Subgraph topology and dynamic graph topology enhanced graph learning and pairwise feature context relationship integration for predicting disease-related miRNAs

Ping Xuan,[†,‡] Xiaoying Qi,[†] Sentao Chen,[‡] Jing Gu,[†] Xiuju Wang,[†] Hui Cui,[¶] Jun Lu,[†] and Tiangang Zhang[*,§]

[†]School of Computer Science and Technology, Heilongjiang University, Harbin 150080, China

[‡]Department of Computer Science and Technology, Shantou University, Shantou 515063, China

[¶]Department of Computer Science and Information Technology, La Trobe University, Melbourne 3083, Australia

[§]School of Cyberspace Security, Hainan University, Haikou 570228, China

E-mail: zhang@hlju.edu.cn

Abstract

As an increasing number of miRNAs have become the biomarkers of various human diseases, prediction of the candidate disease-related miRNAs is helpful for facilitating the early diagnosis of diseases. Most of the recent prediction models concentrated on learning of the features from the heterogeneous graph composed of the miRNAs and diseases. However, they failed to fully exploit the subgraph structures consisting of multiple miRNA and disease nodes, and they also did not completely integrate the context relationships among the pairwise features. We proposed a prediction model,

SFPred, to integrate and encode the local topologies from neighborhood subgraphs, the dynamically evolved heterogeneous graph topology, and the context among pairwise features. First, the importance of a miRNA (disease) node to another node is formulated according to their subgraphs composed of their neighbors. Second, the features of each miRNA (disease) node continuously change when the graph encoding gradually deepens for the miRNA-disease heterogeneous network. A strategy based on MLP is designed to estimate the edge weights according to the changed node features and form the dynamic graph topology. Third, considering the context relationships among the features of a pair of miRNA and disease nodes, a context relationship sensitive transformer is constructed to integrate these relationships. Finally, since the previous encoding layer of the transformer contains more detailed features of the pairwise, we present a multi-perspective residual strategy to supplement the detailed features to the following encoding layer from the channel perspective and the feature one, respectively. The extensive experiments confirmed that SFPred outperforms eight state-of-the-art methods for prediction of miRNA-disease associations, and the ablation experiments validate the effectiveness of the proposed innovations. The recall rates for the top ranked candidate miRNAs related to the diseases and the case studies on three diseases indicate SFPred's ability in screening the reliable candidates for the subsequent biological experiments.

Introduction

MicroRNAs (miRNAs) are endogenous non-coding RNAs, typically 22 to 24 nucleotides long, that can regulate the expression of messenger RNAs (mRNAs)[1–3] . An increasing number of studies indicate a significant correlation between miRNAs and the onset and progression of many diseases. For example, the expression of miR-509 is markedly decreased in colorectal cancer tissues and cells compared to normal tissues.[4] Additionally, miR-451a and miR-139 show promise as biomarkers for diagnosing lung squamous cell carcinoma.[5] Therefore, uncovering potential miRNA-disease associations is crucial for understanding disease pathogenesis and advancing diagnosis and treatment.

Computerized methods for predicting disease-related miRNAs can offer biologists more dependable candidates for screening true associations between miRNAs and diseases. These

methods can be broadly categorized into three groups. The first category relies on the biological premise that miRNAs with similar functions are typically associated with similar diseases.[6] The functional similarity between two miRNAs is determined by their associated diseases,[7] and a miRNA graph is constructed using this information. HDMP[8] predicts the association between miRNAs and diseases by aggregating information from the k-nearest neighbors within the miRNA graph. Additionally, the random walk with restart method has been employed to infer miRNA-disease associations.[9] However, these methods are challenging to apply to diseases that lack known associated miRNAs. Another approach involves a model based on the hypergeometric distribution of multiple miRNA similarities.[10] However, this model does not take into account disease similarity.

The second type of methods integrate disease similarity and miRNA-disease association information into the miRNA functional graph, resulting in a miRNA-disease heterogeneous graph. KATZ[11] algorithm is used to capture neighbor information within this graph, along with network consistency projection to determine the probability of miRNA-disease associations. Various approaches, such as matrix completion,[12–15] random walks,[16–19] random forests,[20,21] regression trees,[22] and probabilistic matrix factorization,[23,24] are utilized to predict new miRNA candidates for diseases. However, these methods are considered shallow prediction models and do not extensively learn the feature representation of miRNA and disease nodes.

Deep learning is able to learn the deep representative features of objects, and achieve decent performance in image reconstruction for magnetic particle imaging[25] and disease-related miRNA prediction. Some approaches utilize neural networks to learn the features of miRNAs and diseases and predict candidate miRNAs through neural-induced matrix completion models.[26,27] For example, the GCNA-MDA[28] prediction model learns the feature representations of miRNA and disease nodes using fully-connected neural networks and autoencoders. However, these models typically focus on individual miRNA similarity graphs and disease similarity graphs. For the miRNA-disease heterogeneous network, several prediction methods have been proposed, which are based on contrastive learning,[29–32] generative adversarial strategies,[33,34] graph convolution networks,[35] stacked autoencoders,[36] and graph

attention networks.[37–40] Nevertheless, these methods do not fully take into account the dynamic changes in graph topology during multi-layer graph inference processes.

In this paper, we designed a method based on **s**ubgraph and dynamic topologies enhanced graph convolutional network and **f**eature context relationship sensitive transformer for **pred**iction of miRNA-disease associations (SFPred). It intelligently combines local topological information from subgraphs within a heterogeneous graph, continuously evolving graph topology structures, and takes into account the relationships between features of pairwise. The advantages of our approach are summarized as follows.

First, the subgraphs composed of the neighborhood nodes were constructed for each miRNA (disease) node to exploit the local graph structures within the miRNA-disease heterogeneous graph. Most of the previous prediction methods focused on aggregating the features of the neighbor nodes and they did not fully integrate the subgraph structures consisting these neighbors. We designed a novel strategy to estimate the importance between two nodes by integrating the overlapped vertices and edges in their subgraphs.

Second, the graph learning process includes multiple graph convolution layers, and the features of miRNA and disease nodes are evolved dynamically with the increasement of layers. We presented a strategy based on MLP to update the weights of the connections among all the miRNAs and disease nodes and form the dynamic heterogeneous graph topology.

Third, the features of a pair of miRNA and disease nodes contain the similarities and associations between one of these nodes and all the miRNA and disease nodes, thus these features have context relationships. A learning strategy based on both convolution neural network and context sensitive transformer was presented to integrate the context relationships.

Finally, the original dual-channel features of the pairwise imply more pairwise details. The residual strategies from both channel and feature perspectives are designed to fuse the representative features from the deep encoding layer of the transformer and the detailed features from the shallow encoding layer. The contributions of SFPred's major innovations for the improved prediction performance, and SFPred's superior prediction performance were showed by the ablation experiments, the comparison with the advanced prediction methods,

and the case studies.
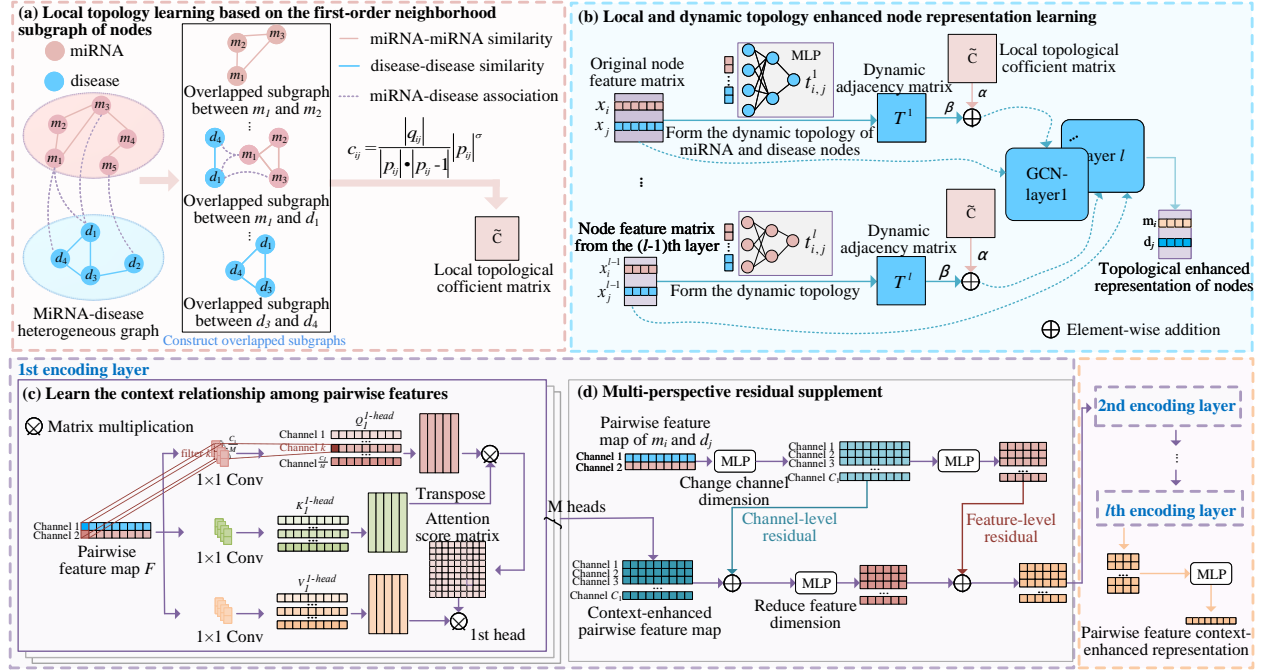
Figure 1: Framework of the proposed SFPred model. **(a)** generate the local topology according to the first-order neighborhodd subgraphs of nodes. **(b)** construct the dynamic graph topology and learn the topology enhanced representation of nodes. **(c)** encode the context relationships among the features of a pair of miRNA and disease nodes. **(d)** supplement the detailed features to convolutional transformer learning from the channel and feature levels.

Materials and methods

We propose a model called SFPred (Figure 1) to predict potential miRNAs associated with a given disease. Within the miRNA-disease heterogeneous graph, we construct subgraphs of the first-order neighborhood for any pair of nodes. This allows us to create a locally topological enhanced heterogeneous graph adjacency matrix. As the node features continuously change during the multi-layer graph inferring process, the topology between nodes also dynamically evolves. To address this, we incorporate a mechanism for dynamic topological embedding to capture and adapt to topology changes throughout the graph reasoning process. Furthermore, we employ a feature learning module based on transformer to capture contextual relationships between the features of pairwise miRNA-disease.

Dataset

We gathered a dataset consisting of 7,908 confirmed associations between miRNAs and diseases from the Human MicroRNA Disease Database.[41] This dataset includes 793 miRNAs

and 341 diseases. A directed acyclic graph (DAGs) for a disease typically embraces all terms related to that disease, sourced from the American Library of Medicine.[42] To measure semantic similarity between diseases, we followed previous studies[7] and calculated the similarity based on the DAGs of the two diseases.

Construction of the miRNA-disease heterogeneous graph

We constructed a heterogeneous graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ incorporating nodes representing miRNAs and diseases, along with their association and similarity connections. The node set $\mathcal{V} = \mathcal{V}^{micro} \cup \mathcal{V}^{disea}$ consists of the miRNA node set $\mathcal{V}^{micro}$ and the disease node set $\mathcal{V}^{disea}$. The edge set $\mathcal{E}$ includes edges representing miRNA-disease associations, miRNA similarities, and disease similarities. An edge $e_{i,j} \in \mathcal{E}$ is used to connect pairs of nodes $v_i, v_j \in \mathcal{V}$. The miRNA-disease association matrix $\mathcal{K}^{micro-disea}$ is formulated as,

$$\mathcal{K}^{micro-disea} \in \mathcal{R}^{\mathcal{N}_m \times \mathcal{N}_d}, if \ v_i \in \mathcal{V}^{micro}, v_j \in \mathcal{V}^{disea} \tag{1}$$

where $\mathcal{N}_m$ and $\mathcal{N}_d$ signify the quantities of miRNAs and diseases, respectively. In the matrix $\mathcal{K}^{micro-disea}$, the i-th row $\mathcal{K}_{i,*}^{micro-disea}$ is the associations between the i-th miRNA $v_i \in \mathcal{V}^{micro}$ and all diseases. The miRNAs associated with the disease are recorded in the j-th column $\mathcal{K}_{*,j}^{micro-disea}$ of matrix $\mathcal{K}^{micro-disea}$. If $v_i \in \mathcal{V}^{micro}$, $v_j \in \mathcal{V}^{disea}$ already have a known association, $\mathcal{K}_{ij}^{micro-disea}$ is 1; otherwise, $\mathcal{K}_{ij}^{micro-disea}$ is 0.

The similarity matrix O incorporates the miRNA similarity matrix O^{micro} and the disease similarity matrix O^{disea},

$$O = \begin{cases} O^{micro} \in \mathcal{R}^{\mathcal{N}_m \times \mathcal{N}_m} \\ O^{disea} \in \mathcal{R}^{\mathcal{N}_d \times \mathcal{N}_d}. \end{cases} \tag{2}$$

The degree of similarity between two miRNA (disease) nodes v_i, v_j is reflected by $O_{ij}^{micro}(O_{ij}^{disea})$, which has a value ranging from 0 to 1. A higher value of $O_{ij}^{micro}(O_{ij}^{disea})$ symbolizes greater functional(semantic) similarity. The semantic similarity of two diseases is calculated using their DAGs, following the method proposed by Wang et al.[7] O_{ij}^{micro} represents the similarity between miRNA nodes v_i and v_j, which is computed based on the two sets of diseases

associated with v_i, v_j.
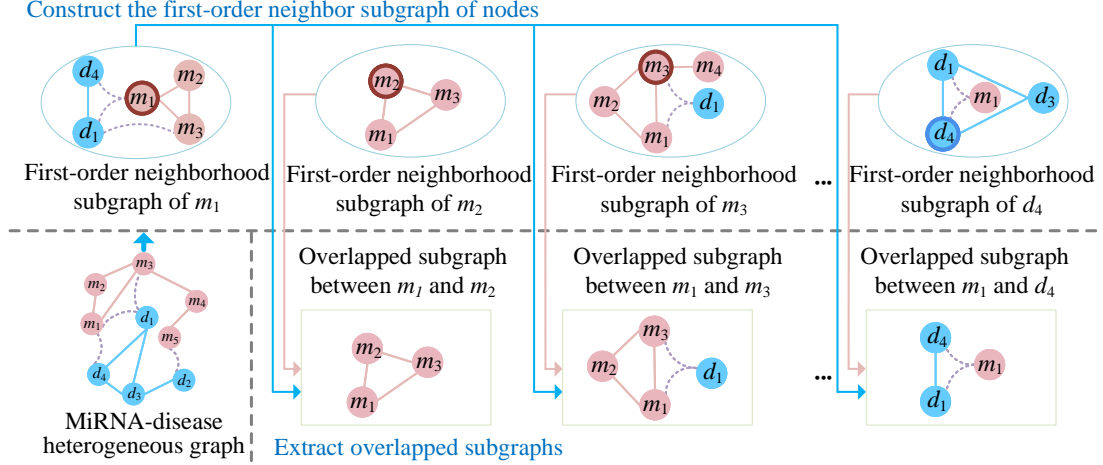
Figure 2: Illustration of how to construct the subgraphs composed of the first-order neighbors and get the overlapped subgraph between two nodes.

Based on the association and similarity matrix, the adjacency matrix $\mathcal{A}$ of the miRNA-disease heterogeneous graph can be constructed as follows,

$$\mathcal{A} = \begin{bmatrix} O^{micro} & \mathcal{K}^{micro-disea} \\ \mathcal{K}^{micro-disea^T} & O^{disea} \end{bmatrix} \in \mathcal{R}^{\mathcal{N}_\mathcal{V} \times \mathcal{N}_\mathcal{V}} \tag{3}$$

where $\mathcal{K}^{micro-disea^T}$ represents the transpose of $\mathcal{K}^{micro-disea}$, and $\mathcal{N}_\mathcal{V} = \mathcal{N}_m + \mathcal{N}_d$. The i-th row of $\mathcal{A}$, denoted as $\mathcal{A}_i$, contains the associations and similarities between node v_i and all miRNAs and diseases, serving as the attribute vector for v_i. Therefore, $\mathcal{A}$ can be considered as the node feature matrix of the miRNA-disease heterogeneous graph and is denoted as X.

Subgraph and dynamic topologies enhancement for node feature learning

The first-order neighbors of each miRNA (disease) node v_i and the edges among them form the first-order neighbor subgraph. Most of previous methods did not fully leverage the local topology within a subgraph. During the process of graph convolutional learning, the node features change gradually, resulting in a dynamical changing graph topology. We proposed the subgraph and dynamic topology enhanced graph convolutional network (SDGCN) to integrate the topological representations of miRNA and disease nodes. SDGCN comprises L^{te}

7
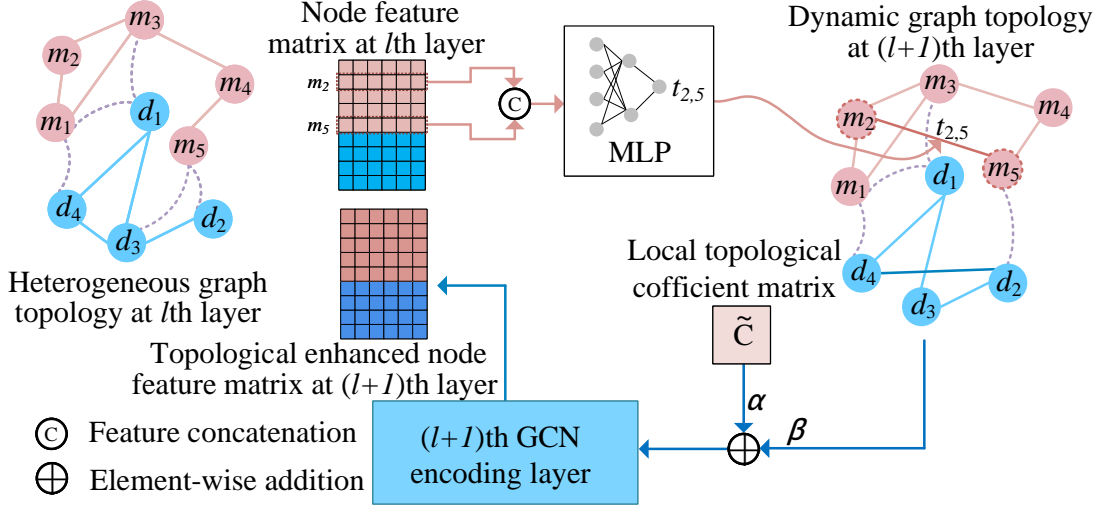
Figure 3: Construction of the overlapped subgraphs of m_1 with the remaining nodes as an example.

graph convolutional layers, with each block containing importance evaluation block between nodes and a dynamic topological embedding block. We describe the learning process of the l-th graph convolutional layer to illustrate.

In the l-th layer, the first-order neighbor set of miRNA (disease) nodes v_i is represented as $\phi(i) = \{v_j \in \mathcal{V} | (v_i, v_j) \in \mathcal{E}\}$, where $\mathcal{E}$ is the set of edges of graph $\mathcal{G}$. Furthermore, the neighbor nodes in $\phi(i)$ include both miRNA and disease node types. The first-order neighborhood subgraph of v_i, denoted as $\mathcal{S}_i$, covers the similarity and association connections between $v_j \in \phi(i) \cup v_i$, reflecting their respective local topology. For miRNA (disease) nodes v_i and v_j, $\mathcal{S}_{ij} = \mathcal{S}_i \cup \mathcal{S}_j$ represents their overlapped subgraph (Figure 2).

To obtain the importance of miRNA (disease) node v_i to v_j at the local topological level, we calculate its local topological coefficient c_{ij} based on $\mathcal{S}_{ij}$,

$$c_{ij} = \frac{|q_{ij}|}{|p_{ij}| \cdot |p_{ij} - 1|} \times |p_{ij}|^\sigma \tag{4}$$

where p_{ij} and q_{ij} represent the number of vertices and edges of $\mathcal{S}_{ij}$, respectively, and σ is a hyperparameter. By calculating c_{ij} between any two miRNA (disease) nodes, we obtain the local topological cofficient matrix C. We normalize C using the Laplacian to obtain $\tilde{C}$,

$$\tilde{C} = (D)^{-\frac{1}{2}}(C + I)(D)^{-\frac{1}{2}} \tag{5}$$

where $D_{ii} = \sum_j C_{ij}$, and I represents the identity matrix.

SDGCN consists of multiple graph convolutional layers. As the number of encoding layers increases, the features of miRNA and disease nodes evolve gradually, leading to a dynamic change of the heterogeneous graph topology. Most previous methods did not fully integrate the dynamic topologies during the graph convolution process. We designed a dynamic topology encoding strategy to capture the topological changes of heterogeneous graph (Figure 3). The miRNA (disease) nodes v_i and v_j, their node features at the (l-1)-th layer are articulated as x_i^{l-1} and x_j^{l-1} respectively. The connection between x_i^{l-1} and x_j^{l-1} forms the edge weight t_{ij}^l between nodes v_i and v_j,

$$t_{ij}^l = Relu(W_t^l \left[x_i^{l-1}; x_j^{l-1} \right] + b_t^l) \tag{6}$$

where W_t^l and b_t^l represent the weight matrix and bias vector of the l-th layer, respectively, and $[;]$ denotes the concatenation operation. The connection weights between all pairs of miRNA (disease) nodes are computed to obtain the dynamic adjacency matrix T^l.

$\tilde{C}$ and T^l represent the adjacency matrices for node importance and dynamic topology, respectively. $\tilde{C}$ and T^l contribute differently to the topological representation learning of nodes, we designed an adaptive fusion strategy. α and β are the learnable parameters to balance the contribution of two matrices, resulting in the fused adjacency matrix $\mathcal{H}$,

$$\mathcal{H} = \alpha\tilde{C}\oplus\beta T^l \tag{7}$$

α and β are initialized randomly. The topological enhancement representation of the miRNA node v_i is denoted as x_i^l,

$$x_i^l = W_x^l(\lambda(\sum_{v_j\in\phi(i)} \mathcal{H}_{ij}^l + 1)x_i^{l-1} + \sum_{v_j\in\phi(i)} (\mathcal{H}_{ij}^l + 1)x_j^{l-1}) + b_x^l$$
$$l = 1, ...L^{te} \tag{8}$$

where W_x^l and b_x^l are the weight matrix and bias vector of the l-th layer, and L^{te} symbol-

9

izes the total number of graph convolutional encoding layers. x_i^{l-1} and x_j^{l-1} represent the feature vectors of node v_i and its neighbors v_j at the (l-1)-th layer, respectively. They are supplemented to enhance the features of these nodes themselves. λ is a learnable parameter used to adaptively adjust the weighting of feature v_i. Similarly, the topological enhancement representation $x_{\mathcal{N}_m+j}^l$ of disease node v_j can be computed. The topological enhancement representations x_i^{te} and x_j^{te} obtained after L^{te} layers of graph convolution are concatenated to form x_{ij}^{te}.

Feature context learning based on transformer

A miRNA has the neighboring miRNAs which have similar functions with it and it also has the neighboring diseases which are associated with it. Similarly, a disease also has its neighboring diseases and miRNAs. Thus, the features of a miRNA (disease) include these similarities and associations. As the miRNA with similar functions are usually associated with similar diseases, the pairwise features have contextual relationships. The previous methods did not fully utilize these contextual relationships. We designed a feature context sensitive transformer (FCTransformer) to capture these relationships. FCTransformer consists of L^{fcl} encoding layers. We illustrate the encoding process using the l-th layer as an example.

We stack the feature vectors x_i and x_{N_m+j} of m_i and d_j along the channel dimension to obtain the original pairwise feature embedding for m_i-d_j $F \in R^{(\mathcal{N}_\mathcal{V},2)}$ (as shown in Figure 1c). The k-th column contains a pair of features, and it is worth to be focused on. Assuming k is 5, if m_5 has similar function to m_i and m_5 is associated with d_j, m_i is more likely to be associated with d_j. The l-th layer's input feature map $F^{l-1} \in R^{(\mathcal{N}_{l-1},C_{l-1})}$ has C_{l-1} channels. In the i-th attention head, we conduct $\frac{C_{l-1}}{M}$ 1×1 convolutions on F^{l-1}, and then we get $Q_l^i \in R^{(\mathcal{N}_{l-1},\frac{C_l}{M})}$, $K_l^i \in R^{(\mathcal{N}_{l-1},\frac{C_l}{M})}$ and $V_l^i \in R^{(\mathcal{N}_{l-1},\frac{C_l}{M})}$,

$$Q_l^i = P_{l,i}^Q * F^{l-1} \tag{9}$$

$$K_l^i = P_{l,i}^K * F^{l-1} \tag{10}$$

$$V_l^i = P_{l,i}^V * F^{l-1} \tag{11}$$

where $\mathcal{N}_{l-1}$ represents the dimension of the pairwise feature in the $(l\text{-}1)$-th layer, and $\frac{C_l}{M}$ denotes the number of channels per head. $P_{l,i}^Q$, $P_{l,i}^K$, $P_{l,i}^V$ represent the sets containing $\frac{C_{l-1}}{M}$ 1×1 convolution kernels, respectively. $F^0 = F$, '$*$' represents the convolution operation. The dot product operation of Q_l^i and K_l^i yields the attention score matrix $Q_l^i(K_l^i)^T$ at the pairwise feature level. $Q_l^i(K_l^i)^T$ is then normalized and multiplied with V_l^i to obtain the feature matrix D_l^i, which captures the contextual relationships of pairwise for the i-th attention head,

$$D_l^i = softmax(\frac{Q_l^i(K_l^i)^T}{\sqrt{d}})V_l^i \tag{12}$$

where $d = \mathcal{N}_{l-1}$, which is utilized to smooth gradients during the training process. After passing through the l-th encoding layer, the pairwise feature matrix $D_l \in \mathcal{R}^{(\mathcal{N}_{l-1}, C_l)}$ can be obtained,

$$D_l = \coprod_M^{i=1} D_l^i \tag{13}$$

where $\coprod$ denotes the concatenation operation.

For the original dual-channel pairwise feature map, the encoding layers of FCTransformer may alter its number of channels and feature dimension. The previous methods did not completely utilize the detailed information from both the channel and feature perspectives. Therefore, it is necessary to design a multi-perspective residual supplementation strategy to complement the information. The pairwise feature matrix obtained from the $(l\text{-}1)$-th layer learning is denoted as $F^{l-1} \in \mathcal{R}^{(\mathcal{N}_{l-1}, C_{l-1})}$. A fully-connected neural network is applied to adjust the channel number of F^{l-1}, ensuring it matches the number of channels as D_l. The resulting feature matrix is denoted as $\hat{F}^{l-1} \in \mathcal{R}^{(\mathcal{N}_{l-1}, C_l)}$,

$$\hat{F}^{l-1} = \theta(W_F^l F^{l-1} + b_F^l). \tag{14}$$

Here, θ represents the $GELU$ [43] activation function, W_F^l is the weight matrix, and b_F^l is the bias vector.

The feature matrix $\hat{F}^{l-1}$ contains more detailed features of pairwise m_i-d_j, and we com-

bine it with D_l to obtain $\hat{D}_l \in \mathcal{R}^{(\mathcal{N}_{l-1}, C_l)}$ using channel residual connections,

$$\hat{D}_l = \hat{F}^{l-1} \oplus D_l, \tag{15}$$

where $\oplus$ represents element-wise addition. $\hat{D}_l$ is then subjected to a fully-connected neural network to reduce the dimensionality of the pairwise features, forming the pairwise feature matrix $\tilde{D}_l \in \mathcal{R}^{(\mathcal{N}_l, C_l)}$,

$$\tilde{D}_l = Relu(W_d^l \hat{D}_l + b_d^l), \tag{16}$$

where W_d^l and b_d^l represent the weight matrix and bias vector, respectively.

Similarly, $\hat{F}^{l-1}$ also undergoes a fully-connected layer to make it consistent with the feature dimension of $\tilde{D}_l$, resulting in a pairwise feature matrix $\tilde{F}^{l-1} \in \mathcal{R}^{(\mathcal{N}_l, \mathcal{C}_l)}$,

$$\tilde{F}^{l-1} = Relu(W_f^l F^{l-1} + b_f^l) \tag{17}$$

where W_f^l denotes the weight matrix, and b_f^l denotes the bias vector. Feature residual is then applied to $\tilde{F}^{l-1}$ and $\tilde{D}_l$ to obtain the feature matrix F^l for the l-th layer,

$$F^l = \tilde{F}^{l-1} \oplus \tilde{D}_l \tag{18}$$

The enhanced representation of pairwise context for m_i-d_j, obtained after F undergoes learning through L^{fcl} encoding layers, is denoted as $F^{fcl} \in \mathcal{R}^{(\mathcal{N}_{fcl}, C_{fcl})}$. F^{fcl} is then reshaped and renamed as $x_{ij}^{fcl} \in \mathcal{R}^{(1, C_{fcl} \times \mathcal{N}_{fcl})}$.

Association score estimation

The representations of miRNA m_i and disease d_j after topological enhancement, denoted as x_{ij}^{te}, and the representations of pairwise after feature enhancement, denoted as x_{ij}^{fcl}, are concatenated together to form the pairwise feature vector x_{ij}^{mix},

$$x_{ij}^{mix} = \left[x_{ij}^{te} || x_{ij}^{fcl} \right] \tag{19}$$

The fully-connected layer and the softmax layer are employed to calculate the association distributions of m_i and d_j, represented as φ,

$$\varphi = softmax(x_{ij}^{mix}W_{mix} + b_{mix}) \tag{20}$$

where W_{mix} and b_{mix} represent the learnable weight matrices and bias vectors, respectively. $\varphi[0]$ represents the probability that m_i and d_j are not associated, while $\varphi[1]$ represents the probability that they are associated.

We calculate the cross-entropy loss between the true labels of miRNA and disease pairwise and the predicted association probabilities φ,

$$L_{loss} = -\sum_{i=1} [y_{label} \times log(\varphi[1]) + (1 - y_{label}) \times log(1 - \varphi[0])] + \rho||\omega|| \tag{21}$$

where y_{label} denotes the true labels, and $y_{label} = 1$ if the miRNA and disease are already associated, otherwise $y_{label} = 0$. ω represents the collection of all parameters in the model, and its $L2$ regularization is introduced to prevent overfitting, with ρ being a hyperparameter.

Experimental results and discussions
Parameter settings

The number of SDGCN encoding layers and that of FCTransformer encoding layers were selected from {1, 2, 3}. The experiments were conducted for all the combinations of their layer numbers. As shown in the supplementary table S1, the model achieved the highest AUC (AUC = 0.941) and AUPR (AUPR = 0.470) when both of their layer numbers were 2. The possible reason is that a single encoding layer failed to completely learn the deep associations between miRNAs and diseases, while three encoding layer may amplify the useless information within the data about miRNAs and diseases. The dimenstions of the output features from SDGCN's first and second layers were 512 and 128, respectively. The output feature dimensions of FCTransformer's first and second layers were set to 256 and 64, and the number of heads in each layer was 4. The filter size and number in FCTransformer's first encoding layer were set to 1×1 and 16, respectively, and its stride was 1. FCTransformer's

second encoding layer contained 16 filters with size=1×1 and stride=1. The channel number of the first layer, C_1, was selected from {128, 64, 32}, and that of the second layer, C_2, was selected from {64, 32, 16}. The model achieved the best performance when C_1 and C_2 were 64 and 16, respectively (supplementary table S2). After the enhanced pairwise feature representations were obtained, the input and output dimensions of the fully-connected neural network were 1024 and 256, respectively. During the process of estimation of association scores, the dimension of hidden layer of fully-connected neural network was 64. During training, the learning rate was set to 0.001, and weight decay was set to 0.0005.

Evaluate metrics

Five-fold cross-validation is performed to evaluate the performance of SFPred and other comparison models. The observed associations between miRNA and diseases are considered positive samples, while the negative samples consist of all unobserved miRNA-disease pairwise. In each fold of cross-validation, the dataset is divided into training and testing sets. The training set contains 80% of the positive samples and an equal number of randomly chosen negative samples. On the other hand, the testing set comprises the remaining 20% of positive samples and all remaining negative samples.

Receiver Operating Characteristic (ROC) curve area (AUC) and Precision-Recall (PR) curve area (AUPR), as well as the recall of top k candidates, are used as evaluation metrics to assess model performance. For each of the 341 diseases, we obtain the AUC and AUPR values through five-fold cross-validation. The average AUC and AUPR values across all folds are then considered as the predictive results for that disease. Similarly, the average results of AUC and AUPR for all diseases are regarded as the predictive results for the entire model. Biologists typically focus on the top-ranked predictions for further validation. Therefore, we calculate the recall of the *top* $k \in [30, 60, ..., 240]$ candidates, where k ranges from 30 to 240 with a step size of 30, to evaluate the prediction performance. A higher recall indicates that more true associations are found among the top-ranked candidates.

Ablation experiments

Ablation experiments were conducted to evaluate the effectiveness of inter-node importance evaluation (INIE), dynamic topology embedding (DTE), and pairwise feature context learn-

14

ing (PFCL), channel residual strategy (CRS) and feature residual strategy (FRS), as presented in Table 1.

**Table 1: Results of conducting the ablation experiments on SFPred.**

INIE	DTE	PFCL	CRS	FRS	Average AUC	Average AUPR
×	✓	✓	✓	✓	0.937	0.431
✓	×	✓	✓	✓	0.934	0.423
✓	✓	×	✓	✓	0.908	0.296
✓	✓	✓	×	✓	0.937	0.454
✓	✓	✓	✓	×	0.932	0.413
✓	✓	✓	✓	✓	0.941	0.470

When the model lacked INIE, the AUC and AUPR decreased by 0.4% and 3.9%, respectively. Similarly, ignoring DTE resulted in a 0.7% decrease in AUC and a 4.7% decrease in AUPR compared to the final model. In contrast, compared to the model without PFCL, the complete model demonstrated an improvement of 3.3% in AUC and 17.4% in AUPR. After removing CRS, the corresponding AUC and AUPR decreased by 0.4% and 1.6%, respectively. The model without FRS has an AUC and AUPR that are 0.9% and 5.7% lower than the complete model, respectively. This indicated that the supplementation of detailed information from both the channel and feature perspectives is indeed able to enhance the pairwise feature learning. Experimental results suggest that PFCL has the greatest contribution to enhancing prediction performance. This is attributed to its focus on information directly related to miRNA-disease pairwise and its effective integration of contextual relationships between pairwise features.

Comparison with other methods

SFPred is compared with eight state-of-the-art miRNA-disease prediction methods, including Liu's method,[16] NIMCGCN,[27] GMDA,[44] MMGCN,[45] NIMGSA,[26] AGAEMD,[37] AMH-MDA,[46] and MAMFGAT.[40] As depicted in Figure 4, using the optimal parameters for each method, SFPred achieves the best performance in terms of average AUC and AUPR for all 341 diseases (AUC = 0.941, AUPR = 0.470). SFPred's average AUC is 5% better than Liu's method, 8.7% higher than NIMCGCN, 1.3% and 4.7% higher than GMDA and MMGCN, respectively. It outperforms NIMGSA and AGAEMD by 8% and 3.2%, respectively, and is

4.4% and 2.4% better than AMHMDA and MAMFGAT. In terms of average AUPR values, SFPred is 37.1% higher than Liu's methods on average, 22% higher than GMDA, 28.4% higher than NIMCGCN, 30.8% higher than MMGCN, 30.4% higher than NIMGSA, 12.8% higher than AGAEMD, 28.3% higher than AMHMDA, and 9.2% higher than MAMFGAT.
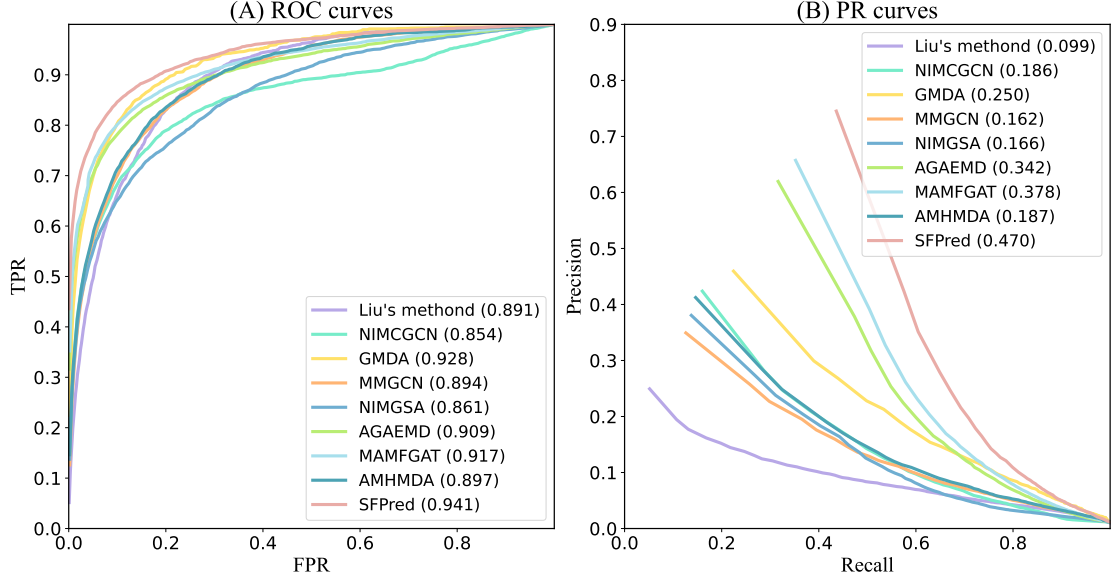
Figure 4: ROC curves and PR curves of SFPred and the compared methods.

We conducted Wilcoxon tests to determine whether SFPred achieved significantly higher AUC (AUPR) values compared to other methods across all 341 diseases. The results of the paired Wilcoxon tests (Table 2) indicate that, with a p-value threshold of 0.05, SFPred consistently achieved significantly better predictive performance.

Table 2: Results of the paired Wilcoxon test for comparing SFPred and several advanced methods.

	p-value of AUC	p-value of AUPR
MAMFGAT	1.4122e-03	2.9366e-02
AMHMDA	4.27631e-4	1.1687e-4
AGAEMDA	2.7696e-03	9.7602e-05
MMGCN	5.5024e-4	1.8804e-4
NIMCGCN	1.0658e-3	1.6420e-2
NIMGSA	2.7657e-3	1.4635e-3
GMDA	5.3443e-03	2.3933e-03
Liu's method	7.8656e-09	3.7836e-04

Liu's method relies on shallow prediction models based on non-negative matrix factorization, which struggle to capture the deep features of miRNA-disease nodes, resulting in

slightly inferior performance. NIMCGCN, MMGCN, and AMHMDA employ graph convolution techniques to learn node latent representations separately from single-layer graphs of miRNA and disease similarity, thereby overlooking known miRNA-disease association. NIMGSA, AGAEMD and MAMFGAT primarily concentrate on graph inference learning on heterogeneous networks. However, these methods overlook the dynamic evolution of graph topology information during the inference process. GMDA, on the other hand, employs generative adversarial networks to learn feature embeddings of miRNA-disease pairwise. However, it falls short in integrating rich contextual information among pairwise features.
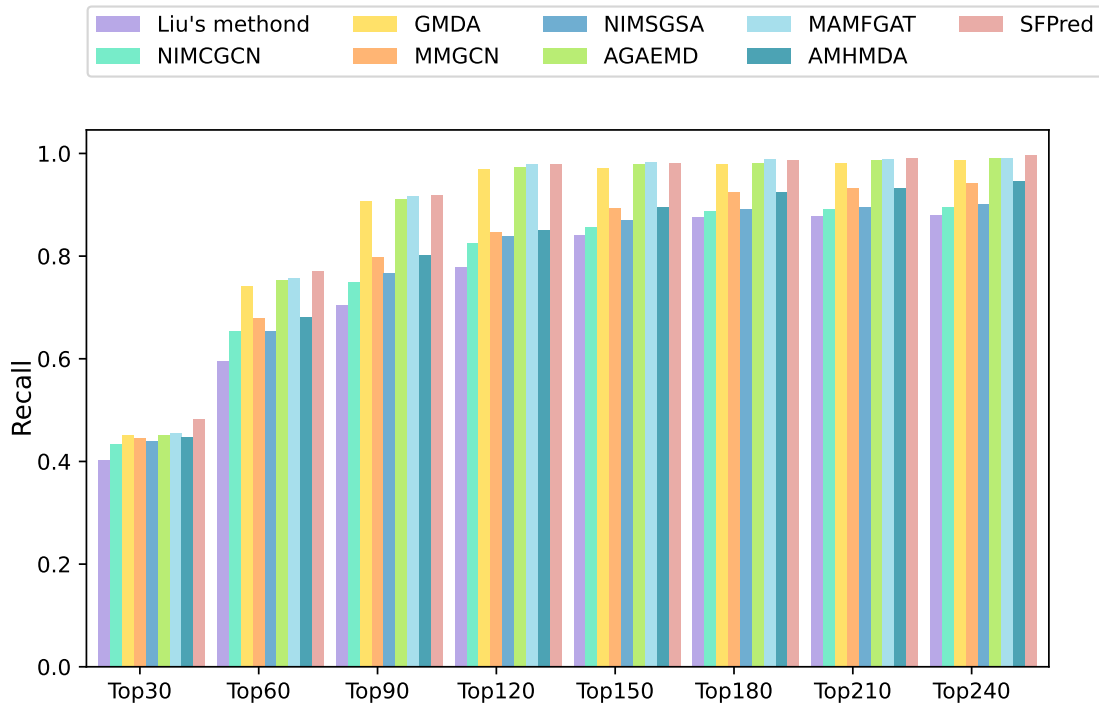
Figure 5: Recall rates of SFPred and the compared methods at different k values.

Figure 5 illustrates the top-k recall rates, where higher recall rates indicate that more true associations are included in the top-k candidates. Across different k values, SFPred consistently outperforms other methods, achieving recall rates of 48.2% for the top 30, 77.1% for the top 60, 91.9% for the top 90, and 97.8% for the top 120. MAMFGAT ranks at 45.5%, 75.6%, 91.6%, and 97.9% for the top 30, 60, 90, and 120,respectively. AGAEMD and GMDA achieve recall rates of 45.1%, 75.3%, 91%, 97.3% and 45%, 74.2%, 90.6%, and 97%, respectively. AMHMDA's recall rate is slightly higher than MMGCN, with values of 44.7%, 68%, 80.2%, and 85.1%, while MMGCN achieves 44.5%, 67.8%, 79.8%, and 84.7%, respectively. Both NIMGSA and NIMCGCN have lower recall rates compared to AMHMDA

17

and MMGCN. NIMGSA achieves recall rates of 43.9%, 65.4%, 76.7%, and 83.8%, while NIMCGCN achieves 43.3%, 65.4%, 74.9%, and 82.5%. Liu's method exhibits relatively lower recall rates, with values of 40.2%, 59.5%, 70.5%, and 77.9%.

Case studies: Breast neoplasms, Hepatocellular carcinoma, and Lymphoma

To demonstrate SFPred's ability to identify promising candidate miRNAs related to diseases, we conducted case studies on breast neoplasms, hepatocellular carcinoma, and lymphoma. For each disease, candidate miRNAs were ranked in descending order based on their association scores, and the top 50 candidates were listed.

Table 3: The top 50 candidate miRNAs associated with breast neoplasms.

Rank	MiRNA name	Description	Rank	MiRNA name	Description
1	hsa-mir-196a	dbDEMC, miRCancer, miRTarBase	26	hsa-mir-1471	dbDEMC, miRTarBase
2	hsa-mir-34a	dbDEMC, miRCancer, miRTarBase	27	hsa-mir-510	dbDEMC, miRCancer, miRTarBase, miR2Disease
3	hsa-mir-146b	miRCancer, miRTarBase, miR2Disease	28	hsa-mir-661	dbDEMC, miRTarBase, miR2Disease
4	hsa-mir-214	dbDEMC, miRCancer, miRTarBase	29	hsa-mir-516a	miRTarBase
5	hsa-mir-605	dbDEMC	30	hsa-mir-381	dbDEMC, miRCancer
6	hsa-mir-200a	dbDEMC, miRCancer, miRTarBase, miR2Disease	31	hsa-mir-218	dbDEMC, miRCancer, miRTarBase
7	hsa-mir-183	dbDEMC, miRCancer, miRTarBase	32	hsa-mir-383	dbDEMC, miRCancer, miRTarBase
8	hsa-mir-132	dbDEMC, miRCancer, miRTarBase	33	hsa-mir-22	dbDEMC, miRCancer, miRTarBase
9	hsa-mir-453	dbDEMC	34	hsa-mir-200c	dbDEMC, miRCancer, miRTarBase, miR2Disease
10	hsa-mir-195	dbDEMC, miRCancer, miRTarBase	35	hsa-mir-564	dbDEMC
11	hsa-mir-199a	dbDEMC, miRTarBase	36	hsa-mir-128b	dbDEMC, miR2Disease
12	hsa-mir-520g	dbDEMC, miRCancer	37	hsa-mir-1323	dbDEMC, miRTarBase
13	hsa-mir-135	literature	38	hsa-mir-298	dbDEMC, miRTarBase
14	hsa-mir-221	dbDEMC, miRCancer, miRTarBase, miR2Disease	39	hsa-mir-224	dbDEMC, miRCancer, miRTarBase
15	hsa-mir-30e	dbDEMC, miRCancer	40	hsa-mir-19	literature
16	hsa-mir-181a	dbDEMC, miRCancer, miRTarBase	41	hsa-mir-192	dbDEMC
17	hsa-mir-141	dbDEMC, miRCancer, miRTarBase, miR2Disease	42	hsa-mir-126	dbDEMC, miRCancer, miRTarBase
18	hsa-mir-133a	dbDEMC, miRCancer, miRTarBase	43	hsa-mir-21	dbDEMC, miRCancer, miRTarBase
19	hsa-mir-621	dbDEMC	44	hsa-mir-513a	literature
20	hsa-mir-873	dbDEMC, miRCancer	45	hsa-mir-148a	dbDEMC, miRCancer, miRTarBase, miR2Disease
21	hsa-mir-19a	dbDEMC, miRCancer, miRTarBase	46	hsa-mir-618	dbDEMC
22	hsa-mir-1274b	dbDEMC	47	hsa-mir-302f	dbDEMC
23	hsa-mir-874	dbDEMC, miRCancer, miRTarBase	48	hsa-mir-19b	dbDEMC, miRCancer, miRTarBase
24	hsa-mir-205	dbDEMC, miRCancer, miRTarBase, miR2Disease	49	hsa-mir-375	dbDEMC, miRCancer, miRTarBase
25	hsa-mir-145	dbDEMC, miRCancer, miRTarBase, miR2Disease	50	hsa-mir-892b	dbDEMC, miRCancer

Firstly, we used several databases for validation. The miRTarBase database contains miRNA-disease interactions that have been experimentally validated. These interactions are obtained through NLP text mining of relevant literature, followed by manual curation and verification. The miRCancer database includes associations between miRNAs and cancer that have been experimentally validated and extracted from published literature. The miR2Disease database focuses on the abnormal expression of miRNAs and their links to diseases. Table 3 presents the top 50 candidates related to breast neoplasms, with 34 candidates recorded in miRTarBase, 32 in miRCancer, and 11 in miR2Disease. This result indicates the

dysregulation of these miRNAs in breast neoplasms and their documented associations with the disease.

Moreover, the dbDEMC database collects miRNA data with differential expression levels between cancer tissues and normal tissues from various sources such as the Gene Expression Omnibus, ArrayExpress, Sequence Read Archive, and the Cancer Genome Atlas. Among the top 50 candidates, 45 were confirmed in this database, indicating significant differences in their expression levels in breast neoplasms compared to normal tissues. Additionally, three candidate miRNAs, hsa-mir-135,[47] hsa-mir-19[48] and hsa-mir-513a,[49] have literature support and are marked as "Literature". Hsa-mir-135 is downregulated in breast neoplasm tissues, hsa-mir-19 and hsa-mir-513a are upregulated compared to normal tissues.

The top 50 candidates associated with hepatocellular carcinoma are presented in supplementary table S3, miRTarBase includes 5 candidates, miRCancer includes 29, and miR2Disease includes 4, underscoring the strong evidence supporting their involvement in hepatocellular carcinoma. These findings suggest significant differences in the expression levels of these candidates between hepatocellular carcinoma and normal tissues. Additionally, dbDEMC reports 48 candidates, indicating significant disparities in their expression levels between hepatocellular carcinoma and normal tissues. Moreover, one candidate, identified as "Literature" exhibits upregulation in expression levels in this disease,[50] as documented in relevant literature.

Supplementary table S4 presents the top 50 candidates associated with lymphoma. Among these candidates, 34 are supported by miRTarBase, indicating their established association with the disease. Moreover, the dbDEMC database includes 49 candidate genes, suggesting the involvement of these miRNAs in lymphoma. Notably, the expression levels of hsa-mir-466[51] are significantly upregulated in lymphoma patients compared to healthy individuals.

Prediction of new disease-associated miRNAs

To identify new miRNAs associated with diseases, we utilize all known miRNA-disease associations, as well as an equal number of randomly selected unobserved pairwise, to train the SFPred model. Afterward, we apply this trained model to predict potential miRNAs linked to 341 diseases. The top 50 candidate scores for each disease prediction can be found

in supplementary table S5.

Conclusions

We presented a miRNA-disease association prediction method to integrate the subgraph topologies, heterogeneous graph topology, and the context relationships among the pairwise features. The constructed subgraphs consisting of a miRNA (disease) node and its first-order neighbors is beneficial for formulating the importance between two nodes from their local topology perspective. The graph topology composed of all the miRNA and disease nodes was dynamically changed by measuring the edge weights based on MLP. The designed convolution transformer module was able to capture the context relationships among the pairwise features. The channel-level and feature-level residual strategies supplement the detailed features to the pairwise feature context learning. The 5-fold cross-validation experimental results indicated that SFPred achieves higher AUC and AUPR than the 8 compared methods. SFPred's higher recall rates demonstrates its ability in discovering the more genuine associations in the top-ranked candidates. Case studies on 3 diseases further show SFPred is able to get the reliable candidate miRNAs for the diseases.

Data and Software Availability

The source code and datasets are freely available at https://github.com/pingxuan-hlju/SFPred.

Acknowledgement

Supporting Information Available

- Prediction results for the different layer numbers of SDGCN and FCTransformer are listed in supplementary table S1.

- Prediction performance when changing the numbers of C_1 and C_2 are listed in supplementary table S2.

- The top 50 candidate miRNAs associated with hepatocellular carcinoma are listed in supplementary table S3.

- The top 50 candidate miRNAs associated with lymphoma are listed in supplementary table S4.

- The top 50 high-quality disease-related lncRNAs are listed in supplementary table S5.

References

(1) Ambros, V. The functions of animal microRNAs. *Nature* **2004**, *431*, 350–355.

(2) Bellina, A.; Malfatti, M. C.; Salgado, G.; Fleming, A. M.; Antoniali, G.; Gualandi, N.; La Manna, S.; Marasco, D.; Dassi, E.; Burrows, C. J.; others The Apurinic/Apyrimidinic Endodeoxyribonuclease 1 is an RNA G-quadruplex binding protein and regulates miR-92b expression in cancer cells. *bioRxiv* **2024**, 2024–02.

(3) Van Meter, E. N.; Onyango, J. A.; Teske, K. A. A review of currently identified small molecule modulators of microRNA function. *European journal of medicinal chemistry* **2020**, *188*, 112008.

(4) Elrebehy, M. A.; Abdelghany, T. M.; Elshafey, M. M.; Gomaa, M. H.; Doghish, A. S. miR-509–5p promotes colorectal cancer cell ferroptosis by targeting SLC7A11. *Pathology-Research and Practice* **2023**, *247*, 154557.

(5) Gao, B.; Li, R.; Song, X.; Hu, S.; Yang, F. miR-139-5p and miR-451a as a diagnostic biomarker in LUSC. *Pharmacogenomics and Personalized Medicine* **2023**, 313–323.

(6) Zhao, Y.; Chen, X.; Yin, J. Adaptive boosting-based computational model for predicting potential miRNA-disease associations. *Bioinformatics* **2019**, *35*, 4730–4738.

(7) Wang, D.; Wang, J.; Lu, M.; Song, F.; Cui, Q. Inferring the human microRNA functional similarity and functional network based on microRNA-associated diseases. *Bioinformatics* **2010**, *26*, 1644–1650.

(8) Xuan, P.; Han, K.; Guo, M.; Guo, Y.; Li, J.; Ding, J.; Liu, Y.; Dai, Q.; Li, J.; Teng, Z.; others Prediction of microRNAs associated with human diseases based on weighted k most similar neighbors. *PloS one* **2013**, *8*, e70204.

(9) Xuan, P.; Han, K.; Guo, Y.; Li, J.; Li, X.; Zhong, Y.; Zhang, Z.; Ding, J. Prediction of potential disease-associated microRNAs based on random walk. *Bioinformatics* **2015**, *31*, 1805–1815.

(10) 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 systems biology* **2010**, *4*, 1–9.

(11) Chen, M.; Deng, Y.; Li, Z.; Ye, Y.; He, Z. KATZNCP: a miRNA–disease association prediction model integrating KATZ algorithm and network consistency projection. *BMC bioinformatics* **2023**, *24*, 229.

(12) Ha, J.; Park, C.; Park, C.; Park, S. Improved prediction of miRNA-disease associations based on matrix completion with network regularization. *Cells* **2020**, *9*, 881.

(13) Chen, X.; Sun, L.-G.; Zhao, Y. NCMCMDA: miRNA–disease association prediction through neighborhood constraint matrix completion. *Briefings in bioinformatics* **2021**, *22*, 485–496.

(14) Peng, L.; Yang, C.; Huang, L.; Chen, X.; Fu, X.; Liu, W. RNMFLP: predicting circRNA–disease associations based on robust nonnegative matrix factorization and label propagation. *Briefings in bioinformatics* **2022**, *23*, bbac155.

(15) Jiang, H.; Yang, M.; Chen, X.; Li, M.; Li, Y.; Wang, J. miRTMC: a miRNA target prediction method based on matrix completion algorithm. *IEEE Journal of Biomedical and Health Informatics* **2020**, *24*, 3630–3641.

(16) Liu, Y.; Zeng, X.; He, Z.; Zou, Q. Inferring microRNA-disease associations by random walk on a heterogeneous network with multiple data sources. *IEEE/ACM transactions on computational biology and bioinformatics* **2016**, *14*, 905–915.

(17) Chen, M.; Liao, B.; Li, Z. Global similarity method based on a two-tier random walk for the prediction of microRNA–disease association. *Scientific reports* **2018**, *8*, 6481.

(18) Xie, X.; Wang, Y.; He, K.; Sheng, N. Predicting miRNA-disease associations based on PPMI and attention network. *BMC bioinformatics* **2023**, *24*, 113.

(19) Yao, H.-b.; Hou, Z.-j.; Zhang, W.-g.; Li, H.; Chen, Y. Prediction of microRNA-disease potential association based on sparse learning and multilayer random walks. *Journal of Computational Biology* **2024**, *31*, 241–256.

(20) Chen, X.; Wang, C.-C.; Yin, J.; You, Z.-H. Novel human miRNA-disease association inference based on random forest. *Molecular Therapy-Nucleic Acids* **2018**, *13*, 568–579.

(21) Liu, W.; Lin, H.; Huang, L.; Peng, L.; Tang, T.; Zhao, Q.; Yang, L. Identification of miRNA–disease associations via deep forest ensemble learning based on autoencoder. *Briefings in Bioinformatics* **2022**, *23*, bbac104.

(22) Chen, X.; Huang, L.; Xie, D.; Zhao, Q. EGBMMDA: extreme gradient boosting machine for MiRNA-disease association prediction. *Cell death & disease* **2018**, *9*, 3.

(23) He, K.; Wu, R.; Zhu, Z.; Li, J.; Lu, X. A probabilistic matrix decomposition method for identifying miRNA-Disease associations. Intelligent Computing Theories and Application: 16th International Conference, ICIC 2020, Bari, Italy, October 2–5, 2020, Proceedings, Part II 16. 2020; pp 399–410.

(24) Ha, J.; Park, C.; Park, C.; Park, S. IMIPMF: Inferring miRNA-disease interactions using probabilistic matrix factorization. *Journal of biomedical informatics* **2020**, *102*, 103358.

(25) Nigam, S.; Gjelaj, E.; Wang, R.; Wei, G.-W.; Wang, P. Machine Learning and Deep

Learning Applications in Magnetic Particle Imaging. *Journal of Magnetic Resonance Imaging* **2024**,

(26) Jin, C.; Shi, Z.; Lin, K.; Zhang, H. Predicting miRNA-disease association based on neural inductive matrix completion with graph autoencoders and self-attention mechanism. *Biomolecules* **2022**, *12*, 64.

(27) 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*, 2538–2546.

(28) Liao, Q.; Ye, Y.; Li, Z.; Chen, H.; Zhuo, L. Prediction of miRNA-disease associations in microbes based on graph convolutional networks and autoencoders. *Frontiers in Microbiology* **2023**, *14*, 1170559.

(29) Ruan, X.; Jiang, C.; Lin, P.; Lin, Y.; Liu, J.; Huang, S.; Liu, X. MSGCL: inferring miRNA–disease associations based on multi-view self-supervised graph structure contrastive learning. *Briefings in Bioinformatics* **2023**, *24*, bbac623.

(30) Liao, Q.; Fu, X.; Zhuo, L.; Chen, H. An efficient model for predicting human diseases through miRNA based on multiple-types of contrastive learning. *Frontiers in Microbiology* **2023**, *14*, 1325001.

(31) Peng, W.; He, Z.; Dai, W.; Lan, W. MHCLMDA: multihypergraph contrastive learning for miRNA–disease association prediction. *Briefings in Bioinformatics* **2024**, *25*, bbad524.

(32) Ouyang, D.; Liang, Y.; Wang, J.; Li, L.; Ai, N.; Feng, J.; Lu, S.; Liao, S.; Liu, X.; Xie, S. Hgclamir: Hypergraph contrastive learning with attention mechanism and integrated multi-view representation for predicting mirna-disease associations. *PLOS Computational Biology* **2024**, *20*, e1011927.

(33) Xu, L.; Li, X.; Yang, Q.; Tan, L.; Liu, Q.; Liu, Y. Application of bidirectional generative adversarial networks to predict potential miRNAs associated with diseases. *Frontiers in Genetics* **2022**, *13*, 936823.

(34) Wang, S.; Li, Y.; Zhang, Y.; Pang, S.; Qiao, S.; Zhang, Y.; Wang, F. Generative adversarial matrix completion network based on multi-Source data fusion for miRNA–disease associations prediction. *Briefings in Bioinformatics* **2023**, *24*, bbad270.

(35) Guo, Y.; Zhou, D.; Ruan, X.; Cao, J. Variational gated autoencoder-based feature extraction model for inferring disease-miRNA associations based on multiview features. *Neural Networks* **2023**, *165*, 491–505.

(36) Wang, C.-C.; Li, T.-H.; Huang, L.; Chen, X. Prediction of potential miRNA–disease associations based on stacked autoencoder. *Briefings in bioinformatics* **2022**, *23*, bbac021.

(37) Zhang, H.; Fang, J.; Sun, Y.; Xie, G.; Lin, Z.; Gu, G. Predicting miRNA-disease associations via node-level attention graph auto-encoder. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* **2022**, *20*, 1308–1318.

(38) Wang, W.; Chen, H. Predicting miRNA-disease associations based on graph attention networks and dual Laplacian regularized least squares. *Briefings in Bioinformatics* **2022**, *23*, bbac292.

(39) Li, G.; Fang, T.; Zhang, Y.; Liang, C.; Xiao, Q.; Luo, J. Predicting miRNA-disease associations based on graph attention network with multi-source information. *BMC bioinformatics* **2022**, *23*, 244.

(40) Jin, Z.; Wang, M.; Tang, C.; Zheng, X.; Zhang, W.; Sha, X.; An, S. Predicting miRNA-disease association via graph attention learning and multiplex adaptive modality fusion. *Computers in Biology and Medicine* **2024**, *169*, 107904.

(41) Li, Y.; Qiu, C.; Tu, J.; Geng, B.; Yang, J.; Jiang, T.; Cui, Q. HMDD v2. 0: a database for experimentally supported human microRNA and disease associations. *Nucleic acids research* **2014**, *42*, D1070–D1074.

(42) Kim, S.; Yeganova, L.; Wilbur, W. J. Meshable: searching PubMed abstracts by utilizing MeSH and MeSH-derived topical terms. *Bioinformatics* **2016**, *32*, 3044–3046.

(43) Lee, M. Mathematical analysis and performance evaluation of the gelu activation function in deep learning. *Journal of Mathematics* **2023**, *2023*, 4229924.

(44) Xuan, P.; Wang, D.; Cui, H.; Zhang, T.; Nakaguchi, T. Integration of pairwise neighbor topologies and miRNA family and cluster attributes for miRNA–disease association prediction. *Briefings in Bioinformatics* **2022**, *23*, bbab428.

(45) Wei, Y.; Wang, X.; Nie, L.; He, X.; Hong, R.; Chua, T.-S. MMGCN: Multi-modal graph convolution network for personalized recommendation of micro-video. Proceedings of the 27th ACM international conference on multimedia. 2019; pp 1437–1445.

(46) 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*, bbad094.

(47) Xu, L.-M.; Zhang, J.; Ma, Y.; Yuan, Y.-J.; Yu, H.; Wang, J.; Cao, X.-C.; Zhu, L.; Wang, P. MicroRNA-135 inhibits initiation of epithelial-mesenchymal transition in breast cancer by targeting ZNF217 and promoting m6A modification of NANOG. *Oncogene* **2022**, *41*, 1742–1751.

(48) Milioli, H. H.; Tishchenko, I.; Riveros, C.; Berretta, R.; Moscato, P. Basal-like breast cancer: molecular profiles, clinical features and survival outcomes. *BMC medical genomics* **2017**, *10*, 1–17.

(49) Fan, Y.; Wang, J.; Jin, W.; Sun, Y.; Xu, Y.; Wang, Y.; Liang, X.; Su, D. CircNR3C2 promotes HRD1-mediated tumor-suppressive effect via sponging miR-513a-3p in triple-negative breast cancer. *Molecular cancer* **2021**, *20*, 1–22.

(50) Nagy, Á.; Lánczky, A.; Menyhárt, O.; Győrffy, B. Validation of miRNA prognostic power in hepatocellular carcinoma using expression data of independent datasets. *Scientific reports* **2018**, *8*, 9227.

(51) Dawidowska, M.; Jaksik, R.; Drobna, M.; Szarzyńska-Zawadzka, B.; Kosmalska, M.; Sedek, Ł.; Machowska, L.; Lalik, A.; Lejman, M.; Ussowicz, M.; others Comprehen-

sive investigation of miRNome identifies novel candidate miRNA-mRNA interactions implicated in T-cell acute lymphoblastic leukemia. *Neoplasia* **2019**, *21*, 294–310.

TOC Graphic