

RESEARCH

Open Access



DGCLCMI: a deep graph collaboration learning method to predict circRNA-miRNA interactions

Chao Cao^{1,2}, Mengli Li², Chunyu Wang³, Lei Xu⁴, Quan Zou^{1,2}, Yansu Wang¹ and Wu Han^{5*}

Abstract

Background Numerous studies have shown that circRNA can act as a miRNA sponge, competitively binding to miRNAs, thereby regulating gene expression and disease progression. Due to the high cost and time-consuming nature of traditional wet lab experiments, analyzing circRNA-miRNA associations is often inefficient and labor-intensive. Although some computational models have been developed to identify these associations, they fail to capture the deep collaborative features between circRNA and miRNA interactions and do not guide the training of feature extraction networks based on these high-order relationships, leading to poor prediction performance.

Results To address these issues, we innovatively propose a novel deep graph collaboration learning method for circRNA-miRNA interaction, called DGCLCMI. First, it uses word2vec to encode sequences into word embeddings. Next, we present a joint model that combines an improved neural graph collaborative filtering method with a feature extraction network for optimization. Deep interaction information is embedded as informative features within the sequence representations for prediction. Comprehensive experiments on three well-established datasets across seven metrics demonstrate that our algorithm significantly outperforms previous models, achieving an average AUC of 0.960. In addition, a case study reveals that 18 out of 20 predicted unknown CMI data points are accurate.

Conclusions The DGCLCMI improves circRNA and miRNA feature representation by capturing deep collaborative information, achieving superior performance compared to prior methods. It facilitates the discovery of unknown associations and sheds light on their roles in physiological processes.

Keywords CircRNA-miRNA interaction, Graph neural networks, Collaborative filtering, Word2vec, LSTM

Background

Because of its unique closed-loop structure, circRNA has a high degree of stability in vivo and is resistant to degradation by RNA enzymes [1, 2]. However, as a special non-coding RNA, due to technological limitations, its important physiological functions were not recognized in the early stages of discovery, and it was generally regarded as a by-product of RNA processing and the result of abnormal gene splicing, which did not attract widespread attention. With the development of high-throughput sequencing technologies in recent years, the physiological function of circRNA has been gradually uncovered [3–5], and many computational models have

*Correspondence:

Wu Han
kevinwh@stanford.edu

¹ Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, China

² Yangtze Delta Region Institute (Quzhou), University of Electronic Science and Technology of China, Quzhou, Zhejiang 324003, China

³ Faculty of Computing, Harbin Institute of Technology, Harbin, Heilongjiang 150001, China

⁴ School of Electronic and Communication Engineering, Shenzhen Polytechnic University, Shenzhen, Guangdong 518055, China

⁵ Department of Statistics, Stanford University, Stanford, CA 94043, USA



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

been proposed accordingly, such as circRNA identification [6–8], circRNA and protein interactions [9–14], circRNA and disease association prediction [15–20], circRNA and drug discovery [21–26]. Among them, one of the most typical areas is the interaction between circRNA and miRNA, which plays a critical role in gene expression, cellular function regulation, and the pathological processes of diseases [13, 27–32]. For example, in tumorigenesis and development [33], circRNA ciRS-7 removes the inhibition of miR-7 on its target genes by adsorption of miR-7, thus promoting the proliferation and migration of tumor cells. In rheumatoid arthritis [34], circRNA hsa_circ_0005198 regulates miR-145, affecting the proliferation and migration of fibroblast-like synovial cells. These studies fully indicate that analyzing the regulatory mechanism mediated by circRNA-miRNA can help people understand the causes of diseases and carry out targeted prevention and treatment. Therefore, the analysis of circRNA-miRNA interactions has great research significance and potential clinical application value.

However, traditional wet lab experiments for validating circRNA-miRNA associations often require long periods for sample preparation and experimental analysis, along with expensive reagents, instruments, and equipment. In addition, they place high demands on the technical level and expertise of researchers. As a result, traditional experiments are often low-throughput, making large-scale, comprehensive, and systematic analyses difficult, which limits further research on circRNA. With the rapid development of big data analysis technologies, many association prediction algorithms have emerged [35–42], which can analyze and model existing interaction data to predict potential unknown associations with high confidence. In subsequent experimental validation, this approach effectively narrows down the most likely candidate objects, reducing the cost and trial-and-error time for experimental verification. The current CMI prediction algorithms are summarized in chronological order below.

In 2021, Lan et al. [43] introduced Gaussian interaction profile (GIP) kernels to calculate the similarity of circRNA and miRNA respectively, and construct a heterogeneous network by associating them with known CMIs. Then, they applied the DeepWalk algorithm [44] based on matrix factorization to extract the latent features of circRNA and miRNA in the heterogeneous network, ultimately predicting unknown associations via regularization, neighborhood information-based logical matrix factorization, and inner-product inference. This algorithm is named NECMA. Qian et al. [45] employed singular value decomposition (SVD) to capture the linear features of the corresponding circRNA and miRNA molecules in CMI and then calculated the multi-similarity of

these molecules using Levenshtein distance and GIP kernels. Nonlinear features were extracted using the Variational Graph Autoencoder (VGAE) [46], and LightGBM was used to predict based on both linear and nonlinear features. This algorithm is named CMIVGSD.

In 2022, Guo et al. [47] constructed a new prediction model, WSCD, which used the continuous bag of words (CBOW) model and structural deep network embedding (SDNE) [48] respectively to train word embedding representations as molecular attribute features and graph low-dimensional embeddings as behavioral features. The fused features were then inputted into convolutional neural networks (CNN) and deep neural network (DNN) for circRNA-miRNA association prediction. He et al. [49] proposed a latent interaction prediction model based on graph convolutional networks (GCN) [50], named GCNCMI. Through graph convolution operations, this algorithm effectively mines and propagates complex relationships between nodes, providing deep information for predictions. Yu et al. [51] devised a universal prediction architecture SGCNCMI, which combines multimodal information. Specifically, they first construct fused features using k -mer representation to capture RNA's attribute features and introduce GIP kernels and sigmoid kernels to capture circRNA-miRNA similarity features. Then, the Sparse Autoencoder (SAE) is used to further extract deeper information and serve as feature representations for the corresponding nodes. Finally, a GCN model aggregates adjacent information based on node attributes and association networks to predict potential CMIs. Wang et al. [52] developed an improved algorithm, KGDCMI. They input RNA attribute information, obtained from sequences and similarities via k -mer and GIP kernels, into SAE for representative feature extraction, and use HOPE graph embeddings to mine behavioral information in CMI associations. Finally, a DNN model is applied to fuse features and predict unknown CMIs.

In 2023, Wang et al. [53] introduced a denoising multi-view feature fusion prediction algorithm called JSND-CMI. In detail, it calculates the Jaccard distance between sequences as structural features and uses sigmoid kernels to compute similarities as sequence attribute features. Struc2vec [54] is applied to extract local topological structures from the association network. The multi-view features are then trained with denoising autoencoders (DAE) to obtain more robust feature representations, followed by prediction using GBDT [55]. Zhou et al. [56] designed the SPBCMI model, which combines structural features captured by graph embedding and sequence features extracted by BERT [57]. These features are input into a GBDT classifier for training to complete the CMI interaction prediction task. Wang et al. [58] proposed the

KS-CMI algorithm, which constructs a circRNA-miRNA-cancer interaction network based on balance theory to extract molecular behavioral features. Subsequently, DAE is employed to enhance feature robustness, and the CatBoost classifier is used for prediction. Li et al. [59] introduced the DeepCMI model, which integrates molecular similarity matrices and topological information from GIP kernels in biomedical graphs. Multi-source information features are obtained and mapped into the same vector space using local linear embedding, and topological information features are further extracted using text-associated DeepWalk. Finally, an XGBoost [60] classifier is employed to judge whether circRNAs and miRNAs interact. Wei et al. [61] proposed the BCMCMI model, which combines semantic features of sequences captured by BERT networks, features from cosine similarity, and topological features of heterogeneous networks captured by Metapath2vec [62], training an XGBoost classifier to predict potential CMIs.

In 2024, Guo et al. [63] put forward a new prediction algorithm, CA-CMA, combining text embedding representation and convolutional autoencoders. Firstly, Skip-Gram is used to obtain RNA embedding features, which are further refined using Convolutional Autoencoders (CAE). Meanwhile, Doc2Vec is employed to capture the semantic features of the sequences. Finally, CMI prediction is performed based on feature fusion using a DNN. Soon after, Guo et al. [64] proposed an improved algorithm BGF-CMAP, which utilized GBDT and graph embedding methods. RNA word embeddings were obtained via Word2Vec [65], and CMI topological features were extracted using graph factorization (GF) and large-scale information network embedding (LINE). These features were fused and input into GBDT for CMI classification.

Although existing CMI prediction models have achieved relatively good prediction performance through various feature embedding algorithms and efficient neural network architectures, they generally suffer from the following issues that require further improvement: First, these models overlook the exploration of deep collaborative features in circRNA and miRNA interactions. Secondly, they fail to guide the training of the underlying feature extraction network based on deep collaborative information, making it difficult to obtain representative feature embeddings, which affects the algorithm's performance.

To address these issues, we introduce and extend the neural collaborative filtering algorithm from the field of recommender systems to circRNA-miRNA association prediction. Specifically, we innovatively construct a neural graph collaborative filtering model (NGCF) combined with a joint training framework for the underlying

feature extraction network. An optimized loss function is designed to explicitly guide the training direction of the underlying feature extraction network based on deep collaborative information from circRNA-miRNA interactions. These features are then stored in their respective embeddings as representative features, and the association prediction score can be obtained by calculating the inner product of the corresponding vectors.

Experimental results on three well-established datasets demonstrate that our DGCLCMI algorithm achieves outstanding performance compared to previous methods. Additionally, the ablation study and case analysis of the two main improvements proposed in this paper validate the effectiveness of our algorithm. Therefore, DGCLCMI is an innovative and high-performance circRNA-miRNA association prediction tool (the model architecture diagram is shown in Fig. 1), which is expected to advance research in this field. The contributions of this paper are summarized as follows:

- (1) For the first time, the neural collaborative filtering algorithm is introduced and improved to mine deep interaction features of circRNA-miRNA.
- (2) An innovative joint optimization framework for deep collaborative feature mining and underlying feature extraction is constructed.
- (3) Experiments on three well-recognized datasets show that DGCLCMI achieves superior performance compared to existing methods.

Results and discussion

Performance of the proposed algorithm

We summarized publicly available datasets commonly used in previous research, including CMI-20208, CMI-9589, and CMI-9905, and conducted fivefold cross-validation on these datasets using the proposed algorithm. The obtained performance is measured using seven indicators from various aspects, and the results are shown in Fig. 2 and Table 1. At the same time, for the convenience of intuitive evaluation, we have also visualized the results, as shown in Fig. 3.

It can be observed that our algorithm achieved good performance in the fivefold cross-validation of both the 9589 and 20,208 CMI correlated pairs, demonstrating excellent performance across all metrics. Notably, the algorithm showed particularly strong performance in the specificity and precision metrics. As is well-known, higher specificity indicates a stronger ability of the model to recognize negative samples, leading to a lower misdiagnosis rate, while higher precision reflects better accuracy in identifying positive cases. Additionally, the algorithm demonstrates reasonably good performance in the sensitivity metric, which measures the recall of

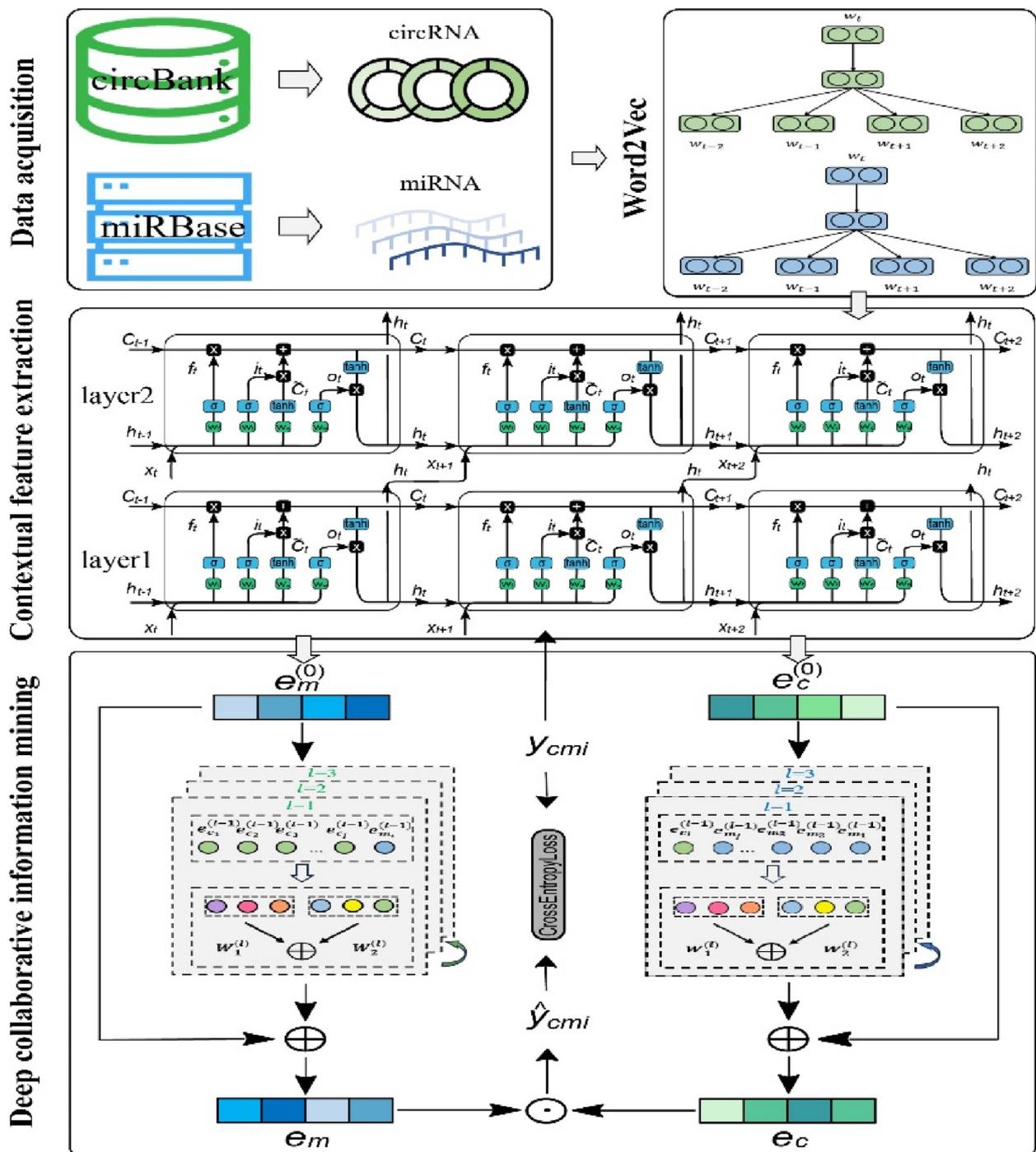


Fig. 1 DGCLCMI model overall architecture diagram

positive samples. Consequently, the algorithm achieves outstanding overall performance in the comprehensive AUC metric, with fivefold average AUCs of 0.9546, 0.9610, and 0.9645 for the CMI-20208, CMI-9589, and CMI-9905 datasets, respectively, yielding an average AUC of 0.9600. A high AUC value suggests that the

model can effectively differentiate between positive and negative samples across different thresholds.

As is known, a higher specificity value indicates a stronger ability of the model to identify negative samples, resulting in a lower misdiagnosis rate; a higher precision value reflects the model's accuracy in

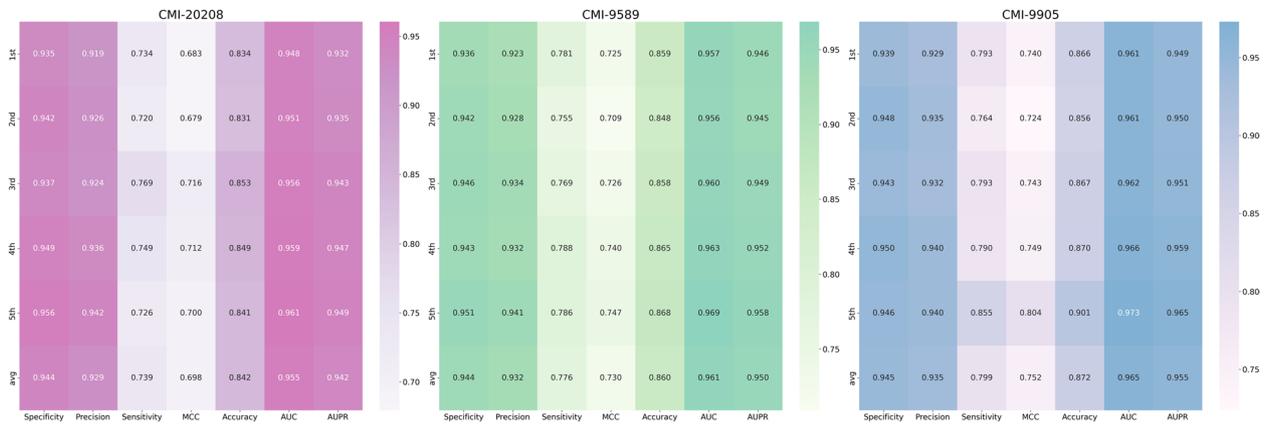


Fig. 2 Heatmap showing fivefold cross-validation results of our algorithm across three datasets

Table 1 Average performance of our algorithm on the CMI-20208, CMI-9589, and CMI-9905 datasets

Dataset	Specificity	Precision	Sensitivity	MCC	Accuracy	AUC	AUPR
CMI-20208	0.9439	0.9294	0.7394	0.6981	0.8417	0.9546	0.9415
CMI-9589	0.9437	0.9318	0.7758	0.7297	0.8596	0.9610	0.9501
CMI-9905	0.9452	0.9352	0.7988	0.7521	0.8719	0.9645	0.9546
Average	0.9443	0.9321	0.7713	0.7266	0.8577	0.9600	0.9487

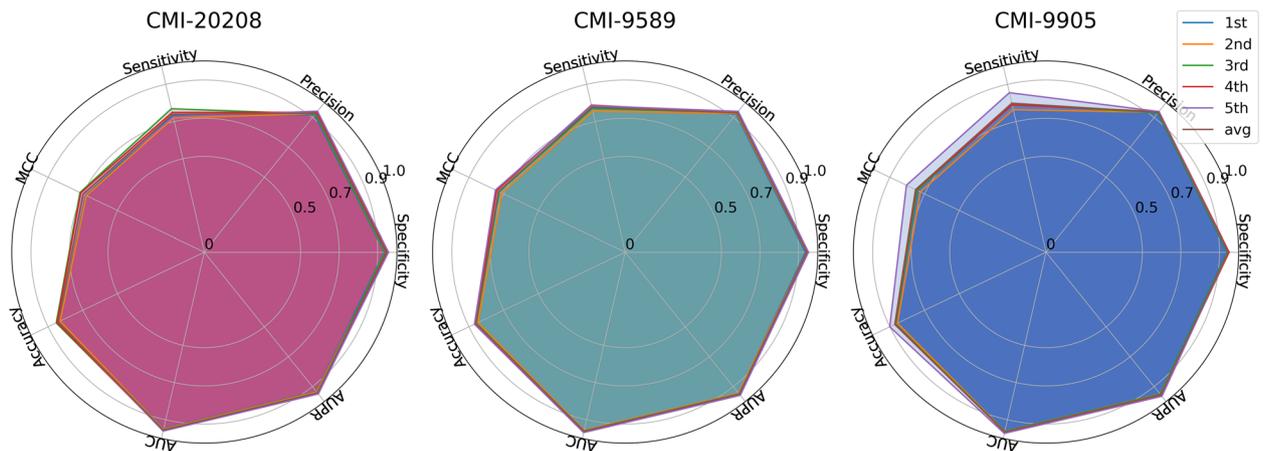


Fig. 3 Radar chart showing fivefold cross-validation results of our algorithm across three datasets

identifying positive cases. Additionally, the algorithm displayed satisfactory performance in the sensitivity metric, which measures the recall of positive samples. Therefore, the algorithm achieved excellent overall performance in the AUC (area under the Curve) metric, with the fivefold average AUC on the CMI-20208, CMI-9589, and CMI-9905 datasets being 0.9543, 0.9611, and 0.9647, respectively, with an average of 0.9600. A high AUC value indicates that the model was able to

effectively distinguish between positive and negative samples at different thresholds.

Of course, compared to other metrics, the MCC result may not stand out as much. As a metric that considers true positives, false positives, false negatives, and true negatives, MCC imposes higher demands on the algorithm's performance. Nevertheless, our algorithm demonstrates significant improvement in this metric compared to state-of-the-art methods (related

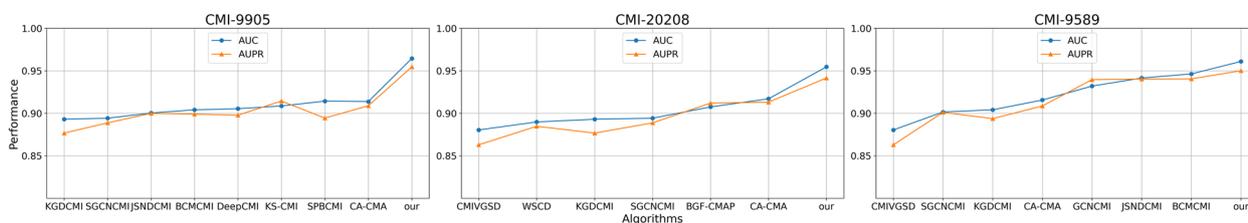


Fig. 4 AUC and ACPR curve comparison of our algorithm with existing prediction algorithms on three datasets

comparison experiments are presented in the next section). In summary, the proposed algorithm exhibits strong performance in the CMI prediction task and offers valuable insights for the exploration of potential CMIs.

Performance comparison with existing prediction algorithms

In this section, we compare the proposed method with several advanced CMI prediction models using AUC and AUPR indicators on three datasets, as shown in Fig. 4. Our method outperforms previous algorithms in terms of AUC across all datasets, with significant improvements on certain datasets. Specifically, on the CMI-9905, CMI-20208, and CMI-9589 datasets, it surpasses the second-best method by 5.07%, 3.76%, and 1.47% in AUC, and 4.58%, 2.84%, and 0.96% in AUPR, respectively.

We observed that most of the previous algorithms, such as BGF-CMAP, SPBCMI, KS-CMI, DeepCMI, BCMCMI, and JSNDCMI, use the gradient boosting tree (GBT) algorithm or its variants as classifiers for the captured features. Although these classifiers iteratively reduce prediction errors by constructing multiple decision trees and making final predictions using weighted averages, they generally only handle “static” features. These methods cannot dynamically generate new features or adjust them in real time, limiting their ability to extract deep dynamic features. For time-series or dynamic features, corresponding feature creation must occur during the data preprocessing stage. This limitation hinders the algorithm’s capacity to extract the most representative features that would maximize performance, ultimately affecting the prediction results. In contrast, our model utilizes a neural graph collaborative filtering framework combined with a bottom-layer feature extraction network. This collaborative approach captures deep information in circRNA-miRNA interactions, which is then stored in their respective embeddings. As a result, this dynamic training mechanism enables comprehensive exploration of association patterns in CMI, leading to significant performance improvements over static methods (related ablation experiments and evaluations of different classifiers are discussed in the next section).

Table 2 Comparison of four performance metrics between the proposed and existing algorithms on CMI-9905

Algorithm	Accuracy	Precision	Specificity	MCC
KGDCMI	0.8265	0.8435	0.8510	0.6538
JSNDCMI	0.8231	0.8232	0.8249	-
DeepCMI	0.8244	0.8290	-	0.6496
BCMCMI	0.8316	0.8083	-	0.6670
KS-CMI	0.8343	0.8366	-	-
SPBMCI	0.8405	-	-	0.6864
CA-CMA	0.8399	0.8328	0.8289	0.6804
Our	0.8719	0.9352	0.9452	0.7521

Table 3 Comparison of four performance metrics between the proposed and existing algorithms on CMI-20208

Algorithm	Accuracy	Specificity	Precision	MCC
WSCD	0.8161	0.8132	0.8143	0.6323
BGF-CMAP	0.8290	0.8384	0.8353	0.6581
Our	0.8417	0.9439	0.9294	0.6981

We also found that algorithms using autoencoders or their variants, such as SGCNMI, KGDCMI, CMIVGSD, and CA-CMA, which perform further feature extraction, generally fail to achieve optimal performance. While autoencoders can automatically extract low-dimensional feature representations from high-dimensional data without labels, the features they generate often lack clear practical or physical significance. Moreover, since there are no explicit supervisory signals, the extracted features are difficult to directly apply to downstream tasks. To obtain useful feature embeddings for specific tasks, significant manual intervention (e.g., tuning, adding constraints) is usually required. In contrast, the joint optimization and adaptive feature-capturing network proposed in this paper offer notable practical advantages for task-oriented feature extraction.

For a more comprehensive comparison, we present additional metrics on the CMI-9905 and CMI-20208 datasets, as shown in Tables 2 and 3. It is clear that,

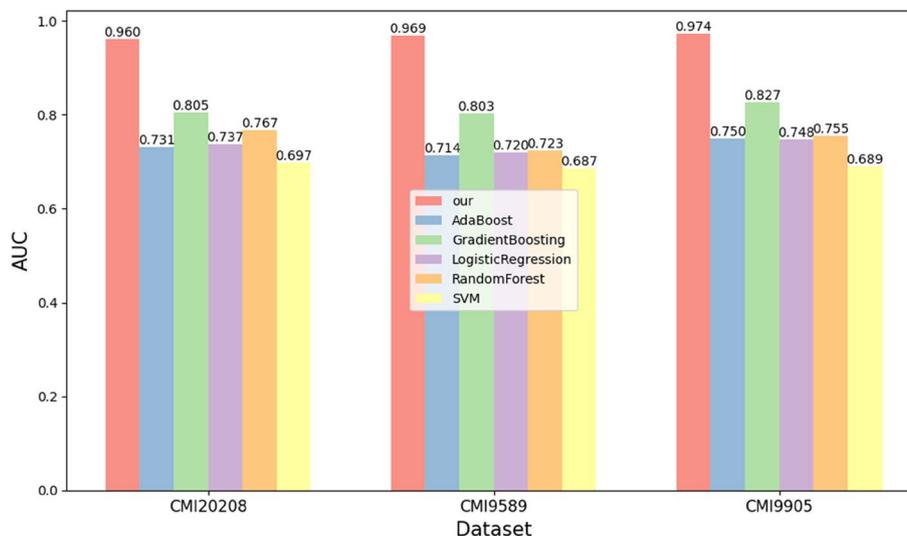


Fig. 5 Performance comparison of DGCLCMI and other classifiers on the same features numerically processed

compared to existing models, our algorithm exhibits superior performance across all metrics. In CMI-9905, compared to the CA-CMA model proposed in 2024, our method achieves approximately a 12% improvement in Specificity, a 10% improvement in Precision, and a 6% increase in overall MCC. In the larger CMI interaction network (CMI-20208), compared to the BGF-CMAP model proposed in 2024, our algorithm still achieves about an 11% improvement in Specificity, a 10% improvement in Precision, and about a 4% increase in overall MCC. These results further highlight the robustness of our algorithm, demonstrating the effectiveness of the deep graph collaboration method proposed in this paper.

Ablation experiments and analysis

Performance evaluation of different classifier algorithms

To highlight the superiority of the proposed algorithm, which introduces and improves the neural collaborative filtering algorithm to mine deep interaction features of CMI, we perform an ablation analysis comparing it with commonly used classifier algorithms in previous studies such as AdaBoost, Gradient Boosting, Logistic Regression, Random Forest, and SVM. This comparison is based on the same numerically processed sequence features, as shown in Fig. 5. The results show that our dynamic and unified model performs exceptionally well across all datasets, standing out clearly. This further validates the earlier point that our algorithm surpasses “static” classifiers, which are based on decision tree ensemble learning algorithms and cannot dynamically adjust features during training.

AdaBoost improves performance by adjusting sample weights, while Random Forest uses Bagging (Bootstrap

Aggregating) to train multiple decision trees in parallel, with the final result determined by the weighted sum of all trees. Gradient Boosting optimizes the model by gradually reducing residuals. However, these classifiers can only perform classification based on existing features and cannot adapt features based on label information, leaving room for performance improvement. Logistic Regression, being a linear model, is suitable for linearly separable data, but for the highly nonlinear CMI prediction task, it struggles to capture such complexity effectively.

We also observed that SVM did not perform well in this experiment. This might be because SVM typically excels in small sample sizes and high-dimensional, linearly separable scenarios. However, in large-scale datasets with more noise, it may not be as efficient as other algorithms, such as Random Forest or neural networks.

Performance evaluation of different feature extraction algorithms

In this section, to validate the contribution of the deep collaborative feature mining and bottom-layer feature extraction joint optimization framework proposed in this paper, as well as the chosen feature extraction algorithm LSTM, we compare and perform an ablation analysis on various feature extraction algorithms used in previous studies, such as CNN, RNN, GRU, and CAE, using the proposed neural graph collaborative filtering model (Fig. 6). We observe that the feature extraction algorithms paired with the neural graph collaborative filtering model (except for CAE) all achieved excellent performance under the joint optimization framework. However, the performance of the non-jointly optimized CAE is underwhelming, likely because the low-dimensional features

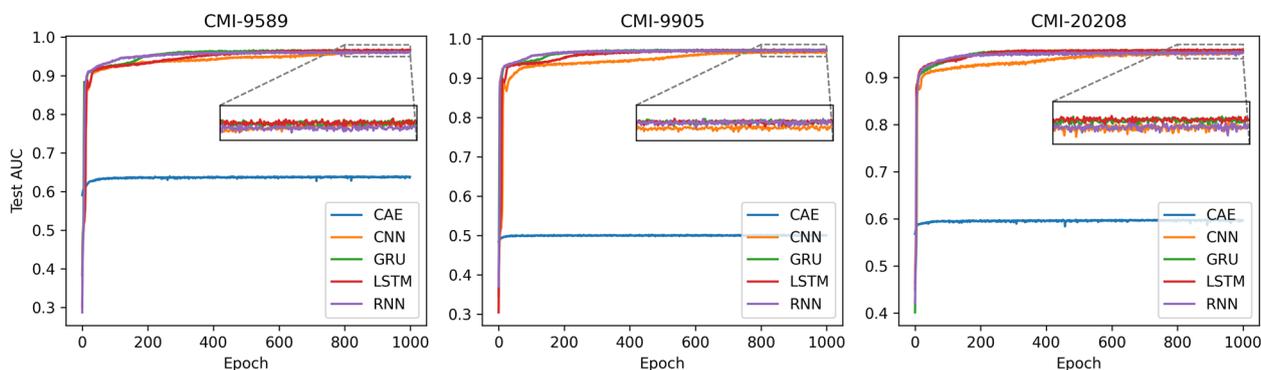


Fig. 6 Performance evaluation of various feature extraction algorithms within the same deep collaborative information mining framework

captured by the autoencoder lack clear meaning and do not have direct constraints for optimization based on the final task, resulting in suboptimal performance.

Although CNN, RNN, LSTM, and GRU show very similar performance across the three datasets, the red curve stands out slightly, achieving the best performance. The reason may be that, in time-series data, CNN extracts patterns within local time windows using one-dimensional convolutional kernels, similar to how local spatial features are extracted in images. It can capture the local correlations between different time steps in the sequence. However, for long-term dependency problems, it typically requires increasing the number of convolution layers or using a larger receptive field (filter size), but this does not fully address the long-range dependency issue. As a result, CNN's performance ranks lower among the four algorithms. The remaining three algorithms—RNN, LSTM, and GRU—are commonly used for sequence data processing. RNN is the most basic recurrent neural network structure, processing data step by step in a sequence by sharing the same weights. Each time step receives the current input and the hidden state from the previous time step. However, RNN suffers from gradient vanishing or explosion problems when processing long sequences, making it difficult to capture long-range dependencies. LSTM was specifically designed to address the long-range dependency issue in standard RNNs. It introduces a “cell state” to directly pass information and uses a “gate mechanism” to control the flow of information. LSTM effectively remembers long-term information and selectively “forgets” irrelevant data, making it suitable for long-sequence tasks. However, due to its complex gating structure, LSTM has a higher computational cost and longer training time (as shown in the figure, LSTM is the slowest to converge among the three). GRU is a simplified version of LSTM, merging some of the gating mechanisms in LSTM to reduce the number of parameters.

Although the performance difference between GRU and LSTM is not significant in the experiments of this paper, the simpler structure of GRU may result in slightly weaker performance when handling long-range dependencies in very long sequences. Therefore, the model in this paper selects LSTM, in combination with the neural graph collaborative filtering model, as the primary model for capturing deep collaborative information.

Case study

In order to further validate the proposed algorithm, which aids in the exploration of unknown associations, we conduct a case study in this section. First, we train the algorithm on CMI data with known labels and then predict whether there is an interaction between unknown circRNA and miRNA pairs. Unknown samples with high confidence were selected and further tested by consulting relevant literature and the CircInteractome database. If relevant literature or database data existed, the sample was labeled as “Confirmed”; otherwise, it was labeled as “Unconfirmed”. The results are shown in Table 4. It can be seen that, of the 20 unknown correlated sample pairs in the table, 18 have been confirmed, and the remaining 2 may also be confirmed in the future through practical testing. Overall, the proposed algorithm is able to provide potential interaction pairs with high confidence, effectively narrowing down the scope of candidates and reducing the cost of experimental trial and error, demonstrating a strong ability to identify potential CMIs.

Conclusions

Circular RNAs (circRNAs) facilitate the expression of specific target genes by modulating miRNA activity and alleviating miRNA-mediated suppression, thereby influencing critical cellular processes including proliferation, differentiation, and apoptosis. Therefore, studying circRNA-miRNA interactions is crucial for deciphering intracellular regulatory networks and understanding

Table 4 Twenty high-confidence CMIs predicted by DGCLCMI

No	circRNA	miRNA	Evidence
1	hsa_circ_0009566	hsa-miR-646	Confirmed
2	hsa_circ_0010085	hsa-miR-658	Confirmed
3	hsa_circ_0023020	hsa-miR-326	Confirmed
4	hsa_circ_0024899	hsa-miR-663b	Confirmed
5	hsa_circ_0030874	hsa-miR-658	Confirmed
6	hsa_circ_0041605	hsa-miR-361-3p	Confirmed
7	hsa_circ_0049497	hsa-miR-1183	Confirmed
8	hsa_circ_0049497	hsa-miR-3182	Unconfirmed
9	hsa_circ_0050270	hsa-miR-1289	Confirmed
10	hsa_circ_0050501	hsa-miR-658	Confirmed
11	hsa_circ_0056345	hsa-miR-1299	Confirmed
12	hsa_circ_0063366	hsa-miR-1224-3p	Confirmed
13	hsa_circ_0063366	hsa-miR-6504-5p	Unconfirmed
14	hsa_circ_0069972	hsa-miR-326	Confirmed
15	hsa_circ_0073951	hsa-miR-450b-3p	Confirmed
16	hsa_circ_0075975	hsa-miR-361-3p	Confirmed
17	hsa_circ_0078807	hsa-miR-1183	Confirmed
18	hsa_circ_0079478	hsa-miR-1224-3p	Confirmed
19	hsa_circ_0081151	hsa-miR-587	Confirmed
20	hsa_circ_0081156	hsa-miR-326	Confirmed

complex gene expression mechanisms. Although existing computational models have been proposed for predicting these interactions, they predominantly suffer from two fundamental limitations:

They overlook the extraction of interaction collaboration features and fail to train the feature extraction network based on this information, which affects the performance of the algorithms. To resolve these drawbacks, we introduce DGCLCMI, a new deep graph collaborative learning framework. Specifically, we enhanced the NGCF model to capture deep collaborative features of CMIs and employed these signals to guide the extraction of representative features from biological sequences for prediction. In comprehensive evaluations across three benchmark CMI datasets, our algorithm demonstrates superior performance over state-of-the-art methods. In conclusion, our framework exhibits strong CMI prediction performance, which facilitates the exploration of unknown CMIs, thereby revealing the underlying disease-regulating networks and advancing the development of early diagnosis and targeted therapies.

Methods

Datasets

In this study, to better evaluate the performance of the proposed algorithm and facilitate comparison with existing prediction models, we use the publicly available, experimentally validated datasets CMI-9905,

Table 5 Datasets details

Dataset	CMI-9589	CMI-9905	CMI-20208
CircRNA	2115	2346	3569
MiRNA	821	962	1152
CMI	9589	9905	20,208

CMI-9589, and CMI-20208, which have been employed in previous studies [63]. This allows for a performance evaluation under the same benchmark datasets. The detailed data are presented in Table 5. These datasets were sourced from the circBank [66] (<http://www.circbank.cn/>) and miRBase [67] (<https://www.mirbase.org/>) databases, containing experimentally verified circRNA-miRNA interaction data. For model training and evaluation, we also randomly selected an appropriate number of negative samples to construct the circRNA-miRNA association matrix D and performed a five-fold cross-validation to randomly split the dataset into training and testing sets.

Proposed model architecture

The architecture diagram of our prediction model is shown in Fig. 1, which primarily consists of four main components: preliminary sequence feature extraction, sequence time dependence capturing, circRNA-miRNA deep collaboration information mining, and CMI interaction predicting. In the following sections, we will describe each of these four modules in detail to more clearly illustrate the internal structure and functionality of the algorithm proposed in this paper.

Sequence numerical processing

The original circRNA and miRNA sequences are only composed of four bases: A, G, C, and U, making the data highly specialized and difficult to interpret. To convert these sequences into numerical features that can be understood by machines and facilitate further analysis, we adopt the Skip-gram model, which is commonly used for word representation in natural language processing (NLP), for the initial extraction of sequence numerical features.

Skip-gram [65], as a training form of Word2Vec, uses the center word w_t of a given sentence to maximize the corresponding context word w_{t+j} of the predicted position to train the word representation (the formula is expressed as follows). Skip-gram is used in many NLP tasks and has achieved significant performance improvement.

$$\frac{1}{T} \sum_{t=1}^T \sum_{-k \leq j \leq k, j \neq 0} \log P(w_{t+j} | w_t) \quad (1)$$

For the encoding of circRNA and miRNA, consistent with CA-CMA, we treated each base as a word, set the dimension of the word vector to 64, and used the Skip-gram model to train the embedding feature representations for the four bases of the corresponding RNA sequences.

Sequence contextual feature extraction

In order to capture the temporal contextual features in the sequence, we employ the well-known LSTM network in the field of NLP to regulate information flow through a gating mechanism. This mechanism filters valuable features, removes noise data, and effectively retains long-range dependencies. The key components of LSTM include the forget gate, input gate, and output gate, which determine how information is retained or discarded at each time step.

Specifically, the forget gate controls which information should be discarded from the cell state, and its computation is as follows:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2)$$

where f_t is the output of the forget gate. σ is the sigmoid activation function. W_f is the weight matrix of the forget gate. h_{t-1} is the hidden state of the previous time step. x_t is the current input. b_f is the bias term.

The input gate determines which new information should be added to the cell state. This process consists of two steps: (1) generating new candidate information; (2) updating the cell state based on the input gate. First, the candidate cell states are calculated:

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (3)$$

where \tilde{C}_t represents the new candidate information, and \tanh is the hyperbolic tangent function.

Input gate calculation:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (4)$$

Updated cell status:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (5)$$

Output gate calculation:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (6)$$

Updated hidden status:

$$h_t = o_t \cdot \tanh(C_t) \quad (7)$$

To effectively capture the semantic features of circRNA and miRNA sequences, we set the input feature

dimension of the LSTM network to 64 and the hidden state dimension to 256. Additionally, we stack two layers of LSTM networks to enhance the model's capability to capture dependencies. The final output is compressed and refined using a fully connected layer, yielding 128-dimensional feature representations.

circRNA-miRNA deep collaborative information mining

Inspired by the NGCF model proposed by Wang et al. [68] and the MLNGCF model [69], we apply the NGCF model with improvements to extend its applicability to CMI prediction tasks. Moreover, in previous studies employing NGCF, feature extraction, and collaborative information mining were treated as independent modules and trained separately. For example, the MLNGCF model first applied a deep autoencoder network (DAE) to extract low-dimensional embeddings from circRNA-circRNA and disease-disease similarity graphs. These embeddings were then fed into NGCF to mine collaborative information. However, the feature extraction and collaborative information mining modules were trained separately, preventing NGCF from leveraging the mined collaborative information to refine feature embeddings, thereby limiting their representativeness.

In this study, we integrate the feature extraction module and the deep interactive information mining module into a unified optimization framework. By capturing the sequence context in the previous section, we obtained initial representative features of circRNA and miRNA sequences. We then propose a deep collaboration model that further explores circRNA-miRNA interactions to refine sequence embeddings based on CMI interaction data. Additionally, the mined collaboration information is leveraged to guide the extraction of sequence context features through gradient backpropagation. Specifically, we employ GNN to construct a multi-layer message passing mechanism, capturing circRNA-miRNA collaborative signals based on the CMI graph structure and optimizing the learned embeddings of circRNA and miRNA.

Construction of multi-layer message propagation mechanism

Drawing inspiration from recommendation systems, we generalize a similar concept by treating miRNAs interacting with circRNA as features of circRNA, thereby measuring the similarity between different miRNAs. Additionally, the interaction between circRNA and miRNA reflects the binding preference of circRNA for specific miRNAs.

Message transfer

To achieve this, we design a message passing mechanism for information exchange between circRNA and miRNA. Given a set of CMI data (c,m) and their corresponding embeddings e_c and e_m , the message propagation mechanism encoding $m \rightarrow c$ integrates the embedding information of e_m and interactive encoding between e_c and e_m , formulated as follows:

$$m_{m \rightarrow c} = f(e_m, e_c, p_{mc}) \quad (8)$$

$$m_{m \rightarrow c} = \frac{1}{\sqrt{|\mathcal{N}_c| |\mathcal{N}_m|}} (W_1 e_m + W_2 (e_m \odot e_c)) \quad (9)$$

where p_{mc} represents a loss factor in the message passing process, N_c and N_m denote the respective first-order neighborhoods, and W_1 and W_2 are the weight matrices for information propagation, responsible for extracting useful information from the corresponding elements.

Message aggregation

According to the miRNA message transmission $m_{m \rightarrow c}$, the original embedded information e_c is aggregated to obtain a new representation $e_c^{(1)}$. The weight matrix W_1 remains consistent with the one used in previous layers.

$$e_c^{(1)} = \text{LeakyReLU} \left(W_1 e_c + \sum_{i \in \mathcal{N}_c} m_{m \rightarrow c} \right) \quad (10)$$

By stacking the aforementioned l message propagation layers, circRNA and miRNA can receive collaborative signals propagated from their l -order neighbors, thereby capturing higher-order interaction information of CMI. Moreover, the model explicitly encodes deep collaborative information into the sequence representations, which is crucial for the subsequent evaluation of the association strength between circRNA and miRNA.

$$\begin{cases} m_{m \rightarrow c}^{(l)} = \frac{1}{\sqrt{|\mathcal{N}_c| |\mathcal{N}_m|}} (W_1^{(l)} e_m^{(l-1)} + W_2^{(l)} (e_m^{(l-1)} \odot e_c^{(l-1)})) \\ e_c^{(l)} = \text{LeakyReLU} \left(W_1^{(l)} e_c^{(l-1)} + \sum_{i \in \mathcal{N}_c} m_{m \rightarrow c}^{(l)} \right) \end{cases} \quad (11)$$

To facilitate parallelization, we employ a matrix representation for layer-wise message propagation based on the GNN computational framework. Specifically, the interaction matrix R is derived from the training set of CMI data, and we construct the adjacency matrix $A \in R^{(n+m) \times (n+m)}$ of the collaboration graph, where n and m denote the numbers of miRNAs and circRNAs, respectively. The Laplacian matrix is then computed and normalized. Based on the computational framework of GNN, the matrix formulation for message propagation can be derived accordingly.

$$A_{cm} = \begin{bmatrix} 0 & R \\ R^T & 0 \end{bmatrix} \quad (12)$$

$$\mathcal{L} = \mathbf{D}^{-\frac{1}{2}} (A_{cm} + I) \mathbf{D}^{-\frac{1}{2}} \quad (13)$$

$$\mathbf{E}^{(0)} = [e_{m_1}, \dots, e_{m_n}, e_{c_1}, \dots, e_{c_m}] \in R^{(n+m) \times d} \quad (14)$$

$$\mathbf{E}^{(l)} = \text{LeakyReLU} \left(\mathcal{L} \mathbf{E}^{(l-1)} \mathbf{W}_1^{(l)} + (\mathbf{E}^{(l-1)} \odot \mathcal{L} \mathbf{E}^{(l-1)}) \mathbf{W}_2^{(l)} \right) \quad (15)$$

where I denotes the identity matrix and D represents the degree matrix. Through the l -layer message passing process, we can obtain high-order circRNA-miRNA collaborative signals and encode them into the corresponding sequence embeddings as representative sequence features.

Prediction of CMI interaction

After sequence semantic feature extraction and deep interaction information mining, we can obtain the representative features of circRNA and miRNA sequences in a shared feature space. Thus, to evaluate the interaction between circRNAs and miRNAs, we directly calculate the inner product of their corresponding feature embeddings, e_c and e_m , from which interaction scores can be derived. To closely fit the training data and capture the underlying CMI mechanism, we use cross-entropy loss as a measure of the difference between the model's predicted and true values and apply gradient backpropagation to update the model parameters.

$$\hat{y}_{\text{CMI}} = e_m^T e_c \quad (16)$$

$$\text{loss} = \frac{1}{k} \sum_i^k [y_i \log \hat{y}_{\text{CMI}} + (1 - y_i) \log (1 - \hat{y}_{\text{CMI}})] \quad (17)$$

Experimental setup and evaluation metrics

The experimental results of the model proposed in this paper are based on the following settings: the learning rate lr is set to $1e-4$, the batch_size is 128, and the graph collaboration network employs three message-passing layers, each with a size of 64. The node_dropout and mess_dropout are both set to 0.1. The Adam optimizer is used to train the entire model with betas=(0.9, 0.999), eps= $1e-8$. In addition, the performance of the algorithm is evaluated using a variety of metrics, including precision (Prec.), specificity (Spec.), sensitivity (Sens.), accuracy (Accu), Matthews correlation coefficient (MCC), area under the curve (AUC), and area under the precision-recall curve (AUPR), providing a comprehensive evaluation of the model's performance.

Abbreviations

circRNA	Circular RNA
miRNA	MicroRNA
CMI	CircRNA-miRNA interaction
CNN	Convolutional neural network
RNN	Recurrent neural network
GNN	Graph neural network
LSTM	Long short-term memory
GRU	Gated recurrent unit
GIP	Gaussian interaction profile
SVD	Singular value decomposition
VGAE	Variational graph autoencoder
CBOW	Continuous bag of words
SDNE	Structural deep network embedding
DNN	Deep neural network
GCN	Graph convolutional networks
SAE	Sparse autoencoder
DAE	Denoising autoencoders
GBDT	Gradient boosting decision tree
BERT	Bidirectional Encoder Representations from Transformers
CatBoost	Categorical boosting
XGBoost	Extreme gradient boosting
CAE	Convolutional autoencoders
GF	Graph factorization
LINE	Large-scale information network embedding
NGCF	Neural graph collaborative filtering model
GBT	Gradient boosting tree
AUC	Area under the curve
MCC	Matthews correlation coefficient
AUPR	Area under the precision-recall curve
NLP	Natural language processing

Acknowledgements

We would like to thank the three anonymous reviewers and the relevant journal staff, whose constructive feedback has been very helpful in enhancing the presentation of this paper.

Authors' contributions

C.C. conceived and designed the experiment. C.C. and M.L. performed the experiment. C.W. and L.X. analyzed the results and revised the experimental process. Y.W. and Q.Z. revised the manuscript. Q.Z. and W.H. provided funding, resources, and project administration. All authors reviewed and approved the final manuscript.

Funding

Our work was supported by the National Natural Science Foundation of China (62231013, 62422113, 62271329, 62373080), Shenzhen Polytechnic University Research Fund (nos. 6024310027K, 6022310036K), Shenzhen Science and Technology Program (20231129091450002), and Sichuan Tianfu Emei Plan.

Data availability

All data generated or analysed during this study are included in this published article, its supplementary information files and publicly available repositories: Zenodo [70] (<https://zenodo.org/records/15063000>) and GitHub (<https://github.com/cc646201081/DGCLCMI>).

Declarations**Ethics approval and consent to participate**

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 15 January 2025 Accepted: 25 March 2025

Published: 23 April 2025

References

- Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, Maier L, Mackowiak SD, Gregersen LH, Munschauer M. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature*. 2013;495(7441):333–8.
- Conn SJ, Pillman KA, Toubia J, Conn VM, Salamanidis M, Phillips CA, Roslan S, Schreiber AW, Gregory PA, Goodall GJ. The RNA binding protein quaking regulates formation of circRNAs. *Cell*. 2015;160(6):1125–34.
- Huang G, Li S, Yang N, Zou Y, Zheng D, Xiao T. Recent progress in circular RNAs in human cancers. *Cancer Lett*. 2017;404:8–18.
- Feng X-Y, Zhu S-X, Pu K-J, Huang H-J, Chen Y-Q, Wang W-T. New insight into circRNAs: characterization, strategies, and biomedical applications. *Exp Hematol Oncol*. 2023;12(1):91.
- Verduci L, Tarcitano E, Strano S, Yarden Y, Blandino G. CircRNAs: role in human diseases and potential use as biomarkers. *Cell Death Dis*. 2021;12(5):468.
- Zeng X, Lin W, Guo M, Zou Q. Details in the evaluation of circular RNA detection tools: Reply to Chen and Chuang. *PLoS Comput Biol*. 2019;15(4): e1006916.
- Niu M, Wang C, Chen Y, Zou Q, Qi R, Xu L. CircRNA identification and feature interpretability analysis. *BMC Biol*. 2024;22(1):44.
- Liu Y, Li R, Ding Y, Hei X, Wu F-X. P4PC: a portal for bioinformatics resources of piRNAs and circRNAs. *Curr Bioinform*. 2024;19(9):873–8.
- Niu M, Zou Q, Lin C. CRBPD: identification of circRNA-RBP interaction sites using an ensemble neural network approach. *PLoS Comput Biol*. 2022;18(1): e1009798.
- Cao C, Wang C, Yang S, Zou Q. CircSI-SSL: circRNA-binding site identification based on self-supervised learning. *Bioinformatics*. 2024;40(1):btac004.
- Xu Z, Song L, Liu S, Zhang W. Deepcrbp: improved predicting function of circrna-rbp binding sites with deep feature learning. *Front Comp Sci*. 2024;18(2): 182907.
- Guo Y, Lei X, Liu L, Pan Y. circ2CBA: prediction of circRNA-RBP binding sites combining deep learning and attention mechanism. *Front Comp Sci*. 2023;17(5): 175904.
- Ogunjobi TT, Ohaeri PN, Akintola OT, Atanda DO, Orji FP, Adebayo JO, Abdul SO, Eji CA, Asebebe AB, Shodipe OO: Bioinformatics applications in chronic diseases: a comprehensive review of genomic, transcriptomics, proteomic, metabolomics, and machine learning approaches. *Medinformatics*. 2024.
- Zulfiqar H, Guo Z, Ahmad RM, Ahmed Z, Cai P, Chen X, Zhang Y, Lin H, Shi Z. Deep-STP: a deep learning-based approach to predict snake toxin proteins by using word embeddings. *Front Med*. 2024;10:1291352.
- Chen Y, Wang J, Wang C, Liu M, Zou Q. Deep learning models for disease-associated circRNA prediction: a review. *Brief Bioinform*. 2022;23(6):bbac364.
- Niu M, Wang C, Zhang Z, Zou Q. A computational model of circRNA-associated diseases based on a graph neural network: prediction and case studies for follow-up experimental validation. *BMC Biol*. 2024;22(1):24.
- Tian Y, Zou Q, Wang C, Jia C. MAMLCDA: a meta-learning model for predicting circRNA-disease association based on MAML combined with CNN. *IEEE J Biomed Health Inform*. 2024;28(7):4325–35.
- Zhu H, Hao H, Yu L. Identification of microbe–disease signed associations via multi-scale variational graph autoencoder based on signed message propagation. *BMC Biol*. 2024;22(1):172.
- Zhang W, Wei H, Zhang W, Wu H, Liu B. Multiple types of disease-associated RNAs identification for disease prognosis and therapy using heterogeneous graph learning. *Sci China Inf Sci*. 2024;67(8): 189103.
- Wu S, Feng J, Liu C, Wu H, Qiu Z, Ge J, Sun S, Hong X, Li Y, Wang X. Machine learning aided construction of the quorum sensing communication network for human gut microbiota. *Nat Commun*. 2022;13(1):3079.
- Chen Y, Wang J, Wang C, Zou Q. AutoEdge-CCP: a novel approach for predicting cancer-associated circRNAs and drugs based on automated edge embedding. *PLoS Comput Biol*. 2024;20(1): e1011851.

22. Liu M, Li C, Chen R, Cao D, Zeng X. Geometric deep learning for drug discovery. *Expert Syst Appl.* 2024;240: 122498.
23. Huang Z, Xiao Z, Ao C, Guan L, Yu L. Computational approaches for predicting drug-disease associations: a comprehensive review. *Front Comp Sci.* 2025;19(5):1–15.
24. Xiang H, Zeng L, Hou L, Li K, Fu Z, Qiu Y, Nussinov R, Hu J, Rosen-Zvi M, Zeng X. A molecular video-derived foundation model for scientific drug discovery. *Nat Commun.* 2024;15(1):9696.
25. Ren X, Wei J, Luo X, Liu Y, Li K, Zhang Q, Gao X, Yan S, Wu X, Jiang X. HydrogelFinder: a foundation model for efficient self-assembling peptide discovery guided by non-peptidal small molecules. *Advanced Science.* 2024;11:2400829.
26. Ai C, Yang H, Liu X, Dong R, Ding Y, Guo F. MTMol-GPT: de novo multi-target molecular generation with transformer-based generative adversarial imitation learning. *PLoS Comput Biol.* 2024;20(6): e1012229.
27. Cheng J, Zhuo H, Xu M, Wang L, Xu H, Peng J, Hou J, Lin L, Cai J. Regulatory network of circRNA-miRNA-mRNA contributes to the histological classification and disease progression in gastric cancer. *J Transl Med.* 2018;16:1–14.
28. Ma B, Wang S, Wu W, Shan P, Chen Y, Meng J, Xing L, Yun J, Hao L, Wang X. Mechanisms of circRNA/lncRNA-miRNA interactions and applications in disease and drug research. *Biomed Pharmacother.* 2023;162: 114672.
29. Wei M, Wang L, Li Y, Li Z, Zhao B, Su X, Wei Y, You Z. BioKG-CMI: a multi-source feature fusion model based on biological knowledge graph for predicting circRNA-miRNA interactions. *Science China Inf Sci.* 2024;67(8): 189104.
30. Qiao J, Jin J, Yu H, Wei L. Towards retraining-free RNA modification prediction with incremental learning. *Inf Sci.* 2024;660: 120105.
31. Zou Q, Xing P, Wei L, Liu B. Gene2vec: gene subsequence embedding for prediction of mammalian N6-methyladenosine sites from mRNA. *RNA.* 2019;25(2):205–18.
32. Li H, Liu B. BioSeq-Diablo: biological sequence similarity analysis using Diabolo. *PLoS Comput Biol.* 2023;19(6): e1011214.
33. Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK, Kjems J. Natural RNA circles function as efficient microRNA sponges. *Nature.* 2013;495(7441):384–8.
34. Li Z, Huang C, Bao C, Chen L, Lin M, Wang X, Zhong G, Yu B, Hu W, Dai L. Exon-intron circular RNAs regulate transcription in the nucleus. *Nat Struct Mol Biol.* 2015;22(3):256–64.
35. Chen X, Li TH, Zhao Y, Wang CC, Zhu CC. Deep-belief network for predicting potential miRNA-disease associations. *Briefings in bioinformatics.* 2021;22(3):bbaa186.
36. Ha J, Park C, Park C, Park S. IMIPMF: inferring miRNA-disease interactions using probabilistic matrix factorization. *J Biomed Inform.* 2020;102: 103358.
37. Ha J, Park S. NCMD: Node2vec-based neural collaborative filtering for predicting miRNA-disease association. *IEEE/ACM Trans Comput Biol Bioinf.* 2022;20(2):1257–68.
38. Ha J. MDMF: predicting miRNA-disease association based on matrix factorization with disease similarity constraint. *Journal of Personalized Medicine.* 2022;12(6): 885.
39. Ha J. SMAP: similarity-based matrix factorization framework for inferring miRNA-disease association. *Knowl-Based Syst.* 2023;263: 110295.
40. Ha J, Park C. MLMD: metric learning for predicting miRNA-disease associations. *IEEE Access.* 2021;9:78847–58.
41. Ha J. LncRNA expression profile-based matrix factorization for identifying lncRNA-disease associations. *IEEE Access.* 2024.
42. Ha J. Graph convolutional network with neural collaborative filtering for predicting miRNA-disease association. *Biomedicines.* 2025;13(1): 136.
43. Lan W, Zhu M, Chen Q, Chen J, Ye J, Liu J, Peng W, Pan S. Prediction of circRNA-miRNA associations based on network embedding. *Complexity.* 2021;2021(1):6659695.
44. Perozzi B, Al-Rfou R, Skiena S. Deepwalk: online learning of social representations. In: *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining.* 2014:701–10.
45. Qian Y, Zheng J, Zhang Z, Jiang Y, Zhang J, Deng L. CMIVGSD: circRNA-miRNA interaction prediction based on Variational graph auto-encoder and singular value decomposition. In: *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM).* Houston: IEEE; 2021. p. 205–10.
46. Kipf TN, Welling M. Variational graph auto-encoders. 2016. arXiv preprint arXiv:161107308.
47. Guo LX, You ZH, Wang L, Yu CQ, Zhao BW, Ren ZH, Pan J. A novel circRNA-miRNA association prediction model based on structural deep neural network embedding. *Brief Bioinform.* 2022;23(5):bbac391.
48. Wang D, Cui P, Zhu W. Structural deep network embedding. In: *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining.* 2016:1225–34.
49. He J, Xiao P, Chen C, Zhu Z, Zhang J, Deng L. GCNCMI: a graph convolutional neural network approach for predicting circRNA-miRNA interactions. *Front Genet.* 2022;13: 959701.
50. Kipf TN, Welling M. Semi-supervised classification with graph convolutional networks. 2017.
51. Yu C-Q, Wang X-F, Li L-P, You Z-H, Huang W-Z, Li Y-C, Ren Z-H, Guan Y-J. SGCNCMI: a new model combining multi-modal information to predict circRNA-related miRNAs, diseases and genes. *Biology.* 2022;11(9): 1350.
52. Wang X-F, Yu C-Q, Li L-P, You Z-H, Huang W-Z, Li Y-C, Ren Z-H, Guan Y-J. KGDCMI: a new approach for predicting circRNA-miRNA interactions from multi-source information extraction and deep learning. *Front Genet.* 2022;13: 958096.
53. Wang XF, Yu CQ, You ZH, Li LP, Huang WZ, Ren ZH, Li YC, Wei MM. A feature extraction method based on noise reduction for circRNA-miRNA interaction prediction combining multi-structure features in the association networks. *Brief Bioinform.* 2023;24(3):bbad111.
54. Ribeiro LF, Saverese PH, Figueiredo DR. struc2vec: Learning node representations from structural identity. In: *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining.* 2017. p. 385–94.
55. Friedman JH. Greedy function approximation: a gradient boosting machine. *Annals of statistics.* 2001;29:1189–232.
56. Zhou J, Wang X, Niu R, Shang X, Wen J. Predicting circRNA-miRNA interactions utilizing transformer-based RNA sequential learning and high-order proximity preserved embedding. *Iscience.* 2024;27:27(1).
57. Devlin J, Chang M-W, Lee K, Toutanova K: Bert: Pre-training of deep bidirectional transformers for language understanding. In: *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies.* 2019;1(long and short papers):4171–86.
58. Wang XF, Yu CQ, You ZH, Qiao Y, Li ZW, Huang WZ, Zhou JR, Jin HY. KS-CMI: a circRNA-miRNA interaction prediction method based on the signed graph neural network and denoising autoencoder. *Iscience.* 2023;26(8):107478.
59. Li Y-C, You Z-H, Yu C-Q, Wang L, Hu L, Hu P-W, Qiao Y, Wang X-F, Huang Y-A. DeepCMI: a graph-based model for accurate prediction of circRNA-miRNA interactions with multiple information. *Brief Funct Genomics.* 2024;23(3):276–85.
60. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. 2016. p. 785–94.
61. Wei M-M, Yu C-Q, Li L-P, You Z-H, Wang L. BCMCMI: a fusion model for predicting circRNA-miRNA interactions combining semantic and meta-path. *J Chem Inf Model.* 2023;63(16):5384–94.
62. Dong Y, Chawla NV, Swami A. metapath2vec: Scalable representation learning for heterogeneous networks. In: *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining.* 2017. p. 135–44.
63. Guo LX, Wang L, You ZH, Yu CQ, Hu ML, Zhao BW, Li Y. Likelihood-based feature representation learning combined with neighborhood information for predicting circRNA-miRNA associations. *Brief Bioinform.* 2024;25(2):bbae020.
64. Guo LX, Wang L, You ZH, Yu CQ, Hu ML, Zhao BW, Li Y. Biolinguistic graph fusion model for circRNA-miRNA association prediction. *Briefings in Bioinformatics.* 2024;25(2):bbae058.
65. Mikolov T, Chen K, Corrado G, Dean J. Efficient estimation of word representations in vector space. 2013.
66. Liu M, Wang Q, Shen J, Yang BB, Ding X. Circbank: a comprehensive database for circRNA with standard nomenclature. *RNA Biol.* 2019;16(7):899–905.
67. Griffiths-Jones S, Grocock RJ, Van Dongen S, Bateman A, Enright AJ. miR-Base: microRNA sequences, targets and gene nomenclature. *Nuc Acids Res.* 2006;34(suppl_1):D140–4.

68. Wang X, He X, Wang M, Feng F, Chua TS. Neural graph collaborative filtering. In: Proceedings of the 42nd international ACM SIGIR conference on research and development in information retrieval. 2019. p. 165–74.
69. Wu Q, Deng Z, Zhang W, Pan X, Choi KS, Zuo Y, Shen HB, Yu DJ. MLNGCF: circRNA–disease associations prediction with multilayer attention neural graph-based collaborative filtering. *Bioinformatics*. 2023;39(8):btad499.
70. Cao C. DGCLCMI (code). Zenodo. 2025. <https://doi.org/10.5281/zenodo.15063000>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.