Briefings in Bioinformatics, 22(6), 2021, 1–12

https://doi.org/10.1093/bib/bbab319
Problem Solving Protocol

Drug repositioning based on the heterogeneous information fusion graph convolutional network

Lijun Cai, Changcheng Lu, Junlin Xu, Yajie Meng, Peng Wang, Xiangzheng Fu, Xiangxiang Zeng and Yansen Su

Corresponding authors: Junlin Xu, College of Computer Science and Electronic Engineering, Hunan University, Changsha, Hunan, 410082, China, Tel.: 86-18273118685; E-mail: xjl@hnu.edu.cn; Peng Wang, College of Computer Science and Electronic Engineering, Hunan University, Changsha, Hunan, 410082, China, Tel.: 86-15700770926; E-mail: wangpenglw@126.com; Yansen Su, Key Laboratory of Intelligent Computing and Signal Processing of Ministry of Education, School of Computer Science and Technology, Anhui University, Hefei 230601, China, Tel.: 86-15755146125; E-mail: suyansen@ahu.edu.cn

Abstract

In silico reuse of old drugs (also known as drug repositioning) to treat common and rare diseases is increasingly becoming an attractive proposition because it involves the use of de-risked drugs, with potentially lower overall development costs and shorter development timelines. Therefore, there is a pressing need for computational drug repurposing methodologies to facilitate drug discovery. In this study, we propose a new method, called DRHGCN (Drug Repositioning based on the Heterogeneous information fusion Graph Convolutional Network), to discover potential drugs for a certain disease. To make full use of different topology information in different domains (i.e. drug–drug similarity, disease–disease similarity and drug–disease association networks), we first design inter- and intra-domain feature extraction modules by applying graph convolution operations to the networks to learn the embedding of drugs and diseases, instead of simply integrating the three networks into a heterogeneous network. Afterwards, we parallelly fuse the inter- and intra-domain embeddings to obtain the more representative embeddings of drug and disease. Lastly, we introduce a layer attention mechanism to combine embeddings from multiple graph convolution layers for further improving the prediction performance. We find that DRHGCN achieves high performance (the average AUROC is 0.934 and the average AUPR is 0.539) in four benchmark datasets, outperforming the current approaches. Importantly, we conducted molecular docking experiments on DRHGCN-predicted candidate drugs, providing several novel approved drugs for Alzheimer's disease (e.g. benzatropine) and Parkinson's disease (e.g. trihexyphenidyl and haloperidol).

Key words: drug; drug–disease association prediction; drug repositioning; graph convolutional network; heterogeneous information fusion

Introduction

With advancements in genomics, proteomics, life sciences and technologies, pharmaceutical research and development technology has developed rapidly over the past decades. However, *de novo* drug discovery is still time-consuming, costly and laborious [1]. Approximately 90% of experimental drugs fail to pass phase I clinical trials because drugs with new structures tend to induce unpredictable adverse reactions [2]. Meanwhile, the drug development pipeline requires billions of dollars, and introducing a new drug to the market takes an average of about 9–12 years

Lijun Cai is a professor at Hunan University. His research interests include bioinformatics, artificial intelligence and data mining.
Changcheng Lu is a graduate student at Hunan University. His research interests include bioinformatics and deep learning.
Junlin Xu is a doctoral student at Hunan University. His research interests include clustering, single cell, bioinformatics and computational biology.
Yajie Meng is a doctoral student at Hunan University. His research interests include bioinformatics and data mining.
Peng Wang is a post-doctoral researcher at Hunan University. His research interests include bioinformatics and data mining.
Xiangzheng Fu is a post-doctoral researcher at Hunan University. His research interest is classification of proteins in bioinformatics.
Xiangxiang Zeng is a professor at Hunan University. His research interests include bio-computing and bioinformatics.
Yansen Su is an associate professor at Anhui University. Her research interests include bioinformatics and systems biology.
Submitted: 5 May 2021; **Received (in revised form):** 30 June 2021

[3]. Thus, improving the productivity of pharmaceutical research and development is challenging and meaningful. Drug repositioning (also known as drug repurposing) is gaining increased attention because of its capability to dramatically reduce the temporal and monetary costs of drug development, lower the risk of unexpected toxicity, and shorten the overall development cycle [4–6]. Instead of adopting an exhausting wet-lab experiments search, drug repositioning identifies new indications for existing approved drugs and then experimentally verifies possible candidates [7, 8]. In recent years, drug repositioning has been widely applied in disease and related therapeutic areas, such as anticancer drug discovery [5], identification of novel therapies for orphan and rare diseases [9], overcoming of drug resistance [10] and advancement of personalized medicine [11]. These successful applications have shown that drug repositioning is increasingly becoming an attractive proposition [12].

The computational drug repositioning method has achieved rapid development due to rapidly accumulating high-throughput data (e.g. genomic data, protein structures and phenotypes) and the progress made in computer sciences [6, 13]. It narrows down the search space for the existing drugs by suggesting potential candidate drugs for validation via wet-lab experiments. In recent years, many computational technologies have been utilized to predict the potential or new indications of small-molecule compounds [14, 15]. Cao et al. [16] developed a computational method for the drug–target interactions prediction by combining the information from chemical, biological and the drug–target interaction network. They formulate the drug–target interactions prediction problem as a binary classification problem where the information about drugs, targets and confirmed drug–target interactions are translated into features that are used to train a random forest (RF) model, which is then utilized to predict the drug–target interactions. However, the method faces the problem of selecting effective negative samples. Chen et al. [17] proposed the network-based random walk with restart on the heterogeneous network method (termed NRWRH) to predict potential drug–target interactions. NRWRH integrates protein–protein similarity, drug–drug similarity and known drug–target interaction networks into a heterogeneous network. Alaimo et al. [18] presented a novel network-based inference method called domain tuned-hybrid (DT-hybrid) for the identification of drug–target interactions. DT-hybrid considers important features within the drug–target domain by adding the domain-based knowledge through a similarity matrix. To integrate more prior information flexibly, Dai et al. [19] used a matrix factorization method with the integration of genomic space to predict novel drug indications. They fuse the topological characteristics of gene interaction network into the features of drugs and diseases. Luo et al. [20] offered a drug repositioning recommendation system called DRRS to predict novel drug indications. DRRS assumes that the drug–disease matrix is low-rank and utilizes a fast singular value thresholding algorithm to complete the drug–disease matrix. Yang et al. [21] utilized a matrix completion algorithm, called bounded nuclear norm regularization (BNNR), to construct low-rank drug–disease matrix approximations consistent with known associations. In order to optimize the fusion process of multiple similarities, Yang et al. [22] further designed a multi-similarities bilinear matrix factorization method, called MSBMF, to identify promising drug-associated indications for drugs. Different from the linear multiplication of latent features in conventional matrix factorization, Zeng et al. [23] developed a network-based deep learning approach, called deepDR. deepDR learns high-level features of drugs from heterogeneous networks by using a multi-modal deep autoencoder and then infers new applications for existing drugs by using a collective variational autoencoder. Li et al. [24] fused drug–drug similarity and disease–disease similarity into a gray-scale image, and then deep convolution neural network was utilized to identifying potential drug–disease associations.

The above models can be roughly classified into four principal categories: (i) classical machine learning model-based approaches, such as Cao's method; (ii) network propagation-based approaches, such as NRWRH, DT-hybrid and Luo's method; (iii) matrix factorization and completion approaches, such as Dai's method, DRRS, BNNR; (iv) deep learning, such as deepDR and Li's method.

Recently, the graph convolutional network (GCN) has attracted increasing attention and been applied to many networks-related prediction tasks to learn network topology-preserving node-level vector embedding [25, 26]. For example, Huang et al. [26] adopted a graph convolution for miRNA-drug resistance associations prediction (called GCMDR). GCMDR focuses on the associations between miRNA and drug and overlooks the neighborhood associations among a small group of closely related miRNAs or drugs. Li et al. [27] presented a novel method of neural inductive matrix completion with graph convolutional network (NIMCGCN) for predicting miRNA-disease associations. NIMCGCN performs the graph convolution to the miRNA–miRNA and disease–disease networks, respectively, but the drug-miRNA associations are ignored in the feature extraction process. Similar to [27], the input graph in the study of Zhao et al. [28] only considers the drug–drug associations and target–target associations, while ignoring the interaction between drug and target. Yu et al. [29] proposed a layer attention graph convolutional network (LAGCN) for predicting the drug–disease associations. LAGCN integrates the known drug–disease associations, drug–drug similarities and disease–disease similarities into a heterogeneous network and utilizes the graph convolutional operations to the heterogeneous network to learn the drug and disease embeddings. However, network topology information from different domains (i.e. drug and disease domains) is mixed without distinction in such a heterogeneous network, which may cause a substantial loss of network-specific information.

To overcome the mentioned limitations, we proposed a novel drug repositioning method of heterogeneous information fusion graph convolutional network (DRHGCN) to address the problem of drug repositioning in this study. The basic idea of DRHGCN is as follows. We used three networks i.e. the drug–drug similarity, disease–disease similarity and observed drug–disease associations networks. First, considering the different network topology information in different domains, we designed an intra-domain feature extraction module to extract drug and disease intra-domain embeddings based on drug–drug and disease–disease similarity networks, and developed an inter-domain feature extraction module for message passing between drugs and diseases, which is composed of a bilinear aggregator and a traditional GCN aggregator. Then, the inter- and intra-domain embeddings were parallelly fused to obtain the embedding representation of the drug and disease. Third, to enrich the representation capability of drugs and diseases, we adopted a layer attention mechanism to combine embeddings from multiple graph convolutional layers. Benchmarking comparison results show that DRHGCN achieves substantial performance improvement performs over the existing models. In summary, DRHGCN is a useful tool to prioritize existing drugs for further investigation, which has the potential to accelerate

Table 1. Statistics of datasets used in this study

Datasets	No. of drugs	No. of diseases	No. of associations	Sparsity*
Fdataset	593	313	1933	0.0104
Cdataset	663	409	2352	0.0087
Ldataset	598	269	18 416	0.1145
LRSSL	763	681	3051	0.0059

*The sparsity is defined as the ratio of the number of known associations to the number of all possible associations.

drug discovery and therapeutic development for understudied diseases.

Methods

Datasets

The performance of DRHGCN was evaluated on four benchmark datasets, namely Fdataset [30], Cdataset [31], LRSSL [32] and Ldataset [29]. Table 1 summarizes the statistical data, including (i) number of drugs, (ii) number of diseases, (iii) number of the known drug–disease associations and (iv) sparsity of the drug–disease association matrix, of the three datasets. The first Fdataset consists of 593 drugs, 313 diseases and 1933 proven drug–disease associations. Drugs are extracted from the DrugBank (DB) database [33], which is a comprehensive database containing extensive information about drugs and their targets. Diseases are collected from human phenotypes defined in the Online Mendelian Inheritance in Man (OMIM) database [34], which is a public resource for information on human genes and diseases. Therefore, we opted to use Fdataset for a comprehensive test of our algorithm performance. The second dataset Cdataset contains 2352 associations between 663 drugs in DrugBank database and 409 diseases in OMIM database. The third dataset LRSSL includes 763 drugs, 681 diseases and 3051 drug–disease associations from the study of Wang et al. [35]. In order to better present the relationship between these three datasets, we have drawn a Venn diagram (Supplementary Figure S1). The last dataset Ldataset contains 598 drugs, 269 diseases and 18 416 known drug–disease associations derived from the Comparative Toxicogenomics Database (CTD) [36].

In this study, the similarity matrix S^r of drug pairs was calculated based on the chemical structure of SMILES [37], which measures the similarity between the drugs via the Chemical Development Kit [38] calculating the Tanimoto score [39] of 2D chemical fingerprints of the drug pairs. The similarity matrix S^d of disease pairs was conventionally calculated based on disease phenotype by using MimMiner [40], which measures the similarity between diseases through text mining analysis of their medical descriptions information in the OMIM database.

Construction of three networks

We constructed three networks, namely, known drug–disease association, drug–drug similarity and disease–disease similarity, to improve the prediction performance by completely infusing heterogeneous information. The known drug–disease association network was denoted as a graph G with N drugs and M diseases, and its adjacent matrix is $A \in \{0,1\}^{N\times M}$. $A_{ij} = 1$ if a drug r_i is associated with a disease d_j. $A_{ij} = 0$ if the association between drug r_i and disease d_j is unknown or unobserved.

For drug–drug similarity network, we let graph G_r denote the drug–drug similarity network with N drugs, and its adjacency matrix $A^r \in \mathbb{R}^{N\times N}$ is composed of the drug similarity matrix S^r. Specifically, if drug r_j is the *topk* nearest neighbor of drug r_i based on drug similarity matrix S^r, then the (i,j) th entry of A^r is S^r_{ij}; otherwise $A^r_{ij} = 0$. *topk* denotes the number of K nearest neighbors of each drug or each disease. Similarly, for the disease–disease similarity network, we let graph G_d denote the disease–disease similarity network with M diseases, and its adjacency matrix $A^d \in \mathbb{R}^{M\times M}$ is composed of disease similarity matrix S^d. Specifically, if disease d_j is the *topk* nearest neighbor of disease d_i based on disease similarity matrix S^d, then the (i,j) th entry of A^d is S^d_{ij}; otherwise $A^d_{ij} = 0$.

Model structure

This section describes DRHGCN in detail to predict the potential drug–disease associations. The workflow of DRHGCN is briefly shown in Figure 1.

Encoder

GCN [25] is a multilayer connected neural network architecture used to learn low-dimensional representations of nodes from graph-structured data. The encoder of DRHGCN is mainly based on drug–drug and disease–disease similarity networks and uses GCN to extract drug and disease intra-domain embeddings, respectively. Meanwhile, the drug–disease association network is utilized to extract drug and disease inter-domain embeddings. Subsequently, the inter- and intra-domain embeddings are fused to obtain the final embedding representation of the drug and disease.

First, we initialize the embeddings of drugs and diseases as follows:

$$H^0 = \begin{bmatrix} H^0_r \\ H^0_d \end{bmatrix} = \begin{bmatrix} S^r & 0 \\ 0 & S^d \end{bmatrix} \in \mathbb{R}^{(N+M)\times(N+M)}. \quad (1)$$

Second, the intra-domain feature extraction module is defined as follows:

$$\hat{H}^{l+1} = \begin{bmatrix} \hat{H}^{l+1}_r \\ \hat{H}^{l+1}_d \end{bmatrix} = \begin{bmatrix} GCN\left(A^r, H^l_r, W^l_r\right) \\ GCN\left(A^d, H^l_d, W^l_d\right) \end{bmatrix}, \quad (2)$$

where $\hat{H}^{l+1}_r \in \mathbb{R}^{N\times k}$ is the drug intra-domain output features at the l th-layer, $\hat{H}^{l+1}_d \in \mathbb{R}^{M\times k}$ is the disease intra-domain output features at the l th-layer, H^l_r is the drug input embeddings at the l th-layer, H^l_d is the disease input embeddings at the l th-layer and $W^l_r \in \mathbb{R}^{k\times k}$ and $W^l_d \in \mathbb{R}^{k\times k}$ are trainable matrices of the l th-layer intra-domain feature extraction module. Graph convolution operation is denoted as $GCN(A, H, W)$, and is formulated as:

$$\mathrm{GCN}(\mathrm{A},\mathrm{H},\mathrm{W}) = \sigma\left(D^{-\frac{1}{2}}AD^{-\frac{1}{2}}\mathrm{HW}\right), \quad (3)$$

where $D = \mathrm{diag}(\sum_j A_{ij})$ and $\sigma(\cdot)$ is a ReLU [41] activation function.

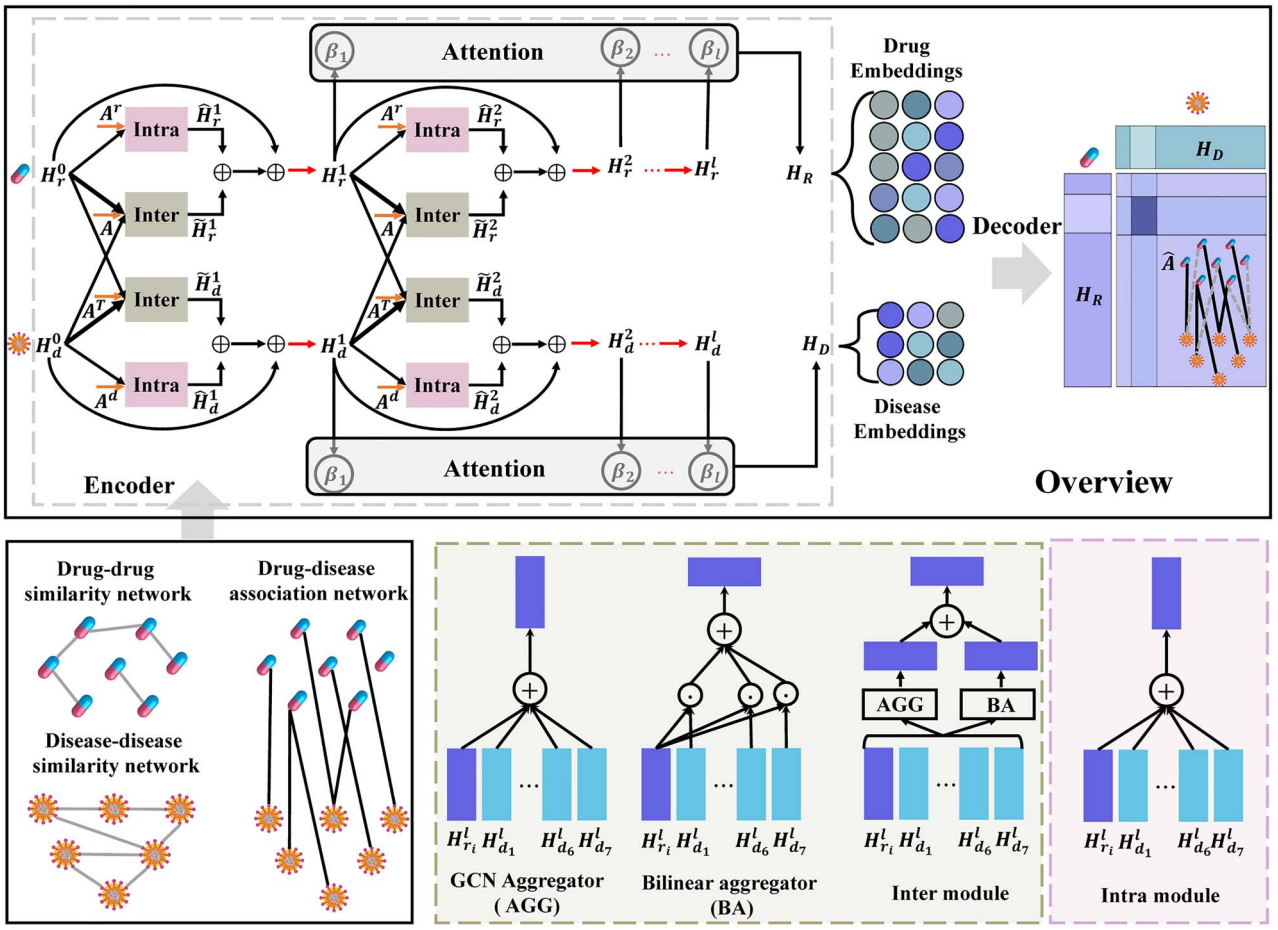

Figure 1. Overview of the DRHGCN architecture. DRHGCN mainly consists of a GCN-based encoder and a decoder, and it completely fuses heterogeneous information of drug–drug similarity, disease–disease similarity and drug–disease association networks. The GCN-based encoder utilizes two inter- and intra-domain modules to parallelly extract drug and disease features. The intra-domain module extracts drug- and disease-specific intra-domain features, and the inter-domain module emphasizes the common properties and dilutes the inconsistent information between different biological domains to extract drug and disease inter-domain features. The decoder predicts the association probability score based on pairs of drug and disease embeddings.

Inspired by Bilinear Graph Neural Network (BGNN) [42], the multiplication of two vectors is an effective means to model the interaction, and it can emphasize the common properties and dilute the discrepant information. Our proposed inter-domain feature extraction module for message passing between drugs and diseases is composed of a bilinear aggregator (BA) and a traditional GCN aggregator (AGG). The bilinear aggregator is suitable for modeling the neighbor interactions in the local structure by summation and element-wise product operations. Specifically, for a drug r_i, its drug inter-domain feature extraction module is defined as:

$$\tilde{H}_{r_i}^{l+1} = \sigma\left(\alpha^l \frac{\sum_j \left(H_{d_j}^l W^l \odot H_{r_i}^l W^l\right) A_{ij}}{\sum_j A_{ij}} + \left(1-\alpha^l\right)\frac{\sum_j H_{d_j}^l W^l A_{ij}}{\sum_j A_{ij}}\right), \quad (4)$$

where $\odot$ is an element-wise product, $\tilde{H}_{r_i}^{l+1} \in \mathbb{R}^k$ is the l th-layer drug inter-domain output feature of drug r_i, $W^l \in \mathbb{R}^{k\times k}$ is a trainable matrix and $\alpha^l \in \mathbb{R}$ is a trainable scalar used to balance the importance between BA and the traditional GCN aggregator. In the same manner, for a disease d_j, its disease inter-domain feature extraction module is defined as:

$$\tilde{H}_{d_j}^{l+1} = \sigma\left(\alpha^l \frac{\sum_i \left(H_{r_i}^l W^l \odot H_{d_j}^l W^l\right) A_{ij}}{\sum_i A_{ij}} + \left(1-\alpha^l\right)\frac{\sum_i H_{r_i}^l W^l A_{ij}}{\sum_i A_{ij}}\right), \quad (5)$$

where $\tilde{H}_{d_j}^{l+1} \in \mathbb{R}^k$ is the l th-layer disease inter-domain output feature of disease d_j.

For simplicity, we merged intra-domain and inter-domain features as follows:

$$\begin{bmatrix} H_r^{l+1} \\ H_d^{l+1} \end{bmatrix} = \begin{bmatrix} \hat{H}_r^{l+1} + \tilde{H}_r^{l+1} \\ \hat{H}_d^{l+1} + \tilde{H}_d^{l+1} \end{bmatrix}, \quad (6)$$

where H_r^{l+1} denotes the drug infusion features, and H_d^{l+1} represents the disease infusion features. Stacking multiple GCN layers leads to the common vanishing gradient problem [43]. This condition means that back-propagating through these networks would cause over-smoothing [44], eventually leading to the features of graph vertices within each connected component converging to the same value. Hence, we added a residual skip

Algorithm 1 DRHGCN

Input: Drug adjacency matrix A^r; disease adjacency matrix A^d; drug–disease associations matrix A; drug embedding matrix H_r^0; disease embedding matrix H_d^0; number of layers L; maximum training epochs T.
Output: Predicted probability score matrix $\hat{A}$.
Initialize trainable parameters $\theta = \{W_d^l, W_r^l, W^l, \alpha^l, \beta^l, \dots\}$;
$t \leftarrow 1$;
repeat
  for $l = 1, 2, \dots, L$ do
    Computer intra-domain features $\hat{H}^l$ with Eq. (2);
    Computer inter-domain features $\tilde{H}^l$ with Eqs (4) and (5);
    Obtain nodes embeddings H^l by merge $\hat{H}^l$ and $\tilde{H}^l$ with Eq. (7);
  end for
    Combine nodes embeddings H^l with Eq. (8), and obtain the final embeddings of drugs H_R and the final embeddings of diseases H_D;
    Obtain the prediction matrix $\hat{A}$ with Eq. (9);
    Update θ by optimizing Eq. (10);
    $t \leftarrow t + 1$;
until $t > T$ or Eq. (10) is converged;
return $\hat{A}$;

connection [43, 45] between each module's input and output layers, the Equation (6) is modified as follows:

$$H^{l+1} = \hat{H}^{l+1} + \tilde{H}^{l+1} + H^l = \begin{bmatrix} \hat{H}_r^{l+1} + \tilde{H}_r^{l+1} + H_r^l \\ \hat{H}_d^{l+1} + \tilde{H}_d^{l+1} + H_d^l \end{bmatrix}, \tag{7}$$

where H^{l+1} is the l th-layer output embeddings of the nodes (drugs and diseases). The embeddings at different GCN layers capture different levels of information of the input graphs. After the L layer, we obtained L k-dimensional drug and disease embeddings, respectively. Similar to LAGCN [29], we introduced layer attention as one component of the network architecture of DRHGCN, which adaptively combines embeddings at different graph convolution layers with an attention mechanism to further improve the prediction performance. Specifically, we paid different attention to convolution layers to integrate embeddings and obtained the final embeddings of drugs and diseases as follows:

$$\begin{bmatrix} H_R \\ H_D \end{bmatrix} = \sum_{l=1} \beta^l H^l, \tag{8}$$

where $\beta^l \in \mathbb{R}$ is auto-learned by neural networks and initialized as $1/L$, which denotes the contributions of the embeddings at different convolution layers to the final embeddings. $H_R \in \mathbb{R}^{N \times k}$ is the final embeddings of drugs, and $H_D \in \mathbb{R}^{M \times k}$ is the final embeddings of diseases.

Decoder

To reconstruct the associations between drugs and diseases, our decoder $f(H_R, H_D)$ is formulated as follows:

$$\hat{A} = f(H_R, H_D) = sigmoid\left(H_R H_D^T\right), \tag{9}$$

where $\hat{A} \in \mathbb{R}^{N \times M}$ is the predicted probability score matrix. The predicted score for the association between drug r_i and disease d_j is given by the corresponding (i, j) th entry of $\hat{A}$. The detailed steps of DRHGCN to predict novel drug–disease associations are described in Algorithm 1.

Optimization

Given that known drug–disease associations have been validated manually, they are highly reliable and important for improving prediction performance. However, the number of known drug–disease associations is far less than the number of unknown or unobserved drug–disease pairs. Hence, our DRHGCN learns parameters by minimizing the weighted binary cross-entropy loss as follows:

$$loss = -\frac{1}{N \times M}\left(\gamma \times \sum\nolimits_{(i,j) \in S^+} \log \hat{A}_{ij} + \sum\nolimits_{(i,j) \in S^-} \left(1 - \log \hat{A}_{ij}\right)\right), \tag{10}$$

where (i, j) denotes the pair of drug r_i and disease d_j, S^+ denotes the set of all known drug–disease association pairs and S^- represents the set of all unknown or unobserved drug–disease association pairs. The balance factor $\gamma = \frac{|S^-|}{|S^+|}$ is used to reduce the impact of data imbalance, where $|S^-|$ and $|S^+|$ are the number of pairs in S^+ and S^-.

Following previous studies [29], we optimized the model through the Adam optimizer [46] and initialized weights as described in [47]. To generalize effectively to the unobserved data, we trained the model in a denoising setup by randomly dropping out edges with a fixed probability. We also applied regular dropout [48] to the graph convolution layers.

Results and discussions

Baseline methods

To evaluate the performance of our proposed model, we compared DRHGCN with the six state-of-the-art drug repositioning methods listed below.

- DRIMC [49] is an improved drug repositioning method based on Bayesian induction matrix completion.
- DRRS [20] models the drug repositioning problem as recommending novel treatments based on known drug–disease associations by using a low-rank matrix approximation and random algorithm.
- LAGCN [29] is a layered attention graph convolutional network, which is used for the drug–disease associations prediction.

Table 2. The AUROC and AUPR obtained under the 10-fold cross-validation

Datasets	NRLMF	DRIMC	DRRS	SCMFDD	NIMCGCN	LAGCN	DRHGCN
AUROC							
Fdataset	0.936 ±0.002*	0.913±0.002*	0.928±0.001*	0.776±0.001*	0.832±0.004*	0.884±0.026*	0.944±0.002*
Cdataset	0.949 ±0.002*	0.933±0.001*	0.948±0.002*	0.793±0.001*	0.855±0.004*	0.920±0.005*	0.960±0.001*
Ldataset	0.776±0.003*	0.757±0.001*	0.841±0.001*	0.850 ±0.001*	0.826±0.002*	0.812±0.087*	0.876±0.001*
LRSSL	0.923±0.002*	0.932±0.001*	0.927±0.002*	0.769±0.001*	0.833±0.004*	0.935 ±0.001*	0.957±0.001*
Avg.	0.896	0.884	0.911	0.797	0.837	0.888	0.934
AUPR							
Fdataset	0.522 ±0.010*	0.314±0.004*	0.475±0.006*	0.051±0.001*	0.344±0.006*	0.130±0.012*	0.543±0.006*
Cdataset	0.616 ±0.006*	0.392±0.003*	0.572±0.005*	0.052±0.001*	0.441±0.009*	0.191±0.006*	0.640±0.005*
Ldataset	0.359 ±0.003*	0.306±0.001*	0.397±0.003*	0.495±0.002*	0.440±0.004*	0.501±0.060*	0.555±0.001*
LRSSL	0.462±0.005*	0.267±0.001*	0.342±0.006*	0.036±0.001*	0.274±0.007*	0.114±0.003*	0.417 ±0.005*
Avg.*	0.490	0.320	0.447	0.159	0.374	0.234	0.539

indicates that DRHGCN significantly outperforms the other compared methods with P-values <0.05 using the paired t-test. Avg. shows the average AUROC/AUPR over four datasets. The best results in each row are in bold faces and the second best results are underlined.

- SCMFDD [50] is a similarity constraint matrix factorization to predict drug–disease associations.
- NIMCGCN [27] is a new neural induction matrix completion method of the graph convolutional network, and it was first used to predict miRNA–disease associations. Later on, it was proven to have great potential in drug repositioning.
- NRLMF [51] focuses on modeling the probability of drug–target interaction through logistic matrix factorization.

Parameter setting

Our proposed DRHGCN algorithm uses a three-layer architecture with 64 hidden units in each layer, a regular dropout rate of 0.4, an edge dropout rate of 0.2, learning rate of 0.05 and a maximum training epoch of 400 in all experiments. We empirically set *topk* to 15 (**Supplementary Figures S2-S3**). The hyperparameters of DRRS, LAGCN, SCMFDD and NIMCGCN were chosen as their optimal values provided by their publications. For NRLMF, following [49], we set the dimensionality of the subspace $R = 200$, regularization parameters $\lambda_d = \lambda_t = 0.125$, $\alpha = 0.25$, $\beta = 0.125$ and learning rate $\gamma = 0.5$.

Performance of DRHGCN in the cross-validation

To evaluate the performance of DRHGCN, we performed 10-fold cross-validation in four benchmark datasets. During the 10-fold cross-validation, we randomly selected 10% of the known drug–disease association pairs in the dataset and 10% of the random unknown drug–disease association pairs as the testing set; the remaining 90% of clinically reported drug–disease association pairs and unknown drug–disease associations pairs were used to train the model. The area under the receiver operating characteristic curve (AUROC) and the area under the precision–recall curve (AUPR) has been widely used in bioinformatics research [52, 53], and are adopted to evaluate the overall performance of DRHGCN. Considering the potential data bias of cross-validation, we performed 10 times 10-fold cross-validation on four datasets and calculated the variance of the results using 10-cross validation repeated 10 times to show the stability of the results. We repeated each test 10 times to obtain an average result. As shown in Table 2, the final average AUROC obtained by DRHGCN is 0.934, which is 2.36% higher than the second best method DRRS, and the average AUPR obtained by DRHGCN is 0.539, which is 4.93% higher than that obtained by the second method NRLMF. It should be noted that DRHGCN achieves the highest AUPR over three datasets (i.e. Fdataset, Cdataset and Ldataset) and obtains the second best AUPR on the LRSSL dataset, where NRLMF outperforms DRHGCN (0.462 for NRLMF versus 0.417 for DRHGCN). In addition, we performed a paired t-test to evaluate whether DRHGCN's performance is significantly better than the other methods. The statistical results in Supplementary Tables S1 and S2 indicate that DRHGCN yields the significantly better performance under the *P*-value threshold of 0.05 in terms of not only AUROCs but AUPRs as well. Benchmarking comparison results on four datasets show that DRHGCN performs better than the six state-of-the-art prediction models. In particular, the result of each 10-fold cross-validation is in general consistent, indicating our model exhibits convincing performance and is highly robust (Figure 2, Supplementary Figures S4–S7).

When the number of negative samples in the dataset is much larger than the number of positive samples, the number of true positives that are usually correctly predicted reflects the distinguishing capability of a prediction method that distinguishes true positives. Biologists usually select the top prediction results for further verification through wet laboratory experiments. The top k candidates are accurate and can discover potential drug–disease associations. As shown in Supplementary Figure S8, the number of positive samples successfully recovered by DRHGCN in the top 600 samples was significantly better than that of the other algorithms in addition to the LRSSL dataset.

Discovering candidates for new diseases or new drugs

To evaluate the capability of DRHGCN for predicting potential drugs for new diseases, we conducted a new experiment on Fdataset. For each disease d_i, we deleted all the known

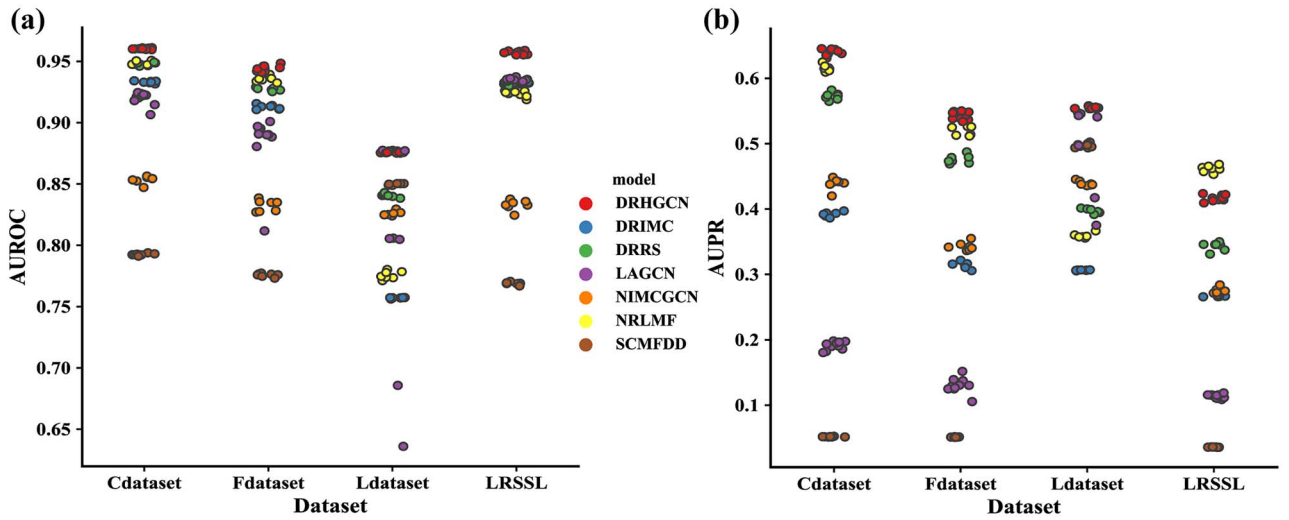

Figure 2. The performance of all methods in 10 times 10-fold-cross-validation of Fdataset, Cdataset, Ldataset and LRSSL dataset, respectively. (A) Area under the receiver operating characteristic curves (AUROC) of prediction results. (B) Area under the precision-recall curves (AUPR) of prediction results.

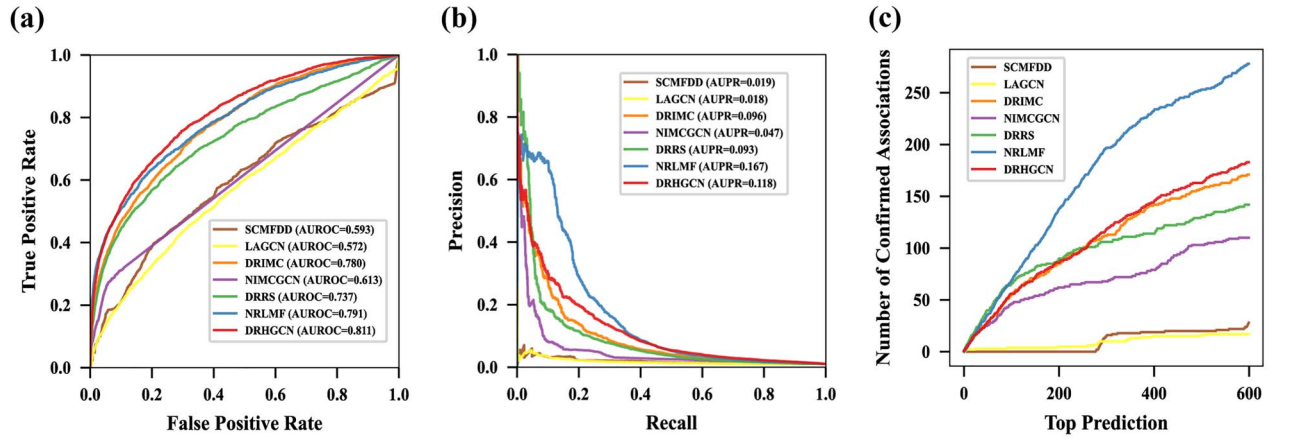

Figure 3. The performance of all methods in predicting potential drugs for new diseases on Fdataset. (A) Receiver operating characteristic (ROC) curves of prediction results obtained by applying DRHGCN and other competitive methods. (B) Precision-recall (PR) curves of prediction results obtained by applying DRHGCN and other competitive methods. (C) The confirmed associations in top 600 predictions obtained by applying DRHGCN and other competitive methods.

drug–disease associations about disease d_i as the testing set and used all the remaining associations as the training samples. Without any known associations of a new disease, DRHGCN could utilize the similarity information of the new disease to predict the potential disease-related drugs. Compared with the six other methods, DRHGCN achieved excellent performance (AUROC = 0.811). NRLMF, DRIMC, DRRS, SCMFDD, NIMCGCN and LAGCN had AUROCs of 0.791, 0.780, 0.737, 0.593, 0.613 and 0.572, respectively (Figure 3A, Supplementary Table S3 **and Figure S9**). In addition, we further verified the number of positive samples successfully recovered by the top k candidates. As shown in Figure 3C, under the highest threshold of 600, the number of positive samples successfully recovered by DRHGCN was second only to that of NRLMF, which again shows that DRHGCN has a good capability for accurate prioritization of potential disease-related drugs.

For a new drug without any known associations, DRHGCN is able to predict the potential indications for new drugs by taking advantage of the similarity information of new drugs. We also add the opposite case in which all relationships for each drug are removed to predict indications for new drugs. As shown in Supplementary Figure S10, we find that DRHGCN achieves the second best performance (AUROC is 0.147 and AUPR is 0.808) in Fdataset. The reason is that the input of NRLMF only contains drug–drug similarity and disease–disease similarity, while the input of DRHGCN contains known drug–disease association.

Performance of DRHGCN by ablation analysis

According to Figure 1, DRHGCN mainly consists of two parts: inter-domain and intra-domain feature extraction modules. To check the contribution of each component, we compared DRHGCN with four variants based on Fdataset. The model variants are summarized as follows:

- DRHGCN-Inter: It mainly consists of a GCN-based encoder and a decoder. Specifically, the GCN-based encoder only utilizes the inter-domain feature extraction module to extract drug and disease features. The decoder predicts the association probability score based on pairs of drug and disease embeddings (Supplementary Figure S11).

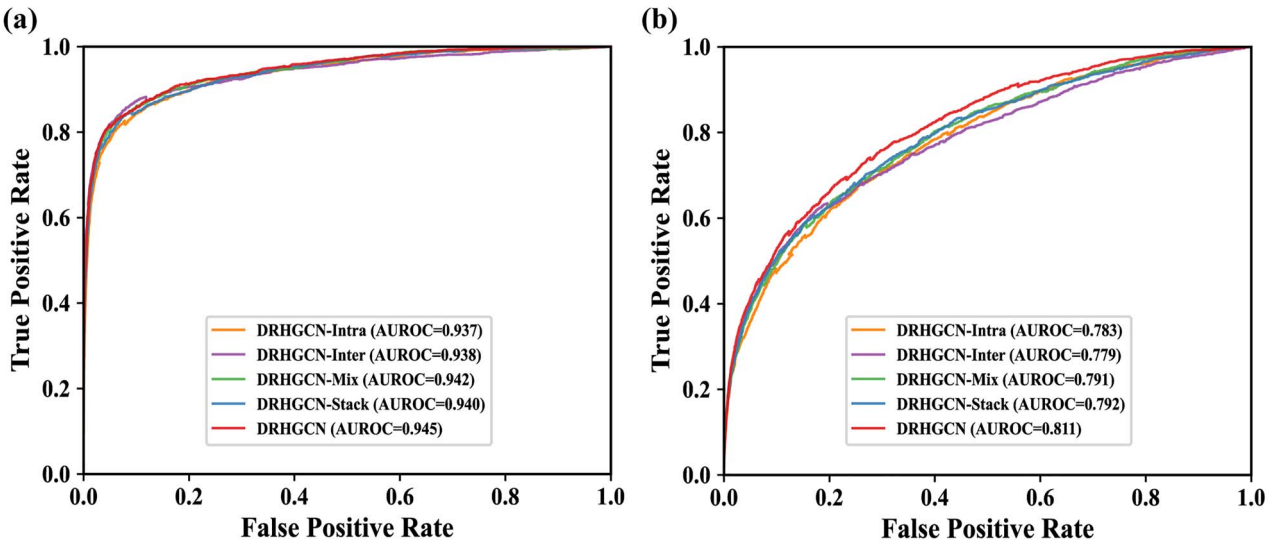

Figure 4. Performance of DRHGCN and various variants on Fdataset. (A) ROC curve of prediction results based on 10-fold cross-validation. (B) ROC curve of prediction results based on local LOOCV.

- DRHGCN-Intra: It is similar to DRHGCN-Inter, which mainly consists of a GCN-based encoder and a decoder. The GCN-based encoder only utilizes the intra-domain feature extraction module to extract drug and disease features. The decoder predicts the association probability score based on pairs of drug and disease embeddings (Supplementary Figure S12).
- DRHGCN-Mix: It ignores the differences between different networks and extracts drug and disease feature based on a big heterogeneous network which is constructed based on drug–disease associations, drug–drug similarities and disease–disease similarities (Supplementary Figure S13).
- DRHGCN-Stack: It takes the output of the intra-domain feature extraction module as the input of the inter-domain feature extraction module (Supplementary Figure S14).

Figure 4 shows the performance of DRHGCN and the various variants on Fdataset. In terms of 10-fold cross-validation, the performance of using the inter-domain and intra-domain feature extraction modules at the same time was better than that of using only a single module. This result shows that parallelly fusing heterogeneous information helps improve the prediction performance of DRHGCN. In addition, we verified the prediction capability of DRHGCN and the various variants in discovering candidate drugs for new diseases via local leave-one-out-cross-validation (LOOCV). We found that the performance of DRHGCN is better than that of the other variants. This finding indicates that DRHGCN using parallel aggregation not only preserves the inter-domain and intra-domain feature information to the greatest extent, but also eliminates the influence of differences in different networks.

Case study: Identified drugs for Alzheimer's disease and Parkinson's disease

To further verify the reliability capability of DRHGCN, we conducted detailed case studies on the computationally predicted candidate drugs for two neurodegenerative diseases, namely, Alzheimer's disease (AD) and Parkinson's disease (PD), which have no efficacious medications available yet. In the process of identifying potential drugs for AD and PD, we treated all the known drug–disease associations in the Fdataset as the training set and regarded the missing drug–disease associations as the candidate set. After the interaction probability of all candidate drug–disease associations are predicted by DRHGCN, we subsequently rank the candidate drugs by the predicted probabilities for each disease, so that the top-ranked associations are the most likely to interact.

Table 3. The top 10 DRHGCN-predicted candidate drugs for AD

Rank	Candidate drugs (DrugBank ID)	Pieces of evidence
1	Amantadine (DB00915)	CTD
2	Carbidopa (DB00190)	CTD
3	Ropinirole (DB00268)	CTD/DrugCentral
4	Haloperidol (DB00502)	CTD/ClinicalTrials
5	Scopolamine (DB00747)	CTD/ClinicalTrials
6	Benzatropine (DB00245)	NA
7	Pramipexole (DB00413)	CTD/ClinicalTrials
8	Tetrabenazine (DB04844)	[54]
9	Levocarnitine (DB00583)	PubChem
10	Levodopa (DB01235)	CTD/PubChem/ClinicalTrials

AD. We focused on the top 10 potential drugs predicted by DRHGCN (Table 3) and adopted highly reliable sources and clinical trials (i.e. DB, CTD [36], PubChem [55], DrugCentral [56] and ClinicalTrials) to check the predicted drug–disease associations. Amantadine was initially used for the prophylaxis and treatment of signs and symptoms of infection caused by various strains of influenza A virus. It was also used to treat PD and drug-induced extrapyramidal reactions. Herein, amantadine is the first predicted potential drug for treating AD. A previous study reported that amantadine significantly improved the mental state of two patients with AD [57]. Haloperidol, an effective first-generation antipsychotic drug and one of the most commonly used antipsychotics in the world, was predicted by DRHGCN to have an associate with AD. Such an association can be supported by CTD and ClinicalTrials. In addition, scopolamine

and pramipexole predicted by DRHGCN have also been confirmed by CTD and ClinicalTrials for AD treatment. In conclusion, among the top 10 predictive drugs based on the confidence score, nine drugs (90% success rate) have been verified by the reliable sources, clinical trials and other published studies.

PD. For PD, we focused on analyzing the top 10 candidate drugs predicted by DRHGCN. As shown in Table 4, we found that 10 out of 10 drugs (100% success rate) have been verified by the reliable sources and clinical trials. For example, rivastigmine, a parasympathomimetic or cholinergic drug that can be used to treat mild to moderate dementia, was predicted by DRHGCN to have an association with PD. This drug–disease association is recorded by DB, CTD, PubChem, DrugCentral and ClinicalTrials. Besides, memantine, an NMDA receptor antagonist, was predicted by DRHGCN to also have an effect on Alzheimer's disease. This prediction is supported by DB, PubChem and ClinicalTrials.

In addition to the analysis of the DRHGCN-predicted drugs, we also conducted molecular docking simulation experiments to demonstrate the practical application capability of DRHGCN. Taking AD as an example, we computed the molecular binding energies between the top 10 predicted drugs and five common target proteins (Supplementary Table S4) and characterized the ligand-protein binding mode between benzatropine and the target protein (Figure 5) through docking modeling, which is executed by using AutoDock Vina [58] and DS visualizer software. For the un-confirmed benzatropine-AD association, we utilized the Acetylcholinesterase (AChE, PDB code: 4EY5) as the target, the molecular binding energy between benzatropine and AChE is −9.0 kcal/mol. In addition, from Figure 5, we can observe that Var der Waals interactions exist between the small molecule and amino acid residues Tyr124, Tyr72, Phe297, Phe338, Phe295, Val294, Leu76, Gln291, Glu292 and Ser293. Moreover, there are some hydrophobic interactions such as pi–pi interaction between benzene ring and residues Trp286 and Tyr341, alkyl interactions between the compound and residue Leu289. The results imply that DRHGCN-predicted drug candidates have associations with AD. We expect that the predicted candidate drugs targeting AD will provide a meaningful reference to assist

Table 4. The top 10 DRHGCN-predicted candidate drugs for PD

Rank	Candidate drugs (DrugBank IDs)	Pieces of evidence
1	Rivastigmine (DB00989)	DB/CTD/PubChem/DrugCentral/ClinicalTrials
2	Trihexyphenidyl (DB00376)	DB/CTD/PubChem/DrugCentral
3	Biperiden (DB00810)	DB/CTD/PubChem/DrugCentral
4	Levodopa (DB01235)	DB/CTD/PubChem/DrugCentral/ClinicalTrials
5	Amantadine (DB00915)	DB/CTD/PubChem/DrugCentral/ClinicalTrials
6	Bromocriptine (DB01200)	DB/CTD/PubChem/DrugCentral/ClinicalTrials
7	Haloperidol (DB00502)	DB/PubChem/ClinicalTrials
8	Memantine (DB01043)	DB/PubChem/ClinicalTrials
9	Donepezil (DB00843)	DB/PubChem/ClinicalTrials
10	Vitamin E (DB00163)	CTD/PubChem/ClinicalTrial

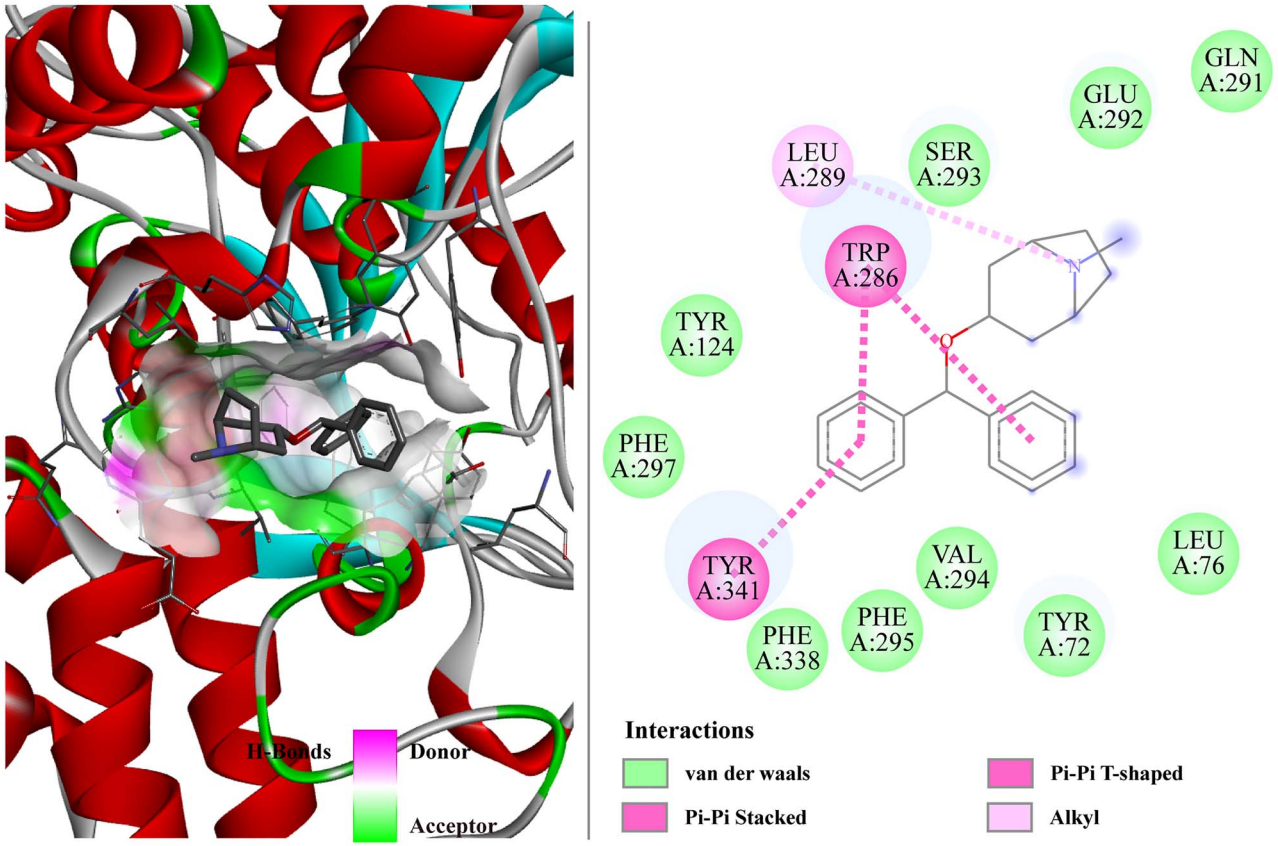

Figure 5. Docked Benzatropine (DrugBank ID: DB00245) with AChE (PDB code: 4EY5) as well as their interactions.

clinicians. In summary, DRHGCN provides a useful tool for predicting potential drugs for AD, PD.

Conclusion

We proposed a heterogeneous information-fusion, GCN-based, deep-learning methodology (called DRHGCN) to precisely discover candidate drugs for diseases. On account of network-specific topology information, we design intra- and inter-domain feature extraction modules to obtain intra- and inter-domain embeddings and parallelly fuse them to obtain the move representative embeddings of the drug and disease. Then, we introduce an attention mechanism to combine the embeddings from multiple graph convolutional layers. Extensive experiments demonstrated that DRHGCN is superior to current prediction methods and various variants of DRHGCN, exhibiting convincing performance in the rapid discovery of repurposable drugs for understudied diseases.

Although satisfactory results have been achieved from DRHGCN, there are still some limitations to this approach. First, despite the abundance of biological data, DRHGCN only uses drug–drug and disease–disease similarity from a single data source. The fusion of multiple data sources can tolerate noise effectively. In the future, we will consider more biological entities involved in drug–disease associations, such as genes, miRNAs and targets, and build a knowledge graph network with numerous entity types and links to assist in drug repositioning. Second, our proposed model is based on the assumption that similar diseases exhibit similar association patterns with drugs, but it could not be applied to all novel diseases. This is because we fail to obtain the features for novel diseases whose disease–disease associations are not available. In future work, we can collect more prior biological knowledge, such as disease phenotype-based similarity and disease semantic similarity to overcome this limitation.

In summary, DRHGCN that completely fuses the heterogeneous information of the drug–drug similarity, disease–disease similarity and observed drug–disease associations networks based on GCN can help pharmacologists or biologists effectively narrow down the search space of candidate drugs. It may further guide them to conduct wet-lab experiments and thus reduce costs and time. We expect that DRHGCN is useful for biological researchers and complementary to existing methods for accelerating drug repurposing and therapeutic development for understudied diseases.

Key Points

- We propose a heterogeneous information-fusion, graph convolutional network-based, fully end-to-end deep-learning methodology called DRHGCN for drug repositioning.
- To the best of our knowledge, the inter- and intra-domain feature extraction modules are the first proposed to fully learn the embedding of drugs and diseases by considering the different network topology information in different domains.
- We parallelly fuse the inter- and intra-domain embeddings to obtain the more representative embeddings of drug and disease;
- We introduce a layer attention mechanism to combine embeddings from multiple graph convolution layers for further improving the prediction performance.

Data Availability

The implementation of DRHGCN and the preprocessed data is available at: https://github.com/TheWall9/DRHGCN.

Supplementary data

Supplementary data are available online at *Briefings in Bioinformatics*.

Funding

The work was supported by the National Key Research and Development Program of China (Grant No. 2021YFE0102100); the Hunan Provincial Innovation Foundation for Postgraduate under (No. CX20200434); the Postdoctoral Science Foundation of China (No. 2020M672487); the National Natural Science Foundation of China (Grant No. 62006074).

References

1. Li J, Zheng S, Chen B, *et al*. A survey of current trends in computational drug repositioning. *Brief Bioinform* 2016;**17**:2–12.
2. Krantz A. Diversification of the drug discovery process. *Nat Biotechnol* 1998;**16**:1294–4.
3. Dickson M, Gagnon JP. The cost of new drug discovery and development. *Discov Med* 2009;**4**:172–9.
4. Chen H, Zhang H, Zhang Z, *et al*. Network-based inference methods for drug repositioning. *Comput Math Methods Med* 2015;**2015**:1–7.
5. Ye H, Liu Q, Wei J. Construction of drug network based on side effects and its application for drug repositioning. *PLoS One* 2014;**9**:e87864.
6. Zou J, Zheng M-W, Li G, *et al*. Advanced systems biology methods in drug discovery and translational biomedicine. *Biomed Res Int* 2013;**2013**:1–8.
7. Barratt MJ, Frail DE. *Drug repositioning: Bringing new life to shelved assets and existing drugs*. Hoboken, New Jersey, USA: John Wiley & Sons, Inc, 2012.
8. Xue H, Li J, Xie H, *et al*. Review of drug repositioning approaches and resources. *Int J Biol Sci* 2018;**14**:1232.
9. Setoain J, Franch M, Martínez M, *et al*. NFFinder: an online bioinformatics tool for searching similar transcriptomics experiments in the context of drug repositioning. *Nucleic Acids Res* 2015;**43**:W193–9.
10. Younis W, Thangamani S, Seleem MN. Repurposing non-antimicrobial drugs and clinical molecules to treat bacterial infections. *Curr Pharm Des* 2015;**21**:4106–11.
11. Li YY, Jones SJ. Drug repositioning for personalized medicine. *Genome Med* 2012;**4**:1–14.
12. Lotfi Shahreza M, Ghadiri N, Mousavi SR, *et al*. A review of network-based approaches to drug repositioning. *Brief Bioinform* 2018;**19**:878–92.
13. Lavecchia A, Cerchia C. In silico methods to address polypharmacology: current status, applications and future perspectives. *Drug Discov Today* 2016;**21**:288–98.
14. Zhou X, Dai E, Song Q, *et al*. In silico drug repositioning based on drug-miRNA associations. *Brief Bioinform* 2020;**21**:498–510.
15. Chen M, Zhang Y, Li A, *et al*. Bipartite heterogeneous network method based on co-neighbor for MiRNA-disease association prediction. *Front Genet* 2019;**10**:385.
16. Cao DS, Zhang LX, Tan GS, *et al*. Computational prediction of drug target interactions using chemical, biological, and network features. *Molecular informatics* 2014;**33**:669–81.
17. Chen X, Liu M-X, Yan G-Y. Drug–target interaction prediction by random walk on the heterogeneous network. *Mol Biosyst* 2012;**8**:1970–8.

18. Alaimo S, Pulvirenti A, Giugno R, *et al*. Drug–target interaction prediction through domain-tuned network-based inference. *Bioinformatics* 2013;**29**:2004–8.
19. Dai W, Liu X, Gao Y, *et al*. Matrix factorization-based prediction of novel drug indications by integrating genomic space. *Comput Math Methods Med* 2015;**2015**:1–9.
20. Luo H, Li M, Wang S, *et al*. Computational drug repositioning using low-rank matrix approximation and randomized algorithms. *Bioinformatics* 2018;**34**:1904–12.
21. Yang M, Luo H, Li Y, *et al*. Drug repositioning based on bounded nuclear norm regularization. *Bioinformatics* 2019;**35**:i455–63.
22. Yang M, Wu G, Zhao Q, *et al*. Computational drug repositioning based on multi-similarities bilinearmatrix factorization. *Brief Bioinform* 2020;**22**:1–14.
23. Zeng X, Zhu S, Liu X, *et al*. deepDR: a network-based deep learning approach to in silico drug repositioning. *Bioinformatics* 2019;**35**:5191–8.
24. Li Z, Huang Q, Chen X, *et al*. Identification of drug-disease associations using information of molecular structures and clinical symptoms via deep convolutional neural network. *Front Chem* 2020;**7**:924.
25. Kipf TN, Welling M. Semi-supervised classification with graph convolutional networks. In: International Conference on Learning Representations (ICLR), 2017. https://iclr.cc/archive/www/2017.html.
26. Huang Y-A, Hu P, Chan KC, *et al*. Graph convolution for predicting associations between miRNA and drug resistance. *Bioinformatics* 2020;**36**:851–8.
27. Li J, Zhang S, Liu T, *et al*. Neural inductive matrix completion with graph convolutional networks for miRNA-disease association prediction. *Bioinformatics* 2020;**36**:2538–46.
28. Zhao T, Hu Y, Valsdottir LR, *et al*. Identifying drug–target interactions based on graph convolutional network and deep neural network. *Brief Bioinform* 2021;**22**:2141–50.
29. Yu Z, Huang F, Zhao X, *et al*. Predicting drug–disease associations through layer attention graph convolutional network. *Brief Bioinform* 2020;**22**:1–11.
30. Gottlieb A, Stein GY, Ruppin E, *et al*. PREDICT: a method for inferring novel drug indications with application to personalized medicine. *Mol Syst Biol* 2011;**7**:496.
31. Luo H, Wang J, Li M, *et al*. Drug repositioning based on comprehensive similarity measures and bi-random walk algorithm. *Bioinformatics* 2016;**32**:2664–71.
32. Liang X, Zhang P, Yan L, *et al*. LRSSL: predict and interpret drug–disease associations based on data integration using sparse subspace learning. *Bioinformatics* 2017;**33**:1187–96.
33. Wishart DS, Knox C, Guo AC, *et al*. DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res* 2006;**34**:D668–72.
34. Hamosh A, Scott AF, Amberger J, *et al*. Online Mendelian inheritance in man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids Res* 2002;**30**:52–5.
35. Wang F, Zhang P, Cao N, *et al*. Exploring the associations between drug side-effects and therapeutic indications. *J Biomed Inform* 2014;**51**:15–23.
36. Davis AP, Grondin CJ, Johnson RJ, *et al*. The comparative toxicogenomics database: update 2017. *Nucleic Acids Res* 2017;**45**:D972–8.
37. Weininger D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J Chem Inf Comput Sci* 1988;**28**:31–6.
38. Steinbeck C, Han Y, Kuhn S, *et al*. The Chemistry Development Kit (CDK): an open-source Java library for chemo-and bioinformatics. *J Chem Inf Comput Sci* 2003;**43**:493–500.
39. Tanimoto T. An elementary mathematical theory of classification and prediction, IBM Report (November, 1958), cited in: G. Salton, *Automatic Information Organization and Retrieval*. McGraw-Hill New York, 1968.
40. Van Driel MA, Bruggeman J, Vriend G, *et al*. A text-mining analysis of the human phenome. *Eur J Hum Genet* 2006;**14**:535–42.
41. Nair V, Hinton GE. Rectified linear units improve restricted boltzmann machines. In: Proceedings of the 27th International Conference on Machine Learning (ICML-10). Haifa, Israel: Omnipress, 2010, 807–14.
42. Zhu H, Feng F, He X, *et al*. Bilinear graph neural network with neighbor interactions. In: Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI). Virtual, Japan: IJCAI, 2020, 1452–58.
43. Li G, Muller M, Thabet A, *et al*. Deepgcns: Can gcns go as deep as cnns? In: Proceedings of the IEEE International Conference on Computer Vision. Seoul, Korea (South): IEEE, 2019, 9266–75.
44. Li Q, Han Z, Wu X-M. Deeper insights into graph convolutional networks for semi-supervised learning. In: Proceedings of the AAAI Conference on Artificial Intelligence. New Orleans, Louisiana, USA: AAAI, 2018, 3538–45.
45. He K, Zhang X, Ren S, *et al*. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. Washington, DC, USA: IEEE, 2016, 770–8.
46. Kingma DP, Ba J. Adam: a method for stochastic optimization. In: International Conference on Learning Representations (ICLR), 2015. https://iclr.cc/archive/www/2015.html.
47. Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks. In: Proceedings of the thirteenth international conference on artificial intelligence and statistics. Sardinia, Italy: JMLR, 2010, 249–56.
48. Srivastava N, Hinton G, Krizhevsky A, *et al*. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research* 2014;**15**:1929–58.
49. Zhang W, Xu H, Li X, *et al*. DRIMC: an improved drug repositioning approach using Bayesian inductive matrix completion. *Bioinformatics* 2020;**36**:2839–47.
50. Zhang W, Yue X, Lin W, *et al*. Predicting drug-disease associations by using similarity constrained matrix factorization. *BMC bioinformatics* 2018;**19**:1–12.
51. Liu Y, Wu M, Miao C, *et al*. Neighborhood regularized logistic matrix factorization for drug-target interaction prediction. *PLoS Comput Biol* 2016;**12**:e1004760.
52. Zhang Y, Chen M, Cheng X, *et al*. MSFSP: a novel miRNA–disease association prediction model by federating multiple-similarities fusion and space projection. *Front Genet* 2020;**11**:389.
53. Zhang Y, Chen M, Li A, *et al*. LDAI-ISPS: LncRNA–disease associations inference based on integrated space projection scores. *Int J Mol Sci* 2020;**21**:1508.
54. Kilbourn MR, DaSilva JN, Frey KA, *et al*. In vivo imaging of vesicular monoamine transporters in human brain using [11C] tetrabenazine and positron emission tomography. *J Neurochem* 1993;**60**:2315–8.
55. Kim S, Thiessen PA, Bolton EE, *et al*. PubChem substance and compound databases. *Nucleic Acids Res* 2016;**44**:D1202–13.

56. Avram S, Bologa CG, Holmes J, *et al*. DrugCentral 2021 supports drug discovery and repositioning. *Nucleic Acids Res* 2021;**49**:D1160–9.
57. Erkulwater S, Pillai R. Amantadine and the end-stage dementia of Alzheimer's type. *South Med J* 1989;**82**:550–4.
58. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* 2010;**31**:455–61.