Applied Intelligence (2023) 53:27177–27190
https://doi.org/10.1007/s10489-023-04977-8

CSDTI: an interpretable cross-attention network with GNN-based drug molecule aggregation for drug-target interaction prediction

Yaohua Pan[1] · Yijia Zhang[1] · Jing Zhang[1] · Mingyu Lu[1]

Accepted: 19 August 2023 / Published online: 4 September 2023
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract
Drug-target interaction (DTI) is a critical and complex process that plays a vital role in drug discovery and design. In deep learning-based DTI methods, graph neural networks (GNNs) are employed for drug molecule modeling, attention mechanisms are utilized to simulate the interaction between drugs and targets. However, existing methods still face two limitations in these aspects. First, GNN primarily focus on local neighboring nodes, making it difficult to capture the global 3D structure and edge information. Second, the current attention-based methods for modeling drug-target interactions lack interpretability and do not fully utilize the deep representations of drugs and targets. To address the aforementioned issues, we propose an interpretable network architecture called CSDTI. It utilizes a cross-attention mechanism to capture the interaction features between drugs and targets. Meanwhile, we design a drug molecule aggregator to capture high-order dependencies within the drug molecular graph. These features are then utilized simultaneously for downstream tasks. Through rigorous experiments, we have demonstrated that CSDTI outperforms state-of-the-art methods in terms of performance metrics such as AUC, precision, and recall in DTI prediction tasks. Furthermore, the visualization mapping of attention weights indicates that CSDTI can provide chemical insights even without external knowledge.

Keywords Drug discovery · Interpretability analysis · Cross-Attention network · Drug-target interaction

1 Introduction

Drug discovery and development is very complex and expensive research requiring large amounts of capital, human resources and technical expertise [1]. It costs an average of $1.8 billion to develop a new drug and takes 13.5 years to bring a drug to the market [2]. This process is clearly struggling to meet the needs of rapidly evolving diseases such as COVID-19 [3]. On the other hand, finding new indications for existing drugs, known as drug repurposing, has proven to be the best way to accelerate drug discovery. The identification of drug-target interactions(DTIs) is a committed step of drug repurposing, but traditional biological experiments are often expensive and arduous [4]. Therefore, with the increasing availability of big biomedical data, in-silico methods for predicting drug-target interactions have gained significant attention due to their efficiency and relatively low cost. In recent years, deep learning has achieved performance beyond traditional methods in many areas, such as computer vision and natural language processing. With the large amount of biomedical data generated in recent years, deep learning has enabled the rapid development of drug-target interaction (DTI) prediction. In the DTI task, a significant challenge lies in obtaining high-level hidden representations from drug molecules. Öztürk et al. [5] proposed a deep learning based model named DeepDTA, the method learned latent representations from drug SMILES(Simplified Molecular Input Line Entry System) strings and protein sequences via 1D convolutional neural networks(CNNs). Similar to DeepDTA, WideDTA [6] also exploited the 1D text sequence information of drug molecules and proteins to mine potential representations. A similar approach, Zhao et al. [7], utilizes three consecutive 1D-CNN layers to learn

✉ Yijia Zhang
zhangyijia@dlmu.edu.cn

Yaohua Pan
panyaohua@dlmu.edu.cn

Jing Zhang
1120211397@dmu.edu.cn

Mingyu Lu
lumingyu@dlmu.edu.cn

[1] School of Information Science and Technology, Dalian Maritime University, Dalian, Liaoning 116026, China

the feature matrix of drugs and proteins. Although the aforementioned sequence-based approach yielded promising results, it overlooked the crucial three-dimensional structural information of the molecule. In recent research, graph neural networks (GNNs) have been employed to tackle this problem in drug-target interaction (DTI) prediction. Nguyen et al. [8] proposed a model called GraphDTA to capture the spatial structure information of drugs by modeling the drug molecules as a graph, and four different variants of GNNs were tried. Cheng et al. [9] and Li et al. [10] employed graph attention networks(GATs) to extract drug features. Compared to sequence-based methods, graph-based methods offer the advantage of capturing the three-dimensional structural information of drug molecules, leading to richer information and successful learning of drug molecule representations. However, current graph-based methods still face limitations in molecular representation learning. One such limitation is that graph neural networks (GCNs) primarily focus on local neighbor nodes and struggle to capture global 3D structure and edge information, impeding their ability to fully reflect the molecule's characteristics. Drug-target interaction is a multifaceted process that encompasses biological and chemical knowledge [11]. Hence, it is essential and valuable to comprehend the substructural interactions within drug-target interactions. Therefore, obtaining concise, efficient, and interpretable information about drug-target interactions is another crucial challenge in the DTI task. In recent years, there has been a growing focus on research dedicated to understanding the interaction properties of local substructures [7, 9, 12]. Zhao et al. [7] introduced the HyperAttention model, which combines CNN and attention mechanisms to capture interactions between atoms and amino acids, but the sequence-based CNN approach overlooks the 3D structural information of the drug molecule. Chen et al. [12] proposed a transformer-based architecture for compound-protein prediction named TransformerCPI, which addressed some of the shortcomings of previous sequence-based approaches and used label reversal experiments to verify whether the model learns the interaction processes between compounds and proteins. While the aforementioned approaches that emphasize the interaction process have demonstrated remarkable performance, it is crucial to recognize that drug-target interaction is a complex phenomenon that encompasses both biological and chemical knowledge. Extracting independent features from drug-target pairs can offer additional discriminative information for prediction. Consequently, it is inadequate to solely separate the representations learned from drug-target pairs and the information about their interactions. Instead, a systematic extraction of effective features from both sources is necessary. To tackle the aforementioned challenges, we propose CSDTI, a novel interpretable deep learning-based model in this study. CSDTI utilizes the deep representations of drugs and targets, along with their interaction information, to accurately predict drug-target interactions (DTIs). More specifically, we first model the actual hypothesis from biological inspiration and learn the interaction information of the local substructure of the drug-target pair. For this process, cross-attention structures are chosen as extractors with rich semantic extraction capabilities. Next, we model deep representations of drugs and targets and fuse them with the interaction information to be used together for downstream tasks. Among them, the drug molecule aggregator is used to extract global features of the drug, which captures global topological information between atoms. Similarly, convolutional structures at different nuclear scales are used to extract deep representations of the targets. Finally, the three feature vectors are fused and fed into a fully connected dense layer to predict the DTI. Multiple experimental results demonstrate that CSDTI outperforms state-of-the-art methods across four benchmark datasets. In summary, our work makes the following main contributions:

- We propose a novel deep learning-based model, CSDTI, that incorporates the deep representations of drugs and targets, as well as their interaction information, for accurate prediction of drug-target interactions.
- In drug molecule representation learning, we acknowledged the significance of higher-order dependencies and boundary information. To overcome the limitations of previous GNN methods in capturing global structural information for molecule representation, we developed an effective drug molecule aggregator, demonstrated through ablation studies.
- Our method outperforms state-of-the-art methods on multiple benchmark datasets, demonstrating its superior performance. Additionally, the visualization of attention weights and case studies confirm the model's ability to provide valuable biological insights.

2 Related work

The computational prediction of Drug-Target Interactions (DTIs) has garnered significant attention due to its potential to provide valuable guidance for drug discovery while reducing overall costs. There are three main approaches to DTI prediction: molecular docking [13, 14], similarity-based methods [15, 16], and deep learning-based methods. In this paper, we primarily focus on deep learning-based methods, because compared to other approaches, they offer timely results and achieve considerable accuracy. We have organized the related work based on the typical workflow of DTI tasks.

2.1 Feature extraction of drug-target pairs

2.1.1 Sequence based methods

Recently, there have been rapid advancements in deep learning (DL) technologies. A class of DL methods specifically designed for handling sequential information has shown remarkable success in various applications such as language translation and speech recognition. Given that chemical compounds and target proteins can be represented in a sequential format such as SMILES and amino acid sequences, respectively, sequential deep learning methods have been extensively explored for predicting compound-protein interactions. CNN has been used for structure-based binding affinity prediction, taking inspiration from its success in computer vision, as demonstrated in [17]. Zhao et al. [7] employed deep CNN to learn the feature matrix of drugs and proteins. Wu et al. [18] utilized CNN to learn the representation of local views of drug molecules. In addition, the Transformer, a sequence-based method, has also been widely used in DTI prediction tasks, as seen in studies such as [19, 20].

2.1.2 Graph based methods

Proteins and compounds can be represented as graphs, with nodes corresponding to their constituent amino acids or chemical elements, and edges connecting them to represent interactions or relationships. This enables the use of graph-based deep learning techniques, such as graph neural networks (GNNs), for predicting DTIs. The basic approach involves learning embedding vectors separately for the compound and protein graphs, and then combining them using a late integration strategy for DTI prediction. GCN has been used to learn embedding vectors of molecular graphs in [20–22], while Lim et al. [23] employed a similar approach to embed the 3D graph representation of protein-ligand complexes as input. However, a limitation of GNN is that it only considers local neighboring nodes and may not fully capture global 3D structure and edge information.

2.2 Attention methods in drug-target interaction feature learning

The attention mechanism was originally invented to improve the quality of machine translation by aligning two different representations(e.g., source and target languages) [24]. The utilization of the attention mechanism offers several benefits. First, it enables neural models to effectively consider distant relationships between features, thereby enhancing task performance [25, 26]. Second, the attention mechanism provides interpretability to model predictions, allowing for better understanding and insight into the reasoning behind the model's decisions [27]. Regarding the prediction of drug-target interactions, most studies have demonstrated the benefits of attention mechanisms in generating superior representations [12, 28, 29]. Among these studies, two of them are particularly similar to our work. Kim et al. [28] introduced a gated cross-attention mechanism that explicitly models the interaction between drugs and targets to attend to their features. Kurata et al. [29] encoded the SMILES representations of drugs and the amino acid sequences of target proteins into embedding matrices, which were directly input into a cross-attention network to generate drug-protein contextual matrices. However, there are still two limitations of attention mechanisms when applied to DTI tasks. First, these studies face the limitation of GNNs in capturing the global structure of drug molecules while extracting drug features. Second, existing attention-based models tend to become overly complex, which hampers the interpretability of the models. In contrast to previous works, our study introduces a novel approach where we directly apply the cross-attention mechanism to the interaction process between drugs and targets. Through extensive experiments, we have successfully demonstrated the effectiveness of this method. Additionally, we provide model explanations through case studies and attention visualization, further enhancing our understanding of the model's behavior. Summary of recent classical related works is provided in Table 1.

3 Methodologies

3.1 Overview

Figure 1 shows the network architecture of CSDTI. The system mainly consists of an embedding module, a representation learning module and an interaction module. In the representation learning module, the drug molecule aggregator and the multiscale 1D convolution-based protein encoder learn the deep representations of drugs and proteins, respectively, and the interaction module focuses on capturing the substructure interaction process between drug-target pairs.. Finally, the deep representations of the drug-target pairs and the interaction features between them are input to the output module to obtain the prediction results.

3.2 Drug molecule aggregator

In GNN-based DTI methods, GNN is employed for learning the representation of drug molecules, where a drug molecule can be viewed as a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{X})$, where $\mathcal{V}$ is the set of atoms in the molecular graph $\mathcal{G}$ and $\mathcal{E} \in \mathcal{V} \times \mathcal{V}$ represents the set of chemical bonds that exist between the atoms in $\mathcal{V}$. $\mathcal{X} \in \mathbb{R}^{|\mathcal{V}| \times n}$ is the feature matrix where each row represents the n-dimensional feature encoding of an atom. The set

Table 1 Summary of the related work based on deep learning

Method	Type	Datasets	Compounds	Proteins	The mian Network Architecture	Advantage	Disadvantage
DeepDTA [5]	DTA	Davis, KIBA	Molecular graph	Amino acid sequences	CNN	Only utilized the sequence information of targets and drugs to predict DTA	Ignored the structural information of drugs and the interaction information between drugs and targets
WideDTA [6]	DTA	Davis, KIBA	Ligand SMILES, Ligand Maximum Common Substructures	Protein sequence, Protein domains and motifs	CNN	Leveraged the sequence information of multiple drugs and targets to the fullest extent	Ignored the structural information of drugs and the interaction information between drugs and targets
DeepAffinity [19]	CPA	BindingDB, STITCH, UniRef	SMILES	structural property sequence	RNN, CNN, Attention Mechanism	Introduced secondary structure information for each residue	Ignored the structural information of drugs and the interaction information between drugs and targets
DeepConv-DTI [30]	DTI	DrugBank, IUPHAR, KEGG	Fingerprint	Amino acid sequences	CNN, DNN	Captured local residue patterns of generalized protein categories	Ignored the structural information of drugs and the interaction information between drugs and targets
GraphDTA [8]	DTA	Davis, KIBA	Molecular graph	Amino acid sequences	CNN, GNN	Captured structural	Ignored the interaction

Table 1 continued

Method	Type	Datasets	Compounds	Proteins	The mian Network Architecture	Advantage	Disadvantage
						information in molecules	process between drugs and targets
TransformerCPI [12]	CPI	Human, C.elegans, BindingDB	Molecular graph	Amino acid sequences	Transformer	Simulated interactions between atoms and amino acids	Shallow GNNs are insufficient to capture the global 3D structure of molecules
HyperAttentionDTI [7]	DTI	DrugBank, Davis, KIBA	SMILES	Amino acid sequences	CNN, Attention Mechanism	Simulated interactions between atoms and amino acids	Neglected the 3D structural information of drug molecules
MolTrans [31]	DTI	BIOSNAP, Davis, BindingDB	SMILES	Amino acid sequences	Sub-structural pattern mining algorithm, Transformer	Captured semantic relationships between substructures	Neglected the 3D structural information of drug molecules
MGraphDTA [32]	DTA	Davis, Filtered davis, KIBA, Metz, Huamn, C.elegans, ToxCast	Molecular graph	Amino acid sequences	GNN, CNN	Visualized the importance of molecular substructures by employing gradient-weighted affinity activation mapping	Ignored the interaction process between drugs and targets
CSDTI	DTI	DrugBank, Huamn, Davis, KIBA	Molecular graph	Amino acid sequences	GNN-based Drug Molecule Aggregator, GLU, Cross-Attention	Capture high-order dependencies in drug molecules, learn the interaction features between drugs and targets	Unable to learn protein structure information

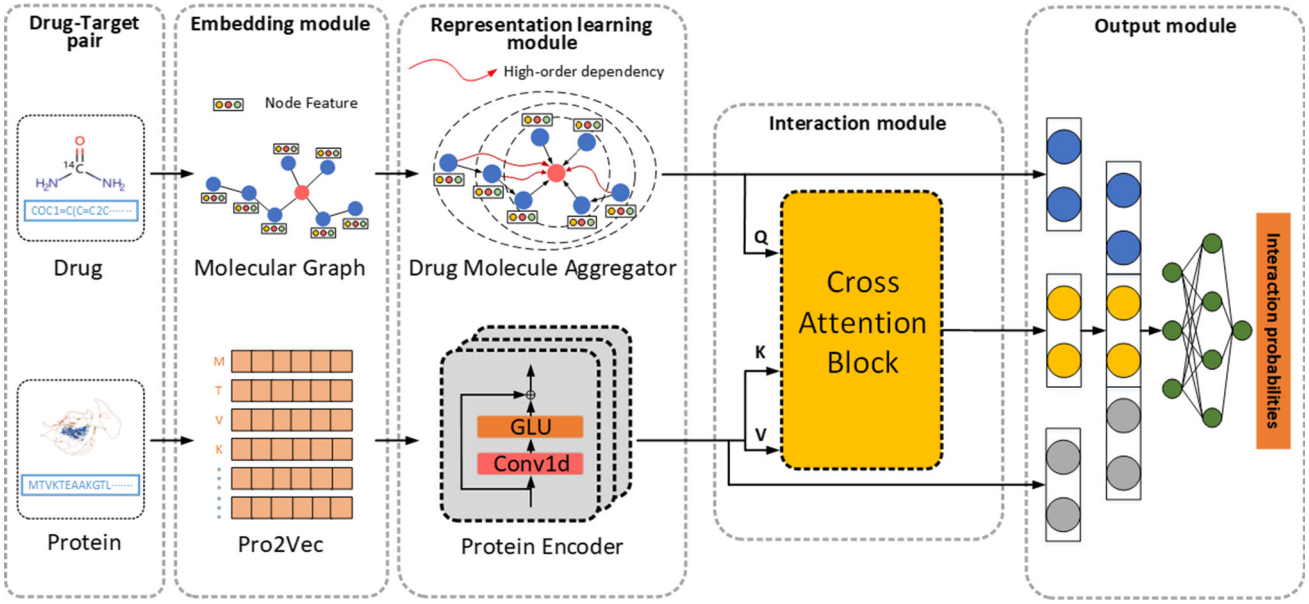

Fig. 1 The network architecture of CSDTI. The representation learning module is used to learn deep representations of drugs and proteins, the interaction module learns interaction features between drugs and proteins, and the output module combines these three features for prediction

of edges $\mathcal{E}$ can also be represented as an adjacency matrix $A \in \{0, 1\}^{|\mathcal{V}|\times|\mathcal{V}|}$ where $A_{u,v} = 1$ if there is a chemical bond from atom u to atom v. The goal of our method is to learn a representation of the entire molecular graph. Graph neural networks take the graph structure data as input and updates the node features through the message a_v^k aggregated from the neighboring nodes. After k iterations of aggregation, the neighbor features and spatial structure information within the k-hop neighborhood are aggregated into a node's representation. Formally, the k-th layer of a GNN is

$$a_v^{(k)} = AGGREGATE^{(k)} h_u^{(k-1)} \tag{1}$$

$$\begin{aligned} a_v^{(k)} &= AGGREGATE^{(k)}(\{h_u^{(k-1)} | u \in \mathcal{N}(v)\}), \\ h_v^{(k)} &= COMBINE^{(k)}(h_v^{k-1}, a_v^{(k)}) \end{aligned} \tag{2}$$

where $h_v^{(k)}$ is the represent of node v at the k-th iteration. Especially, $h_v^{(0)} = X_v$, and $\mathcal{N}(v)$ is a set of nodes adjacent to v. Therefore, in order to capture the structure composed of k-hop neighbors, the GNN model needs to stack k layers. Therefore, when used for learning drug molecule features, GNN primarily focuses on local neighboring atoms, which makes it challenging to capture the global 3D structure and edge information, thus limiting its ability to fully learn molecular features. When learning the features of drug molecules, we have taken into account the importance of high-order dependencies and graph edge information, and designed a drug molecule aggregator, that can simultaneously aggregate all the information within the k-hop neighborhood of each atomic node. During the message passing phase, the hidden states of node v at l-th layer, i.e. h_v^l, is updated by message m_v^{l+1} according to

$$m_v^{l+1} = ||_{k=1}^{k} f_k^l(\{h_u^l | u \in \mathcal{N}_k(v)\}) \tag{3}$$

$$h_v^{l+1} = U_l(h_v^l, m_v^{l+1}) \tag{4}$$

where $||$ is a concatenation operation, and $\mathcal{N}_k(v)$ denotes the set of k-hop neighbors of v. $f_k^l(\cdot)$ denotes the aggregate functions of the l-th iteration on the k-th hop and U_l denotes the vertex update functions of the l-th iteration. The readout phase obtains the representation h_G of the entire graph from the node features obtained from the final aggregation:

$$h_G = R(\{h_v^L | v \in G\}) \tag{5}$$

where $R(\cdot)$ is a graph-level readout function, L is the number of the layers.

3.3 Protein encoder

The high-level structure of a protein determines the biological function of the protein, so tertiary information is very important for the representation of protein molecules. However, the restricted protein structure does not exist stably in a certain form, so the acquisition of the protein geometric structure still faces great challenges. Therefore, we used the most

stable primary structure of the protein as the amino acid representation. We used this one-hot encoding scheme for protein sequences, mainly because it is the simplest method to construct a unified representation (UniRep), which is broadly applicable and generalized to unseen regions of sequence space [33]. First, we build a dictionary to map amino acids to an integer, so that the protein sequence is represented as an integer sequence. We set the maximum length of the protein sequence to 1200, as this length covers at least 80% of the protein [5]. We then map each integer to an n-dimensional vector via the embedding layer (i.e., each amino acid is represented as an n-dimensional vector). The protein encoder consists of a multilayer combination of one-dimensional convolutional and gated linear units [34]. To allow the model to learn deeper representations in protein sequences, we increase the depth of the network by skip connections and change the weights of the network layer by layer [35]. Skip connections jump some layers in the neural network and use the output of one layer as input to the next, deepening the network while avoiding network degradation [36]. The model learns a deeper representation while avoiding overfitting due to the complex structure. Ultimately, the output of the protein encoder at layer i can be expressed as:

$$h_{i+1} = \sigma(h_i * W1_i + b1_i) \otimes (h_i * W2_i + b2_i) \tag{6}$$

where σ is a nonlinear activation function. * represents the one-dimensional convolution operation. $\otimes$ denotes the element-wise product. $W1_i, W2_i \in R^{k \times m_1 \times m_2}$, $b1_i, b2_i \in R^{m_2}$ are the learnable parameters, m_1 and m_2 are the dimensions of input feature and hidden feature, k is the filter size. $h_i \in R^{n \times m_1}$ is the input of ith protein encoder.

3.4 Drug -target interaction module

In the real biological DTI process, the interaction between drug and target is a complex process involving knowledge biology and chemistry, so we hope that the model can consider the process of drug and target inter-reaction and thus provide more discriminative information for DTI prediction. On the one hand, after obtaining a deep representation of the drug and the target, we need a reasonable method to integrate the two kinds of information. On the other hand, we need our model to be able to learn information about the interaction between each drug-target pair, rather than just using the respective features of the drug and target for downstream classification or regression tasks. Therefore, we combine potential representations of drugs and targets through a cross-attention module and use it to model the process of drug-target interaction. The process of interaction can be represented as follows:

$$Q = f_Q(D); K = f_K(P); V = f_V(P); \tag{7}$$

$$Attention\_energy = Softmax(\frac{QK^T}{\sqrt{C/d}}) \tag{8}$$

$$\begin{aligned} Interaction_{D,P} &= CrossAttention(Q, K, V) \\ &= Attention\_energy \times V \end{aligned} \tag{9}$$

where Q is created from the output of the drug molecule aggregator D, K and V are created from the output of the protein encoder P by the projection functions $f = W^T x + b$(where w denotes the weight and b denotes the bias). C is the embedding dimension and d is the number of heads. Figure 2 shows the detailed structure of the attention module. By modeling the interaction process, the drug and target are no longer isolated parts, and the model learns the local substructure interaction information of the drug-target pair and uses it as input for downstream tasks.

3.5 Output block

In contrast to previous work, we evaluated our model in both the DTI and DTA tasks in the study. We simultaneously apply the deep representation of drugs and targets and the interaction features(D, P and $Interaction_{D,P}$, respectively) between them learned in the previous section (Sections 3.2 to 3.4) to downstream tasks.

$$\hat{y} = \sigma(W_{out}[(D; Interaction_{D,P}; P)] + b_{out}) \tag{10}$$

where σ is the sigmoid function, W_{out} and b_{out} are the learnable parameters, $\hat{y}$ is the predicted label, D and P are the outputs of the drug molecule aggregator and the protein encoder, respectively. For the DTI tasks, the cross-entropy loss function is used as the loss function for backpropagation.

$$loss(\Theta) = \frac{1}{N} \sum_i -[y_i \cdot \log(\hat{y_i}) + (1-y_i) \cdot log(1-\hat{y_i})] + \frac{\lambda}{2} ||\Theta||_2^2 \tag{11}$$

where $\hat{y_i}$ is the predicted label, y_i is the actual label, and Θ is the set of parameters in the model. N is the number of samples, λ is the $L2$ regularization coefficient. For the DTA tasks, mean squared error(MSE) is used to measure the prediction error of the model:

$$MSE = \frac{1}{n} \sum_{k=1}^{n} (p_k - y_k)^2 \tag{12}$$

where p is the predicted values and y is the actual values, n is the number of samples. The complete pseudo-code of the CSDTI algorithm is illustrated in Algorithm 1.

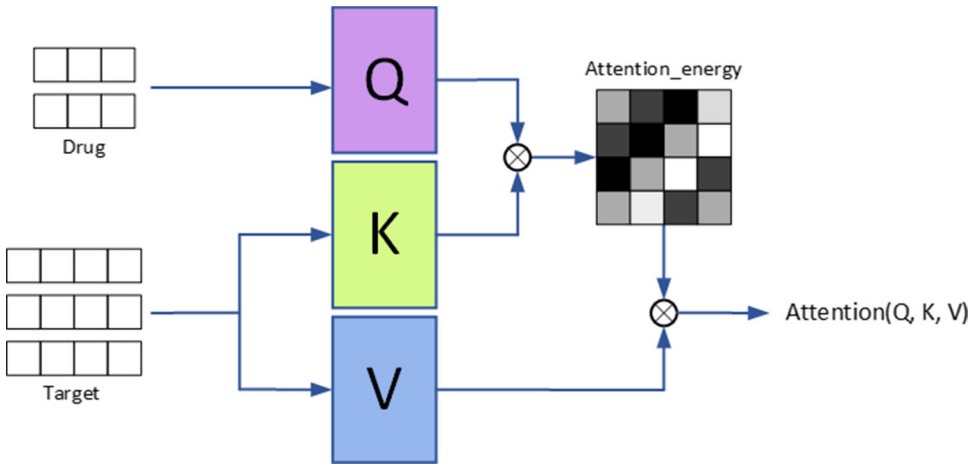

Fig. 2 Visualization of attention blocks reveals their functionality. While the self-attention block accepts inputs from a single source, cross-attention blocks operate by incorporating information from two distinct sources

Algorithm 1 The CSDTI algorithm.

Input: Molecular graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{X})$, protein initialized embedding emb_{in}, The number of 1DCNN layers M, Constant λ, layer normalization function LN
for $u \in \mathcal{V}$ **do**
  for $l = 1, 2, ...L$ **do**
    for $k = 1, 2, ...K$ **do**
      Calculate $f_k^l(\{h_u^l | u \in \mathcal{N}_k(v)\})$
    end for
    Update h_v^l via (1)
  end for
end for
for m in M **do**
  $emb_{out} \leftarrow LN(Conv(emb_{in}) + emb_{in} * \lambda)$
  $emb_{in} \leftarrow GLU(emb_{out})$
end for
$a, b, c = linear\_q(h_u^l), linear\_k(emb_{in}), linear\_v(emb_{in})$
$weight = matmul(a, b.transpose(1, 2))$
$Attention = softmax(weight, dim = -1)$
$Interaction_{D,P} = Attention \times c$
$P = Classifier(Concate(h_v^l, Interaction_{D,P}, emb_{in}))$
Output: Drug-target pairs' interaction probability P

4 Experiments

4.1 Experimental setup

In this study, we evaluated our model from both DTI prediction and DTA prediction tasks. In the experiments, our proposed model is implemented in PyTorch. The dataset is split into a training set, a validation set and a test set in the ratio of 8:1:1. We train the model on the training set and select the best hyperparameters according to the performance on the validation set, and finally evaluate the model on the test set.

4.2 Dataset

In this study, we formulated DTA prediction as a binary classification problem. We extracted drug and target data from the DrugBank database [37] to build the experimental dataset. After manually discarding some drugs with strings not recognized by the RDkit python package [38], we ended up with 6645 drugs, 4254 targets, and 17511 positive DTIs. Moreover, we also apply our model on some previous benchmark datasets, Human [39]. In detail, the human dataset consists of 6728 positive interactions between 2726 unique compounds and 2001 unique proteins. At the same time, we also compared CSDTI with the classical approach in DTA tasks. We evaluated our proposed model for DTA tasks using the benchmark datasets Davis [40] and KIBA [41]. The Davis dataset is measured as K_d constants and consists of 442 proteins and 68 ligands. The KIBA dataset is measured as KIBA scores and consists of 229 proteins and 2111 drugs, the details of the data are shown in Table 2.

4.3 Comparison of results

4.3.1 Comparison results in the DTI tasks

For the classification task, we used the area under curve (AUC), precision, and recall as performance metrics to evaluate the model following the previous studies. In this section, we compared our proposed model CSDTI with DeepDTA [5], DeepConv-DTI [30], MolTrans [31] and Transformer-CPI [12]. DeepDTA [5] consists of two 3-layer CNNs and was originally designed for predicting binding affinity, so we changed its fully connected last layer and changed the loss function accordingly to make it suitable for the DTI tasks. As

Table 2 Summary of the datasets

Dataset	Task type	Compounds	Proteins	Interactions
DrugBank	DTI	6645	4254	35022
Human	DTI	2726	2001	6728
Davis	DTA	68	442	30056
KIBA	DTA	2111	229	118254

shown in Table 3, for the DrugBank dataset, CSDTI achieved a relatively significant improvement compared to baselines. Compared with the second best method, CSDTI achieves an improvement of 3.5% and 3.7% in terms of AUC and recall. For small public datasets, Human, CSDTI still showed better performance than baselines. Compared with the second best method, CSDTI achieves an improvement of 1.23% and 0.96% in terms of AUC and AUPR, respectively.

4.3.2 Comparison results in the DTA tasks

For the regression task, we used mean square error (MSE) and concordance index (CI) as performance measures. We compared our proposed model CSDTI with the SOTA DeepDTA [5], WideDTA [6], GraphDTA [8] and DeepAffinity [19]. Furthermore, we compared our work with the work of Kim et al. [28]. (referred to as GCADTI for convenience) which is closely related to our method. The experimental results of baselines and CSDTI are shown in Table 4. On the Davis dataset, CSDTI exceeds the baseline in all evaluation metrics. Compared to GraphDTA, MSE and CI all achieve great improvement, MSE decreased from 0.229 to 0.220 and CI increased from 0.893 to 0.899. On the KIBA dataset, our proposed model CSDTI achieves a greater improvement, especially in the MSE metric, which is 8.6% lower than the second best method of 0.139.

4.3.3 Analysis of experimental results

In the DTI task, CSDTI achieves significant improvements over the baseline in all metrics on the DrugBank dataset. Methods that focus on interaction features such as CSDTI, MolTrans, and TransformerCPI are more competitive than earlier methods that solely extract representations of drug-target pairs, such as DeepDTA and DeepConv-DTI. This demonstrates the necessity of modeling drug-target interactions. Compared to methods such as MolTrans and TransformerDTI, which also focus on simulating drug-target interactions, our architecture is more efficient in extracting and utilizing deep representations of drug-target pairs. On the small-scale public dataset Human, our proposed method shows improvements in terms of precision and recall, which aligns with our expectations. Although this advantage is not as pronounced as in the DrugBank dataset, considering the limited room for improvement in AUC on these datasets already exceeding 95%, such progress remains quite substantial. MolTrans, which is based on a substructure pattern mining algorithm, outperforms our method in terms of precision on the Human dataset, suggesting that there is still room for improvement in protein representation learning within our model. In the DTA task, these two datasets demonstrate the effectiveness of CSDTI in regression tasks. In the Davis dataset, we observe an imbalanced label distribution, leading most models to predict affinities biased toward smaller values. As a result, all methods perform worse in terms of MSE on the Davis dataset compared to the KIBA dataset. However, our proposed method still achieves a performance gain of 0.9% in terms of MSE compared to the suboptimal methods. For the KIBA dataset, we notice a highly concentrated

Table 3 Comparison results of the proposed CSDTI and baselines on the DrugBank and Human datasets for DTI prediction tasks

Dataset	Method	Precision	Recall	AUC
DrugBank	DeepDTA	0.786	0.798	0.871
	DeepConv-DTI	0.736	0.767	0.836
	MolTrans	0.809	0.767	0.862
	TransformerCPI	0.774	0.821	0.865
	CSDTI	0.835	0.852	0.902
Human	DeepDTA	0.938	0.935	0.972
	DeepConv-DTI	0.939	0.907	0.967
	MolTrans	0.955	0.933	0.974
	TransformerCPI	0.911	0.937	0.970
	CSDTI	0.937	0.946	0.982

Table 4 Comparison results of the proposed CSDTI and baselines on the Davis and KIBA datasets for DTA prediction tasks

Datasets	Davis		KIBA	
Method	MSE ↓	CI ↑	MSE ↓	CI ↑
DeepDTA	0.261	0.878	0.194	0.863
WideDTA	0.262	0.886	0.179	0.875
GraphDTA(GCN)	0.254	0.880	0.139	0.889
GraphDTA(GAT)	0.232	0.892	0.179	0.866
GraphDTA(GIN)	0.229	0.893	0.147	0.882
GraphDTA(GAT-GCN)	0.245	0.881	0.139	0.891
DeepAffinity(RNN+RNN)	0.253	0.900	0.188	0.842
DeepAffinity(RNN+GCN)	0.260	0.881	0.288	0.797
DeepAffinity(CNN+GCN)	0.657	0.737	0.680	0.576
DeepAffinity(HRNN+GCN)	0.252	0.881	0.201	0.842
DeepAffinity(HRNN+GIN)	0.436	0.822	0.445	0.689
GCADTI	0.242	0.891	0.152	0.883
CSDTI	0.220	0.899	0.128	0.901

Table 5 Results of ablation experiments on the DrugBank dataset

Dataset	Precision	Recall	AUC
$CSDTIw/oI$	0.806	0.796	0.871
$CSDTIw/oH$	0.797	0.810	0.878
$CSDTIw/oG$	0.819	0.825	0.884
CSDTI	0.835	0.852	0.902

label distribution, making it difficult to predict affinity trends and creating challenges for most models to perform well in terms of the CI metric. However, our proposed method still achieves a performance gain of 1% in terms of CI compared to the suboptimal methods. Compared to methods that solely focus on learning representations of drug-target pairs, our proposed model demonstrates better performance, indicating the ability of the interaction module to effectively learn interaction features. Our method outperforms GCADTI, which also focuses on drug-target interaction features, highlighting the capability of the drug molecule aggregator to learn high-level representations of drug molecules.

4.4 Ablation study

To further investigate the importance of components, we design the following variants of CSDTI:

- $CSDTIw/oI$: We removed the interaction module, which means that CSDTI uses only the representations of drugs and proteins learned by the global encoder and the protein encoder to predict DTIs.
- $CSDTIw/oH$: We removed the aggregation of high-order dependency in the drug molecule aggregator, which means CSDTI only aggregates immediate neighbors.
- $CSDTIw/oG$: We removed the linear gating unit from the protein coding and used only CNN as the protein encoder.

Table 5 summarizes the experimental results on the DrugBank dataset. The results of study $CSDTIw/oI$ demonstrate that the interaction module is effective in simulating drug-target interactions. Additionally, removing the aggregation of higher-order neighbors in the drug molecule aggregator leads to a decrease in performance, indicating that the global molecular aggregator can effectively aggregate high-order neighborhood information when learning the representation of drug molecules. Finally, the results of $CSDTIw/oG$ show that a protein encoder consisting of a multilayer combination of one-dimensional convolution and gated linear units is capable of extracting effective features while avoiding overfitting of the model due to structural complexity when learning the representation of proteins.

4.5 Case study

To further validate the validity of the proposed model, we applied CSDTI for de novo predictions of important drugs and targets, namely Diacerein (Drugbank ID: DB11994) and Aromatic-L-amino-acid decarboxylase (Uniprot ID: P20711). In this section, we predict the interaction probability between these two important drugs and targets and known drugs or targets in the dataset by the pre-trained model. After the prediction, we ranked the drugs and targets based on the prediction results and validated the top ten candidate drugs and targets through the DrugBank database [37]. The top ten candidate targets predicted by CSDTI for Diacerein are shown in Table 6, we can find that 4 of the top ten predicted candidate targets were successfully predicted. Meanwhile, Table 7 shows the top ten candidate drugs predicted by CSDTI for the target Aromatic-L-amino-acid decarboxylase, 3 of the top ten predicted candidate drugs were successfully predicted. Accurate screening of targets (or drugs) against a particular drug (or targets) from a large amount of data is challenging. Therefore, the results of the above two case studies show that CSDTI can accurately predict potential drug-target interactions from a large number of samples, which has significant implications for drug screening and drug repurposing.

4.6 Model interpretation

To demonstrate that the attention mechanism not only enhances the performance of the model but also brings more interpretability, we conducted a case study on Riboflavin kinase(PDB: 1NB9) and its corresponding ligand. First, we

Table 6 The predicted candidate targets for drug Diacerein

Rank	Target name	Uniprot ID	Reference
1	Dihydrofolate reductase	P00374	Unknown
2	Nuclear receptor subfamily 1 group I member 2	O75469	Unknown
3	Cytochrome P450 2C8	P10632	Unknown
4	Oxysterols receptor LXR-alpha	Q13313	Sheng et al. [42]
5	Cytochrome P450 11B2, mitochondrial	P19099	Unknown
6	Cytochrome P450 1A2	P05177	Tang et al. [43]
7	Prostaglandin G/H synthase 2	P35354	Unknown
8	Cytochrome P450 2E1	P05181	Tang et al. [43]
9	Oxysterols receptor LXR-beta	P55055	Sheng et al. [42]
10	Retinoic acid receptor alpha	P10276	Unknown

Table 7 The predicted candidate drugs for target Aromatic-L-amino-acid decarboxylase

Rank	Drug name	DrugBank ID	Reference
1	N-Methyl-Pyridoxal-5'-Phosphate	DB01639	Unknown
2	Pyridoxine phosphate	DB02209	Unknown
3	Carbidopa	DB00190	Chen et al. [44]
4	Metyrosine	DB00765	Unknown
5	Foscarbidopa	DB16171	Unknown
6	Fluorodopa (18F)	DB13848	Unknown
7	Amantadine	DB00915	Li et al. [45]
8	Droxidopa	DB06262	Goldstein et al. [46]
9	Melevodopa	DB13313	Unknown
10	Levonordefrin	DB06707	Unknown

input the drug SMILES and the amino acid sequence into our model and obtained the protein attention matrix. Applying the mean operator to this matrix, we derived the protein attention vector, which reflects the distribution of attention values on the amino acid sequence. As shown in Table 8, the actual binding sites received closer attention from the model (the actual binding positions are amino acid positions 26, 45, 64, 78, 103, 106, 108, 114, and 121). Next, we mapped the attention vector to the 3D structure of the complex to visualize which regions in the protein have more significant contributions to the interaction. As depicted in Fig. 3, the left side of the figure displays the 3D structure of the ligand, with the highlighted red color indicating the binding sites, while the right side shows the amino acid sequence of the protein along with the attention weights learned by the model for each amino acid. The attention weights of each amino acid are represented by a heat map with different colors, and the actual binding positions on the amino acid sequence are highlighted in red. From Table 8 and Fig. 3, it can be observed that the sequence segments with higher attention weights in the heat map indeed cover multiple actual binding sites (such as amino acid positions 26, 103, 106, and 108 in the sequence). The analysis results indicate that our model can learn biological insights from the raw data. Additionally, we noticed that some non-binding sites are also highlighted, which may provide some suggestions for further investigations by relevant researchers.

4.7 Limitations and hypotheses

4.7.1 Limitations

In this study, we have identified certain limitations and proposed hypotheses to address them, aiming to enhance the understanding of drug-target interactions and improve the interpretability of our model.

- Unidirectional Focus: Our current reaction module only considers the influence of drugs on targets. However, drug and target interactions are mutually influential. To overcome this limitation, future work will explore more reasonable interaction mechanisms to investigate the bidirectional effects between drugs and targets.
- Interpretability of Attention Blocks: The interpretability of the attention blocks in our CSDTI model is limited by the dimensionality reduction performed by the MLP layer. Overcoming this limitation will help us gain deeper insights into the interaction between drugs and targets and enhance the interpretability of our model.

Table 8 Attention weights of the amino acid sequence of Riboflavin kinase(PDB: 1NB9) learned by the model

Protein sequnce	Attention Weights																			
1-20	0.53	0.54	0.46	0.61	0.58	0.41	0.47	0.53	0.58	0.59	0.46	0.58	0.61	0.35	0.37	0.48	0.54	0.61	0.54	0.43
21-40	0.49	0.39	0.45	0.47	0.78	0.81	0.74	0.48	0.41	0.57	0.35	0.38	0.45	0.43	0.42	0.51	0.39	0.34	0.58	0.43
41-60	0.61	0.58	0.51	0.48	0.41	0.38	0.47	0.51	0.39	0.58	0.41	0.42	0.37	0.56	0.54	0.45	0.47	0.39	0.46	0.55
61-80	0.58	0.54	0.38	0.43	0.84	0.81	0.79	0.86	0.76	0.47	0.39	0.46	0.44	0.45	0.49	0.53	0.52	0.54	0.57	0.47
81-100	0.51	0.53	0.55	0.61	0.49	0.47	0.49	0.61	0.58	0.47	0.39	0.45	0.61	0.47	0.43	0.51	0.53	0.51	0.51	0.57
101-120	0.52	0.76	0.88	0.87	0.69	0.81	0.81	0.85	0.61	0.39	0.58	0.54	0.78	0.89	0.44	0.41	0.57	0.42	0.58	0.54
121-140	0.87	0.81	0.75	0.47	0.61	0.45	0.45	0.46	0.52	0.38	0.39	0.55	0.37	0.57	0.61	0.6	0.39	0.4	0.47	0.51
141- 147	0.42	0.47	0.51	0.43	0.39	0.38	0.61													

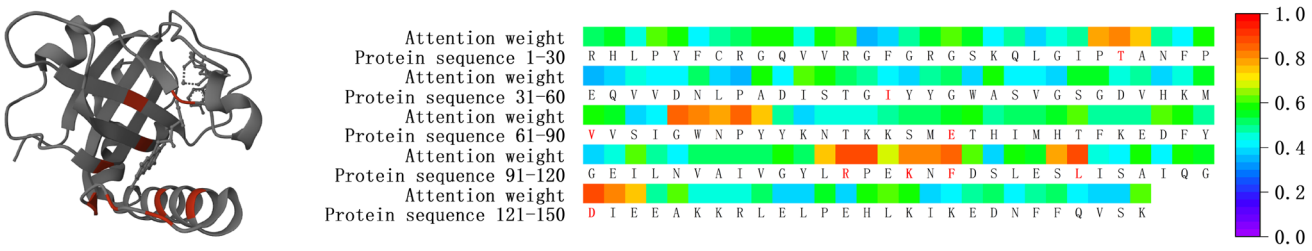

Fig. 3 Visualization of attention weights. The left side of the figure represents the three-dimensional structure of the ligand. The right side shows the amino acid sequence of the protein, with the attention weights of each amino acid represented by a heatmap using different colors

4.7.2 Hypotheses

- We hypothesize that the interaction between drug molecules and targets significantly impacts drug targeting prediction. We assume that effective modeling of drug-target interactions using deep learning models can improve the accuracy of drug targeting predictions.
- We assume that integrating the deep representations of drug molecules with the interaction features between drugs and targets can capture essential information related to drug targeting, leading to enhanced prediction accuracy.

It is important to note that these hypotheses are based on the identified limitations in our study and can be further refined and expanded to align with the specific objectives and methodology of your research.

5 Conclusion

Identifying potential drug-target interactions is a crucial step in drug repositioning. While existing work has been successful, significant challenges remain in improving DTI prediction performance and enhancing model interpretability. In this study, we propose an interpretable end-to-end deep learning architecture for DTI prediction. Our proposed network architecture can learn the interaction features between drugs and targets while effectively integrating deep representations of drugs and targets for downstream tasks. Additionally, our designed drug molecular aggregator can capture higher-order dependencies in drug molecules, overcoming the limitations of previous GNN-based approaches in molecular representation learning. Extensive experimental results demonstrate that CSDTI achieves competitive results compared to existing models on multiple benchmark datasets. For the DrugBank dataset, CSDTI shows significant improvements compared to the second-best method, with a 3.5% increase in AUC and a 3.7% increase in recall. Furthermore, the visualization of attention weights demonstrates that our approach can provide biological insights. Drug-target interaction is a highly complex process, and it is evident that there is room for improvement in CSDTI. For instance, in this study, we used the primary structure of proteins in learning their representations, which hinders the model's ability to capture protein structural information. The tertiary structure of proteins determines their biological functions, making the inclusion of higher-level structural information crucial for protein molecular representation. Therefore, in future work, we will focus on representing protein tertiary structures in deep learning-based models.

Acknowledgements All authors would like to thank the reviewers for the valuable comments

Author Contributions Both YZ and YP designed the method and experiments. YP performed the experiments and analyzed the results. YP and JZ wrote the manuscript. YZ and ML provided suggestions and feedback. All authors have read and approved the final manuscript

Funding This work is supported by a grant from the Natural Science Foundation of China (No. 62072070)

Availability of data and materials The datasets underlying this article are available in GitHub at https://github.com//ziduzidu/CSDTI

Declarations

Conflict of interest The authors declare that they have no conflicts of interest to report regarding the present study

Ethics approval No ethics approval was required for the study

References

1. Dickson M, Gagnon JP (2004) Key factors in the rising cost of new drug discovery and development. Nat Rev Drug Discov 3(5):417–429
2. Paul SM, Mytelka DS, Dunwiddie CT, Persinger CC, Munos BH, Lindborg SR, Schacht AL (2010) How to improve r&d productivity: the pharmaceutical industry's grand challenge. Nat Rev Drug Discov 9(3):203–214
3. Gordon PM, Hamid F, Makeyev EV, Houart C (2020) A conserved role for sfpq in repression of pathogenic cryptic last exons. bioRxiv, 2020–03

4. Noble ME, Endicott JA, Johnson LN (2004) Protein kinase inhibitors: insights into drug design from structure. Science 303(5665):1800–1805
5. Öztürk H, Özgür A, Ozkirimli E (2018) Deepdta: deep drug-target binding affinity prediction. Bioinformatics 34(17):821–829
6. Öztürk H, Ozkirimli E, Özgür A (2019) Widedta: prediction of drug-target binding affinity. arXiv:1902.04166
7. Zhao Q, Zhao H, Zheng K, Wang J (2022) Hyperattentiondti: improving drug-protein interaction prediction by sequence-based deep learning with attention mechanism. Bioinformatics 38(3):655–662
8. Nguyen T, Le H, Quinn TP, Nguyen T, Le TD, Venkatesh S (2021) Graphdta: predicting drug-target binding affinity with graph neural networks. Bioinformatics 37(8):1140–1147
9. Cheng Z, Yan C, Wu F-X, Wang J (2022) Drug-target interaction prediction using multi-head self-attention and graph attention network. IEEE/ACM Trans Comput Biol Bioinformatics 19(4):2208–2218. https://doi.org/10.1109/TCBB.2021.3077905
10. Li M, Lu Z, Wu Y, Li Y (2022) Bacpi: a bi-directional attention neural network for compound-protein interaction and binding affinity prediction. Bioinformatics 38(7):1995–2002
11. Schenone M, Dančík V, Wagner BK, Clemons PA (2013) Target identification and mechanism of action in chemical biology and drug discovery. Nat Chem Biol 9(4):232–240
12. Chen L, Tan X, Wang D, Zhong F, Liu X, Yang T, Luo X, Chen K, Jiang H, Zheng M (2020) Transformercpi: improving compound-protein interaction prediction by sequence-based deep learning with self-attention mechanism and label reversal experiments. Bioinformatics 36(16):4406–4414
13. Trott O, Olson AJ (2010) Autodock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem 31(2):455–461
14. Luo H, Mattes W, Mendrick DL, Hong H (2016) Molecular docking for identification of potential targets for drug repurposing. Curr Topics Med Chem 16(30):3636–3645
15. Pahikkala T, Airola A, Pietilä S, Shakyawar S, Szwajda A, Tang J, Aittokallio T (2015) Toward more realistic drug-target interaction predictions. Briefings Bioinformatics 16(2):325–337
16. He T, Heidemeyer M, Ban F, Cherkasov A, Ester M (2017) Simboost: a read-across approach for predicting drug-target binding affinities using gradient boosting machines. J Cheminformatics 9(1):1–14
17. MacLean F (2021) Knowledge graphs and their applications in drug discovery. Expert Opinion Drug Discov 16(9):1057–1069
18. Wu Y, Gao M, Zeng M, Zhang J, Li M (2022) Bridgedpi: a novel graph neural network for predicting drug-protein interactions. Bioinformatics 38(9):2571–2578
19. Karimi M, Wu D, Wang Z, Shen Y (2019) Deepaffinity: interpretable deep learning of compound-protein affinity through unified recurrent and convolutional neural networks. Bioinformatics 35(18):3329–3338
20. Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, Li Q, Shoemaker BA, Thiessen PA, Yu B et al (2019) Pubchem 2019 update: improved access to chemical data. Nucleic Acids Res 47(D1):1102–1109
21. Zheng S, Li Y, Chen S, Xu J, Yang Y (2020) Predicting drug-protein interaction using quasi-visual question answering system. Nat Mach Intell 2(2):134–140
22. Zu S, Chen T, Li S (2015) Global optimization-based inference of chemogenomic features from drug-target interactions. Bioinformatics 31(15):2523–2529
23. Lim J, Ryu S, Park K, Choe YJ, Ham J, Kim WY (2019) Predicting drug-target interaction using a novel graph neural network with 3d structure-embedded graph representation. J Chem Inf Model 59(9):3981–3988
24. Zhang B, Xiong D, Su J (2018) Neural machine translation with deep attention. IEEE Trans Pattern Anal Mach Intell 42(1):154–163
25. Yang Z, Yang D, Dyer C, He X, Smola A, Hovy E (2016) Hierarchical attention networks for document classification. In: Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp 1480–1489
26. Anderson P, He X, Buehler C, Teney D, Johnson M, Gould S, Zhang L (2018) Bottom-up and top-down attention for image captioning and visual question answering. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp 6077–6086
27. Seo S, Huang J, Yang H, Liu Y (2017) Interpretable convolutional neural networks with dual local and global attention for review rating prediction. In: Proceedings of the Eleventh ACM Conference on Recommender Systems, pp 297–305
28. Kim Y, Shin B (2021) An interpretable framework for drug-target interaction with gated cross attention. In: Machine Learning for Healthcare Conference, pp 337–353. PMLR
29. Kurata H, Tsukiyama S (2022) Ican: Interpretable cross-attention network for identifying drug and target protein interactions. Plos one 17(10):0276609
30. Lee I, Keum J, Nam H (2019) Deepconv-dti: Prediction of drug-target interactions via deep learning with convolution on protein sequences. PLoS Comput Biol 15(6):1007129
31. Huang K, Xiao C, Glass LM, Sun J (2021) Moltrans: molecular interaction transformer for drug-target interaction prediction. Bioinformatics 37(6):830–836
32. Yang Z, Zhong W, Zhao L, Chen CY-C (2022) Mgraphdta: deep multiscale graph neural network for explainable drug-target binding affinity prediction. Chem Sci 13(3):816–833
33. Alley EC, Khimulya G, Biswas S, AlQuraishi M, Church GM (2019) Unified rational protein engineering with sequence-based deep representation learning. Nat Methods 16(12):1315–1322
34. Dauphin YN, Fan A, Auli M, Grangier D (2017) Language modeling with gated convolutional networks. In: International Conference on Machine Learning, pp 933–941
35. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp 770–778
36. Wang H, Cao P, Wang J, Zaiane OR (2022) Uctransnet: rethinking the skip connections in u-net from a channel-wise perspective with transformer. Proceedings of the AAAI Conference on Artificial Intelligence 36:2441–2449
37. Wishart DS, Knox C, Guo AC, Shrivastava S, Hassanali M, Stothard P, Chang Z, Woolsey J (2006) Drugbank: a comprehensive resource for in silico drug discovery and exploration. Nucleic Acids Res 34(suppl_1):668–672
38. Bento AP, Hersey A, Félix E, Landrum G, Gaulton A, Atkinson F, Bellis LJ, De Veij M, Leach AR (2020) An open source chemical structure curation pipeline using rdkit. J Cheminformatics 12:1–16
39. Liu H, Sun J, Guan J, Zheng J, Zhou S (2015) Improving compound-protein interaction prediction by building up highly credible negative samples. Bioinformatics 31(12):221–229
40. Davis MI, Hunt JP, Herrgard S, Ciceri P, Wodicka LM, Pallares G, Hocker M, Treiber DK, Zarrinkar PP (2011) Comprehensive analysis of kinase inhibitor selectivity. Nat Biotechnol 29(11):1046–1051
41. Tang J, Szwajda A, Shakyawar S, Xu T, Hintsanen P, Wennerberg K, Aittokallio T (2014) Making sense of large-scale kinase inhibitor bioactivity data sets: a comparative and integrative analysis. J Chem Inf Model 54(3):735–743
42. Sheng X, Zhu X, Zhang Y, Cui G, Peng L, Lu X, Zang YQ (2012) Rhein protects against obesity and related metabolic disorders through liver x receptor-mediated uncoupling protein 1 upregulation in brown adipose tissue. Int J Biol Sci 8(10):1375–1384

43. Tang J-c, Yang H, Song X-y, Song X-h, Yan S-l, Shao J-q, Zhang T-l, Zhang J-n (2009) Inhibition of cytochrome p450 enzymes by rhein in rat liver microsomes. Phytotherapy Res: An Int J Devoted Pharmacol Toxicol Evaluation Nat Product Derivatives 23(2):159–164
44. Chen X, Ji ZL, Chen YZ (2002) Ttd: therapeutic target database. Nucleic Acids Res 30(1):412–415
45. Li X-M, Juorio AV, Qi J, Boulton AA (1998) Amantadine increases aromatic 1l-amino acid decarboxylase mrna in pc12 cells. J Neurosci Res 53(4):490–493
46. Goldstein DS (2006) l-dihydroxyphenylserine (l-dops): a norepinephrine prodrug. Cardiovascular Drug Rev 24(3–4):189–203

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Yaohua Pan received a bachelor's degree from Qingdao Agricultural University in 2021. He is currently studying for a master's degree in the Computer Science and Technology Department of Dalian Maritime University. His research interests include natural language processing and bioinformatics.

Yijia Zhang received the BSc, MSc and PhD degrees from the Dalian University of Technology, China, in 2003, 2009 and 2014. He is a professor in the College of Information Science and Technology at the Dalian Maritime University. He has published more than 60 research papers on topics in text mining and bioinformatics.

Jing Zhang received a bachelor's degree from Zaozhuang University in 2019. Currently, she is studying for a master's degree at the School of Information Science and Technology, Dalian Maritime University.

Mingyu Lu born in 1963, received his doctor'sdegree from Tsinghua University in 2002, and is now a Profession and doctoral supervisor at Dalian Maritime University. His research interests include data mining, pattern recognition, machine learning, and natural language processing.