

Original Paper

# Heterogeneous Network With Multiview Path Aggregation: Drug–Target Interaction Prediction Study Design

Haixue Zhao<sup>1,2\*</sup>, MEng; Kui Yao<sup>1\*</sup>, MEng; Yunjiong Liu<sup>3,4,5</sup>, PhD; Chao Che<sup>2</sup>, PhD; Lin Tang<sup>2</sup>, PhD

<sup>1</sup>Key Laboratory of Advanced Design and Intelligent Computing, Ministry of Education, Dalian University, Dalian, China

<sup>2</sup>School of Software Engineering, Dalian University, Dalian, China

<sup>3</sup>Institute of Artificial Intelligence, School of Informatics, Xiamen University, Xiamen, China

<sup>4</sup>National Institute of Diagnostics and Vaccine Development in Infectious Diseases, National Innovation Platform for Industry-Education Integration in Vaccine Research, NMPA Key Laboratory for Research and Evaluation of Infectious Disease Diagnostic Technology, Xiamen University, Xiamen, China

<sup>5</sup>State Key Laboratory of Vaccines for Infectious Diseases, Xiang An Biomedicine Laboratory, School of Public Health, Xiamen University, Xiamen, China

\*these authors contributed equally

## Corresponding Author:

Lin Tang, PhD

School of Software Engineering

Dalian University

No. 10 Xuefu Street, Dalian Economic and Technological Development Zone, Liaoning

Dalian 116622

China

Phone: 86 18940965015

Email: [tanglin@dlu.edu.cn](mailto:tanglin@dlu.edu.cn)

## Abstract

**Background:** Drug–target interaction (DTI) prediction is crucial in drug repositioning, as it can significantly reduce research and development costs and shorten the development cycle. Most existing deep learning–based approaches employ graph neural networks for DTI prediction. However, these approaches still face limitations in capturing complex biochemical features, integrating multilevel information, and providing interpretable model insights.

**Objective:** This study proposes a heterogeneous network model based on multiview path aggregation, aiming to predict interactions between drugs and targets.

**Methods:** This study employed a molecular attention transformer to extract 3D conformation features from the chemical structures of drugs and utilized Prot-T5, a protein-specific large language model, to deeply explore biophysically and functionally relevant features from protein sequences. By integrating drugs, proteins, diseases, and side effects from multisource heterogeneous data, we constructed a heterogeneous graph model to systematically characterize multidimensional associations between biological entities. On this foundation, a meta-path aggregation mechanism was proposed, which dynamically integrates information from both feature views and biological network relationship views. This mechanism effectively learned potential interaction patterns between biological entities and provided a more comprehensive representation of the complex relationships in the heterogeneous graph. It enhanced the model's ability to capture sophisticated, context-dependent relationships in biological networks. Furthermore, we integrated multiscale features of drugs and proteins within the heterogeneous network, significantly improving the prediction accuracy of DTIs and enhancing the model's interpretability and generalization ability.

**Results:** In the DTI prediction task, the proposed model achieves an AUPR (area under the precision-recall curve) of 0.901 and an AUROC (area under the receiver operating characteristic curve) of 0.966, representing improvements of 1.7% and 0.8%, respectively, over the baseline methods. Furthermore, a case study on the KCNH2 target demonstrates that the proposed model successfully predicts 38 out of 53 candidate drugs as having interactions, which further validates its reliability and practicality in real-world scenarios.

**Conclusions:** The proposed model shows marked superiority over baseline methods, highlighting the importance of integrating heterogeneous information with biological knowledge in DTI prediction.

**Keywords:** drug-target interactions; drug repurposing; heterogeneous networks; graph neural networks

## Introduction

### Background

Drugs play a crucial role in treating diseases by interacting with multiple targets and modulating their functions. Accurately predicting drug-target interactions (DTI) is essential for understanding drug mechanisms [1], discovering new targets, and facilitating drug repositioning. Traditional methods rely heavily on large amounts of labeled data and often suffer from label noise in negative samples. In recent years, positive-unlabeled learning has been effectively applied to alleviate this issue [2]. Meanwhile, self-supervised learning strategies integrated with heterogeneous biomedical networks have improved the accuracy of DTI prediction by leveraging multimodal information [3]. With the rapid advancement of large language models (LLM), their applications in the biomedical domain have expanded significantly. Due to their powerful sequence representation and semantic understanding capabilities, LLMs are driving DTI prediction methods toward a higher level of performance [4].

Currently, DTI prediction methods are primarily categorized into three types [5,6]: ligand similarity-based methods [7,8], network-based methods [9,10], and structure-based methods [11]. Ligand similarity-based methods predict DTIs by comparing the structural similarity of drugs. These methods are computationally efficient but often overlook complex biochemical properties and molecular characteristics, which may lead to inaccurate predictions [12]. Network-based methods rely on large amounts of high-quality interaction data and are computationally intensive [13-15]. They typically lack structural information about drugs and targets, resulting in poor performance on sparse networks. Structure-based methods usually depend on the three-dimensional structures of target proteins and drugs [16]. However, these methods exhibit limited efficacy for proteins with unknown structures.

To address the above issues, we propose a heterogeneous network (HN) model based on multiview path aggregation for DTI (MVPA-DTI) prediction. MVPA-DTI enhances prediction capability through a multiview feature extraction and fusion mechanism. First, the molecular attention Transformer model is used to extract the three-dimensional spatial structure information of drugs, while the protein-specific LLM Prot-T5 is leveraged to deeply explore the biophysical properties of protein sequences, forming two feature views based on structure and sequence. Subsequently, a HN is constructed by incorporating multisource heterogeneous information from proteins, drugs, diseases, and side effects. This approach effectively integrates the previously defined feature views into the message-passing framework of the HN, thereby forming a comprehensive biological network relationship view. The multiview path aggregation

mechanism allows the model to dynamically synthesize feature information from multiple perspectives. By introducing a meta-path aggregation mechanism, the model dynamically synthesizes feature information across different perspectives, capturing underlying drug-target associations at multiple hierarchical levels through both feature-oriented views and biological network relational contexts. During the message-passing process, MVPA-DTI optimizes the weight distribution by incorporating both network topology and biological prior knowledge. By extracting node-level feature embeddings from heterogeneous data, the model is capable of accurately predicting new DTI. Experimental results show that MVPA-DTI outperforms existing advanced methods across multiple evaluation metrics. Additionally, for the voltage-gated inward-rectifying potassium channel KCNH2 target, which is related to cardiovascular diseases, MVPA-DTI was used for candidate drug screening. Among 53 candidate drugs, 38 are predicted to interact with this target (10 of which are already used in clinical treatment). This finding not only validates the effectiveness of MVPA-DTI in predicting DTI but also further demonstrates its application potential and practical value in real drug development. The main contributions of this paper are as follows:

1. We employ a molecular attention Transformer network that extracts three-dimensional structural information from molecular graphs through a physics-informed attention mechanism, establishing a structural feature perspective of the drug.
2. The sequence-level features are extracted from protein sequences using Prot-T5, a protein-specific LLM. The model maps sequence features to functional relevance, establishing a sequence feature perspective of proteins.
3. We integrate the drug structural view and protein sequence view into a multi-entity HN to construct a biological network relationship view. The meta-path information aggregation mechanism captures higher-order interaction patterns among different types of nodes.
4. We propose a multiview path aggregation-enhanced HN model that integrates protein and drug biological knowledge to accurately identify critical drug-target relationships. Benchmark tests show that the model exhibits a significant advantage in prediction performance.

### Related Work

DTI prediction plays a significant role in drug discovery and drug repositioning. With the rapid advancement of deep learning technologies, an increasing number of deep learning-based DTI prediction methods emerge. These methods, particularly those combining big data with deep learning, gradually overcome the limitations of traditional approaches. Based on different prediction strategies, DTI prediction methods are generally categorized into three types [5,6]: ligand similarity-based methods [7,8], network-based

methods [9,10], and structure-based methods [11]. In recent years, prediction methods based on LLMs also bring remarkable breakthroughs to the DTI field, opening new pathways for its development.

### **Ligand Similarity–Based Methods**

In DTI prediction, ligand similarity–based methods estimate potential interactions by comparing the structures or chemical similarity of drug molecules. These methods typically rely on the SMILES (simplified molecular input line entry system) representation or molecular fingerprints of drugs to calculate drug–drug similarities and infer potential target interactions based on these similarities. For instance, Thafar et al [17] proposed the DTiGEMS model, which integrates drug–drug similarities and employs a similarity selection and fusion algorithm to enhance the accuracy of DTI predictions. Additionally, Shim et al [18] introduced a similarity-based convolutional neural network method. This method calculates the outer product of the similarity matrix between drugs and targets and employs a two-dimensional convolutional neural network to capture potential relationships between them. These methods have improved DTI prediction accuracy to some extent, but they often overlook the dynamic interactions and complex spatial structures between molecules. Moreover, ligand similarity–based methods operate under the assumption that similar drug molecules share similar targets. However, this assumption does not always hold, especially when dealing with molecules that exhibit significant structural differences.

### **Network-Based Methods**

Network-based methods typically rely on large amounts of known DTI data and utilize graph algorithms for modeling. In these methods, drugs and targets are represented as nodes in a graph, while the interactions between them are represented as edges. Additionally, drug–drug similarity networks, protein–protein similarity networks, and known DTI networks are often integrated into a HN. Examples of such approaches include DTI networks, multisimilarity collaborative matrix factorization, and HN models. On this basis, researchers used graph structures to explore potential interaction patterns and promoted the development of tasks such as drug repositioning.

In recent years, with the development of graph representation learning [19,20], network-based methods have achieved deep fusion modeling of structural and semantic information. For example, Zhao et al proposed a regulation-aware graph learning method that enhances drug representation by integrating gene regulation information, thereby improving the prediction effect of drug repositioning [21]. GNN methods mostly rely on the direct adjacency structure of nodes, which easily ignores the rich high-order semantic information and cross-modal associations in biological networks. To this end, some studies have proposed information fusion methods that combine low-order and high-order graph structures. For example, Zhao et al jointly modeled direct neighbor relationships and high-order network path features to effectively enhance the discriminability of drug and target representation [22]. At the same time, in order

to more comprehensively model the heterogeneous relationships between multiple entities such as drugs, proteins, and diseases, Zhao et al constructed a heterogeneous information network using semantic paths and attention mechanisms to achieve the integration and discrimination of multimodal information, thereby improving the accuracy and generalization ability of DTI prediction [23]. The iGRLDTI introduced edge weight regulation mechanisms and regularization terms in GNN, further improving the model's ability to learn potential interactions in heterogeneous biological networks [24].

### **Structure-Based Methods**

Traditional structure-based methods typically use molecular docking techniques, combined with molecular dynamics simulations, to accurately predict the binding patterns between drugs and targets. With the development of deep learning methods, an increasing number of studies have integrated structural information with neural networks. For example, DeepDTA [25] and DeepDrug3D [26] have enhanced the modeling capability of the drug–target binding process by incorporating three-dimensional structural information of proteins and drugs, further improving prediction accuracy. Although deep learning models can handle structural information, they generally require large amounts of labeled data for effective training, and obtaining high-quality drug–target binding data remains a challenge.

The application of language models is particularly crucial in the absence of molecular structural information. Drawing on techniques from the field of Natural Language Processing, pretrained language models (such as BERT [27] and MolBERT [28]) are used to deeply characterize drug and protein sequence data. By automatically learning latent features and semantic information within sequences, these models enhance the accuracy of DTI predictions. A key advantage of LLMs is their ability to process large-scale, unlabeled data and improve model generalization through transfer learning. For instance, MolBERT [28] and ChemBERTa [29] achieve high-quality molecular representations by pretraining on massive drug molecular datasets. Transformer-based protein language models, such as ProtBERT [30], TAPE [31], and ProtT5 [32], are pretrained on large protein sequence databases. This enables them to capture hierarchical and context-rich sequence representations, enhancing the understanding of protein characteristics. These models can efficiently and accurately characterize proteins using only sequence data, without relying on three-dimensional structural information.

## **Methods**

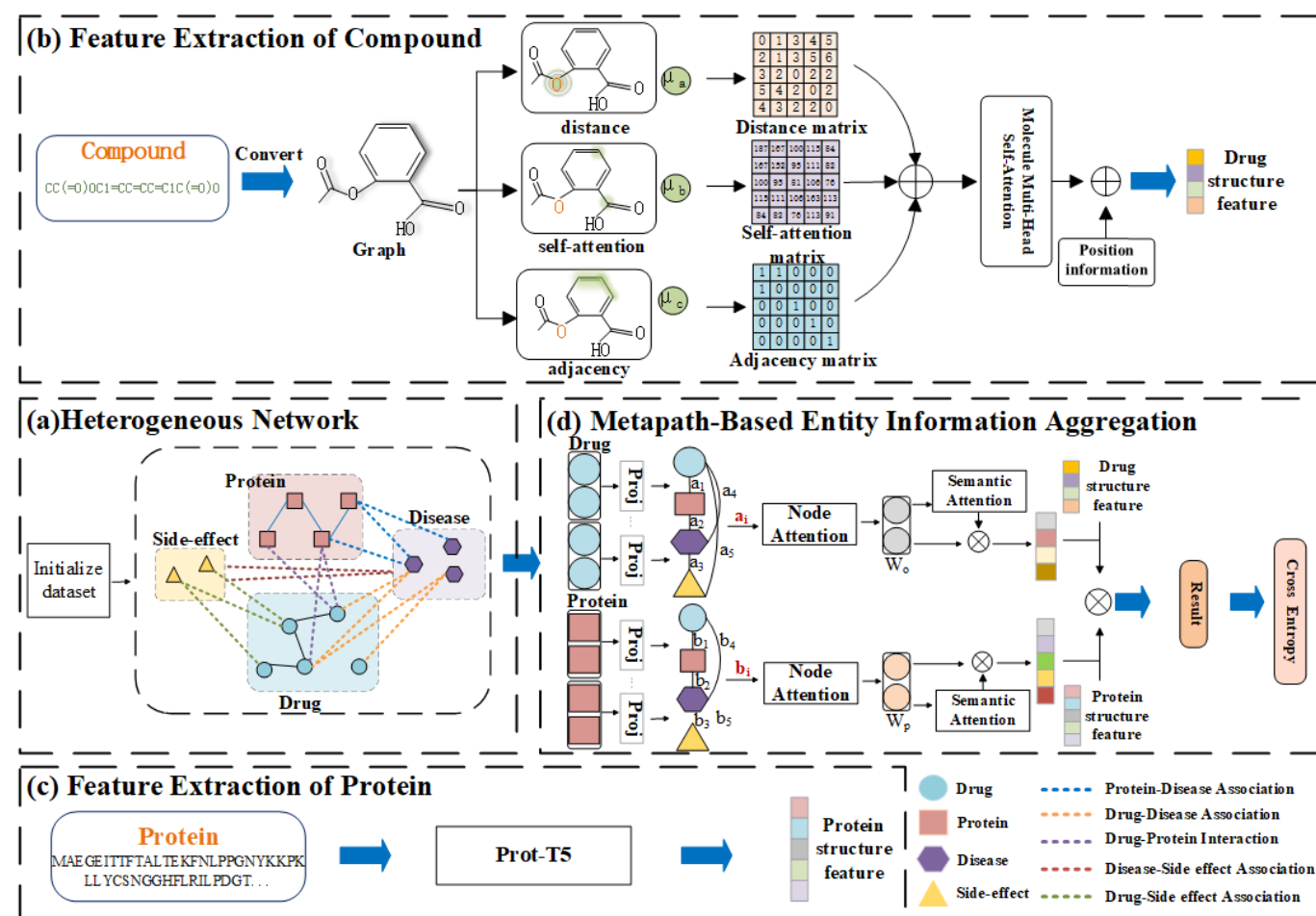
### **Framework**

MVPA-DTI combined multiple types of heterogeneous data. It extracted structural and sequential features from drug molecules and protein sequences. It also optimized the weight distribution for information propagation in a heterogeneous graph neural network. As illustrated in Figure 1, the process consisted of four steps: (a) First, an HN is constructed,

incorporating drugs, proteins, diseases, and drug side effects. (b) Next, a molecular attention Transformer [33] network is used to extract 3D structural features from the SMILES [34] representations of drugs, which are then assigned to drug nodes. (c) Then, the Prot-T5 model is employed to analyze key biophysical features of protein data, which are subsequently assigned to target nodes in the graph. (d) In

the final step, a Heterogeneous Graph Attention Network (HAN) [35] is utilized to analyze network topology, unify node embeddings, determine node importance using meta-paths, and integrate node features through semantic attention. Ultimately, MVPA-DTI predicted potential DTI by optimizing the drug-protein reconstruction loss and the cross-entropy loss.

**Figure 1.** Architecture of the MVPA-DTI model: (A) Construction of the HN; (B) extraction of structural features from the SMILES sequences of drugs to reconstruct drug representations in the HN; (C) extraction of key biophysical properties from protein sequences using Prot-T5 to reconstruct protein representations in the HN; (D) extraction of topological features through the HAN and message-passing mechanism, followed by DTI prediction. DTI: drug-target interaction; HAN: Heterogeneous Graph Attention Network; HN: heterogeneous network; MVPA-DTI: multi-view path aggregation for drug-target interaction; Prot-T5: ProtTrans T5-XL-U50; SMILES: Simplified Molecular Input Line Entry System.



## Feature View

### Feature Extraction of Compound Structures

The MVPA-DTI model utilized SMILES sequences as input for drug molecules to learn effective representations of their structural features. Since SMILES sequences shared structural similarities with natural language, contextual information can be effectively leveraged to analyze molecular features. However, traditional methods often struggled to capture long-range atomic relationships when processing molecules. To overcome this limitation, this study employed an improved Transformer architecture.

First, we employed the RDKit toolkit [36] to convert SMILES sequences into molecular graph structures. Since the lengths of SMILES sequences vary across different

molecules, a maximum length of 100 is selected to construct reasonable molecular representations. This setting covered at least 90% of the compounds in the dataset. For sequences exceeding this length, truncation is applied, while shorter sequences are padded with zeros to maintain consistent input formatting.

For feature extraction, we employed the Molecule Attention Transformer (MAT) [33] for encoding. Its core idea is to replace the self-attention layer of the traditional Transformer [37] with an enhanced molecular self-attention mechanism. By integrating adjacency information from molecular graphs and interatomic distance information, MAT comprehensively captured the feature representation  $S_{drug}$  of drug structures. The attention mechanism integrated interatomic distance and molecular topology.



This enhancement allowed the attention distribution to more precisely capture internal molecular relationships. The calculation formula for molecular multihead self-attention is as follows:

$$A^{(i)} = \left( \mu_b \rho \left( \frac{Q_i K_i^T}{\sqrt{d_k}} \right) + \mu_a g(D) + \mu_c A \right) V_i \quad (1)$$

where  $A \in \{0, 1\}^{N_{atoms} \times N_{atoms}}$  represented the adjacency matrix of the molecular graph and  $D \in R^{N_{atoms} \times N_{atoms}}$  denoted the distances between atoms. The query, key, and value vector matrices are defined as  $Q_i = XW_i^q$ ,  $K_i = XW_i^k$ , and  $V_i = XW_i^v$ , respectively, where  $W$  represented learnable parameters. In the attention calculation,  $\mu_a$ ,  $\mu_b$ , and  $\mu_c$  corresponded to the weights of different attention components, with specific meanings as follows:  $\mu_a$  measured the importance of interatomic distances in the attention mechanism;  $\mu_b$  adjusted the influence of self-attention on the overall attention mechanism; and  $\mu_c$  controlled the role of the adjacency matrix in the attention computation.

## Feature Extraction of Protein Structures

The MVPA-DTI model employed Prot-T5 to extract features from protein sequences. By treating sequences as natural language, it used Transformer-based self-supervised learning to capture biological insights. During the data preprocessing stage, protein sequences are treated as text data composed of 20 standard amino acids (along with a few unknown amino acids, such as X). To adapt to the input format of Natural Language Processing tasks, each protein sequence is treated as a token sequence of individual amino acids, with spaces inserted between each amino acid to ensure the Transformer can correctly parse the sequence information. Additionally, to enhance the model's generalization capability, all uncommon or unresolved amino acid symbols (eg, B, O, U, Z) are mapped to the universal token X, ensuring data consistency.

Prot-T5 is based on the Text-to-Text Transfer Transformer (T5) [38] architecture, which employs an encoder-decoder structure. However, for the task of protein feature extraction, only the encoder part of T5 is utilized to map protein sequences into a high-dimensional feature space. Prot-T5 adopted the Span Masking language modeling strategy, where during training, continuous segments of amino acids in the sequence are randomly masked, and the model is tasked with reconstructing the masked portions based on the context. This enabled the model to learn long-range dependencies and semantic information within protein sequences. As shown in Equation 2, given a protein sequence  $S = (S_1, S_2, \dots, S_n)$ , it is first embedded into a high-dimensional space  $E(S)$  and then fed into the Transformer to compute the hidden representations:

$$H = \text{Transformer}(E(S)) \quad (2)$$

Finally, the global representation of the entire sequence is obtained through average pooling, as calculated by the following formula:

$$S_{protein} = \frac{1}{n} \sum_{i=1}^n H_i \quad (3)$$

where  $H_i$  represented the hidden states output by the Transformer and  $S_{protein}$  served as the global feature of the protein.

## Relational View of Biological Networks

### Heterogeneous Networks

The HN is described as an undirected graph  $G = (V, E)$ , where  $V = \{v_1, v_2, \dots, v_n\}$  represented the set of nodes and  $E = \{E(1), E(2), \dots, E(k)\}$  denoted the set of ktypes of edges, with each  $E^{(i)} = \{e_{i1}, e_{i2}, \dots, e_{im}\}$  representing a specific type of edge.

As shown in Figure 1A, this paper constructed a comprehensive heterogeneous information network, which included drug-drug interactions, drug-protein interactions, drug-disease associations, drug-side effect associations, protein-protein interactions, and protein-disease associations. In the network, entities such as drugs, targets, diseases, and side effects are represented as nodes, while the relationships and interactions between nodes are represented as edges. In the proposed framework, each node belonged to only one type of entity, and all edges are undirected and have nonnegative weights. Node messages are first sent to their first-order neighbors and then propagated to higher-order neighbors through the network edges, a process known as message passing.

The message-passing process in the HN is divided into multiple stages. First, nodes transmitted their information to first-order neighbors. This information is then propagated to higher-order neighbors through network edges, forming more complex information dissemination. Message passing is not limited to directly connected nodes but also includes information transmitted through multihop paths. This enabled the model to integrate features across a broader neighborhood, thereby better capturing long-range dependencies and multilevel network structures.

### Metapath-Based Entity Information Aggregation

In heterogeneous graph representation learning, our goal is to learn effective feature representations for each node in the HN. However, the challenge of this task lies in not only integrating information from different types of nodes and edges in the heterogeneous graph but also considering the heterogeneous features and content of each node. To address this issue, we employed HAN [35] as the topological feature extraction method for the heterogeneous information network. HAN modeled the structural relationships in the HN through a hierarchical attention mechanism. Specifically, node-level attention learned meta-path-based neighborhood

weights and aggregates them to generate semantically specific node embeddings, while semantic-level attention assigned weights across different meta-paths to obtain the optimal task-specific representation, as illustrated in Figure 1D. In terms of meta-path selection, we selected two biologically meaningful paths: D-P-D (Drug-Protein-Drug) and P-D-P-D (Protein-Disease-Protein-Drug). D-P-D is used to capture the semantic relationship between drugs connected by common targets, which helped the model identify drug pairs with similar functions or related potential mechanisms of action, while P-D-P-D further explored the multi-hop indirect connections between proteins formed by disease associations, thereby reflecting a more complex biological network structure. We selected these two meta-paths based on their good semantic interpretability and biological relevance, which helps the model more effectively model the potential interactions between drugs and targets.

After obtaining the node embeddings, we further proceeded with the message-passing process. Assuming the initial node embeddings are defined as  $f^0: V \rightarrow R^d$ , where  $f^0(v)$  represented the mapping of node  $v$  in the  $d$ -dimensional space, the information aggregation process for node  $v$  can be expressed as:

$$f_v = \sigma \left( \frac{1}{K} \sum_{k=1}^K \sum_{u \in N_v} \alpha_{uv} W f^0(u) \right) \quad (4)$$

$$\alpha_{uv} = \frac{\exp(\text{AconC}(\alpha^T [W f^0(u) \parallel W f^0(v)]))}{\sum_{k \in N_u} \exp(\text{AconC}(\alpha^T [W f^0(u) \parallel W f^0(k)]))} \quad (5)$$

where  $\sigma$  represents a nonlinear activation function,  $K$  denotes the number of attention layers,  $N_v$  represents the set of neighboring nodes of node  $v$ ,  $W$  is the shared weight parameter,  $\alpha_{uv}$  is the edge weight computed by the attention mechanism, and  $\text{AconC}(\cdot)$  [39] is an adaptive activation function.

Based on obtaining the graph structure representation, we further integrated the structural information of each node to construct the final drug and protein representation. For drug nodes, its representation integrates the meta-path structure features, the original embedding representation, and the molecular structure features, which is specifically expressed as:

$$f_{\text{drug}} = \text{L2 Norm}(\text{ReLU}([ \text{mate}_{\text{DPPD}}, E_{\text{drug}} + \text{L2 Norm}(S_{\text{drug}})] W_0 + b_0)) \quad (6)$$

Similarly, the protein node representation combined its meta-path structure features, original embeddings, and sequence representations and is expressed as:

$$f_{\text{protein}} = \text{L2 Norm}(\text{ReLU}([ \text{mate}_{\text{PDPD}}, E_{\text{protein}} + \text{L2 Norm}(S_{\text{protein}})] W_0 + b_0)) \quad (7)$$

Among them,  $W_0$  and  $b_0$  are learnable linear transformation parameters, which are used to uniformly map the concatenated multiple features to a common low-dimensional semantic space, and implicitly model the importance weights of each

feature subspace during the training process to improve the discriminative ability of the final representation.

After completing the representations of drugs and proteins, we used the inner product to calculate the interaction probability between them, as shown in Figure 1D. For a drug node  $u$  and a protein node  $v$ , their interaction probability is calculated as follows:

$$p^{uv} = \sigma((f_u)^T f_v) \quad (8)$$

where  $\sigma$  is the sigmoid function and  $p^{uv}$  represents the interaction score between node  $u$  and node  $v$ .

To optimize the model, we employed multiple loss terms for joint training. Among them, the cross-entropy loss is used to supervise the interaction prediction task, defined as follows:

$$\mathcal{L}_{\text{CEL}} = \sum_{r \in R} -y_r \log(p_r) - (1 - y_r) \log(1 - p_r) \quad (9)$$

where  $R$  is the set of relationships,  $y_r$  is the ground truth label of sample  $r$  (1 for positive samples and 0 for negative samples), and  $p_r$  is the predicted probability.

In addition, to enhance the model's structural perception of node features, we introduced reconstruction losses for drugs and proteins, which are defined as follows:

$$\mathcal{L}_{\text{rec}}^{\text{drug}} = \|\hat{A}_{\text{drug}} - A_{\text{drug}}\|_F^2 \quad (10)$$

$$\mathcal{L}_{\text{rec}}^{\text{protein}} = \|\hat{A}_{\text{protein}} - A_{\text{protein}}\|_F^2 \quad (11)$$

Among them,  $A_{\text{drug}}$  and  $A_{\text{protein}}$  are the adjacency matrices of drug and protein nodes in the original HN, respectively.  $\hat{A}$  represents the adjacency matrix reconstructed by the model through node embedding and  $\|\cdot\|_F$  represents the Frobenius norm. The final optimization target is the weighted sum of three losses:

$$\mathcal{L} = \mathcal{L}_{\text{CEL}} + \lambda_1 \mathcal{L}_{\text{rec}}^{\text{drug}} + \lambda_2 \mathcal{L}_{\text{rec}}^{\text{protein}} \quad (12)$$

where  $\mathcal{L}_{\text{rec}}^{\text{drug}}$  and  $\mathcal{L}_{\text{rec}}^{\text{protein}}$  represent the reconstruction losses for drugs and proteins, respectively, while  $\lambda_1$  and  $\lambda_2$  are weighting coefficients used to balance the contributions of different loss terms. Through joint optimization, we can improve the accuracy of DTI prediction while ensuring that the generated feature representations possess robust structural information and discriminative capabilities.

## Results

### Dataset

This study conducted DTI prediction experiments on the Luo dataset [9], which was assembled by Luo et al. To

better capture biological characteristics, we incorporated the SMILES sequences of drug molecules and the structural sequence data of proteins into the dataset. [Tables 1](#) and [2](#) present the relevant statistical information of nodes and edges

in the dataset. The dataset comprises six independent drug–protein interaction networks, with all edge weights being binary values.

**Table 1.** Statistics of node information in the dataset.

Node type	Count
Drug	708
Protein	1512
Disease	5603
Side effect	4192

**Table 2.** Statistics of association information between nodes in the dataset.

Edge type	Count	Source
Drug–protein interaction	1923	DrugBank version 3.0 <a href="#">[40]</a>
Drug–drug interaction	10,036	DrugBank version 3.0 <a href="#">[40]</a>
Protein–protein interaction	7363	HPRD Release 9 <a href="#">[41]</a>
Drug–disease association	199,214	Comparative Toxicogenomics Database <a href="#">[42]</a>
Protein–disease association	1,596,745	Comparative Toxicogenomics Database <a href="#">[42]</a>
Drug–side-effect association	80,164	SIDER Release 2 <a href="#">[43]</a>

Experimental Settings

Building upon the study by Luo et al [\[9\]](#), 90% of the positive and negative samples from the dataset were used to construct the HN and train MVPA-DTI, while the remaining 10% were reserved for testing. In this study, we adopted a data preprocessing process consistent with NeoDTI, including negative sample sampling strategy, drug–target similarity calculation method, and similarity threshold setting

to eliminate redundancy. Subsequently, the model’s performance was evaluated through 10-fold cross-validation, with the effectiveness of the method being measured using AUROC and AUPR metrics.

MVPA-DTI consists of three modules: Prot-T5, MAT, and HAN, with the parameter settings detailed in [Table 3](#). During the training process, we employed the Adam optimizer to update the network weights.

**Table 3.** Parameter settings in the experiment.

Parameter	Value
Number step	3000
Batch size	128
Learning rate	0.001
Dropout rate	0.1
HAN <sup>a</sup> input size	1024
MAT <sup>b</sup> multi-head attention number	16
MAT stack number	8

<sup>a</sup>HAN: Heterogeneous Graph Attention Network.

<sup>b</sup>MAT: Heterogeneous Graph Attention Network.

Comparison With Prior Work

In this study, to better simulate real-world scenarios, we randomly sampled negative samples while retaining all positive samples, setting the ratio of positive to negative samples at 10:1. This ratio aims to reflect the imbalance between positive and negative samples in the real world. Subsequently, we compared the performance of MVPA-DTI with several benchmark models to evaluate its effectiveness in DTI prediction tasks; they are briefly described as follows:

- MSCMF [\[44\]](#): MSCMF effectively predicts DTI by utilizing multiple similarity matrices of drugs and targets through collaborative matrix factorization.
- HNM [\[45\]](#): HNM integrates disease, drug, and target information via a HN, aiming to enhance the efficiency and accuracy of drug repositioning.
- DTINet [\[9\]](#): DTINet predicts interactions between new drugs and targets by integrating heterogeneous data and learning low-dimensional feature representations.

- NeoDTI [46]: NeoDTI automatically generates feature representations by integrating diverse information from HN.
- EEG-DTI [47] : EEG-DTI is an end-to-end learning framework based on heterogeneous graph convolutional networks, capable of effectively learning features from multiple biological entities.
- SHGCL-DTI [48]: SHGCL-DTI combines semi-supervised learning with graph contrastive learning, enhancing the model’s adaptability.

The performance of MVPA-DTI was compared with various models, and the experimental results are shown in Table 4. MVPA-DTI achieves the best performance, demonstrating significant improvement over other DTI models. It achieves an AUPR of 0.901, which is 1.7% higher than the second-best method, and an AUROC of 0.967, representing a 0.9% improvement over SHGCL-DTI. MSCMF employs matrix transformation to optimize prediction results through network inference and the topological structure of data. However, compared to the recent outstanding performance

of deep learning methods, MSCMF does not fully exploit the latent information in data matrix embeddings or the features of adjacent nodes, thereby limiting its predictive performance. HNM does not adopt mainstream heterogeneous data embedding methods for feature representation and information integration, resulting in insufficient generalization ability and lower prediction accuracy. In contrast, DTI-Net, NeoDTI, and EEG-DTI further extract hidden features by combining matrix transformation with neural networks, enabling more accurate modeling of node relationships and improving prediction performance. SHGCL-DTI employs a graph contrastive learning strategy to capture the structural information of heterogeneous graphs. This is achieved by enhancing the similarity of positive sample pairs and reducing the similarity of negative sample pairs. However, SHGCL-DTI fails to fully utilize the rich semantic information and complex interaction patterns in heterogeneous graphs, presenting certain limitations when processing biological data.

**Table 4.** The AUPR and AUROC evaluation results for each model.

Model	AUPR <sup>a</sup>	AUROC <sup>b</sup>
HNM <sup>c</sup>	0.579	0.834
MSCMF <sup>d</sup>	0.603	0.856
DTINet <sup>e</sup>	0.818	0.916
NeoDTI <sup>f</sup>	0.855	0.943
EEG-DTI <sup>g</sup>	0.847	0.952
SHGCL-DTI <sup>h</sup>	0.884	0.958
MVPA-DTI (ours) <sup>i</sup>	0.901	0.966

<sup>a</sup> AUPR: area under the precision–recall curve.

<sup>b</sup> AUROC: area under the receiver operating characteristic curve.

<sup>c</sup> HNM: heterogeneous network model.

<sup>d</sup> MSCMF: multiple similarities collaborative matrix factorization.

<sup>e</sup> DTINet: drug–target interaction prediction network.

<sup>f</sup> NeoDTI: neural integration of neighbor information for DTI prediction.

<sup>g</sup> EEG-DTI: end-to-end graph for drug–target interaction prediction

<sup>h</sup> SHGCL-DTI: semi-supervised heterogeneous graph contrastive learning for drug–target interaction prediction.

<sup>i</sup> MVPA-DTI: multi-view path aggregation for drug–target interaction.

Existing methods fail to fully exploit critical biochemical information in drug molecular structures and protein sequences, potentially leading to information loss during the node embedding process. In contrast, MVPA-DTI integrates a graph attention network that incorporates all types of adjacent nodes, enhancing the extraction of composite structural information. Furthermore, it employs HAN to model the potential relationships between different types of nodes. Additionally, the method dynamically fuses features extracted from sequences and graph to update the embedding representations of drug and protein nodes. Through this process, it progressively strengthens the weight of node information

and optimizes the feature representation of different types of nodes, thereby improving the predictive accuracy.

### Robustness Experiment

To further validate the stability of MVPA-DTI, considering the potential presence of redundant information in the dataset, additional experiments were conducted to evaluate the model’s predictive performance. First, we removed some samples from the dataset, including DTIs with similar drugs or targets and DTIs involving drugs with similar drug interactions. The details of the removed data are provided in Table 5.

**Table 5.** Number of drugs, targets, and DTIs included in the redundant data.

Redundant data	Drug number	Target number	DTI number
DTI of similar drugs or targets	146	78	955



Redundant data	Drug number	Target number	DTI number
DTI of drugs with similar side effects	17	0	51

<sup>a</sup>DTI: drug–target interactions.

After removing similar drugs and targets, as shown in Table 6, MVPA-DTI achieves a 15.9% improvement over the second-best method, NeoDTI. Meanwhile, Table 6 also demonstrates that after removing drugs with similar side effects, MVPA-DTI outperforms the second-best method by 3.1%.

**Table 6.** AUPR performance of each model after adjustment of the dataset.

Model	AUPR <sup>a</sup>		
	DTIs <sup>b</sup> with similar drugs and targets were removed	DTIs with drugs with similar side effects were removed	Trained on non-unique dataset and tested on unique dataset
HNM <sup>c</sup>	0.547	0.581	0.233
MSCMF <sup>d</sup>	0.265	0.593	0.206
DTINet <sup>e</sup>	0.611	0.803	0.389
NeoDTI <sup>f</sup>	0.694	0.848	0.432
EEG-DTI <sup>g</sup>	0.686	0.846	0.431
SHGCL-DTI <sup>h</sup>	0.617	0.729	0.443
MVPA-DTI <sup>i</sup>	0.849	0.873	0.46

<sup>a</sup>AUPR: area under the precision–recall curve.

<sup>b</sup>DTI: drug–target interaction.

<sup>c</sup>HNM: heterogeneous network model.

<sup>d</sup>MSCMF: multiple similarities collaborative matrix factorization.

<sup>e</sup>DTINet: drug–target interaction prediction network.

<sup>f</sup>NeoDTI: neural integration of neighbor information for DTI prediction.

<sup>g</sup>EEG-DTI: end-to-end graph for drug–target interaction prediction.

<sup>h</sup>SHGCL-DTI: semi-supervised heterogeneous graph contrastive learning for drug–target interaction prediction.

<sup>i</sup>MVPA-DTI: multi-view path aggregation for drug–target interaction.

The experimental results indicate that although the model's performance declined after removing a significant portion of specific DTI data, MVPA-DTI still maintains the best AUPR, demonstrating its strong robustness. Additionally, this study treats drug–protein interactions as a specific case for experimentation. To objectively evaluate the predictive capability of MVPA-DTI, both special drug–target relationships and conventional drug–target relationships were separately processed. Specifically, the model was first trained on a dataset in which the relationships between drugs and proteins were nonunique and then tested on a dataset with unique interactions. As shown in Table

6, MVPA-DTI significantly outperforms the second-best method, with a 1.7% improvement in AUPR. This result suggests that MVPA-DTI exhibits stronger generalization ability in predicting DTI.

## Ablation Experiment

To verify the contribution of each module in MVPA-DTI, we conducted ablation experiments. The experimental results are presented in Table 7, where ProSF represents the protein sequence feature extraction module and DruSF denotes the drug compound structure feature extraction module.

**Table 7.** Evaluation of the impact of MVPA-DTI modules on performance.

Method	AUROC <sup>a</sup>	AUPR <sup>b</sup>	$F_1$ -score	MCC <sup>c</sup>
w/o <sup>d</sup> ProSF <sup>e</sup>	0.963	0.886	0.839	0.828
w/o DruSF <sup>f</sup>	0.964	0.891	0.840	0.827
w/o ProSF and DruSF	0.957	0.875	0.813	0.804
protBert-MVPA-DTI <sup>g</sup>	0.964	0.89	0.831	0.822
MVPA-DTI <sup>h</sup>	0.966	0.901	0.848	0.839

<sup>a</sup>AUROC: Area Under the Receiver Operating Characteristic curve.

<sup>b</sup>AUPR: Area Under the Precision-Recall Curve.

<sup>c</sup>MCC: Matthews Correlation Coefficient.

<sup>d</sup>w/o denotes corresponding module was removed.

<sup>e</sup>ProSF: protein sequence feature extraction module.

<sup>f</sup>DruSF: drug compound structure feature extraction module.

<sup>g</sup>protBert-MVPA-DTI replaces the ProSF module with ProtBert processing.

<sup>h</sup>MVPA-DTI: multi-view path aggregation for drug–target interaction.

## The Effectiveness of ProSF

As shown in Table 7, the ProSF module improves the model performance by 0.3% in AUROC and 1.5% in AUPR. ProSF effectively extracts deep semantic information from proteins, capturing features related to key physicochemical properties such as secondary structure and solubility, thereby enhancing the model's predictive capability.

## The Effectiveness of DruSF

As shown in Table 7, the DruSF module enhances the model's performance, with improvements of 0.2% and 1.0% in AUROC and AUPR metrics, respectively. This enhancement is attributed to DruSF's ability to thoroughly explore the structural features of drugs, thereby assigning higher weights to drugs during the message-passing process and further optimizing the prediction outcomes.

The experimental results demonstrate that the MVPA-DTI model exhibits high effectiveness in extracting the biological structural features of drugs and proteins. By deeply exploring the chemical structures, physical properties, and biological functions of drugs and proteins, the model can accurately capture the complex interaction relationships between them, thereby enhancing DTI prediction performance. Further analysis reveals that protein structural information plays a more critical role than drug structural information in the prediction process. This phenomenon may be attributed to the protein-specific LLM Prot-T5, which more effectively captures evolutionary conservation and functional relevance in protein sequences, thereby providing more discriminative feature representations for DTI prediction.

## Discussion

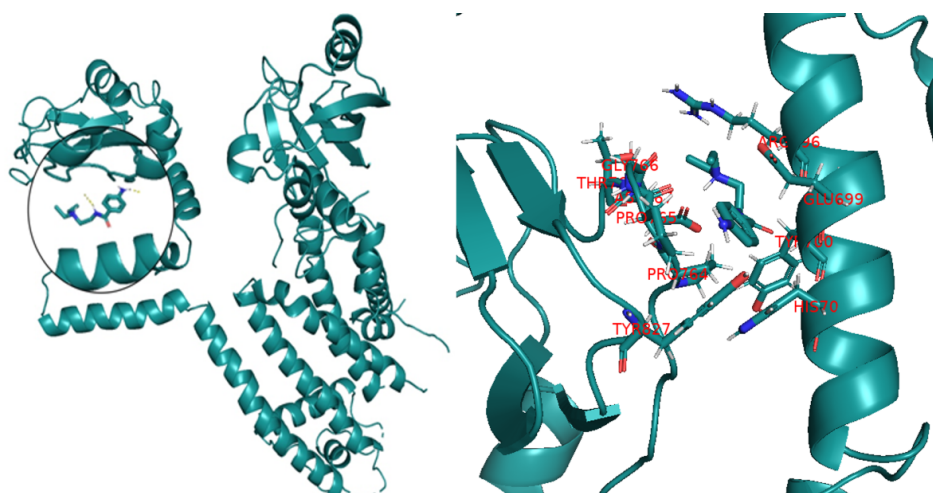
### Case Study

In evaluating the practical application of MVPA-DTI, we predicted 53 candidate drugs targeting the voltage-gated inwardly rectifying potassium channel KCNH2 (hERG). The KCNH2 channel, a critical protein in cardiac electrophysiology, plays a central role in regulating ventricular repolarization. Dysfunction or dysregulation of this channel delays ventricular repolarization, manifesting as QT interval

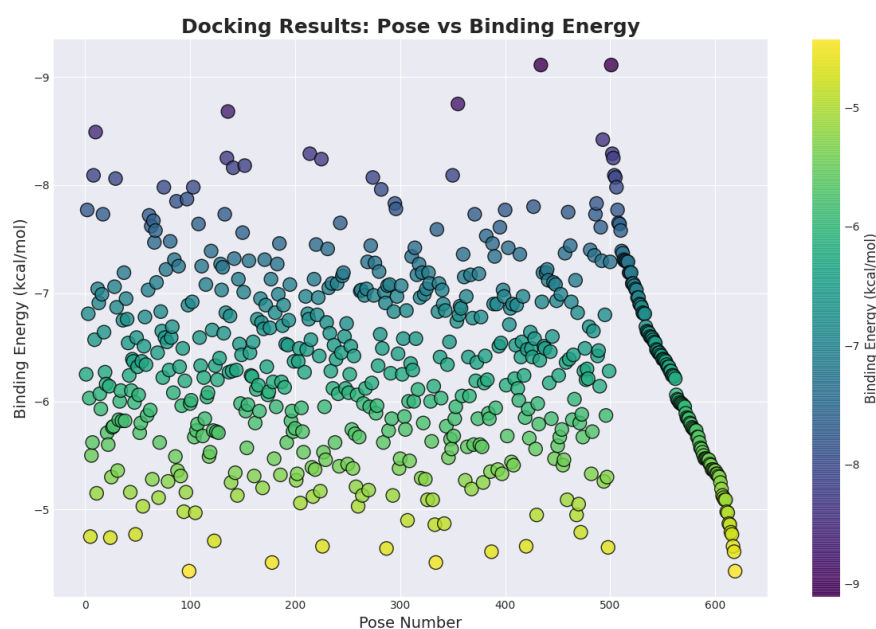
prolongation on electrocardiograms. This phenomenon has been extensively documented to correlate with the pathogenesis of various cardiovascular diseases [49]. The results demonstrate that 38 out of the 53 candidate drugs exhibited potential interactions with KCNH2. Among these, 10 drugs have been experimentally validated in published studies to interact with the KCNH2 channel. Although the remaining candidates have not yet been experimentally confirmed, their established associations with cardiovascular pathologies suggest potential therapeutic relevance in cardiovascular disease management. Studies have demonstrated that procainamide interacts with the KCNH2 channel, primarily by inhibiting its function. This inhibition prolongs the duration of potassium ion efflux, leading to QT interval prolongation [50]. Nicotine, as a blocker of the KCNH2 channel, significantly affects the electrophysiological properties of the heart. This interaction may have important clinical implications for assessing the impact of nicotine on cardiac health [51]. Ranolazine is considered to have potential therapeutic effects, as it improves cardiac electrophysiological abnormalities caused by genetic variations by modulating the function of the KCNH2 channel [52].

In this study, AutoDock was used to perform molecular docking simulations on the interaction between procainamide and KCNH2 channel protein. As shown in Figures 2 and 3, procainamide can stably bind to the central cavity region of the KCNH2 channel protein and is embedded in a hydrophobic pocket surrounded by the S6 helix and the lamellar structure. The binding site involves multiple key amino acid residues, including THR 768, TYR 827, etc. These residues have been widely reported in the literature as core sites for regulating drug binding and gating behavior of KCNH2 channels. These binding residues are visually annotated in Figure 2. Binding mode analysis showed that hydrophobic interactions and potential hydrogen bonds were formed between procainamide and the above residues, which may interfere with the open state of the channel and affect its function. As shown in Figure 3, binding energy analysis showed that procainamide interacted strongly with KCNH2, and the binding energy of some conformations was as low as -9 kcal/mol, indicating that it has high binding stability and potential biological activity.

**Figure 2.** Visualization of the interaction between procainamide and the KCNH2 channel. KCNH2: Potassium Voltage-Gated Channel Subfamily H Member 2.



**Figure 3.** Binding energy distribution of procainamide docking with KCNH2. KCNH2: Potassium Voltage-Gated Channel Subfamily H Member 2.



## Conclusions

This study proposes a novel DTI prediction method MVPA-DTI, which extracts key biological features from protein and drug sequences and reconstructs them into a heterogeneous graph, enabling the model to capture the most critical biological information during each iteration, thereby optimizing the weight assignment for drugs and targets. Experimental results demonstrate that MVPA-DTI outperforms existing methods across multiple benchmark tests. Although MVPA-DTI effectively captures DTI, the biological mechanisms underlying these interactions are complex,

involving multiple factors, which are not yet fully considered by MVPA-DTI. Future improvements should focus on in-depth exploration of the biological details of DTI to enhance prediction accuracy and applicability. To enhance the ability to model the complex relationship between drugs and targets, in the future, it is possible to consider introducing new graph neural network structures such as FCGCN [53] into the drug–target graph modeling process to more effectively integrate molecular structure, pharmacological properties, and network topology information.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grants No. 62076045), the 111 Project (Grants No. D23006), the National Foreign Expert Project of China (Grants No. D20240244), and the Interdisciplinary Project of Dalian University (Grants No. DLUXK-2023-YB-003 and DLUXK-2023-YB-009).

## Conflicts of Interest

None declared.

## References

1. Chu Y, Shan X, Chen T, et al. DTI-MLCD: predicting drug-target interactions using multi-label learning with community detection method. *Brief Bioinform.* May 20, 2021;22(3):bbaa205. [doi: [10.1093/bib/bbaa205](https://doi.org/10.1093/bib/bbaa205)] [Medline: [32964234](https://pubmed.ncbi.nlm.nih.gov/32964234/)]
2. Lan W, Wang J, Li M, et al. Predicting drug-target interaction using positive-unlabeled learning. *Neurocomputing.* Sep 2016;206:50-57. [doi: [10.1016/j.neucom.2016.03.080](https://doi.org/10.1016/j.neucom.2016.03.080)]
3. Liu Z, Chen Q, Lan W, Lu H, Zhang S. SSLDTI: A novel method for drug-target interaction prediction based on self-supervised learning. *Artif Intell Med.* Mar 2024;149:102778. [doi: [10.1016/j.artmed.2024.102778](https://doi.org/10.1016/j.artmed.2024.102778)] [Medline: [38462280](https://pubmed.ncbi.nlm.nih.gov/38462280/)]
4. Lan W, Tang Z, Liu M, et al. The large language models on biomedical data analysis: a survey. *IEEE J Biomed Health Inform.* Jun 2025;29(6):4486-4497. [doi: [10.1109/JBHI.2025.3530794](https://doi.org/10.1109/JBHI.2025.3530794)] [Medline: [40031690](https://pubmed.ncbi.nlm.nih.gov/40031690/)]
5. Shi W, Yang H, Xie L, Yin XX, Zhang Y. A review of machine learning-based methods for predicting drug-target interactions. *Health Inf Sci Syst.* Dec 2024;12(1):30. [doi: [10.1007/s13755-024-00287-6](https://doi.org/10.1007/s13755-024-00287-6)] [Medline: [38617016](https://pubmed.ncbi.nlm.nih.gov/38617016/)]
6. Thafar M, Raies AB, Albaradei S, Essack M, Bajic VB. Comparison study of computational prediction tools for drug-target binding affinities. *Front Chem.* 2019;7:782. [doi: [10.3389/fchem.2019.00782](https://doi.org/10.3389/fchem.2019.00782)] [Medline: [31824921](https://pubmed.ncbi.nlm.nih.gov/31824921/)]
7. Mathai N, Kirchmair J. Similarity-based methods and machine learning approaches for target prediction in early drug discovery: performance and scope. *Int J Mol Sci.* May 19, 2020;21(10):3585. [doi: [10.3390/ijms21103585](https://doi.org/10.3390/ijms21103585)] [Medline: [32438666](https://pubmed.ncbi.nlm.nih.gov/32438666/)]
8. Lian M, Wang X, Du W. Integrated multi-similarity fusion and heterogeneous graph inference for drug-target interaction prediction. *Neurocomputing.* Aug 2022;500:1-12. [doi: [10.1016/j.neucom.2022.04.104](https://doi.org/10.1016/j.neucom.2022.04.104)]
9. Luo Y, Zhao X, Zhou J, et al. A network integration approach for drug-target interaction prediction and computational drug repositioning from heterogeneous information. *Nat Commun.* Sep 18, 2017;8(1):573. [doi: [10.1038/s41467-017-00680-8](https://doi.org/10.1038/s41467-017-00680-8)] [Medline: [28924171](https://pubmed.ncbi.nlm.nih.gov/28924171/)]
10. Yuan Q, Gao J, Wu D, Zhang S, Mamitsuka H, Zhu S. DrugE-Rank: improving drug-target interaction prediction of new candidate drugs or targets by ensemble learning to rank. *Bioinformatics.* Jun 15, 2016;32(12):i18-i27. [doi: [10.1093/bioinformatics/btw244](https://doi.org/10.1093/bioinformatics/btw244)] [Medline: [27307615](https://pubmed.ncbi.nlm.nih.gov/27307615/)]
11. Zhang YF, Wang X, Kaushik AC, et al. SPVec: a Word2vec-inspired feature representation method for drug-target interaction prediction. *Front Chem.* 2019;7(1-11):895. [doi: [10.3389/fchem.2019.00895](https://doi.org/10.3389/fchem.2019.00895)] [Medline: [31998687](https://pubmed.ncbi.nlm.nih.gov/31998687/)]
12. Öztürk H, Ozkirimli E, Özgür A. A comparative study of SMILES-based compound similarity functions for drug-target interaction prediction. *BMC Bioinformatics.* Mar 18, 2016;17:128. [doi: [10.1186/s12859-016-0977-x](https://doi.org/10.1186/s12859-016-0977-x)] [Medline: [26987649](https://pubmed.ncbi.nlm.nih.gov/26987649/)]
13. Wu Z, Pan S, Chen F, Long G, Zhang C, Yu PS. A comprehensive survey on graph neural networks. *IEEE Trans Neural Netw Learn Syst.* Jan 2021;32(1):4-24. [doi: [10.1109/TNNLS.2020.2978386](https://doi.org/10.1109/TNNLS.2020.2978386)] [Medline: [32217482](https://pubmed.ncbi.nlm.nih.gov/32217482/)]
14. Zhang Z, Cui P, Zhu W. Deep learning on graphs: a survey. *IEEE Trans Knowl Data Eng.* 2020;34(1):249-270. [doi: [10.1109/TKDE.2020.2981333](https://doi.org/10.1109/TKDE.2020.2981333)]
15. Zhang Z, Chen L, Zhong F, et al. Graph neural network approaches for drug-target interactions. *Curr Opin Struct Biol.* Apr 2022;73:102327. [doi: [10.1016/j.sbi.2021.102327](https://doi.org/10.1016/j.sbi.2021.102327)] [Medline: [35074533](https://pubmed.ncbi.nlm.nih.gov/35074533/)]
16. Singh S, Malik BK, Sharma DK. Molecular drug targets and structure based drug design: a holistic approach. *Bioinformation.* Dec 23, 2006;1(8):314-320. [doi: [10.6026/97320630001314](https://doi.org/10.6026/97320630001314)] [Medline: [17597912](https://pubmed.ncbi.nlm.nih.gov/17597912/)]
17. Thafar MA, Olayan RS, Ashoor H, et al. DTiGEMS+: drug-target interaction prediction using graph embedding, graph mining, and similarity-based techniques. *J Cheminform.* Jun 29, 2020;12(1):44. [doi: [10.1186/s13321-020-00447-2](https://doi.org/10.1186/s13321-020-00447-2)] [Medline: [33431036](https://pubmed.ncbi.nlm.nih.gov/33431036/)]
18. Shim J, Hong ZY, Sohn I, Hwang C. Prediction of drug-target binding affinity using similarity-based convolutional neural network. *Sci Rep.* Feb 24, 2021;11(1):4416. [doi: [10.1038/s41598-021-83679-y](https://doi.org/10.1038/s41598-021-83679-y)] [Medline: [33627791](https://pubmed.ncbi.nlm.nih.gov/33627791/)]
19. Zhang P, Yuan J, Li L, et al. Key substructure learning with chemical intuition for material property prediction[c]. Presented at: Proceedings of the International Conference on Database Systems for Advanced Applications; 87-103; Pohang, Korea. 2024.[doi: [10.1007/978-981-97-5575-2\\_6](https://doi.org/10.1007/978-981-97-5575-2_6)]
20. Zhang P, Che C, Jin B, Yuan J, Li R, Zhu Y. NCH-DDA: Neighborhood contrastive learning heterogeneous network for drug-disease association prediction. *Expert Syst Appl.* Mar 2024;238:121855. [doi: [10.1016/j.eswa.2023.121855](https://doi.org/10.1016/j.eswa.2023.121855)]



21. Zhao BW, Su XR, Yang Y, et al. Regulation-aware graph learning for drug repositioning over heterogeneous biological network. *Inf Sci (Ny)*. Jan 2025;686:121360. [doi: [10.1016/j.ins.2024.121360](https://doi.org/10.1016/j.ins.2024.121360)]
22. Zhao BW, Wang L, Hu PW, et al. Fusing higher and lower-order biological information for drug repositioning via graph representation learning. *IEEE Trans Emerg Topics Comput*. 2023;12(1):163-176. [doi: [10.1109/TETC.2023.3239949](https://doi.org/10.1109/TETC.2023.3239949)]
23. Su X, Hu P, Yi H, You Z, Hu L. Predicting drug-target interactions over heterogeneous information network. *IEEE J Biomed Health Inform*. 2022;27(1):562-572. [doi: [10.1109/JBHI.2022.3219213](https://doi.org/10.1109/JBHI.2022.3219213)]
24. Zhao BW, Su XR, Hu PW, Huang YA, You ZH, Hu L. iGRLDTI: an improved graph representation learning method for predicting drug-target interactions over heterogeneous biological information network. *Bioinformatics*. Aug 1, 2023;39(8):btad451. [doi: [10.1093/bioinformatics/btad451](https://doi.org/10.1093/bioinformatics/btad451)] [Medline: [37505483](https://pubmed.ncbi.nlm.nih.gov/37505483/)]
25. Öztürk H, Özgür A, Ozkirimli E. DeepDTA: deep drug-target binding affinity prediction. *Bioinformatics*. Sep 1, 2018;34(17):i821-i829. [doi: [10.1093/bioinformatics/bty593](https://doi.org/10.1093/bioinformatics/bty593)] [Medline: [30423097](https://pubmed.ncbi.nlm.nih.gov/30423097/)]
26. Pu L, Govindaraj RG, Lemoine JM, Wu HC, Brylinski M. DeepDrug3D: classification of ligand-binding pockets in proteins with a convolutional neural network. *PLoS Comput Biol*. Feb 2019;15(2):e1006718. [doi: [10.1371/journal.pcbi.1006718](https://doi.org/10.1371/journal.pcbi.1006718)] [Medline: [30716081](https://pubmed.ncbi.nlm.nih.gov/30716081/)]
27. Devlin J, Chang MW, Lee K, Toutanova K. Bert: pre-training of deep bidirectional transformers for language understanding. Presented at: Proceedings of the Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers); Jun 2, 2019;4171-4186; Minneapolis, Minnesota. 2019.[doi: [10.18653/v1/N19-1423](https://doi.org/10.18653/v1/N19-1423)]
28. Li J, Jiang X. Mol-BERT: an effective molecular representation with BERT for molecular property prediction. *Wireless Commun Mobile Comput*. Jan 2021;2021(1):7181815. URL: <https://onlinelibrary.wiley.com/toc/6302/2021/1> [doi: [10.1155/2021/7181815](https://doi.org/10.1155/2021/7181815)]
29. Chithrananda S, Grand G, Ramsundar B. ChemBERTa: large-scale self-supervised pretraining for molecular property prediction. *arXiv*. Preprint posted online on 2020. [doi: [10.48550/arxiv.2010.09885](https://doi.org/10.48550/arxiv.2010.09885)]
30. Brandes N, Ofer D, Peleg Y, Rappoport N, Linial M. ProteinBERT: a universal deep-learning model of protein sequence and function. *Bioinformatics*. Apr 12, 2022;38(8):2102-2110. [doi: [10.1093/bioinformatics/btac020](https://doi.org/10.1093/bioinformatics/btac020)] [Medline: [35020807](https://pubmed.ncbi.nlm.nih.gov/35020807/)]
31. Rao R, Bhattacharya N, Thomas N, et al. Evaluating protein transfer learning with TAPE. *Adv Neural Inf Process Syst*. Dec 2019;32:9689-9701. [doi: [10.1101/676825](https://doi.org/10.1101/676825)] [Medline: [33390682](https://pubmed.ncbi.nlm.nih.gov/33390682/)]
32. Elnaggar A, Heinzinger M, Dallago C, et al. ProtTrans: toward understanding the language of life through self-supervised learning. *IEEE Trans Pattern Anal Mach Intell*. Oct 2022;44(10):7112-7127. [doi: [10.1109/TPAMI.2021.3095381](https://doi.org/10.1109/TPAMI.2021.3095381)] [Medline: [34232869](https://pubmed.ncbi.nlm.nih.gov/34232869/)]
33. Maziarka Ł, Danel T, Mucha S, et al. Molecule attention transformer. *arXiv*. Preprint posted online on 2020. [doi: [10.48550/arXiv.2002.08264](https://doi.org/10.48550/arXiv.2002.08264)]
34. Toropov AA, Toropova AP, Mukhamedzhanova DV, Gutman I. Simplified molecular input line entry system (SMILES) as an alternative for constructing quantitative structure-property relationships (QSPR). Preprint posted online on 2005. URL: <http://nopr.niscair.res.in/bitstream/123456789/18068/1/IJCA%2044A%288%29%201545-1552.pdf> [Accessed 2025-09-15]
35. Wang X, Ji H, Shi C, et al. Heterogeneous graph attention network. Presented at: WWW '19; May 13, 2019; San Francisco CA USA. May 13, 2019.URL: <https://dl.acm.org/doi/proceedings/10.1145/3308558> [doi: [10.1145/3308558.3313562](https://doi.org/10.1145/3308558.3313562)]
36. Bento AP, Hersey A, Félix E, et al. An open source chemical structure curation pipeline using RDKit. *J Cheminform*. Sep 1, 2020;12(1):51. [doi: [10.1186/s13321-020-00456-1](https://doi.org/10.1186/s13321-020-00456-1)] [Medline: [33431044](https://pubmed.ncbi.nlm.nih.gov/33431044/)]
37. Vaswani A. Attention is all you need. Presented at: Advances in neural information processing systems; Dec 4-9, 2017; Long Beach, California, USA.
38. Raffel C, Shazeer N, Roberts A, et al. Exploring the limits of transfer learning with a unified text-to-text transformer. *J Mach Learn Res*. 2020;21(1):5485-5551.
39. Ma N, Zhang X, Liu M, Sun J. Activate or not: learning customized activation. Presented at: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); Jun 20-25, 2021; Nashville, TN, USA. [doi: [10.1109/CVPR46437.2021.00794](https://doi.org/10.1109/CVPR46437.2021.00794)]
40. Knox C, Law V, Jewison T, et al. DrugBank 3.0: a comprehensive resource for “omics” research on drugs. *Nucleic Acids Res*. Jan 2011;39(Database issue):D1035-41. [doi: [10.1093/nar/gkq1126](https://doi.org/10.1093/nar/gkq1126)] [Medline: [21059682](https://pubmed.ncbi.nlm.nih.gov/21059682/)]
41. Wang Z, Zhang Q, Shuang-Wei H, et al. Multi-level protein structure pre-training via prompt learning. Presented at: Proceedings of the The Eleventh International Conference on Learning Representations; 2023; Kigali, Rwanda. URL: <https://openreview.net/forum?id=XGagtiJ8XC> [Accessed 2025-09-15]

42. Davis AP, Murphy CG, Johnson R, et al. The comparative toxicogenomics database: update 2013. *Nucleic Acids Res.* Jan 2013;41(Database issue):D1104-14. [doi: [10.1093/nar/gks994](https://doi.org/10.1093/nar/gks994)] [Medline: [23093600](https://pubmed.ncbi.nlm.nih.gov/23093600/)]
43. Kuhn M, Campillos M, Letunic I, Jensen LJ, Bork P. A side effect resource to capture phenotypic effects of drugs. *Mol Syst Biol.* 2010;6(1):343. [doi: [10.1038/msb.2009.98](https://doi.org/10.1038/msb.2009.98)] [Medline: [20087340](https://pubmed.ncbi.nlm.nih.gov/20087340/)]
44. Zheng X, Ding H, Mamitsuka H, Zhu S. Collaborative matrix factorization with multiple similarities for predicting drug-target interactions. Presented at: KDD' 13; Aug 11, 2013; Chicago Illinois USA. Aug 11, 2013. URL: <https://dl.acm.org/doi/proceedings/10.1145/2487575> [doi: [10.1145/2487575.2487670](https://doi.org/10.1145/2487575.2487670)]
45. Wang W, Yang S, Zhang X, Li J. Drug repositioning by integrating target information through a heterogeneous network model. *Bioinformatics.* Oct 15, 2014;30(20):2923-2930. [doi: [10.1093/bioinformatics/btu403](https://doi.org/10.1093/bioinformatics/btu403)] [Medline: [24974205](https://pubmed.ncbi.nlm.nih.gov/24974205/)]
46. Wan F, Hong L, Xiao A, Jiang T, Zeng J. NeoDTI: neural integration of neighbor information from a heterogeneous network for discovering new drug-target interactions. *Bioinformatics.* Jan 1, 2019;35(1):104-111. [doi: [10.1093/bioinformatics/bty543](https://doi.org/10.1093/bioinformatics/bty543)] [Medline: [30561548](https://pubmed.ncbi.nlm.nih.gov/30561548/)]
47. Peng J, Wang Y, Guan J, et al. An end-to-end heterogeneous graph representation learning-based framework for drug-target interaction prediction. *Brief Bioinform.* Sep 2, 2021;22(5):bbaa430. [doi: [10.1093/bib/bbaa430](https://doi.org/10.1093/bib/bbaa430)] [Medline: [33517357](https://pubmed.ncbi.nlm.nih.gov/33517357/)]
48. Yao K, Wang X, Li W, et al. Semi-supervised heterogeneous graph contrastive learning for drug-target interaction prediction. *Comput Biol Med.* Sep 2023;163:107199. [doi: [10.1016/j.combiomed.2023.107199](https://doi.org/10.1016/j.combiomed.2023.107199)] [Medline: [37421738](https://pubmed.ncbi.nlm.nih.gov/37421738/)]
49. Itoh T, Tanaka T, Nagai R, et al. Genomic organization and mutational analysis of HERG, a gene responsible for familial long QT syndrome. *Hum Genet.* Apr 1998;102(4):435-439. [doi: [10.1007/s004390050717](https://doi.org/10.1007/s004390050717)] [Medline: [9600240](https://pubmed.ncbi.nlm.nih.gov/9600240/)]
50. Chiu PJS, Marcoe KF, Bounds SE, et al. Validation of a [3H]astemizole binding assay in HEK293 cells expressing HERG K<sup>+</sup> channels. *J Pharmacol Sci.* Jul 2004;95(3):311-319. [doi: [10.1254/jphs.fpe0040101](https://doi.org/10.1254/jphs.fpe0040101)] [Medline: [15272206](https://pubmed.ncbi.nlm.nih.gov/15272206/)]
51. Yang B feng, Xu D hui, Xu C qian, et al. Inactivation gating determines drug potency: a common mechanism for drug blockade of HERG channels. *Acta Pharmacol Sin.* May 2004;25(5):554-560. [Medline: [15132818](https://pubmed.ncbi.nlm.nih.gov/15132818/)]
52. Smith JL, Reloj AR, Nataraj PS, et al. Pharmacological correction of long QT-linked mutations in KCNH2 (hERG) increases the trafficking of Kv11.1 channels stored in the transitional endoplasmic reticulum. *Am J Physiol Cell Physiol.* Nov 1, 2013;305(9):C919-30. [doi: [10.1152/ajpcell.00406.2012](https://doi.org/10.1152/ajpcell.00406.2012)] [Medline: [23864605](https://pubmed.ncbi.nlm.nih.gov/23864605/)]
53. Yang Y, Li G, Li D, Zhang J, Hu P, Hu L. Integrating fuzzy clustering and graph convolution network to accurately identify clusters from attributed graph. *IEEE Trans Netw Sci Eng.* 2024;12(2):1112-1125. [doi: [10.1109/TNSE.2024.3524077](https://doi.org/10.1109/TNSE.2024.3524077)]

## Abbreviations

**DruSF:** drug compound structure feature extraction module  
**DTI:** drug-target interaction  
**HAN:** Heterogeneous Graph Attention Network  
**HN:** Heterogeneous Network  
**LLM:** large language model  
**MAT:** Molecule Attention Transformer  
**ProSF:** protein sequence feature extraction module.

*Edited by Qiao Jin; peer-reviewed by Dongjiang Niu, Lun Hu, Wei Lan, Zhizheng Wang; submitted 27.03.2025; final revised version received 11.07.2025; accepted 25.07.2025; published 02.10.2025*

*Please cite as:*

Zhao H, Yao K, Liu Y, Che C, Tang L

Heterogeneous Network With Multiview Path Aggregation: Drug-Target Interaction Prediction Study Design

*JMIR Med Inform* 2025;13:e74974

URL: <https://medinform.jmir.org/2025/1/e74974>

doi: [10.2196/74974](https://doi.org/10.2196/74974)

© Haixue Zhao, Kui Yao, Yunjong Liu, Chao Che, Lin Tang. Originally published in *JMIR Medical Informatics* (<https://medinform.jmir.org>), 02.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Medical Informatics*, is properly cited. The complete bibliographic information, a link to the original publication on <https://medinform.jmir.org>, as well as this copyright and license information must be included.