



---

*Research article*

## **A novel drug-drug interactions prediction method based on a graph attention network**

**Xian Tan<sup>1</sup>, Shijie Fan<sup>1</sup>, Kaiwen Duan<sup>1</sup>, Mengyue Xu<sup>1</sup>, Jingbo Zhang<sup>1</sup>, Pingping Sun<sup>1,\*</sup> and Zhiqiang Ma<sup>2,\*</sup>**

<sup>1</sup> School of Information Science and Technology, Northeast Normal University, Changchun, China

<sup>2</sup> School of Sciences Changchun Humanities and Sciences College, Changchun, China

\* **Correspondence:** Email: [sunpp567@nenu.edu.cn](mailto:sunpp567@nenu.edu.cn), [mazq@nenu.edu.cn](mailto:mazq@nenu.edu.cn).

**Abstract:** With the increasing need for public health and drug development, combination therapy has become widely used in clinical settings. However, the risk of unanticipated adverse effects and unknown toxicity caused by drug-drug interactions (DDIs) is a serious public health issue for polypharmacy safety. Traditional experimental methods for detecting DDIs are expensive and time-consuming. Therefore, many computational methods have been developed in recent years to predict DDIs with the growing availability of data and advancements in artificial intelligence. In silico methods have proven to be effective in predicting DDIs, but detecting potential interactions, especially for newly discovered drugs without an existing DDI network, remains a challenge. In this study, we propose a predicting method of DDIs named HAG-DDI based on graph attention networks. We consider the differences in mechanisms between DDIs and add learning of semantic-level attention, which can focus on advanced representations of DDIs. By treating interactions as nodes and the presence of the same drug as edges, and constructing small subnetworks during training, we effectively mitigate potential bias issues arising from limited data availability. Our experimental results show that our method achieves an F1-score of 0.952, proving that our model is a viable alternative for DDIs prediction. The codes are available at: <https://github.com/xtnenu/DDIFramework>.

**Keywords:** drug-drug interaction; graph attention network; machine learning; graph embedding; computational biology

---

### **1. Introduction**

Drug-drug interactions (DDIs) are a change in the effect of one drug due to the presence of another drug [1]. It can promote the efficacy or reduce the side effects and affect the drug absorption or produce adverse side effects. With the development of drugs, combination therapy is widely used clinically, and

DDIs play an important part [2]. Therefore, the research of DDIs has great importance for new drugs and clinical pharmacy treatment. The medical research methods of DDIs are very diverse, including *in vitro* experiments, animal experiments and pharmaceutical experiments, as well as the research according to the clinical results. However, the above methods also have limitations which cannot predict DDIs on large scale datasets and it is important development of low-cost and high-efficiency DDIs research methods [3]. Hence, *In silico* methods provide a possibility and can provide certain references for clinical experiments. There are two types of computer experiments for DDI prediction [4]. The first type uses medical literature, databases and clinical records as research objects, and analyzes them using natural language processing [5,6] or data mining methods [7–14]. Deep learning methods have been widely employed in various studies [5, 6, 8, 13]. The second type directly uses drug features to predict whether there is a DDI between two drugs [15–22]. In order to focus on the experimental results of predicting the potential interactions of new drugs with limited information, we will focus on the second type.

Conventional machine learning methods of DDI predictions use data features flexibly, and generally do not require high-level experimental environments. Kastrin et al. [15] took the prediction of DDI as a link prediction problem, and used data from five databases to train five classifiers. Yan et al. [16] developed DDIGIP based on Gaussian interaction profile kernels. Qian et al. [17] developed a gradient boosting-based classifier and demonstrate that targets of adversely DDIs are significantly more likely to have synergistic genetic interactions than non-interacting drug pairs.

The deep learning methods of DDI predictions have higher requirements on computing power of the experimental equipment. Compared with conventional machine learning, deep learning can learn more abstract data representation. Rohani et al. [18] developed a deep learning model based on drug substructure, target, side effect, off-label side effect, pathway, transporter and indication data, making full use of the computing power of deep learning. Ryu et al. [19] developed a deep learning framework that can simultaneously predict DDIs and drug-food interactions. Deng et al. [20] developed an architecture that integrates four deep learning sub-models that learn different features. Liu et al. [21] developed an autoencoder-based deep learning framework that can predict new drugs with unknown interaction relationships.

The network-based methods of DDI predictions compute the graph structure of data. Because many data in biology and medicine exist in the network structure, the network methods can more easily reflect the similarity between data features. Chen et al. [22] applied the Laplace regularized least squares method to the synergistic drug combination to develop the model NLLSS. Tripodi et al. [23] proposed a semantic-reasoning-based approach that can infer DDIs through network computing through biological knowledge bases. Yan et al. [24] proposed IDNDDI, which uses a cosine similarity calculation tool to calculate the similarity of drug features and infer whether a DDI exists. Huang et al. [25] developed a prediction method based on the S-score calculation mechanism.

In recent years, with the advancement of deep learning technology, the integration of deep learning and network analysis methods for DDI research has been on the rise. Karim et al. [26], Wang et al. [27] and Xu et al. [28] have each proposed their own deep network model to address various issues encountered in previous DDI studies. Graph convolutional neural networks and graph attention networks, widely used algorithms in the field of bioinformatics, have also been applied to DDI research. Graph neural networks [29] have been proposed as a powerful method for processing graph representations based on deep learning. The graph structure can effectively represent various complex

network structure data. Although graph neural networks can predict known graph structures, handling unknown graph structures remains challenging. To overcome this limitation, researchers have developed graph attention networks [30]. In DDI research, Nyamabo et al. [31] proposed a graph attention network model based on drug substructures, while Feng et al. [32] proposed a graph attention network model based on chemical molecular graph calculations. Both studies have demonstrated the potential of graph attention networks in DDI research.

Although methods based on drug features have made progress in DDI research and have been confirmed feasible in *in silico* methods, there are still some limitations. First, deep learning methods are mostly trained based on independent samples, and a large amount of data is required to discover the similarity and correlation between samples. Second, the network-based models is lacking a mining ability for advanced representations and some methods using graph structure cannot predict for a drug outside the network. Third, for new drugs, many methods are unable to extract their features and predict their related interactions.

In this study, we propose a novel DDI prediction model based on heterogeneous graph attention networks named HGA-DDI. The HGA-DDI uses DDIs as nodes and the same drugs as edges. To accommodate predictions for new drugs, we use only substructure molecular fingerprints of Pubchem as features. To strengthen the attention of the model to advanced features, we use the node-level attention and the semantic-level attention mechanisms originally used in heterogeneous graph attention networks. Our experimental results show that this new model achieves good performance.

In summary, the major contributions of this work are:

- We develop a better performing graph attention DDI prediction method. This is the first attempt to apply heterogeneous graph attention network algorithm to predict DDIs based on drug molecular fingerprints. Moreover, the performance of the model in this paper is relatively good, which can help with research on new drugs.
- In terms of algorithm innovation, this work successfully applies the method of heterogeneous network to a homogeneous network. Although the graph structure of drug interactions is a homogeneous network, multiple different mechanisms exist between drug interactions, making it critical to pay attention to advanced representations that conventional graph attention network methods cannot. The semantic-level attention mechanism of this method becomes a solution.
- The HGA-DDI only utilizes structural features from drugs, and innovatively treats each interaction as a node and the existence of same drugs as edges. This allows for the construction of a graph structure based on a large amount of existing known data when predicting interactions between new drugs, without compromising the performance of the model.

## 2. Materials and methods

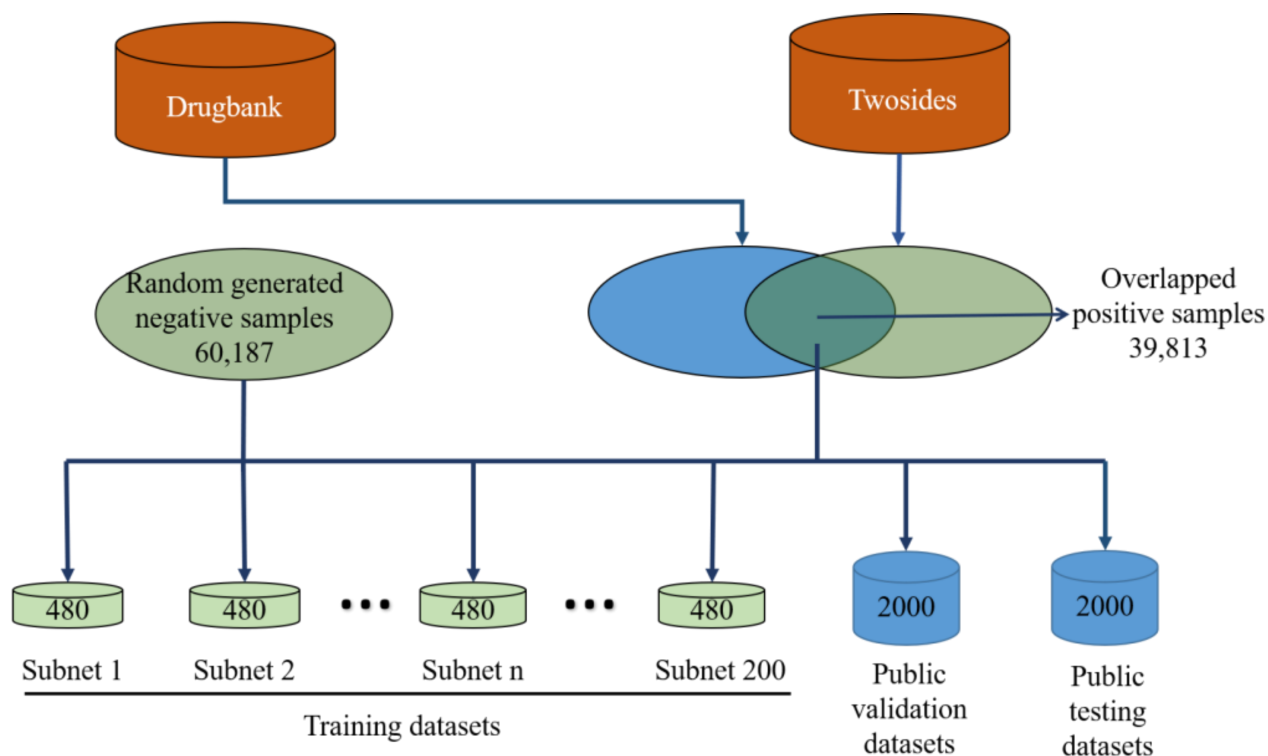
### 2.1. Datasets

Our data is sourced from two databases, Drugbank and Twosides. Drugbank [33] consists of two parts: bioinformatics data and cheminformatics data. It integrates a vast amount of drug biochemical data, target structure and other information for drug research. Twosides [34–36] database collects only DDIs and is a sub-database of adverse DDIs derived from the FAERS (FDA Adverse Event Reporting System) database. We screened 1017 small molecule drugs and 202,304 DDIs that fit the

FDA standards and feature extraction requirements of this work from Drugbank version 5.1.7. Subsequently, we selected 39,813 intersections recorded in Twosides as positive samples. To facilitate experimental grouping and address the imbalance of actual drug effects, we randomly generated 60,187 negative samples that did not appear in both databases, which brings the dataset total to 100,000 by selecting records that do not appear in Drugbank and Twosides.

Molecular fingerprint [37] encodes molecule information into a bit string where each bit represents a molecular feature. In this study, we used molecular fingerprints to represent drugs. Through the Pubchem database, we extracted the substructure fingerprint of drugs as the learning features. The substructure fingerprint has 881 bits, covering a wide range of different substructures and functional groups. To build the graph data structure, we used the DDIs as nodes of the graph, and edges represent whether a drug is involved in two DDIs. For each DDI, we integrate the features of the relationship by comparing the features of each position between the two corresponding drugs. If a position is the same and equal to 1 for both drugs, it is set as 1 in the integrated feature. If the position is the same and equal to 0 for both drugs, it is set as 0 in the integrated feature. Otherwise, it is set as 0.5. Finally, each DDI is encoded as a 881-dimensional vector.

To verify the model on unknown drugs with limited data, we divided the data into 200 random sub-nets for batch training and randomly selected 2% of each sub-net as public validation and testing datasets. The final data distribution is shown in Figure 1:

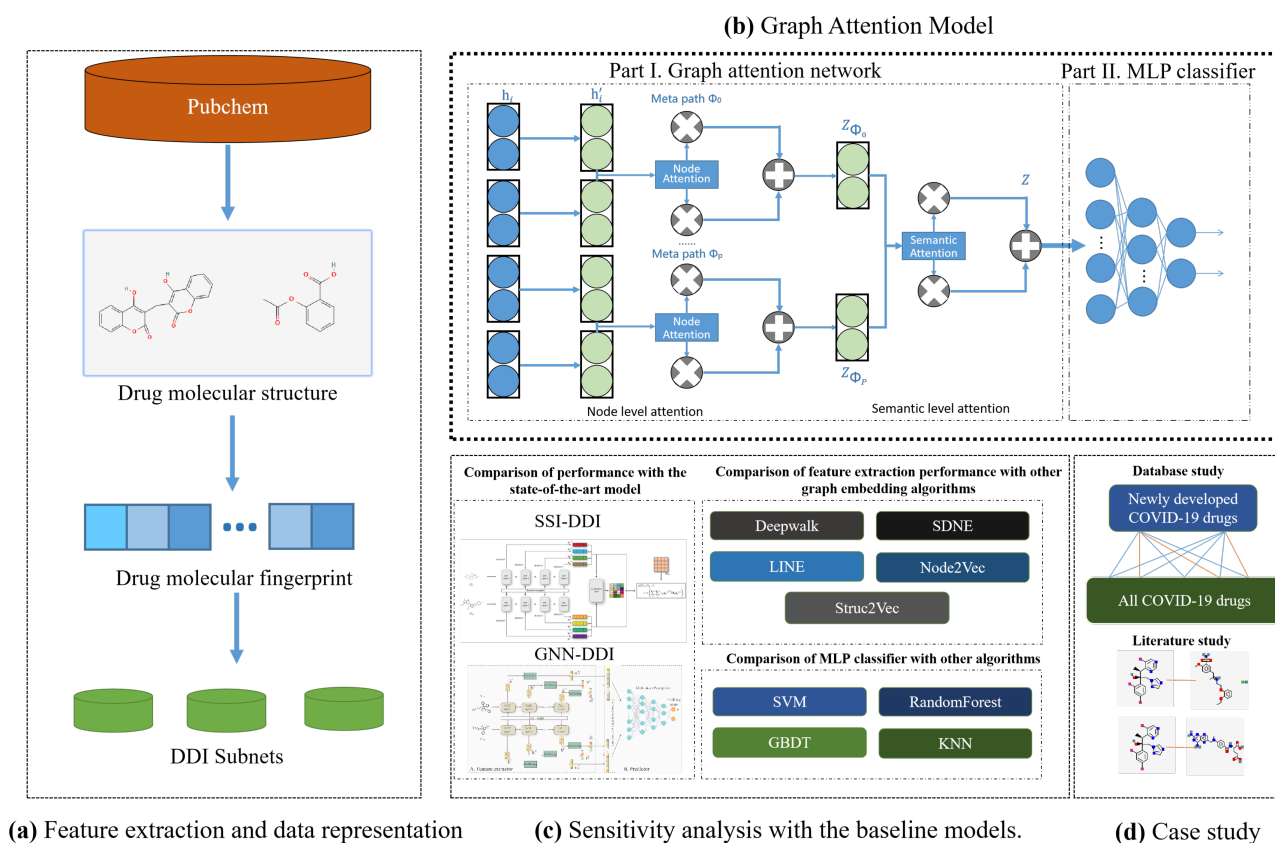


**Figure 1.** Construction of datasets.



## 2.2. Prediction framework

The HGA-DDI includes two parts as shown in Figure 2.



**Figure 2.** Overview of HGA-DDI. (a) The subnets retrieve drug molecular fingerprints from the PubChem database. (b) The heterogeneous graph attention network layer extracts graph embeddings of DDIs, which are then fed into the MLP classifier for classification. (c) We compare the performance of our proposed graph attention network layer with baseline graph embedding algorithms. (d) We conduct an analysis of the model to evaluate its effectiveness.

Because DDIs networks are complex graph structures, in this work, we applied graph attention network [30], which is suitable for graph problems, as the core algorithm. The graph attention network employs an attention mechanism as its main algorithm, eliminating the need for complex calculations involving matrices like Laplace. Instead, it updates node features through the representation of neighboring nodes. In the graph attention network, the learning weights from target nodes to neighbor nodes differ. The adjacency matrix defines the representation of the relevant node, and the calculation of the relationship weights depends on the features of both the node and its neighbor. Specifically, the weight of the neighbor to the node is calculated as follows:

$$e_{ij} = a(Wh_i, Wh_j) \quad (2.1)$$

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\exp(\sum_K e_{ik})} \quad (2.2)$$

where  $e$  is attention coefficient,  $a$  is attention mechanism,  $h_i$  is input feature vector and  $h_j$  is output feature vector,  $\alpha_{ij}$  is the importance from node  $j$  to node  $i$ .

In this work, we added the semantic-level attention which has applied in heterogeneous network by Wang et al. [38]. Meta-path is defined by Wang et al. as different connection modes of nodes, and each meta-path can represent a semantic-level information of nodes. Learning the attention mechanism node embeddings of different meta-paths can provide the different importance. The following formula is represent features from graph attention layer output as the input of semantic level attention:

$$(\beta_{\Phi_0}, \beta_{\Phi_1}, \dots, \beta_{\Phi_p}) = attsem(Z_{\Phi_0}, Z_{\Phi_1}, \dots, Z_{\Phi_p}) \quad (2.3)$$

The importance  $W_{\Phi_i}$  of each meta-path  $\Phi$  is calculated as following:

$$W_{\Phi_i} = \frac{1}{|V|} \sum_{i \in V} q^T \cdot \tan h(Wz_i^\Phi + b) \quad (2.4)$$

where  $V$  is node,  $W$  is the weight matrix,  $b$  is the bias vector and  $q$  is the semantic-level attention vector.

After normalization, the final embedding of model is:

$$Z = \sum_{i=1}^P \beta_{\Phi_i} \cdot Z_{\Phi_i} \quad (2.5)$$

where  $\beta_{\Phi_i}$  is the weight of meta path  $\Phi_i$ ,  $Z_{\Phi_i}$  is the embedding of meta-path  $\Phi_i$ .

The overall process of graph attention network is as follows:

---

**Algorithm 1:** The overall process of graph attention network.

---

**Input:** The graph  $G = (V, E)$ . The node feature  $\{h_i, \forall i \in V\}$ . The meta-path set  $\{\phi_0, \phi_1, \dots, \phi_p\}$ .

**Output:** The final embedding  $Z$ . The node-level attention weight  $\alpha$ . The semantic-level attention weight  $\beta$

**for**  $\Phi_i \in \{\Phi_0, \Phi_1, \dots, \Phi_p\}$  **do**

**for**  $k = 1, 2, \dots, K$  **do**

    Type-specific transformation  $h_i - M_{\phi_i} \cdot h_i$ ;

**for**  $i \in V$  **do**

      Find the meta-path based neighbors  $N_i^\Phi$ ;

**for**  $j \in N_{\phi_i}$  **do**

        Calculate the weight coefficient  $a_{ij}^\phi$ ;

**end for**

      Calculate the semantic-specific node embedding  $Z_i^\Phi = \sigma(\sum_{j \in N_{\phi_i}} \alpha_{ij}^\Phi \cdot h'_j)$ ;

**end for**

    Concatenate the learned embedding from all attention head  $Z_i^\Phi = \parallel_{k=1}^K \sigma(\sum_{j \in N_{\phi_i}} \alpha_{ij}^\Phi \cdot h'_j)$ ;

**end for**

  Calculate the weight of meta-path  $\beta_{\phi_i}$ ;

  Fuse the semantic-specific embedding  $Z_{\phi_i=1}^p \beta_{\phi_i} \cdot Z_\phi$ ;

**end for**

Calculate Cross-Entropy  $L = - \sum_{l \in y_L} Y_l \ln(C \cdot Z_l)$ ;

Back propagation and update parameters in this model;

return  $Z, \alpha, \beta$ .

---

After computing the embeddings of drug nodes, the second part of HGA-DDI is a Multilayer Perceptron (MLP) classifier by using the embeddings as input. To train the MLP classifier, we performed ten-fold cross-validation. In each fold, 10% of the data from the test set was randomly selected as the testing portion for evaluating the performance of the MLP classifier, while the remaining 90% of the data was used for training the MLP classifier. We repeated this process ten times, each time using a different 10% portion for testing and the rest for training. The final performance of the model was determined by averaging the results of these ten folds, taking into account the performance of each individual model.

In this study, we propose two meta-paths: the Interaction Independent Feature Meta-Path (IIFM) and the Interaction-Drug-Interaction Meta-Path (IDIM). IIFM is represented by a diagonal matrix, which indicates that the network uses its own features. The mathematical expression of IIFM is as follows:

$$IIFM = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \dots & 1 \end{bmatrix} \quad (2.6)$$

IDIM refers to the Interaction-Drug-Interaction Matrix. For a given node, if there exists a shared drug between that node and another node, the corresponding element in IDIM is set to 1. Conversely, if there is no shared drug, the element is set to 0. Notably, all diagonal elements in the matrix are set to 1.

### 2.3. Baseline models

To validate the overall effectiveness of the model, we compared two state-of-the-art open-source drug-drug interaction (DDI) models based on graph attention networks. The first model, SSI-DDI [31], has released all its training code, while the second model, GNN-DDI [32], has shared its model architecture. We successfully reproduced both models and compared the results by extracting the SMILES representations of drug molecules from the test set used in this study.

Additionally, in order to test the ability of the graph attention network used in HGA-DDI to extract embeddings, we compared the embedding features extracted by five baseline models based on graph algorithms, and also used MLP classifier for comparison. The introductions of the five algorithms are as follows:

#### 1) Deepwalk

DeepWalk [39] is a network-based language modeling algorithm that utilizes local information obtained from truncated random walks to learn latent representations. It treats walks as the equivalent of sentences and consists of a random walk generator and an update procedure.

#### 2) SDNE

Structural Deep Network Embedding (SDNE) [40] is a semi-supervised deep learning algorithm that incorporates two orders of similarity. The first-order similarity primarily reflects the local characteristics of the graph and is used as supervised information in the supervised component. The second-order similarity mainly reflects the global characteristics of the graph, which is used by the unsupervised component.

#### 3) LINE

Large-scale Information Network Embedding (LINE) [41] optimizes an objective function and proposes an edge-sampling algorithm that improves both the effectiveness and efficiency of stochastic gradient descent.

#### 4) Node2Vec

Node2Vec [42] learns continuous feature representations of networks and maps nodes to low-dimensional feature representations to maximize the likelihood representation of network neighbor nodes. It defines a flexible notion of a node's network neighborhood, designs a biased random walk procedure and learns to explore a variety of neighbor representations.

#### 5) Struc2Vec

Struc2Vec [43] uses a hierarchy to measure node similarity at different scales and constructs a multi-layer graph to encode structural similarities and generate structural context for nodes.

To validate the rationality of the MLP classifier of HGA-DDI, we compared it with several machine learning methods. They are Support Vector Machines (SVM), Random Forests (RF), Gradient Boosting Decision Tree (GBDT) and K-Nearest Neighbor (KNN) Classifier.

## 3. Results and discussion

### 3.1. Metrics

In order to evaluate the performance, we use precision (PRE), sensitivity (SEN), specificity (SPE), accuracy (ACC), F1 score and Matthews correlation coefficient (MCC) as metrics, and their formulas are as follows:

$$PRE = \frac{TP}{TP + FP} \quad (3.1)$$

$$SEN = \frac{TP}{TP + FN} \quad (3.2)$$

$$SPE = \frac{TN}{TN + FP} \quad (3.3)$$

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (3.4)$$

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}} \quad (3.5)$$

$$F1 = 2 \times \frac{PRE \times SEN}{PRE + SEN} \quad (3.6)$$

where TP is the number of the true positive, TN is the number of the true negative, FP is the number of the false positive and FN is the number of the false negative.

### 3.2. Analysis of the comparison with other methods

To validate the effectiveness of the model, we compared the performance of HGA-DDI with SSI-DDI and GNN-DDI models on the same test dataset. The comparison results are in Table 1.

**Table 1.** Comparison with other algorithm models.

Model	ACC	PRE	SEN	SPE	F1	MCC
SSI-DDI	0.931	0.920	0.943	0.918	0.931	0.862
GNN-DDI	0.908	0.913	0.903	0.912	0.908	0.816
HGA-DDI	0.952	0.964	0.939	0.965	0.952	0.904

In comparison to two state-of-the-art graph attention networks, HGA-DDI demonstrated the best performance across all metrics, showcasing the overall computational superiority of the model. However, during the comparison, it was noted that one limitation of HGA-DDI is its inability to predict the types of drug interactions. This will be addressed and improved upon in our future work.

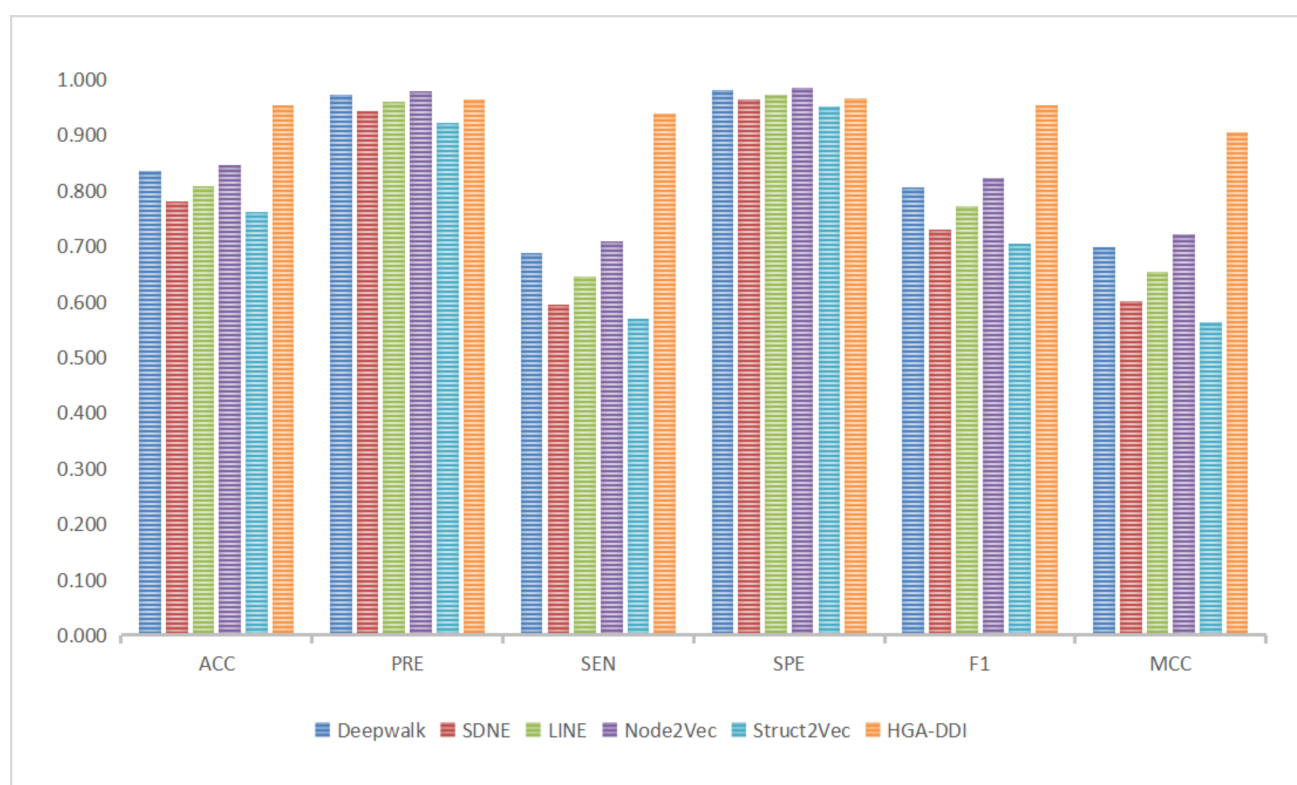
### 3.3. Sensitivity analysis of graph embedding methods

To verify the graph attention network, we conduct five baseline models which used to calculate drug graph embedding on testing datasets of this work. The results are in Table 2.

**Table 2.** Comparison with other graph embedding algorithms.

Model	ACC	PRE	SEN	SPE	F1	MCC
Deepwalk	0.834	0.972	0.688	0.980	0.806	0.698
SDNE	0.779	0.941	0.595	0.963	0.729	0.600
LINE	0.809	0.958	0.645	0.972	0.771	0.653
Node2Vec	0.846	0.978	0.708	0.984	0.821	0.720
Struct2Vec	0.761	0.921	0.570	0.951	0.704	0.564
HGA-DDI	0.952	0.964	0.939	0.965	0.952	0.904

The Table 2 presents the comparison results, demonstrating the superior performance of HGA-DDI in terms of ACC, SEN, F1-score and MCC. These results indicate the competence of HGA-DDI in DDI prediction. While other graph embedding algorithms may exhibit a bias towards encoding data towards positive samples due to data imbalance, this results in HGA-DDI not achieving the best performance in terms of PRE and SPE metrics. However, when considering comprehensive metrics such as F1 and MCC, it becomes evident that HGA-DDI possesses better capability in distinguishing between positive and negative samples. This discriminative ability highlights the advantage of the heterogeneous graph attention network employed by HGA-DDI. Furthermore, the highest ACC value obtained by our model suggests its accuracy in identifying DDI samples and its effective extraction of graph embeddings using the graph attention network. Figure 3 provides a visual representation of the comparison results.



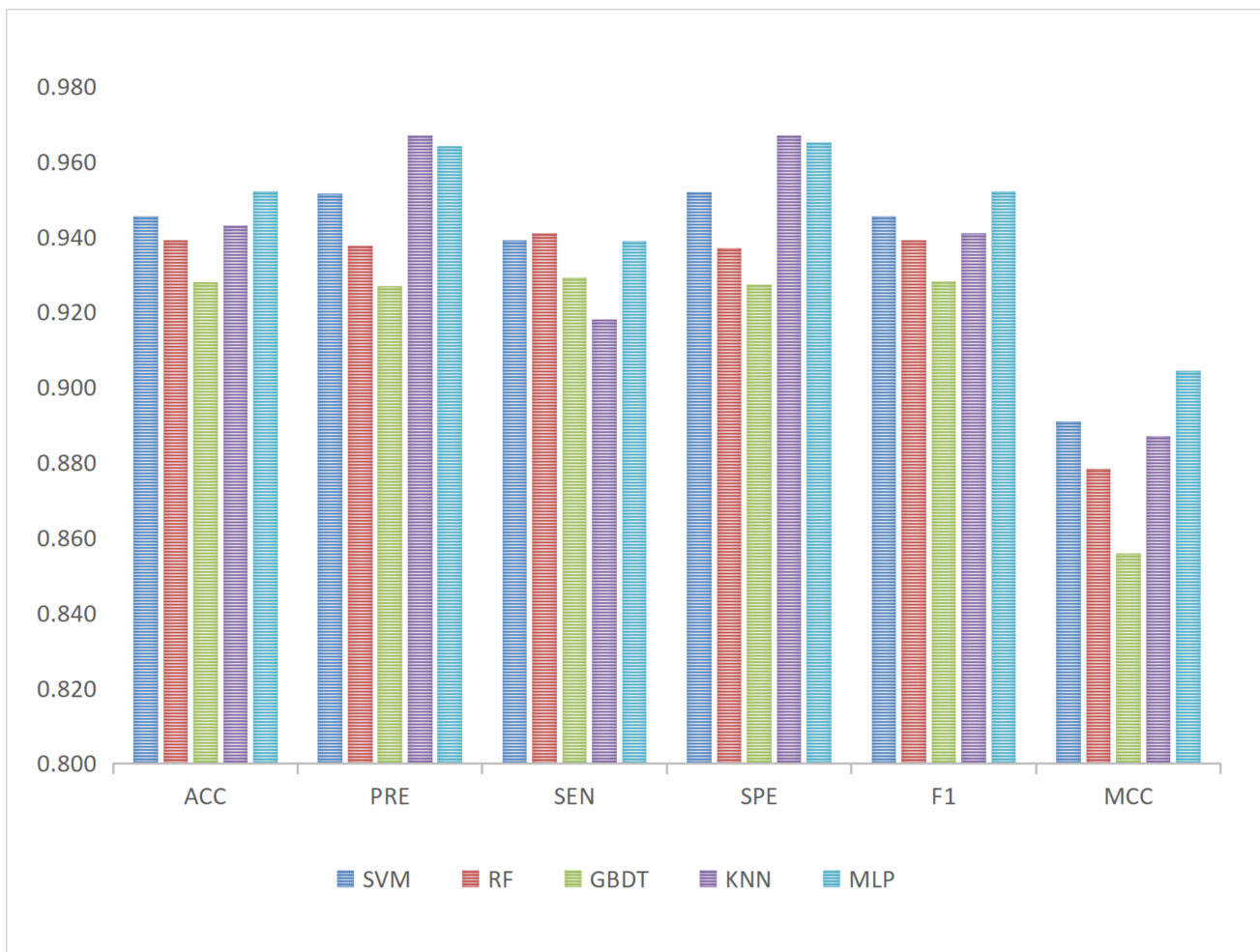
**Figure 3.** Visualization of graph embedding methods comparison results.

#### 3.4. Sensitivity analysis of classifier algorithms

To validate the rationale and performance of the MLP classifier model, we compared it with several machine learning methods. The results are in Table 3.

**Table 3.** Comparison with other classifier algorithms.

Model	ACC	PRE	SEN	SPE	F1	MCC
SVM	0.946	0.951	0.939	0.952	0.945	0.891
RF	0.939	0.938	0.941	0.937	0.939	0.878
GBDT	0.928	0.927	0.929	0.927	0.928	0.856
KNN	0.943	0.967	0.918	0.967	0.941	0.887
MLP	0.952	0.964	0.939	0.965	0.952	0.904

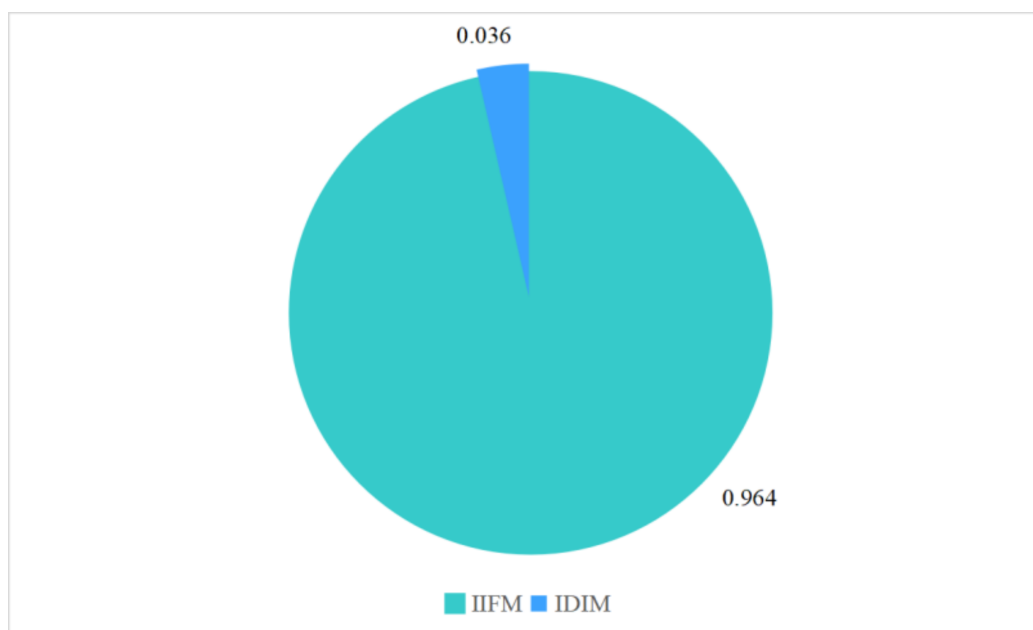
**Figure 4.** Visualization of classifier algorithms comparison results.

In Table 3, similar to the comparison with other graph embedding algorithms, the MLP algorithm did not outperform in all metrics. These slight differences may be attributed to algorithmic errors and the learning tendency of the classifier itself. However, the MLP classifier performed the best in terms of ACC, F1, and MCC, which are comprehensive measures of the model's predictive ability for both positive and negative samples. This aligns with our expectations, and thus we can consider MLP as the

most suitable classifier algorithm for HGA-DDI among the current options. Figure 4 provides a visual representation of the comparison results.

### 3.5. Analysis of meta-paths

In this work, we propose and use two meta-paths, IIFM and IDIM. The weights of the two meta-paths which indicate the learning of the importance assigned by our method are 0.964 and 0.036 respectively. The visualized results as shown in Figure 5:



**Figure 5.** Visualization of classifier algorithms comparison results.

As shown in the Figure 5, the meta-path IDIM is given a higher weight on our training datasets, which means that the method regards IDIM as the most critical meta-path for identifying drug interactions. The experimental results also reflect that IDIM has more effective features than IIFM. It also further confirms the validity of semantic-level attention and the difference in the effectiveness of the meta-paths.

### 3.6. Case study

To verify the ability of this method to predict real data, we conduct database and literature studies as case studies. In the database study, to demonstrate the advancement of the model in considering interaction relationships with shared drugs as edges, we focused on studying newly developed drugs related to COVID-19. We collected a total of 1734 drugs related to Covid-19 from PubChem, differentiating them by whether they were included in PubChem between 2021 and 2022, resulting in 57 new drugs. We predicted a total of 98,718 potential relationships between each drug and all drugs. Using HGA-DDI for prediction, a total of 19,055 interactions were predicted as positive, with 8128 interactions classified as high-confidence samples (predicted probability of being positive greater than 95%).



We statistically analyzed these high-confidence samples to identify the top 20 sensitive new drugs in terms of drug interactions. The results are in Table 4 (as some compounds were not named, they are represented by their PubChem ID and molecular formula):

**Table 4.** The top 20 most sensitive new drugs in terms of drug interactions.

PubChem ID	PubChem inclusion date	Molecular formula	Number of related DDIs
155803731	20210305	C <sub>27</sub> H <sub>18</sub> Cl <sub>2</sub> N <sub>4</sub> O <sub>4</sub>	1734
156599206	20210922	C <sub>96</sub> H <sub>126</sub> N <sub>12</sub> Na <sub>6</sub> O <sub>21-6</sub>	1652
155803730	20210305	C <sub>27</sub> H <sub>19</sub> CIN <sub>4</sub> O <sub>4</sub>	1444
155801622	20210226	C <sub>26</sub> H <sub>30</sub> N <sub>4</sub> O <sub>2</sub> S	1418
155803732	20210305	C <sub>27</sub> H <sub>20</sub> CIN <sub>3</sub> O <sub>2</sub>	1314
155801623	20210226	C <sub>28</sub> H <sub>34</sub> N <sub>4</sub> O <sub>2</sub> S	1208
155294426	20210114	C <sub>8</sub> H <sub>16</sub> N <sub>2</sub> O <sub>3</sub> S	889
155294427	20210114	C <sub>12</sub> H <sub>17</sub> NO <sub>2</sub>	839
155804576	20210316	C <sub>29</sub> H <sub>44</sub> N <sub>4</sub> O <sub>6</sub>	793
155294420	20210114	C <sub>9</sub> H <sub>18</sub> N <sub>2</sub> O <sub>3</sub> S	768
155294419	20210114	C <sub>8</sub> H <sub>14</sub> N <sub>2</sub> O <sub>2</sub> S	723
155801621	20210226	C <sub>29</sub> H <sub>36</sub> N <sub>4</sub> O <sub>2</sub> S	683
156592231	20210908	C <sub>31</sub> H <sub>53</sub> N <sub>3</sub> NaO <sub>49</sub> S <sub>8</sub>	601
155804534	20210316	C <sub>9</sub> H <sub>19</sub> BrN <sub>5</sub> O <sub>4</sub> P	577
155804577	20210316	C <sub>29</sub> H <sub>46</sub> N <sub>4</sub> O <sub>7</sub>	542
155804568	20210316	C <sub>25</sub> H <sub>33</sub> N <sub>3</sub> O <sub>4</sub>	430
162396309	20220119	C <sub>23</sub> H <sub>26</sub> CIN <sub>5</sub> O <sub>3</sub>	359
165360157	20221011	C <sub>54</sub> H <sub>83</sub> N <sub>15</sub> O <sub>21</sub> S <sub>2</sub>	339
155804583	20210316	C <sub>13</sub> H <sub>18</sub> N <sub>4</sub> O	288
155804575	20210316	C <sub>29</sub> H <sub>38</sub> N <sub>4</sub> O <sub>6</sub>	256

Through database research, we have demonstrated the potential of HGA-DDI in predicting interactions for new drugs. Additionally, in the literature study, we collected the reports on DDIs from PubMed in recent three years, and found a total of 36 DDIs, including clinical, pharmaceutical and in vitro experimental methods. Twenty DDIs are not all composed of small molecule drugs, and among 16 DDIs that are composed with two small molecule drugs and have substructure molecular fingerprints on PubChem, only 2 of them not be included by Drugbank. They are interaction between voriconazole and tamsulosin hydrochloride [44] and interaction between voriconazole and methotrexate [45], which are also predicted to be positive samples. The results show that our method has the ability to correctly predict data outside our datasets.

#### 4. Conclusions

Drug-drug interactions have high value for medical and clinical studies, especially drug development. Capturing more and richer comprehensive information about DDIs is one of the key tasks in public health and drug development. In silico methods to predict drug interaction can effectively guide the medical experiment, and modeling DDIs as a graph structure can effectively

analyze the correlation. In this work, we propose an interaction prediction method based on graph attention mechanism, and the learning of semantic-attention mechanism is effectively used in the method. Finally, the prediction performance of this model is better than five comparison models on our testing datasets. Moreover, through the analysis of the meta-paths selection, the importance of the reference neighbor node weight of this problem is verified. Finally, through several testing cases, it demonstrated the availability of our method.

### Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

### Acknowledgments

This work is supported Natural Science Funds of Jilin Province (Grant No. 20200201158JC), the science and technology research project of “13th Five-Year” of the Education Department of Jilin province under (Grant No. JJKH20211297KJ).

### Conflict of interest

The authors declare there is no conflict of interest.

### References

1. K. Baxter, *Stockley's Drug Interactions : A Source Book of Interactions, Their Mechanisms, Clinical Importance and Management*, Pharmaceutica Press, 2010.
2. D. M. Qato, J. Wilder, L. P. Schumm, V. Gillet, G. C. Alexander, Changes in prescription and over-the-counter medication and dietary supplement use among older adults in the united states, 2005 vs 2011, *JAMA Intern. Med.*, **176** (2016), 473–482. <https://doi.org/10.1001/jamainternmed.2015.8581>
3. Y. Chen, T. Ma, X. Yang, J. Wang, B. Song, X. Zeng, Muffin: Multi-scale feature fusion for drug–drug interaction prediction, *Bioinformatics*, **37** (2021), 2651–2658. <https://doi.org/10.1093/bioinformatics/btab169>
4. Y. Qiu, Y. Zhang, Y. Deng, S. Liu, W. Zhang, A comprehensive review of computational methods for drug-drug interaction detection, *IEEE/ACM Trans. Comput. Biol. Bioinf.*, **19** (2022), 1968–1985. <https://doi.org/10.1109/TCBB.2021.3081268>
5. Z. Zhao, Z. Yang, L. Luo, H. Lin, J. Wang, Drug drug interaction extraction from biomedical literature using syntax convolutional neural network, *Bioinformatics*, **32** (2016), 3444–3453. <https://doi.org/10.1093/bioinformatics/btw486>
6. R. Kavuluru, A. Rios, T. Tran, Extracting drug-drug interactions with word and character-level recurrent neural networks, in *2017 IEEE International Conference on Healthcare Informatics*, IEEE, (2017), 5–12.

7. S. Kim, H. Liu, L. Yeganova, W. J. Wilbur, Extracting drug–drug interactions from literature using a rich feature-based linear kernel approach, *J. Biomed. Inf.*, **55** (2015), 23–30. <https://doi.org/10.1016/j.jbi.2015.03.002>
8. I. N. Dewi, S. Dong, J. Hu, Drug-drug interaction relation extraction with deep convolutional neural networks, in *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, IEEE, (2017), 1795–1802.
9. Y. Shen, K. Yuan, Y. Li, B. Tang, M. Yang, N. Du, et al., Drug2vec: Knowledge-aware feature-driven method for drug representation learning, in *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, IEEE, (2018), 757–800.
10. J. S. Almenoff, W. DuMouchel, L. A. Kindman, X. Yang, D. Fram, Disproportionality analysis using empirical bayes data mining: A tool for the evaluation of drug interactions in the post-marketing setting, *Pharmacoepidemiol. Drug Saf.*, **12** (2003), 517–521. <https://doi.org/10.1002/pds.885>
11. G. N. Norén, A. Bate, R. Orre, I. R. Edwards, Extending the methods used to screen the who drug safety database towards analysis of complex associations and improved accuracy for rare events, *Stat. Med.*, **25** (2006), 3740–3757. <https://doi.org/10.1002/sim.2473>
12. A. Suzuki, N. Yuen, K. Ilic, R. T. Miller, M. J. Reese, H. R. Brown, et al., Comedications alter drug-induced liver injury reporting frequency: Data mining in the who vigibase™, *Regul. Toxicol. Pharm.*, **72** (2015), 481–490. <https://doi.org/10.1001/jamaneurol.2015.0365>
13. R. Harpaz, H. S. Chase, C. Friedman, Mining multi-item drug adverse effect associations in spontaneous reporting systems, *BMC Bioinf.*, **11** (2010), 1–8.
14. Y. Noguchi, A. Ueno, M. Otsubo, H. Katsuno, I. Sugita, Y. Kanematsu, et al., A new search method using association rule mining for drug–drug interaction based on spontaneous report system, *Front. Pharmacol.*, **9** (2018), 197.
15. A. Kastrin, P. Ferik, B. Leskošek, Predicting potential drug–drug interactions on topological and semantic similarity features using statistical learning, *PLoS One*, **13** (2018), e0196865. <https://doi.org/10.1371/journal.pone.0196865>
16. C. Yan, G. Duan, Y. Pan, F. X. Wu, J. Wang, Ddigip: Predicting drug–drug interactions based on gaussian interaction profile kernels, *BMC Bioinf.*, **20** (2019), 1–10. <https://doi.org/10.1186/s12859-018-2565-8>
17. S. Qian, S. Liang, H. Yu, Leveraging genetic interactions for adverse drug–drug interaction prediction, *PLoS Comput. Biol.*, **15** (2019), e1007068. <https://doi.org/10.1371/journal.pcbi.1007068>
18. N. Rohani, C. Eslahchi, Drug-drug interaction predicting by neural network using integrated similarity, *Sci. Rep.*, **9** (2019), 13645.
19. J. Y. Ryu, H. U. Kim, S. Y. Lee, Deep learning improves prediction of drug–drug and drug–food interactions, *Proc. Natl. Acad. Sci.*, **115** (2018), E4304–E4311.
20. Y. Deng, X. Xu, Y. Qiu, J. Xia, W. Zhang, S. Liu, A multimodal deep learning framework for predicting drug–drug interaction events, *Bioinformatics*, **36** (2020), 4316–4322. <https://doi.org/10.1093/bioinformatics/btaa501>

21. S. Liu, Z. Huang, Y. Qiu, Y. P. P. Chen, W. Zhang, Structural network embedding using multi-modal deep auto-encoders for predicting drug-drug interactions, in *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, IEEE, (2019), 445–450. <https://doi.org/10.1109/BIBM47256.2019.8983337>
22. X. Chen, B. Ren, M. Chen, Q. Wang, L. Zhang, G. Yan, Nllss: Predicting synergistic drug combinations based on semi-supervised learning, *PLoS Comput. Biol.*, **12** (2016), e1004975. <https://doi.org/10.1371/journal.pcbi.1004975>
23. I. Tripodi, K. B. Cohen, L. Hunter, A semantic knowledge-base approach to drug-drug interaction discovery, in *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, IEEE, (2017), 1123–1126.
24. C. Yan, G. Duan, Y. Zhang, F. X. Wu, Y. Pan, J. Wang, IDNDDI: An integrated drug similarity network method for predicting drug-drug interactions, in *Bioinformatics Research and Applications: 15th International Symposium, ISBRA 2019*, Springer International Publishing, (2019), 89–99.
25. J. Huang, C. Niu, C. D. Green, L. Yang, H. Mei, J. D. J. Han, Systematic prediction of pharmacodynamic drug-drug interactions through protein-protein-interaction network, *PLoS Comput. Biol.*, **9** (2013), e1002998. <https://doi.org/10.1371/journal.pcbi.1002998>
26. R. Karim, M. Cochez, J. Jares, M. Uddin, O. Beyan, S. Decker, Drug-drug interaction prediction based on knowledge graph embeddings and convolutional-lstm network, in *Proceedings of the 10th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics*, ACM, (2019), 113–123.
27. F. Wang, X. Lei, B. Liao, F. X. Wu, Predicting drug–drug interactions by graph convolutional network with multi-kernel, *Briefings Bioinf.*, **23** (2022), bbab511.
28. N. Xu, P. Wang, L. Chen, J. Tao, J. Zhao, Mr-gnn: Multi-resolution and dual graph neural network for predicting structured entity interactions, *arXiv preprint*, (2019), arXiv:1905.09558. <https://doi.org/10.48550/arXiv.1905.09558>
29. F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, G. Monfardini, The graph neural network model, *IEEE Trans. Neural Networks*, **20** (2009), 61–80. <https://doi.org/10.1109/TNN.2008.2005605>
30. P. Velickovi, G. Cucurull, A. Casanova, A. Romero, P. Liò, Y. Bengio, Graph attention networks, *arXiv preprint*, (2017), arXiv:1710.10903. <https://doi.org/10.48550/arXiv.1710.10903>
31. A. K. Nyamabo, H. Yu, J. Y. Shi, Ssi-ddi: Substructure-substructure interactions for drug-drug interaction prediction, *Briefings Bioinf.*, **22** (2021), bbab133.
32. Y. H. Feng, S. W. Zhang, Prediction of drug-drug interaction using an attention-based graph neural network on drug molecular graphs, *Molecules*, **27** (2022), 3004. <https://doi.org/10.3390/molecules27093004>
33. M. Zitnik, B. Zupan, Collective pairwise classification for multi-way analysis of disease and drug data, in *Biocomputing 2016: Proceedings of the Pacific Symposium*, (2016), 81–92.
34. S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, et al., Pubchem in 2021: New data content and improved web interfaces, *Nucleic Acids Res.*, **49** (2021), D1388–D1395.

35. N. P. Tatonetti, P. P. Ye, R. Daneshjou, R. B. Altman, Data-driven prediction of drug effects and interactions, *Sci. Transl. Med.*, **4** (2012), 125ra31.
36. E. Poluzzi, E. Raschi, U. Moretti, F. D. Ponti, Drug-induced torsades de pointes: data mining of the public version of the fda adverse event reporting system (aers), *Pharmacoepidemiol. Drug Saf.*, **18** (2009), 512–518. [https://doi.org/10.1016/S0262-1762\(09\)70172-5](https://doi.org/10.1016/S0262-1762(09)70172-5)
37. N. Salim, J. D. Holliday, P. Willett, Combination of fingerprint-based similarity coefficients using data fusion, *J. Chem. Inf. Comput. Sci.*, **43** (2002), 435–442. <https://doi.org/10.1021/ci025596j>
38. X. Wang, H. Ji, C. Shi, B. Wang, Y. Ye, P. Cui, et al., Heterogeneous graph attention network, *World Wide Web Conf.*, **2019** (2019), 2022–2032.
39. B. Perozzi, R. Al-Rfou, S. Skiena, Deepwalk: Online learning of social representations, in *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, (2014), 701–710.
40. D. Wang, P. Cui, W. Zhu, Structural deep network embedding, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, (2016), 1225–1234.
41. J. Tang, M. Qu, M. Wang, M. Zhang, J. Yan, Q. Mei, Line: Large-scale information network embedding, in *Proceedings of the 24th International Conference on World Wide Web*, (2015), 1067–1077.
42. A. Grover, J. Leskovec, Node2vec: Scalable feature learning for networks, *arXiv preprint*, (2016), arXiv:1607.00653. <https://doi.org/10.48550/arXiv.1607.00653>
43. L. Ribeiro, P. Saverese, D. Figueiredo, Struc2vec: Learning node representations from structural identity, in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, (2017), 385–394.
44. D. J. Bogaert, L. Verlinden, E. Vandecruys, G. Laureys, E. Verhaeghe, T. Bauters, Severe phototoxicity associated with concomitant use of methotrexate and voriconazole, an overlooked drug-drug interaction, *Pediatr. Blood Cancer*, **67** (2020), e28246.
45. C. J. Parramón-Teixidó, A. Pau-Parra, J. Burgos, D. Company, Voriconazole and tamsulosin: A clinically relevant drug–drug interaction, *Enferm. Infecciosas y Microbiol. Clin.*, **39** (2021), 361–363. <https://doi.org/10.1016/j.eimce.2021.05.005>



AIMS Press

©2023 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)