BMC Bioinformatics

RESEARCH

Open Access

MFCADTI: improving drug-target interaction prediction by integrating multiple feature through cross attention mechanism

Na Quan[1], Shicheng Ma[2], Kai Zhao[1], Xuehua Bi[3*] and Linlin Zhang[2,4*]

*Correspondence:
bxh0327@foxmail.com;
zllnadasha@xju.edu.cn

[1] College of Computer Science and Technology, Xinjiang University, Urumqi 830046, China
[2] College of Software, Xinjiang University, Urumqi 830046, China
[3] Medical Engineering and Technology College, Xinjiang Medical University, Urumqi 830017, China
[4] Center of Network and Information Technology, Xinjiang University, Urumqi 830046, China

Abstract

Accurately identifying potential drug-target interactions (DTIs) is a critical step in drug discovery. Multiple heterogeneous biological data provide abundant features for DTI prediction. Many computational methods have been proposed based on these data. However, most of these methods either extract features from sequences or from networks, utilizing only one aspect of the characteristics of drugs and targets, neglecting the complementary information between these two types of features. In fact, integrating different types of features will provide more valuable information for DTI prediction. In this article, we propose a novel method to improve the predictive capability for DTIs, named MFCADTI, by integrating multi-source feature through cross-attention mechanisms. The method extracts network topological features from the heterogeneous network and attribute features from sequences of drugs and targets. Considering the complementarity and heterogeneity between network and attribute features, cross-attention mechanisms are used to integrate the network and attribute features of drugs and targets. To capture the correlations between drugs and targets, cross-attention is used to learn the interaction features of each drug-target pair. We evaluate MFCADTI on two datasets and experimental results demonstrate a significant improvement in the performance of MFCADTI compared to state-of-the-art methods. Finally, case studies illustrate that MFCADTI is an effective DTI prediction way that provides valuable guidance for drug development. The data and source code used in this study are available at: https://github.com/Dejavun/MFCADTI.

Keywords: Drug-target interaction prediction, Network feature, Attribute feature, Cross-attention, Feature fusion

Introduction

Drug-target interaction (DTI) refers to the binding of a drug to a target location that results in a change in its behavior/function [1]. The prediction of drug-target interactions (DTIs) is a crucial step in the process of drug development or repositioning, aiming to identify potential new drugs or new targets for existing drugs [2]. Accurately identifying potential DTIs through a computational approach can reduce the cost and time required for drug development. Early DTI prediction methods can be classified into

ligand-based methods and docking-based methods [3]. The former uses the similarity between ligands and target proteins to predict DTIs [4], while the latter relies on the 3D structure and dynamic simulations of target proteins to estimate the probability of DTI [5]. However, these methods face challenges in practical applications due to the lack of ligands and the 3D structures of targets [6]. Advances in computational methods for DTI prediction is essential for accelerating drug discovery. Machine learning-based methods have been proposed with striking results, which can be further categorized into feature-based methods and network-based methods.

In feature-based methods [7–12], researchers aim to extract the inherent biological characteristics of drug targets and identify DTIs by utilizing machine learning models. Drug compounds can be represented as 1D sequences, chemical graphs, or fingerprints like Extended-Connectivity Fingerprints (ECFPs). Target sequences are described as 1D sequences and transformed into various descriptors, such as Protein Sequence Composition (PSC) descriptors. For instance, idti-MLKdr [7] applied three-dimensionality reduction techniques to extract features from the sequences of target proteins and the Simplified Molecular Input Line Entry System (SMILES), and employed Support Vector Machines (SVM) to predict DTIs. Wang et al. [8] adopted stacked autoencoders to learn features based on the Position-Specific Scoring Matrix (PSSM) of target sequences. They combined these features with drug molecular substructure fingerprint data to construct feature descriptors and inputted them into a Random Forest (RF) model for precise prediction. PUDT [9] integrated multiple target resources including target structure information, target function category information and target function annotation information. DeepConv-DTI [10] used a MultiLayer Perceptron (MLP) on molecular fingerprints and a 1D-CNN on sequences of target proteins to obtain features of drugs and targets, and then predicted their association probabilities by using Fully Connected Layers (FCL). The strengths of the feature-based methods lie in their foundation on biologically relevant features, displaying robust scalability. Feature-based methods are well suited to the situation where the amount of data is small, but the representation of the drug and target is limited, as the descriptors tend to focus on only one aspect of the property.

Network-based methods [13–17] refer to methods that utilize graph-based techniques on networks containing drugs, targets and other biological entities to predict DTIs based on existing connections in the networks. Cheng et al. [13] used Network-Based Inference (NBI) to derive novel drug-target interactions solely based on the bipartite network topology similarity of drugs and targets. Recently, an increasing number of data sources related to drugs and targets have emerged. Researchers have constructed bioinformatics networks by integrating heterogeneous data from drugs, targets, diseases, side effects, and more. These networks provide diverse insights and multiple perspectives for predicting DTIs. For instance, DTINet [14] employed Inductive Matrix Completion (IMC) to predict DTIs based on the learned features of drugs and targets by integrating data from heterogeneous data sources including drugs, targets, disorders, and side effects. SSLDTI [18]combined graph autoencoder and self-supervised learning to accurately encode multilevel features of graphs using only a small number of labelled samples. DeepDTnet [16] presented a novel network-based method for target identification and drug repurposing, which systematically

embedded 15 types of chemical, genomic, phenotypic, and cellular networks to predict new molecular targets among known drugs under a PU-learning framework successfully. Based on the heterogeneous network related to drug and target, GCNDTI [17] adopted the graph convolutional neural network to learn the low-dimensional topological vector of drug and target nodes, and obtained the probability score between them through vector space projection. Network-based methods enhance the accuracy of DTI prediction by capturing similar information across different kinds of biological networks as the features of drugs and targets [19]. However, they heavily rely on the topological structure of nodes in heterogeneous networks for feature representation, without delving into the biological structural information of drugs and targets, which is the determining factors in drug-target interactions.

Recent studies [20–23] have shown that considering additional biological knowledge is essential for network-based interaction prediction, expecially in DTI prediction. For example, Ji et al. [20] used network representation methods to obtain behavioral features of drug and protein nodes, and then combined their intrinsic attribute features (e.g., drug molecular fingerprints and protein sequence) using a random forest classifier for training and prediction. Kg-mtl [21] extracts drug and target features from knowledge graphs and molecular graphs in a collaborative manner, designing an effective shared unit that integrates semantic relationships between drug entities and compound structures within both knowledge graphs and molecular graphs for the prediction of drug-target interactions. Nevertheless, these studies struggle to capture the intricate interactions between drugs and targets by simply concatenating their features or merely calculating the distances and lack of adequate modeling on the underlying mechanisms of DTI. Cross-attention is initially introduced as an attention mechanism in Transformers, allowing each position in the decoder to cover the whole positions in the input sequence [24], enabling better utilization of the input sequence data. Subsequently, it has found widespread applications in various fields, including natural language processing (NLP) [25], computer vision (CV) [26], and bioinformatics [27, 28]. These applications have demonstrated the efficacy of cross-attention in establishing interactions between different input sequences to fully capitalize on their correlations and enhance model performance in handling diverse tasks involving multimodal data.

To address the above problems and inspired by cross-attention, we propose MFCADTI, a method that fuses the network feature and attribute feature of drugs and targets using cross-attention to improve the predictive capability for DTIs. Firstly, MFCADTI extracts multiple features of drugs and targets from heterogeneous data sources, which are network features from heterogeneous network and attribute features from sequence. Specifically, we integrate the known associations between drugs, targets, diseases, and side effects in the dataset to construct a heterogeneous network. Secondly, we employ the Large-scale Information Network Embedding (LINE) [29] to extract network features of drugs and targets from the heterogeneous network and use the Frequent Continuous Subsequence (FCS) [30] to obtain attribute features based on the SMILES of drugs and the (AA) sequences of the targets. Thirdly, Cross-attention mechanisms are used respectively to integrate the network and attribute features of drugs and targets. Subsequently, cross-attention is also applied to capture the interaction features between drugs and targets. At last, the final interaction feature representations are fed into FCL to predict the

DTIs. Experimental results demonstrate that MFCADTI achieves better performance compared to baseline methods.

Materials and methods

Datasets

We adopt two publicly available heterogeneous network datasets proposed by Luo et al. [14] and Zeng et al. [31] (named Luo_data and Zeng_data, respectively) as experiment datasets. We describe the construction of the datasets in detail in the Supplementary materials. After removing duplicated and isolated nodes, the experiment data for MFCADTI are obtained, and the details are shown in Table 1.

To extract the attribute features of drugs and targets, we further obtain the SMILES sequences of drugs and the amino acid (AA) sequences of target proteins in the heterogeneous graph from the Drugbank database using Drugbank IDs and UniProt IDs. Sequence can be downloaded from the PubChem [32] and UniProt [33] databases when such data is not found in the Drugbank database.

The proposed method

The MFCADTI framework consists of three parts: network feature extraction, attribute feature extraction, and cross-attention feature fusion and prediction. As shown in Fig. 1, the network and attribute features of drugs and targets are obtained through network feature extraction and attribute feature extraction, respectively. Subsequently, cross-attention feature fusion is applied to these two feature types. Finally, the resulting interaction feature representations are used as input for a FCL to perform prediction.

Construction of heterogeneous network

We construct a heterogeneous biological network $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the node sets, $\mathcal{E}$ is the edge sets, edge $(v_i, v_j) \in \mathcal{E}$ connects a pair of nodes $v_i, v_j \in \mathcal{V}$. In the heterogeneous network, there are four type nodes which are the drug, target, disease, and side effect, and six type edges including drug-drug interactions, drug-target interactions, drug-disease associations, drug-side effect associations, target-target interactions, and target-disease associations.

Network feature extraction of drugs and targets

We employ the LINE to learn the network feature representations of drugs and targets. LINE can map closely connected nodes in large networks to similar positions in a low-dimensional vector space by considering the first-order and second-order similarity of the node.

When LINE computes the first-order similarity, the probability of node v_j being a neighbor of v_i for each edge (v_i, v_j) is defined as follows:

$$P_1(v_j|v_i) = \frac{1}{1 + exp\big(-\vec{u}_j^T \cdot \vec{u}_i\big)}, \tag{1}$$

where $\vec{u}_i \in \mathbb{R}^e$ and $\vec{u}_j \in \mathbb{R}^e$ are low-dimensional vector representation of node v_i and node v_j, respectively. e is the feature dimension, which is defined by experiment.

According to the weights of the edges, the empirical distribution is obtained:

Table 1 Number of nodes and edges in the two datasets used for the MFCADTI

Datasets	Nodes					Edges						
	Drug	Target	Disease	Se	Total	E_1	E_2	E_3	E_4	E_5	E_6	Total
Luo_data	708	1493	5442	3828	11,471	1923	10,036	7363	199,214	1,596,745	801,624	1,895,445
Zeng_data	726	1894	435	11,848	14,903	4978	132,768	16,133	1208	23,080	263,805	441,972

Se: Side effect, E_1: Drug-target interactions, E_2: Drug-drug interactions, E_3: Target-target interactions, E_4: Drug-disease associations, E_5: Target-disease associations, E_6: Drug-side effect associations

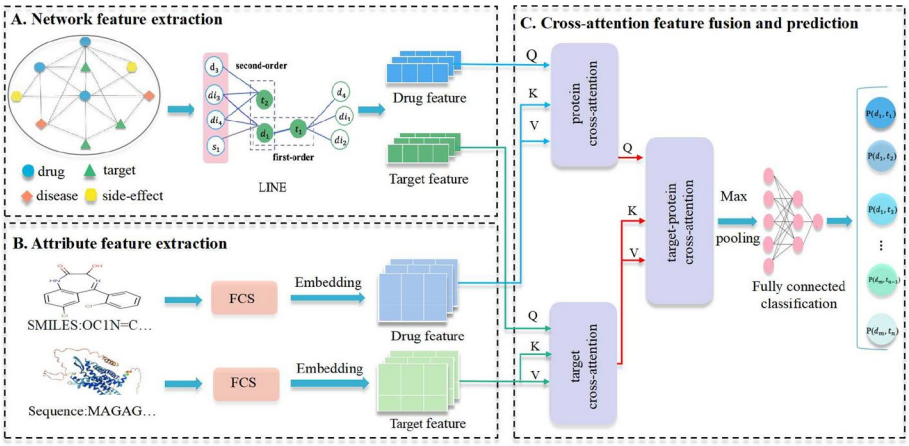

Fig. 1 The framework of MFCADTI. The MFCADTI framework consists of three parts: network feature extraction, attribute feature extraction, and cross-attention feature fusion and prediction

$$\hat{P}_1(v_i, v_j) = \frac{\omega_{ij}}{W}, \tag{2}$$

where ω_{ij} is the weight of the edge (v_i, v_j), and W is the sum of edge weights. To ensure the first-order similarity of nodes, the Kullback–Leibler (KL) divergence is used to measure the similarity between the empirical and the probabilistic distribution. The objective function is defined as follows:

$$L_1 = -\sum_{(v_i,v_j)\in\mathcal{E}} \omega_{ij} \log P_1(v_j|v_i), \tag{3}$$

every node can be represented in the e-dimensional space by finding the $\vec{u}_i \in \mathbb{R}^e$ that minimizes the objective function.

When considering the second-order similarity of nodes, the probability that the node v_j is a neighbor of v_i is:

$$P_2(v_j|v_i) = \frac{exp(-\vec{u}_j^{'T} \cdot \vec{u}_i)}{\sum_{k=1}^{|N|} exp(-\vec{u}_k^{'T} \cdot \vec{u}_i)}, \tag{4}$$

where $|N|$ is the number of vertices in the heterogeneous network, $\vec{u}_i^{'}$ is the embedding vector representation of v_i when it serves as a specific "context".

$$\hat{P}_2(v_i, v_j) = \frac{\omega_{ij}}{d_i}, \tag{5}$$

where d_i is the output-degree of the node v_i. By optimizing the following objective function, the second-order similarity of nodes in the network is preserved:

$$L_2 = -\sum_{(v_i,v_j)\in\mathcal{E}} \omega_{ij} \log\left(P_2(v_j|v_i)\right), \tag{6}$$

every node can be represented a e-dimensional vector $\vec{u}_i$ by learning the $\vec{u}_i$ and $\vec{u}_i^{'}$ that minimizes this objective.

Therefore, we obtain every drug's network feature $drug_{net} \in \mathbb{R}^e$ and every target's network feature $target_{net} \in \mathbb{R}^e$ after using LINE algorithm, here e represents the feature dimension.

Attribute feature extraction of drugs and targets

Following the work [11], we employ the FCS to extract high-quality and appropriately sized substructure embeddings of drugs and targets that contribute to their interactions.

Driven by the knowledge that DTI occurs at the substructure level, FCS decomposes drug and target sequences into sub-sequences, individual atoms, and amino acid sequences according to the identification method of the sub-word based on NLP [34]. We use the FCS algorithm to convert the input drugs' SMILES and targets' AA into a series of explicit subsequences. The generated drug and target subsequences are then embedded into attribute feature matrices using nn.Embedding with the addition of positional information during the embedding process. We follow the work [30] to set the maximum length of sequence for the drug to be 50 and the target to be 545. By using the nn.Embedding, we obtain an attribute feature matrix $Drug_{attr} \in \mathbb{R}^{50\times e}$ for each drug, and an attribute feature matrix $Target_{attr} \in \mathbb{R}^{545\times e}$ for each target. We conduct a statistical analysis on the distribution of the SMILES lengths of drugs and the sequence lengths of targets in the two datasets (see Supplementary Figs. S1–S4). The statistical results also indicate that the choice of 50 for drug SMILES length and 545 for AA sequence length is appropriate for our work. Here the attribute feature dimension e is the same as the network feature dimension, we set the dimension e as 64 and the analysis result is shown in Supplementary Fig. S5.

Cross-attention feature fusion

To obtain efficient and unified representations of drugs and targets, we integrate $drug_{net}$ with $Drug_{attr}$ and $target_{net}$ with $Target_{attr}$ using the cross-attention mechanism, respectively. The detailed structure of the cross-attention module is illustrated in Fig. 1(C).

In drug cross-attention, an enhanced feature representation of the drug is derived by fusing the network features and attribute features. It can be formulated as follows:

$$Q = drug_{net} W_X^Q, K = Drug_{attr} W_X^K, V = Drug_{attr} W_X^V, \tag{7}$$

$$Drug_{enh} = CrossAttetion(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d/h}}\right) \cdot V, \tag{8}$$

where the query matrix Q is computed from the drug's network feature matrix $drug_{net}$, while the key matrix K and value matrix V are provided by the drug's attribute feature $Drug_{attr}$. W_X^Q, W_X^K, and W_X^V are the learnable weight matrices. h and d are the number of heads of attention and the embedding dimension, respectively. Here we set the number of heads h as 6, which is experimentally validated to be the optimal (the result is available in Supplementary Fig. S6).

Similarly, in the target cross-attention, an enhanced feature representation of the target can be expressed as follows:

$$Q = target_{net} W_X^Q, K = Target_{attr} W_X^K, V = Target_{attr} W_X^V, \tag{9}$$

$$Target_{enh} = CrossAttetion\big(Q, K, V\big). \tag{10}$$

Finally, we fuse the feature representation $Drug_{enh}$ and $Target_{enh}$ using a cross-attention mechanism. The final interaction matrix can be expressed as follows:

$$Q = Drug_{enh} W_X^Q, K = Target_{enh} W_X^K, V = Target_{enh} W_X^V, \tag{11}$$

$$Interaction_{(d,t)} = CrossAttetion\big(Q, K, V\big). \tag{12}$$

DTI prediction

We initially apply the max-pooling to $Interaction_{(d,t)}$ to obtain the final interaction features. Subsequently, these features are fed into a FCL comprising two sets of linear transformation layers (with ReLU activation and dropout layers) followed by a sigmoid function. The final prediction score for drug-target interactions can be obtained after training the model with the cross-entropy loss. The cross-entropy loss function is as follows:

$$loss = -[y \log(\hat{y}) + (1 - y) \log(1 - (\hat{y}))], \tag{13}$$

where y is the ground truth label, $\hat{y}$ is the predicted label.

Experiments and results

Experiment setup

In this study, we investigate the performance of MFCADTI under varying experiment settings to identify optimal parameter configurations. The experiments are conducted on Luo_data dataset to evaluate the model's ability to predict DTIs. We use the known drug-target associations in the dataset as positive samples and randomly select the same number of unknown associations of DTIs as negative samples, and Liu et al's method [35] is adopted to further divide the DTI dataset into three subsets: the training set, validation set, and testing set in a 0.855: 0.045: 0.1, respectively. The key parameters explored in our experiments include the dimension of embeddings and the number of attention heads. We describe the performances of MFCADTI with different parameter settings in detail in the Supplementary materials (Parameter analysis).

To evaluate the predictive performance of the proposed model, we use five performance metrics: Accuracy (Acc), Precision, F_1, Area Under the Receiver Operating Characteristic curve (AUC), and Area Under the Precision-Recall curve (AUPR). We perform five independent runs to train the model and evaluate its performance on the test dataset. The statistical metrics are calculated by taking the average and standard deviation over the seven models, the best result is displayed in bold.

Performance comparison with baselines

We compare our method with seven baselines: drugBAN [36], MCL-DTI [37], ICAN [11], HyperAttentionDTI [12], Ji et al.'s model [20], DeepDTnet [16] and drugMAN [38],

the AUC and AUPR results are shown in Fig. 2. The figure shows that MFCADTI obtains the highest average AUC of 0.948 on Luo_data, which is higher than the second best method drugMAN by 2.6%. The average AUPR of MFCADTI is 0.927 in Luo_data. It is higher than that of drugMAN, DeepDTnet, Ji et al.'s model, HyperAttentionDT, ICAN, MCL-DTI and drugBAN by 0.1%, 1.3%, 1.7%, 2.4%, 3.0%, 3.7% and 4.4%, respectively. Compared with the other competitive baselines in Zeng_data, MFCADTI also improves at least 4% and 2.8% on average in terms of AUC and AUPR. It is worth mentioning that, compared with the method HyperAttentionDTI that also uses the attention mechanism, our method has achieved significant improvement on two datasets. This indicates that our adoption of the cross-attention mechanism between the sequence and network features of drugs and targets is a very effective feature fusion strategy. The comparison results on ACC, Precision, and F1 are listed in Tabels S1 and S2 of the Supplementary material.

From the results, it can be seen that our proposed method, along with methods MultiDTI and Ji et al.'s model, takes into account both network and attribute features, generally outperform methods that rely solely on either feature-based or network-based alone. This is because the features of drugs and targets are enriched comprehensively. However, it should be noted that despite considering both network and attribute features, Ji et al.'s model exhibits lower AUC values compared to deepDTnet, a model solely based on network features. This difference can be attributed to the fact that Ji et al.'s model simply converts drug SMILES and target sequences into features without effectively learning rich attribute features. Furthermore, it merely concatenates drug and target features without effectively capturing the intricate relationships between drugs and targets compared with MFCADTI. The above experimental results highlight MFCADTI's superiority in predicting DTIs based on both network and attribute features.

Impact of the cross-attention module

To explore the impact of the cross-attention module on MFCADTI, we explore five different variants of joint drug-targe representation on Luo_data dataset. These variants include *concate+cross*, which combines attributes and network features of drugs/targets through concatenation followed by applying cross-attention to capture the interaction between the features of drug and target; *cross+concate*, which directly employs cross-attention on drugs'/targets' attributes and network features before concatenating

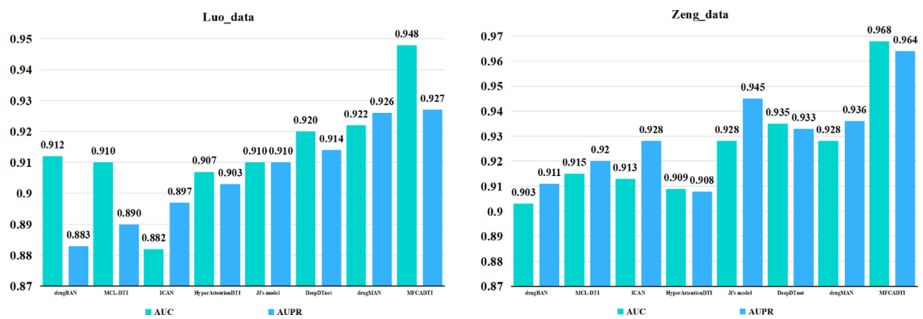

Fig. 2 Performance comparison with baselines

Table 2 Impact of the cross-attention module on Luo_data dataset

Variants	Acc	Precision	F_1	AUC	AUPR
Concate+cross	0.859 ± 0.007	0.797 ± 0.797	0.866 ± 0.008	0.934 ± 0.008	0.910 ± 0.013
Cross+concate	0.869 ± 0.003	0.814 ± 0.011	0.873 ± 0.004	0.938 ± 0.001	0.919 ± 0.013
Concate+concate	0.874 ± 0.007	0.808 ± 0.004	0.880 ± 0.007	0.936 ± 0.001	0.914 ± 0.003
Weight+weight	0.866 ± 0.001	0.784 ± 0.001	0.876 ± 0.002	0.914 ± 0.001	0.881 ± 0.004
Weight+concate	0.870 ± 0.003	0.798 ± 0.009	0.878 ± 0.001	0.930 ± 0.002	0.909 ± 0.004
MFCADTI	**0.886 ± 0.008**	**0.855 ± 0.017**	**0.885 ± 0.005**	**0.948 ± 0.007**	**0.927 ± 0.011**

Table 3 Comparison with different feature combinations on Luo_data dataset

Feature	Acc	Precision	F_1	AUC	AUPR
Attribute	0.819 ± 0.020	0.824 ± 0.041	0.808 ± 0.021	0.878 ± 0.015	0.885 ± 0.013
Network	0.870 ± 0.005	0.835 ± 0.010	0.870 ± 0.006	0.934 ± 0.001	0.911 ± 0.003
Net+SMILES	**0.886 ± 0.008**	**0.855 ± 0.017**	**0.885 ± 0.005**	**0.948 ± 0.007**	**0.927 ± 0.011**
Net+Mol	0.865 ± 0.007	0.830 ± 0.017	0.865 ± 0.007	0.927 ± 0.002	0.893 ± 0.009
Net+Mol+SMILES	0.863 ± 0.003	0.835 ± 0.005	0.861 ± 0.004	0.938 ± 0.001	0.899 ± 0.012

the features of drug and target; *concate+concate*, which concatenates drugs'/targets' attributes and network features before concatenating the features of drug and target; *weight+weight* is used to fuse the network and attribute features of the drug/target with dynamic weighting, and the drug-target pair features are learned with dynamic weighting; and *weight+concate* is used to fuse the network and attribute features of the drug/target with dynamic weighting, and the drug-target pair features are splice with concate. The experiment results in Table 2 indicate that our model achieves the best results, with 88.6%, 85.5%, 88.5%, 94.8%, and 92.7% on ACC, Precision, F_1, AUC, and AUPR, respectively, outperforming the performance of the other five variants. Using cross-attention is more effective in learning the complementary features between network and attribute features, thereby enhancing the representation of drugs and targets. Furthermore, cross-attention enhances the interactions between drugs and targets when compared to the concatenation strategy for drug-target pairs with better experiment results.

Comparison with different feature combinations

To test the impact of different feature combinations on the classification results, we conduct experiments with various feature combinations on Luo_data dataset. Specifically, we compare the performance of MFCADTI using only network features, only attribute features, and both network and attribute features. Furthermore, the model with different attribute features are also compared. Attribute features of protein targets are extracted from AA sequences, while attribute features of drugs are extracted from molecular structures and SMILES sequences, respectively. The final experimental results are presented in Table 3. We can see, in the Luo_data dataset, using only network features is better than using only attribute features in terms of Acc, Precision, F_1, AUC, and AUPR, and improves the performance by 5.1%, 1.1%, 6.2%, 5.6%, and 2. 6%, while the performance of using both network and attribute features in Acc,

Precision, F_1, AUC, and AUPR is better than using only network features, and the performance is improved by 1.6%, 2.0%, 1.5%, 1.4%, and 1.6% respectively compared to using only network features. It is evident that both network and attribute features significantly contribute to the DTI prediction. In particular, the integration of network and attribute features enhances MFCADTI's ability to predict DTIs. Therefore, incorporating complementary knowledge from both network topology and biological attributes allows for a more comprehensive understanding of DTIs. The experimental results also show that there is a difference in model performance when different attribute features are used. Among them, extracting drug attribute features using SMILES sequences can provide high-quality feature representations for MFCADTI.

Comparison of different network features

In biological networks, various graph representation learning methods have shown promising capabilities in learning network representations of biological molecules. We compare six well-known graph representation learning methods on Luo_data dataset to examine how various graph representation methods perform when integrated into MFCADTI. These methods include Variational Graph Autoencoder (VGAE) [39] based on Graph Convolutional Networks (GCN) [40], DeepWalk [41], Node2vec [42], heterogeneous graph attention(HAN) [43], Simple-HGN [44], and LINE which is used by our work. In the Luo_data dataset, the method LINE achieved 88.6% for ACC, 85.5% for Precision, 88.5% for F_1, 94.8% for AUC, and 92.7% for AUPR, all of which are better than the other five graph representation learning methods. The experiment results are presented in Fig. 3, where we notice that LINE performs better than the other methods, indicating that it is a good candidate for learning network representations of drugs and targets in heterogeneous networks. On the other hand, VAGE performs poorly may be due to excessive smoothing of the representations caused by averaging different types of node features during the learning of representations, limiting its ability to capture the interactions in the network.

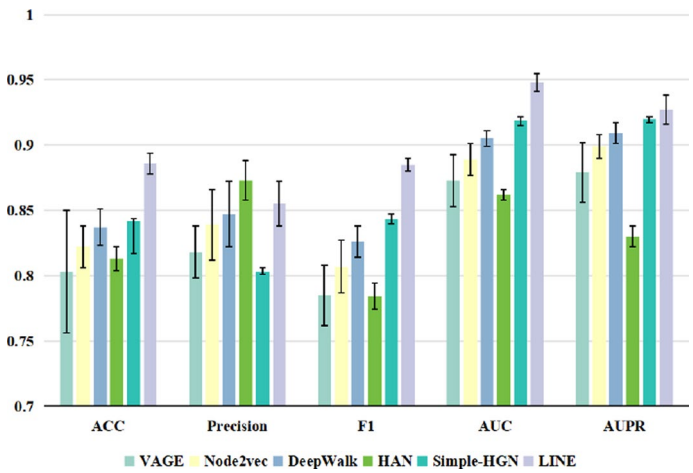

Fig. 3 Comparison of different network features

Comparison of different sequence encoding methods

Based on the drug's SMILES and the target's AA sequences, various deep-learning methods can be employed to extract the attribute features of drugs and targets. To demonstrate the superiority of FCS with nn.Embedding in our work, we conduct experiments on two datasets with other three different sequence encoding methods: LSTM, CNN, and GRU. The experiment results, as shown Fig. 4, indicate that FCS with nn.Embedding is the most optimal one. In the Luo_data dataset, the methods FCS with nn.Embedding achieves an accuracy, precision, F_1-score, AUC and AUPR of 88.6%, 85.5%, 88.5%, 94.8% and 92.7%, respectively, which are all better than the other three sequence encoding methods. This indicates FCS with nn.Embedding excels at extracting fundamental and meaningful biomedical semantics from drug and target sequences, allowing it to preserve the semantic relevance of the sequences while capturing essential substructure. We also conduct ablation experiments with different loss functions, which shows that the nn.BCELoss loss function used in MFCADTI is optimal, and the results of the experiments are in Supplementary Table S3.

Cold start analysis

In real-world applications, there are few known interaction data between a drug and a target node, which greatly degrades model generalization capability. Therefore, we further conduct the cold start experiments on two datasets to analyze the robustness of MFCADTI.

We simulate the so-called cold start problem by artificially creating isolated vertices through three experimental settings corresponding to the cases when: (a) the drugs are new, (b) the targets are new, and (c) both the drugs and their targets are new. We name settings (a), (b), and (c) as C1, C2 and C3, respectively. Under setting C1, we randomly select 20% of drugs and all DTI pairs associated with these drugs as the test set, the remaining drugs' related DTIs are used as the training set and validation set. Under setting C2, we randomly select 20% of targets and all DTI pairs associated with these targets as the test set, the remaining targets' related DTIs are used as the training set and validation set. Under setting C3, we randomly select 20% of drugs and targets and

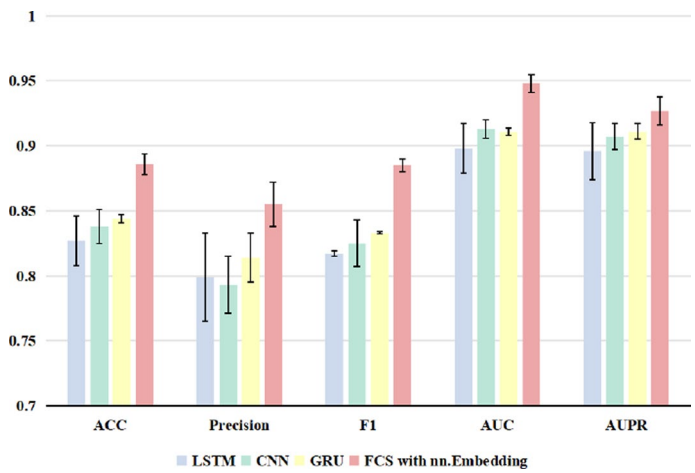

Fig. 4 Comparison of different sequence encoding methods

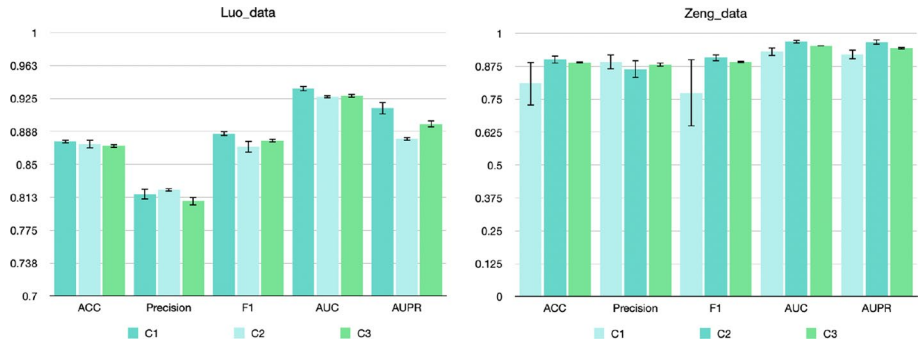

Fig. 5 Prediction performance with different cold start experimental settings on two datasets

Table 4 The top 10 targets predicted to be associated with Pravastatin in the Luo_data

Uniport ID	Target name	Evidence
P41143	Delta-type opioid receptor	CHEMBL
P18507	Gamma-aminobutyric acid receptor subunit gamma-2	*
Q01959	Sodium-dependent dopamine transporter	CHEMBL
P08908	5-hydroxytryptamine receptor 1A	*
O76074	cGMP-specific 3′,5′-cyclic phosphodiesterase	CHEMBL
P41594	Metabotropic glutamate receptor 5	*
Q9Y233	cAMP and cAMP-inhibited cGMP 3′,5′-cyclic. phosphodiesterase 10A	*
P08588	Beta-1 adrenergic receptor	Super PRED
P07550	Beta-2 adrenergic receptor	CHEMBL
P14867	Gamma-aminobutyric acid receptor subunit alpha-1	*

all DTI pairs composed of these drugs and targets as the test set. The DTIs are unrelated to these drugs and targets are divided into the training set and validation set. The results in Fig. 5 show that MFCADTI can still achieve promising performances in cold-start scenarios.

Case study

In addition to the cold start analysis, we conduct case studies on two datasets to further evaluate the real-world performance of MFCADTI. Specifically, we select the drugs Pravastatin and Masoprocol in Luo_data and Zeng_data, respectively. Firstly, we remove known interactions between corresponding drugs and targets from the training set to prove the effectiveness of our approach for new drugs (drugs with no known interacted targets). Secondly, the trained model is used to predict the probability of association between the selected drugs and the candidate targets, and the top 10 targets are chosen according to the predicted score. Finally, we perform analysis and validation using ChEMBL [45] and SuperPred [46] databases. The top 10 candidate targets for pravastatin in Luo_data are shown in Table 4, and the candidate targets of Masoprocol in Zeng_data are listed in Supplementary Table S4.

There are 5 candidate targets (50%) of Pravastatin's top 10 prediction results in the Luo_data that have been validated by the reference database accurately. Among the newly predicted candidate targets (50%), one interesting finding is the Gamma-aminobutyric

acid receptor subunit gamma-2 (GABRG2), which is a subunit of the GABA-A receptor involved in the signal transmission of gamma-aminobutyric acid (GABA), a critical inhibitory neurotransmitter in the central nervous system [47]. Recent research [48] has revealed that GABA plays a significant role in the medullary vasomotor center, which can regulate cardiovascular functions such as blood pressure and heart rate and lead to increased oxygen supply, activation of cerebral blood flow, or even stroke. On the other hand, Pravastatin, as an HMG-CoA reductase inhibitor, is commonly used to lower lipid levels and reduce the risk of cardiovascular events including myocardial infarction and stroke. The potential association between Pravastatin and the GABRG2 target suggests a promising direction for further investigations into the drug's mechanisms of action and its potential therapeutic applications in cardiovascular disease. These case study results suggest that MFCADTI could be a useful way to discover new drug-target associations to improve the virtual screening phase of drug discovery, especially for novel drugs without any known interactions.

Conclusion

In this article, we propose a prediction method, named MFCADTI, which considers the complementary relationship between the network topological features from the heterogeneous network and attribute features from sequences of drugs and targets. We leverage the cross-attention mechanism to mitigate the heterogeneity between network and attribute features and enhance the correlation between drug and target features to improve prediction accuracy. According to experiment results on two benchmark datasets, the accuracy and robustness of MFCADTI are superior to those of the state-of-the-art DTI prediction algorithms. Further research into case study also reveals that MFCADTI can forecast probable DTI and identify drug-related targets.

Our work highlights the benefits of fusing network and attribute features for predicting DTIs. As the presence of many other types of biological entities in biological networks, learning meaningful and impactful feature representations of nodes in heterogeneous biological networks remains a challenge. In future research, we plan to explore the use of heterogeneous network-based graph neural methods such as HetGNN [49]. This would enable the model to mine the relationships and interactions between different biological entities, enhancing its ability to learn feature representations and provide more accurate predictions.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12859-025-06075-7.

Supplementary file 1.

Author contributions

XB provided research ideas on the algorithm framework, supervised the research work, and revised the whole manuscript. NQ designed the model framework, implemented experiments and analysis, and wrote this manuscript. LZ guided the experimental process and supervised the completion of this study. SM and KZ provided advice on model. All authors read and approved the final version of this manuscript.

Funding

This work was supported by the Natural Science Foundation of China (62366052), the Key R&D Program of Xinjiang Uygur Autonomous Region (2022B03023), and the Natural Science Foundation of Xinjiang Uygur Autonomous Region (2024D01C126, 2022D01C427).

Data availability
The code and data used in this study are freely downloadable at https://github.com/LabBioMedCoder/MFCADTI.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Received: 23 December 2024 Accepted: 3 February 2025
Published online: 18 February 2025

References

1. Sachdev K, Gupta MK. A comprehensive review of feature based methods for drug target interaction prediction. J Biomed Inf. 2019;93:103159.
2. Chen X, Yan CC, Zhang X, et al. Drug-target interaction prediction: databases, web servers and computational models. Brief Bioinf. 2016;17(4):696–712.
3. Wu L, Shen Y, Li M, Wu FX. Network output controllability-based method for drug target identification. IEEE Trans Nanobiosci. 2015;14(2):184–91.
4. Keiser MJ, Setola V, Irwin JJ, et al. Predicting new molecular targets for known drugs. Nature. 2009;462(7270):175–81.
5. Morris GM, Huey R, Lindstrom W, et al. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. J Comput Chem. 2009;30(16):2785–91.
6. Li Y, Huang YA, You ZH, et al. Drug-target interaction prediction based on drug fingerprint information and protein sequence. Molecules. 2019;24(16):2999.
7. Mahmud SMH, Chen W, Jahan H, et al. Dimensionality reduction based multi-kernel framework for drug-target interaction prediction. Chemometr Intell Lab Syst. 2021;212:104270.
8. Wang L, You ZH, Chen X, et al. A computational-based method for predicting drug-target interactions by using stacked autoencoder deep neural network. J Comput Biol. 2018;25(3):361–73.
9. Lan W, Wang J, Li M, Liu J, Li Y, Wu F-X, Pan Y. Predicting drug-target interaction using positive-unlabeled learning. Neurocomputing. 2016;206:50–7.
10. Lee I, Keum J, Nam H. DeepConv-DTI: prediction of drug-target interactions via deep learning with convolution on protein sequences. PLoS Comput Biol. 2019;15(6):1007129.
11. Kurata H, Tsukiyama S. ICAN: interpretable cross-attention network for identifying drug and target protein interactions. PLoS One. 2022;17(10):0276609.
12. Zhao Q, Zhao H, Zheng K, Wang J. Hyperattentiondti: improving drug-protein interaction prediction by sequence-based deep learning with attention mechanism. Bioinformatics. 2022;38(3):655–62.
13. Cheng F, Liu C, Jiang J, et al. Prediction of drug-target interactions and drug repositioning via network-based inference. PLoS Comput Biol. 2012;8(5):1002503.
14. 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. 2017;8(1):573.
15. Wan F, Hong L, Xiao A, et al. Neodti: neural integration of neighbor information from a heterogeneous network for discovering new drug-target interactions. Bioinformatics. 2019;35(1):104–11.
16. Zeng X, Zhu S, Lu W, Liu Z, Huang J, Zhou Y, Fang J, Huang Y, Guo H, Li L, et al. Target identification among known drugs by deep learning from heterogeneous networks. Chem Sci. 2020;11(7):1775–97.
17. Gao C, Li J, Ji X, et al. Drug target interaction prediction method based on graph convolutional neural networks. Comput Sci. 2021;48(10):127–34.
18. 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. 2024;149:102778.
19. Zhao BW, Hu L, You ZH, et al. Hingrl: predicting drug-disease associations with graph representation learning on heterogeneous information networks. Brief Bioinf. 2022;23(1):515.
20. Ji BY, You ZH, Jiang HJ, et al. Prediction of drug-target interactions from multi-molecular network based on line network representation method. J Transl Med. 2020;18(1):347.
21. Ma T, Lin X, Song B, Philip SY, Zeng X. Kg-mtl: knowledge graph enhanced multi-task learning for molecular interaction. IEEE Trans Knowl Data Eng. 2022;35(7):7068–81.
22. Tao W, Lin X, Liu Y, Zeng L, Ma T, Cheng N, Jiang J, Zeng X, Yuan S. Bridging chemical structure and conceptual knowledge enables accurate prediction of compound-protein interaction. BMC Biol. 2024;22(1):248.
23. Ma T, Chen Y, Tao W, Zheng D, Lin X, Pang C-I, Liu Y, Wang Y, Wang L, Song B, et al. Learning to denoise biomedical knowledge graph for robust molecular interaction prediction. IEEE Trans Knowl Data Eng. 2024.
24. Ashish V, Noam S, Niki P et al. Attention is all you need. arXiv preprint arXiv:1706.03762 2017.
25. Zhang L, Cheng Y. A densely connected criss-cross attention network for document-level relation extraction. arXiv preprint arXiv:2203.13953 2022.
26. Lin H, Cheng X, Wu X, et al. Cat: cross attention in vision transformer. arXiv preprint arXiv:2106.05786v1 2022.

27. Nguyen NQ, Jang G, Kim H, Kang J. Perceiver CPI: a nested cross-attention network for compound-protein interaction prediction. Bioinformatics. 2023;39(1):731.
28. Zhang L, Lu D, Bi X, et al. Predicting disease genes based on multi-head attention fusion. BMC Bioinf. 2023;24(1):162.
29. Tang J, Qu M, Wang M, et al. Line: large-scale information network embedding. In: Proceedings of the 24th international conference on world wide web, 2015;pp. 1067–1077.
30. Huang K, Xiao C, Glass LM, Sun J. Moltrans: molecular interaction transformer for drug-target interaction prediction. Bioinformatics. 2021;37(6):830–6.
31. Zeng X, Zhu S, Lu W, et al. Target identification among known drugs by deep learning from heterogeneous networks. Chem Sci. 2020;11(7):1775–97.
32. Kim S, Chen J, Cheng T, et al. Pubchem in 2021: new data content and improved web interfaces. Nucleic Acids Res. 2021;49(D1):1388–95.
33. Boutet E, Lieberherr D, Tognolli M, et al. Uniprotkb/swiss-prot. Methods Mol Biol. 2007;406:89–112.
34. Sennrich R, Haddow B, Birch A. Neural machine translation of rare words with subword units. In: ACL, 2016;pp. 1715–1725.
35. Liu Z, Chen Q, Lan W, et al. GADTI: graph autoencoder approach for DTI prediction from heterogeneous network. Front Genet. 2021;12:650821.
36. Bai P, Miljković F, John B, Lu H. Interpretable bilinear attention network with domain adaptation improves drug-target prediction. Nat Mach Intell. 2023;5(2):126–36.
37. Qian Y, Li X, Wu J, Zhang Q. MCL-DTI: using drug multimodal information and bi-directional cross-attention learning method for predicting drug-target interaction. BMC Bioinf. 2023;24(1):323.
38. Zhang Y, Wang Y, Wu C, Zhan L, Wang A, Cheng C, Zhao J, Zhang W, Chen J, Li P. Drug-target interaction prediction by integrating heterogeneous information with mutual attention network. BMC Bioinf. 2024;25(1):361.
39. Thomas NK, Welling M. Variational graph auto-encoders. arXiv preprint arXiv:1611.07308 2016.
40. Thomas NK, Welling M. Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907 2016.
41. Perozzi B, Al-Rfou R, Skiena S. Deepwalk: online learning of social representations. In: Proceedings of the 20th ACM SIGKDD international conference on knowledge discovery and data mining, 2014;pp. 701–710.
42. Grover A, Leskovec J. node2vec: scalable feature learning for networks. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, 2016;pp. 855–864.
43. Wang Xiao, Ji Houye, Shi Chuan, Wang Bai, Ye Yanfang, Cui Peng, Yu Philip S. Heterogeneous graph attention network. In: The world wide web conference, 2019;pp. 2022–2032.
44. Lv Q, Ding M, Liu Q, Chen Y, Feng W, He S, Zhou C, Jiang J, Dong Y, Tang J. Are we really making much progress? revisiting, benchmarking and refining heterogeneous graph neural networks. In: Proceedings of the 27th ACM SIGKDD conference on knowledge discovery & data mining, 2021;pp. 1150–1160.
45. Bento AP, Gaulton A, Hersey A, et al. The chembl bioactivity database: an update. Nucleic Acids Res. 2014;42:1083–90.
46. Nickel J, Gohlke BO, Erehman J, et al. Superpred: update on drug classification and target prediction. Nucleic Acids Res. 2014;42:26–31.
47. Huang J, Chen J, Wang W, et al. Birthdate study of GABAergic neurons in the lumbar spinal cord of the glutamic acid decarboxylase 67-green fluorescent protein knock-in mouse. Front Neuroanat. 2013;7:42.
48. Sarasa SB, Mahendran R, Muthusamy G, et al. A brief review on the non-protein amino acid, gamma-amino butyric acid (GABA): its production and role in microbes. Curr Microbiol. 2020;77(4):534–44.
49. Zhang C, Song D, Huang C, et al. Heterogeneous graph neural network. In: Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining, 2019;pp. 793–803.

Publisher's Note