

Received 19 June 2025, accepted 18 July 2025, date of publication 23 July 2025, date of current version 29 July 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3591823

RESEARCH ARTICLE

Pharmacophore-Aware Dual-View Learning With Bidirectional Cross-Attention for Drug-Drug Interaction Prediction

WENXIAO ZHANG¹, SEONG YOON SHIN¹, AND HAILIANG TANG²

¹School of Computer Science and Engineering, Kunsan National University, Gunsan 54150, Republic of Korea

²School of Software, Kunsan National University, Gunsan 54150, Republic of Korea

Corresponding authors: Wenxiao Zhang (kk19856482@163.com) and Seong Yoon Shin (1450718153@qq.com)

ABSTRACT Accurate drug-drug interaction (DDI) prediction is critical for ensuring patient safety and guiding clinical decision-making. Existing methods often rely on single-view molecular representations, limiting their ability to capture the complex structural and spatial properties of drugs. In this study, we propose a novel pharmacophore-aware dual-view learning framework (PharmaDual) that integrates both 2D and 3D representations of pharmacophores for enhanced DDI prediction. Specifically, we first extract pharmacophore fragments as key substructures, and independently encode their 2D and 3D spatial information using specialized graph-and geometry-based encoders. To effectively combine the complementary views, we introduce a bidirectional cross-attention fusion module that dynamically aligns and integrates 2D and 3D pharmacophore representations. Extensive experiments on benchmark DDI datasets demonstrate that our method consistently outperforms existing approaches, highlighting the benefit of dual-view modeling and cross-attentive fusion in capturing nuanced pharmacophore-level interactions. The code is available at <https://github.com/ZWX1289/PharmaDual>

INDEX TERMS Drug-drug interactions, graph neural network, Transformer, bidirectional cross-attention, learning latent representations.

I. INTRODUCTION

Drug-drug interactions (DDIs) refer to the pharmacological or clinical consequences that may occur when multiple drugs are co-administered. Identifying and characterizing DDIs is crucial in both drug development and routine medical care, as it plays a central role in optimizing treatment regimens and ensuring patient safety [1]. In the treatment of multifactorial diseases, polypharmacy is often unavoidable. However, combining several medications may not only influence drug efficacy but also increase the likelihood of adverse reactions. Although multi-drug therapies can harness synergistic effects to improve clinical outcomes, they simultaneously introduce a higher risk of harmful interactions, some of which can be severe or even fatal [2], [3], [4]. Consequently, precise DDI prediction has become an essential requirement in clinical workflows [5], [6].

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Wang¹.

Reliable identification of DDIs contributes to the reduction of drug-induced complications [7], ultimately decreasing hospitalization rates, healthcare costs, and the probability of therapeutic failure. It also enables healthcare professionals to avoid unsafe drug combinations and supports more informed prescribing practices. With the rapid advancement of Artificial Intelligence (AI) and Deep Learning (DL), these technologies have shown remarkable success in solving challenging tasks in bioinformatics [8]. Leveraging these computational approaches allows for efficient extraction and modeling of complex biomedical knowledge, making them highly effective for DDI prediction [8], [9]. Unlike conventional approaches that often suffer from limited biochemical insight and poor scalability, deep learning-based models have demonstrated superior capability in capturing intricate drug relationships, thereby drawing increasing interest in the field.

We have witnessed the emergence of diverse computational models for predicting drug-drug interactions (DDIs).

These models aim to mitigate the costs associated with pharmaceutical research by providing feasible outcomes for biological experiments. In recent years, an increasing number of studies have represented drug molecules as graphs to extract informative structural features [10], [11]. For example, Deac et al. [12] introduced a GNN-based framework that utilizes molecular structural data for DDI prediction. Likewise, Wang et al. [13] integrated both pharmaceutical and genomic data with GCN and attention mechanisms to identify promising drug combinations. In another study, Zhang et al. [14] explored the incorporation of node centrality, spatial encoding, and edge descriptors alongside a lightweight attention mechanism to capture important molecular structural features. Additionally, drugs can be decomposed into bioactive substructures, such as functional groups or atom clusters, which are critical for DDI modeling. Several studies have focused on learning interactions at the substructure level to enhance prediction performance [15], [16], [17], [18]. For instance, Nyamabo et al. [16] used GAT to process molecular graphs of drug pairs, extracting substructure representations from the receptive fields of each GAT layer. Similarly, Yu et al. [19] proposed substructure-aware embeddings based on predefined functional groups and developed a tensor neural network specifically designed for DDI prediction. Furthermore, Li et al. [20] introduced the pharmacokinetics (PK) model as one of the pioneering computational approaches for predicting DDIs. More recently, Vilar et al. [21], [22], [23] utilized drug similarity measures derived from both 2D and 3D molecular structures, interaction profiles, and their combinations to predict potential drug interactions.

Molecular structure-driven approaches have gained considerable traction for predicting drug-drug interactions (DDIs), leveraging concepts rooted in medicinal chemistry [24]. These methods examine the functional groups and chemical substructures that constitute drugs to explore their pharmacokinetic and pharmacodynamic behaviors and to identify potential DDIs. Unlike knowledge-based models, structure-based strategies [25], [26], [27], [28], [29], [30] represent drugs as standalone entities and perform DDI prediction directly based on drug pairs, without incorporating external biomedical knowledge. A central design of these models is the focus on localized chemical components rather than entire molecular graphs, since DDIs are typically driven by interactions among specific substructures [26], [31]. The underlying assumption is that representations learned from these substructures are generalizable across drugs sharing similar chemical motifs [25], [30]. For instance, MR-GNN [29] applies graph neural networks to derive multi-scale features from molecular graphs. CASTER [25] employs sequential pattern mining to extract frequent substructures, followed by an auto-encoding and dictionary learning process to enhance generalization and interpretability. Models like SSI-DDI [28], MHCADDI [27], and CMPNN-CS [30] introduce co-attention mechanisms that align substructures across drug pairs for interactive learning. Specifically, CMPNN-CS

treats chemical bonds as gates that regulate message passing in the GNN, enabling substructure encoding in a self-supervised manner.

Despite significant progress in DDI prediction, several key challenges remain: (1) Many existing approaches neglect pharmacophore-level features—critical chemical substructures that directly influence drug interactions—resulting in a loss of essential chemical insights needed for precise and interpretable predictions. (2) Conventional methods treat molecular representations independently, failing to exploit complementary information between views, thus limiting the effectiveness of multi-view learning. (3) Current fusion techniques, such as simple concatenation or basic attention mechanisms, often overlook the complex relationships between different views, hindering the model's ability to capture fine-grained inter-view dependencies.

To address the aforementioned challenges, we propose PharmaDual, a novel DDI prediction framework that integrates dual-view learning with bidirectional cross-attention fusion to effectively capture key pharmacophore features and enhance molecular representation learning. The main contributions of this work are summarized as follows:

- We propose PharmaDual, a pharmacophore-aware encoding strategy that enhances molecular representations by capturing key chemical substructures, providing crucial insights into drug activity and interactions.
- We introduce a novel framework that integrates dual-view learning to effectively capture both 2D and 3D molecular representations, allowing for a more comprehensive understanding of drug structures.
- We develop a bidirectional cross-attention mechanism to fuse the 2D and 3D molecular representations, enabling the model to learn complex dependencies between structural and spatial features, improving the accuracy of DDI predictions.
- We conduct extensive experiments on benchmark datasets to evaluate the performance of PharmaDual, and the results demonstrate that PharmaDual outperforms state-of-the-art methods in DDI prediction, achieving superior accuracy and robustness.

II. RELATED WORK

This section offers an in-depth review of prior studies on DDI prediction, focusing particularly on two key areas: (1) drug representation, and (2) Multi-View based DDI Prediction.

A. DRUG REPRESENTATION

The majority of existing methods for DDI prediction employ molecular fingerprints [32], [33], [34] or other drug profiles, including side effects [32], [33], binding targets [34], transporters, enzymes, pathways, and combinations thereof [33], [35], [36] to predict potential DDIs. Molecular fingerprints [37], [38] are binary vectors that indicate the presence (i.e., 1) or absence (i.e., 0) of specific chemical substructures. Similarly, other profiles are also represented as binary vectors, denoting the presence or absence of particular

characteristics such as side effects or binding targets. Certain approaches [21], [39], [40], [41], [42], [43] further preprocess the drug representation by using similarity vectors, which quantify the similarity between a drug and others within the aforementioned representation spaces using measures like cosine similarity or Jaccard similarity. This assumption is predicated on the premise that drugs with similar or dissimilar profiles demonstrate corresponding biological activities [21]. However, these representations exhibit inherent limitations due to these approaches being manually crafted and constrained by expert knowledge, which restricts their ability to uncover newly emerged DDI information, especially when dealing with unknown drugs. Moreover, the availability of some features may be limited during the early stages of drug development, thereby impeding the applicability of methods reliant on such features.

B. MULTI-VIEW BASED DDI PREDICTION

Drugs are represented in multiple forms, such as molecular graphs, SMILES strings, and 3D structures, and leveraging these diverse features simultaneously can improve DDI prediction. He et al. [44] introduced MFFGNN, which integrates topological information from molecular graphs and SMILES through feature extraction modules (MGFEM and SSFEM), followed by aggregation and fusion to enhance drug representations. Similarly, Pang et al. [45] proposed AMDE, which encodes drug features in multiple dimensions. Their method uses two channels to process drug SMILES sequences, extracting two-dimensional atom map features and one-dimensional sequence features using Rdkit and FCS, respectively. These features are then sent to a 2D feature graph encoder and a 1D feature sequence encoder for further encoding. On the other hand, He et al. [46] proposed 3DGT-DDI, which combines 3D structural features with textual information using a 3D GNN and textual attention mechanism. SCIBERT is employed for extracting text features, while SchNet [47] captures 3D geometric data, enhancing prediction accuracy and model interpretability.

III. PROPOSED PHARMADUAL FRAMEWORK

As depicted in Figure 1, we introduce PharmaDual, a cross-modal contrastive learning framework designed for DDI prediction, consisting of three core modules: (a) 2D view module, (b) 3D view module, (c) dual-view fusion module, and (d) DDI prediction module. In the 2D view module, drug molecules are represented by pharmacophores and encoded using a 2D neural network, followed by a readout operation to extract structural representations. In the 3D view module, the coordinates of pharmacophores are converted to 3D coordinates, and the spatial relationships within the pharmacophores are captured using a Transformer-based encoder to encode 3D molecular information. A bidirectional cross-attention mechanism is then applied to fuse the 2D and 3D representations, enabling the model to accurately predict potential DDIs by integrating complementary structural

and spatial information. In the DDI prediction module, the fused molecular representation is utilized to infer potential drug-drug interactions, enabling accurate and robust prediction.

A. 2D VIEW MODULE

The molecular graph is denoted $G = \{V, E\}$, V refers to the set of nodes, and E denotes the set of edges. We use the BRICS algorithm [48] implemented in RDKit to decompose each molecule into a set of chemically meaningful fragments, which serve as pharmacophores. The resulting pharmacophore set is denoted as $G = \{(V^1, E^1), (V^2, E^2), (V^3, E^3), \dots, (V^N, E^N)\}$, where N denotes the total number of pharmacophores.

To ensure reproducibility and consistency, we apply the following preprocessing steps after BRICS decomposition: (1) **Minimum size filtering**: fragments containing fewer than four heavy atoms are discarded to avoid trivial or uninformative structures; (2) **Removal of generic linker atoms**: generic BRICS linkers (dummy atoms, denoted as $[*]$) are removed, and any resulting disconnected or invalid fragments are excluded; (3) **Canonicalization**: each remaining fragment is canonicalized into a SMILES representation, and duplicate fragments are removed to ensure uniqueness and chemical diversity. After filtering, these retained fragments are treated as subgraphs for downstream 2D encoding.

Next, we employ a graph neural network (GNN) to encode the initial features of the first pharmacophore, aiming to obtain higher-level feature representations. In GNN, the node aggregation operation typically involves a summation of neighbor node features combined with the node's own features. Through multiple rounds of node aggregation operations, GNN gradually updates the node representations, capturing complex relationships and patterns in the graph data. Furthermore, we utilize mean pooling for the readout operation to obtain the representation of the entire pharmacophore.

Taking the first pharmacophore (V^1, E^1) as an example. The message-passing process in the k -th layer of GNNs is shown as follows:

$$a_i^{(k)} = \text{AGGREGATE}^{(k)}(\{h_u^{(k-1)} : u \in N(i)\}) \quad (1)$$

$$h_i^{(k)} = \text{COMBINE}^{(k)}(h_i^{(k-1)}, a_i^{(k)}) \quad (2)$$

where $a_i^{(k)}$ represents the feature formed by aggregating the features of neighboring nodes for node i within pharmacophore 1, $N(v)$ is the set of neighboring nodes for node i , and $h_i^{(k)}$ is the feature of node i at the k -th layer. In order to further extract the pharmacophore-level feature $h_{P(1)}$, the readout operation integrates all node features in pharmacophore 1, as shown below:

$$h_{P(1)} = \text{READOUT}(\{h_i^{(k)} : i \in P(1)\}) \quad (3)$$

By repeating the same process, we obtain the features of all pharmacophores in the molecule: $h_{P(2)}, h_{P(3)}, \dots, h_{P(N)}$.

Next, we perform a readout operation on the features of all pharmacophores contained in the drug molecule to obtain

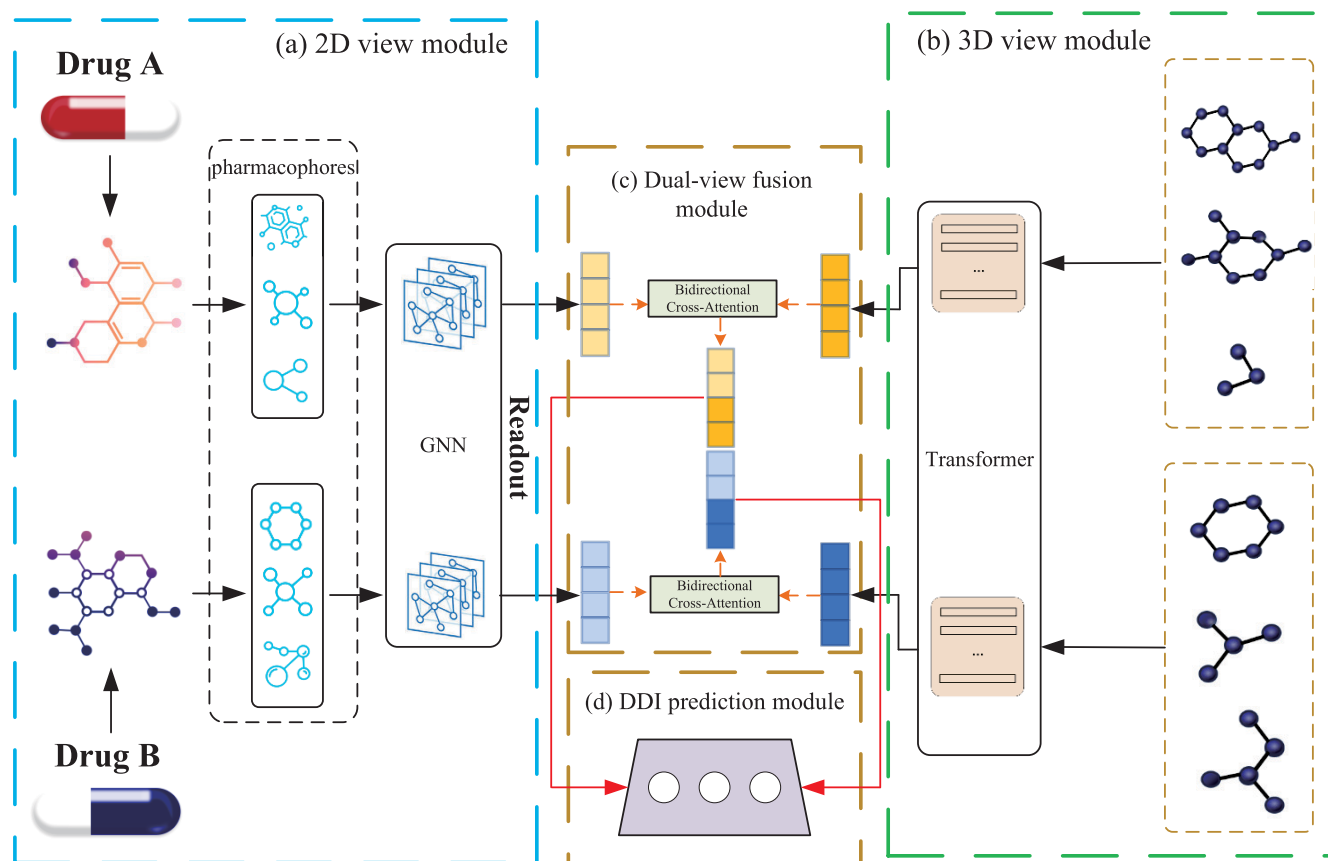


FIGURE 1. The overall architecture of PharmaDual. The framework comprises four modules: (a) the 2D view module encodes pharmacophore-based substructures using a 2D neural network to extract structural features; (b) the 3D view module captures spatial relationships among pharmacophores through a Transformer-based encoder using 3D coordinates; (c) the dual-view fusion module employs a bidirectional cross-attention mechanism to integrate 2D and 3D representations; (d) the DDI prediction module utilizes the fused representations to predict potential drug-drug interactions.

its 2D molecular feature. The detailed process is described as follows.

$$z_{2D} = \text{Readout}(h_{p(1)}, h_{p(2)}, h_{p(3)}, \dots, h_{p(N)}) \quad (4)$$

B. 3D VIEW MODULE

We utilize RDKit to generate 3D conformations by converting the 2D atomic coordinates within each pharmacophore into 3D spatial structures. Subsequently, a Transformer-based encoder is employed to model the spatial dependencies among atoms, enabling the extraction of rich pharmacophore-level representations from the 3D molecular view.

To ensure structural consistency and efficiency, we retain only the lowest-energy conformer for each drug. This conformer is selected from those generated by RDKit's ETKDG algorithm, which balances physical plausibility with diversity. The lowest-energy conformer is generally considered the thermodynamically most stable and biologically relevant structure under ambient conditions [49], and has been widely used in prior structure-based molecular learning tasks [50], [51]. This strategy not only ensures computational efficiency but also provides a stable geometric basis for downstream representation learning.

Taking the first pharmacophore (V^1, E^1) as an example, $v_{(i)}^1 \in V^1$ represents the i -th atom ($1 \leq i \leq |V^1|$) and $e_{(j)}^1 \in E^1$ denotes the j -th bond ($1 \leq j \leq |E^1|$). The node feature matrix of the first pharmacophore is denoted as $X_V^1 \in \mathbb{R}^{|V^1| \times D_V}$, and the edge feature matrix is denoted as $X_E^1 \in \mathbb{R}^{|E^1| \times D_E}$, where D_V and D_E are the dimensions of node and edge features, respectively.

We first project the initial node and edge features into a unified dimensional space, resulting in updated features $X_{V'}^1$ and $X_{E'}^1$. These updated features are then concatenated to form the initial pharmacophore representation, as illustrated below:

$$X_1' = \text{CONCAT}(X_{V'}^1, X_{E'}^1) \quad (5)$$

Next, we employ a Transformer to encode the initial features of the first functional group, aiming to obtain higher-level feature representations. The Transformer architecture is shown in Figure 2.

$$Q_i = X_1' W_Q^{(i)} \quad (6)$$

$$K_i = X_1' W_K^{(i)} \quad (7)$$

$$V_i = X_1' W_V^{(i)} \quad (8)$$

$$Attention_i = \text{softmax}(\frac{Q_i K_i^T}{\sqrt{d_k}} + q_{ij}) V_i \quad (9)$$

$$X_1'' = \text{MultiHead}(Q, K, V) \\ = \text{Concat}(Attention_1, \dots, Attention_h) W^O \quad (10)$$

where $W_Q^{(i)}$, $W_K^{(i)}$, $W_V^{(i)}$, W^O represent learnable matrices. q_{ij} represents the 3D distance between atom i and atom j .

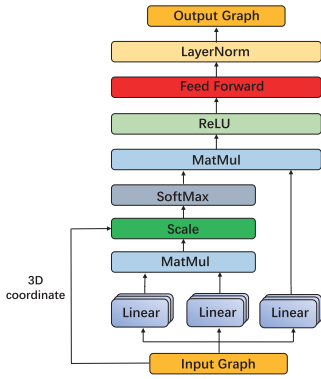


FIGURE 2. The architecture of Transformer.

By repeating the same process, we obtain the features of all pharmacophores in the molecule: $X_2'', X_3'', \dots, X_N''$.

Next, we perform a readout operation on the features of all pharmacophores contained in the drug molecule to obtain its 3D molecular feature. The detailed process is described as follows.

$$z_{3D} = \text{Readout}(X_1'', X_2'', X_3'', \dots, X_N'') \quad (11)$$

C. DUAL-VIEW FUSION MODULE

To fully exploit the complementary information between the 2D and 3D drug molecular representations, we design a bidirectional cross-attention mechanism, where both views alternately serve as queries to attend to each other. The attention computations are defined as follows:

1) 2D-TO-3D CROSS-ATTENTION

$$Q_{2D}^i = z_{2D} W_{Q_{2D}}^i \quad (12)$$

$$K_{3D}^i = z_{3D} W_{K_{3D}}^i \quad (13)$$

$$V_{3D}^i = z_{3D} W_{V_{3D}}^i \quad (14)$$

$$Attention_i = \text{softmax} \frac{Q_{2D}^i (K_{3D}^i)^T}{\sqrt{d}} V_{3D}^i \quad (15)$$

$$H_{2D} = \text{MultiHead}(Q_{2D}, K_{3D}, V_{3D}) \\ = \text{Concat}(Attention_i, \dots, Attention_N) W^O \quad (16)$$

where H_{2D} captures 3D-aware 2D features. All projection matrices $W_{Q_{2D}}$, $W_{K_{3D}}$, and $W_{V_{3D}}$ are of shape $\mathbb{R}^{d \times d}$, where d is the hidden feature dimension (set to 64 in our experiments). The number of attention heads N is set to 4.

2) 3D-TO-2D CROSS-ATTENTION

$$Q_{3D}^i = z_{3D} W_{Q_{3D}}^i \quad (17)$$

$$K_{2D}^i = z_{2D} W_{K_{2D}}^i \quad (18)$$

$$V_{2D}^i = z_{2D} W_{V_{2D}}^i \quad (19)$$

$$Attention_i = \text{softmax} \frac{Q_{3D}^i (K_{2D}^i)^T}{\sqrt{d}} V_{2D}^i \quad (20)$$

$$H_{3D} = \text{MultiHead}(Q_{3D}, K_{2D}, V_{2D}) \\ = \text{Concat}(Attention_i, \dots, Attention_N) W^O \quad (21)$$

where H_{3D} denotes 2D-enhanced 3D features. All projection matrices $W_{Q_{3D}}$, $W_{K_{2D}}$, and $W_{V_{2D}}$ are of shape $\mathbb{R}^{d \times d}$, where d is the hidden feature dimension (set to 64 in our experiments). The number of attention heads N is set to 4.

Finally, the fused representation H_{FUSED} is constructed by aggregating H_{2D} and H_{3D} , such as by concatenation and subsequent transformation:

$$H_{FUSED} = \text{FFN}([H_{2D} || H_{3D}]) \quad (22)$$

where FFN denotes a feed-forward network, and $||$ indicates concatenation. This fused representation is then used for DDI prediction.

D. DDI PREDICTION MODULE

In the DDI prediction module, we obtain the fused representations H_{FUSED}^A and H_{FUSED}^B for the drug pair (A, B) using the bidirectional cross-attention mechanism over their 2D and 3D features. These representations are then concatenated and fed into a multilayer perceptron (MLP) to compute the interaction score:

$$H = \text{MLP}([H_{FUSED}^A || H_{FUSED}^B]) \quad (23)$$

$$\hat{y}_{A,B} = \sigma(w^T H + b) \quad (24)$$

where $||$ denotes vector concatenation, $\sigma(\cdot)$ is the sigmoid function, and w, b are learnable parameters. The model is trained using binary cross-entropy loss, which minimizes the discrepancy between the predicted interaction score $\hat{y}_{A,B}$ and the true label $y_{A,B}$:

$$L_{CE} = - \sum_{i=1}^N [y_{A,B}^{(i)} \log(\hat{y}_{A,B}^{(i)}) + (1 - y_{A,B}^{(i)}) \log(1 - \hat{y}_{A,B}^{(i)})] \quad (25)$$

where N is the number of training samples, and $y_{A,B}^{(i)}$ is the ground truth label for the i -th drug pair, indicating the presence or absence of a drug-drug interaction. This training strategy enables the model to optimize its parameters to predict potential DDIs accurately.

IV. RESULTS

A. DATASET

We evaluated the effectiveness of PharmaDual on two widely used real-world datasets: DrugBank and TWOSIDES. The DrugBank database integrates resources from bioinformatics, chemoinformatics, and other related domains to offer extensive drug-related information [52]. It includes 86 interaction types that capture various metabolic influences between drugs, comprising 1,706 drugs and 191,808 DDI triplets.

TABLE 1. Hyperparameter configurations of model experiments.

HyperParameters	Value
Epoch	150
Learning rate	0.0001
Batch size	64
Weight decay	0.00005
2D embedding dimension	64
3D embedding dimension	64
Depth	3
Num_heads	4

In contrast, the TWOSIDES dataset [53] contains 645 drugs, 963 categories of interactions, and 4,576,287 DDI triplets, obtained through systematic filtering and preprocessing of the original TWOSIDES data. Unlike DrugBank, TWOSIDES focuses on phenotypic-level interactions.

B. EXPERIMENTAL SETTINGS

To assess the computational cost of our model, we report the average training time and memory consumption on the DrugBank dataset. Using an NVIDIA RTX 3090 GPU, the average training time is 42 seconds per epoch, and the peak memory usage is 4.8 GB. Despite incorporating dual-view encoding and a cross-attention fusion module, the architecture remains efficient for datasets of DrugBank scale. For larger compound libraries, the model structure supports further parallelization or view simplification to maintain scalability.

To rigorously evaluate the performance of the DDI prediction model, we employ a 5-fold cross-validation strategy. The DDI prediction task is framed as a binary classification problem, where each instance comprises a pair of drugs annotated as either interacting or non-interacting. In the training phase, positive instances are assigned a label of “1,” while negative instances are labeled as “0.” Model training is conducted in accordance with the hyperparameter configurations detailed in Table 1. To select the hyperparameters listed in Table 1, we performed a grid search on the training set and evaluated performance on a held-out validation set. The search ranges were as follows:

- **Epoch:** {100, 150, 200}
- **Learning rate:** {1e-4, 5e-4, 1e-3}
- **Batch size:** {32, 64, 128}
- **Weight decay:** {0, 1e-5, 5e-5}
- **2D embedding dimension:** {32, 64, 128}
- **3D embedding dimension:** {32, 64, 128}
- **Depth:** {2, 3, 4}
- **Num_heads:** {3, 4, 5}

C. EVALUATION METRICS

In this section, we utilize three primary evaluation metrics—AUROC, AUPRC, and F1 score—to assess the performance of PharmaDual. The confusion matrix presented in Table 2 serves as the foundation for computing these metrics.

- (1) Recall reflects the proportion of true positive instances correctly identified by a classification model.

TABLE 2. Confusion matrix for prediction results.

	Actual Positive (P)	Actual Negative (N)
Predicted Positive (P)	True Positive (TP)	False Negative (FN)
Predicted Negative (N)	False Positive (FP)	True Negative (TN)

This metric becomes particularly crucial when the cost associated with false negatives (missed positive cases) is significant, as it aims to minimize the occurrence of such errors.

$$Recall = \frac{TP}{TP + FN} \quad (26)$$

- (2) Accuracy is the proportion of correctly classified instances, including both true positives and true negatives, relative to the total number of instances in the dataset. This metric is particularly informative when the dataset is balanced, with an approximately equal distribution of positive and negative cases.

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \quad (27)$$

- (3) Precision quantifies the proportion of true positive instances among all instances predicted as positive by a classification model. This metric is particularly significant when the cost of false positives (incorrectly identified positive cases) is high, as it seeks to minimize such errors.

$$Precision = \frac{TP}{TP + FP} \quad (28)$$

- (4) The ROC curve is constructed on a coordinate system defined by the false positive rate (FPR) and the true positive rate (TPR). The area under the curve, referred to as AUROC, serves as a key metric for evaluating the model’s performance. A higher AUROC value indicates superior classification performance. The definitions of TPR and FPR are provided below.

$$TPR = \frac{TP}{TP + FN} \quad (29)$$

$$FPR = \frac{FP}{FP + TN} \quad (30)$$

- (5) The Precision-Recall Curve (PRC) is generated by plotting the recall rate against the precision rate on a coordinate plane. The area under the PRC curve (AUPRC) serves as a quantitative measure of the model’s performance and is commonly used to evaluate the effectiveness of the classifier.
- (6) F1 score is a metric that takes into account both Precision and Recall simultaneously. Its definition can be expressed as Equation 31.
- (7) The t-test is a statistical method used to determine whether there is a significant difference between the means of two related or independent groups.

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (31)$$

D. BASELINES

To comprehensively assess the performance of PharmaDual, we compared it with several state-of-the-art baselines, including both substructure-driven approaches and dual-view representation learning frameworks.

- **MHCADDI** [12]: leverages a co-attention mechanism to integrate the joint representations of drug pairs, thereby improving individual drug feature learning.
- **SSI-DDI** [16]: applies a multi-layer Graph Attention Network (GAT) to extract substructural features and estimate interaction likelihoods between substructures for DDI prediction.
- **MR-GNN** [54]: constructs a graph neural network with a multi-resolution design and integrates a dual-graph state long short-term memory (LSTM) module to model interactions among biomedical entities.
- **GMPNN-CS** [55]: explores substructures at different granularities and explicitly models their interactions to predict drug-drug interactions.
- **GAT-DDI** [56]: utilizes a graph attention network to learn complex relational patterns within drug graphs for DDI prediction.
- **DGNN-DDI** [57]: incorporates a substructure-aware attention mechanism within a graph neural network framework to enhance DDI prediction accuracy.
- **MM-GANN-DDI** [58]: is a GNN framework that combines molecular graphs, SMILES sequences, and pharmacological features via graph attention, using graph-agnostic meta-training (GAMT) for novel drug DDI type prediction.
- **KITE-DDI** [59]: integrates molecular SMILES and biomedical knowledge graphs via a Transformer architecture for end-to-end DDI prediction, delivering superior accuracy.
- **Taco-DDI** [60]: employs a graph transformer with dynamic co-attention to learn molecular representations for DDI risk prediction, achieving higher accuracy and offering interpretable insights.

E. PERFORMANCE EVALUATION

In the warm-start setting, the training and testing datasets include overlapping drugs. Each experiment is conducted five times, and the average results are reported to ensure robustness. For each run, the dataset is randomly divided into training, validation, and testing subsets using stratified sampling to preserve the distribution of interaction types. To facilitate fair model comparisons, all data partitioning is carried out prior to training, ensuring that each model is evaluated on the same dataset splits. The average performance across five repetitions is summarized in Table 3. As shown, PharmaDual consistently surpasses all baseline approaches on both DrugBank and TWOSIDES datasets across all evaluation metrics. While previous state-of-the-art methods achieved strong ACC scores of 96.33% and 86.96% on DrugBank and TWOSIDES, respectively, PharmaDual further advances the performance, achieving 98.39%

on DrugBank and 91.38% on TWOSIDES. Furthermore, PharmaDual obtains excellent AUPRC values of 99.62% on DrugBank and 92.19% on TWOSIDES, underscoring its effectiveness in identifying positive interaction instances. These results validate the superior predictive capability of PharmaDual in DDI tasks involving known drugs.

In the cold-start setting, the training and testing sets are mutually exclusive with respect to drug identities, ensuring that no drugs overlap between them. This configuration is designed to assess the model's capability to predict DDIs involving entirely novel drugs. Since no structural or contextual information about the test drugs is available during training, this task poses a greater challenge and requires enhanced generalization ability [13], [52]. Formally, let G denote the complete set of drugs, with G_{old} and G_{new} representing the subsets used for training and testing, respectively. These satisfy $G = G_{old} \cup G_{new}$ and $G_{old} \cap G_{new} = \emptyset$. In each run, we randomly select 20% of the unique drugs as *unseen drugs* (G_{new}) to form the test set, and use the remaining 80% (G_{old}) for training. All DDIs in the training set are restricted to those where both drugs belong to G_{old} , while test DDIs include only those involving at least one drug from G_{new} . This guarantees that the model has never encountered the test drugs during training, thereby eliminating any structural or label-level leakage. Moreover, negative samples are generated independently within each subset based only on drugs available in that subset, preserving the cold-start constraint.

To assess whether drug-level splitting causes the disappearance of rare interaction types, we analyzed the distribution of the least frequent DDI labels across all five folds. Table 4 shows the train/test counts of the bottom-5 interaction types. Although these types are rare, we observe that none of them disappear entirely from the test sets, confirming that class diversity is preserved even under cold-start constraints.

Table 6 summarizes the average performance of all compared models across three independent runs. Under the cold-start scenario, all methods experience a noticeable drop in performance; nevertheless, PharmaDual consistently outperforms the baselines. In particular, it achieves AUROC gains of 3.01% on DrugBank and 4.45% on TWOSIDES over the current best-performing method, alongside notable improvements in F1 score by 7.11% and 10.46%, respectively. These results underscore PharmaDual's robustness in predicting interactions involving previously unseen drugs. Overall, PharmaDual establishes a new state-of-the-art under both warm-start and cold-start evaluation protocols.

In real-world DDI prediction tasks, the number of positive interactions is typically far fewer than negative pairs, making the class distribution highly imbalanced. To account for this, we report both AUROC and AUPRC scores, with the latter being more informative under imbalanced conditions. Additionally, we assess model robustness under different positive-to-negative sampling ratios (e.g., 1:1, 1:5, 1:10) in the cold-start setting. Results are summarized in Table 7,

TABLE 3. The performance of PharmaDual and baselines on two datasets in the warm-start setting (%).

Model	DrugBank						Twosides					
	ACC	AUROC	AUPRC	F1	p(AUROC)	p(AUPRC)	ACC	AUROC	AUPRC	F1	p (AUROC)	p(AUPRC)
MHCADDI	83.80	91.16	89.26	85.06	0.072	0.083	-	88.20	-	-	0.063	0.071
SSI-DDI	96.33	98.95	98.57	96.38	0.063	0.073	78.20	85.85	82.71	79.81	0.036	0.043
MR-GNN	93.23	97.31	96.45	93.39	0.079	0.056	85.39	91.93	89.32	86.46	0.025	0.039
GMPNN-CS	95.31	98.45	97.91	95.40	0.063	0.043	86.96	92.94	90.38	87.85	0.048	0.039
GAT-DDI	92.03	96.28	94.64	92.29	0.061	0.046	67.32	75.16	72.50	63.70	0.028	0.035
DGNN-DDI	96.09	98.94	98.51	96.16	0.059	0.043	85.29	91.92	89.41	86.12	0.063	0.072
MM-GANN-DDI	95.29	97.20	95.38	94.13	0.054	0.021	85.63	91.83	90.86	88.69	0.045	0.037
KITE-DDI	91.28	95.16	94.81	91.33	0.053	0.056	70.62	79.95	73.68	62.97	0.038	0.092
Taco-DDI	94.86	96.49	95.76	95.93	0.041	0.041	84.95	90.88	88.94	85.93	0.056	0.079
PharmaDual(ours)	98.39	99.15	99.62	98.09	0.022	0.013	91.38	94.17	92.19	89.52	0.015	0.018

Note. p-values are obtained from one-sided paired t-tests over five independent runs between each baseline and PharmaDual. All AUROC and AUPRC improvements of PharmaDual are statistically significant ($p < 0.05$).

TABLE 4. Train/test counts of the 5 rarest interaction types across all folds on the DrugBank dataset (cold-start setting).

Interaction Type	Total	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
Type 85	23	18 / 5	19 / 4	18 / 5	20 / 3	17 / 6
Type 0	30	22 / 8	21 / 9	24 / 6	24 / 6	23 / 7
Type 82	32	26 / 6	22 / 10	25 / 7	24 / 8	26 / 6
Type 61	53	45 / 8	44 / 9	42 / 11	45 / 8	46 / 7
Type 38	76	61 / 15	62 / 14	60 / 16	55 / 21	63 / 13

Note: Each cell shows train / test count per fold. Even rare types appear in both sets.

TABLE 5. Train/test counts of the 5 rarest interaction types across all folds on the Twosides dataset (cold-start setting).

Interaction Type	Total	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
Type 63	152	118 / 34	121 / 31	138 / 14	130 / 22	115 / 37
Type 126	187	150 / 37	155 / 32	140 / 47	162 / 25	153 / 34
Type 109	216	173 / 43	160 / 56	180 / 36	182 / 34	165 / 51
Type 218	258	210 / 48	201 / 57	195 / 63	190 / 68	223 / 35
Type 482	329	263 / 66	260 / 69	271 / 211	253 / 229	275 / 54

Note: Each cell shows train / test count per fold. Even rare types appear in both sets.

showing that the model maintains stable performance across varying class priors.

F. ABLATION STUDY

To further investigate the contribution of each module in PharmaDual, we perform an ablation study under the cold-start setting on DrugBank dataset, which more effectively differentiates the performance of the models. The variants considered in this study are as follows:

- w/o 2D view module: This variant removes the 2D molecular representation and relies solely on the 3D view for encoding drug information, aiming to assess the role of geometric information in DDI prediction.
- w/o 3D view module: This variant removes the 3D molecular representation and utilizes only the 2D view module to encode drug information, aiming to evaluate the contribution of graph features to the overall performance.

The results of the ablation study in Table 8 show that each module contributes significantly to the overall performance of PharmaDual. When the 2D view module is removed (w/o 2D view module), there is a marked decline in both AUROC

and F1 scores, with AUROC dropping by 3.99% and F1 by 4.88%. This indicates that the 2D molecular representation, which captures structural information, is crucial for accurately modeling the interactions between drugs. On the other hand, removing the 3D view module (w/o 3D view module) also leads to a significant performance reduction, with AUROC and F1 decreasing by 4.19% and 5.31%, respectively. This suggests that 3D molecular features, which capture geometric patterns in the drug structure, provide complementary information that enhances the predictive accuracy.

To quantitatively evaluate the benefit of pharmacophore integration over traditional fingerprint features (e.g., ECFP), we conducted additional ablation experiments. PharmaDual_ESFP extracts drug features using ESFP, while PharmaDual obtains drug features based on pharmacophores. The results, summarized in Table 9, demonstrate that pharmacophore-based representations consistently outperform ECFP across multiple metrics. This confirms the effectiveness of integrating pharmacophore information in our framework.

G. PARAMETER SENSITIVITY STUDIES

In this work, we perform a comprehensive analysis of the influence of key hyperparameters on the effectiveness of the proposed PharmaDual framework. Specifically, we examine the effects of 2D embedding dimensionality, 3D embedding dimensionality, the number of self-attention heads in the Transformer, and the depth of the GNN. The hyperparameter sensitivity experiments are conducted on the DrugBank dataset under the warm-start evaluation setting.

1) THE DIMENSION OF 2D FEATURE EMBEDDINGS

To investigate the influence of 2D feature embedding dimensions on model performance, we conducted experiments using five different settings: 16, 32, 64, 128, and 256. As illustrated in Figure 3, the model achieves optimal performance when the embedding dimension is set to 64. Lower dimensions, such as 16 and 32, provide limited representational capacity, which hinders the model's ability to effectively capture complex drug-related features, leading to inferior predictive performance. Conversely, higher dimensions (128 and 256) offer increased expressiveness but

TABLE 6. The performance of PharmaDual and baselines on two datasets in the cold-start setting (%).

Model	DrugBank						Twosides					
	ACC	AUROC	AUPRC	F1	p(AUROC)	p(AUPRC)	ACC	AUROC	AUPRC	F1	p(AUROC)	p(AUPRC)
MHCADDI	70.58	77.84	76.16	72.74	0.079	0.085	66.50	72.53	71.06	67.21	0.065	0.073
SSI-DDI	76.38	84.23	84.94	73.54	0.081	0.085	65.40	73.43	75.03	54.12	0.059	0.065
MR-GNN	75.99	84.85	84.89	72.30	0.072	0.073	67.33	76.52	75.25	59.71	0.068	0.079
GMPNN-CS	79.95	89.34	89.25	77.22	0.065	0.062	71.57	81.90	82.90	63.83	0.058	0.062
GAT-DDI	77.94	86.58	85.81	75.28	0.046	0.069	71.55	80.71	80.44	65.91	0.043	0.052
DGNN-DDI	77.07	86.35	86.97	73.03	0.055	0.043	70.31	85.12	83.71	59.41	0.039	0.035
MM-GANN-DDI	81.62	85.26	86.39	72.82	0.065	0.053	72.36	84.29	82.16	70.69	0.059	0.037
KITE-DDI	80.93	86.13	87.01	73.96	0.066	0.092	70.15	83.62	73.68	71.34	0.038	0.063
Taco-DDI	79.65	85.61	85.41	75.15	0.079	0.081	73.75	84.29	88.94	73.56	0.056	0.061
PharmaDual(ours)	85.29	92.35	90.18	84.33	0.018	0.016	81.95	89.57	86.95	77.67	0.021	0.036

Note. p-values are obtained from one-sided paired t-tests over five independent runs between each baseline and PharmaDual. All AUROC and AUPRC improvements of PharmaDual are statistically significant ($p < 0.05$).

TABLE 7. Performance under different positive-to-negative ratios in the cold-start setting on the DrugBank dataset.

Pos:Neg Ratio	AUROC	AUPRC
1:1	92.35	90.18
1:5	92.06	89.53
1:10	91.89	88.98

TABLE 8. Ablation study performance of PharmaDual.

Model	ACC	AUROC	AUPRC	F1
PharmaDual(w/o 2D view module)	81.25	80.33	81.43	77.28
PharmaDual(w/o 3D view module)	81.16	80.13	80.92	76.85
PharmaDual	85.29	84.32	84.29	82.16

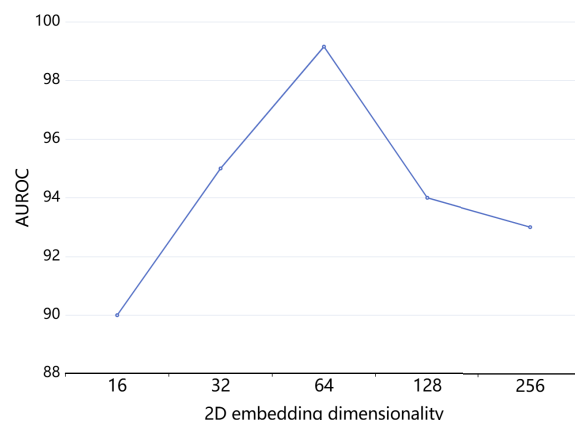
introduce feature redundancy and a higher risk of overfitting, ultimately compromising the model's generalization capability. These results indicate that an embedding dimension of 64 provides a favorable trade-off between representation richness and generalization, thereby yielding the most robust performance for DDI prediction.

2) THE DIMENSION OF 3D FEATURE EMBEDDINGS

To assess how varying the dimensionality of 3D feature embeddings affects model efficacy, we evaluated five configurations: 16, 32, 64, 128, and 256. As shown in Figure 4, the highest predictive accuracy is attained when the embedding size is set to 64. Embeddings of lower dimensionality (i.e., 16 and 32) fail to adequately capture the intricate structural and spatial characteristics of drugs, thereby limiting the model's capacity to learn informative representations. On the other hand, excessively large dimensions (128 and 256) tend to introduce redundant or noisy information, which not only inflates the model complexity but also increases susceptibility to overfitting. Taken together, these findings suggest that a 3D embedding dimension of 64 achieves an optimal balance

TABLE 9. Performance comparison of PharmaDual and PharmaDual_ESFP on the DrugBank dataset.

Model	ACC	AUROC	AUPRC	F1
PharmaDual_ESFP	93.51	93.62	92.86	91.71
PharmaDual	98.39	99.15	99.62	98.09

**FIGURE 3.** Sensitivity analysis on the dimension of 2D feature embeddings.

between expressive power and generalization, leading to superior DDI prediction performance.

3) THE NUMBER OF ATTENTION HEADS

To explore how varying the number of attention heads affects the performance of the PharmaDual model, we evaluate four configurations: 2, 4, 6, and 8 heads. As depicted in Figure 5, the model achieves its highest performance when employing 4 attention heads, reaching an ACC of 98.39% and an AUROC of 99.15%. The subpar results observed with 6 heads likely stem from limited capacity to capture complex and heterogeneous drug interactions. On the other hand, increasing the number of heads to 8 slightly diminishes performance, which may be attributed to increased model complexity leading to overfitting or reduced training stability. Overall, the configuration with 4 attention heads appears to strike an effective trade-off between expressiveness and robustness, making it the most suitable choice in our experimental setup.

4) THE DEPTH OF THE GNN

To investigate the effect of GNN depth on model performance, we conduct experiments with networks consisting of 2, 3, 4, and 5 layers. As shown in Figure 6, the best

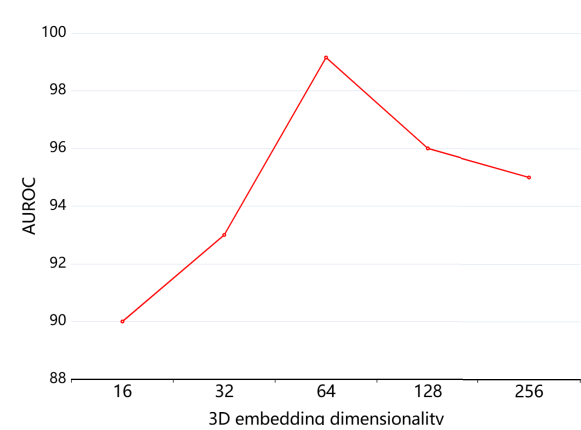


FIGURE 4. Sensitivity analysis on the dimension of 3D feature embeddings.

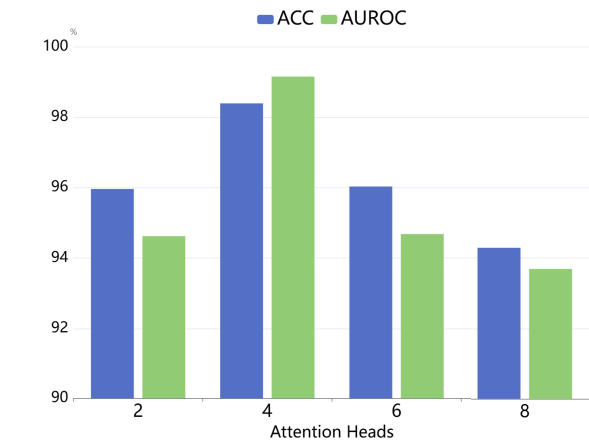


FIGURE 5. Sensitivity analysis of the number of attention heads.

results are achieved with a 3-layer GNN, indicating that this configuration provides a sufficient receptive field to capture meaningful topological and semantic information without introducing excessive noise. A 4-layer GNN yields slightly inferior performance, suggesting that additional layers may offer marginal gains at the cost of increased complexity. In contrast, both shallow (2-layer) and overly deep (5-layer) architectures exhibit suboptimal performance. The 2-layer GNN likely lacks the capacity to model long-range dependencies within molecular graphs, while the 5-layer configuration may suffer from over-smoothing or gradient vanishing issues. These findings suggest that a 3-layer GNN offers an effective trade-off between model expressiveness and stability in our setting.

V. CASE STUDY

In this section, we further verify the DDIs prediction effectiveness of our proposed framework in the real-world scenario by investigating case studies. Particularly, we randomly selected 5 pairs of predicted DDIs generated by PharmaDual. Table 10 records the predicted results and the real DDIs

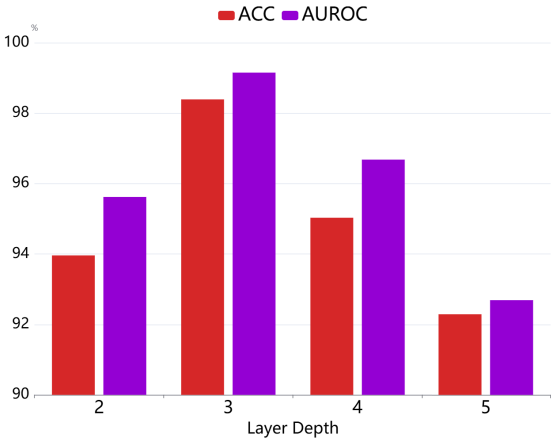


FIGURE 6. Sensitivity analysis of the depth of the GNN.

(i.e., evidence). By analyzing Table 10 and the corresponding references [61], [62], [63], [64], [65], we can find that all 5 pairs DDIs prediction results are in agreement with the evidence found in precious works [61], [62], [63], [64], [65]. For instance, Stage et al. found that when Pyrimethamine was combined with Aliskiren, it produced the side effects of Sarcoma by competitively inhibiting MATE1 and MATE2-K [61]. Banakh et al. verified that the catechol-Omethyltransferase produced by the interaction between Pyrimethamine and Tolcapone caused Breast disorder [62]. Parving et al. proved through clinical experiments that the combination of Atorvastatin and Amlodipine can produce Muscle inflammation [63]. The presented evidence demonstrates the promising and practical predictive performance of our proposed RPNAnet framework in real-world DDI prediction tasks.

TABLE 10. Case study on the DDI prediction results of PharmaDual.

DrugA	DrugB	Interaction	Evidence
Pyrimethamine	Aliskiren	Sarcoma	[61]
Tolcapone	Pyrimethamine	Breast disorder	[66]
Amlodipine	Atorvastatin	Muscle inflammation	[67]
Omeprazole	Amoxicillin	Renal tubular acidosis	[62]
Aliskiren	Tioconazole	Breast inflammation	[63]

To further demonstrate the interpretability of our model, we examined three drug pairs with known DDIs: Lopinavir–Arbidol, Mefloquine–Rupintrivir, and Triflupromazine hydrochloride–Lopinavir (Figure 7). For Lopinavir–Arbidol, the model identified substructures linked to CYP3A4 inhibition, aligning with their known metabolic competition. In Mefloquine–Rupintrivir, it captured shared antiviral features that may interfere with viral replication pathways. For Triflupromazine hydrochloride–Lopinavir, the model focused on CNS-related regions, suggesting possible pharmacodynamic interactions. These results highlight the model’s ability to provide mechanistic insights alongside accurate DDI predictions.

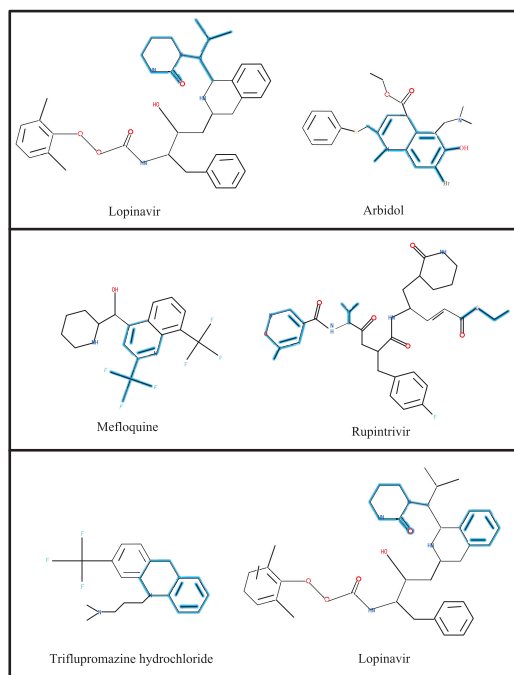


FIGURE 7. Case studies demonstrating pharmacophores identified by PharmaDual in clinically confirmed DDIs.

To further quantify interpretability, we evaluated two standard explanation metrics on the DrugBank dataset: fidelity and sparsity.

- **Fidelity** measures how consistently the model maintains its prediction when only the pharmacophore substructures are used instead of the full molecular graphs. Specifically, let $f(\cdot)$ denote the DDI prediction function, G_A and G_B be the original molecular graphs of the drug pair, and G_A^{pharma} , G_B^{pharma} be their corresponding pharmacophore subgraphs. Fidelity is then defined as:

$$Fidelity = \frac{1}{N} \sum_{i=1}^N \mathbb{1}[f(G_A, G_B) = f(G_A^{pharma}, G_B^{pharma})] \quad (32)$$

where N is the number of test samples and $\mathbb{1}[\cdot]$ is the indicator function.

- **Sparsity** reflects the conciseness of the explanation and is defined as the ratio of atoms in the pharmacophore to the total number of atoms in the full molecule:

$$Sparsity = \frac{|V^{pharma}|}{|V|} \quad (33)$$

where $|V^{pharma}|$ is the number of atoms in the pharmacophore subgraph and $|V|$ is the number in the full molecular graph.

On the DrugBank test set, our model achieved an average fidelity of 93.5% and sparsity of 14.2%, indicating that the pharmacophores are both highly informative and compact. These metrics quantitatively support the mechanistic relevance of the highlighted substructures shown in Figure 7.

VI. CONCLUSION

In this study, we presented PharmaDual, a dual-view learning framework for DDI prediction that jointly leverages the 2D structural and 3D spatial representations of drug molecules. By employing pharmacophore-based substructure encoding in both 2D and 3D views, PharmaDual captures rich chemical semantics and spatial configurations that are critical for modeling drug interactions. The use of a cross-attention fusion mechanism further enables effective integration of complementary multi-view features, resulting in more accurate and generalizable DDI predictions. Experimental evaluations demonstrate that PharmaDual achieves superior performance compared to existing state-of-the-art methods. Nevertheless, our current approach does not yet incorporate dynamic or temporal interaction information, which could reflect time-dependent drug effects. Future research will explore integrating temporal data and real-world pharmacovigilance signals to enhance the practical utility of PharmaDual in clinical settings.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their efforts to improve the manuscript.

REFERENCES

- [1] Y. Zhong, G. Li, J. Yang, H. Zheng, Y. Yu, J. Zhang, H. Luo, B. Wang, and Z. Weng, "Learning motif-based graphs for drug-drug interaction prediction via local-global self-attention," *Nature Mach. Intell.*, vol. 6, no. 9, pp. 1094–1105, Aug. 2024.
- [2] T. Takeda, M. Hao, T. Cheng, S. H. Bryant, and Y. Wang, "Predicting drug-drug interactions through drug structural similarities and interaction networks incorporating pharmacokinetics and pharmacodynamics knowledge," *J. Cheminformatics*, vol. 9, no. 1, pp. 1–9, Dec. 2017.
- [3] D. Huang, Z. Jiang, L. Zou, and L. Li, "Drug-drug interaction extraction from biomedical literature using support vector machine and long short term memory networks," *Inf. Sci.*, vols. 415–416, pp. 100–109, Nov. 2017.
- [4] X. Chen, B. Ren, M. Chen, Q. Wang, L. Zhang, and G. Yan, "NLLSS: Predicting synergistic drug combinations based on semi-supervised learning," *PLOS Comput. Biol.*, vol. 12, no. 7, Jul. 2016, Art. no. e1004975.
- [5] K. Han, E. E. Jeng, G. T. Hess, D. W. Morgens, A. Li, and M. C. Bassik, "Synergistic drug combinations for cancer identified in a CRISPR screen for pairwise genetic interactions," *Nature Biotechnol.*, vol. 35, no. 5, pp. 463–474, May 2017.
- [6] X. Sun, K. Dong, L. Ma, R. Sutcliffe, F. He, S. Chen, and J. Feng, "Drug-drug interaction extraction via recurrent hybrid convolutional neural networks with an improved focal loss," *Entropy*, vol. 21, no. 1, p. 37, Jan. 2019.
- [7] J. L. Gulikers, L.-S. Otten, L. E. L. Hendriks, K. Winckers, Y. Henskens, J. Leentjens, M. M. van den Heuvel, R. ter Heine, S. Croes, B. Piet, and R. M. J. M. van Geel, "Proactive monitoring of drug-drug interactions between direct oral anticoagulants and small-molecule inhibitors in patients with non-small cell lung cancer," *Brit. J. Cancer*, vol. 131, no. 3, pp. 481–490, Aug. 2024.
- [8] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," *Science*, vol. 349, no. 6245, pp. 255–260, Jul. 2015.
- [9] Z. Wu, D. Shangguan, Q. Huang, and Y.-K. Wang, "Drug metabolism and transport mediated the hepatotoxicity of pleuropterus multiflorus root: A review," *Drug Metabolism Rev.*, vol. 56, no. 4, pp. 349–358, Oct. 2024.
- [10] M. Li, A. Micheli, Y. G. Wang, S. Pan, P. Lió, G. S. Gnecco, and M. Sanguineti, "Guest editorial: Deep neural networks for graphs: Theory, models, algorithms, and applications," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 4, pp. 4367–4372, Apr. 2024.
- [11] J. Li, R. Zheng, H. Feng, M. Li, and X. Zhuang, "Permutation equivariant graph framelets for heterophilous graph learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 9, pp. 11634–11648, Sep. 2024.

- [12] A. Deac, Y.-H. Huang, P. Veličković, P. Liò, and J. Tang, "Drug-drug adverse effect prediction with graph co-attention," 2019, *arXiv:1905.00534*.
- [13] J. Wang, X. Liu, S. Shen, L. Deng, and H. Liu, "DeepDDS: Deep graph neural network with attention mechanism to predict synergistic drug combinations," *Briefings Bioinf.*, vol. 23, no. 1, Jan. 2022, Art. no. bbab390.
- [14] X. Zhang, G. Wang, X. Meng, S. Wang, Y. Zhang, A. Rodriguez-Paton, J. Wang, and X. Wang, "Molormer: A lightweight self-attention-based method focused on spatial structure of molecular graph for drug-drug interactions prediction," *Briefings Bioinf.*, vol. 23, no. 5, Sep. 2022, Art. no. bbac296.
- [15] H. Wang, D. Lian, Y. Zhang, L. Qin, and X. Lin, "GoGNN: Graph of graphs neural network for predicting structured entity interactions," 2020, *arXiv:2005.05537*.
- [16] A. K. Nyamabo, H. Yu, and J.-Y. Shi, "SSI-DDI: Substructure-substructure interactions for drug-drug interaction prediction," *Briefings Bioinf.*, vol. 22, no. 6, Nov. 2021, Art. no. bbab133.
- [17] Z. Yang, W. Zhong, Q. Lv, and C. Yu-Chian Chen, "Learning size-adaptive molecular substructures for explainable drug-drug interaction prediction by substructure-aware graph neural network," *Chem. Sci.*, vol. 13, no. 29, pp. 8693–8703, 2022.
- [18] Z. Li, S. Zhu, B. Shao, X. Zeng, T. Wang, and T.-Y. Liu, "DSN-DDI: An accurate and generalized framework for drug-drug interaction prediction by dual-view representation learning," *Briefings Bioinf.*, vol. 24, no. 1, Jan. 2023, Art. no. bbac597.
- [19] H. Yu, S. Zhao, and J. Shi, "STNN-DDI: A substructure-aware tensor neural network to predict Drug-Drug interactions," *Briefings Bioinf.*, vol. 23, no. 4, Jul. 2022, Art. no. bbac209.
- [20] L. Li, M. Yu, R. Chin, A. Lucksiri, D. A. Flockhart, and S. D. Hall, "Drug-drug interaction prediction: A Bayesian meta-analysis approach," *Statist. Med.*, vol. 26, no. 20, pp. 3700–3721, Sep. 2007.
- [21] S. Vilar, R. Harpaz, E. Uriarte, L. Santana, R. Rabadan, and C. Friedman, "Drug-Drug interaction through molecular structure similarity analysis," *J. Amer. Med. Inform. Assoc.*, vol. 19, no. 6, pp. 1066–1074, Nov. 2012.
- [22] S. Vilar, E. Uriarte, L. Santana, N. P. Tatonetti, and C. Friedman, "Detection of drug-drug interactions by modeling interaction profile fingerprints," *PLoS ONE*, vol. 8, no. 3, Mar. 2013, Art. no. e58321.
- [23] S. Vilar, E. Uriarte, L. Santana, T. Lorberbaum, G. Hripcsak, C. Friedman, and N. P. Tatonetti, "Similarity-based modeling in large-scale prediction of drug-drug interactions," *Nature Protocols*, vol. 9, no. 9, pp. 2147–2163, Sep. 2014.
- [24] M. W. Harrold and R. M. Zavod, "Basic concepts in medicinal chemistry," Tech. Rep., 2014.
- [25] S. Cao, W. Lu, and Q. Xu, "Grarep: Learning graph representations with global structural information," in *Proc. 24th ACM Int. Conf. Inf. Knowl. Manage.*, 2015, pp. 891–900.
- [26] B. Perozzi, R. Al-Rfou, and S. Skiena, "DeepWalk: Online learning of social representations," in *Proc. 20th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2014, pp. 701–710.
- [27] A. Grover and J. Leskovec, "node2vec: Scalable feature learning for networks," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2016, pp. 855–864.
- [28] J. Tang, M. Qu, M. Wang, M. Zhang, J. Yan, and Q. Mei, "LINE: Large-scale information network embedding," in *Proc. 24th Int. Conf. World Wide Web*, May 2015, pp. 1067–1077.
- [29] D. Wang, P. Cui, and W. Zhu, "Structural deep network embedding," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2016, pp. 1225–1234.
- [30] M. Nickel, K. Murphy, V. Tresp, and E. Gabrilovich, "A review of relational machine learning for knowledge graphs," *Proc. IEEE*, vol. 104, no. 1, pp. 11–33, Jan. 2016.
- [31] N. Guan, D. Song, and L. Liao, "Knowledge graph embedding with concepts," *Knowl.-Based Syst.*, vol. 164, pp. 38–44, Jan. 2019.
- [32] P. Zhang, F. Wang, J. Hu, and R. Sorrentino, "Label propagation prediction of drug-drug interactions based on clinical side effects," *Sci. Rep.*, vol. 5, no. 1, p. 12339, Jul. 2015.
- [33] H. Yu, K.-T. Mao, J.-Y. Shi, H. Huang, Z. Chen, K. Dong, and S.-M. Yiu, "Predicting and understanding comprehensive drug-drug interactions via semi-nonnegative matrix factorization," *BMC Syst. Biol.*, vol. 12, no. S1, pp. 101–110, Apr. 2018.
- [34] J.-Y. Shi, K.-T. Mao, H. Yu, and S.-M. Yiu, "Detecting drug communities and predicting comprehensive drug-drug interactions via balance regularized semi-nonnegative matrix factorization," *J. Cheminformatics*, vol. 11, no. 1, pp. 1–16, Dec. 2019.
- [35] R. Masumshah, R. Aghdam, and C. Eslahchi, "A neural network-based method for polypharmacy side effects prediction," *BMC Bioinf.*, vol. 22, no. 1, pp. 1–17, Dec. 2021.
- [36] Y. Deng, X. Xu, Y. Qiu, J. Xia, W. Zhang, and S. Liu, "A multimodal deep learning framework for predicting drug-drug interaction events," *Bioinformatics*, vol. 36, no. 15, pp. 4316–4322, Aug. 2020.
- [37] J. L. Durant, B. A. Leland, D. R. Henry, and J. G. Nourse, "Reoptimization of MDL keys for use in drug discovery," *J. Chem. Inf. Comput. Sci.*, vol. 42, no. 6, pp. 1273–1280, Nov. 2002.
- [38] E. E. Bolton, Y. Wang, P. A. Thiessen, and S. H. Bryant, "PubChem: Integrated platform of small molecules and biological activities," in *Annual Reports in Computational Chemistry*. Amsterdam, The Netherlands: Elsevier, 2008, pp. 217–241.
- [39] A. Gottlieb, G. Y. Stein, Y. Oron, E. Ruppin, and R. Sharan, "INDI: A computational framework for inferring drug interactions and their associated recommendations," *Mol. Syst. Biol.*, vol. 8, no. 1, p. 592, Jan. 2012.
- [40] W. Zhang, Y. Chen, F. Liu, F. Luo, G. Tian, and X. Li, "Predicting potential drug-drug interactions by integrating chemical, biological, phenotypic and network data," *BMC Bioinf.*, vol. 18, no. 1, pp. 1–12, Dec. 2017.
- [41] R. Ferdousi, R. Safdari, and Y. Omid, "Computational prediction of drug-drug interactions based on drugs functional similarities," *J. Biomed. Informat.*, vol. 70, pp. 54–64, Jun. 2017.
- [42] W. Zhang, Y. Chen, D. Li, and X. Yue, "Manifold regularized matrix factorization for drug-drug interaction prediction," *J. Biomed. Informat.*, vol. 88, pp. 90–97, Dec. 2018.
- [43] T. Ma, C. Xiao, J. Zhou, and F. Wang, "Drug similarity integration through attentive multi-view graph auto-encoders," 2018, *arXiv:1804.10850*.
- [44] C. He, Y. Liu, H. Li, H. Zhang, Y. Mao, X. Qin, L. Liu, and X. Zhang, "Multi-type feature fusion based on graph neural network for drug-drug interaction prediction," *BMC Bioinf.*, vol. 23, no. 1, p. 224, Dec. 2022.
- [45] S. Pang, Y. Zhang, T. Song, X. Zhang, X. Wang, and A. Rodriguez-Patón, "AMDE: A novel attention-mechanism-based multidimensional feature encoder for drug-drug interaction prediction," *Briefings Bioinf.*, vol. 23, no. 1, Jan. 2022, Art. no. bbab545.
- [46] H. He, G. Chen, and C. Y.-C. Chen, "3DGT-DDI: 3D graph and text based neural network for drug-drug interaction prediction," *Briefings Bioinf.*, vol. 23, no. 3, May 2022, Art. no. bbac134.
- [47] K. T. Schütt, P.-J. Kindermans, H. E. Sauceda, S. Chmiela, A. Tkatchenko, and K. Müller, "SchNet: A continuous-filter convolutional neural network for modeling quantum interactions," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017.
- [48] J. Degen, C. Wegscheid-Gerlach, A. Zaliani, and M. Rarey, "On the art of compiling and using 'drug-like' chemical fragment spaces," *ChemMedChem*, vol. 3, no. 10, pp. 1503–1507, Oct. 2008.
- [49] H. P. G. Thompson and G. M. Day, "Which conformations make stable structures? Mapping crystalline molecular geometries to the conformational energy landscape," *Chem. Sci.*, vol. 5, no. 8, pp. 3173–3182, 2014.
- [50] M. Salha, H. Adenusi, J. H. Dupuis, E. Bodo, B. Botta, I. McKenzie, R. Y. Yada, D. H. Farrar, J. Magolan, K. V. Tian, and G. A. Chass, "Bioactivity of the cannabigerol cannabinoid and its analogues—The role of 3-dimensional conformation," *Organic Biomolecular Chem.*, vol. 21, no. 22, pp. 4683–4693, 2023.
- [51] H. Luo, P. Zhang, H. Huang, J. Huang, E. Kao, L. Shi, L. He, and L. Yang, "DDI-CPI, a server that predicts drug-drug interactions through implementing the chemical-protein interactome," *Nucleic Acids Res.*, vol. 42, no. W1, pp. W46–W52, Jul. 2014.
- [52] D. S. Wishart, "DrugBank 5.0: A major update to the DrugBank database for 2018," *Nucleic Acids Res.*, vol. 46, no. D1, pp. D1074–D1082, Jan. 2018.
- [53] M. Zitnik, M. Agrawal, and J. Leskovec, "Modeling polypharmacy side effects with graph convolutional networks," *Bioinformatics*, vol. 34, no. 13, pp. i457–i466, Jul. 2018.
- [54] N. Xu, P. Wang, L. Chen, J. Tao, and J. Zhao, "MR-GNN: Multi-resolution and dual graph neural network for predicting structured entity interactions," 2019, *arXiv:1905.09558*.

- [55] A. K. Nyamabo, H. Yu, Z. Liu, and J.-Y. Shi, "Drug-drug interaction prediction with learnable size-adaptive molecular substructures," *Briefings Bioinf.*, vol. 23, no. 1, Jan. 2022, Art. no. bbab441.
- [56] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, "Graph attention networks," 2017, *arXiv:1710.10903*.
- [57] M. Ma and X. Lei, "A dual graph neural network for drug-drug interactions prediction based on molecular structure and interactions," *PLOS Comput. Biol.*, vol. 19, no. 1, Jan. 2023, Art. no. e1010812.
- [58] J. Feng, Y. Liang, and T. Yu, "MM-GANN-DDI: Multimodal graph-agnostic neural networks for predicting Drug-Drug interaction events," *Comput. Biol. Med.*, vol. 166, Nov. 2023, Art. no. 107492.
- [59] A. Tamir and J.-S. Yuan, "KITE-DDI: A knowledge graph integrated transformer model for accurately predicting drug-drug interaction events from drug SMILES and biomedical knowledge graph," *IEEE Access*, vol. 13, pp. 40028–40043, 2025.
- [60] J. Qiao, X. Guo, J. Jin, D. Wang, K. Li, W. Gao, F. Cui, Z. Zhang, H. Shi, and L. Wei, "Taco-DDI: Accurate prediction of drug-drug interaction events using graph transformer-based architecture and dynamic co-attention matrices," *Neural Netw.*, vol. 189, Sep. 2025, Art. no. 107655.
- [61] M. G. Russo, M. I. Sancho, L. M. A. Silva, H. A. Baldoni, T. Venancio, J. Ellena, and G. E. Narda, "Looking for the interactions between omeprazole and amoxicillin in a disordered phase. An experimental and theoretical study," *Spectrochimica Acta A, Mol. Biomolecular Spectrosc.*, vol. 156, pp. 70–77, Mar. 2016.
- [62] I. Banakh, K. Haji, R. Kung, S. Gupta, and R. Tiruvoipati, "Severe rhabdomyolysis due to presumed drug interactions between atorvastatin with amlodipine and ticagrelor," *Case Rep. Crit. Care*, vol. 2017, pp. 1–4, Jan. 2017.
- [63] H.-H. Parving, B. M. Brenner, J. J. V. McMurray, D. de Zeeuw, S. M. Haffner, S. D. Solomon, N. Chaturvedi, F. Persson, A. S. Desai, M. Nicolaidis, A. Richard, Z. Xiang, P. Brunel, and M. A. Pfeffer, "Cardiorenal end points in a trial of aliskiren for type 2 diabetes," *New England J. Med.*, vol. 367, no. 23, pp. 2204–2213, Dec. 2012.
- [64] S. Z. Alparslan-Gök, S. Miquel, and S. H. Tijs, "Cooperation under interval uncertainty," *Math. Methods Oper. Res.*, vol. 69, no. 1, pp. 99–109, Mar. 2009.
- [65] A. Çevik, "Computer-aided diagnosis of Alzheimer's disease and mild cognitive impairment with MARS/CMARS classification using structural mr images," Tech. Rep., 2017.
- [66] T. B. Stage, K. Brøsen, and M. M. H. Christensen, "A comprehensive review of Drug-Drug interactions with metformin," *Clin. Pharmacokinetics*, vol. 54, no. 8, pp. 811–824, Aug. 2015.
- [67] J. Bicker, A. Fortuna, G. Alves, P. Soares-Da-Silva, and A. Falcão, "Elucidation of the impact of P-glycoprotein and breast cancer resistance protein on the brain distribution of Catechol-O-Methyltransferase inhibitors," *Drug Metabolism Disposition*, vol. 45, no. 12, pp. 1282–1291, Dec. 2017.



WENXIAO ZHANG received the M.S. degree in business administration from Shandong Normal University, China, in 2020. She is currently pursuing the Ph.D. degree with the School of Computer and Engineering, Kunsan National University. Her main research interests include educational data mining, deep learning, and causal inference.



SEONG YOON SHIN received the M.S. and Ph.D. degrees from the Department of Computer Information Engineering, Kunsan National University, Gunsan, Republic of Korea, in 1997 and 2003, respectively. Since 2006, he has been a Professor with the School of Computer Science and Engineering. His research interests include educational data mining, deep learning, and image processing.



HAILIANG TANG received the M.S. degree from the School of Information Science and Engineering, Shandong Normal University, China, in 2017. He is currently pursuing the Ph.D. degree with the School of Software, Kunsan National University. His main research interests include educational data mining, deep learning, and causal inference.

...