

Closer to Biological Mechanism: Drug-Drug Interaction Prediction from the Perspective of Pharmacophore

Mingliang Dou^{1*}, Linfeng Wen^{2,5*}, Jinyang Xie¹, Jijun Tang^{3,5}, Shiqiang Ma^{3†}, Fei Guo^{4†}

¹College of Computer Science and Technology, Taiyuan University of Technology, Taiyuan, China

²Southern University of Science and Technology, Shenzhen, China

³Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China

⁴School of Computer Science and Engineering, Central South University, Changsha, China

⁵Faculty of Computer Science and Control Engineering, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China

doumingliang@tyut.edu.cn, 12433482@mail.sustech.edu.cn, jiejinyang7219@link.tyut.edu.cn

(sq.ma, jj.tang)@siat.ac.cn, guofei@csu.edu.cn

Abstract

Drug combinations are widely used in modern medicine but may cause severe adverse drug reactions. Therefore, making effective drug-drug interactions (DDI) prediction is crucial for pharmacovigilance. Existing DDI prediction models are typically built from a structural perspective, assuming that drugs with similar molecular structures may exhibit similar interactions. However, such approaches overlook the biological mechanisms underlying DDI in the human body. This not only weakens the generalization ability of the model, but also makes its interpretability less convincing. Inspired by this, we propose a new method called PC-DDI. Unlike structure-based models, PC-DDI utilizes pharmacophores as basic unit, and designs a complete pharmacophore feature processing framework. It further constructs a pharmacophore-based bipartite graph to model interactions between pharmacophores. This approach allows us to explore the underlying mechanisms of DDI from a functional perspective. We also design a spatial attention weight graph convolution module to optimize the message passing process by integrating pharmacophore position features with node features. Furthermore, we apply causal inference to identify key pharmacophores in pharmacophore bipartite graph, enhancing the interpretability. Compared with the SOTA, PC-DDI achieves an accuracy improvement of 1.84% under the transductive setting and consistently outperforms others in all other experiments.

Introduction

Combination therapy leads to improved therapeutic outcomes for complex diseases. However, DDI can cause severe adverse reactions, posing irreversible risks to patients (Wang et al. 2021; Huang et al. 2020). Therefore, studying DDI is crucial for minimizing harms caused by drug combinations and guiding pharmaceutical development. In vitro and in vivo experiments are crucial for understanding drug

action, but their high cost, time demands, and labor intensity limit large-scale evaluations (Hao et al. 2023). Thus, it is essential to design computational models that not only predict DDI with high accuracy but also deliver interpretable explanations.

In recent years, most DDI prediction methods have been built from a structural perspective, assuming that drugs with similar substructures are likely to exhibit similar interactions (Li et al. 2023; Niu et al. 2024) (see Figure 1b). These methods employ GNN to extract substructures and identify those with high co-attention scores as key substructures responsible for DDI. However, they often identify incomplete or pharmacologically inactive substructures, and the co-attention scores they use may not provide meaningful explanations (Jain and Wallace 2019). This limits the understanding of DDI mechanisms and may mislead further research. Similarly, a few recent studies have explored the role of functional groups in DDI (Zhang et al. 2025). Yet, these approaches remain structure-based and overlook the fact that the same functional group can have different effects depending on its molecular context. Such ambiguity can confuse the model during training and impair its ability to learn meaningful, generalizable patterns. In summary, these models merely capture statistical correlations based on structural similarity without uncovering the underlying causal mechanisms of DDIs, which constrains both their performance and interpretability.

The mechanisms of most DDI are associated with pharmacokinetic or pharmacodynamic processes, they are typically indirect effects mediated by endogenous substances in the body (Gottlieb et al. 2012)(as shown in Figure 1a). Compared to drug substructures or functional groups, pharmacophores—defined as the three-dimensional features essential for molecular binding to biological targets—can directly interact with human targets and are not restricted to specific molecular scaffolds. As a result, pharmacophore-based models not only better reveal the causal mechanisms underlying DDI (as illustrated in Figure 1c), but also capture the complete functional patterns of drug molecules involved

*Equal contribution

†Shiqiang Ma and Fei Guo are corresponding authors.

Copyright © 2026, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

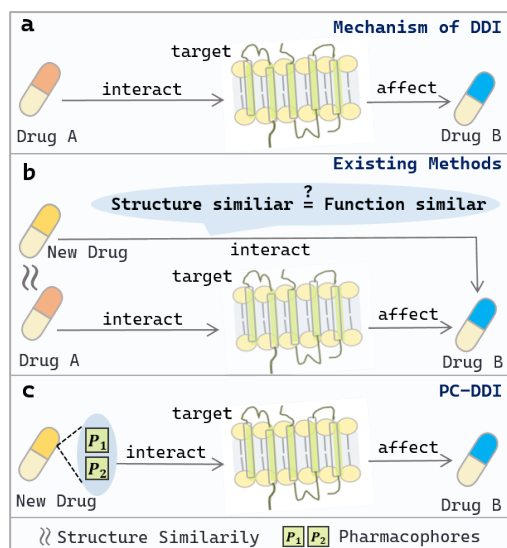


Figure 1: **a:** DDI mechanism: DDI is not a direct interaction between drugs, but an indirect effect mediated through targets in the body, which may alter physiological states and affect co-administered drugs. **b:** Most existing methods determine whether drugs interact based on structural similarity. **c:** PC-DDI takes a functional perspective, exploring the associations between pharmacophores and DDI, to reveal the functional patterns underlying how drugs interact in the body.

in biochemical reactions, thereby improving model generalizability and predictive accuracy—particularly for newly developed drugs.

Building on this foundation, we propose a novel DDI prediction method that treats pharmacophores as the fundamental modeling unit. PC-DDI extracts pharmacophore features from drug molecular graphs and organizes them into a bipartite pharmacophore graph, where pharmacophores from one drug are connected to those from the other. We design a spatial attention weight graph convolution module (SAWGCN) to optimize the message passing process by integrating pharmacophore position features with node features. In addition, we introduce a causal learning approach to identify the pharmacophores that have the greatest impact on DDI prediction.

Our major contributions can be summarized as follows:

- We propose PC-DDI, which to the best of our knowledge is the first pharmacophore-based framework for DDI prediction, enhancing model generalization by capturing interactions among pharmacophores.
- We design SAWGCN to incorporate spatial information into pharmacophore graph features, yielding more comprehensive graph representations.
- We introduce a causal inference approach to identify key pharmacophores with causal influence on DDI, improving the interpretability of the model.

Related Works

Drug-Drug Interaction Prediction

Effective DDI prediction methods hold immense significance for the clinical application of medications (Deng et al. 2022; Ren et al. 2025). In recent years, machine learning approaches have become widely adopted in this domain. Similarity-based methods (Takeda et al. 2017; Yan et al. 2021) assume that if drug i interacts with drug j , then drugs similar to i may also interact with j . Knowledge graph-based methods (Hong et al. 2022; Chen et al. 2024) integrate biomedical entities like proteins, genes, and enzymes, leveraging rich semantic graph information to improve prediction. Recently, deep learning based on molecular structures has advanced rapidly, evolving from molecular sequences (Huang et al. 2020) to molecular graphs (Wang et al. 2021; Chen et al. 2021) and now focusing on molecular substructures in DDI (Nyamabo, Yu, and Shi 2021; Niu et al. 2024; Zhang et al. 2025). Despite improved accuracy, existing methods may capture incomplete or inactive substructures, weakening mechanistic interpretability, limiting generalization to new drugs, and potentially misleading downstream research.

Causal Learning

Causal learning is a type of learning that aims to understand how one event produces or influences another event (Gopnik and Schulz 2007). In recent years, causal learning has gained increasing attention for enhancing the interpretability and transparency of GNN. Conditional probability-based method model feature dependencies and generate subgraph-based explanations, enabling local interpretation of model outputs (Luo et al. 2020). Methods based on Granger causality treat the explanation of model predictions as a statistical causal inference task, identifying cause-effect relationships between input features and output decisions by measuring temporal or structural dependencies (Lin, Lan, and Li 2021). Additionally, certain strategies apply reinforcement learning to identify the most influential substructures in molecular graphs, producing more accurate and concise explanations of model behavior (Wang et al. 2022; Hu, Wu, and Qian 2025). In our research, we aim to introduce causal learning methods to extract the most influential pharmacophores for DDI prediction while enhancing the model’s interpretability.

Method

In this study, we propose PC-DDI, a DDI prediction model from the perspective of pharmacophore. The overall framework is shown in Figure 2a, and the detailed method is described below.

Problem Definition

The goal of DDI prediction is to determine if a pair of drugs has a specific interaction. Given a DDI triplet (G_h, G_t, r) , where G_h and G_t are molecule graphs of the drugs h and t , and r is DDI type, we need to judge whether the interaction exists. This is treated as a binary classification problem. The output of DDI can be expressed as $f: G_h \times G_t \times r \rightarrow [0,$

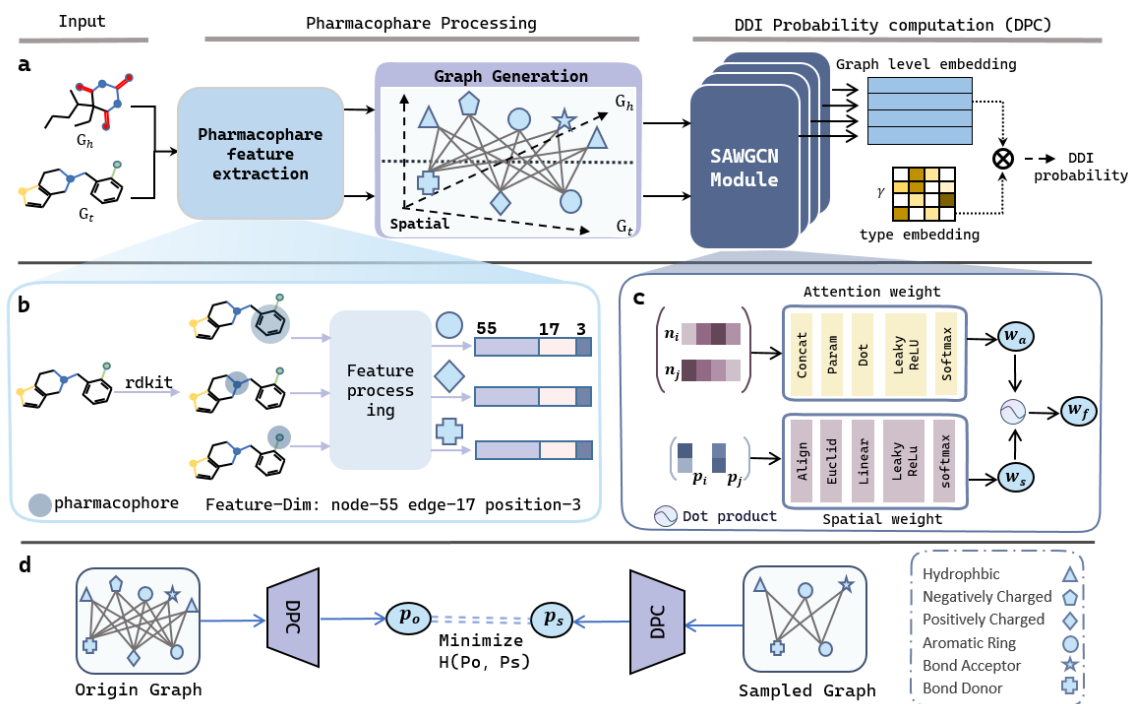


Figure 2: Pipeline of PC-DDI. **(a)** The overall framework of PC-DDI. The input molecular graphs are processed to extract pharmacophore features, based on which a pharmacophore graph is constructed. SAWGCN is then applied to generate a graph level embedding with spatial information. Finally, the graph level embedding is combined with the type embedding to predict the probability of DDI. \otimes in graph denote the DDI interaction scoring function defined in Equation (13). **(b)** Pharmacophore feature extraction module. **(c)** Spatial attention weight graph convolution module. **(d)** Causal learning module.

1]. f gives the prediction probability, using 0.5 as the threshold, outputting 1 if $f > 0.5$, else 0. 1 stands for interaction exists, 0 stands for no interaction.

Pharmacophore Feature Extraction

Our model takes a DDI triplet (G_h, G_t, r) as input, where G_h and G_t are molecular graphs derived from SMILES strings using RDKit, and r denotes the interaction type. Each molecular graph is represented as $G = (V, O)$, with atom features $\mathbf{v}_i \in \mathbb{R}^{1 \times 55}$ and bond features $\mathbf{o}_{ij} \in \mathbb{R}^{1 \times 17}$.

To extract pharmacophore features, we first identify different kinds of pharmacophores from each drug, including hydrophobic centers, aromatic rings, hydrogen bond donors, hydrogen bond acceptors, positive charged groups and negative charged groups. For each type of pharmacophore, we use RDKit to obtain the atom sequence, atom features, and spatial coordinates associated with that pharmacophore (as shown in Figure 2b).

The pharmacophore node feature \mathbf{n}_p is obtained by applying sum pooling over all atom features:

$$\mathbf{n}_p = \sum_{i=1}^N \mathbf{v}_i \quad (1)$$

where $\mathbf{n}_p \in \mathbb{R}^d$ is the pharmacophore node feature vector, $\mathbf{v}_i \in \mathbb{R}^d$ is the feature vector of the i -th atom, and N is the total number of atoms in the pharmacophore.

The spatial position of each pharmacophore is defined as the centroid (mean) of its constituent atomic coordinates—a coarse yet conformation-robust representation—and since all pharmacophores within the same molecule are derived from a single molecular conformation, spatial consistency among them is preserved. To extract pharmacophore edge features, we first aggregate all edge features associated with the atoms within each pharmacophore. These edge features are summed and then divided by the number of atoms in the sequence to obtain a unified edge feature for that pharmacophore. When two pharmacophores from different drugs are connected, their respective edge features are summed to form the final edge feature between them.

We thereby obtain a pharmacophore graph $G = (N, E, S)$, where pharmacophore node features $\mathbf{n}_i \in \mathbb{R}^{1 \times 55}$, edge features $\mathbf{e}_{ij} \in \mathbb{R}^{1 \times 17}$, and spatial features $\mathbf{s}_i \in \mathbb{R}^{1 \times 3}$.

Pharmacophore Graph Generation

During the generation of pharmacophore node features, pharmacophores of the same type with identical atom sequences will produce the same feature representations. To avoid redundant computations on repeated features in subsequent model processing, we perform deduplication of pharmacophores with identical node feature.

Then we construct a bipartite graph (as shown in the pharmacophore graph in Figure 2), where nodes represent phar-

macophores and edges are generated by fully connecting pharmacophores from different drugs.

Spatial Attention Weight Graph Convolution

We propose the SAWGCN to effectively aggregate neighborhood information in pharmacophore graphs by combining node feature with spatial feature. Node features are first projected into a lower-dimensional space:

$$\mathbf{n}_i^{\text{proj}} = \mathbf{W}_1 \mathbf{n}_i^{(l)}, \quad \mathbf{n}_j^{\text{proj}} = \mathbf{W}_2 \mathbf{n}_j^{(l)} \quad (2)$$

where $\mathbf{n}_i^{(l)}$ and $\mathbf{n}_j^{(l)}$ denote the features of nodes i and j at layer l , and $\mathbf{W}_1, \mathbf{W}_2$ are learnable projection matrices.

An attention coefficient w_a is computed by concatenating the projected features and applying a learnable attention vector **ATTN**, followed by LeakyReLU and Softmax:

$$w_a = \text{Softmax} \left(\text{LeakyReLU} \left(\mathbf{ATTN}^\top [\mathbf{n}_i^{\text{proj}} \parallel \mathbf{n}_j^{\text{proj}}] \right) \right) \quad (3)$$

To incorporate spatial structure, we compute the Euclidean distance $d_{ij} = \|\mathbf{s}_i - \mathbf{s}_j\|_2$, where $\mathbf{s}_i, \mathbf{s}_j \in \mathbb{R}^3$ are the 3D coordinates, and apply exponential decay with a learnable weight \mathbf{W}_3 :

$$w_s = \text{Softmax} \left(\text{LeakyReLU} \left(\mathbf{W}_3 e^{-d_{ij}} \right) \right) \quad (4)$$

The final attention weight w_f combines both feature and spatial components, used to adjust the edge weights in the bipartite graph. The node feature is then updated by aggregating weighted messages from neighbors and edge features.

$$w_f = w_a \cdot w_s \quad (5)$$

$$\mathbf{n}_i^{(l+1)} = \mathbf{n}_i^{(l)} + \text{ELU} \left(\sum_{j \in \mathcal{N}(i)} w_f \cdot \mathbf{n}_j^{\text{proj}} + \mathbf{W}_{\text{edge}} \mathbf{e}_{ij} \right) \quad (6)$$

where $\mathcal{N}(i)$ is the set of neighbors of node i , $\mathbf{e}_{ij} \in \mathbb{R}^{d_e}$ is the edge feature, and \mathbf{W}_{edge} is a learnable transformation matrix. The output $\mathbf{n}_i^{(l+1)}$ is the updated node representation at the next layer.

Graph Level Embedding Generation

Graph level embeddings integrate information from the most representative pharmacophoric subgraph to achieve a comprehensive representation of the global features of pharmacophore graphs:

$$\mathcal{G}_{\text{sub}} = \text{SAGPooling}(\mathcal{G}), \quad \mathbf{g}_{\text{node}} = \sum_{i \in \mathcal{G}_{\text{sub}}} \mathbf{n}_i^s \quad (7)$$

$$\mathbf{g}'_{\text{edge}} = \text{ELU} \left(\mathbf{W}_{\text{edge}} \cdot \sum_{e \in \mathcal{G}_{\text{sub}}} \mathbf{e}_e \right) \quad (8)$$

where \mathcal{G} is the input pharmacophore graph and \mathcal{G}_{sub} is the subgraph extracted by SAGPooling, containing key pharmacophore nodes. \mathbf{n}_i^s and \mathbf{e}_e denote the feature vectors of the i -th node and e -th edge in \mathcal{G}_{sub} , respectively. The global node feature \mathbf{g}_{node} is obtained by summing all node embeddings, while the global edge feature $\mathbf{g}'_{\text{edge}}$ is computed by applying

a linear transformation followed by ELU activation to the summed edge embeddings.

The pharmacophore subgraph feature is then obtained by fusing node and edge embeddings via the Hadamard product:

$$\mathbf{g} = \mathbf{g}_{\text{node}} \odot \mathbf{g}'_{\text{edge}} \quad (9)$$

which effectively combines both node-level and edge-level information to enhance the overall graph representation.

Finally, graph level embeddings from multiple GNN layers are normalized and stacked to form a multi-layer feature representation \mathbf{G} :

$$\mathbf{G} = \left[\frac{\mathbf{g}^{(1)}}{\|\mathbf{g}^{(1)}\|_2} \parallel \dots \parallel \frac{\mathbf{g}^{(L)}}{\|\mathbf{g}^{(L)}\|_2} \right] \quad (10)$$

where $\mathbf{g}^{(l)}$ denotes the graph embedding at the l -th GNN layer, and $\|\cdot\|_2$ is the L2 norm. The resulting $\mathbf{G} \in \mathbb{R}^{B \times L \times D}$ serves as input for subsequent modules, with B being the batch size, L the number of GNN layers, and D the feature dimension.

DDI Prediction and Loss Function

The \mathbf{Q} and \mathbf{K} vectors are obtained via linear projections, where $\mathbf{G} \in \mathbb{R}^{B \times L \times D}$ is the input graph feature, and $\mathbf{W}_q, \mathbf{W}_k \in \mathbb{R}^{D \times d_k}$ are learnable parameters with $d_k = D/2$. The attention weight matrix is computed as:

$$\mathbf{E} = \tanh(\mathbf{Q}\mathbf{W}_q \oplus \mathbf{K}\mathbf{W}_k + \mathbf{b}) \quad (11)$$

$$\mathbf{A} = \text{Softmax}(\mathbf{E}\mathbf{a}) \quad (12)$$

where $\mathbf{b} \in \mathbb{R}^{d_k}$ is a bias term, $\mathbf{a} \in \mathbb{R}^{d_k}$ is a projection vector, and \oplus represents appropriate tensor expansion and addition operations. The resulting attention weights $\mathbf{A} \in \mathbb{R}^{B \times L \times L}$ capture the importance between different GNN layers. The result is normalized using Softmax to obtain the final attention matrix.

The final DDI prediction score is obtained by aggregating graph-level features across different layers, where $\mathbf{g}^{(i)}$ and $\mathbf{g}^{(j)} \in \mathbb{R}^{B \times d_g}$ denote the embeddings of the graph at different layers for a batch of B samples. The interaction score p_o is defined as:

$$p_o = \sum_{i,j} \alpha_{ij} \cdot \mathbf{g}^{(i)} \mathbf{M}_r (\mathbf{g}^{(j)})^\top \quad (13)$$

where α_{ij} denotes the attention coefficient for layer i and layer j from $\mathbf{A} \in \mathbb{R}^{B \times L \times L}$ (with L being the number of layers), and $\mathbf{M}_r \in \mathbb{R}^{d_g \times d_g}$ denotes relation-specific interactions.

The model is trained using a pairwise loss based on sigmoid activation and binary cross-entropy. For each positive sample p_i and its corresponding negative sample p'_i , the loss is:

$$\mathcal{L} = -\frac{1}{q} \sum_{i=1}^q (\log \sigma(p_i) + \log \sigma(-p'_i)) \quad (14)$$

\mathcal{L} encourages higher scores for interacting pairs and lower scores for non-interacting ones.

Model	Transductive (Known-Known)				Inductive-S1 (Unknown-Unknown)				Inductive-S2 (Unknown-Known)			
	ACC	AUC-ROC	AP	F1	ACC	AUC-ROC	AP	F1	ACC	AUC-ROC	AP	F1
DeepDDI	0.9080	0.9625	0.9426	0.9321	0.5904	0.6809	0.6903	0.3770	0.7288	0.8273	0.8319	0.6702
SSI-DDI	0.8807	0.9366	0.8842	0.9115	0.6540	0.7343	0.5412	0.7503	0.7638	0.8423	0.8494	0.7354
GATDDI	0.9436	0.9862	0.9426	0.9251	0.6631	0.7275	0.6868	0.7161	0.6983	0.7729	0.7579	0.7301
GMPNN-CS	0.9494	0.9823	0.9501	0.9640	0.6857	0.7496	0.6532	0.7544	0.7772	0.8484	0.8487	0.7829
DSN-DDI	0.9344	0.9762	0.9360	0.9518	0.7342	0.8179	0.7034	0.8182	0.8192	0.9101	0.9109	0.8018
SRR-DDI	0.9663	0.9898	0.9669	0.9733	0.7025	0.7650	0.6850	0.7725	0.8015	0.8900	0.8950	0.7950
PEB-DDI	0.9683	0.9924	0.9688	0.9743	0.7565	0.8308	0.7588	0.8260	0.8301	0.9194	0.9205	0.8152
HDN-DDI	0.9708	0.9978	0.9722	0.9776	0.7711	0.8582	0.8124	0.8043	0.8617	0.9196	0.9169	0.8611
Ours	0.9832	0.9987	0.9928	0.9843	0.8049	0.8904	0.8254	0.8531	0.8802	0.9429	0.9404	0.8931

Table 1: Performance comparison under the D1. The best results are highlighted in **bold**.

Causal Learning to Identify Key Pharmacophore

Inspired by Granger causality (Granger 1969), we introduce a causal learning approach to identify the key pharmacophores that have a significant causal influence on DDI prediction.

Our goal is to minimize the cross entropy $H(p_o, p_s)$, where p_o is the prediction score using the full graph G_o , and p_s is the prediction score using a sampled subgraph of G_o . By minimizing this entropy, we aim to find subgraphs whose presence significantly influences the DDI prediction. To quantify the causal impact of each pharmacophore node, we define an **Importance Score** $I(v_j)$ for node v_j as:

$$I(v_j) = \frac{p_o - p_o(G_o \setminus \{v_j\})}{p_o} \quad (15)$$

where $p_o(G_o \setminus \{v_j\})$ denotes the prediction score obtained after removing node v_j from the original graph G_o . The score $I(v_j)$ reflects the relative change in prediction confidence upon node removal, serving as an indicator of its causal contribution to the prediction.

In PC-DDI, we iteratively prune pharmacophore nodes with the smallest $I(v_j)$ values, i.e., those have the least impact on the prediction, until the importance scores $I(v_j)$ of the remaining pharmacophores exceed a predefined threshold. The resulting subgraph G_s contains the pharmacophores that are most critical to DDI prediction.

Experiment

Datasets. We evaluate PC-DDI on five datasets, denoted as D1 to D5. D1, D2, and D3 correspond to the public DDI datasets from DrugBank 5.0.3 (Ryu, Kim, and Lee 2018), from Marinka et al. (Marinka Zitnik, Sosič, and Leskovec 2018), and from Lin et al. (Lin et al. 2023), respectively. To assess generalization to update interactions, we construct two additional datasets, D4 and D5. Both D4 and D5 use the full DeepDDI data for training. The test set of D4 contains newly emerged DDI triplets from DrugBank 5.0.5, while the test set of D5 consists of those from DrugBank 5.0.7.

Configuration. We evaluate the model under three tasks: transductive (Known–Known), where all test drugs are seen during training; inductive-S1 (Unknown–Unknown), where both drugs in a pair are unknown (unseen in training); and inductive-S2 (Unknown–Known), where one drug is unknown and the other is known.

Evaluation Metrics. To verify the performance of DDI prediction, four widely used metrics are selected: ACC, AUC-ROC, AP and F1-score. To reduce the impact of random errors, we report the average values of the four metrics across five repeated experiments.

Baseline Methods. We compare with 8 representative baseline models as follows: Deep-DDI (Ryu, Kim, and Lee 2018), GAT (Velickovic et al. 2017), SSI-DDI (Nyamabo, Yu, and Shi 2021), GMPNN-CS (Nyamabo et al. 2022), and DSN-DDI (Li et al. 2023), PEB-DDI (Shen et al. 2023), SRR-DDI (Niu et al. 2024), HDN-DDI (Sun and Zheng 2025).

Implementation Details. We train the model for 100 epochs with a fixed learning rate of 0.005. Parameters are initialized using Xavier initialization and optimized with Adam and a weight decay of 0.005. We use a batch size of 1024 and generate one negative sample per positive interaction.

Overall Performance

To comprehensively evaluate PC-DDI, we conduct experiments on D1, D2 and D3. For D1 and D3 we test under transductive and inductive setting. For D2, we perform experiments under the transductive setting. The results are summarized in Table 1, 2 and 3.

In the transductive experiments, PC-DDI outperforms the SOTA on both the D1, D2 and D3. Specifically, PC-DDI achieves an accuracy of 0.9832 on D1 and 0.9922 on D2. In D3, PC-DDI outperform the second-best model by 1.44%. These results demonstrate that the pharmacophore-based approach can fully exploit all functional patterns of known drugs and enable high-precision inference of the interaction relationships among DDI triples.

Model	ACC	AUC-ROC	AP	F1
DeepDDI	0.8388	0.9263	0.9851	0.8912
SSI-DDI	0.9219	0.9809	0.9897	0.9398
GATDDI	0.8475	0.9224	0.9898	0.8813
GMPNN-CS	0.9262	0.9812	0.9896	0.9357
DSN-DDI	0.8889	0.9669	0.9634	0.8812
SRR-DDI	0.9055	0.9768	0.9855	0.9355
PEB-DDI	0.9846	0.9929	0.9945	0.9882
HDN-DDI	0.9445	0.9742	0.9802	0.9503
Ours	0.9922	0.9966	0.9936	0.9912

Table 2: Performance comparison on D2.

Model	Transductive (Known-Known)				Inductive-S1 (Unknown-Unknown)				Inductive-S2 (Unknown-Known)			
	ACC	AUC-ROC	AP	F1	ACC	AUC-ROC	AP	F1	ACC	AUC-ROC	AP	F1
DeepDDI	0.8720	0.9402	0.9353	0.8720	0.6012	0.6875	0.6766	0.5986	0.6698	0.7346	0.7218	0.6954
SSI-DDI	0.9023	0.9571	0.9459	0.9038	0.6540	0.7343	0.7503	0.5412	0.7638	0.8423	0.8494	0.7354
GATDDI	0.8781	0.9398	0.9266	0.8827	0.6631	0.7275	0.7161	0.6868	0.6983	0.7729	0.7579	0.7301
GMPNN-CS	0.9077	0.9564	0.9413	0.9096	0.6857	0.7496	0.7544	0.6532	0.7772	0.8484	0.8487	0.7892
DSN-DDI	0.9430	0.9841	0.9830	0.9430	0.6683	0.7324	0.7312	0.6314	0.7648	0.8296	0.8192	0.7648
SRR-DDI	0.8906	0.9477	0.9297	0.8936	0.6712	0.7559	0.7368	0.6488	0.7432	0.8159	0.8018	0.7508
PEB-DDI	0.9613	0.9914	0.9906	0.9618	0.6994	0.7790	0.7453	0.6819	0.7850	0.8395	0.8645	0.7996
HDN-DDI	0.9602	0.9924	0.9916	0.9455	0.6773	0.7602	0.7650	0.6139	0.7216	0.8013	0.7933	0.7364
Ours	0.9746	0.9972	0.9974	0.9739	0.7350	0.8412	0.8564	0.7212	0.8123	0.8972	0.8892	0.8455

Table 3: Performance comparison on D3.

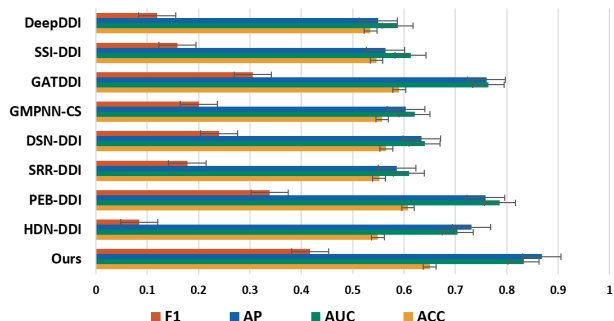


Figure 3: Performance comparison on the D4.

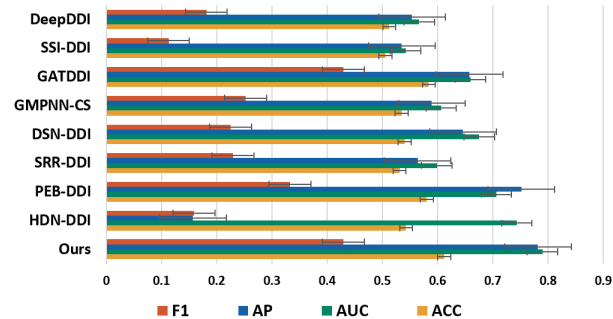


Figure 4: Performance comparison on the D5.

New Version Dataset

Under the inductive settings on datasets D1 and D3, PC-DDI demonstrates strong generalization capability. In the inductive-S1 scenario, PC-DDI achieves an AUC-ROC of 0.8904 and an F1 score of 0.8531 on D1. On D3, it outperforms the second-best method by 3.56% in accuracy and 9.14% in AP. In the inductive-S2 scenario, PC-DDI surpasses the second-best approach by 1.85% in accuracy and 3.20% in F1 on D1, and achieves the best overall performance on D3 with an accuracy of 0.8123 and an F1 score of 0.8455.

This demonstrates the effectiveness of PC-DDI in predicting interactions involving newly introduced drugs as well as entirely novel drug pairs. The performance improvement is attributed to the pharmacophore-based modeling approach, which captures essential interaction patterns between drugs, thereby enhancing both generalization and interpretability. By emphasizing core pharmacological features rather than detailed molecular structures, the model enables efficient knowledge transfer to previously unseen drugs.

Most existing studies are based on the DrugBank 5.0.3 dataset, but the DrugBank database has continued to evolve through regular updates. Newer versions include more recent drug interaction records, corrected annotations, and expanded drug information. To better evaluate the generalization and robustness of our model on updated and more realistic data, we conduct experiments DrugBank 5.0.5 and 5.0.7. The reason for selecting these two versions is that the number of drugs and DDI types they contain is closer to that of version 5.0.3, which facilitates a fair comparison.

As shown in table Figure 3 and 4, PC-DDI consistently outperforms existing DDI prediction methods across all evaluation metrics. On the D4, our model attains the highest accuracy (ACC = 0.6499) among all compared methods, outperforming the second-best model by 4.22%. On D5, PC-DDI also demonstrates superior performance, with an accuracy improvement of 2.78% over the runner-up.

Experimental results show that despite the greater challenges posed by the D4 and D5, PC-DDI still demonstrates superior performance compared to structure-based methods on these updated datasets. These results highlight that the pharmacophore-based approach offers superior robustness and generalization in modeling complex and realistic DDI data.

Ablation Study

As shown in Table 4, we conduct ablation studies to evaluate the impact of key components in PC-DDI. We remove SAG-Pooling (*w/o SAG*), self-attention (*w/o S-att*), and spatial attention weight graph convolution (*w/o SAWconv*) to assess their contributions. For pharmacophore feature aggregation, we compare mean pooling (*w/mean*), max pooling (*w/max*), and attention-based pooling (*w/att*) with default (*w/sum*). Additionally, we test a fully connected approach for pharmacophore graph generation, denoted as *w/full*.

The results in Table 4 show that removing key components of PC-DDI leads to performance drops across all metrics, confirming their importance. Specifically, *w/o SAG*, *w/o S-att*, and *w/o SAWconv* all result in significant performance degradation: *w/o SAG* substantially reduces accuracy and

Model Variant	SAG	Self-attention	SAW	Aggregation	Full-conncet	ACC	AUC-ROC	AP	F1
w/o SAG	✗	✓	✓	Sum	✗	0.9751	0.9911	0.9871	0.9798
w/o S-att	✓	✗	✓	Sum	✗	0.9568	0.9967	0.9968	0.9563
w/o SAW	✓	✓	✗	Sum	✗	0.9710	0.9925	0.9945	0.9713
w/ mean	✓	✓	✓	Mean	✗	0.9814	0.9984	0.9983	0.9826
w/ max	✓	✓	✓	Max	✗	0.9643	0.9935	0.9927	0.9653
w/ att	✓	✓	✓	Attention	✗	0.9762	0.9973	0.9973	0.9719
w/ full	✓	✓	✓	Sum	✓	0.9504	0.9983	0.9883	0.9487
Default	✓	✓	✓	Sum	✗	0.9832	0.9987	0.9988	0.9843

Table 4: Ablation study under D1 transductive setting.

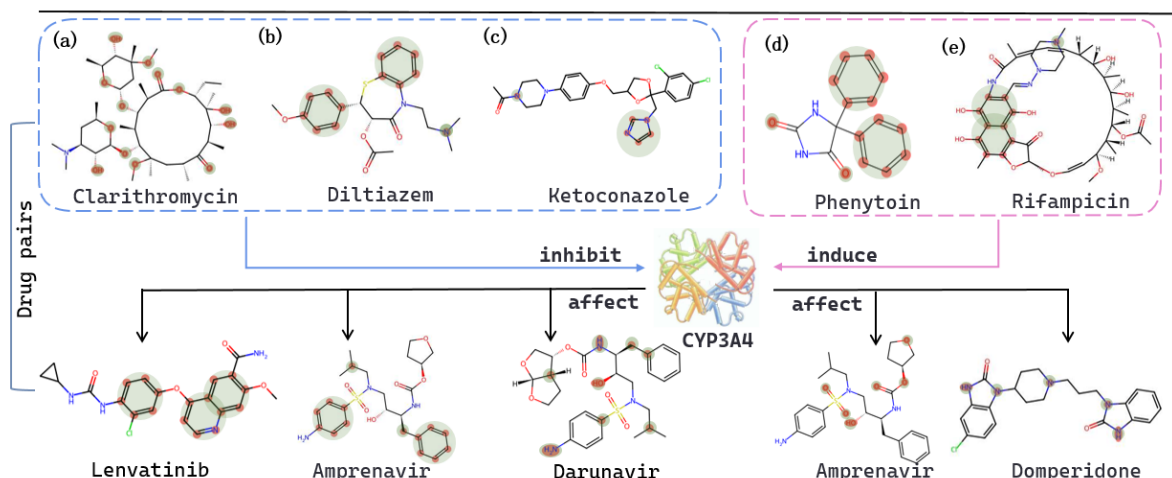


Figure 5: (a)–(e) are five CYP3A4-mediated DDI pairs from DrugBank. These interactions occur when one drug modulates CYP3A4 activity—through inhibition or induction—thereby affecting the metabolism of the co-administered drug. Key pharmacophore atoms are highlighted in red, complete pharmacophores are shaded in green for clarity.

F1-score, demonstrating that SAGPooling is essential for effective graph-level feature extraction; *w/o S-att* highlights the critical role of global attention; and *w/o SAWconv* underscores the necessity of local spatial modeling.

Case Study

Cytochrome P450 3A4 (CYP3A4) is a key enzyme in drug metabolism, involved in processing about half of all clinical drugs. Some drugs can either inhibit or induce CYP3A4 activity, thereby altering the efficacy and safety of other medications. In this case study, we investigate how CYP3A4 inhibitors and inducers work by analyzing key pharmacophores in DDI pairs (Figure 5).

Clarithromycin binds to the active site of CYP3A4 via its macrolide ring, inhibiting the enzyme through hydrogen bonds and non-covalent interactions (Ahlin et al. 2009; Zhou 2008). The aromatic ring of Lenvatinib (Figure 5(a)) serves as the primary binding site for the enzyme. By occupying the same or adjacent sites, Clarithromycin competitively inhibits the metabolism of Lenvatinib, increasing its plasma concentration and potentially enhancing its therapeutic effect or the risk of toxicity.

Diltiazem occupies the CYP3A4 active site via its aromatic ring and hydrogen-bond acceptor. Ketoconazole coordinates with the heme iron through the nitrogen atom

of its imidazole ring, aided by hydrophobic interactions. Phenytoin activates PXR through its aromatic ring and hydrogen-bond donor/acceptor, inducing CYP3A4 expression and accelerating the metabolism of drugs like domperidone. Rifampicin, containing multiple aromatic rings and polar groups, interacts more strongly with PXR, leading to markedly enhanced CYP3A4 activity (Chen and Raymond 2006; Nallani et al. 2003). The pharmacophores shown in Figure 5 align well with the actual biological mechanisms.

Conclusion

In this paper, we propose PC-DDI, a novel DDI prediction method that integrates pharmacophore modeling with causal learning. PC-DDI is the first to explore the biological mechanisms of DDIs from the functional perspective of drug molecules, providing a novel angle for DDI prediction. The key pharmacophores identified through causal learning play a crucial role in explaining these mechanisms. Extensive experiments demonstrate that our method achieves high predictive performance and enhances model generalization, especially in scenarios involving new drugs. Increasing the practical value of DDI prediction algorithms. In future work, we will focus on refining pharmacophore representations and extending the framework to more complex scenarios.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (NSFC 62322215, 62532017, 62402488, 62502336, and U24A20257) and the Shenzhen Science and Technology Program (JCYJ20241202130212016). This study was also supported in part by the High-Performance Computing Center of Central South University and by the Shanxi Provincial Department of Science and Technology Basic Research Project under Grant No. 202403021222056.

References

- Ahlin, G.; Hilgendorf, C.; Karlsson, J.; Szigyarto, C. A.-K.; Uhlen, M.; and Artursson, P. 2009. Endogenous gene and protein expression of drug-transporting proteins in cell lines routinely used in drug discovery programs. *Drug metabolism and disposition*, 37(12): 2275–2283.
- Chen, J.; and Raymond, K. 2006. Roles of rifampicin in drug-drug interactions: underlying molecular mechanisms involving the nuclear pregnane X receptor. *Annals of clinical microbiology and antimicrobials*, 5(1): 3.
- Chen, S.; Semenov, I.; Zhang, F.; Yang, Y.; Geng, J.; Feng, X.; Meng, Q.; and Lei, K. 2024. An effective framework for predicting drug–drug interactions based on molecular substructures and knowledge graph neural network. *Computers in Biology and Medicine*, 169: 107900.
- Chen, Y.; Ma, T.; Yang, X.; Wang, J.; Song, B.; and Zeng, X. 2021. MUFFIN: multi-scale feature fusion for drug–drug interaction prediction. *Bioinformatics*, 37(17): 2651–2658.
- Deng, Y.; Qiu, Y.; Xu, X.; Liu, S.; Zhang, Z.; Zhu, S.; and Zhang, W. 2022. META-DDIE: predicting drug–drug interaction events with few-shot learning. *Briefings in bioinformatics*, 23(1): bbab514.
- Gopnik, A.; and Schulz, L. 2007. *Causal learning: Psychology, philosophy, and computation*. Oxford University Press.
- Gottlieb, A.; Stein, G. Y.; Oron, Y.; Ruppín, E.; and Sharan, R. 2012. INDI: a computational framework for inferring drug interactions and their associated recommendations. *Molecular Systems Biology*, 8(1): 592.
- Granger, C. W. J. 1969. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, 424–438.
- Hao, X.; Chen, Q.; Pan, H.; Qiu, J.; Zhang, Y.; Yu, Q.; Han, Z.; and Du, X. 2023. Enhancing drug–drug interaction prediction by three-way decision and knowledge graph embedding. *Granular Computing*, 8(1): 67–76.
- Hong, Y.; Luo, P.; Jin, S.; and Liu, X. 2022. LaGAT: link-aware graph attention network for drug–drug interaction prediction. *Bioinformatics*, 38(24): 5406–5412.
- Hu, W.; Wu, J.; and Qian, Q. 2025. CiRLExpainer: Causality-Inspired Expainer for Graph Neural Networks via Reinforcement Learning. *IEEE Transactions on Neural Networks and Learning Systems*, 36(6): 9970–9984.
- Huang, K.; Xiao, C.; Hoang, T.; Glass, L.; and Sun, J. 2020. Caster: Predicting drug interactions with chemical substructure representation. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, 702–709.
- Jain, S.; and Wallace, B. C. 2019. Attention is not Explanation. *arXiv preprint arXiv:1902.10186*.
- Li, Z.; Zhu, S.; Shao, B.; Zeng, X.; Wang, T.; and Liu, T.-Y. 2023. DSN-DDI: an accurate and generalized framework for drug–drug interaction prediction by dual-view representation learning. *Briefings in Bioinformatics*, 24(1): bbac597.
- Lin, J.; Wu, L.; Zhu, J.; Liang, X.; Xia, Y.; Xie, S.; Qin, T.; and Liu, T.-Y. 2023. R2-DDI: relation-aware feature refinement for drug–drug interaction prediction. *Briefings in bioinformatics*, 24(1).
- Lin, W.; Lan, H.; and Li, B. 2021. Generative causal explanations for graph neural networks. In *International Conference on Machine Learning*, 6666–6679. PMLR.
- Luo, D.; Cheng, W.; Xu, D.; Yu, W.; Zong, B.; Chen, H.; and Zhang, X. 2020. Parameterized explainer for graph neural network. *Advances in neural information processing systems*, 33: 19620–19631.
- Marinka Zitnik, S.; Sosič, R.; and Leskovec, J. 2018. BioSNAP datasets: Stanford biomedical network dataset collection. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 8(1): 1.
- Nallani, S. C.; Glauser, T. A.; Hariparsad, N.; Setchell, K.; Buckley, D. J.; Buckley, A. R.; and Desai, P. B. 2003. Dose-dependent induction of cytochrome P450 (CYP) 3A4 and activation of pregnane X receptor by topiramate. *Epilepsia*, 44(12): 1521–1528.
- Niu, D.; Xu, L.; Pan, S.; Xia, L.; and Li, Z. 2024. SRR-DDI: a drug–drug interaction prediction model with substructure refined representation learning based on self-attention mechanism. *Knowledge-Based Systems*, 285: 111337.
- Nyamabo, A. K.; Yu, H.; Liu, Z.; and Shi, J. Y. 2022. Drug-drug interaction prediction with learnable size-adaptive molecular substructures. *Briefings in Bioinformatics*, 23(1): bbab441.
- Nyamabo, A. K.; Yu, H.; and Shi, J.-Y. 2021. SSI-DDI: substructure–substructure interactions for drug–drug interaction prediction. *Briefings in Bioinformatics*, 22(6): bbab133.
- Ren, Z.; Zeng, X.; Lao, Y.; You, Z.; Shang, Y.; Zou, Q.; and Lin, C. 2025. Predicting rare drug-drug interaction events with dual-granular structure-adaptive and pair variational representation. *Nature Communications*, 16(1): 3997.
- Ryu, J. Y.; Kim, H. U.; and Lee, S. Y. 2018. Deep learning improves prediction of drug–drug and drug–food interactions. *Proceedings of the national academy of sciences*, 115(18): E4304–E4311.
- Shen, X.; Li, Z.; Liu, Y.; Song, B.; and Zeng, X. 2023. PEB-DDI: a task-specific dual-view substructural learning framework for drug–drug interaction prediction. *IEEE Journal of Biomedical and Health Informatics*, 28(1): 569–579.
- Sun, J.; and Zheng, H. 2025. HDN-DDI: a novel framework for predicting drug-drug interactions using hierarchical molecular graphs and enhanced dual-view representation learning. *BMC bioinformatics*, 26(1): 28.
- Takeda, T.; Hao, M.; Cheng, T.; Bryant, S. H.; and Wang, Y. 2017. Predicting drug–drug interactions through drug

structural similarities and interaction networks incorporating pharmacokinetics and pharmacodynamics knowledge. *Journal of cheminformatics*, 9: 1–9.

Velickovic, P.; Cucurull, G.; Casanova, A.; Romero, A.; Lio, P.; and Bengio, Y. 2017. Graph attention networks. *arXiv preprint arXiv:1710.10903*.

Wang, X.; Wu, Y.; Zhang, A.; Feng, F.; He, X.; and Chua, T.-S. 2022. Reinforced causal explainer for graph neural networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(2): 2297–2309.

Wang, Y.; Min, Y.; Chen, X.; and Wu, J. 2021. Multi-view graph contrastive representation learning for drug-drug interaction prediction. In *Proceedings of the web conference 2021*, 2921–2933.

Yan, X.-Y.; Yin, P.-W.; Wu, X.-M.; and Han, J.-X. 2021. Prediction of the drug–drug interaction types with the unified embedding features from drug similarity networks. *frontiers in Pharmacology*, 12: 794205.

Zhang, R.; Wang, X.; Liu, G.; Wang, P.; Zhou, Y.; and Wang, P. 2025. Motif-Oriented Representation Learning with Topology Refinement for Drug-Drug Interaction Prediction. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, 1102–1110.

Zhou, S.-F. 2008. Drugs behave as substrates, inhibitors and inducers of human cytochrome P450 3A4. *Current Drug Metabolism*, 9(4): 310–322.