

MCOAN: multimodal contrastive representation learning for cross-omics adaptive disease regulatory network prediction

Junqi Long^{1,†, }, Bo Liu^{2,†}, Jianqiang Li¹, Shuangtao Zhao^{3,*}

¹School of Computer Science, Beijing University of Technology, Beijing 100124, China

²School of Mathematical and Computational Sciences, Massey University, Auckland 0745, New Zealand

³Department of Thoracic Surgery, Beijing Tuberculosis and Thoracic Tumor Research Institute/Beijing Chest Hospital, Capital Medical University, Beijing 101149, China

*Corresponding author. Department of Thoracic Surgery, Beijing Tuberculosis and Thoracic Tumor Research Institute/Beijing Chest Hospital, Capital Medical University, No. 9 Beiguan Street, Tongzhou District, Beijing 101149, China. E-mail: zst-1981@163.com

[†]These authors contributed equally to this work.

Associate Editor: Jonathan Wren

Abstract

Motivation: Interactions among long noncoding RNAs, circular RNAs, microRNAs, and messenger RNAs form complex gene expression regulatory networks, which are of great significance for the diagnosis, prevention, and treatment of complex diseases. Although existing computational methods have been developed to predict interactions among certain molecular types, they are generally limited to single-modality perspectives, overlooking competitive specificity and co-target cooperativity across multi-omics molecules, and thereby limiting their ability to elucidate cross-omics regulatory mechanisms.

Results: We proposed a novel cross-omics adaptive multimodal contrastive learning framework (MCOAN) that learns multimodal regulatory mechanisms and effectively predicts disease-associated molecular regulatory networks. Specifically, we first constructed a five-layer heterogeneous graph architecture to comprehensively integrate the complex regulatory associations among multi-omics nodes. Then, we proposed an unsupervised multimodal contrastive learning strategy that maximizes mutual information across distinct regulatory views, thereby enhancing node representations by efficiently capturing local neighborhood structure and global semantic information. Meanwhile, we also proposed a cross-omics adaptive learning mechanism that captures complex competitive specificity and co-target cooperativity across distinct regulatory networks, thereby further enhancing the structural awareness in node representations. Furthermore, we evaluated multiple downstream classifiers to accurately predict multimodal molecular regulatory networks. Finally, extensive experiments show that MCOAN consistently outperforms existing methods, achieving strong predictive accuracy and generalization (max AUC=0.9881; max AUPR=0.9826), and further confirm its real-world predictive performance through case studies.

Availability and implementation: All resources are available at <https://github.com/JunqiLab/MCOAN.git>.

1 Introduction

Long noncoding RNAs (lncRNAs), circular RNAs (circRNAs), microRNAs (miRNAs), and messenger RNAs (mRNAs) exhibit complex multilayered mechanisms regulating gene expression and play an indispensable role in the initiation and progression of diseases (Tay et al. 2014, Yan et al. 2025). Substantial experimental evidence indicates that interactions among these molecules exert both synergistic and competitive influences on disease regulation (Schmitz et al. 2014, Tay et al. 2014, Denzler et al. 2016). However, existing biological experimental

approaches exhibit limited predictive range and accuracy in elucidating multimolecular regulatory associations in disease, and their high costs preclude them from meeting the rapidly growing demand for predictions (Wang et al. 2022, Chai et al. 2025). It is, therefore, essential to develop efficient computational methods to complement multimolecular biological experiments.

With the rapid advancement in machine learning and deep learning in bioinformatics, numerous computational methods have been developed to predict intermolecular regulatory interactions. Early studies predominantly focused on machine learning approaches, which relied on handcrafted feature

representations derived from molecular functional annotations, expression profiles, and sequences, and fed these features into supervised learning models for association prediction. For example, prior studies have also employed principal component analysis to project lncRNA and disease similarity features into a low-dimensional space, followed by random forest to predict molecular association scores (Zhu et al. 2021). These approaches offer high computational efficiency and a degree of interpretability, with tree-based models in particular revealing putative key factors via feature-importance measures. However, such methods rely heavily on representative features, struggle to automatically capture complex nonlinear regulatory patterns, and often overlook much of the topological information encoded within molecular networks, thereby compromising generalization and robustness.

Recently, many studies have adopted matrix factorization methods to decompose high-dimensional data matrices into the product of several low-rank matrices, thereby yielding more informative feature representations. For example, previous studies (Xiao et al. 2018, Li et al. 2021, Ding et al. 2022, Gao et al. 2022, Qin et al. 2024) employed decomposition techniques such as singular value decomposition (SVD), non-negative matrix factorization, and Laplacian Eigenmaps to compress high-dimensional and sparse molecular interactions into consistent latent-factor representations, effectively reducing feature dimensionality and alleviating data sparsity. However, such methods are essentially bilinear approximation methods that primarily rely on feature similarity and first-order topological structure, often focusing on single-relation prediction and struggling to capture nonlinear, multi-hop semantic regulatory information and higher-order topological structure, while their generalization performance is susceptible to data noise (Fu et al. 2018, Xiao et al. 2018, Wu et al. 2020). Meanwhile, other studies have also directly leveraged network topology for modeling regulatory interactions. Methods such as random walk with restart and label propagation (Nguyen et al. 2021, Qu et al. 2021, Chen et al. 2022, Yu et al. 2022, Chen et al. 2024) learn reachability probabilities among molecular nodes via multistep propagation on the normalized adjacency matrix, thereby enabling effective prediction of molecular regulatory links in sparse networks. Nevertheless, their heavy reliance on linear propagation mechanisms renders them sensitive to hyperparameters and prior graph structure, thereby constraining generalization across multimodal heterogeneous edges and hindering the effective capture of higher-order topology and high-dimensional heterogeneous interactions.

Moreover, graph neural networks have progressively incorporated attention mechanisms and message passing for heterogeneous graphs, with training generally relying on large numbers of labeled samples and substantial computational resources to supervise representation learning across multiple molecular types to improve structural awareness and semantic discriminability in deep learning. For example, recent studies (Sheng et al. 2022, Sheng et al. 2023, Chai et al. 2025) employed a graph contrastive learning strategy to learn unified molecular representations and designed a multichannel attention mechanism to capture molecular representations across inter-graph, intra-graph, and complex-graph aggregations. Further work (Wu et al. 2021, He et al. 2024) learned molecular node representations using graph auto-encoders and incorporated discriminators to determine

their associative interactions. Nevertheless, these methods are prone to overfitting in weakly labeled multi-omics molecular regulatory networks. Their representation learning remains largely confined to local receptive fields and lacks effective alignment of global multimolecular regulatory semantics, thereby making it difficult to accurately capture the complex dynamic regulatory mechanisms across multimodal molecular regulatory networks.

To address the above limitations, the proposed MCOAN framework introduces innovations across three parts—structure, learning, and fusion. In the structural part, we comprehensively integrate interlayer interactions and intralayer multidimensional similarities among multi-omics molecules, constructing a five-layer heterogeneous graph comprising diseases, lncRNAs, circRNAs, miRNAs, and mRNAs, and uniformly mapping competitive specificity and co-target cooperativity across multi-omics molecules onto a learnable higher-order graph representation. In the learning part, we design multimodal unsupervised contrastive objectives to align representations between the local neighborhood structural view and the global semantic view, thereby effectively reducing dependence on strong supervision and improving the robustness of the learned representations. In the fusion part, we propose a cross-omics adaptive fusion strategy to achieve fine-grained dynamic modeling of multimodal molecular regulatory networks, yielding unified embeddings that are both discriminative and generalizable and providing a solid foundation for subsequent multi-omics prediction. In summary, the main contributions of MCOAN are as follows:

- 1) This study constructs a five-layer heterogeneous graph by comprehensively integrating inter-layer interactions and intra-layer multidimensional similarities among multi-omics molecules, enabling effective low-dimensional topological representations of multi-omics molecules.
- 2) MCOAN proposes a cross-omics adaptive learning mechanism to effectively learn competitive specificity and co-target cooperativity across multi-omics regulatory networks.
- 3) MCOAN proposes a multimodal contrastive representation learning strategy across multiple views to effectively capture global dependencies and local neighborhood structure in multi-omics molecular representations.
- 4) Extensive experiments demonstrate that the MCOAN framework outperforms existing methods, and case studies further validate its practical effectiveness for identifying multi-omics molecular regulatory networks.

2 Materials and methods

The overview of the proposed cross-omics adaptive disease regulatory network prediction framework based on multimodal contrastive learning was presented in Fig. 1. First, the framework comprehensively integrated multimodal inter-layer interactions and multidimensional intra-layer similarities among lncRNAs, circRNAs, miRNAs, mRNAs, and diseases to construct a five-layer heterogeneous disease regulatory network that encapsulates comprehensive information on intermolecular interactions (Fig. 1A). Next, the framework employed a multimodal self-supervised contrastive learning strategy with a cross-omics

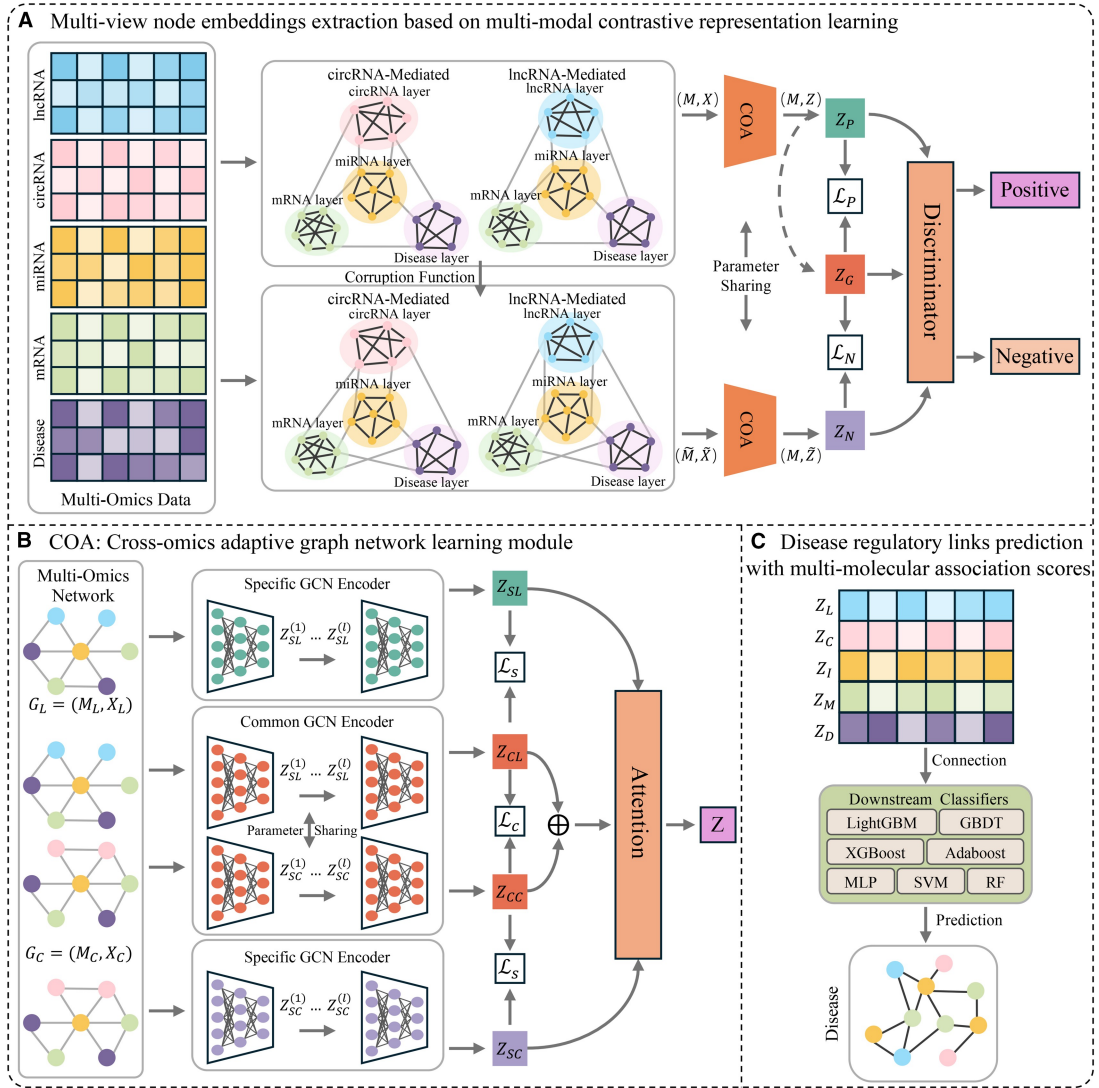


Figure 1 The overall framework of MCOAN. (A) Multi-view embedding feature extraction based on multimodal contrastive representation learning strategy. (B) Cross-omics adaptive regulatory learning mechanism. (C) Multi-molecule disease regulatory link prediction.

adaptive regulatory mechanism to effectively learn representative features of multi-omics molecular nodes (Fig. 1A and B). Finally, we constructed pairwise node embeddings and selected the optimal classifier to predict their molecular interaction probabilities within the multimolecular disease regulatory network (Fig. 1C).

2.1 Multi-modal molecular regulatory graph construction

To obtain more representative multi-omics regulatory interaction matrices, we constructed a five-layer multimodal regulatory heterogeneous graph (MORHG) based on inter-layer interactions and multidimensional intra-layer similarities among lncRNAs, circRNAs, miRNAs, mRNAs, and diseases. MORHG was defined as an undirected graph $G = (V, E)$, where $V = \{V_{\text{lncRNA}} \cup V_{\text{circRNA}} \cup V_{\text{miRNA}} \cup V_{\text{mRNA}} \cup V_{\text{disease}}\}$ is the set of all nodes, and $e_{ij} = (v_i, v_j) \in E$ denotes an edge between nodes v_i and v_j . We defined $M = (A, S)$ as the composite adjacency matrix

of G , where A is the multi-omics interactions matrix (1) and S is the multidimensional intra-layer similarity matrix (2).

$$A = \begin{cases} A^{LC} \in \mathbb{R}^{N_l \times N_c}, (v_i, v_j) \in V_{\text{lncRNA}} \times V_{\text{circRNA}} \\ A^{LI} \in \mathbb{R}^{N_l \times N_i}, (v_i, v_j) \in V_{\text{lncRNA}} \times V_{\text{miRNA}} \\ A^{LM} \in \mathbb{R}^{N_l \times N_m}, (v_i, v_j) \in V_{\text{lncRNA}} \times V_{\text{mRNA}} \\ A^{LD} \in \mathbb{R}^{N_l \times N_d}, (v_i, v_j) \in V_{\text{lncRNA}} \times V_{\text{disease}} \\ A^{CI} \in \mathbb{R}^{N_c \times N_i}, (v_i, v_j) \in V_{\text{circRNA}} \times V_{\text{miRNA}} \\ A^{CM} \in \mathbb{R}^{N_c \times N_m}, (v_i, v_j) \in V_{\text{circRNA}} \times V_{\text{mRNA}} \\ A^{CD} \in \mathbb{R}^{N_c \times N_d}, (v_i, v_j) \in V_{\text{circRNA}} \times V_{\text{disease}} \\ A^{IM} \in \mathbb{R}^{N_i \times N_m}, (v_i, v_j) \in V_{\text{miRNA}} \times V_{\text{mRNA}} \\ A^{ID} \in \mathbb{R}^{N_i \times N_d}, (v_i, v_j) \in V_{\text{miRNA}} \times V_{\text{disease}} \\ A^{MD} \in \mathbb{R}^{N_m \times N_d}, (v_i, v_j) \in V_{\text{mRNA}} \times V_{\text{disease}} \end{cases} \quad (1)$$

In the inter-layer interactions matrix A , N_l, N_c, N_i, N_m and N_d denote the numbers of lncRNAs, circRNAs, miRNAs, mRNAs, and diseases, respectively. For a node pair (v_i, v_j) , if an association e_{ij}

exists, the corresponding element was set to $A_{ij} = 1$. Otherwise, $A_{ij} = 0$.

$$S = \begin{cases} S^L \in \mathbb{R}^{N_l \times N_l}, (v_i, v_j) \in V_{\text{IncRNA}} \times V_{\text{IncRNA}} \\ S^C \in \mathbb{R}^{N_c \times N_c}, (v_i, v_j) \in V_{\text{circRNA}} \times V_{\text{circRNA}} \\ S^I \in \mathbb{R}^{N_i \times N_i}, (v_i, v_j) \in V_{\text{miRNA}} \times V_{\text{miRNA}} \\ S^M \in \mathbb{R}^{N_m \times N_m}, (v_i, v_j) \in V_{\text{mRNA}} \times V_{\text{mRNA}} \\ S^D \in \mathbb{R}^{N_d \times N_d}, (v_i, v_j) \in V_{\text{disease}} \times V_{\text{disease}} \end{cases} \quad (2)$$

In the intra-layer multidimensional similarity matrix S , considering that RNA molecules generally exhibit functional homology and regulatory similarity patterns during disease progression, we defined a unified RNA intra-layer similarity representation S^K ($K \in \{L, C, I, M\}$) following Equation (3) (Sheng et al. 2023), which integrates the functional similarity F and the Gaussian interaction profile kernel (GIPK) similarity GK for lncRNAs, circRNAs, miRNAs, and mRNAs. Meanwhile, we integrated the disease semantic similarity SD and GIPK similarity GK^D (4) to construct the intra-layer similarity of disease nodes. In these formulations, $GK_{R,ij}^K$ denote the GIPK similarity between nodes v_i and v_j of type K within modality R .

$$S_{ij}^K = \begin{cases} F_{ij}^K, & F_{ij}^K \neq 0 \\ \frac{GK_{L,ij}^K + GK_{C,ij}^K + GK_{I,ij}^K + GK_{M,ij}^K}{4}, & F_{ij}^K = 0 \end{cases} \quad (3)$$

$$S_{ij}^D = \frac{SD_{ij} + (GK_{L,ij}^D + GK_{C,ij}^D + GK_{I,ij}^D + GK_{M,ij}^D)/4}{2} \quad (4)$$

Accordingly, the MORHG adjacency matrix that comprehensively integrates inter-layer interactions and multidimensional intra-layer similarities was given by Equation (5), where $N_v = N_l + N_c + N_i + N_m + N_d$ and A^T denotes the transpose of A . In addition, we defined the normalized form of M as the attribute matrix of MORHG, denoted $X \in \mathbb{R}^{N_v \times N_v}$.

$$M = \begin{bmatrix} S^L & A^{LC} & A^{LI} & A^{LM} & A^{LD} \\ A^{LC^T} & S^C & A^{CI} & A^{CM} & A^{CD} \\ A^{LI^T} & A^{CI^T} & S^I & A^{IM} & A^{ID} \\ A^{LM^T} & A^{CM^T} & A^{IM^T} & S^M & A^{MD} \\ A^{LD^T} & A^{CD^T} & A^{ID^T} & A^{MD^T} & S^D \end{bmatrix} \in \mathbb{R}^{N_v \times N_v} \quad (5)$$

2.2 Cross-omics adaptive regulatory learning mechanism

First, to effectively align the two distinct biological regulatory networks, we construct two independent cross-omics regulatory subgraphs $G_L(M_L, X_L)$ and $G_C(M_C, X_C)$ on the basis of the unified MORHG representation by masking molecular node relations that fall outside their respective regulatory scopes. Specifically, for the lncRNA-mediated (or circRNA-mediated) graph G_L (or G_C), all matrix contributions associated with circRNA (or lncRNA) nodes are masked, thereby ensuring that every retained relation in M_L (or M_C) and X_L (or X_C) belongs to the lncRNA (or circRNA) regulatory domain, with the non-zero interactions in M_L (or M_C) defined as the edge set of G_L (or G_C), ensuring that the two subgraphs form cross-omics aligned regulatory structures.

Then, to effectively learn the biological mechanisms by which multi-omics molecules regulate disease, we designed a cross-omics adaptive graph learning module to dynamically capture the multidimensional specific and cooperative molecular representations in multi-omics disease regulatory networks. The module comprises three parts: specificity encoder, common encoder, and attention mechanism (Fig. 1B). These two types of encoders are both defined on the lncRNA-mediated regulatory networks G_L and the circRNA-mediated regulatory networks G_C , where the specificity encoder use independently parameterized graph convolutional networks to learn the distinct competitive regulatory patterns inherent to each network and produce the specificity embeddings Z_{SL} and Z_{SC} ; whereas the common encoder shares parameters across both networks to capture their cross-omics co-targeting regulatory pattern and produce the cooperative embeddings Z_{CL} and Z_{CC} . The multi-dimensional attention mechanism (6) computes a weighted fusion of Z_{SL} , Z_{SC} , and Z_C (the mean of Z_{CL} and Z_{CC}) to obtain the final representation Z (7), which are parameterized by attention weights α_{SL} , α_{SC} and α_C , a learnable projection matrix W that maps embeddings into the attention space, a bias vector b , and a shared attention query vector q .

$$\alpha_s = \frac{\exp(q^T \cdot \tanh(W \cdot Z_s + b))}{\sum_{s' \in S} \exp(q^T \cdot \tanh(W \cdot Z_{s'} + b))} \quad (6)$$

$$Z = \alpha_{SL} \cdot Z_{SL} + \alpha_{SC} \cdot Z_{SC} + \alpha_C \cdot Z_C \quad (7)$$

Next, to further ensure that the module effectively learns both competitive specificity and co-target cooperativity in multi-omics molecular regulation networks, we introduced a specificity regularization term $\mathcal{L}_S$ and a consistency regularization term $\mathcal{L}_C$. The specificity constraint $\mathcal{L}_S$ is based on the Hilbert-Schmidt Independence Criterion (HSIC), which measures the discrepancy between the lncRNA-specific and circRNA-specific embedding spaces. For embeddings Z_C and Z_S , HSIC is defined as:

$$\text{HSIC}(Z_C, Z_S) = (n-1)^{-2} \cdot \text{tr}(RK_C RK_S) \quad (8)$$

where K_C and K_S are Gram matrices with entries $K(i, j) = Z_i Z_j^T$, and $R = I - 1/n\mathbf{e}\mathbf{e}^T$ is the centering matrix with identity matrix I and $\mathbf{e}$ is the all-one column vector. Based on this formulation, the specificity regularization is computed as:

$$\mathcal{L}_S = \text{HSIC}(Z_{CL}, Z_{SL}) + \text{HSIC}(Z_{CC}, Z_{SC}) \quad (9)$$

Finally, although the common encoder shares a unified parameter matrix, we introduce a consistency constraint $\mathcal{L}_C$ to align its cross-omics cooperative embeddings by comparing the L2-normalized embeddings from the two common-encoder branches:

$$\mathcal{L}_C = \|Z_{CL} \cdot Z_{CL}^T - Z_{CC} \cdot Z_{CC}^T\|_2^2 \quad (10)$$

2.3 Multi-modal contrastive representation learning strategy

To comprehensively learn local neighborhood structural features and global semantic information in multimodal molecular regulatory networks, we adopted an unsupervised graph

contrastive learning approach (Teng *et al.* 2025) together with binary cross-entropy (BCE) losses $\mathcal{L}_P$ and $\mathcal{L}_N$ (11 and 12) to capture salient multimodal molecular features (Fig. 1A). The learning objective employed a linear discriminator to maximize mutual information across distinct regulatory views (Fig. 2), thereby extracting more discriminative high-level node representations (13). In this formulation, both Z_N and Z_P are computed by the COA module directly from the unified MORHG representation, where Z_N denotes the negative sample features extracted from the corrupted multimodal data ($\tilde{M}, \tilde{X}$) generated by a random corruption function, Z_P denotes the local view features of positive samples, α and β are the regularization coefficients of the biologically positive COA mechanism, and Z_G denotes the global semantic view computed following DGI (Veličković *et al.* 2019) by aggregating the Z_P representations into a graph-level global embedding (Fig. 1A).

$$\mathcal{L}_P = -\frac{1}{N} \sum_{i=1}^N \mathbb{E}_{(M,X)} [\log D(Z_{P,i}, Z_G)] \quad (11)$$

$$\mathcal{L}_N = -\frac{1}{N'} \sum_{i=1}^{N'} \mathbb{E}_{(\tilde{M}, \tilde{X})} [\log(1 - D(Z_{N,i}, Z_G))] \quad (12)$$

$$\mathcal{L} = \frac{N}{N+N'} \mathcal{L}_P + \frac{N'}{N+N'} \mathcal{L}_N + \alpha \mathcal{L}_S + \beta \mathcal{L}_C \quad (13)$$

3 Results

3.1 Comparison with other baseline methods

To systematically assess the advantages of the proposed model for multi-omics regulatory network prediction, we conducted multiple performance comparisons between MCOAN and all baseline models across three independent datasets (Fig. 3). The results show that MCOAN consistently achieves the best performance across all datasets, exhibiting significant superiority over existing methods in both discriminative ability (AUC) and robustness to multimodal data imbalance (AUPR). Specifically, MCOAN achieves strong predictive performance (AUC=0.9635 and

AUPR=0.9610) on Dataset 1. Similarly, MCOAN achieves the highest predictive performance (AUC=0.9881 and AUPR=0.9826) on the largest dataset (Dataset 2), not only outperforming typical matrix factorization methods (e.g. SVDNVLDA) but also significantly outperforming deep learning models (e.g. CERDA). Moreover, MCOAN maintains robust predictive performance (AUC=0.9830, AUPR=0.9767) on Dataset 3, achieving ~2%–4% performance gains over the second-best model, exhibiting strong discriminative and generalization capabilities. Finally, to verify the model's capability of learning from highly imbalanced multimodal data, we further evaluated performance on the three least frequent edge types in the more skewed Datasets 2 and 3. The results show that the AUC and AUPR remain at 0.9796 and 0.9805 in Dataset 2, respectively, while the AUC and AUPR reach 0.9351 and 0.9332 in Dataset 3, consistently exhibiting stable and significant advantages over all baseline methods.

3.2 Ablation study

3.2.1 The impact of multi-dimensional similarity matrices

To investigate the impact of different similarity matrices, we performed fivefold cross-validation to compare ablated variants that remove each similarity source. The results show that MCOAN achieves the best performance on both AUC and AUPR, significantly outperforming all variants (Fig. 4A). Specifically, removing the multi-omics GIPK similarity (W/O G) or the disease semantic similarity (W/O SD) from MCOAN leads to marked drops in both performance metrics. Moreover, removing the functional similarity (W/O F) from MCOAN results in the largest degradation, with the most pronounced decline observed for AUPR. Therefore, these results demonstrate that the multi-dimensional similarity matrices are indispensable for link prediction, with functional similarity contributing most to predictive performance.

3.2.2 The impact of feature learning strategies

To assess the effectiveness of the feature learning strategies, we conducted fivefold cross-validation to compare the ablated variants that each key strategy. The results show that MCOAN

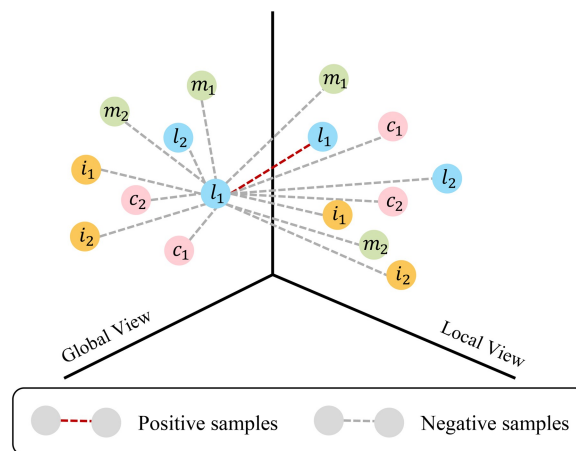


Figure 2 Multi-modal contrastive representation learning.

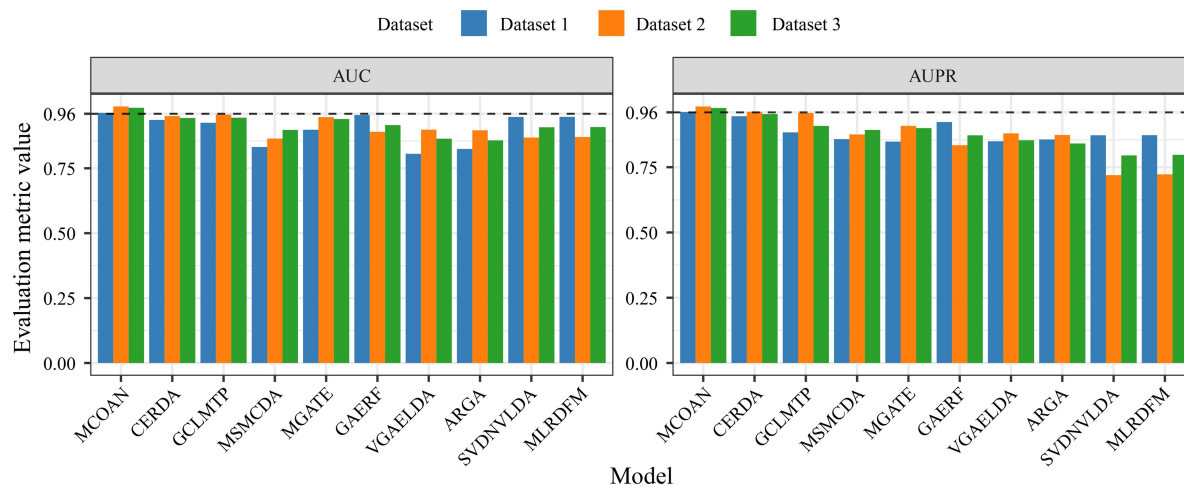


Figure 3 Performance comparison with other baseline methods.

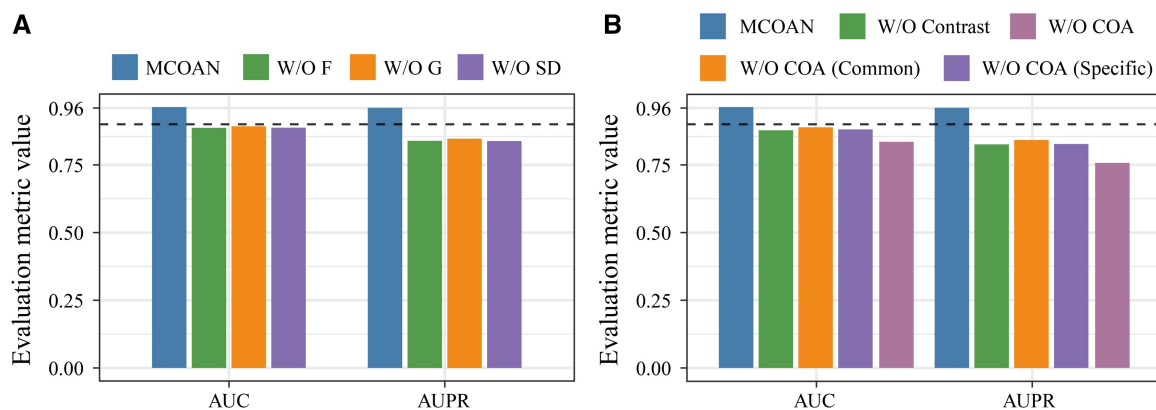


Figure 4 Multi-perspective ablation validation. (A) Ablation validation of the multi-dimensional similarity matrices. (B) Ablation validation of the feature learning strategies.

significantly outperforms all ablated variants (Fig. 4B). Specifically, removing the multimodal contrastive learning strategy leads to marked performance degradation, with AUC and AUPR decreasing by 8.56% and 13.52%, respectively. Similarly, ablating different regulatory mechanisms within the COA module also causes noticeable performance declines, especially when the specificity mechanism is removed. Furthermore, removing the COA module yields the greatest performance reduction, reducing AUC and AUPR by 12.83% and 20.39%, respectively. Overall, these results effectively highlight the critical contributions of different feature learning strategies to the representation learning on multimodal molecular regulatory data.

3.3 Data sensitivity analysis

To evaluate the robustness of MCOAN under varying training data sizes, we conducted a sensitivity analysis by gradually reducing the proportion of each training set from 80% to 20%. The results show that the model's predictive performance remains high overall across all three independent datasets, exhibiting only minor fluctuations (Fig. 5). Specifically, AUC decreases slightly from 0.9635 to 0.9564 and AUPR from 0.9610 to 0.9532

on Dataset 1. Moreover, AUC and AUPR remain ~ 0.9881 and 0.9826 on Dataset 2. Similarly, AUC and AUPR also remain around 0.9830 and 0.9767 on Dataset 3. These findings indicate that MCOAN is minimally sensitive to training data size, maintaining strong predictive performance even with only 20% of the training samples and thereby effectively mitigating the impact of data limitations in practical applications.

3.4 Time complexity analysis

To systematically evaluate the computational scalability of MCOAN, we conducted a time complexity analysis. MCOAN comprises three core parts: the COA mechanism for cross-omics adaptive feature learning, contrastive learning for multimodal representation extraction, and a downstream classifier for molecular regulatory interaction prediction (Fig. 1). The COA module has overall complexity $O(|E| \cdot V)$, primarily driven by the specific GCN encoder and the shared GCN encoder. The multimodal contrastive learning and prediction stages have complexities $O(V^2)$ and $O(|E| \cdot \log|E|)$, respectively. In summary, given the sparsity of real biological regulatory networks, the overall complexity can be expressed as $O(V^2)$.

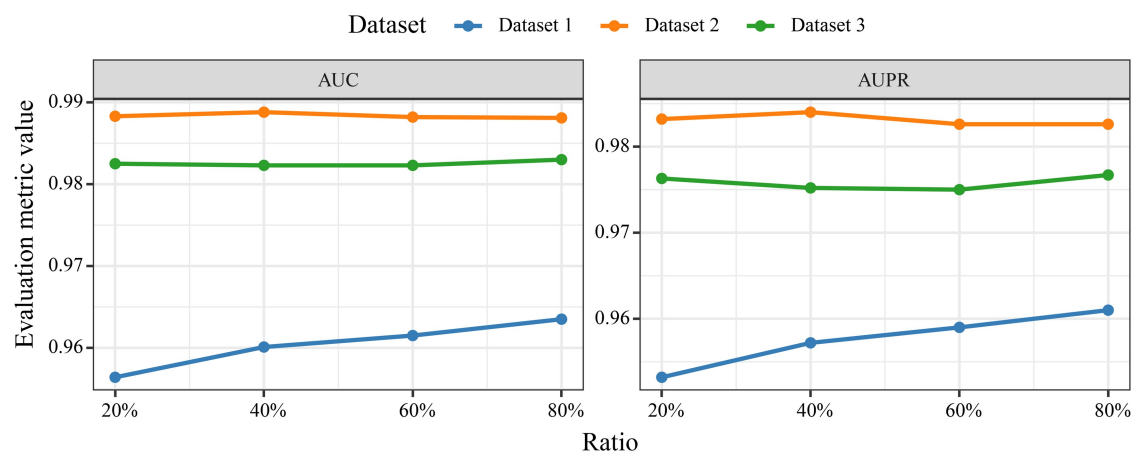


Figure 5 Data sensitivity analysis.

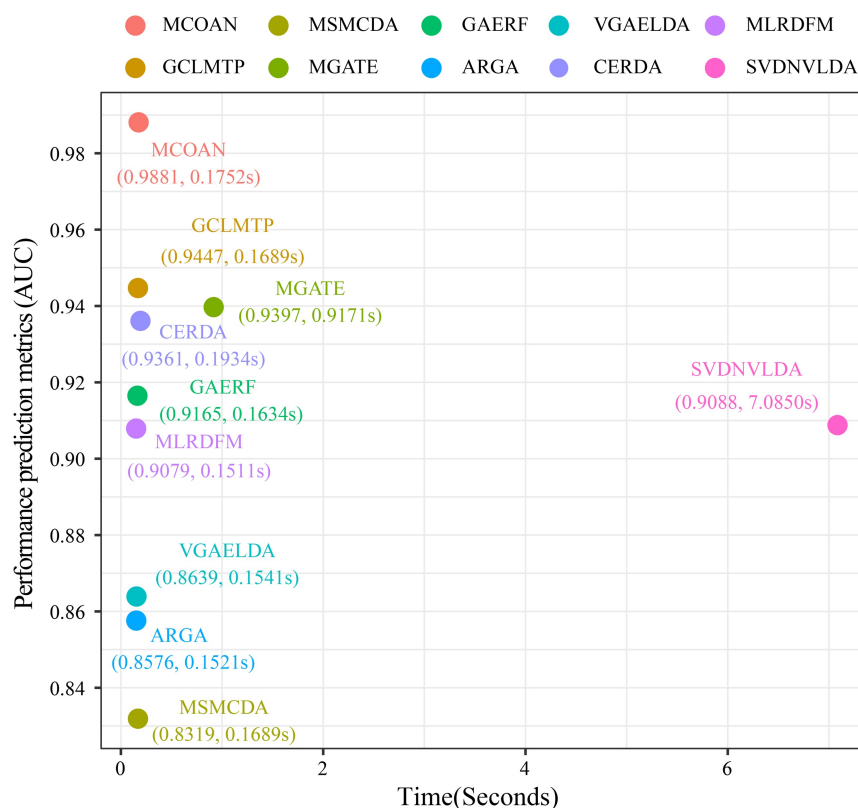


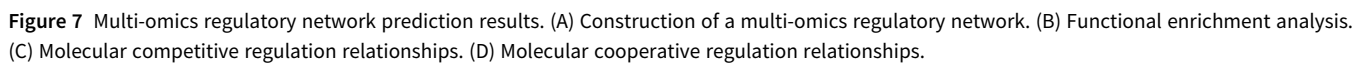
Figure 6 Computational complexity analysis.

To present the analysis results more intuitively, we further compared the training time and predictive performance of each baseline method on the largest dataset (Fig. 6). The results show that MCOAN achieves the highest predictive performance (AUC=0.9881) while maintaining a low training cost (0.1752s). In contrast, the traditional SVD-based methods require significantly more time for each matrix decomposition step. Moreover, although other baseline methods are close to MCOAN in efficiency, but they still exhibit significant gaps in predictive performance. Taken together, these results indicate that MCOAN strikes an effective balance between computational cost and

predictive performance, demonstrating strong scalability and practical value.

3.5 Case study

To further validate the practicality of MCOAN, we conducted a case study on lung cancer, which has high global incidence and mortality (Siegel *et al.* 2025), to further explore potential multi-omics regulatory mechanisms. Specifically, we first used MCOAN to predict the top 10 cancer-related candidate genes within the multi-omics regulatory network (Table 5, available as supplementary data at



Additionally, the network can robustly capture competitive specificity and co-target cooperativity, as evidenced by the visualized correlations and dependencies among top-ranked molecular features across multi-omics (Fig. 7C and D). The results showed significant correlations within each molecular layer and between the two mediator layers (lncRNA and circRNA) in the competitive regulatory heatmap, indicating that the model can effectively capture competitive relationships both across upstream regulatory

4 Discussion

This study proposed a cross-omics adaptive framework for disease regulatory network prediction based on multimodal

contrastive representation learning (MCOAN). The framework combined cross-omics adaptive learning and multimodal contrastive representation learning strategies to dynamically learn competitive specificity and co-target cooperativity across multimodal regulatory networks, enabling systematic modeling of complex multi-omics disease mechanisms. Extensive experiments showed that MCOAN outperformed multiple baselines while remaining computationally efficient. Furthermore, the sensitivity analysis indicated MCOAN's low dependence on training set size and robustness to data imbalance, reflecting strong discriminative power and generalization. In addition, MCOAN not only identified candidate genes with potential biological significance but also revealed signaling pathways and molecular interaction patterns closely associated with tumorigenesis and progression in the case study, thereby further validating the model's clinical reliability and translational potential.

In summary, MCOAN demonstrated excellent predictive performance and scalability for cross-omics regulatory prediction tasks. It not only maintained robust modeling and prediction advantages across diverse imbalanced data distributions but also provided new avenues to interrogate complex multimodal molecular regulatory mechanisms and discover potential therapeutic targets. Future research will further integrate complex pathological structures, gene expression profiles, and other multidimensional clinical modalities to better evaluate and optimize the model, enhancing its applicability in real-world medical settings.

Acknowledgements

We sincerely thank the reviewers for their constructive and insightful comments during the revision process, which significantly helped us improve both the methods and the manuscript.

Author contributions

Junqi Long (Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Validation, Visualization, Writing—original draft, Writing—review & editing), Bo Liu (Conceptualization, Funding acquisition, Project administration, Supervision), Jianqiang Li (Conceptualization, Project administration, Supervision), Shuangtao Zhao (Conceptualization, Project administration, Supervision)

Supplementary material

[Supplementary material](#) is available at *Bioinformatics* online.

Conflict of interests

None declared.

Funding

This study was partially supported by the National Natural Science Foundation of China (Grant No: 62076015), the Beijing

Natural Science Foundation—Haidian Original Innovation Joint Fund (Project No. L252088), and the REaDI Fund of Massey University.

Data availability

The data used in this study were obtained from multiple publicly available databases, including LncACTdb v3.0 (<http://bio-big-data.hrbmu.edu.cn/LncACTdb>), RNADisease v4.0 (<http://www.rnadisease.org>), NPInter v5.0 (<http://bigdata.ibp.ac.cn/npinter5>), LncRNA2Target v2.0 (<https://bio-computing.hrbmu.edu.cn/lncrna2target>), Circ2Disease (<http://bioinformatics.zju.edu.cn/Circ2Disease>), miRTarBase v10.0 (<https://mirtarbase.cuhk.edu.cn>), and CircBank (<https://www.circbank.cn>).

References

- Chai Z, Su Y, Tian X *et al*. Predicting disease associations based on the higher order structure of ceRNA networks. *Brief Bioinform* 2025;**26**:bbaf518.
- Chen M, Deng Y, Li A *et al*. Inferring latent disease-lncRNA associations by label-propagation Algorithm and random projection on a heterogeneous network. *Front Genet* 2022;**13**:798632.
- Chen M, Deng Y, Li Z *et al*. SCPLPA: an miRNA-disease association prediction model based on spatial consistency projection and label propagation algorithm. *J Cell Mol Med* 2024;**28**:e18345.
- Denzler R, McGeary SE, Title AC *et al*. Impact of microRNA levels, target-site complementarity, and cooperativity on competing endogenous RNA-regulated gene expression. *Mol Cell* 2016;**64**:565–79.
- Ding Y, Lei X, Liao B *et al*. MLRDFM: a multi-view laplacian regularized DeepFM model for predicting miRNA-disease associations. *Brief Bioinform* 2022;**23**:bbac079.
- Fu G, Wang J, Domeniconi C *et al*. Matrix factorization-based data fusion for the prediction of lncRNA–disease associations. *Bioinformatics* 2018;**34**:1529–37.
- Gao Z, Wang Y-T, Wu Q-W *et al*. A new method based on matrix completion and non-negative matrix factorization for predicting disease-associated miRNAs. *IEEE/ACM Trans Comput Biol Bioinform* 2022;**19**:763–72.
- Gu X, Huang Q, Huang Y *et al*. Identification of dynamic network biomarker ITGB1 for erlotinib pre-resistance using single-cell differential covariance entropy. *Mol Ther Oncol* 2025;**33**:200993.
- Haake SM, Plosa EJ, Kropski JA *et al*. Ligand-independent integrin β 1 signaling supports lung adenocarcinoma development. *JCI Insight* 2022;**7**:e154098.
- He L, Zou Q, Dai Q *et al*. Adversarial regularized autoencoder graph neural network for microbe-disease associations prediction. *Brief Bioinform* 2024;**25**:bbae584.
- Li J, Li J, Kong M *et al*. SVDNLDA: predicting lncRNA-disease associations by singular value decomposition and node2vec. *BMC Bioinformatics* 2021;**22**:538.
- Liang L, Luo Y, Li D *et al*. LncRNA OIP5-AS1 modulates the biological behaviour of lung cancer cells by regulating the hsa-miR-29b-3p/ZIC5 axis. *J Cell Mol Med* 2025;**29**:e70596.
- Liu Q, Zhang J, Guo C *et al*. Proteogenomic characterization of small cell lung cancer identifies biological insights and subtype-specific therapeutic strategies. *Cell* 2024;**187**:184–203.e28.

- Nguyen VT, Le TTK, Than K *et al.* Predicting miRNA–disease associations using improved random walk with restart and integrating multiple similarities. *Sci Rep* 2021;**11**:21071.
- Qin C, Zhang J, Ma L. EMCMDA: predicting miRNA-disease associations via efficient matrix completion. *Sci Rep* 2024;**14**:12761.
- Qu J, Wang C-C, Cai S-B *et al.* Biased random walk with restart on multilayer heterogeneous networks for MiRNA-Disease association prediction. *Front Genet* 2021;**12**:720327.
- Rocca A, Crinò L, Braga L *et al.* Refining treatment strategies for non-small cell lung cancer lacking actionable mutations: insights from multi-omics studies. *Br J Cancer* 2025; **133**:1405–27.
- Schmitz U, Lai X, Winter F *et al.* Cooperative gene regulation by microRNA pairs and their identification using a computational workflow. *Nucleic Acids Res* 2014;**42**:7539–52.
- Sheng N, Huang L, Wang Y *et al.* Multi-channel graph attention autoencoders for disease-related lncRNAs prediction. *Brief Bioinform* 2022;**23**:bbab604.
- Sheng N, Wang Y, Huang L *et al.* Multi-task prediction-based graph contrastive learning for inferring the relationship among lncRNAs, miRNAs and diseases. *Brief Bioinform* 2023; **24**:bbad276.
- Siegel RL, Kratzer TB, Giaquinto AN *et al.* Cancer statistics, 2025. *CA Cancer J Clin* 2025;**75**:10–45.
- Song KJ, Choi S, Kim K *et al.* Proteogenomic analysis reveals non-small cell lung cancer subtypes predicting chromosome instability, and tumor microenvironment. *Nat Commun* 2024; **15**:10164.
- Tay Y, Rinn J, Pandolfi PP. The multilayered complexity of ceRNA crosstalk and competition. *Nature* 2014;**505**:344–52.
- Teng Z, Tian Z, Zhou M *et al.* Enhancing lncRNA-miRNA interaction prediction with multimodal contrastive representation learning. *Brief Bioinform* 2025;**26**:bbaf281.
- Veličković P, Fedus W, Hamilton WL *et al.* Deep graph infomax. In: *International Conference on Learning Representations*, 2019.
- Wang W, Lu Y, Qin G-M *et al.* lncRNA RP11-297P16.4 promotes the invasion and metastasis of non-small-cell lung carcinoma by targeting the miR-145-5p/MMP-2/9 axis. *Biomedicines* 2025a;**13**:617.
- Wang W, Zhang L, Sun J *et al.* Predicting the potential human lncRNA-miRNA interactions based on graph convolution network with conditional random field. *Brief Bioinform* 2022;**23**:bbac463.
- Wang X, Song X, Wei H *et al.* hsa_circ_0002872 functions in lung cancer prognostic and tumorigenesis through regulating ZBTB46 via sponging hsa-miR-29b-1-5p. *Clinics (Sao Paulo)* 2025b;**80**:100733.
- Wu Q-W, Xia J-F, Ni J-C *et al.* GAERF: predicting lncRNA-disease associations by graph auto-encoder and random Forest. *Brief Bioinform* 2021;**22**:bbaa391.
- Wu T-R, Yin M-M, Jiao C-N *et al.* MCCMF: collaborative matrix factorization based on matrix completion for predicting miRNA-disease associations. *BMC Bioinformatics* 2020;**21**:454.
- Xiao Q, Luo J, Liang C *et al.* A graph regularized non-negative matrix factorization method for identifying microRNA-disease associations. *Bioinformatics* 2018;**34**:239–48.
- Yan S, Fu P, Li H *et al.* Comprehensive analysis of circRNA, lncRNA, miRNA and mRNA expression profiles and their competing endogenous RNA networks in hepatitis B virus-related hepatocellular carcinoma. *Mol Biotechnol* 2025;**67**:329–41.
- Yu N, Liu ZP, Gao R. Predicting multiple types of MicroRNA-disease associations based on tensor factorization and label propagation. *Comput Biol Med* 2022;**146**:105558.
- Zhu R, Wang Y, Liu J-X *et al.* IPCARF: improving lncRNA-disease association prediction using incremental principal component analysis feature selection and a random Forest classifier. *BMC Bioinformatics* 2021;**22**:175.