Interaction Module**: With the addition of the local interaction module to the GLI-1-cg model, the predicted result in the PDBbind v2016 dataset is pretty satisfied with the RMSE of 1.294 (decreased 0.098) and the T-test result of 6.946 (P-value< 0.1%). Compareable results are also obtained in GLI-0, GLI-2 and PDBbind v2020, CSAR-HiQ datasets, demonstrating that in statistics, the local interaction module can significantly improve binding affinity prediction by reducing predicted error by 0.102 at most. This result also demonstrates the effectiveness and importance of the local interaction module.

---

**Note:**

- The standard deviation of each index is indicated in brackets.
- It is difficult to apply them to macromolecular systems like proteins. Additionally, these models make poor use of long-range information.
- The SIGN model is a comprehensive model that exploits the chemical element pair as the long-range interaction effect and extracts the angle and spatial info. The results of this model are superior to the simple spatial model and topological model. However, like the spatial model, this model consumes a substantial amount of time and computing resources.
- Finally, the results of our GLI-based models are better than the above SOTA models. For PDBbind v2016, GLI-based models are the first model to make RMSE less than 1.3. For CSAR-HiQ as test data, GLI-based models are the first model to make RMSE less than 1.55 and MAE less than 1.2. These results demonstrate our GLI framework has encouraging performance in predicting protein-ligand affinity.

2) **Ablation Study of GLI**: As shown in Table I, II, we conduct ablation and T-test experiments on our GLI framework. The corresponding analysis is as follows:

- **Chemical Info Module**: When only the chemical info module is included, our GLI framework is equivalent to using only biochemical information of data. We use (GAT+GCN), GIN, and GCN2 as the network structures to implement the chemical info module, respectively. The experimental results indicate the chemical info module is not sufficient to make...
TABLE II
T-test results of different modules in GLI.\textsuperscript{4}

| GLI     | PDBbind v2016 | PDBbind v2020 | CSAR-HiQ |
|---------|---------------|---------------|----------|
|         | RMSE↓ | ΔRMSE↓ | T-test↑ | RMSE↓ | ΔRMSE↓ | T-test↑ | RMSE↓ | ΔRMSE↓ | T-test↑ |
| GLI-0-c | 1.599 |         |         | 1.593 |         |         | 1.936 |         |         |
| GLI-0-cg | 1.399 | -0.199  | 15.616\textsuperscript{**} | 1.457 | -0.136  | 15.728\textsuperscript{***} | 1.646 | -0.290  | 3.949\textsuperscript{***} |
| GLI-0-cgl | 1.321 | -0.077  | 5.768\textsuperscript{**} | 1.354 | -0.102  | 10.426\textsuperscript{***} | 1.595 | -0.051  | 2.310\textsuperscript{**} |
| GLI-1-c | 1.554 |         |         | 1.568 |         |         | 1.849 |         |         |
| GLI-1-cg | 1.392 | -0.162  | 12.742\textsuperscript{***} | 1.417 | -0.151  | 19.721\textsuperscript{***} | 1.638 | -0.211  | 5.724\textsuperscript{**} |
| GLI-1-cgl | 1.294 | -0.098  | 6.946\textsuperscript{**} | 1.353 | -0.064  | 5.808\textsuperscript{**} | 1.597 | -0.040  | 2.014\textsuperscript{*} |
| GLI-2-c | 1.604 |         |         | 1.578 |         |         | 2.316 |         |         |
| GLI-2-cg | 1.407 | -0.197  | 16.506\textsuperscript{**} | 1.438 | -0.141  | 16.797\textsuperscript{**} | 1.587 | -0.729  | 15.749\textsuperscript{***} |
| GLI-2-cgl | 1.347 | -0.060  | 5.127\textsuperscript{**} | 1.405 | -0.032  | 3.582\textsuperscript{**} | 1.549 | -0.038  | 2.447\textsuperscript{a} |

\textsuperscript{a}*, **, *** means the P-value is less than 5%, 1%, 0.1%.

interactions for the binding affinity prediction.

In addition, based on GLI-0-cgl model, we add spatial graph-based models (SchNet, DimeNet) to update the node embedding of the local interaction graph. The comparison results show that these spatial models provide few or even negative improvements. This could be because the existing spatial model intends to update the node embedding, which would reduce the gap between protein and ligand nodes.

V. CONCLUSION

In this paper, concentrating on the prediction of protein-ligand binding affinity, we propose a global-local interaction framework called GLI from the perspective of multi-level inter-molecular interactions in the binding process. In addition to superior prediction accuracy and comparable computational efficiency, the experimental results demonstrate that our framework is highly generalizable and scalable.

REFERENCES

[1] U. Ryde and P. S'Derhjelm, “Ligand-binding affinity estimates supported by quantum-mechanical methods,” Chemical Reviews, vol. 116, no. 9, pp. 5520–5566, 2016.
[2] D. B. Kitchen, H. Decornez, J. R. Furr, and J. Bajorath, “Docking and scoring in virtual screening for drug discovery: methods and applications.” Nature Reviews Drug Discovery, vol. 3, no. 11, pp. 935–949, 2004.
[3] J. Tang, F. Wang, and F. Cheng, “Artificial intelligence for drug discovery,” in Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, 2021, pp. 4074–4075.
[4] Ballester, P., J., Mitchell, J., and B., “A machine learning approach to predicting protein-ligand binding affinity with applications to molecular docking.” Bioinformatics Oxford, 2010.
[5] H. Li, K. S. Leung, M. H. Wong, and P. J. Ballester, “Improving autodock vina using random forest: The growing accuracy of binding affinity prediction by the effective exploitation of larger data sets,” Molecular Informatics, vol. 34, no. 2-3, pp. 115–126, 2015.
[6] B. R. Katherine Kantardjieff, “Structural bioinformatic approaches to the discovery of new antibacterial drugs.” Curr Pharm Des, vol. 10, no. 26, pp. –, 2004.
[7] Guoping, Hu, Guanglin, Kuang, Wen, Xiao, Weihua, Li, Guixia, and Liu, “Performance evaluation of 2d fingerprint and 3d shape similarity methods in virtual screening,” Journal of Chemical Information and Modeling, vol. 52, no. 5, pp. 1103–1113, 2012.
[8] I. Wallach, M. Dzamba, and A. Heifets, “Atomnet: A deep convolutional neural network for bioactivity prediction in structure-based drug discovery,” Mathematische Zeitschrift, vol. 47, no. 1, pp. 34–46, 2015.
[9] I. Lee, J. Keum, and H. Nam, “Deepeconv-dti: Prediction of drug-target interactions via deep learning with convolution on protein sequences,” PLOS Computational Biology, vol. 15, no. 6, pp. e1007129–, 2019.
[10] L. Zheng, F. J., and Y. Mu, “Molecular representation of inter-molecular contact based convolutional neural network for protein-ligand binding affinity prediction,” ACS Omega, vol. 4, no. 14, 2019.
[11] Z. Wang, L. Zheng, Y. Liu, Y. Qu, Y. Q. Li, M. Zhao, Y. Mu, and W. Li, “Onionnet-2: A convolutional neural network model for predicting protein-ligand binding affinity based on residue-atom contacting shells,” 2021.
[12] A. Narayanan, M. Chandramohan, R. Venkatesan, L. Chen, and S. Jaiswal, “graph2vec: Learning distributed representations of graphs,” 2017.
[13] K. Xu, W. Hu, J, Leskovec, and S. Jegelka, “How powerful are graph neural networks?” 2018.
[14] M. M. Stepniewska-Dziubinska, Z. Piotr, and S. Pawel, “Development and evaluation of a deep learning model for protein-ligand binding affinity prediction,” Bioinformatics, no. 21, p. 21.
[15] D. Greenspan, “Long and short range forces: Gravitation and molecular attraction and repulsion.” Elsevier Ltd, 1980.
[16] M. K. Felner, M. Schneider, and A. Salam, “Theory of long-range intermolecular interactions,” Wiley, 2009.
[17] M. Sun, J. Xing, H. Wang, B. Chen, and J. Zhou, “Mocl: Contrastive learning on molecular graphs with multi-level domain knowledge,” 2021.
[18] K. T. Schütt, P. J. Kindermans, H. E. Sauceda, S. Chmiela, A. Tkatchenko, and K.-R. Müller, “SchNet: A continuous-filter convolutional neural network for modeling quantum interactions,” in Advances in Neural Information Processing Systems, 2017.
[19] J. Klüwer, J. Oro, and S. Gimmernann, “Directional message passing for molecular graphs,” 2020.
[20] Y. Liu, L. Wang, M. Liu, X. Zhang, B. Oztekin, and S. Ji, “Spherical message passing for 3d graph networks,” 2021.
[21] S. Li, J. Zhou, T. Xu, L. Huang, F. Wang, H. Xiong, W. Huang, D. Dou, and H. Xiong, “Structure-aware interactive graph neural networks for the prediction of protein-ligand binding affinity,” 2021.
[22] H. Cho, E. K. Lee, and I. S. Choi, “Interactionnet: Modeling and explaining of noncovalent protein-ligand interactions with noncovalent graph neural network and layer-wise relevance propagation,” 2020.
[23] http://www.pdbbind.org.cn/.
[24] http://csarbind.org/.
[25] Y. Li, D. Tarlow, M. Brockschmidt, and R. Zemel, “Gated graph convolutional networks,” 2016.
[26] P. Velicković, G. Cucurull, A. Casanova, A. Romero, P. Liabed, and Y. Bengio, “Graph attention networks,” in ICLR, 2017.
[27] A. Tkatchenko, and K.-R. Müller, “SchNet: A continuous-filter convolutional neural network for modeling quantum interactions,” in Advances in Neural Information Processing Systems, 2017.
[28] A. Tkatchenko, and K.-R. Müller, “SchNet: A continuous-filter convolutional neural network for modeling quantum interactions,” in Advances in Neural Information Processing Systems, 2017.
[29] M. Chen, Z. Wei, Z. Huang, B. Ding, and Y. Li, “Simple and deep graph convolutional networks,” 2020.
[30] https://github.com/ningyuewuya/global_local_interaction_framework.