Methods 207 (2022) 81–89

Contents lists available at ScienceDirect

Methods

journal homepage: www.elsevier.com/locate/ymeth

Toward drug-miRNA resistance association prediction by positional encoding graph neural network and multi-channel neural network

Chengshuai Zhao [a,1], Haorui Wang [b,1], Weiwei Qi [c], Shichao Liu [a,*]

[a] College of Informatics, Huazhong Agricultural University, Wuhan 430070, China
[b] School of Computer Science, Wuhan University, Wuhan 430072, China
[c] Hubei Bailianhe Pumped-storage Power Station, Wuhan 430074, China

ARTICLE INFO

Keywords:
Drug-miRNA resistance association
Representation learning
Graph neural network
Positional encoding
Multi-channel neural network

ABSTRACT

Drug discovery is a costly and time-consuming process, and most drugs exert therapeutic efficacy by targeting specific proteins. However, there are a large number of proteins that are not targeted by any drug. Recently, miRNA-based therapeutics are becoming increasingly important, since miRNA can regulate the expressions of specific genes and affect a variety of human diseases. Therefore, it is of great significance to study the associations between miRNAs and drugs to enable drug discovery and disease treatment. In this work, we propose a novel method named DMR-PEG, which facilitates drug-miRNA resistance association (DMRA) prediction by leveraging positional encoding graph neural network with layer attention (LAPEG) and multi-channel neural network (MNN). LAPEG considers both the potential information in the miRNA-drug resistance heterogeneous network and the specific characteristics of entities (i.e., drugs and miRNAs) to learn favorable representations of drugs and miRNAs. And MNN models various sophisticated relations and synthesizes the predictions from different perspectives effectively. In the comprehensive experiments, DMR-PEG achieves the area under the precision-recall curve (AUPR) score of 0.2793 and the area under the receiver-operating characteristic curve (AUC) score of 0.9475, which outperforms the most state-of-the-art methods. Further experimental results show that our proposed method has good robustness and stability. The ablation study demonstrates each component in DMR-PEG is essential for drug-miRNA drug resistance association prediction. And real-world case study presents that DMR-PEG is promising for DMRA inference.

1. Introduction

Most drugs exert therapeutic efficacy by targeting specific proteins. Drug development is an extremely costly and time-consuming process. It is estimated that it will spend 2.6 billion dollars as well as 12 years developing a new drug [1]. The key to drug discovery is target identification. Research shows that there are about 20,000 to 25,000 protein-coding genes identified in the human genome [2], whereas all approved drugs only target around 600 disease-modifying proteins [3]. There are a large number of proteins that are not targeted by any drug (i.e., they are undruggable), which impedes the drug discovery process. In this case, some researchers turn their eyes to other biomedical entities, such as miRNA.

MiRNA is a category of single-stranded, endogenous, evolutionally conserved RNA, which plays a critical role in many biological processes [4,5]. Recently, some scientists discover that drug sensitivity and resistance are greatly influenced by miRNA profiling in patients [6,7]. Excessive miRNA expression can downregulate genes with protein products targeted by drugs. By contrast, insufficient miRNA expression can upregulate genes with protein products inhibiting drug function [8]. Furthermore, the intervention with miRNAs may allow specific manipulation of proteins. For example, miRNA inhibitors can help with selective upregulation of a target protein, while miRNA mimics can induce downregulation of the target protein by gene silencing [9]. Most importantly, miRNAs are therapeutic targets for a large number of diseases [10], including diabetes [11], cardiovascular diseases [12], and lung cancer [13]. Consequently, the study on drug-miRNA resistance associations (DMRAs) helps to target undruggable proteins and further enables drug discovery and disease treatment [14,15].

To expedite the identification of DMRAs, it is common practice to perform in silico prediction to refine the candidate list for further validation experiments. As far as we know, however, there are only a few

* Corresponding author.
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.ymeth.2022.09.005
Received 19 April 2022; Received in revised form 1 September 2022; Accepted 18 September 2022
Available online 24 September 2022

computational tools proposed for drug-miRNA resistance association prediction. For instance, GCMDR [16] predicts drug-miRNA associations by developing a graph convolution model. In the light of the significance of DMRA identification, more research is needed.

Drug-miRNA resistance association prediction is a challenging task, and there exist some difficulties. The first is how to learn good representations of drugs and miRNAs. There are some features available involving drugs and miRNAs, e.g., SMILES, fingerprint, and similarity. But the question is how to merge them effectively. The second is how to predict the DMRAs precisely given the representation. Because of the sophisticated relations between drugs and miRNAs, ordinary predictors are not powerful enough for accurate prediction.

To overcome the above challenges, we propose a novel method named DMR-PEG for drug-miRNA resistance association prediction. For the first question, DMR-PEG leverages a positional encoding graph neural network with layer attention (LAPEG) to extract favorable representations of drugs and miRNAs from a drug-miRNA association heterogeneous network. Specifically, LAPEG employs a positional encoding graph neural network which updates the node and positional features by separate channels and keeps permutation and rotation equivariance, where a layer attention (LA) mechanism is utilized to combine the representation from different hops. To fully exploit property information, we further apply LAPEG to molecular graphs for drug representation learning and design a task-specific feature extractor for miRNA representation learning. For the second one, a well-designed multi-channel neural network (MNN) in DMR-PEG is built to make and synthesize the predictions for association from various channels. In MNN, three modules are considered: multi-layer perceptron (MLP), generalized tensor factorization (GTF), and compressed tensor network (CTN) to capture insights from different perspectives and model sophisticated drug-miRNA relations. Our main contributions can be summarized as follows:

- We propose a novel computational method named DMR-PEG, which can precisely predict the resistance associations between the drugs and miRNAs.
- We design a positional encoding neural network with layer attention that considers both the potential information in the miRNA-drug resistance heterogeneous network and the specific characteristics (properties) of drugs and miRNAs.
- We construct a multi-channel neural network that models various sophisticated relations and synthesizes the predictions from different perspectives effectively.
- We conduct comprehensive experiments to compare DMR-PEG with the most state-of-the-art methods (where DMR-PEG achieves the most competitive results), discuss the robustness and sensitivity, validate the effectiveness of components in our proposed model, and finally testify its practical value in real-world data.

2. Materials

In this section, we will introduce the materials used in our experiments.

2.1. Datasets

In the experiments, the data could be divided into 3 categories by their types: drug properties, drug-miRNA resistance associations, and miRNA characteristics. These data are collected from different biomedical databases or papers:

2.1.1. ncDR

ncDR [17] (http://www.jianglab.cn/ncDR) is comprehensive cheminformatics and bioinformatics resource that collects curated and predicted drug resistance-related non-coding RNA (ncRNA). Obtained from nearly 3,300 pieces of literature in about 900 published papers, the dataset contains 5,864 validated relationships between 145 drug compounds and 1,039 ncRNAs which consist of 877 miRNAs and 162 lncRNAs. Moreover, the dataset also provides 226,109 unverified drug-miRNA resistance associations, which are predicted by drug response data, miRNA expressions, and lncRNA expressions.

2.1.2. PubChem

PubChem [18] is the largest open database of freely accessible chemical information in the world, which is maintained by the United States National Institute of Health (NIH), including chemical and physical properties, biological activities, safety and toxicity information, patents, literature citations and more. More than 80 database vendors contribute to the growing PubChem database. So far, PubChem has contained formulas, structures, and identifiers of more than 111 M chemicals, 281 M substances of mixtures, extracts, complexes, and uncharacterized, and 295 M bioactivities from high-throughput screening programs.

2.1.3. miroRNA.org

miroRNA.org [19] (http://www.microrna.org) is a comprehensive database of miRNA target prediction and miRNA expression files. Acquired by miRanda algorithm [20] and a comprehensive sequencing project, miroRNA.org has been one of the most important resources.

2.1.4. Paper source

Paper [21] computes functional 2589-dimension similarity features of miRNAs based on Gene Ontology terms.

The detailed information about the data used in the experiment are listed in Table 1.

2.2. Graph neural networks

Nowadays, graphs are playing an increasingly important role in biomedical representation learning. Many challenging tasks can be solved due to the introduction of the networks (or graphs) [22–27], such as cancer drug response prediction based on heterogeneous bipartite networks [28], lncRNA-miRNA interaction inference base on lncRNA-miRNA interaction networks [29,30], and drug repositioning based on bio-entities knowledge graphs [31]. By utilizing graphs, we can easily model biomedical entities and relations between them, which are usually embedded into a low-dimensional space via graph embedding methods.

Recently, graph neural network (GNN) [32–34] is among the most popular methods to learn node embeddings in the graph. The key of GNN is the message passing framework [35]. Given a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, X)$, where $\mathcal{V}$ is the vertex set, $\mathcal{E}$ represents edge set and X denotes node features. Message passing framework learns latent representation vectors of nodes by aggregating the node features through existing links, which can be illustrated as:

(a) initialize node representations with node features:

$$\mathbf{h}_v^{(0)} \leftarrow X_v, \ \forall v \in \mathcal{V} \tag{1}$$

Table 1
Dataset.

Data	Number	Dimension	Source
Drug SMILES	106	N.A.	PubChem [18]
Drug fingerprint	106	920	PubMed
MiRNA functional similarity	754	2,589	Paper [21]
MiRNA expression	754	172	microRNA.org [19]
Drug-miRNA resistance association	3338	N.A.	ncDR [17]

(b) update node representations by neighborhood aggregation, which can be denoted by:

$$\mathbf{m}_{uv}^{(l)} = \phi_e\left(\mathbf{h}_u^{(l-1)}, \mathbf{h}_v^{(l-1)}\right), \ \forall (u,v) \in \mathscr{E} \tag{2}$$

$$\mathbf{a}_v^{(l)} = \sum_{u \in N_v}\left(\mathbf{m}_{uv}^{(l)}\right), \ \forall v \in \mathscr{V} \tag{3}$$

$$\mathbf{h}_v^{(l)} = \phi_h\left(\mathbf{h}_v^{(l-1)}, \mathbf{a}_v^{(l)}\right) \tag{4}$$

where N_v represents the set of neighbors of node v. ϕ_e and ϕ_h are map functions representing the node and edge operations respectively, which are usually implemented by neural networks. Specifically, ϕ_e firstly map the node representations to message m via edge (u, v). After aggregation, the message along with the original representation is utilized to produce the new node representation by function ϕ_h.

Inspired by the powerful capacity of the GNN framework, we consider it to learn the representations of drugs and miRNAs in our architecture.

2.3. Positional encoding

However, conventional GNN has its limitations, since it can not distinguish isomorphic nodes in the graph. As shown in Fig. 1, it indicates that GNN may generate the same numerical representations for many non-isomorphic sets of nodes.

To address this issue, related works consider feature augmentations, which add some precomputed node/edge features before running GNN on the graph. These precomputed node features could be node IDs, random features (RFs) [36], and distance encodings (DEs) [37]. Node IDs can help GNN distinguish nodes, but the one-hot encoding imposes order on nodes and thus violates the permutation invariance principle. RF enhances the capacity of GNN by introducing some randomness. Specifically, random features need to be resampled from the same distribution in the training process, thus if the distribution is invariant to node IDs, the obtained model is permutation invariant [38], but in practice, this setting makes the predictive model difficult to converge. DE aims to measure some relative distance between nodes by using some structural features, e.g., the shortest path distance between nodes. DE, as extra features, is theoretically more powerful and empirically performs well [39–41], but suffers high computational complexity.

PEG [42] employs absolute positional encoding (PE) to replace (relative) distance encoding by measuring the distance between positional encoding, since the absolute positions may be shared across different queries, which can achieve better scalability. The positional encoding is usually extracted by graph embedding techniques, such as Laplacian eigenmaps [43], deepwalk [44], LINE [45], etc., which compute PE of nodes by matrix factorization. Moreover, PEG keeps the GNN layer permutation equivariant rotation equivariant concerning node features and positional encodings, which can contribute to favorable results. Therefore, we follow this setting in the proposed method to learn powerful representations in this paper.

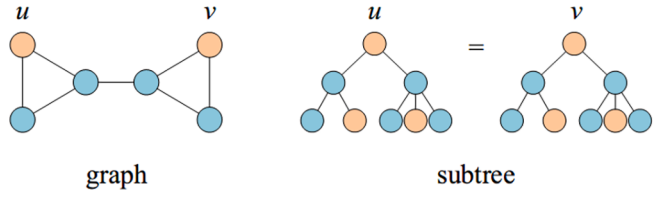

Fig. 1. Isomorphic nodes. Nodes that get matched under graph automorphism will be associated with the same representation by GNN. For example, Node u and v have the same subtree during the aggregation process, they can not be disguised by GNN.

2.4. Molecular graph

There exists a combination of chemical rings, chains, and functional groups [46], which play a significant role in the properties and structures of a drug. The relationship between atoms, which supports the pharmacophoric elements of drugs, can orient them in the right direction for optimal interaction with the receptor. Therefore, the molecular graph, which considers atoms and bonds as nodes and edges, is a suitable tool to represent the properties and structures of the drug.

Since we have collected the SMILES of the drug, the next step is to construct the drug molecular graph. In this paper, we acquire the feature of atoms following DeepChem[47]. The final feature of atoms is denoted by concatenating five types of atom features. Detailed information is illustrated in Table 2. The construction process can be viewed in Fig. 2.

3. Methods

In this section, we will introduce our proposed method to identify drug-miRNA resistance associations. Firstly, we construct a drug-miRNA association heterogeneous network, which is illustrated in Section 3.1. Then, we formulate a positional encoding graph neural network with layer attention (LAPEG) in Section 3.2. Later, we illustrate the representation learning of drugs and miRNAs with LAPEG as well as a task-specific feature extractor in Section 3.3. In the end, we enable DMRA prediction by leveraging a multi-channel neural network in Section 3.4.

3.1. Heterogeneous network

To fully exploit the information of known drug-miRNA resistance associations, we first construct the drug-miRNA heterogeneous network. Given some drugs and miRNAs, associations between them can be denoted by an adjacent matrix A_{dm}. Further, we introduce the matrix S_d and S_m, which indicate drug-drug and miRNA-miRNA similarity, to make the best of the properties of entities. In this work, we consider drug fingerprints and miRNA expressions to calculate similarity. There are many similarity measures, such as Cosine similarity, Jaccard similarity, and Pearson similarity. From our preliminary study, Pearson similarity [48], a simple but effective measure, can lead to satisfactory performance. Thus, the similarity can be formulated as:

$$S = \frac{\sum (X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum (X_i - \overline{X})^2 \sum (Y_i - \overline{Y})^2}} \tag{5}$$

where X, Y denotes two vectors of entity (i.e., drug or miRNA).

Then, the heterogeneous network can be constructed based on the adjacent matrix and similarity matrix, which can be denoted by:

$$A_{het} = \begin{bmatrix} S_d & A_{dm} \\ A_{dm}^T & S_m \end{bmatrix} \tag{6}$$

The drug-miRNA association heterogeneous network contains information on drug-miRNA associations and similarities between entities. Thus, it is a desirable resource for DMRA prediction.

Table 2
The features of atom in the drug molecular graph.

Feature	Dimension
Element type	44
Number of neighbors	11
Number of Hydrogen elements	11
Number of implicit Hydrogen elements	11
Aromatic or not	1
Initial feature	88

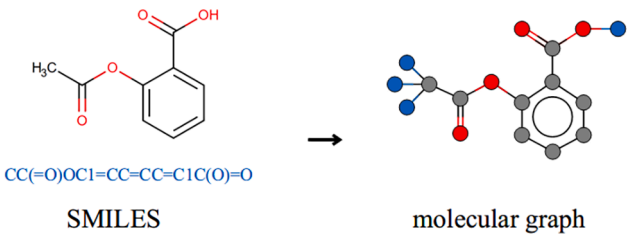

Fig. 2. Construction of molecular graph. A molecular graph considers atoms and bonds as nodes and edges, which can be constructed from SMILES.

3.2. Positional encoding graph neural network with layer attention

In this selection, we design a positional encoding graph neural network with layer attention (LAPEG) to learn representations in the network. LAPEG, as an essential component of DMR-PEG, has two modules: positional encoding graph neural network (PEG) and layer attention (LA) mechanism.

3.2.1. Positional encoding graph neural network

Positional encoding graph neural network utilizes separate channels to update the node and positional features and keeps permutation and rotation equivariance, which is an effective tool to extract representations from both drug-miRNA heterogeneous network and drug molecular graphs. In this work, Laplacian eigenmap (LE), preserving the eigenvectors that correspond to the p smallest eigenvalues of the normalized Laplacian matrix L, is considered to be the PE. Note that the reason why we choose LE is that LE is simple but effective, which could be denoted as:

$$L \triangleq I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}} = U\Lambda U^T \tag{7}$$

$$Z(v|A) = U_{v,1:p} \tag{8}$$

Later, the initial PE $Z^{(0)}$ along with the initial node feature $H^{(0)}$ is fed to a PEG layer for refinement. Intuitively, we can stack several PEG layers to learn high-order representations. After the l-th layer, the representations can be denoted by:

$$H^{(l)}, Z^{(l)} = \left(\sigma\left[\left(\widehat{A} \odot E\right)H^{(l-1)}W\right], Z^{(l-1)}\right) \tag{9}$$

where σ is a non-linear activation. $\widehat{A} = D^{-\frac{1}{2}}\left(A + I\right)D^{\frac{1}{2}}$, where $D = diag(\Sigma_{j=1}A_{ij})$ is the degree matrix of graph G and I is an identity matrix. E is an edge weight matrix, $E_{uv} = MLP(||Z_u - Z_v||)$, $\forall u, v \in V$, and $\odot$ is the Hadamard product. W is a linear transformation.

3.2.2. Layer attention mechanism

In the previous section, L-layer PEG is utilized to learn node representations while keeping equivariance. Since different layers in PEG capture information of various hops (shown in Fig. 3), we combine representations of different layers with a layer attention mechanism [49] to learn better representations:

$$H = \sum_{l=1}^{L} a^{(l)}H^{(l)} \tag{10}$$

where $a^{(l)}$ is the attention weight with respect to $H^{(l)}$.

Thus, our proposed LAPEG can be formulated by:

$$H, Z = LAPEG\left(A, H^{(0)}, Z^{(0)}\right) \tag{11}$$

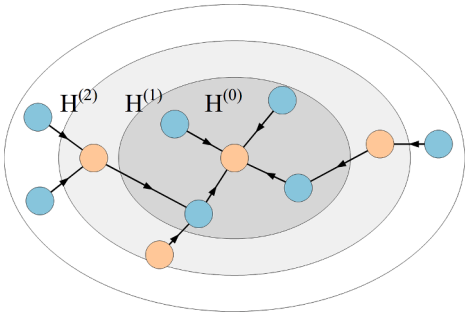

Fig. 3. Layer attention mechanism. $H^{(l)}$ is the embeddings of the nodes from l-th hop. The final representations are computed by weighted summation of those on various hops.

3.3. Representation learning

In this section, we will introduce the representation learning process of drugs and miRNAs.

Since we have constructed the drug-miRNA association heterogeneous network (Section 3.1) and formulated the positional encoding graph neural network with layer attention model (Section 3.2), it is natural to apply LAPEG to learn representations of drugs and miRNA from heterogeneous networks. The above process can be defined as:

$$H_{het}, Z_{het} = LAPEG\left(A_{het}, H_{het}^{(0)}, Z_{het}^{(0)}\right) \tag{12}$$

Note that the representation of nodes H_{het} can be categorized as H_{drug} (i.e., drug representations) and H_{miRNA} (i.e., miRNA representations).

With respect to drugs, we have obtained molecular graphs. Samely, we employ LAPEG to learn node representation (molecule representation):

$$H_{mol}, Z_{mol} = LAPEG\left(A_{mol}, H_{mol}^{(0)}, Z_{mol}^{(0)}\right) \tag{13}$$

To acquire the representation of the drug from the molecular graph, here, we leverage a pooling operation by summarizing the representations of atoms:

$$F_{drug} = Pooling\left(H_{mol}\right) \tag{14}$$

Given $H_{mol} \in \mathbb{R}^{(n,d)}$, after pooling operation, the drug representation can be $F_{drug} \in \mathbb{R}^{d}$. Different operations (e.g., max, mean, and sum) can be considered in the pooling process, further discussions can be viewed in Section 4.3.1.

As to miRNAs, we have collected the functional similarity $F_{function}$ and miRNA expressions $F_{expressions}$. Then, we consider refining the representations of miRNAs with a task-specific feature extractor where the concatenation of two features is the input. In this paper, we implement the task-specific feature extractor with a multi-layer perceptron (MLP):

$$F_{miRNA} = MLP\left(F_{function} \oplus F_{expressions}\right) \tag{15}$$

where $\oplus$ is concatenation.

Ultimately, the representation of drug and miRNA can be computed by the synthesis of the graph embedding of heterogeneous network and identical features of entities:

$$D = H_{drug} \oplus F_{drug} \tag{16}$$

$$M = H_{miRNA} \oplus F_{miRNA} \tag{17}$$

The whole process of the LAPEG model can be viewed in Fig. 4.

3.4. Multi-channel neural network

In this section, we develop a multi-channel neural network (MNN) to refine the representation of drugs and miRNAs and predict drug-miRNA resistance associations.

The MNN contains three modules: multi-layer perceptron (MLP) [50,51], generalized tensor factorization (GTF), and compressed tensor

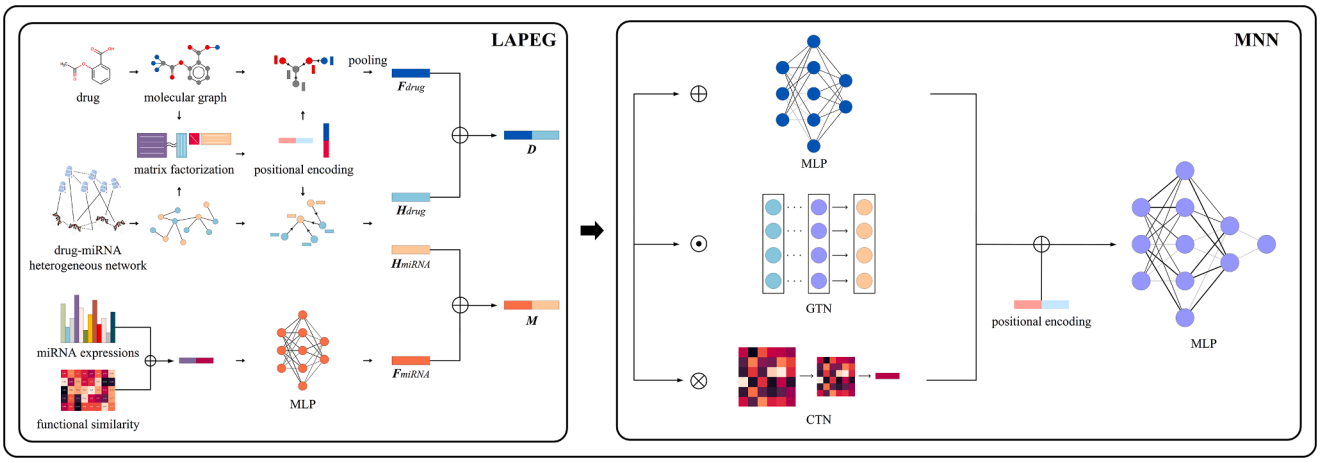

Fig. 4. The workflow of DMR-PEG. DMR-PEG includes two main components: positional encoding graph neural network with layer attention (LAPEG) and multi-channel neural network (MNN). LAPEG learns favorable graph embeddings of drugs and miRNAs from a heterogeneous network. Later, the embeddings along with representations extracted from their identical structures and characteristics are refined by MNN. MNN makes the best of MLP, GTF, and CTN to model the sophisticated relations between the drugs and miRNAs. Ultimately, The prediction scores are given by synthesizing information from the three channels and the positional encoding.

network (CTN) [52]. These three modules are heterogeneous, indicating they can capture relations from different perspectives and then make relatively independent predictions. To further model the sophisticated relations between the drug and miRNA, we design various operators for these modules.

For MLP, we concatenate the representations of drug and miRNA as the input to model the plain and ordinary relation between drug and miRNA:

$$\Gamma = MLP(D \oplus M) \tag{18}$$

In GTF, the double-wise relation is considered, which is denote by:

$$\Lambda = GTF(D \odot M) \tag{19}$$

With respect to CTN, we adopt outer-product to capture implicit spatial relation, which can be defined as:

$$\Omega = CTN(D \otimes M) \tag{20}$$

To enable drug-miRNA resistance association prediction, we need to synthesize various sources of information. Given a drug i and a miRNA j, at last, the outputs of MLP, GTF, and CTN along with the position encoding refined from a heterogeneous network are fed into an MLP to produce the prediction score:

$$P_{ij} = MLP\left(\Gamma_{ij} \oplus \Lambda_{ij} \oplus \Omega_{ij} \oplus \left|\left|Z_i - Z_j\right|\right|\right) \tag{21}$$

4. Results and discussion

In this section, we will first give experiment setting and evaluation metrics in Section 4.1. Then, we design comprehensive experiments to present the start-of-the-art and robustness performance of DMR-PEG in Section 4.2 and Section 4.3. And we discuss the sensitivity of hyper-parameters in Section 4.4. Later, we validate the effectiveness of each component in DMR-PEG in Section 4.5. Lastly, we apply our proposed model to real-world data to investigate its practical value in Section 4.6.

4.1. Experimental setting and evaluation metrics

In this section, we conduct comprehensive experiments to evaluate the performance of our proposed model. The experimental setting and evaluation metrics can be summarized by the following:

4.1.1. Evaluation metrics

We randomly select 85%, 5%, and 10% of the existing drug-miRNA associations as positive training, validation, and testing data, along with 10 times, 100 times, and all non-existing associations which are considered to be negative samples. The following metrics are adopted to evaluate different methods in our experiments: the area under the precise-recall curve (AUPR) and the area under the receiver-operating characteristic curve (AUC), F1-measure (F1), accuracy (ACC), recall (REC), specificity (SPEC), and precision (PRE). All experiments are run 10 times, and average performance is calculated to avoid bias.

4.1.2. Experimental setting

We implement DMR-PEG with 3-layers attention PEG with 128-dimension initial positional encodings, which learns embedding vectors of 128 dimensions. Adam optimizer with learning rate 1e-3 is selected to optimize the model for 30 epochs. Moreover, dropout layers with a rate of 0.2 are utilized to alleviate overfitting. In practice, we also try other parameter settings, which leads to inferior performance (or efficiency) of DMR-PEG.

4.2. Comparative experiment

In this section, we will compare DMR-PEG with state-of-the-art drug-miRNA resistance association prediction methods, including two traditional link prediction methods: collaborative filtering[53] and label propagation[54], four graph embedding methods: graph factorization [55], SDNE[56], LINE[45], and GCN[33], and one task specified methods: GCMDR [16]. Note that we also consider our previous work DMR-GCN [57], which employs the traditional graph neural network layers rather than the PEG layers to extract node representations. Comparative results are demonstrated in Table 3.

From the table, we can observe that (1) in general, our methods (i.e., DMR-GCN and DMR-PEG) achieve the AUPR score of 0.2747 and 0.2793, the AUC score of 0.9393 and 0.9475, which both outperform all above state-of-the-art methods. It suggests both DMR-GCN and DMR-PEG are well-designed and suitable for DMRA prediction. (2) in detail, compared with traditional link prediction methods, the improvement is more significant, which indicates the effective representation learning capacity of the GNN framework. (3) then, both DMR-GCN and DMR-PEG present superior performance compared with single graph embedding methods, which suggests that it is significant to consider the properties and characteristics of entities, and only the drug-miRNA associations are not enough for DMRA prediction. (4) further, by the comparison of

Table 3
Performance of state-of-the-art methods and DMR-PEG.

Methods	AUPR	AUC	F1	ACC	REC	SPEC	PRE
Collaborative fitering	0.2046 ± 0.0058	0.8618 ± 0.0058	0.2873 ± 0.0042	0.9856 ± 0.0007	0.3314 ± 0.0157	0.9913 ± 0.0008	0.2662 ± 0.0134
Label propagation	0.2262 ± 0.0060	0.8610 ± 0.0039	0.2875 ± 0.0059	**0.9886 ± 0.0010**	0.3176 ± 0.0220	**0.9945 ± 0.0012**	0.2554 ± 0.0294
Graph factorization	0.1546 ± 0.0121	0.8712 ± 0.0151	0.2274 ± 0.0127	0.9818 ± 0.0015	0.3009 ± 0.0301	0.9878 ± 0.0017	0.1911 ± 0.0126
SDNE	0.2264 ± 0.0135	0.8869 ± 0.0047	0.2884 ± 0.0116	0.9877 ± 0.0008	0.3191 ± 0.0176	0.9936 ± 0.0010	0.2657 ± 0.0216
LINE	0.1716 ± 0.0065	0.8605 ± 0.0106	0.2427 ± 0.0083	0.9830 ± 0.0024	0.2932 ± 0.0139	0.9896 ± 0.0035	0.1995 ± 0.0082
GCN	0.1180 ± 0.0095	0.7828 ± 0.0138	0.1895 ± 0.0018	0.9835 ± 0.0018	0.2185 ± 0.0103	0.9902 ± 0.0018	0.1748 ± 0.0026
GCMDR	0.2274 ± 0.0020	0.9295 ± 0.0006	0.2813 ± 0.0020	0.9857 ± 0.0004	0.3189 ± 0.0127	0.9916 ± 0.0005	0.2597 ± 0.0072
DMR-GCN	0.2747 ± 0.0034	0.9393 ± 0.0019	**0.3241 ± 0.0049**	0.9885 ± 0.0004	0.3188 ± 0.0182	0.9943 ± 0.0012	**0.3120 ± 0.0098**
DMR-PEG	**0.2793 ± 0.0068**	**0.9475 ± 0.0195**	0.3196 ± 0.0042	0.9864 ± 0.0003	**0.3513 ± 0.0132**	0.9919 ± 0.0022	0.2995 ± 0.0079

DMR-PEG with DMR-GCN, we can conclude that employing LAPEG is favorable for representation learning in the DMRA prediction task.

Thus, our proposed method is well designed and can predict drug-miRNA resistance associations precisely.

4.3. Robustness experiment

To evaluate the generalization ability of our proposed model, we design an experiment on different sparsity of the heterogeneous network by removal of a certain proportion of links. In the experiments, we first randomly retain 90%, 80%, and 70% of drug-miRNA associations in the heterogeneous network and build our proposed method in these settings respectively.

As shown in Fig. 5, we can find that generally, the performance of DMR-PEG slightly decreases as more associations are removed. The AUC score falls from 0.9475 to 0.9394, 0.9332, and 0.9238 respectively when the 90%, 80%, and 70% drug-miRNA associations are retained. Specifically, in the case where 20 % associations in the datasets are removed, the AUPR score and the AUC score fall by 9.70% and 1.51 %. It is worth mentioning that even if only 70 % of associations are retained, the proposed method achieves the AUPR score of 0.2423 and the AUC score of 0.9238 which are still the most competitive among all the state-of-the-art methods.

In conclusion, our proposed DMR-PEG has good robustness.

4.4. Sensitivity analysis

In this selection, we examine the influence of several key hyper-parameters on the performance of the proposed model.

4.4.1. Impact of pooling methods

In DMR-PEG, the global pooling method is utilized to acquire graph embedding from molecular graphs (Section 3.3). Here, we discuss the impact of three different pooling methods, i.e., max, mean, and sum. As we can observe in Fig. 6, the predictive model achieves the best performance when adopting the mean pooling operation, which is because compared with max and sum, the mean operation considers all the nodes in the molecular graph and treat those equally by the average. The max operation based pooling which only considers the most important nodes in the graph thus produces the second performance. The operation of sum is a simple addition of all the nodes, which neglects the effect of the number of nodes on the gradients during the training process. Moreover, the mean pooling operation contributes to the most stable performance due to the same reason.

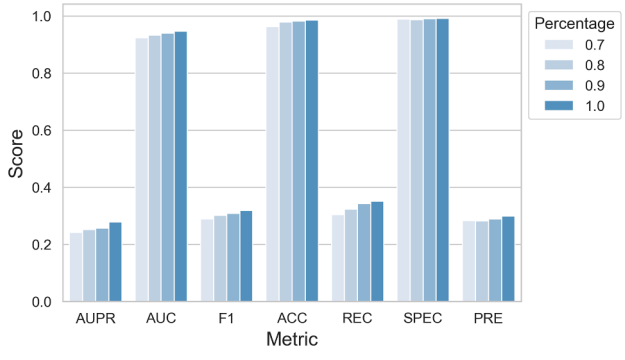

Fig. 5. Performance of DMR-PEG on the network of different sparsity. It is presented with a bar plot where a darker color or higher bar indicates a higher score. It is demonstrated that when the drug-miRNA associated are removed, the performance of DMR-PEG is slightly decreasing, which suggests better robustness.

4.4.2. Impact of dimensionality

Then, we investigate the influence of the dimension of node embeddings and positional encodings by varying from 32 to 256, which is demonstrated in Fig. 7. Generally, node embedding vectors and positional encodings have different dimension sensitivity. The performance of positional encodings is more stable compared to that of node

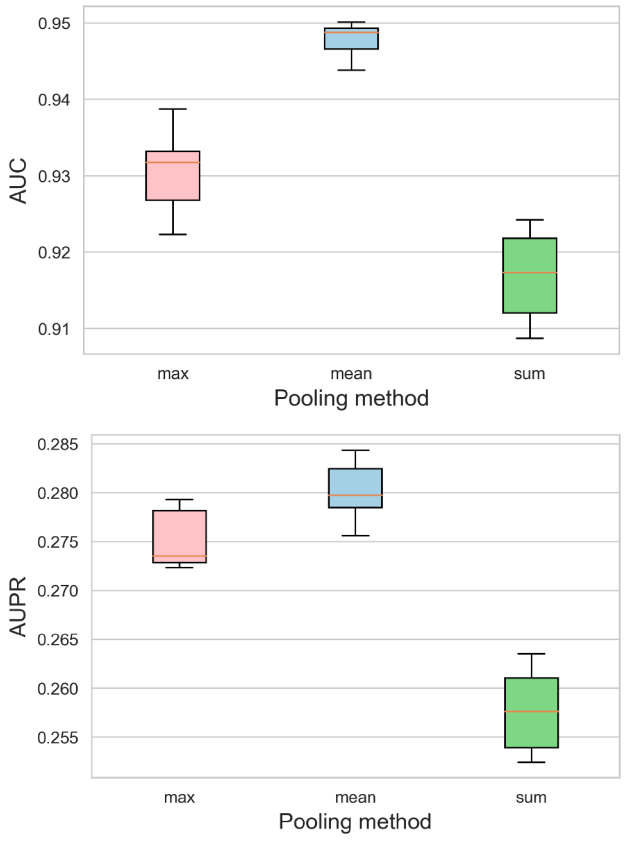

Fig. 6. Impact of pooling methods. It is demonstrated with a box plot that the AUC and AUPR score is highest and stablest when adopting meaning pooling operation. The max operation is the secondary and the sum operation is the last..

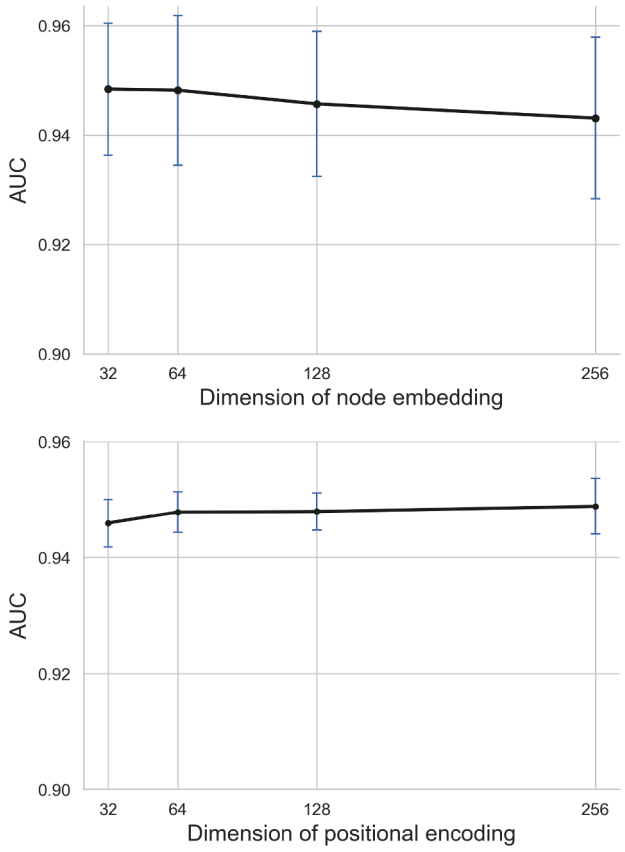

Fig. 7. Impact of dimensionality. It is demonstrated with a line plot that the performance of positional encodings is more stable than embedding vectors proved by the low changes and stand errors. Generally, DMR-PEG achieves stable performance under different hyperparameter settings.

embedding vectors, which can be proved from two aspects. On the one hand, the change of performance is more slight under different dimension settings. On the other hand, under the same dimension setting, the stand error of performance is quite low. And DMR-PEG achieves better as the dimension of positional encodings increases. But the performance falls when it comes to node embedding. Specifically, the AUC score falls from 0.9484 to 0.9482, 0.9457, and 0.9431 when the dimensions of node embedding increase from 32 to 64, 128, and 256. It is because higher dimensionality contains much more information, but suffers a lot more from overfitting.

Overall, our proposed model shows stable performance under different hyperparameter settings.

4.5. Ablation study

To further validate the effect of each component in DMR-PEG, we design several variants and conduct ablation in this section:

- **Default** is a complete implementation of DMR-PEG where none of the modules is removed.
- **w/o PE** removes the positional encoding in the LAPEG. Then, the DMR-PEG degenerates into a graph neural network with layer attention.
- **w/o MLP** removes the MLP in the MNN.
- **w/o GTF** removes the GTF in the MNN.
- **w/o CTN** removes the CTN in the MNN.
- **w/o LA** removes the layer attention mechanism in the LAPEG. Then, the DMR-PEG degenerated into a positional encoding graph neural network (PEG).

The comparison of our model and its variants are listed in Fig. 8. It is clear that DMR-PEG with all components achieves superior performances in most scenarios, and the removal of any will undermine the predictive capacity of the model. Specifically, no matter what modules in the multi-channel neural network are deleted (i.e., MLP, GTF, CTN), the performance falls accordingly, which indicates the necessity of the multi-channel architecture and the relations between drugs and miRNAs are very sophisticated. Overall, the layer attention mechanism is more efficient than other modules. Specifically, compared with the default setting, the AUPR score and AUC of the PEG model fall by 20.55 % and 2.01 % respectively. Moreover, the effectiveness of the PE can be testified due to its significant contribution to DMR-PEG, which suggests the introduction of PE facilitates the representation learning in the drug-miRNA resistance association prediction.

Therefore, each component in our proposed model is essential to achieve state-of-the-art performance for DMRA prediction.

4.6. Case study

The primary goal of computational methods is to screen the drug-miRNA resistance associations to guide wet experiments. Here, we utilize the DMR-PEG model to discover potential drug-miRNA resistance associations in the heterogeneous network. Firstly, we construct the DMR-PEG model using all associations in our dataset. Then, all non-associated drug-miRNA pairs are re-evaluated by the predictive model. We check on the top 10 drug-miRNA resistance associations scored by DMR-PEG and try to find scientific evidence to support our findings.

As shown in Table 4, four of them are confirmed to be new associations. Geng et al. [58] suggests that upregulation of miR-23b can sensitize U87 GSCs to TMZ-induced proliferation inhibition, and Liao et al. [59] reveals the association between the polymorphism of miR-146a and the clinical features of the patients who receives adjuvant chemotherapy of oxaliplatin and fluoropyrimidines. Gemcitabine has a major influence on miR145 levels in kidney cancer treatment [60]. Cisplatin, which is usually used in the treatment of cancer of the esophagus, has relations with the expression level of hsa-mir-425 [61]. The case studies show that the DMR-PEG is promising for novel drug-miRNA resistance associations discovery.

5. Conclusion

The identification of drug-miRNA resistance association can help to facilitate miRNA-targeted drug discovery and further treat some diseases. In this work, we focus on important but rarely studied computational methods to promote drug-miRNA resistance association prediction. Specifically, we propose a computational tool named DMR-PEG. DMR-PEG leverages positional encoding graph neural network with layer attention (LAPEG) and multi-channel neural network (MNN). Comprehensive experiments demonstrate that DMR-PEG is competitive,

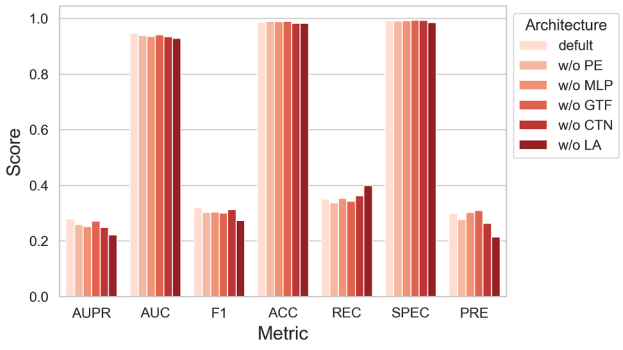

Fig. 8. Ablation study of LAPEG. It is demonstrated with a bar plot that each component in DMR-PEG model is essential and the removal of any will undermine the predictive capacity of the model.

Table 4
Top 10 drug-miRNA resistance associations predicted by DMR-PEG.

NO.	Drug	miRNA	Evidence
1	Temozolomide	hsa-miR-23b	YES [58]
2	Gemcitabine	hsa-miR-30b	N.A.
3	Oxaliplatin	hsa-miR-146a	YES [59]
4	Gemcitabine	hsa-miR-145	YES [60]
5	Gemcitabine	hsa-miR-197	N.A.
6	Doxorubicin	hsa-miR-363	N.A.
7	Gemcitabine	hsa-miR-320a	N.A.
8	5-Fluorouracil	hsa-miR-100	N.A.
9	Cisplatin	hsa-miR-425	YES [61]
10	Gemcitabine	hsa-miR-23b	N.A.

robust, stable, and promising for drug-miRNA resistance association inference. And further discussion validates the effectiveness of each component in the DMR-PEG.

In the future, we will take more biomedical entities into account and explore more powerful representative learning strategies.

CRediT authorship contribution statement

Chengshuai Zhao: Methodology, Writing - review & editing. **Haorui Wang:** Conceptualization, Methodology, Software. **Weiwei Qi:** Writing - review & editing. **Shichao Liu:** Conceptualization, Methodology, Supervision, Project administration.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No.62102158), the Fundamental Research Funds for the Central Universities (Grant No.2662022JC004), 2021 Foshan support project for promoting the development of university scientific and technological achievements service industry(zc03040000014). Huazhong Agricultural University Scientific & Technological Self-innovation Foundation. The funders have no role in study design, data collection, data analysis, data interpretation, or writing of the manuscript.

References

[1] H.S. Chan, H. Shan, T. Dahoun, H. Vogel, S. Yuan, Advancing drug discovery via artificial intelligence, Trends in pharmacological sciences 40 (8) (2019) 592–604.
[2] F. Collins, E. Lander, J. Rogers, R. Waterston, I. Conso, Finishing the euchromatic sequence of the human genome, Nature 431 (7011) (2004) 931–945.
[3] A.L. Hopkins, C.R. Groom, The druggable genome, Nature reviews Drug discovery 1 (9) (2002) 727–730.
[4] D.M. Dykxhoorn, C.D. Novina, P.A. Sharp, Killing the messenger: short rnas that silence gene expression, Nature reviews Molecular cell biology 4 (6) (2003) 457–467.
[5] M.R. Fabian, N. Sonenberg, The mechanics of mirna-mediated gene silencing: a look under the hood of mirisc, Nature structural & molecular biology 19 (6) (2012) 586–593.
[6] G.A. Calin, C.M. Croce, Microrna signatures in human cancers, Nature reviews cancer 6 (11) (2006) 857–866.
[7] D. Kazmierczak, K. Jopek, K. Sterzynska, M. Nowicki, M. Rucinski, R. Januchowski, The profile of microrna expression and potential role in the regulation of drug-resistant genes in cisplatin-and paclitaxel-resistant ovarian cancer cell lines, International journal of molecular sciences 23 (1) (2022) 526.
[8] M. Matsui, D.R. Corey, Non-coding rnas as drug targets, Nature reviews Drug discovery 16 (3) (2017) 167–179.
[9] M.F. Schmidt, Drug target mirnas: chances and challenges, Trends in biotechnology 32 (11) (2014) 578–585.
[10] W. Zhang, M.E. Dolan, Emerging role of micrornas in drug response, Current opinion in molecular therapeutics 12 (6) (2010) 695.
[11] C. Zhang, Novel functions for small rna molecules, Current opinion in molecular therapeutics 11 (6) (2009) 641.
[12] Y. Lu, Y. Zhang, H. Shan, Z. Pan, X. Li, B. Li, C. Xu, B. Zhang, F. Zhang, D. Dong, et al., Microrna-1 downregulation by propranolol in a rat model of myocardial infarction: a new mechanism for ischaemic cardioprotection, Cardiovascular research 84 (3) (2009) 434–441.
[13] L. Du, A. Pertsemlidis, micrornas and lung cancer: tumors and 22-mers, Cancer and Metastasis Reviews 29 (1) (2010) 109–122.
[14] Y. Yamanishi, M. Araki, A. Gutteridge, W. Honda, M. Kanehisa, Prediction of drug–target interaction networks from the integration of chemical and genomic spaces, Bioinformatics 24 (13) (2008) i232–i240.
[15] Z. Chu, S. Liu, and W. Zhang, Hierarchical graph representation learning for the prediction of drug-target binding affinity, arXiv preprint arXiv:2203.11458, 2022.
[16] Y.-A. Huang, P. Hu, K.C. Chan, Z.-H. You, Graph convolution for predicting associations between mirna and drug resistance, Bioinformatics 36 (3) (2020) 851–858.
[17] E. Dai, F. Yang, J. Wang, X. Zhou, Q. Song, W. An, L. Wang, W. Jiang, ncdr: a comprehensive resource of non-coding rnas involved in drug resistance, Bioinformatics 33 (24) (2017) 4010–4011.
[18] E.E. Bolton, Y. Wang, P.A. Thiessen, S.H. Bryant, Pubchem: integrated platform of small molecules and biological activities, in Annual reports in computational chemistry, Elsevier 4 (2008) 217–241.
[19] D. Betel, M. Wilson, A. Gabow, D.S. Marks, C. Sander, The microrna. org resource: targets and expression, Nucleic acids research vol. 36 (suppl_1) (2008) D149–D153.
[20] B. John, A.J. Enright, A. Aravin, T. Tuschl, C. Sander, D.S. Marks, J.C. Carrington, Human microrna targets, PLoS biology 2 (11) (2004), e363.
[21] Y. Yang, X. Fu, W. Qu, Y. Xiao, H.-B. Shen, Mirgofs: a go-based functional similarity measurement for mirnas, with applications to the prediction of mirna subcellular localization and mirna–disease association, Bioinformatics 34 (20) (2018) 3547–3556.
[22] C. Zhao, S. Liu, F. Huang, S. Liu, W. Zhang, Csgnn: Contrastive self-supervised graph neural network for molecular interaction prediction, in: Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, 2021, pp. 19–27.
[23] F. Cheng, C. Liu, J. Jiang, W. Lu, W. Li, G. Liu, W.-X. Zhou, J. Huang, and Y. Tang, Prediction of drug-target interactions and drug repositioning via network-based inference. PLoS Comput. Biol., vol. 8, no. 5, 2012. [Online]. Available: http://dblp.uni-trier.de/db/journals/ploscb/ploscb8.html#ChengLJLLLZHT12.
[24] Z. Yu, F. Huang, X. Zhao, W. Xiao, and W. Zhang, Predicting drug–disease associations through layer attention graph convolutional network, Briefings in Bioinformatics, 2020. [Online]. Available: doi: 10.1093/bib/bbaa243.
[25] W. Zhang, Y. Chen, F. Liu, F. Luo, G. Tian, X. Li, Predicting potential drug-drug interactions by integrating chemical, biological, phenotypic and network data, Bmc Bioinformatics 18 (1) (2017) 18.
[26] I.A. Kovács, K. Luck, K. Spirohn, Y. Wang, C. Pollis, S. Schlabach, W. Bian, D.K. Kim, N. Kishore, and T. Hao, Network-based prediction of protein interactions, Nature Communications, vol. 10, no. 1, 2019.
[27] F. Huang, X. Yue, Z. Xiong, Z. Yu, S. Liu, and W. Zhang, Tensor decomposition with relational constraints for predicting multiple types of microRNA-disease associations, Briefings in Bioinformatics, 2020. [Online]. Available: doi: 10.1093/bib/bbaa140.
[28] X. Liu, C. Song, F. Huang, H. Fu, W. Xiao, W. Zhang, Graphcdr: a graph neural network method with contrastive learning for cancer drug response prediction, Briefings in Bioinformatics vol. 23 (1) (2022) p. bbab457.
[29] C. Zhao, Y. Qiu, S. Zhou, S. Liu, W. Zhang, Y. Niu, Graph embedding ensemble methods based on the heterogeneous network for lncrna-mirna interaction prediction, BMC genomics 21 (13) (2020) 1–12.
[30] S. Zhou, X. Yue, X. Xu, S. Liu, W. Zhang, Y. Niu, Lncrna-mirna interaction prediction from the heterogeneous network through graph embedding ensemble learning, in: 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), IEEE, 2019, pp. 622–627.
[31] Z. Xiong, F. Huang, Z. Wang, S. Liu, W. Zhang, A multimodal framework for improving in silico drug repositioning with the prior knowledge from knowledge graphs, IEEE/ACM Transactions on Computational Biology and Bioinformatics (2021).
[32] J. Bruna, W. Zaremba, A. Szlam, and Y. LeCun, Spectral networks and locally connected networks on graphs, arXiv preprint arXiv:1312.6203, 2013.
[33] T.N. Kipf and M. Welling, Semi-supervised classification with graph convolutional networks, arXiv preprint arXiv:1609.02907, 2016.
[34] M. Defferrard, X. Bresson, P. Vandergheynst, Convolutional neural networks on graphs with fast localized spectral filtering, Advances in neural information processing systems 29 (2016).
[35] J. Gilmer, S.S. Schoenholz, P.F. Riley, O. Vinyals, G.E. Dahl, Neural message passing for quantum chemistry, in: International conference on machine learning, PMLR, 2017, pp. 1263–1272.
[36] R. Sato, M. Yamada, H. Kashima, Random features strengthen graph neural networks, in: Proceedings of the 2021 SIAM International Conference on Data Mining (SDM), SIAM, 2021, pp. 333–341.
[37] P. Li, J. Leskovec, The expressive power of graph neural networks, in: L. Applications, P. Wu, J. Pei Cui, L. Zhao (Eds.), Graph Neural Networks: Foundations, Frontiers, ch. 5, Springer, Singapore, 2021, pp. 63–98.
[38] R. Murphy, B. Srinivasan, V. Rao, B. Ribeiro, Relational pooling for graph representations, in: International Conference on Machine Learning, PMLR, 2019, pp. 4663–4673.
[39] P. Li, Y. Wang, H. Wang, J. Leskovec, Distance encoding: Design provably more powerful neural networks for graph representation learning, Advances in Neural Information Processing Systems 33 (2020) 4465–4478.

[40] M. Zhang, Y. Chen, Link prediction based on graph neural networks, Advances in neural information processing systems 31 (2018).
[41] M. Zhang, P. Li, Y. Xia, K. Wang, and L. Jin, Revisiting graph neural networks for link prediction, 2020.
[42] H. Wang, H. Yin, M. Zhang, and P. Li, Equivariant and stable positional encoding for more powerful graph neural networks, arXiv preprint arXiv:2203.00199, 2022.
[43] M. Belkin, P. Niyogi, Laplacian eigenmaps for dimensionality reduction and data representation, Neural computation 15 (6) (2003) 1373–1396.
[44] B. Perozzi, R. Al-Rfou, S. Skiena, Deepwalk: Online learning of social representations, in: Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining, 2014, pp. 701–710.
[45] J. Tang, M. Qu, M. Wang, M. Zhang, J. Yan, Q. Mei, Line: Large-scale information network embedding, in: Proceedings of the 24th international conference on world wide web, 2015, pp. 1067–1077.
[46] J.B. Taylor, Comprehensive medicinal chemistry II, Elsevier, 2007.
[47] B. Ramsundar, P. Eastman, P. Walters, V. Pande, Deep learning for the life sciences: applying deep learning to genomics, microscopy, drug discovery, and more, O'Reilly Media Inc, 2019.
[48] D. Wang, J. Wang, M. Lu, F. Song, Q. Cui, Inferring the human microrna functional similarity and functional network based on microrna-associated diseases, Bioinformatics 26 (13) (2010) 1644–1650.
[49] Z. Yu, F. Huang, X. Zhao, W. Xiao, W. Zhang, Predicting drug–disease associations through layer attention graph convolutional network, Briefings in Bioinformatics vol. 22 (4) (2021) 243.
[50] X. Wu, B. Shi, Y. Dong, C. Huang, N.V. Chawla, Neural tensor factorization for temporal interaction learning, in: Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining, 2019, pp. 537–545.
[51] X. He, L. Liao, H. Zhang, L. Nie, X. Hu, T.-S. Chua, Neural collaborative filtering, in: Proceedings of the 26th international conference on world wide web, 2017, pp. 173–182.
[52] H. Chen, J. Li, Learning data-driven drug-target-disease interaction via neural tensor network, in: IJCAI, 2020, pp. 3452–3458.
[53] X. Su, T.M. Khoshgoftaar, A survey of collaborative filtering techniques, Advances in artificial intelligence 2009 (2009).
[54] F. Wang, C. Zhang, Label propagation through linear neighborhoods, IEEE Transactions on Knowledge and Data Engineering 20 (1) (2007) 55–67.
[55] A. Ahmed, N. Shervashidze, S. Narayanamurthy, V. Josifovski, A.J. Smola, Distributed large-scale natural graph factorization, in: Proceedings of the 22nd international conference on World Wide Web, 2013, pp. 37–48.
[56] D. Wang, P. Cui, W. Zhu, Structural deep network embedding, in: Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining, 2016, pp. 1225–1234.
[57] H. Wang, S. Khan, S. Liu, F. Zheng, W. Zhang, Predicting drug-mirna resistance with layer attention graph convolution network and multi channel feature extraction, in: 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), IEEE, 2021, pp. 1083–1089.
[58] J. Geng, H. Luo, Y. Pu, Z. Zhou, X. Wu, W. Xu, Z. Yang, Methylation mediated silencing of mir-23b expression and its role in glioma stem cells, Neuroscience letters 528 (2) (2012) 185–189.
[59] Y.-Q. Liao, Y.-L. Liao, J. Li, L.-X. Peng, Y.-Y. Wan, R. Zhong, Polymorphism in mir-146a associated with clinical characteristics and outcomes in gastric cancer patients treated with adjuvant oxaliplatin and fluoropyrimidines, OncoTargets and therapy 8 (2015) 2627.
[60] E.I. Papadopoulos, G.M. Yousef, A. Scorilas, Gemcitabine impacts differentially on bladder and kidney cancer cells: distinct modulations in the expression patterns of apoptosis-related micrornas and bcl2 family genes, Tumor Biology 36 (5) (2015) 3197–3207.
[61] R. Hummel, T. Wang, D.I. Watson, M.Z. Michael, M. Van der Hoek, J. Haier, D. J. Hussey, Chemotherapy-induced modification of microrna expression in esophageal cancer, Oncology reports 26 (4) (2011) 1011–1017.