

RESEARCH

Open Access



MvGraphDTA: multi-view-based graph deep model for drug-target affinity prediction by introducing the graphs and line graphs

Xin Zeng¹, Kai-Yang Zhong¹, Pei-Yan Meng¹, Shu-Juan Li², Shuang-Qing Lv³, Meng-Liang Wen⁴ and Yi Li^{1*}

Abstract

Background Accurately identifying drug-target affinity (DTA) plays a pivotal role in drug screening, design, and repurposing in pharmaceutical industry. It not only reduces the time, labor, and economic costs associated with biological experiments but also expedites drug development process. However, achieving the desired level of computational accuracy for DTA identification methods remains a significant challenge.

Results We proposed a novel multi-view-based graph deep model known as MvGraphDTA for DTA prediction. MvGraphDTA employed a graph convolutional network (GCN) to extract the structural features from original graphs of drugs and targets, respectively. It went a step further by constructing line graphs with edges as vertices based on original graphs of drugs and targets. GCN was also used to extract the relationship features within their line graphs. To enhance the complementarity between the extracted features from original graphs and line graphs, MvGraphDTA fused the extracted multi-view features of drugs and targets, respectively. Finally, these fused features were concatenated and passed through a fully connected (FC) network to predict DTA.

Conclusions During the experiments, we performed data augmentation on all the training sets used. Experimental results showed that MvGraphDTA outperformed the competitive state-of-the-art methods on benchmark datasets for DTA prediction. Additionally, we evaluated the universality and generalization performance of MvGraphDTA on additional datasets. Experimental outcomes revealed that MvGraphDTA exhibited good universality and generalization capability, making it a reliable tool for drug-target interaction prediction.

Keywords Multi-view, Drug-target affinity, Graph deep model, Line graphs, Data augmentation

Background

Target proteins represent an important category of biological macromolecules in cells, interacting with various molecules such as drugs (small molecules), proteins, nucleic acids, peptides, and other substrates. Accurately detecting drug-target interaction (DTI) or quantifying drug-target binding strength, known as drug-target affinity (DTA), holds significant importance. This forms the foundation for identifying target proteins [1], discovering drugs [2], and treating unknown diseases [3] at the molecular level. Currently, two primary methods are employed to identify DTA: biological experiments [4] and computational methods [5, 6]. However,

*Correspondence:

Yi Li

yili@dali.edu.cn

¹ College of Mathematics and Computer Science, Dali University, Dali 671003, China

² Yunnan Institute of Endemic Diseases Control & Prevention, Dali 671000, China

³ Institute of Surveying and Information Engineering, West Yunnan University of Applied Science, Dali 671000, China

⁴ State Key Laboratory for Conservation and Utilization of Bio-Resources in Yunnan, Yunnan University, Kunming 650000, China



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

biological experiments entail lengthy cycles and substantial resource investments, both in labor and finances. Hence, the development of computational methods for accurately predicting DTA stands as an essential research pursuit in the field of drug discovery.

After nearly two decades of research, the field of predicting DTA has witnessed a transformation from structure-based molecular docking methods [7] to computational methods based on machine learning [5, 8, 9]. While molecular docking methods have proven effective, they suffer from notable drawbacks [10], including high demands on computational resources, lengthy processing times, and a requirement for structure files of drugs and targets. Recent successes in applying machine learning models to computational biology, such as toxicity monitoring [11], drug-drug interaction prediction [12], and protein–protein interaction sites prediction [13], have opened up new avenues for DTA research. The emergence of numerous computational methods based on machine learning has firmly established them as the predominant computational methods for DTA prediction. Despite the impressive performance of traditional machine learning methods, they heavily rely on user-provided features for drugs and targets, lacking the ability to autonomously extract the hidden features from input raw data. The advent of deep learning has effectively addressed this limitation by extracting high-level features directly from the original data of drugs and targets, enabling the creation of end-to-end prediction models of DTA. Furthermore, deep learning methods can achieve performance comparable to or even superior to traditional machine learning methods [14]. As a result, the pursuit of efficient DTA prediction methods using deep learning has emerged as a prominent research trend.

Currently, research on predicting DTI using deep learning can be categorized into two main types: classification task and regression task. In the classification task, the goal is to predict whether there is an interaction between drug and target. For instance, deep learning models were firstly employed to extract implicit crucial features from diverse data modalities such as sequences and structures in some methods like MCANet [15], DeepConv-DTI [16], TransformerCPI [17], HyperAttentionDTI [18], MolTrans [19], BridgeDPI [20], CoaDTI [21], and iGRLDTI [22]. Then fully connected (FC) networks were utilized for binary classification predictions. On the other hand, the regression task aims to predict DTA. Methods like DeepDTA [14], DeepDTAF [23], Pafnucy [24], BAPA [25], FAST [26], OnionNet [27], IMCPSF [28], GLI [29], CAPLA [30], and GraphscoreDTA [31] fall into this category. Among them, DeepDTA, DeepDTAF, and CAPLA primarily predicted DTA by leveraging the sequence features of drugs and targets. Models, such

as Convolutional Neural Network (CNN) [32], Recurrent Neural Network (RNN) [33], and attention mechanism [34], were employed to capture hidden high-level sequence features from drug SMILES (Simplified Molecular Input Line Entry System) [35] and target sequences. The remaining methods predominantly predicted DTA based on the structural features of drugs and targets. These structural features were derived from the structures of drugs, targets, or their complexes using Graph Neural Network (GNN) [36] or 3D-CNN (3D Convolutional Neural Network).

While many methods employ GNN to extract the structural features from drug and target graphs, and some methods aggregate the edge features, they may not fully capture the features of drugs and targets, particularly the interaction relationship features between vertices, i.e., the high-level features hidden in the edges. This is because they often focus solely on using edges for aggregating vertex features or enhancing these features with edge-based information, without fully exploring the potential impact of the interaction relationship features on the model's performance.

To address this issue, we proposed a novel graph deep model, MvGraphDTA, which leveraged multiple views such as graphs and line graphs of drugs and targets for DTA prediction. Initially, MvGraphDTA employed GCN to extract the structural features and interaction relationship features from graphs and line graphs of drugs and targets, respectively. It then employed a fusion strategy to fuse the structural features and interaction relationship features of drugs and targets, enhancing the complementarity between two types of features. Finally, the fused features of drugs and targets were concatenated and fed into a FC network to predict DTA. Simultaneously, we also proposed a data augmentation strategy to expand the training dataset and provide more comprehensive training for the model. Experimental results showed that MvGraphDTA outperformed the competitive state-of-the-art methods. Additionally, we also assessed the universality and generalization performance of MvGraphDTA on other datasets, and these results revealed that MvGraphDTA had good practical performance and was a reliable multitask prediction tool for drug-target interaction.

Results and discussion

Performance comparison of MvGraphDTA with competitive state-of-the-art methods

To evaluate the performance of MvGraphDTA, we initially trained the model on the PDBbind_2016 augmented training set using 5-fold cross-validation and assessed its performance on the test set. Subsequently, we compared the performance of MvGraphDTA with

eight deep learning-based methods. These included DeepDTA, DeepDTAF, and CAPLA, which were sequence-based methods, while Pafnucy, OnionNet, FAST, IMCP-SF, and GLI leveraged the structures of drugs and targets. Experimental results (Table 1) showed that MvGraphDTA outperformed the competitive state-of-the-art methods across all evaluation metrics. Specifically, there were improvements of 6.4% (MAE), 4.8% (RMSE), 1.2% (PCCs), and 1% (CI) in comparison to the optimal values achieved by the state-of-the-art methods, respectively.

To thoroughly assess the model's performance, we trained MvGraphDTA on PDBbind_2019 and subsequently carried out independent testing under CASF2013 and CASF2016. We benchmarked MvGraphDTA against six competitive state-of-the-art methods based on deep learning. During the preprocessing phase of PDBbind_2019, we initially applied data augmentation on the dataset derived from drug similarity clustering. Following this, the model underwent training on the augmented dataset. Finally, we evaluated MvGraphDTA's performance on CASF2013 and CASF2016 and compared it with state-of-the-art methods. Experimental results (Table 2) revealed that under CASF2013, MvGraphDTA showed improvements across all evaluation metrics (MAE, 1.4%; RMSE, 3.9%; PCCs, 3.3%; CI, 0.1%), compared to the optimal values of state-of-the-art methods. In CASF2016, the performance metrics of MvGraphDTA were slightly lower than those of GraphscoreDTA. For example, MvGraphDTA exhibited a lower MAE by 0.006, RMSE by 0.002, PCCs by 0.007, and CI by 0.012. One potential reason for this was that CASF2013 was a subset of CASF2016. In CASF2016, additional drug-target pairs were introduced based on drug similarity clustering, resulting in a greater number of pairs compared to CASF2013. However, within these pairs in CASF2016, there existed

Table 1 Performance comparison of MvGraphDTA with competitive state-of-the-art methods under the PDBbind_2016 dataset

Methods	MAE↓	RMSE↓	PCCs↑	CI↑
DeepDTA	1.148	1.443	0.749	0.771
DeepDTAF	1.073	1.355	0.789	0.799
Pafnucy	1.129	1.418	0.775	0.789
OnionNet	—	1.278	0.816	—
FAST	1.019	1.308	0.810	—
IMCP-SF	1.155	1.452	0.791	—
GLI	1.026	1.294	—	—
CAPLA	0.966	1.200	0.843	0.820
MvGraphDTA	0.904	1.142	0.853	0.828

Table 2 Performance comparison of MvGraphDTA with state-of-the-art methods under drug similarity clustering dataset

Dataset	Methods	MAE↓	RMSE↓	PCCs↑	CI↑
CASF2013	DeepDTA	1.319	1.608	0.710	0.754
	DeepDTAF	1.506	2.010	0.616	0.702
	Pafnucy	1.376	1.716	0.662	0.738
	BAPA	1.648	2.014	0.620	0.719
	FAST	1.387	1.720	0.662	0.737
	GraphscoreDTA	1.179	1.486	0.757	0.782
	MvGraphDTA	1.162	1.428	0.782	0.783
CASF2016	DeepDTA	1.177	1.458	0.745	0.771
	DeepDTAF	1.130	1.426	0.765	0.780
	Pafnucy	1.253	1.527	0.754	0.776
	BAPA	1.582	1.936	0.615	0.714
	FAST	1.207	1.513	0.746	0.773
	GraphscoreDTA	0.981	1.249	0.831	0.819
	MvGraphDTA	0.987	1.251	0.824	0.807

significant differences between the targets. These differences hindered MvGraphDTA from effectively extracting deep structural features of targets.

On the other hand, we first applied data augmentation on the dataset derived from target similarity clustering, based on PDBbind_2019. We then trained the model using the augmented dataset. Finally, we also evaluated the performance of MvGraphDTA on CASF2013 and CASF2016, comparing it with competitive state-of-the-art methods. Experimental results (Table 3) indicated that, under CASF2013, MvGraphDTA improved MAE and RMSE by 1.5% and 1.7%, respectively, compared to GraphscoreDTA, which achieved the best performance.

Table 3 Performance comparison of MvGraphDTA with state-of-the-art methods under target similarity clustering dataset

Dataset	Methods	MAE↓	RMSE↓	PCCs↑	CI↑
CASF2013	DeepDTA	1.417	1.785	0.654	0.731
	DeepDTAF	1.514	1.791	0.672	0.732
	Pafnucy	1.335	1.666	0.687	0.748
	BAPA	1.571	1.923	0.625	0.721
	FAST	1.252	1.554	0.741	0.777
	GraphscoreDTA	1.235	1.542	0.758	0.780
	MvGraphDTA	1.216	1.516	0.742	0.772
CASF2016	DeepDTA	1.264	1.578	0.728	0.762
	DeepDTAF	1.179	1.445	0.782	0.792
	Pafnucy	1.231	1.494	0.756	0.779
	BAPA	1.528	1.886	0.628	0.719
	FAST	1.170	1.467	0.757	0.781
	GraphscoreDTA	1.053	1.349	0.810	0.811
	MvGraphDTA	1.005	1.283	0.810	0.807

However, in terms of PCCs and CI, it experienced a decrease of 2.1% and 1%, respectively, compared to GraphscoreDTA. For CASF2016, MvGraphDTA exhibited improvements in MAE and RMSE by 4.6% and 4.9%, respectively, compared to GraphscoreDTA. Both methods achieved an optimal value of 0.81 in PCCs, but in CI, MvGraphDTA exhibited a decrease of 0.49% compared to GraphscoreDTA.

To sum up, MvGraphDTA outperformed competitive state-of-the-art methods in the test set of PDBbind_2016, and it also exhibited comparable or superior performance in the independent test sets CASF2013 and CASF2016. These results highlighted the reliability of MvGraphDTA as a tool for predicting DTA.

Ablation experiment

Line graphs for multi-view feature fusion

Based on the utilization of GNN for feature extraction from drug and target graphs, we introduced a multi-view feature extraction approach incorporating the line graphs to achieve better performance compared to the single-view method. Firstly, we constructed single-view methods SVM and ESVM, leveraging only drug and target graphs. SVM comprised three consecutive graph convolutional layers to aggregate adjacent vertex information for extracting hidden high-level features from drug and target graphs. These features were then concatenated into a FC network for DTA prediction. Building upon SVM, ESVM further integrated the edge information of drug and target graphs, thereby enhancing the feature extraction. Subsequently, we introduced a multi-view method (MVM) that considered drug and target graphs alongside line graphs, although

it excluded edge information. Finally, SVM, ESVM, MVM, and MvGraphDTA were trained and tested on PDBbind_2016. Experimental results (Table 4) showed that ESVM outperformed SVM across all evaluation metrics, highlighting the beneficial impact of edge information aggregation. MVM achieved superior performance compared to SVM, indicating that utilizing multi-view features can help improve the model's performance. And MvGraphDTA exhibited better performance than both single-view model (SVM, ESVM) and multi-view model (MVM), with improvements in MAE, RMSE, PCCs, CI, and R^2 by 1.8%, 0.8%, 0.4%, 0%, and 0.8%, respectively. These findings underscored the effectiveness of multi-view feature fusion based on graphs and line graphs composition in enhancing the model's performance.

Number of graph convolutional layers

In Table 4, we opted for the common practice of using 3 graph convolutional layers for MvGraphDTA. Generally, as the number of graph convolutional layers increases, the disparities among the vertex features extracted by the network diminish gradually, resulting in a certain degree of transitional smoothness in these features. This smoothness can pose a challenge for the model in terms of convergence and effective parameter learning. To further refine the model's architecture, we conducted experiments with varying numbers of graph convolutional layers. Experimental results (Table 5) revealed that MvGraphDTA exhibited improved performance when employing 2 to 4 graph convolutional layers, particularly excelling with 2 or 4 layers. Key evaluation metrics such as MAE, RMSE, PCCs, CI, and R^2 approached performance levels of approximately 0.93, 1.19, 0.83, 0.82, and 0.69, respectively. However, when utilizing 5 graph convolutional layers, the model began to experience issues with over-smoothing in the extracted vertex features, leading to a significant decline in performance. Following a comparative analysis of performance, we ultimately decided to set the number of graph convolutional layers for MvGraphDTA to 2. This choice not only exhibited the optimal performance, but also reduced the number of learning parameters and accelerated execution speed.

Table 4 Performance comparison of single-view and multi-view methods

Methods	MAE↓	RMSE↓	PCCs↑	CI↑	R^2 ↑
SVM	1.034	1.321	0.798	0.799	0.631
ESVM	0.999	1.279	0.809	0.804	0.654
MVM	1.012	1.259	0.817	0.812	0.665
MvGraphDTA	0.981	1.249	0.820	0.812	0.670

Table 5 Performance comparison of MvGraphDTA with different numbers of graph convolutional layers

Methods	Layers	MAE↓	RMSE↓	PCCs↑	CI↑	R^2 ↑
MvGraphDTA	2	0.934	1.192	0.839	0.821	0.699
	3	0.981	1.249	0.820	0.812	0.670
	4	0.935	1.195	0.837	0.818	0.697
	5	1.004	1.308	0.779	0.800	0.638

Table 6 Performance comparison of MvGraphDTA under data augmentation

Method	Data augmentation	MAE↓	RMSE↓	PCCs↑	CI↑	R ² ↑
MvGraphDTA	No	0.934	1.192	0.839	0.821	0.699
	Yes	0.904	1.142	0.853	0.828	0.724

Data augmentation

Firstly, we employed data augmentation to expand the samples of the training set in PDBbind_2016, effectively generating two new drug-target pairs based on each original pair. During the generation of these new pairs, if traversing all amino acid residues of the target fails to yield two new complete target subgraphs, we adjusted the number of new drug-target pairs generated (up to 1) based on the actual complete target subgraph formation. To maintain consistency between the newly generated complete target subgraph and the original target graph, minimizing potential impacts on drug-target binding affinity, we opted to randomly remove adjacent amino acid residues at a rate of 10%. This augmentation expanded the training set of PDBbind_2016 from 12,993 to 38,905 drug-target pairs. Subsequently, we trained MvGraphDTA through 5-fold cross-validation on this augmented training set. Finally, we evaluated MvGraphDTA's performance on the test set of PDBbind_2016. Experimental results (Table 6) showed that MvGraphDTA exhibited improvements of 3.2%, 4.2%, 1.7%, 0.9%, and 3.6% across MAE, RMSE, PCCs, CI, and R², respectively. Although the approach of selectively removing amino acid residues to form new complete target subgraphs while preserving the original drug-target pair affinities may seem to defy biological intuition, the resulting target subgraphs retained a high degree of similarity to the original target graph, thus preserving crucial information. Comparative analysis of experimental results further affirmed that data augmentation of the training set facilitated the extraction of implicit crucial features, consequently enhancing the model's performance.

Universal analysis of MvGraphDTA for drug-target interaction prediction

Apart from exhibiting good performance in predicting DTA, we further assessed the universality of MvGraphDTA in predicting DTI. Firstly, we expanded the training set of Enzymes using data augmentation. Then, MvGraphDTA underwent training using 5-fold cross-validation. Finally, we conducted a comparative analysis of MvGraphDTA's performance with those of competitive state-of-the-art methods on the test set of Enzymes. Experimental findings (Table 7) indicated that MvGraphDTA not only achieved similar optimal values

Table 7 Performance comparison of MvGraphDTA in drug-target interaction prediction

Dataset	Methods	Accuracy↑	AUC↑	AUPRC↑
Enzymes	DeepConv-DTI	0.747	0.802	0.757
	TransformerCPI	0.501	0.484	0.491
	HyperAttentionDTI	0.451	0.418	0.456
	MolTrans	0.727	0.789	0.735
	MCANet	0.760	0.804	0.730
	MvGraphDTA	0.790	0.791	0.838

Table 8 Generalization performance comparison of MvGraphDTA with state-of-the-art methods under CSAR-HIQ_51 and CSAR-HIQ_36 datasets

Dataset	Methods	MAE↓	RMSE↓	PCCs↑	CI↑
CSAR-HIQ_51	DeepDTAF	1.862	2.272	0.606	0.710
	Pafnucy	1.667	1.944	0.622	0.698
	IMCP-SF	1.278	1.629	0.769	0.780
	CAPLA	1.550	1.848	0.686	0.727
	MvGraphDTA	1.412	1.681	0.706	0.711
CSAR-HIQ_36	DeepDTAF	2.318	2.765	0.543	0.670
	Pafnucy	1.291	1.658	0.566	0.566
	IMCP-SF	1.205	1.560	0.631	0.748
	CAPLA	1.160	1.454	0.704	0.760
	MvGraphDTA	1.096	1.366	0.734	0.779

to state-of-the-art methods in terms of AUC, but also exhibited improvements of 3.9% and 14.8% in Accuracy and AUPRC, compared to the optimal values obtained with state-of-the-art methods. These results suggested that MvGraphDTA can effectively serve as a multi-tasking tool for both predicting DTA and DTI.

Analysis of generalization performance of MvGraphDTA

Following the completion of performance and universality evaluation of MvGraphDTA, we proceeded to assess its generalization performance on two independent test sets, CSAR-HIQ_51 and CSAR-HIQ_36, and compared its performance with four state-of-the-art methods. Experimental results (Table 8) showed that while MvGraphDTA exhibited good generalization performance on CSAR-HIQ_51, its various evaluation

metrics were slightly lower than those of the competitive method IMCP-SF, which achieved the optimal performance. The reason for this may be that drug-target pairs in CSAR-HIQ_51 exhibited notable diversity compared to the training set of MvGraphDTA, thereby limiting the ability of MvGraphDTA to accurately extract their structure features. However, On CSAR-HIQ_36, MvGraphDTA outperformed all competitive methods, with the key evaluation metrics such as RMSE and PCCs increasing by 6.1% and 4.3%, respectively, compared to the best-performing competitive method CAPLA. These findings suggested that MvGraphDTA exhibited robust generalization performance and was reliable for practical applications.

Interpretability analysis of MvGraphDTA

In the process of training deep learning models, understanding the importance of each input feature is crucial for interpretability analysis and performance optimization of model. In this study, we employed a gradient-based approach to evaluate the importance of features. This method involved calculating the gradients of each feature during training to quantify its impact on the model's performance. The gradient reflected how sensitive the model's loss function is to changes in input features. Larger absolute gradients indicated greater impact on performance. To perform an explanatory analysis of the effectiveness of MvGraphDTA, we randomly selected two drug-target pairs, sample 1eby and sample 3e05, from the test set. We calculated the gradients of the features of drug and target in each sample and averaged these gradients to quantify the importance of each feature. Our analysis of experimental results revealed that atoms and amino acid residues involved in drug-target interaction obtained higher scores (Figs. 1a and 2a). Furthermore, we visualized the atoms (Figs. 1b and 2b) and amino acid residues (Figs. 1c and 2c) that scored higher. The visualization results indicated that MvGraphDTA effectively identified key atoms and amino acid residues in the drug and target, especially the identified amino acid residues were distributed around the pocket where drug and target bound.

Conclusions

MvGraphDTA was a novel deep learning method based on graph convolutional networks for predicting DTA. It integrated multi-view features of graphs and line graphs from drugs and targets, and employed data augmentation to achieve superior performance compared to competitive state-of-the-art methods across multiple datasets. Furthermore, we conducted a comprehensive evaluation of MvGraphDTA's universality and generalization capabilities. Experimental results showed MvGraphDTA's good

universality in DTI and effectiveness in real-world applications. Despite numerous experimental validations confirming MvGraphDTA's reliability as a DTA prediction tool, there are still some limitations and shortcomings that need to be addressed:

- (1) MvGraphDTA predominantly relies on the three-dimensional structures of targets, yet many target sequences lack experimentally determined structures. While some computational tools like AlphaFold2 offer a solution to predict these structures, their accuracy is higher for monomeric targets, leaving room for improvement in predicting the structures of multi-body targets.
- (2) Although MvGraphDTA exhibited good performance leveraging the structural aspects of drugs and targets, it overlooked the sequence features, which also played a significant role in predicting DTA.

In the future, to further improve the performance of DTA prediction, especially the generalization capabilities, we plan to delve deeper into two key aspects:

- (1) Introducing cutting-edge deep learning technologies such as contrastive learning, heterogeneous information networks [22], and dual-channel hypergraph convolutional network [37] into the architecture of MvGraphDTA: by incorporating these technologies, MvGraphDTA can learn more comprehensive and in-depth features of both drugs and targets. This approach aims to improve the performance and generalization abilities of MvGraphDTA.
- (2) Expansion beyond structural features: we intend to moderately incorporate the sequence features of drugs and targets and motif features of targets using attention mechanisms [38] to improve the model's performance.

Methods

Datasets

In this study, PDBbind_2016 and PDBbind_2019 [39] served as the benchmark datasets for predicting DTA, while CASF2016 [40] and CASF2013 [41] were utilized as independent test sets. PDBbind_2016 comprises three segments: standard set, refined set, and core set, containing 9226, 4057, and 290 drug-target pairs, respectively. We selected the core set as the test set. However, it is worth noting that the core set is derived from the standard and refined sets. Therefore, by merging the standard and refined sets and subsequently removing the samples of the core set, we obtained

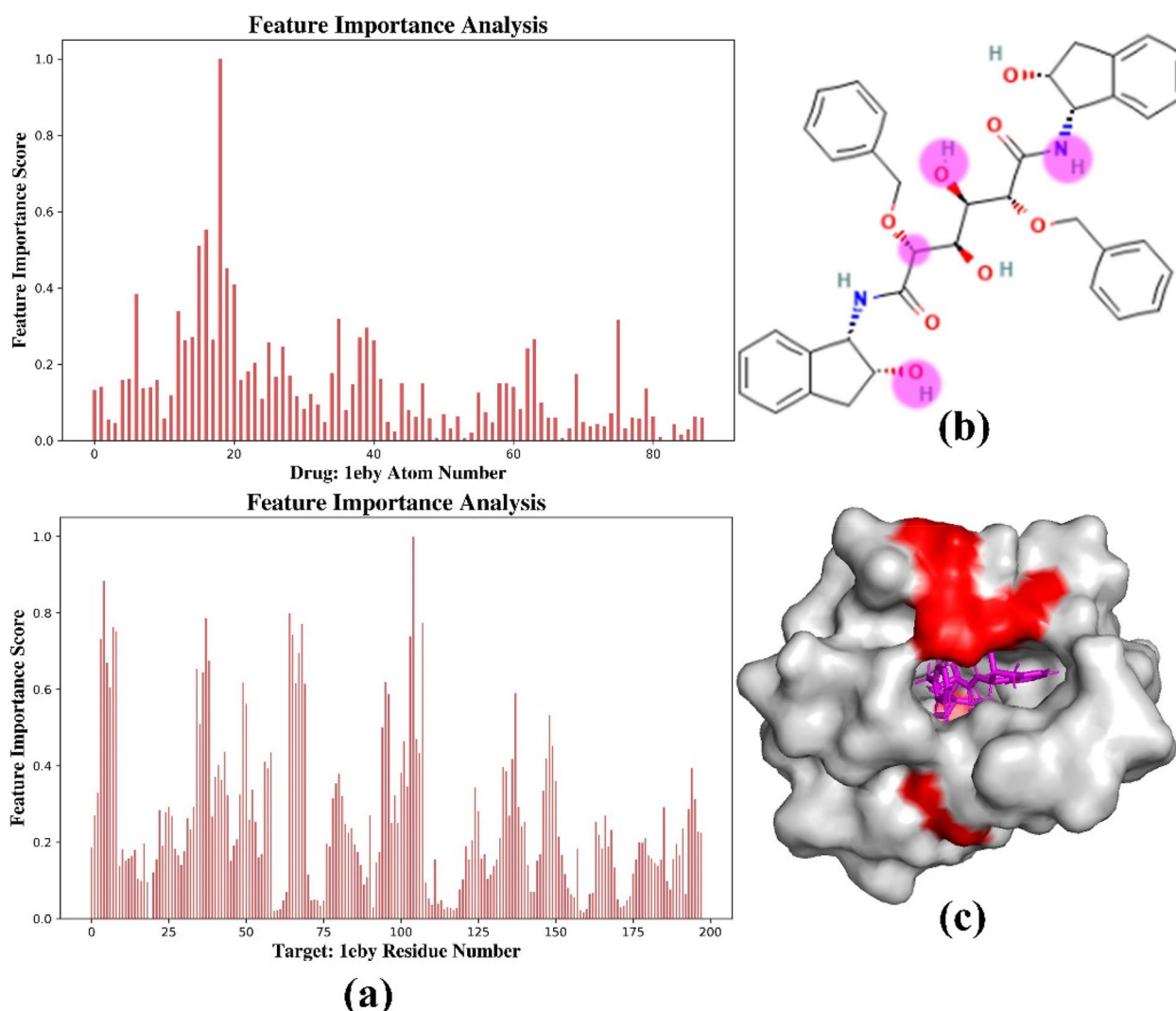


Fig. 1 Interpretability analysis of case 1eby based on MvGraphDTA. **a** Scores for the importance of features of atoms and amino acid residues. **b** Atoms with higher scores in drug (purple). **c** Distribution of amino acid residues with higher scores in target (red)

12,993 samples for model training and validation. As for PDBbind_2019, CASF2016, and CASF2013, they consist of 17,652, 285, and 192 samples, respectively. Employing identical data preprocessing and clustering methods as GraphscoreDTA [31], we eventually retained 13,851, 279, and 182 samples, respectively. The clustering results based on drug and target similarity for PDBbind_2019 are presented in Table 9.

In addition to evaluating the model's performance in predicting DTA for regression task, we also assessed the model's universality in predicting DTI for classification task using Enzymes dataset, as mentioned in MCANet [15]. Enzymes dataset consists of 2920 pairs of drug-target positive samples and 2920 pairs

of drug-target negative samples, derived from 660 targets and 444 drugs. We split the Enzymes dataset in an 8:2 ratio to allocate the number of training and testing samples, as detailed in Table 10.

Finally, we further evaluated the generalization performance of MvGraphDTA using the CSAR-HiQ dataset [42]. CSAR-HiQ dataset comprises two subsets, containing 176 and 167 drug-target pairs. Due to sample overlap between these subsets and the training set of PDBbind_2016, we excluded the overlapped drug-target pairs based on PDBID of target. This process resulted in obtaining 51 and 36 drug-target pairs, which we named as CSAR-HiQ_51 and CSAR-HiQ_36 datasets, respectively.

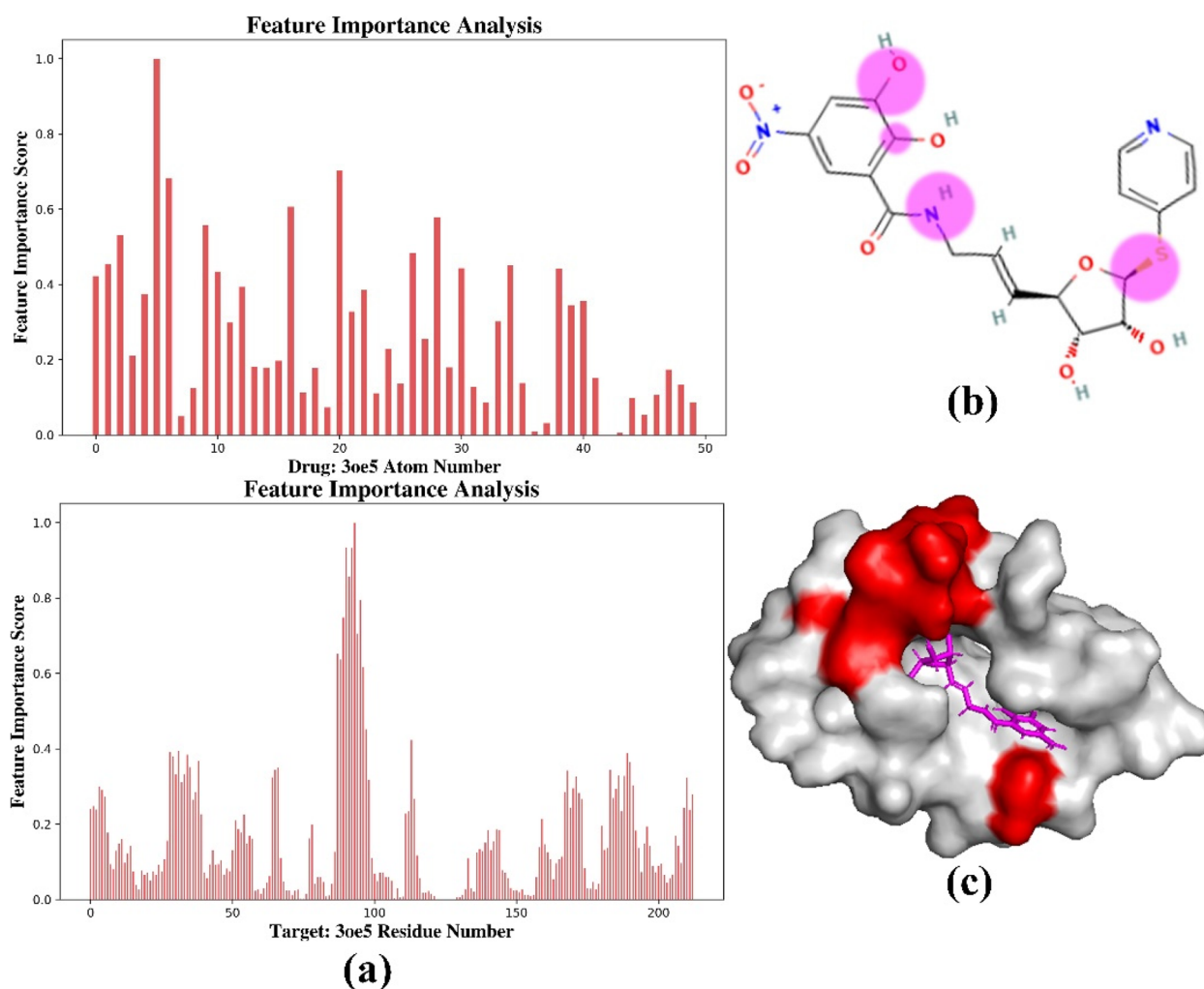


Fig. 2 Interpretability analysis of case 3oe5 based on MvGraphDTA. **a** Scores for the importance of features of atoms and amino acid residues. **b** Atoms with higher scores in drug (purple). **c** Distribution of amino acid residues with higher scores in target (red)

Table 9 Partition results of PDBbind_2019 dataset based on clustering method

Clustering method	Training set	Validation set	Test set
Drug similarity-based	9869	1328	2654
Target similarity-based	9674	1158	3019

Table 10 Enzymes dataset and partitioning results

Dataset	Total samples	Positive samples	Negative samples	Training set	Test set
Enzymes	5840	2920	2920	4672	1168

Multi-view representation of drugs

Multi-view representation of drugs can be categorized into two main types: drug graphs and their corresponding line graphs (Fig. 3). For PDBbind_2016 and PDBbind_2019, we constructed a drug graph using the atoms from the provided structure files (.mol2 or.sdf) as vertices, and the bonds between these atoms as edges. Simultaneously, to capture more comprehensive features of drugs, we built the line graphs based on line graph theory. Here, we treated the edges of drug graph as vertices, and vice versa. For the samples lacking structure files of drugs, we employed RDKit tool [43] to convert drug SMILES into their mol objects. From these, we constructed the graphs and line graphs of drugs. Vertex features were represented using one-hot encoding

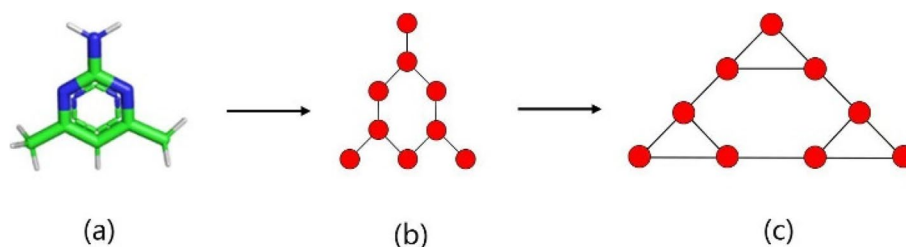


Fig. 3 Graph and line graphs of drug. **a** Drug. **b** Graph of drug. **c** Line graph of drug

based on the count of different atomic types obtained from the dataset statistics. Regarding the edge features, they were categorized into two types based on the bond types (Table 11): (1) The edge features with provided three-dimensional structures of drugs were encoded using a 5-dimensional one-hot coding; (2) For drugs lacking three-dimensional structures, the edge features were encoded using a 4-dimensional one-hot coding.

Multi-view representation of targets

Multi-view representation of targets also included two distinct types: target graphs and their line graphs. In PDBbind_2016 and PDBbind_2019, we initially used BioPython [44] to directly read the carbon α atom coordinates of amino acid residues from the provided structure files (.pdb) of targets, employing them as vertices of target graphs. Subsequently, we calculated the Euclidean distance between every pair of carbon α atoms. If the distance fell below the threshold of 8 Å, we established an edge between two vertices to form target graph. Finally, we employed the edges of target graph as vertices, and vice versa, to construct the line graph of target. The vertex features were represented via one-hot encoding, typically encompassing 20 different amino acid residue types constituting target. As for the edge features, we adopted the same approach as AGAT-PPIS [13], involving the Euclidean distance between carbon α atoms, and the cosine value of the angle formed between carbon α atoms at both ends of edge, relative to the reference point (carbon α atom of the first amino acid residue of target). For datasets lacking three-dimensional structures of targets, we referred to the UniprotID provided to retrieve the corresponding three-dimensional structures from

the Uniprot database [45]. For a small subset of targets for which three-dimensional structures could not be obtained from the Uniprot database, we employed AlphaFold2 [46] to predict their three-dimensional structures.

Architecture of model

In this study, we proposed a novel graph neural network architecture called MvGraphDTA based on multi-view representations of drugs and targets for predicting DTA (Fig. 4a). The network architecture (Algorithm 1) mainly comprised three key modules: multi-view feature extraction module, multi-view feature fusion module, and drug-target affinity prediction module.

Multi-view feature extraction module: For feature extraction of each view, we employed two consecutive graph convolutional layers, with the number of hidden units set to 128 in each layer. The vertex features of graph and line graph, obtained through the aggregation process of the preceding layer, were utilized as input for the subsequent layer (Fig. 4b). To enhance information aggregation among vertices in graphs of drug and target, as well as their line graphs, we incorporated edge features. Simultaneously, we applied Min-Max Normalization to normalize the edge features of the target graph for preventing significant discrepancies for the values across different dimensions. For the features update process (Formula 1, Fig. 4c) of vertex h_i in graph, we aggregated the features of the vertex itself, its neighboring vertices $h_j (j \in N(i))$, and the edges connected $e_k (k \in E(i))$ between vertices. This process remained consistent when updating vertex features in the line graph (Formula 2, Fig. 4c).

Table 11 Bond types of drugs

Three-dimensional structures of drugs	Bond types
Provided by dataset	Single bond, double bond, triple bond, aromatic bond (benzene ring bond), and unknown bond
Not provided by dataset	Single bond, double bond, triple bond, and aromatic bond (benzene ring bond)

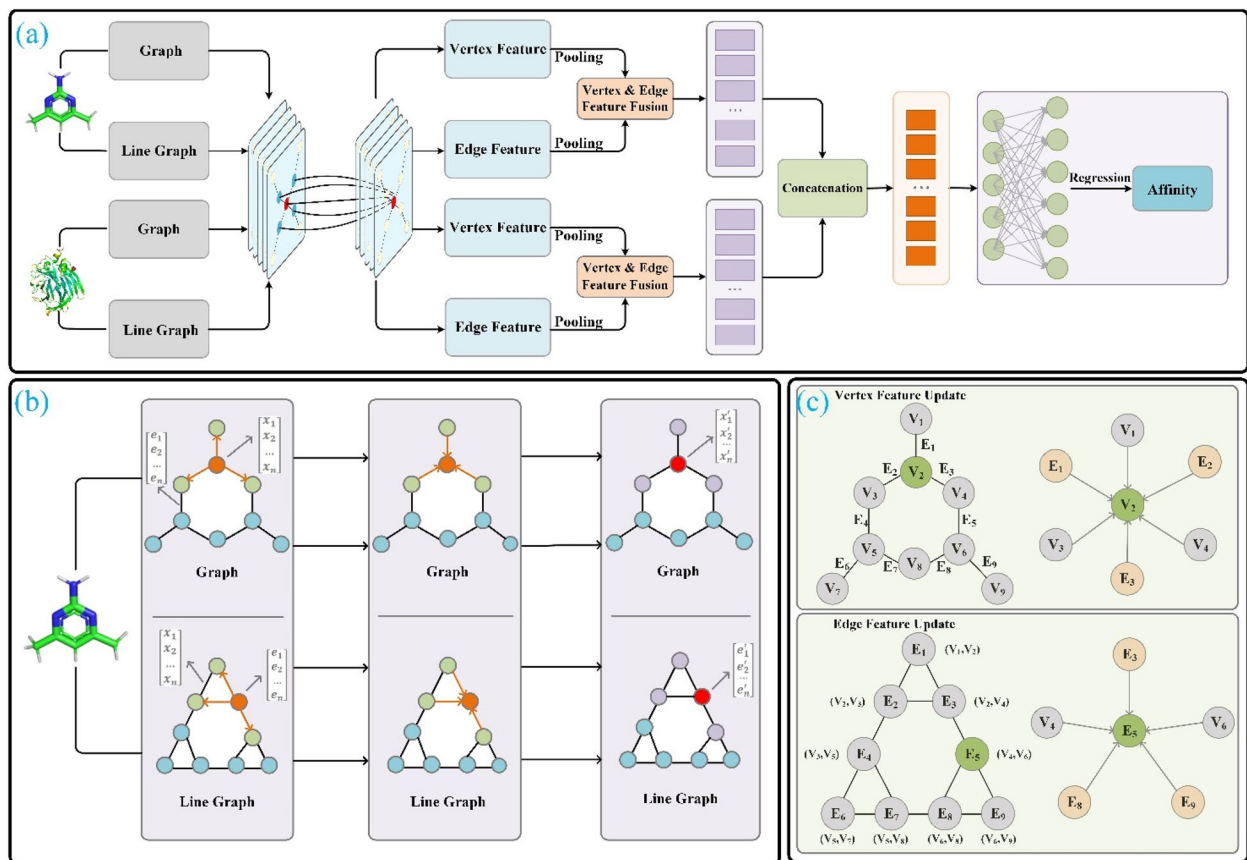


Fig. 4 Architecture of MvGraphDTA. **a** Firstly, graph convolutional networks were employed to extract the structural and interaction relationship features of drugs and targets from their respective graphs and line graphs. Subsequently, the extracted structural and interaction relationship features were fused separately. Finally, the fused features of drugs and targets were concatenated, and drug-target affinity was predicted through a fully connected network. **b** The vertex features in graph and line graph, obtained through the aggregation process of the preceding layer, were utilized as input for the subsequent layer. **c** We showed the updating process of vertex features in graph and line graph

$$h_i' = \sigma(W_1 h_i + W_2 \cdot AGGR(h_j(j \in N(i))) + W_3 \cdot AGGR(e_k(k \in E(i)))) \tag{1}$$

$$e_i' = \sigma(W_1 e_i + W_2 \cdot AGGR(e_j(j \in E(i))) + W_3 \cdot AGGR(h_k(k \in N(i)))) \tag{2}$$

Algorithm 1: procedure of MvGraphDTA

For epochs	/* Number of training iterations
Multi-view feature extraction of drugs by GNN	
Multi-view feature extraction of targets by GNN	
Multi-view feature fusion of drugs	
Multi-view feature fusion of targets	
Concatenation between the fusion features of drugs and the fusion features of targets	
Drug-target affinity prediction by fully connected network	

Here, W_1 , W_2 , and W_3 are weight matrices. $AGGR(\cdot)$ stands for the aggregation function based on summation. σ represents the LeakyReLU activation function. e_i , e_j ($j \in E(i)$), h_k ($k \in N(i)$) refer to the vertex, neighboring vertices, and edges in line graph, respectively. h_i' and e_i' denote the updated features of vertices in graph and line graph, respectively.

Multi-view feature fusion module: For drug, through the multi-view feature extraction module, we acquired the vertex features in graph and line graph. These features were then compressed using graph max pooling operation. The resultant compressed vertex features were added to derive the fusion features of drug. The procedure for obtaining the fusion features of target followed the same approach.

Drug-target affinity prediction module: The fused features of drug and target were concatenated and fed into a FC network consisting of two hidden layers and one output layer to predict DTA. The neurons in the two hidden layers were set to 256 and 128, respectively. LeakyReLU activation function was applied after the hidden layers, and the output layer contained 1 neuron.

Data augmentation

In publicly reported ColdDTA [47], the original drug graph served as a template. A portion of vertices in the graph were selectively removed based on specific rules to construct some new drug graphs. These new drugs were then paired with the original target, maintaining the affinity value between the drug and target unchanged. This process resulted in a substantial increase in the number of drug-target pairs, leading to more comprehensive model training and significantly improved the model's performance, as anticipated. Drawing inspiration

from this approach, we proposed a similar data augmentation strategy for the training dataset of MvGraphDTA. For each drug-target pair, we randomly selected a carbon α atom from the target graph as the starting vertex. Subsequently, this vertex and its neighboring vertices were removed until a certain proportion of vertices in the graph were removed. This process formed a new complete target subgraph, with the affinity value of new target with the original drug remaining unchanged (Fig. 5). In cases where multiple incomplete target subgraphs were generated during the formation of a new target subgraph, the new vertex was selected as starting point until a complete target subgraph was obtained. It was important to note that data augmentation was solely employed to expand the training data of the model. In other words, the newly generated drug-target pairs were incorporated into the dataset for model training alongside the original drug-target pairs. The test set of the model remained unchanged throughout this process.

Training details

MvGraphDTA was implemented under the PyTorch framework (<https://pytorch.org/>). We trained it using fivefold cross-validation on PDBbind_2016, with 100 epochs for each training iteration. Furthermore, we also trained it for 100 epochs on PDBbind_2019 and performed independent testing. The optimizer utilized was AdamW (Formula 3), with a learning rate set to 0.0001. For DTA prediction (regression), mean square error (MSE) was employed as loss function (Formula 4). Meanwhile, to evaluate the universality of MvGraphDTA for DTI prediction (classification), cross-entropy was used as loss function (Formula 5). Additionally, during the training phase, we implemented the cosine annealing algorithm (Formula 6) to dynamically adjust the learning rate.

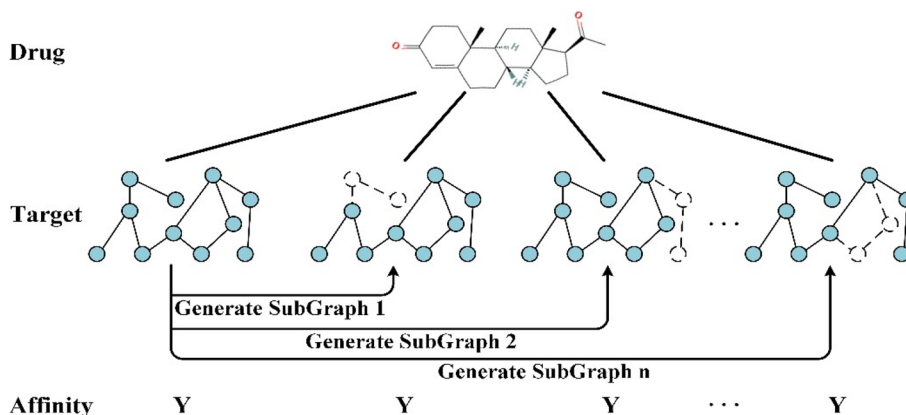


Fig. 5 Data augmentation process. Firstly, we randomly selected an amino acid residue from the target graph as the starting vertex. Subsequently, this vertex and its adjacent vertices were removed until a certain proportion of vertices in the graph were removed. This process formed a new complete target subgraph, with the affinity value of a new target with the original drug remaining unchanged

$$\theta_{t+1} = \theta_t - \gamma \left(\frac{\widehat{m}_t}{\sqrt{\widehat{v}_t + \varepsilon}} + \lambda \theta_t \right) \quad (3)$$

Here, θ_{t+1} denotes the updated parameter, where γ indicates the learning rate. ε is a constant added to ensure numerical stability, and λ represents the weight decay coefficient. Additionally, \widehat{m}_t and \widehat{v}_t represent the first and second-order moment estimates for bias correction, respectively.

$$Loss_{MSE} = \frac{\sum_{i=1}^n (y_{pred} - y)^2}{n} \quad (4)$$

Among them n , y , and y_{pred} represent the number of training samples, true value, and predictive value, respectively.

$$Loss_{CE} = - \sum_{i=1}^n y_{real}^{(i)} \cdot \ln y_{pred}^{(i)} + (1 - y_{real}^{(i)}) \cdot \ln(1 - y_{pred}^{(i)}) \quad (5)$$

In formula 5, n represents the total number of samples, while $y_{real}^{(i)}$ and $y_{pred}^{(i)}$ denote the predictive and true affinity values of the i -th drug-target pair, respectively.

$$\eta_t = \eta_{min} + \frac{1}{2} (\eta_{max} - \eta_{min}) (1 + \cos(\frac{T_{cur}}{T_{max}} \pi)) \quad (6)$$

Here, η_{min} and η_{max} represent the minimum and maximum values of learning rate, respectively. T_{cur} signifies the number of executions of cosine annealing algorithm, while T_{max} represents the number of iterations needed to complete half a cycle of periodic cosine function. η_t represents the value of learning rate after each execution of cosine annealing algorithm.

Evaluation metrics

In this study, the model's evaluation metrics for DTA prediction included mean absolute error (MAE), root mean square error (RMSE), Pearson correlation coefficients (PCCs), Consistency Index (CI), and R -square (R^2). For assessing the universality of model in DTI prediction, the evaluation metrics utilized were accuracy, area under the receiver operating characteristic curve (AUC), and area under the precision-recall curve (AUPRC).

Abbreviations

GCN	Graph convolutional network
FC	Fully connected
DTI	Drug-target interaction
DTA	Drug-target affinity
CNN	Convolutional Neural Network
RNN	Recurrent Neural Network
GNN	Graph Neural Network
3D-CNN	3D Convolutional Neural Network
MVM	Multi-view method

MSE	Mean square error
MAE	Mean absolute error
RMSE	Root mean square error
PCCs	Pearson correlation coefficients
CI	Consistency Index
R^2	R -Square
AUC	Area under the receiver operating characteristic curve
AUPRC	Area under the precision-recall curve

Acknowledgements

Not applicable.

Authors' contributions

All authors contributed to the study conception and design. X.Z.: Ideas, writing—original manuscript, writing—review and editing, writing—revised manuscript. K.-Y.Z.: Creation of models, data curation, implementation of the computer code and supporting algorithms. P.-Y.M.: Data analysis. And visualization. S.-J.L.: Data analysis. S.-Q.L.: Data analysis. M.-L.W.: Supervision. Y.L.: Development of methodology, formal analysis, modification guidance. All authors reviewed and approved the final manuscript.

Funding

This work was supported by the National Natural Sciences Foundation of China (No. 62366002), Yunnan Young and Middle-aged Academic and Technical Leaders Reserve Talent Project in China (No. 202405AC350023), and a grant (No. 2023KF005) from State Key Laboratory for Conservation and Utilization of Bio-Resources in Yunnan, Yunnan University.

Availability of data and materials

The datasets and code supporting the conclusions of this article are available in the Zenodo repository, [<https://doi.org/https://doi.org/10.5281/zenodo.13299715>]. All data generated or analyzed during this study are included in this published article, its supplementary information files, and publicly available repositories.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 23 May 2024 Accepted: 13 August 2024

Published online: 26 August 2024

References

1. Yu J-L, Dai Q-Q, Li G-B. Deep learning in target prediction and drug repositioning: recent advances and challenges. *Drug Discov Today*. 2022;27:1796–814.
2. Elbadawi M, Gaisford S, Basit AW. Advanced machine-learning techniques in drug discovery. *Drug Discov Today*. 2021;26:769–77.
3. Liao J, Chen H, Wei L, Wei L. GSAML-DTA: An interpretable drug-target binding affinity prediction model based on graph neural networks with self-attention mechanism and mutual information. *Comput Biol Med*. 2022;150:106145.
4. Begg AC. Cisplatin and radiation: Interaction probabilities and therapeutic possibilities. *Int J Radiat Oncol*. 1990;19:1183–9.
5. Dhakal A, McKay C, Tanner JJ, Cheng J. Artificial intelligence in the prediction of protein–ligand interactions: recent advances and future directions. *Brief Bioinform*. 2022;23:bbab476.
6. Zhang Y, Hu Y, Han N, Yang A, Liu X, Cai H. A survey of drug-target interaction and affinity prediction methods via graph neural networks. *Comput Biol Med*. 2023;163:107136.

7. Lauria A, Ippolito M, Fazzari M, Tutone M, Di Blasi F, Mingoia F, et al. IKK- β inhibitors: An analysis of drug–receptor interaction by using molecular docking and pharmacophore 3D-QSAR approaches. *J Mol Graph Model*. 2010;29:72–81.
8. Xu L, Ru X, Song R. Application of machine learning for drug-target interaction prediction. *Front Genet*. 2021;12:680117.
9. Lim S, Lu Y, Cho CY, Sung I, Kim J, Kim Y, et al. A review on compound-protein interaction prediction methods: Data, format, representation and model. *Comput Struct Biotechnol J*. 2021;19:1541–56.
10. Stebliankin V, Shirali A, Baral P, Chapagain P, Narasimhan G. PlsToN: Evaluating protein binding interfaces with transformer networks. Preprint. *Bioinformatics*; 2023.
11. Basile AO, Yahi A, Tatonetti NP. Artificial intelligence for drug toxicity and safety. *Trends Pharmacol Sci*. 2019;40:624–35.
12. Huang A, Xie X, Yao X, Liu H, Wang X, Peng S. HF-DDI: Predicting Drug-Drug Interaction events based on multimodal hybrid fusion. *J Comput Biol*. 2023;30:961 cmb.2023.0068.
13. Zhou Y, Jiang Y, Yang Y. AGAT-PPIS: a novel protein–protein interaction site predictor based on augmented graph attention network with initial residual and identity mapping. *Brief Bioinform*. 2023;24:bbad122.
14. Öztürk H, Özgür A, Ozkirimli E. DeepDTA: deep drug–target binding affinity prediction. *Bioinformatics*. 2018;34:821–9.
15. Bian J, Zhang X, Zhang X, Xu D, Wang G. MCANet: shared-weight-based MultiheadCrossAttention network for drug–target interaction prediction. *Brief Bioinform*. 2023;24:bbad082.
16. 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:e1007129.
17. Chen L, Tan X, Wang D, Zhong F, Liu X, Yang T, et al. TransformerCPI: improving compound–protein interaction prediction by sequence-based deep learning with self-attention mechanism and label reversal experiments. *Bioinformatics*. 2020;36:4406–14.
18. Zhao Q, Zhao H, Zheng K, Wang J. HyperAttentionDTI: improving drug–protein interaction prediction by sequence-based deep learning with attention mechanism. *Bioinformatics*. 2022;38:655–62.
19. Huang K, Xiao C, Glass LM, Sun J. MolTrans: Molecular Interaction Transformer for drug–target interaction prediction. *Bioinformatics*. 2021;37:830–6.
20. Wu Y, Gao M, Zeng M, Zhang J, Li M. BridgeDPI: a novel graph neural network for predicting drug–protein interactions. *Bioinformatics*. 2022;38:2571–8.
21. Huang L, Lin J, Liu R, Zheng Z, Meng L, Chen X, et al. CoaDTI: multi-modal co-attention based framework for drug–target interaction annotation. *Brief Bioinform*. 2022;23:bbac446.
22. Zhao B-W, Su X-R, Hu P-W, Huang Y-A, You Z-H, Hu L. iGRDLTI: an improved graph representation learning method for predicting drug–target interactions over heterogeneous biological information network. *Bioinformatics*. 2023;39:btad451.
23. Wang K, Zhou R, Li Y, Li M. DeepDTAF: a deep learning method to predict protein–ligand binding affinity. *Brief Bioinform*. 2021;22:bbab072.
24. Stepniowska-Dziubinska MM, Zielenkiewicz P, Siedlecki P. Development and evaluation of a deep learning model for protein–ligand binding affinity prediction. *Bioinformatics*. 2018;34:3666–74.
25. Seo S, Choi J, Park S, Ahn J. Binding affinity prediction for protein–ligand complex using deep attention mechanism based on intermolecular interactions. *BMC Bioinformatics*. 2021;22:542.
26. Jones D, Kim H, Zhang X, Zemla A, Stevenson G, Bennett WFD, et al. Improved protein–ligand binding affinity prediction with structure-based deep fusion inference. *J Chem Inf Model*. 2021;61:1583–92.
27. Zheng L, Fan J, Mu Y. OnionNet: a multiple-layer intermolecular-contact-based convolutional neural network for protein–ligand binding affinity prediction. *ACS Omega*. 2019;4:15956–65.
28. Wang DD, Chan M-T. Protein–ligand binding affinity prediction based on profiles of intermolecular contacts. *Comput Struct Biotechnol J*. 2022;20:1088–96.
29. Zhang Y, Zhou G, Wei Z, Xu H. Predicting protein–ligand binding affinity via joint global-local interaction modeling. 2022.
30. Jin Z, Wu T, Chen T, Pan D, Wang X, Xie J, et al. CAPLA: improved prediction of protein–ligand binding affinity by a deep learning approach based on a cross-attention mechanism. *Bioinformatics*. 2023;39:btad049.
31. Wang K, Zhou R, Tang J, Li M. GraphscoreDTA: optimized graph neural network for protein–ligand binding affinity prediction. *Bioinformatics*. 2023;39:btad340.
32. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. 2015.
33. Zaremba W, Sutskever I, Vinyals O. Recurrent neural network regularization. 2015.
34. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. 2017.
35. Weininger D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J Chem Inf Model*. 1988;28:31–6.
36. Xu K, Hu W, Leskovec J, Jegelka S. How powerful are graph neural networks? *Int Conf Learn Represent*. 2019.
37. Hu L, Zhang M, Hu P, Zhang J, Niu C, Lu X, et al. Dual-channel hypergraph convolutional network for predicting herb–disease associations. *Brief Bioinform*. 2024;25:bbae067.
38. Zhao B-W, He Y-Z, Su X-R, Yang Y, Li G-D, Huang Y-A, et al. Motif-aware miRNA-disease association prediction via hierarchical attention network. *IEEE J Biomed Health Inform*. 2024;28:4281–94.
39. Wang R, Fang X, Lu Y, Wang S. The PDBbind Database: collection of binding affinities for protein–ligand complexes with known three-dimensional structures. *J Med Chem*. 2004;47:2977–80.
40. Su M, Yang Q, Du Y, Feng G, Liu Z, Li Y, et al. Comparative Assessment of Scoring Functions: The CASF-2016 update. *J Chem Inf Model*. 2019;59:895–913.
41. Li Y, Su M, Liu Z, Li J, Liu J, Han L, et al. Assessing protein–ligand interaction scoring functions with the CASF-2013 benchmark. *Nat Protoc*. 2018;13:666–80.
42. Wang DD, Xie H, Yan H. Proteo-chemometrics interaction fingerprints of protein–ligand complexes predict binding affinity. *Bioinformatics*. 2021;37:2570–9.
43. Landrum G. RDKit: A software suite for cheminformatics, computational chemistry, and predictive modeling. 2013.
44. Cock PJA, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, et al. Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics*. 2009;25:1422–3.
45. The UniProt Consortium, Bateman A, Martin M-J, Orchard S, Magrane M, Agivetova R, et al. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res*. 2021;49:D480–9.
46. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596:583–9.
47. Fang K, Zhang Y, Du S, He J. ColdDTA: Utilizing data augmentation and attention-based feature fusion for drug–target binding affinity prediction. *Comput Biol Med*. 2023;164:107372.

Publisher's Note

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