WILEY | Hindawi

Research Article

Prediction of circRNA-miRNA Associations Based on Network Embedding

Wei Lan [iD],[1,2] **Mingrui Zhu,**[1] **Qingfeng Chen** [iD],[1] **Jianwei Chen** [iD],[1] **Jin Ye** [iD],[1] **Jin Liu** [iD],[3] **Wei Peng** [iD],[4] **and Shirui Pan** [iD][5]

[1]*School of Computer, Electronics and Information, Guangxi University, Nanning 530004, China*
[2]*Guangxi Key Laboratory of Multimedia Communications and Network Technology, Guangxi University, Nanning 530004, China*
[3]*Hunan Provincial Key Lab on Bioinformatics, School of Computer Science and Engineering, Central South University, Changsha 410083, China*
[4]*Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650050, China*
[5]*Faculty of Information Technology, Monash University, Melbourne 3800, Australia*

Correspondence should be addressed to Qingfeng Chen; qingfeng@gxu.edu.cn

circRNA is a novel class of noncoding RNA with closed-loop structure. Increasing biological experiments have shown that circRNAs play an important role in many diseases by acting as a miRNA sponge to indirectly regulate the expression of miRNA target genes. Therefore, predicting associations between circRNAs and miRNAs can promote the understanding of pathogenesis of disease. In this paper, we propose a new computational method, NECMA, based on network embedding to predict potential associations between circRNAs and miRNAs. In our method, the Gaussian interaction profile (GIP) kernel similarities of circRNA and miRNA are calculated based on the known circRNA-miRNA associations, respectively. Then, the circRNA-miRNA association network, circRNA GIP kernel similarity network, and miRNA GIP kernel similarity network are utilized to construct the heterogeneous network. Furthermore, the network embedding algorithm is used to extract potential features of circRNA and miRNA from the heterogeneous network, respectively. Finally, the associations between circRNAs and miRNAs are predicted by using neighborhood regularization logic matrix decomposition and inner product. The performance of NECMA is evaluated by using ten-fold cross-validation. The results show that this method has better prediction accuracy than other state-of-the-art methods.

1. Introduction

circRNA is a new group of endogenous noncoding RNA that is highly represented in the mammalian transcriptome [1]. Compared with other noncoding RNAs (such as lncRNAs and miRNAs), circRNAs did not receive extensive attention in the early stage. With the development of high-throughput biological sequencing technology, more and more circRNA molecules have been discovered. Increasing studies have shown that circRNA does not have $5'$-terminal cap and $3'$-terminal poly (A) tail and can form a closed ring structure with covalent bonds [2]. Like other noncoding RNAs,

circRNAs are also widely found in eukaryotes' brains [3], stomachs [4], and mammary glands [5]. Meanwhile, circRNAs are more stable than other linear noncoding RNAs due to their unique circular structure [6]. In addition, the unique structure of circRNA enables it to regulate gene transcription and expression [7]. For example, ciRs7 can bind related miRNAs and act as a miR-7 sponge to affect miR-7 binding to the target gene [8]. In addition, it has been found that circHIPK3 can sponge miR-124 and inhibit the activity of miR-124 in malignant tumors to achieve the purpose of regulating cell growth [9]. Numerous evidences have shown that miRNAs are closely related to a variety of

diseases [10]. For example, miR-145 inhibits colon cancer cell growth by targeting the insulin receptor substrate [11]. Therefore, predicting the potential associations between circRNAs and miRNAs can help biologists to understand complex pathogenesis of disease and further contribute to disease diagnoses.

With the continuous development of high-throughput sequencing technology, a large number of circRNAs have been discovered. Simultaneously, plenty of databases are developed to store circRNA-related information such as circBase [12], circR2Disease [13], circRNADisease [14], and circ2Disease [15]. circBase is an online database that provides users with a variety of basic circRNA information such as circRNA ID, sequence, gene description, and location [12]. circR2Disease is a public database that stores experimentally verified circRNA-related disease information. The database contains 793 circRNA-disease associations including 661 circRNAs and 100 diseases [13]. The circRNADisease database contains 354 circRNA-disease interactions, 330 circRNAs, and 48 diseases [14]. Similar to circR2Disease and circRNADisease, the circ2Disease database is used to store a vitro-proved circRNA-disease association database through which users can obtain circRNA-disease associations and the associations between miRNAs and its targets [15]. These databases enable users to identify potential associations between circRNA and miRNA by using computational methods.

Compared with traditional biological experiment methods, the circRNA-miRNA association prediction based on computational methods can maintain high accuracy and be less time-consuming. Therefore, more and more attention has been paid to circRNA-miRNA association prediction based on computational methods. At present, a large number of computational prediction models have been applied in many fields of biology, for example, predicting associations between diseases and genes, miRNA-disease associations [16, 17], circRNA-disease associations [18, 19], lncRNA-disease associations [20, 21], protein function [22, 23], drug-target interactions [24, 25], and lncRNA-miRNA associations [26, 27]. Compared with other fields, there are few prediction models based on the computational method in the circRNA-miRNA association prediction. Therefore, it is urgent to develop an effective computational method to infer circRNA-miRNA associations.

For the above purposes, in this study, we propose a new computational algorithm based on network embedding, NECMA, to predict circRNA-miRNA association. In our method, the circRNA-miRNA network is constructed based on experimental verified circRNA-miRNA associations. Then, based on the circRNA-miRNA associations, the GIP kernel similarities of circRNA and miRNA are calculated by using the Gaussian interaction profile kernel similarity, respectively. Furthermore, the circRNA GIP kernel similarity network, miRNA GIP kernel similarity network, and circRNA-miRNA association network are integrated to construct the circRNA-miRNA heterogeneous network. In addition, the network embedding model is employed to learn the features of circRNA and miRNA based on the circRNA-miRNA heterogeneous network, respectively.

Finally, the weighted neighborhood regularized logistic matrix factorization and inner product are combined to predict potential circRNA-miRNA associations. The tenfold cross-validation is used to evaluate the performance of our method. The experimental results show that NECMA achieves better performance than other state-of-the-art methods. In addition, the case study shows that NECMA could effectively infer potential circRNA-miRNA associations which are confirmed by the latest literature.

2. Related Work

Numerous experiments have shown that circRNA and miRNA have a close association with diseases. The current circRNA-disease association prediction algorithms are divided into the following categories. (1) Network-based circRNA-disease association prediction method: Fan et al. [28] used known circRNA-disease associations, circRNA expression profile similarities, and disease phenotype similarities to construct the circRNA-disease heterogeneous network and then used KATZ to predict potential associations between circRNAs and diseases. Li et al. [29] integrated known circRNA-disease associations, circRNA functional similarities, and disease semantic similarities and utilized network-consistent projections to identify potential circRNA-disease associations. Zhao et al. [30] developed an ensemble learning algorithm to predict the potential association between circRNA and diseases. In this method, the circRNA-disease heterogeneous network is constructed from known circRNA-disease association network, circRNA similarity network, and disease similarity network and circRNA-disease association is predicted by using KATZ and bipartite network projections. Lei and Bian [31] used random walk with restart and KNN algorithms to identify potential associations between circRNAs and diseases based on known circRNA-disease associations, circRNA similarities, and disease similarities. Li et al. [32] predicted potential circRNA-disease associations based on known circRNA-disease association networks, circRNA similarity, and disease similarity by using inductive matrix completion. Wei and Liu [18] reconstructed the circRNA-disease association network using circRNA similarities and disease similarities and then used nonnegative matrix factorization to predict potential associations. (2) Machine learning-based circRNA-disease association prediction method: Lei and Fang [33] fused circRNA expression profile similarity network, circRNA sequence similarity network, and circRNA functional annotation similarity network to construct the circRNA similarity network. The disease similarity network is constructed by integrating the disease functional similarity network and the disease semantic similarity network. Finally, based on the known circRNA-disease association network, circRNA similarity network, and disease similarity network, the potential feature of circRNA and diseases were extracted, respectively, and then the gradient boosting decision tree algorithm was used to predict the potential circRNA-disease associations.

Similarly, miRNA-disease association prediction algorithms can also be classified into the similar categories. Peng

et al. [34] developed a miRNA-disease association prediction model (ThrWRDE) that performs a restart random walk algorithm on a variety of miRNA-related biological data and then integrates the results obtained from multiple restart random walk models. You et al. [35] integrated known miRNA-disease associations, miRNA similarity, and disease similarity to construct a miRNA-disease heterogeneous network and then used a depth-first search algorithm to count the path between miRNA and diseases. Finally, the different pathways between miRNA and diseases are integrated to obtain the predicted association score between miRNA and diseases. Chen et al. [36] constructed miRNA similarity and disease similarity by integrating miRNA functional similarity and miRNA Gaussian interaction profile kernel similarity and disease Gaussian interaction profile kernel similarity and disease semantic similarity, respectively. Finally, the inductive matrix completion is used to obtain the final predicted miRNA-disease association. Chen et al. [37] extracted the potential representations of miRNA and disease, respectively, using a stacked autoencoder and then obtained the predicted score of miRNA-disease association by using support vector machine (SVM).

3. Materials and Methods

3.1. Materials.
The circRNA-miRNA associations are downloaded from circR2Cancer database (http://www.biobdlab.cn:8000/). The circR2Cancer database [38] is a manually curated database which contains not only circRNA-cancer association data but also circRNA-miRNA association data and miRNA-cancer association data. After removing redundant data, 130 circRNAs, 412 miRNAs, and 477 associations are extracted in final. Furthermore, the adjacent matrix $CM_{m \times n}$ is constructed to represent circRNA-miRNA association, where m represents the number of circRNAs and n represents the number of miRNAs. The value of element $CM(i, j)$ is equal to 1 when circRNA i is related to miRNA j, otherwise 0.

3.2. circRNA and miRNA Similarity Calculation.
In this study, the Gaussian interaction profile (GIP) kernel similarity is used to calculate similarities of circRNA and miRNA. Based on the assumption that circRNAs with similar functions are often associated with similar miRNAs, circRNA GIP kernel similarity and miRNA GIP kernel similarity are calculated based on the circRNA-miRNA interaction network, respectively. For pairwise circRNAs c_i and c_j, the GIP kernel similarity $CS(c_i, c_j)$ between circRNAs c_i and c_j is defined as follows:

$$CS(c_i, c_j) = \exp\left(-\gamma_c \left\| CM(c_i, :) - CM(c_j, :) \right\|^2\right), \quad (1)$$

where $CM(c_i, :)$ represents the c_i row in the matrix CM and γ_c represents the kernel bandwidth, which is defined as follows:

$$\gamma_c = \frac{1}{\left(1/n_c \sum_{i=1}^{n_c} \left\| CM(c_i, :) \right\|^2\right)}, \quad (2)$$

where n_c represents the number of rows in matrix CM.

Similarly, the miRNA GIP kernel similarity $MS(m_i, m_j)$ between miRNA m_i and miRNA m_j is defined as follows:

$$MS(m_i, m_j) = \exp\left(-\gamma_m \left\| CM(:, m_i) - CM(:, m_i) \right\|^2\right), \quad (3)$$

where $CM(:, m_i)$ represents m_i column of matrix CM and γ_m represents the kernel bandwidth, which is defined as follows:

$$\gamma_m = \frac{1}{\left(1/n_m \sum_{i=1}^{n_m} \left\| CM(:, m_i) \right\|^2\right)}, \quad (4)$$

where n_m represents the number of columns in the interaction matrix CM.

3.3. Construction of Heterogeneous Network.
The heterogeneous network used for circRNA-miRNA association prediction is composed of three subnetworks including circRNA-miRNA interaction network, circRNA GIP kernel similarity network, and miRNA GIP kernel similarity network. Based on the above three subnetworks, the heterogeneous network H is constructed as follows:

$$H = \begin{bmatrix} CS & CM \\ CM^T & MS \end{bmatrix}, \quad (5)$$

where CM represents the circRNA-miRNA interaction network, CM^T represents the transpose of the circRNA-miRNA interaction network, CS represents the circRNA GIP kernel similarity network, and MS represents the miRNA GIP kernel similarity network.

3.4. The Feature Extraction Based on NetMF.
After obtaining the heterogeneous network H, the network embedding as matrix factorization (NetMF) algorithm [39] is used to extract the potential features of circRNA and miRNA on the heterogeneous network, respectively. NetMF is a matrix factorization framework based on the original DeepWalk algorithm. To be specific, the NetMF model is the matrix factorization form of DeepWalk algorithm derived from the implicit decomposition model of the skip-gram with negative-sampling model (SGNS) [40, 41]. It can reduce the noise information in the matrix H and improve the performance of the prediction model. The NetMF model is defined as the probability distribution of truncated random walk, which is calculated as

$$\text{Pro} = D^{-1} \cdot H, \quad (6)$$

where D denotes a diagonal matrix and the elements in D represent the generalized degree of nodes in the matrix H.

Then, we conducted w times random walk on the heterogeneous network H according to the probability

distribution calculated before. It is used to sample the nodes in the heterogeneous network to obtain the transfer matrix trans_H which is defined as follows:

$$\text{trans\_}H = \sum_{m=1}^{w} \text{Pro}. \tag{7}$$

After obtaining the transition matrix trans_H, the DeepWalk matrix is obtained as

$$H_{\text{DeepWalk}} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} H(i,j)}{bw} \cdot \text{trans\_}H \cdot D^{-1}, \tag{8}$$

where n represents the dimension of heterogeneous network H and b represents the number of negative samples.

Since the density of H_{DeepWalk} increases the time complexity of subsequent calculations, the approximate matrix H'_{DeepWalk} is defined as follows:

$$H'_{\text{DeepWalk}} = \max\left(H_{\text{DeepWalk}}, 1\right). \tag{9}$$

After obtaining the matrix H_{DeepWalk} of the circRNA-miRNA heterogeneous network H, the low-dimensional space feature vectors of circRNA and miRNA are obtained by using the singular value decomposition (SVD) model [42–44], which is defined as follows:

$$U_d \sum_d V_d^T = \log H'_{\text{DeepWalk}}, \tag{10}$$

where d represents the dimension of a low-dimensional space.

Finally, the eigenmatrix eigen_matrix is calculated:

$$\text{eigen\_matrix} = U_d \sum_d. \tag{11}$$

The dimension of eigen_matrix is $m + n$ and d. This matrix is composed of circRNA feature vectors u and miRNA feature vectors v, in which the dimensions are m and n, respectively.

3.5. circRNA-miRNA Association Prediction.

The potential eigenvectors of circRNA and miRNA are obtained by NetMF on heterogeneous network H. Then, the weighted neighborhood regularized logistic matrix factorization [45] and inner product are utilized to reconstruct the circRNA-miRNA association matrix.

The weighted neighborhood regularized logistic matrix factorization is defined as follows:

$$\text{matrix\_pre}(i,j) = \frac{e^{\left((\alpha)u_i(1-\alpha)v_j^T\right)}}{1 + e^{\left((\alpha)u_i(1-\alpha)v_j^T\right)}}. \tag{12}$$

The inner product is defined as follows:

$$\text{inner\_pre}(i,j) = u_i v_j^T, \tag{13}$$

where u_i represents the feature vector of circRNA i and v_j represents the feature vector of miRNA j. α represents the weight coefficient to balance the influence of two feature vectors on the reconstructed matrix.

Finally, the score of circRNA-miRNA association matrix is defined as follows:

$$\text{Pre}(i,j) = \max(\text{matrix\_pre}(i,j), \text{inner\_pre}(i,j)), \tag{14}$$

where $\text{Pre}(i,j)$ denotes the predicted score between circRNA i and miRNA j.

The flowchart of NECMA is shown in Figure 1. It mainly contains the following steps: first, the Gaussian interaction profile kernel similarity is utilized to calculate circRNA similarity and miRNA similarity based on the known circRNA-miRNA associations, respectively. Then, the heterogeneous network H is constructed based on the circRNA-miRNA association network, miRNA similarity network, and circRNA similarity network. Furthermore, the NetMF is used to extract the low-dimensional features of circRNA and miRNA on heterogeneous network H, respectively. Finally, the weighted neighborhood regularized logistic matrix factorization and the inner product are utilized to reconstruct the circRNA-miRNA association matrix based on the circRNA feature vector and miRNA feature vector (Algorithm 1).

4. Result

4.1. Ten-Fold Cross-Validation.

In order to evaluate the performance of NECMA, we conduct the ten-fold cross-validation in the experiment. In the ten-fold cross-validation, the known circRNA-miRNA associations are randomly divided into ten subsets. Then, in each round of cross-validation experiment, one set is selected as the test samples and the other nine sets are treated as the training samples which are used in model training. The final score of circRNA-miRNA association is predicted by using the model. The higher the score of the association, the higher the probability of circRNA-miRNA interaction. Then, we rearranged the score of circRNA-miRNA association in descending order. Furthermore, the true positive rate (TPR) and false positive rate (FPR) are calculated by altering the threshold. The TPR and FPR are defined as follows:

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}},$$
$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}}, \tag{15}$$

where TP and FP represent true positive and false positive, respectively, and TN and FN represent true negative and false negative, respectively. Finally, the receiver operating characteristics (ROC) curve is plotted based on TPR and FPR, and the area under ROC curve (AUROC) is calculated to evaluate the predictive power of the model. The higher the AUROC value, the better the performance of the model.

Similarly, the area under precision-recall (AUPR) curve based on precision and recall is used to evaluate the performance of prediction model. The precision and recall are defined as follows:
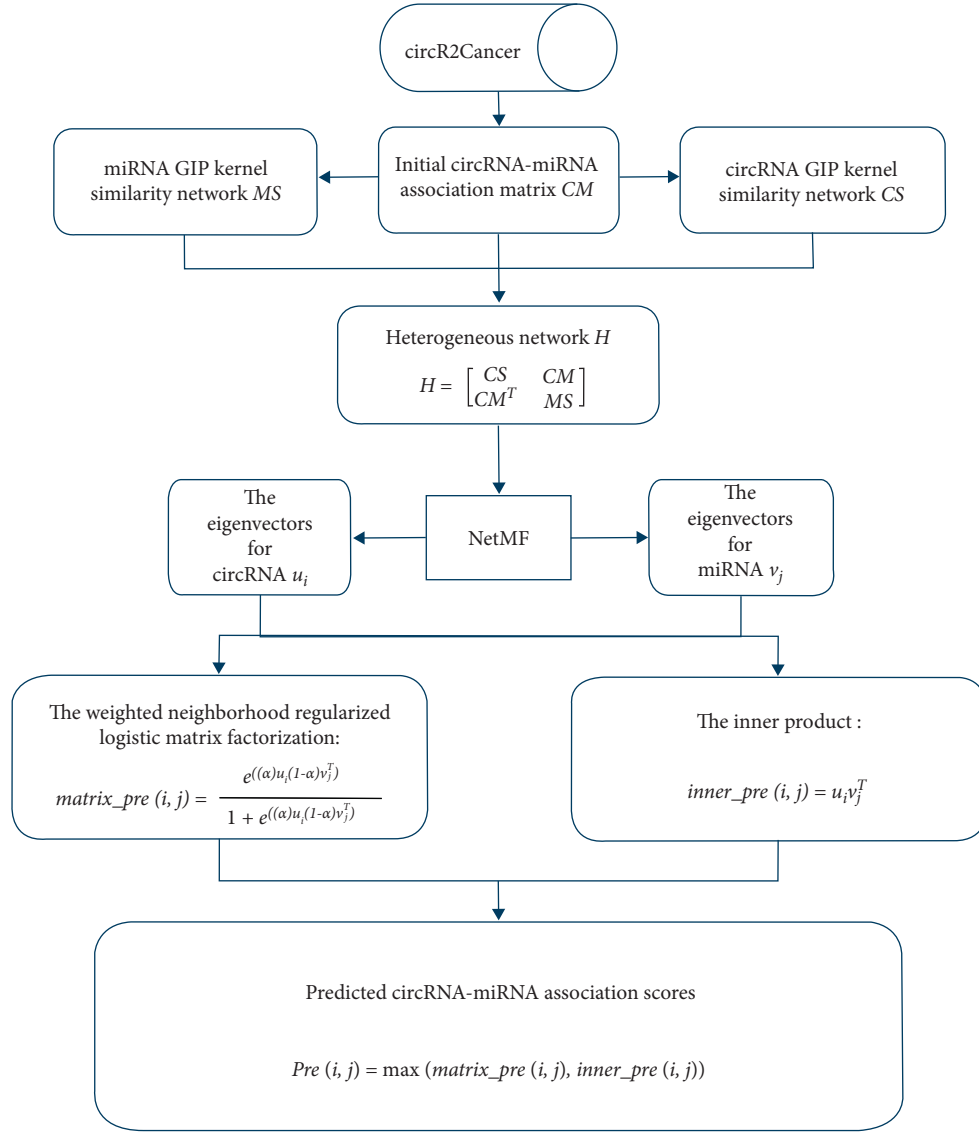
Figure 1: The flowchart of NECMA.

Input: circRNA-miRNA association matrix CM
Output: Predicted association matrix Pre
(1) Calculate the circRNA GIP kernel similarity CS based on known circRNA-miRNA associations
(2) Calculate the miRNA GIP kernel similarity MS based on known circRNA-miRNA associations
(3) Construct a heterogeneous network H based on CM, CS, and MS
(4) Calculate generalized degree matrix D based on heterogeneous network H
(5) Calculate the probability of truncated random walk Pro $= D^{-1} \cdot H$
(6) Calculate the transition matrix trans$\_H = \sum_{m=1}^{w}$ Pro
(7) Calculate $H_{\text{DeepWalk}} = (\sum_{i=1}^{n} \sum_{j=1}^{n} H(i,j)/bw) \cdot \text{trans}\_H \cdot D^{-1}$
(8) Calculate $H'_{\text{DeepWalk}} = \max(H_{\text{DeepWalk}}, 1)$
(9) Dimensionality reduction through SVD $U_d \sum_d V_d^T = \log H'_{\text{DeepWalk}}$
(10) Calculate circRNA-miRNA feature matrix eigen$\_$matrix $= U_d \sum_d$
(11) Extract circRNA feature u and miRNA feature v based on eigen$\_$matrix
(12) Calculate matrix$\_$pre$(i,j) = e^{((\alpha)u_i(1-\alpha)v_j^T)}/1 + e^{((\alpha)u_i(1-\alpha)v_j^T)}$
(13) Calculate inner$\_$pre$(i,j) = u_i v_j^T$
(14) **Return:** Predicted association matrix: Pre$(i,j) = \max(\text{matrix\_pre}(i,j), \text{inner\_pre}(i,j))$

Algorithm 1: NECMA.

$$precision = \frac{TP}{TP + FP},$$

$$recall = \frac{TP}{TP + FN},$$

(16)

where precision represents the proportion of positive examples in the predicted results to the actual positive examples and recall represents the proportion of all true positive cases divided into positive cases, which measures the classifier's ability to recognize positive cases.

In addition, in order to demonstrate the superiority of NECMA in predicting the potential association of circRNA-miRNA. We compare NECMA with three state-of-the-art algorithms including RWRLncD [46], NCPLDA [47], and LRLSLDA [48]. Figures 2 and 3 show the AUROC and AUPR values obtained by different prediction models in ten-fold cross-validation, respectively. The results of ten-fold cross-validation show that the AUROC and AUPR of the NECMA are better than other three prediction algorithms. It can be found from Figure 2 that the AUROC value of NECMA is 0.8264 which is higher than RWRLncD (0.5243), NCPLDA (0.6985), and LRLSLDA (0.7661). Simultaneously, it can be observed from Figure 3 that the AUPR value of NECMA is 0.0048 which is higher than RWRLncD (0.0016), NCPLDA (0.0011), and LRLSLDA (0.0026). The overall results of ten-fold cross-validation are shown in Table 1. It can be concluded that NECMA is an effective method in identifying association between circRNA and miRNA.

4.2. Effect of Parameters. There are three parameters contained in the NetMF model (context window w, negative sampling number b, and embedded dimension d). According to the previous study [39], both the context window w and the negative sampling number b are set to 1. To test the effect of embedded dimension d, we set embedded dimension d ranging from 8 to 128. The result is shown in Figure 4. It can be found that the AUROC value of NECMA is the highest when the embedded dimension d of the NetMF is set to 8. In addition, we also test the effect of parameter weight coefficient α in neighborhood regularization logistic matrix factorization which is used to balance the influence of two eigenvectors in the process of matrix reconstruction. The parameter α ranges from 0.1 to 0.9 with 0.1 increasing in each time. The influence of parameter α on the prediction performance is shown in Figure 5. It can be observed that the AUROC value obtained is the highest when $\alpha = 0.6$.

4.3. Case Study. To further illustrate the ability of NECMA to predict potential circRNA-miRNA associations, we conduct a case study on miR-130a-3p. We select the top 10 circRNAs predicted by NECMA and prove these associations by manually retrieving related databases and literature.

The numerous experiments have shown that miR-130a-3p is associated with proliferation and migration of many cancer cells [49]. For example, miR-130a-3p can regulate its target Smad4 to inhibit migration and invasion of gemcitabine-resistant (GR) hepatocellular carcinoma (HCC)
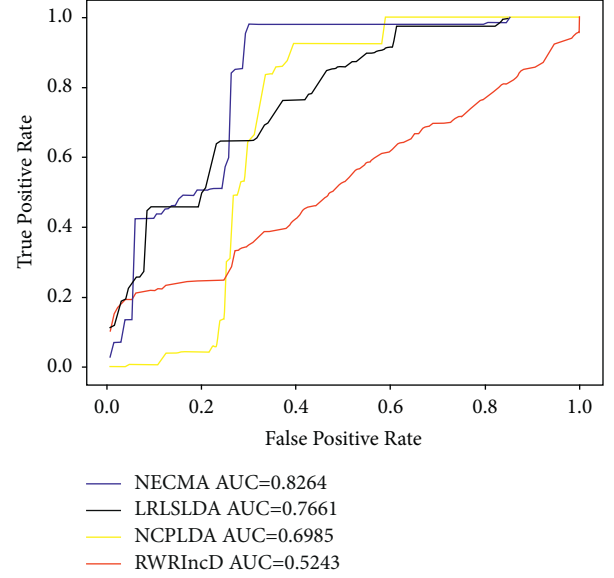

FIGURE 2: The AUROC values of the NECMA model and the other three algorithms on the ten-fold cross-validation.
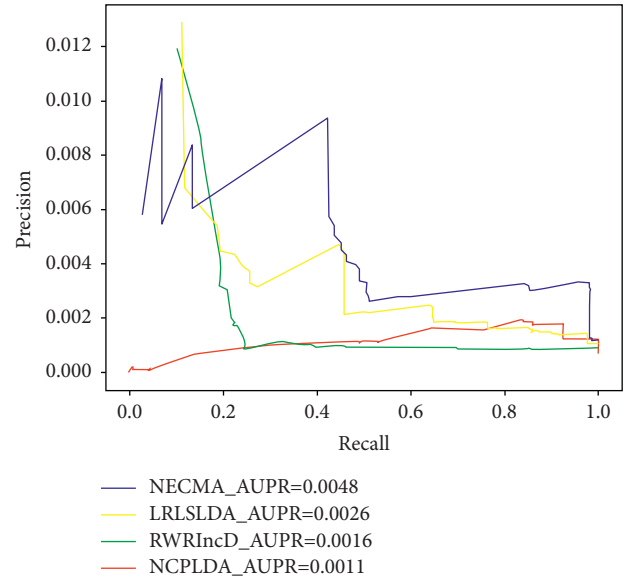

FIGURE 3: The AUPR values of the NECMA model and the other three algorithms on the ten-fold cross-validation.

cells [50]. Therefore, the correct prediction of the circRNAs associated with miR-130a-3p is useful for understanding complex disease mechanisms. The top 10 predicted circRNAs of miR-130a-3p are shown in Table 2. The results showed that nine circRNAs (hsa_circ_0068942, hsa_circ_0089378, hsa_circ_0083357, hsa_circ_0006323, hsa_circ_0032970, hsa_circ_0051172, hsa_circ_0054537, hsa_circ_0057576, and hsa_circ_0082824) have been confirmed in the literature. It has been confirmed that the moderately upregulated hsa_circ_0068942 ranked at top 1 can serve as miR-130a-3p sponge and the disease marker for coronary artery disease (CAD) [51]. It has been demonstrated that hsa_circ_0089378 ranked at top 2 can act as

TABLE 1: Performance comparison of NECMA with RWRLncD, LRLSLDA, and NCPLDA.

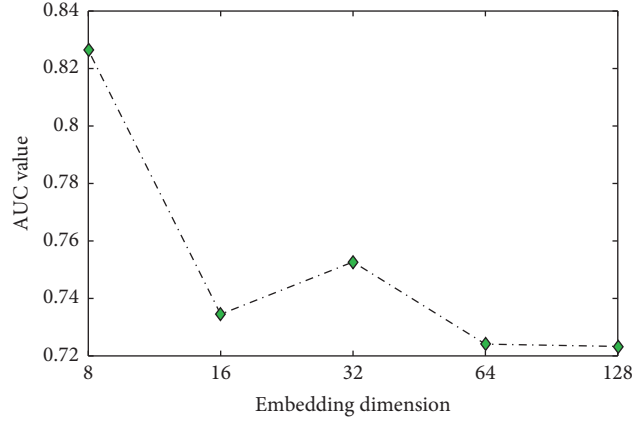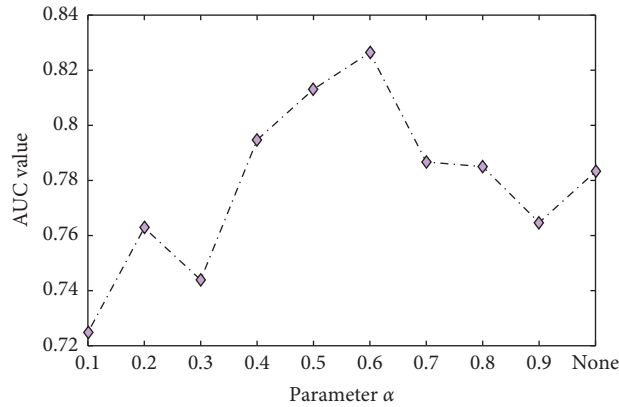| Methods	AUC	AUPR	Precision	Recall	Accuracy	F1-score	Specificity
RWRLncD	0.5243	0.0016	0.0013	0.5321	0.4922	0.0026	0.4922
LRLSLDA	0.7661	0.0026	0.0019	0.7720	0.4926	0.0039	0.4924
NCPLDA	0.6958	0.0011	0.0009	0.7005	0.4981	0.0018	0.4979
NECMA	**0.8264**	**0.0048**	**0.0026**	**0.8277**	**0.4993**	**0.0053**	**0.4989**

The bold values indicate best performance.

FIGURE 4: The impact of embedded dimension d.

FIGURE 5: The influence of parameter α.

TABLE 2: Top ten candidate circRNAs for miR-130a-3p.

Rank	circRNA	Evidence
1	hsa_circ_0068942	PMID: 30159442
2	hsa_circ_0089378	PMID: 31119072
3	hsa_circ_0083357	PMID: 30368217
4	hsa_circ_0006323	PMID: 29730164
5	hsa_circ_0032970	PMID: 28947970
6	hsa_circ_0051172	PMID: 29182528
7	hsa_circ_0054537	PMID: 28947970
8	hsa_circ_0057576	PMID: 28947970
9	hsa_circ_0082824	PMID: 29182528
10	hsa_circ_0005986	Uknown

the sponge for miR-130a-3p to affect its target mRNA expression in coronary artery disease [52]. It has been demonstrated that hsa_circ_0083357 ranked at top 3 can play an important role in coronary artery disease through miR-130a-3p-mediated circRNA-mRNA-competitive endogenous RNA (ceRNA) networks [53]. It has been found that hsa_circ_0006323 ranked at top 4 can inhibit the expression of miR-130a-3p in coronary artery disease cells

[54]. It has been demonstrated that hsa_circ_0032970 ranked at top 5 can bind to miR-130a-3p binding sites in coronary artery disease cells [55]. It has been demonstrated that hsa_circ_0051172 ranked at top 6 can regulate the expression of TRPM3 by targeting miR-130a-3p in coronary artery disease [56]. In addition, it has been confirmed that differential expression of hsa_circ_0054537 ranked at top 7 and hsa_circ_0057576 ranked at top 8 can not only inhibit miR-130a-3p but also lead to upregulation of TRPM3 [55]. It has been discovered that hsa_circ_0082824 ranked at top 9 can promote the expression of TRPM3 in target cells in coronary artery disease by inhibiting miR-130a-3p [54].

5. Conclusion

Accumulating experiments have shown that predicting associations between circRNAs and miRNAs not only helps to understand complex disease mechanisms but also contributes to prevent and diagnose diseases [57]. In this study, we propose a computational method, NECMA, to infer circRNA-miRNA associations. In this model, we first construct the circRNA-miRNA association matrix based on known circRNA-miRNA associations. Then, the Gaussian interaction profile kernel similarity is used to calculate circRNA similarity and miRNA similarity based on known circRNA-miRNA associations, respectively. Furthermore, the heterogeneous network is constructed based on three subnetworks (circRNA-miRNA association network, circRNA similarity network, and miRNA similarity network). In addition, the NetMF is employed to extract the subspace features of circRNA and miRNA from the heterogeneous network, respectively. Finally, the scores of circRNA-miRNA associations are predicted by using weighted neighborhood regularized logistic matrix factorization and inner product. In order to show the performance of NECMA, we compare NECMA with three state-of-the-art methods (RWRlncD, NCPLDA, and LRLSLDA) in terms of ten-fold cross-validation. The experimental results show that the NECMA achieves a higher AUROC value (0.8264) than the other three prediction models. In addition, it is demonstrated that NECMA could correctly identify potential associations between circRNA and miRNAs by constructing a case study on miR-130a-3p.

Although the NECMA model can effectively predict the potential circRNA-miRNA association, there are still many limitations. First, the NECMA model mainly relies on known circRNA-miRNA association data and the imbalance of positive and negative samples will greatly affect the prediction accuracy of the model. Second, the setting of parameters will also affect the prediction results of the model. In addition, the integration of various circRNA and miRNA information can further improve the predictive power of the model [58–60]. Moreover, the NECMA model cannot predict new circRNA-miRNA without any known association. Therefore, we will integrate more biological data of circRNA and miRNA in the future, which will make it more reliable [61–63].

Data Availability

The underlying data supporting the results of our study can be found at http://www.biobdlab.cn:8000/.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this study.

Acknowledgments

References

[1] S. Qu, X. Yang, X. Li et al., "Circular RNA: a new star of noncoding RNAs," *Cancer Letters*, vol. 365, no. 2, pp. 141–148, 2015.

[2] L. H. Jiang, D. W. Sun, J. C. Hou, H.-d. Zhang, and Z. L. Ji, "CircRNA: a novel type of biomarker for cancer," *Breast Cancer*, vol. 25, no. 1, pp. 1–7, 2018.

[3] M. Hanan, H. Soreq, and S. Kadener, "CircRNAs in the brain," *RNA Biology*, vol. 14, no. 8, pp. 1028–1034, 2017.

[4] L.-L. Chen, "The biogenesis and emerging roles of circular RNAs," *Nature Reviews Molecular Cell Biology*, vol. 17, no. 4, pp. 205–211, 2016.

[5] A. A. Nair, N. Niu, X. Tang et al., "Circular RNAs and their associations with breast cancer subtypes," *Oncotarget*, vol. 7, no. 49, pp. 80967–80979, 2016.

[6] S. Memczak, M. Jens, A. Elefsinioti et al., "Circular RNAs are a large class of animal RNAs with regulatory potency," *Nature*, vol. 495, no. 7441, pp. 333–338, 2013.

[7] G. Huang, S. Li, N. Yang, Y. Zou, D. Zheng, and T. Xiao, "Recent progress in circular RNAs in human cancers," *Cancer Letters*, vol. 404, pp. 8–18, 2017.

[8] W. Weng, Q. Wei, S. Toden et al., "Circular RNA ciRS-7-A promising prognostic biomarker and a potential therapeutic target in colorectal cancer," *Clinical Cancer Research*, vol. 23, no. 14, pp. 3918–3928, 2017.

[9] G. Chen, Y. Shi, M. Liu, and J. Sun, "circHIPK3 regulates cell proliferation and migration by sponging miR-124 and regulating AQP3 expression in hepatocellular carcinoma," *Cell Death & Disease*, vol. 9, no. 2, pp. 175–213, 2018.

[10] W. Lan, J. Wang, M. Li, W. Peng, and F. Wu, "Computational approaches for prioritizing candidate disease genes based on PPI networks," *Tsinghua Science and Technology*, vol. 20, no. 5, pp. 500–512, 2015.

[11] J. Zhang, H. Guo, H. Zhang et al., "Putative tumor suppressor miR-145 inhibits colon cancer cell growth by targeting oncogene friend leukemia virus integration 1 gene," *Cancer*, vol. 117, no. 1, pp. 86–95, 2011.

[12] P. Glažar, P. Papavasileiou, and N. Rajewsky, "circBase: a database for circular RNAs," *RNA*, vol. 20, no. 11, pp. 1666–1670, 2014.

[13] C. Fan, X. Lei, Z. Fang, Q. Jiang, and F. X. Wu, "CircR2-Disease: a manually curated database for experimentally supported circular RNAs associated with various diseases," *Database*, vol. 2018, pp. 1–6, 2018.

[14] Z. Zhao, K. Wang, F. Wu et al., "circRNA disease: a manually curated database of experimentally supported circRNA-disease associations," *Cell Death & Disease*, vol. 9, no. 5, pp. 475–482, 2018.

[15] D. Yao, L. Zhang, M. Zheng, X. Sun, Y. Lu, and P. Liu, "Circ2Disease: A manually curated database of experimentally validated circRNAs in human disease," *Scientific Reports*, vol. 8, no. 1, pp. 1–6, 2018.

[16] W. Lan, J. Wang, M. Li, J. Liu, F.-X. Wu, and Y. Pan, "Predicting microRNA-disease associations based on improved microRNA and disease similarities," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 15, no. 6, pp. 1774–1782, 2018.

[17] W. Peng, W. Lan, J. Zhong, J. Wang, and Y. Pan, "A novel method of predicting microRNA-disease associations based on microRNA, disease, gene and environment factor networks," *Methods*, vol. 124, pp. 69–77, 2017.

[18] H. Wei and B. Liu, "iCircDA-M. F.: iCircDA-MF: identification of circRNA-disease associations based on matrix factorization," *Briefings in Bioinformatics*, vol. 21, no. 4, pp. 1356–1367, 2020.

[19] C. Yan, J. Wang, and F. X. Wu, "DWNN-RLS.: regularized least squares method for predicting circRNA-disease associations," *BMC Bioinformatics*, vol. 19, no. 19, Article ID 520, 2018.

[20] W. Lan, M. Li, K. Zhao et al., "LDAP: a web server for lncRNA-disease association prediction," *Bioinformatics*, vol. 33, no. 3, pp. 458–460, 2017.

[21] X. Wu, W. Lan, Q. Chen, Y. Dong, J. Liu, and W. Peng, "Inferring LncRNA-disease associations based on graph autoencoder matrix completion," *Computational Biology and Chemistry*, vol. 87, Article ID 107282, 2020.

[22] R. Cao and J. Cheng, "Integrated protein function prediction by mining function associations, sequences, and protein-protein and gene-gene interaction networks," *Methods*, vol. 93, pp. 84–91, 2016.

[23] W. Peng, J. Du, L. Li, W. Dai, and W. Lan, "Predicting protein functions by using non-negative matrix factorisation with multi-networks co-regularisation," *International Journal of Data Mining and Bioinformatics*, vol. 23, no. 4, pp. 318–342, 2020.

[24] W. Lan, J. Wang, M. Li et al., "Predicting drug-target interaction using positive-unlabeled learning," *Neurocomputing*, vol. 206, pp. 50–57, 2016.

[25] C. Yan, J. Wang, W. Lan, F. X. Wu, and Y. Pan, "SDTRLS: predicting drug-target interactions for complex diseases based on chemical substructures," *Complexity*, vol. 2017, Article ID 2713280, 10 pages, 2017.

[26] H. Liu, G. Ren, H. Chen, Q. Liu, Y. Yang, and Q. Zhao, "Predicting lncRNA-miRNA interactions based on logistic matrix factorization with neighborhood regularized," *Knowledge-Based Systems*, vol. 191, Article ID 105261, 2020.

[27] W. Lan, L. Huang, D. Lai, and Q. Chen, "Identifying interactions between long noncoding RNAs and diseases based on computational methods," *Methods in Molecular Biology*, vol. 1754, pp. 205–221, 2018.

[28] C. Fan, X. Lei, and F.-X. Wu, "Prediction of CircRNA-disease associations using KATZ model based on heterogeneous networks," *International Journal of Biological Sciences*, vol. 14, no. 14, pp. 1950–1959, 2018.

[29] G. Li, Y. Yue, C. Liang, Q. Xiao, P. Ding, and J. Luo, "NCPCDA: network consistency projection for circRNA-disease association prediction," *RSC Advances*, vol. 9, no. 57, pp. 33222–33228, 2019.

[30] Q. Zhao, Y. Yang, G. Ren, E. Ge, and C. Fan, "Integrating bipartite network projection and KATZ measure to identify novel CircRNA-disease associations," *IEEE Transactions on Nanobioscience*, vol. 18, no. 4, pp. 578–584, 2019.

[31] X. Lei and C. Bian, "Integrating random walk with restart and k-nearest neighbor to identify novel circRNA-disease association," *Scientific Reports*, vol. 10, no. 1, pp. 1943–1949, 2020.

[32] M. Li, M. Liu, Y. Bin, and J. Xia, "Prediction of circRNA-disease associations based on inductive matrix completion," *BMC Medical Genomics*, vol. 13, no. 5, pp. 42–13, 2020.

[33] X. Lei and Z. Fang, "GBDTCDA: predicting circRNA-disease associations based on gradient boosting decision tree with multiple biological data fusion," *International Journal of Biological Sciences*, vol. 15, no. 13, pp. 2911–2924, 2019.

[34] W. Peng, W. Lan, Z. Yu, J. Wang, and Y. Pan, "A framework for integrating multiple biological networks to predict MicroRNA-disease associations," *IEEE Transactions on Nanobioscience*, vol. 16, no. 2, pp. 100–107, 2016.

[35] 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 ID e1005455, 2017.

[36] 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.

[37] X. Chen, Y. Gong, D.-H. Zhang, Z.-H. You, and Z.-W. Li, "DRMDA: deep representations-based miRNA-disease association prediction," *Journal of Cellular and Molecular Medicine*, vol. 22, no. 1, pp. 472–485, 2018.

[38] W. Lan, M. Zhu, Q. Chen et al., "CircR2Cancer: a manually curated database of associations between circRNAs and cancers," *Database*, vol. 2020, 2020.

[39] J. Qiu, Y. Dong, H. Ma, J. Li, K. Wang, and J. Tang, "Network embedding as matrix factorization: unifying deepwalk, line, pte, and node2vec," in *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining (WSDM'18)*, pp. 459–467, Marina Del Rey, CA, USA, February 2018.

[40] J. Wu, X. Zhu, C. Zhang, and P. S. Yu, "Bag constrained structure pattern mining for multi-graph classification," *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 10, pp. 2382–2396, 2014.

[41] J. Wu, S. Pan, X. Zhu, and Z. Cai, "Boosting for multi-graph classification," *IEEE Transactions on Cybernetics*, vol. 45, no. 3, pp. 416–429, 2014.

[42] G. H. Golub and C. Reinsch, "Singular value decomposition and least squares solutions," *Numerische Mathematik*, vol. 14, no. 5, pp. 403–420, 1970.

[43] C. Y. Liu, C. Zhou, J. Wu, Y. Hu, and L. Guo, "Social recommendation with an essential preference space," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, no. 1, New Orleans, LA, USA, Febraury 2018.

[44] L. Gao, J. Wu, Z. Qiao, C. Zhou, H. Yang, and Y. Hu, "Collaborative social group influence for event recommendation," in *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, pp. 1941–1944, Indianapolis, IA, USA, October 2016.

[45] Y. Liu, M. Wu, C. Miao, P. Zhao, and X. L. Li, "Neighborhood regularized logistic matrix factorization for drug-target interaction prediction," *PLoS Computational Biology*, vol. 12, no. 2, 2016.

[46] J. Sun, H. Shi, Z. Wang et al., "Inferring novel lncRNA-disease associations based on a random walk model of a lncRNA functional similarity network," *Molecular BioSystems*, vol. 10, no. 8, pp. 2074–2081, 2014.

[47] G. Li, J. Luo, C. Liang, Q. Xiao, P. Ding, and Y. Zhang, "Prediction of LncRNA-disease associations based on network consistency projection," *IEEE Access*, vol. 7, pp. 58849–58856, 2019.

[48] X. Chen and G.-Y. Yan, "Novel human lncRNA-disease association inference based on lncRNA expression profiles," *Bioinformatics*, vol. 29, no. 20, pp. 2617–2624, 2013.

[49] X. Chen, B. Yue, C. Zhang et al., "MiR-130a-3p inhibits the viability, proliferation, invasion, and cell cycle, and promotes apoptosis of nasopharyngeal carcinoma cells by suppressing BACH2 expression," *Bioscience Reports*, vol. 37, no. 3, 2017.

[50] Y. Liu, Y. Li, R. Wang et al., "MiR-130a-3p regulates cell migration and invasion via inhibition of Smad4 in gemcitabine resistant hepatoma cells," *Journal of Experimental & Clinical Cancer Research: Climate Research*, vol. 35, no. 1, pp. 19–11, 2016.

[51] L. M. Holdt, A. Kohlmaier, and D. Teupser, "Molecular functions and specific roles of circRNAs in the cardiovascular system," *Non-Coding RNA Research*, vol. 3, no. 2, pp. 75–98, 2018.

[52] W. Wang, Y. Wang, H. Piao et al., "Circular RNAs as potential biomarkers and therapeutics for cardiovascular disease," *PeerJ*, vol. 7, Article ID e6831, 2019.

[53] Y. Bei, T. Yang, L. Wang et al., "Circular RNAs as potential theranostics in the cardiovascular system," *Molecular Therapy-Nucleic Acids*, vol. 13, pp. 407–418, 2018.

[54] K. Lei, H. Bai, Z. Wei et al., "The mechanism and function of circular RNAs in human diseases," *Experimental Cell Research*, vol. 368, no. 2, pp. 147–158, 2018.

[55] R.-Y. Pan, P. Liu, H.-T. Zhou et al., "Circular RNAs promote TRPM3 expression by inhibiting hsa-miR-130a-3p in coronary artery disease patients," *Oncotarget*, vol. 8, no. 36, pp. 60280–60290, 2017.

[56] S. Haque and L. Harries, "Circular RNAs (circRNAs) in health and disease," *Genes*, vol. 8, no. 12, Article ID 353, 2017.

[57] X. Lei, T. B. Mudiyanselage, Y. Zhang et al., "A comprehensive survey on computational methods of non-coding RNA and disease association prediction," *Briefings in Bioinformatics*, vol. 22, no. 4, Article ID bbaa350, 2020.

[58] W. Lan, D. Lai, Q. Chen et al., "LDICDL: LncRNA-disease association identification based on collaborative deep learning," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, pp. 1–9, 2020.

[59] W. Lan, Q. Chen, T. Li, C. Yuan, S. Mann, and B. Chen, "Identification of important positions within miRNAs by integrating sequential and structural features," *Current Protein & Peptide Science*, vol. 15, no. 6, pp. 591–597, 2014.

[60] Q. Chen, W. Lan, and J. Wang, "Mining featured patterns of MiRNA interaction based on sequence and structure similarity," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 10, no. 2, pp. 415–422, 2013.

[61] R. Zheng, M. Li, X. Chen et al., "An ensemble method to reconstruct gene regulatory networks based on multivariate adaptive regression splines," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 18, no. 1, pp. 347–354, 2021.

[62] J. Liu, D. Zeng, R. Guo, M. Lu, F. X. Wu, and J. Wang, "MMHGE: detecting mild cognitive impairment based on multi-atlas multi-view hybrid graph convolutional networks and ensemble learning," *Cluster Computing*, vol. 24, no. 3, pp. 103–113, 2020.

[63] J. Liu, M. Li, W. Lan, F. X. Wu, Y. Pan, and J. Wang, "Classification of Alzheimer's disease using whole brain hierarchical network," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 15, no. 2, pp. 624–632, 2016.