Hierarchical Cross-level Graph Contrastive Learning for Drug-Drug Interaction Prediction

Yuhan Ye[1], Jingbo Zhou[2,⋆], Shuangli Li[1], Congxi Xiao[1], Haochao Ying[3], and Hui Xiong[4,⋆]

[1]The University of Science and Technology of China
[2]Business Intelligence Lab, Baidu Research [3]Zhejiang University
[4]The Hong Kong University of Science and Technology (Guangzhou)
[1]xyqqj@mail.ustc.edu.cn, [2]{zhoujingbo, lishuangli, xiaocongxi}@baidu.com
[3]haochaoying@zju.edu.cn, [4]xionghui@ust.hk

Abstract. Drug-Drug Interaction (DDI) prediction is crucial for various biomedical applications like polypharmacy. Recently, some graph learning-based methods achieved promising performance in DDI prediction. However, limited attention has been given to the integration of substructure information and drug relationships to capture complex DDI patterns using self-supervised learning techniques. To this end, we propose a novel hierarchical cross-level graph contrastive learning framework named HCC, aimed at capturing hierarchical structural information and hidden DDI patterns. Firstly, we construct a drug-motif interaction graph to extract semantic motifs and model complex connections among drugs and motifs. Then, we design motif- and molecule-level self-supervised tasks. One task learns the motif-driven connectivity of the drug-motif graph, while the other learns global similarity of molecular graphs. Finally, a cross-level contrastive learning module is introduced to align multi-view information. Extensive evaluation on real-world datasets demonstrates that our method outperforms existing competitors.

Keywords: Drug-Drug Interaction · Graph Contrastive learning.

1 Introduction

Drug Drug Interactions (DDI) can occur when two or more drugs interact pharmacologically in the human body. As shown in Figure 1(a), these interactions can result in various biological consequences affecting drug efficacy. While polypharmacy is necessary for many medical applications, such as multiple health conditions and cancer treatment [13], the attendant potential DDI problem may raise serious risk of adverse effects [12]. Hence, it's essential to identify drug interactions to improve healthcare outcomes.

Efforts in DDI detection have intensified, as shown in Figure 1(b), with traditional methods relying on costly wet chemical experiments [18], prompting increased attention toward automated computational DDI prediction. Early

⋆ Corresponding authors. This work was done when the first author was an intern at Baidu Research under the supervision of Jingbo Zhou.

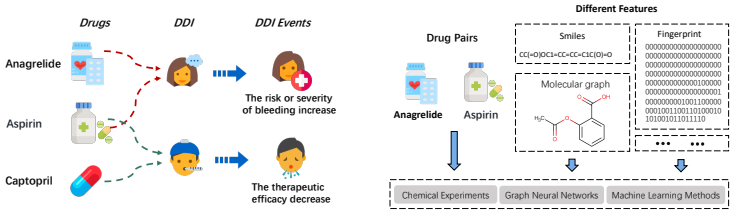


Fig. 1. (a) An example of DDI events. (b) Existing features and methods in DDI.

machine learning methods explored various drug similarities using heterogeneous features like fingerprint [16], assuming that structurally similar drugs might share similar DDI. Recently, graph learning-based methods have emerged, leveraging drug molecular graphs [17], knowledge graphs [10], and interaction graphs [17] for DDI prediction. Despite these advancements, a compelling demand remains for more accurate prediction. First, while self-supervised learning methods have gained utilization for representation learning in molecular property prediction [22], the development of specialized methods tailored to the abundant relations and structural patterns for DDI prediction remains scarce. Second, while internal chemical substructures and external drug relations have proven effective, integrating both to improve hierarchical contrastive learning for DDI prediction has received insufficient focus.

Therefore, in this work, we propose a novel Hierarchical Cross-level graph Contrastive learning framework (HCC) for DDI prediction, aiming to handle complex hidden DDI patterns through motif- and molecule-level self-supervised learning tasks, along with a cross-level contrastive learning component. At the molecule level, drugs are depicted as molecular graphs for detailed structure modeling, with atoms as nodes and chemical bonds as edges. We utilize a global similarity approximation task on these molecular graphs, leveraging enriched structural and semantic domain knowledge. Simultaneously, we extract special substructures called motifs and construct a drug-motif interaction graph at the motif level as an augmentation perspective to facilitate comprehensive substructure modeling and relationship learning, where nodes represent drugs or motifs. The motif-level task employs a motif-driven edge reconstruction strategy, integrating connectivity information from the graph. To ensure cohesive information integration across various perspectives, we introduce a cross-level contrastive learning component that maximizes mutual information between representations from two perspectives of the same molecule. Finally, drug embeddings derived from the well-pretrained model facilitate downstream DDI prediction. The key contributions of our work are summarized as follows:

– We introduce a novel Hierarchical Cross-level graph Contrastive learning framework (HCC) for DDI prediction, aiming to capture comprehensive information about drugs from diverse perspectives.

– The proposed cross-level contrastive learning not only enriches the representation learning process but also enhances our ability to predict DDIs accurately.
– Extensive experiments on different scale real-world datasets demonstrate the effectiveness of the proposed HCC over state-of-the-art approaches.

2 Related Work

DDI prediction. The increase in biomedical data availability has catalyzed early machine learning applications. DeepDDI [12] constructed a structural similarity profile based on fingerprint, while DDIMDL [2] combined heterogeneous drug features. Recently, graph neural networks (GNNs) have been pivotal in enhancing DDI prediction. CASTER [6] developed a deep auto-encoding method to embed functional drug representations. SSI-DDI [11] utilized GNNs to identify DDI through drug substructure interactions. MIRACLE [17] and DSN-DDI [9] considered both inter- and intra-view information of drug molecules. MDNN [10] collected drug-related entities to conduct drug knowledge graphs for learning.

Self-supervised Learning of molecular graphs. Self-supervised learning techniques have been widely used for molecular representation learning, incorporating domain knowledge. Hu et al. [5] introduced multiple pretraining strategies, i.e., context prediction and node masking for node-level and graph property prediction for graph-level learning. Moreover, MGSSL [22] applied motif label predictions task at the graph level. MoCL [14] proposed a substructure substitution augmentation strategy, incorporating local and global domain knowledge for richer representations. KCL [4] constructed a chemical element KG to guide graph augmentation and enhance molecular graph contrastive learning.

3 Preliminaries

In this section, we introduce the notations and formulate the DDI problem. **Definition 1: Molecular Graph**. Given a drug set D, we define each drug molecular graph as $\mathcal{D}_i = (\mathcal{A}, \mathcal{B})$, where $\mathcal{A}$ denotes the atom set and $\mathcal{B}$ denotes the chemical bonds set. Moreover, we extract meaningful subgraph as motif:$\mathcal{M} = \{\mathcal{A}_M, \mathcal{B}_M\}(\mathcal{A}_M \subseteq \mathcal{A}, \mathcal{B}_M \subseteq \mathcal{B})$.

Definition 2: Drug-Motif Interaction Graph. We denote the constructed drug-motif interaction graph as $\mathcal{G}(\mathcal{V}, \mathcal{E})$. There are two types of nodes on this graph: drug nodes $\mathcal{V}^D = \{v_1^D, ..., v_m^D\}$ and motif nodes $\mathcal{V}^M = \{v_1^M, ..., v_n^M\}$, with $\mathcal{V} = \mathcal{V}^D \cup \mathcal{V}^M$. Each drug node v_i^D and motif node v_i^M are associated with their graph structure $\mathcal{D}_i$ and $\mathcal{M}_i$. We will detail the construction in Section 4.1.

Problem: Drug-Drug Interaction Prediction. The DDI prediction task aims to predict potential interactions between drug pairs. Given a DDI dataset $\{(D_i, D_j, r)_k\}_{k=1}^N$ and interaction types $\mathcal{I} = \{I_i\}_{i=1}^C$ where C is the number of possible types, $\{D_i, D_j\} \in D$ denote a drug pair with interaction r of type I_i. Our objective is to learn a model $\mathcal{M}_{ddi} : D \times D \times \mathcal{I} \rightarrow y_{ddi} \in [0, 1]$ that predicts the probability of a given interaction type occurring between input drug pairs.

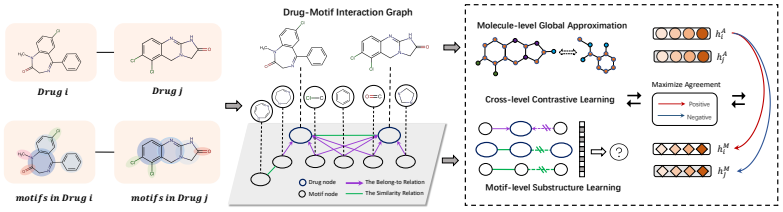


Fig. 2. Illustration of our proposed framework. Left: Inputs. Middle: The drug-motif interaction graph. Right: The hierarchical cross-level contrastive learning strategy.

4 Model Framework

As illustrated in Figure 2, we introduce the hierarchical cross-level contrastive learning framework for DDI prediction. At first, we construct a drug-motif interaction graph, modeling interactive structural information between drugs and motifs as an augmentation view to complement drug information. To capture complex DDI patterns, we design hierarchical self-supervised tasks, including a Motif-level substructure learning strategy and Molecule-level global similarity matrix approximation, facilitating representation learning from diverse perspectives. Moreover, we introduce a Cross-level contrastive learning module to fuse representations, promoting the mutual supplementation of diverse information.

4.1 Drug-motif interaction graph construction

To construct the drug-motif interaction graph, we first extract meaningful motifs to build a motif vocabulary, forming two types of nodes: drug and motif nodes. Then, we establish three types of edges representing drug-drug, drug-motif, and motif-motif interactions to capture various interactive relationships.

Motif extraction. We use a molecular fragmentation process to extract motifs and generate motif vocabulary. To ensure the chemical validity of the extracted motifs, we follow these steps based on chemical domain knowledge. Firstly, for coordination compound molecules, we break coordination bonds and add the ligand to the vocabulary, as ligands naturally function as effective motifs. Then, following [1], we apply the Breaking of Retrosynthetically Interesting Chemical Substructures (BRICS) algorithm to decompose molecules into diverse high-quality motifs based on chemical reactions. Finally, we include post-processing procedures from [22] to improve the occurrence frequency of motifs.

Graph construction. After creating the motif vocabulary, we construct the node set of the graph. Then, we introduce three types of edges to model diverse dependencies, including drug-motif inclusion relationships, motif-motif structural similarity, and drug-drug structural similarity. The motivation for building drug-motif edges is that different motifs can result in different reactions between drugs [13], emphasizing the importance of modeling the motifs contained within a drug. Formally, if a drug i contains a motif j, we add a drug-motif edge to represent this inclusion relationship. Moreover, considering the

different importance of motifs to drug interactions, we follow [21] to compute TF-IDF value between connected nodes as the edge weight. The inclusion edge enhances the local connections and enables the capture of higher-order relations indirectly. Additionally, to directly leverage global relationships between two drugs or motifs, we establish drug-drug edges and motif-motif edges based on their chemical structure similarity. Specifically, we first extract MACCS keys [3] to represent molecular structure features. Then, we compute the similarity between each pair of drugs or motifs using the Tanimoto coefficient through their MACCS keys, obtaining a similarity score. Next, drug or motif pairs with a similarity score exceeding threshold γ will be linked by an edge, enhancing global connections. The constructed drug-motif interaction graph effectively captures diverse relationships, augmenting the molecular graph with additional information, including internal meaningful motifs and external drug relationships.

4.2 Hierarchical cross-level contrastive learning

The hierarchical cross-level contrastive learning framework HCC aims to capture DDI patterns from Motif-level, Molecule-level, and Cross-level perspectives.

Motif-level substructure learning. As the constructed drug-motif interaction graph contains rich structural information, our main idea is to develop a self-supervised task based on the motif-driven structural relations of different nodes to exploit potential interactions. We introduce the edge masking task focused on reconstructing motif-motif, motif-drug, and drug-motif edges. Explicitly learning the inclusion relations between drugs and motifs enables the model to understand intrinsic structural characteristics and identify correlations relevant to potential drug reactions, enhancing insight into local connectivity. Implicitly reconstructing relationships among drugs involving multiple motifs provides aggregation-level interactive dependencies, enhancing motif-view structure learning for DDI prediction and contributing to global connectivity learning.

Specifically, considering the drug-motif graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, we initially randomly mask n_e edges of three types to create the set $\mathcal{E}_M \in \mathcal{E}$. Meanwhile, we adopt a negative sampling strategy to select negative node pairs and obtain the negative set $\overline{\mathcal{E}_M} = \{(v_i, v_j) | v_i, v_j \in \mathcal{V}, (v_i, v_j) \notin \mathcal{E}\}$. Formally, for node pairs $\{v_i, v_j\}$ in our sample set $\widetilde{\mathcal{E}^M} = \mathcal{E}^M \cup \overline{\mathcal{E}^M}$, we employ a GNN encoder to gain motif-view embeddings $\{h_i^M, h_j^M\}$. We then utilize a classifier to learn the existence of relationships between specific nodes, with binary cross-entropy loss:

$$\mathcal{L}_{motif} = \sum -y_e log\, \mathcal{M}_e((h_i^M, h_j^M)) - (1 - y_e) log(1 - \mathcal{M}_e((h_i^M, h_j^M))), \quad (1)$$

where $\mathcal{M}_e(\cdot)$ denotes the classifier, a 2-layer Multi-Layer Perceptron (MLP).

Molecule-level global approximation. To complement our hierarchical pretraining scheme, we further focus on the internal structure and semantics of drug molecules from a global perspective. It is expected that structure-correlated

graphs that share similar semantics should be closer in the latent space [7]. To integrate molecular-level chemical domain knowledge, we design a similarity approximation task, which guides the model to retain global semantic information via pairwise comparisons, improving the ability to capture drug interrelations.

Specifically, we first leverage MACCS keys as inherent structural features of molecular graphs and calculate the original global similarity matrix Y^s through Tanimoto coefficient to measure pairwise drug relevance. Formally, given graph pair $\{\mathcal{D}_i, \mathcal{D}_j\}$, we generate molecule-view embeddings $\{h_i^A, h_j^A\}$ via another GNN encoder. The model $\mathcal{M}s$, combined with proximity loss, aligns estimated and original global similarity matrices for molecular-level approximation as:

$$\mathcal{L}_{mole} = \sum \left(\mathcal{M}_s((h_i^A, h_j^A)) - Y_{ij}^s\right)^2, \tag{2}$$

where Y_{ij}^s denotes the inherent global similarity between drugs i and j.

Cross-level contrastive learning for alignment. Considering the valuable insights from both motif- and molecule-level information in modeling drug interactions, integrating these cross-level embeddings is beneficial for DDI prediction. The molecule-view drug embeddings, derived from molecular graphs, include raw atom and bond attributes along with structural details. While the motif-view drug embeddings obtained from the drug-motif graph encapsulate internal meaningful motifs, external relationship information, and domain knowledge semantics. On one hand, the representations of the same molecule in both views can mutually enrich information from diverse perspectives. On the other hand, graph convolution in the drug-motif graph may cause over-smoothing, while relying solely on molecular graphs for drug representations risks overfitting to existing knowledge. To address these challenges, we introduce graph contrastive learning to integrate and balance information for multi-level representation alignment.

Formally, given drug molecular graphs and the drug-motif interaction graph, we use different graph encoders to obtain multi-view drug representations $\{h^A, h^M\}$. Then, we employ a contrastive objective that encourages the molecule-level representations to align with the representations of positive samples while being distinguishable from negative samples at the motif level, as follows:

$$l_{cross}(h_i^A, h_i^M) = -log\frac{e^{s(h_i^A, h_i^M)/\tau}}{e^{s(h_i^A, h_i^M)/\tau} + \sum_{i,j\neq 1}^{m} e^{s(h_i^A, h_j^M)/\tau}}, \tag{3}$$

where s calculates cosine similarity, τ is the temperature parameter. Negative samples include all other nodes from another view. The final contrastive learning loss $\mathcal{L}_{cross}$ is the mean across all samples, enabling a greater scope of representing rich but distinguished semantics for each drug molecule.

4.3 Pre-Training and Fine-tuning Framework

Pre-training. We combine the self-supervised pre-training tasks and the contrastive learning component to form the following overall objective function:

$$\mathcal{L} = \mathcal{L}_{motif} + \lambda_1 \mathcal{L}_{mole} + \lambda_2 \mathcal{L}_{cross}, \tag{4}$$

where λ_1, λ_2 are tuning parameters that control the importance.

Drug-drug interaction prediction. Given a DDI tuple (D_i, D_j, r), we obtain the drug pair embeddings $\{h_i, h_j\}$ from well-pretrained graph encoder, the DDI prediction can be represented as the probability as follows:

$$P(D_i, D_j, r) = \mathcal{M}_{ddi}(h_i, h_j, R_i), \tag{5}$$

where $\mathcal{M}_{ddi}$ is the predictive model, specifically a logistic regression function, and R_i is the learnable representation of the interaction r. The model is fine-tuned by minimizing the cross-entropy loss, as expressed below:

$$\mathcal{L}_{ddi} = -\frac{1}{N}\sum_{k=1}^{N} y_{ddi} log(p_k) + (1 - y_{ddi}) log(1 - p_k), \tag{6}$$

where N is the tuple number, p_k denotes the predicted probability of interaction for a DDI tuple, y_{ddi} is the presence of interaction r between drug D_i and D_j.

5 Experiments

We conduct extensive experiments on two DDI datasets to evaluate the effectiveness of our model. The code of HCC is available at `https://github.com/PaddlePaddle/PaddleHelix/tree/dev/apps/drug_drug_interaction/HCC`.

5.1 Experimental Settings

Datasets. We conducted evaluations on two kinds of datasets: a small-scale dataset **DeepDDI** and a large-scale dataset **DrugBankDDI**. DeepDDI is released by [12] which contains 1,704 drugs and 191,511 pair-wise DDI across 86 interaction types, each described by a general sentence structure. DrugBankDDI is crawled and parsed the verified DDI from DrugBank (V5.1.9) [19] by ourselves. In preprocessing, we excluded drugs that can not be converted into graphs from SMILES strings by RDKit[1]. After preprocessing, the dataset contains 3,643 drugs and 1,151,039 pairwise samples classified into 174 interaction types.

Baselines. We compare our framework with several comparative methods: three GNN models (GCN [8], GAT [15], GIN [20]), four state-of-the-art (SOTA) DDI prediction methods (SSP-MLP [12], SSI-DDI [11], MIRACLE [17], DSN-DDI [9]), and four SOTA molecular representation learning methods (MoCL [14], KCL [4], MGSSL [22] and HM-GNN [21]).

Experimental Configurations. We divided the dataset into training, validation, and test sets with a 6:2:2 ratio and repeated this process five times using different random seeds. The similarity threshold γ was set to 0.7 for small-scale and 0.65 for large-scale datasets. For pre-training, we used $\lambda_1 = 100, \lambda_2 = 10$ for small-scale datasets and $\lambda_1 = 10, \lambda_2 = 1$ for large-scale ones. We used ACC, F1-score, and AUC metrics as evaluation metrics.

[1] http://www.rdkit.org/

Table 1. Performance comparison on all datasets.

Method		DeepDDI			DrugbankDDI		
		ACC	F1	AUC	ACC	F1	AUC
GNN-based Methods	GCN	0.782(.003)	0.800(.002)	0.862(.002)	0.769(.001)	0.792(.001)	0.862(.001)
	GAT	0.811(.002)	0.819(.002)	0.883(.002)	0.802(.002)	0.813(.002)	0.887(.001)
	GIN	0.840(.001)	0.848(.002)	0.909(.001)	0.866(.002)	0.872(.001)	0.939(.001)
DDI-based Methods	SSP-MLP	0.802(.013)	0.811(.009)	0.885(.009)	0.876(.006)	0.880(.005)	0.939(.004)
	SSI-DDI	0.844(.007)	0.851(.005)	0.920(.006)	0.882(.003)	0.885(.003)	0.945(.002)
	MIRACLE	0.859(.009)	0.885(.013)	0.930(.006)	0.923(.002)	0.926(.001)	0.968(.001)
	DSN-DDI	0.880(.012)	0.882(.013)	0.946(.010)	0.923(.008)	0.925(.008)	0.973(.005)
Molecule-based Methods	MGSSL	0.859(.001)	0.865(.001)	0.921(.002)	0.934(.000)	0.936(.000)	0.967(.000)
	MoCL	0.856(.002)	0.862(.003)	0.919(.002)	0.923(.001)	0.926(.001)	0.963(.000)
	KCL	0.878(.005)	0.880(.004)	0.941(.004)	0.866(.002)	0.870(.002)	0.930(.001)
	HM-GNN	0.872(.002)	0.878(.002)	0.934(.001)	0.927(.000)	0.930(.000)	0.966(.000)
	HCC (Ours)	**0.894(.002)**	**0.899(.002)**	**0.948(.001)**	**0.949(.001)**	**0.950(.001)**	**0.981(.001)**

5.2 Performance Comparison

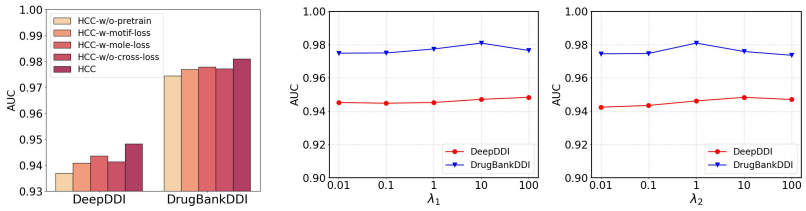
As shown in Table 1, our framework outperforms all baselines, demonstrating the effectiveness of our strategies. Furthermore, we can draw more detailed conclusions: (1) MIRACLE and DSN-DDI excel by effectively integrating multi-view information. (2) MGSSL achieves suboptimal results on DrugBankDDI, which incorporates motif generation tasks, highlighting the significance of motif compositional information in molecular representation learning. (3) KCL outperforms MoCL on DeepDDI, while MoCL excels on DrugBankDDI. Both chemical element and substructure information provide valuable enhancements for contrastive learning. Compared with these solutions, our framework performs the best thanks to its following advantages: (1) HCC utilizes complex relationships in drug-motif interaction graph, accessing rich contextual insight beyond single drugs. (2) It maximizes the benefits of motifs to communicate different molecules, beyond influencing the physicochemical properties within individual molecules. (3) The multi-level self-supervised tasks effectively utilize semantic and connective information, improving hierarchical understanding. The contrastive learning further enhances multi-view fusion and alignment for accurate DDI predictions.

5.3 Ablation Study

We conduct ablation studies to further validate the contributions of different designs in our framework. We compare our proposed HCC with its following variants: (i) **w/o pretrain**: w/o the pretaining process. (ii) **w only motif-loss**: with only motif-level loss in Eq.1. (iii) **w only mole-loss**: with only molecule-level loss in Eq.2; (iv) **w/o cross-loss**: w/o the contrastive learning loss in Eq.3. As shown in Figure 3(a), each component plays a critical role in DDI prediction.

5.4 Hyper-parameter Sensitivity Analysis

In this section, we analyze the effect of loss coefficient λ_1 and λ_2 when varying them from 0.01 to 100 on all datasets. As depicted in Figure 3(b), performance



(a) (b)

Fig. 3. (a) Comparison of HCC with its variants. (b) Parameter sensitivity analysis.

tends to increase with the higher values, while too large values may bring a slight drop. Appropriate contrastive learning can assist the model in integrating and balancing information of different levels and enhancing representation learning.

6 Conclusion

In this study, we proposed a hierarchical cross-level contrastive learning framework HCC for DDI prediction. We first constructed a drug-motif interaction graph, and then designed hierarchical self-supervised tasks from multi-level perspectives. Extensive experiments on real-world datasets demonstrated that our HCC outperforms state-of-the-art methods. In future work, we will expand our framework's application to additional molecular graph contexts.

Acknowledgment

This work was partially supported by Guangzhou-HKUST(GZ) Joint Funding ProgramGrant No.2023A03J0008), Education Bureau of Guangzhou Municipality, Guangdong Science and Technology Department, and the Foshan HKUST Projects (FSUST21-FYTRI01A).

References

1. Degen, J., Wegscheid-Gerlach, C., Zaliani, A., Rarey, M.: On the art of compiling and using'drug-like'chemical fragment spaces. ChemMedChem: Chemistry Enabling Drug Discovery **3**(10), 1503–1507 (2008)
2. Deng, Y., Xu, X., Qiu, Y., Xia, J., Zhang, W., Liu, S.: A multimodal deep learning framework for predicting drug–drug interaction events. Bioinformatics **36**(15), 4316–4322 (2020)
3. Durant, J.L., Leland, B.A., Henry, D.R., Nourse, J.G.: Reoptimization of mdl keys for use in drug discovery. Journal of chemical information and computer sciences **42**(6), 1273–1280 (2002)
4. Fang, Y., Zhang, Q., Yang, H., Zhuang, X., Deng, S., Zhang, W., Qin, M., Chen, Z., Fan, X., Chen, H.: Molecular contrastive learning with chemical element knowledge graph. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 36, pp. 3968–3976 (2022)

5. Hu, W., Liu, B., Gomes, J., Zitnik, M., Liang, P., Pande, V., Leskovec, J.: Strategies for pre-training graph neural networks. arXiv preprint arXiv:1905.12265 (2019)
6. Huang, K., Xiao, C., Hoang, T., Glass, L., Sun, J.: Caster: Predicting drug interactions with chemical substructure representation. In: Proceedings of the AAAI conference on artificial intelligence. vol. 34, pp. 702–709 (2020)
7. Jin, W., Derr, T., Liu, H., Wang, Y., Wang, S., Liu, Z., Tang, J.: Self-supervised learning on graphs: Deep insights and new direction. arXiv preprint arXiv:2006.10141 (2020)
8. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907 (2016)
9. Li, Z., Zhu, S., Shao, B., Zeng, X., Wang, T., Liu, T.Y.: Dsn-ddi: an accurate and generalized framework for drug–drug interaction prediction by dual-view representation learning. Briefings in Bioinformatics **24**(1), bbac597 (2023)
10. Lyu, T., Gao, J., Tian, L., Li, Z., Zhang, P., Zhang, J.: Mdnn: A multimodal deep neural network for predicting drug-drug interaction events. In: IJCAI. pp. 3536–3542 (2021)
11. Nyamabo, A.K., Yu, H., Shi, J.Y.: Ssi–ddi: substructure–substructure interactions for drug–drug interaction prediction. Briefings in Bioinformatics **22**(6), bbab133 (2021)
12. Ryu, J.Y., Kim, H.U., Lee, S.Y.: Deep learning improves prediction of drug–drug and drug–food interactions. Proceedings of the National Academy of Sciences **115**(18), E4304–E4311 (2018)
13. Silverman, R.B., Holladay, M.W.: The organic chemistry of drug design and drug action. Academic press (2014)
14. Sun, M., Xing, J., Wang, H., Chen, B., Zhou, J.: Mocl: Data-driven molecular fingerprint via knowledge-aware contrastive learning from molecular graph. In: Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining. pp. 3585–3594 (2021)
15. Velickovic, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., Bengio, Y., et al.: Graph attention networks. stat **1050**(20), 10–48550 (2017)
16. Vilar, S., Harpaz, R., Uriarte, E., Santana, L., Rabadan, R., Friedman, C.: Drug—drug interaction through molecular structure similarity analysis. Journal of the American Medical Informatics Association **19**(6), 1066–1074 (2012)
17. Wang, Y., Min, Y., Chen, X., Wu, J.: Multi-view graph contrastive representation learning for drug-drug interaction prediction. In: Proceedings of the Web Conference 2021. pp. 2921–2933 (2021)
18. Whitebread, S., Hamon, J., Bojanic, D., Urban, L.: Keynote review: in vitro safety pharmacology profiling: an essential tool for successful drug development. Drug discovery today **10**(21), 1421–1433 (2005)
19. Wishart, D.S., Feunang, Y.D., Guo, A.C., Lo, E.J., Marcu, A., Grant, J.R., Sajed, T., Johnson, D., Li, C., Sayeeda, Z., et al.: Drugbank 5.0: a major update to the drugbank database for 2018. Nucleic acids research **46**(D1), D1074–D1082 (2018)
20. Xu, K., Hu, W., Leskovec, J., Jegelka, S.: How powerful are graph neural networks? arXiv preprint arXiv:1810.00826 (2018)
21. Yu, Z., Gao, H.: Molecular representation learning via heterogeneous motif graph neural networks. In: International Conference on Machine Learning. pp. 25581–25594. PMLR (2022)
22. Zhang, Z., Liu, Q., Wang, H., Lu, C., Lee, C.K.: Motif-based graph self-supervised learning for molecular property prediction. Advances in Neural Information Processing Systems **34**, 15870–15882 (2021)