**Supplementary Information for**

**Twin Cross-scale Contrastive Learning with Multi-modality Fusion for Drug-Target Affinity Prediction**

Linna Zhang1,2,Zhaowei Wang3,Wuhao Liu1,2, Xiaodong Duan1,2 ,Quan Zou4 and Qiguo Dai1,2,\*

1College of Computer Science and Engineering, Dalian Minzu University, 116600, Dalian, China

2SEAC Key Laboratory of Big Data Applied Technology, Dalian Minzu University, 116600, Dalian, China

3College of Computer Science and Engineering, Dalian University of Technology, 116024, Dalian, China

4Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, 611730, Chengdu, China

\*Corresponding author:daiqiguo@dlnu.edu.cn(Q.D.)

**1 Graph Representation of Node Feature**

For drug atom nodes, the initial feature is represented by a 78-dimensional vector(Supplementary Table 1). For drug motif nodes, the detailed meanings of the initial feature are shown in Supplementary Table 2, with the node features represented by a 92-dimensional vector. The node features for target proteins are represented as a 54-dimensional vector, with further details provided in Supplementary Table 3.

**Supplementary Table** **1. the node features for the drug atom-level graph.**

|  |  |
| --- | --- |
| **Descriptor** | **Dimensions** |
| Atom symbol (one-hot) | 44 |
| Degree of the atom (one-hot) | 11 |
| Total number of H atoms bound to the heavy atom (one-hot) | 11 |
| Number of implicit H atoms bound to the heavy atom (one-hot) | 11 |
| Whether the atom is aromatic | 1 |
| **All** | **78** |

**Supplementary Table** **2. the node features for the drug motif-level graph.**

|  |  |
| --- | --- |
| **Descriptor** | **Dimensions** |
| Atomic symbols contained in the motif | 44 |
| Number of atoms in the motif | 11 |
| Number of edges connecting to other motifs | 11 |
| Total number of hydrogen atoms connected by motif | 12 |
| Implicit valence of motif | 12 |
| Whether the motif is a simple ring | 1 |
| Whether the motif is chemically bonded or not | 1 |
| **All** | **92** |

**Supplementary Table** **3. the node features for the target protein graph.**

|  |  |
| --- | --- |
| **Descriptor** | **Dimensions** |
| Position-Specific Scoring Matrix(PSSM) | 21 |
| Residue symbol (one-hot) | 21 |
| Whether the residue is aliphatic | 1 |
| Whether the residue is aromatic | 1 |
| Whether the residue is polar neutral | 1 |
| Whether the residue is acidic (negatively charged) | 1 |
| Whether the residue is basic (positively charged) | 1 |
| Residue weight | 1 |
| Dissociation constant for the –COOH group (-log) | 1 |
| Dissociation constant for the –NH3 group (-log) | 1 |
| pH at the isoelectric point | 1 |
| Hydrophobicity of the residue (pH = 2) | 1 |
| Hydrophobicity of the residue (pH = 7) | 1 |
| Dissociation constant for any other group in the molecule (-log) | 1 |
| **All** | **78** |

**2 Detail of Twin Cross-Scale Contrastive Learning**

**The Strategy of Positive and Negative Sample Selection** As illustrated in the TCCL model, it implements a contrastive learning strategy to enhance feature representations. Within this framework, drugs act as anchors, and their features must remain consistent with positive samples across various scales while differing from negative samples. A primary challenge in cross-scale contrastive learning is the identification of positive and negative samples for a given anchor to generate self-supervised signals. Generally, the embedding of a drug derived at the network scale serves to define its embedding at the molecular scale as positive samples. To facilitate the generation of these positive samples, we propose an innovative sampling technique that combines molecular feature-based sampling with network structure-based sampling strategies.

In molecular feature-based sampling, nodes with similar molecular structures are likely to exhibit high correlation and can thus be considered as positive sample pairs. We utilize the PubChem structure clustering tool to compute drug similarities, represented by the matrix , and the Smith-Waterman algorithm to determine protein sequence similarities, stored in the matrix . These similarity computations are conducted at the molecular level. Conversely, network structure-based sampling considers that nodes sharing common neighbors are likely to be related. We use a meta-path-based method to compute drug similarities by counting the number of meta-paths (drug-target-drug) between drugs. These similarity measures are normalized and stored as and , representing network-level similarities between drugs and targets, respectively.

By integrating both molecular-level and network-level perspectives, we select positive samples for each drug and target. Specifically, to identify positive samples for drug , we sum and , then sort the i-th row in descending order, selecting the top K samples as positive samples, with the remaining designated as negative samples.

**The Explanation of the Contrastive Learning Formula** For Drug , the contrastive learning loss at the network scale are as follows:

where typically represents cosine similarity and the parameter serves as a temperature coefficient, which is instrumental in modulating the similarity distribution of embeddings, thereby enhancing the discernibility between positive and negative samples. The value of is generally determined through empirical methods. The sets and represent the positive and negative sample sets corresponding to , respectively.

Similarly, the contrastive learning loss at the network scale are as follows:

In the similar way, we can define , , and for the target as follows:

**3 Hyperparameter Settings**

The hyperparameter settings in our experiment are shown in the Supplementary Table 4.

**Supplementary Table** **4. Hyperparameter Settings of the Model. \* indicates that the hyperparameter settings differ on the Davis and KIBA datasets.**

|  |  |
| --- | --- |
| **Hyperparameter** | **Value(s)** |
| Epoch | 2000 |
| Batch size | 512 |
| Optimizer | Adam |
| Learning rate | 0.0005 |
| Length of the drug SMILES | 100 |
| Length of the target sequence | 1200 |
| *Dropout* rate of the drug/protein graph\* | 0.2/0 |
| Number of GraphSAGE layers on the molecular scale | 3 |
| Number of GNN layers on the network scale | 2 |
| Pooling method for the sequence | Max |
| Pooling method for the graph | Mean |
| The temperature parameter | 0.8 |
| Balancing the bias terms of the two scales | 0.5 |
| The number of positive samples for each node\* | 3/10 |

**4 Ablation Study**

This section displays the detailed experimental data from the ablation studies conducted on the Davis and KIBA datasets, as shown in Table 5.

**Supplementary Table** **5. Ablation Study Performance on Davis and KIBA Datasets..**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Dataset** | **Variants** |  |  |  |  |
|  | without Semantic Information | 0.159 | 0.911 | 0.795 | 0.895 |
|  | without Structural Information | 0.159 | 0.910 | 0.797 | 0.895 |
| Davis | without Network Scale | 0.198 | 0.902 | 0.736 | 0.868 |
|  | without Contrastive Learning | 0.168 | 0.908 | 0.783 | 0.889 |
|  | with Triple Contrastive Learning | 0.162 | 0.910 | 0.791 | 0.893 |
|  | TCCL | **0.154** | **0.915** | **0.801** | **0.899** |
|  | without Semantic Information | 0.126 | 0.900 | 0.815 | 0.905 |
|  | without Structural Information | 0.130 | 0.898 | 0.806 | 0.902 |
| KIBA | without Network Scale | 0.130 | 0.898 | 0.806 | 0.902 |
|  | without Contrastive Learning | 0.129 | 0.900 | 0.810 | 0.903 |
|  | with Triple Contrastive Learning | 0.129 | 0.898 | 0.806 | 0.902 |
|  | TCCL | **0.122** | **0.904** | **0.821** | **0.908** |