MVIC: Multi-View Information Collaborative Fusion for Drug-Drug Interaction Prediction

Xianxian Zhao[1], Chengxin He[2], and Lei Duan[1(✉)]

[1]School of Computer Science, Sichuan University, Chengdu, China
xianx.zhao@gmail.com, leiduan@scu.edu.cn
[2]National Engineering Research Center for Biomaterials, Sichuan University, Chengdu, China
cxinhe@foxmail.com

Abstract. Drug-drug interaction (DDI) prediction is always a focal point in order to furnish more effective and safer therapy in cases of multiple or complex diseases. We consider this problem from both a functional and a structural solution as drug direct-indirect association information and drug structure information are important and unignorable during this process. Therefore, we propose a method called MVIC (short for **m**ulti-**v**iew **i**nformation **c**ollaborative fusion for drug-drug interaction prediction) which incorporating **drug direct-indirect association** and **molecular structure information** for DDI prediction. MVIC extracts information from two views, i.e., the network view and the molecular structure view. For the former, we construct a drug information network, which then undergoes meta-path-specific representation learning and a designed transformer-like semantic fusion module to obtain the corresponding representations. For the latter, we encode the molecular structure via graph neural network (GNN). In the end, we introduce a multi-view collaborative information fusion module for predicting DDIs. The experiments prove that our method outperforms baselines across all four metrics on three datasets. We also conduct a case study to show the capability of MVIC to predict new prospective DDIs. Our source code is available at https://github.com/scu-kdde/Bioinfo-MVIC-2025.

Keywords: Drug-drug interaction prediction · Drug information network · Molecular structure · Graph neural network.

1 Introduction

Due to the complex or co-existing diseases among people, especially in the elderly, the use of drug combinations is needed [6]. However, the concurrent use of multiple medications increases the risk of unexpected drug-drug interactions (DDIs), which are one of the commonest causes of adverse drug reactions (ADRs) [17]. DDIs are also one of the major reasons for drugs being withdrawn from the market [19]. Previously, DDIs detection through in vitro and in vivo experiments are useful but costly, time-consuming, and labour-intensive [35]. With the development of artificial intelligence, deep learning methods are used to alleviate this situation and perform DDI prediction.

Depending on the information considered, we divide the current state-of-the-art methods for DDI prediction into three categories, *biomedical network-based* method, *molecular structure-based* method, and *hybrid* method. Biomedical network-based methods[12,34] tend to explore the network which consists of biomedical entities as nodes and interactions or associations information as edges. They pay more attention to the properties of entities (manifested as interactions or associations) while neglecting the structure of the drugs to varying degrees. Molecular structure-based methods capture DDIs information from the chemical structure of the drugs. Some works[14,30] attempt to learn the implicit substructure from the 2D structure, i.e., atomic graph converted from SMILES [26], and GNNs are often used during this process. This category is devoted to exploring the relationship between the occurrence of drug interactions and drug structure. Hybrid methods often utilize DDI network and drug structure together. Some practices[18,25,31] unify DDI network and drug structure into a framework and carry out optimization for predicting DDIs.

Although these methods have made significant progress in the field of drug-drug interaction (DDI) prediction, we argue that there are still some limitations that need to be addressed. First, as far as we know, existing methods fail to simultaneously and comprehensively consider drug-direct and drug-indirect association information from the network view, as well as drug structure information from the molecular view. Integrating these three types of information from two views can lead to better DDI predictions. However, most of the current multi-view information fusion methods are insufficient and inefficient, failing to fully leverage the potential of multi-view information and achieve the collaborative fusion [25,28]. Second, in their work, the structure features from the molecular view are hierarchically encoded and used as the initial representations of drugs in the network view during each iteration of the training process. This kind of mode places molecular structure information in a minor position, which may lead to the blurring of drug structural information that contributed to DDI. This processing of structural information is unreasonable. For example, a common cause for DDI occurrence is the altered drug absorption [7]. Among it, tetracycline analogues are susceptible to DDIs with antacids mainly because of their structures [16].

To overcome the aforementioned limitations, we propose an end-to-end method MVIC, which endeavors to deal with DDI prediction from two views, i.e., the network view and the molecular structure view, to incorporate drug direct-indirect association and drug structure information. For the association information, we construct a heterogeneous graph, containing drugs and other biomedical entities. Meta-path is utilized to capture rich semantics among composition connection [20]. Then the meta-path-specific representation learning enables us to observe constructed graph from multiple semantic channels, which makes us gain drug-direct representations firstly. And the transformer-like semantic fusion is designed to acquire the drug-indirect representations. As for drug molecular structure, we use the GIN as an encoder. After the information extraction from two views, we design a information collaborative fusion module. We calculate a

coefficient matrix from different combinations across multiple perspectives (i.e., drug-direct, drug-indirect and molecular structure perspectives) via an attention mechanism to achieve sufficient information fusion. Finally, we make DDI prediction between a drug pair. In summary, the main contributions of this work are listed below:

- We propose a novel method to perform drug-drug interaction prediction called MVIC, which extracts drug direct-indirect association and molecular structure information.
- For the association information, we introduce a transformer-like module under the instruction of drug-direct associations, which achieves drug-indirect semantic fusion. As for multi-view information after extraction, we design a coefficient matrix from different combinations by co-attention to achieve collaborative fusion.
- Extensive experiments are conducted to evaluate the performance of MVIC in DDI prediction on three datasets of different scales. Experiments show the effectiveness of the proposed method.

2 Preliminaries

In this section, some formal definitions of the terms used are given firstly.
Definition 1: Drug Information Network. A drug information network (DIN) is defined as $\mathcal{G} = (\mathcal{V},\ \mathcal{E})$, where $\mathcal{V}$ denotes the biomedical entity set and $\mathcal{E}$ denotes the link set. And $\mathcal{G}$ is also associated with the biomedical entity type mapping function $\phi :\ \mathcal{V} \rightarrow \mathcal{A}$ and the link type mapping function $\psi :\ \mathcal{E} \rightarrow \mathcal{R}$. $\mathcal{A}$ and $\mathcal{R}$ denote the sets of predefined biomedical entity types and link types, respectively, where $|\mathcal{A}| + |\mathcal{R}| >\ 2$.
Definition 2: Meta-path. A meta-path P is defined as a path in the form of $P = A_1 \xrightarrow{R_1} A_2 \xrightarrow{R_2} \cdots \xrightarrow{R_l} A_{l+1}$ (abbreviated as $A_1A_2 \cdots A_{l+1}$), which describes a composite relation $R = R_1 \circ R_2 \circ \cdots \circ R_l$ between node types A_1 and A_{l+1}, where $\circ$ denotes the composition operator on relations.
Definition 3: Meta-path based Neighbors. Given a meta-path P and a node v in the DIN, the meta-path based neighbors N_v^P of node v are defined as the set of nodes that connect with node v via meta-path P. Note that N_v^P includes v itself if P is symmetric.

3 Methodology

In this section, we provide the details of MVIC whose architecture is shown in Figure 1. We sequentially elaborate on the design for extracting information from the network view and the molecular structure view. Subsequently, we present the multi-view collaborative fusion module which integrates drug direct-indirect associations and drug structure information.

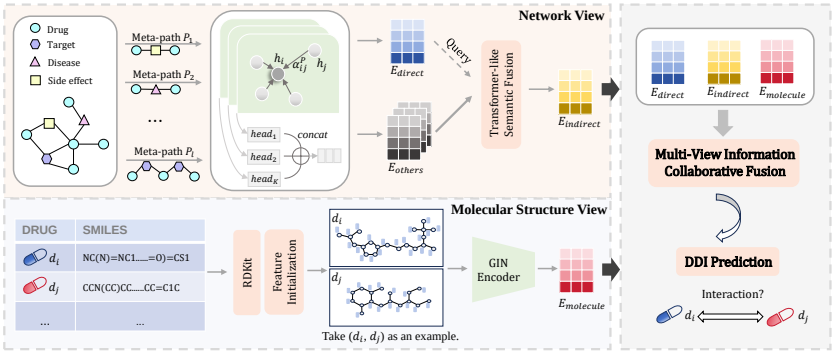

Fig. 1. The overall architecture of MVIC.

3.1 Information Extraction of Network View

Construction of the Drug Information Network. To obtain the biomedical functional information of drugs, we construct a DIN $\mathcal{G}$ which is a drug-related heterogeneous graph, rather than merely capturing knowledge from a homogeneous drug interaction network.

The DIN $\mathcal{G}$ consists of four types of nodes and four edge types. Specifically, $\mathcal{G}$ includes the following node types: drug, target, disease, and side effect. Based on these nodes, We establish the following edge types: drug-drug interactions, drug-target interactions, drug-disease connections and drug-side effect connections. Each type of associations is defined by the corresponding matrix.

Meta-path-specific Representation Learning. Based on the above organization of $\mathcal{G}$, we predefine seven meta-paths $\mathcal{M}$ to explore the relevant structural and semantics information to improve DDI prediction, which are drug-drug, drug-target-drug, drug-side effect-drug, drug-disease-drug, drug-target-drug-target-drug, drug-side effect-drug-side effect-drug, drug-disease-drug-disease-drug. Through the meta-path drug-drug $P^{dd} \in \mathcal{M}$, we derive the drug interaction network, which is a subgraph of $\mathcal{G}$, allowing us to observe $\mathcal{G}$ from the perspective of drug-direct associations. According to the other meta-paths, we can observe $\mathcal{G}$ from the perspective of drug-indirect associations.

Inspired by [24], we apply an attention mechanism [22] on each meta-path-based subgraph to achieve neighbor aggregation. For node i, we aim to determine the significance of each of its neighbors to it. The coefficients can be computed as follows:

$$e_{ij}^P = LeakyReLU(a_P^T \, [Wh_i \, || Wh_j \,]), \tag{1}$$

which represents the importance of node j to node i on meta-path P, a_P is a learnable attention vector for meta-path P and $||$ is the concatenation operator.

After that, we normalize coefficients using the $softmax$ function:

$$\alpha_{ij}^{P} = softmax\left(e_{ij}^{P}\right) = \frac{\exp\left(e_{ij}^{P}\right)}{\sum_{k \in \mathcal{N}_i^P} \exp\left(e_{ik}^{P}\right)}, \tag{2}$$

where N_i^P denotes the neighbors of node i via the meta-path P, and e_{ij}^P is only calculated for node $j \in N_i^P$. The embedding of node i is then updated by aggregating its neighbors with the learned attention vector. Similar to [22], the multi-head attention mechanism is adopted to make the learning process more stable. When K attention heads are used, the output feature representation is as follows:

$$z_i^P = ||_{k=1}^{K} ELU\left(\sum_{j \in N_i^P} \alpha_{ij}^{Pk} h_j\right). \tag{3}$$

In this process, for a drug i, we acquire its meta-path-specific representations, denoted as $\{z_i^P \mid P \in \mathcal{M}\}$, where $\mathcal{M}$ is the meta-path set mentioned before. Then on the meta-path P^{dd}, all drugs features form $\mathbf{E}_{direct}$. Through other meta-paths except P^{dd}, these representations constitute $\mathbf{E}_{others}$.

Transformer-like Semantic Fusion. Although we have obtained the meta-path-specific representations of every drug, each channel of $\mathbf{E}_{others}$ only reflects a drug indirectly from one aspect. To learn more comprehensive drug-indirect representations, semantic fusion is conducted.

In general, some previous studies [21] build heterogeneous graph, extract semantics based on the defined meta-paths, and achieve semantics fusion to perform DDI prediction. However, they do not notice the characteristic of DDI prediction problem. Drug-direct representations play a primary role while drug-indirect information holds a secondary position. In that case, we propose a transformer-like semantic fusion module instructed by $\mathbf{E}_{direct}$ to fuse $\mathbf{E}_{others}$ automatically. It can be formulated as:

$$Q = Q_0 W_Q,\ K = K_0 W_K,\ V = V_0 W_V, \tag{4}$$

$$Att\left(Q,\ K,\ V\right) = softmax\left(\frac{QK^T}{\sqrt{d}}\right)V, \tag{5}$$

$$O = \beta \cdot Att\left(Q,\ K,\ V\right) + Q_0, \tag{6}$$

where Q, K, V denote the queries, keys, values matrices respectively, d denotes the dimension of keys and W_Q, W_K, W_V, β are learnable parameters. Among them, $\mathbf{E}_{direct}$ acts as Q_0, while $\mathbf{E}_{others}$ is taken as K_0 and V_0. This is because we try to identify the useful parts of $\mathbf{E}_{others}$ for the queries. $Att(\cdot)$ denotes the vanilla attention [22]. The output O is the element-wise product of β and $Att(\cdot)$, in addition, plus a residual connection.

With this designed module, the main messages $\mathbf{E}_{indirect}$ related to drug interactions are screened and extracted from $\mathbf{E}_{others}$.

Table 1. Feature components based on chemical knowledge.

Feature Component	# Dimensions
Atomic symbol	44
Degree of the atom	1
Implicit valence	1
Formal charge	1
Number of radical electrons	1
Hybridization	5
Whether or not the atom is aromatic	1
Total number of hydrogens on the atom	1

3.2 Information Extraction of Molecular Structure View

Duvenaud et al. [3] and Kearnes et al. [8] show that representations derived from the molecular graphs are more useful and competitive than handcrafted molecular features such as molecular fingerprints in the follow-up tasks [15]. From SMILES [26], the textual sequences, drugs are converted into 2D (2-dimensional) molecular graph forms by open-source library RDKit[1]. In the molecular graph, the atom serves as a node and the chemical bond serves as an edge.

Relying on some chemical knowledge, the feature for each atom can be initialized as a 55-dimensional vector. The components of the vector $h_v \in \mathbb{R}^{1\times 55}$ are listed in Table 1.

After feature initialization, we leverage GNN to encode molecular graph. GNN mainly adopts the message passing mechanism [5], where the representation of the center node is iteratively updated by aggregating information of its neighbors. The expression can be formalized as:

$$m_v^{(k)} = AGGREGATE^{(k)}\left(\left\{h_u^{(k-1)} : u \in N_v\right\}\right), \tag{7}$$

$$h_v^{(k)} = COMBINE^{(k)}\left(h_v^{(k-1)},\ m_v^{(k)}\right), \tag{8}$$

where $h_v^{(k)}$ denotes the embedding of atom v at the k-th layer. $AGGREGATE^{(k)}(\cdot)$ and $COMBINE^{(k)}(\cdot)$ differ in GNN variants.

Among them, graph isomorphism network (GIN) [29] is a simple neural architecture and as powerful as the Weisfeiler-Lehman test. In this work, we use 3-layer GIN to model drug molecular graphs. GIN updates node representations as

$$h_v^{(k)} = MLP^{(k)}\left(\left(1+\epsilon^{(k)}\right)\cdot h_v^{(k-1)} + \sum_{u\in N_v} h_u^{(k-1)}\right), \tag{9}$$

where MLP denotes multi-layer perceptron, and ϵ could be a learnable parameter or a fixed scalar.

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

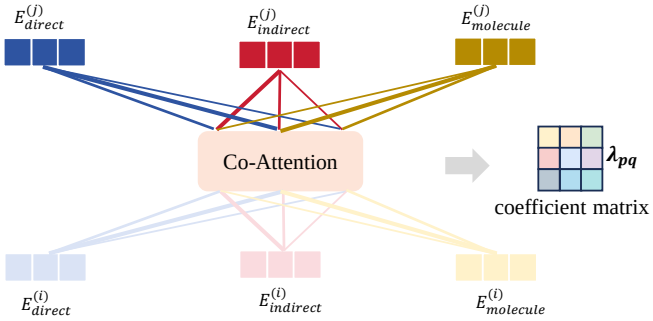

Fig. 2. The module architecture of multi-view information collaborative fusion.

In order to get entire representation of each molecular graph, the function $READOUT$, is used to calculate the sum of all atom embeddings. Given a drug D, the process can be formulated as:

$$h_D = READOUT\left(\left\{h_v^{(L)} \middle| v \in D\right\}\right). \tag{10}$$

Deriving all drugs representations from Equation (10), we acquire the representations $\mathbf{E}_{molecule}$ of the molecular graph view.

3.3 Collaborative Fusion for Multi-View Information

After the process of DIN and molecular structure, we get $\mathbf{E}_{direct}$, $\mathbf{E}_{indirect}$ and $\mathbf{E}_{molecule}$ depicting the drug properties related to DDIs from two views. It is noteworthy that the multi-view information flowing have a complementarity and similarity to some extent. As shown in Figure 2, we propose a collaborative fusion module via co-attention [13], which has been widely applied to visual question answering (VQA), to fuse these representations.

Aiming to fuse them and make final DDI prediction between drug i and drug j, we obtain

$$\mathbf{E}^{(i)} = \left\{\mathbf{E}^{(i)}_{direct},\ \mathbf{E}^{(i)}_{indirect},\ \mathbf{E}^{(i)}_{molecule}\right\}, \tag{11}$$

$$\mathbf{E}^{(j)} = \left\{\mathbf{E}^{(j)}_{direct},\ \mathbf{E}^{(j)}_{indiect},\ \mathbf{E}^{(j)}_{molecule}\right\}, \tag{12}$$

respectively. We calculate a coefficient matrix using different combinations. And the co-attention mechanism is used to explore the importance among drug direct-indirect association and molecular structure information, which can be formulated as follows:

$$\lambda_{pq} = b^T \tanh\left(W_i\mathbf{E}^{(i)}_p + W_j\mathbf{E}^{(j)}_q\right)\ \ p,\ q \in \{direct,\ indirect,\ molecule\}, \tag{13}$$

where b is a learnable weight vector, W_i and W_j are learnable weight matrices. In that case, nine importance coefficients are calculated and weight matrices vary from combination to combination.

Table 2. Statistics of integrated multi-source data.

	Type	Dataset			Resource
		ZhangDDI	*ChCh-Miner*	*DeepDDI*	
# Nodes	Target	1,968	1,686	1,968	UniProtKB [1]
	Disease	7,205	7,205	7,205	CTD [2]
	Side effect	3,935	3,935	3,935	SIDER [9]
# Edges	Drug-Target	2,895	3,418	6,314	DrugBank [27]
	Drug-Disease	237,807	278,325	447,187	CTD [2]
	Drug-Side effect	67,783	75,581	100,425	SIDER [9]

3.4 Drug-Drug Interaction Prediction

By accumulating the interaction importance coefficients, we are able to detect the probability of the interaction occurring between two drugs. And it can be formalized as:

$$F\left(d_i,\ d_j\right) = \sigma\left(\sum_p \sum_q \lambda_{pq} \mathbf{E}_p^{(i)^T} \mathbf{E}_q^{(j)}\right), \tag{14}$$

where σ is the sigmoid function. For each pair of drugs, cross-information interactions weighted and aggregated contribute to the final DDI prediction.

The overall model could be trained in an end-to-end way. Considering the fact that DDI prediction can be regarded as a binary classification task, we employ the binary cross-entropy loss function:

$$\mathcal{L} = -\frac{1}{N} \sum_{i=1}^{N} \left[y \log(\hat{y}) + (1-y) \log(1-\hat{y})\right], \tag{15}$$

where N is the total number of training samples, y denotes the label of training samples and $\hat{y}$ is computed by Equation (15).

4 Experiments

In this section, we first give a description of the datasets, evaluation metrics, experiment settings, and baselines to compare with. And then we make comparisons with baselines. Finally, we give a comprehensive analysis of MVIC's performance under different experiment conditions.

4.1 Datasets

To verify the scalability and robustness of MVIC, the benchmark datasets of this study are *ZhangDDI*, *ChCh-Miner* and *DeepDDI*, which are provided by Wang et al. [25]. They are small-scale, medium-scale and large-scale, respectively. During the preprocessing, some abnormal data items are removed, which are nonstandard SMILES strings and cannot be converted to graphs. The statistics of the datasets after preprocessing are as follows: (1) *ZhangDDI* [32]: This dataset

Table 3. The hyper-parameters settings of experiments.

Hyper-parameter	Learning rate	Weight decay	Embedding dimension	# Heads
ZhangDDI	0.015	1e-4	128	4
ChCh-Miner	0.03	1e-6	128	8
DeepDDI	0.07	1e-6	128	8

is of small-scale, and it contains 544 drugs and 45,720 pairwise DDIs. (2) *ChCh-Miner*[2]: This dataset is of medium-scale, and it contains 997 drugs and 21,486 pairwise DDIs. (3) *DeepDDI* [19]: This dataset is of large-scale, and it contains 1,704 drugs and 191,870 pairwise DDIs.

Like [31], we split train-validation-test in 6:2:2 ratio for every dataset. Meanwhile, we derive the integrated multi-source dataset [1,2,9,27] composed of different biomedical entities from [10]. It is worth noting that, in experiments on each DDI dataset, we only input the drug-specific data which comes from [10] corresponding to the particular dataset. The statistics are shown in Table 2. To prevent label leakage, just the drug-drug interaction pairs from the training data are used to construct the network for validation and prediction.

4.2 Evaluation Metrics and Experiment Settings

To evaluate the performance of our method MVIC, four common metrics are adopted in the experiments, including the area under the receiver operating characteristic curve (AUROC), average precision (AP), accuracy (ACC) and F1-score (F1). We report the mean values of these four metrics over five repetitions.

We train the model with the Adam optimizer. The implement of our method is based on Pytorch (version 2.1.0) and DGL (version 1.1.2). We initialize our method via Xavier initialization. All experiments are conducted on a NVIDIA GeForce RTX 3090 GPU. The hyper-parameter settings for the experiments are shown in Table 3.

4.3 Baselines

We prove the effectiveness of MVIC by comparison with several baselines, which can be briefly introduced as follows:

CSGNN [33] levarages a deep mix-hop GNN undertaking the molecular interaction prediction task and a contrastive GNN to implement the self-supervised learning task.

DeepDDS [23] proposes a deep learning model based on GNN and attention mechanism to predict the synergistic effect of drug combinations.

MIRACLE [25] proposes a multi-view method for DDI prediction, which leverages 2D structures of drugs and DDI network.

HTCL-DDI [31] utilizes the information of the molecular view. Meanwhile, structure and semantics features are captured based on DDI network.

[2] `https://snap.stanford.edu/biodata/datasets/10001/10001-ChCh-Miner.html`

Table 4. Performance comparison with baselines on three datasets.

Dataset	Metric	CSGNN	DeepDDS	MIRACLE	HTCL-DDI	SSI-DDI	DSN-DDI	MUSE	**MVIC**
ZhangDDI	AUROC	0.9171	0.9320	0.9874	0.9864	0.9314	0.9113	0.9797	**0.9908**
	AP	0.8902	0.9208	0.9630	0.9689	0.9209	0.8642	0.9467	**0.9755**
	ACC	0.8414	0.8563	0.9557	0.9519	0.8535	0.8665	0.9068	**0.9610**
	f1	0.8360	0.8279	0.9074	0.8957	0.8196	0.8768	0.7433	**0.9127**
ChCh-Miner	AUROC	0.9768	0.9710	0.9620	0.9761	0.9809	0.9669	0.9943	**0.9982**
	AP	0.9756	0.9851	0.9950	0.9960	0.9897	0.9634	0.9993	**0.9997**
	ACC	0.9254	0.9038	0.9077	0.9473	0.9219	0.8889	0.8362	**0.9867**
	F1	0.9247	0.9221	0.9455	0.9697	0.9398	0.8812	0.8978	**0.9925**
DeepDDI	AUROC	0.9401	0.9438	0.9276	0.9449	0.9179	0.9322	0.9669	**0.9969**
	AP	0.9417	0.9568	0.9677	0.9741	0.9347	0.9287	0.9887	**0.9989**
	ACC	0.8633	0.8887	0.9033	0.9208	0.8538	0.8541	0.7041	**0.9828**
	F1	0.8601	0.9127	0.9354	0.9478	0.8823	0.8560	0.7542	**0.9886**

SSI-DDI [15] adopts 4-layer graph attention network to derive implicit substructures of a drug and uses a co-attention for DDI prediction.

DSN-DDI [11] learns drug substructures from the single drug and the drug pair in each encoder.

MUSE [18] integrates drugs structure and DDI network into a unified multi-scale framework, using a variant expectation maximization to alleviate the optimization imbalance.

4.4 Comparison with Baselines

The performance comparison results with baselines on three benchmark datasets are shown in Table 4. Some baselines results are implemented and provided by [31]. The best performance on each dataset has been highlighted in **bold** and the runner-up results are underlined.

The results demonstrate that MVIC outperforms all the other baselines under four metrics among three datasets. Specifically, our method achieves AUROC scores of 0.9908, 0.9982, 0.9969, AP scores of 0.9755, 0.9997, 0.9989, ACC sores of 0.9610, 0.9867, 0.9828 and F1 scores of 0.9127, 0.9925, 0.9886 among all datasets. Next, we report the improvements that our method achieves over the strongest competitors across three datasets. Compared with MIRACLE, we make improvements of 0.34%, 1.25%, 0.53%, 0.53% on the *ZhangDDI* dataset. As for the *ChCh-Miner* dataset, we surpass HTCL-DDI by 2.21%, 0.47%, 3.94%, 2.28% under AUROC, AP, ACC, F1. Although MUSE achieves decent scores under AUROC and AP, MUSE significantly underperforms in terms of ACC and F1 metrics. On the *DeepDDI* dataset, we make improvements of 3.00%, 1.02% compared with MUSE under AUROC and AP, and 6.20%, 4.08% compared with HTCL-DDI under ACC and F1.

Moreover, we obtain the following observations: (1) Compared with CSGNN, DeepDDS, SSI-DDI, and DSN-DDI merely considering the network view or the molecule view individually, our method MVIC exhibits better performance. (2) As for the comparison with multi-view methods MIRACLE, HTCL-DDI and MUSE, it is likely due to our consideration of drug-indirect information and the

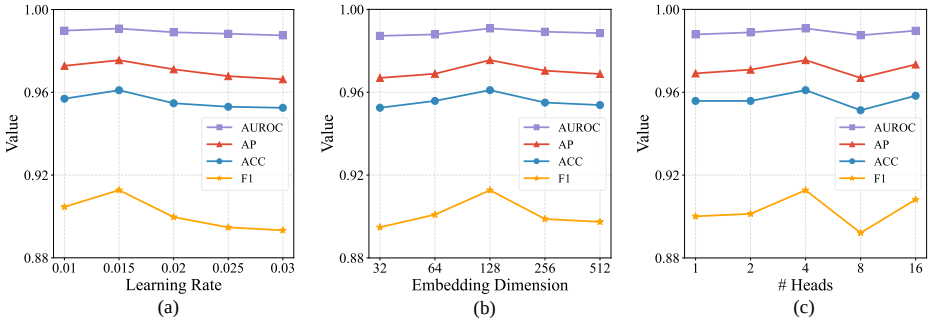

Fig. 3. Parameter sensitivity analysis on the *ZhangDDI* dataset.

designed collaborative fusion module that our method MVIC surpasses all of them.

4.5 Parameter Sensitivity Analysis

We conduct a parameter sensitivity analysis on the *ZhangDDI* dataset to explore the impact of key parameters on the performance of MVIC. Due to the relatively simple and straightforward impact of weight decay on our model, we prioritize focusing on hyper-parameters that have a more significant impact. Specifically, we only investigate the effects of learning rate, embedding dimension, and the number of heads in the optimization process, without conducting further analysis on weight decay. Figure 3 illustrates the results obtained by varying a single parameter while keeping the other parameters fixed.

Effect of the learning rate. We vary the learning rate to 0.01, 0.015, 0.02, 0.025 and 0.03. Results in Figure 3 (a) indicate that when the learning rate is set to 0.015, MVIC achieves the best performance. Afterwards, the scores of MVIC decreases as the learning rate increases. Therefore, we set 0.015 as the default learning rate on the *ZhangDDI* dataset.

Effect of the dimension of embeddings. We vary the dimension of embeddings from 32 to 512. As shown in Figure 3 (b), MVIC exhibits its robustness to changes in dimensionality. Initially, as dimensionality increases, the embeddings capture more information, achieving the optimal performance at 128 dimensions. After this point, further increasing the dimensionality may introduce redundant information, leading to a decrease in performance.

Effect of the number of heads. We vary the number of heads from 1 to 16 to confirm the impact of multi-head attention. The results are presented in Figure 3 (c). Note that the dimension of embeddings is kept in 128 during this process. The inclusion of multi-head mechanism does not always improve performance. As can be seen in Figure 3 (c), when the number of heads is 8, the value is even lower than when the number of heads is 1, which can also be regarded as not introducing the multi-head mechanism. Obviously, when the number of heads is 4, MVIC performs the best.

Table 5. The experimental results of ablation study on three datasets.

Dataset	Metric	MVIC-w/o mol	MVIC-w/o ind	MVIC-w/o asso	MVIC-w/o cfu	MVIC-w/o se	**MVIC**
ZhangDDI	AUROC	0.9832	0.9902	0.9642	0.9857	0.9863	**0.9908**
	AP	0.9584	0.9735	0.9131	0.9635	0.9655	**0.9755**
	ACC	0.9471	0.9580	0.9079	0.9472	0.9524	**0.9610**
	F1	0.8807	0.9073	0.8088	0.8838	0.8927	**0.9127**
ChCh-Miner	AUROC	0.9969	0.9967	0.9824	**0.9987**	0.9945	0.9982
	AP	0.9996	0.9995	0.9977	**0.9998**	0.9992	0.9997
	ACC	0.9693	0.9832	0.9423	0.9755	0.9796	**0.9867**
	F1	0.9823	0.9905	0.9667	0.9859	0.9885	**0.9925**
DeepDDI	AUROC	0.9946	0.9942	0.9247	0.9312	0.9919	**0.9969**
	AP	0.9980	0.9977	0.9687	0.9642	0.9969	**0.9989**
	ACC	0.9743	0.9763	0.8783	0.9356	0.9715	**0.9828**
	F1	0.9830	0.9843	0.9185	0.9590	0.9812	**0.9886**

4.6 Ablation Study

To verify the contribution of each component to MVIC, we design the following variants of it: (1) MVIC-w/o mol: MVIC without considering molecular structure view; (2) MVIC-w/o ind: MVIC without considering drug-indirect association information; (3) MVIC-w/o asso: MVIC without considering drug-indirect and drug-direct information; (4) MVIC-w/o cfu: MVIC removes designed multi-view information collaborative fusion module and then the drug-direct, drug-indirect, and molecular representations, after concatenation, pass through a MLP module; (5) MVIC-w/o se: MVIC does not consider the drug-direct and the drug-indirect association separately and removes the transformer-like semantic fusion module based on that. We do not set a variant, MVIC without considering drug-direct association, because drug-indirect association relies on the guidance of drug-direct association.

The results are reported in Table 5. The effects of all variants, except for MVIC-w/o cfu on the *ChCh-Miner* dataset, have decreased. This demonstrates the effectiveness of MVIC's designed architecture. As for the marginal growth of MVIC-w/o cfu in AUROC and AP on the *ChCh-Miner* dataset, this perhaps can attributed to the sparsity of the subgraphs defined by the meta-path P^{dd} is so high that concatenation preserves information by enlarging the dimensionality. It is worth noting that the designed information fusion module still makes an improvement in ACC and F1 on this dataset.

4.7 Performance under Different Positive-to-Negative Ratios

Empirically, experiments under the default $6:2:2$ data split provided by Wang et al. [25] have witnessed MVIC's effectiveness. Exploring further, we figure out the impact of different positive-to-negative sample ratios on the results to better understand the generalizability of MVIC. Following [4] and considering the scale of three datasets, we do experiments under the positive and the negative ratios of 1:1, 1:3, 1:6 on the *DeepDDI* dataset. The known DDI pairs are viewed as positive samples and the unlabeled drug pairs could be selected as potential negative samples. The results are shown in Figure 4. We find that with the

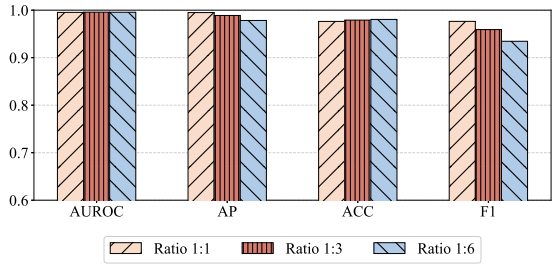

Fig. 4. MVIC under different positive and negative ratios on the *DeepDDI* dataset.

increase of negative samples, the AP and F1 scores show a declining trend. This implies that the ratio of positive to negative samples has an impact on our method. Overall, however, MVIC still demonstrates high performance, which confirms the generalizability of MVIC.

4.8 Case Study

We assess the performance of MVIC by predicting new DDIs. And this experiment is conducted on the *DeepDDI* dataset. According to the composition of the *DeepDDI* dataset, the remaining drug pairs, i.e., 1,259,086 pairs, apart from the known drug-drug interactions, may contain potential DDIs. We do not split the *DeepDDI* dataset; instead, all data is treated as the training set. After training, we test the model on unknown drug pairs and obtain the corresponding probability via Equation (15). Then we select the top 10 drug pairs with the highest probability values to identify the presence of new drug-drug interactions.

The results are demonstrated in Table 6. Mostly, the confirmation of drug-drug interactions is with the 5.1.12 version of the DrugBank database [27] (accessed on October 28, 2024). As shown in Table 6, 10 out of 10 drug pairs have an interaction between each other. For example, according to the description, the metabolism of phenobarbital can be decreased when combined with primidone if patients take in primidone and phenobarbital together.

5 Conclusion

In this study, we propose a method called MVIC for DDI prediction via multi-view information collaborative fusion, which incorporates drug direct-indirect association and drug molecular structure information. For the drug direct-indirect association, we extract information from the constructed DIN, which then passes meta-path-specific representation learning and our proposed transformer-like semantic fusion module. For the drug structure, we adopt GIN to derive molecule information. Finally, the designed multi-view collaborative fusion achieves final DDI prediction. Experiments prove the effectiveness of our proposed MVIC, which outperforms all the baselines across four metrics on three datasets.

Table 6. The top 10 drug pairs with the highest probabilities for predicting new DDIs.

Rank	Drug1	Drug2	Description
1	Primidone	Phenobarbital	The metabolism of Phenobarbital can be decreased when combined with Primidone.
2	Vemurafenib	Amiodarone	The metabolism of Vemurafenib can be decreased when combined with Amiodarone.
3	Warfarin	Acenocoumarol	The metabolism of Acenocoumarol can be increased when combined with Warfarin.
4	Phenprocoumon	Acenocoumarol	The risk or severity of bleeding can be increased when Phenprocoumon is combined with Acenocoumarol.
5	Fluvoxamine	Venlafaxine	The risk or severity of serotonin syndrome can be increased when Fluvoxamine is combined with Venlafaxine.
6	Phenprocoumon	Phenindione	The risk or severity of bleeding can be increased when Phenindione is combined with Phenprocoumon.
7	Phenindione	Acenocoumarol	The risk or severity of bleeding can be increased when Phenindione is combined with Acenocoumarol.
8	Fluvoxamine	Ziprasidone	The risk or severity of adverse effects can be increased when Fluvoxamine is combined with Ziprasidone.
9	Dicoumarol	Phenindione	The risk or severity of bleeding can be increased when Dicoumarol is combined with Phenindione.
10	Dicoumarol	Acenocoumarol	The risk or severity of bleeding can be increased when Dicoumarol is combined with Acenocoumarol.

In the future work, we will focus on extracting meta-paths automatically, which could make our method reduce dependence on domain knowledge. We are also interested in the interpretability of DDI prediction.

Acknowledgments. This work was supported in part by the National Natural Science Foundation of China (62472294).

References

1. Consortium, U.: Uniprot: a worldwide hub of protein knowledge. Nucleic Acids Research **47**(D1), D506–D515 (2019)
2. Davis, A.P., Grondin, C.J., Johnson, R.J., Sciaky, D., Wiegers, J., Wiegers, T.C., Mattingly, C.J.: Comparative toxicogenomics database (ctd): update 2021. Nucleic Acids Research **49**(D1), D1138–D1143 (2021)
3. Duvenaud, D., Maclaurin, D., Aguilera-Iparraguirre, J., Gómez-Bombarelli, R., Hirzel, T., Aspuru-Guzik, A., Adams, R.P.: Convolutional networks on graphs for learning molecular fingerprints. In: NIPS. pp. 2224–2232 (2015)
4. Feng, Y.H., Zhang, S.W., Shi, J.Y.: DPDDI: a deep predictor for drug-drug interactions. BMC Bioinformatics **21**(1), 419 (2020)
5. Gilmer, J., Schoenholz, S.S., Riley, P.F., Vinyals, O., Dahl, G.E.: Neural message passing for quantum chemistry. In: ICML. pp. 1263–1272 (2017)
6. Han, K., Jeng, E.E., Hess, G.T., Morgens, D.W., Li, A., Bassik, M.C.: Synergistic drug combinations for cancer identified in a crispr screen for pairwise genetic interactions. Nature Biotechnology **35**(5), 463–474 (2017)
7. Hochman, J., Tang, C., Prueksaritanont, T.: Drug–drug interactions related to altered absorption and plasma protein binding: theoretical and regulatory considerations, and an industry perspective. Journal of Pharmaceutical Sciences **104**(3), 916–929 (2015)
8. Kearnes, S., McCloskey, K., Berndl, M., Pande, V., Riley, P.: Molecular graph convolutions: moving beyond fingerprints. Journal of Computer-Aided Molecular Design **30**, 595–608 (2016)

9. Kuhn, M., Letunic, I., Jensen, L.J., Bork, P.: The SIDER database of drugs and side effects. Nucleic Acids Research **44**(D1), D1075–D1079 (2016)
10. Li, M., Cai, X., Xu, S., Ji, H.: Metapath-aggregated heterogeneous graph neural network for drug–target interaction prediction. Briefings in Bioinformatics **24**(1), bbac578 (2023)
11. Li, Z., Zhu, S., Shao, B., Zeng, X., Wang, T., Liu, T.Y.: DSN-DDI: an accurate and generalized framework for drug–drug interaction prediction by dual-view representation learning. Briefings in Bioinformatics **24**(1), bbac597 (2023)
12. Liu, S., Zhang, Y., Cui, Y., Qiu, Y., Deng, Y., Zhang, Z., Zhang, W.: Enhancing drug-drug interaction prediction using deep attention neural networks. IEEE/ACM Transactions on Computational Biology and Bioinformatics **20**(2), 976–985 (2022)
13. Lu, J., Yang, J., Batra, D., Parikh, D.: Hierarchical question-image co-attention for visual question answering. In: NIPS. pp. 289–297 (2016)
14. Nyamabo, A.K., Yu, H., Liu, Z., Shi, J.Y.: Drug–drug interaction prediction with learnable size-adaptive molecular substructures. Briefings in Bioinformatics **23**(1), bbab441 (2022)
15. Nyamabo, A.K., Yu, H., Shi, J.Y.: SSI–DDI: substructure–substructure interactions for drug–drug interaction prediction. Briefings in Bioinformatics **22**(6), bbab133 (2021)
16. Ogawa, R., Echizen, H.: Clinically significant drug interactions with antacids: an update. Drugs **71**, 1839–1864 (2011)
17. Palleria, C., Di Paolo, A., Giofrè, C., Caglioti, C., Leuzzi, G., Siniscalchi, A., De Sarro, G., Gallelli, L.: Pharmacokinetic drug-drug interaction and their implication in clinical management. Journal of Research in Medical Sciences **18**(7), 601 (2013)
18. Rao, J., Xie, J., Yuan, Q., Liu, D., Wang, Z., Lu, Y., Zheng, S., Yang, Y.: A variational expectation-maximization framework for balanced multi-scale learning of protein and drug interactions. Nature Communications **15**(1), 4476 (2024)
19. Ryu, J.Y., Kim, H.U., Lee, S.Y.: Deep learning improves prediction of drug–drug and drug–food interactions. Proceedings of the National Academy of Sciences **115**(18), E4304–E4311 (2018)
20. Shi, C., Li, Y., Zhang, J., Sun, Y., Philip, S.Y.: A survey of heterogeneous information network analysis. IEEE Transactions on Knowledge and Data Engineering **29**(1), 17–37 (2016)
21. Tanvir, F., Saifuddin, K.M., Islam, M.I.K., Akbas, E.: DDI prediction with heterogeneous information network - meta-path based approach. IEEE/ACM Transactions on Computational Biology and Bioinformatics **21**(5), 1168–1179 (2024)
22. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A., Kaiser, L., Polosukhin, I.: Attention is all you need. In: NIPS. pp. 6000–6010 (2017)
23. Wang, J., Liu, X., Shen, S., Deng, L., Liu, H.: DeepDDS: deep graph neural network with attention mechanism to predict synergistic drug combinations. Briefings in Bioinformatics **23**(1), bbab390 (2022)
24. Wang, X., Ji, H., Shi, C., Wang, B., Ye, Y., Cui, P., Yu, P.S.: Heterogeneous graph attention network. In: WWW. pp. 2022–2032 (2019)
25. Wang, Y., Min, Y., Chen, X., Wu, J.: Multi-view graph contrastive representation learning for drug-drug interaction prediction. In: WWW. pp. 2921–2933 (2021)
26. Weininger, D.: Smiles, a chemical language and information system. 1. introduction to methodology and encoding rules. Journal of Chemical Information and Computer Sciences **28**(1), 31–36 (1988)

27. Wishart, D.S., Knox, C., Guo, A.C., Cheng, D., Shrivastava, S., Tzur, D., Gautam, B., Hassanali, M.: DrugBank: a knowledgebase for drugs, drug actions and drug targets. Nucleic Acids Research **36**(suppl_1), D901–D906 (2008)
28. Xiong, Z., Liu, S., Huang, F., Wang, Z., Liu, X., Zhang, Z., Zhang, W.: Multi-relational contrastive learning graph neural network for drug-drug interaction event prediction. In: AAAI. pp. 5339–5347 (2023)
29. Xu, K., Hu, W., Leskovec, J., Jegelka, S.: How powerful are graph neural networks? In: ICLR (2019)
30. Yang, Z., Zhong, W., Lv, Q., Chen, C.Y.C.: Learning size-adaptive molecular substructures for explainable drug–drug interaction prediction by substructure-aware graph neural network. Chemical Science **13**(29), 8693–8703 (2022)
31. Zhang, R., Wang, X., Wang, P., Meng, Z., Cui, W., Zhou, Y.: HTCL-DDI: a hierarchical triple-view contrastive learning framework for drug–drug interaction prediction. Briefings in Bioinformatics **24**(6), bbad324 (2023)
32. Zhang, W., Chen, Y., Liu, F., Luo, F., Tian, G., Li, X.: Predicting potential drug-drug interactions by integrating chemical, biological, phenotypic and network data. BMC Bioinformatics **18**, 1–12 (2017)
33. Zhao, C., Liu, S., Huang, F., Liu, S., Zhang, W.: CSGNN: contrastive self-supervised graph neural network for molecular interaction prediction. In: IJCAI. pp. 3756–3763 (2021)
34. Zhao, W., Yuan, X., Shen, X., Jiang, X., Shi, C., He, T., Hu, X.: Improving drug–drug interactions prediction with interpretability via meta-path-based information fusion. Briefings in Bioinformatics **24**(2), bbad041 (2023)
35. Zhong, Y., Li, G., Yang, J., Zheng, H., Yu, Y., Zhang, J., Luo, H., Wang, B., Weng, Z.: Learning motif-based graphs for drug–drug interaction prediction via local–global self-attention. Nature Machine Intelligence **6**(9), 1094–1105 (2024)