scientific reports

Check for updates

OPEN

Multi-layer graph attention neural networks for accurate drug-target interaction mapping

Qianwen Lu[2], Zhiheng Zhou[3,4] & Qi Wang[1]✉

In the crucial process of drug discovery and repurposing, precise prediction of drug-target interactions (DTIs) is paramount. This study introduces a novel DTI prediction approach—Multi-Layer Graph Attention Neural Network (MLGANN), through a groundbreaking computational framework that effectively harnesses multi-source information to enhance prediction accuracy. MLGANN not only strides forward in constructing a multi-layer DTI network by capturing both direct interactions between drugs and targets as well as their multi-level information but also amalgamates Graph Convolutional Networks (GCN) with a self-attention mechanism to comprehensively integrate diverse data sources. This method exhibited significant performance surpassing existing approaches in comparative experiments, underscoring its immense potential in elevating the efficiency and accuracy of DTI predictions. More importantly, this study accentuates the significance of considering multi-source data information and network heterogeneity in the drug discovery process, offering new perspectives and tools for future pharmaceutical research.

Drug-Target Interaction (DTI) prediction is a computational task aimed at identifying potential interactions between chemical compounds (drugs) and biological molecules (targets, typically proteins). Prediction of DTI is a crucial component in modern drug research and development. Accurate DTI prediction can provide valuable clues for drug design and expedite drug repurposing, with approximately 75% of drugs being eligible for repurposing[1]. Traditional in vitro experimental tests can validate DTI, but this approach is both expensive and time-consuming. In recent years, computer-assisted methods for DTI prediction have garnered widespread attention, including matrix factorization[2], kernel-based methods[3], graph embedding techniques[4], and more. The introduction of these methods narrows the search space, reduces the workload of in vitro experiments, and accelerates the drug development process.

Although the importance of DTI prediction is widely recognized, most current research in this area primarily relies on the chemical structure of drugs and the protein sequences of targets for DTI prediction[5–7]. This approach neglects other crucial multi-source information about drugs and targets, such as the physicochemical properties of drugs and the relationships between targets and diseases[8,9], which are equally important for DTI prediction. With deep learning demonstrating excellent performance across various domains, some researchers have also applied it to DTI prediction and have achieved promising results[10–12]. However, methods based on neural networks or machine learning, when integrating information about drugs and targets, can only utilize the information from individual drugs or targets[3,10]. The result of this approach is a loss of the unique information regarding the similarity between drugs and targets. In contrast, network-based methods can model the similarity between drugs and targets and various features as a DTI network. This network is used to represent multidimensional information about drugs and targets and is applied to DTI prediction.

Due to the superior performance of graph neural network (GNN) methods in network data analysis, some researchers have proposed GNN-based models for DTI prediction on heterogeneous DTI networks[13–15]. However, most of the work separately considers drug and target networks, neglecting the information regarding their interactions. Additionally, some models treat heterogeneous DTI networks as homogeneous graphs for neighbor aggregation[13], which results in the inability to capture the heterogeneous information within the DTI network, leading to suboptimal network node representations.

To address the above issues, in this paper, we propose a multi-layer graph attention neural network (MLGANN) to capture multi-source information on drugs and targets for DTI prediction. Specifically, we leverage collected multi-source information about drugs and targets to construct a multi-layer DTI network, which is then utilized

[1]College of Science, China Agricultural University, Beijing 100083, China. [2]SDU-ANU Joint Science College, Shandong University, Weihai 264209, Shandong, China. [3]Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China. [4]School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100190, China. ✉email: wangqi_math@cau.edu.cn

for DTI prediction. The multi-layer DTI network not only encompasses interaction information between drugs and targets but also captures multi-level information about drugs and targets. Furthermore, we design a graph neural network-based model that employs Graph Convolutional Networks (GCN) to capture multi-source information about drugs and targets within the multi-layer DTI network. Subsequently, a self-attention mechanism layer is employed to integrate the multi-source information and obtain representations for drugs and targets. Finally, through comparisons with multiple models on DTI datasets, the results demonstrate that the proposed approach outperforms state-of-the-art methods.

In sum, the major contributions of our proposed method can be summarized as follows:

- We construct a multi-layer DTI network for DTI prediction, which not only encapsulates multi-level information of drugs and targets but also captures the interactions between these levels. This comprehensive integration enhances the accuracy of DTI predictions.
- We design a multi-layer graph attention neural network (MLGANN) that effectively captures multi-source drug and target information within the multi-layer DTI network, as well as the interactions between them, further improving DTI prediction outcomes.
- Through comparative analysis with multiple models on the DTI dataset, our experimental results demonstrate the effectiveness and superiority of the proposed method in addressing the DTI prediction problem.

Related work

Drug-target interaction (DTI) prediction is a critical task in drug discovery and development, where the goal is to identify potential interactions between drugs and target proteins. Over the years, various computational approaches have been developed to address this problem, ranging from similarity-based methods to more advanced graph-based approaches.

Traditional methods for DTI prediction often rely on similarity measures, where similar drugs are assumed to interact with similar targets. These methods typically compute drug-drug and target-target similarity matrices, which are then used in conjunction with machine learning models to predict interactions[16]. However, these approaches have limitations, such as their inability to integrate multi-source information about drugs and targets. For example, methods like DTINet and deepDTnet fall into this category but struggle with capturing the full complexity of drug-target relationships.

Recent advancements have seen the application of knowledge graphs to DTI prediction[17], where biological knowledge is structured as a graph and used to infer new interactions. Methods such as ComplEx[18] and KGE_NFM[19] have been developed to leverage knowledge graph embeddings for DTI prediction. While these methods provide a way to incorporate diverse biological information, they often require significant domain-specific knowledge to construct the knowledge graph, which can limit their applicability.

Graph neural networks (GNNs) have gained popularity for their ability to model complex relationships in network data. Several GNN-based methods have been proposed for DTI prediction, including IMCHGAN[15], SGCL-DTI[14], and MHGNN[20]. These methods treat drug-target pairs as nodes in a heterogeneous network and use GNNs to learn embeddings that capture the intricate relationships between drugs and targets. For instance, SGCL-DTI employs a supervised graph co-contrastive learning approach, while MHGNN uses heterogeneous graph attention networks to enhance prediction accuracy.

In addition to the methods mentioned above, several recent works have further advanced the field. SGCLDGA[21] and NGCN[22] are notable examples. SGCLDGA enhances DTI prediction by incorporating graph neural networks and contrastive learning that focus on graph structures. NGCN, on the other hand, leverages integrated heterogeneous network to extract relevant biological properties and association information while maintaining the topology information.

Furthermore, methods like TripletMultiDTI[23], DeepTraSynergy[24], and HGTDR[25] represent the cutting edge of DTI prediction. TripletMultiDTI uses a triplet loss to improve the distinction between interacting and non-interacting pairs. DeepTraSynergy is designed to predict drug synergy, which is closely related to DTI, using deep learning techniques. HGTDR employs heterogeneous graph transformers to advance drug repurposing efforts by effectively modeling the multi-relational nature of drug-target interactions.

While existing methods have made significant strides in DTI prediction, they often face challenges in integrating multi-source information or dealing with the heterogeneity of drug-target interactions. To address these challenges, our proposed Multi-Layer Graph Attention Neural Network (MLGANN) not only captures multi-level interaction information between drugs and targets but also integrates diverse data sources through a novel multi-layer attention mechanism. This approach allows for a more comprehensive and accurate prediction of DTIs, as demonstrated in our experimental results.

Methods

Multi-layer DTI network

Give a drug set $D = \{d_i\}_{i=1}^{n_d}$ and a target set $T = \{t_j\}_{j=1}^{n_t}$, the similarity between drugs (targets) can be assessed in various perspectives, which are represented by a set of matrices $\{A^{D,k}\}_{k=1}^{m_d}$, $(\{A^{T,l}\}_{l=1}^{m_t})$,where $A^{D,k} \in \mathbb{R}^{n_d \times n_d}$, $(A^{T,l} \in \mathbb{R}^{n_t \times n_t})$ and $m_d(m_t)$ is the number of similarity types for drugs (targets). Let the binary matrix $A^Y \in \{0,1\}^{n_d \times n_t}$ indicate the interactions between drugs in D and targets in T, where $A_{ij}^Y = 1$ denotes that d_i and t_j interact with each other, and $A_{ij}^Y = 0$ otherwise. A multi-layer DTI network $G^M = (V^M, E^M)$ as shown in Fig. 1 for D and T consists of $\{A^{D,k}\}_{k=1}^{m_d}$, $\{A^{T,l}\}_{l=1}^{m_t}$ and A^Y, with its adjacency matrices represented as follows:

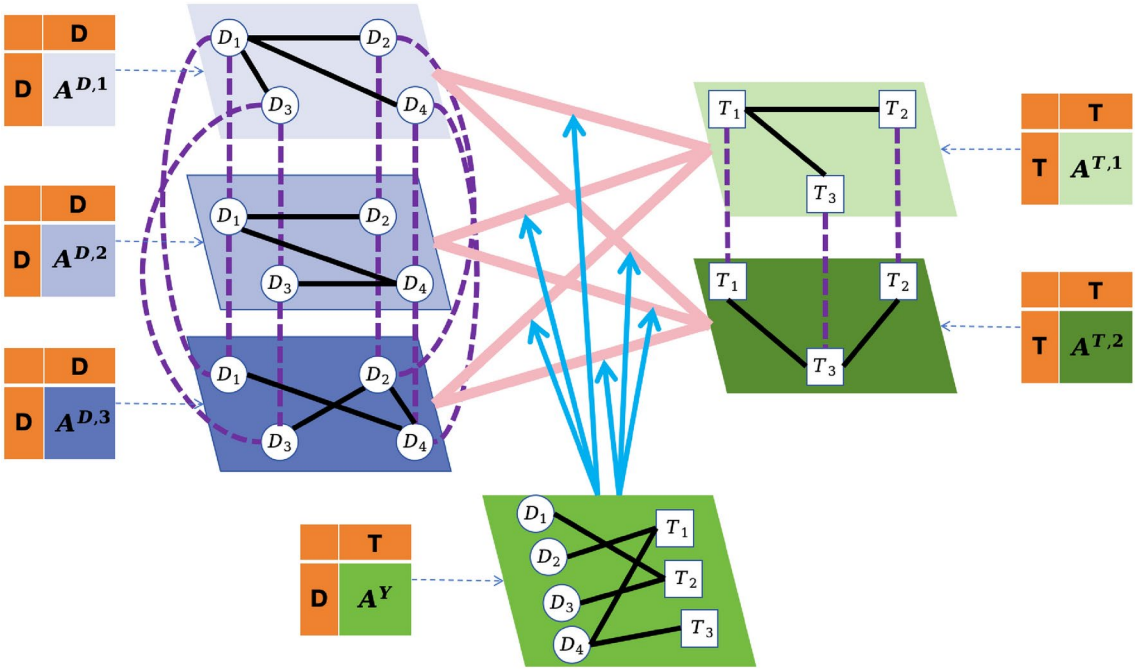


Fig. 1. The figure illustrates a multiplex layer drug-target interaction (DTI) network, which integrates multi-level information from drugs (D) and targets (T) across several layers. On the left, different layers of drug associations are shown (labeled as $A^{D,1}$, $A^{D,2}$, $A^{D,3}$), representing various relationships among drugs D_1, D_2, D_3 and D_4. On the right, target associations are depicted in similar layers ($A^{T,1}$ and $A^{T,2}$) with targets T_1, T_2 and T_3. The central part of the diagram displays the interaction between drugs and targets, where a multi-layer network structure is used to capture the complex interplay between different layers of information. This multi-layered approach enables for more comprehensive DTI prediction by considering both intra-layer and inter-layer interactions.

$$A^M = \begin{pmatrix} A^{D,1} & I & I & \cdots & A^Y & A^Y \\ I & A^{D,2} & I & \cdots & A^Y & A^Y \\ I & I & A^{D,3} & \cdots & A^Y & A^Y \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ A^Y & A^Y & A^Y & \cdots & A^{T,m_t-1} & I \\ A^Y & A^Y & A^Y & \cdots & I & A^{T,m_t} \end{pmatrix}. \tag{1}$$

where $||V^M|| = N = (n_d \times m_d + n_t \times m_t)$.

Multi-layer attention graph neural network

We propose a model called the Multi-Layer Graph Attention Neural Network (MLGANN) for DTI prediction. In the multi-layer network of DTI, apart from the interaction between drugs and targets, there is also interaction information among various properties within drugs and targets themselves. Therefore, we utilize the designed MLGANN to capture both the interaction information between drugs and targets and the multi-source information within drugs and targets.

Multi-layer neighbor aggregation

Let $X \in \mathbb{R}^{N \times f}$ represent the initial features of nodes in the multi-layer DTI network, where f denotes the dimension of the embedding space. We apply graph neural networks to learn embeddings for drugs and targets on the multi-layer DTI network. Specifically, in our model, we employ Graph Convolutional Networks (GCN), as they are both simple and effective. These embeddings can be refined by applying P layers of GCN across the entire multi-layer DTI network:

$$X^{(p)} = \sigma\left(\hat{D}^{-\frac{1}{2}}\hat{A}\hat{D}^{-\frac{1}{2}}X^{(p-1)}W^{(p)}\right), \quad p = 1, 2, \ldots, P, \tag{2}$$

where $X^{(0)} = X$, $\hat{A} = A^M + I^M$, and A^M is the adjacency matrix of the multi-layer DTI network, I^M is an identity matrix of the same size as A^M, $\hat{D}$ is a diagonal matrix with $\hat{D}_{ij} = \sum_{j=1} \hat{A}_{ij}$, $W^{(p)} \in \mathbb{R}^{f \times f}$ is a trainable weight matrix, σ is a nonlinear activation function ReLU.

For a node $v \in G^M$ (representing either a drug or a target), Eq. (2) updates the embedding of that node as follows:

$$x_v^{(p)} = \sigma\left(W^{(p)} \sum_u \frac{1}{\alpha_{vu}} x_u^{(p-1)}\right), \quad u \in \{v \cup N_v \cup C_v\} \tag{3}$$

where α_{vu} is the normalized weight, N_v is the set of neighbors of node *v* in layer of G^M, and C_v is the set of nodes that correspond to the same drug/target as node *v*. Therefore, MLGANN not only aggregates the neighbors of node *v* in layer of G^M (similar to what GCN does) but also embeds nodes corresponding to the same drug/target in different layers of G^M. This allows information to be transmitted across different layers of G^M. By leveraging information from different layers of G^M, MLGANN can learn better representations for each node, especially for nodes with limited interactions in a particular layer of G^M. This is the main distinction between MLGANN and existing other network embedding methods.

Multi-layer attention pooling
We concatenate all representations learned by the *P*-layer GCN to obtain the final node embedding:

$$\begin{aligned} z_i^{D,k} &= \left[z_i^{D,k(0)}, z_i^{D,k(1)}, \cdots, z_i^{D,k(P)}\right] \\ z_j^{T,l} &= \left[z_j^{T,l(0)}, z_j^{T,l(1)}, \cdots, z_j^{T,l(P)}\right], \end{aligned} \tag{4}$$

where $z_i^{D,k}$ denotes final embedding of *i*th drug in *k* layer of G^M and $z_i^{D,k(p)}$ denotes G^M's *k*th layer embedding of *i*th drug in GCN *p*th layer, $z_j^{T,l}$ represents final embedding of *j*th target in *l*th layer of G^M, $z_j^{T,l(p)}$ denotes G^M's *l*th layer embedding of *j*th drug in GCN *p*th layer.

To obtain the final representations of drugs and targets, we have designed a self-attention mechanism to aggregate the representation vectors of drugs and targets across different layers for DTI prediction in the G^M graph. The computer process is as follows:

$$\begin{aligned} e_i^{D,k} &= q^D \cdot LeakyReLU\left(W^D z_i^{D,k}\right), \quad e_j^{T,l} = q^T \cdot LeakyReLU\left(W^T z_j^{T,l}\right) \\ \alpha_i^k &= \frac{e_i^{D,k}}{\sum_{k'=1}^{m_d} e_i^{D,k'}}, \quad z_i^D = \sum_{k=1}^{m_d} \alpha_i^k z_i^{D,k}, \quad \beta_j^l = \frac{e_i^{T,l}}{\sum_{l'=1}^{m_t} e_j^{T,l'}}, \quad z_j^T = \sum_{l=1}^{m_t} \beta_j^l z_j^{T,l}, \end{aligned} \tag{5}$$

where $z_i^D \in \mathbb{R}^{f'}$ and $z_i^D \in \mathbb{R}^{f'}$ are the final representations of drugs and targets, $W^D \in \mathbb{R}^{f' \times f'}$ and $W^T \in \mathbb{R}^{f' \times f'}$ are trainable parameter matrices, $q^D \in \mathbb{R}^{f'}$ and $q^T \in \mathbb{R}^{f'}$ are trainable vectors.

DTI prediction

Let G^Y be the DTI network derived from the adjacency matrix A^Y. For an edge $d_i t_j$ in G^Y, where z_i^D and z_j^T are final representation vectors of drug d_i and target t_j, respectively. we sample a non-existing edge $d_u t_v$ in G^Y, where z_u^D and z_v^T are final representation vectors of drug d_u and target t_v, respectively. We consider DTP $d_i t_j$ as a positive sample and $d_u t_v$ as a negative sample. Therefore, we design the loss function based on cross-entropy as follows:

$$\mathscr{L} = -\log\left(\sigma\left(< z_i^D, z_j^T >\right)\right) - \log\left(\sigma\left(- < z_u^D, z_v^T >\right)\right) \tag{6}$$

where σ is a nonlinear activation function Sigmoid, $< \cdot, \cdot >$ is the inner product in Euclidean space.

Experiments

Experimental setup

Datasets We collected data on DTIs from DrugBank (Version5.1.8), as well as relevant data on drug chemical structural similarity, drug-side effect relationships, drug-disease relationships, target interaction relationships, and target-disease relationships. All of this data is available on DrugBank. By taking the intersection of these datasets, we obtained a total of 201 drugs, 252 targets, and 907 interactions between them. We used 607 of these interactions for training and 300 interactions for testing. The statistical information for these datasets is summarized in Table 1, where D denotes drug, T denotes target, D-D-1 denotes drug chemical structural similarity, D-D-2 denotes similarity based on drug-side effect relationships, D-D-3 denotes similarity based on drug-disease relationships, T-T-1 denotes target interaction relationships, T-T-2 denotes similarity based on target-disease relationships.

Baselines The baselines can be categorized into three categories: (1) Similarity-based methods, mainly including DTINet[12], deepDTnet[26] and NEDTP[4]; (2) KG-based models which use TransE[27], ComplEx[18] and KGE_NFM[19]; (3) Graph neural network (GNN)-based methods, such as DTI-MGNN[8], SGCL-DTI[14], MHGNN[20] and IMCHGAN[15].

D	T	DTI	D-D-1	D-D-2	D-D-3	T-T-1	T-T-2
201	252	907	5142	1263	1598	12123	23545

Table 1. Statistics for the drug target interaction dataset.

Models		AUC	Precision	Recall	F1 score	AUPR
Similarity-based	DTNet	0.6811±0.0050	0.6503±0.0042	0.5837±0.0063	0.6287±0.0055	0.6924±0.0037
	deepDTNet	0.6704±0.0078	0.6619±0.0099	0.5736±0.0079	0.5878±0.0108	0.6544±0.0049
	NEDTP	*0.8027± 0.0041*	*0.7662± 0.0072*	0.7018±0.0037	0.7240±0.0061	0.8191±0.0031
KG-based	TransE	0.6245±0.0045	0.5535±0.0044	0.4980±0.0068	0.5122±0.0073	0.6674±0.0055
	KGE-NFM	0.7584±0.0033	0.6973±0.0030	0.6901±0.0046	0.6784±0.0065	0.7688±0.0041
	ComplEx	0.7061±0.0083	0.6229±0.0057	0.6417±0.0032	0.6527±0.0044	0.7477±0.0067
GNN-based	DTI-MGNN	0.7408±0.0069	0.6894±0.0057	0.6427±0.0073	0.6339±0.0045	0.7540±0.0058
	SGCL-DTI	0.6567±0.0167	0.6639±0.0112	0.5422±0.0089	0.5847±0.0095	0.7111±0.0098
	IMCHGAN	0.6864±0.0071	0.6675±0.0074	0.6149±0.0056	0.6482±0.0044	0.7001±0.0063
	MHGNN	0.7980±0.0056	0.7531±0.0035	*0.7618± 0.0066*	*0.7252± 0.0052*	*0.8263± 0.0036*
	MLGANN	**0.8699±0.0023**	**0.8122±0.0037**	**0.7989±0.0041**	**0.8018±0.0035**	**0.8552±0.0026**
	P-value	***$P<0.001$	***$P<0.001$	**$P<0.01$	***$P<0.001$	**$P<0.01$

Table 2. Comparison of our model with baseline on dataset DTI. The highest and the second highest results over each measurement are in bold and italics, respectively. Best model indicate significant improvement at *P*-value level when compared to the second highest performance, using one-sided paired t-test. **, *** represent significant levels of 1% and 0.1% respectively.

Parameter Configuration The hidden node embedding dimension of our models is set to 256 and the output embedding dimension is set to 64. We stack two convolutional layers. For the model optimization, our model will be trained 200 epochs with a learning rate of 0.01 and a batch size of 256. For HAN and MAGNN, the dimension of the attention vector in the feature fusion seted to 128. Other parameters are set to the same as these in MLGANN. For fair comparison, all these baselines adopt the same training set and test set as MLGANN. All models use Adam optimizer. In the experiments, all models are conducted with DGL platform and PyTorch framework.

Evaluation Metrics We conduct the link prediction task to validate the performance of all models. For training and testing stage, positive sample consists of a DTI, and the corresponding negative samples consist of an unknown DTI. We use the precision score, the recall score, the F1 score, the area under the curve (AUC) and the area under the precision-recall curve (AUPR) as evaluation metrics.

Overall performance

The experimental results for DTI prediction are shown in Table 2, where all results are the averaged across 5 repeated runs. As we can see, MLGANN significantly outperforms all baseline methods. It shows improvements of 0.0672, 0.0460, 0.0371, and 0.0766 over metrics of AUC, the precision score, the recall score, and the F1 score, demonstrating the superiority of our proposed model. Compared to similarity-based methods, MLGANN can harness multi-source data for drugs and targets, capturing various feature information of drugs and targets in the multi-layer DTI network for prediction. The end-to-end modeling approach also proves advantageous for DTI prediction. However, similarity-based methods rely on strong assumptions in modeling, leading to suboptimal final prediction results. Furthermore, two-stage models separate representation learning from the prediction

Models	AUC	Precision	Recall	F1 score	AUPR
w/o D-D-1	0.7832±0.0037	0.7478±0.0062	0.7118±0.0043	0.7281±0.0047	0.8069±0.0034
w/o D-D-2	0.8393±0.0082	0.7973±0.0067	0.7525±0.0056	0.7490±0.0077	0.8413±0.0050
w/o D-D-3	0.8078±0.0057	0.7775±0.0033	0.7402±0.0048	0.7415±0.0072	0.8098±0.0066
w/o T-T-1	0.8114±0.0058	0.7726±0.0044	0.7567±0.0039	0.7385±0.0051	0.8284±0.0053
w/o T-T-2	0.8235±0.0077	0.8004±0.0031	0.7531±0.0062	0.7306±0.0068	0.8316±0.0047
Ours	**0.8699±0.0023**	**0.8122±0.0037**	**0.7989±0.0041**	**0.8018±0.0035**	**0.8552±0.0039**
P-value	**$P<0.01$	**$P<0.01$	**$P<0.01$	***$P<0.001$	**$P<0.01$

Table 3. Results of ablation study (where w/o = without). Best model indicate significant improvement at *P*-value level when compared to the second highest performance, using one-sided paired t-test. The best results are marked in bold (the higher the better). **, *** represent significant levels of 1% and 0.1% respectively.

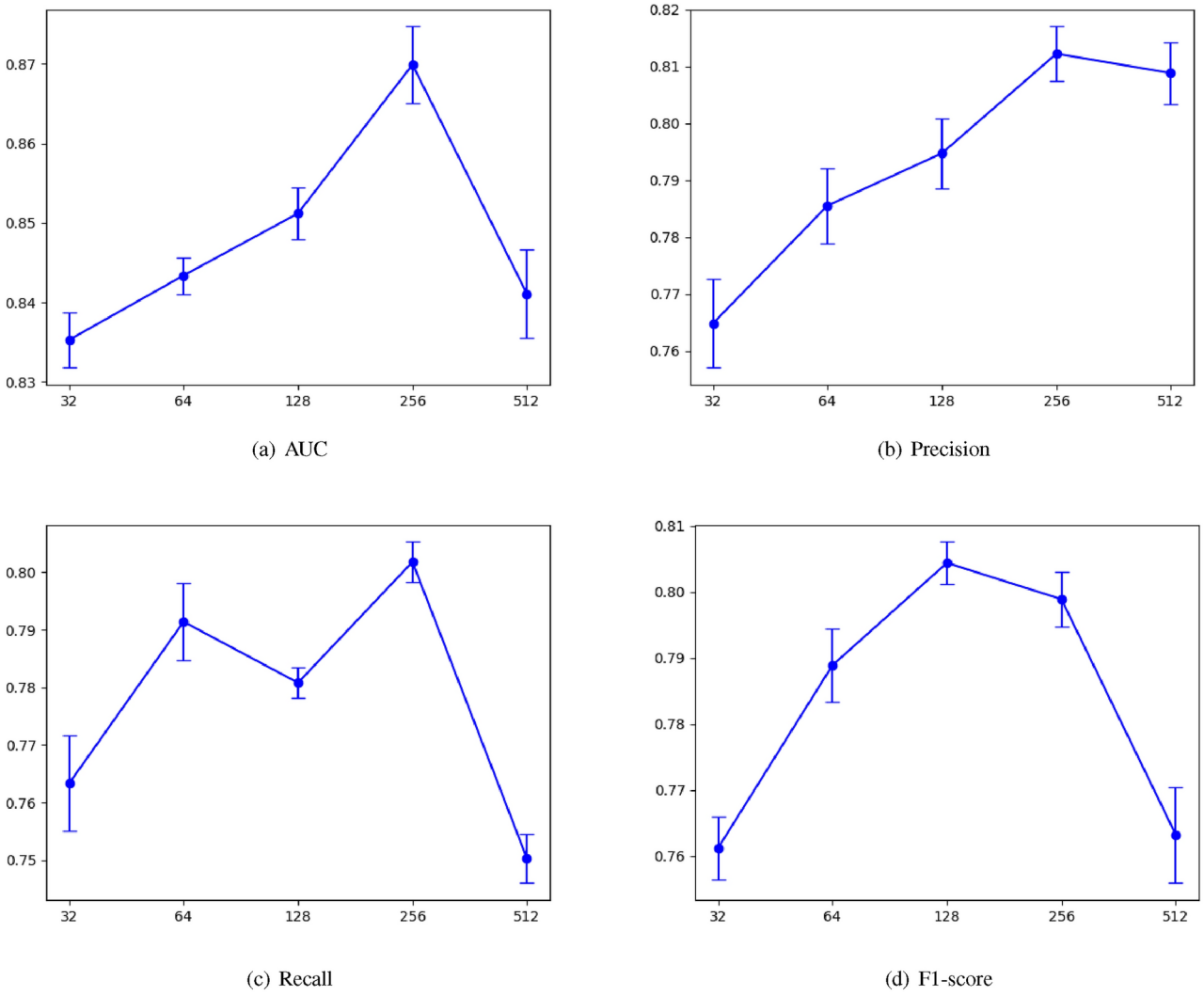


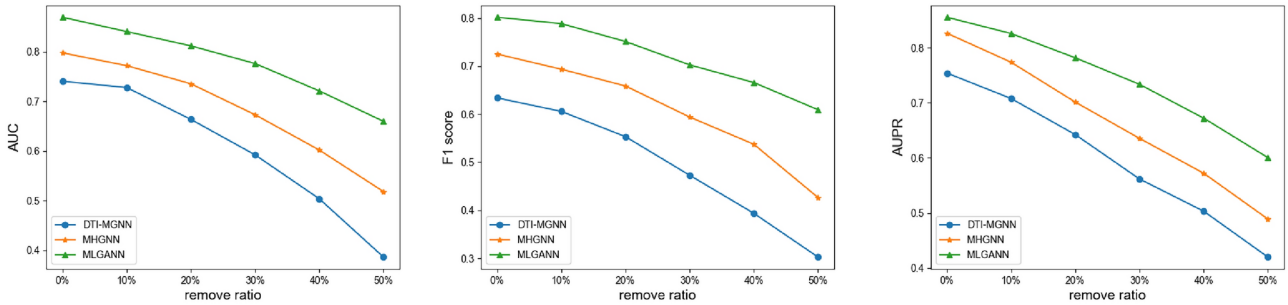
Fig. 2. Impacts of different hidden feature dimensions.

task, resulting in learned representation vectors that contradict DTI prediction outcomes. Compared to KG-based methods, MLGANN can learn better representations of drugs and targets.

Models based on GNNs approach DTI prediction by modeling the multi-source data of drugs and targets as a heterogeneous network and using metapaths and GCN for the task. SGCL-DTI constructs two heterogeneous networks, a topological network and a semantic network, with DTPs as nodes. The edges between nodes in the network are determined by the similarity of DTPs, as DTPs that share common drugs or targets have higher similarity than those without shared drugs or targets. Therefore, the topological and semantic networks exhibit significant overlap. Additionally, SGCL-DTI has access to labels of DTPs during training. DTI-MGNN constructs a DTP graph with DTPs as nodes, separating the representation learning of drugs and targets from the DTI prediction task. By comparing MLGANN with MHGNN, we can conclude that directly applying heterogeneous graph embedding methods to DTI prediction tasks does not yield optimal DTI prediction results.

Ablation study

To clearly demonstrate the effectiveness of the multi-layer graph we constructed, we conducted ablation experiments by progressively removing each layer in the multi-layer DTI network. The results are presented in Table 3. The experimental findings indicate that the results obtained with the multi-layer DTI network, utilizing 3 layers for drugs and 2 layers for targets, outperform the results obtained by removing any single layer within the multi-layer DTI network. Furthermore, the removal of D-D-1, D-D-3, and T-T-1 has a significant impact on the experimental results, suggesting that information related to drug chemical structural similarity, similarity based on drug-disease relationship, and target interaction relationship are more important than similarity based on drug-disease relationship and target-disease relationships.



(a) AUC (b) F1 score (c) AUPR

Fig. 3. The experiment result of robust analysis.

Parameter analysis

To analyze the performance of DTI prediction, we altered the dimension of the hidden layer from 32 to 512. The results for different metrics are shown in Fig. 2. From Fig. 2, we observe that the values of AUC, precision, and F1 score initially increase and then decrease with the increment of feature dimensions, with AUC and precision peaking at a dimension of 256, and F1 score peaking at a dimension of 128. The recall score reaches its peak result at the dimension of 256. This indicates that too small a dimension of the hidden layer cannot capture the characteristic information of the data, while too large a dimension causes the model to be prone to overfitting, affecting the model's generalization performance.

Robust analysis

To analyze the robustness of model, we randomly remove 10–50% train data. The results for different metrics are shown in Fig. 3. From Fig. 3, we observe that the values of AUC, F1 score, and AUPR decrease with the increment of remove ratio. However, among the three models, MLGANN not only maintains the highest performance in all metrics, but also has the smallest decline. This shows that MLGANN has good robustness in DTI prediction tasks.

Conclusion

In this study, we have successfully developed the Multi-Layer Graph Attention Neural Network (MLGANN) aimed at elucidating the intricate structures and nuanced semantics inherent in multi-layer Drug-Target Interaction (DTI) networks. MLGANN leverages GCN to capture multi-source information about drugs and targets, and further employs a self-attention mechanism to obtain representations for drugs and targets. We conducted model training and validation using data obtained from DrugBank, and the results demonstrate the effectiveness of MLGANN for DTI prediction. Looking ahead, we are poised to expand our research horizons by incorporating additional datasets, thereby enriching our training and validation frameworks. This strategic augmentation is anticipated to further refine the predictive capabilities of our model, making it a more potent tool in the realm of drug discovery and development.

Data availability

The data on drug-target interactions underpinning this study were sourced from DrugBank (Version 5.1.8). For more details, please visit https://go.drugbank.com/.

Received: 2 February 2024; Accepted: 8 October 2024
Published online: 30 October 2024

References

1. Nosengo, N. et al. Can you teach old drugs new tricks? *Nature* **534**, 314–316 (2016).
2. Ezzat, A., Zhao, P., Wu, M., Li, X.-L. & Kwoh, C.-K. Drug-target interaction prediction with graph regularized matrix factorization. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **14**, 646–656 (2016).
3. Ding, Y., Tang, J. & Guo, F. Identification of drug-target interactions via dual Laplacian regularized least squares with multiple kernel fusion. *Knowl.-Based Syst.* **204**, 106254 (2020).
4. An, Q. & Yu, L. A heterogeneous network embedding framework for predicting similarity-based drug-target interactions. *Brief. Bioinform.* **22**, bbab275 (2021).
5. Gao, K. Y. et al. Interpretable drug target prediction using deep neural representation. In *IJCAI* Vol. 2018, 3371–3377 (2018).
6. Huang, K., Xiao, C., Glass, L. M. & Sun, J. Moltrans: Molecular interaction transformer for drug-target interaction prediction. *Bioinformatics* **37**, 830–836 (2021).
7. Chu, Y. et al. DTI-MLCD: Predicting drug-target interactions using multi-label learning with community detection method. *Brief. Bioinform.* **22**, bbaa205 (2021).
8. Liu, B., Pliakos, K., Vens, C. & Tsoumakas, G. Drug-target interaction prediction via an ensemble of weighted nearest neighbors with interaction recovery. *Appl. Intell.* **52**, 1–23 (2022).
9. Pliakos, K., Vens, C. & Tsoumakas, G. Predicting drug-target interactions with multi-label classification and label partitioning. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **18**, 1596–1607 (2019).
10. Olayan, R. S., Ashoor, H. & Bajic, V. B. DDR: Efficient computational method to predict drug-target interactions using graph mining and machine learning approaches. *Bioinformatics* **34**, 1164–1173 (2018).

11. Zhang, Z.-C. et al. A graph regularized generalized matrix factorization model for predicting links in biomedical bipartite networks. *Bioinformatics* **36**, 3474–3481 (2020).
12. Luo, Y. et al. A network integration approach for drug-target interaction prediction and computational drug repositioning from heterogeneous information. *Nat. Commun.* **8**, 573 (2017).
13. Wan, F., Hong, L., Xiao, A., Jiang, T. & Zeng, J. NeoDTI: Neural integration of neighbor information from a heterogeneous network for discovering new drug-target interactions. *Bioinformatics* **35**, 104–111 (2019).
14. Li, Y., Qiao, G., Gao, X. & Wang, G. Supervised graph co-contrastive learning for drug-target interaction prediction. *Bioinformatics* **38**, 2847–2854 (2022).
15. Li, J., Wang, J., Lv, H., Zhang, Z. & Wang, Z. IMCHGAN: Inductive matrix completion with heterogeneous graph attention networks for drug-target interactions prediction. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **19**, 655–665 (2021).
16. Adasme, M. F., Parisi, D., Sveshnikova, A. & Schroeder, M. Structure-based drug repositioning: Potential and limits. In *Seminars in cancer biology* Vol. 68 192–198 (Elsevier, 2021).
17. Mohamed, S. K., Nounu, A. & Nováček, V. Drug target discovery using knowledge graph embeddings. In *Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing*, 11–18 (2019).
18. Trouillon, T., Welbl, J., Riedel, S., Gaussier, É. & Bouchard, G. Complex embeddings for simple link prediction. In *International Conference on Machine Learning*, 2071–2080 (PMLR, 2016).
19. Ye, Q. et al. A unified drug-target interaction prediction framework based on knowledge graph and recommendation system. *Nat. Commun.* **12**, 6775 (2021).
20. Li, M., Cai, X., Xu, S. & Ji, H. Metapath-aggregated heterogeneous graph neural network for drug-target interaction prediction. *Brief. Bioinform.* **24**, bbac578 (2023).
21. Fan, Y. et al. SGCLDGA: Unveiling drug-gene associations through simple graph contrastive learning. *Brief. Bioinform.* **25**, bbae231 (2024).
22. Cao, J. et al. NGCN: Drug-target interaction prediction by integrating information and feature learning from heterogeneous network. *J. Cell Mol. Med.* **28**, e18224 (2024).
23. Dehghan, A., Razzaghi, P., Abbasi, K. & Gharaghani, S. TripletMultiDTI: Multimodal representation learning in drug-target interaction prediction with triplet loss function. *Expert Syst. Appl.* **232**, 120754 (2023).
24. Rafiei, F. et al. DeepTraSynergy: Drug combinations using multimodal deep learning with transformers. *Bioinformatics* **39**, btad438 (2023).
25. Gharizadeh, A., Abbasi, K., Ghareyazi, A., Mofrad, M. R. & Rabiee, H. R. HGTDR: Advancing drug repurposing with heterogeneous graph transformers. Preprint at arXiv:2405.08031 (2024).
26. Zeng, X. et al. Target identification among known drugs by deep learning from heterogeneous networks. *Chem. Sci.* **11**, 1775–1797 (2020).
27. Bordes, A., Usunier, N., Garcia-Duran, A., Weston, J. & Yakhnenko, O. Translating embeddings for modeling multi-relational data. *Adv. Neural Inf. Processi. Syst.* **26** (2013).

Author contributions

Q.L. conceived the project, developed the prediction method, designed the experiments, implemented the experiments, and wrote the paper. Z.Z. conceived the project, designed the experiments, analyzed the result, and wrote the paper. Q.W. conceived the project, analyzed the result, and revised the paper. All authors have reviewed and approved the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Q.W.

Reprints and permissions information is available at www.nature.com/reprints.

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