

RESEARCH ARTICLE

Molecular Substructure-Aware Network With Reinforcement Pooling and Deep Attention Mechanism for Drug-Drug Interaction Prediction

KE ZHAO¹, HAILIANG TANG¹, HAILIN ZHU¹, AND WENXIAO ZHANG²¹School of Information Science and Engineering, Qilu Normal University, Jinan 250200, China²School of Finance and Economics, Shandong University of Engineering and Vocational Technology, Jinan 250200, China


Corresponding author: Hailin Zhu (China.zhuhl_qlnu@163.com)

ABSTRACT Predicting potential drug-drug interactions (DDIs) can effectively mitigate unforeseen interactions throughout the entire drug development process, playing a pivotal role in ensuring drug safety. However, traditional methods are laborious and require specific expert knowledge. This paper proposes RPDAnet, a novel molecular substructure-aware network based on Reinforced Pooling and Deep Attention mechanism, to investigate the interactive relationships between drugs and predict the potential DDIs. Particularly, RPDAnet leverages reinforcement learning to dynamically select informative molecular fragments, thus enhancing its generalization capacity without relying on prior knowledge. Subsequently, RPDAnet develops Communicative Message Massing Neural Network (CMPNN) to enhance the representation of molecular structures by reinforcing message interactions between nodes and edges through a communicative kernel. Finally, RPDAnet aggregates the interactions between substructures of drugs to predict the DDI between a pair of drugs. The experimental results on two real-world datasets demonstrate that our proposed RPDAnet outperforms the state-of-the-art methods with more than 5% performance gains in DDI prediction.

INDEX TERMS Drug-drug interactions, reinforcement learning, deep attention neural networks, learning latent representations, graph algorithms.

I. INTRODUCTION

Drug-drug interactions (DDIs) frequently arise when one drug induces a pharmacokinetic (PK) or pharmacodynamic (PD) effect in the presence of another, and they are primary factors leading to medical injuries [1], [2]. DDIs commonly contribute to adverse drug reactions (ADRs) and elevated healthcare expenditures, which pose substantial threats to patients and public health [3]. For instance, acetylsalicylic acid, more commonly known as aspirin, is a medication employed to alleviate pain and fever stemming from diverse etiologies. Aspirin possesses anti-inflammatory and antipyretic properties, as well as the ability to inhibit

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

platelet aggregation, making it valuable in the prevention of blood clots and myocardial infarction. However, the potential for elevated risk or increased severity of hypertension exists, as exemplified by negative drug-drug interactions when acetylsalicylic acid is co-administered with 1-benzylimidazole [4]. Recently, the study [5] revealed that approximately 6.7% of hospital readmissions in the USA in 2021 were attributable to DDIs, with a fatality rate of 0.32%.

The prediction of potential DDIs not only diminishes unexpected drug interactions but also reduces the expense associated with drug development, thereby facilitating the optimization of the drug design process. Consequently, investigating DDI holds significance in both drug development and clinical practice, particularly concerning co-administered medications. To further economize and enable the analysis of

extensive interaction datasets, automated techniques for DDI identification are imperative.

The identification of drug-drug interactions (DDIs) is typically conducted in pharmaceutical research and clinical settings through extensive experimental testing, encompassing *in vitro* experiments and clinical trials. However, the extensive manual analysis involved in studying DDIs is both time-consuming and laborious, even when employing high-throughput methods [6]. In recent years, computational methods (*in silico*) have emerged as a cost-effective, efficient, and rapid alternative to tackle this challenge. By harnessing existing knowledge of known DDIs, these computational methods can predict potential interactions, thereby alleviating the necessity for exhaustive experimental testing.

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. Li et al. [7] proposed the pharmacokinetics (PK) model, one of the pioneering computational approaches, to predict the DDIs. Afterward, Vilar et al. employed drug similarities computed from 2D and 3D molecular structures, interaction profiles, and their combinations to predict potential drug pairs [8], [9], [10]. Moreover, Sridhar et al. [11] devised a DDI prediction framework based on probabilistic programming and drug similarities, enabling the identification of drug pairs with a heightened probability of interaction. Although similarity-based methods have exhibited commendable performance, they encounter challenges in investigating complex interactive relationships due to their inability to capture high-order connective features. To address this limitation, Rohani and Eslahchi [12] proposed a neural network-based approach that incorporates non-linear similarity fusion, enabling deep learning of high-dimensional drug features and extending the model's capacity from local to a global perspectives. In addition, the MRMF model [13] based on manifold regularization matrix factorization is presented, which employs manifold regularization to embed multiple drug features and utilizes matrix factorization techniques for predicting potential DDIs. The MRMF takes advantage of the inherent structure within drug data by incorporating manifold regularization, thereby providing an enhanced framework for DDI prediction.

Recently, molecule structure-based methods for predicting DDIs have gained significant popularity, these approaches rely on analyzing medicinal chemistry knowledge [14], drugs are composed of various functional groups and chemical substructures to explore pharmacokinetic and pharmacodynamic properties and predict potential DDIs. Molecular structure-based methods [15], [16], [17], [18], [19], [20] treat drugs as independent entities and predict DDIs solely based on drug pairs, without the need for external biomedical knowledge. These methods focus on local chemical structures (substructures) rather than the entire molecular structure, as DDIs primarily arise from chemical

reactions among these substructures [16], [21]. Molecular structure-based methods assume that the learned information about chemical substructures can be generalized to different drugs with similar substructures [15], [20]. For instance, MR-GNN [19] utilized graph neural networks (GNNs) to extract multi-scale representations of chemical substructures from molecular graphs. CASTER [15] employed a chemical sequential pattern mining algorithm to generate recurring chemical substructures as representations of drugs. This is followed by an auto-encoding module and dictionary learning to enhance the model's generalizability and interpretability. SSI-DDI [18], MHCADDI [17], and CMPNN-CS [20] incorporated the co-attention mechanism between the learned substructures of drug pairs, allowing for communication between the drugs. In CMPNN-CS, bonds are viewed as gates that control the flow of messages passing through the GNN, effectively representing substructures in a self-supervised manner.

However, the aforementioned methods present the following challenges: (1) Firstly, these methods ignore the passing messages from bonds to atoms, which limits the ability to efficiently capture complementary information between atoms and bonds. (2) Secondly, current approaches treat the molecule fragments as equally important, but the fact is that each fragment contributes differently to the properties of the molecule. (3) Thirdly, current approaches primarily focus on predicting drug-drug interactions (DDIs) at a macroscopic level, neglecting the prediction of DDIs at the molecular substructure level.

To address the above challenges, we propose RPDAnet, a novel DDIs prediction framework based on reinforced pooling and deep attention mechanism, which can effectively model molecular fragments and accurately predict potential DDIs. The main contributions of this work are summarized as follows:

- We propose a novel graph communication neural network to interactively update edge and node representations of drugs, which can effectively capture high-quality molecular fragment representations.
- We present a reinforcement pooling mechanism (RPM) to adaptively screen most related neighboring nodes for target nodes through a reinforcement learning process, which helps RPDAnet to aggregate more informative molecular fragments to obtain a more precise molecular representation.
- We develop a deep attention neural network (DANN) to acquire comprehensive insights into the interactive relationships between drugs, enabling accurate prediction of DDIs through weighted aggregation of interaction scores between substructures.
- We conduct extensive experiments on real-world datasets to evaluate the performance of RPDAnet, experimental results demonstrate that the RPDAnet outperforms the state-of-the-art methods in DDIs prediction.

II. RELATED WORK

This section provides a comprehensive overview of previous research in the field of DDI prediction, with particular emphasis on two pivotal aspects: (1) drug representation, and (2) the DDI prediction methodologies.

A. DRUG REPRESENTATION

The majority of existing methods for DDI prediction employ molecular fingerprints [22], [23], [24] or other drug profiles, including side effects [22], [23], binding targets [24], transporters, enzymes, pathways, and combinations thereof [23], [25], [26] to predict potential DDIs. Molecular fingerprints [27], [28] 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 [8], [13], [29], [30], [31], [32] 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 [8]. 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.

In recent years, there has been an increasing application of graph neural networks (GNNs) [33], [34], [35], [36], which are deep learning models designed specifically for graph-structured data. These models have shown promising performance in various tasks related to chemical molecules and DDIs [37], [38], [39], [40]. However, the majority of existing approaches primarily focus on acquiring representations of drugs as a whole entity, ignoring the crucial interactive relationships within DDIs, specifically the functional groups and chemical substructures that constitute the drug molecule. Although some recent methods [41], [42] have been proposed to address the involvement of substructures in DDIs, they treat the hidden representations of nodes (referred to as patch representations) at each GNN layer as substructure representations of drugs. Consequently, this approach generates substructures with regular shapes whose sizes are determined by the receptive field of the GNN layer, resulting in their generalization ability being compromised. Different from the existing methods, we propose a novel framework to extract substructures from molecular graphs, enabling the direct learning of substructures with diverse sizes and shapes within the molecule.

The existing DDI prediction methods can be classified into three main categories: machine learning-based methods,

deep learning-based methods, and GNN-based methods [43]. Machine learning-based methods utilize natural language processing techniques to extract annotated DDIs from diverse biomedical literature sources, including medical reports, electronic medical records [44], scientific literature [45], and insurance claim databases. For instance, Liu et al. [46] employed large-scale text mining and statistical inference techniques to capture latent semantics and enhance prediction performance. Although such methods are effective in identifying approved DDIs, they are unable to detect unlabeled or potential DDIs before drug combination treatment. Deep learning-based methods typically train deep neural networks to end-to-end extract drug properties features, such as substructures, pathways, targets, and structural similarity profiles, to predict the potential DDIs. Particularly, Deep-CCI [47] employs convolutional neural networks (CNNs) and fully connected layers to extract hidden representations from SMILES encoding, thereby effectively predicting drug-drug interactions by assembling various neural units. However, a major limitation of these methods is that they require a large number of parameters to guarantee the performance of the model, greatly reducing the interpretability of the models. In contrast, GNN-based methods [48], [49], [50] represent biomedical data as heterogeneous graphs and employ various graph-specific approaches, such as graph auto-encoders [32], matrix factorization [23], and label propagation [22], to analyze drug-drug interconnections on the graphs. By incorporating external biomedical knowledge, GNN-based methods enhance prediction performance. However, their performance has not been fully exploited as they only leverage chemical structure information but ignore the fine-grained interactive relationships between drugs.

III. PROPOSED RPDAnet FRAMEWORK

The presented RPDAnet framework, as depicted in *Figure 1*, comprises three primary modules. More specifically, the feature encoding learning module employs the CMPNN encoder to acquire molecular representations, which are enhanced through the reinforcement of message interactions between nodes and edges via a communicative kernel. The reinforcement pooling module employs the RPM to select informative subgraphs, enhancing the model's generalization capacity without relying on prior knowledge. Finally, the DDIs prediction module employs a DAN network to accurately predict Drug-Drug Interactions (DDIs).

A. MODULE FEATURE ENCODING LEARNING MODULE

Firstly, we rely on RDKit to draw functional groups from SMILES of drug molecules. Then, a functional group can be seen in a small molecular graph. In this study, a small molecular graph is denoted as a graph $G = (V, R)$, where V refers to the set of nodes, and R denotes the set of bidirectional edges. For each node v , x_v represents its initial features, while for each edge $r(u, v)$, $x_{r(u,v)}$ represents the initial features. Notably, we extract distinct initial features for atoms and bonds within the original molecular graph G .

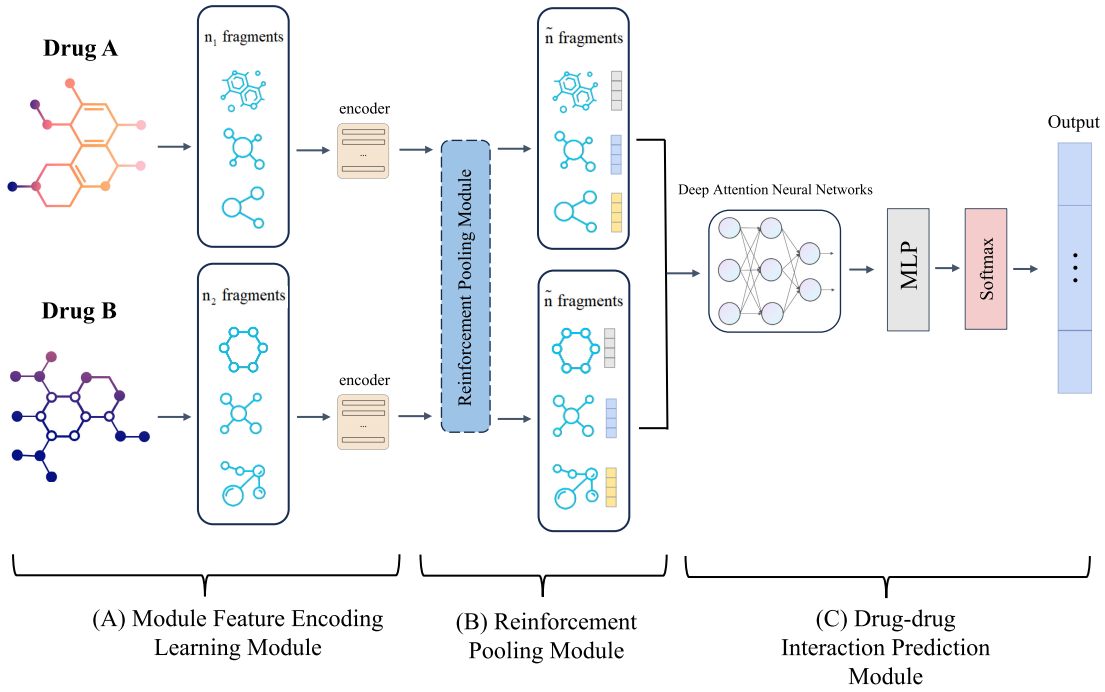


FIGURE 1. Overview of RPDAnet.

Here, we develop a CMPNN network as the graph encoder, which enhances graph representation by reinforcing message interactions between edges and nodes.

Firstly, to update the node hidden states, each node $v \in V$ aggregates representations of its incoming edges rather than its neighboring nodes in G . The intermediate message vector is obtained as follows:

$$\begin{aligned} m^k(v) &= \text{AGGREGATE}(\{h^{k-1}(r_{(u,v)}), \forall u \in N_v\}) \\ &= \sum_{u \in N(v)} h^{k-1}(r_{(u,v)}) \odot \text{pooling}(\sum_u h^{k-1}(r_{(u,v)})), \end{aligned} \quad (1)$$

where k represents the current layer, $r(u, v)$ denotes the edge between nodes u and v , while the pooling operator is a max pooling function, and \odot indicates an element-wise multiplication operator. We employ the max pooling technique to emphasize edges with the most significant information intensity, i.e., a node's hidden state predominantly relies on the most influential message from its incoming edges. Subsequently, the node's current hidden state $h^{k-1}(v)$ is concatenated with the message vector $m^k(v)$ and passed through a communicative function, which facilitates the update of the node's hidden state $h^k(v)$:

$$\begin{aligned} h^k(v) &= \text{COMMUNICATE}(m^k(v), h^{k-1}(v)) \\ &= \sigma(W^k \cdot \text{CONCAT}(h^{k-1}(v), m^k(v))), \end{aligned} \quad (2)$$

where the hidden state $h^k(v)$ serves as a pivotal message transfer station, responsible for receiving incoming messages, and subsequently integrating and transmitting them to the next station. σ is an activation function, and CONCAT

represents a concatenation operation. To achieve this, a specific communication function is employed, which entails feeding both the node and edge features into a Multi-Layer Perceptron (MLP) followed by a rectified linear unit (ReLU) activation.

Secondly, the message of the edge $r_{(v,m)}$ is derived by subtracting its inverse edge information from the hidden state $h^k(v)$:

$$m^k(r_{(v,w)}) = h^k(v) - h^{k-1}(r_{(v,w)}), \quad (3)$$

where $r_{(w,v)}$ represents the inverse edge of $r_{(v,w)}$. To update the edge hidden states, RPDAnet inputs the edge intermediate message $m^k(r_{(v,m)})$ into a fully connected layer, then incorporating them into the initial edge feature $x_{r_{(u,v)}}$. Subsequently, a ReLU activation function is applied to the output layer, which serves as the intermediate message vector for the subsequent iteration. This process can be concisely represented as follows:

$$h^k(r_{(v,w)}) = \sigma(x_{r_{(u,v)}} + W \cdot m^k(r_{(v,w)})), \quad (4)$$

Then, conducting k iterations to aggregate the first $k - 1$ layers hidden feature vectors:

$$m(v) = \text{AGGREGATE}(\{h^k(r_{(u,v)}), \forall u \in N(v)\}), \quad (5)$$

where the final node representation $h(v)$ of the graph is derived by communicating the information from incoming edges, the current node representation, and the initial node feature:

$$h(v) = \text{COMMUNICATE}(m(v), h^k(v), x_v), \quad (6)$$

Finally, we employ a readout operator to obtain the functional group representation:

$$h_G = \sum_{v \in V} \text{GRU}(h(v)), \quad (7)$$

where GRU [42] is the gated recurrent network.

B. REINFORCEMENT POOLING MODULE

DDIs primarily arise from interactions between functional groups. Therefore, we utilize the RPDAnet to model DDIs, which can highlight significant functional groups while downplaying the minor ones in DDI prediction tasks.

We propose a reinforced pooling mechanism (RPM) based on reinforcement learning to select significant functional groups in DDIs. Particularly, RPM introduces a top- k sampling tactic with an adaptive pooling ratio $k \in (0, 1]$ to select more informative nodes. RPM employs a trainable vector, denoted as p , which projects all functional group features into 1D footprints $\{vail_i | g_i \in G\}$. Here, $vail$ represents the amount of information preserved from a functional group g_i when projected onto the direction of p . Subsequently, we derive the importance values of functional groups $vail$ and rank the groups in descending order based on these values. To obtain a subset of relevant functional groups for the current batch, we select the top $n' = [k \cdot n]$ groups and exclude all other non-relevant functional groups. During the training phase, the computation of $vail$ for functional group g_i on p is expressed as follows:

$$vail_i = \frac{z_i p}{\|p\|}, \quad idx = \text{rank}(\{vail_i\}, n'), \quad (8)$$

where $(\{vail_i, n'\})$ represents the operation of functional group ranking, the resulting idx from $(\{vail_i\}, n')$ represents the indices corresponding to the selected functional groups. The parameter k is dynamically updated at the end of each epoch via reinforcement learning iterations.

Due to the pooling ratio k in $tok - k$ sampling is not directly involved in the optimization of the objective function, thus it cannot be tuned by the backpropagation algorithm. To overcome this limitation, RPM employs a reinforcement learning (RL) algorithm to autonomously explore the optimal k value within the range of $(0, 1]$, instead of treating it as a hyper-parameter subject to fine-tuning. RPM models the procedure of k updating as a finite Markov decision process (MDP) that includes states, actions, transitions, rewards, and termination, which are formulated as follows:

- **State.** At epoch e , the state s_e is denoted by the indices of the selected functional groups idx (i.e., Eq 8), the index labeling process is shown in Eq 9 :

$$s_e = idx_e. \quad (9)$$

- **Action.** The RL agent updates the value of k according to each action a_e , which is determined by the reward function. The action a is defined as either adding or subtracting a fixed value ϵ , where $\epsilon \in [0, 1]$.

- **Reward.** RPM introduces a discrete reward function, $\text{reward}(s_e, a_e)$, associated with each action a_e at state s_e , which guides the k to increase or decrease ϵ based on the accuracy loss between epoch e and epoch $e - 1$:

$$\text{reward}(s_e, a_e) \begin{cases} +1, & \text{if } acc_e > acc_{e-1} \\ 0, & \text{if } acc_e = acc_{e-1} \\ -1, & \text{if } acc_e < acc_{e-1} \end{cases}. \quad (10)$$

- **Transition.** Once the k value has been updated based on the reward function (i.e., Eq 9), RPM employs the $top - k$ sampling to select a new set of functional groups in the next epoch.

- **Termination.** The reinforcement learning process of RPM will terminate if the variation of k across ten consecutive epochs does not exceed a certain threshold ϵ , meaning that the optimal threshold value k has been found and the $top - k$ informative subgraphs can be selected. The terminal condition is mathematically expressed as follows :

$$\text{Range}(\{k_{e-10}, \dots, k_e\}) \leq \epsilon. \quad (11)$$

Here, RPM leverages Q-learning [51] as the learning strategy to train the Markov Decision Process (MDP). Q-learning is an off-policy reinforcement learning algorithm that aims at identifying the optimal action to be taken in the current state. The objective is achieved by approximating the Bellman optimality equation, which can be formulated as follows:

$$Q^*(s_e, a_e) = \text{reward}(s_e, a_e) + \gamma \cdot \max_{a'} Q^*(s_{e+1}, a'), \quad (12)$$

Herein, $\gamma \in [0, 1]$ signifies the discount factor applied to future rewards. To train our policy, we employ a ϵ -greedy approach to obtain the exploration probability as follows :

$$\pi(a_e | s_e; Q^*) = \begin{cases} \text{random action}, & w.p. \epsilon \\ \arg \max Q^*(s_e, a), & \text{otherwise} \end{cases}. \quad (13)$$

Consequently, the RL agent employs an exploration strategy to find an optimal action at each epoch according to Eq 13.

C. DDIs PREDICTION MODULE

Taking drug d_x and drug d_y as an example, we first utilize RPM to select the $top - k$ important functional groups, and the selected functional groups contained in drug d_x are represented as V_x , in drug d_y is represented as V_y . As shown in Figure 2, given a DDI tuple (d_x, r, d_y) , the DDI prediction is established based on the joint probability of the tuple.

$$P(\widetilde{d_x}, r, d_y) = \sigma \left(\sum_i \sum_j \gamma_{ij} \widehat{h}(i)^{(x)T} M_r \widehat{h}(j)^{(y)} \right), \quad (14)$$

where $\sigma(\cdot)$ represents the Sigmoid function. $\widehat{h}(i)^{(x)}$ and $\widehat{h}(j)^{(y)}$ represent linear transformations of the substructure $h(i)^{(x)}$ and $h(j)^{(y)}$, respectively.

$$\widehat{h}(i)^{(x)} = W_{(x)} \cdot h(i)^{(x)}, \quad i = 1, \dots, |V_x|, \quad (15)$$

$$\widehat{h}(j)^{(y)} = W_{(y)} \cdot h(j)^{(y)}, \quad i = 1, \dots, |V_y|, \quad (16)$$

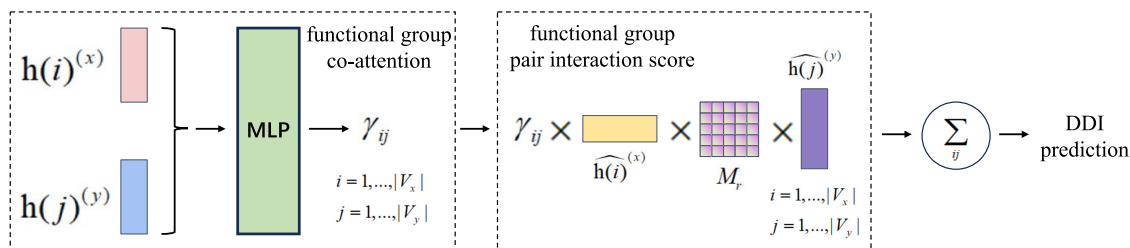


FIGURE 2. Functional group-based deep attention neural network.

where $W_{(x)} \in \mathbb{R}^{b \times b}$ and $W_{(y)} \in \mathbb{R}^{b \times b}$ are learnable transformation matrices.

• The cross-functional group interaction weight between functional groups $h(i)^{(x)}$ of drug d_x and $h(j)^{(y)}$ of d_y is denoted as $\gamma_{ij} \in [0, 1]$, it is formulated as follows:

$$\gamma_{ij} = \text{softmax}\left(\text{MLP}_{\gamma}\left(\mathbf{h}(i)^{(x)} \parallel \mathbf{h}(j)^{(y)}\right)\right), \quad i = 1, \dots, |V_x| \\ j = 1, \dots, |V_y|. \quad (17)$$

D. DDI PREDICTION

• The learnable representation matrix $M_r \in \mathbb{R}^{b \times b}$ corresponds to the interaction type of r . To reduce the number of parameters, we impose a constraint on it that ensures its diagonal matrix:

$$M_r = \text{diag}(m_r), \quad (18)$$

where $\text{diag}(\cdot)$ represents the diagonal matrix, while $m_r \in \mathbb{R}^b$ is a learnable vector that pertains to the specific type of interaction r .

Finally, the final objective function of RPDAnet can be expressed as follows:

$$L = -\sum P_{i,j} \log(\widetilde{P}_{i,j}) + (1 - P_{i,j}) \log(1 - \widetilde{P}_{i,j}), \quad (19)$$

where $P_{i,j}$ is real DDI tuple and $\widetilde{P}_{i,j}$ is the predicted drug tuple.

IV. EXPERIMENTAL ANALYSIS

A. DATASETS AND SETTINGS

In this section, we conduct extensive experiments on real-world datasets to evaluate the performance of our proposed RPDAnet. The experimental datasets encompass DrugBank and Twosides, with their respective details presented below:

DrugBank: The dataset was sourced from FDA/Health Canada drug labels, comprising 191,808 DDI tuples involving 1,706 distinct drugs. Each drug is represented in SMILES notation, and their molecular graphical representations were generated using the RDKit python library [14]. Within this dataset, there are 86 defined interaction types that elucidate the influence of one drug on the metabolism of another. Notably, each DDI tuple in this dataset represents a unique interaction between a pair of drugs, ensuring no duplication with different interactions for the same drug combination.

Twosides: The dataset utilized was proposed by [52] and obtained after filtering the original TWO SIDES side effects data [53]. It comprises 4,649,441 DDI triplets involving 645 drugs, with a total of 1,317 distinct interaction types. Notably, unlike the DrugBank dataset, these interactions primarily pertain to phenotypic effects rather than metabolic interactions. In this context, the interactions are characterized as adverse effects, such as headaches or throat pain. Following the approach in [52], the dataset was further preprocessed to remove interaction types occurring in less than 500 DDI tuples, ensuring that only commonly occurring types were considered. As a result, a dataset comprising 963 interaction types and 4,576,287 DDI tuples is obtained.

For the two public datasets, we select drug data and corresponding SMILES information from the DrugBank dataset and select DDI information from the Twosides dataset. RPDAnet is a GNN-based DDI prediction framework, we need to construct graph G to model the interdependent relationship among drugs. RPDAnet takes the instances of drugs as nodes within graph G , and utilizes the adjacent matrices A to depict their connections (i.e., edges). For example, if there is a DDI between dru_i and dru_j , the element of $A_{i,j}$ is set to 1; otherwise, it is set to 0. The graph G is constructed by continuously injecting nodes and edges. We finally constructed the DDIs graph with 548 drugs and 64 DDIs where each DDI having no less than 150,000 interactions. For the constructed DDIs graph, we randomly select 80% nodes as training data to train our model, 10% nodes as validation data to fine-tune our model, and the rest of the nodes as a test set to test our model.

B. BASELINES

We conducted a comprehensive comparison of our model against state-of-the-art methods to verify the effectiveness of RPDAnet. The baseline methods include:

- **MHCADDI [53]:** MHCADDI employs a co-attention mechanism to effectively integrate joint drug-drug information while conducting representation learning for individual drugs.
- **GAT-DDI [54]:** GAT-DDI serves as the baseline model in our implementation, utilizing graph attention networks (GAT) to generate drug representations directly utilized for DDI prediction.

TABLE 1. 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)

- **MR-GNN [42]:** MR-GNN adopts a strategy where the representations obtained at each graph convolution layer of nodes capture substructures of varying sizes for each drug. These captured representations are then collectively fed into a recurrent neural network to generate a joint representation for a drug pair, facilitating DDI prediction.
- **SSI-DDI [18]:** SSI-DDI considers the hidden features of each node as substructures and subsequently calculates the interactions between these substructures to derive the ultimate DDI prediction.
- **MDF-SA-DDI [55]:** MDF-SA-DDI explores various drug combinations by merging two drugs in four distinct manners. Subsequently, the combined drug feature representation is fed into four distinct drug fusion networks, including a Siamese network, a convolutional neural network, and two auto-encoders. The latent feature fusion is achieved using transformer blocks.
- **DDIMDL [26]:** DDIMDL involves the creation of deep neural network-based sub-models utilizing four distinct drug features: chemical substructures, targets, enzymes, and pathways. Subsequently, a joint deep neural network (DNN) framework is employed to integrate these sub-models and acquire comprehensive cross-modality representations of drug pairs. These representations are then utilized to predict drug-drug interaction (DDI) events.
- **Lee's approach [56]:** Lee utilizes autoencoders and a deep feed-forward network, both trained on structural similarity profiles, Gene Ontology term similarity profiles, and target gene similarity profiles of established drug pairs. This methodology aims to predict the pharmacological effects of drug-drug interactions (DDIs).
- **DeepDDI [57]:** DeepDDI comprises two main components: the structural similarity profile (SSP) generation pipeline and the deep neural network (DNN). In this approach, the two SSPs associated with each input drug structure pair are merged and then input into the DNN to predict the specific interaction type between the drugs.

C. EVALUATION METRICS

In this section, we adopt three key metrics including AUROC, AUPRC, and F1 score to evaluate the performance of RPDAnet. Table 1 shows the confusion matrix for prediction results, which are the basic elements that makeup AUROC, AUPRC, and F1 score.

(1) Recall measures the proportion of actual positive cases that were correctly identified by a classification model. It is important when the cost of false negatives (missed positive cases) is high, and we want to minimize the number of false

negatives:

$$\text{Recall} = \frac{TP}{TP + FN}. \quad (20)$$

(2) Accuracy is the ratio of correctly predicted data points (both true positives and true negatives) to the total number of data points in the dataset. It is a general metric that is useful when the dataset is balanced (roughly equal number of positive and negative cases):

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + FP + TN}. \quad (21)$$

(3) Precision measures the proportion of true positive cases among all the positive predictions made by a classification model. Precision is important when the cost of false positives (incorrectly classified positive cases) is high, and we want to minimize the number of false positives:

$$\text{Precision} = \frac{TP}{TP + FP}. \quad (22)$$

(4) The ROC curve is plotted on a coordinate system that is constructed using the false positive rate (FPR) and the true positive rate (TPR). The area under this curve, known as AUROC, serves as a performance metric for the model. A larger AUROC value indicates a better recognition performance. The definitions of TPR and FPR are as follows:

$$\text{TPR} = \frac{TP}{TP + FN}, \quad (23)$$

$$\text{FPR} = \frac{FP}{FP + TN}. \quad (24)$$

(5) The Precision-Recall Curve (PRC) is constructed by plotting the recall rate (i.e., Recall) against the precision rate (i.e., Precision) in a coordinate system. The area under the PRC curve (AUPRC) quantifies the performance and is utilized as a metric to assess the classifier's effectiveness. F1 score is a metric that takes into account both Precision and Recall simultaneously. Its definition can be expressed as follows:

$$\text{F1} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \quad (25)$$

D. EXPERIMENTAL RESULTS

Table 2 and Table 3 display the prediction results of RPDAnet alongside those of the baseline methods. Remarkably, RPDAnet outperforms all the baseline models, exhibiting the most favorable performance. Specifically, RPDAnet demonstrates at least a 5% improvement in AUROC, a 7% enhancement in AUPRC, and a notable 6% increase in F1 score when compared to the baseline models.

To summarize, among the baseline methods, DeepDDI utilizing a single similarity feature and DNN exhibits the poorest performance. However, DDIMDL and Lee's method show their strengths in AUPRC and F1 scores. The commonality between these two approaches lies in their utilization of four distinct drug similarities as input features

TABLE 2. Performance on the Drugbank dataset.

Model	AUROC	AUPRC	F1	Recall	Accuracy	Precision
MHCADDI	91.16	89.26	85.36	93.21	86.32	84.29
GAT-DDI	89.21	91.56	80.56	92.69	87.13	82.32
MR-GNN	92.87	90.57	89.36	94.35	85.29	83.21
SSI-DDI	86.95	88.26	80.62	90.63	80.31	81.32
MDF-SA-DDI	91.35	86.16	83.92	89.32	80.16	79.31
DDIMDL	93.16	87.31	85.19	91.67	82.39	86.93
Lee	89.24	86.31	83.17	92.54	82.34	79.23
DeepDDI	91.34	87.26	89.35	93.63	83.16	73.29
RPDAnet	98.36	97.54	95.34	99.21	94.36	94.01

TABLE 3. Performance on the twosides dataset.

Model	AUROC	AUPRC	F1	Recall	Accuracy	Precision
MHCADDI	86.24	88.20	83.16	88.31	86.51	84.69
GAT-DDI	83.62	85.49	83.24	85.39	81.36	82.61
MR-GNN	85.00	84.32	80.63	82.67	80.31	81.69
SSI-DDI	85.85	82.71	79.62	81.92	79.31	80.34
MDF-SA-DDI	87.28	83.69	81.64	83.21	85.14	82.49
DDIMDL	84.31	80.62	81.64	84.16	80.31	79.31
Lee	82.62	84.39	76.32	83.29	82.16	80.69
DeepDDI	87.21	86.39	82.46	85.62	82.35	80.19
RPDAnet	97.28	96.13	94.26	95.10	93.15	96.73



FIGURE 3. Comparative experimental results of different encoders.

for the predictor. MDF-SA-DDI creates a multi-structure neural network to refine the similarity features used by DDIMDL and employs the self-attention mechanism to combine the processed features. In contrast to drug similarity, our proposed RPDAnet introduces heterogeneous graphs to capture drug interaction features from extensive biomedical data. Furthermore, the captured information regarding both the graph structure and sequence structure from the inherent chemical makeup of the drug, thereby ensuring comprehensive representations of the drugs. Based on the characteristics of these various features, we have developed three deep neural network channels dedicated to processing the raw features. As a result, our method significantly surpasses other advanced approaches and demonstrates enhanced predictive performance.

TABLE 4. Experimental results for the effectiveness of RPM.

Dataset	Model	AUROC	AUPRC	F1
DrugBank	RPDAnet(without RPM)	93.16	94.68	92.41
	RPDAnet	98.36	97.54	95.34
Twosides	RPDAnet(without RPM)	92.42	91.63	90.24
	RPDAnet	97.28	96.13	94.26

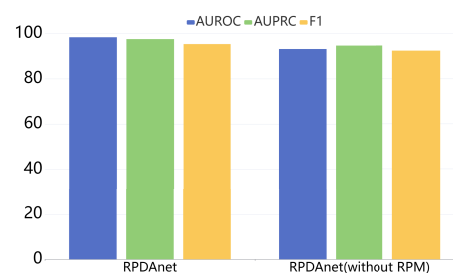


FIGURE 4. Comparative results for the effectiveness of RPM.

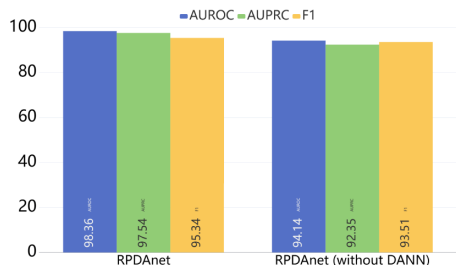
E. ABLATION EXPERIMENT

In this section, we conduct comprehensive ablation experiments to explore the impact of our three innovations on model performance.

First, to verify the effectiveness of CMPNN, we conduct comparative experiments on encoders. Specifically, we replaced the encoders with GAT, GCN, GMPNN, MPNN, and conducted comparative experiments on these models containing different encoders. As shown in Figure 3. the

TABLE 5. Experimental results for the functional group-based DANN.

Dataset	Model	AUROC	AUPRC	F1
DrugBank	RPDAnet (without DANN)	94.14	92.35	93.51
	RPDAnet	98.36	97.54	95.34
Twosides	RPDAnet (without DANN)	91.39	92.17	91.35
	RPDAnet	97.28	96.13	94.26

**FIGURE 5.** Comparative results for the effectiveness of DANN.**TABLE 6.** Case study on the DDI prediction results of RPNAnet.

DrugA	DrugB	Interaction	Evidence
Pyrimethamine	Aliskiren	Sarcoma	[58]
Tolcapone	Pyrimethamine	Breast disorder	[63]
Amlodipine	Atorvastatin	Muscle inflammation	[64]
Omeprazole	Amoxicillin	Renal tubular acidosis	[59]
Aliskiren	Tioconazole	Breast inflammation	[60]

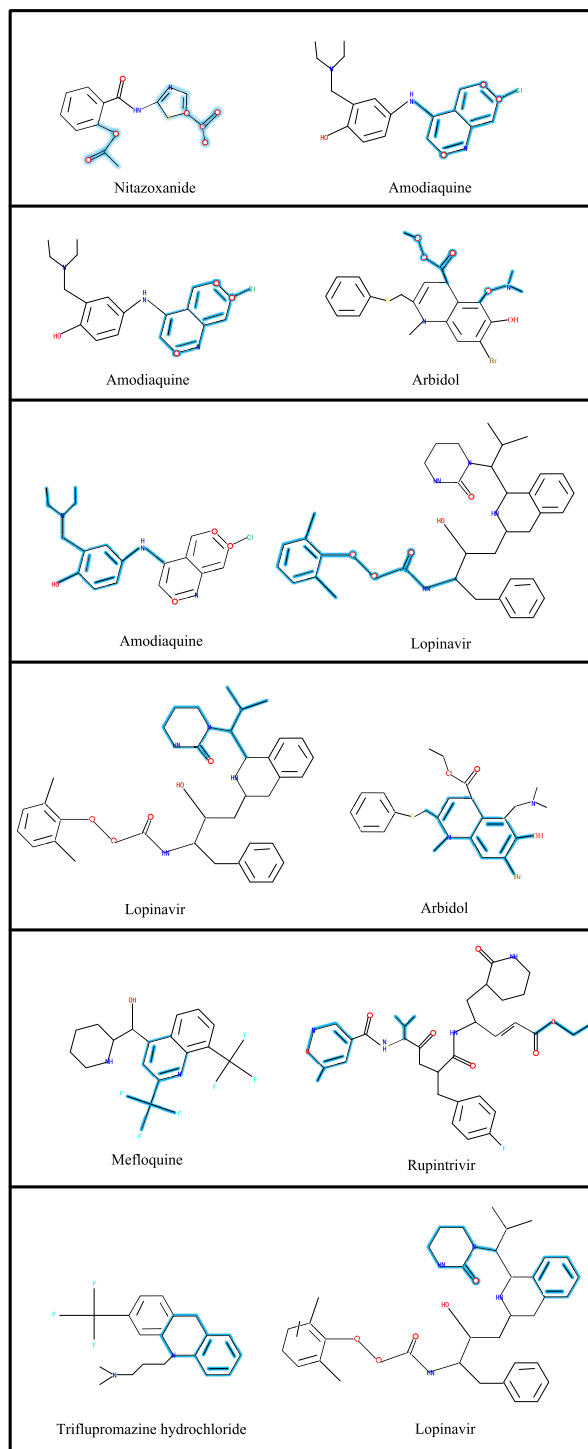
experimental results prove that the quality of embedding is improved because CMPNN strengthens the interaction process between bonds and atoms, which is an important factor in determining RPDAnet outperforms the state-of-the-art models.

Then, we investigate the effectiveness of RPM on RPDAnet. We conducted comparative experiments on RPDAnet with and without the RPM module. The experimental results are shown in Table 4 on DrugBank and Twosides datasets. Figure 4 shows the comparison results with and without RPM on the DrugBank dataset. Experimental results demonstrate that the RPM module improves model generalization and predictive performance by adaptively selecting more informative functional group subgraphs.

Finally, we explore the significance of functional group-based DANN on RPNAnet. We compare the proposed RPDAnet with its variant without the DANN module on DrugBank and Twosides datasets. As demonstrated in Table 5 and Figure 5, the experimental results prove that the functional group-based DANN plays a key role in predicting DDIs due to it is capable of considering the microscopic interactions between functional groups.

V. CASE STUDY

In this section, we further validate the predictive efficacy of our proposed framework for drug-drug interaction (DDI) by conducting case studies in real-world scenarios. Particularly, we randomly selected five pairs of predicted drug-drug interactions (DDIs) generated by our proposed RPNAnet model. Table 6 presents the predicted results alongside the corresponding evidence from previous studies [58], [59],

**FIGURE 6.** Key substructures of headache-causing drug combinations. The most important substructures are shown in blue.

[60], [61], [62]. Upon analysis, we found that all five pairs of DDI predictions were consistent with the evidence reported in these works. For example, Stage et al. reported that the combination of Pyrimethamine and Aliskiren competitively inhibited MATE1 and MATE2-K, leading to the occurrence of Sarcoma as a side effect [58]. Banakh et al. confirmed

that the interaction between Pyrimethamine and Tolcapone resulted in the production of catechol-O-methyltransferase, causing Breast disorder [59]. Parving et al. demonstrated through clinical experiments that Muscle inflammation can be induced by combining Atorvastatin with Amlodipine [60]. The presented evidence showcases the promising and practical predictive performance of our proposed RPNAnet framework in real-world DDI prediction tasks.

VI. VISUALIZATION STUDY

The visualization results of our model are presented in Figure 6, where the contributions of substructures are depicted as a heat map with blue fill on the molecular graph [65]. It can be observed that the drug amodiaquine, nitazoxanide, arbidol, and lopinavir all have potential side effects related to headaches, which is consistent with findings reported in a previous study [65]. Furthermore, when combined with different drugs, the key substructures of amodiaquine drugs remain essentially unchanged and consistently contain phenylacetate ethyl ester (" $C_6H_5CH_2COOCH_3$ ") or its structural components. The visualization results have effectively demonstrated the promising interpretability of our model and proved its effectiveness in real-world application scenarios.

VII. CONCLUSION

This paper proposes RPDAnet, a novel DDIs prediction framework based on reinforced pooling and deep attention mechanisms. Particularly, RPDAnet introduces reinforcement learning to dynamically select informative molecular fragments. RPDAnet then develops a CMPNN module to obtain more robust functional group embedding and infer the potential DDIs. We empirically evaluate the effectiveness of RPDAnet on two real-world datasets. The experimental results prove that our proposed RPDAnet framework exhibits superior performance in DDI prediction tasks.

In future work, we aim to enhance reinforcement learning by incorporating two-person cooperative games with interval uncertainty, thereby effectively selecting more informative functional groups [61]. In addition, we will introduce more advanced optimization algorithms to fine-tune optimal parameters to further improve model performance [66], [67]. Meanwhile, the relevant technologies of image processing can also be introduced to assist the model to achieve optimal performance [62]. Research works [68], [69] also provide new ideas for our future work.

ACKNOWLEDGMENT

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

REFERENCES

- [1] 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.
- [2] M. Wang, "Predicting rich drug–drug interactions via biomedical knowledge graphs and text jointly embedding," 2017, *arXiv:1712.08875*.
- [3] F. Cheng and Z. Zhao, "Machine learning-based prediction of drug–drug interactions by integrating drug phenotypic, therapeutic, chemical, and genomic properties," *J. Amer. Med. Inform. Assoc.*, vol. 21, no. 2, pp. 278–286, Oct. 2014.
- [4] D. S. Wishart, Y. D. Feunang, A. C. Guo, E. J. Lo, A. Marcu, J. R. Grant, T. Sajed, D. Johnson, C. Li, and Z. Sayeeda, "DrugBank 5.0: A major update to the DrugBank database for 2018," *Nucleic Acids Res.*, vol. 46, no. 1, pp. 1074–1082, Jan. 2018.
- [5] G. Shtar, L. Rokach, and B. Shapira, "Detecting drug–drug interactions using artificial neural networks and classic graph similarity measures," *PLoS ONE*, vol. 14, no. 8, Aug. 2019, Art. no. e0219796.
- [6] X. Sun, S. Vilar, and N. P. Tatonetti, "High-throughput methods for combinatorial drug discovery," *Sci. Transl. Med.*, vol. 5, no. 205, Oct. 2013, Art. no. 205rv1.
- [7] 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.
- [8] 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.
- [9] 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.
- [10] S. Vilar, E. Uriarte, L. Santana, T. Lorberbaum, G. Hripscak, 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.
- [11] D. Sridhar, S. Fakhraei, and L. Getoor, "A probabilistic approach for collective similarity-based drug–drug interaction prediction," *Bioinformatics*, vol. 32, no. 20, pp. 3175–3182, Oct. 2016.
- [12] N. Rohani and C. Eslahchi, "Drug–drug interaction predicting by neural network using integrated similarity," *Sci. Rep.*, vol. 9, no. 1, p. 13645, Sep. 2019.
- [13] 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.
- [14] M. W. Harrold and R. M. Zavod, "Basic concepts in medicinal chemistry," Tech. Rep., 2014.
- [15] S. Cao, W. Lu, and Q. Xu, "GraRep: Learning graph representations with global structural information," in *Proc. 24th ACM Int. Conf. Inf. Knowl. Manag.*, Oct. 2015, pp. 891–900.
- [16] 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.
- [17] 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.
- [18] 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.
- [19] 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.
- [20] 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.
- [21] N. Guan, D. Song, and L. Liao, "Knowledge graph embedding with concepts," *Knowl.-Based Syst.*, vol. 164, pp. 38–44, Jan. 2019.
- [22] 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.
- [23] 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. 1, pp. 101–110, Apr. 2018.
- [24] 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.

- [25] 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.
- [26] 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.
- [27] 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.
- [28] 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*, vol. 4, Amsterdam, The Netherlands: Elsevier, 2008, pp. 217–241.
- [29] A. Gottlieb, G. Y. Stein, Y. Oron, E. Ruppim, 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.
- [30] 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.
- [31] 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.
- [32] T. Ma, C. Xiao, J. Zhou, and F. Wang, "Drug similarity integration through attentive multi-view graph auto-encoders," 2018, *arXiv:1804.10850*.
- [33] P. Velickovic, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, "Graph attention networks," 2017, *arXiv:1710.10903*.
- [34] M. Defferrard, X. Bresson, and P. Vandergheynst, "Convolutional neural networks on graphs with fast localized spectral filtering," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 29, 2016, pp. 1–9.
- [35] J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl, "Neural message passing for quantum chemistry," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2017, pp. 1263–1272.
- [36] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," 2016, *arXiv:1609.02907*.
- [37] D. K. Duvenaud, D. Maclaurin, J. Iparraguirre, R. Bombarell, T. Hirzel, A. Aspuru-Guzik, and R. P. Adams, "Convolutional networks on graphs for learning molecular fingerprints," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 28, 2015, pp. 1–9.
- [38] S. Kearnes, K. McCloskey, M. Berndl, V. Pande, and P. Riley, "Molecular graph convolutions: Moving beyond fingerprints," *J. Comput.-Aided Mol. Des.*, vol. 30, no. 8, pp. 595–608, Aug. 2016.
- [39] Z. Wu, B. Ramsundar, E. N. Feinberg, J. Gomes, C. Geniesse, A. S. Pappu, K. Leswing, and V. Pande, "MoleculeNet: A benchmark for molecular machine learning," *Chem. Sci.*, vol. 9, no. 2, pp. 513–530, 2018.
- [40] K. Yang, K. Swanson, W. Jin, C. Coley, P. Eiden, H. Gao, A. Guzman-Perez, T. Hopper, B. Kelley, M. Mathea, A. Palmer, V. Settels, T. Jaakkola, K. Jensen, and R. Barzilay, "Analyzing learned molecular representations for property prediction," *J. Chem. Inf. Model.*, vol. 59, no. 8, pp. 3370–3388, Aug. 2019.
- [41] 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.
- [42] 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*.
- [43] 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.
- [44] J. D. Duke, X. Han, Z. Wang, A. Subhadarshini, S. D. Karnik, X. Li, S. D. Hall, Y. Jin, J. T. Callaghan, and M. J. Overhage, "Literature based drug interaction prediction with clinical assessment using electronic medical records: Novel myopathy associated drug interactions," *PLoS Comput. Biol.*, pp. 178–189, Aug. 2012, Art. no. e1002614.
- [45] H. Wu, Y. Xing, W. Ge, X. Liu, J. Zou, C. Zhou, and J. Liao, "Drug–drug interaction extraction via hybrid neural networks on biomedical literature," *J. Biomed. Informat.*, vol. 106, Jun. 2020, Art. no. 103432.
- [46] N. Liu, C.-B. Chen, and S. Kumara, "Semi-supervised learning algorithm for identifying high-priority drug–drug interactions through adverse event reports," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 1, pp. 57–68, Jan. 2020.
- [47] S. Kwon and S. Yoon, "DeepCCI: End-to-end deep learning for chemical-chemical interaction prediction," in *Proc. 8th ACM Int. Conf. Bioinf., Comput. Biology, Health Informat.*, Aug. 2017, pp. 203–212.
- [48] F. Wang, X. Lei, B. Liao, and F.-X. Wu, "Predicting drug–drug interactions by graph convolutional network with multi-kernel," *Briefings Bioinf.*, vol. 23, no. 1, Jan. 2022, Art. no. bbab511.
- [49] X. Su, Z. You, D. Huang, L. Wang, L. Wong, B. Ji, and B. Zhao, "Biomedical knowledge graph embedding with capsule network for multi-label drug–drug interaction prediction," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 6, pp. 5640–5651, Jun. 2023.
- [50] Y. Dai, C. Guo, W. Guo, and C. Eickhoff, "Drug–drug interaction prediction with Wasserstein adversarial autoencoder-based knowledge graph embeddings," *Briefings Bioinf.*, vol. 22, no. 4, Jul. 2021, Art. no. bbab256.
- [51] K. Huang, T. Fu, W. Gao, Y. Zhao, Y. Roohani, J. Leskovec, C. W. Coley, C. Xiao, J. Sun, and M. Zitnik, "Therapeutics data commons: Machine learning datasets and tasks for drug discovery and development," 2021, *arXiv:2102.09548*.
- [52] M. Zitnik, M. Agrawal, and J. Leskovec, "Modeling polypharmacy side effects with graph convolutional networks," *Bioinformatics*, vol. 34, no. 13, pp. 457–466, Jul. 2018.
- [53] N. P. Tatonetti, P. P. Ye, R. Daneshjou, and R. B. Altman, "Data-driven prediction of drug effects and interactions," *Sci. Transl. Med.*, vol. 4, no. 125, Mar. 2012, Art. no. 125ra31.
- [54] A. Deac, Y.-H. Huang, P. Velickovic, P. Lio, and J. Tang, "Drug–drug adverse effect prediction with graph co-attention," 2019, *arXiv:1905.00534*.
- [55] S. Lin, Y. Wang, L. Zhang, Y. Chu, Y. Liu, Y. Fang, M. Jiang, Q. Wang, B. Zhao, Y. Xiong, and D.-Q. Wei, "MDF-SA-DDI: Predicting drug–drug interaction events based on multi-source drug fusion, multi-source feature fusion and transformer self-attention mechanism," *Briefings Bioinf.*, vol. 23, no. 1, Jan. 2022, Art. no. bbab421.
- [56] G. Lee, C. Park, and J. Ahn, "Novel deep learning model for more accurate prediction of drug–drug interaction effects," *BMC Bioinf.*, vol. 20, no. 1, pp. 1–8, Dec. 2019.
- [57] J. Y. Ryu, H. U. Kim, and S. Y. Lee, "Deep learning improves prediction of drug–drug and drug–food interactions," *Proc. Nat. Acad. Sci. USA*, vol. 115, no. 18, pp. 4304–4311, May 2018.
- [58] 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.
- [59] 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.
- [60] 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, "Cardiorespiratory end points in a trial of aliskiren for type 2 diabetes," *New England J. Med.*, vol. 367, no. 23, pp. 2204–2213, Dec. 2012.
- [61] 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.
- [62] A. Cevik, "Computer-aided diagnosis of Alzheimer's disease and mild cognitive impairment with MARS/CMARS classification using structural MR images," *Middle East Technol. Univ.*, Sep. 2017, pp. 1–114.
- [63] 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.
- [64] 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.
- [65] M. Jozefowicz-Korczynska, A. Pajor, and W. L. Grzelczyk, "The ototoxicity of antimalarial drugs—A state of the art review," *Frontiers Neurol.*, vol. 12, Apr. 2021, Art. no. 661740.
- [66] A. Özmen, E. Kropat, and G.-W. Weber, "Robust optimization in spline regression models for multi-model regulatory networks under polyhedral uncertainty," *Optimization*, vol. 66, no. 12, pp. 2135–2155, Dec. 2017.
- [67] G. Kara, A. Özmen, and G.-W. Weber, "Stability advances in robust portfolio optimization under parallelpiped uncertainty," *Central Eur. J. Oper. Res.*, vol. 27, no. 1, pp. 241–261, Mar. 2019.

- [68] G. Nalcaci, A. Özmen, and G. W. Weber, "Long-term load forecasting: Models based on MARS, ANN and LR methods," *Central Eur. J. Oper. Res.*, vol. 27, no. 4, pp. 1033–1049, Dec. 2019.
- [69] B. Kalayci and V. P. Gazi, "Modelling the mutual interaction of finance and human factor by using machine learning techniques," in *Proc. 8th Int. Conf. Econ. Turkish Econ. Assoc.*, Ankara, Turkey, Sep. 2022, pp. 1–14.



KE ZHAO received the M.S. degree in software engineering from Shandong University, China, in 2006.

He is currently a Lecturer with the College of Information Science and Engineering, Qilu Normal University, China. His current research interests include machine learning, pattern recognition, and intelligent systems.



HAILIANG TANG received the M.A. degree in computer application from Shandong Normal University, China, in 2017.

He is currently a Lecturer with the College of Information Science and Engineering, Qilu Normal University, China. His current research interests include machine learning, neural networks, and data mining.



HAILIN ZHU received the M.A. degree in computer software and theory from Shandong Normal University, China, in 2009.

He is currently a Lecturer with the College of Information Science and Engineering, Qilu Normal University, China. His current research interests include machine learning, pattern recognition, and the Internet of Things.



WENXIAO ZHANG received the M.S. degree in business administration from Shandong Normal University, China, in 2020.

She is currently a Lecturer with the College of Digital Finance, Shandong University of Engineering and Vocational Technology, China. Her current research interests include machine learning, deep learning, pattern recognition, and data mining.

...