BMC Genomics

RESEARCH

Open Access

Disentangled similarity graph attention heterogeneous biological memory network for predicting disease-associated miRNAs

Yinbo Liu[1,2†], Qi Wu[1†], Le Zhou[2], Yuchen Liu[2,3], Chao Li[2], Zhuoyu Wei[1], Wei Peng[1], Yi Yue[1*] and Xiaolei Zhu[1*]

Abstract

Background The association between MicroRNAs (miRNAs) and diseases is crucial in treating and exploring many diseases or cancers. Although wet-lab methods for predicting miRNA-disease associations (MDAs) are effective, they are often expensive and time-consuming. Significant advancements have been made using Graph Neural Network-based methods (GNN-MDAs) to address these challenges. However, these methods still face limitations, such as not considering nodes' deep-level similarity associations and hierarchical learning patterns. Additionally, current models do not retain the memory of previously learned heterogeneous historical information about miRNAs or diseases, only focusing on parameter learning without the capability to remember heterogeneous associations.

Results This study introduces the K-means disentangled high-level biological similarity to utilize potential hierarchical relationships fully and proposes a Graph Attention Heterogeneous Biological Memory Network architecture (DiGAMN) with memory capabilities. Extensive experiments were conducted across four datasets, comparing the DiGAMN model and its disentangling method against ten state-of-the-art non-disentangled methods and six traditional GNNs. DiGAMN excelled, achieving AUC scores of 96.35%, 96.10%, 96.01%, and 95.89% on the Data1 to Data4 datasets, respectively, surpassing all other models. These results confirm the superior performance of DiGAMN and its disentangling method. Additionally, various ablation studies were conducted to validate the contributions of different modules within the framework, and's encoding statuses and memory units of DiGAMN were visualized to explore the utility and functionality of its modules. Case studies confirmed the effectiveness of DiGAMN's predictions, identifying several new disease-associated miRNAs.

Conclusions DiGAMN introduces the use of a disentangled biological similarity approach for the first time and successfully constructs a Disentangled Graph Attention Heterogeneous Biological Memory Network model. This network can learn disentangled representations of similarity information and effectively store the potential biological entanglement information of miRNAs and diseases. By integrating disentangled similarity information with a heterogeneous attention memory network, DiGAMN enhances the model's ability to capture and utilize complex underlying biological data, significantly outperforming many existing models. The concepts used in this method also provide new perspectives for predicting miRNAs associated with diseases.

Keywords MiRNA discovery, Disentangled similarity, Heterogeneous biological memory network, MDAs prediction

†Yinbo Liu and Qi Wu contributed equally to this work.

*Correspondence:
Yi Yue
yyyue@ahau.edu.cn
Xiaolei Zhu
xlzhu_mdl@hotmail.com
Full list of author information is available at the end of the article

Introduction

MicroRNAs (miRNAs) are non-coding RNA molecules that regulate gene expression post-transcriptionally, impacting various biological processes [1–3]. Numerous studies have indicated that their dysregulation is associated with various diseases [4, 5]. For example, miR-636 has been shown to regulate cystic fibrosis inflammation [6]. Additionally, the 3p axis of miR-124 has been found to regulate the progression of gastric cancer cells [7]. Therefore, identified miRNA-disease associations (MDAs) can be important biomarkers for disease diagnosis [8, 9]. MDAs also contribute to understanding the underlying pathological mechanisms of various complex diseases [10, 11]. Conventional methods for identifying disease-associated miRNAs typically involve wet lab experiments [6, 7, 12]. However, wet lab validation of unknown MDAs is labor-intensive and limited in scope [12]. Therefore, efficient and cost-effective computational models provide the means for large-scale prediction of potential MDAs [12, 13].

Previous scholars have comprehensively reviewed tasks related to miRNA-disease associations (MDAs). For example, Huang et al. [14] reviewed 29 state-of-the-art models and classified computational models for MDA prediction into categories based on whether they integrated various data sources. They discussed the progress made since 2017 in overcoming the challenges of effective MDA prediction and analyzed the advantages and disadvantages of each model category within the proposed classification. Further, they recommended a feasible evaluation workflow for MDAs tasks based on their analysis results and also discussed the potential challenges encountered in MDAs tasks [15]. Additionally, they summarized recent experimental findings related to common miRNA-associated diseases and provided updates on miRNA-related databases, including new releases since 2017 [16]. Chen et al. [17] discussed in detail the functions of miRNAs, miRNA-target interactions, miRNA-disease associations, and some significant publicly available miRNA-related databases. They reviewed 20 state-of-the-art computational models for predicting miRNA-disease associations from various perspectives, offering many unique insights.

To rapidly identify potential associations between miRNAs and diseases, many scholars have proposed computational methods to predict MDAs [18–20]. Perez-Iratxeta et al. [18] developed an algorithm that prioritizes genes based on their potential associations with genetic diseases in specific chromosomal regions. The algorithm integrates data mining from biomedical databases and gene sequence analysis. Xu et al. [19] proposed a miRNA prioritization approach that leverages the functional similarities between miRNA target genes, derived from matched miRNA and mRNA expression datasets, and known disease genes. Chen et al. [20] developed the Within and Between Score for MiRNA-Disease Association (WBSMDA) model to predict potential miRNAs associated with various complex diseases. WBSMDA integrates known miRNA-disease associations and multiple similarity networks, making it particularly suitable for diseases without any known miRNA associations. These methods generally suffer from low accuracy. To improve accuracy, many studies have proposed similarity-based methods to predict miRNA-disease associations [21–30]. In these methods, miRNAs' sequence and functional information are used to measure the similarity between miRNAs. In contrast, semantic and functional details on diseases are used to measure the similarity between diseases. Chen et al. [21] developed a method using global network similarity measures with a restarted random walk to predict new miRNA-disease associations. Xuan et al. [22] designed a new method that combines a novel miRNA similarity and disease similarity measurement with weighted k-nearest neighbors to explore potential miRNA-disease associations. Shi et al. [23] built a heterogeneous network based on known miRNA-disease association information to identify co-regulated parts of miRNA-disease associations from a protein–protein interaction network perspective, obtaining potential scores for miRNA-disease associations. Furthermore, Lan et al. [24] introduced kernelized Bayesian matrix factorization as a method for identifying miRNA-disease associations. More recently, More recently, Chen et al. [27] combined neighborhood constraints with matrix completion algorithms (NCMCMDA) to assist in predicting unknown associations. Liu et al. [30] integrated multi-channel encoding similarities to predict MDAs. Jiao et al. [12] proposed using a multi-kernel learning (MKL) algorithm to construct comprehensive miRNA and disease similarities.

However, these methods do not consider the deep similarity connections between miRNA and disease nodes or the model's hierarchical learning of biological nodes. Therefore, we introduce disentangled similarity representations of miRNAs and diseases to obtain deep similarity connections and enable the model to adaptively learn node information hierarchically [31–33].

Methods for modeling MDAs based on machine learning and graph deep learning have achieved significant success [34]. Using traditional machine learning methods, Zhao et al. [35] developed an adaptive boosting method that integrates weak classifiers into a robust classifier for reliable predictions. Chen et al. [36] another ensemble learning algorithm was proposed using decision tree ensembles to improve MDA predictive performance. These methods have somewhat promoted MDA

prediction tasks, but they were quickly replaced by deep learning. In deep autoencoding and matrix decomposition, Ji et al. [37] developed a predictive framework named AEMDA by using regression models to extract features in the miRNA-disease space and inputting them into deep autoencoders. Zhou et al. [38] demonstrated the advantages of multi-kernel learning by predicting miRNA-disease associations through deep autoencoders with multiple-kernel learning. They also proposed a bipartite graph-based collaborative matrix factorization method to predict miRNA-disease associations, effectively addressing the problem of data sparsity [39]. Ding et al. [40] enhanced model robustness by predicting miRNA-disease associations based on multi-view variational graph autoencoder and matrix factorization. Li et al. [41] designed a model with neural inductive matrix completion capabilities to identify candidate MDAs. Li et al. [42] proposed a graph autoencoder model for miRNA-disease association prediction, effectively capturing information in graph structures. For multimodal information fusion and deep learning methods, Lou et al. [43] improved the utilization of multimodal data by predicting miRNA-disease associations through learning multimodal networks and fusing mixed neighborhood information. Bai et al. [44] demonstrated the effectiveness of multimodal information fusion by combining miRNA-disease associations and graph convolutional networks to predict miRNA subcellular localization (DAmiRLocGNet). Luo et al. [45] proposed a new model named TDMDA by introducing tensor decomposition-based methods and auxiliary information to integrate multi-type data for MDA identification, effectively predicting MDA tasks.

Based on Graph Convolutional Networks (GCN) and attention mechanisms, Tang et al. [46] proposed the MMGCN model, using a multi-channel attention mechanism to fuse key features extracted from various information sources, enhancing model performance. Zhong et al. [47] demonstrated the integration capabilities of multiple graph models by predicting miRNA-disease associations based on graph random propagation networks and attention networks. Attention mechanisms can automatically assign weights to different nodes, thus improving the results. Using hypergraph networks, some scholars have proposed many models to predict MDAs. For example, Wang et al. [48] demonstrated the potential of hypergraph learning by predicting miRNA-disease associations based on hypergraph learning with high-dimensional features. Liang et al. [49] further improved predictive performance by combining graph and hypergraph convolutional networks to predict miRNA-disease associations. Chang et al. [50] introduced Hyper GCN, using supernodes to construct a miRNA-disease heterogeneous hypergraph, and proposed the HGSMDA method, using GCNs to extract information from different perspectives to predict associations between miRNAs and diseases. Ning et al. [51] achieved higher prediction accuracy by introducing multi-view similarity networks and hypergraph learning to identify miRNA-disease associations (AMHMDA). In summary, these hypergraph methods improve model performance by introducing new supernodes and hyperedges. In heterogeneous graph learning, Zhou et al. [52] proposed a universal prediction method for potential miRNA-disease associations based on heterogeneous graph learning, proving the broad applicability of heterogeneous graph learning. Chen et al. [53] proposed the MDHGI model for miRNA-disease association prediction based on matrix decomposition and heterogeneous graph inference, effectively integrating heterogeneous data sources. Wang et al. [54] developed an MSHGAN-MDA model based on meta-subgraph heterogeneous graph attention networks for miRNA-disease association prediction, further improving predictive performance. Li et al. [55] proposed a new model for miRNA-disease association prediction based on heterogeneous graph convolutional networks, demonstrating the potential applications of heterogeneous graphs. Peng et al. [56] introduced a gene layer to construct a miRNA-gene-disease heterogeneous network. They designed a multi-relational graph convolutional network model (HGCNMDA) to predict miRNA-disease associations from the heterogeneous network. These heterogeneous graph learning methods can learn heterogeneous information about miRNA-disease, enhancing prediction accuracy. Combining graph neural networks and traditional machine learning methods, Yang et al. [57] proposed an effective prediction method based on multiple graph convolutional networks and random forests. Jiao et al. [12] used graph attention-deep autoencoders to learn similarity information between miRNAs and diseases, which can also predict MDA well.

Additionally, based on some unique deep learning methods, Chen et al. [58] proposed a potential miRNA-disease association prediction model based on deep belief networks, proving the applicability of deep belief networks. Ha et al. [59] proposed a node2vec-based neural collaborative filtering method for predicting miRNA-disease associations, demonstrating the advantages of collaborative filtering. Huang et al. [60] demonstrated the powerful functions of tensor decomposition by predicting multiple types of miRNA-disease associations through relationally constrained tensor decomposition.

Despite the significant achievements of these deep learning methods in exploring MDAs, they often overlook the complex historical dependencies between the encoded latent factors of miRNAs and diseases captured

by parameterized memory. Many studies have shown that deep memory networks have significant advantages in specific association prediction tasks [33, 61, 62]. For example, the Social Attentional Memory Network can capture the attention shared by nodes at various aspects [61]. Moreover, memory-augmented message passing can automatically extract latent semantic associations [33]. New memory-augmented message propagation and aggregation schemes in graph neural architectures automatically incorporate semantic associations into the representations of nodes [62]. Our proposed Heterogeneous Memory Network can effectively capture the interdependencies between latent hierarchical biological factors by utilizing aggregated memory in multiple latent representation spaces. Meanwhile, considering the problem of assigning different weights to different nodes, we introduced a graph attention mechanism to solve this issue.

In conclusion, this study constructs the Disentangled Graph Attention Heterogeneous Biological Memory Network (DiGAMN), which uses the K-means method [29] to disentangle the high-order hierarchical similarity information of miRNAs and diseases and then learns the dependencies of latent hierarchical biological factors using a heterogeneous memory network with aggregated graph attention embeddings. DiGAMN effectively learns relevant node information and disentangles similarity information, thus capturing the interdependencies between latent biological factors. This method enhances the model's ability to reconstruct subtle historical relationships, thereby helping to predict MDAs more accurately. We comprehensively compared DiGAMN with various MDA models based on different methods, validated the effectiveness of the DiGAMN model, and conducted a detailed analysis of each part of DiGAMN. Figure 1 details the workflow of DiGAMN.

Materials and methods

miRNA Sequence Similarity

The *Needleman-Wunsch* algorithm [63] is employed to construct the initial miRNA sequence similarity matrix, denoted as R_s. We eliminate the randomization of similarity values through the typical similarity correlation matrix $NR_s(mR_a, mR_b)$ [57].

$$SM(mR_a, mR_b) = \frac{R_s(mR_a, mR_b) - R_s^{min}}{1 - R_s^{min}} \tag{1}$$

where the notation mR represents miRNA. R_s^{min} represent the minimum similarity values in R_s, $SM \in R^{m \times m}$ is represented by the sequence similarity among m different miRNAs, mR_a and mR_b represent two different miRNA sequences, denoted as a and b.

Disease semantic similarity

Disease semantic similarity assesses the similarity between diseases by comparing their semantic relationships. Inspired by Yang et al. [57], disease semantic similarity can be calculated using Directed Acyclic Graphs (DAGs). The DAGs of disease K are represented as D_k, and the semantic block is denoted as m.

$$\begin{cases} D_k(m) = 1 & m = K \\ D_k(m) = max\{\Delta \times D_k(m') | m' \in childrenofm\} & m \neq K \end{cases} \tag{2}$$

$$SD(Dis_a, Dis_b) = \frac{\sum_{m \in D_a \cap D_b} (Dis_a(m) + Dis_b(m))}{\sum_{m \in D_a} Dis_a(m) + \sum_{m \in D_b} Dis_b(m)} \tag{3}$$

In Eq. 2, a semantic attenuation factor Δ is set to 0.5. Equation 3 calculates the similarity between diseases Dis_a and Dis_b. Where n represents the number of diseases, $SD \in R^{n \times n}$ is the semantic similarity matrix.

In addition to sequence and semantic similarity, other researchers also employ Gaussian similarity, functional similarity, and other measures [12, 48–50].

K-means disentangle similarity

After capturing the similarity information of miRNA sequences (SM) and disease semantic similarity information (SD), we use the K-means [64] method to disentangle the similarity information. Initially, K initial similarity cluster centroids are selected, and each sample is assigned to the disentangled category represented by the nearest cluster centroid. Then, the K-means algorithm is trained. The K-means clustering categories and their corresponding miRNA or disease are constructed into a graph (as shown in Fig. 1(a)). The disentangling categories (DimR) of miRNAs are constructed as nodes N_{DimR}, and miRNA is connected to its corresponding category as edges E_{DimR}, N_{mR} represents the node of miRNA:

$$G_{DimR} =< N_{mR}, N_{DimR}, E_{DimR} > \tag{4}$$

Similarly, for Disease, the disentangling categories ($DiDis$) are constructed as nodes N_{DiDis}, and they are connected to their corresponding categories as edges E_{DiDis}, N_{Dis} represents the node of Disease:

$$G_{DiDis} =< N_{Dis}, N_{DiDis}, E_{DiDis} > \tag{5}$$

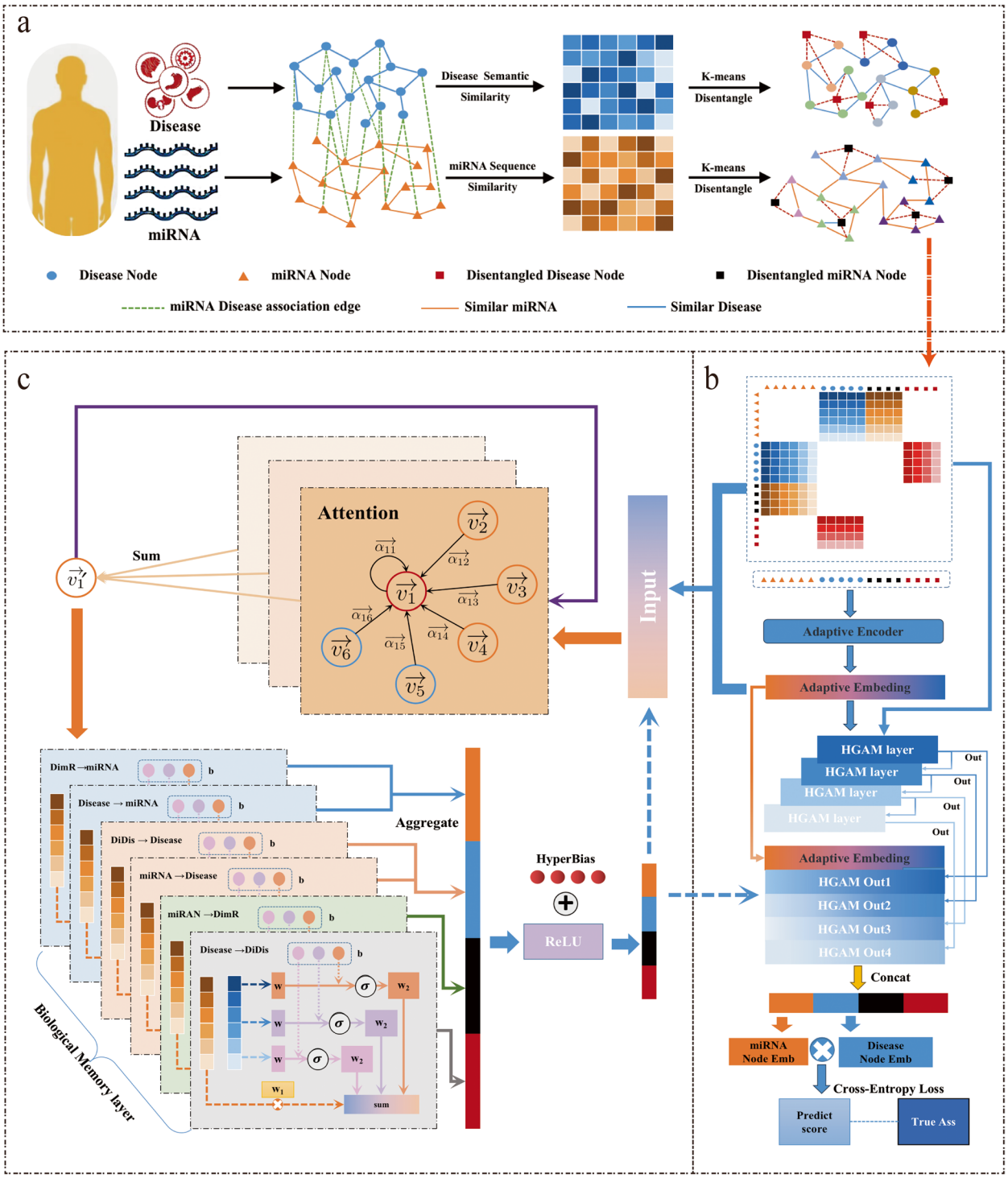

Fig. 1 Flowchart of DiGAMN for MDAs prediction. **a** Process and schematic diagram of data processing and disentangling similarity. **b** Overview of the entire model workflow, including the Adaptive Encoder module, multi-layer GAMN module, and the loss function used for training. **c** The GAMN module comprises graph attention and its Sum Aggregation, memory units for biological information (miRNA-Disease, etc.), and the HyperBias module

Adaptive disentangled encoding layer

The Adaptive Coding Layer is a technique used in neural networks for adaptive feature representation. It aims to learn an efficient feature representation adapted to the input data and task by encoding and decoding it. Traditional neural network models typically use fixed feature representations, which may work well for some functions but may be less effective when faced with different data types or tasks. The introduction of the Adaptive Coding Layer aims to address this issue by automatically learning feature representations of the data. By utilizing an adaptive network to encode the node information of miRNA and Disease, all types of nodes need to be constructed together to form an encodable combination of all nodes:

$$N_{all} = [N_{mR} \cup N_{DiS} \cup N_{DiMR} \cup N_{DiDis}], \tag{6}$$

Additionally, to integrate disentangled information with associative information, it is necessary to construct a comprehensive matrix:

$$A' == \begin{bmatrix} A & E_{DimR} & \\ A^T & & E_{DiDis} \\ E_{DimR}^T & & \\ & E_{DiDis}^T & \end{bmatrix} \tag{7}$$

Then the comprehensive graph G_{all}:

$$G_{all} =< N_{all}, A' > \tag{8}$$

Perform adaptive encoding operations on all nodes:

$$V_{(node)} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ \cdots \\ v_n \end{bmatrix} = W_{N_{all}^T}^h = Adapt\left(N_{all}^T\right) \tag{9}$$

In the above formula, W represents learnable parameters, h represents the length of the autoencoder, which is also the length of v. v_i represents the encoding of the *i-th* node.

Aggregate graph attention layer

For each node in the graph G_{all}, represented as $v_i \in V_{(node)}^L$, the Attention update formula [64] based on the adjacency matrix is as follows:

$$Attention^l(v_i^l) = \sum_{j=1}^{N} \alpha_{ij}^l W^l v_j^l \tag{10}$$

$$\alpha_{ij}^l = \frac{\exp(LeakyReLU(a_l^T \left[W v_i^l || v_j^l \right])}{\sum_{k \in N(i)} \exp(LeakyReLU(a_l^T \left[W v_i^l || v_j^l \right])} \tag{11}$$

In the above formula, α is a learnable weight vector, *LeakyReLU* is a rectified linear unit activation function with a negative slope, "||" denotes vector concatenation operation, W is a learnable weight matrix, l denotes the layer of aggregated attention, v_i^l represents the encoding for the node, where $v_i^1 \in V_{(node)}$.

Subsequently, a multi-head graph Attention is aggregated by summation operation:

$$v_{att,i}^{l+1} = \sum_{h=1}^{H} ReLU(Attention^{(l,h)}\left(v_i^l\right)) \tag{12}$$

$$v_{att,i}^{l+1} = \sum_{h=1}^{H} ReLU(\sum_{j=1}^{N} \alpha_{ij}^{(l,h)} \cdot W^{(l,h)} v_i^l) \tag{13}$$

$$V_{all}^{Att} = v_{att,1}^{-1}, v_{att,2}^{-1}, v_{att,3}^{-1}, \ldots, v_{att,i}^{-1}, \ldots, v_{att,n}^{-1}] \tag{14}$$

where $l = -1$, which indicates the last layer of aggregated Attention.

Heterogeneous biological memory network layer

The memory network layer can store and retrieve past information, enabling the network to handle association data and long-term dependencies [33, 61, 62] better. The computation formula for the Heterogeneous Memory Network is as follows:

$$\nabla(v_t, \mathrm{m}) = LeakyReLU(v_t \cdot W_{\nabla,m} + b_m) \tag{15}$$

$$\}_{t \leftarrow s}(v_t, v_s, \mathcal{M}) = (\sum_{m=1}^{|\mathcal{M}|} \nabla(v_t, \mathrm{m}) W_{2,m}) \cdot (v_s W_{1,m}) \tag{16}$$

where, t stands for the target node of edges in the biological association network, s represents the source node, and ∇ denotes the encoding of source nodes on the *m-th* memory unit. $\}_{t \leftarrow s}$ indicates the function of aggregating memory from source to target. $\mathcal{M}$ represents the number of memory units, where m denotes the memory layer, m belongs to the set {1, 2, ..., $\mathcal{M}$}, where $W_{\nabla,m} \in R^{d \times d}$, $W_{1,m} \in R^{d \times d}$, $W_{2,m} \in R^d$, $b_m \in R$. $W_{1,m}$ and $W_{2,m}$ represent two corresponding learnable parameter matrices, b_m represents the bias corresponding

to each memory unit. The aggregate biological memory encoding for miRNA and disease is given by:

$$V_x = \frac{\sum_{x \neq z} \}_{x \leftarrow z}\left(V_x, V_z^{Att}, \mathcal{M}_x^e\right)}{\sum_{x \neq z} \mathcal{N}_{x \leftarrow z}} \tag{17}$$

$$\begin{cases} x \in \{miRNA, Disease\} \\ z \in \{mRNA, Disease, DimR, DiDis\} \end{cases} \tag{18}$$

In the above formula, $\mathcal{M}_x^e$ represents the type of edge corresponding to x, and $\mathcal{N}$ represents the number of associated neighboring objects. $\mathcal{N}_{x \leftarrow z}$ represents the number of neighboring objects associated with $x \leftarrow z$. V represents the encoding matrix.

For *DimR* and *DiDis*, the memory encoding is as follows:

$$V_x = \frac{\}_{x \leftarrow z}\left(V_x, V_z^{Att}, \mathcal{M}_x^e\right)}{\mathcal{N}_{x \leftarrow z}} \tag{19}$$

$$\begin{cases} x \in \{DimR, DiDis, \} \\ z \in \{mRNA, Disease\} \end{cases} \tag{20}$$

Finally, adding bias to the model to obtain the final encoding:

$$v_{\mathrm{i}} = \max(0, \mathrm{v_i}) + \mathrm{HB_i},\ \mathrm{v_i} \in \mathrm{V_x}, \tag{21}$$

where, v represents the encoding vector, HB_i represents the HyperBiase. The *max* function means that if v_i is greater than 0, the result is v_i; otherwise, the result is 0.

The Attention memory referred to in the above process is called *GAMN* (as shown in Fig. 1(c)). The calculation formula for the ℓ-th layer of *GAMN* is expressed as:

$$V_{(\mathrm{node})}^{\ell+1} = \mathrm{GAMN}^{\ell}\left(\mathrm{V}_{(\mathrm{node})}^{\ell}\right) \tag{22}$$

Multi-layer modules and decoding miRNA–disease associations

By performing concatenation operations, we can effectively retain the richness of information across different layers:

$$V_{(node)}^{Out} = V_{(node)}^{1} || [||]_{\ell=1}^{L} GAMN^{\ell}(V_{(node)}^{\ell})] \tag{23}$$

$$[V_{mR}^{Out} \bigcup V_{Dis}^{Out} \bigcup V_{DimR}^{Out} \bigcup V_{DiDis}^{Out}]^{T} \leftarrow V_{(node)}^{Out} \tag{24}$$

In the above equation, "||" denotes the concatenation operation, and L represents the maximum number of layers in GAMN.

Using these embeddings, we compute the scores of the association matrix as shown below formula:

$$A_{\mathrm{score}} = \frac{\mathrm{V_{mR}^{Out}} \times \mathrm{V_{Dis}^{Out\,T}}}{\mathrm{d(v_i^{Out})}}, \mathrm{v_i^{Out}} \in \mathrm{V^{out}}. \tag{25}$$

In Eq. 25, $d(v_i^{Out})$ represents the dimensionality after concatenation, V_{mR}^{Out} and V_{Dis}^{Out} respectively represent the node embeddings of the final encoding, A_{score} represents the predicted miRNA-disease association score matrix.

Optimization objective

In this model, we utilize the Cross-Entropy Loss [65] function as the model's loss function. The Cross-Entropy Loss function is a commonly used loss function that measures the difference between the predicted probability distribution of the model and the true labels.

$$\mathcal{L}_{<mR,Dis>} = -\frac{1}{n}\left(\sum_{A(i,j)=1} \log A_{score}(i,j) + \sum_{A(i,j)=0} \log(1 - \log A_{score}(i,j)\right) \tag{26}$$

where $\mathcal{L}_{<mR,Dis>}$ represents the loss function for miRNA and Disease, where n denotes the total number of samples. $A(i,j) = 1$ indicates an association between miRNA and Disease for samples i and j. $A(i,j) = 0$ indicates that there is no association between sample i and j in terms of miRNA and Disease. Ultimately, the parameters of our model are learned using the Adam optimizer [66].

Evaluation metrics

We utilized a comprehensive set of evaluation metrics, including AUC, AUPRC, accuracy, precision, recall, F1-score, and specificity, to evaluate the performance of our model.

$$Accuary = \frac{TP + TN}{TP + TN + FP + FN} \tag{27}$$

$$Precision = \frac{TP}{TP + FP} \tag{28}$$

$$Recall = \frac{TP}{TP + FN} \tag{29}$$

$$F1_{Score} = 2 \times \frac{Precision \bullet Recall}{Precision + Recall} \tag{30}$$

$$Specificity = \frac{TN}{TN + FP} \tag{31}$$

In the above equation, TP, TN, FP, and FN represent different elements of a binary classification model's performance: TP (True Positives): The number of correctly predicted positive instances (i.e., instances that are positive and are predicted as positive). TN (True Negatives): The number of correctly predicted negative instances (i.e., instances that are negative and are predicted as negative). FP (False Positives): The number of instances that are predicted as positive but are negative (i.e., instances that are negative but are predicted as positive). FN (False Negatives): The number of instances that are predicted as negative but are positive (i.e., instances that are positive but are predicted as negative).

The Receiver Operating Characteristic (ROC) curve plots the True Positive Rate (TPR) on the y-axis against the False Positive Rate (FPR) on the x-axis. The Area Under the Curve (AUC) is beneath the ROC curve. A larger AUC indicates a better model performance, representing a larger area under the ROC curve.

$$TPR = \frac{TP}{TP + FN} \tag{32}$$

$$FPR = \frac{FP}{FP + TN} \tag{33}$$

The Precision-Recall Curve (PRC) describes the relationship between precision and recall. In the PRC plot, classifiers represented by curves closer to the upper-right corner indicate better performance, as higher precision and recall values signify better model effectiveness. We also used area under the PRC (AUPRC) as an evaluation indicator.

Experiments and results

Data collection and preparation

In this study, we obtained four different miRNA-disease association datasets, named Data1, Data2, Data3, and Data4, by integrating the work of Dai et al. [34], Bai et al. [44], Ning et al. [51] and Ding et al. [40]. The datasets utilized the HMDD v3.0 or v3.2 database [67], and the data obtained after removing duplicate entities from databases such as mRBase [44] are presented in Table 1. Additionally, we collected associated data from the HMDD v4.0 database [68] as supplementary validation data. The entire database is scheduled to be released in June 2023.

Figure 2(b) illustrates the local association relationships of miRNA-disease, while Fig. 3(b) compares the actual association relationships and the predicted scores across the four datasets. The visualization of the top 200 diseases associated with the number of miRNAs is shown in Supplementary Figure S1 across the four datasets. Among the four datasets, the number of miRNAs associated with diseases does not exceed 98 within the 95% range. According to the dataset statistics, 472 diseases (19.69%) were associated with only one miRNA across all datasets, indicating a high specificity of these miRNAs to their corresponding diseases. Figure 2(c) illustrates the local-specificity network in Data1. Figure 2(d) describes the information entropy of miRNA molecules at different nucleotide positions. As shown in Fig. 2(f), the disentangled K-means method takes a similarity matrix (Fig. 2(e)) of nodes as input and initializes the number of clusters. Figure 2(g) displays the visualization of the disentangled node results on the four datasets.

Comparison with Other SOTA Methods

To validate the comprehensive performance of the DiGAMN model, we conducted a fivefold cross-validation in this section. We compared the performance of

Table 1 Details of the 4 collected datasets. "*# of miRNAs*" indicates the number of miRNA, while "*# of Diseases*" indicates the number of Diseases

Dataset	Data1	Data2	Data3	Data4
# of miRNAs	917	1041	853	788
# of Diseases	792	640	591	374
miRNA-Disease Association	14,550	15,547	12,446	8968
Association Density Degree	0.0200	0.0233	0.0247	0.0304

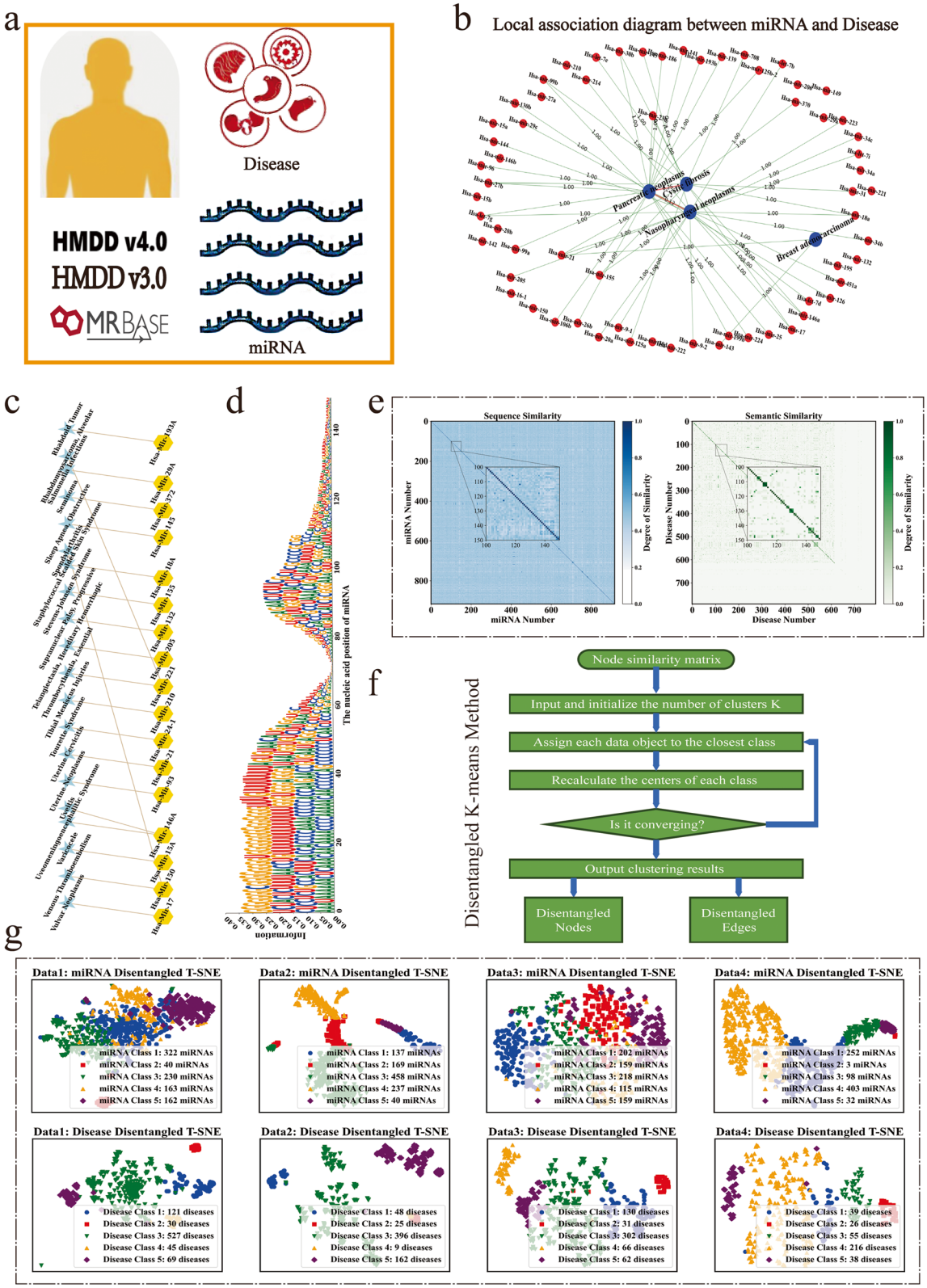

Fig. 2 Data Collection and Preparation displays: (**a**) Database sources used in this study. **b** Local association graph of disease-related miRNAs on Data1. **c** Partially specific diseases and their corresponding miRNA connections in Data1. **d** Display of sequence distribution information for all miRNAs in the data1. **e** Similarity of miRNA sequences and semantic similarity of diseases in Data1 (Data2, Data3, and Data4 are presented in Supplementary Figure S2. **f** Process of Disentangled K-means Method. **g** T-SNE [69] visualization of disentangled miRNA and disease similarities

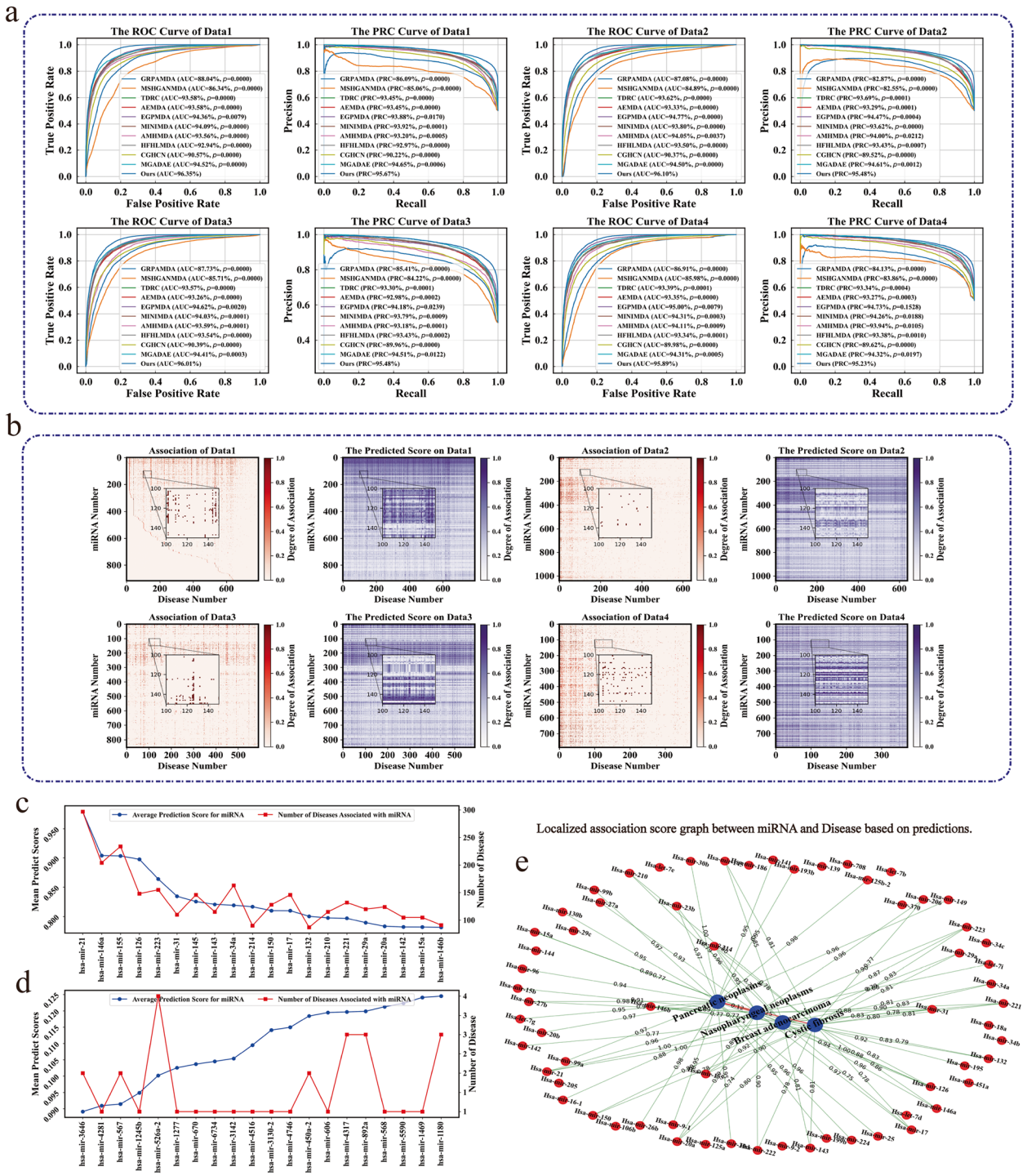

Fig. 3 The performance of DiGAMN and the presentation of its prediction results. **a** fivefold cross-validation ROC curves and PRC curves of DiGAMN method on 10 SOTA methods across four datasets. The *p*-values in the figure represent the t-test results for the differences between our method and the method from the fivefold cross-validation. More detailed experimental evaluation metric results can be found in Supplementary Table S3. **b** Visualization of scores predicted by the DiGAMN method against true miRNA-disease associations. **c** Top 20 average prediction scores of a miRNA and the number of its true associated diseases. **d** A miRNA's bottom 20 average prediction scores are the number of its true associated diseases. **e** Local association graph of predicted miRNA-disease associations in Data1

DiGAMN in miRNA-disease associations (MDA) with ten current state-of-the-art methods. The ten methods used for comparison were MINIMDA [43], TDRC [60], AEMDA [37], GRPAMDA [47], HFHLMDA [48], CGHCN [49], AMHMDA [51], EGPMDA [52], MSHGANMDA [54] and MGADAE [12]. These ten methods are representative in the current field and each possesses unique characteristics. They cover different computational frameworks, including similarity-based, tensor decomposition, deep autoencoders, graph neural networks, and heterogeneous graph networks. The details are as follows:

Similarity-based methods:

(i) Integrated Similarity Network ***MINIMDA*****:** Combines multiple data sources to construct similarity networks, using mixed neighborhood information and MLP models to predict miRNA-disease associations.

Tensor decomposition methods:

(ii) Tensor Decomposition Approach ***TDRC*****:** Utilizes biological feature constraints in tensor decomposition to enhance efficiency and accuracy in predicting miRNA-disease associations.

Combined applications of deep autoencoder methods:

(iii) Matrix Decomposition and Deep Autoencoder Integration ***AEMDA*****:** Integrates disease semantic similarity and miRNA functional similarity, using deep autoencoders for association prediction.

Graph neural network methods combined application:

(iv) Combination of Attention Mechanism and Deep Learning ***GRPAMDA*****:** Integrates graph random propagation and attention mechanisms, predicting miRNA-disease associations through a fully connected layer.

(v) Combination of High-dimensional Features and Hypergraph Learning ***HFHLMDA*****:** Uses high-dimensional features and hypergraph learning techniques to enhance the prediction performance of miRNA-disease associations significantly.

(vi) Hypergraph Convolutional Network Learning ***CGHCN*****:** Merges graph convolution and hypergraph convolution techniques, reducing overfitting and extracting effective information.

(vii) Hypergraph and Graph Attention Network Combination ***AMHMDA*****:** Combines GCN and attention mechanisms to effectively aggregate information from different similarity perspectives, enhancing prediction accuracy.

Combined applications of heterogeneous graph network methods:

(viii) Multimodal Heterogeneous Graph ***EGPMDA*****:** Integrates multiple data sources using a graph transformer model to optimize the description of miRNA-disease associations**.**

(ix) Multisubgraph and Heterogeneous Graph Attention Network ***MSHGANMDA*****:** Defines meta-subgraphs, extracting features through meta-subgraph attention techniques, and trains the model end-to-end using a fully connected layer and cross-entropy loss.

(x) Multikernel Graph Attention and Heterogeneous Graph Network Integration **MGADAE**: Combines multikernel graph attention and deep autoencoders, deeply learning and decoding miRNA and disease features through graph convolution and attention mechanisms.

In this section, we considered miRNA-disease pairs with recorded associations as positive samples. In the experiments on balanced datasets, we randomly selected an equal number of miRNA-disease pairs with no recorded associations as negative samples to ensure sample balance. We masked 20% of the association information during cross-validation in the test set to prevent information leakage. The comparative results of the experiments are shown in Fig. 3(a). The results indicate that our model performed well regarding AUC and PRC across all datasets. From the analysis of Data1 to Data4, our model achieved AUC scores of 96.35%, 96.10%, 96.01%, and 95.89%, respectively, surpassing all ten other models. This demonstrates an essential advantage with an average improvement of 2 percentage points over the highest results from other models. The t-test results were all less than 0.01, indicating highly significant results. Figure 3(b) shows the relationship and distribution between real associations and predicted scores for each dataset, indicating certain similarities. Figures 3(c) and 3(d) present the average predicted scores and the trend of association counts for the top 20 miRNAs with the highest and lowest association numbers in Data1. Figure 3(e) illustrates the local association network of prediction results and scores, clearly showing the predicted scores of miRNAs associated with each disease.

Furthermore, considering the imbalance in the dataset where unknown miRNA-disease associations significantly outnumber known associations, we selected all known association edges in the dataset as positive samples and chose N times the number of unknown association edges as negative samples to create an imbalanced dataset. A fivefold cross-validation was conducted to verify the robustness of the experimental results. During model training, 20% of the positive samples and 20% of the negative samples were randomly selected as the

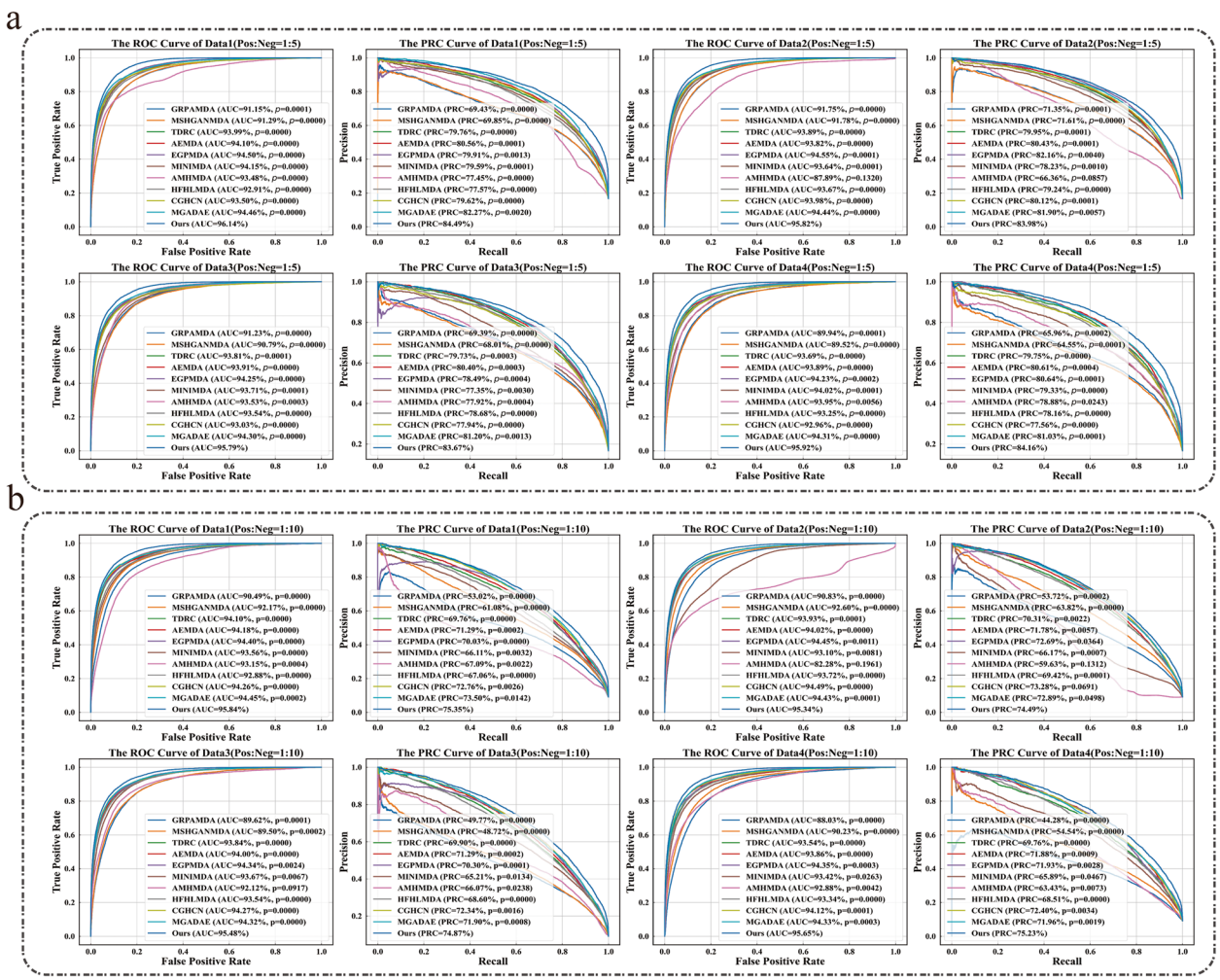

Fig. 4 Performance of DiGAMN on imbalanced datasets. **a** ROC and PRC curves when the positive-to-negative sample ratio is 1:5. **b** ROC and PRC curves when the positive-to-negative sample ratio is 1:10. More detailed experimental evaluation metric results can be found in Supplementary Table S4. The *p*-values in the figure represent the t-test results for the differences between our method and the method from the fivefold cross-validation

test set, while the remaining data were used as the training set. Figure 4 illustrates the performance of DiGAMN and ten non-disentangled similarity SOTA methods on imbalanced datasets with positive-to-negative sample ratios of 1:5 (Fig. 4(a)) and 1:10 (Fig. 4(b)). On the dataset with a positive-to-negative sample ratio of 1:5, the DiGAMN method achieved AUC values of 96.14%, 95.82%, 95.79%, and 95.92% across four datasets, with PRC values of 84.49%, 83.98%, 83.67%, and 84.16%, respectively. On the dataset with a positive-to-negative sample ratio of 1:10, the AUC values were 95.84%, 95.34%, 95.48%, and 95.65%, with PRC values of 75.35%, 74.49%, 74.87%, and 75.23%. The DiGAMN model achieved the highest AUC and PRC values across all datasets compared to ten other methods. Compared to the other ten methods, the *p*-values from the t-tests between the DiGAMN model and the other models were all less than 0.01, indicating that the results are significantly better than those of the other models, further demonstrating its predictive performance on imbalanced datasets.

Case study

Based on our model, we conducted a case analysis using Data1, randomly selecting four diseases with varying degrees of associated miRNA abundance. These diseases are cystic fibrosis, pancreatic neoplasms, breast adenocarcinoma, and nasopharyngeal neoplasms. In Data1, the numbers of associated miRNAs for cystic fibrosis, nasopharyngeal neoplasms, breast adenocarcinoma, and pancreatic neoplasms are 14, 53, 2, and 175, respectively. We selected 50% of the associated miRNA data for each disease to be masked while retaining the remaining 50% of associated data and other associated data as the training set. The top 10 miRNAs, excluding

Table 2 Case analysis table for four diseases. The *Scores* represent the predictions made by the DiGAMN model. *Data 1* indicates whether the miRNA appeared in the Data1 dataset, where *1* represents appearing, and *0* represents not appearing. "*New e.d.*" indicates whether the top miRNA has been validated by the HMDD v4.0 (H v4) database or other recent studies. H v4', which stands for highly homologous RNA has been confirmed

Cystic Fibrosis				**Nasopharyngeal Neoplasms**			
RNA names	**Scores**	**Data1**	**New e.d**	**RNA names**	**Scores**	**Data1**	**New e.d**
hsa-mir-223	0.8318	0	H v4'	hsa-mir-126	1.000	0	
hsa-mir-34a	0.8312	0	H v4	hsa-mir-206	0.9628	0	H v4
hsa-mir-221	0.8148	0		hsa-mir-23b	0.9599	0	Ref [75]
hsa-mir-143	0.8114	0	H v4	hsa-mir-214	0.9548	1	H v4
hsa-mir-29a	0.8108	0		hsa-mir-27a	0.9306	0	
hsa-mir-145	0.8065	0	H v4	hsa-mir-195	0.9180	0	
hsa-mir-31	0.8042	1	H v4	hsa-mir-15b	0.9068	0	
hsa-mir-20a	0.7956	0		hsa-mir-222	0.9007	0	
hsa-mir-210	0.7955	0		hsa-mir-20a	0.8879	1	H v4
hsa-mir-132	0.7868	0		hsa-mir-146b	0.8858	1	Ref [76]
Breast Adenocarcinoma				**Pancreatic Neoplasms**			
RNA names	Scores	Data1	New e.d	RNA names	Scores	Data1	New e.d
hsa-mir-146a	0.9582	0		hsa-mir-99b	0.9738	1	H v4
hsa-mir-155	0.9169	0		hsa-mir-20b	0.9688	0	H v4'
hsa-mir-126	0.8757	0		hsa-mir-23b	0.9686	1	H v4
hsa-mir-223	0.8671	0		hsa-mir-25	0.9620	1	H v4
hsa-mir-145	0.8528	0		hsa-mir-149	0.9565	0	
hsa-mir-31	0.8334	0		hsa-mir-141	0.9549	1	H v4
hsa-mir-132	0.8267	0		hsa-mir-30d	0.9474	0	H v4'
hsa-mir-143	0.8142	1		hsa-mir-9–1	0.9469	0	
hsa-mir-34a	0.8132	0		hsa-mir-199b	0.9457	1	H v4
hsa-mir-150	0.8050	0		hsa-mir-130b	0.9455	1	H v4

the ones involved in training, are provided in Table 2. The partial correlation graph of the prediction results is shown in Fig. 3(e).

The prediction of miRNA associations with cystic fibrosis is a compelling example demonstrating the effectiveness of the DiGAMN model. In the top 10 predictions, three novel associated genes, hsa-mir-223 [70], hsa-mir-143 [71], and hsa-mir-145 [72], were discovered by DiGAMN. Research has shown that the function miR-223 is related to inflammation and immune regulation, which are implicated in cystic fibrosis [70]. hsa-mir-143-5p has a moderate but significant inhibitory effect on the Cystic Fibrosis Transmembrane Conductance Regulator (CFTR), which may impact the development and progression of cystic fibrosis [71]. miR-145-5p enhances the expression of CFTR in Calu-3 cells, thereby playing a regulatory role in cystic fibrosis [72].

Additionally, DiGAMN identified associations between nasopharyngeal neoplasms and hsa-mir-206 [73, 74], hsa-miR-23b [75], and hsa-mir-146b [76]. hsa-mir-206 may exert its tumor-suppressive effect by regulating specific signaling pathways, targeting essential genes, or modulating cell proliferation and metastasis mechanisms [73]. The pronounced tumor-inhibiting effect of hsa-mir-206 in C666-1 cells further suggests its potential as a new therapeutic approach for differentiating nasopharyngeal carcinoma [74]. A recent study in 2024 revealed that hsa-miR-23b-3p, a miRNA not included in HMDD v4.0, may regulate the expression of NACC1 in nasopharyngeal cancer cells [75]. Additionally, this miRNA was found to impact biological processes such as cell proliferation and apoptosis.

For breast adenocarcinoma, another miRNA gene, hsa-mir-143, was successfully predicted from Data1. Unfortunately, breast adenocarcinoma has been removed from HMDD v4.0, and there is limited literature available on this specific disease. In the top 10 predictions for pancreatic neoplasms, most of the miRNAs have already been considered associated with the disease by Data1.

Comparative analysis with traditional graph neural network methods

To individually compare the effectiveness and capability of disentangled similarity, in this section, we conducted comparative experiments among six baseline

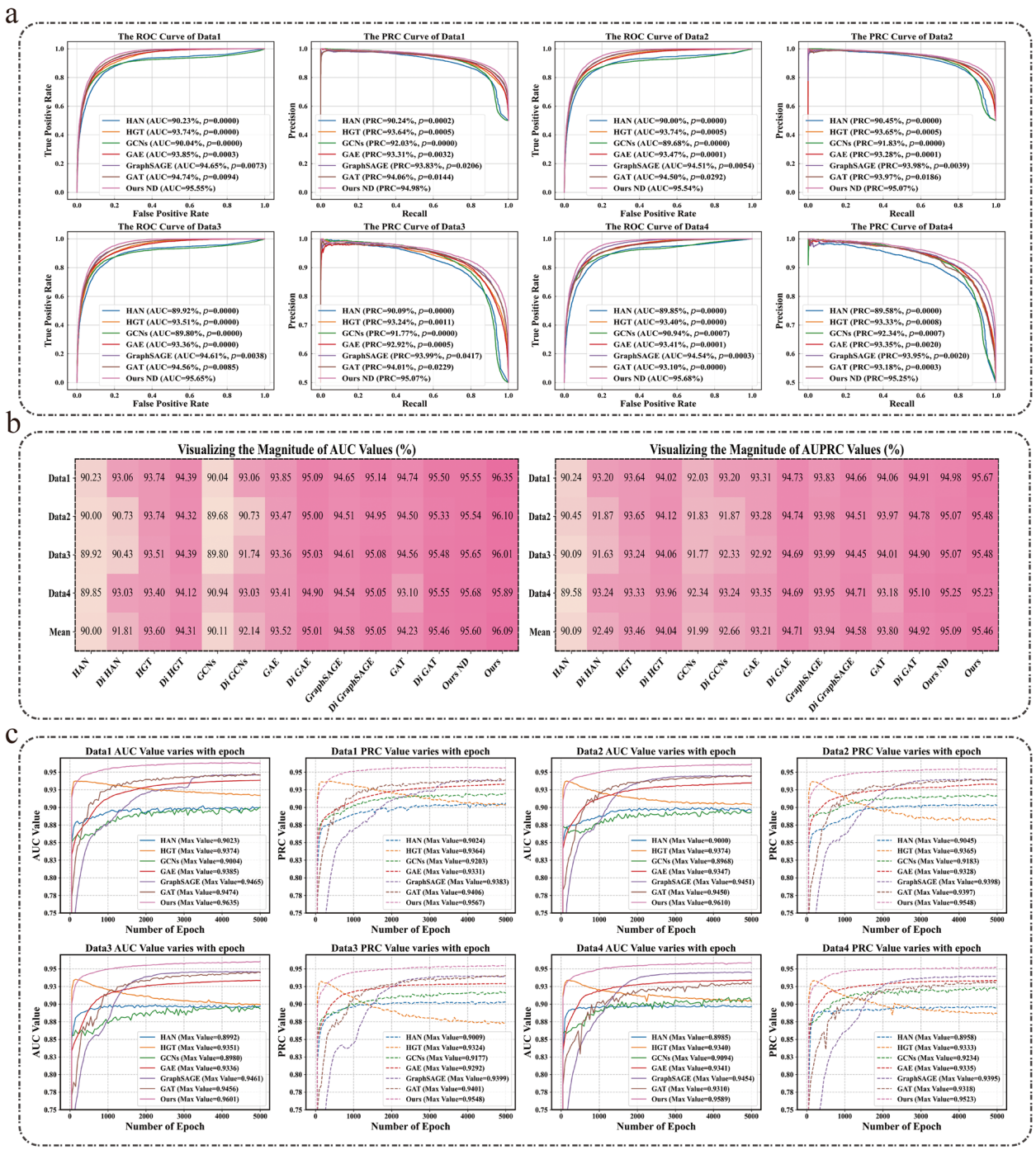

Fig. 5 Comparative analysis with baseline GNNs. **a** ROC and PRC curves of DiGAMN and six baseline GNNs. The *p*-values in the figure represent the t-test results for the differences between our method and the method from the fivefold cross-validation. **b** Comparison with six baseline GNNs on four datasets, where Di* denotes the disentangling similarity method, Ours ND denotes non-disentangled DiGAMN. **c** The trend of AUC and PRC values with epochs for our method and baseline GNNs. The specific data about this section is shown in Supplementary Table S7

Graph Neural Network methods (GNNs), including HGT [77], HAN [78], GCNs [79], GAE [80], GraphSAGE [81], and GAT [64]. These methods uniformly utilized fused miRNA sequence similarity and Disease semantic similarity, along with their Gaussian similarity [40], as encoding inputs. Figure 5(b) shows that our model outperformed the baseline GNNs on all datasets. Additionally, among all models, the disentangled

versions of the baseline GNNs consistently exhibited higher AUC and AUPRC values, indicating that disentangling similarity is an effective approach to improve model performance. The non-disentangled DiGAMN method also achieved high AUC and AUPRC values, surpassing even the disentangled versions of the baseline GNNs. This indicates that the non-disentangled DiGAMN model itself has high utility. The AUC and PRC values of DiGAMN exhibit a growth rate second only to HGT over epochs, surpassing the other five baseline GNNs (Fig. 5(c)).

Parameter sensitivity analysis

To validate the impact of different hyperparameters on the performance of the DiGAMN model, we conducted a parameter analysis of the DiGAMN model's attention layer count, hidden channel count, memory unit size, and hidden layer count. The analysis results are shown in Fig. 6, where (a)-(d) respectively present the parameter sensitivity for the Data1-Data4 datasets.

Exploration of model encoding embeddings status

To better explore the effectiveness of the adaptive encoder layer and GAMN model in learning feature representations and discovering associations in the data, we conducted a visual analysis (Fig. 7(a), (b)). In this section, we configured 4 GAMN hidden layers, 4 graph attention layers, 8 hidden channels, and 8 memory units. Figures 7(a) and 4(b) demonstrate that *Adapt Encoding* allows the model to capture the features of miRNA and Disease more comprehensively, resulting in richer results in feature representation.

Figure 7(c) provides a visualization analysis of the GAMN layers, examining the richness of information extraction for miRNA and Disease at different levels (Adaptive Encoder and 4 GAMN layers). In DiGAMN, each layer learns different levels of abstraction from the data. Layers 1–3 perform better in learning rich representations. Regarding the complex relationship between miRNA and Disease, each layer of the DiGAMN model tends to have different learning tendencies. Therefore,

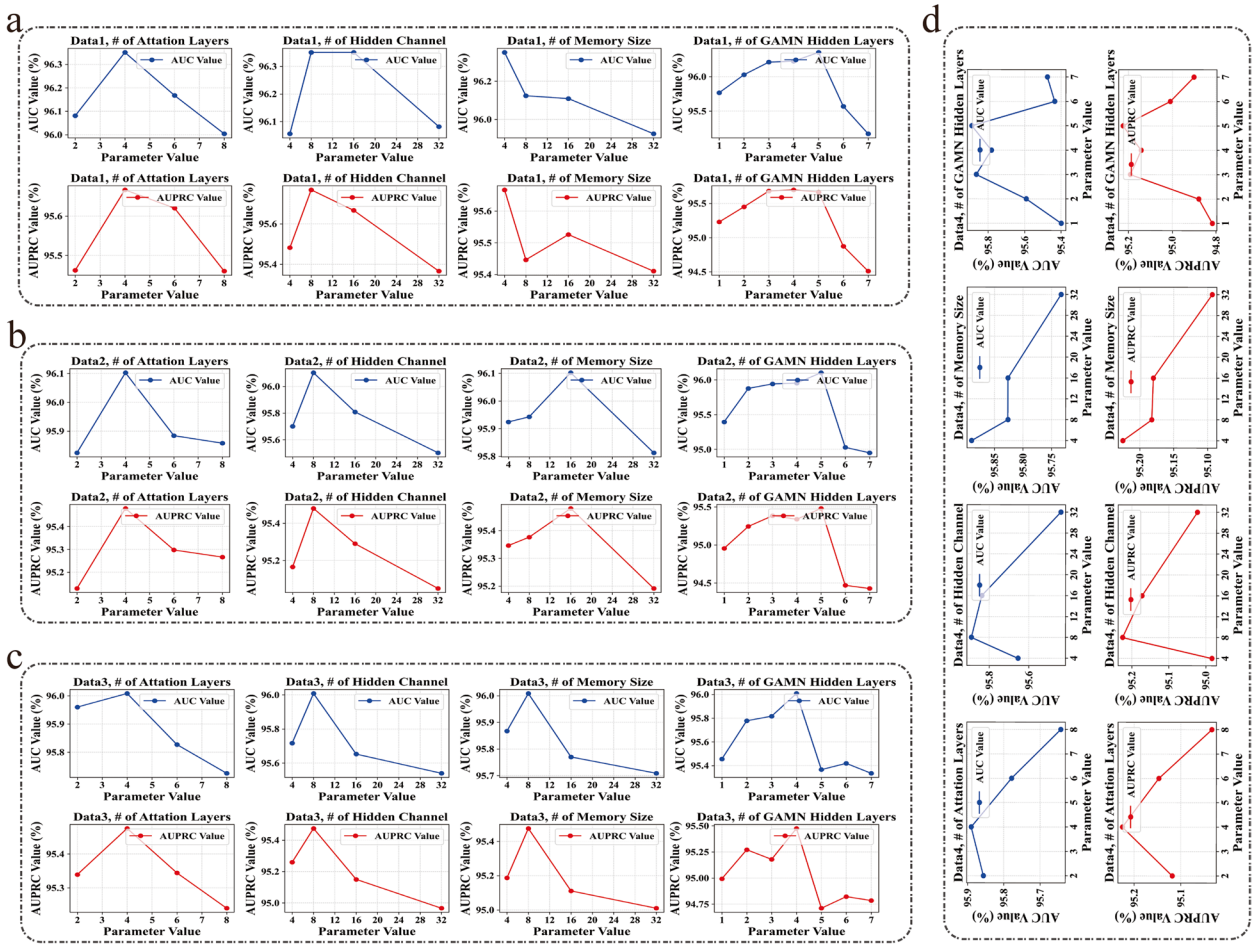

Fig. 6 Sensitivity analysis of hyperparameters, where (**a**-**d**) respectively represent the AUC or PRC values corresponding to different hyperparameters on Data1-Data4. The optimal parameters for the four datasets can be found in Supplementary Table 2

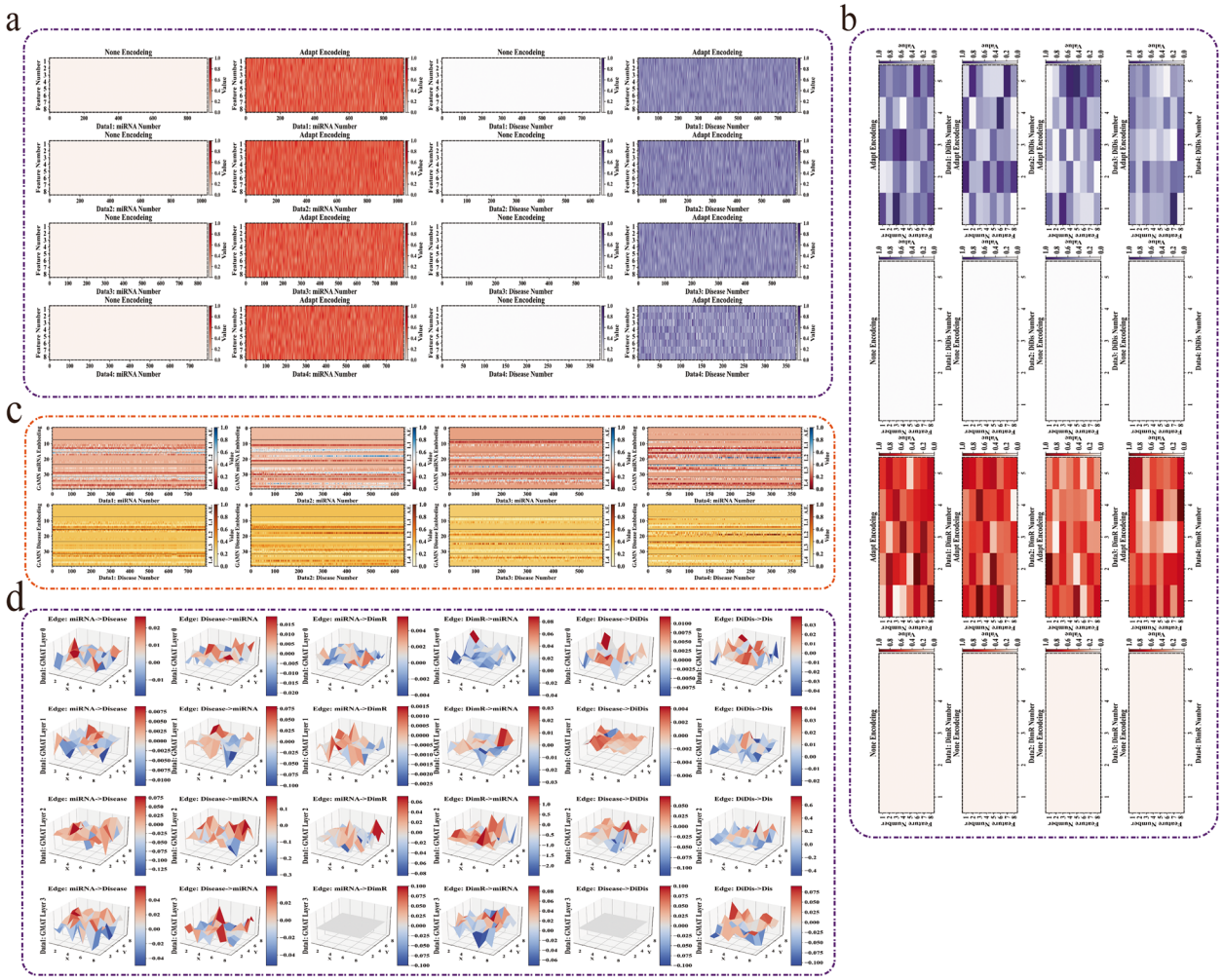

Fig. 7 Exploration of Model Encoding Embeddings. **a** Visualization of adaptive encoding features on miRNA nodes and Disease nodes. **b** Visualization of adaptive encoding features on DimR nodes and DiDis nodes. **c** Visualization of multi-layer features on multi-layer DiGAMT, where A.E. represents the adaptive encoding layer, and L.1–4 represents four GAMT layers. **d** Investigation of memory layers in Data1, where each subgraph represents an 8*8 memory unit, and each layer includes edges aggregated from all directions. The visualization results of the memory layers for the other three datasets can be found in Supplementary Figure S4-S6

concatenating all embedded data can preserve the richness of the information.

Exploration of memory network functionality

To investigate the effectiveness of memory layers in the GAMN model for handling long sequences and tasks with sequential characteristics, as well as to improve the stability and performance of the model during the training and inference stages, the Data1 dataset was analyzed. The visualization results are shown in Fig. 7(d). Based on the analysis, it can be observed that in the first three layers, there are significant fluctuations in the associations between *miRNA* → *Disease*, *Disease* → *miRNA*, *miRNA* → *DimR*, *DimR* → *miRNA*, *Disease* → *DiDis* and *DiDis* → *Disease*. These fluctuations reflect substantial parameter adjustments made by the model during the early stages of training, indicating that the memory layers retain and propagate information about these association relationships. In contrast, in the fourth layer, there are no fluctuations between *miRNA* → *DimR* and *Disease* → *DiDis*, suggesting that these relationships are no longer involved in the subsequent decoding operations of the model. Therefore, the model relies mainly on the memory capacity provided by the intermediate layers to retain disentangled node information.

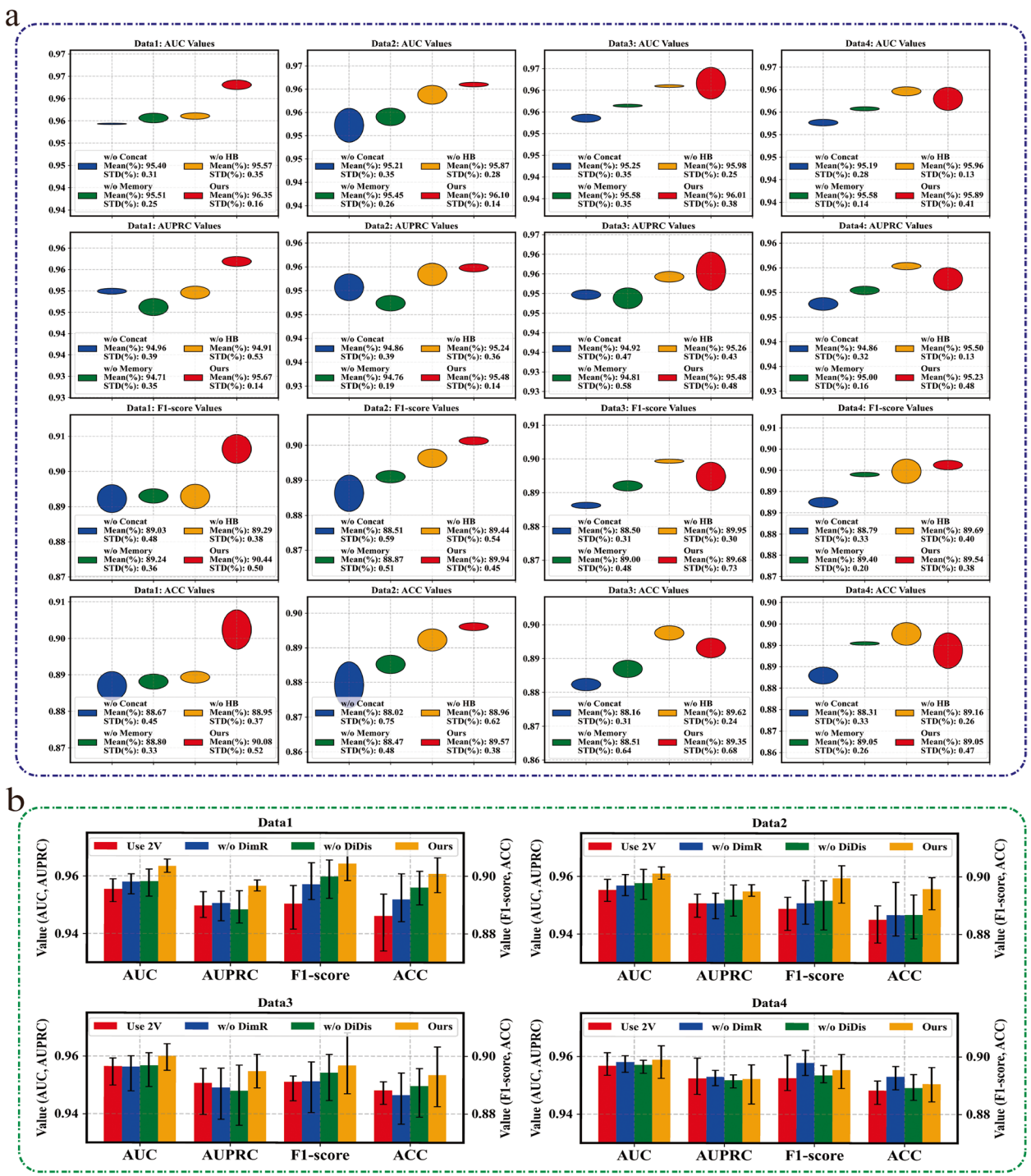

Fig. 8 Two types of ablation experiments. **a** Ablation experiments on different modules of DiGAMN. The higher the position of the circle, the larger the overall value and the better the model's performance. The flatter the circle, the smaller the range of data distribution and the higher the stability of the model. "Mean (%)" represents the mean, and "STD (%)" represents the standard deviation. **b** Ablation experiments on disentangling information of DiGAMN. The higher the bar chart, the larger the mean value. The shorter the line on the bar chart, the smaller the range of data distribution

Ablation experiments

We conducted ablation experiments to investigate the roles of different components in our designed GAMN model. We created three variants for the GAMN component: "w/o Concat" removed the concatenation of multiple GAMN operations, "w/o Memory" removed the memory layer, and "w/o HB" removed the HyperBias layer to examine the influence of bias on data distribution adjustment.

Figure 8(a) presents the experimental results, showing that when the concatenation layer and memory layer, which we focused on, were excluded, the main evaluation metrics AUC and AUPRC decreased, indicating the effectiveness of these components. The detailed data regarding ablation experiments is shown in Supplementary Table S9. However, some evaluation metrics of specific datasets improved when the "HB" module was removed, suggesting that HyperBias has certain limitations.

In addition, we created three variants of DiGAMN to explore the functionality of disentangling miRNA similarity and disease similarity: "Use 2 V" did not disentangle miRNA and diseases and instead utilized miRNA sequence similarity, disease semantic similarity and Gaussian similarity directly; "w/o DimR" only disentangled diseases without disentangling miRNA, and "w/o DiDis" only disentangled miRNA without disentangling diseases.

Figure 8(b) shows that our method outperformed other variants regarding the main evaluation metric, AUC, across four datasets. This demonstrates the improved performance of disentangling similarity in miRNA-disease association prediction tasks. The detailed data regarding ablation experiments is shown in Supplementary Table 10.

Conclusion

miRNAs have been proven to be closely related to the occurrence of complex human diseases and cancers. Many scholars have proposed various computational methods to address the limitations of traditional experimental validation methods in MDAs, such as time-consuming issues. However, these methods ignored the potential characteristics of high-order hierarchical node associations in disentangling the similarity of miRNAs or diseases. They did not focus on preserving the memory of potential heterogeneous historical biological associations. Therefore, this study proposes a Disentangling similarity Graph Attention Heterogeneous Biological Memory Network (DiGAMN) model based on disentangled miRNA similarity and disease similarity data for MDAs prediction to explore the importance of disentangling high-order hierarchical nodes and heterogeneous memory.

In DiGAMN, we first use the k-means method to disentangle biological similarity information and introduce an adaptive encoding layer to automatically learn input encoding information adapted to four types of nodes (miRNA, disease, DimR, and DiDis). DiGAMN utilizes graph attention layers to aggregate the encoding information of nodes and their neighboring nodes and assign different attention weights to other nodes. It employs a heterogeneous biological memory network to store and retrieve potential dependencies in miRNA-disease association data. Finally, multiple layers of modules are concatenated to aggregate richer and more comprehensive feature representations.

This paper compares DiGAMN's performance with ten state-of-the-art models based on different methods and six traditional GNNs, verifying DiGAMN's excellent performance and the superiority of disentangled similarity information. The contribution of each module is validated through ablation experiments. Visualization of encoding status and memory units demonstrates the richness of encoding information in different encoding layers and the mechanism of the biological memory module. In case studies, several newly identified disease-associated miRNAs are presented, proving the effectiveness of DiGAMN. However, DiGAMN method has some limitations, such as the lack of a specific perception layer for cold-start MDAs tasks, and its adaptability to other association prediction tasks, such as drug-disease interaction (DDI) prediction and protein–protein interaction (PPI) prediction, still needs further investigation. Looking ahead, we will combine disentangling and heterogeneous memory units with other similar methods to explore other biologically related tasks and make further valuable contributions.

Additionally, we will investigate using heterogeneous biological memory units as representation encoders in combination with traditional machine learning methods to explore their predictive performance. In the future, the focus on using large models for predicting MDAs and enhancing the interpretability of these predictions will become critical research hotspots. Large models can leverage vast datasets to improve prediction accuracy but often require significant computational resources and pose challenges in interpretation. Addressing interpretability is crucial for validating predictions and facilitating clinical acceptance, driving the need for methods that provide clear, biologically plausible explanations. These advancements will enhance our understanding and treatment of diseases linked to miRNAs.

Key points

This paper first proposes applying disentangled higher-order hierarchical similarity methods in graph neural network models for predicting MDAs. It verifies the effectiveness of the disentangled similarity method through comparative experiments with six traditional

GNNs. This method will provide new insights into biological association and interaction predictions.

The study introduces a novel heterogeneous graph attention biological memory network for predicting disease-associated miRNAs. This network effectively captures and memorizes potential biological disentanglement information and the associations between disease-associated miRNAs.

In this study, the performance of the DiGAMN model is validated through extensive comparative experiments with ten state-of-the-art models employing different construction methods across four datasets.

The application of the DiGAMN model successfully identifies novel Disease-associated miRNAs related to diseases such as Cystic Fibrosis and Nasopharyngeal Carcinoma [73, 76], highlighting the utility and effectiveness of the model in discovering new miRNA-disease associations.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-024-11078-4.

Supplementary Material 1

Supplementary Material 2

Acknowledgements
We would like to express our sincere gratitude to all who contributed to this study. Special thanks to the colleagues at the School of Information and Artificial Intelligence, Anhui Agricultural University, for their valuable support and collaboration. We appreciate the financial support provided by these programs, which has been essential for completing this research. Furthermore, we would like to acknowledge the anonymous reviewers for their insightful comments and suggestions, which have significantly improved the quality of this manuscript. Lastly, we thank all the participants in this study for their cooperation and contribution.

Authors' contributions
Y.L. provided experimental conceptualization and methodological design of the model and conducted visualization of the results. Q.W. analyzed and organized the data and drafted the first manuscript. L.Z. tested the program. Y.L. guided the concept. C.L. made adjustments and revisions. Z.W. gave good advice on drawing figures. W.P. conceived the ideas for this paper. Y.Y. provided good guidance for algorithm design. X.Z. reviewed and revised the first manuscript and provided financial support. All authors read and approved the final manuscript.

Funding
This work was supported by the Young Wanjiang Scholar Program of Anhui Province and the Research Program of the Education Department of Anhui Province (2023AH050998).

Data availability
All the code and data related to this paper have been publicly available on our GitHub Repository (https://github.com/yinboliu-git/GAMN). Detailed definitions of these training parameters and evaluation metric metrics can be found in the Supplementary Materials and Supplementary Tables. Supplementary tables include Detailed metrics tables for ten current state-of-the-art methods on balanced and imbalanced datasets, including but not limited to AUC, PRC, F1 score, ACC, etc.; Supplementary tables for case studies; Comparison of six basic graph neural network models using non-disentangled and disentangled data; The impact of the number of disentanglements on model performance across four datasets under the same default hyperparameters; Performance comparison of the model under different ablation experiment settings, demonstrating the impact of each component on overall model performance; Hyperparameters of ten current state-of-the-art methods and six basic graph neural network models; Experiments on the effects of the k value in k-means on the results.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no conflict of interest.

Author details
[1]School of Information and Artificial Intelligence, Anhui Agricultural University, Hefei, Anhui 230036, China. [2]Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China. [3]China University of Petroleum, Beijing, Beijing 102249, China.

Received: 27 June 2024 Accepted: 21 November 2024
Published online: 02 December 2024

References

1. Ambros V. The functions of animal microRNAs. Nature. 2004;431(7006):350–5.
2. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell. 2004;116(2):281–97.
3. Verleih M, Visnovska T, Nguinkal JA, Rebl A, Goldammer T, Andreassen R. The Discovery and Characterization of Conserved and Novel miRNAs in the Different Developmental Stages and Organs of Pikeperch (Sander lucioperca). Int J Mol Sci. 2023;25(1):189.
4. Farazi TA, Hoell JI, Morozov P, Tuschl T. MicroRNAs in human cancer. MicroRNA cancer regulation: advanced concepts. Bioinform Syst Biol Tools. 2013;774:1–20.
5. Farazi TA, Spitzer JI, Morozov P, Tuschl T. miRNAs in human cancer. J Pathol. 2011;223(2):102–15.
6. Bardin P, Foussigniere T, Rousselet N, Rebeyrol C, Porter JC, Corvol H, Tabary O. miR-636: a newly-identified actor for the regulation of pulmonary inflammation in cystic fibrosis. Front Immunol. 2019;10:2643.
7. Bo W, Wang XG, Zhang M, Zhang Z. ZNF655 mediated by LINC01210/miR-124-3p axis promotes the progression of gastric cancer. Kaohsiung J Med Sci. 2023;39(3):200–8.
8. Jiang Q, Hao Y, Wang G, Juan L, Zhang T, Teng M, Liu Y, Wang Y. Prioritization of disease microRNAs through a human phenome-microRNAome network. BMC Syst Biol. 2010;4:1–9.
9. Li J, Zhang Y, Wang Y, Zhang C, Wang Q, Shi X, Li C, Zhang R, Li X. Functional combination strategy for prioritization of human miRNA target. Gene. 2014;533(1):132–41.
10. Lai Z, Liang J, Zhang J, Mao Y, Zheng X, Shen X, Lin W, Xu G. Exosomes as a delivery tool of exercise-induced beneficial factors for the prevention and treatment of cardiovascular disease: a systematic review and meta-analysis. Front Physiol. 2023;14:1190095.
11. Wilson RA, Deasy W, Hayes A, Cooke MB. High fat diet and associated changes in the expression of micro-RNAs in tissue: lessons learned from animal studies. Mol Nutr Food Res. 2017;61(6):1600943.
12. Jiao C-N, Zhou F, Liu B-M, Zheng C-H, Liu J-X, Gao Y-L: Multi-Kernel Graph Attention Deep Autoencoder for MiRNA-Disease Association Prediction. IEEE Journal of Biomedical and Health Informatics 2023.
13. Yu S, Wang H, Li J, Zhao J, Liang C, Sun Y: A Multi-Relational Graph Encoder Network for Fine-Grained Prediction of MiRNA-Disease Associations. IEEE/ACM Transactions on Computational Biology and Bioinformatics 2023.

14. Huang L, Zhang L, Chen X. Updated review of advances in microRNAs and complex diseases: towards systematic evaluation of computational models. Briefings Bioinformatics. 2022;23(6):bbac407.
15. Huang L, Zhang L, Chen X. Updated review of advances in microRNAs and complex diseases: taxonomy, trends and challenges of computational models. Briefings Bioinformatics. 2022;23(5):bbac358.
16. Huang L, Zhang L, Chen X. Updated review of advances in microRNAs and complex diseases: experimental results, databases, webservers and data fusion. Briefings in bioinformatics. 2022;23(6):397.
17. Chen X, Xie D, Zhao Q, You Z-H. MicroRNAs and complex diseases: from experimental results to computational models. Brief Bioinform. 2019;20(2):515–39.
18. Perez-Iratxeta C, Wjst M, Bork P, Andrade MA. G2D: a tool for mining genes associated with disease. BMC Genet. 2005;6:1–9.
19. Xu C, Ping Y, Li X, Zhao H, Wang L, Fan H, Xiao Y, Li X. Prioritizing candidate disease miRNAs by integrating phenotype associations of multiple diseases with matched miRNA and mRNA expression profiles. Mol BioSyst. 2014;10(11):2800–9.
20. Chen X, Yan CC, Zhang X, You Z-H, Deng L, Liu Y, Zhang Y, Dai Q. WBSMDA: within and between score for MiRNA-disease association prediction. Sci Rep. 2016;6(1):21106.
21. Chen X, Liu M-X, Yan G-Y. RWRMDA: predicting novel human microRNA–disease associations. Mol BioSyst. 2012;8(10):2792–8.
22. Xuan P, Han K, Guo M, Guo Y, Li J, Ding J, Liu Y, Dai Q, Li J, Teng Z. Prediction of microRNAs associated with human diseases based on weighted k most similar neighbors. PLoS ONE. 2013;8(8): e70204.
23. Shi H, Xu J, Zhang G, Xu L, Li C, Wang L, Zhao Z, Jiang W, Guo Z, Li X. Walking the interactome to identify human miRNA-disease associations through the functional link between miRNA targets and disease genes. BMC Syst Biol. 2013;7:1–12.
24. Lan W, Wang J, Li M, Liu J, Wu F-X, Pan Y. Predicting microRNA-disease associations based on improved microRNA and disease similarities. IEEE/ACM Trans Comput Biol Bioinf. 2016;15(6):1774–82.
25. Chen X, Yan G-Y. Semi-supervised learning for potential human microRNA-disease associations inference. Sci Rep. 2014;4(1):5501.
26. Yin M-M, Liu J-X, Gao Y-L, Kong X-Z, Zheng C-H. NCPLP: a novel approach for predicting microbe-associated diseases with network consistency projection and label propagation. IEEE Transactions on Cybernetics. 2020;52(6):5079–87.
27. Chen X, Sun L-G, Zhao Y. NCMCMDA: miRNA–disease association prediction through neighborhood constraint matrix completion. Brief Bioinform. 2021;22(1):485–96.
28. Ha J. SMAP: Similarity-based matrix factorization framework for inferring miRNA-disease association. Knowl-Based Syst. 2023;263: 110295.
29. You Z-H, Huang Z-A, Zhu Z, Yan G-Y, Li Z-W, Wen Z, Chen X. PBMDA: A novel and effective path-based computational model for miRNA-disease association prediction. PLoS Comput Biol. 2017;13(3): e1005455.
30. Liu Y, Yan X, Li J, Ren X, Wu Q, Wang G-A, Chen Y, Zhu X: miRNA-Disease Association Prediction based on Heterogeneous Graph Transformer with Multi-view similarity and Random Auto-encoder. In: 2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM): 2023: IEEE; 2023: 885–888.
31. Li Z, Shen X, Jiao Y, Pan X, Zou P, Meng X, Yao C, Bu J: Hierarchical bipartite graph neural networks: Towards large-scale e-commerce applications. In: 2020 IEEE 36th International Conference on Data Engineering (ICDE): 2020: IEEE; 2020: 1677–1688.
32. Epasto A, Perozzi B: Is a single embedding enough? learning node representations that capture multiple social contexts. In: The world wide web conference: 2019; 2019: 394–404.
33. Xia L, Shao Y, Huang C, Xu Y, Xu H, Pei J: Disentangled graph social recommendation. In: 2023 IEEE 39th International Conference on Data Engineering (ICDE): 2023: IEEE; 2023: 2332–2344.
34. Dai Q, Chu Y, Li Z, Zhao Y, Mao X, Wang Y, Xiong Y, Wei D-Q. MDA-CF: predicting miRNA-disease associations based on a cascade forest model by fusing multi-source information. Comput Biol Med. 2021;136: 104706.
35. Zhao Y, Chen X, Yin J. Adaptive boosting-based computational model for predicting potential miRNA-disease associations. Bioinformatics. 2019;35(22):4730–8.
36. Chen X, Wang L, Qu J, Guan N-N, Li J-Q. Predicting miRNA–disease association based on inductive matrix completion. Bioinformatics. 2018;34(24):4256–65.
37. Ji C, Gao Z, Ma X, Wu Q, Ni J, Zheng C. AEMDA: inferring miRNA–disease associations based on deep autoencoder. Bioinformatics. 2021;37(1):66–72.
38. Zhou F, Yin M-M, Jiao C-N, Zhao J-X, Zheng C-H, Liu J-X. Predicting miRNA–Disease Associations Through Deep Autoencoder With Multiple Kernel Learning. IEEE Transactions on Neural Networks and Learning Systems. 2021;34(9):5570–9.
39. Zhou F, Yin M-M, Jiao C-N, Cui Z, Zhao J-X, Liu J-X. Bipartite graph-based collaborative matrix factorization method for predicting miRNA-disease associations. BMC Bioinformatics. 2021;22:1–16.
40. Ding Y, Lei X, Liao B, Wu F-X. Predicting miRNA-disease associations based on multi-view variational graph auto-encoder with matrix factorization. IEEE J Biomed Health Inform. 2021;26(1):446–57.
41. Li J, Zhang S, Liu T, Ning C, Zhang Z, Zhou W. Neural inductive matrix completion with graph convolutional networks for miRNA-disease association prediction. Bioinformatics. 2020;36(8):2538–46.
42. Li Z, Li J, Nie R, You Z-H, Bao W. A graph auto-encoder model for miRNA-disease associations prediction. Briefings in bioinformatics. 2021;22(4):bbaa240.
43. Lou Z, Cheng Z, Li H, Teng Z, Liu Y, Tian Z. Predicting miRNA–disease associations via learning multimodal networks and fusing mixed neighborhood information. Briefings in Bioinformatics. 2022;23(5):bbac159.
44. Bai T, Yan K, Liu B. DAmiRLocGNet: miRNA subcellular localization prediction by combining miRNA–disease associations and graph convolutional networks. Briefings in Bioinformatics. 2023;24(4):bbad212.
45. Luo J, Liu Y, Liu P, Lai Z, Wu H. Data integration using tensor decomposition for the prediction of miRNA-disease associations. IEEE J Biomed Health Inform. 2021;26(5):2370–8.
46. Tang X, Luo J, Shen C, Lai Z. Multi-view multichannel attention graph convolutional network for miRNA–disease association prediction. Briefings in Bioinformatics. 2021;22(6):bbab174.
47. Zhong T, Li Z, You Z-H, Nie R, Zhao H. Predicting miRNA–disease associations based on graph random propagation network and attention network. Briefings in Bioinformatics. 2022;23(2):bbab589.
48. Wang Y-T, Wu Q-W, Gao Z, Ni J-C, Zheng C-H. MiRNA-disease association prediction via hypergraph learning based on high-dimensionality features. BMC Med Inform Decis Mak. 2021;21:1–13.
49. Liang X, Guo M, Jiang L, Fu Y, Zhang P, Chen Y. Predicting miRNA–disease associations by combining graph and hypergraph convolutional network. Interdiscip Sci: Comput Life Sci. 2024;16:1–15.
50. Chang Z, Zhu R, Liu J, Shang J, Dai L. HGSMDA: miRNA–Disease Association Prediction Based on HyperGCN and Sørensen-Dice Loss. Non-coding RNA. 2024;10(1):9.
51. Ning Q, Zhao Y, Gao J, Chen C, Li X, Li T, Yin M. AMHMDA: attention aware multi-view similarity networks and hypergraph learning for miRNA–disease associations identification. Briefings in Bioinformatics. 2023;24(2):bbad094.
52. Zhou Y, Wu M, Ouyang C, Wang X, Zhu M: Generalizable prediction of potential miRNA-disease associations based on heterogeneous graph learning. arXiv preprint arXiv:230707957 2023.
53. Chen X, Yin J, Qu J, Huang L. MDHGI: matrix decomposition and heterogeneous graph inference for miRNA-disease association prediction. PLoS Comput Biol. 2018;14(8): e1006418.
54. Wang S, Wang F, Qiao S, Zhuang Y, Zhang K, Pang S, Nowak R, Lv Z. Mshganmda: Meta-subgraphs heterogeneous graph attention network for mirna-disease association prediction. IEEE J Biomed Health Inform. 2022;27(10):4639–48.
55. Li C, Liu H, Hu Q, Que J, Yao J. A novel computational model for predicting microRNA–disease associations based on heterogeneous graph convolutional networks. Cells. 2019;8(9):977.
56. Peng W, Che Z, Dai W, Wei S, Lan W. Predicting miRNA-disease associations from miRNA-gene-disease heterogeneous network with multi-relational graph convolutional network model. IEEE/ACM Trans Comput Biol Bioinf. 2022;20(6):3363–75.
57. Yang Y, Sun Y, Li F, Guan B, Liu J-X, Shang J: MGCNRF: Prediction of disease-related miRNAs based on multiple graph convolutional networks and random forest. IEEE Transactions on Neural Networks and Learning Systems 2023.

58. Chen X, Li T-H, Zhao Y, Wang C-C, Zhu C-C. Deep-belief network for predicting potential miRNA-disease associations. Briefings in Bioinformatics. 2021;22(3):bbaa186.
59. 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.
60. Huang F, Yue X, Xiong Z, Yu Z, Liu S, Zhang W. Tensor decomposition with relational constraints for predicting multiple types of microRNA-disease associations. Briefings in bioinformatics. 2021;22(3):bbaa140.
61. Chen C, Zhang M, Liu Y, Ma S: Social attentional memory network: Modeling aspect-and friend-level differences in recommendation. In: *Proceedings of the twelfth ACM international conference on web search and data mining: 2019*; 2019: 177–185.
62. Santoro A, Bartunov S, Botvinick M, Wierstra D, Lillicrap T: Meta-learning with memory-augmented neural networks. In: International conference on machine learning: 2016: PMLR; 2016: 1842–1850.
63. Needleman SB, Wunsch CD. A general method applicable to the search for similarities in the amino acid sequence of two proteins. J Mol Biol. 1970;48(3):443–53.
64. Veličković P, Cucurull G, Casanova A, Romero A, Lio P, Bengio Y: Graph attention networks. arXiv preprint arXiv:171010903 2017.
65. Cover TM: Elements of information theory: John Wiley & Sons; 1999.
66. Kingma DP: Adam: A method for stochastic optimization. arXiv preprint arXiv:14126980 2014.
67. Huang Z, Shi J, Gao Y, Cui C, Zhang S, Li J, Zhou Y, Cui Q. HMDD v3. 0: a database for experimentally supported human microRNA–disease associations. Nucleic Acids Res. 2019;47(D1):D1013–7.
68. Cui C, Zhong B, Fan R, Cui Q. HMDD v4. 0: a database for experimentally supported human microRNA-disease associations. Nucleic Acids Res. 2024;52(D1):D1327–32.
69. Van der Maaten L, Hinton G. Visualizing data using t-SNE. J Mach Learn Res. 2008;9(11):2579–2605.
70. Stachowiak Z, Wojsyk-Banaszak I, Jończyk-Potoczna K, Narożna B, Langwiński W, Kycler Z, Sobkowiak P, Bręborowicz A, Szczepankiewicz A. MiRNA expression profile in the airways is altered during pulmonary exacerbation in children with cystic fibrosis—a preliminary report. J Clin Med. 2020;9(6):1887.
71. De Santi C, Gadi S, Swiatecka-Urban A, Greene CM. Identification of a novel functional miR-143-5p recognition element in the Cystic Fibrosis Transmembrane Conductance Regulator 3'UTR. AIMS genetics. 2018;5(01):053–62.
72. Fabbri E, Tamanini A, Jakova T, Gasparello J, Manicardi A, Corradini R, Sabbioni G, Finotti A, Borgatti M, Lampronti I. A peptide nucleic acid against microRNA miR-145-5p enhances the expression of the cystic fibrosis transmembrane conductance regulator (CFTR) in Calu-3 cells. Molecules. 2017;23(1):71.
73. Khasraghi LB, Nouri M, Vazirzadeh M, Hashemipour N, Talebi M, Zarch FA, Majidpoor J, Kalhor K, Farnia P, Najafi S. MicroRNA-206 in human cancer: mechanistic and clinical perspectives. Cell Signal. 2023;101: 110525.
74. Yang L, Lu P, Yang X, Li K, Chen X, Qu S. Excavating novel diagnostic and prognostic long non-coding RNAs (lncRNAs) for head and neck squamous cell carcinoma: an integrated bioinformatics analysis of competing endogenous RNAs (ceRNAs) and gene co-expression networks. Bioengineered. 2021;12(2):12821–38.
75. Romero-López MJ, Jiménez-Wences H. Cruz-De La Rosa MI, Alarcón-Millán J, Mendoza-Catalán MÁ, Ortiz-Sánchez E, Tinajero-Rodríguez JM, Hernández-Sotelo D, Valente-Niño GW, Martínez-Carrillo DN: miR-218-5p, miR-124-3p and miR-23b-3p act synergistically to modulate the expression of NACC1, proliferation, and apoptosis in C-33A and CaSki cells. Non-coding RNA Research. 2024;9(3):720–31.
76. Zhang E, Li X. LncRNA SOX2-OT regulates proliferation and metastasis of nasopharyngeal carcinoma cells through miR-146b-5p/HNRNPA2B1 pathway. J Cell Biochem. 2019;120(10):16575–88.
77. Hu Z, Dong Y, Wang K, Sun Y: Heterogeneous graph transformer. In: Proceedings of the web conference 2020: 2020; 2020: 2704–2710.
78. Wang X, Ji H, Shi C, Wang B, Ye Y, Cui P, Yu PS: Heterogeneous graph attention network. In: The world wide web conference: 2019; 2019: 2022–2032.
79. Kipf TN, Welling M: Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:160902907 2016.
80. Kipf TN, Welling M: Variational graph auto-encoders. arXiv preprint arXiv:161107308 2016.
81. Hamilton W, Ying Z, Leskovec J: Inductive representation learning on large graphs. Advances in neural information processing systems 2017, 30.

Publisher's Note