




MDGF-MCEC: a multi-view dual attention embedding model with cooperative ensemble learning for CircRNA-disease association prediction

Qunzhuo Wu, Zhaohong Deng , Xiaoyong Pan , Hong-Bin Shen, Kup-Sze Choi, Shitong Wang, Jing Wu and Dong-Jun Yu 

Corresponding author: Zhaohong Deng, Jiangnan University, Wuxi, Jiangsu 214012, China. E-mail: dengzhaohong@jiangnan.edu.cn

Abstract

Circular RNA (circRNA) is closely involved in physiological and pathological processes of many diseases. Discovering the associations between circRNAs and diseases is of great significance. Due to the high-cost to verify the circRNA-disease associations by wet-lab experiments, computational approaches for predicting the associations become a promising research direction. In this paper, we propose a method, MDGF-MCEC, based on multi-view dual attention graph convolution network (GCN) with cooperative ensemble learning to predict circRNA-disease associations. First, MDGF-MCEC constructs two disease relation graphs and two circRNA relation graphs based on different similarities. Then, the relation graphs are fed into a multi-view GCN for representation learning. In order to learn high discriminative features, a dual-attention mechanism is introduced to adjust the contribution weights, at both channel level and spatial level, of different features. Based on the learned embedding features of diseases and circRNAs, nine different feature combinations between diseases and circRNAs are treated as new multi-view data. Finally, we construct a multi-view cooperative ensemble classifier to predict the associations between circRNAs and diseases. Experiments conducted on the CircR2Disease database demonstrate that the proposed MDGF-MCEC model achieves a high area under curve of 0.9744 and outperforms the state-of-the-art methods. Promising results are also obtained from experiments on the circ2Disease and circRNADisease databases. Furthermore, the predicted associated circRNAs for hepatocellular carcinoma and gastric cancer are supported by the literature. The code and dataset of this study are available at <https://github.com/ABard0/MDGF-MCEC>.

Keywords: circRNA-disease association, attention mechanism, graph convolution network, multi-view cooperation learning, ensemble learning

Introduction

Circular RNAs (circRNAs) are a type of single-stranded RNA that are widely present at eukaryotic cells. They have low expression levels and relatively stable structures compared with messenger RNAs (mRNAs) [1]. In addition, circRNAs are expressed in a tissue-specific way [2]. Especially, they can function as a sponge for microRNAs (miRNAs) to regulate gene expression [3–5] or directly act as a carrier to bind specific RNA-binding proteins [6–8]. In addition, circRNAs can also be directly involved in encoding some short proteins [9]. Due to

the advancement and breakthrough of bioinformatics and RNA sequencing technologies, a large volume of circRNAs have been identified in cells and tissues [10].

Increasing evidence [11, 12] shows that circRNAs expression is often abnormal in many diseases [13, 14], such as cancer [15]. Thus, circRNAs can be used as a reliable disease biomarker [16]. In recent years, the associations between circRNAs and diseases have become a hotspot in biomedical research. A number of relevant databases have thus been established, which can be divided into three main types: databases

Qunzhuo Wu is a master student at Jiangnan University. His research interests include text data mining, bioinformatics and machine learning.

Dr Zhaohong Deng is a full professor in the School of Artificial Intelligence and Computer Science of Jiangnan University, Key Laboratory of Computational Neuroscience and Brain-Inspired Intelligence (LCNBI) and ZJLab. His research interests include bioinformatics, pattern recognition, computational intelligence and their applications.

Dr Xiaoyong Pan is an assistant professor in the Department of Automation of Shanghai Jiao Tong University. His research interests include pattern recognition, image processing and data mining in bioinformatics.

Dr Hong-Bin Shen is a full professor at Shanghai Jiao Tong University. His research interests include basic theory of pattern recognition and artificial intelligence, protein engineering and complex network in bioinformatics.

Dr Kup-Sze Choi is a full professor at the Hong Kong Polytechnic University. His research interests include pattern recognition, data mining and image processing.

Dr Shitong Wang is a full professor in the School of Artificial Intelligence and Computer Science of Jiangnan University. His research interests include pattern recognition, computer application and light industry information technology.

Dr Jing Wu is a full professor of the School of Biotechnology and the Key Laboratory of Industrial Biotechnology Ministry in Jiangnan University. Her research interests include the location of enzyme-encoded genes, the construction of high-efficiency expression systems for genetic engineering, the protein engineering of enzyme preparations and the genetic engineering of metabolic engineering strains.

Dr Dong-Jun Yu is a full professor of the School of Computer Science and Engineering. His research interests include bioinformatics, pattern recognition and intelligent computing.

Received: May 3, 2022. **Revised:** June 19, 2022. **Accepted:** June 26, 2022

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

depositing circRNAs information, such as Circ2Traits [17] and circBase [18], databases containing diseases information, such as Mesh [19, 20], and databases recording the associations between circRNAs and diseases, such as CircR2Disease [21], Circ2Disease [22] and circRNADisease [23].

Although some associations between circRNAs and diseases have been detected by wet-lab experiments, a large number of potential disease-circRNA associations still remain unknown. Due to the high-cost and time-consuming of wet-lab experiments, machine learning technology in recent years provides a feasible solution to address this challenge [24]. The idea is to train an intelligent prediction model based on the known circRNA-disease associations and then use the trained model to predict the unknown associations. In this way, wet-lab experiments can be further carried out under the guidance of the prediction results. This solution improves the efficiency in discovering the associations between circRNAs and diseases, becoming an important research direction in the field of bioinformatics.

Recently, many computational methods based on machine learning were proposed to predict unknown circRNA-disease associations. For example, Wang et al. [25] proposed a deep convolution neural network model based on multi-source data by establishing a similarity matrix. Deepthi and Jereesh [26] proposed the AE-RF model by using the combination of autoencoder and random forest. Zheng et al. [27] proposed the ICDA-CGR model, which introduced circRNA sequence information and quantified the nonlinear relationship in circRNA sequences by chaos game representation (CGR) based on the sequence position information. Wang et al. [28] proposed the IMS-CDA method, which combined the disease semantic similarity, disease Jaccard similarity, Gaussian interaction profile kernel similarity and circRNA similarity information to extract the hidden features using stacked auto-encoder (SAE). Wang et al. [29] proposed a semi-supervised generative adversarial network (GAN) model, called SGANRDA, which fused natural language features of circRNA sequences to pre-train the GAN network and then fine-tuned the network parameters using labeled samples. Li et al. [30] proposed the SIMCCDA based on matrix completion to predict the associations between circRNAs and diseases by Speedup Inductive Matrix Completion (SIMC). Wei and Liu [31] proposed iCircDA-MF model to introduce gene information into the limited training data and construct the circRNA-gene-disease relation network to expand the data sources, this model also used matrix decomposition and completion techniques to reduce the feature noise. Niu et al. [32] proposed GMNN2DC method by using a graph Markov neural network (GMNN) for circRNA-disease association prediction. In this work, multi-sourced database was integrated and a graph Markov convolutional neural network was employed to make prediction by integrating a graph autoencoder and variational inference. Chen et al. [33] proposed RGCNCDA method for circRNA-disease association prediction.

This method first integrated multi-similarity networks to construct a global heterogeneous network. Then, it employed the random walk with restart to learn low-dimensional and high-order information as the topological features. A prediction model based on an R-GCN encoder and a DistMult decoder was also built to make prediction.

Despite the progress of computational circRNA-disease association predictions, major challenges still remain. First, most of the existing methods are limited to the extraction of raw features from diseases and circRNAs, where new discriminant features are learned for subsequent classifier training. An advanced feature learning strategy needs to be developed to improve the effectiveness of the existing methods. Second, existing classifier learning methods rely on the fusion data of disease features and circRNA features to train the classifier directly, which is rather simple. Thus, it is necessary to develop more comprehensive classifier learning techniques for predicting circRNA-disease associations.

To address the challenges mentioned above, a novel method MDGF-MCEC integrating multi-view dual attention embedding learning with cooperative ensemble learning is proposed for predicting circRNA-disease associations. In the MDGF-MCEC, different disease relation graphs and circRNA relation graphs are first constructed by using similarities of circRNAs and similarities of diseases from different views. Then, multi-view graphs are fed into graph convolution networks (GCNs), respectively, to extract the deep features of different views. In particular, a GCN with a dual-attention mechanism is adopted to learn the embedded features of circRNAs and diseases from different relation graphs. The dual-attention mechanism adjusts and balances the weights for different views of features at the channel level and spatial level to enhance the discriminate ability of the learned embedded features. In order to make full use of embedded features of circRNAs and diseases, nine circRNA-disease feature combinations from different views are obtained, and then a multi-view cooperative ensemble classifier is constructed by introducing a multi-view cooperative decision and ensemble classification mechanism.

The proposed method MDGF-MCEC achieves a high area under curve (AUC) of 0.9744, 0.9571 and 0.9725 on benchmark datasets from the circR2disease, circ2Disease and circRNADisease databases using 5-fold cross-validation, respectively. The results from case studies of gastric cancer and hepatoma carcinoma show that the predicted circRNA-disease associations by MDGF-MCEC are confirmed by biological experiments in the literature.

Materials and methods

Overview of MDGF-MCEC

The framework of the proposed MDGF-MCEC is shown in Figure 1. MDGF-MCEC mainly consists of four modules from Parts A to D. Part A is the module of data sources,

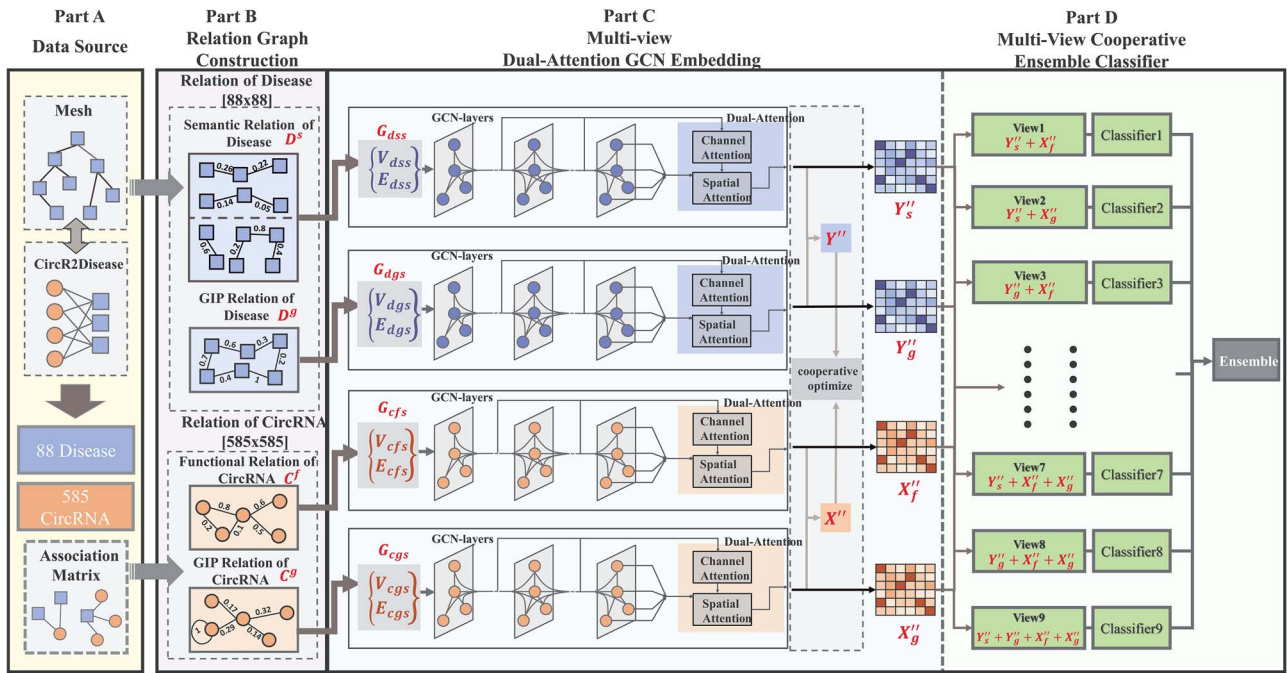


Figure 1. The framework of the proposed MDGF-MCEC. (A) The related databases are shown. (B) circRNA relation graphs and disease relation graphs are constructed by using different measurements in different views. (C) The multi-view relation graphs are fed into a multi-view dual attention GCN for feature learning of graph nodes. (D) The extracted circRNA features and disease features are used to generate different kinds of circRNA-disease feature combinations as different views, and a multi-view cooperative ensemble learning classifier is further constructed for making prediction.

consisting of disease database, circRNA database and circRNA-disease association database. Part B is a relation graph construction module, which generates both the disease relation graphs and the circRNA relation graphs by using different similarity measurements. Part C is a graph embedding learning module, which is used to learn embedded features of nodes in both types of relation graphs based on the multi-view dual-attention GCN. Part D is a multi-view cooperative ensemble classification module, which is a classifier using both multi-view collaboration and ensemble learning.

Benchmark datasets

The benchmark datasets used in this study are obtained from the two types of databases.

(i) CircRNA-disease association databases

Three circRNA-disease association databases are used in this study. The first is CircR2Disease [21] database (<http://bioinfo.snnu.edu.cn/CircR2Disease/>). This database contains 739 circRNA-disease associations that have been experimentally verified, involving 661 circRNAs and 100 diseases. After removing redundant data, 650 pairs of circRNA-disease associations are obtained, involving 585 circRNAs and 88 diseases.

In addition, we use the Circ2Disease [22] database (<http://bioinfo.snnu.edu.cn/CircR2Disease/>) and CircRNADisease [23] database (<http://cgga.org.cn:9091/circRNADisease/>) to verify the reliability of the proposed model. The Circ2Disease database contains 270 disease-circRNA associations between 249 circRNAs and 60 diseases, and the CircRNADisease database contains 351

disease-circRNA associations between 330 circRNAs and 48 diseases.

(ii) Disease database

The Mesh disease database (<https://www.nlm.nih.gov/mesh/meshhome.html>) is integrated with the circRNA-disease association database for model training. The database represents diseases as directed acyclic graphs based on medical and biological associations. It contains 4818 diseases and 11 648 associations between diseases.

In this study, 650 pairs of known circRNA-diseases associations from CircR2Disease database are used as the positive set. In order to construct a negative set that is balanced with the positive set, the strategy of Wang et al. [25] is adopted, where 650 pairs of circRNA-disease not in the positive set between circRNAs and diseases are randomly selected. Although this strategy may mislabel some unconfirmed circRNA-disease association pairs (i.e. potential positive examples) as negative samples, the error rate $[650 \div (585 \times 88) \approx 1.26\%]$ is very low. Finally, the dataset used in this study consists of 1300 samples with an equal number of positive and negative samples.

To verify the reliability of our prediction model, case studies are conducted using the confirmed circRNA-disease associations from the literature, which is a gold standard to assess the consistency of the predicted results.

Multi-view relation graph construction

Disease relation graphs and circRNA relation graphs from different views are constructed to mine the information of circRNAs and diseases. Based on the constructed

relation graphs, features of circRNAs and diseases are then exploited to identify the associations between them. To build the relation graphs, the adjacency matrix of circRNAs and diseases is first constructed based on the CircR2Disease database. The size of the adjacency matrix is 585×88 , corresponding to 585 circRNAs and 88 diseases. If an association exists between a circRNA and a disease, the value of the corresponding element in the adjacency matrix is 1; otherwise 0. Then, two disease relation graphs and two CircRNA relation graphs are constructed as follows:

- (i) Disease relation graph based on the semantic similarity

$$G_{dss} = (V_{dss}, E_{dss}) \quad (1)$$

- (ii) Disease relation graph based on the disease GIP kernel similarity

$$G_{dgs} = (V_{dgs}, E_{dgs}) \quad (2)$$

- (iii) CircRNA relation graph based on the functional similarity

$$G_{cfs} = (V_{cfs}, E_{cfs}) \quad (3)$$

- (iv) CircRNA relation graph based on the GIP kernel similarity

$$G_{cgs} = (V_{cgs}, E_{cgs}) \quad (4)$$

where V_* is a set of nodes and E_* is a set of edges. The details for the construction of the above relation graphs are described in Parts A–D of the Supplementary Material.

Finally, as shown in Figure 2, multiple relation graphs are constructed for 88 diseases and 585 circRNAs in the CircR2Disease database. Graph embedding techniques are then used to extract discriminative features for the diseases and circRNAs.

Embedding learning based on the multi-view dual attention GCN

Model construction

To effectively identify circRNA-disease associations, highly discriminative circRNA features and disease features need be learned from the circRNA relation graphs and disease relation graphs, respectively. For this purpose, we propose a multi-view dual attention GCN for feature learning as shown in Figure 3. The core components of the network are four parallel dual attention GCN modules associated with different circRNA relation graphs or disease relation graphs. These modules are trained in a cooperative manner.

Dual attention GCN module

Figure 4 gives the details of the four parallel dual attention GCN modules; here, we take the GIP kernel relation graph of circRNAs as an example. The corresponding modules associated with other relation graphs can be constructed in the same way. There are four modules in Figure 4, they are GCN component, the dual-attention mechanism, the channel attention mechanism and the spatial attention mechanism.

GCN component

The dual-attention GCN module contains a traditional GCN component at the bottom of Figure 4. In this part, the traditional GCN component directly performs on the circRNA relation graph G_{cgs} and uses the structural information to capture the dependency between the nodes by aggregating the neighbors' information and generates the embeddings of the nodes in the relation graph [34–36].

For the multi-view dual attention GCN in this study, cooperative learning is conducted on the nodes of two disease relation graphs and two circRNA relation graphs to generate different types of embeddings.

Let A_c be the adjacency matrix of G_{cgs} , V_c be the node set of G_{cgs} , N_c be the number of nodes in G_{cgs} and D_c be the dimension of the embedding. Thus, the features of the nodes on the entire graph G_{cgs} can be expressed as a matrix $X_g \in R^{N_c \times D_c}$, where $x[i] \in R^{D_c}$ ($i \in \{1, 2, \dots, N_c\}$) is the embedding of the i th node. For the GCN module in Figure 4A, the initial embedding matrix $X_g^{(0)}$ is randomly initialized in the layer 0 and the embeddings in the t th layer ($t > 0$) are learned by the following forward propagation formula:

$$X_g^{(t)} = \hat{S}_c^{-1/2} \hat{A}_c \hat{S}_c^{-1/2} X_g^{(t-1)} \Theta^{(t-1)}, \quad (5)$$

where $X_g^{(t)} \in R^{N_c \times D_c}$, $\hat{A}_c = A_c + I_c$, $[\hat{S}_c]_{ii} = \sum_j [\hat{A}_c]_{ij}$, I_c is an identity matrix and \hat{S}_c is a symmetric matrix, and $\Theta^{(t-1)} \in R^{D_c \times D_c}$ is a filter, which is a learnable matrix. To simplify the formula, we suppose that $\hat{L}_c = \hat{S}_c^{-1/2} \hat{A}_c \hat{S}_c^{-1/2}$. Thus, the forward propagation formula of graph embedding in GCN is further simplified as follows:

$$X_g^{(t)} = \hat{L}_c X_g^{(t-1)} \Theta^{(t-1)}. \quad (6)$$

To enhance the power of the model, activation functions are introduced in each layer ($t > 0$). To obtain more information after convolution, not only the embedding obtained from the first layer but also that of the second layer are retained in the module shown in Figure 4. Then, the two-layer embeddings of the nodes in the GIP kernel

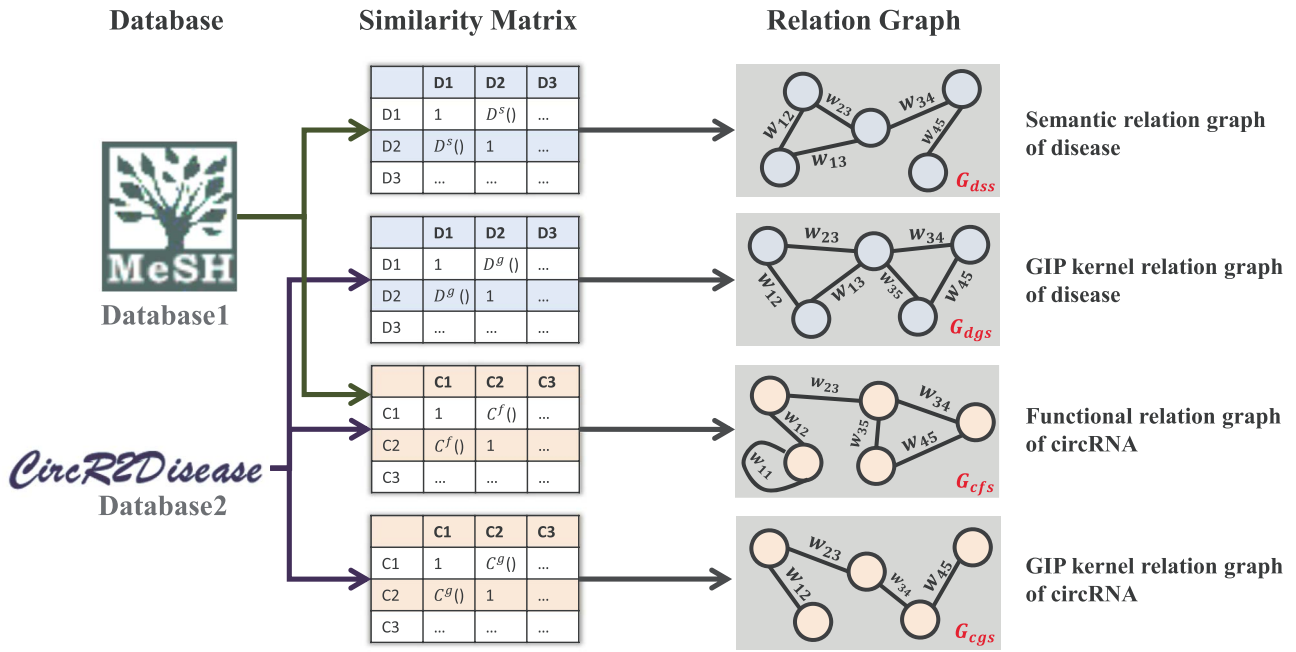


Figure 2. Relation graphs constructed for diseases circRNAs, both diseases and circRNAs have two graphs (semantic relation graph of the disease, GIP kernel relation graph of the disease, functional relation graph of circRNAs, GIP kernel relation graph of circRNAs).

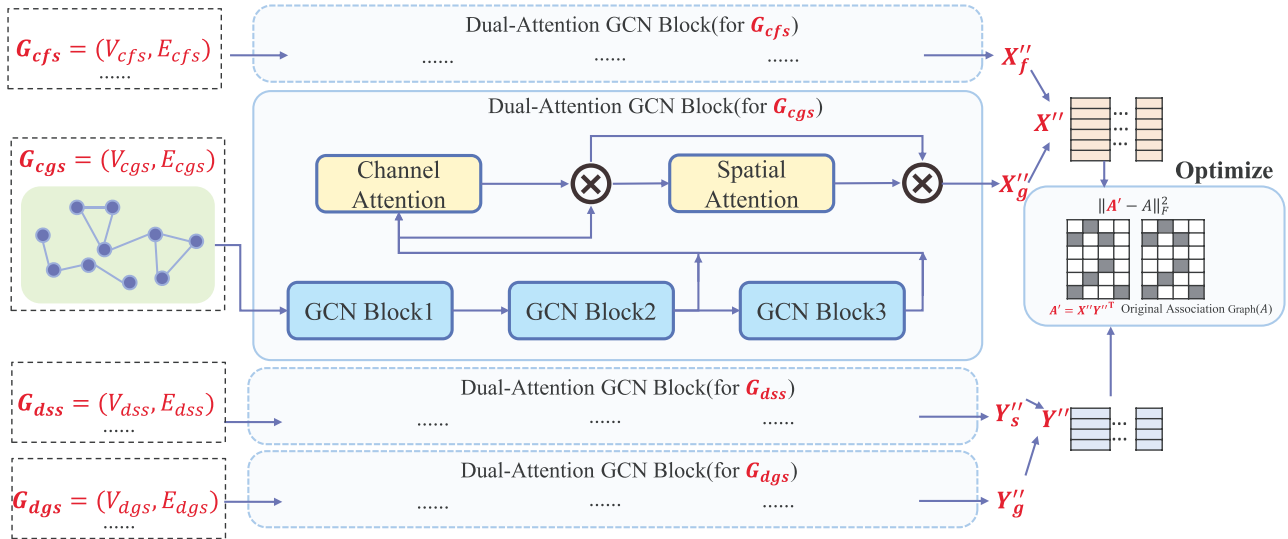


Figure 3. Multi-view dual attention GCN on four relation graphs.

similarity relation graph of circRNA is obtained as

$$X_g^{(1)} = \text{active}(\hat{L}_c X_g^{(0)} \Theta^{(0)}) \quad (7)$$

$$X_g^{(2)} = \text{active}(\hat{L}_c \text{active}(\hat{L}_c X_g^{(0)} \Theta^{(0)}) \Theta^{(1)}). \quad (8)$$

where **active** denotes an activation function.

Similarly, the two-layer embeddings for the other three relation graphs can be obtained by the GCN components in the dual attention GCN module shown in Figure 3. Finally, the two-layer embeddings of the two circRNA relation graphs and the two disease relation graphs are obtained as follows: $X_f = \{X_f^{(1)}; X_f^{(2)}\}$, $X_g = \{X_g^{(1)}; X_g^{(2)}\}$, $Y_s = \{Y_s^{(1)}; Y_s^{(2)}\}$, $Y_g = \{Y_g^{(1)}; Y_g^{(2)}\}$, where $X_f, X_g \in \mathbb{R}^{2 \times N_c \times D_c}$ and $Y_s, Y_g \in \mathbb{R}^{2 \times N_d \times D_d}$. These embeddings can be further used

in the subsequent dual attention learning blocks. The experimental setting of the number of graph convolution layers in our model (default 2) is analyzed in Figure S2 in Part H of the Supplementary Material.

Dual attention mechanism

In the middle of Figure 4B, the multi-view dual attention GCN [37] is performed on each relation graph; here, we take the circRNA relation graph G_{cgs} as an example.

First, the two-layer embeddings $X_g = \{X_g^{(1)}; X_g^{(2)}\}$, $X_g \in \mathbb{R}^{C_c \times N_c \times D_c}$ are taken as the input of the channel attention block, where C_c is the number of channels or the number of embedding layers (the default number is two), D_c is the dimension of embeddings obtained by GCN blocks and N_c is the number of nodes in a circRNA relation graph. The

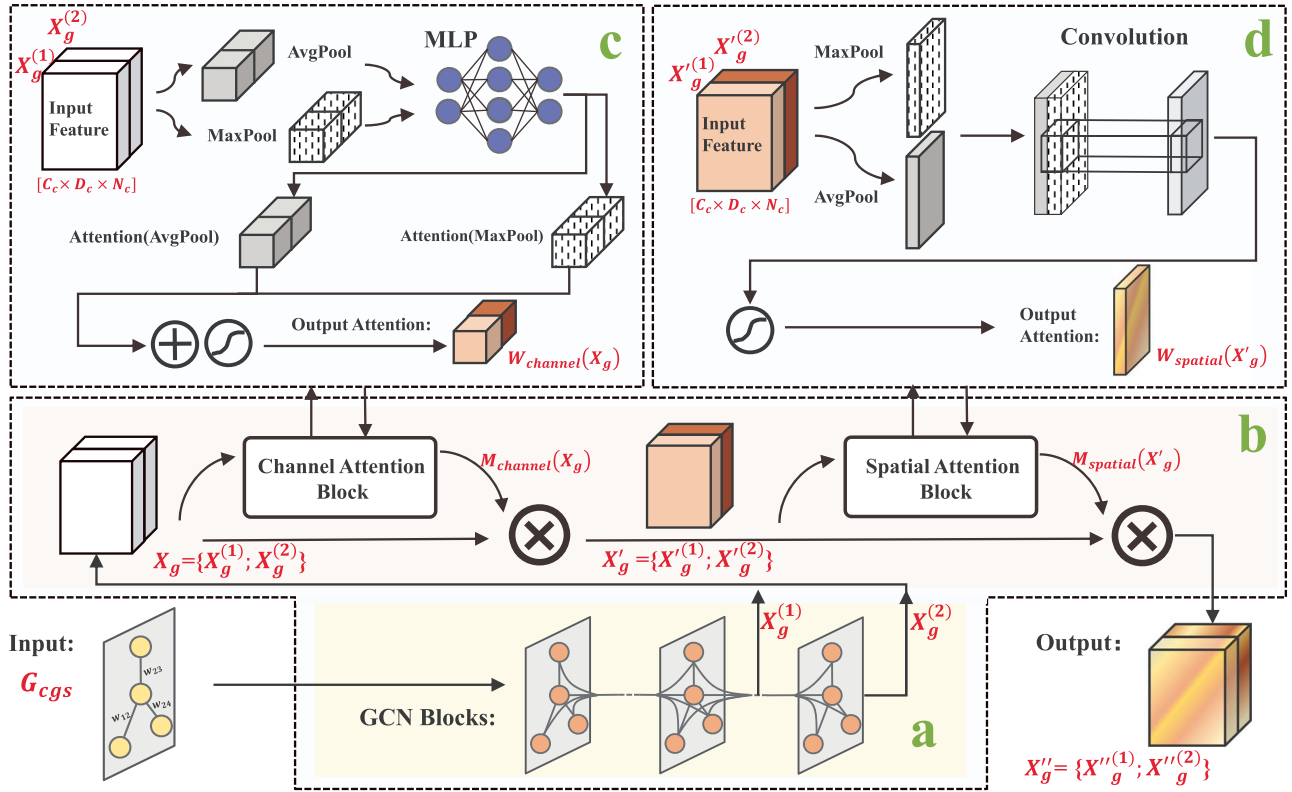


Figure 4. The structure of dual attention GCN. (A) The GCN component; (B) the dual-attention mechanism; (C) the channel attention mechanism; (D) the spatial attention mechanism.

weights of each channel are generated from the block of channel attention, which can be expressed as a vector $W_{channel}(X_g) \in R^C$. Based on $W_{channel}(X_g)$ and the size of X_g , the channel attention map $M_{channel}(X_g) \in R^{C \times N_c \times D_c}$ can be obtained, and the values of the elements in the map $M_{channel}(X_g)(i, :, :)$ are set as $W_{channel}(X_g)(i)$, $i \in \{1, \dots, C_c\}$ accordingly. Then, the embeddings obtained from the channel attention block are expressed as

$$X'_g = M_{channel}(X_g) \otimes X_g, \quad (9)$$

where \otimes denotes the dot product.

Furthermore, the embedding X'_g is used as the input of the spatial attention block, where a spatial attention weight matrix $W_{spatial}(X'_g) \in R^{N_c \times D_c}$ is generated. Based on $W_{spatial}(X'_g)$ and the size of X'_g , the spatial attention map $M_{spatial}(X'_g) \in R^{C_c \times N_c \times D_c}$ is obtained, where $M_{spatial}(X'_g)(i, :, :)$ = $W_{spatial}(X'_g)$ and $i \in \{1, \dots, C_c\}$. Then, the embeddings obtained from the spatial attention block are expressed as

$$X''_g = M_{spatial}(X'_g) \otimes X'_g, \quad (10)$$

where X''_g is the final embedding of circRNAs obtained from the circRNA relation graph based on the GIP kernel similarity by the dual attention GCN.

Channel attention block

We take the circRNA relation graph G_{cgs} in Figure 4C as an example, the mechanism of the channel attention block is explained as follows. First, the multi-layer embedding X_g is fed into the channel attention block, and each embedding layer can be regarded as a representation of a circRNA. In order to extract the influence of different layers on the whole embeddings, a multi-layer perceptron is used to compress and restore the input embeddings. To aggregate the attention information [37], max-pooling and average-pooling are added simultaneously to obtain a more accurate channel attention. The process of the channel attention block is expressed as

$$W_{channel}(X_g) = \sigma(\text{MLP}(\text{AvgPool}(X_g)) + \text{MLP}(\text{MaxPool}(X_g))) \\ = \sigma(\text{MLP}(X_g^{\text{avg}}) + \text{MLP}(X_g^{\text{max}})) \quad (11)$$

where $X_g^{\text{avg}} \in R^C$ and $X_g^{\text{max}} \in R^C$ are the feature maps obtained from the average pooling and max pooling, respectively; σ is the sigmoid activation function; MLP is the multi-layer perceptron with shared parameters. Then, the channel attention vector $W_{channel}(X_g) \in R^C$ is generated after sending X_g^{avg} and X_g^{max} to the shared MLP layer. Based on $W_{channel}(X_g)$ and the size of X_g , the channel attention map $M_{channel}(X_g) \in R^{C_c \times N_c \times D_c}$ and the embedding X'_g can be obtained using Equation (9).

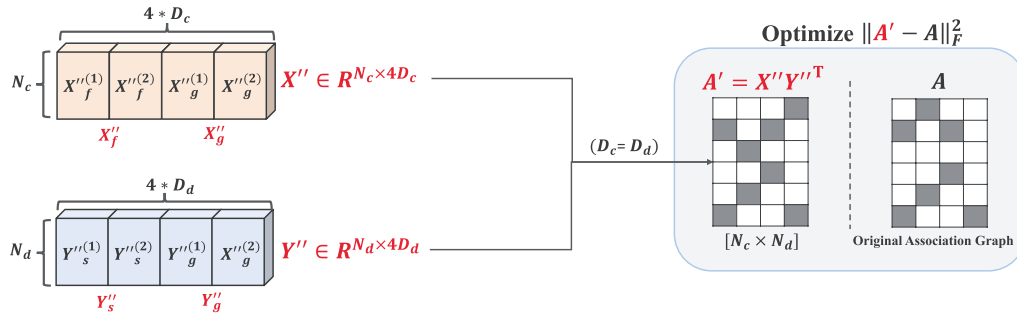


Figure 5. Parameter optimization of multi-view dual attention GCN with collaborative learning.

Spatial attention block

The mechanism of the spatial attention block is shown in Figure 4D, where the embeddings' spatial relations are used to construct the spatial attention map. The spatial attention block attempts to improve the channel attention by paying more attention to the information of a certain part in a channel. To calculate the spatial attention, average-pooling and max-pooling are applied along the channel axis for inputs and the results are concatenated to obtain feature descriptors. After concatenation, we apply convolution operation to generate the spatial attention matrix $W_{\text{spatial}}(X'_g)$, which is used to represent the regions of enhanced or weakened attentions. The process of the spatial attention block is expressed as

$$W_{\text{spatial}}(X'_g) = \sigma \left(f_{\text{conv}} \left(\left[\text{AvgPool}(X'_g); \text{MaxPool}(X'_g) \right] \right) \right), \\ = \sigma \left(f_{\text{conv}} \left([X'_g{}^{\text{avg}}; X'_g{}^{\text{max}}] \right) \right), \quad (12)$$

where $X'_g{}^{\text{avg}} \in \mathbb{R}^{N_c \times D_c}$ and $X'_g{}^{\text{max}} \in \mathbb{R}^{N_c \times D_c}$ represent the feature maps obtained by the average-pooling and max-pooling along the channel axis, respectively. Then, the feature maps are combined and convolved by the standard convolution layer f_{conv} to generate the two-dimension spatial attention matrix $W_{\text{spatial}}(X'_g)$. Based on the $W_{\text{spatial}}(X'_g)$ and the size of X'_g , the spatial attention map $M_{\text{spatial}}(X'_g) \in \mathbb{R}^{C_c \times N_c \times D_c}$ and the embedding X''_g can be obtained using Equation (10).

Parameter optimization of multi-view dual attention GCN with collaborative learning

To train the proposed multi-view dual attention GCN, a multi-view cooperative objective function is designed to optimize the model parameters. The collaboration learning mechanism is shown in Figure 5. The details for parameter optimization are described in Part E of the Supplementary Material.

Multi-view cooperative ensemble classifier

After the multi-view dual attention GCN, the embeddings of two circRNA views, the functional similarity-based embedding X''_f and the GIP kernel similarity-based embedding X''_g , are obtained. In addition, the embeddings of two disease views, the semantic similarity-based graph embedding Y''_s and the GIP kernel similarity-based graph embedding Y''_g , are also obtained. How to effectively

use the above multi-view circRNA and disease information to construct an efficient model for predicting the associations between circRNAs and diseases is a challenging task. Traditional approaches usually concatenate the multi-view features of both circRNAs and diseases and then train the classifier. However, the utilization of multi-view information in traditional approaches is inadequate. To address this problem, we propose a multi-view cooperative ensemble classifier.

First, various combinations of circRNA and disease embeddings, i.e. different feature sets from different views, are obtained. Then, XGBoost (Extreme Gradient Boosting) [38] classifiers are trained using different circRNA-disease feature combinations and their labels. For different combinations of circRNA features and disease features, different XGBoost classifiers are trained. Based on these classifiers, the secondary ensemble is created by voting mechanism, and finally, a multi-view cooperative ensemble classifier is implemented. Due to highly correlated and complementary relationship between the multiple views obtained from the diverse combinations, the circRNA features and disease features obtained from the multi-view dual attention GCN can be fully utilized.

Specifically, to fully use the learned multi-view circRNA and disease information, we construct nine different views for different combinations of circRNA and disease features: $\{X''_f, Y''_s\}, \{X''_f, Y''_g\}, \{X''_g, Y''_s\}, \dots, \{X''_f, X''_g, Y''_g\}, \{X''_f, X''_g, Y''_s, Y''_g\}$. For each combined view, we train an ensemble learning classifier with XGBoost and then create a secondary ensemble for the nine ensemble classifiers to obtain the multi-view cooperative ensemble classifier. The process of the proposed multi-view cooperative ensemble learning classifier is shown in Figure 6.

Results

Evaluation metrics

To evaluate the performance of the proposed method, six metrics, i.e. accuracy (Acc), sensitivity (Sen), precision (Pre), F1-score (F1), Matthews correlation coefficient (MCC) and AUC [39], are used as the evaluation metrics. The details of each metric are described in Part F of the Supplementary Material.

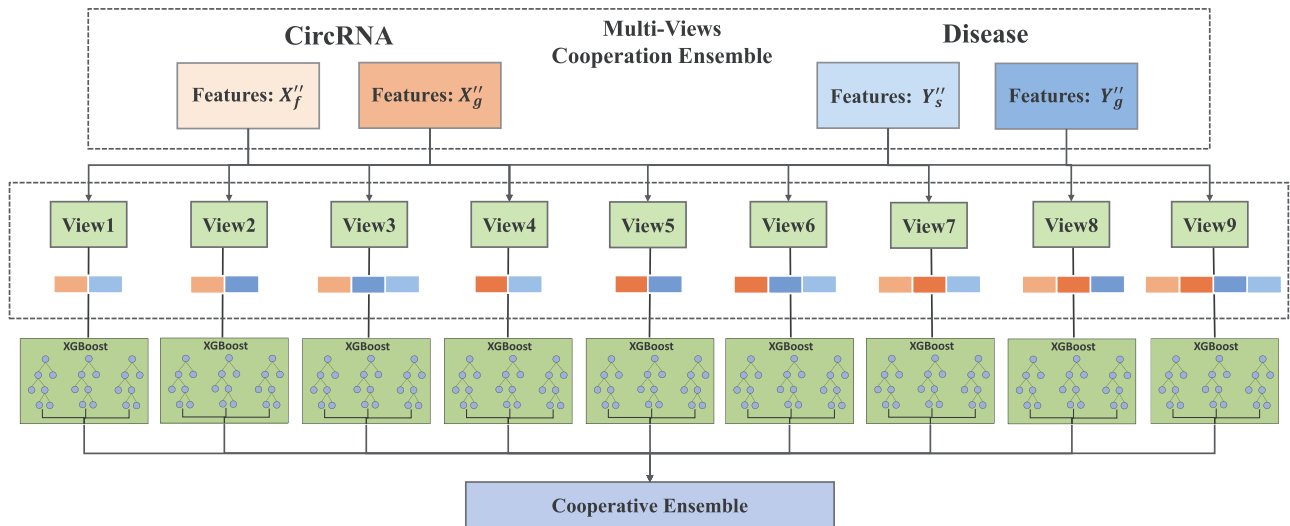


Figure 6. The construction of multi-view cooperative ensemble classifier.

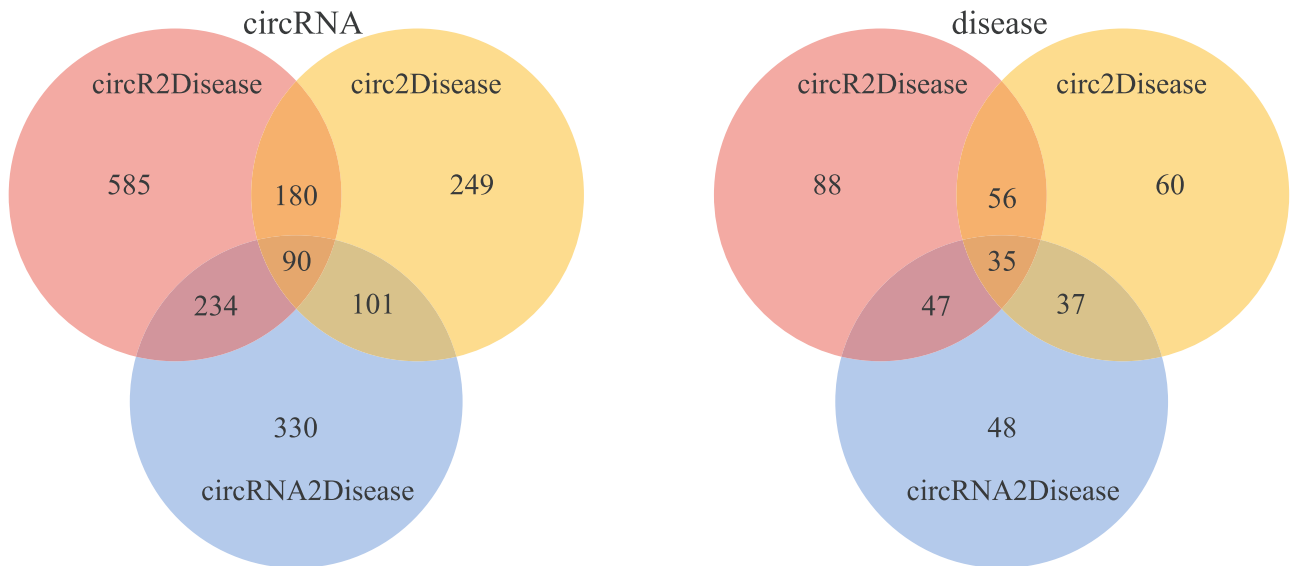


Figure 7. Distribution of circRNAs and diseases in three databases.

Experiment data sources

Three different databases are used to comprehensively evaluate the performance of the proposed MDGF-MCEC for predicting circRNA-disease associations; they are circR2Disease [21], circ2Disease [22] and circRNADisease [23]. Figure 7 shows the overlap of circRNAs and diseases in the three databases. There are 249 circRNAs and 60 diseases in the circ2Disease database, 585 circRNAs and 88 diseases in the circR2Disease database and 330 circRNAs and 48 diseases in the circRNADisease database. In particular, 90 circRNAs and 35 diseases are shared by the three databases. Since the known associations in the circR2Disease database are more sufficient than those in circRNADisease and circ2Disease, the experiments in this study are conducted mainly on the circR2Disease database. Circ2Disease and circRNADisease databases are further used to verify the prediction ability of MDGF-MCEC.

Performance on circR2Disease

We first evaluate the prediction performance of the proposed MDGF-MCEC on the dataset from the circR2Disease database using 5-fold cross-validation. The whole dataset is evenly divided into k subsets, where each subset takes turn to be the test set, while the other subsets are used as the training set. The ROC curves of each fold on the validation set from the circR2Disease database are shown in Figure 8. The corresponding results of other metrics are shown in Table 1. It can be seen from the results that the proposed method MDGF-MCEC has achieved promising results.

Comparison with state-of-the-art methods

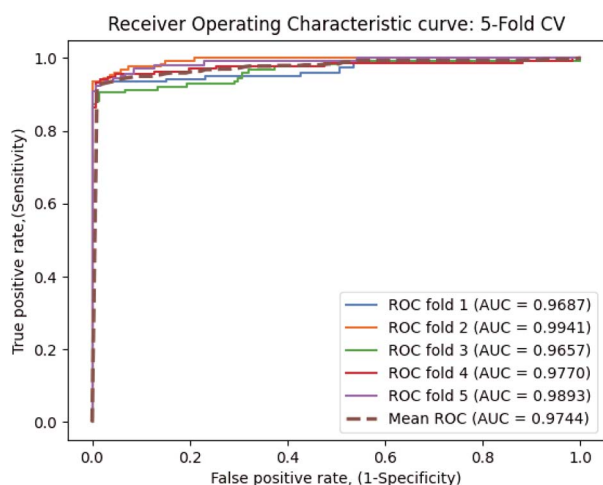
To demonstrate the advantages of the proposed MDGF-MCEC, we compare it with six state-of-the-art methods. The six methods are RGCNDA [33], Wang's method [25], NCPCDA [40], iCircDA-MF [31], GCNDA [41] and

Table 1. Results of 5-fold cross-validation obtained by the proposed MDGF-MCEC on circR2Disease

Fold	Acc (%)	Sen (%)	F1 (%)	MCC (%)	AUC (%)
1	95.76	91.86	95.35	91.70	98.77
2	93.84	88.72	93.65	88.22	97.11
3	95.38	91.72	95.31	91.04	98.47
4	93.84	87.09	93.10	88.27	95.62
5	94.23	89.78	94.25	88.93	97.25
Average	94.61 \pm 0.65	89.83 \pm 3.29	94.33 \pm 0.79	89.63 \pm 2.12	97.44 \pm 1.26

Table 2. Performance comparison of MDGF-MCEC with six state-of-the-art methods in terms of AUC

Methods	MDGF-MCEC (our method)	RGCNCDA	NCPCDA	iCircDA-MF	GCNCDA	PWCDA	Wang's method
AUC	0.9740	0.9507	0.9201	0.9178	0.9090	0.8900	0.8667

**Figure 8.** The ROC curves of 5-fold cross-validation obtained by the proposed MDGF-MCEC on the circR2Disease database.

PWCDA [42]. The results of 5-fold cross-validation on the circR2Disease database are compared. Since the evaluation metrics adopted by different methods are different, the metric AUC that is used in common is presented here for the comparison in Table 2. The results show that the proposed method MDGF-MCEC outperforms other methods. It should be noted that although the methods under the comparison are all evaluated on the circRNA-disease associations from the circR2Disease database, the data that they used are not completely the same. For example, the methods RGCNCDA, NCPCDA and iCircDA-MF only use human data, PWCDA uses both human and mouse data, and Wang's method and GCNCDA use not only the human and mouse data but also data of other species.

Performance on circ2Disease and circRNADisease

Further experiments are conducted to assess the performance of the proposed model MDGF-MCEC on the Circ2Disease and CircRNADisease databases. The

results of 5-fold cross-validation in terms of five metrics are shown in Tables 3 and 4. The ROC curves and the corresponding AUC scores are shown in Figure 9. The trends of F1-score with different probability thresholds in three databases are shown in Figure S1 in Part G of the Supplementary Material. The results show that the proposed MDGF-MCEC also achieves good performance on the Circ2Disease and CircRNADisease, suggesting that MDGF-MCEC could be generally applied to different databases with promising performance.

Effectiveness of dual attention mechanism

To demonstrate the effectiveness of the dual attention mechanism in GCN in MDGF-MCEC, we compare it with three other variants of the method MDGF-MCEC, where (i) only the channel attention module, (ii) only the spatial attention module and (iii) no attention module. The performance of the three variants and the proposed method MDGF-MCEC on the circR2Disease database are shown in Figure 10. The results show that the performance of the variant method without any attention blocks is the worst, and the introduction of the channel attention or spatial attention can improve the performance. The three variants are inferior to MDGF-MCEC with both channel attention and spatial attention, demonstrating the added value of the attention mechanism. In addition, the results also show that the effectiveness of the spatial attention is more significant than that of channel attention.

The performance of multi-view learning

To make full use of the circRNA and disease features of different views learned by the multi-view dual-attention GCN, we constructed a multi-view cooperative ensemble learning classifier based on nine different combined views for circRNA-disease associations. Here, the combined views of disease features and circRNA features are used for model training, and the results based on the different combined views are then evaluated.

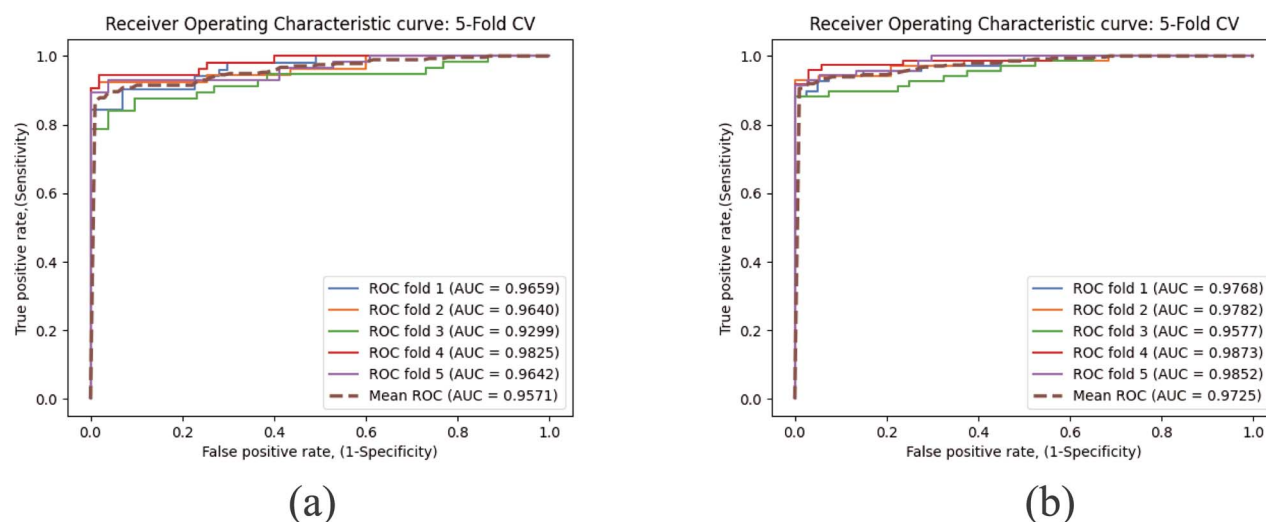
Table 5 lists the performance of XGBoost classifiers based on different combined views and the proposed

Table 3. Results of 5-fold cross-validation for the proposed MDGF-MCEC on Circ2Disease

Fold	Acc (%)	Sen (%)	F1 (%)	MCC (%)	AUC (%)
1	91.66	84.21	91.42	84.60	96.59
2	95.37	92.45	95.14	90.86	96.40
3	88.89	78.57	88.00	79.89	92.99
4	95.37	90.57	95.05	91.11	98.25
5	92.59	84.31	91.49	85.99	96.42
Average	92.77 \pm 5.86	86.02 \pm 24.75	92.22 \pm 7.10	86.49 \pm 17.56	95.71 \pm 2.94

Table 4. Results of 5-fold cross-validation for the proposed MDGF-MCEC on CircRNADisease

Fold	Acc (%)	Sen (%)	F1 (%)	MCC (%)	AUC (%)
1	92.59	94.36	94.36	83.55	93.77
2	95.37	95.94	96.59	89.37	97.25
3	90.74	88.23	92.31	81.28	92.41
4	89.81	94.28	92.30	77.41	98.19
5	91.66	94.12	93.43	82.06	96.93
Average	92.77 \pm 3.69	86.02 \pm 7.07	92.22 \pm 2.54	86.49 \pm 15.12	97.25 \pm 4.93

**Figure 9.** The ROC curves and AUC of the MDGF-MCEC on (A) Circ2Disease database and (B) CircRNADisease database.

multi-view cooperative ensemble classifier. Of the nine views, the performance of the classifier constructed using the ninth view is generally superior to those using the other combined views. However, the prediction ability of the models using one combined view is weaker than that of the cooperative ensemble learning classifier based on multiple combined views. It also indicates that the combined views are complementary to each other. As a result of the cooperation of multiple combined views, features of circRNAs and diseases learned by the dual-attention GCN are fully exploited for making classification.

Case study

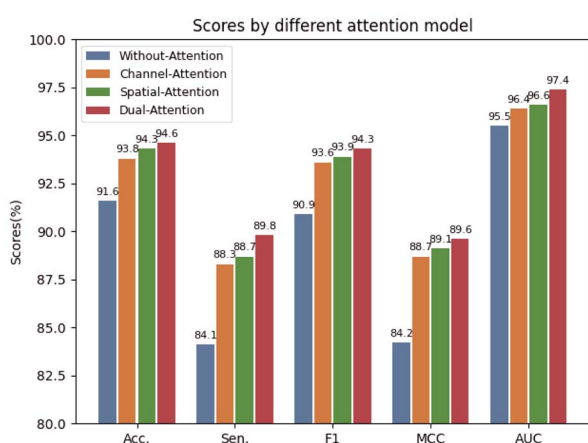
A case study is conducted to verify the effectiveness of the proposed method MDGF-MCEC. First, a MDGF-MCEC model is trained on the dataset from the circR2Disease database. Then, MDGF-MCEC is used to predict the

associations of 50 830 circRNA-disease pairs with unknown relationships in circRNA-disease association matrix. In this case study, we use the trained MDGF-MCEC to investigate the circRNAs that are related to hepatocellular carcinoma (HCC) and gastric cancer, respectively. Top 10 circRNAs ranked by the scores from MDGF-MCEC are listed in Tables 6 and 7.

In addition, we verify the predicted results against the circRNA-disease associations published in the literature. For example, the following associations have been reported in the literature: circPVT1 regulates cell proliferation, apoptosis and glycolysis in HCC via miR-377/TRIM23 [43]; circZNF609 enhances HCC cell proliferation, metastasis and stemness by activating the hedgehog pathway through the regulation of miR-15a-5p/15b-5p and GLI2 expressions [44], and Cir-ITCH [45] inhibits gastric cancer migration, invasion and proliferation by regulating the Wnt/ β -catenin pathway. The literature

Table 5. The performance of the proposed method MDGF-MCEC with different combined views

View of CircRNA-disease combination	Acc (%)	Sen (%)	Pre (%)	F1 (%)	MCC (%)	AUC (%)
View1: $X_f'' + Y_s''$	93.76	87.95	91.67	93.48	88.41	96.96
View2: $X_f'' + Y_g''$	83.53	83.96	83.33	83.47	67.27	90.19
View3: $X_f'' + Y_s'' + Y_g''$	83.69	83.85	83.49	83.61	67.34	90.65
View4: $X_g'' + Y_s''$	85.23	87.53	83.65	85.52	70.55	92.62
View5: $X_g'' + Y_g''$	82.07	84.43	80.49	82.38	64.19	90.13
View6: $X_g'' + Y_s'' + Y_g''$	83.30	83.84	83.08	83.35	66.72	91.58
View7: $X_f'' + X_g'' + Y_s''$	89.19	86.25	87.57	86.89	73.95	93.61
View8: $X_f'' + X_g'' + Y_g''$	84.76	85.56	84.32	84.92	79.51	91.11
View9: $X_f'' + X_g'' + Y_s'' + Y_g''$	94.00	86.03	85.95	85.96	72.00	93.13
Multi-view cooperation ensemble (our method)	94.61	89.83	99.33	94.33	89.63	97.40

**Figure 10.** Effectiveness of attentions mechanism (dual-attention mechanism, only the channel attention module, only the spatial attention module, no attention module).**Table 6.** The top 10 predicted circRNA-disease associations related to HCC

Rank	CircRNA	Supporting evidence (PMID)
1	circPVT1	33 364 841
2	circZNF609	32 398 664
3	hsa_circ_103110	Not available
4	hsa_circ_0007534	Not available
5	hsa_circ_0000520	27 258 521
6	hsa_circRNA_103809	31 147 216
7	hsa_circ_0072088	34 568 335
8	circHIAT1	31 108 351
9	hsa_circ_0006404	Not available
10	hsa_circ_0050898	Not available

that supports the prediction results is given Tables 6 and 7.

Furthermore, the top 20 circRNA-disease association pairs of all the predicted 50 830 circRNA-disease pairs, in descending order of the predicted scores by MDGF-MCEC, are shown in Table 8. It should be pointed out that although some predicted associations of circRNA-disease pairs are currently not supported by the literature, they may suggest potential associations that biological experiments can be conducted for verification in future.

Table 7. The top 10 predicted circRNA-disease associations related to gastric cancer

Rank	CircRNA	Supporting evidence (PMID)
1	Cir-ITCH	33 060 778
2	circZNF609	31 632 070
3	hsa_circ_0007534	32 419 229
4	hsa_circ_0085154	Not available
5	hsa_circ_0000615	34 049 561
6	hsa_circ_0067531	Not available
7	circZKSCAN1	Not available
8	circRNA_000839	Not available
9	hsa_circ_0005986	33 620 018
10	circ-Foxo3/hsa_circ_0006404	33 833 780

Discussion and conclusion

In this study, a novel approach MDGF-MCEC based on a multi-view dual-attention GCN and cooperative ensemble learning classifier is proposed to fully exploit various information of circRNAs and diseases for inferring the potential circRNA-disease associations. In order to analyze the effectiveness of the proposed method MDGF-MCEC, experiments are conducted on data from the circR2Disease, circRNADisease and circ2Disease databases using 5-fold cross-validation. In addition, the method is compared with six state-of-the-art methods. Experiments are also performed to verify the effectiveness and reliability of MDGF-MCEC. First, we investigate the added value of the attention mechanism. The results show that the model with a dual attention mechanism is superior to the models with a single attention module (channel or spatial) or without any attention module. Then, in order to verify the effectiveness of the multi-view cooperative ensemble classification in MDGF-MCEC, the circRNA and disease embeddings learned by the multi-view dual attention GCN module are combined in different ways to investigate the performance of the multi-view cooperative ensemble learning classifier under different feature combinations. The results show that MDGF-MCEC makes full use of the multi-view circRNA features and disease features. Finally, we evaluate the predicted results of MDGF-MCEC against the circRNA-disease associations reported in the literature.

Table 8. The top twenty prediction results of circRNA-disease associations by MDGF-MCEC

Rank	CircRNA	Disease	Supporting evidence (PMID)
1	hsa_circRNA_100782	Glioma	Not available
2	circHIPK3	Major depressive disorder	Not available
3	hsa_circ_0001649	Bladder cancer	Not available
4	Cir-ITCH	Gastric cancer	33 060 778
5	hsa_circ_0000284	Repeated implantation failure	Not available
6	circHIPK3	Lung cancer	31 232 177
7	circHIPK3	Myocardial senescence	31 799 682
8	circHIPK3	Glioblastoma multiforme	33 194 066
9	circHIPK3	Cervical cancer	33 177 072
10	circHIPK3	Radiation-induced liver fibrosis	Not available
11	circHIPK3	Prostate cancer	30 863 152
12	hsa_circ_0000518	Osteosarcoma	Not available
13	hsa_circRNA_100782	Lupus nephritis	Not available
14	circHIPK3	Pancreatic cancer	32 104 074
15	hsa_circRNA_100782	Multiple system atrophy	Not available
16	circHIPK3	Immunosenescence	34 326 381
17	hsa_circ_0000284	Coronary artery disease	31 619 168
18	hsa_circRNA_100782	Rheumatoid arthritis	Not available
19	hsa_circ_0000284	Primary hepatic carcinoma	Not available
20	hsa_circ_0001649	Osteosarcoma	32 954 926

The results show that more than half of the predicted circRNA-disease associations are supported by the literature.

Although the proposed method MDGF-MCEC achieves promising results for predicting circRNA-disease associations, there are still some challenges that require further investigation. At present, circRNA-disease association prediction is a relatively new research field, where the annotated data are limited to enhance and assess the generalization ability of the proposed method. In addition, in MDGF-MCEC, circRNA-disease pairs with unknown associations are regarded as the negative samples. This is a feasible strategy but it may lead to mislabeling of circRNA-disease pairs, resulting in the deviations of the training process. A potential solution is to introduce positive-unlabeled learning strategy, which builds a classifier with only positive and unlabeled examples, but no negative examples [46–48]. In fact, there have been many positive-unlabeled learning strategies that were developed for positive-unlabeled learning tasks, such as ‘two-step approach’ [49] and ‘weighted strategy’ [50]. In future, these methods will be investigated for model training of circRNA-disease association predictors.

Key Points

- The proposed approach generates four different relation graphs (circRNA functional similarity relation graph, circRNA GIP similarity relation graph, disease semantic similarity relation graph, disease GIP similarity relation graph) as four initial multi-views.
- We use multi-view dual-attention GCN as feature learning method of graph embedding process. And then, the

better deep representations can be obtained for subsequent classification.

- In this work, the proposed cooperative ensemble learning classifier can make a full and accurate classification, in which different multi-view circRNA-disease feature combinations are used generate various single-view ensemble classifiers and then these classifiers are integrated as a multi-view cooperative ensemble classifier with secondary ensemble.

Supplementary data

Supplementary data are available online at [https://academic.oup.com/bib](https://academic.oup.com/bib/article/23/5/bbac289/6652197).

Funding

National Key Research and Development Program of China (2021YFE010178); National Natural Science Foundation of China (62176105, 62073219, 61725302, 61903248); The Six Talent Peaks Project in Jiangsu Province (XYDXX-056); The Hong Kong Innovation and Technology Fund (MRF/015/18); RGC GRF project (PolyU 12006/19E); Science and Technology Commission of Shanghai Municipality (20S11902100); Shanghai Municipal Science and Technology Major Project (2018SHZDZX01).

References

1. Shang Q, Yang Z, Jia R, et al. The novel roles of circRNAs in human cancer. *Mol Cancer* 2019;**18**:1–10.
2. Zeng X, Lin W, Guo M, et al. A comprehensive overview and evaluation of circular RNA detection tools. *PLoS Comput Biol* 2017;**13**:e1005420.

3. Tao H, Xiong Q, Zhang F, et al. Circular RNA profiling reveals chi_circ_0008219 function as microRNA sponges in pre-ovulatory ovarian follicles of goats (*Capra hircus*). *Genomics* 2018;**110**:257–66.
4. Pan X, Wenzel A, Jensen LJ, et al. Genome-wide identification of clusters of predicted microRNA binding sites as microRNA sponge candidates. *PLoS One* 2018;**13**:e0202369.
5. Hansen TB, Jensen TI, Clausen BH, et al. Natural RNA circles function as efficient microRNA sponges. *Nature* 2013;**495**:384–8.
6. Huang A, Zheng H, Wu Z, et al. Circular RNA-protein interactions: functions, mechanisms, and identification. *Theranostics* 2020;**10**:3503.
7. Pan X, Fang Y, Li X, et al. RBPsuite: RNA-protein binding sites prediction suite based on deep learning. *BMC Genomics* 2020;**21**:1–8.
8. Niu M, Zou Q, Lin C. CRBPDL: identification of circRNA-RBP interaction sites using an ensemble neural network approach. *PLoS Comput Biol* 2022;**18**:e1009798.
9. Legnini I, Di Timoteo G, Rossi F, et al. Circ-ZNF609 is a circular RNA that can be translated and functions in myogenesis. *Mol Cell* 2017;**66**:22–37.e9.
10. Jeck WR, Sorrentino JA, Wang K, et al. Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA* 2013;**19**:141–57.
11. Shao Y, Yang Y, Lu R, et al. Identification of tissue-specific circRNA hsa_circ_0000705 as an indicator for human gastric cancer. *Int J Clin Exp Pathol* 2017;**10**:3151–6.
12. Wang M, Gu B, Yao G, et al. Circular RNA expression profiles and the pro-tumorigenic function of circRNA_10156 in hepatitis B virus-related liver cancer. *Int J Med Sci* 2020;**17**:1351.
13. Danan M, Schwartz S, Edelheit S, et al. Transcriptome-wide discovery of circular RNAs in archaea. *Nucleic Acids Res* 2012;**40**:3131–42.
14. Kristensen LS, Okholm TLH, Venø MT, et al. Circular RNAs are abundantly expressed and upregulated during human epidermal stem cell differentiation. *RNA Biol* 2018;**15**:280–91.
15. Hsiao K-Y, Lin Y-C, Gupta SK, et al. Noncoding effects of circular RNA CCDC66 promote colon cancer growth and metastasis. *Cancer Res* 2017;**77**:2339–50.
16. Zhang H-D, Jiang L-H, Sun D-W, et al. CircRNA: a novel type of biomarker for cancer. *Breast Cancer* 2018;**25**:1–7.
17. Ghosal S, Das S, Sen R, et al. Circ2Traits: a comprehensive database for circular RNA potentially associated with disease and traits. *Front Genet* 2013;**4**:283.
18. Glažar P, Papavasileiou P, Rajewsky N. circBase: a database for circular RNAs. *RNA* 2014;**20**:1666–70.
19. Ono T, Kuhara S. A novel method for gathering and prioritizing disease candidate genes based on construction of a set of disease-related MeSH® terms. *BMC Bioinform* 2014;**15**:1–12.
20. Xiang Z, Qin T, Qin ZS, et al. A genome-wide MeSH-based literature mining system predicts implicit gene-to-gene relationships and networks. *BMC Syst Biol* 2013;**7**:1–15.
21. Fan C, Lei X, Fang Z, et al. CircR2Disease: a manually curated database for experimentally supported circular RNAs associated with various diseases. *Database* 2018;**2018**:bay044.
22. Yao D, Zhang L, Zheng M, et al. Circ2Disease: a manually curated database of experimentally validated circRNAs in human disease. *Sci Rep* 2018;**8**:1–6.
23. Zhao Z, Wang K, Wu F, et al. circRNA disease: a manually curated database of experimentally supported circRNA-disease associations. *Cell Death Dis* 2018;**9**:1–2.
24. 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.
25. Wang L, You Z-H, Huang Y-A, et al. An efficient approach based on multi-sources information to predict circRNA-disease associations using deep convolutional neural network. *Bioinformatics* 2020;**36**:4038–46.
26. Deepthi K, Jereesh A. Inferring potential CircRNA-disease associations via deep autoencoder-based classification. *Mol Diagn Ther* 2021;**25**:87–97.
27. Zheng K, You Z-H, Li J-Q, et al. iCDA-CGR: identification of circRNA-disease associations based on chaos game representation. *PLoS Comput Biol* 2020;**16**:e1007872.
28. Wang L, You Z-H, Li J-Q, et al. IMS-CDA: prediction of CircRNA-disease associations from the integration of multisource similarity information with deep stacked autoencoder model. *IEEE Trans Cybern* 2020;**51**:5522–31.
29. Wang L, Yan X, You Z-H, et al. SGANRDA: semi-supervised generative adversarial networks for predicting circRNA-disease associations. *Brief Bioinform* 2021;**22**:bbab028.
30. Li M, Liu M, Bin Y, et al. Prediction of circRNA-disease associations based on inductive matrix completion. *BMC Med Genomics* 2020;**13**:1–13.
31. Wei H, Liu B. iCircDA-MF: identification of circRNA-disease associations based on matrix factorization. *Brief Bioinform* 2020;**21**:1356–67.
32. Niu M, Zou Q, Wang C. GMNN2CD: identification of circRNA-disease associations based on variational inference and graph Markov neural networks. *Bioinformatics* 2022;**38**:2246–53.
33. Chen Y, Wang Y, Ding Y, et al. RGCNCDA: relational graph convolutional network improves circRNA-disease association prediction by incorporating microRNAs. *Comput Biol Med* 2022;**143**:105322.
34. Kipf TN, Welling M. Semi-supervised classification with graph convolutional networks. In: *International Conference on Learning Representations*. Palais des Congrès Neptune, Toulon, France: ICLR, 2017, 1–14.
35. Li C, Qin X, Xu X, et al. Scalable graph convolutional networks with fast localized spectral filter for directed graphs. *IEEE Access* 2020;**8**:105634–44.
36. Li J, Zhang S, Liu T, et al. Neural inductive matrix completion with graph convolutional networks for miRNA-disease association prediction. *Bioinformatics* 2020;**36**:2538–46.
37. Woo S, Park J, Lee J-Y, et al. Cbam: Convolutional block attention module. In: *Proceedings of the European Conference on Computer Vision*. Munich, Germany: ECCV, 2018, 8–14.
38. Chen T, He T, Benesty M, et al. Xgboost: extreme gradient boosting. *R package version 04-2* 2015;**1**:1–4.
39. Myerson J, Green L, Warusawitharana M. Area under the curve as a measure of discounting. *J Exp Anal Behav* 2001;**76**:235–43.
40. Li G, Yue Y, Liang C, et al. NCPCDA: network consistency projection for circRNA-disease association prediction. *RSC Adv* 2019;**9**:33222–8.
41. Wang L, You Z-H, Li Y-M, et al. GCNCDA: a new method for predicting circRNA-disease associations based on graph convolutional network algorithm. *PLoS Comput Biol* 2020;**16**:e1007568.
42. Lei X, Fang Z, Chen L, et al. PWCDCA: path weighted method for predicting circRNA-disease associations. *Int J Mol Sci* 2018;**19**:3410.

43. Bu N, Dong Z, Zhang L, et al. CircPVT1 regulates cell proliferation, apoptosis and glycolysis in hepatocellular carcinoma via miR-377/TRIM23 axis. *Cancer Manag Res* 2020;**12**:12945.
44. He Y, Huang H, Jin L, et al. CircZNF609 enhances hepatocellular carcinoma cell proliferation, metastasis, and stemness by activating the hedgehog pathway through the regulation of miR-15a-5p/15b-5p and GLI2 expressions. *Cell Death Dis* 2020;**11**:1–12.
45. Peng Y, Wang HH. Cir-ITCH inhibits gastric cancer migration, invasion and proliferation by regulating the Wnt/ β -catenin pathway. *Sci Rep* 2020;**10**:1–13.
46. Zeng X, Zhong Y, Lin W, et al. Predicting disease-associated circular RNAs using deep forests combined with positive-unlabeled learning methods. *Brief Bioinform* 2020;**21**:1425–36.
47. Zhang Y, Qiu Y, Cui Y, et al. Predicting drug-drug interactions using multi-modal deep auto-encoders based network embedding and positive-unlabeled learning. *Methods* 2020;**179**:37–46.
48. Wu Y, Zhu D, Wang X, et al. An ensemble learning framework for potential miRNA-disease association prediction with positive-unlabeled data. *Comput Biol Chem* 2021;**95**:107566.
49. Kaboutari A, Bagherzadeh J, Kheradmand F. An evaluation of two-step techniques for positive-unlabeled learning in text classification. *Int J Comput Appl Technol Res* 2014;**3**:592–4.
50. Lee WS, Liu B. Learning with positive and unlabeled examples using weighted logistic regression. In: *Proceedings of the Twentieth International Conference on International Conference on Machine Learning*. Washington, DC USA: ICML, 2003, 448–55.