



GCNCMI: A Graph Convolutional Neural Network Approach for Predicting circRNA-miRNA Interactions

Jie He¹, Pei Xiao¹, Chunyu Chen¹, Zeqin Zhu¹, Jiaxuan Zhang² and Lei Deng^{1*}

¹School of Computer Science and Engineering, Central South University, Changsha, China, ²Department of Electrical Engineering, University of California, San Diego, San Diego, CA, United States

OPEN ACCESS

Edited by:

Pingjian Ding,
Case Western Reserve University,
United States

Reviewed by:

Guanghui Li,
East China Jiaotong University, China
Weidun Xie,
City University of Hong Kong, Hong
Kong SAR, China

*Correspondence:

Lei Deng
leideng@csu.edu.cn

Specialty section:

This article was submitted to
RNA,
a section of the journal
Frontiers in Genetics

Received: 02 June 2022

Accepted: 23 June 2022

Published: 05 August 2022

Citation:

He J, Xiao P, Chen C, Zhu Z, Zhang J
and Deng L (2022) GCNCMI: A Graph
Convolutional Neural Network
Approach for Predicting circRNA-
miRNA Interactions.
Front. Genet. 13:959701.
doi: 10.3389/fgene.2022.959701

The interactions between circular RNAs (circRNAs) and microRNAs (miRNAs) have been shown to alter gene expression and regulate genes on diseases. Since traditional experimental methods are time-consuming and labor-intensive, most circRNA-miRNA interactions remain largely unknown. Developing computational approaches to large-scale explore the interactions between circRNAs and miRNAs can help bridge this gap. In this paper, we proposed a graph convolutional neural network-based approach named GCNCMI to predict the potential interactions between circRNAs and miRNAs. GCNCMI first mines the potential interactions of adjacent nodes in the graph convolutional neural network and then recursively propagates interaction information on the graph convolutional layers. Finally, it unites the embedded representations generated by each layer to make the final prediction. In the five-fold cross-validation, GCNCMI achieved the highest AUC of 0.9312 and the highest AUPR of 0.9412. In addition, the case studies of two miRNAs, hsa-miR-622 and hsa-miR-149-5p, showed that our model has a good effect on predicting circRNA-miRNA interactions. The code and data are available at <https://github.com/csuhjhjhj/GCNCMI>.

Keywords: circRNA, miRNA, deep learning, graph convolution neural network, circRNA-miRNA interaction

1 INTRODUCTION

Non-coding RNA (ncRNA) refers to various RNA molecules that will not translate into a protein. There has been much agreement through numerous studies that ncRNA has monumental biological functions though it only part a small fraction of the genomes. Since the discovery of RNA and ribosomal RNA in the 1950s, non-coding RNA that plays a biological role has been known for 60 years (Palazzo and Lee, 2015). As well as their roles at the transcriptional and post-transcriptional levels, ncRNA plays a critical role in epigenetic regulation of gene expression. The recent finding suggests that some of these RNAs are also involved in translation and splicing (Steitz and Moore, 2003; Butcher and Brow, 2005; Gesteland et al., 2006).

MicroRNA (miRNA) was discovered in 1993 by the Ambros and Ruvkun groups in *Caenorhabditis elegans* (Lee et al., 1993) and brought a revolution to molecular biology. They are small single-stranded molecules that derive from transcripts' unique hairpin structures called pre-miRNA. Most miRNAs are transcribed from DNA sequences into primary miRNAs, then processed into precursor miRNAs and become mature miRNAs finally (O'Brien et al., 2018; Liu et al., 2021). Furthermore, miRNAs have been found to regulate gene expression post-transcriptionally by

affecting mRNA translation, implying that dysregulation of miRNAs may be associated with various diseases by affecting gene expression (Bartel, 2004). For instance, recent studies showed approximately 50% of annotated human miRNAs are located in cancer-associated regions of the genome called fragile sites. This indicated that miRNA plays a crucial role in cancer progression (Calin et al., 2004).

Circular RNA consists of large non-coding RNAs produced by a non-canonical splicing event called back splicing. They are ubiquitous in species ranging from viruses to mammals during post-transcriptional processes. Viroids are the first circRNA to be discovered, though they are not produced by a back splicing mechanism (Sanger et al., 1976). A few years later, most circRNAs are observed in the cytoplasm and some small fractions in the nucleus. Circular forms of RNAs were observed or synthesized in diverse species such as viruses (Kos et al., 1986), prokaryotes (Ford and Ares, 1994), unicellular eukaryotes (Grabowski et al., 1981), and mammals (Capel et al., 1993). Most circRNA are expressed from known encoding proteins, composed of single or multiple exons. With the progress of high-throughput RNA-sequencing and bioinformatics tools, scientists have found the human transcriptome's general feature ubiquitous in many other metazoans.

A diverse set of circRNAs have been identified as having functions such as sponges, decoys, or translatable elements that alter gene or protein expression. Biological functions of circRNAs have only been investigated for a small fraction, while most of which are proposed as miRNA sponges (Hansen et al., 2013; Memczak et al., 2013; Deng et al., 2022). Sponging up miRNA and interacting with RNA-binding proteins (RBP), circRNA plays many pathological functions like regulating miRNA activity. He et al. (He et al., 2022) performed circRNA microarray analysis and found its expression profile in diabetes. By acting as microRNA sponges for miR-7 (ciRS-7) and miR-124-3p and miR-338-3p (circHIPK3), ciRS-7 and circHIPK3 promote insulin secretion. circRNAs were identified in cancers, so it also proposed to play a crucial role in the intimation and development of tumors. (Ashwal-Fluss et al., 2014; Dong et al., 2017; Soslaw, 2018). Most studies focus on the role of circRNA in tumors. circRNA was described as oncogenes. Diverse cellular functions of circRNA suggest their potential for cancer treatment as biomarkers and therapeutic targets (Chen and Huang, 2018; Li et al., 2019).

The interactions between circRNA and miRNA have been gradually discovered in recent years, and some related databases have been established. The CircR2Cancer database (Lan et al., 2020) contains 1,439 interactions between 1,135 circRNAs and 82 cancers. In addition, the database also includes basic information such as detection methods and expression patterns of circRNAs. However, there are few datasets on direct circRNA-miRNA interactions. Moreover, the known interactions are only a tiny part. Discovering the interactions between circRNAs and miRNAs is beneficial to understanding the interactions between circRNA and miRNA and disease. Using biological experiments to verify the interactions between circRNA-miRNA is time-consuming and labor-intensive. Computational methods can be used to mine the interactions between circRNA-miRNA more effectively. Still, there is little work to predict the circRNA-miRNA interactions.

As far as we know, GCNCMI is the first method to predict the circRNA-miRNA interactions, but other methods in the field of bioinformatics are still worth reference. Many methods based on computational interactions have recently achieved good results in predicting microbe-disease interactions and ncRNA-disease interactions. AE-RF (Deepthi and Jereesh, 2021) build an autoencoder to mine potential interaction features and then train a random forest model to predict circRNA-disease interactions. The DMFMDA (Liu et al., 2020) uses one-hot encoding of diseases and microorganisms to convert a vector representation in a low-dimensional space by embedding the propagation layer. The obtained vector representation is then input into a multi-layer neural network, and the parameters of the neural network are continuously optimized through Bayesian sorting to achieve accurate prediction. Deng et al. (Deng et al., 2020) constructed a meta-pathway-based circRNA-disease feature vector. This vector representation combines multiple similarities such as circRNA similarity, disease similarity, etc. The prediction is finally achieved using a random forest classifier. KATZHMDA (Chen et al., 2017) predicts the interactions between unknown microbes and disease by the Gaussian kernel similarity between known microbes and disease. NTSHMDA (Luo and Long, 2018) constructs a disease-microbe heterogeneity network based on the known similarity between microorganisms and diseases and assigns equal weights to known disease-microbe interactions according to the different contributions of diseases and microorganisms, which is conducive to reducing prediction error. Liu et al. (Dayun et al., 2021) established a multi-component graph attention network, which first passed a decomposer to identify node-level feature vectors, then combined the feature vectors to obtain a unified embedding vector, which was finally input into a fully connected network to predict microorganisms unknown interactions with the disease. SDLDA (Zeng et al., 2020) extract the linear and nonlinear interactions between lncRNA and diseases through singular value decomposition and neural network and finally unites the linear and nonlinear features into a new feature vector, which is input to the fully connected layer to realize prediction.

Although the above methods have achieved good prediction results, there are still some problems that will affect mining efficiency. Some existing association prediction methods rely on known similarities, but it is difficult to construct such similarities with the increasing number of miRNAs and circRNAs. There are far fewer known associations than unknown associations. Therefore, these methods are unsuitable when the circRNA and miRNA data increase. When the scale of data increases, how to mine the higher-order interactions of circRNA-miRNA is an urgent problem to be solved. In this paper, we construct a bipartite graph to describe the interaction information between circRNA and miRNA using known relationship pairs of them. Then we develop a graph convolutional network method to mine the deep semantic information that carries collaborative signal in the bipartite graph. We propagate the information flow recursively over the graph structure and continuously aggregate the interactive information between nodes to refine the embedding of each

TABLE 1 | The number of circRNAs, miRNAs, and circRNA-miRNA interactions included in the dataset.

circRNA	miRNA	interactions	unlabeled interactions
2,115	821	9,589	9,589

node. Finally, We concatenate the embeddings generated by each layer to predict the relationship of unknown circRNA-miRNA pairs. Experimental results show that our GCNCMI model outperforms the other six state-of-the-art methods.

2 MATERIALS AND METHODS

2.1 Datasets

We built the benchmark dataset from the circBank database (Liu et al., 2019). circBank contains 140,790 circRNAs. Each circRNA collects information such as miRNA binding sites, protein-coding ability, etc. We removed redundant parts of the dataset and extracted 2,115 circRNAs and 821 miRNAs from the circBank database, including 9,589 known circRNA-miRNA interactions. It now can be downloaded on the website <http://www.circbank.cn/downloads.html>. In addition, we randomly selected 9,589 unlabeled samples from the benchmark dataset. The detailed information can be seen in **Table 1**.

2.2 Problem Description

Our work aims to predict unknown relationships based on known circRNA and miRNA relationships. We use $U = \{u_1, u_2 \dots u_n\}$ and $V = \{v_1, v_2, \dots, v_m\}$ to respectively represent the collection of n circRNAs and m miRNAs, and use the interaction matrix $R \in \mathbb{R}^{n \times m}$ to represent the relationship between them. If the circRNA u_i is related to miRNA v_j , then the $R_{ij} = 1$, otherwise $R_{ij} = 0$. It should be noted that $R_{ij} = 0$ can only indicate that the two RNAs have not yet found a relationship, but may actually be related.

2.3 Graph Construction

We use a bipartite graph $G(U \cup V, E)$ constructed by the interaction matrix $R \in \mathbb{R}^{n \times m}$ to show the relationship between circRNAs and miRNAs, where U, V are the vertex sets denoting the circRNAs and miRNAs, and E is the edge sets constructed from the association matrix $R \in \mathbb{R}^{n \times m}$. This bipartite graph can be expanded into a complex interaction graph as shown in **Figure 1**. This interaction graph contains the higher-order interaction information of circRNA and miRNA, from which we can mine deep semantic information that carry collaborative signal. For example, the path $u_1 - v_1 - u_2$ and $u_1 - v_2 - u_2$ indicate the behavior similarity between u_1 and u_2 , as both circRNAs have interacted with v_1 and v_2 . Then, the interaction between u_2 and v_3 suggests that u_1 and v_3 are likely to be related.

2.4 GCNCMI

To capture the deep interaction information embedded in the interaction graph, we model the high-order interaction information of circRNA-miRNA in the embedding function.

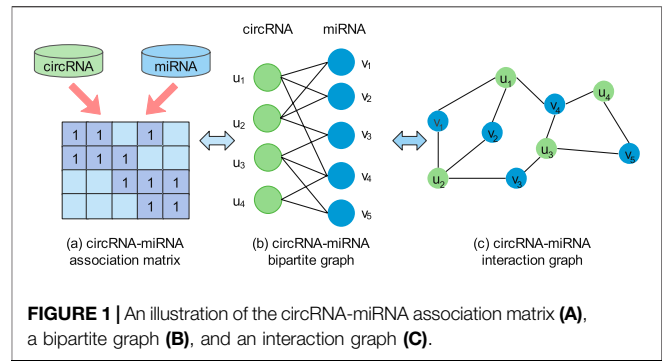


FIGURE 1 | An illustration of the circRNA-miRNA association matrix (A), a bipartite graph (B), and an interaction graph (C).

We propagate the information flow recursively over the graph structure and continuously aggregate the information of neighboring nodes to refine the embedding representation of the nodes (Hamilton et al., 2017; Xu et al., 2018; Wang et al., 2019). The architecture of our proposed GCNCMI model is shown in **Figure 2**. There are three parts to the framework: 1) An embedding layer that offers initialized circRNA embeddings and miRNA embeddings from the input data; 2) multiple embedding propagation layers that refine the embeddings by aggregating higher-order interaction information; 3) the prediction layer that concatenates the embeddings from different propagation layers and outputs the prediction score of a circRNA-miRNA pair.

2.4.1 Embedding Layer

We use the embedding vector $e_{u_i}^k \in \mathbb{R}^s$ ($e_{v_i}^k \in \mathbb{R}^s$) to describe the circRNA u (miRNA v) in k -th layer, where s is the embedding size. The initial state of circRNA embeddings and miRNA embeddings in embedding layer can be abstracted as:

$$E_u^0 = [e_{u_1}^0, e_{u_2}^0, \dots, e_{u_n}^0] \tag{1}$$

$$E_v^0 = [e_{v_1}^0, e_{v_2}^0, \dots, e_{v_m}^0] \tag{2}$$

Where E_u^0 is the initial embedding of circRNAs, and E_v^0 is the initial embedding of miRNAs. The initial embedding will be continuously optimized and improved end-to-end, which will be mentioned in the next section.

2.4.2 Embedding Propagation Layers

Next, we continuously aggregate the information of the node itself and its adjacent nodes to refine the embeddings of miRNAs and circRNAs. This is based on the GNN message-passing architecture (Hamilton et al., 2017; Xu et al., 2018). During an embedding update, the message aggregated by each node consists of two parts: the messages from the neighbor nodes of the previous layer and the messages inherited from the node itself.

As shown in **Figure 3**, in the k -th propagation layer, the embedding of circRNA u can be recursively formulated as:

$$e_u^k = \sigma \left(m(u, k) + \sum_{v \in N_u} m(u, v, k) \right) \tag{3}$$

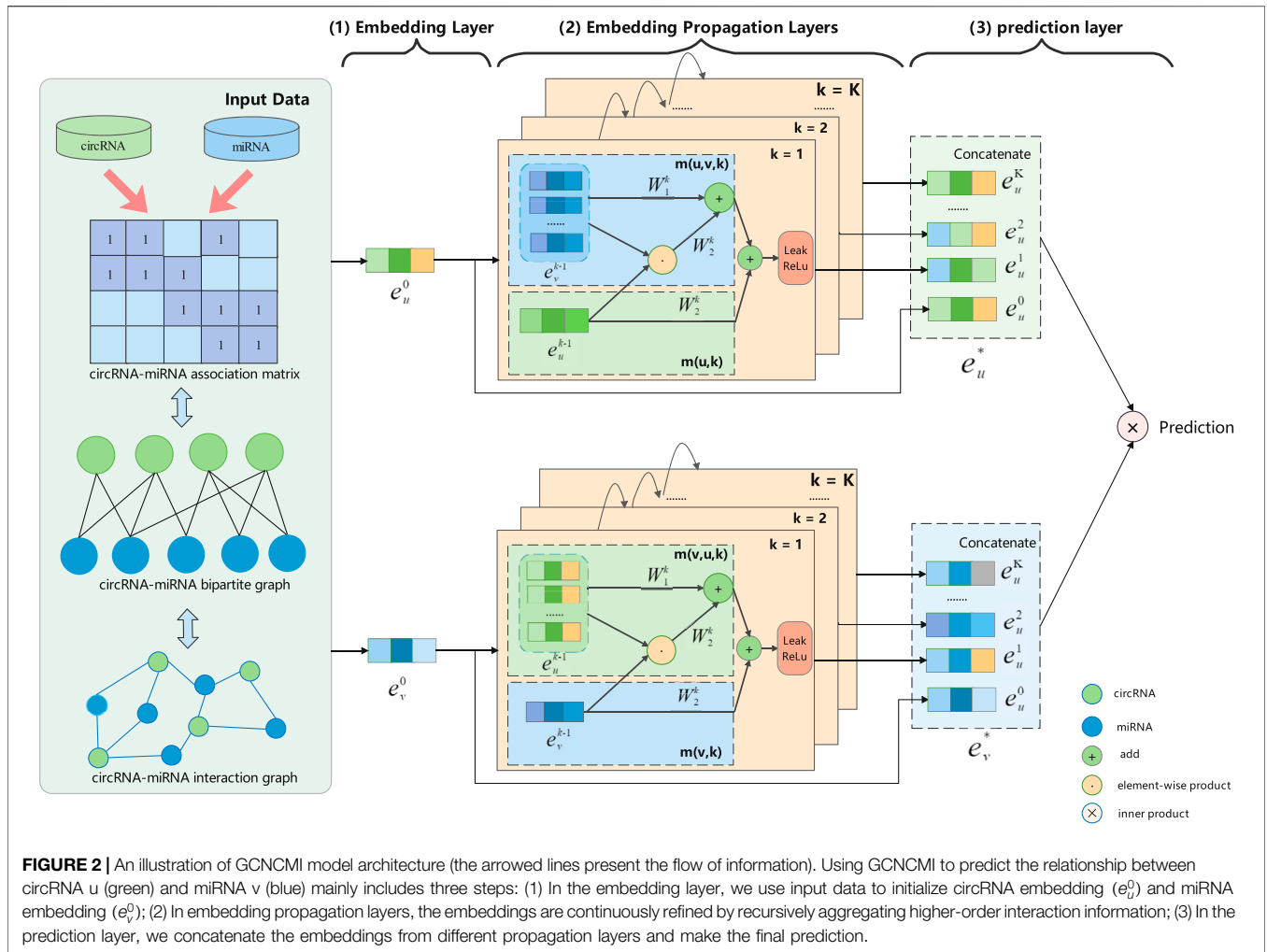


FIGURE 2 | An illustration of GCNCMI model architecture (the arrowed lines present the flow of information). Using GCNCMI to predict the relationship between circRNA u (green) and miRNA v (blue) mainly includes three steps: (1) In the embedding layer, we use input data to initialize circRNA embedding (e_u^0) and miRNA embedding (e_v^0); (2) In embedding propagation layers, the embeddings are continuously refined by recursively aggregating higher-order interaction information; (3) In the prediction layer, we concatenate the embeddings from different propagation layers and make the final prediction.

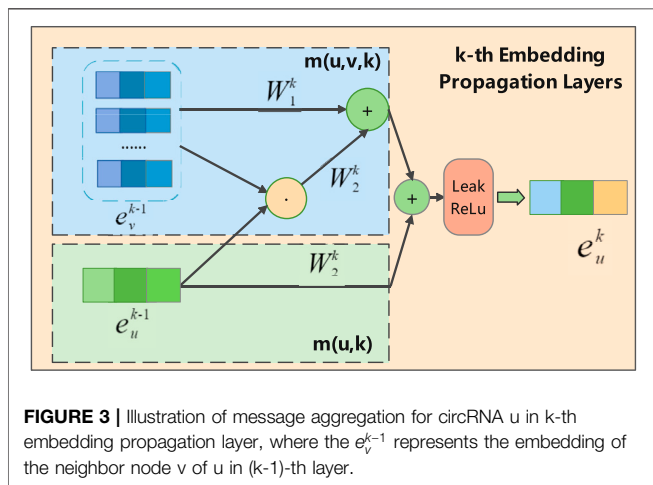


FIGURE 3 | Illustration of message aggregation for circRNA u in k -th embedding propagation layer, where the e_v^{k-1} represents the embedding of the neighbor node v of u in $(k-1)$ -th layer.

Where e_u^k represents the embedding of circRNA obtained in the k -th embedding propagation layer, $\sigma(\cdot)$ is the activation function LeakyReLU (Nikolakopoulos and Karypis, 2019), v denotes the neighbor nodes of u , and $m(u, k)$ represent the

messages delivered from the previous layer itself, while $m(u, v, k)$ representing the messages delivered by all neighbor nodes from the previous layer. The $m(u, k)$ and $m(u, v, k)$ can be formulated as follows:

$$m(u, k) = W_1^k e_u^{k-1} \tag{4}$$

$$m(u, v, k) = \frac{W_1^k e_v^{k-1} + W_2^k (e_v^{k-1} \odot e_u^{k-1})}{\sqrt{N(u) \|N(v)\|}} \tag{5}$$

Where $W_1^k, W_2^k \in \mathbb{R}^{d_k \times d_{k-1}}$ are the trainable transformation matrices used to extract propagation information, and d_k is the transformation size; e_u^{k-1} is the circRNA embedding representation generated from the $(k-1)$ -th propagation layer, which will further contribute its information to the circRNA embedding u at layer k . We use the graph Laplacian norm $1/\sqrt{N(u) \|N(v)\|}$ to control how much the propagating message decays as the path length increases, where $N(u)$ represent the first-hop neighbors of circRNA u (miRNA v). In Eq. 4, we consider the self-connection of nodes, which can effectively retain the original feature information to avoid information variation when the number of layers increases. For the neighbor nodes of node u , we aggregate

not only the information of node v but also aggregate the interaction information between the u and v . It is encoded via $e_v^{k-1} \odot e_u^{k-1}$, where \odot is element-wise product operation. In this way, more information from similar nodes can be passed, which enhances the representation ability of the model and helps to improve the accuracy of prediction results. Eqs 3–5 represent the calculation process of the embedding circRNA u at the k -th layer. Analogously, the embedded representation of miRNA can be obtained.

2.4.3 Model Prediction

After multi-layer propagation, we can obtain multiple embedding representations of miRNAs and circRNAs. The embeddings obtained by different propagation layers contain different orders of interaction information, so they have different contributions to reflecting the relationship between circRNAs and miRNAs. Therefore, we concatenate all embeddings to express the final embedding. The following formula shows the final embedding representation of circRNA u and miRNA v through K embedding propagation layers:

$$e_u^* = e_u^0 \| e_u^1 \| \dots \| e_u^K, \quad e_v^* = e_v^0 \| e_v^1 \| \dots \| e_v^K \quad (6)$$

Where $\|$ denotes concatenation operation, this simple concatenation operation can makes our final embeddings contain richer semantic information without increasing the learning parameters. Finally, we perform an inner product operation on the final embedding to obtain the interaction prediction between circRNA u and miRNA v :

$$\hat{y}_{GCNCMI}(u, v) = e_u^* \otimes e_v^* \quad (7)$$

Algorithm 1 shows the pseudocode description for predicting the interaction between circRNA u and miRNA v using GCNCMI.

2.4.4 Model Optimization

Pointwise loss and pairwise loss are two common methods used to update model parameters (He et al., 2016). The pointwise learning emphasizes the loss between the predicted value \hat{y}_{uv} and target value y_{uv} . Still, we prefer to address predicting the interactions between circRNA and miRNA from the perspective of ranking. Therefore, we choose pairwise loss optimization to update model parameters. Bayesian Personalized Ranking (BPR) is a matrix factorization-based pairwise loss function that is often used to optimize recommendation tasks similar to our prediction task (Rendle, 2010). Specifically, it can be formulated as follows:

$$\min_{\Theta} L = \sum_{(u,i,j) \in D} -\ln s(\hat{y}_{ui} - \hat{y}_{uj}) + \lambda \|\Theta\|_2^2 \quad (8)$$

where $s(\dots)$ is the sigmoid function; $D = \{(u, i, j) | (u, i) \in R^+, (u, j) \in R^-\}$ is the pairwise training sample containing positive samples R^+ (i.e., circRNA u has interacted with miRNA v_i) and negative samples R^- (i.e., the interactions between circRNA u and miRNA v_j is unknown). \hat{y}_{ui} denotes the prediction score of

u and v_i . \hat{y}_{uj} denotes the prediction score of u and v_j . $\Theta = \{E, \{W_1^k, W_2^k\}_{k=1}^K\}$ represents all model parameters that will be trained. λ is a parameter used to control the strength of L_2 regularization. We use Adam as the optimizer to update the model parameters. Additionally, we use message dropout and node dropout to avoid model overfitting during training. Message dropout means that we will drop the message in Eq. 3 with a certain probability during the propagation, while node dropout randomly drops a specific node and discards all its outgoing messages. Dropout operations can reduce the influence of specific RNAs, making the model more robust.

Algorithm 1. GCNCMI algorithm to predict the interaction between circRNA u and miRNA v

Input: interaction matrix $R \in \mathbb{R}^{n \times m}$;
 bipartite graph $G(U \cup V, E)$; depth K ;
 dimension D ; weight matrices $W_1^k, W_2^k, \forall k \in \{1, \dots, K\}$;
 neighborhood function N ; activation function LeakyReLU σ

Output: interaction prediction between circRNA u and miRNA v

- 1 Initialize Embeddings:
 $E^0 = [\underbrace{e_{u_1}^0, \dots, e_{u_n}^0}_{\text{circRNA embeddings}}, \underbrace{e_{v_1}^0, \dots, e_{v_m}^0}_{\text{miRNA embeddings}}]$;
- 2 **for** layer $k = 1$ to K **do**
- 3 **for** all node embeddings $e_i^k \in \mathbb{E}^k$ **do**
- 4 $m(i, k) = W_1^k e_i^{k-1}$;
- 5 $m(i, j, k) = \frac{W_1^k e_j^{k-1} + W_2^k (e_j^{k-1} \odot e_i^{k-1})}{\sqrt{N(i) \| N(j) |}}$;
- 6 $e_i^k \leftarrow \sigma \left(m(i, k) + \sum_{j \in N_i} m(i, j, k) \right)$;
- 7 **end**
- 8 **end**
- 9 $e_u^* \leftarrow e_u^0 \| e_u^1 \| \dots \| e_u^K$;
- 10 $e_v^* \leftarrow e_v^0 \| e_v^1 \| \dots \| e_v^K$;
- 11 $\hat{y}_{GCNCMI}(u, v) \leftarrow e_u^* \otimes e_v^*$;
- 12 **return** $\hat{y}_{GCNCMI}(u, v)$;

3 EXPERIMENT

3.1 Experimental Settings

To evaluate the performance of our model in predicting circRNA-miRNA interactions, we combined the known 9,589 interactions used as positive samples, and 9,589 unlabeled interactions were randomly selected from the benchmark dataset as negative samples. We performed five-fold cross-validation on the constructed dataset. The

TABLE 2 | The five-fold cross-validation results of GCNCMI.

No.	AUPR	AUC	ACC	Pre	Recall	F1
1	0.9293	0.9288	0.8508	0.9390	0.8289	0.8805
2	0.9428	0.9352	0.8531	0.9424	0.8440	0.8905
3	0.9453	0.9372	0.8578	0.9450	0.8357	0.8870
4	0.9396	0.9282	0.8532	0.9392	0.8341	0.8835
5	0.9412	0.9312	0.8503	0.9408	0.8298	0.8818
Average	0.9396	0.9320	0.8530	0.9413	0.8345	0.8847

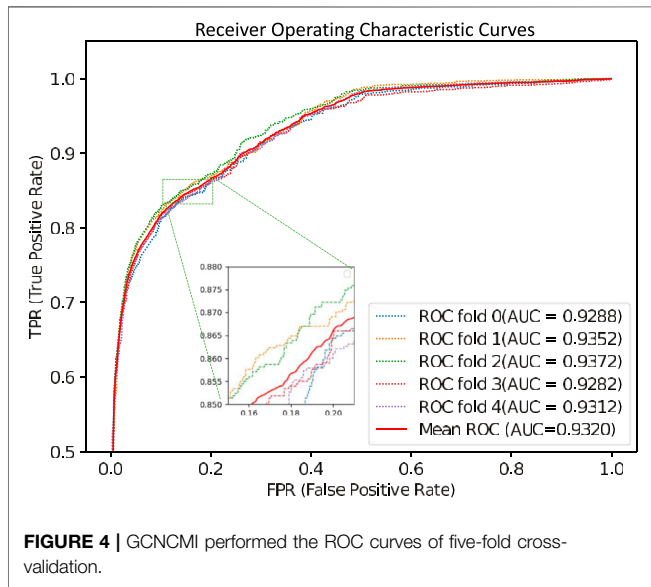


TABLE 3 | The performance of GCNCMI on different layers.

K	AUPR	AUC	Acc	Pre	Recall	F1
1	0.9283	0.9198	0.8368	0.9280	0.8319	0.8773
2	0.9412	0.9312	0.8503	0.9408	0.8298	0.8818
3	0.9393	0.9301	0.8480	0.9390	0.8340	0.8834
4	0.9374	0.9272	0.8446	0.9371	0.8444	0.8883
5	0.9361	0.9244	0.8295	0.9358	0.8371	0.8837

validated circRNA-miRNA interactions were randomly divided into five parts. Take each part as a positive sample and an equal number of unlabeled samples from the benchmark data as negative samples to form a test set. At the same time, perform the same operation on the remaining four parts to obtain a training set. This operation is performed until the loop is completed five times.

To measure the performance of GCNCMI more comprehensively, we used AUC, AUPR, Recall, Accuracy (Acc), precision (Pre), and F1 Scores. The definitions of each indicator are as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)$$

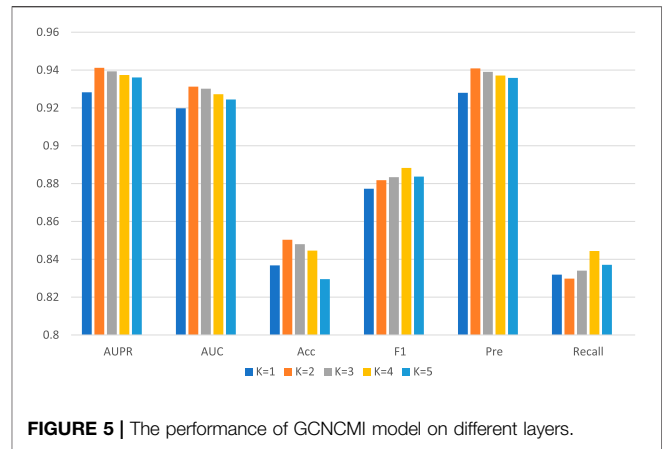
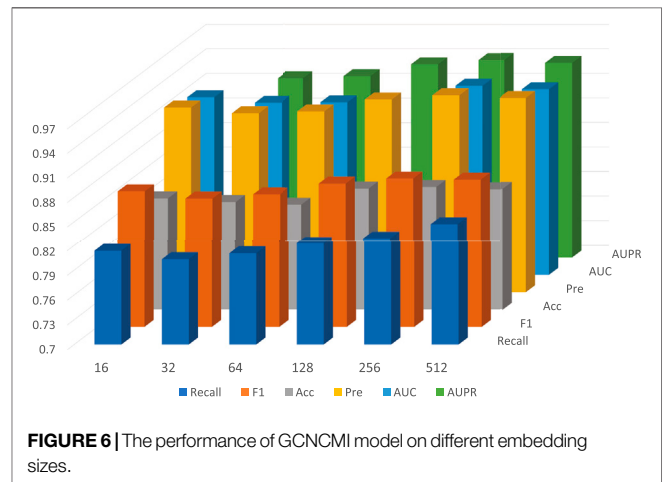


TABLE 4 | The performance of GCNCMI model on different embedding sizes.

D	AUPR	AUC	Acc	Pre	Recall	F1
16	0.9260	0.9170	0.8360	0.9257	0.8136	0.8660
32	0.9190	0.9102	0.8316	0.9187	0.8032	0.8571
64	0.9215	0.9110	0.8287	0.9212	0.8105	0.8623
128	0.9361	0.9265	0.8485	0.9357	0.8230	0.8757
256	0.9412	0.9312	0.8503	0.9409	0.8298	0.8819
512	0.9376	0.9268	0.8475	0.9373	0.8303	0.8806



$$Precision = \frac{TP}{TP + FP} \quad (10)$$

$$Recall = \frac{TP}{TP + FN} \quad (11)$$

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall} \quad (12)$$

Where *TP* and *FP* represent the number of correctly classified samples and the number of misclassified samples in known circRNA-miRNA interactions, respectively, *TN* represents the number of correctly predicted unrelated circRNA-miRNA interactions, and *FN* represents the

TABLE 5 | Performance comparison of different methods under five-fold cross validation.

Methods	AUC	AUPR	Acc	Pre	Recall	F1
AE-RF	0.7662 ± 0.0050	0.8239 ± 0.0042	0.8333 ± 0.0013	0.8923 ± 0.0019	0.9279 ± 0.0019	0.9097 ± 0.0010
DMFCDA	0.7321 ± 0.0240	0.7115 ± 0.0171	0.6975 ± 0.0112	0.8160 ± 0.0265	0.7729 ± 0.1112	0.7938 ± 0.0707
DMFMDA	0.7922 ± 0.0057	0.8230 ± 0.0089	0.7307 ± 0.0049	0.7030 ± 0.0080	0.7246 ± 0.0116	0.7136 ± 0.0065
KATZHMDA	0.8469 ± 0.0017	0.8647 ± 0.0019	0.8073 ± 0.0030	0.8511 ± 0.0055	0.7227 ± 0.0106	0.7816 ± 0.0071
NTSHMDA	0.8526 ± 0.0016	0.8772 ± 0.0018	0.6276 ± 0.0083	0.7556 ± 0.0518	0.4040 ± 0.0531	0.5264 ± 0.0486
SDLDA	0.7875 ± 0.0307	0.8286 ± 0.0189	0.6693 ± 0.0019	0.8287 ± 0.0108	0.7891 ± 0.0809	0.8084 ± 0.0706
GCNCMI	0.9320 ± 0.0014	0.9396 ± 0.0406	0.8530 ± 0.0134	0.9413 ± 0.0204	0.8345 ± 0.0301	0.8846 ± 0.0068

number of prediction errors in unrelated miRNA-circRNA interactions. *F1* is a weighted average of model precision and Recall.

3.2 Cross-Validation Results

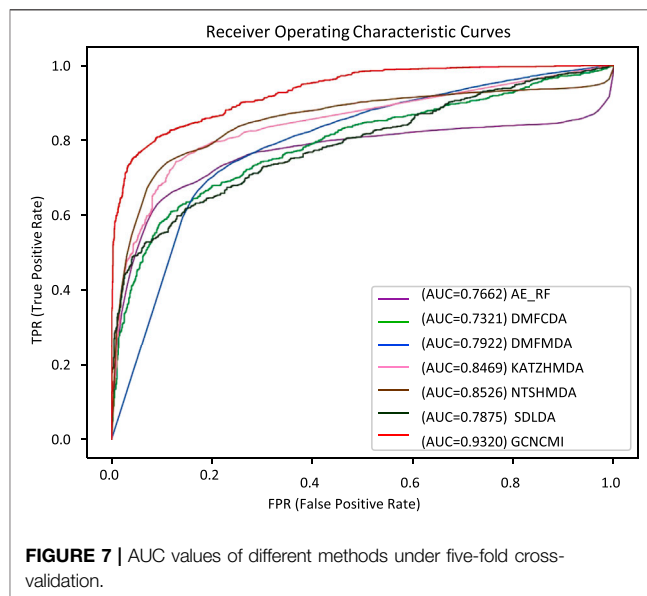
We performed five-fold cross-validations to evaluate the performance of the GCNCMI model in predicting circRNA-miRNA interactions. The experimental results of the five-fold cross-validation are shown in **Table 2**. As shown in the table, the AUC of the five-fold cross-validations are: 0.9288, 0.9352, 0.9372, 0.9282, 0.9312. On the AUPR, the AUPR of the five-fold cross-validations are 0.9293, 0.9428, 0.9453, 0.9396, 0.9412, respectively. In addition, we also plotted the ROC curve of GCNCMI, as shown in **Figure 4**. The above experimental results show that GCNCMI has good performance in predicting unknown circRNA-miRNA interactions.

3.3 Parameter Influence

For GCNCMI, two essential parameters affect its performance: *K* (the number of layers) and *D* (the dimension of the embedding vector). When *K* is 2, and *D* is 256, our model GCNCMI achieves the best performance under five-fold cross-validation.

The setting of the number of layers *K* indicates that our final embedding model incorporates the information of *K*-hop neighbor nodes in the bipartite graph, which can learn more hidden interaction information between nodes for the neural network. **Table 3** lists the detailed values, and **Figure 5** shows the trend chart for different layers. We tried from 1 to 5 layers for the number of layers of the model and found that the model's accuracy at the beginning will increase with the increase of the number of layers. The best performance of the model is when the layer is 2. As the number of network layers increases, the hidden feature pairs of nodes tend to converge to the same value, which leads to an over-smoothing problem in the network.

On the other hand, under the framework of five-fold cross-validation, we conducted experiments for *D* in 16, 32, 64, 128, 256, 512, and other 6 cases; the detailed data is shown in **Table 4**. In general, as the dimension of the embedding vector increases, the expressive power of the model increases. But as can be seen from the **Figure 6**, from 16, 32, 64, 128, 256, the model's performance has been increasing at first, but at 256, the commission has reached the maximum value. As *D* continues to grow, it will adversely affect the model's performance.

**TABLE 6** | The number of circRNAs, miRNAs, and circRNA-miRNA interactions included in the independent test dataset.

circRNA	miRNA	interactions	unlabeled interactions
1,502	494	9,386	9,386

3.4 Compared With State-Of-The-Art Methods

Since circRNA and miRNA interaction is a relatively new field, GCNCMI is the first method we know to predict the interaction between circRNA and miRNA, but other advanced methods in bioinformatics still provide us with reference. To better verify the performance of GCNCMI in inferring the interaction between circRNA and miRNA. We compare GCNCMI with six other state-of-the-art methods in bioinformatics.

Considering the scarcity of related biological resources, in calculating biological similarity, we only calculated Gaussian interaction profile biological similarity (GIP). In addition, since the adjacency matrix initialized each time is different, it requires us to re-mine the information in the bipartite graph. Strictly speaking, in similarity-based methods [AE-RF (Deepthi and Jereesh, 2021), KATZHMDA (Chen et al., 2017),



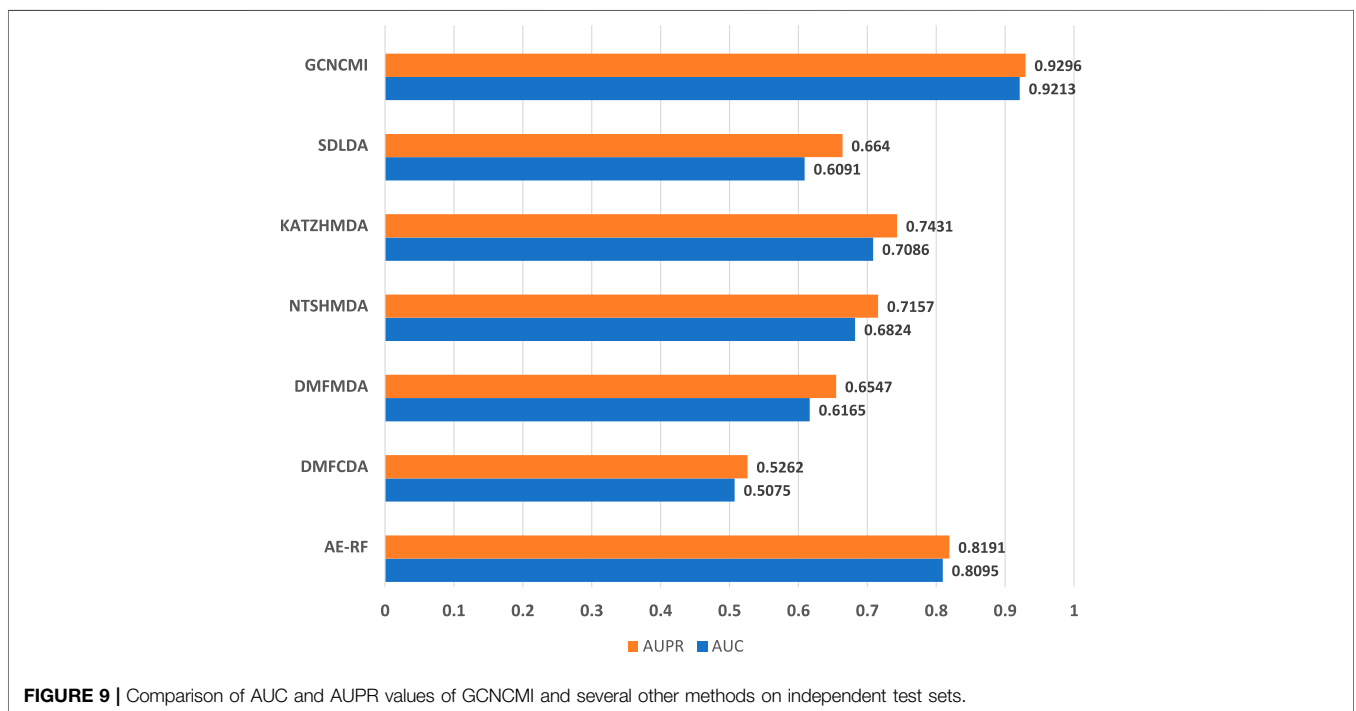
NTSHMDA (Luo and Long, 2018)], the similarity matrix is recalculated each time during the cross-validation process. In the SDLDA method, we used SVD singular value decomposition to obtain linear features of circRNAs and miRNAs. The DMFMDA method chooses a Bayesian loss function over the loss function instead of the mean squared error.

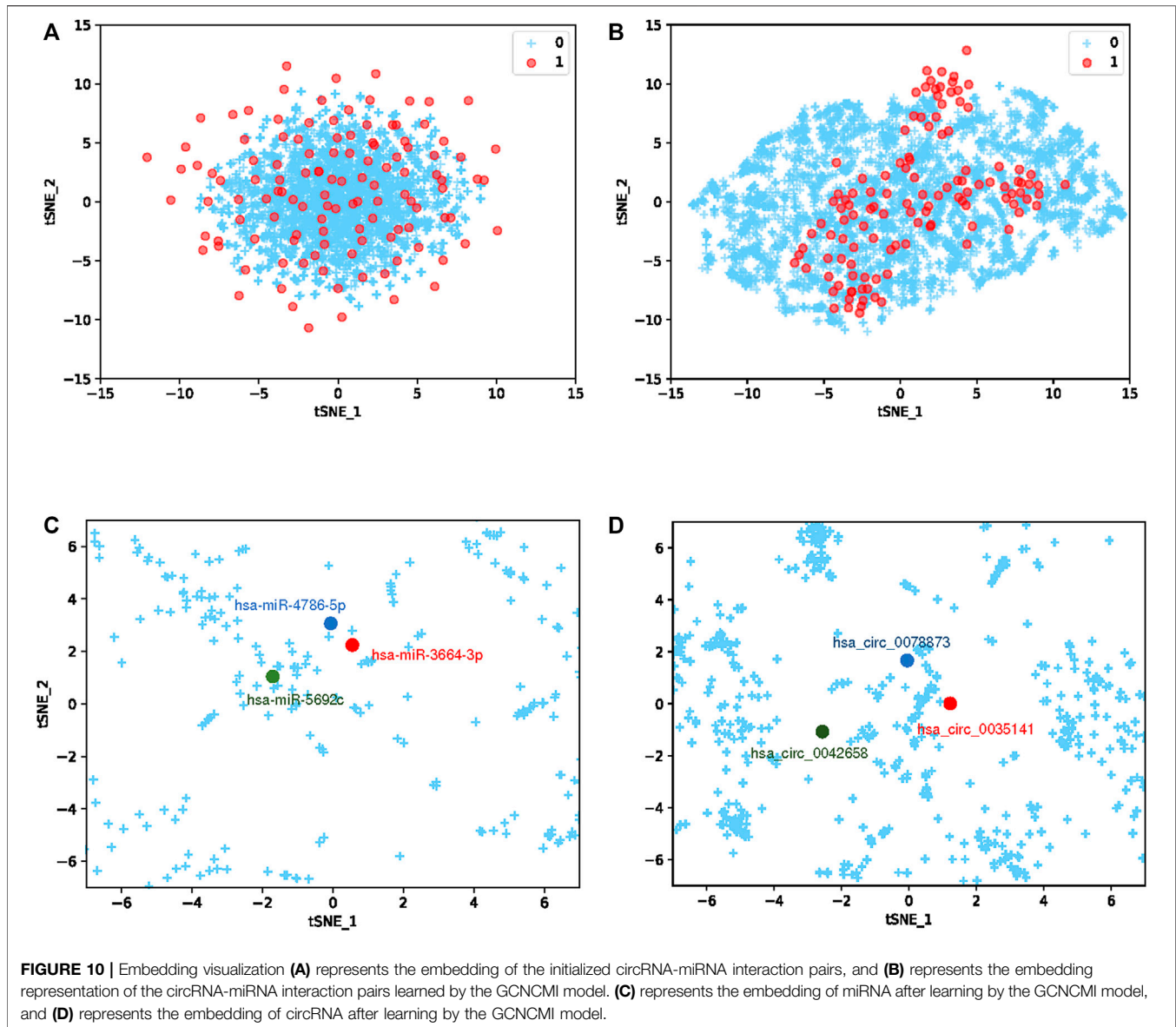
We performed a ten-times, five-fold cross-validation of GCNCMI with six advanced methods, changing the random number seed each time, and calculated the mean and standard deviation of 10 experiments. **Table 5** lists several methods such as

AE-RF (Deepthi and Jereesh, 2021), DMFCDA (Liu et al., 2020), DMFMDA (Liu et al., 2020), KATZHMDA (Chen et al., 2017), NTSHMDA (Luo and Long, 2018), SDLDA (Zeng et al., 2020), and compared with the GCNCMI model. **Figure 7** plots the AUC curves to compare the seven methods. As can be seen from **Table 5** and **Figure 7**, GCNCMI mines the high-order interactions between circRNA and miRNA; GCNCMI is higher than other methods in most indicators, among which the AUC value of GCNCMI is 0.9320, and the highest among different methods is NTSHMDA, whose AUC value is 0.8526, which is 7.94% lower than GCNCMI. GCNCMI value of AUPR is 0.9396, which is 6.24% higher than the second-best method, NTSHMDA. The above experimental results show that our model performs well in predicting the relationship between circRNA and miRNA.

The radar **Figure 8** shows the performance of GCNCMI on AUC, AUPR, ACC, Recall, F1, Pre. The evaluation index is set from 0 to 1. As shown from **Figure 8**, the distance between the point and the center of the circle reflects the level of the value. It is evident that GCNCMI is better than other methods in predicting the circRNA-miRNA relationship.

To further verify the accuracy of the GCNCMI model in circRNA-miRNA association prediction, we retrieved the data from the PubMed database, removed the known relationships that overlapped with the training dataset, and established a 9,386 miRNA-circRNA association relationship, 494 miRNAs, an independent test set of 1,502 circRNAs, and 9,386 unlabeled interactions were randomly selected from the benchmark dataset as negative samples. The specific information of the independent test set can be found in **Table 6**. Although there may be a small part of the independent test set and the unknown overlapping relationship in the training set, it can be ignored because it





occupies a small proportion of the entire unvalidated sample set. The basic model for predicting circRNA-miRNA associations was obtained by training on our data set and tested on the independent test set. The test results are as **Figure 9**. The AUC of the GCNCMI model reached 0.9213, and the AUPR value reached 0.9296, which is higher than several other methods of comparison. The independent test results further showed that GCNCMI is an effective tool for inferring miRNA-circRNA associations.

3.5 Embedding Visualization

To more clearly demonstrate the learning ability of the GCNCMI, We use T-SNE (Van der Maaten and Hinton, 2008) to visualize the embedding of circRNA-miRNA interaction pairs. Because the number of unknown relationships is much larger than the number of known associations, and to better visualize the overall

mining of higher-order relationships by GCNCMI, we choose to visualize more unlabeled samples than labeled samples. The main goal of T-SNE is to convert multi-dimensional datasets into low-dimensional datasets. Compared with other dimensionality reduction algorithms, T-SNE is the most effective technique in data visualization. Since T-SNE is not a linear dimensionality reduction technique, it can capture the complex manifold structure of high-dimensional data. We initially a 32-dimensional vector to represent miRNA and circRNA. To explore the similarity between vector representations, we used the T-SNE algorithm to reduce the vector to 2-dimensional, as shown in **Figure 10A**. The blue + represents unknown miRNA-circRNA interaction pairs, and the red dots represent the known circRNA-miRNA interaction pairs. **Figure 10B** shows the embedding of the circRNA-miRNA interactions learned by the GCNCMI model. Comparing **Figures 10A,B**, it can be seen that GCNCMI has a good effect on mining

TABLE 7 | The top 10 circRNAs with the closest relationship to hsa-miR-622 predicted by GCNCMI model.

Rank	CircRNA	Evidence(PMID)	Score
1	hsa_circ_0000231	34183076	0.8822
2	hsa_circ_0101432	Unconfirmed	0.8820
3	hsa_circ_0119872	33579337	0.8815
4	hsa_circ_0008574	32616043	0.8798
5	hsa_circ_0000211	31668923	0.8796
6	hsa_circ_0001273	35567340	0.8712
7	hsa_circ_0086902	Unconfirmed	0.8592
8	hsa_circ_KCNQ5	35413218	0.8542
9	hsa_circ_0101432	35297300	0.8498
10	hsa_circ_0006000	Unconfirmed	0.8469

TABLE 8 | The top 10 circRNAs with the closest relationship to hsa-miR-149-5p predicted by GCNCMI model.

Rank	CircRNA	Evidence(PMID)	Score
1	hsa_circ_0061140	32224273	0.8737
2	hsa_circ_0075341	31706100	0.8722
3	hsa_circ_0008956	34153672	0.8702
4	hsa_circ_0000654	31778020	0.8693
5	hsa_circ_0051239	Unconfirmed	0.8689
6	hsa_circ_ROBO2	34649241	0.8673
7	hsa_circ_0011385	34720052	0.8672
8	hsa_circ_0087352	35286916	0.8671
9	hsa_circ_0123996	32707301	0.8661
10	hsa_circ_0031059	Unconfirmed	0.8648

high-order interactions between miRNAs and circRNAs, and the GCNCMI can better use the known interaction pairs to mine potential miRNA-circRNA interaction pairs. In addition, we also visualized the learned circRNA embeddings and miRNA embeddings. **Figure 10C** shows the learned miRNA embeddings. We used the GCNCMI model to predict the top 30 circRNAs most closely associated with each miRNA, and also predicted the top 30 miRNAs most closely associated with each circRNA. The hsa-miR-4786-5p and hsa-miR-3664-3p were associated with nine similar circRNAs, and hsa-miR-4786-5p and hsa-miR-5692c were associated with five similar circRNAs. Therefore, the hsa-miR-4786-5p is more similar to hsa-miR-3664-3p. It can also be seen from **Figure 10C** that the distance between hsa-miR-4786-5p and hsa-miR-3664-3p is closer. **Figure 10D** shows the visualization of the embedding of circRNAs after model learning. The hsa-circ-0078873 and hsa-circ-0042658 were associated with three similar miRNAs, and hsa-circ-0035141 and hsa-circ-0078873 were associated with seven similar miRNAs. Therefore, hsa-circ-0078873 is closer to hsa-circ-0035141, and it can be seen from **Figure 10D** that hsa-circ-0078873 is closer to hsa-circ-0035141. The experimental results show that GCNCMI can effectively learn the potential higher-order interactions between miRNAs and circRNAs.

3.6 Case Studies

It is of great significance to discover unknown associations between circRNAs and miRNAs. We selected two miRNAs, hsa-miR-622 and hsa-miR-149-5p, for case studies. Specifically, we first delete the circRNAs that have been

experimentally validated for the selected miRNAs. Then, the remaining circRNAs were sorted in descending order according to the values predicted by the GCNCMI model. The following shows the results of the normalized prediction scores of the GCNCMI model. Finally, we screened the top 10 circRNAs and collected evidence in the published literature for testing.

miR-622 (Lu et al., 2022) is a miRNA of 13q31.3 in the eukaryotic genome, and its expression is mainly in the nucleus. In recent years, studies have found that miR-622 can functionally inhibit the malignant proliferation of cells, which is helpful for cancer treatment. In recent years, miR-149-3p (Yang et al., 2017) can effectively inhibit the proliferation and apoptosis of malignant tumors. Recent studies have found that miR-149-3p can increase the sensitivity of drugs. **Table 7** and **Table 8** list the top 10 candidate circRNAs of hsa-miR-622 and hsa-miR-149-5p. We selected the top 10 candidate circRNAs as our predicted circRNAs, respectively, and finally, we compared the predicted results with the experimentally validated interactions. It can be seen that 7 of hsa-miR-622 were confirmed by existing evidence, and 8 of hsa-miR-149-5p were confirmed by existing evidence. It should be noted that unproven associations may exist and require further experimental verification.

4 CONCLUSION

CircRNAs are circular non-coding RNAs with regulatory functions, most of which exist in eukaryotic excerpts, and most circRNAs are composed of exons. Because circRNAs are less affected by nucleases, circRNAs are more stable than linear RNAs. Current studies have shown that circRNAs can competitively adsorb miRNAs, and circRNAs can bind to proteins to inhibit the activity. Therefore, there is an urgent need to explore the relationship between circRNA and miRNA. However, because traditional biological experiments are time-consuming and labor-intensive, a more efficient method is needed to explore the potential relationship between circRNA and miRNA.

In this paper, we proposed a graph convolutional neural network prediction model for circRNA and miRNA interactions. To fully exploit the potential high-order interactions between circRNAs and miRNAs, we designed a graph convolutional neural network method to propagate the interaction's relation recursively without computing the similarity of circRNAs and miRNAs. The experimental results demonstrated the excellent performance of GCNCMI in predicting the interactions between circRNAs and miRNAs. The results of independent tests indicate that the GCNCMI model has good generalization performance in predicting unknown circRNA and miRNA relationships. Finally, a case study compared our predictions with those validated by biological experiments, further demonstrating the model's excellent predictive performance. The above results indicate that GCNCMI is an excellent method for predicting the potential interactions between circRNAs and miRNAs.

While GCNCMI has excellent performance, it also has some limitations. First, due to the scarcity of biological resources,

GCNCMI only uses the association data of circRNAs and miRNAs, and the quality of the data will affect the performance of GCNCMI model training. In the future, using heterogeneous data from multiple perspectives will be considered to improve the model's performance further.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

JH and LD designed and implemented the prediction method. JH, PX, CC, and JZ analyzed the data and wrote the manuscript. LD reviewed and revised the manuscript.

REFERENCES

- Ashwal-Fluss, R., Meyer, M., Pamudurti, N. R., Ivanov, A., Bartok, O., Hanan, M., et al. (2014). Circrna Biogenesis Competes with Pre-mrna Splicing. *Mol. Cell* 56, 55–66. doi:10.1016/j.molcel.2014.08.019
- Bartel, D. P. (2004). MicroRNAs. *Cell* 116, 281–297. doi:10.1016/s0092-8674(04)00045-5
- Butcher, S. E., and Brow, D. A. (2005). Towards Understanding the Catalytic Core Structure of the Spliceosome. *Biochem. Soc. Trans.* 33, 447–449. doi:10.1042/bst0330447
- Calin, G. A., Sevignani, C., Dumitru, C. D., Hyslop, T., Noch, E., Yendamuri, S., et al. (2004). Human MicroRNA Genes Are Frequently Located at Fragile Sites and Genomic Regions Involved in Cancers. *Proc. Natl. Acad. Sci. U.S.A.* 101, 2999–3004. doi:10.1073/pnas.0307323101
- Capel, B., Swain, A., Nicolis, S., Hacker, A., Walter, M., Koopman, P., et al. (1993). Circular Transcripts of the Testis-Determining Gene Sry in Adult Mouse Testis. *Cell* 73, 1019–1030. doi:10.1016/0092-8674(93)90279-y
- Chen, B., and Huang, S. (2018). Circular Rna: An Emerging Non-Coding Rna as a Regulator and Biomarker in Cancer. *Cancer Lett.* 418, 41–50. doi:10.1016/j.canlet.2018.01.011
- Chen, X., Huang, Y. A., You, Z. H., Yan, G. Y., and Wang, X. S. (2017). A Novel Approach Based on Katz Measure to Predict Associations of Human Microbiota with Non-Infectious Diseases. *Bioinformatics* 33, 733–739. doi:10.1093/bioinformatics/btw715
- Dayun, L., Junyi, L., Yi, L., Qihua, H., and Deng, L. (2021). Mgmtmda: Predicting Microbe-Disease Associations via Multi-Component Graph Attention Network. *IEEE/ACM Trans. Comput. Biol. Bioinforma.* doi:10.1109/tcbb.2021.3116318
- Deepthi, K., and Jereesh, A. S. (2021). Inferring Potential CircRNA-Disease Associations via Deep Autoencoder-Based Classification. *Mol. Diagn. Ther.* 25, 87–97. doi:10.1007/s40291-020-00499-y
- Deng, L., Huang, Y., Liu, X., and Liu, H. (2022). Graph2MDA: A Multi-Modal Variational Graph Embedding Model for Predicting Microbe-Drug Associations. *Bioinformatics* 38, 1118–1125. doi:10.1093/bioinformatics/btab792
- Deng, L., Yang, J., and Liu, H. (2020). Predicting Circrna-Disease Associations Using Meta Path-Based Representation Learning on Heterogenous Network. In 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) (IEEE), 5–10. doi:10.1109/bibm49941.2020.9313215
- Dong, R., Ma, X.-K., Chen, L.-L., and Yang, L. (2017). Increased Complexity of Circrna Expression During Species Evolution. *RNA Biol.* 14, 1064–1074. doi:10.1080/15476286.2016.1269999

FUNDING

This work was supported by the National Natural Science Foundation of China under Grant No. 61972422.

ACKNOWLEDGMENTS

We are grateful for resources from the High Performance Computing Center of Central South University. The work was carried out at National Supercomputer Center in Tianjin, and the calculations were performed on TianHe.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.959701/full#supplementary-material>

- Ford, E., and Ares, M. (1994). Synthesis of Circular Rna in Bacteria and Yeast Using Rna Cyclase Ribozymes Derived from a Group I Intron of Phage T4. *Proc. Natl. Acad. Sci. U.S.A.* 91, 3117–3121. doi:10.1073/pnas.91.8.3117
- Gesteland, R., Cech, T., and Atkins, J. (2006). *The Rna World*. 3rd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press. [Google Scholar].
- Grabowski, P. J., Zaug, A. J., and Cech, T. R. (1981). The Intervening Sequence of the Ribosomal Rna Precursor Is Converted to a Circular Rna in Isolated Nuclei of Tetrahymena. *Cell* 23, 467–476. doi:10.1016/0092-8674(81)90142-2
- Hamilton, W., Ying, Z., and Leskovec, J. (2017). Inductive Representation Learning on Large Graphs. *Adv. Neural Inf. Process. Syst.* 30.
- Hansen, T. B., Jensen, T. I., Clausen, B. H., Bramsen, J. B., Finsen, B., Damgaard, C. K., et al. (2013). Natural Rna Circles Function as Efficient MicroRNA Sponges. *Nature* 495, 384–388. doi:10.1038/nature11993
- He, H., Zhang, J., Gong, W., Liu, M., Liu, H., Li, X., et al. (2022). Involvement of Circrna Expression Profile in Diabetic Retinopathy and its Potential Diagnostic Value. *Front. Genet.* 13, 833573. doi:10.3389/fgene.2022.833573
- He, X., Zhang, H., Kan, M. Y., and Chua, T. S. (2016). Fast Matrix Factorization for Online Recommendation with Implicit Feedback. *ACM*. doi:10.1145/2911451.2911489
- Kos, A., Dijkema, R., Arnberg, A. C., Van der Meide, P. H., and Schellekens, H. (1986). The Hepatitis Delta (δ) Virus Possesses a Circular RNA. *Nature* 323, 558–560. doi:10.1038/323558a0
- Lan, W., Zhu, M., Chen, Q., Chen, B., Liu, J., Li, M., et al. (2020). Circr2cancer: A Manually Curated Database of Associations Between Circrnas and Cancers. *Database (Oxford)* 2020. doi:10.1093/database/baaa085
- Lee, R. C., Feinbaum, R. L., and Ambros, V. (1993). The *c. elegans* Heterochronic Gene *Lin-4* Encodes Small Rnas with Antisense Complementarity to *Lin-14*. *Cell* 75, 843–854. doi:10.1016/0092-8674(93)90529-y
- Li, Z., Zhou, Y., Yang, G., He, S., Qiu, X., Zhang, L., et al. (2019). Using Circular Rna Smarca5 as a Potential Novel Biomarker for Hepatocellular Carcinoma. *Clin. Chim. Acta* 492, 37–44. doi:10.1016/j.cca.2019.02.001
- Liu, D., Huang, Y., Nie, W., Zhang, J., and Deng, L. (2021). Smlaf: Mirna-Disease Associations Prediction Based on Stacked Autoencoder and Xgboost. *BMC Bioinforma.* 22, 1–18. doi:10.1186/s12859-021-04135-2
- Liu, M., Wang, Q., Shen, J., Yang, B. B., and Ding, X. (2019). Circbank: A Comprehensive Database for Circrna with Standard Nomenclature. *RNA Biol.* 16, 899–905. doi:10.1080/15476286.2019.1600395
- Liu, Y., Wang, S., Zhang, J., Zhang, W., Zhou, S., and Li, W. (2020). Dmfmda: Prediction of Microbe-Disease Associations Based on Deep Matrix Factorization Using Bayesian Personalized Ranking. *IEEE/ACM Trans. Comput. Biol. Bioinform* PP, 1763–1772. doi:10.1109/TCBB.2020.3018138

- Lu, J., Xie, Z., Xiao, Z., and Zhu, D. (2022). The Expression and Function of Mir-622 in a Variety of Tumors. *Biomed. Pharmacother.* 146, 112544. doi:10.1016/j.biopha.2021.112544
- Luo, J., and Long, Y. (2020). Ntshmda: Prediction of Human Microbe-Disease Association Based on Random Walk by Integrating Network Topological Similarity. *IEEE/ACM Trans. Comput. Biol. Bioinform* 17, 1341–1351. doi:10.1109/TCBB.2018.2883041
- Memczak, S., Jens, M., Elefsinioti, A., Torti, F., Krueger, J., Rybak, A., et al. (2013). Circular Rnas Are a Large Class of Animal Rnas with Regulatory Potency. *Nature* 495, 333–338. doi:10.1038/nature11928
- Nikolakopoulos, A. N., and Karypis, G. (2019). Recwalk: Nearly Uncoupled Random Walks for Top-N Recommendation. *Proc. Twelfth ACM Int. Conf. Web Search Data Min.*, 150–158.
- O'Brien, J., Hayder, H., Zayed, Y., and Peng, C. (2018). Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. *Front. Endocrinol. (Lausanne)* 9, 402. doi:10.3389/fendo.2018.00402
- Palazzo, A. F., and Lee, E. S. (2015). Non-Coding Rna: What Is Functional and What Is Junk? *Front. Genet.* 6, 2. doi:10.3389/fgene.2015.00002
- Rendle, S. (2010). Factorization Machines. In *ICDM 2010, The 10th IEEE International Conference on Data Mining, Sydney, Australia, 14–17. December 2010*. doi:10.1109/icdm.2010.127
- Sanger, H. L., Klotz, G., Riesner, D., Gross, H. J., and Kleinschmidt, A. K. (1976). Viroids Are Single-Stranded Covalently Closed Circular Rna Molecules Existing as Highly Base-Paired Rod-Like Structures. *Proc. Natl. Acad. Sci. U.S.A.* 73, 3852–3856. doi:10.1073/pnas.73.11.3852
- Soslau, G. (2018). Circular Rna (Circrna) Was an Important Bridge in the Switch from the Rna World to the Dna World. *J. Theor. Biol.* 447, 32–40. doi:10.1016/j.jtbi.2018.03.021
- Steitz, T. A., and Moore, P. B. (2003). Rna, the First Macromolecular Catalyst: The Ribosome Is a Ribozyme. *Trends Biochem. Sci.* 28, 411–418. doi:10.1016/s0968-0004(03)00169-5
- Van der Maaten, L., and Hinton, G. (2008). Visualizing Data Using T-Sne. *J. Mach. Learn. Res.* 9.
- Wang, X., He, X., Cao, Y., Liu, M., and Chua, T.-S. (2019). Kgat: Knowledge Graph Attention Network for Recommendation. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 950–958.
- Xu, K., Li, C., Tian, Y., Sonobe, T., Kawarabayashi, K.-i., and Jegelka, S. (2018). Representation Learning on Graphs with Jumping Knowledge Networks. *International Conference on Machine Learning*. PMLR, 5453–5462.
- Yang, D., Du, G., Xu, A., Xi, X., and Li, D. (2017). Expression of Mir-149-3p Inhibits Proliferation, Migration, and Invasion of Bladder Cancer by Targeting S100a4. *Am. J. Cancer Res.* 7, 2209–2219.
- Zeng, M., Lu, C., Zhang, F., Li, Y., Wu, F.-X., Li, Y., et al. (2020). Sldda: Lncrna-Disease Association Prediction Based on Singular Value Decomposition and Deep Learning. *Methods* 179, 73–80. doi:10.1016/j.ymeth.2020.05.002

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

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 He, Xiao, Chen, Zhu, Zhang and Deng. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.