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

https://doi.org/10.1093/bib/bbab174
Problem Solving Protocol

Multi-view Multichannel Attention Graph Convolutional Network for miRNA–disease association prediction

Xinru Tang, Jiawei Luo, Cong Shen and Zihan Lai

Corresponding author: Jiawei Luo, College of Computer Science and Electronic Engineering, Hunan University, Changsha 410083, China.
E-mail: luojiawei@hnu.edu.cn

Abstract

Motivation: In recent years, a growing number of studies have proved that microRNAs (miRNAs) play significant roles in the development of human complex diseases. Discovering the associations between miRNAs and diseases has become an important part of the discovery and treatment of disease. Since uncovering associations via traditional experimental methods is complicated and time-consuming, many computational methods have been proposed to identify the potential associations. However, there are still challenges in accurately determining potential associations between miRNA and disease by using multisource data.
Results: In this study, we develop a Multi-view Multichannel Attention Graph Convolutional Network (MMGCN) to predict potential miRNA–disease associations. Different from simple multisource information integration, MMGCN employs GCN encoder to obtain the features of miRNA and disease in different similarity views, respectively. Moreover, our MMGCN can enhance the learned latent representations for association prediction by utilizing multichannel attention, which adaptively learns the importance of different features. Empirical results on two datasets demonstrate that MMGCN model can achieve superior performance compared with nine state-of-the-art methods on most of the metrics. Furthermore, we prove the effectiveness of multichannel attention mechanism and the validity of multisource data in miRNA and disease association prediction. Case studies also indicate the ability of the method for discovering new associations.

Key words: miRNA–disease associations; multiview; graph convolutional networks; deep learning

Introduction

MicroRNAs (miRNAs) are a kind of small endogenous noncoding RNAs that play critical roles in multiple biological processes, including cell development, proliferation differentiation, immune reaction and so on [1–3]. Their abnormal expression can directly or indirectly cause changes in the expression level of their regulated target genes, leading to complex diseases [4–6]. Therefore, to understand the pathogenesis of diseases, it is necessary to identify more associations between miRNAs and diseases.

So far, several wet experiment methods, such as qPT-PCR [7] and Northern blot [8], can identify the disease-associated miRNA by reflecting the true expression level of miRNA in cells. However, determining the associations between miRNAs

Xinru Tang is a master's student in the College of Computer Science and Electronic Engineering, Hunan University. Her research interests include deep learning and bioinformatics.
Jiawei Luo is a full-time professor in the College of Computer Science and Electronic Engineering, Hunan University. Her research interests include graph theory, data mining, computational biology and bioinformatics.
Cong Shen is a PhD student at Hunan University. His current research interests include pharmacogenomics, biological networks and computational systems biology.
Zihan Lai is a master's student in the College of Computer Science and Electronic Engineering, Hunan University. Her research interests include network pharmacology.
Submitted: 1 March 2021; **Received (in revised form):** 8 April 2021

and diseases by biological experiments is time-consuming and expensive [9]. As a result, it is necessary to develop computational methods to identify disease-related miRNAs. According to previous studies [10], existing computational methods can be divided into two categories, i.e. similarity measure-based methods and machine learning-based methods.

The similarity-based approach is based on the assumption that miRNAs with similarity functions are more likely to be associated with similar diseases and vice versa. To well utilize the known information, Jiang *et al.* [11] considered the miRNA–disease association prediction problem as the link prediction problem on miRNA–disease networks. The scholars proposed the first computational method that used a discrete hypergeometric probability distribution to identify the miRNA–disease association. Later on, to achieve better prediction results, Xuan *et al.* [12] designed a computational model HDMP that evaluated the *k* most functionally similar neighbors to explore the possibility that miRNAs are associated with the disease. However, the above methods only considered the direct edge information in the involved networks, neglecting the global structure of these networks. Based on global network information, some researchers [13–15] predict miRNA–disease associations by using the random walk with restart algorithm. Although many attempts have achieved significant performances in the aspect of finding novel associations, for those diseases where there are few or no known related miRNAs, most of the methods cannot infer potentially related miRNAs. To solve this problem, many new methods have been proposed, NetGS [9] explored the global network similarity and used diffusion profile consistency to capture the relationship between diseases and miRNAs. It can solve the new miRNA–disease association prediction problem with reliable accuracy. Moreover, Xiao *et al.* [16] proposed a graph regularized non-negative matrix factorization approach, which is effective to infer unknown miRNA–disease associations for those novel diseases and miRNAs. Nevertheless, it has the problem of being sensitive to neighborhood information and strongly dependent on the input data source. To this end, Zhang *et al.* [17] formulated a model that could learn the neighbor information for each node adaptively to update the latent factor of the node, and further promote the effectiveness of prediction on potential links. Although these similarity-based models perform well, many complex and nonlinear relationships between miRNAs and diseases networks remain difficult to be captured.

Recently, machine learning-based models have been proposed to better extract features, and thereby improve prediction performance. Chen *et al.* put forward inductive matrix completion (IMCMDA) [18] and ensemble learning method (EDTMDA) [19] successively to learn complicated nonlinear relationships between networks. However, with the rapid growth of associated data in recent years, traditional machine learning algorithms cannot adapt to complex and changeable data. Deep learning methods have been widely used in bioinformatics applications, for their ability to process unstructured data quickly and efficiently. For example, CNNMDA [20] utilized dual convolutional neural networks (CNN) to learn the original and global representation of a miRNA–disease pair. To overcome the problem of 'catastrophic forgetting', MISSIM [21] introduced incremental learning into the field of bioassociation prediction for the first time and achieved superior performance on the miRNA disease association prediction task. While graph-based methods, such as GCN [22], GAT [23], etc., can use contextual information to make the model have splendid performance. Now there are some studies that exploited graph neural networks to learn the context representation of nodes, such as GCN-MF [24], IDDkin [25], GAERF [26], GCNCDA [27], and they all achieved outstanding performance. Nevertheless, most of the models simply employ single information or perform simple information filling to integrate multiple sources of information. It is still a challenge on how to effectively use multisource information to improve the accuracy of prediction. Recent studies [28–30] have shown that the incorporation of multisource information is widely embraced in the biological fields. For example, MDA-SKF [31] is a similarity-kernel-fusion-based methodology for integrating multiple similarity kernels of miRNA and disease. However, MDA-SKF ignored contextual information in the graph. To better learn the node features, NeoDTI [32] integrated diverse information from heterogeneous network data and used graph neural network to automatically learn topology-preserving representations of drugs and targets. IMDA-BN [33] integrated nine relationship types and utilized graph embedding algorithm to obtain the network embedded representation of nodes for miRNA–disease interactions prediction. Integrating multisource information into miRNA–disease association prediction may improve the performance. Besides, the contribution of different information to the prediction task is not necessarily the same [34]. Thus, proposing a method that can learn multiview features adaptively may provide a new perspective for miRNA disease association prediction.

In this study, we present a Multi-view Multichannel Attention Graph Convolutional Network named MMGCN for miRNA–disease association prediction. First, we use the graph convolutional network (GCN) to learn the multiview features of miRNA and disease from their multiple similarity networks, respectively. Second, instead of simply integrate multisource information, we utilize the attention mechanism on multiple features to learn the importance of different features adaptively. Then the CNN combiner is used to obtain the final embedding. Finally, we regard the problem of disease-related miRNA identification as a recommender task and use matrix completion to predict potential miRNA–disease associations. To evaluate the effectiveness of MMGCN, we compare it with some state-of-the-art approaches on two benchmark datasets under 10 times 5-fold cross-validation. Furthermore, to demonstrate the necessity of each stage, we conduct ablation study and compare the results of our methods under different views. Finally, case studies are used to illustrate the strong predictive power of MMGCN in discovering novel associations between diseases and miRNAs.

Methods

Model framework

In this section, we propose a novel multiview deep learning framework named MMGCN to predict miRNA–disease associations. We frame the miRNA–disease prediction problem as a recommendation task by fusing multisource similarity views of miRNA and disease. More specifically, as shown in Figure 1, MMGCN is an end-to-end model that consists of the following three modules: (i) A multiview GCN encoder for encoding miRNAs and disease multisource data (Figure 1A). (ii) A multichannel attention mechanism to separately learn the importance of the different channel information for miRNA and disease learned from the GCN encoder (Figure 1B and C). (iii) A CNN combiner that recombines the features from the multichannel attention module to obtain a unified embedding of miRNA and disease (Figure 1D). After obtaining the feature embeddings of miRNA

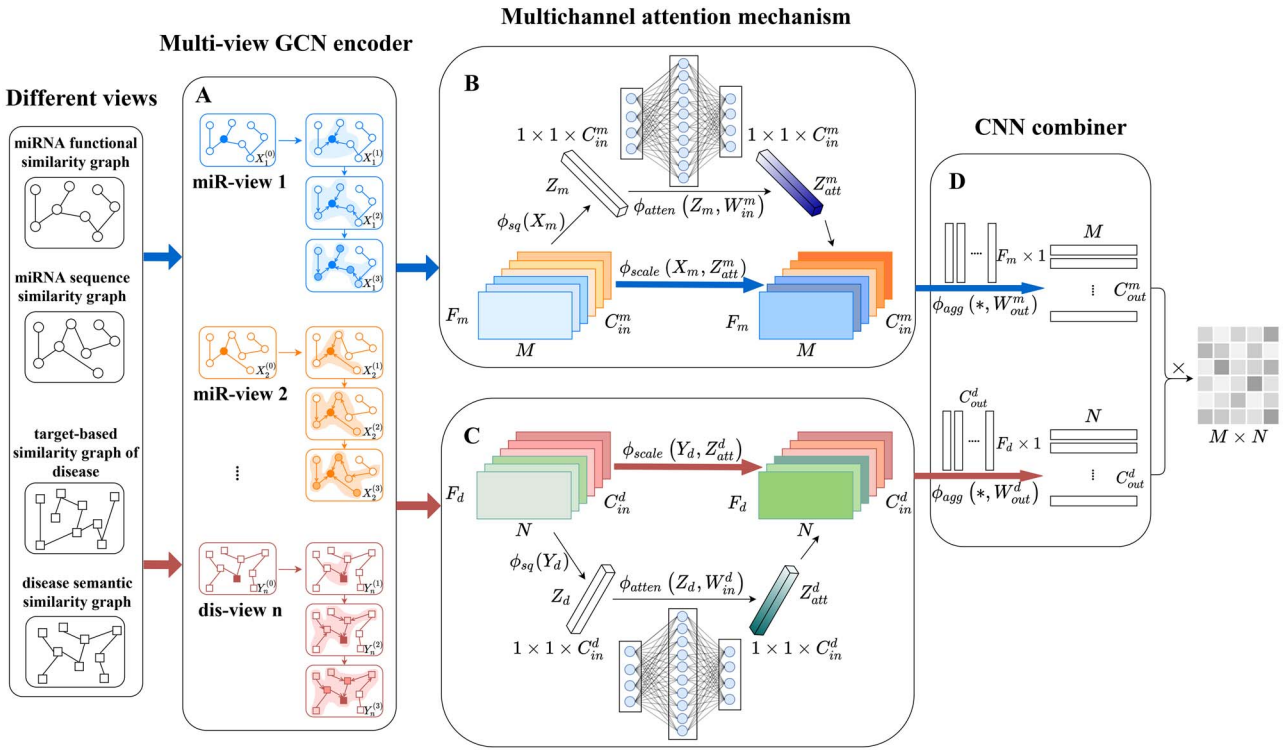

Figure 1. Overview of MMGCN model architecture. (A) Multi-view GCN encoder. The GCN encoder takes multiple similarity graphs of miRNA and disease nodes as input, fuses their neighbor information, and generates miRNA and disease embeddings under different views, respectively. (B) Multichannel attention mechanism on miRNA. It is leveraged to focus on the more important miRNA channel embedding, and get normalized channel attention features. (C) Multichannel attention mechanism on disease. Its function is similar to that of (B), which is used to capture important information on the disease feature channel. (D) CNN combiner. The multichannel attention features of miRNA and disease are convolved respectively to obtain their corresponding representations for association prediction.

and disease, respectively, we use matrix factorization to obtain the potential correlation matrix of miRNA and disease.

Similarity measures

In this section, we utilize a variety of biological data sources to comprehensively characterize similarity views of miRNAs and diseases, which are calculated in detail below.

MiRNA sequence similarity

We figure out miRNA seed region sequence similarity matrix M_s by using Needleman–Wunsch Algorithm [35]. To ensure global consistency, we get the normalized relation $NM_s(m_i, m_j)$ between m_i and m_j as follows:

$$NM_s(m_i, m_j) = \frac{M_s(m_i, m_j) - M_{s_{min}}}{M_{s_{max}} - M_{s_{min}}} \quad (1)$$

where $M_{s_{min}}$ and $M_{s_{max}}$ represent the minimum and maximum score in miRNA sequence similarity matrix M_s, respectively. Then, the sequence similarity between miRNAs m_i and m_j can be written as follows:

$$G_m^s(m_i, m_j) = \begin{cases} 1 & m_i = m_j \\ NM_s(m_i, m_j) & m_i \neq m_j \end{cases} \quad (2)$$

$G_m^s \in R^{M \times M}$ can be regarded as a sequential view of miRNAs, where M denotes the number of miRNAs.

MiRNA functional similarity

To avoid reliance on existing associations between miRNA and disease, we follow previous works [16], using associations between miRNAs and genes to calculate miRNA functional similarity.

The gene functional interaction network is obtained from HumanNet [36] which contains the probability of a functional linkage between genes. It is called log-likelihood scores (*LLS*). We first obtain the similarity between genes LLS_N by applying min–max normalization to *LLS*. And the gene similarity graph is defined as follows:

$$S(g_i, g_j) = \begin{cases} 1 & g_i = g_j \\ 0 & g_i \neq g_j \cap e(g_i, g_j) \notin S_{HumanNet} \\ LLS_N(g_i, g_j) & g_i \neq g_j \cap e(g_i, g_j) \in S_{HumanNet} \end{cases} \quad (3)$$

where $S_{HumanNet}$ contains all links in HumanNet; $e(g_i, g_j)$ is the edge between gene g_i and g_j. Then, we calculate the similarity score between a gene g_s and a set of genes $GS = \{g_{s1}, g_{s2}, \ldots, .g_{sk}\}$. as follows:

$$S(g_s, GS) = \max_{1 \le i \le k}(S(g_s, g_{si})) \quad (4)$$

Subsequently, we can get the functional similarity between miRNAs m_i and m_j by Equation (5).

$$G_m^f(m_i, m_j) = \frac{\sum_{g \in G_i} S(g, GS_j) + \sum_{g \in G_j} S(g, GS_i)}{|G_i| + |G_j|} \tag{5}$$

where G_i and G_j represents the corresponding gene sets that are related to miRNAs m_i and m_j, respectively. Finally, we obtain the view of miRNA functional similarity $G_m^f \in R^{M \times M}$.

Disease semantic similarity

To describe the relationship between diseases, we use the Directed Acyclic Graph (DAG) to compute their semantic similarities in the same way as Wang *et al.* [37]. Consequently, the DAG graph of disease d can be defined as $DAG(d) = (d, T(d), E(d))$, in which $T(d)$ denotes a set of diseases that contains all the ancestors of d including d, and $E(d)$ is the link set of the DAG graph of d. Therefore, the semantic contribution of a disease t in $DAG(d)$ to disease d can be computed as follows:

$$\begin{cases} D_d(t) = 1 \text{ if } t = d \\ D_d(t) = \max\{\Delta * D_d(t') | t' \in \text{children of } t\} \text{ if } t \neq d \end{cases} \tag{6}$$

where Δ is the semantic contribution factor ($\Delta = 0.5$) [37]. It means that as the distance between disease t and its ancestor is farther, the semantic contribution of disease t to disease d is lower. After that, the similarity between different disease d_i and d_j can be evaluated as follows:

$$G_d^s(d_i, d_j) = \frac{\sum_{t \in T(d_i) \cap T(d_j)} \left(D_{d_i}(t) + D_{d_j}(t)\right)}{\sum_{t \in T(d_i)} D_{d_i}(t) + \sum_{t \in T(d_j)} D_{d_j}(t)} \tag{7}$$

where $D_{d_i}(t)$ and $D_{d_j}(t)$ denotes the semantic contribution of disease t to disease d_i and d_j, respectively. Matrix $G_d^s \in R^{N \times N}$ can be considered as the semantic view of disease, where N denotes the number of diseases.

Target-based similarity measure for diseases

The target-based similarity measure for diseases is similar to the calculation of miRNA functional similarity, using associations between diseases and genes. Analogously, the formula of similarity between disease d_i and d_j can be defined as:

$$G_d^f(d_i, d_j) = \frac{\sum_{g \in G_i} S(g, GS_j) + \sum_{g \in G_j} S(g, GS_i)}{|GS_i| + |GS_j|} \tag{8}$$

Similar to miRNA functional similarity, GS_i and GS_j represents gene sets that are related to diseases d_i and d_j, respectively. Then we obtain the view of disease target-based similarity $G_d^f \in R^{N \times N}$.

Multiview GCN encoder

GCN is a powerful graph neural network, which acts directly on the graph and utilizes the graph structural information. It can aggregate information about neighbors, captures dependencies between data and generates a useful representation of nodes in the graph. To capture the structure information of multiple similarity views, we use GCN to encode different views respectively. As shown in Figure 1A, GCN is leveraged to obtain the embeddings of miRNA and disease in different views. We define the sth view of miRNA and the qth view of disease as G_m^s and G_d^q, respectively. Given $x_i \in R^{1 \times p}$ as the feature of miRNA i, the features $\{x_{i_1}, x_{i_2}, \ldots, x_{i_k}\}$ belongs to its neighbor in view s: $\{i_1, i_2, \ldots, i_k\}$, respectively. In addition to the neighbor information, that of the node itself cannot be ignored when learning the embedding. For the message passing from miRNA i and its neighbors, we can get the representation of i in view s as follows:

$$x_i' = \sigma\left(\left(\tilde{s}_{i,i} x_i + \sum_{j=1}^{k} \tilde{s}_{i,j} x_{i_j}\right) W_i\right) \tag{9}$$

where $\tilde{s}_{i,j}$ means the corresponding normalized similarity weight between miRNA i and its neighbor i_j in view s; $W_i \in R^{p \times F_m}$ is a learnable parameter to project the features into the latent space and $\sigma(\bullet)$ denotes the nonlinear activation function.

After getting the propagation rule for a single node using GCN on view s, we can obtain the miRNA nodes embedding on the entire graph G_m^s:

$$X_s^{(l+1)} = \sigma\left(\tilde{D}_s^{-\frac{1}{2}} \tilde{S} \tilde{D}_s^{-\frac{1}{2}} X_s^{(l)} W_s^{(l)}\right) \tag{10}$$

where $X_s^{(l)} \in R^{M \times F_m}$ denotes the F_m dimension embeddings of M miRNA nodes that obtained by the lth GCN layer on view s, in particular, $X_s^{(0)}$ is a randomly initialized embedding; $W_s^{(l)} \in R^{F_m \times F_m}$ is a learnable matrix, $\tilde{D}_s$ denotes the diagonal matrix, $\tilde{D}_s(i,i) = \sum_j \tilde{S}(i,j)$. $\tilde{S}$ is the corresponding normalized similarity weight matrix on view s. Since we add a self-loop to ensure the importance of the node itself, the formula is constructed as follows:

$$\tilde{S} = I + S \tag{11}$$

Similarly, we can get disease nodes embedding on the qth view G_d^q:

$$Y_q^{(l+1)} = \sigma\left(\tilde{D}_q^{-\frac{1}{2}} \tilde{Q} \tilde{D}_q^{-\frac{1}{2}} Y_q^{(l)} W_q^{(l)}\right) \tag{12}$$

analogously, $Y_q^{(l)} \in R^{N \times F_d}$ denotes the F_d dimension embeddings of N disease nodes that attained by the lth GCN layer on view q, in particular, $Y_q^{(0)}$ is a randomly initialized embedding; $W_q^{(l)} \in R^{F_d \times F_d}$ is a learnable matrix, $\tilde{D}_q$ denotes the diagonal matrix, $\tilde{D}_q(i,i) = \sum_j \tilde{Q}(i,j)$. $\tilde{Q}$ represents the similarity weight matrix on view q after adding the self-loop:

$$\tilde{Q} = I + Q \tag{13}$$

For miRNAs and diseases in a single view, we can get their embedding as follows:

$$\left\{X_s^{(1)}, X_s^{(2)}, \ldots, X_s^{(l)}\right\} \tag{14}$$

$$\left\{Y_q^{(1)}, Y_q^{(2)}, \ldots, Y_q^{(l)}\right\} \tag{15}$$

where l represents the number of layers of GCN.

By using multiple layers of GCN to encode multiviews of miRNAs and diseases, we can obtain the features of miRNAs and diseases from different perspectives. For miRNAs of S views and disease of Q views, their features can be shown as:

$$\left\{\left\{X_1^{(1)}, X_1^{(2)}, \ldots, X_1^{(l)}\right\}, \left\{X_2^{(1)}, X_2^{(2)}, \ldots, X_2^{(l)}\right\}, \ldots, \left\{X_S^{(1)}, X_S^{(2)}, \ldots, X_S^{(l)}\right\}\right\} \quad (16)$$

$$\left\{\left\{Y_1^{(1)}, Y_1^{(2)}, \ldots, Y_1^{(l)}\right\}, \left\{Y_2^{(1)}, Y_2^{(2)}, \ldots, Y_2^{(l)}\right\}, \ldots, \left\{Y_Q^{(1)}, Y_Q^{(2)}, \ldots, Y_Q^{(l)}\right\}\right\} \quad (17)$$

After getting multiple embeddings under different views, we shall introduce how to focus on more important feature information in the next step.

Multichannel attention

In the user–item recommendation task of social networks, the user's purchasing behavior is driven by one or more motives, and the features of the item often include multiple aspects. A single perspective embedding may not achieve the expected effect for the recommendation task. Therefore, in such tasks, researchers tend to obtain the embeddings of users and items from multiple perspectives, and different perspectives should have different contributions when learning their embeddings [38]. With the same way, for miRNAs and diseases, node features extracted from different similarity views may contain distinct context information, thus we introduce attention mechanism to focus on more important features.

As shown in Figure 1B, the multiple feature matrices of miRNA are stacked to form a feature tensor.We regard the stacked miRNA feature tensor as an image, and each feature matrix of miRNA can be seen as a channel on the image. Then the task of acquiring the importance of features from multiple views can be transformed into getting the importance of multiple channels in the task of image feature extraction. Inspired by Hu *et al.* [39], we use attention on the channel to assign weights to features.

To obtain the importance of different channels, we first utilize global average pooling to generate channel-based statistics. For miRNAs with C_{in}^m channels, a statistic $Z \in R^{1\times 1\times C_{in}^m}$ is generated by squeezing miRNA's features $X_m = [x_1, x_2, \ldots, x_{C_{in}^m}], X_m \in R^{F_m\times M\times C_{in}^m}$. Specifically, for the cth feature matrix of miRNA x_c, the channel statistic z_c is calculated as:

$$z_c = \phi_{sq}(x_c) = \frac{1}{F_m \times M}\sum_{i=1}^{F_m}\sum_{j=1}^{M} x_c(i,j) \quad (18)$$

To fully capture channel importance, we employ attention mechanism to compute the attention weights of channels:

$$Z_{att} = \phi_{atten}\left(Z, W_{in}^m\right) = \delta\left(W_2\sigma\left(W_1 Z\right)\right) \quad (19)$$

where $\delta(\bullet)$ is Sigmoid activation, $\sigma(\bullet)$ is Relu activation, $W_{in}^m = \{W_1, W_2\}$ is the training parameter. Finally, multichannel attention can be defined as $Z_{att} = [z_1^{att}, z_2^{att}, \ldots, z_{C_{in}^m}^{att}]$.

After getting the attention weight of different channels, we combine the channel features with attention to normalize it, which is defined as follows:

$$\tilde{x}_c = \phi_{scale}\left(x_c, z_c^{att}\right) = z_c^{att} \bullet x_c \quad (20)$$

Through the above steps, we can obtain the normalized miRNA channel information $\tilde{X}_m = [\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_{C_{in}^m}]$. Similarly, the disease channel information $\tilde{Y}_d = [\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_{C_{in}^d}]$.

Next, we will introduce how to combine the channel information to learn the final miRNA and disease embedding.

CNN combiner

Since multiple convolution kernels of CNN can help us learn complex nonlinear relations in the feature of nodes, after obtaining the multichannel features of miRNA, we utilize CNN to extract features and generate the final miRNA embeddings that integrate multiple view information. Given miRNA channel information $\tilde{X}_m = [\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_{C_{in}^m}]$, the final embedding $X'_m \in R^{C_{out}^m\times M}$ is defined as follows:

$$Cout_k = \phi_{agg}\left(\tilde{X}_m\right) = bias_k + \sum_{i=1}^{C_{in}^m}\tilde{x}_i * w_k^m \quad (21)$$

$$X'_m = stack\left(Cout_k\right) \quad (22)$$

where $w_k^m \in R^{F_m\times 1}$ denotes convolution filter, it belongs to set $W_{out}^m = \{w_1^m, w_2^m, \ldots, w_{C_{out}^m}^m\}$, and $*$ denotes the convolution operator. $Cout_k \in R^{1\times M}$ means the embedding from the kth output channel, where k= $1, 2\ldots, C_{out}^m$. The final miRNA embedding X'_m is obtained by stacking the embeddings of multiple output channels.

Similarly, we can obtain disease final embedding Y'_d. These learned embeddings will be used as the input in the final miRNA–disease association prediction task.

Association prediction for miRNA and disease

We regard miRNA and disease association prediction as a recommendation task, and the preference matrix $U \in R^{M\times N}$ can be computed as:

$$U = {X'_m}^T Y'_d \quad (22)$$

for the elements in U, the higher U_{ij} is, the more likely miRNA i is to be associated with disease j, otherwise, the less likely miRNA i is to be associated with disease j.

We use mean square error by minimizing the Frebious norm of the difference between preference matrix U and label matrix U' as the loss function of our model. The loss is formally formulated as follows:

$$L = \left\| U - U' \right\|_F^2 \quad (23)$$

We test our codes on a machine equipped with one NVIDIA 2060 GPU and one 3.60 GHz Intel(R) Core(TM) i7-9700K CPU with 32 GB memory, and the running time of MMGCN on HMDDV3.2 dataset is 753 s.

Experiments

Data preparation

We use multiple biomedical datasets to extract the multisource information of miRNAs and diseases, and build a network of multiple views for subsequent association prediction tasks. Below we describe the process of building a multiview network in our study.

Known human miRNA–disease associations verified by experimental evidences in literature are obtained from HMDD v3.2 [40], which includes 35 547 associations between 1206

miRNAs and 894 diseases. The miRNA sequence information is downloaded from miRBase Release 22 [41], which contains 1917 human miRNAs. The miRTarBase Release 8.0 [42] contains 502 652 pairs of miRNA–gene relationships, including 2599 miRNAs. As done in previous studies [34, 37], we merge multiple miRNA transcripts with the same mature miRNA. Then we intersect the datasets and obtain 905 disease-associated miRNAs that have sequence information. The semantic trees of diseases and disease–gene relations are downloaded from MESH (https://www.nlm.nih.gov/mesh/) and DisGeNET v7.0 [43]. The two databases contain 29 638 and 30 170 diseases, respectively. After merging duplicates and removing the irregular data, we get 12 446 experimentally validated miRNA–disease associations among 853 miRNAs and 591 diseases. The weighted gene–gene association network is got from HumanNet v2 [36]. For miRNAs and diseases we obtain above, there are 14 971 and 19 258 genes associated with them, respectively.

Besides, to better evaluate the performance of the model, we use the miR2disease database [44] to build a small multiview network. MiR2disease contains 2925 experimentally validated associations between 438 miRNAs and 174 diseases. We use the same method as above to perform intersection and deduplication of the datasets, and finally get 701 miRNA–disease associations, including 193 miRNAs and 54 diseases.

Experiment settings

We evaluate the performance of MMGCN to predict potential disease-related miRNAs by performing 5-fold cross-validation. In each round, the known miRNA–disease associations are considered as positive samples and randomly divided into five disjoint subsets. One subset is considered as the testing set and the remaining is utilized as the training set. The number of GCN layer is selected in {2, 3, 4}, embedding size from miRNA and disease is selected from {32, 64, 128, 256}, the number of filters is set from {32, 64, 128, 256} and learning rate is chosen in {0.1, 0.01, 0.001}. We evaluate the performance of MMGCN and the baseline models on HMDD and miR2disease dataset, and the parameters used by the other baselines are in the supplementary materials. The analysis experiments of the MMGCN model are based on the HMDD dataset. All experiments are repeated 10 times to obtain a sound estimate of prediction results.

Evaluation metrics include area under the receiver operating characteristic (ROC) curve (AUC), area under the precision/recall (PR) curve (AUPRC), accuracy, precision, recall and F1-score. Moreover, it is important to guide wet experiments through the top-ranked disease-associated miRNAs obtained based on the computational model. Therefore, we use precision@N and recall@N to verify the performance of MMGCN and other methods within the top N. We set $N = \{5\%, 10\%\}$ in all experiments. The results reported are average of the 10 runs.

Baselines

To demonstrate the effectiveness of our method, we compare it with the following baselines. In this study, the comparison methods and the MMGCN model are run on the same dataset.

IMCMDA [18]: The method integrated miRNA functional similarity, disease semantic similarity, and Gaussian interaction profile kernel similarity, and used the inductive matrix completion to generate associations.

SPM [34]: Zeng *et al*. constructed a bilayer network by integrating the miRNA–disease association network and their similarity networks and then utilized the structural perturbation method to discover potential associations between miRNA and disease.

MDHGI [45]: It is a computational model of matrix decomposition and heterogeneous graph inference for discovering new miRNA–disease associations.

GRGMF [17]: The method utilized similarity information as Laplacian regularized terms, and proposed graph regularized generalized matrix factorization to identify potential links in biomedical bipartite networks.

NIMCGCN [46]: Li *et al*. performed GCN on miRNA similarity network and disease similarity network, respectively, and introduced a neural inductive matrix completion method to predict miRNA–disease associations.

MDA-SKF [31]: The method proposed a novel similarity kernel fusion to integrate multiple miRNA and disease similarity kernels, respectively. Then utilized the Laplacian regularized least squares method to uncover potential miRNA–disease associations.

KBMFMDA [47]: It is a neoteric Bayesian model that combined kernel-based nonlinear dimensionality reduction, matrix factorization and binary classification. The method projected miRNAs and diseases into a unified subspace then predict potential miRNAs for diseases.

SNMDA [48]: It utilized the sparsity of the miRNA disease association network and presented a novel approach to predict the potential microRNA disease associations based on sparse neighborhood.

LLCMDA [49]: It is an approach based on Locality-constrained Linear Coding (LLC). The method reconstructed similarity networks for both miRNAs and diseases using LLC and then applied label propagation on the similarity networks to get reliable relevant scores for the final output.

Performance comparison

Here, we use 10 times 5-fold cross-validation to evaluate the performance of MMGCN with the other baselines on HMDD and miR2disease datasets.

As shown in Figure 2 and 3 and Tables 1 and 2, our proposed MMGCN method can achieve competitive performance on both two datasets. For HMDD dataset, MMGCN outperforms all the other methods under all evaluation metrics. Due to the sparsity of the data, the precision and recall values of MMGCN are relatively low. But MMGCN is still better than the comparison method, which may attribute to the use of the attention mechanism in multiple perspectives, i.e. the Multichannel attention mechanism enables MMGCN to capture richer and more important features. It demonstrates the reasonable design of our model. Compared with the single-view approach (GRGMF), our MMGCN uses multiple similar views to enable the model to learn more information. Compared with the methods that simply combine different similar views into one (IMCMDA, SPM, MDHGI, NIMCGCN, KBMFMDA, SNMDA and LLCMDA), MMGCN adaptively learns the weight of different views' features through a well-designed channel attention mechanism to focus on more important features. Although MDASKF is a multikernel fusion method that uses three similar information of miRNAs and diseases to make association prediction, which combines more information than MMGCN, the performance is still not as good as that of MMGCN. It may be because MMGCN uses graph neural network to capture the context information of nodes and has better representation ability. For miR2disease dataset, Figure 3 and Table 2 show that MMGCN can also perform well on a small dataset. It is better than other comparison methods in most

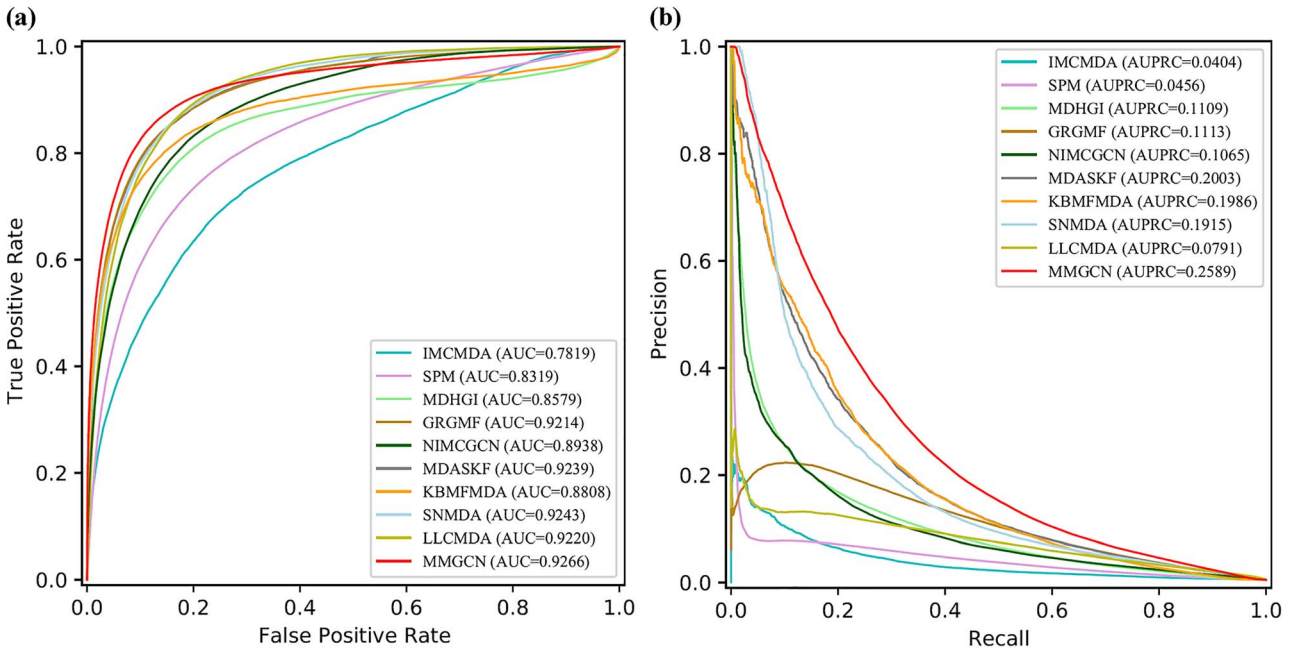

Figure 2. (a) Comparison of ROC curve and AUC values between MMGCN and other existing methods on HMDD dataset. (b) Comparison of precision/recall curve and AUPRC value between MMGCN and other existing methods on HMDD dataset.

Table 1. Performance of compared methods on HMDD dataset

Method	Accuracy	Precision	Recall	F1	precision@5%	precision@10%	recall@5%	recall@10%
IMCMDA	0.9885	0.0889	0.1365	0.1074	0.0350	0.0236	0.3474	0.4705
SPM	0.9815	0.0714	0.2168	0.1068	0.0437	0.0294	0.4338	0.5847
MDHGI	0.9901	0.1624	0.2224	0.1861	0.0555	0.0342	0.5512	0.6791
GRGMF	0.9894	0.1747	0.2904	0.2174	0.0659	0.0394	0.6544	0.7835
NIMCGCN	0.9909	0.1681	0.1979	0.1809	0.0544	0.0347	0.5406	0.6893
MDASKF	0.9929	0.2798	0.2575	0.2677	0.0657	0.0393	0.6526	0.7806
KBMFMDA	0.9929	0.2881	0.2562	0.2680	0.0635	0.0375	0.6304	0.7452
SNMDA	0.9924	0.2474	0.2470	0.2471	0.0630	0.0283	0.6261	0.7741
LLCMDA	0.9851	0.1141	0.2811	0.1613	0.0597	0.0381	0.5936	0.7576
MMGCN	0.9934	0.3338	0.2989	0.3146	0.0705	0.0413	0.7005	0.8200

cases. Overall, these experimental results indicate the effectiveness of MMGCN in predicting potential relations of miRNA and disease.

Ablation study

To analyze the necessity of components of our model, we adopt two variants of MMGCN (MMGCN-nl and MMGCN-noatten) as comparison methods. Specifically, MMGCN-nl means we only use the output of GCN at the *n*th layer as the feature. MMGCN-noatten removes the multichannel attention and regards the different channel information as equally important, which only uses the CNN combiner to fuse the features.

Table 3 shows the average evaluation metrics obtained with MMGCN and its variant models. For MMGCN-nl and MMGCN, after incorporating the feature output by each layer of GCN, compared with a single layer of information, the model obtains more disease and miRNA structural features. This result shows that the neighbors at different distances in the network may contain various information. For MMGCN-noatten and MMGCN, after adding adaptive attention mechanism to different features, compared with a simple combination, MMGCN can get better final embedding of miRNA and disease. The result of this experiment proves that the importance of the feature information in different views is different. Using a multichannel attention mechanism can enhance the representation of nodes to improve prediction performance.

Performance of different view fusion

To demonstrate the effectiveness of multiview learning, we compare the results of our approaches leveraging different views.

As shown in Table 4, MMGCN can achieve the best performance using all views on miRNA and disease similarity networks. Moreover, in the case that miRNA similarity views are consistent, the approach would have a better performance by adding more views than it with only a single view on the disease similarity network, vice versa. A possible reason is that the addition of different view information enriches the characteristics of the nodes. However, the model may also have the problem of performance degradation caused by adding one more view. For instance, the accuracy of MMGCN under the view of miRNA functional similarity (G_m^f) and disease semantic similarity (G_d^s) is 0.9932, while the result decreases a little (0.9931) after adding disease target-based similarity (G_d^f). It may be caused by the

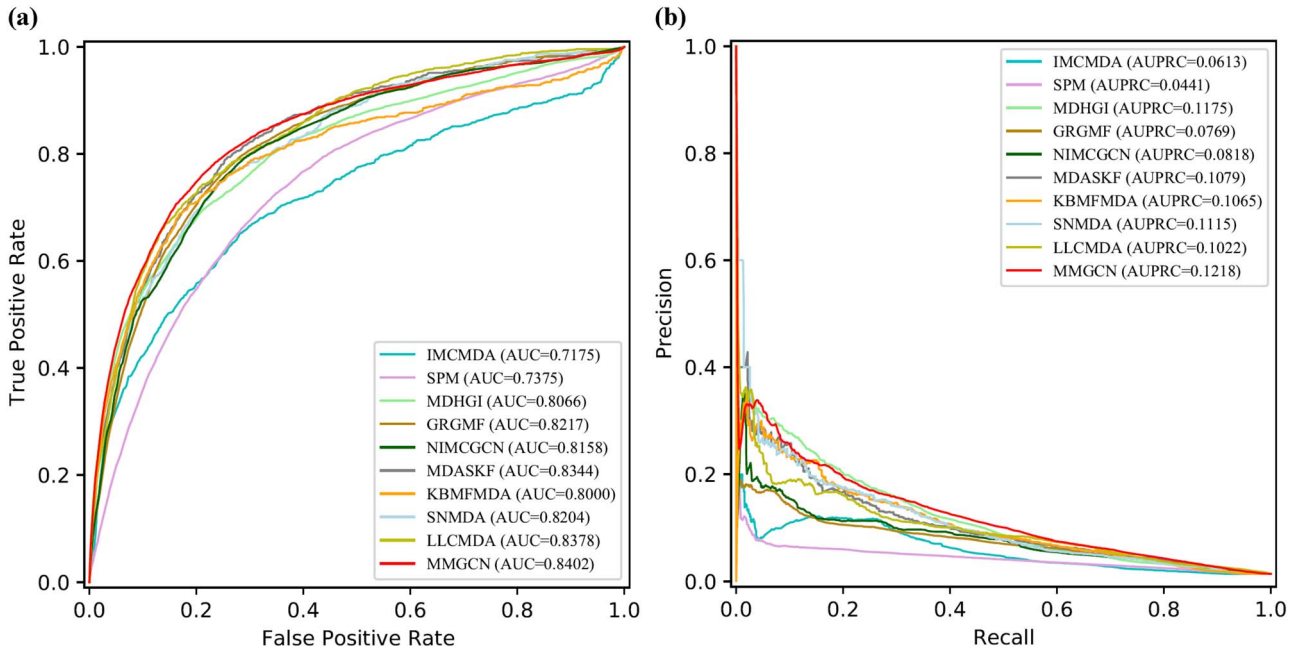

Figure 3. (a) Comparison of ROC curve and AUC values between MMGCN and other existing methods on miR2disease dataset. (b) Comparison of precision/recall curve and AUPRC value between MMGCN and other existing methods on miR2disease dataset.

Table 2. Performance of compared methods on miR2disease dataset

Method	Accuracy	Precision	Recall	F1	precision@5%	precision@10%	recall@5%	recall@10%
IMCMDA	0.9607	0.1180	0.2734	0.1652	0.0885	0.0594	0.3130	0.4192
SPM	0.9377	0.0697	0.2447	0.1032	0.0592	0.0496	0.2094	0.3505
MDHGI	0.9728	0.1989	0.2712	0.2207	0.1146	0.0753	0.4052	0.5314
GRGMF	0.9571	0.1213	0.2757	0.1567	0.0906	0.0704	0.3205	0.4974
NIMCGCN	0.9592	0.1192	0.2868	0.1668	0.0968	0.0730	0.3425	0.5155
MDASKF	0.9741	0.1794	0.2185	0.1932	0.1042	0.0753	0.3685	0.5314
KBMFMDA	0.9707	0.1890	0.2685	0.2096	0.1090	0.0771	0.3857	0.5442
SNMDA	0.9700	0.1770	0.2857	0.2126	0.1094	0.0732	0.3871	0.5171
LLCMDA	0.9724	0.1653	0.2261	0.1886	0.1030	0.0797	0.3642	0.5630
MMGCN	0.9706	0.1882	0.2930	0.2215	0.1195	0.0811	0.4228	0.5724

Table 3. Performance of MMGCN and its variants

Models	MMGCN-nl	MMGCN-noatten	MMGCN
AUPRC	0.1121	0.2279	0.2589
AUC	0.8816	0.9195	0.9266
Accuracy	0.9914	0.9933	0.9934
Precision	0.1833	0.3148	0.3338
Recall	0.1984	0.2738	0.2989
F1	0.1892	0.2919	0.3146
precision@5%	0.0544	0.0684	0.0705
precision@10%	0.0342	0.0403	0.0413
recall@5%	0.5406	0.6794	0.7005
recall@10%	0.6790	0.8017	0.8200

incorporated noise when adding views. But in most cases, the addition of multiple views can effectively improve the performance of the model.

In summary, adding multiple views makes MMGCN more efficient, especially when all views are added to the model, which can achieve the best performance.

Parameter analysis of MMGCN

In this section, we will analyze some of the parameters in our model to illustrate their impact. All experiments are repeated 10 times in 5-fold cross-validation on known miRNA–disease associations to obtain a more accurate result. In the following,

Table 4. Performance of MMGCN with different views

Views	$G_m^f+G_d^f$	$G_m^f+G_d^s$	$G_m^s+G_d^s$	$G_m^s+G_d^f$	$G_m^sG_m^f+G_d^f$	$G_m^sG_m^f+G_d^s$	$G_m^f+G_d^sG_d^f$	$G_m^s+G_d^sG_d^f$	$G_m^sG_m^f+G_d^sG_d^f$
AUPRC	0.1810	0.2074	0.2152	0.1895	0.2131	0.2464	0.2128	0.2310	0.2589
AUC	0.9032	0.9172	0.9156	0.9021	0.9126	0.9228	0.9182	0.9178	0.9266
Accuracy	0.9927	0.9932	0.9932	0.9931	0.9933	0.9933	0.9931	0.9932	0.9934
Precision	0.2695	0.3012	0.3002	0.2889	0.3099	0.3220	0.3007	0.3118	0.3338
Recall	0.2504	0.2555	0.2550	0.2372	0.2606	0.2883	0.2720	0.2797	0.2989
F1	0.2589	0.2754	0.2742	0.2592	0.2826	0.3033	0.2844	0.2930	0.3146
precision@5%	0.0604	0.0665	0.0671	0.0609	0.0638	0.0698	0.0672	0.0683	0.0705
precision@10%	0.0374	0.0397	0.0398	0.0375	0.0389	0.0409	0.0400	0.0401	0.0413
recall@5%	0.6000	0.6608	0.6669	0.6046	0.6334	0.6931	0.6672	0.6786	0.7005
recall@10%	0.7439	0.7898	0.7915	0.7446	0.7729	0.8134	0.7949	0.7980	0.8200

we vary one parameter to test its effect while the others are fixed.

GCN layer

We use GCN to obtain the feature of miRNA and disease from different views. Figure 4(a) indicates that the number of GCN layers has little impact on the model performance, which may be attributed to the multichannel attention module. The deepening of the GCN layer will make the node lose the diversity of features, but the multichannel attention can make the model focus on more important features and reduce the impact of useless features. In this paper, the number of GCN layers is set to 2.

Embedding size

The embedding size of nodes in GCN is an important factor of MMGCN, which can directly affect the performance of the model. In this comparison, we changed the embedding size in {32,64,128,256}. As can be seen from Figure 4(b), within a certain range, the larger the embedding dimension, the higher the AUC value. Besides, the convergence rate of the model is similar under different embedding dimensions, which can prove that our model has good stability. In this paper, the embedding size is set to 256.

Number of filters

The number of filters in the CNN determines the final feature dimension of miRNA and disease node. After calculating the final feature, the model will carry out the subsequent matrix completion task. From Figure 4(c), it can be found that the model with different number of filters produces little fluctuation in results and does not affect the convergence speed of the model. In this paper, the number of filters is set to 128.

Learning rate

Learning rate is a parameter controlling the step size of gradient descent, and its value is related to whether the algorithm can obtain the optimal solution. A large learning rate will cause the model to be unstable and unable to converge. Conversely, a small learning rate may make it easily stuck in a poor local minimum and converge slowly. Figure 4(d) shows that there is an optimal value for the initial learning rate, thus in this paper, we set the learning rate to 0.001.

Case studies

In this section, we conduct two case studies to demonstrate the effectiveness of our MMGCN in predicting novel associations between diseases and miRNAs and the applicability to new diseases.

In the first case study, we focus on predicting potential miRNA–disease associations using our method. In particular, we use MMGCN with all the known miRNA–disease associations from HMDD V3.2 as the training dataset to make the prediction. Three diseases are selected, namely, breast neoplasms, lymphoma and esophageal neoplasms, for each disease, we prioritize the top 10 candidate miRNAs based on the predicted association scores. As shown in Table 5, we list the top 10 predicted miRNAs associated with the above three diseases, and almost all of them are confirmed by independent external databases dbDEMC [50].

Besides, to illustrate the applicability of MMGCN to new diseases, that is, diseases with no known associated miRNAs, we carry out the second case study on Lung Neoplasms. The known associations with the Lung Neoplasms are all removed from HMDDV3.2 so that it can be regarded as a new disease. After implementing MMGCN and ranking the candidate miRNAs of Lung Neoplasms according to their association scores, 10 out of the top 10 predicted Lung Neoplasms-related miRNAs are confirmed by HMDD V3.2(shown in Table 6).

The results present in the above two case studies show that MMGCN can achieve excellent predictive performance. It can not only discover new associations, but is also reliable in predicting miRNAs associated with new diseases.

Conclusion

In this study, we propose a novel multiview-based graph neural network approach MMGCN, for predicting potential associations in miRNA–disease networks. MMGCN can not only effectively make use of known associations to predict potential associations, but also adaptively extract useful information from multiple similarity views to enhance prediction performance. Experimental results on two datasets show that in most cases MMGCN outperforms state-of-the-art models under 5-fold cross-validation. Moreover, subsequent experiments demonstrate the effectiveness of different modules of our model. Case studies also verify the ability of MMGCN in finding new associations between diseases and miRNAs and its applicability to new diseases.

However, there is still some work that requires further research. First, the experiment in different view fusion shows that the structural information of the similarity network

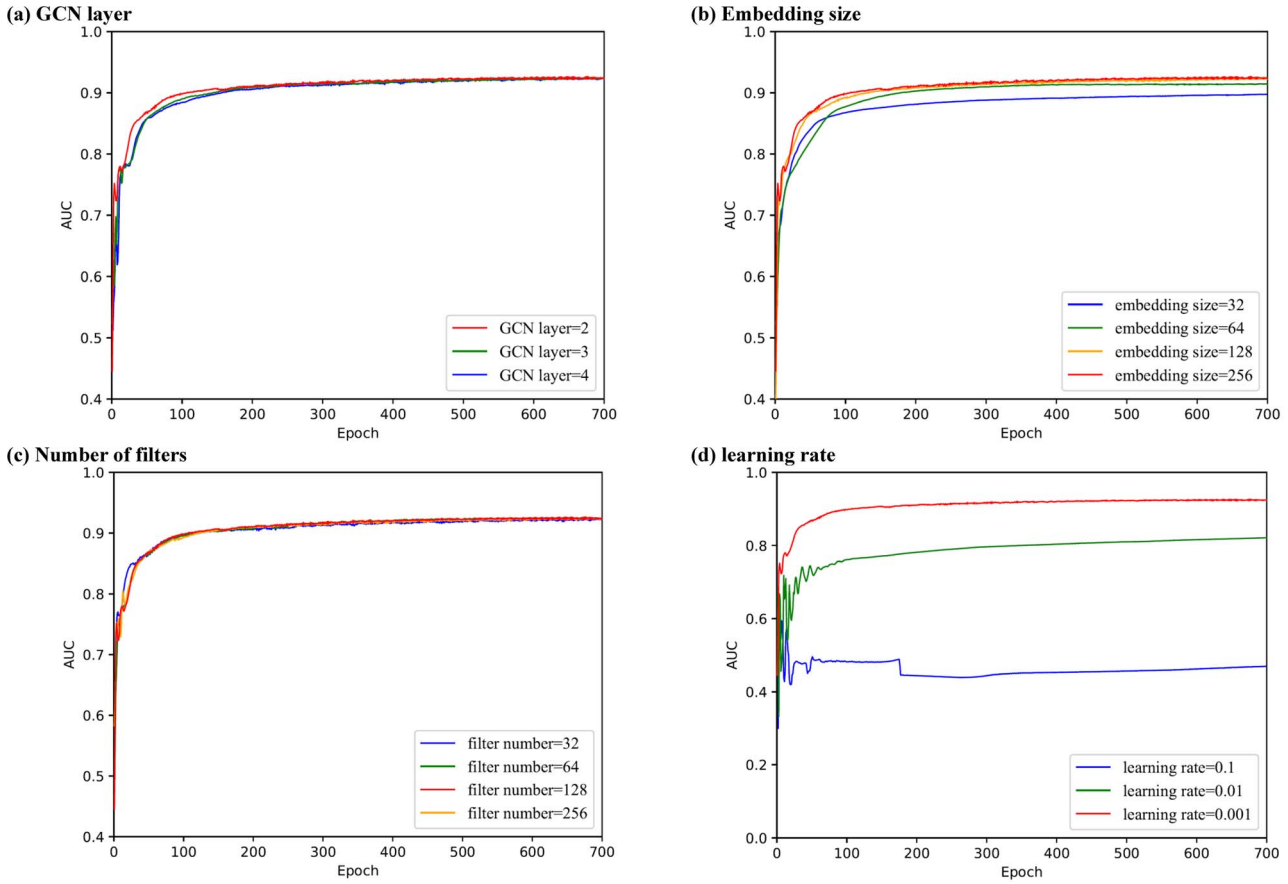

Figure 4. Parameter analysis of MMGCN.

Table 5. The top 10 associated miRNAs for breast neoplasms, lymphoma and esophageal neoplasms predicted by MMGCN based on known associations in HMDD V3.2 database. The prediction results are verified in dbDEMC

Cancer	No. Confirmed	Top 10 prediction					
		Rank	miRNA	Evidence	Rank	miRNA	Evidence
Breast neoplasms	10	1	hsa-mir-19b-2	dbDEMC	6	hsa-mir-99b	dbDEMC
		2	hsa-mir-186	dbDEMC	7	hsa-mir-211	dbDEMC
		3	hsa-mir-181a-1	dbDEMC	8	hsa-mir-454	dbDEMC
		4	hsa-mir-615	dbDEMC	9	hsa-mir-138-1	dbDEMC
		5	hsa-mir-6838	Unconfirmed	10	hsa-mir-28	dbDEMC
Lymphoma	10	1	hsa-mir-34a	dbDEMC	6	hsa-mir-146b	dbDEMC
		2	hsa-mir-29a	dbDEMC	7	hsa-mir-15b	dbDEMC
		3	hsa-mir-223	dbDEMC	8	hsa-mir-27a	dbDEMC
		4	hsa-mir-106b	dbDEMC	9	hsa-mir-125b-1	dbDEMC
		5	hsa-mir-145	dbDEMC	10	hsa-mir-148a	dbDEMC
Esophageal neoplasms	10	1	hsa-mir-17	dbDEMC	6	hsa-mir-23b	dbDEMC
		2	hsa-mir-29a	dbDEMC	7	hsa-mir-142	dbDEMC
		3	hsa-mir-222	dbDEMC	8	hsa-mir-15b	dbDEMC
		4	hsa-mir-200b	dbDEMC	9	hsa-mir-23a	dbDEMC
		5	hsa-mir-18a	dbDEMC	10	hsa-mir-125b-1	dbDEMC

between miRNAs and diseases significantly affects the learned feature representation, which further affects the final prediction result. How to collect a variety of valuable biological information to build more effective miRNA and disease similar networks is a topic worth studying in the future. Second, further research is also needed on integrating different biological information more reasonably and improving prediction performance.

Table 6. The top 10 associated miRNAs for lung neoplasms predicted by MMGCN based on known associations in HMDD V3.2 database. All known associations contain lung neoplasms in HMDD V3.2 database are removed before the prediction process. The prediction results are verified in HMDD V3.2

Cancer	No. Confirmed	Top 10 prediction					
		Rank	miRNA	Evidence	Rank	miRNA	Evidence
Lung neoplasms	10	1	hsa-mir-146a	HMDD	6	hsa-mir-146b	HMDD
		2	hsa-mir-155	HMDD	7	hsa-mir-221	HMDD
		3	hsa-mir-21	HMDD	8	hsa-mir-210	HMDD
		4	hsa-mir-150	HMDD	9	hsa-mir-34a	HMDD
		5	hsa-mir-126	HMDD	10	hsa-mir-223	HMDD

Key Points

- MiRNAs play significant roles in the development of human complex diseases.
- There are still challenges in accurately determining potential associations between miRNA and disease by using multisource data.
- In this paper, we designed a Multi-view Multichannel Attention Graph Convolutional Network (MMGCN) to predict potential miRNA–disease associations.
- The experiments verified the superior performance of MMGCN and showed the effectiveness of multiple views in the prediction of miRNA and disease association.

Data availability

Our work can be download in https: //github.com/Txinru/MMGCN

Supplementary data

Supplementary data are available online at https://academic.oup.com/bib.

Funding

National Natural Science Foundation of China (Grant No. 61873089, 62032007), Hunan Provincial Innovation Foundation for Postgraduate (Grant No. CX20200436).

References

1. Ambros V. The functions of animal microRNAs. *Nature* 2004;**431**(7006):350–5.
2. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 2004;**116**(2):281–97.
3. Taganov KD, Boldin MP, Chang K-J, *et al*. NF- B-dependent induction of microRNA miR-146, an inhibitor targeted to signaling proteins of innate immune responses. *Proc Natl Acad Sci* 2006;**103**(33):12481–6.
4. Croce CM, Calin GA. miRNAs, cancer, and stem cell division. *Cell* 2005;**122**(1):6–7.
5. Lu M, Zhang Q, Deng M, *et al*. An analysis of human microRNA and disease associations. *PLoS One* 2008;**3**(10):e3420.
6. Reinhart BJ, Slack FJ, Basson M, *et al*. The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature* 2000;**403**(6772):901–6.
7. Wilkening S, Bader A. Quantitative real-time polymerase chain reaction: methodical analysis and mathematical model. *J Biomol Tech* 2004;**15**:107.
8. Pall GS, Hamilton AJ. Improved northern blot method for enhanced detection of small RNA. *Nat Protoc* 2008;**3**(6):1077.
9. Chen M, Lu X, Liao B, *et al*. Uncover miRNA-disease association by exploiting global network similarity. *PLoS One* 2016;**11**(12):e0166509.
10. Zeng X, Zhang X, Zou Q. Integrative approaches for predicting microRNA function and prioritizing disease-related microRNA using biological interaction networks. *Brief Bioinform* 2016;**17**(2):193–203.
11. Jiang Q, Hao Y, Wang G, *et al*. Prioritization of disease microRNAs through a human phenome-microRNAome network. *BMC Syst Biol* 2010;**4**:1–9.
12. Xuan P, Han K, Guo M, *et al*. Prediction of microRNAs associated with human diseases based on weighted k most similar neighbors. *PLoS One* 2013;**8**(8):e70204.
13. Chen X, Liu MX, Yan. RWRMDA: predicting novel human microRNA-disease associations. *Mol Biosyst* 2012;**8**(10):2792–8.
14. Shi H, Xu J, Zhang G, *et al*. Walking the interactome to identify human miRNA-disease associations through the functional link between miRNA targets and disease genes. *BMC Syst Biol* 2013;**7**(1):101.
15. Chen X, Yan GY. Semi-supervised learning for potential human microRNA-disease associations inference. *Sci Rep* 2014;**4**:5501.
16. Xiao Q, Luo J, Liang C, *et al*. A graph regularized non-negative matrix factorization method for identifying microRNA-disease associations. *Bioinformatics* 2018;**34**(2):239–48.
17. Zhang ZC, Zhang XF, Wu M, *et al*. A graph regularized generalized matrix factorization model for predicting links in biomedical bipartite networks. *Bioinformatics* 2020;**36**(11):3474–81.
18. Chen X, Wang L, Qu J, *et al*. Predicting miRNA-disease association based on inductive matrix completion. *Bioinformatics* 2018;**34**(24):4256–65.
19. Chen X, Zhu CC, Yin J. Ensemble of decision tree reveals potential miRNA-disease associations. *PLoS Comput Biol* 2019;**15**(7):e1007209.
20. Xuan P, Sun H, Wang X, *et al*. Inferring the disease-associated miRNAs based on network representation learning and convolutional neural networks. *Int J Mol Sci* 2019;**20**:3648.
21. You Z-H, Li Y-R, Zhou J-R, *et al*. MISSIM: an incremental learning-based model with applications to the prediction of miRNA-disease association. *IEEE/ACM Trans Comput Biol Bioinform* 2020.

22. Kipf TN, Welling M. Semi-Supervised Classification with Graph Convolutional Networks. arXiv e-prints. 2016, arXiv:1609.02907.
23. Veličković P, Cucurull G, Casanova A *et al*. Graph attention networks. arXiv e-prints. 2017, arXiv:1710.10903.
24. Han P, Yang P, Zhao P *et al*. GCN–MF: Disease–gene Association Identification by Graph Convolutional Networks and Matrix Factorization. In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. Anchorage, AK, USA: Association for Computing Machinery, 2019, 705–13.
25. Shen C, Luo J, Ouyang W, *et al*. IDDkin: network-based influence deep diffusion model for enhancing prediction of kinase inhibitors. *Bioinformatics* 2020.
26. Wu Q-W, Xia J-F, Ni J-C, *et al*. GAERF: predicting lncRNA-disease associations by graph auto-encoder and random forest. In: *Briefings in bioinformatics*, 2021.
27. Wang L, You Z-H, Li Y-M, *et al*. GCNCDA: a new method for predicting circRNA-disease associations based on Graph Convolutional Network Algorithm. *PLoS Comput Biol* 2020;**16**(5):e1007568.
28. Liu P, Luo J, Chen X. miRCom: tensor completion integrating multi-view information to deduce the potential disease-related miRNA-miRNA pairs. *IEEE/ACM Trans Comput Biol Bioinform* 2020.
29. Wang L, You ZH, Chen X, *et al*. LMTRDA: using logistic model tree to predict MiRNA-disease associations by fusing multi-source information of sequences and similarities. *PLoS Comput Biol* 2019;**15**(3):e1006865.
30. Shen C, Luo J, Lai Z, *et al*. Multiview joint learning-based method for identifying small-molecule-associated MiRNAs by integrating pharmacological, genomics, and network knowledge. *J Chem Inf Model* 2020;**60**(8):4085–97.
31. Jiang L, Ding Y, Tang J, *et al*. MDA-SKF: similarity kernel fusion for accurately discovering miRNA-disease association. *Front Genet* 2018;**9**:618.
32. Wan F, Hong L, Xiao A, *et al*. NeoDTI: neural integration of neighbor information from a heterogeneous network for discovering new drug–target interactions. *Bioinformatics* 2019;**35**(1):104–11.
33. Zheng K, You Z-H, Wang L, *et al*. iMDA-BN: identification of miRNA-disease associations based on the biological network and graph embedding algorithm. *Comput Struct Biotechnol J* 2020;**18**:2391–400.
34. Zeng X, Liu L, Lü L, *et al*. Prediction of potential disease-associated microRNAs using structural perturbation method. *Bioinformatics* 2018;**34**(14):2425–32.
35. Needleman SB, Wunsch CD. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol* 1970;**48**(3):443–53.
36. Hwang S, Kim CY, Yang S, *et al*. HumanNet v2: human gene networks for disease research. *Nucleic Acids Res* 2019;**47**(D1):D573–80.
37. Wang D, Wang J, Lu M, *et al*. Inferring the human microRNA functional similarity and functional network based on microRNA-associated diseases. *Bioinformatics* 2010;**26**(13):1644–50.
38. Wang X, Wang R, Shi C *et al*. Multi-component graph convolutional collaborative filtering. 2019. arXiv preprint arXiv:1911.10699.
39. Hu J, Shen L, Sun G. Squeeze-and-excitation networks. In: *IEEE transactions on pattern analysis and machine intelligence*. 2020;**42**:2011–2023.
40. Huang Z, Shi J, Gao Y, *et al*. HMDD v3. 0: a database for experimentally supported human microRNA–disease associations. *Nucleic Acids Res* 2019;**47**(D1):D1013–7.
41. Kozomara A, Birgaoanu M, Griffiths-Jones S. miRBase: from microRNA sequences to function. *Nucleic Acids Res* 2019;**47**(D1):D155–62.
42. Huang H-Y, Lin Y-C-D, Li J *et al*. miRTarBase 2020: updates to the experimentally validated microRNA–target interaction database, *Nucleic Acids Res* 2020;**48**: D148-54.
43. Piñero J, Ramírez-Anguita JM, Saüch-Pitarch J, *et al*. The DisGeNET knowledge platform for disease genomics: 2019 update. *Nucleic Acids Res* 2020;**48**(D1):D845–55.
44. Jiang Q, Wang Y, Hao Y, *et al*. miR2Disease: a manually curated database for microRNA deregulation in human disease. *Nucleic Acids Res* 2009;**37**:D98–104.
45. Chen X, Yin J, Qu J, *et al*. MDHGI: matrix decomposition and heterogeneous graph inference for miRNA-disease association prediction. *PLoS Comput Biol* 2018;**14**(8): e1006418.
46. Li J, Zhang S, Liu T, *et al*. Neural inductive matrix completion with graph convolutional networks for miRNA-disease association prediction. *Bioinformatics* 2020;**36**(8):2538–46.
47. Chen X, Li SX, Yin J, *et al*. Potential miRNA-disease association prediction based on kernelized Bayesian matrix factorization. *Genomics* 2020;**112**(1):809–19.
48. Qu Y, Zhang H, Liang C, *et al*. SNMDA: a novel method for predicting microRNA-disease associations based on sparse neighbourhood. *J Cell Mol Med* 2018;**22**(10): 5109–20.
49. Qu Y, Zhang H, Lyu C, *et al*. LLCMDA: a novel method for predicting miRNA gene and disease relationship based on locality-constrained linear coding. *Front Genet* 2018;**9**: 576.
50. Yang Z, Wu L, Wang A, *et al*. dbDEMC 2.0: updated database of differentially expressed miRNAs in human cancers. *Nucleic Acids Symp Ser* 2017;**45**(D1):D812–8.