

RESEARCH

Open Access



Graph embedding ensemble methods based on the heterogeneous network for lncRNA-miRNA interaction prediction

Chengshuai Zhao¹, Yang Qiu¹, Shuang Zhou², Shichao Liu¹, Wen Zhang^{1*} and Yanqing Niu^{3*}

From 2019 IEEE International Conference on Bioinformatics and Biomedicine (IEEE BIBM 2019)
San Diego, CA, USA. 18-21 November 2019

Abstract

Background: Researchers discover lncRNA-miRNA regulatory paradigms modulate gene expression patterns and drive major cellular processes. Identification of lncRNA-miRNA interactions (LMIs) is critical to reveal the mechanism of biological processes and complicated diseases. Because conventional wet experiments are time-consuming, labor-intensive and costly, a few computational methods have been proposed to expedite the identification of lncRNA-miRNA interactions. However, little attention has been paid to fully exploit the structural and topological information of the lncRNA-miRNA interaction network.

Results: In this paper, we propose novel lncRNA-miRNA prediction methods by using graph embedding and ensemble learning. First, we calculate lncRNA-lncRNA sequence similarity and miRNA-miRNA sequence similarity, and then we combine them with the known lncRNA-miRNA interactions to construct a heterogeneous network. Second, we adopt several graph embedding methods to learn embedded representations of lncRNAs and miRNAs from the heterogeneous network, and construct the ensemble models using two ensemble strategies. For the former, we consider individual graph embedding based models as base predictors and integrate their predictions, and develop a method, named GEEL-PI. For the latter, we construct a deep attention neural network (DANN) to integrate various graph embeddings, and present an ensemble method, named GEEL-FI. The experimental results demonstrate both GEEL-PI and GEEL-FI outperform other state-of-the-art methods. The effectiveness of two ensemble strategies is validated by further experiments. Moreover, the case studies show that GEEL-PI and GEEL-FI can find novel lncRNA-miRNA associations.

Conclusion: The study reveals that graph embedding and ensemble learning based method is efficient for integrating heterogeneous information derived from lncRNA-miRNA interaction network and can achieve better performance on LMI prediction task. In conclusion, GEEL-PI and GEEL-FI are promising for lncRNA-miRNA interaction prediction.

Keywords: lncRNA-miRNA interactions, Graph embedding, Ensemble learning, Attention mechanism

* Correspondence: zhangwen@mail.hzau.edu.cn;
niuyanqing@mail.scuec.edu.cn

¹College of Informatics, Huazhong Agricultural University, Wuhan 430070, China

³School of Mathematics and Statistics, South-Central University for Nationalities, Wuhan 430074, China

Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

Non-coding RNAs (ncRNAs), including long non-coding RNA (lncRNA), miRNA, snRNA, are a category of RNAs that are not translated into functional proteins. A surge of studies has betrayed that ncRNAs have regulatory functions in biological processes [1–4]. lncRNAs are a class of ncRNAs with more than 200 nucleotides (nt), playing important roles in gene imprinting, immune response, and chromatin remodeling [1, 2]. MiRNAs are a category of single-stranded, endogenous, evolutionally conserved ncRNAs with 20–25 nt, which are involved in diverse biological processes, such as the regulation of metabolism, cell differentiation, gene expression, embryonic development, and apoptosis [3–5]. lncRNA-miRNA regulatory paradigms modulate gene expression patterns that drive major cellular processes (e.g., cell proliferation, cell differentiation, and cell death) which are central to mammalian physiologic and pathologic processes [6]. Furthermore, it has been found that both lncRNAs and miRNAs relate closely to severe diseases [7, 8]. Therefore, a critical key to reveal the mechanism of associated biological processes and diseases is to characterize various functions of lncRNAs and miRNAs.

lncRNAs and miRNAs produce complicated effects through their interactions with other biological molecules such as DNAs, RNAs, and proteins, thus conducting researches on lncRNA-biomolecule interactions contributes to portraying the functions of lncRNAs and miRNAs [9–11]. Lately, some studies have demonstrated that lncRNAs can be used as a decoy or sponge to regulate miRNAs' behavior [12], indicating that identifying lncRNA-miRNA interactions (LMIs) helps to understand the functions of lncRNAs and miRNAs.

In earlier researches, unknown LMIs were identified through wet experiments. However, due to the laborious, costly, and time-consuming process of wet methods, it is more common to refine the candidate list in silico prediction for further validation experiments, in order to accelerate the identification of LMIs.

Recently, plenty of computational approaches have been proposed to predict LMIs. Huang et al. [13] propose a two-way diffusion model EPLMI for lncRNA-miRNA interaction prediction, which considers the known LMIs as a bipartite network. Huang et al. [14] develop GBCF, which builds a Bayesian collaborative filtering model using sequence, expression profiles, and target genes. Hu et al. [15] introduce a model, namely INLMI, which is based on the sequence similarity network and the expression similarity network. Zhang et al. [16] propose SLNPM which constructs the integrated similarity-based graph exploiting LMIs and genomic sequences, and implement a label propagation process on graphs for LMI prediction. These pioneers have produced good performances, but there still exist some

limitations. On the one hand, some of the existing methods (e.g., EPLMI, GBCF and INLMI) heavily rely on biological features of lncRNAs and miRNAs, such as target gene information or expression profiles, which are not obtainable for all lncRNAs (miRNAs). On the other hand, the structure of the LMI network cannot be fully in pervious methods; nevertheless, it is fairly crucial to effectively utilize the structural and topological information of the LMI network for link inference.

Graph embedding learning (a.k.a. network representation learning), can be employed to preserve the structural property of the graph and map nodes of the graph into low-dimensional space, attracting widespread attention recently. To the best of our knowledge, some graph embedding methods have been exploited to reveal unknown associations between biomedical entities [17–19]. Motivated by the previous work in bioinformatics, we use graph embedding methods to capture information from LMI network.

Ensemble learning is one of the research hotspots in machine learning and pattern recognition. To date, ensemble learning methods have been increasingly used in computational biology because of their unique advantages in managing small samples, complex data structures, and high dimensionality [20]. Ensemble learning is an efficient technique that aggregates multiple machine learning models to achieve overall high prediction accuracy and good generalization [21]. It usually performs better than individual methods. Inspired by pioneering works [22–27], we adopt ensemble strategies to integrate individual predictions and embeddings to enhance the performance of LMI prediction.

In this paper, we propose novel LMI prediction methods based on graph embedding and ensemble strategies. Firstly, we calculate similarity based on lncRNA sequences and miRNA sequences and construct a heterogeneous network by combining them with the known LMIs. Secondly, we utilize five graph embedding methods (i.e., Laplacian Eigenmaps [28], HOPE [29], GraRep [30], DeepWalk [31], and GAE [32]) to capture structural information from the heterogeneous network, and learn the representation of lncRNAs and miRNAs. Later, we represent the lncRNA-miRNA pairs by merging lncRNA's representation with miRNA's representation, and build ensemble models based on pair features. As the extension of our previous work [33], we consider two ensemble strategies. For the former, we consider all the individual graph embedding based models as base predictors and integrate their predictions to develop a prediction method, named GEEL-PI. As for the latter, we construct a deep attention neural network (DANN) to learn lncRNA-miRNA pair representations by combining various graph embeddings, and develop a method, named GEEL-FI. The experimental results demonstrate that the

proposed methods GEEL-PI and GEEL-FI can predict lncRNA-miRNA interactions with higher accuracy compared with other state-of-the-art methods. Moreover, the effectiveness of the prediction integration and attention network is proved by extensive experiments. Furthermore, we conduct case studies to validate the predicted LMIs which do not exist in our dataset. In conclusion, both GEEL-PI and GEEL-FI are useful for predicting LMIs. Our contribution can be summarized as:

- (1) We consider a variety of graph embedding methods to learn the embedded representations from the lncRNA-miRNA heterogeneous network.
- (2) We introduce a deep attention neural network to learn high-level sophisticated representations by focusing on different aspects of embedded representations.
- (3) We consider two different ensemble strategies in this work. Then we design comprehensive experiments to compare them and analyze their effectiveness.

Results and discussion

Evaluation metrics

In this paper, we implement 5-fold cross-validation (5-CV) to evaluate our models. The following metrics are adopted in our experiments: the area under the precision-recall curve (AUPR), the area under the receiver-operating characteristic curve (AUC), F-measure (F1), accuracy (ACC), recall (REC), specificity (SPEC), and precision (PRE).

Parameter settings

In this study, both GEEL-PI and GEEL-FI have two major components: graph embedding and ensemble learning. Here, we introduce parameter settings.

Parameter settings for graph embedding methods

In this study, both GEEL-PI and GEEL-FI adopt five graph embedding methods: LE, GraRep, HOPE, DeepWalk, and GAE to learn representations of lncRNAs and miRNAs. The graph embedding methods are implemented by BioNEV [19].

Here, we discuss the parameter settings of five graph embedding methods. Firstly, we fix the representation dimension of all the graph embedding methods θ as 120 and consider other specified parameters of each graph embedding method. For GraRep, we consider the k th transition probability matrix $k\text{-step} \in \{1, 2, 3, 4\}$. For DeepWalk, we fix the walk length t as 80, and consider the combinations of window size $w \in \{10, 20, 30, 40\}$ and walk per vertex $\gamma \in \{10, 20, 30, 40\}$. For GAE, we consider autoencoder and

variational autoencoder respectively, and select the size of hidden layers $\beta \in \{32, 64, 128, 256, 512, 1024\}$. For the aforementioned graph embedding methods, we adopt the optimal parameters which achieve the highest AUPR scores.

Parameter settings for ensemble methods

In this paper, we propose two ensemble strategies: prediction combination for GEEL-PI and attention neural network for GEEL-FI. The detailed parameter settings are described below.

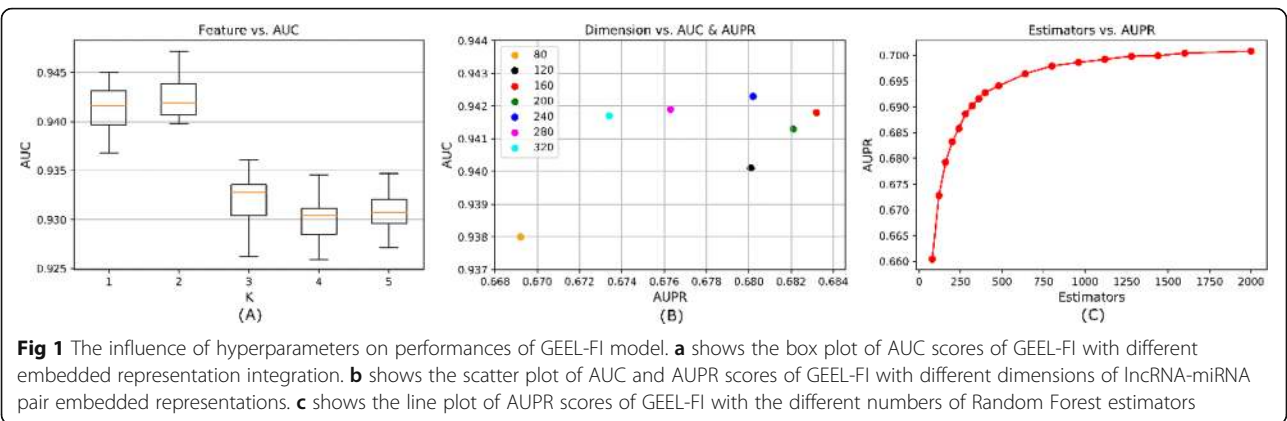
For GEEL-PI, Random Forest and Logistic Regression are implemented by “scikit-learn” [34] where default hyperparameters are adopted. For the logistic regression, we additionally adopt L2 regulation with default parameters.

For GEEL-FI, we tune the following parameter settings: (1) the number of hidden layers μ and the size of hidden layers β in DANN (2) the embedded representation vectors ε involved in the feature fusion (3) the dimension of lncRNA-miRNA pair features θ (4) the number of estimators η in Random Forest classifier.

In the attention layer of DANN, we design two groups of attention weights for individual lncRNA-miRNA pair features. For fully-connected layers, we consider different combinations of the parameters: number of hidden layers $\mu \in \{1, 2, 3, 4\}$, size of hidden layers $\beta \in \{480, 240, 120, 60, 30\}$. Then we use the grid search to optimize these parameters according to their performances on 5-CV. Finally, we design a two-hidden-layer neural network, and the size of each layer is 120 and 60 respectively.

As for the embedded representation vectors ε , we consider combinations of embedded representation vectors for merged lncRNA-miRNA pair features. For individual graph embedding methods, we implement 5-CV for 20 times. In the light of AUC and AUPR scores, we reorder five graph embedding methods as GraRep, LE, GAE, HOPE, DeepWalk. And then we select the top K features as the candidates for lncRNA-miRNA pair features. Here we visualize the trend of AUC scores over the combination of top K features in Fig. 1 (a). The fused feature based on the top 2 graph embedding methods (i.e. GraRep and LE) owns the best performances. Hence, we adopt $\varepsilon = \{\text{GraRep}, \text{LE}\}$.

We consider the dimension of lncRNA-miRNA pair features $\theta \in \{80, 120, 160, 240, 280, 320\}$ with the consideration of the AUPR and AUC scores. As presented in Fig. 1 (b), fused features of 160 dimensions have a higher AUPR score and that of 240 dimensions has a higher AUC score. In the subsequent experiment, pair features of 160 dimensions achieve better performance, thus we set $\theta = 160$.



Eventually, we consider the number of estimators η in Random Forest from 80 to 2000. In Fig. 1 (c), when the number of estimators equals to 2000, the AUPR score has little improvement. Considering computational efficiency and time costs, we set $\eta = 2000$.

After analysis above, we adopt $\mu = 2$, $\beta = \{240, 120\}$, $\varepsilon = \{GraRep, LE\}$, $\theta = 160$ and $\eta = 2000$ for GEEL-FI. All the parameters used in graph embedding ensemble methods are summarized in Table 1.

Comparison with state-of-the-art methods

Here, we compare our models with several state-of-the-art methods including EPLMI [13], INLMI [15], and SLNPM [35]. EPLMI infers link probability according to the similarity between lncRNA and miRNA expression profiles. Specifically, EPLMI constructs a bipartite network using known lncRNA-miRNA interactions and exploits lncRNA (miRNA) expression profile information via the network for LMI prediction. INLMI integrates the sequence similarity and the expression similarity, and adopts a two-way diffusion algorithm to infer LMIs. SLNPM predicts LMIs by implementing a label propagation algorithm on two biomedical entities similarity graphs respectively. EPLMI and SLNPM are

implemented according to the descriptions in the publications, then we evaluate the above models on our dataset by using 5-fold cross-validation experiments.

As shown in Table 2, GEEL-FI achieves the best AUPR score (0.7011), and the best AUC score (0.9578), and GEEL-PI achieves the second-best AUPR score (0.7004) and AUC score (0.9537), which significantly outperform other state-of-art methods. The substantial improvement of our models could be attributed to two factors: (1) GEEL-PI and GEEL-FI make the best of the structural properties implied in the lncRNA-miRNA heterogeneous network by employing graph embedding. (2) GEEL-PI and GEEL-FI adopt ensemble strategies (i.e. prediction integration and feature integration) to integrate multi-view information.

In computational experiments, the top-ranked predictions are critical to reflect the performances of models. Here, we calculate the recall and precision of the aforementioned models on top-ranked predictions ranging from the top 100 to the top 1000. As presented in Fig. 2 (a), both GEEL-PI and GEEL-FI achieve best recall scores over all thresholds. For instance, when checking the top 500 predictions, GEEL-PI and GEEL-FI achieve recall scores of 0.5719 and 0.5706, nevertheless, the

Table 1 Parameter settings for proposed methods

Methods	Components	Parameters
Graph embedding methods	Representation vector	dimension θ : 120
	GraRep	k -step: 1
	DeepWalk	walk length t : 80, walk per vertex γ : 30, window size w : 30
	GAE	variational Autoencoder, hidden size β : 512
GEEL-PI	Random Forest	default parameters
	Logistic Regression	L2 regulation with default parameters
GEEL-FI	DANN	hidden layers μ : 2, hidden size β : {240,120}
	Representation vector	embeddings ε : {GraRep, LE}
	Pair feature	dimension θ : 160
	Random Forest	estimators η : 2000

recall scores for SLNPM, EPLMI, INLMI remain 0.5283, 0.0921, 0.0884 respectively. Similarly, both GEEL-PI and achieve better precision scores than other benchmark methods as given in Fig. 2 (b). For example, both GEEL-PI and GEEL-FI can infer 86% real interactions in the top 500 predictions, whereas SLNPM, EPLMI, INLMI can only find 80, 10, 10% real interactions. Therefore, both GEEL-PI and GEEL-FI are preferable for LMI prediction compared with other state-of-the-art methods.

Effect of ensemble learning

In this paper, we adopted two ensemble strategies to integrate heterogeneous information and develop our methods: GEEL-PI and GEEL-FI. In the following, we evaluate the performances of base predictors and our methods by 20 runs of 5-CV and discuss how the ensemble strategies improve performances.

As demonstrated in Table 3, generally, these graph embedding based models could produce satisfactory performances, achieving AUPR scores > 0.65 and AUC scores > 0.92. In terms of the standard deviations of 20 runs of experiments, all these prediction models could lead to stable results. The experimental results indicate that graph embedding methods can efficiently capture inherent properties from the lncRNA-miRNA heterogeneous network for LMI inference.

Further, we integrate above five graph embedding based methods by ensemble strategies to enhance the accuracy of the model. GEEL-PI integrates different prediction scores from five graph embedding-based predictors, achieving AUPR score of 0.7004 and AUC score of 0.9537. GEEL-FI attentively integrates lncRNA and miRNA representations to obtain distinctive lncRNA-miRNA pair features, achieving AUPR score of 0.7011 and AUC score of 0.9578. Both GEEL-PI and GEEL-FI achieve superior performances compared with base predictors, which indicates our ensemble strategies can contribute to higher accuracy for LMI prediction.

To evaluate the generalization ability of our ensemble models, we design an experiment on different sparsity of the heterogeneous network by removal of a certain proportion of links. In the experiments, we randomly delete 10, 20, 30, and 40% of LMIs in the heterogeneous network. Then, we build the base predictors and the ensemble models on the networks with fewer interactions.

Table 4 reports the AUPR scores of different prediction methods. As we can observe, the ensemble models GEEL-PI and GEEL-FI produce higher AUPR scores than all the base predictors as the ratios of removed links ranging from 10 to 40%. More importantly, when the network becomes sparser, the performances of the ensemble models are less affected than other individual predictors. For instance, when the number of removed interactions ranging from 10 to 20%, the AUPR scores of LE, GraRep, HOPE, DeepWalk, GAE, GEEL-PI and GEEL-FI reduce by 2.7, 2.1, 2.1, 2.3, 4.3, 1.7, and 1.7% respectively, which verifies the generalization ability and robustness of our ensemble models.

In conclusion, integrating individual graph embedding based models with ensemble learning can effectively improve accuracy, generalization ability, and robustness in LMI prediction.

Effect of attention network

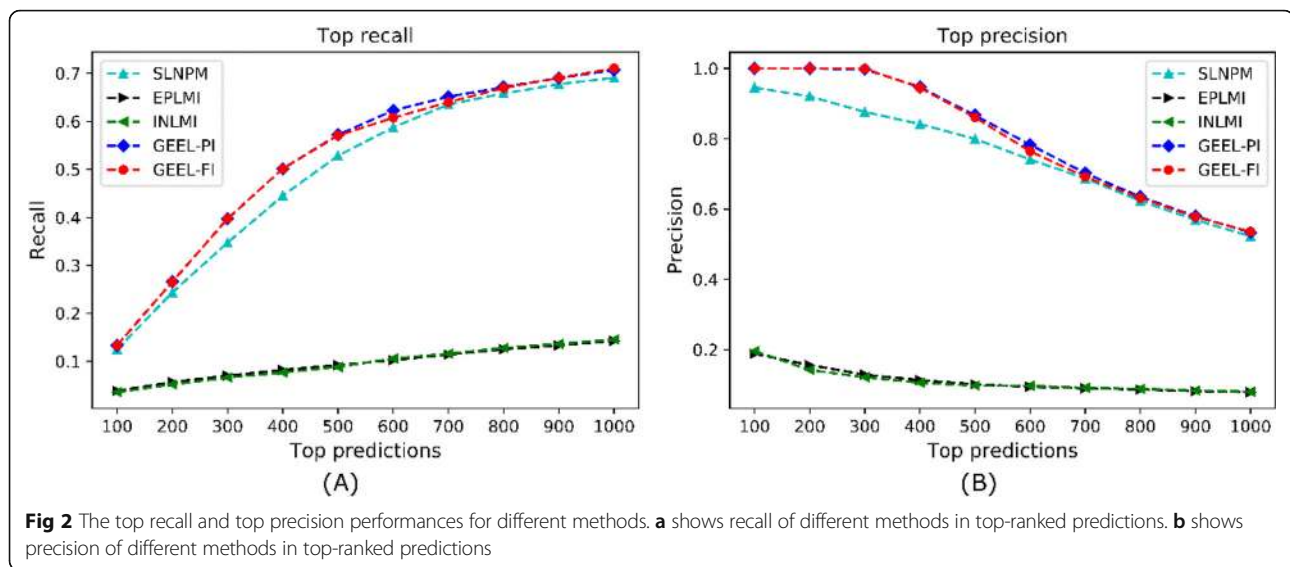
In the design of GEEL-FI, we consider a deep attention neural network to integrate graph embeddings as the ensemble strategy. DANN learn lncRNA-miRNA pair features by capturing the different aspects of representation vectors. To validate the effectiveness of the attention mechanism, we evaluate the performances of GEEL-FI and our designed comparison method on LMI prediction.

To validate the effect of attention network on feature fusion, we design the comparison variant as GEEL-F, which merges diverse embedded lncRNA and miRNA representations directly, without considering the different importance of embedded representations. For i th lncRNA and j th miRNA, the merged representation of lncRNA is defined as $L_i = \sum_{k \in S} l_i^k$ and the merged representation of miRNA is defined as $M_j = \sum_{k \in S} m_j^k$, where S

is a set of lncRNA and miRNA representations learned by graph embedding methods. And the lncRNA and miRNA pair feature is computed as $F_{ij} = [L_i; M_j]$. We construct GEEL-FI and GEEL-F based on learned graph embeddings. To validate the effectiveness of our attention mechanism at a larger scale, we choose the K embeddings for the fused feature. Here we respectively adopt $S = \{\text{GraRep}\}$, $\{\text{GraRep}, \text{GAE}\}$, $\{\text{GraRep}, \text{HOPE}, \text{DeepWalk}\}$, $\{\text{GraRep}, \text{HOPE}, \text{DeepWalk}, \text{LE}\}$ and $\{\text{LE}, \text{GraRep}, \text{HOPE}, \text{DeepWalk}, \text{and GAE}\}$ with respect to $K = \{1, 2, 3, 4, 5\}$ as our benchmarks to compare the performances of GEEL-F and GEEL-FI for LMI prediction. As shown in Fig. 3, given $K = \{1, 2, 3, 4, 5\}$, GEEL-FI achieves AUPR scores of 0.6810, 0.6838, 0.6539, 0.6538 and 0.6670 which outperforms 0.6805, 0.6725, 0.6493, 0.6487 and 0.6541 respectively. The experimental result demonstrates the utilization of attention mechanism can contribute to better performance for LMI prediction.

Table 2 Performances of different methods

Methods	AUPR	AUC	F1	ACC	REC	SPEC	PRE
EPLMI	0.0706	0.8494	0.1055	0.9939	0.1373	0.9962	0.0883
INLMI	0.0723	0.8477	0.1086	0.9935	0.1531	0.9956	0.0867
SLNPM	0.6207	0.9165	0.6652	0.9972	0.6331	0.9988	0.7016
GEEL-PI	0.7004	0.9537	0.6933	0.9977	0.5945	0.9995	0.8342
GEEL-FI	0.7011	0.9578	0.6915	0.9977	0.5790	0.9996	0.8604



Therefore, we can conclude that our deep attention neural network can effectively merge multiple embedded lncRNA and miRNA representations and learn better lncRNA-miRNA pair features for LMI prediction.

To further probe into how the attention network captures different aspects of embedded representations, we fix K as 5 and implement 5-CV for 20 times. Then we visualize the attention weights of lncRNA representations and miRNA representations learned by attention neural network. In Fig. 4, we can observe that (1) for lncRNAs, DANN generally pays much attention to the GAE-based embeddings, and for miRNA, it assigns higher attention weights to GraRep-based embeddings, which indicates the graph embedding based on neural network and matrix factorization method are efficient in LMI prediction. (2) furthermore, attention weights vary with lncRNA sequences and miRNA sequences in each fold, which validates DANN can adaptively adjust its attention to learn distinctive lncRNA-miRNA pair features according to specific lncRNA and miRNA data.

Consequently, our deep attention neural network can learn high-level sophisticated representations of lncRNA-

miRNA pairs and enhance the performances of GEEL-FI on LMI prediction.

Case studies

The primary goal of computational methods is to refine the candidate list and guide further validation experiments. Here, we conduct case studies to demonstrate the practical capability of the proposed method for unknown LMI inference. Firstly, we train the model on our dataset. Then, we employ our model to score unlabeled lncRNA-miRNA pairs. Later, we validate the prediction result by a comprehensive datasets starBase [36]. Here, we list the top 10 LMIs in Table 5. As we can observe, both GEEL-PI and GEEL-FI can correctly infer 8 LMIs among their top 10 predictions. For instance, our proposed model can accurately predict that lncRNA lnc-ACER2-1:1 can interact with miRNA hsa-miR-106a-5p. ACER2 is one of the human alkaline ceramidases, and can produce lncRNA lnc-ACER2-1. MiRNA hsa-miR-106a-5p can participate in various biological processes, and are involved in severe diseases (e.g., gastric carcinoma and glioblastoma) [37, 38]. Some researchers have

Table 3 Performances of based predictors and the ensemble models

Embedding	AUPR	AUC	F1	ACC	REC	SPEC	PRE
LE	0.6654 ± 0.0033	0.9430 ± 0.0017	0.6592 ± 0.0040	0.9976 ± 0.0001	0.5429 ± 0.0079	0.9995 ± 0.0001	0.8420 ± 0.0144
GraRep	0.6805 ± 0.0037	0.9417 ± 0.0019	0.6818 ± 0.0036	0.9977 ± 0.0001	0.5703 ± 0.0066	0.9996 ± 0.0001	0.8498 ± 0.0137
HOPE	0.6573 ± 0.0036	0.9281 ± 0.0022	0.6796 ± 0.0035	0.9976 ± 0.0001	0.5813 ± 0.0087	0.9994 ± 0.0001	0.8198 ± 0.0134
DeepWalk	0.6511 ± 0.0037	0.9383 ± 0.0018	0.6463 ± 0.0051	0.9974 ± 0.0001	0.5452 ± 0.0133	0.9994 ± 0.0001	0.7993 ± 0.0248
GAE	0.6664 ± 0.0031	0.9292 ± 0.0023	0.6754 ± 0.0033	0.9976 ± 0.0001	0.5666 ± 0.0086	0.9995 ± 0.0001	0.8395 ± 0.0185
GEEL-PI	0.7004 ± 0.0035	0.9537 ± 0.0022	0.6933 ± 0.0032	0.9977 ± 0.0001	0.5945 ± 0.0063	0.9995 ± 0.0001	0.8342 ± 0.0128
GEEL-FI	0.7011 ± 0.0030	0.9578 ± 0.0013	0.6915 ± 0.0029	0.9977 ± 0.0001	0.5790 ± 0.0063	0.9996 ± 0.0001	0.8604 ± 0.0124

Table 4 Performances on the network of different sparsity

Removal ratio	LE	GraRep	HOPE	DeepWalk	GAE	GEEL-PI	GEEL-FI
10%	0.6496	0.6666	0.6448	0.6341	0.6537	0.6858	0.6838
20%	0.6323	0.6524	0.6311	0.6192	0.6254	0.6744	0.6719
30%	0.6124	0.6355	0.6171	0.5982	0.6206	0.6561	0.6579
40%	0.5884	0.6156	0.5959	0.5761	0.6009	0.6347	0.6372

discovered that the expression of hsa-miR-106a-5p is down-regulated in breast tissues, and ACER2 could serve as a target gene of hsa-miR-106a-5p [39]. Whereas, the interaction between lnc-COL6A3-5:1 and hsa-miR-4500 is to be confirmed in the future. In general, both GEEL-PI and GEEL-FI are effective tools to indicate novel interactions between lncRNA and miRNA.

Conclusions

lncRNAs and miRNAs are critical to cellular processes, and inferring their interactions contributes to betraying the mechanism of complicated disease. In this paper, we propose novel graph embedding ensemble learning methods: GEEL-PI and GEEL-FI. Comparison with other state-of-art methods demonstrates both GEEL-PI and GEEL-FI achieve higher accuracy performances for LMI prediction. The adoption of graph embedding methods overcomes the limitation of traditional features, and makes our model efficiently capture the inherent structural properties of LMI heterogeneous network. Further experiments indicate that ensemble learning and

attention mechanism are powerful to enhance accuracy, generalization ability, and robustness of LMI prediction model. Moreover, the case studies are also performed to prove the practical capability of our methods. In conclusion, both GEEL-PI and GEEL-FI are promising for LMI prediction.

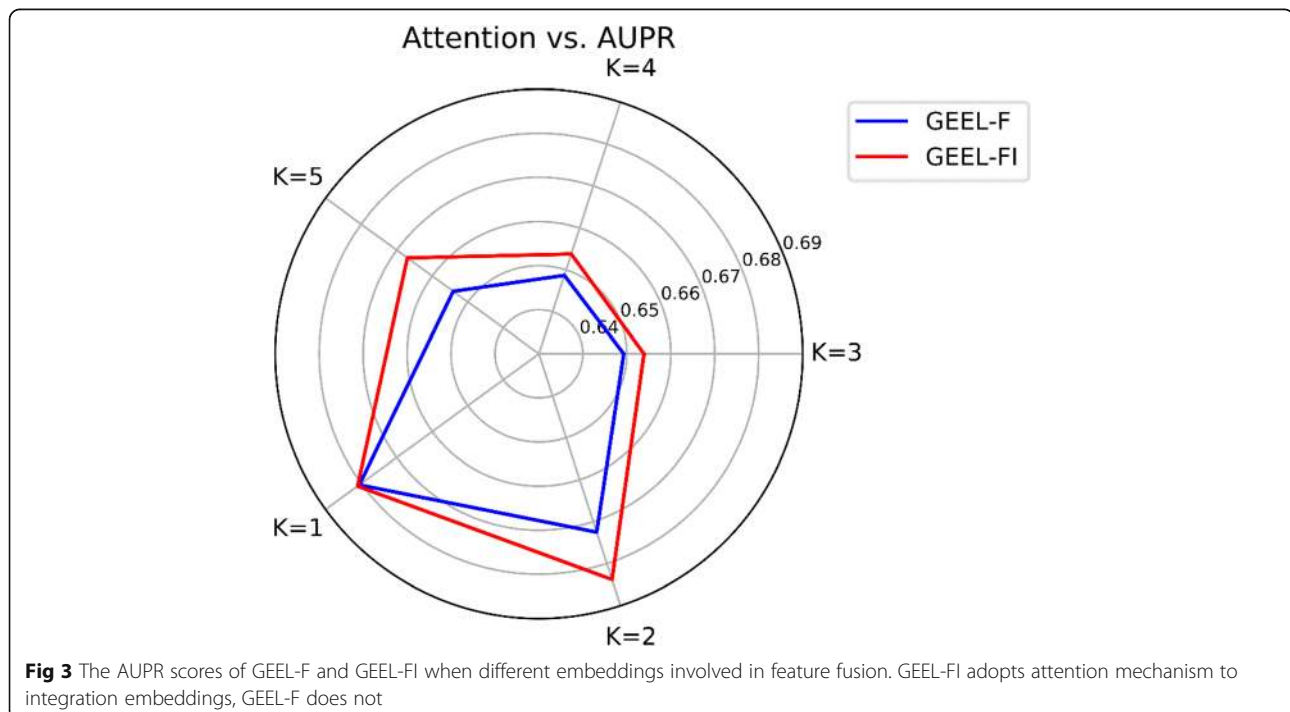
Datasets and methods

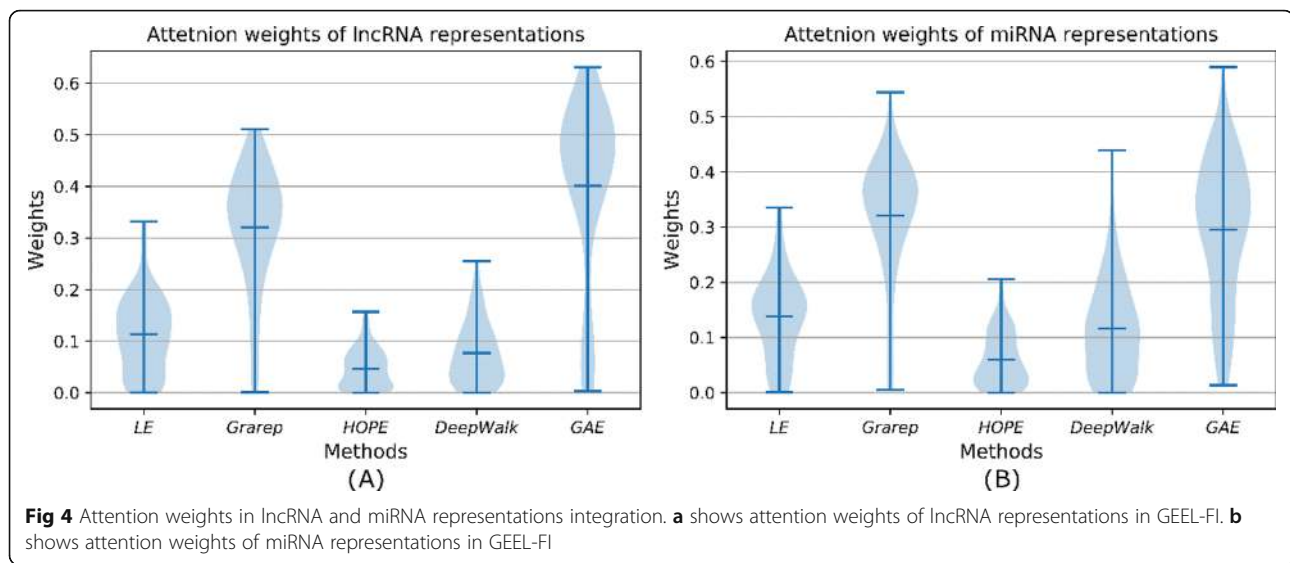
Datasets

We collect 8091 experimentally verified lncRNA-miRNA interactions from the lncRNASNP dataset [40]. After removing duplicated interactions, we obtain 5118 interactions between 780 lncRNAs and 275 miRNAs. We then download lncRNA sequences from NONCODE dataset [41] and miRNA sequences from miRBase dataset [42] separately. Ultimately, we compile our dataset with 3784 interactions between 642 lncRNAs and 275 miRNAs.

Heterogeneous network

To model the complicated relationship between biomedical entities, we design a lncRNA-miRNA heterogeneous





network by integrating the known LMIs with the sequence similarity, as shown in Fig. 5 (a).

Given r lncRNAs and t miRNAs, the interaction matrix can be denoted by $A \in \mathbb{R}^{r \times t}$, where $A(i, j) = 1$ if i th lncRNA and j th miRNA are interacting, otherwise $A(i, j) = 0$. Our previous work [35] indicates that the pairwise similarity between biomedical entities (i.e. lncRNA and miRNA sequence similarity) can help to infer interactions. Therefore, same as our previous work, we extract 5-spectrum feature [43] from lncRNA (miRNA) sequence and then calculate similarity by linear neighborhood similarity measure (LNS) [35]. In this way, we acquire lncRNA similarity matrix $S_l \in \mathbb{R}^{r \times r}$ and miRNA similarity matrix $S_m \in \mathbb{R}^{t \times t}$, where $S(i, j)$ is the similarity score between i th and j th lncRNAs (miRNAs). Further, for a single biomedical entity, we consider the top 10 most similar

entities as its immediate neighborhoods, and obtain adjacency matrix $W_l \in \mathbb{R}^{r \times r}$ and $W_m \in \mathbb{R}^{m \times m}$ from S_l and S_m separately. Ultimately, we regard biomedical entities (i.e. a lncRNAs and a miRNAs) as nodes and their relationships (i.e. LMs, lncRNA-lncRNA similarity and miRNA-miRNA similarity) as edges to construct the heterogeneous network H :

$$H = \begin{bmatrix} W_l & A \\ A^T & W_m \end{bmatrix} \in \mathbb{R}^{(r+t) \times (r+t)} \quad (1)$$

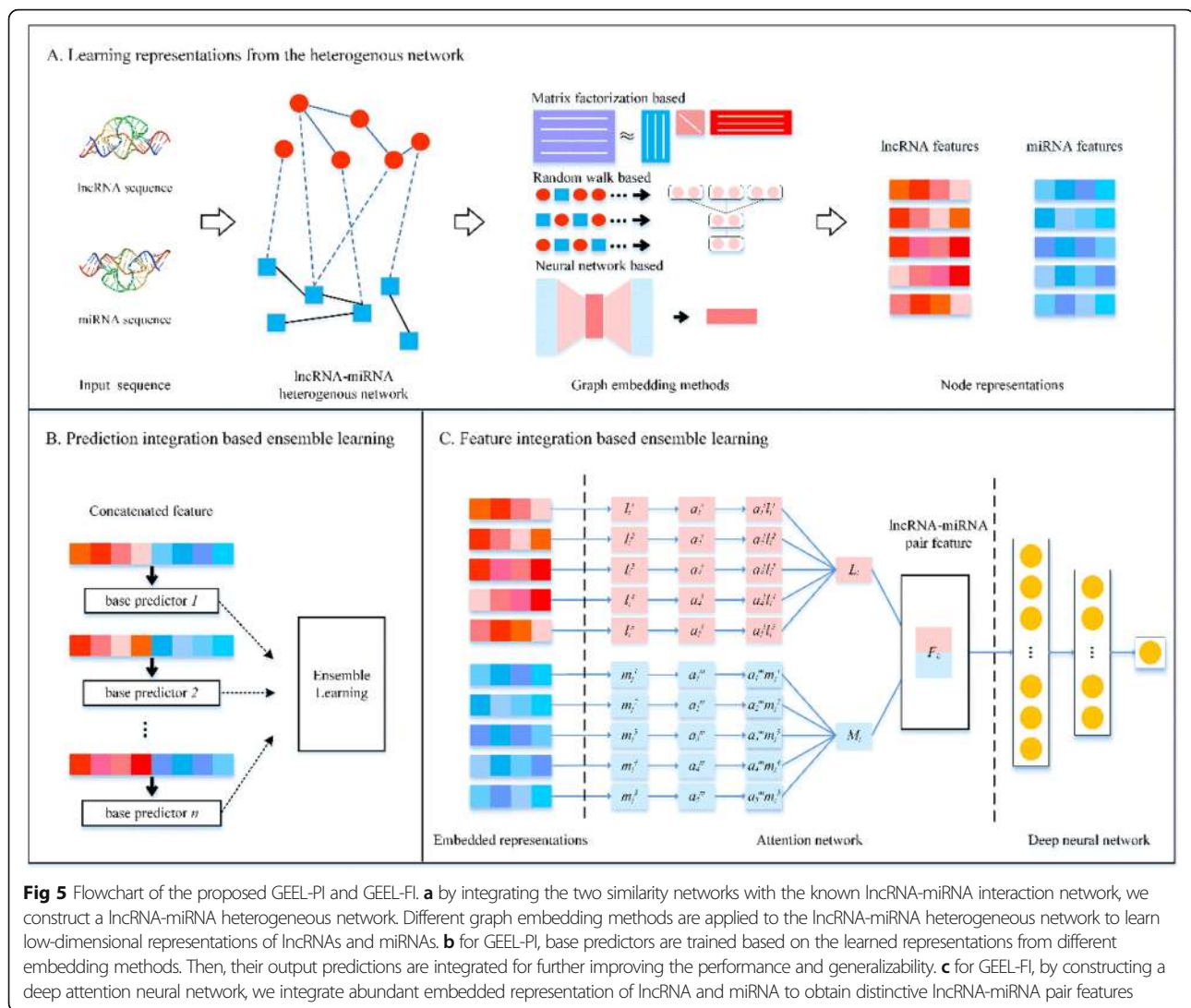
where A^T denotes the transpose of the matrix A .

Graph embedding methods

To fully exploit the topological properties of the heterogeneous network, we choose graph embedding methods

Table 5 Top 10 prediction of GEEL-PI and GEEL-FI

Rank	GEEL-PI			GEEL-FI		
	LncRNAs	MiRNAs	Evidence	LncRNAs	MiRNAs	Evidence
1	lnc-COL6A3-5:1	hsa-miR-4500	×	lnc-COL6A3-5:1	hsa-miR-4500	×
2	lnc-ACER2-1:1	hsa-miR-17-5p	✓	lnc-ALYREF-1:1	hsa-miR-372-3p	✓
3	lnc-FAS-1:1	hsa-miR-302b-3p	✓	MIR17HG:2	hsa-miR-520a-3p	✓
4	lnc-PDK3-1:1	hsa-miR-93-5p	✓	lnc-PDK3-1:1	hsa-miR-302d-3p	✓
5	lnc-ACER2-1:1	hsa-miR-106a-5p	✓	USP2-AS1:10	hsa-miR-302b-3p	✓
6	lnc-ALYREF-1:1	hsa-miR-372-3p	✓	lnc-PDK3-1:1	hsa-miR-93-5p	✓
7	MIR17HG:2	hsa-miR-520a-3p	✓	lnc-NMRK1-1:1	hsa-miR-520d-3p	✓
8	lnc-NMRK1:1	hsa-miR-520d-3p	✓	lnc-ACER2-1:1	hsa-miR-17-5p	✓
9	lnc-RPE-1:1	hsa-miR-130a-3p	×	lnc-ACER2-1:1	hsa-miR-106a-5p	✓
10	lnc-PDK3-1:1	hsa-miR-302d-3p	✓	lnc-RPE-1:1	hsa-miR-130a-3p	×



from three categories [19] (i.e. matrix factorization, random walk, and neural network).

From the matrix factorization-based category, we adopt Laplacian Eigenmaps (LE) [28], GraRep [30] and HOPE [29]. LE computes a low-dimensional representation of the dataset, optimally preserving local neighborhood information by using the Laplacian of the graph [28]. GraRep integrates global structural information of the graph into the learning process and learns high-order proximity [30]. HOPE can preserve high-order proximities of large scale graphs and is capable of capturing the asymmetric transitivity [29].

From the random walk-based category, We select DeepWalk [31]. DeepWalk uses local information obtained from truncated random walks to learn latent representations by treating walks as the equivalent of sentences [31].

We consider Graph Auto Encoder (GAE) [32] as a representative of the neural network-based methods.

GAE obtains low-dimensional node representations by reconstructing the heterogeneous network with consideration of the first-order and second-order of proximities.

By employing the aforementioned graph embedding methods, the topological and inherent properties of the heterogeneous network are acquired, then the learned distinctive representations will be further used in the downstream task, as shown in Fig. 5 (a).

Graph embedding ensemble learning based on prediction integration

In this section, we introduce a graph embedding ensemble learning method based on prediction integration (GEEL-PI). We build base predictors based on individual graph embedding methods, and further combine their predictions with ensemble strategy to infer LMIs.

To build a base predictor, firstly, we acquire the low-dimensional representations of miRNAs and

lncRNAs using the corresponding graph embedding method. Then we denote lncRNA-miRNA pairs as the concatenation of two kinds of embeddings and further build a Random Forest predictor based on pairs. The reason why we adopt Random Forest lies in its high-efficiency.

Following the steps outlined above, we can construct five base predictors based on corresponding graph embedding methods. The five graph embedding methods are heterogeneous, which captures inherent structure properties from different aspects, thus they may demonstrate different generalization abilities on datasets. Therefore, it is natural to integrate several predictors by using ensemble strategies. Theoretically, ensemble learning is to build a model $\phi: (f_1(x), f_2(x), \dots, f_n(x)) \rightarrow \{0, 1\}$, which maps the outcome of n base predictors to a label. Specifically, we consider logistic regression as the mapping function ϕ , which is simple but can model the non-linear relationship between base predictors and labels. In this way, we construct GEEL-PI for LMI prediction as described in Fig. 5 (B).

Graph embedding ensemble learning based on feature integration

In this section, we introduce a graph embedding ensemble learning method based on feature integration (GEEL-FI). We construct a deep attention neural network to learn lncRNA-miRNA pair representations, and further develop a classifier for LMI prediction.

The deep attention neural network contains attention layer and deep fully-connected neural layers, as given in Fig. 5(c). First, we consider attention mechanism to integrate different embedded representations. Because heterogeneous lncRNA and miRNA features could be correlated and have redundant information, if directly merge them, it may affect the performances of conventional classifiers negatively. Attention mechanism can be used to assign importance weights to different representations which can determine the most relevant aspects, disregarding noise and redundancies in the input [44]. Motivated by its successful applications in many fields [45–51], we adopt an attention mechanism to integrate heterogeneous genomic representations. Then we consider the deep neural network (DNN) for feature refinement. DNN allows computational models with multiple processing layers to learn representations of lncRNAs and miRNAs with multiple levels of abstraction. Moreover, deep learning discovers intricate structure in large data sets by using the backpropagation algorithm to indicate how a machine should change its internal parameters that are used to compute the representation in each layer from the representation in the previous layer [52]. Therefore, we construct a DANN to adaptively capture the importance of each embedding feature and

learn distinctive high-level representations for LMI prediction.

Specifically, given i th lncRNA and j th miRNA, by using five embedding methods, we obtain five lncRNA representations and five miRNA representations, let l_i^k and m_j^k ($k = 1, 2, 3, 4, 5$) denote embeddings from *LE*, *GraRep*, *HOPE*, *DeepWalk* and *GAE*, $i = 1, 2, \dots, r$ and $j = 1, 2, \dots, t$. Then these representations are fed into attention networks. Let L_i denotes the integrated feature for i th lncRNA, and M_j denotes the integrated feature for j th miRNA. The merged representation of lncRNA and miRNA are defined as:

$$L_i = \sum_k a_k^l l_i^k \quad (2)$$

$$M_j = \sum_k a_k^m m_j^k \quad (3)$$

where a_k^l denotes an attention weight measuring the importance of embedded representation k with respect to i th lncRNA, and a_k^m is an attention weight measuring the importance of embedded representation k with respect to j th miRNA.

Next, we concatenate i th lncRNA representation L_i and j th miRNA representation M_j to obtain lncRNA-miRNA pair feature F_{ij} , which indicates the interaction between i th lncRNA and j th miRNA:

$$F_{ij} = [L_i; M_j] \quad (4)$$

where $[L_i; M_j]$ is the concatenation of the two vectors.

To learn preferable representations of lncRNA-miRNA interactions, we consider the interacting lncRNA-miRNA pairs as positive instances and non-interacting lncRNA-miRNA pairs as negative instances to build a deep neural network. For i th lncRNA and j th miRNA, the lncRNA-miRNA pair feature F_{ij} is fed into deep fully connected layers as following:

$$Z_L = \text{ReLU}(\mathbf{W}_L(\text{ReLU}(\mathbf{W}_{L-1} \cdots \text{ReLU}(\mathbf{W}_1 F_{ij} + \mathbf{b}_1)) + \mathbf{b}_{L-1}) + \mathbf{b}_L) \quad (5)$$

where L denotes the number of hidden layers; *ReLU* is an activation function [53], and \mathbf{W}_l and \mathbf{b}_l are the weight matrix and bias vector for the l th layer, respectively.

And the prediction score between i th lncRNA and j th miRNA \hat{p}_{ij} is computed as:

$$\hat{p}_{ij} = \text{Sigmoid}(\mathbf{W}Z_L + \mathbf{b}) \quad (6)$$

where *Sigmoid* is an activation function; \mathbf{W} and \mathbf{b} are the weight matrix and bias vector, respectively.

And we adopt the following binary cross entropy as the loss function:

$$\mathcal{L} = -\frac{1}{r \cdot t} \sum_{i=1}^r \sum_{j=1}^t [p_{ij} \log \hat{p}_{ij} + (1 - p_{ij}) \log (1 - \hat{p}_{ij})] \quad (7)$$

where \mathcal{L} denotes loss function; r and t are total numbers of lncRNAs and miRNAs respectively. p_{ij} is a label, $p_{ij} = 1$ if i th lncRNA and j th miRNA are interacting, otherwise $p_{ij} = 0$;

Therefore, the attention weights a_k^l and a_k^m can be updated through the backpropagation algorithm [54] and gradient descent algorithm according to the above loss function \mathcal{L} . The update procedure can be described as:

$$a_k^l = a_k^l - \alpha \frac{\partial \mathcal{L}}{\partial a_k^l} \quad (8)$$

$$a_k^m = a_k^m - \alpha \frac{\partial \mathcal{L}}{\partial a_k^m} \quad (9)$$

where α is the learning rate of the neural network.

Here, to improve performances of LMI prediction, we build a Random Forest classifier based on pair features.

Therefore, we utilize DANN to integrate multiple features obtained by graph embedding methods to learn better representations of lncRNA-miRNA pairs, and construct GEEL-FI.

Abbreviations

GEEL-PI: Graph embedding ensemble learning based on prediction integration; GEEL-FI: Graph embedding ensemble learning based on feature integration; LMIs: lncRNA-miRNA interactions; nt: Nucleotides; DANN: Deep attention neural network; DNN: Deep neural network; 5-CV: 5-Fold cross validation; AUC: Area under ROC curve; AUPR: Area under precision-recall curve; SEN: Sensitivity; SPEC: Specificity; PREC: Precision; ACC: Accuracy; F: F-measure

Acknowledgments

Not applicable.

About this supplement

This article has been published as part of *BMC Genomics Volume 21 Supplement 13, 2020: Selected articles from the 2019 IEEE International Conference on Bioinformatics and Biomedicine (IEEE BIBM 2019): genomics (part 1)*. The full contents of the supplement are available online at <https://bmcbgenomics.biomedcentral.com/articles/supplements/volume-21-supplement-13>.

Authors' contributions

WZ designed the study. CSZ and SZ prepared data, implemented the algorithm. CSZ, YQ, SZ, SCL, WZ, and YQN drafted the manuscript. All of the authors reviewed and approved the manuscript.

Funding

Publication costs are funded by the National Natural Science Foundation of China (61772381, 62072206, 61572368), National Key Research and Development Program (2018YFC0407904), Huazhong Agricultural University Scientific & Technological Self-innovation Foundation. The funding bodies are not involved in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Availability of data and materials

The dataset lncRNASNP used in this study are freely available at <http://bioinfo.life.hust.edu.cn/lncRNASNP>. We download lncRNA sequence from

NONCODE is now available at <http://www.noncode.org>. And the main datasets of miRNA sequences are available at <http://www.mirbase.org>.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹College of Informatics, Huazhong Agricultural University, Wuhan 430070, China. ²School of Computer Science, Wuhan University, Wuhan 430072, China. ³School of Mathematics and Statistics, South-Central University for Nationalities, Wuhan 430074, China.

Published: 17 December 2020

References

- Turner M, Galloway A, Vigorito E. Noncoding RNA and its associated proteins as regulatory elements of the immune system. *Nat Immunol*. 2014; 15(6):484–91.
- Fatica A, Bozzoni I. Long non-coding RNAs: new players in cell differentiation and development. *Nat Rev Genet*. 2014;15(1):7–21.
- Miska EA. How microRNAs control cell division, differentiation and death. *Curr Opin Genet Dev*. 2005;15(5):563–8.
- Xu P, Guo M, Hay BA. MicroRNAs and the regulation of cell death. *Trends Genet*. 2004;20(12):617–24.
- Lu M, Zhang Q, Deng M, Miao J, Guo Y, Gao W, Cui Q. An analysis of human MicroRNA and disease associations. *PLoS One*. 2008;3(10):e3420.
- Yoon J-H, Abdelmohsen K, Gorospe M. Functional interactions among microRNAs and long noncoding RNAs. In: *Seminars in cell & developmental biology*. 2014. Amsterdam: Elsevier; 2014. p. 9–14.
- Chakravarty D, Sboner A, Nair SS, Giannopoulou E, Rubin MA. The oestrogen receptor alpha-regulated lncRNA NEAT1 is a critical modulator of prostate cancer. *Nat Commun*. 2014;5:5383.
- Latronico MVG, Catalucci D, Condorelli G. Emerging role of MicroRNAs in cardiovascular biology. *Circ Res*. 2007;101(12):1225–36.
- Qian L, Jianguo H, Nanjiang Z, Ziqiang Z, Ali Z, Zhaohui L, Fangting W, Yin-Yuan M. lncRNA loc285194 is a p53-regulated tumor suppressor. *Nucleic Acids Res*. 2013;41(9):4976–87.
- Xu MD, Wang Y, Weng W, Wei P, Qi P, Zhang Q, Tan C, Ni SJ, Dong L, Yang Y. A positive feedback loop of lncRNA-PVT1 and FOXM1 facilitates gastric Cancer growth and invasion. *Clin Cancer Res*. 2016;23(8):2071.
- Berghoff EG, Clark MF, Sean C, Ivelisse C, Leib DE, Kohtz JD. Evf2 (Dlx6as) lncRNA regulates ultraconserved enhancer methylation and the differential transcriptional control of adjacent genes. *Development*. 2013;140(21):4407–16.
- Gong J, Liu W, Zhang J, Miao X, Guo A-Y. lncRNASNP: a database of SNPs in lncRNAs and their potential functions in human and mouse. *Nucleic Acids Res*. 2015;43(Database issue):D181.
- Huang Y-A, Chan KCC, You Z-H. Constructing prediction models from expression profiles for large scale lncRNA-miRNA interaction profiling. *Bioinformatics*. 2018;34(5):812–9.
- Huang Z-A, Huang Y-A, You Z-H, Zhu Z, Sun Y. Novel link prediction for large-scale miRNA-lncRNA interaction network in a bipartite graph. *BMC Med Genomics*. 2018;11(6):17–27.
- Hu P, Huang YA, Chan KCC, You ZH: Discovering an Integrated Network in Heterogeneous Data for Predicting lncRNA-miRNA Interactions; 2018.
- Zhang W, Tang G, Zhou S, Niu Y. lncRNA-miRNA interaction prediction through sequence-derived linear neighborhood propagation method with information combination. *BMC Genomics*. 2019;20(Suppl 11):946.
- Wang YB, You ZH, Li X, Jiang TH, Chen X, Zhou X, Wang L. Predicting protein-protein interactions from protein sequences by a stacked sparse autoencoder deep neural network. *Mol Biosyst*. 2017;13(7):1336–44.
- Zitnik M, Agrawal M, Leskovec J. Modeling polypharmacy side effects with graph convolutional networks. *Bioinformatics*. 2018;34(13):i457–66.

19. Yue X, Wang Z, Huang J, Parthasarathy S, Moosavinasab S, Huang Y, Lin SM, Zhang W, Zhang P, Sun H. Graph embedding on biomedical networks: methods, applications and evaluations. *Bioinformatics*. 2020;36(4):1241–51.
20. Yang P, Hwa Yang Y, Zhou BB, Zomaya AY. A review of ensemble methods in bioinformatics. *Curr Bioinformatics*. 2010;5(4):296–308.
21. Polikar R. Ensemble based systems in decision making. *IEEE Circuits and systems magazine*. 2006;6(3):21–45.
22. Chen X, Yan CC, Zhang X, Zhang X, Dai F, Yin J, Zhang Y. Drug-target interaction prediction: databases, web servers and computational models. *Brief Bioinform*. 2016;17(4):696–712.
23. Zhang W, Niu Y, Xiong Y, Zhao M, Yu R, JJPo L. Computational prediction of conformational B-cell epitopes from antigen primary structures by ensemble learning. *PLoS one*. 2012;7(8):e43575.
24. Zhang W, Liu F, Luo L, JJBb Z. Predicting drug side effects by multi-label learning and ensemble learning. *BMC Bioinformatics*. 2015;16(1):365.
25. Zhang W, Yue X, Tang G, Wu W, Huang F, Zhang X. SFPEL-LPI: sequence-based feature projection ensemble learning for predicting lncRNA-protein interactions. *PLoS Comput Biol*. 2018;14(12):e1006616.
26. Gong Y, Niu Y, Zhang W, Li X. A network embedding-based multiple information integration method for the miRNA-disease association prediction. *BMC Bioinformatics*. 2019;20(1):468.
27. Zhang W, Jing K, Huang F, Chen Y, Li B, Li J, Gong J. SFLLN: a sparse feature learning ensemble method with linear neighborhood regularization for predicting drug–drug interactions. *Inf Sci*. 2019;497:189–201.
28. BELKIN M. Laplacian eigenmaps and spactral techniques for embedding and clustering. *Adv Neural Inf Proces Syst*. 2001;14(6):585–91.
29. Ou M, Peng C, Jian P, Zhang Z, Zhu W. Asymmetric transitivity preserving graph embedding, vol. 2016. New York City: Acm Sigkdd International Conference; 2016.
30. Cao S, Wei L, Xu Q. GraRep: learning graph representations with global structural information, vol. 2015. New York City: Acm International on Conference on Information & Knowledge Management; 2015.
31. Perozzi B, Al-Rfou R, Skiena S. DeepWalk: online learning of social representations, vol. 2014. New York City: Acm Sigkdd International Conference on Knowledge Discovery & Data Mining; 2014.
32. Kipf TN, Welling M. Variational graph auto-encoders. *arXiv preprint arXiv: 161107308* 2016.
33. Zhou S, Yue X, Xu X, Liu S, Zhang W, Niu Y. LncRNA-miRNA interaction prediction from the heterogeneous network through graph embedding ensemble learning. In: 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM): 2019. New York: IEEE; 2019. p. 622–7.
34. Pedregosa F, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J. Scikit-learn: machine learning in python. *J Mach Learn Res*. 2013;12(10):2825–30.
35. Zhang W, Tang G, Wang S, Chen Y, Zhou S, Li X. Sequence-derived linear neighborhood propagation method for predicting lncRNA-miRNA interactions. In: 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM): 2018; 2018.
36. Jun-Hao L, Shun L, Hui Z, Liang-Hu Q, Jian-Hua Y. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res*. 2014;42(Database issue):D92.
37. Xiao B, Guo J, Miao Y, Jiang Z, Rong H, Zhang Y, Li D, Zhong J. Detection of miR-106a in gastric carcinoma and its clinical significance. *Clin Chim Acta*. 2009;400(1):97–102.
38. Yang G, Zhang R, Chen X, Mu Y, Jing A, Chen S, Liu Y, Shi C, Sun L, Rainov NG. MiR-106a inhibits glioma cell growth by targeting E2F1 independent of p53 status. *J Mol Med-Jmm*. 2011;89(10):1037–50.
39. Sabit H, Cevik E, Tombuloglu H, Farag K, Said O. miRNA profiling in MCF-7 breast Cancer cells: seeking a new biomarker. *J Biomed Sci*. 2019;8:3.
40. Jing G, Wei L, Jiayou Z, Xiaoping M, An-Yuan G. lncRNASNP: a database of SNPs in lncRNAs and their potential functions in human and mouse. *Nucleic Acids Res*. 2015;43(Database issue):D181.
41. Changning L, Baoyan B, Geir S, Lun C, Wei D, Yong Z, Dongbo B, Yi Z, Runsheng C. NONCODE: an integrated knowledge database of non-coding RNAs. *Nucleic Acids Res*. 2005;33(Database issue):D112–5.
42. Sam GJ, Grocock RJ, Stijn VD, Alex B, Enright AJ. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res*. 2006; 34(suppl 1):140–4.
43. Leslie C, Eskin E, Noble WS. The spectrum kernel: A string kernel for SVM protein classification. In: *Biocomputing 2002*. World Scientific; 2001. p. 564–75.
44. Chaudhari S, Polatkan G, Ramanath R, Mithal V: An attentive survey of attention models. *arXiv preprint arXiv:190402874* 2019.
45. Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:14090473* 2014.
46. Cheng Z, Ding Y, He X, Zhu L, Song X, Kankanalli MS. A³ 3NCF: an adaptive aspect attention model for rating prediction, vol. 2018. California: IJCAI; 2018. p. 3748–54.
47. Maharjan S, Montes M, González FA, Solorio T: A genre-aware attention model to improve the likability prediction of books. In: *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*: 2018; 2018: 3381–3391.
48. Han C, Shen F, Liu L, Yang Y, Shen HT. Visual spatial attention network for relationship detection. In: *Proceedings of the 26th ACM international conference on multimedia*. Seoul, Republic of Korea: Association for Computing Machinery; 2018. p. 510–8.
49. Hong Z, Zeng X, Wei L, Liu X. Identifying enhancer-promoter interactions with neural network based on pre-trained DNA vectors and attention mechanism. *Bioinformatics*. 2020;36(4):1037–43.
50. Shen T, Zhou T, Long G, Jiang J, Pan S, Zhang C. DiSAN: Directional Self-Attention Network for RNN/CNN-Free Language Understanding. In: *AAAI*. 2018;2018.
51. Ying H, Zhuang F, Zhang F, Liu Y, Xu G, Xie X, Xiong H, Wu J: Sequential recommender system based on hierarchical attention network. *IJCAI International Joint Conference on Artificial Intelligence 2018*, 2018-July: 3926–3932.
52. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436–44.
53. Maas AL, Hannun AY, Ng AY: Rectifier nonlinearities improve neural network acoustic models. In: *Proc icml*: 2013; 2013: 3.
54. Hecht-Nielsen R. Theory of the backpropagation neural network. In: *Neural networks for perception*: Elsevier; 1992. p. 65–93.

Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

