

## Research Article

# DEMLP: DeepWalk Embedding in MLP for miRNA-Disease Association Prediction

Xun Wang <sup>1</sup>, Fuyu Wang,<sup>1</sup> Xinzeng Wang <sup>2</sup>, Sibao Qiao,<sup>1</sup> and Yu Zhuang<sup>1</sup>

<sup>1</sup>College of Computer Science and Technology, China University of Petroleum, Qingdao 266500, China

<sup>2</sup>College of Mathematics and Systems Science, Shandong University of Science and Technology, Qingdao 266500, China

Correspondence should be addressed to Xinzeng Wang; wanglxz@126.com

Received 7 May 2021; Revised 10 August 2021; Accepted 7 September 2021; Published 16 October 2021

Academic Editor: Haibin Lv

Copyright © 2021 Xun Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

miRNAs significantly affect multifarious biological processes involving human disease. Biological experiments always need enormous financial support and time cost. Taking expense and difficulty into consideration, to predict the potential miRNA-disease associations, a lot of high-efficiency computational methods by computer have been developed, based on a network generated by miRNA-disease association dataset. However, there exist many challenges. Firstly, the association between miRNAs and diseases is intricate. These methods should consider the influence of the neighborhoods of each node from the network. Secondly, how to measure whether there is an association between two nodes of the network is also an important problem. In our study, we innovatively integrate graph node embedding with a multilayer perceptron and propose a method DEMLP. To begin with, we construct a miRNA-disease network by miRNA-disease adjacency matrix (MDA). Then, low-dimensional embedding representation vectors of nodes are learned from the miRNA-disease network by DeepWalk. Finally, we use these low-dimensional embedding representation vectors as input to train the multilayer perceptron. Experiments show that our proposed method that only utilized the miRNA-disease association information can effectively predict miRNA-disease associations. To evaluate the effectiveness of DEMLP in a miRNA-disease network from HMDD v3.2, we apply fivefold crossvalidation in our study. The ROC-AUC computed result value of DEMLP is 0.943, and the PR-AUC value of DEMLP is 0.937. Compared with other state-of-the-art methods, our method shows good performance using only the miRNA-disease interaction network.

## 1. Introduction

More and more evidence shows that miRNA impacts biological processes in a very significant way, like cell development and cell proliferation [1–4]. Therefore, studying the nosogenesis, medical diagnosis, and manual interventions of human diseases has a significant influence on predicting miRNA-disease associations [5–22]. Biological experiments work well but cost a lot of time and money. However, to predict the potential miRNA-disease association by cheap and efficient ways, many kinds of methods have been generated [23].

Over the past years, to predict the potential association of miRNA disease, a lot of models based on the known miRNA-disease association network have been developed. These methods mainly consisted of two categories: the score function-based algorithms and machine learning.

Some methods based on matrix completion also have good results in the prediction of miRNA-disease association. Jiang et al. (2010) generated a model by giving priority to the entire human microRNA for diseases of interest [24]. Differing from common local network similarity measures, a method named RWRMDA is generated by Chen et al. (2012), which employ overall measurement of network similarity and adopt node embedding presentation method random walk to infer the potential association of miRNA disease [25]. Another method that extends the process of the Random Walk algorithm is generated by Xuan et al. (2015) [26]. Ji et al. (2015) propose a method called SVAEMDA, which considers the miRNA-disease association prediction as a semisupervised learning problem, then trains a variational autoencoder based predictor to solve the problem [27]. By updating the association adjacency

matrix of miRNA disease through the matrix completion algorithm, a method named MCMCDA to predict the associations of miRNA disease is generated by Li et al. (2017) [28]. By integrating known human miRNA-disease associations, miRNA functional similarity, disease semantic similarity, and Gaussian interaction profile kernel similarity for miRNAs and diseases, a method named PBMDA is proposed by You et al. (2017). This method constructs a heterogeneous graph consisting of three interlinked subgraphs and further adopts a depth first search algorithm to infer potential miRNA-disease associations [29]. According to Matrix Decomposition and Heterogeneous Graph Inference, a method named MDHGI is generated by Chen et al. (2018). This method predicts associations of miRNA disease by predicted association probability [30]. Chen et al. (2018) propose a model of Inductive Matrix Completion for MiRNA-disease Association prediction (IMCMDA), which completes the missing miRNA-disease association based on the known associations and the integrated miRNA similarity and disease similarity [31]. Combining neighborhood constraint with matrix completion, a model named NCMCMA is proposed by Chen et al. (2021) [32].

To predict the potential association of miRNA disease, a lot of machine learning methods have been proposed. To predict miRNA-disease associations, a novel model framework is adopted by Shi et al. (2013). Furthermore, they construct a bipartite network, which is used to analyze the peculiarity of miRNA regulating disease genes. This work provides an original perspective for the discovery of genetic disease associations and may contribute to future research on miRNA involvement in disease pathogenesis [33]. According to the miRNA functional similarity, which is defined by measuring the similarity between genetically related diseases, a model named DHMP is generated by Xuan et al. (2013). The similarity of miRNA is effectively evaluated by measuring the semantic similarity of their associated diseases [34]. With the continuous development of support vector machine and  $k$ -nearest-neighbor technology, a novel method named RKNNMDA is generated by Chen et al. (2017), which combines SVM and KNN and achieves good performance in the prediction task [35]. To predict the potential miRNA-disease association, a model named ABMDA is proposed by Zhao et al. (2019) based on adaptive boosting [36]. By adopting a graph embedding representation learning algorithm and neural network method, an original method named CNNMDA is generated by Xuan et al. (2019) to predict the associations between miRNA and disease [37]. To predict the potential associations of miRNA disease, Peng et al. (2019) propose a novel learning-based framework, MDA-CNN, which constructs a three-layer network and uses an autoencoder and convolutional neural network to catch the essential feature and predict the final label, respectively [38]. Based on inductive matrix completion and graph convolutional network, a model named NIMCGCN is generated by Li et al. (2020). This method generates node embedding feature representations from a network, and then they put the learned features into a matrix completion model to predict miRNA-disease associations [39]. Based on integrating the matrix factoriza-

tion and multilayer perceptron, Liu et al. (2018) propose a method NCFM [40]. Based on node embedding, pair embedding, and multilayer perceptron, Liu et al. (2021) propose a method named CEMDA to predict miRNA-disease associations [41].

It is of great importance to discover the potential miRNA-disease association. So far, most methods use known miRNA-disease and similarity information of disease and miRNA to predict potential associations. However, with the continuous enrichment of human databases and the continuous improvement of high-throughput technology [42], a large number of databases about miRNA-disease association are available for free. More efficient prediction methods only utilized the miRNA-disease association information urgently needed [40]. Many new related methods in the field of drug target prediction have also attracted our attention [43–49]. In this paper, we propose an effective method named DEMLP to predict the miRNA-disease association that only uses the known miRNA-disease association network. In our method, we use DeepWalk to generate the node embedding using the information of their neighborhood and the network structure. After that, we use MLP to calculate the association score of the miRNA disease. In the main experiment, the receiver operating characteristics (ROC) and precision-recall (PR) area under the curve (AUC) of our proposed method are 0.943 and 0.937, respectively. Then, we compare our model with other four known state-of-the-art models, and the receiver operating characteristics (ROC) area under the curve (AUC) of our proposed method is 0.923, which is far more than any other model using the same data HMDD v2.0.

## 2. Materials and Methods

**2.1. Dataset for miRNA-Disease Association Prediction.** Data of miRNA-disease associations that we use in our experiment are obtained from HMDD [50–52]. To construct the interactive network of miRNA disease, we download the whole dataset from the HMDD database of the online website. Moreover, in the case study, we use the data of dbDEMC v3.0 [53] to verify the effect of our model.

**2.2. Data Preparation.** The online website of HMDD v3.2 shows that there are 35547 miRNA-disease association entries which include 1206 miRNA genes and 893 diseases from 19280 papers. We need to clean the original dataset for constructing the miRNA-disease interaction network. Then, we need to cleanse this data set by looking for typography errors, letter capitalization errors (e.g., hsa-mir-200C should be has-mir-200c and Has-mir-93 should be has-mir-93), removing duplicates, and so on. After cleaning the original dataset, we find 1206 miRNAs, 893 diseases, and 18732 miRNA-disease associations. We construct the miRNA-disease network by the miRNA-disease association adjacency matrix ( $MDA$ ) that is generated from the data obtained by cleaning. The element  $MDA_{ij} = 1$ , if miRNA  $v_{m_i}$  is associated with disease  $v_{d_j}$ , on the contrary, if  $MDA_{ij} = 0$  means there is no association between miRNA  $v_{m_i}$  and disease  $v_{d_j}$ ,  $v$  is a vertex of the network. The data cleaning

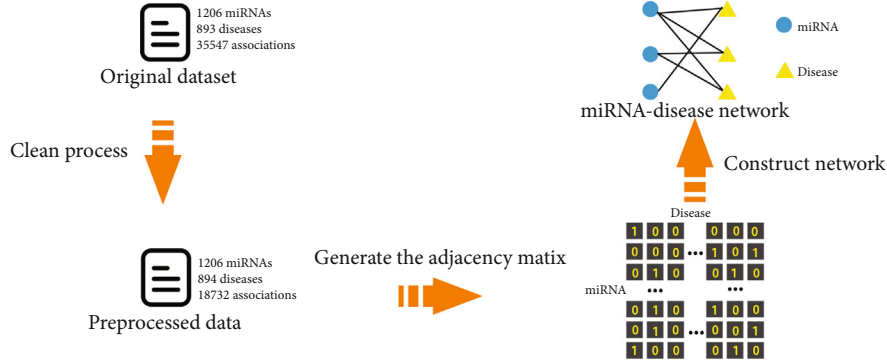


FIGURE 1: Data cleaning and constructing miRNA-disease network process.

and constructing miRNA-disease network process are shown in Figure 1.

**2.3. Prediction of the miRNA-Disease Association.** The goal of DEMLP is to predict potential associations. First, we generate the node embedding by DeepWalk [54], and then we concatenate each miRNA embedding and disease embedding as a new dataset for the MLP (multilayer perceptron) model. To evaluate the effectiveness of DEMLP in the miRNA-disease network from HMDD v3.2, we apply five-fold crossvalidation in our research.

**2.4. DeepWalk on miRNA-Disease Network.** DeepWalk is a method for generating potential representations of nodes in a network. These potential nodes representations consist in a continuous vector space, which could be effectively used in some machine learning methods. DeepWalk considers a set of short truncated random walks on our miRNA-disease network as its corpus and the network vertices as its vocabulary. The random walk generator samples uniformly a random vertex  $vi$  as the root of the random walk  $W_{v_i} \in R^{r \times d}$  ( $\gamma$  is the walks per vertex). A walk sample coequally from the neighbors of the last vertex visited until the hyperparameter length  $t$  that we set before is reached. In our method, we set that the window size  $w$  is 5, walk length  $t$  is 10, walks pervertex  $\gamma$  are 80, and the embedding size  $d$  is 128. We iterate all over the nodes in our miRNA-disease network. For each node vertex  $vi$  of the miRNA-disease network, we generate a random walk  $|W_{v_i}| = t$ , then use it as our corpus to update our node vertex representations. We use the Skip-Gram algorithm to adjust these node vertex representations based on our objective function in Eq. (1).

$$\text{Min} - \log \Pr \left( \frac{(v_i - w, \dots, v_i + w)}{v_i} \mid \Phi(v_i) \right), \quad (1)$$

where  $\Phi v \in V \rightarrow R^{|V| \times d}$  is a mapping function, and this mapping function  $\Phi$  represents the potential node embedding representation of each vertex  $v$  in the network.

**2.5. Multilayer Perceptron.** The multilayer perceptron is the classical machine learning model. We splice the representa-

tions of each miRNA  $\Phi(v_{m_i})$  and disease  $\Phi(v_{d_j})$  as the dataset of our MLP model.

$$X_{ij} = \text{splicing}(\Phi(v_{m_i}), \Phi(v_{d_i})), \quad (2)$$

where  $v_{m_i}$  is a vertex index of miRNA  $i$ ,  $v_{d_i}$  is a vertex index of disease  $j$ , and  $X_{ij} \in R^{1 \times 2d}$  is the concatenating representation of miRNA  $i$  and disease  $j$ .  $|X| = |v_m| \times |v_d|$  is the order of matrix  $X$ .

We use a part of the dataset  $X_{\text{train}}$  as the training data of the MLP model. The dimension of the MLP input layer is 128, and the dimension of the MLP hidden layer is 4. We use tanh as the activation function to compute the hidden layer values [55]. The calculation process is shown in Eq. (3).

$$f_{(1)} = \tanh(X_{\text{train}} W_{d \times 4}^1 + b_{d \times 1}), \quad (3)$$

where tanh is the activation function of the hidden layer, and  $W_{d \times 4}$  and  $b$  are learnable parameters of our MLP model.

There is only one value for the output layer of our MLP model. We choose the *sigmoid* function as the output layer activation function.

$$\text{score}_{ij} = \text{out} = \text{sigmoid}(f_{(1)} W_{4 \times 1}^4 + b_{d \times 1}), \quad (4)$$

where  $\text{score}_{ij}$  represents the correlation score between miRNA vertex  $v_{m_i}$  and disease vertex  $v_{d_j}$ .

We use binary cross entropy between our target and output score. Then, we use stochastic gradient descent to optimize our parameters.

$$\text{Loss} = \sum \text{BCELoss}(\text{score}_{ij}, \text{MDA}_{ij}). \quad (5)$$

The framework of our method is in Figure 2.

**2.6. DEMLP-PLUS.** To further verify the validity of the model in the task of association prediction of miRNA diseases, we add the similarity information of miRNAs and diseases separately into the model to observe whether the prediction results would be improved. The new framework is named DEMLP-PLUS. We refer to the IMCMDA [31] model for the integration network construction process after

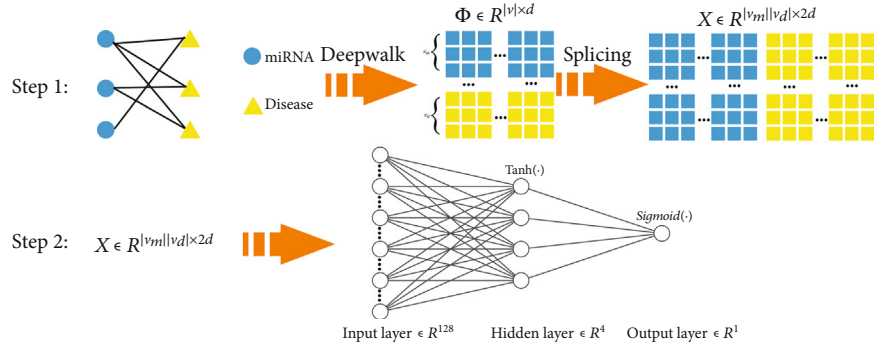


FIGURE 2: DEMLP framework.

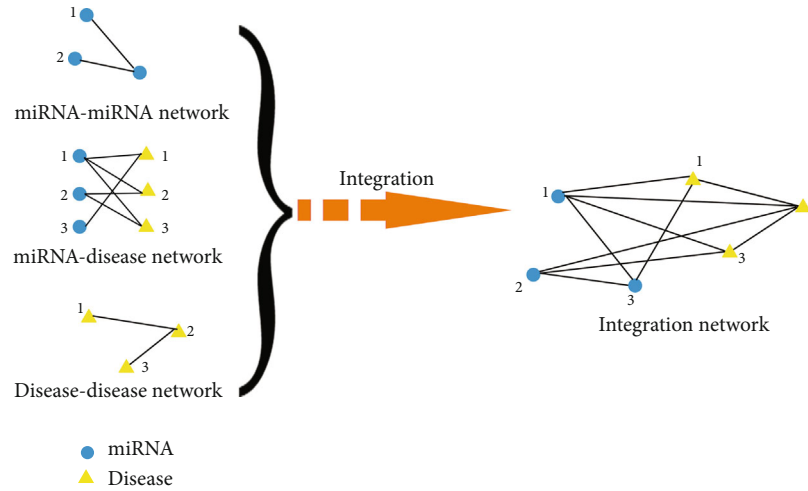


FIGURE 3: Adding more information into a network.

adding the similarity of genes and diseases, respectively, and the integration network is a fusion of genetic associations, disease associations, and gene-disease associations, which construction process is shown in Figure 3.

### 3. Experiments and Result

There are a large number of negative samples in the whole data, and we use the undersampling method to make the positive sample and negative sample reach 1:1 equilibrium. Firstly, to verify the validity of the model in the task of association prediction of gene-disease, we apply fivefold crossvalidation to evaluate the availability of our method in the miRNA-disease network from HMDD v3.2. Then, we compare DEMLP with the other three baseline models: LineMLP, node2vecMLP, and SDNEMLP. In the training process, we applied fivefold crossvalidation on each model and performed 100 iterations to find the optimal parameter with the smallest error. Secondly, we add the similarity information of miRNAs and diseases separately into the model to observe whether the prediction results would be improved. In this section, we use HMDD v2.0 as our dataset. Based on the HMDD v2.0 dataset, we compared the performance of the model in a fusion network with gene-disease similarity information and a normal network. At the same

TABLE 1: PR-AUC and ROC-AUC of each model.

Method	PR-AUC	ROC-AUC
LineMLP	0.910	0.916
node2vecMLP	0.910	0.919
SDNEMLP	0.840	0.861
DEMLP	0.937	0.943

time, we compare our model with other known state-of-the-art models based on the network from HMDD v2.0's associated data set of miRNA disease. Thirdly, we did a case study of lung tumors and breast tumors, and we examine the miRNAs in HMDD v2.0 for these diseases and used dbDEMC v3.0 to verify the top 20 rankings association we predicted.

*3.1. Performance Evaluation.* We use effective indicators ROC-AUC and PR-AUC to estimate the association prediction effect. A test example is labeled as a positive example while the prediction score of miRNA-disease association is more than  $\theta$  ( $\theta$  is a threshold). If not, it is considered as a negative example [56]. We use TN and TP to represent the numbers of correctly identified negative and positive examples, respectively. FN and FP, respectively, represent the

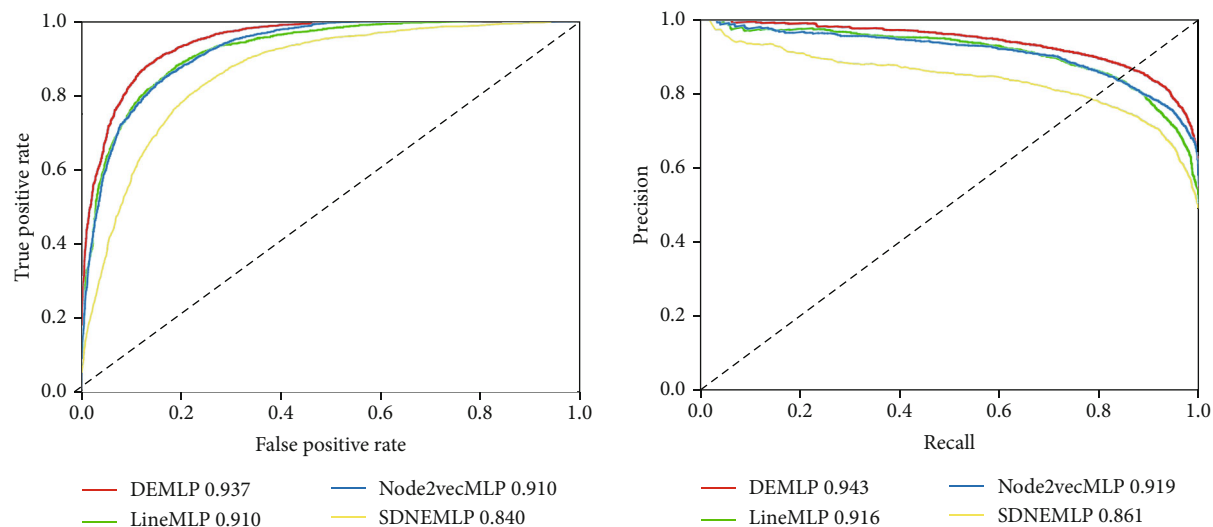


FIGURE 4: Receiver operating characteristic (ROC) and precision-recall (PR) curves of all method.

TABLE 2: PR-AUC and ROC-AUC of DEMLP and DEMLP-PLUS.

Method	PR-AUC	ROC-AUC
DEMLP	0.908	0.923
DEMLP-PLUS	0.921	0.924

TABLE 3: Comparison with other SOTA models.

Method	ROC-AUC
DEMLP	0.923
NCFM	0.912
CGMDA	0.891
Metapath	0.880
BNPMDA	0.889

numbers of positive and negative examples. The formulas for these indicators, TPR (true positive rate), FPR (false positive rate), precision, and recall (recall rate), are defined, respectively, as follows:

$$\begin{aligned} \text{TPR} &= \frac{\text{TP}}{\text{TP} + \text{FN}}, \text{FPR} = \frac{\text{FP}}{\text{TN} + \text{FP}}, \\ \text{Precision} &= \frac{\text{TP}}{\text{TP} + \text{FP}}, \text{Recall} = \frac{\text{TP}}{\text{TN} + \text{FP}}. \end{aligned} \quad (6)$$

Table 1 shows the average ROC-AUC and PR-AUC of the 5-fold crossvalidation of our experiments.

Then, we plot the ROC and PR curve of each method in Figure 3 and the ROC-AUC and PR-AUC bar graph in Figure 4.

**3.2. DEMLP-PLUS Experiments.** To further verify the validity of the model in the task of association prediction of miRNA diseases, we add the similarity information of miRNAs and diseases separately into the model to observe whether the prediction results would be improved. In the

TABLE 4: Prediction results of the top 20 predicted breast neoplasm-related miRNAs based on HMDD v2.0.

TOP 20 miRNAs	Evidence
hsa-mir-1236	dbDEMCM
hsa-mir-592	dbDEMCM
hsa-mir-487a	dbDEMCM
hsa-mir-548c	dbDEMCM
hsa-mir-431	dbDEMCM
hsa-mir-1909	dbDEMCM
hsa-mir-376b	dbDEMCM
hsa-mir-384	dbDEMCM
hsa-mir-1183	dbDEMCM
hsa-mir-596	dbDEMCM
hsa-mir-2861	dbDEMCM
hsa-mir-3196	dbDEMCM
hsa-mir-188	dbDEMCM
hsa-mir-671	dbDEMCM
hsa-mir-519b	dbDEMCM
hsa-mir-572	dbDEMCM
hsa-mir-371a	dbDEMCM
hsa-mir-3148	dbDEMCM
hsa-mir-569	dbDEMCM
hsa-mir-720	dbDEMCM

first step, we construct the integration network using HMDD v2.0; then, we apply fivefold crossvalidation on the model DEMLP and DEMLP-PLUS.

Table 2 indicates that there is only a small improvement in the model after adding the similarity information of miRNAs and diseases separately into the model DEMLP. In the face of higher time and space complexity, our model has better generalization and application advantages when using only the association network of genes and diseases.



TABLE 5: Prediction results of the top 20 predicted lung neoplasm-related miRNAs based on HMDD v2.0.

TOP 20 miRNAs	Evidence
hsa-mir-1236	dbDEMC
hsa-mir-592	dbDEMC
hsa-mir-487a	dbDEMC
hsa-mir-548c	dbDEMC
hsa-mir-431	dbDEMC
hsa-mir-1909	dbDEMC
hsa-mir-376b	dbDEMC
hsa-mir-384	dbDEMC
hsa-mir-1183	dbDEMC
hsa-mir-596	dbDEMC
hsa-mir-2861	dbDEMC
hsa-mir-3196	dbDEMC
hsa-mir-188	dbDEMC
hsa-mir-671	dbDEMC
hsa-mir-519b	dbDEMC
hsa-mir-572	dbDEMC
hsa-mir-371a	dbDEMC
hsa-mir-3148	dbDEMC
hsa-mir-569	dbDEMC
hsa-mir-720	dbDEMC

To verify the validity of the model in the task of association prediction of miRNA diseases, we compare our model with other four known state-of-the-art models (NCFM [40], CGMDA [57], metapath [58], and BNPMMA [59]) based on the network from HMDD v2.0's associated dataset of miRNA disease by the 5-fold crossvalidation.

Table 3 indicates that our model has a good performance in predicting genetic disease association.

**3.3. Case Study.** Studies have shown that microRNA (miRNA) alterations are involved in the initiation and progression of human cancer [6, 8, 60–68]. We conduct a case study of lung tumors [69, 70] and breast tumors [71, 72], where we examine the miRNAs in HMDD v2.0 for these diseases and use dbDEMC v3.0 to verify the top 20 rankings association we predict. The prediction results of lung tumors and breast tumors computed by DEMLP are shown in Tables 4 and 5, respectively.

The prediction results show that our model has a good effect on predicting the unknown association between genes and disease.

## 4. Conclusions

Taking expense and difficulty into consideration, to predict the potential miRNA-disease associations, a lot of high-efficiency computational methods by computer have been developed, based on a network generated by miRNA-disease association dataset. More efficient prediction methods which only utilize miRNA-disease association information are urgently needed. To predict the potential miRNA-disease

associations in this research, we innovatively integrate graph node embedding with multilayer perceptron and propose a method DEMLP. DEMLP can predict the miRNA-disease association effectively utilizing only the miRNA-disease association information. Through the combination of the random walk and the multilayer perceptron, DEMLP can learn the node embedding representation which is rich in network structure information and heighten nonlinear fitting ability. Compared with other models, DEMLP achieves the best result in the task of miRNA-disease association prediction. Moreover, in the future, we will consider referring to models such as EGES [73] to solve the problem of cold start in the prediction of genetic disease association.

## Data Availability

The data used to support the findings of the study are available in the public database <http://www.cuilab.cn/hmdd>.

## Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

## Acknowledgments

This research was supported by the National Science Foundation of China [No. 61873281 and No. 61972416] and the Project of the China University of Petroleum (East China) [2018000003].

## References

- [1] D. P. Bartel, "MicroRNAs: target recognition and regulatory functions," *Cell*, vol. 136, no. 2, pp. 215–233, 2009.
- [2] Y. Li, C. Liang, K. C. Wong, J. Luo, and Z. Zhang, "Mirsynergy: detecting synergistic miRNA regulatory modules by overlapping neighbourhood expansion," *Bioinformatics*, vol. 30, no. 18, pp. 2627–2635, 2014.
- [3] J. Lu, G. Getz, E. A. Miska et al., "MicroRNA expression profiles classify human cancers," *Nature*, vol. 435, no. 7043, pp. 834–838, 2005.
- [4] T. Song, X. Wang, X. Li, and P. Zheng, "A programming triangular DNA origami for doxorubicin loading and delivering to target ovarian cancer cells," *Oncotarget*, vol. 5, 2014.
- [5] V. Ambros, "The functions of animal microRNAs," *Nature*, vol. 431, no. 7006, pp. 350–355, 2004.
- [6] G. A. Calin and C. M. Croce, "MicroRNA signatures in human cancers," *Nature Reviews Cancer*, vol. 6, no. 11, pp. 857–866, 2006.
- [7] A. M. Cheng, M. W. Byrom, J. Shelton, and L. P. Ford, "Anti-sense inhibition of human miRNAs and indications for an involvement of miRNA in cell growth and apoptosis," *Nucleic Acids Research*, vol. 33, no. 4, pp. 1290–1297, 2005.
- [8] A. Esquela-Kerscher and F. J. Slack, "Oncomirs – microRNAs with a role in cancer," *Nature Reviews Cancer*, vol. 6, no. 4, pp. 259–269, 2006.
- [9] X. Karp and V. Ambros, "Developmental biology: enhanced: encountering microRNAs in cell fate signaling," *Science*, vol. 310, no. 5752, pp. 1288–1289, 2005.

- [10] W. P. Kloosterman and R. H. A. Plasterk, "The diverse functions of MicroRNAs in animal development and disease," *Developmental Cell*, vol. 11, no. 4, pp. 441–450, 2006.
- [11] M. Esteller, "Non-coding RNAs in human disease," *Nature Reviews Genetics*, vol. 12, no. 12, pp. 861–874, 2011.
- [12] J. Elmén, M. Lindow, S. Schütz et al., "LNA-mediated microRNA silencing in non-human primates," *Nature*, vol. 452, no. 7189, pp. 896–899, 2008.
- [13] J. T. Mendell and E. N. Olson, "MicroRNAs in stress signaling and human disease," *Cell*, vol. 148, no. 6, pp. 1172–1187, 2012.
- [14] X. Li, X. Wang, T. Song, W. Lu, Z. Chen, and X. Shi, "A novel computational method to reduce leaky reaction in DNA strand displacement," *Journal of Analytical Methods in Chemistry*, vol. 2015, Article ID 675827, 10 pages, 2015.
- [15] X. Shi, Z. Wang, C. Deng, T. Song, L. Pan, and Z. Chen, "A novel bio-sensor based on DNA strand displacement," *Plos One*, vol. 9, no. 10, article e108856, 2014.
- [16] X. Li, T. Song, Z. Chen, X. Shi, C. Chen, and Z. Zhang, "A universal fast colorimetric method for DNA signal detection with DNA strand displacement and gold nanoparticles," *Journal of Nanomaterials*, vol. 2015, 9 pages, 2015.
- [17] X. Shi, C. Chen, X. Li et al., "Size-controllable DNA nanoribbons assembled from three types of reusable brick single-strand DNA tiles," *Soft Matter*, vol. 11, no. 43, pp. 8484–8492, 2015.
- [18] Z. Chen, T. Song, Y. Huang, and X. Shi, "Solving vertex cover problem using DNA tile assembly model," *Journal of Applied Mathematics*, vol. 2013, Article ID 407816, 7 pages, 2013.
- [19] T. Song, X. Wang, Z. Zhang, and L. Hong, "Detecting motifs in DNA sequences by branching from neighbors of qualified potential motifs," *Journal of Computational and Theoretical Nanoscience*, vol. 10, no. 9, pp. 2201–2206, 2013.
- [20] X. Li, Z. Wei, B. Wang, and T. Song, "Stable DNA sequence over close-ending and pairing sequences constraint," *Frontiers in Genetics*, vol. 12, article 644484, 2021.
- [21] X. Shi, X. Wu, T. Song, and X. Li, "Construction of DNA nanotubes with controllable diameters and patterns using hierarchical DNA sub-tiles," *Nanoscale*, vol. 8, no. 31, pp. 14785–14792, 2016.
- [22] Y. Su, S. Wang, E. Li, T. Song, H. Yu, and D. Meng, "Analysis of gene logic networks for arabidopsis," *Current Bioinformatics*, vol. 8, no. 2, pp. 244–252, 2013.
- [23] X. Chen, D. Xie, Q. Zhao, and Z. H. You, "MicroRNAs and complex diseases: from experimental results to computational models," *Briefings in Bioinformatics*, vol. 20, no. 2, pp. 515–539, 2019.
- [24] Q. Jiang, Y. Hao, G. Wang et al., "Prioritization of disease microRNAs through a human phenome-microRNAome network," *BMC Systems Biology*, vol. 4, no. S1, 2010.
- [25] X. Chen, M.-X. Liu, and G.-Y. Yan, "RWRMDA: predicting novel human microRNA-disease associations," *Molecular BioSystems*, vol. 8, no. 10, pp. 2792–2798, 2012.
- [26] P. Xuan, K. Han, Y. Guo et al., "Prediction of potential disease-associated microRNAs based on random walk," *Bioinformatics*, vol. 31, no. 11, pp. 1805–1815, 2015.
- [27] C. Ji, Y. T. Wang, Z. Gao, L. Li, J. C. Ni, and C. H. Zheng, "A semi-supervised learning method for MiRNA-disease association prediction based on variational autoencoder," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, p. 1, 2021.
- [28] J.-Q. Li, Z.-H. Rong, X. Chen, G. Y. Yan, and Z. H. You, "MCMDA: matrix completion for MiRNA-disease association prediction," *Oncotarget*, vol. 8, no. 13, pp. 21187–21199, 2017.
- [29] Z. H. You, Z. A. Huang, Z. Zhu et al., "PBMDA: a novel and effective path-based computational model for miRNA-disease association prediction," *PLoS Computational Biology*, vol. 13, no. 3, article e1005455, 2017.
- [30] X. Chen, J. Yin, J. Qu, and L. Huang, "MDHGI: matrix decomposition and heterogeneous graph inference for miRNA-disease association prediction," *Plos Computational Biology*, vol. 14, no. 8, 2018.
- [31] X. Chen, L. Wang, J. Qu, N. N. Guan, and J. Q. Li, "Predicting miRNA-disease association based on inductive matrix completion," *Bioinformatics*, vol. 34, no. 24, pp. 4256–4265, 2018.
- [32] X. Chen, L.-G. Sun, and Y. Zhao, "NCMCMMDA: miRNA-disease association prediction through neighborhood constraint matrix completion," *Briefings in Bioinformatics*, vol. 22, no. 1, pp. 485–496, 2021.
- [33] H. Shi, J. Xu, G. Zhang et al., "Walking the interactome to identify human miRNA-disease associations through the functional link between miRNA targets and disease genes," *BMC Systems Biology*, vol. 7, no. 1, 2013.
- [34] P. Xuan, K. Han, M. Guo et al., "Correction: prediction of microRNAs associated with human diseases based on weighted k most similar neighbors," *Plos One*, vol. 8, no. 9, 2013.
- [35] X. Chen, Q.-F. Wu, and G.-Y. Yan, "RKNNMDA: ranking-based KNN for MiRNA-disease association prediction," *RNA Biology*, vol. 14, no. 7, pp. 952–962, 2017.
- [36] Y. Zhao, X. Chen, and J. Yin, "Adaptive boosting-based computational model for predicting potential miRNA-disease associations," *Bioinformatics*, vol. 35, no. 22, pp. 4730–4738, 2019.
- [37] P. Xuan, H. Sun, X. Wang, Zhang, and Pan, "Inferring the disease-associated miRNAs based on network representation learning and convolutional neural networks," *International Journal of Molecular Sciences*, vol. 20, no. 15, p. 3648, 2019.
- [38] J. J. Peng, W. W. Hui, Q. Q. Li et al., "A learning-based framework for miRNA-disease association identification using neural networks," *Bioinformatics*, vol. 35, no. 21, pp. 4364–4371, 2019.
- [39] J. Li, S. Zhang, T. Liu, C. Ning, Z. Zhang, and W. Zhou, "Neural inductive matrix completion with graph convolutional networks for miRNA-disease association prediction," *Bioinformatics*, vol. 36, no. 8, pp. 2538–2546, 2020.
- [40] Y. Liu, S.-L. Wang, J.-F. Zhang, W. Zhang, and W. Li, "A neural collaborative filtering method for identifying miRNA-disease associations," *Neurocomputing*, vol. 422, pp. 176–185, 2021.
- [41] B. Liu, X. Zhu, L. Zhang, Z. Liang, and Z. Li, "Combined embedding model for MiRNA-disease association prediction," *BMC Bioinformatics*, vol. 22, no. 1, 2021.
- [42] T. Song, Q. Zou, and P. Zheng, "Guest editorial: bio-inspired computing models and algorithms," *IEEE Transactions on Nanobioscience*, vol. 19, no. 1, pp. 100–101, 2020.
- [43] X. Shi, H. Zhao, X. Li, and T. Song, "Isothermal approach to assemble spatial DNA nanotubes for drug delivery," *Oncotarget*, vol. 5, 2014.
- [44] S. Wang, D. Liu, M. Ding et al., "SE-OnionNet: a convolution neural network for protein-ligand binding affinity prediction," *Frontiers in Genetics*, vol. 11, article 607824, 2020.

- [45] Q. Tian, M. Ding, H. Yang et al., "Predicting drug-target affinity based on recurrent neural networks and graph convolutional neural networks," *Combinatorial Chemistry & High Throughput Screening*, vol. 24, 2021.
- [46] S. Wang, J. Liu, M. Ding et al., "DL-SMILES#: A Novel Encoding Scheme for Predicting Compound Protein Affinity by Deep Learning," *Combinatorial Chemistry & High Throughput Screening*, vol. 24, 2021.
- [47] S. Wang, Z. Du, M. Ding, R. Zhao, A. Rodriguez-Paton, and T. Song, "LDCNN-DTI: a novel light deep convolutional neural network for drug-target interaction predictions," in *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 1132–1136, Seoul, Korea (South), 2020.
- [48] T. Song, Y. Zhong, M. Ding et al., "Repositioning molecules of Chinese medicine to targets of SARS-Cov-2 by deep learning method," in *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 2306–2312, Seoul, Korea (South), 2020.
- [49] X. Wang, Y. Zhong, and M. Ding, "Repositioning drugs to the mitochondrial fusion protein 2 by three-tunnel deep neural network for Alzheimer's disease," *Frontiers in Genetics*, vol. 12, article 638330, 2021.
- [50] M. Lu, Q. Zhang, M. Deng et al., "An analysis of human microRNA and disease associations," *Plos One*, vol. 3, no. 10, 2008.
- [51] Z. Huang, J. Shi, Y. Gao et al., "HMDD v3.0: a database for experimentally supported human microRNA-disease associations," *Nucleic Acids Research*, vol. 47, no. D1, pp. D1013–d1017, 2019.
- [52] Y. Li, C. Qiu, J. Tu et al., "HMDD v2.0: a database for experimentally supported human microRNA and disease associations," *Nucleic Acids Research*, vol. 42, no. D1, pp. D1070–D1074, 2014.
- [53] Z. Yang, L. Wu, A. Wang et al., "DbDEMC 2.0: updated database of differentially expressed miRNAs in human cancers," *Nucleic Acids Research*, vol. 45, no. D1, pp. D812–D818, 2017.
- [54] B. Perozzi, R. Al-Rfou, and S. Skiena, "DeepWalk: online learning of social representations," in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 701–710, New York, New York, USA, 2014.
- [55] W. Malfliet, "The tanh method: a tool for solving certain classes of non-linear PDEs," *Mathematical Methods in the Applied Sciences*, vol. 28, no. 17, pp. 2031–2035, 2005.
- [56] P. Xuan, S. Pan, T. Zhang, Liu, and Sun, "Graph convolutional network and convolutional neural network based method for predicting lncRNA-disease associations," *Cells*, vol. 8, no. 9, p. 1012, 2019.
- [57] K. Zheng, L. Wang, and Z.-H. You, "CGMDA: an approach to predict and validate MicroRNA-disease associations by utilizing chaos game representation and LightGBM," *IEEE Access*, vol. 7, pp. 133314–133323, 2019.
- [58] X. Zhang, Q. Zou, A. Rodriguez-Paton, and X. Zeng, "Meta-path methods for prioritizing candidate disease miRNAs," *Ieee-Acm Transactions on Computational Biology and Bioinformatics*, vol. 16, no. 1, pp. 283–291, 2019.
- [59] X. Chen, D. Xie, L. Wang, Q. Zhao, Z. H. You, and H. Liu, "BNPMDA: bipartite network projection for MiRNA-disease association prediction," *Bioinformatics*, vol. 34, no. 18, pp. 3178–3186, 2018.
- [60] S. Lin and R. I. Gregory, "MicroRNA biogenesis pathways in cancer," *Nature Reviews Cancer*, vol. 15, no. 6, pp. 321–333, 2015.
- [61] P. S. Mitchell, R. K. Parkin, E. M. Kroh et al., "Circulating microRNAs as stable blood-based markers for cancer detection," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 30, pp. 10513–10518, 2008.
- [62] M. Korpai, E. S. Lee, G. Hu, and Y. Kang, "The miR-200 family inhibits epithelial-mesenchymal transition and cancer cell migration by direct targeting of E-cadherin transcriptional repressors ZEB1 and ZEB2," *Journal of Biological Chemistry*, vol. 283, no. 22, pp. 14910–14914, 2008.
- [63] S.-M. Park, A. B. Gaur, E. Lengyel, and M. E. Peter, "The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2," *Genes & Development*, vol. 22, no. 7, pp. 894–907, 2008.
- [64] C. M. Croce, "Causes and consequences of microRNA dysregulation in cancer," *Nature Reviews Genetics*, vol. 10, no. 10, pp. 704–714, 2009.
- [65] M. V. Iorio, R. Visone, G. di Leva et al., "MicroRNA signatures in human ovarian cancer," *Cancer Research*, vol. 67, no. 18, pp. 8699–8707, 2007.
- [66] F. Meng, R. Henson, H. Wehbe-Janek, K. Ghoshal, S. T. Jacob, and T. Patel, "MicroRNA-21 regulates expression of the PTEN tumor suppressor gene in human hepatocellular cancer," *Gastroenterology*, vol. 133, no. 2, pp. 647–658, 2007.
- [67] R. Rupaimoole and F. J. Slack, "MicroRNA therapeutics: towards a new era for the management of cancer and other diseases," *Nature Reviews Drug Discovery*, vol. 16, no. 3, pp. 203–222, 2017.
- [68] S. Volinia, G. A. Calin, C. G. Liu et al., "A microRNA expression signature of human solid tumors defines cancer gene targets," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 7, pp. 2257–2261, 2006.
- [69] N. Yanaihara, N. Caplen, E. Bowman et al., "Unique microRNA molecular profiles in lung cancer diagnosis and prognosis," *Cancer Cell*, vol. 9, no. 3, pp. 189–198, 2006.
- [70] M. Fabbri, R. Garzon, A. Cimmino et al., "MicroRNA-29 family reverts aberrant methylation in lung cancer by targeting DNA methyltransferases 3A and 3B," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 40, pp. 15805–15810, 2007.
- [71] F. Yu, H. Yao, P. Zhu et al., "let-7 regulates self renewal and tumorigenicity of breast cancer cells," *Cell*, vol. 131, no. 6, pp. 1109–1123, 2007.
- [72] M. V. Iorio, M. Ferracin, C. G. Liu et al., "MicroRNA gene expression deregulation in human breast cancer," *Cancer Research*, vol. 65, no. 16, pp. 7065–7070, 2005.
- [73] J. Wang, P. Huang, H. Zhao, Z. Zhang, B. Zhao, and D. L. Lee, "Billion-scale commodity embedding for E-commerce recommendation in Alibaba," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 839–848, London, United Kingdom, 2018.