


Predicting miRNA-disease associations based on lncRNA–miRNA interactions and graph convolution networks

Wengang Wang and Hailin Chen 

Corresponding author. Hailin Chen, School of Software, East China Jiaotong University, Nanchang 330013, China. E-mail: chenhailin@ecjtu.edu.cn

Abstract

Increasing studies have proved that microRNAs (miRNAs) are critical biomarkers in the development of human complex diseases. Identifying disease-related miRNAs is beneficial to disease prevention, diagnosis and remedy. Based on the assumption that similar miRNAs tend to associate with similar diseases, various computational methods have been developed to predict novel miRNA-disease associations (MDAs). However, selecting proper features for similarity calculation is a challenging task because of data deficiencies in biomedical science. In this study, we propose a deep learning-based computational method named MAGCN to predict potential MDAs without using any similarity measurements. Our method predicts novel MDAs based on known lncRNA–miRNA interactions via graph convolution networks with multichannel attention mechanism and convolutional neural network combiner. Extensive experiments show that the average area under the receiver operating characteristic values obtained by our method under 2-fold, 5-fold and 10-fold cross-validations are 0.8994, 0.9032 and 0.9044, respectively. When compared with five state-of-the-art methods, MAGCN shows improvement in terms of prediction accuracy. In addition, we conduct case studies on three diseases to discover their related miRNAs, and find that all the top 50 predictions for all the three diseases have been supported by established databases. The comprehensive results demonstrate that our method is a reliable tool in detecting new disease-related miRNAs.

Keywords: miRNA-disease associations, lncRNA–miRNA interactions, graph convolution networks, multichannel attention mechanism, CNN combiner

Introduction

As one category of endogenous ~22 nt non-coding RNAs, microRNAs (miRNAs) play significant regulatory roles in animals and plants through base pairing with mRNA targets for cleavage or translational repression [1, 2]. More recently, increasing studies have revealed that miRNAs are involved in many important biological processes, such as cell proliferation and signal transduction [3]. The abnormal expression of miRNAs could, therefore, contribute to the progression of complex diseases [4, 5]. Identifying disease-related miRNAs would provide evidence to understand the molecular pathogenesis of diseases.

Biomedical technologies, such as complementary DNA (cDNA) cloning and polymerase chain reaction, have been widely applied to detect disease-related miRNAs [6–8]. Even though, success has been achieved; the biological experiments are costly and time-consuming. To tackle the challenges, computational approaches to predicting the most promising miRNA-disease associations (MDAs) for further biomedical screening are of great importance.

Till now, various computational methods have been developed to predict potential MDAs. These methods are mainly based on the assumption that similar miRNAs tend to be related with similar diseases and vice versa [5]. For example, Chen *et al.* [9] analyzed the effects of similarity measurements on MDA predictions, and proposed a network consistency-based method NetCBI to infer associations between miRNAs and diseases. Experimental results show that integrating similarities from both miRNA and disease

sides could improve prediction accuracy. Chen *et al.* [10] developed a semi-supervised method RLSMDA to predict relationships between diseases and miRNAs using regularized least squares based on miRNA functional similarity and disease semantic similarity. Xuan *et al.* [11] presented a computational method to predict miRNA candidates for diseases of interest by random walk on miRNA functional similarity network. Luo *et al.* [12] proposed a transduction learning-based method CPTL to systematically rank miRNAs related to diseases by combining similarities and known MDAs. Chen *et al.* [13] devised a recommendation-based method HAMDA to predict potential associations between miRNAs and diseases by integrating experimentally verified MDAs and similarity measures. Chen *et al.* [14] presented a novel model IMCMDA based on inductive matrix completion to predict possible MDAs by integrated similarity information. Zeng *et al.* [15] applied structural consistency to prioritize disease-related miRNAs in miRNA-disease bilayer network constructed by association information and similarity measurements. Chen *et al.* [16] developed a computational model MDHGI that applied matrix decomposition and heterogeneous graph inference for MDA predictions. Jiang *et al.* [17] utilize Laplacian regularized least squares (LapRLS) on integrated similarity kernel to discover potential MDAs. Zhang *et al.* [18] proposed a link inference method FLNSNLI to predict MDAs, in which label propagation was implemented to prioritize MDAs after linear neighborhood similarity measures. Xu *et al.* [19] integrated low-rank matrix completion with miRNA and disease

Wengang Wang is a graduate student at School of Software, East China Jiaotong University. His research interest includes deep learning and bioinformatics.

Hailin Chen, PhD, is an associate professor at School of Software, East China Jiaotong University. His research interest includes data mining and bioinformatics.

Received: August 8, 2022. Revised: October 17, 2022. Accepted: October 18, 2022

© The Author(s) 2022. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oup.com

similarity information for MDA inference. Chen *et al.* [20] developed a computational model NCMCMDA using neighborhood constraint matrix completion to recover missing MDAs based on existing MDAs and integrated similarity information.

Satisfactory performance has been achieved in the above methods. Similarity measurements are a key factor in determining the prediction accuracy for these methods. According to our previous study [21], quantifying miRNA-miRNA or disease-disease similarities would be affected because of the incompleteness of biomedical data, which would result in biased predictions or even restrict the application of these methods.

Meanwhile, inspired by the successful applications of machine learning (especially deep learning) techniques in many domains, such as speech recognition, visual object recognition and object detection, biomedical scientists are applying the machine learning algorithms to MDA predictions. For example, Chen *et al.* [22] presented a model of Extreme Gradient Boosting Machine for MiRNA-Disease Association (EGBMMDA) prediction, in which a regression tree under the gradient boosting framework was trained for association prediction. Zeng *et al.* [23] proposed a neural networks-based method NNMDA to identify disease-related miRNAs. Chen *et al.* [24] proposed a novel computational method EDTMDA, which integrated ensemble learning and dimensionality reduction to predict potential MDAs. Ji *et al.* [25] developed a network embedding learning method to learn embeddings of nodes in heterogeneous information network. Random Forest (RF) classifier was then used for predicting potential MDAs. Liu *et al.* [26] proposed a novel combined embedding model to predict miRNA-disease associations (CEMDA), in which gate recurrent unit, multi-head attention mechanism and multi-layer perceptron were used for embedding learning. Liu *et al.* [27] developed a computational framework SMALF, which utilized a stacked autoencoder and XGBoost to predict unknown MDAs. Tang *et al.* [28] developed a Graph Convolutional Network-based inference method MMGCN to predict MDAs, in which multi-view multichannel attention was used for representation learning. Liu *et al.* [29] devised a new computational method via deep forest ensemble learning based on autoencoder to predict MDAs. Yang *et al.* [30] developed a deep learning method PDMMA using graph neural networks (GNNs) and miRNA sequences to predict deep-level MDAs. Wang *et al.* [31] proposed a computational framework MKGAT to predict MDAs, in which graph attention networks (GATs) were used to learn miRNA and disease embeddings and dual LapRLS were for association predictions.

More and more accurate prediction results have been reported from these machine learning-based algorithms. However, there are three major challenges for these methods. First, some machine learning methods still use similarity values as input features for inference. Second, supervised learning methods need negative samples for classification, while experimentally validated miRNA-disease negative samples are not available in reality because of lack of biomedical research interest. Randomly selected negative samples would bring noise to prediction results. Third, setting proper values to parameters for some machine learning methods to obtain optimal results is tricky.

More recently, researchers have made efforts to infer MDAs based on other biological hypotheses. For example, Mørk *et al.* [32] presented a scoring scheme to rank MDAs by coupling miRNA-protein associations with protein-disease associations. Statistical analysis shows significant enrichment for proteins involved in pathways related to diseases. Based on the information of miRNA target genes, Chen *et al.* [33] developed a canonical correlation analysis (CCA)-based computational method to predict MDAs,

in which the extracted correlated sets of genes and diseases provided a biologically relevant interpretation of the formation of some MDAs. Considering the co-regulation relationships between lncRNAs and miRNAs, Huang *et al.* [34] proposed a multiview multitask method MVMTMDA to predict MDAs on a large scale. As both lncRNAs and miRNAs are key regulators, their interactions would provide further knowledge of mechanisms in disease development. These studies provide new perspectives for investigating MDAs. Meanwhile, refined computational methods are required to infer more reliable MDAs.

In this study, we propose a deep learning model MAGCN to predict MDAs. Unlike previous similarity-based methods, our method uses only known MDAs and lncRNA-miRNA interactions (LMIs) for predictions. Specifically, we construct two bipartite networks based on the known MDAs and LMIs. Graph convolution networks (GCN) with multichannel attention mechanism and convolutional neural network (CNN) combiner are applied on the bipartite networks for feature learning. A bilinear decoder is finally developed for association predictions. We test the performance of our method under the condition of cross-validations and compare it with other well-known methods. Results show our method outperforms the existing methods in terms of prediction accuracy. We further conduct case studies on three diseases and discover the top predictions have been well supported by existing databases. The excellent performance demonstrates the usefulness and reliability of our method in inferring novel MDAs.

Materials and methods

Datasets

The datasets used in our study are downloaded from reference [34], in which Huang *et al.* collected experimentally validated LMIs and MDAs from lncRNASNP v2.0 [35] and HMDD v3.0 [36], respectively. After deleting duplicated records and matching IDs in different databases, we finally obtain 541 lncRNAs, 268 miRNAs, 799 diseases, 10 465 LMIs and 11,253 MDAs. We use N_l , N_m and N_d to represent the numbers of lncRNAs, miRNAs and diseases, respectively. An adjacency matrix $A_{l-m} \in \mathbb{R}^{N_l \times N_m}$ is used to describe the LMIs and another adjacency matrix $A_{m-d} \in \mathbb{R}^{N_m \times N_d}$ to denote the MDAs. The value of each element in the two matrices is 1 or 0, indicating known or unknown relationships for LMIs and MDAs, respectively.

Method architecture

In this study, we propose a computational framework named MAGCN to infer potential MDAs based on known LMIs through GCN [37] with multichannel attention mechanism and a CNN combiner. As shown in Figure 1, two bipartite networks are firstly constructed based on known LMIs and MDAs. GCN are then used to learn the embeddings of lncRNAs, miRNAs and diseases. The embedding spaces of lncRNAs, miRNAs and diseases from multiple graph convolution layers are further fused through a CNN combiner with multichannel attention mechanism. Finally, a linear encoder is applied for association predictions based on the obtained features.

Bipartite network construction

The lncRNA-miRNA bipartite network G_{l-m} is defined by the adjacency matrix A_{l-m} and its transpose A_{l-m}^T as:

$$G_{l-m} = \begin{bmatrix} 0 & A_{l-m} \\ A_{l-m}^T & 0 \end{bmatrix} \in \mathbb{R}^{(N_l+N_m) \times (N_l+N_m)} \quad (1)$$

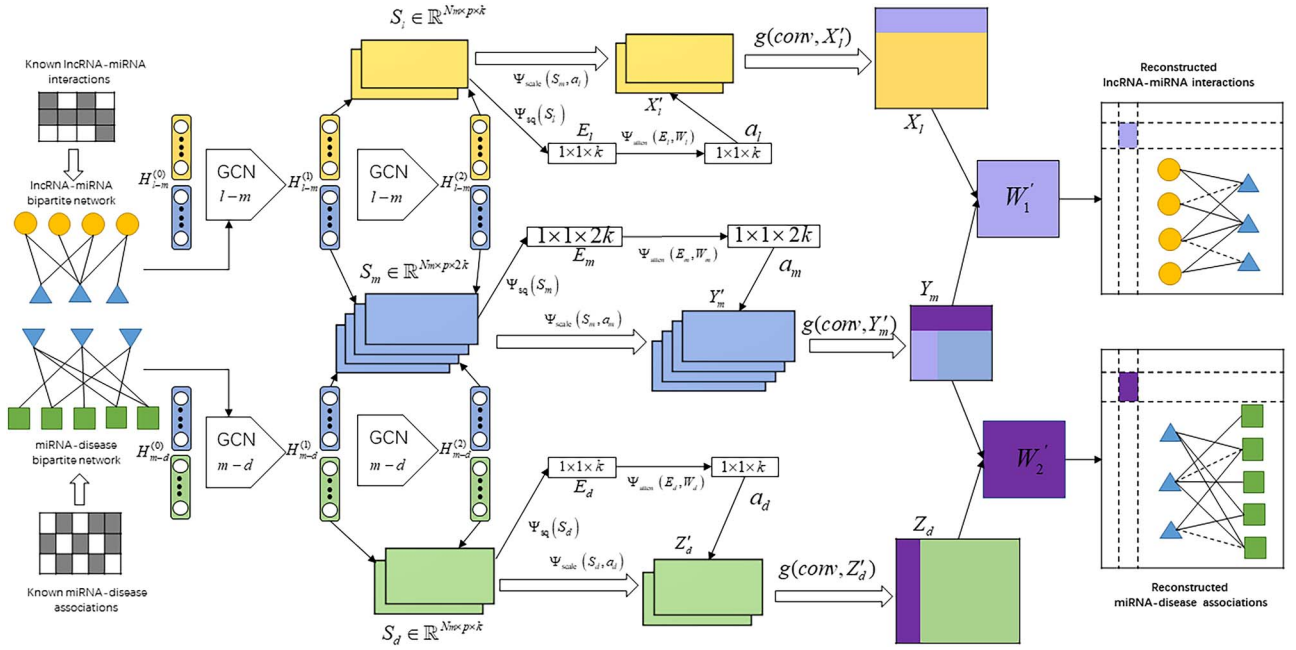


Figure 1. The workflow of our method MAGCN.

Similarly, the miRNA-disease bipartite network G_{m-d} is denoted by the adjacency matrix A_{m-d} and its transpose A_{m-d}^T as:

$$G_{m-d} = \begin{bmatrix} 0 & A_{m-d} \\ A_{m-d}^T & 0 \end{bmatrix} \in \mathbb{R}^{(Nm+Nd) \times (Nm+Nd)} \quad (2)$$

GCN encoder

We use GCN as encoders to obtain multilayer structural information of lncRNAs, miRNAs and diseases based on G_{l-m} and G_{m-d} defined above. We randomly initialize the features of lncRNAs as $X \in \mathbb{R}^{Nl \times f_0}$ ($X_i = \{x_{i1}, x_{i2}, \dots, x_{if_0}\}$), the features of miRNAs as $Y \in \mathbb{R}^{Nm \times f_0}$ and the features of diseases as $Z \in \mathbb{R}^{Nd \times f_0}$, respectively. We then receive the node embedding on G_{m-d} according to the following equation:

$$H_{m-d}^{(k)} = f(H_{m-d}^{(k-1)}, G_{m-d}) = \sigma(D^{-\frac{1}{2}} G_{m-d} D^{-\frac{1}{2}} H_{m-d}^{(k-1)} W_{m-d}^{(k-1)}) \quad (3)$$

where $H_{m-d}^{(k)}$ is the k -layer embedding of GCN, $k = 1, \dots, K$, $D = \text{diag}(\sum_{j=1}^{Nm+Nd} G_{m-d}(i, j))$ is the diagonal node degree matrix of G_{m-d} ,

$W_{m-d}^{(k-1)} \in \mathbb{R}^{(Nm+Nd) \times p}$ is a weight matrix for the k th neural network, p is the dimension of the GCN embedding and $\sigma(\cdot)$ is the non-linear activation function ReLU.

Similarly, we also use GCN as encoders to obtain the embedding $H_{l-m}^{(k)}$ of the nodes on G_{l-m} according to the following equation:

$$H_{l-m}^{(k)} = f(H_{l-m}^{(k-1)}, G_{l-m}) = \sigma(D^{-\frac{1}{2}} G_{l-m} D^{-\frac{1}{2}} H_{l-m}^{(k-1)} W_{l-m}^{(k-1)}) \quad (4)$$

In our study, we construct the initial embeddings $H_{l-m}^{(0)}$ and $H_{m-d}^{(0)}$ for the initial layers of the two bipartite networks G_{l-m} and

G_{m-d} as follows:

$$H_{l-m}^{(0)} = \begin{bmatrix} X \\ Y \end{bmatrix} \in \mathbb{R}^{(Nl+Nm) \times f_0} \quad (5)$$

$$H_{m-d}^{(0)} = \begin{bmatrix} Y \\ Z \end{bmatrix} \in \mathbb{R}^{(Nm+Nd) \times f_0} \quad (6)$$

We obtain the embeddings of lncRNAs, miRNAs and diseases from the two bipartite graphs by using the embeddings of each layer as different feature vectors, respectively. The embedding of each layer on G_{l-m} is denoted as $H_{l-m}^k \in [\begin{smallmatrix} H_{(l-m)_l}^k \\ H_{(l-m)_m}^k \end{smallmatrix}]$, where $H_{(l-m)_l}^k \in \mathbb{R}^{Nl \times p}$ represents the embedding of lncRNAs at the k th layer, and $H_{(l-m)_m}^k \in \mathbb{R}^{Nm \times p}$ denotes the embedding of miRNAs at layer k . The embedding of each layer on G_{m-d} is denoted as $H_{m-d}^k \in [\begin{smallmatrix} H_{(m-d)_m}^k \\ H_{(m-d)_d}^k \end{smallmatrix}]$, where $H_{(m-d)_m}^k \in \mathbb{R}^{Nm \times p}$ represents the embedding of miRNAs at layer k , and $H_{(m-d)_d}^k \in \mathbb{R}^{Nd \times p}$ denotes the embedding of diseases at layer k . Finally, we integrate the node features of the two bipartite graphs to obtain the combined feature spaces of lncRNAs, miRNAs and diseases as follows:

$$S_l = \{H_{(l-m)_l}^1, H_{(l-m)_l}^2, \dots, H_{(l-m)_l}^k\} \quad (7)$$

$$S_m = \{H_{(l-m)_m}^1, H_{(l-m)_m}^2, \dots, H_{(l-m)_m}^k, H_{(m-d)_m}^1, H_{(m-d)_m}^2, \dots, H_{(m-d)_m}^k\} \quad (8)$$

$$S_d = \{H_{(m-d)_d}^1, H_{(m-d)_d}^2, \dots, H_{(m-d)_d}^k\} \quad (9)$$

where S_l , S_m and S_d represent the feature spaces of lncRNAs, miRNAs and diseases, respectively.

Multichannel attention mechanism

The feature spaces of lncRNAs, miRNAs and diseases contain structural information from different layers on the two different bipartite graphs, and different structural information have

different contributions to embedding learning. We therefore use attention mechanism to extract the different features to improve the prediction performance of our model. Inspired by the SENet model [38] proposed by Hu *et al.* in computer vision, we here use the channel attention mechanism to calculate the contribution of different structural information in each space to the final embedding.

To obtain the importance of the different feature matrices, we first use a global average pooling to obtain the representation of each channel. For the miRNA feature space $S_m \in \mathbb{R}^{N_m \times p \times 2k}$ with $2k$ channels, the representation vector $E_m \in \mathbb{R}^{1 \times 1 \times 2k}$ is generated by the squeeze operation and $E_m = \{e_m^1, e_m^2, \dots, e_m^{2k}\}$. More specifically, given the c -th feature matrix S_m^c in the miRNA feature space S_m , the corresponding channel representation e_m^c can be calculated as follows:

$$e_m^c = \psi_{sq}(S_m^c) = \frac{1}{N_m \times p} \sum_{i=1}^{N_m} \sum_{j=1}^p S_m^c(i, j) \quad (10)$$

Correspondingly, the channel attention factor a_m for the miRNA feature space can be computed as follows:

$$a_m = \psi_{atten}(E_m, W_m) = \delta(W_2 \sigma(W_1 E_m)) \quad (11)$$

where $a_m \in \mathbb{R}^{1 \times 1 \times 2k}$ is the attention factor for miRNAs, $W_m = \{W_1, W_2\}$ is trainable parameters, $\delta(\cdot)$ is Sigmoid activation function, and $\sigma(\cdot)$ is ReLU activation function.

After obtaining the attention coefficients for each channel, the attention coefficients are multiplied with the original features to give different weights for each channel, with the following equation:

$$y_c' = \psi_{scale}(S_m^c, a_m^c) = a_m^c \cdot S_m^c \quad (12)$$

Based on the channel attention mechanism, we obtain lncRNA, miRNA and disease channel information as follows:

$$X_l' = [x_1', x_2', \dots, x_k'] \quad (13)$$

$$Y_m' = [y_1', y_2', \dots, y_{2k}'] \quad (14)$$

$$Z_d' = [z_1', z_2', \dots, z_k'] \quad (15)$$

CNN combiner

Inspired by the successful applications of CNNs in computer vision, we use multiple convolutional kernels of CNNs to extract and integrate the final node features from the channel information computed above. Given miRNA channel information Y_m' , the final miRNA features Y_m can be defined as follows:

$$Y_m = g(\text{conv}, Y_m') = \text{conv}(y_1', y_2', \dots, y_{2k}') \quad (16)$$

Similarly, we can obtain the final lncRNA embedding X_l and disease embedding Z_d as

$$X_l = g(\text{conv}, X_l') = \text{conv}(x_1', x_2', \dots, x_k') \quad (17)$$

$$Z_d = g(\text{conv}, Z_d') = \text{conv}(z_1', z_2', \dots, z_k') \quad (18)$$

where conv means convolution operation, the number of convolution kernels is p and kernel size is $p \times 1$.

Bilinear decoder

We use the decoder $A'_{m-d} = f(Y_m, Z_d)$ for reconstructing the miRNA-disease adjacency matrix from the final miRNA features and disease features, which is defined as follows:

$$A'_{m-d} = f(Y_m, Z_d) = \text{sigmoid}(Y_m W_1 Z_d^T) \quad (19)$$

where W_1 is a trainable parameter.

Similarly, the decoder $A'_{l-m} = f(X_l, Y_m)$ is for reconstructing the lncRNA-miRNA adjacency matrix from the final lncRNA features and miRNA features, which is calculated as follows:

$$A'_{l-m} = f(X_l, Y_m) = \text{sigmoid}(X_l W_2 Y_m^T) \quad (20)$$

Optimization

In training, we minimize the following logistic loss function to make the miRNA-disease prediction results accurately close to the true values as follows:

$$\ell_1 = \frac{1}{N_m \times N_d} \sum_{i,j} \left(\frac{A_{m-d}^{train}(i,j) \log A'_{m-d}(i,j) + (1 - A_{m-d}^{train}(i,j)) \log (1 - A'_{m-d}(i,j))}{1} \right) \quad (21)$$

where A_{m-d}^{train} is the adjacency matrix of MDAs in the training set, and A'_{m-d} is the predicted values of MDAs.

Meanwhile, as we use the known LMIs to better train the model, the second loss function is defined as follows:

$$\ell_2 = \|A_{l-m} - A'_{l-m}\|_F^2 \quad (22)$$

where $\|\cdot\|_F$ is the Frobenius norm, A_{l-m} is the original LMIs, and A'_{l-m} is the predicted lncRNA-miRNA values.

The final loss function can be defined as follows:

$$\ell_{total} = \ell_1 + \alpha * \ell_2 + \lambda \|\Theta\|^2 \quad (23)$$

where α and λ are hyperparameters, α is used to balance the weight of the ℓ_2 loss function, λ is for controlling the strength of L2 regularization and Θ represents the trainable parameters in the model.

In our model, we use the Adam optimizer [39] to minimize the loss function. The Adam optimizer iteratively updates the weights of the neural networks based on the training data. To prevent overfitting, we add L2 regularization to the loss function. In addition, the learning rate is adjusted according to the number of epochs trained during the optimization procedure, and we set the learning rate to decrease gradually as the epochs increase to achieve better training results.

Results

We first analyze the effects of the hyperparameters in our model MAGCN using 5-fold cross-validation (5-CV) based on the benchmark datasets in our study. Then, we perform ablation experiments on different components of our model to test the prediction performance. We further use three cross-validation strategies (2-CV, 5-CV and 10-CV) to comprehensively evaluate the performance of MAGCN, and compare it with other existing approaches under 5-CV experimental conditions. Finally, we use MAGCN to

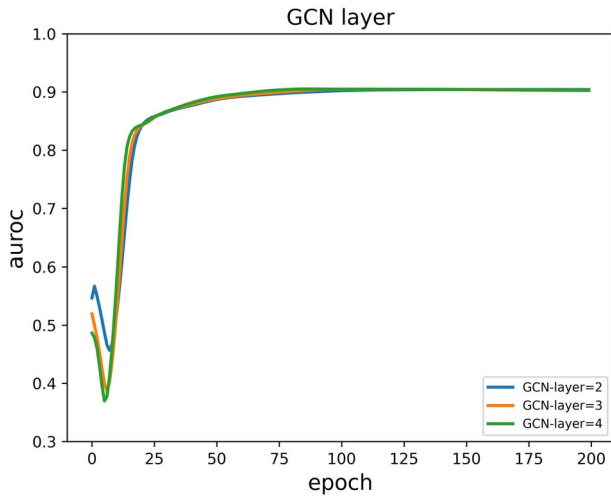


Figure 2. Sensitivity analysis on GCN layer k .

conduct case studies on three diseases to test its applications in reality.

Experimental setting and evaluation metrics

We used k -fold cross-validation ($k = 2, 5, 10$) to evaluate the performance of our method MAGCN by randomly dividing all MDAs into k approximately equal parts, with $k-1$ being used in turn for training and the remaining one for testing. To analyze the performance of our method, we use evaluation metrics including the area under the receiver operating characteristic (ROC) curve (AUROC) and the area under the precision/recall (PR) curve (AUPRC). We also calculate recall (also known as sensitivity), specificity, accuracy, precision and F1-measure (F1-score) for comprehensive comparison.

Parameter sensitivity analysis

There are four important hyperparameters (GCN layer k , initial feature embedding size f_0 , embedding size p and learning rate lr) in our method MAGCN. In this section, we empirically set values to the hyperparameters, and analyze their impacts on inference performance by conducting 5-fold cross-validation experiments on known MDAs. By changing the value of only one parameter, with the others remaining fixed, the following results are received.

GCN layer

Our model uses multiple GCN layers to extract structural information from different layers of lncRNAs, miRNAs and diseases to obtain their features. We set the number of GCN layers to be 2, 3 and 4 for analysis. The received AUC values are shown in Figure 2. We discover that the numbers of GCN layers have little effect on prediction performance. In the following experiments, the number of GCN layers k is therefore set to 2.

Initial feature embedding size

The node features of the model MAGCN are initialized randomly, and the size of the node features f_0 is a hyperparameter. We choose the embedding size for features in the range of {64, 128, 256, 512, 1024} in our experiments, and the results are shown in Figure 3. From Figure 3, we can find that the best result is received when the initial feature embedding size is 512. In this study, we set the embedding size f_0 to be 512.

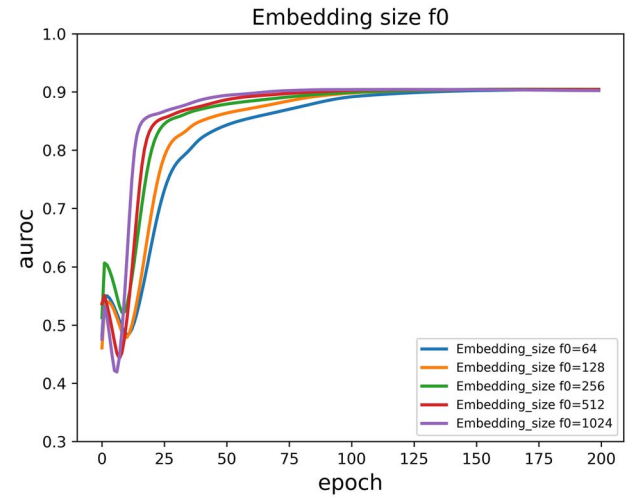


Figure 3. Sensitivity analysis on initial feature embedding size f_0 .

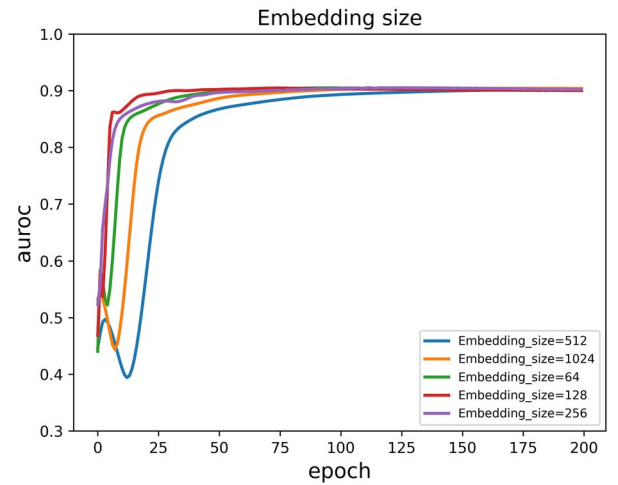


Figure 4. Sensitivity analysis on embedding size p .

Embedding size

In our model MAGCN, we use multiple layers of GCN to obtain the potential embeddings of lncRNAs, miRNAs and diseases, and finally use channel attention and CNN modules to calculate the final embeddings. We analyze the effect of the size of the potential embedding on the model. Specifically, we set the dimensions of the potential embedding to 64, 128, 256, 512 and 1024 for experimental comparison. From Figure 4, we can see that when the potential embedding size is 128, the model achieves the optimal AUC value. Therefore, we set the potential embedding size p to 128 in this study.

Learning rate

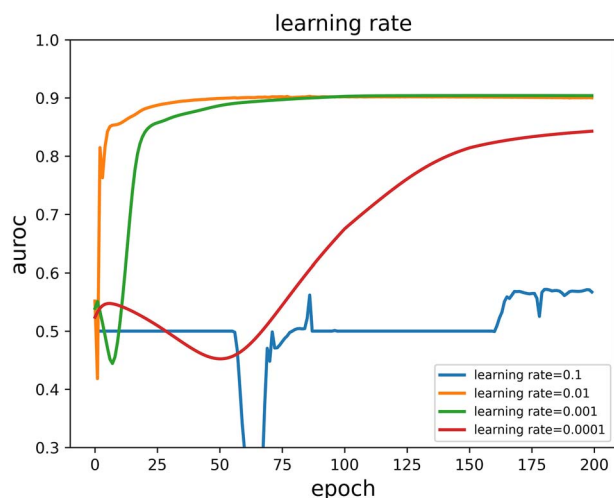
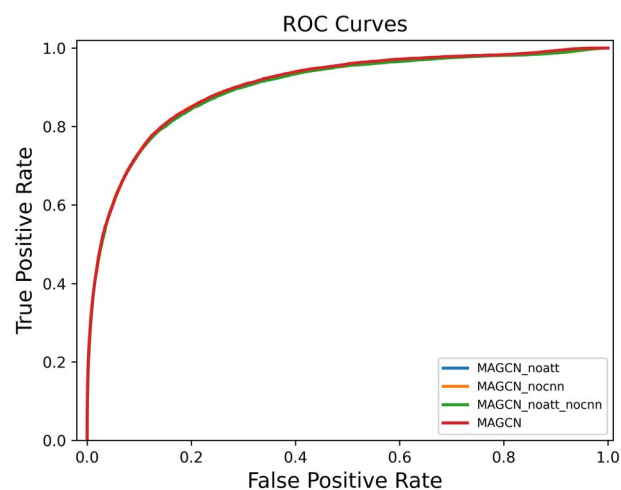
The learning rate is a hyperparameter that is used in the loss function to update the network weights. We vary the value of the learning rate in {0.1, 0.01, 0.001, 0.0001} in our experiments, and from Figure 5 we can see that the AUC value is optimal when the learning rate is 0.001. Therefore, we set the learning rate lr to 0.001.

Finally, the hyperparameters in our model MAGCN are set as follows: the number of training epoch is set to 200, the learning rate is set to 0.001, the loss function scale α is set to 0.0001, the L2 regularization weight loss λ is 0.00005, the number of GCN layers is 2, the initialized embedding size of node features is 512 and

Table 1. Performance of MAGCN and its variants based on 5-fold cross-validations

| Method | AUROC | AUPRC | F1-score | ACC | RECALL | SPEC | PRE |
|--------------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| MAGCN_noatte | 0.9012 | 0.5188 | 0.5054 | 0.9470 | 0.5150 | 0.9710 | 0.4975 |
| MAGCN_nocnn | 0.9009 | 0.5247 | 0.5064 | 0.9477 | 0.5104 | 0.9719 | 0.5036 |
| MAGCN_noatte_nocnn | 0.8988 | 0.5124 | 0.5015 | 0.9480 | 0.4976 | 0.9729 | 0.5065 |
| MAGCN | 0.9032 | 0.5252 | 0.5066 | 0.9471 | 0.5162 | 0.9710 | 0.4981 |

The bold value indicates the highest one in each column.

**Figure 5.** Sensitivity analysis on learning rate lr.**Figure 6.** ROC curves of ablation tests in MAGCN.

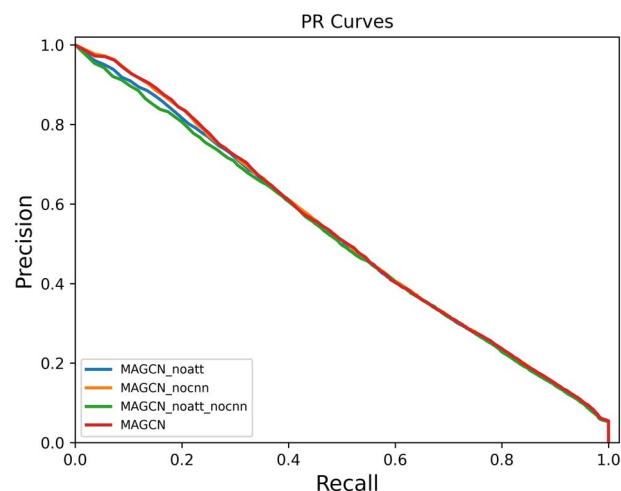
the potential embedding size for lncRNAs, miRNAs and diseases is 128.

Effects of different model components on prediction performance

In MAGCN, we use multichannel attention mechanism and CNN combiner for feature extraction. We conduct ablation experiments on the channel attention mechanism and CNN combiner. We use MAGCN_noatte to represent only using the CNN combiner to combine different channel information to get the final feature information. MAGCN_nocnn means simply adding up the different feature information obtained through the channel attention mechanism to obtain the final result, and not using the CNN combiner to learn the complex non-linear relationships. MAGCN_noatten_nocnn indicates that the final features are obtained by adding up the embeddings of the different layers, while neither the channel attention mechanism nor the CNN combiner is used. Table 1 shows the evaluation metrics on the datasets using MAGCN and its variant models under 5-fold cross-validation experiments. We also plot the corresponding AUC curves and PR curves in Figures 6 and 7, from which we can find that using the multichannel attention mechanism and the CNN combiner can extract more important features in the feature space and learn complex non-linear relationships, making the model's prediction performance improved.

Performance evaluation

In this section, we further evaluate the prediction performance of our model MAGCN based on cross-validations. Since our model can predict both the LMIs and MDAs. We first use LMIs as auxiliary information to predict the potential associations between miRNAs and diseases, with average AUROC values

**Figure 7.** PR curves of ablation tests in MAGCN.

of 0.8984, 0.9032 and 0.9044 under 2-fold, 5-fold and 10-fold cross-validations, respectively. We also use MAGCN to predict potential LMIs based on known MDAs, and the average AUROC values are 0.8973, 0.9605 and 0.9699 under 2-fold, 5-fold and 10-fold cross-validations, respectively. The experimental results demonstrate the reliability of our method in MDAs and LMIs predictions.

Comparison with other methods

In this section, we compare MAGCN with the latest methods that were proposed for MDA predictions. We here select five methods (i.e. MVMTMDA [34], MDA-SKF [17], Zeng *et al.*'s work [15], MDHGI [16] and IMCMDA [14]) for performance comparison. The methods are tested based on 5-fold cross-validations, and the comparative results are shown in Table 2. MAGCN obtains the highest AUROC

Table 2. Performance comparison with other methods based on 5-fold cross-validations

| Prediction task | Method | Average AUROC |
|--|-------------------------|---------------|
| Prediction of miRNA-disease associations | IMCMDA [14] | 0.6233 |
| | MDHGI [16] | 0.6932 |
| | Zeng et al.'s work [15] | 0.7883 |
| | MDA-SKF [17] | 0.8192 |
| | MVMTMDA [34] | 0.8512 |
| | MAGCN | 0.9032 |

Table 3. Comparison of average AUROC values of LMI predictions based on 2-fold, 5-fold and 10-fold cross-validations

| Method | 2-fold | 5-fold | 10-fold |
|---------|--------|--------|---------|
| MVMTMDA | 0.8747 | 0.9014 | 0.9037 |
| MAGCN | 0.8973 | 0.9605 | 0.9699 |

value of 0.9032 in 5-fold cross-validations and its AUROC value is higher than that of other methods by 5.20% (MVMTMDA), 8.4% (MDA-SKF), 11.49% (Zeng et al.'s work), 21% (MDHGI) and 27.99% (IMCMDA), respectively. The experiments indicate the excellent performance of our method.

Meanwhile, both MAGCN and MVMTMDA can be applied to LMI predictions. We therefore use both methods to test their performance in LMI predictions under 2-fold, 5-fold and 10-fold cross-validations. The received AUROC values are available at Table 3. It can be observed from Table 3 that our method MAGCN performs better than MVMTMDA, which further demonstrates the superiority of our method MAGCN.

Case studies

In this section, we conduct case studies to further validate the predictive performance of the model MAGCN in real situations. As tumors are serious illnesses that cause many deaths each year, predicting their related miRNAs is of great interest. We therefore choose to predict disease-related miRNAs with three common cancers (colon tumor, breast cancer and kidney cancer). Specifically, we remove the association information with a specific disease from the known MDA dataset, and train MAGCN with the other information to obtain the prediction results. Since biologists are more interested in the top predictions, we finally choose the top 50 associated miRNAs from the prediction results and validate them on the latest databases, such as HMDD v3.0 [36] and dbDEMC [40]. The validation results are listed in Tables 4, 5 and 6, respectively. We can find from the three tables that all the top 50 predictions in the three diseases have been supported by existing databases. The results from the three case studies suggest that MAGCN is an effective tool in detecting new MDAs.

Conclusion

In this study, we develop an end-to-end GCN-based computational approach MAGCN to predict novel MDAs. Different from previous research, our method MAGCN uses LMIs, instead of similarity measurements, to infer associations between miRNAs and diseases. We apply GCN with multichannel attention mechanism and a CNN combiner as encoders for feature learning. A bilinear decoder is used for association inference. Our method can predict not only MDAs but also LMIs. Extensive experiments including cross-validations and case studies demonstrate the effectiveness and superiority of our method.

It should be noted that the LMIs used in our study are limited and incomplete. The predicted results received by our

Table 4. The top 50 predicted miRNAs associated with colon neoplasms

| Ranking | miRNA | Evidence | Ranking | miRNA | Evidence |
|---------|-----------------|--------------|---------|-----------------|--------------|
| 1 | hsa-miR-21-5p | dbDEMC, HMDD | 26 | hsa-miR-31-5p | dbDEMC, HMDD |
| 2 | hsa-miR-146a-5p | dbDEMC, HMDD | 27 | hsa-miR-1 | dbDEMC, HMDD |
| 3 | hsa-miR-155-5p | dbDEMC, HMDD | 28 | hsa-miR-214-3p | dbDEMC |
| 4 | hsa-miR-223-3p | dbDEMC, HMDD | 29 | hsa-miR-9-5p | dbDEMC |
| 5 | hsa-miR-34a-5p | dbDEMC, HMDD | 30 | hsa-miR-96-5p | dbDEMC, HMDD |
| 6 | hsa-miR-126-3p | dbDEMC, HMDD | 31 | hsa-miR-17-5p | dbDEMC, HMDD |
| 7 | hsa-miR-145-5p | dbDEMC, HMDD | 32 | hsa-miR-125b-5p | dbDEMC, HMDD |
| 8 | hsa-miR-122-5p | dbDEMC | 33 | hsa-miR-92a-3p | dbDEMC, HMDD |
| 9 | hsa-miR-221-3p | dbDEMC, HMDD | 34 | hsa-miR-19a-3p | dbDEMC, HMDD |
| 10 | hsa-miR-132-3p | dbDEMC, HMDD | 35 | hsa-miR-27a-3p | dbDEMC, HMDD |
| 11 | hsa-miR-150-5p | dbDEMC, HMDD | 36 | hsa-miR-124-3p | dbDEMC |
| 12 | hsa-miR-143-3p | dbDEMC, HMDD | 37 | hsa-miR-200b-3p | dbDEMC, HMDD |
| 13 | hsa-miR-183-5p | dbDEMC | 38 | hsa-miR-34c-5p | dbDEMC |
| 14 | hsa-miR-206 | dbDEMC | 39 | hsa-miR-200c-3p | dbDEMC, HMDD |
| 15 | hsa-miR-142-3p | dbDEMC, HMDD | 40 | hsa-miR-30a-5p | dbDEMC, HMDD |
| 16 | hsa-miR-29a-3p | dbDEMC, HMDD | 41 | hsa-miR-15b-5p | dbDEMC, HMDD |
| 17 | hsa-miR-210-3p | dbDEMC, HMDD | 42 | hsa-miR-486-5p | dbDEMC, HMDD |
| 18 | hsa-miR-16-5p | dbDEMC | 43 | hsa-miR-106b-5p | dbDEMC, HMDD |
| 19 | hsa-miR-15a-5p | dbDEMC, HMDD | 44 | hsa-miR-205-5p | dbDEMC, HMDD |
| 20 | hsa-miR-182-5p | dbDEMC | 45 | hsa-miR-93-5p | dbDEMC, HMDD |
| 21 | hsa-miR-222-3p | dbDEMC, HMDD | 46 | hsa-miR-29b-3p | dbDEMC, HMDD |
| 22 | hsa-miR-24-3p | dbDEMC, HMDD | 47 | hsa-miR-192-5p | dbDEMC, HMDD |
| 23 | hsa-miR-133a-3p | dbDEMC, HMDD | 48 | hsa-miR-141-3p | dbDEMC, HMDD |
| 24 | hsa-miR-146b-5p | dbDEMC | 49 | hsa-miR-195-5p | dbDEMC, HMDD |
| 25 | hsa-miR-20a-5p | dbDEMC, HMDD | 50 | hsa-miR-181a-5p | dbDEMC, HMDD |

Table 5. The top 50 predicted miRNAs associated with breast neoplasms

| Ranking | miRNA | Evidence | Ranking | miRNA | Evidence |
|---------|-----------------|--------------|---------|-----------------|--------------|
| 1 | hsa-miR-146a-5p | dbDEMC, HMDD | 26 | hsa-miR-222-3p | dbDEMC, HMDD |
| 2 | hsa-miR-21-5p | dbDEMC, HMDD | 27 | hsa-miR-214-3p | dbDEMC, HMDD |
| 3 | hsa-miR-155-5p | dbDEMC, HMDD | 28 | hsa-miR-320a | dbDEMC, HMDD |
| 4 | hsa-miR-223-3p | dbDEMC, HMDD | 29 | hsa-miR-9-5p | dbDEMC, HMDD |
| 5 | hsa-miR-126-3p | dbDEMC, HMDD | 30 | hsa-miR-200c-3p | dbDEMC, HMDD |
| 6 | hsa-miR-210-3p | dbDEMC, HMDD | 31 | hsa-miR-20a-5p | dbDEMC, HMDD |
| 7 | hsa-miR-132-3p | dbDEMC, HMDD | 32 | hsa-miR-92a-3p | dbDEMC, HMDD |
| 8 | hsa-miR-34a-5p | dbDEMC, HMDD | 33 | hsa-miR-34c-5p | dbDEMC, HMDD |
| 9 | hsa-miR-122-5p | dbDEMC, HMDD | 34 | hsa-miR-143-3p | dbDEMC, HMDD |
| 10 | hsa-miR-145-5p | dbDEMC, HMDD | 35 | hsa-miR-29a-3p | dbDEMC, HMDD |
| 11 | hsa-miR-206 | dbDEMC, HMDD | 36 | hsa-miR-125a-5p | dbDEMC, HMDD |
| 12 | hsa-miR-221-3p | dbDEMC, HMDD | 37 | hsa-miR-182-5p | dbDEMC, HMDD |
| 13 | hsa-miR-183-5p | dbDEMC, HMDD | 38 | hsa-miR-124-3p | dbDEMC, HMDD |
| 14 | hsa-miR-142-3p | dbDEMC, HMDD | 39 | hsa-miR-30a-5p | dbDEMC, HMDD |
| 15 | hsa-miR-96-5p | dbDEMC, HMDD | 40 | hsa-miR-19a-3p | dbDEMC, HMDD |
| 16 | hsa-miR-17-5p | dbDEMC, HMDD | 41 | hsa-miR-205-5p | dbDEMC, HMDD |
| 17 | hsa-miR-133a-3p | dbDEMC, HMDD | 42 | hsa-miR-140-5p | dbDEMC, HMDD |
| 18 | hsa-miR-150-5p | dbDEMC, HMDD | 43 | hsa-miR-486-5p | dbDEMC, HMDD |
| 19 | hsa-miR-146b-5p | dbDEMC, HMDD | 44 | hsa-miR-212-3p | dbDEMC, HMDD |
| 20 | hsa-miR-16-5p | dbDEMC, HMDD | 45 | hsa-miR-15b-5p | dbDEMC, HMDD |
| 21 | hsa-miR-15a-5p | dbDEMC, HMDD | 46 | hsa-miR-192-5p | dbDEMC, HMDD |
| 22 | hsa-miR-125b-5p | dbDEMC, HMDD | 47 | hsa-miR-144-3p | dbDEMC, HMDD |
| 23 | hsa-miR-24-3p | dbDEMC, HMDD | 48 | hsa-miR-106b-5p | dbDEMC, HMDD |
| 24 | hsa-miR-1 | dbDEMC, HMDD | 49 | hsa-let-7b-5p | dbDEMC, HMDD |
| 25 | hsa-miR-31-5p | dbDEMC, HMDD | 50 | hsa-miR-200b-3p | dbDEMC, HMDD |

Table 6. The top 50 predicted miRNAs associated with kidney neoplasms

| Ranking | miRNA | Evidence | Ranking | miRNA | Evidence |
|---------|-----------------|--------------|---------|-----------------|--------------|
| 1 | hsa-miR-146a-5p | dbDEMC | 26 | hsa-miR-96-5p | dbDEMC |
| 2 | hsa-miR-21-5p | dbDEMC, HMDD | 27 | hsa-miR-15a-5p | dbDEMC, HMDD |
| 3 | hsa-miR-155-5p | dbDEMC, HMDD | 28 | hsa-miR-24-3p | dbDEMC |
| 4 | hsa-miR-223-3p | dbDEMC | 29 | hsa-miR-320a | dbDEMC |
| 5 | hsa-miR-126-3p | dbDEMC, HMDD | 30 | hsa-miR-125b-5p | dbDEMC |
| 6 | hsa-miR-210-3p | dbDEMC, HMDD | 31 | hsa-miR-19a-3p | dbDEMC |
| 7 | hsa-miR-122-5p | dbDEMC | 32 | hsa-miR-20a-5p | dbDEMC |
| 8 | hsa-miR-221-3p | dbDEMC | 33 | hsa-miR-92a-3p | dbDEMC |
| 9 | hsa-miR-34a-5p | dbDEMC, HMDD | 34 | hsa-miR-33a-5p | dbDEMC |
| 10 | hsa-miR-206 | dbDEMC | 35 | hsa-miR-486-5p | dbDEMC |
| 11 | hsa-miR-1 | dbDEMC | 36 | hsa-miR-16-5p | dbDEMC |
| 12 | hsa-miR-222-3p | dbDEMC | 37 | hsa-miR-192-5p | dbDEMC, HMDD |
| 13 | hsa-miR-145-5p | dbDEMC | 38 | hsa-miR-29a-3p | dbDEMC |
| 14 | hsa-miR-142-3p | dbDEMC | 39 | hsa-miR-34c-5p | dbDEMC |
| 15 | hsa-miR-183-5p | dbDEMC, HMDD | 40 | hsa-miR-124-3p | dbDEMC |
| 16 | hsa-miR-132-3p | dbDEMC, HMDD | 41 | hsa-miR-194-5p | dbDEMC |
| 17 | hsa-miR-143-3p | dbDEMC | 42 | hsa-miR-15b-5p | dbDEMC |
| 18 | hsa-miR-9-5p | dbDEMC | 43 | hsa-miR-144-3p | dbDEMC |
| 19 | hsa-miR-214-3p | dbDEMC | 44 | hsa-miR-205-5p | dbDEMC |
| 20 | hsa-miR-133a-3p | dbDEMC | 45 | hsa-let-7b-5p | dbDEMC |
| 21 | hsa-miR-182-5p | dbDEMC | 46 | hsa-miR-30a-5p | dbDEMC |
| 22 | hsa-miR-146b-5p | dbDEMC | 47 | hsa-miR-200c-3p | dbDEMC |
| 23 | hsa-miR-150-5p | dbDEMC | 48 | hsa-miR-204-5p | dbDEMC |
| 24 | hsa-miR-31-5p | dbDEMC | 49 | hsa-miR-22-3p | dbDEMC |
| 25 | hsa-miR-17-5p | dbDEMC, HMDD | 50 | hsa-miR-27a-3p | dbDEMC, HMDD |

method may therefore be biased. Integrating more experimentally validated LMIs would provide more reliable predictions. Meanwhile, the expression of lncRNAs and miRNAs are always tissue- or disease-specific. The lncRNA-miRNA associations may not be functionally activated under some conditions. Moreover,

setting proper values to the hyperparameters in our method to obtain optimal prediction results is a challenging task. Besides pathogenic lncRNA-miRNA co-regulations in disease development, miRNAs have also been discovered to lead to translational inhibition or degradation of their target mRNAs.

Incorporating more related biological information would further improve our understanding of the roles of miRNAs in the pathogenesis of human diseases, thus improving the accuracy of MDA predictions.

Key Points

- We propose a GCN-based method MAGCN to predict novel MDAs, in which LMIs instead of similarity measurements are used as initial input features.
- Using multichannel attention mechanism and CNN combiner, our method can learn complex relationships between graph nodes.
- Comprehensive experiments, such as cross-validations and case studies, demonstrate the effectiveness of our method in detecting new MDAs.
- Compared with existing well-known approaches, our method MAGCN shows improvement in prediction accuracy.

Data availability

The datasets and source codes used in this study are freely available at <https://github.com/shine-lucky/MAGCN>.

Authors' contribution

H.C. conceived and designed this study. W.W. implemented the experiments. W.W. and H.C. analyzed the results. W.W. and H.C. wrote the manuscript. Both authors read and approved the final manuscript.

Funding

National Natural Science Foundation of China (61862026).

References

1. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 2004;**116**:281–97.
2. Ambros V. The functions of animal microRNAs. *Nature* 2004;**431**:350–5.
3. Carthew RW, Sontheimer EJ. Origins and mechanisms of miRNAs and siRNAs. *Cell* 2009;**136**:642–55.
4. Croce CM, Calin GA. miRNAs, cancer, and stem cell division. *Cell* 2005;**122**:6–7.
5. Lu M, Zhang Q, Deng M, et al. An analysis of human microRNA and disease associations. *PLoS One* 2008;**3**:e3420.
6. Machová Poláková K, Lopotová T, Klamová H, et al. Expression patterns of microRNAs associated with CML phases and their disease related targets. *Mol Cancer* 2011;**10**:1–13.
7. Le H-B, Zhu W-Y, Chen D-D, et al. Evaluation of dynamic change of serum miR-21 and miR-24 in pre-and post-operative lung carcinoma patients. *Med Oncol* 2012;**29**:3190–7.
8. Pescador N, Pérez-Barba M, Ibarra JM, et al. Serum circulating microRNA profiling for identification of potential type 2 diabetes and obesity biomarkers. *PLoS One* 2013;**8**:e77251.
9. Chen H, Zhang Z. Similarity-based methods for potential human microRNA-disease association prediction. *BMC Med Genomics* 2013;**6**:1–9.
10. Chen X, Yan G-Y. Semi-supervised learning for potential human microRNA-disease associations inference. *Sci Rep* 2014;**4**:1–10.
11. Xuan P, Han K, Guo Y, et al. Prediction of potential disease-associated microRNAs based on random walk. *Bioinformatics* 2015;**31**:1805–15.
12. Luo J, Ding P, Liang C, et al. Collective prediction of disease-associated miRNAs based on transduction learning. *IEEE/ACM Trans Comput Biol Bioinform* 2017;**14**:1468–75.
13. Chen X, Niu YW, Wang GH, et al. HAMDA: hybrid approach for MiRNA-disease association prediction. *J Biomed Inform* 2017;**76**:50–8.
14. Chen X, Wang L, Qu J, et al. Predicting miRNA-disease association based on inductive matrix completion. *Bioinformatics* 2018;**34**:4256–65.
15. Zeng X, Liu L, Lu L, et al. Prediction of potential disease-associated microRNAs using structural perturbation method. *Bioinformatics* 2018;**34**:2425–32.
16. 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**:e1006418.
17. Jiang L, Ding Y, Tang J, et al. MDA-SKF: similarity kernel fusion for accurately discovering miRNA-disease association. *Front Genet* 2018;**9**:618.
18. Zhang W, Li Z, Guo W, et al. A fast linear neighborhood similarity-based network link inference method to predict microRNA-disease associations. *IEEE/ACM Trans Comput Biol Bioinform* 2019;**18**:405–15.
19. Xu J, Zhu W, Cai L, et al. LRMCMDA: predicting miRNA-disease association by integrating low-rank matrix completion with miRNA and disease similarity information. *IEEE Access* 2020;**8**:80728–38.
20. Chen X, Sun L-G, Zhao Y. NCMCMDA: miRNA-disease association prediction through neighborhood constraint matrix completion. *Brief Bioinform* 2021;**22**:485–96.
21. Chen H, Guo R, Li G, et al. Comparative analysis of similarity measurements in miRNAs with applications to miRNA-disease association predictions. *BMC Bioinformatics* 2020;**21**:176.
22. Chen X, Huang L, Xie D, et al. EGBMMDA: extreme gradient boosting machine for MiRNA-disease association prediction. *Cell Death Dis* 2018;**9**:3.
23. Zeng X, Wang W, Deng G, et al. Prediction of potential disease-associated microRNAs by using neural networks. *Mol Ther Nucleic Acids* 2019;**16**:566–75.
24. Chen X, Zhu CC, Yin J. Ensemble of decision tree reveals potential miRNA-disease associations. *PLoS Comput Biol* 2019;**15**:e1007209.
25. Ji BY, You ZH, Cheng L, et al. Predicting miRNA-disease association from heterogeneous information network with GraRep embedding model. *Sci Rep* 2020;**10**:6658.
26. Liu B, Zhu X, Zhang L, et al. Combined embedding model for MiRNA-disease association prediction. *BMC Bioinformatics* 2021;**22**:161.
27. Liu D, Huang Y, Nie W, et al. SMALF: miRNA-disease associations prediction based on stacked autoencoder and XGBoost. *BMC Bioinformatics* 2021;**22**:219.
28. Tang X, Luo J, Shen C, et al. Multi-view multichannel attention graph convolutional network for miRNA-disease association prediction. *Brief Bioinform* 2021;**22**:bbab174.
29. Liu W, Lin H, Huang L, et al. Identification of miRNA-disease associations via deep forest ensemble learning based on autoencoder. *Brief Bioinform* 2022;**23**:bbac104. <https://doi.org/10.1093/bib/bbac104>.

30. Yan C, Duan G, Li N, et al. PDMDA: predicting deep-level miRNA-disease associations with graph neural networks and sequence features. *Bioinformatics* 2022;**38**:2226–34.
31. Wang W, Chen H. Predicting miRNA-disease associations based on graph attention networks and dual Laplacian regularized least squares. *Brief Bioinform* 2022;**23**:bbac292.
32. Mork S, Pletscher-Frankild S, Pallega Caro A, et al. Protein-driven inference of miRNA-disease associations. *Bioinformatics* 2014;**30**: 392–7.
33. Chen H, Zhang Z, Feng D. Prediction and interpretation of miRNA-disease associations based on miRNA target genes using canonical correlation analysis. *BMC Bioinformatics* 2019;**20**: 404.
34. Huang Y-A, Chan KC, You Z-H, et al. Predicting microRNA-disease associations from lncRNA-microRNA interactions via multiview multitask learning. *Brief Bioinform* 2021;**22**:bbaa133.
35. Miao YR, Liu W, Zhang Q, et al. lncRNASNP2: an updated database of functional SNPs and mutations in human and mouse lncRNAs. *Nucleic Acids Res* 2018;**46**:D276–80.
36. 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**:D1013–7.
37. Kipf TN, Welling M. Semi-supervised classification with graph convolutional networks arXiv preprint arXiv:1609.02907. 2016.
38. Hu J, Shen L, Sun G. Squeeze-and-excitation networks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018, p. 7132–41.
39. Kingma DP, Ba J. Adam: a method for stochastic optimization arXiv preprint arXiv:1412.6980. 2014.
40. Yang Z, Wu L, Wang A, et al. dbDEMC 2.0: updated database of differentially expressed miRNAs in human cancers. *Nucleic Acids Res* 2017;**45**:D812–8.