frontiers | Frontiers in Pharmacology

BRIEF RESEARCH REPORT
published: 06 July 2022
doi: 10.3389/fphar.2022.908549

An Integrative Heterogeneous Graph Neural Network–Based Method for Multi-Labeled Drug Repurposing

Shaghayegh Sadeghi, Jianguo Lu and Alioune Ngom*

School of Computer Science, University of Windsor, Windsor, ON, Canada

Drug repurposing is the process of discovering new indications (i.e., diseases or conditions) for already approved drugs. Many computational methods have been proposed for predicting new associations between drugs and diseases. In this article, we proposed a new method, called DR-HGNN, an integrative heterogeneous graph neural network-based method for multi-labeled drug repurposing, to discover new indications for existing drugs. For this purpose, we first used the DTINet dataset to construct a heterogeneous drug–protein–disease (DPD) network, which is a graph composed of four types of nodes (drugs, proteins, diseases, and drug side effects) and eight types of edges. Second, we labeled each drug–protein edge, $dp_{i,j} = (d_i, p_j)$, of the DPD network with a set of diseases, $\{\delta_{i,j,1}, \ldots, \delta_{i,j,k}\}$ associated with both d_i and p_j and then devised multi-label ranking approaches which incorporate neural network architecture that operates on the heterogeneous graph-structured data and which leverages both the interaction patterns and the features of drug and protein nodes. We used a derivative of the GraphSAGE algorithm, HinSAGE, on the heterogeneous DPD network to learn low-dimensional vector representation of features of drugs and proteins. Finally, we used the drug–protein network to learn the embeddings of the drug–protein edges and then predict the disease labels that act as bridges between drugs and proteins. The proposed method shows better results than existing methods applied to the DTINet dataset, with an AUC of 0.964.

Keywords: computational drug repurposing, graph embedding, graphsage, data integration, link prediction, graph neural network

OPEN ACCESS

Edited by:
Xiujuan Lei,
Shaanxi Normal University, China

Reviewed by:
Junliang Shang,
Qufu Normal University, China
Karim Abbasi,
Sharif University of Technology, Iran

****Correspondence:***
Shaghayegh Sadeghi
sadeghi3@uwindsor.ca

Specialty section:
This article was submitted to Experimental Pharmacology and Drug Discovery, a section of the journal Frontiers in Pharmacology

Received: *30 March 2022*
Accepted: *09 May 2022*
Published: *06 July 2022*

Citation:
Sadeghi S, Lu J and Ngom A (2022) An Integrative Heterogeneous Graph Neural Network–Based Method for Multi-Labeled Drug Repurposing. Front. Pharmacol. 13:908549. doi: 10.3389/fphar.2022.908549

1 INTRODUCTION

Drug repurposing (DR) is a process of identifying novel therapeutic purposes for existing drugs. Over the years, computational drug repurposing (CDR), known as *in silico* drug repurposing, has gained considerable popularity in the pharmaceutical industry due to its time and cost efficiency in the drug development process compared to the traditional *de novo* drug discovery process. Drug repurposing can be a promising treatment strategy for a lot of health crises such as COVID-19 since it can shorten the drug development process with much less funding (Sadeghi et al., 2021; Su et al., 2021). In recent years, different computational approaches are suggested for repurposing drugs based on machine learning, network analysis, and text mining (Li et al., 2016). Since network-based methods are capable of using ever-increasing large-scale

biological datasets such as genetic, pharmacogenomics, clinical, and chemical data, they are more desirable for drug repurposing tasks (Sadeghi and Keyvanpour, 2019).

With the recent advances in deep learning methods on graphs due to their promising ability to capture complex and highly non-linear network structures, graph neural network (GNN) method usage on biological networks seems more interesting than ever (Pan et al., 2022), (Yu J.-L. et al., 2021). For example, Yu Z. et al. (2021) proposed a layer attention graph convolutional network (LAGCN) for the drug–disease association prediction. The LAGCN utilizes a GCN to capture structural information from the heterogeneous network composed of drug–disease associations, drug–drug similarities, and disease–disease similarities. The attention mechanism is introduced to combine the embeddings from different convolution layers, which leads to a more informative representation of drugs and diseases. Wang et al. (2020) proposed BiFusion, a bipartite GCN model for CDR through heterogeneous information fusion. BiFusion combines insights from multiscale pharmaceutical information by constructing a multi-relational graph of drug–protein, disease–protein, and protein–protein networks. (Cai et al. (2021) proposed a method, called DRHGCN, based on the heterogeneous information fusion graph convolutional network. deepDR, on the other hand, uses a variational auto-encoder (VAE) to infer candidates for repurposing (Zeng et al., 2019). Zhao et al. (2021) proposed a method called multi-graph representation learning (MGRL), which first uses the graph convolution network to learn the graph representation of drugs and diseases. Then, the graph embedding algorithm represents the relationships between drugs and diseases. Finally, the two kinds of graph representation learning features were put into the random forest classifier for training. The drug repositioning method based on heterogeneous networks and text mining (HeTDR) proposed by Jin et al. (2021) fuses network topology information and text mining information to gain and predicts potential drug-disease associations by an embedding learning method.

The main difference between these aforementioned deep learning-based methods for CDR tasks is that they use different types of network inputs or add extra features and also different GCN structures as decoders. Hence, one way to expand these methods is to include additional biological network types in the equation of the DR task. However, creating the base heterogeneous network for CDR is a challenging task.

This study casted the CDR problem as a link prediction task and proposed DR-HGNN, a novel approach for inferring new drug-disease associations (i.e., new links between drugs and diseases). The main idea is to create a multi-label heterogeneous drug–protein–disease (DPD) network as input for the heterogeneous variation of the GraphSAGE algorithm.

First, DR-HGNN integrates six heterogeneous networks and four homogeneous networks for creating drug and protein side information, which can potentially improve the performance of CDR. Second, DR-HGNN creates a DPD network in which, for each drug and protein in the drug–target interaction (DTI) network, we assume there is at least one disease that connects these two. In other words, we used diseases as our labels in the DTI network. However, this leads to a multi-label problem which means that there can be more than one disease as a bridge for each DTI. Hence, in the third step, we solved this problem with a transformation-based solution. Later, we used a generalized version of GraphSAGE for heterogeneous networks, called HinSAGE. HinSAGE processes the input DPD network for embedding each drug and protein node. Finally, an edge embedding layer will be used for predicting new disease edges between drugs and proteins. This edge embedding scores each edge between drug and proteins and the disease label associated with this edge. DR-HGNN shows a high predictive performance when compared to existing CDR methods.

2 METHODS

Algorithm 1. DR-HGNN

Input : Drug-protein interactions (G_{ds}), drug-drug interactions(G_{dd}), drug-disease associations ($G_{d\delta}$), drug–side-effect associations (G_{ds}), drug-drug similarity (G_d), protein–disease associations ($G_{p\delta}$), protein-protein similarity (G_p), and protein-protein interactions (G_{pp})
Output : Embedding scores for each $\{d_i, p_j, \delta_{k'}\}$ in DPD graph

$F_d \leftarrow$ CompactFeatureLearning (G_{dd}, G_d, $G_{d\delta}$, G_{ds});
$F_p \leftarrow$ CompactFeatureLearning(G_{pp}, G_p, $G_{p\delta}$);
$G_{dp\delta} \leftarrow$ merge(G_{dp}, $G_{p\delta}$);
foreach $\{d_i, p_j, \{\delta_{i,j,1}, \ldots, \delta_{i,j,k}\}\} \in G_{dp\delta}$ **do**
  $\{d_i, p_j, \delta_{k'}\} \leftarrow$ SelectRepresentative($\{d_i, p_j, \{\delta_{i,j,1}, \ldots, \delta_{i,j,k}\}\}$);
$E_{d_i}, E_{p_j} \leftarrow$ HinSAGE(DPD, F_p, F_d) ; /* E_{p_j} is embedding of protein nodes and E_{d_j} drug nodes */
Embedding of $\{d_i, p_j, \delta_{k'}\}$ edges $\leftarrow$ LinkEmbedding(E_{d_i},E_{p_j});

2.1 Construction of the DPD Graph

We constructed a schema of a DPD heterogeneous graph (**Figure 1A**). Diseases in this graph are bridges between drugs and proteins; they are labels on edges connecting drugs and proteins. This graph is constructed from three different heterogeneous sub-networks (i.e., drug–protein interaction, protein–disease association, and drug–disease association).

Each edge in this graph connects drugs and proteins and has diseases as their label. This means we can have a triple of (drugs, proteins, and diseases) for each link $\{d_i, p_j, \{\delta_{i,j,1}, \ldots, \delta_{i,j,k}\}\}$. Each drug and protein has a feature vector constructed by a compact feature learning method (Luo et al., 2017). The input for the compact feature learning method is different types of sub-networks (i.e., drug–protein interactions (G_{ds}), drug–drug interactions (G_{dd}), drug–disease associations ($G_{d\delta}$), drug–side-effect associations (G_{ds}), protein–disease associations ($G_{p\delta}$), protein–protein similarity (G_p), drug–drug similarity (G_d), and protein–protein interactions (G_{pp})) that have side information from a different view of each entity (i.e., drugs and proteins). The output of the compact feature learning method is a matrix representation of the entity (i.e., drugs and proteins) features. The compact feature learning method integrates diverse information from the heterogeneous network by foremost combination of the network diffusion algorithm (random walk with restart) with a dimensionality reduction technique (diffusion component analysis) to obtain informative but low-dimensional vector representations of nodes in the network (Luo et al., 2017).

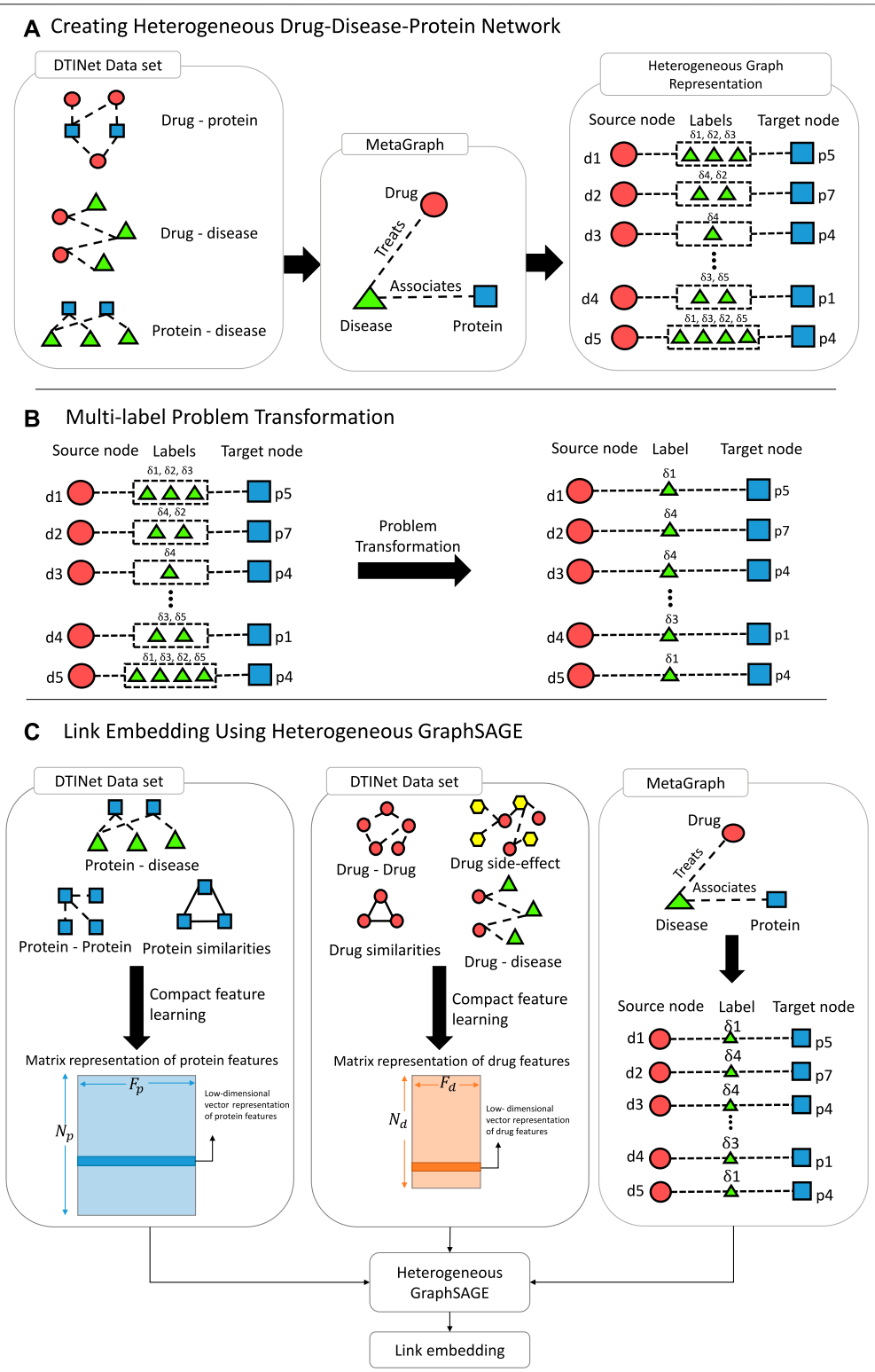

FIGURE 1 | Pipeline of DR-HGNN. **(A)** Creating Heterogeneous Drug-Disease-Protein Network: using the DTINet dataset, a meta-graph is created, which can be presented as a Heterogeneous Graph (on the right). **(B)** Multi-label Problem Transformation: A problem transformation technique is used since the Heterogeneous Graph from step A is multi-labeled. **(C)** Link Embedding Using Heterogeneous GraphSAGE: With matrix representation of each protein and drug and the heterogeneous graph from step B, Heterogeneous GraphSAGE embeds links between nodes of this heterogeneous graph.

TABLE 1 | Comparison of GNN methods.

Method	Handle bipartite graph	Handle heterogeneous graph	Handle different node feature sizes
HinSAGE	Yes	Yes	Yes
GraphSAGE	Yes	No	Yes
GCN	No	No	No
GAT	Yes	No	Yes

Finally, based on drug–protein and protein–disease associations, we labeled the edges the in drug–protein interaction network to create our DPD graph ($G_{dp\delta}$).

$$G_{dp\delta} \leftarrow merge\left(G_{dp}, G_{p\delta}\right), \tag{1}$$

where merge joins adjacency matrices of G_{dp} and $G_{p\delta}$. Here, the joining is carried out on protein names. The $G_{dp\delta}$ graph constructed here can be presented as the triple of (drugs, proteins, and diseases) $\{d_i, p_j, \{\delta_{i,j,1}, \ldots, \delta_{i,j,k}\}\}$.

2.2 Multi-Label Problem Transformation

The constructed $G_{dp\delta}$ has multi-label edges (**Figure 1B**). Not all labels are equally important to the characterization of the edges. Hence, we need a multi-label ranking approach to choose just one of the labels as the labels' representative. One-hot encoding of labels in the $G_{dp\delta}$ graph can be sparse and large. Hence, instead of using a neural network for dealing with our multi-labeled edges, we proposed a method as a prepossessing step to transform our multi-label problem into a single-label problem (Tsoumakas et al., 2009). This method transforms the multi-label learning task into a multi-class or single-label classification task. In other words, LP models the joint distribution of labels. It treats each label subset in the multi-label training set as a class of a multi-class task, and the prediction will be one of these subsets (Chu et al., 2021). For this purpose, for each set of labels for each pair of drug–protein, we used one of the labels as the representative of that set.

For selecting this representative label, we selected the label with less frequency among all sets of labels and more mutual information. This representative label is more informative than other labels since these labels have a more distinctive ability based on the law of frequency. This new DPD network is a compressed version of $G_{dp\delta}$, and every link in this graph can be also presented as unique $\{d_i, p_j, \delta_{k'}\}$. For this purpose, first, we counted the frequency of each disease in the $G_{dp\delta}$ graph and then we selected one disease for each pair of drugs and proteins that has the least appearance in the network. This disease is the new label for the drug–protein association.

2.3 Edge Embedding Using Heterogeneous GraphSAGE

Standard message passing GNNs cannot trivially be applied to heterogeneous graph data as the same functions cannot process node and edge features from different types due to differences in the feature type and size (Fey and Lenssen, 2019). To avoid this problem, here, we used a generalized version of the GraphSAGE algorithm (Hamilton et al., 2017) for heterogeneous graphs called HinSAGE (Data61, 2018).

Looking at **Table 1**, HinSAGE can provide us with the features we want from our GNN, while other methods such as the graph attention network (GAT), graph convolutional network (GCN), and GraphSAGE cannot be performed on heterogeneous networks without implementing message and update functions individually for each edge type.

HinSAGE separate neighborhood weight matrices (W_{neigh}'s) for every unique ordered tuple of (N_1, E, N_2) where N_1 and N_2 are node types (here, N_1 is for drugs, and N_2 is for proteins), and E is an edge type (here, E is for disease) to support heterogeneity of nodes and edges. HinSAGE also will distinct self-feature matrices W_{self} for every node type, where W_{self} is a unique self-edge type for every node type.

As for feature update rules, aggregation (mean) of features from the neighbors of node v *via* edges of type r is being used:

$$h^k_{N_r(v)} = \frac{1}{N_r(v)} \sum_{u \in N_r(v)} D_p\left[h^{k-1}_u\right]. \tag{2}$$

(Data61, 2018).

Meanwhile, forward pass through layer k is as follows:

$$h^k_v = \sigma\left(W^k_{t_v,\text{self}} D_p\left[h^{k-1}_v\right] + W^k_{r,\text{self}} h^k_{N_r(v)} + b^k\right). \tag{3}$$

(Data61, 2018).

Here, $W^k_{t_v,\text{self}}$ is the weight matrix for self-edges for node type t_v and is of shape $d_k \times d_{k-1}$. Also, $W^k_{r,\text{self}}$ is the weight matrix for edges of type r and is of shape $\frac{d_k}{2} \times d_{k-1}(r)$. D_p [.] is a random dropout with probability p applied to its argument vector. σ is the nonlinear activation. b_k is an optional bias, while h^k_v is the output for node v at layer k. r indicates the edge type from node v to node u (r is defined as the unique tuple (t_v, t_e, t_u)), where t_v indicates the type of node v, and t_e indicates the relation type. $N_r(v)$ is a neighbor of the node v *via* the edge type r. $d_{k-1}(r) = \dim(h^k_{N_r(v)})$ is the dimensionality of $(k-1)$-th layer's features of node v's neighbors *via* edge type r. The number of trainable parameters per layer k for this model is as follows:

$$T_v d_k d_{k-1} + R_e d_k d_{k-1} + d_k = (T_v + R_e) d_k d_{k-1} + d_k, \tag{4}$$

(Data61, 2018).

supposing that the dimensionalities of all destination node features for all edge types r are all equivalent, that is, $d_{k-1}(r) = d_{k-1} \forall r$, the number of all node types in the graph is T_v, and the number of all edge types is R_e.

The HinSAGE algorithm requires two types of input: node features and an adjacency matrix of the heterogeneous graph. Here, drugs and proteins are our nodes, and we used the compact feature learning method to obtain their feature vectors. Compact feature learning is a random walk-based algorithm. First, a random walk algorithm with

TABLE 2 | Number of nodes and edges of individual types in the constructed heterogeneous network on DTINet (Luo et al., 2017).

Node	Number of edges			
	Drug	Protein	Disease	Side effect
Drug	10, 036	1, 923	199, 214	80, 164
Protein	1, 923	7, 363	1, 596, 745	–
Number of nodes	708	1, 512	5, 603	4, 192

restart (RWR) is used to compute the diffusion states of individual networks. Then, the low-dimensional representations of feature vectors for each node are obtained by minimizing the difference between the diffusion states s_i and the parameterized multinomial logistic models $\hat{s}_i$. The low-dimensional feature vectors obtained from the previous step encode the relational properties (for example, similarity), association information, and topological context of each node in the heterogeneous network (Luo et al., 2017). As for the adjacency matrix of the heterogeneous graph, we constructed a DPD graph in **Section 2.1**.

After having embeddings of each node, we can use a function that predicts a multi-class edge classification output from (source: drug; destination: protein) node embeddings (node features). For this purpose, this layer combines (source: drug; destination: protein) new embeddings from HinSAGE layers into edge embeddings.

3 METHODOLOGY

This section demonstrates the efficacy and efficiency of the DR-HGNN frameworks for DR tasks and compares them with three state-of-the-art DR algorithms. Five aspects are discussed in the following four subsections: datasets in both our proposed model and the competing models, experiment setting, results of competing methods, and case studies of our proposed model.

3.1 Material and Data

In this study, the DTINet data set from Luo et al. (2017) is used. DTINet is a heterogeneous network with 12,015 nodes and 1,895,445 edges in total and is originally constructed for predicting missing DTI (drug–target interaction) edges. Luo et al. (2017) compiled various curated public drug-related databases (DrugBank (Wishart et al., 2006), the Comparative Toxicogenomics Database (CTD) (Davis et al., 2019), the Human Protein Reference Database (HPRD) (Keshava Prasad et al., 2009), and Side Effect Resource (SIDER) (Kuhn et al., 2010)) (**Table 2**) to create DTINet. The DTINet network integrates four types of nodes (that is, drugs, proteins, diseases, and drug side effects) and six types of edges (that is, drug–protein interactions, drug–drug interactions, drug–disease associations, drug–side effect associations, protein–disease associations, and protein–protein interactions).

Based on chemical structures of drugs and primary sequences of proteins, they also built multiple similarity networks to further augment the network heterogeneity, providing diverse information from a multiple-view perspective. The heterogeneous DPD graph has only 1,923 triples of $\{d_i, p_j, \delta_k\}$ constructed by the DTINet data set. As shown in **Table 2**, the DPD graph has 708 drug nodes and 1,512 protein nodes with 39 diseases, representing 5,603 diseases selected in **Section 2.2**.

3.2 Experimental Setup

CDR can be cast as a link prediction problem, and here, in this study, we predicted the edges between drugs and proteins with the diseases as their label. To evaluate the prediction performance of the DR-HGNN model and the competing methods, we used 5-fold cross-validation (5-CV) since other baseline methods also used 5-CV. We added a matching number of non-interacting triples to the known interacting drug–target–disease triples (DPD graph). Then, data sets were five times shuffled to form five randomly ordered data sets, each of which was divided into training (60%), validation (20%), and test sets (20%) (Luo et al., 2017; Moon et al., 2021).

In experiments, the area under the receiver operating characteristic curve (AUC-ROC) and precision-recall curve (AUPR) are used to measure the performance of results. AUC-ROC and AUPR, as useful measures of accuracy, have been considered, with a meaningful AUC interpretation usually representing the overall performance of the method (Sadeghi and Keyvanpour, 2019). We compared our approach with five other DR methods in the following:

- **DTINet** (Luo et al., 2017), is a low-dimensional vector representation-based method that extracts features from the topology of the nodes in the integrated network and predicts and computes drug–protein target interactions and drug similarity measures through these representations.
- **NMTF** (Ceddia et al., 2020), is a negative matrix factorization-based method that imposes a non-negative constraint on the factorized matrices during multiplication and update operations.
- **LAGCN** (Yu Z. et al., 2021), is a layer attention graph convolutional network-based method for the drug–disease association prediction.
- **deepDR** (Zeng et al., 2019), is an autoencoder-based method for fusing the features and mining new drug disease associations.
- **KBMF** (Gönen et al., 2013), is a kernelized bayesian matrix factorization method that can make use of multiple side information sources and can be applied in recommender systems.

3.3 Performance Comparison

DR-HGNN's results outperform all five methods on the DTINet data set. The parameters (that is, learning rate, dropout, optimizer function, number of layers, and embedding dimensions) in these methods are set to either their optimal values or recommended values reported in the original works. **Figure 2** reports the AUC-ROC of all compared methods on the DTINet data. As shown in **Figure 2**, DR-HGNN outperforms other methods with 0.964 and 0.93 for its AUC-ROC and AUC-PR. An AUC-ROC and a loss history plot of DR-HGNN can also be seen in **Figure 4** for both the training and validation datasets. **Figure 2** also shows LAGCN with 0.94 for AUC-ROC, and 0.92 for AUC-PR is at the second place. deepDR also has the same AUC-PR value as LAGCN, but its AUC-ROC is around 0.91. NMTF as a matrix factorization-based method comes after neural network-based methods with 0.93 and 0.86 for its AUC-ROC and AUC-PR,

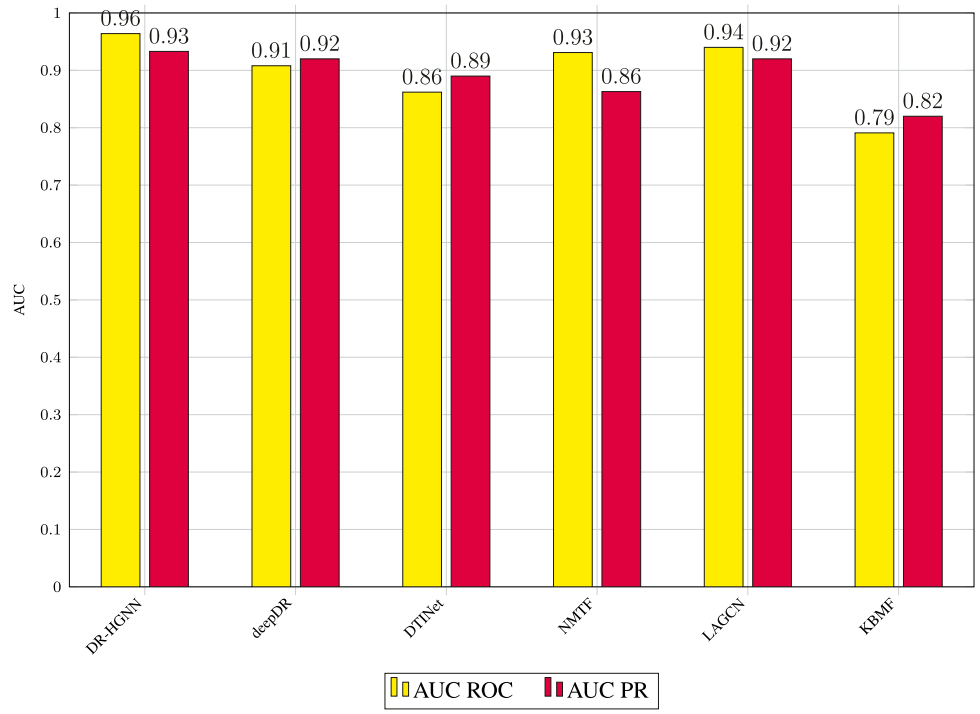

FIGURE 2 | AUC ROC and AUC PR values of prediction results obtained by applying DR-HGNN and other reported methods in 5-fold cross-validation.

TABLE 3 | Results of DR-HGNN on the TL-HBGI dataset (Wang et al., 2014).

Method	AUC	AUPR
TL-HGBI (Wang et al., 2014)	0.95	0.0492
NMF-DR (Sadeghi et al., 2021)	**0.9902**	0.4200
SCMFDD (Zhang et al., 2018)	0.97	0.1500
NTSIM (Zhang et al., 2017)	0.96	0.2631
DR-HGNN	0.9895	**0.4560**

TABLE 4 | AUC ROC results for DR-HGNN based on different parameters.

Adam optimizer		Learning rate				
		1.00E+00	1.00E-01	1.00E-02	1.00E-03	1.00E-04
Dropout	0	0.8245	**0.9566**	0.9487	0.9639	0.8734
	0.1	0.8678	0.9165	0.954	**0.9647**	**0.8868**
	0.2	0.8605	0.9365	**0.9635**	0.8903	0.8778
	0.3	0.9333	0.916	0.9594	0.9167	0.8317
	0.4	0.898	0.9307	0.8955	0.937	0.827
	0.5	0.9492	0.9519	0.8986	0.9606	0.8319
	0.6	0.888	0.9469	0.892	0.9068	0.8687
	0.7	0.8682	0.8923	0.8682	0.961	0.7054
	0.8	**0.9501**	0.93	0.9106	0.9448	0.6968
	0.9	0.8922	0.9457	0.9357	0.9298	0.8472

respectively. Another matrix factorization-based method (KBMF) also has an AUC-ROC of 0.79 and AUC-PR of 0.82.

To illustrate the potential generalization of DR-HGNN, we evaluated another well-known benchmark dataset called the TL-HGBI dataset (Wang et al., 2014) with a 5-fold cross-validation (**Table 3**). The TL-HBGI dataset (Wang et al., 2014) has 1,409 drugs registered by the DrugBank database, 5,080 diseases listed by the OMIM database, and 1,461 known relationships. Drug–drug similarities were calculated based on their chemical structures, and a phenotype-based disease–disease similarity dataset was downloaded from MimMine. **Table 3** shows that DR-HGNN outperforms other methods in both AUC and AUPR metrics. The results show that our approach can also compete with other methods with the AUC-ROC measure and the AUC-PR measure.

3.4 Impact of Parameter Settings

We adjusted the parameters to achieve optimal performances. We showed the effect of using different learning rate parameters and dropout for the Adam optimizer in **Table 4** with 50 epochs. Based on this experiment, the optimal learning rate and dropout for the Adam optimizer with L_2 regularization was for 0.001 learning rate and 0.1 dropout. We also experimented with the size of the embedding. To investigate the effect of layer numbers on model performance, we compared results with a different number of layers in DR-HGNN on the DTINet dataset. **Figure 3** showed the model performance along with the increase in layer numbers and embedding dimension.

We observed that one layer has the lowest performance, suggesting that a shallow network cannot sufficiently propagate the node feature to fuse heterogeneous information, especially for the complex DPD network. Moreover, we found that DR-HGNN achieved significant improvement with two layers' structure. But with more than two layers, the model performance tends to decrease. We believe that GraphSAGE behaves similarly to graph convolutional

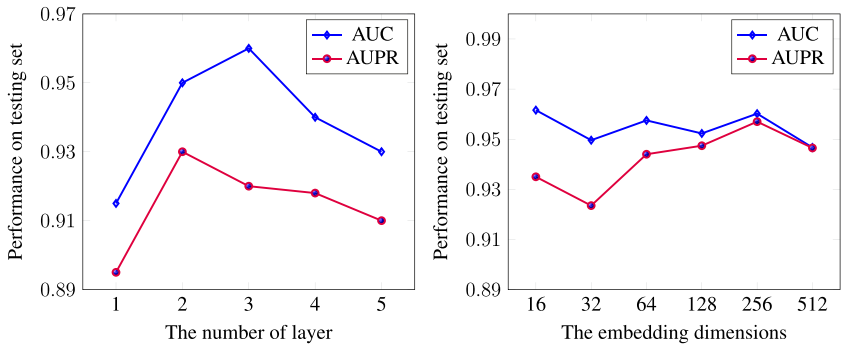

FIGURE 3 | Impact of the number of layers and embedding dimension on the model performance.

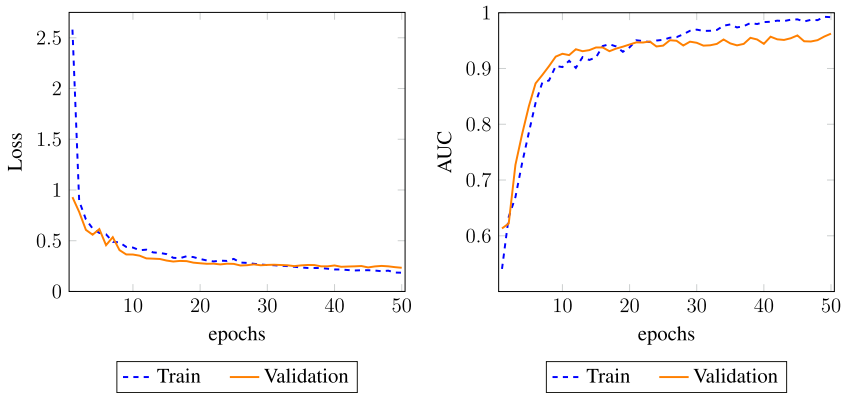

FIGURE 4 | AUC and loss history plot for DR-HGNN on each epoch for the training and validation datasets.

networks (GCNs). A shallow GraphSAGE (1-layer) may not learn sufficient information, and more layers could lead to an over-smoothing issue. Other related works on the GNN also show that two layers are usually enough for capturing the knowledge of the network (Niu et al., 2021; Wang et al., 2020; Li et al., 2019; Chen et al., 2020). **Figure 4** also shows the effect of embedding dimension on the DR-HGNN performance. Based on the results of this experiment, we chose the embedding dimension of 256 since the model has superior AUC-ROC and AUC-PR performance with this dimension size.

3.5 Case Study

In this section, we conducted case studies to evaluate the capability of DR-HGNN in predicting novel drug–disease associations. The relationships between drugs and diseases in the DTINet dataset not only have therapeutic ones but also have drug side effects. Thus, the model predicts both types of relations. Here, we discussed examples of both potential therapeutic relationships and potential side effects.

For verification of the prediction, along with a manual PubMed search, we have examined a publicly available Web server named ChemoText (Capuzzi et al., 2018). For example, hypertension (high blood pressure) has associations with the protein NADH dehydrogenase, subunit 1 (complex I) (UniprotId: P03886). Halothane (DrugBankID: DB01159), a general inhalation anesthetic used for the induction and maintenance of general anesthesia, has also been interacting with P03886. Chemotext and also Pubmed search also validate our connection between hypertension and halothane (Enderby, 1960). The other relationships we validated are between desflurane (DrugBank ID: DB01189), sevoflurane (DrugBank ID: DB01236), and pregabalin (DrugBank ID: DB00230) with hypertension and the shared protein of P03886. As for side effects, we validated nicotine (DrugBank ID: DB00184) associations. We found exciting relationships between nicotine with hyperalgesia. Nicotine is short-term pain relief; however, over time, it may increase pain intensity (Ditre et al., 2018). Nicotine also has an association with diabetic nephropathies and pregnancy complications through proteins P30926 and P36544, respectively.

4 DISCUSSION AND CONCLUSION

In this study, we presented a framework based on the GNN for drug repurposing. We created a heterogeneous drug–disease–protein network using multi-label problem transformation as input for heterogeneous GraphSAGE for repurposing drugs. Although we obtained satisfactory results, DR-HGNN has some limitations. First, we used several networks to create a heterogeneous drug repurposing network. However, in the future, we plan to consider

more networks such as miRNA and genes to make a richer heterogeneous graph. Second, having multi-labeled edges in the drug–disease–protein network in the CDR task is a challenge that should be addressed. DR-HGNN uses the problem transformation approach for handling multi-label edges. MLC in drug repurposing has other challenges, such as label size imbalance. We can propose and use different solutions for this challenge in future work. All in all, DR-HGNN has the potential to be used for predicting edges in other biomedical networks, such as the drug–target interaction.

DATA AVAILABILITY STATEMENT

The datasets analyzed for this study can be found in the DTINet GitHub repository (https://github.com/luoyunan/DTINet). The code for this study is available from the DR-HGNN repository (https://github.com/sshaghayeghs/DR_HGNN).

AUTHOR CONTRIBUTIONS

SS, JL, and AN contributed to the conception and design of the study. SS organized the database. SS performed the statistical analysis. SS wrote the first draft of the manuscript. SS, JL, and AN wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This research is supported by the National Science and Engineering Research Council of Canada (NSERC) (NSERC RGPIN-2016-05017 and NSERC RGPIN-2019-05350).

REFERENCES

Cai, L., Lu, C., Xu, J., Meng, Y., Wang, P., Fu, X., et al. (2021). Drug Repositioning Based on the Heterogeneous Information Fusion Graph Convolutional Network. *Briefings Bioinforma.*

Capuzzi, S. J., Thornton, T. E., Liu, K., Baker, N., Lam, W. I., and O'Banion, C. P. (2018). Chemotext: a Publicly Available Web Server for Mining Drug–Target–Disease Relationships in Pubmed. *J. Chem. Inf. Model.* 58, 212–218. doi:10.1021/acs.jcim.7b00589

Ceddia, G., Pinoli, P., Ceri, S., and Masseroli, M. (2020). Matrix Factorization-Based Technique for Drug Repurposing Predictions. *IEEE J. Biomed. health Inf.* 24, 3162–3172. doi:10.1109/JBHI.2020.2991763

Chen, D., Lin, Y., Li, W., Li, P., Zhou, J., and Sun, X. (2020). Measuring and Relieving the Over-smoothing Problem for Graph Neural Networks from the Topological View. *Proc. AAAI Conf. Artif. Intell.* 34, 3438–3445. doi:10.1609/aaai.v34i04.5747

Chu, Y., Shan, X., Chen, T., Jiang, M., Wang, Y., Wang, Q., et al. (2021). Dti-mlcd: Predicting Drug-Target Interactions Using Multi-Label Learning with Community Detection Method. *Briefings Bioinforma.* 22, bbaa205. doi:10.1093/bib/bbaa205

Data61, C. (2018). Stellargraph Machine Learning Library. *GitHub Repository*. Available at: https://github.com/stellargraph/stellargraph.

Davis, A. P., Grondin, C. J., Johnson, R. J., Sciaky, D., McMorran, R., Wiegers, J., et al. (2019). The Comparative Toxicogenomics Database: Update 2019. *Nucleic acids Res.* 47, D948–D954. doi:10.1093/nar/gky868

Ditre, J. W., Zale, E. L., LaRowe, L. R., Kosiba, J. D., and De Vita, M. J. (2018). Nicotine Deprivation Increases Pain Intensity, Neurogenic Inflammation, and Mechanical Hyperalgesia Among Daily Tobacco Smokers. *J. Abnorm. Psychol.* 127, 578. doi:10.1037/abn0000353

Enderby, G. H. (1960). Halothane and Hypotension. *Anaesthesia* 15, 25–32. doi:10.1111/j.1365-2044.1960.tb13891.x

Fey, M., and Lenssen, J. E. (2019). *Fast Graph Representation Learning with Pytorch Geometric*. *arXiv preprint arXiv:1903.02428*. doi:10.48550/arXiv.1903.02428

Gönen, M., Khan, S., and Kaski, S. (2013). "Kernelized Bayesian Matrix Factorization," in International conference on machine learning (Atlanta, Georgia: PMLR), 864–872.

Hamilton, W. L., Ying, R., and Leskovec, J. (2017). "Inductive Representation Learning on Large Graphs," in Proceedings of the 31st International Conference on Neural Information Processing Systems, 1025–1035.

Jin, S., Niu, Z., Jiang, C., Huang, W., Xia, F., Jin, X., et al. (2021). Hetdr: Drug Repositioning Based on Heterogeneous Networks and Text Mining. *Patterns* 2, 100307. doi:10.1016/j.patter.2021.100307

Keshava Prasad, T., Goel, R., Kandasamy, K., Keerthikumar, S., Kumar, S., Mathivanan, S., et al. (2009). Human Protein Reference Database—2009 Update. *Nucleic acids Res.* 37, D767–D772. doi:10.1093/nar/gkn892

Kuhn, M., Campillos, M., Letunic, I., Jensen, L. J., and Bork, P. (2010). A Side Effect Resource to Capture Phenotypic Effects of Drugs. *Mol. Syst. Biol.* 6, 343. doi:10.1038/msb.2009.98

Li, C., Liu, H., Hu, Q., Que, J., and Yao, J. (2019). A Novel Computational Model for Predicting Microrna–Disease Associations Based on Heterogeneous Graph Convolutional Networks. *Cells* 8, 977. doi:10.3390/cells8090977

Li, J., Zheng, S., Chen, B., Butte, A. J., Swamidass, S. J., and Lu, Z. (2016). A Survey of Current Trends in Computational Drug Repositioning. *Briefings Bioinforma.* 17, 2–12. doi:10.1093/bib/bbv020

Luo, Y., Zhao, X., Zhou, J., Yang, J., Zhang, Y., Kuang, W., et al. (2017). A Network Integration Approach for Drug-Target Interaction Prediction and Computational Drug Repositioning from Heterogeneous Information. *Nat. Commun.* 8, 1–13. doi:10.1038/s41467-017-00680-8

Moon, C., Jin, C., Dong, X., Abrar, S., Zheng, W., Chirkova, R. Y., et al. (2021). Learning Drug-Disease-Target Embedding (Ddte) from Knowledge Graphs to Inform Drug Repurposing Hypotheses. *J. Biomed. Inf.* 119, 103838. doi:10.1016/j.jbi.2021.103838

Niu, Y., Song, C., Gong, Y., and Zhang, W. (2021). Mirna-drug Resistance Association Prediction through the Attentive Multimodal Graph Convolutional Network. *Front. Pharmacol.* 12, 799108. doi:10.3389/fphar.2021.799108

Pan, X., Lin, X., Cao, D., Zeng, X., Yu, P. S., He, L., et al. (2022). *Deep Learning for Drug Repurposing: Methods, Databases, and Applications*. Wiley Interdisciplinary Reviews: Computational Molecular Science, e1597.

Sadeghi, S., Lu, J., and Ngom, A. (2021). A Network-Based Drug Repurposing Method via Non-negative Matrix Factorization. *Bioinformatics* 38, 1369–1377. doi:10.1093/bioinformatics/btab826

Sadeghi, S. S., and Keyvanpour, M. R. (2019). An Analytical Review of Computational Drug Repurposing. *IEEE/ACM Trans. Comput. Biol. Bioinforma.* 18, 472–488. doi:10.1109/TCBB.2019.2933825

Su, X., You, Z., Wang, L., Hu, L., Wong, L., Ji, B., et al. (2021). Sane: A Sequence Combined Attentive Network Embedding Model for Covid-19 Drug Repositioning. *Appl. Soft Comput.* 111, 107831. doi:10.1016/j.asoc.2021.107831

Tsoumakas, G., Katakis, I., and Vlahavas, I. (2009). "Mining Multi-Label Data," in *Data Mining and Knowledge Discovery Handbook* (Springer), 667–685. doi:10.1007/978-0-387-09823-4_34

Wang, W., Yang, S., Zhang, X., and Li, J. (2014). Drug Repositioning by Integrating Target Information through a Heterogeneous Network Model. *Bioinformatics* 30, 2923–2930. doi:10.1093/bioinformatics/btu403

Wang, Z., Zhou, M., and Arnold, C. (2020). Toward Heterogeneous Information Fusion: Bipartite Graph Convolutional Networks for In Silico Drug Repurposing. *Bioinformatics* 36, i525–i533. doi:10.1093/bioinformatics/btaa437

Wishart, D. S., Knox, C., Guo, A. C., Shrivastava, S., Hassanali, M., Stothard, P., et al. (2006). Drugbank: a Comprehensive Resource for In Silico Drug Discovery and Exploration. *Nucleic acids Res.* 34, D668–D672. doi:10.1093/nar/gkj067

Yu, J.-L., Dai, Q.-Q., and Li, G.-B. (2021a). Deep Learning in Target Prediction and Drug Repositioning: Recent Advances and Challenges. *Drug Discov. Today*. doi:10.1016/j.drudis.2021.10.010

Yu, Z., Huang, F., Zhao, X., Xiao, W., and Zhang, W. (2021b). Predicting Drug–Disease Associations through Layer Attention Graph Convolutional Network. *Briefings Bioinforma.* 22, bbaa243. doi:10.1093/bib/bbaa243

Zeng, X., Zhu, S., Liu, X., Zhou, Y., Nussinov, R., and Cheng, F. (2019). Deepdr: a Network-Based Deep Learning Approach to In Silico Drug Repositioning. *Bioinformatics* 35, 5191–5198. doi:10.1093/bioinformatics/btz418

Zhang, W., Yue, X., Chen, Y., Lin, W., Li, B., Liu, F., et al. (2017). "Predicting Drug-Disease Associations Based on the Known Association Bipartite Network," in 2017 IEEE international conference on bioinformatics and biomedicine (BIBM) (IEEE), 503–509. doi:10.1109/bibm.2017.8217698

Zhang, W., Yue, X., Lin, W., Wu, W., Liu, R., Huang, F., et al. (2018). Predicting Drug-Disease Associations by Using Similarity Constrained Matrix Factorization. *BMC Bioinforma.* 19, 1–12. doi:10.1186/s12859-018-2220-4

Zhao, B.-W., You, Z.-H., Wong, L., Zhang, P., Li, H.-Y., and Wang, L. (2021). Mgrl: Predicting Drug-Disease Associations Based on Multi-Graph Representation Learning. *Front. Genet.* 12, 491. doi:10.3389/fgene.2021.657182

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.