

RESEARCH

Open Access



# Interpretable multi-instance heterogeneous graph network learning modelling CircRNA-drug sensitivity association prediction

Mengting Niu<sup>1,2,3</sup>, Chunyu Wang<sup>4</sup>, Yaojia Chen<sup>1,5,6</sup>, Quan Zou<sup>1,2\*</sup> and Ximei Luo<sup>1\*</sup>

## Abstract

**Background** Different expression levels of circular RNAs (circRNAs) affect the sensitivity of human cells to drugs, thus producing different responses to the therapeutic effects of drugs. Using traditional biomedical experiments to discover and confirm sensitivity relationships is not only time-consuming but also costly. Therefore, developing an effective method to accurately predict new associations between circRNAs and drug sensitivity is crucial and urgent. Therefore, we constructed a heterogeneous graph network MiGNN2CDS on the basis of multi-instance learning (MIL).

**Results** We first extracted similar features of circRNAs and drugs and the structural features of drugs to construct a heterogeneous network. To learn the deep embedding features of the heterogeneous network, we designed a heterogeneous graph convolutional network (GCN) architecture. By introducing instance learning, we subsequently designed a pseudo-metapath instance generator and a bidirectional translation embedding projector BiTrans to learn the metapath-level representation of circRNA-drug pairs. Finally, an interpretable multiscale attention network joint predictor was designed to achieve accurate prediction and interpretable analysis of circRNA–drug sensitivity associations.

**Conclusions** MiGNN2CDS achieves better prediction accuracy than many state-of-the-art models do. Case studies show that MiGNN2CDS can effectively predict unknown associations, and the model interpretability of MiGNN2CDS is verified by high-confidence meta-path analysis. The code and data are available at <https://github.com/nmt315320/MiGNN2CDS.git>.

**Keywords** CircRNA–drug sensitivity association, Multi-instance learning, Heterogeneous graph neural network, Interpretable analysis

\*Correspondence:

Quan Zou  
zouquan@nclab.net  
Ximei Luo  
luoximei@uestc.edu.cn

<sup>1</sup> Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu, China

<sup>2</sup> Macao Polytechnic University, Gomes Street, Macau Peninsula, Macau 999078, China

<sup>3</sup> School of Electronic and Communication Engineering, Shenzhen Polytechnic University, Shenzhen 518055, China

<sup>4</sup> Faculty of Computing, Harbin Institute of Technology, Harbin, Heilongjiang 150000, China

<sup>5</sup> Biomedicine Discovery Institute and Department of Biochemistry and Molecular Biology, Monash University, Melbourne, VIC 3800, Australia

<sup>6</sup> Wenzhou Medical University-Monash Biomedicine Discovery Institute Alliance in Clinical and Experimental Biomedicine, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou 325035, China



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

## Background

Circular RNAs (circRNAs) are a group of covalently closed-loop RNA molecules without 3' or 5' ends [1, 2]. An increasing number of drug sensitivity correlation analyses have shown that circRNA expression affects the sensitivity to multiple drugs [3–5]. Recent studies have shown that CircKDM4 C can regulate the miR-548p/PBLD axis to inhibit tumour progression and enhance chemosensitivity to doxorubicin [6]. The overexpression of circRNA-CREIT significantly increases the chemical sensitivity of triple-negative breast cancer (TNBC) cells to DOX [3]. circSMARCA5 is a tumour suppressor that can increase the sensitivity of cancer cells in non-small cell lung cancer (NSCLC) to cisplatin and gemcitabine [7]. Hua reported that circRNACEP128 is upregulated in glioma tissues and that the control of circRNACEP128 can control the proliferation of glioma cells and increase the cytotoxicity of temozolomide by regulating miR-145-5p [8]. In addition, while research on the regulation of glioma angiogenesis by circRNAs was conducted, Meng et al. reported that CircSCAF11 can increase cell resistance to temozolomide through positive regulation of CircSCAF11 [9]. In NSCLC cells, the overexpression of circ\_0001946 induces cisplatin sensitivity and apoptosis. In addition, forced expression of circ\_0002483 and circ\_0030998 significantly inhibited NSCLC and increased paclitaxel sensitivity [10]. The high expression of Circ-HER2 in TNBC can increase sensitivity to pertuzumab, which provides a new treatment direction for TNBC and has important clinical significance [11]. Exosomal circATG4B participates in reducing the sensitivity of CRC cells to chemotherapy, providing a theoretical basis for the treatment of CRCL-OHP drug resistance [12]. Reducing the expression of circPDHK1 in ccRCC cells can increase the sensitivity of cell lines to TKI drugs (such as sunitinib and pazopanib) [13]. These studies indicate that circRNAs can play important roles in regulating sensitivity and resistance to chemotherapy drugs and that the exogenous introduction of tumour suppression-related circRNAs may be a new strategy for cancer treatment.

Researchers can increase sensitivity to chemotherapy drugs by targeting specific circRNA molecules, which can promote the discovery of molecular biomarkers of drug response and help clinicians develop more effective treatment plans [14–16]. Identification of circRNA–drug sensitivity associations (CDSAs) is highly important for drug development. Biological experimental methods for identifying CDSA are usually small-scale, time-consuming and costly. Therefore, to improve efficiency and identification accuracy, current scientific research urgently needs to develop computational methods to predict CDSAs.

CDSA prediction is a new research hotspot. There are currently relatively few research results available. Bo Yang proposed MNGACDA, which uses a node attention graph to explore the embedding of circRNAs and drugs and uses an inner product decoder to predict associations [17]. Shanghui Lu used dual-type multirelation heterogeneous graphs based on a multimodal network to predict associations [18]. Guanghai Li used a deep walk-aware graph attention network (GAN) to extract feature information fully to improve the prediction accuracy [19]. Guanghai Li also proposed MNCLCD, which uses random walk restart to effectively capture useful features and uses a hybrid neighbourhood graph convolutional network (GCN), contrastive learning, and dual Laplace-regularized least squares [20]. Yue Luo used dual-view learning and path mask graph autoencoders to predict associations [21]. Ziqiang introduced the characteristics of the disease, constructed a multimodal network to learn deep features, and used the random forest (RF) algorithm to predict associations [22]. Lei Deng used the graph attention autoencoder algorithm to extract the representation of circRNAs/drugs to improve accuracy [23]. In the above studies, GCN has shown good potential in accurately predicting the CDSA. However, these methods ignore the importance of meta-paths in learning representations on heterogeneous graphs of circRNAs and drugs. In addition, existing studies lack analyses of model interpretability. Multi-instance learning (MIL) can leverage bag-level information for learning, enabling it to perform well in few-shot scenarios while also providing a certain degree of interpretability [24, 25]. Moreover, MIL, a method for aggregating multiple meta-paths corresponding to a specific CDSA pair, has not yet been applied to the CDSA prediction problem.

In this study, we propose a novel heterogeneous graph network MiGNN2 CDS based on MIL for CDSA prediction. The MiGNN2 CDS architecture contains four modules: (1) First, the heterogeneous network is constructed. Similar features of circRNAs and drugs and structural features of drugs are extracted. (2) A heterogeneous graph node embedding extractor, which uses a heterogeneous GCN to learn deep features of circRNAs and drugs. (3) A metapath instance embedding projector, which constructs a new metapath instance generator and bidirectional translation embedding projector BiTrans to learn metapath representations of CDSA. (4) An interpretable association predictor of a multiscale attention network, which realizes accurate prediction and interpretable analysis of CDSA on the basis of a multiscale interpretable joint predictor of the attention mechanism. The experimental results demonstrate that MiGNN2 CDS has higher prediction accuracy than do many state-of-the-art methods. Case studies show that MiGNN2

CDS can effectively predict unknown associations. The experimental results demonstrate the robustness of the MiGNN2 CDS. The model interpretability of MiGNN2 CDS is verified by analysing high-confidence metapaths. The framework diagram of our model MiGNN2 CDS is shown in Fig. 1.

## Results

### Parameter analysis

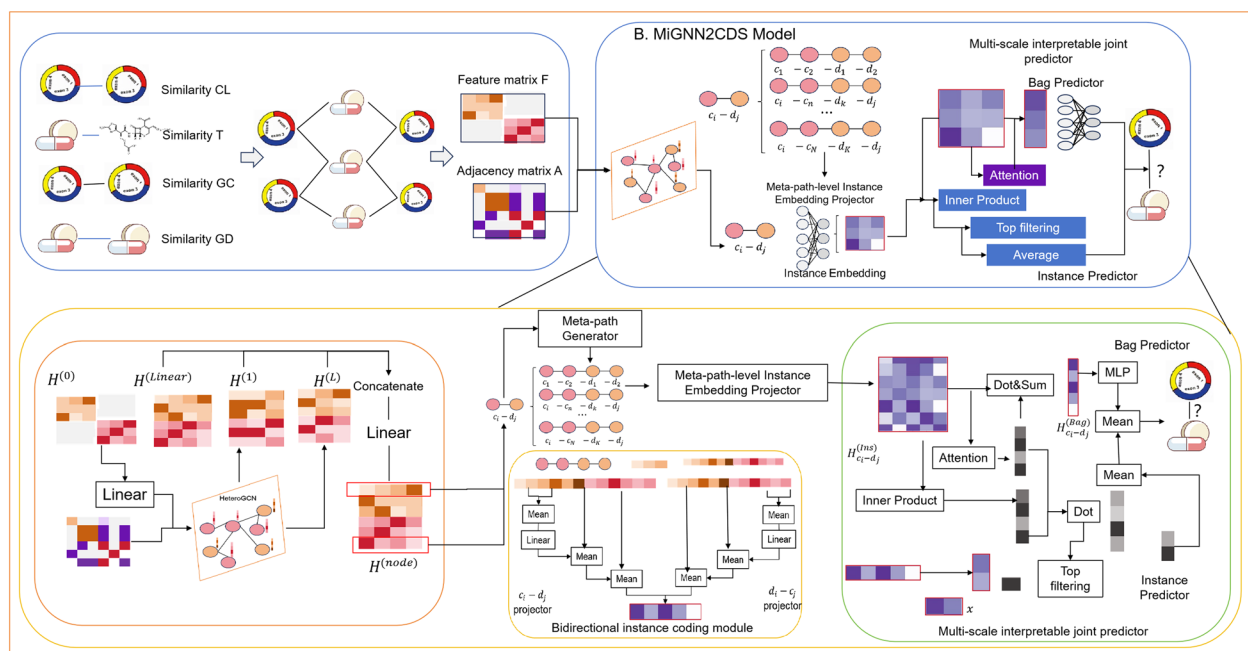
The learning rate (LR) is a hyperparameter that determines the step size for updating model parameters during training. A typical approach is to try various LRs and evaluate the performance of the MiGNN2 CDS at each stage to find the best LR. Through literature review and experience, five common LR strategies were selected. By conducting experimental comparisons, the optimal strategy was chosen based on performance. Therefore, we compared the effects of three common LR decay schemes (step-based scheme, linear, Adam) and two fixed LRs (0.002, 0.0005). Figure 2A shows the performance of the MiGNN2 CDS optimization process under different LRs.

As shown in Fig. 2A, the accuracy, specificity, and recall of the step-based scheme are significantly higher than those of the other four schemes, which are 0.9, 0.4391, and 0.559, respectively. The values of the F1-score, precision, AUC, and AUPR are not much different from those of the other strategies, but they are also improved, and the performance is relatively good.

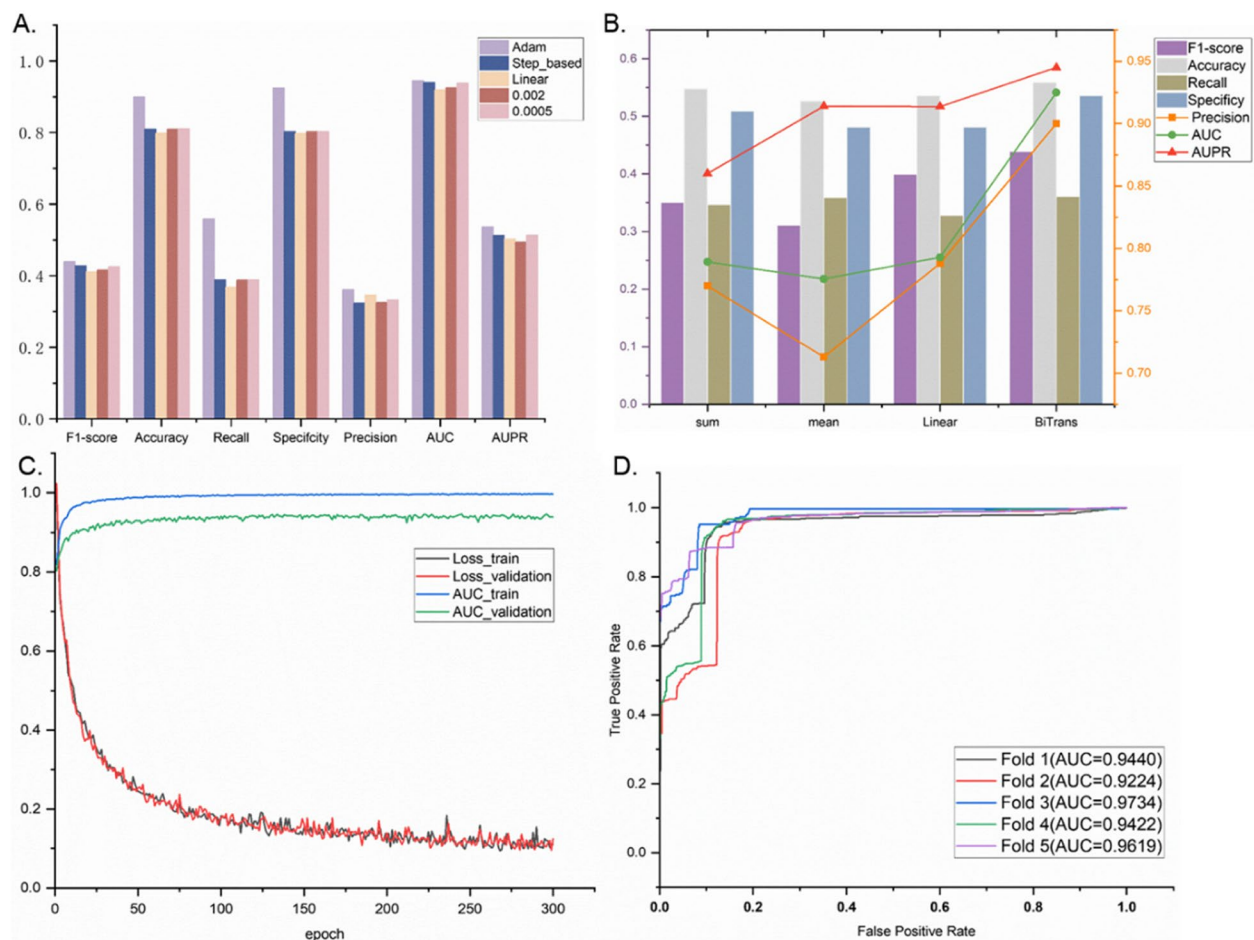
An experimental comparison shows that the step-based scheme is better than the other schemes. Therefore, we choose the step-based scheme as the LR strategy.

In addition, since MIL is a key module of MiGNN2 CDS, it plays a decisive role in instance/bag embedding and downstream prediction. Therefore, we compared the performance of the BiTrans method in MiGNN2 CDS and three simpler alternative methods, namely, sum, mean, and linear processing, and the results are shown in Fig. 2B. MiGNN2 CDS-sum replaces the BiTrans with the sum method; MiGNN2 CDS-mean replaces the BiTrans with the mean method; and MiGNN2 CDS-linear replaces the BiTrans with a linear neural network layer.

As shown in Fig. 2B, compared with MiGNN2 CDS-linear, the MiGNN2 CDS model with the BiTrans has a relative improvement of 4%, 11%, 2%, 13%, 3%, and 5% on average on 7 evaluation indicators; compared with MiGNN2 CDS-sum and MiGNN2 CDS-mean, which do not learn instance representations and only use simple operations to obtain instance representations, the MiGNN2 CDS model yields a relative average improvement of 8%, 13%, 1%, 13%, 1%, 8%, 2% and 12%, 18%, 3%, 14%, 0.2%, 3%, and 5% on 7 evaluation indicators. The results show that BiTrans performs better than the other methods do. We demonstrate through experimental analysis that our carefully designed embedding method for generating metapath instance embeddings can predict the CDSA more accurately.



**Fig. 1** The framework of MiGNN2 CDS



**Fig. 2** **A** Comparison of model performance under different LR schemes. **B** Comparison of model performance under different embedding-based methods. **C** The change curve of the loss and AUC of the MiGNN2 CDS model under different epochs. **D** The ROC curves of the MiGNN2 CDS model

### Model robustness analysis

To prove that our model MiGNN2 CDS does not have overfitting, this section statistically analyses the changes in the loss and area under the receiver operating characteristic (ROC) curve (AUC) of the training and validation sets during the training phase and analyses the impact of the number of epochs on model performance and model convergence. The results are shown in Fig. 2C. The ROC curve of fivefold cross-validation (FFCV) is shown in Fig. 2D.

Figure 2C shows that with increasing epochs, the AUC\_train and AUC\_validation of MiGNN2 CDS both show an upward trend, and the overall Loss\_train and Loss\_validation show a downward trend and gradually stabilize; the model gradually converges, and the training results are gradually optimized. This proves that the model has good robustness and that there is no overfitting phenomenon. In addition, we present the ROC curve and AUC values of the FFCV (Fig. 2D). The AUC values are 0.944, 0.9224, 0.9734, 0.9422, and 0.9619. The ROC

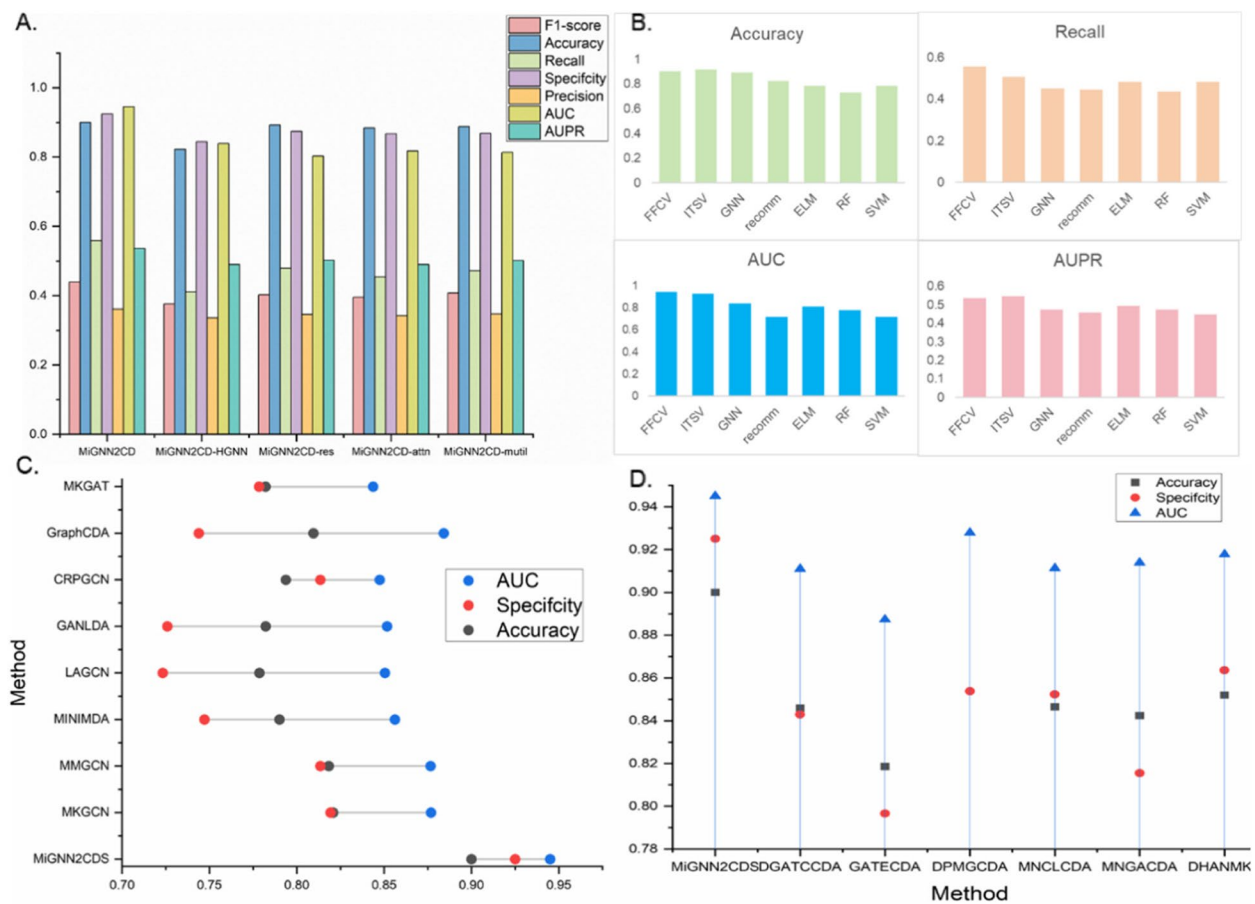
curve of the 5 results is infinitely close to 1 and gradually stabilizes, and the results of the 5 experiments are not much different from the results of the FFCV, which has strong robustness.

### Ablation study of the MiGNN2 CDS

To prove the necessity of the MiGNN2 CDS network architecture, we removed specific modules and designed 4 variants. The 4 variants are MiGNN2 CDS-HGCN, which removes the HeteroGCN layer; MiGNN2 CDS-Res, which removes the residual structure; MiGNN2 CDS-Attn, which replaces the attention-based instance aggregation in MiGNN2 CDS with a simple average package embedding process; and miGNN2 CDS-Multi, which is a MiGNN2 CDS model without an instance predictor. The FFCV results of the 5 network architectures are shown in Fig. 3A.

Fig. 3A shows that, compared with the 4 variant models, MiGNN2 CDS has the best performance in terms of the evaluation indicators. By comparison,





**Fig. 3** **A** Performances of MiGNN2 CDS and its variants. **B** Performance under different classification algorithms. **C** Performance under different classification models. **D** Performance comparison of different classification methods

the performance of the MiGNN2 CDS-HGNN model decreased the most. The performance of MiGNN2 CDS is better than that of any variant model. Compared with the performance of the MiGNN2 CDS-mutil model, it can be observed that simply stacking the representation learning modules (node embedding module, topological subnet embedding module, and graph attention module) leads to severe overfitting, thereby reducing the model's prediction accuracy for unknown CDSA. However, adding a layer attention module can effectively mitigate the degree of model overfitting. Combining it with other modules can further increase the upper limit of the model's representation ability, thereby improving prediction performance. The experimental results show that deleting any of the above specific modules reduces the accuracy of the MiGNN2 CDS. In short, these results prove the rationality of the MiGNN2 CDS architecture

#### Model performance under different algorithms

Traditional machine learning models, although potentially less flexible than deep learning (DL) models

in handling high-dimensional data and non-linear problems, often provide more stable and interpretable results in scenarios with smaller data volumes and lower feature dimensions. Therefore, to demonstrate the effectiveness of our proposed DL model, we compared our MiGNN2 CDS model to GNN, extreme learning machine (ELM), RF, support vector machine (SVM), and recommendation algorithm (recomm). And we compare the results of the FFCV and independent test set validation (ITSV) of the MiGNN2 CDS model (For the comparative algorithms, we directly used the default parameters). The experimental results are shown in Fig. 3B. First, the performance of FFCV and ITSV changed little. Then, the AUC of MiGNN2 CDS was significantly higher than that of RF and SVM. The results of GNN, ELM, and recomm have greatly improved, but the effect is still significantly worse than that of MiGNN2 CDS. These results prove that MiGNN2 CDS has good CDSA prediction ability.

Comparison of the effects of different DL framework

Because CDSA prediction and disease association prediction [26, 27] can be classified into the same problem, we compared the MiGNN2 CDS with 8 other models to verify whether the MiGNN2 CDS model can effectively use biological associations to improve model performance. These models (MKGCN [28], MMGCN [29], MINIMDA [30], LAGCN [31], GANLDA [32], CRPGCN [33], MKGAT) achieved good performance in disease association prediction. The MKGCN is a multimodal knowledge GCN that enhances association recommendations. MMGCN uses a multiview multichannel attention GCN to predict miRNA–disease associations. MINIMDA uses self-multimodal networks and multilayer perceptrons to predict potential associations between miRNAs and disease. LAGCN uses a layer attention GCN to predict drug–disease associations. GANLDA uses a GAN to predict the associations between lncRNA and disease. CRPGCN combines random walk and principal component analysis to construct a GCN to predict the associations between circRNAs and diseases. MKGAT uses a GAN and double Laplace regularized least squares to predict the associations between miRNAs and diseases. Here we present the three indicators of AUC, specificity, and accuracy. The results of other indicators are contained in Additional file 1 Table S1. The results of the FFCV indicators of the 9 methods on the dataset are shown in Fig. 3C.

By comparison, we can see that MiGNN2 CDS has achieved the best performance among all indicators. Compared with the average results of the other 8 models, MiGNN2 CDS yields relative improvements of 10.3, 15.4, and 8.4% in terms of accuracy, specificity, and AUC, respectively. The comparison proves that our proposed model can achieve higher model performance on the benchmark dataset than the previous prediction method can achieve.

Comparisons with state-of-the-art predictors

In this section, to demonstrate the performance of our model, MiGNN2 CDS is compared with the state-of-the-art computational frameworks DGATCCDA [20], GATECDA [23], DPMGCDA [21], MNCLCDA [19], MNGACDA [17], and DHANMKF [18], and the results are shown in Fig. 3D. Here we present the three indicators of AUC, specificity, and accuracy. The results of other indicators are contained in Additional file 1 Table S2.

Our model MiGNN2 CDS achieves good results for all three indicators. The MiGNN2 CDS has an accuracy of 0.9, a specificity of 0.925, and an AUC of 0.945, which are 5.6, 8.8, and 3.3% higher than those of DGATCCDA, GATECDA, DPMGCDA, MNCLCDA, MNGACDA, and DHANMKF. The DGATCCDA, GATECDA, MNCLCDA, MNGACDA, and DHANMKF algorithms use a variety of

DL algorithms to extract deep features of circRNAs and drugs, but the effect still needs to be improved, indicating that although it is important to extract deep features of circRNAs and drugs, it is equally important to build an effective prediction model. In short, the effectiveness of the MiGNN2 CDS model we constructed is proven by comparison with the current optimal algorithm.

Case study

After verifying the superiority of MiGNN2 CDS over previous methods in terms of model performance, this study further explored the application potential of MiGNN2 CDS. We selected two common drugs for the case analysis. First, we used data obtained from the GSDC database to train MiGNNCDS and obtain the prediction score matrix. Then, two drugs were randomly selected, the corresponding prediction scores were extracted, and a CTRP [34] was used to verify the predicted associations. We chose the top 20 candidate circRNAs and displayed them, and the results are shown in Tables 1 and 2.

The drug “methotrexate” (DrugBankID: DB00563) is a broad-spectrum antitumour drug of the folic acid class that is mainly used to treat a variety of cancers and autoimmune diseases [35]. As a test drug, the top 20 potential associated diseases predicted by the MiGNN2 CDS are shown in Table 1. As shown in Table 1, 14 of the top 20 candidates with the highest prediction scores according to MiGNN2 CDS were verified via the CTPR database.

The drug “crizotinib” (DrugBankID: DB11626560) is a tyrosine kinase inhibitor that can inhibit the activity of oncogene receptors such as anaplastic lymphoma kinase, ROS1, and hepatocyte growth factor receptor [36]. As a test drug, the top 20 potential associated circRNAs predicted by the MiGNN2 CDS are shown in Table 1. As shown in Table 1, the top 16 predictions made by MiGNN2 CDS were verified by CTPR.

**Table 1** The Top 20 circRNAs associated with the drug methotrexate

Ranking	circRNA	Evidence	Ranking	circRNA	Evidence
1	HEXIM1	CRPT	11	JUP	CRPT
2	VIM	CRPT	12	GRK2	Nonsignificant
3	COL6 A1	CRPT	13	HSP90B1	CRPT
4	SMAD1	Nonsignificant	14	RGS3	CRPT
5	OAZ1	Nonsignificant	15	POLR2 A	Nonsignificant
6	DNAJB2	CRPT	16	LTBP1	CRPT
7	CTSB	CRPT	17	KTN1	CRPT
8	MATN2	CRPT	18	SYMPK	Nonsignificant
9	RHOA	Nonsignificant	19	FAT1	CRPT
10	STAU1	CRPT	20	ENO2	CRPT

**Table 2** The Top 20 circRNAs associated with the drug crizotinib

Ranking	circRNA	Evidence	Ranking	circRNA	Evidence
1	MGAT4B	CRPT	11	PEA15	CRPT
2	CRIM1	CRPT	12	PNISR	CRPT
3	ASPH	CRPT	13	SPARC	CRPT
4	MUC16	CRPT	14	GSE1	CRPT
5	ANXA2	CRPT	15	KRT19	CRPT
6	CALD1	CRPT	16	PHF21 A	Nonsignificant
7	THBS1	CRPT	17	CTTN	CRPT
8	POLR2 A	Nonsignificant	18	ANP32B	Nonsignificant
9	MEF2D	Nonsignificant	19	PSAP	CRPT
10	PTMS	CRPT	20	PLAU	CRPT

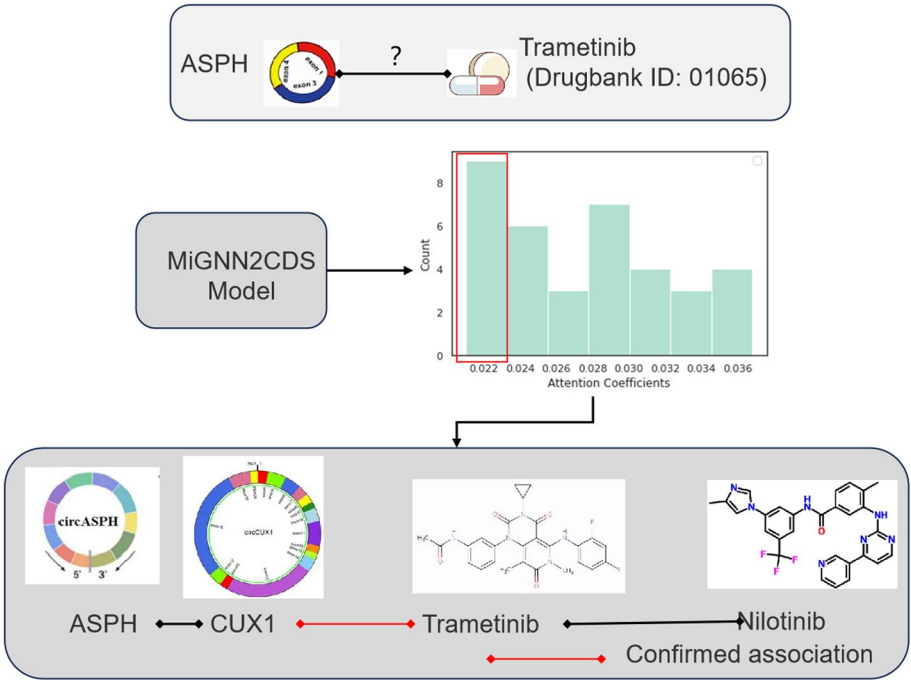
The above results show that the MiGNN2 CDS method can reveal drug-associated diseases that have not appeared in the dataset and has the potential to discover new indications. However, the drug repositioning model based on MiGNN2 CDS prediction requires manual secondary inspection of the prediction results.

**Interpretable analysis**

Using appropriate interpretable methods to explain the prediction results of the MiGNN2 CDS model can help enhance the reliability of drug sensitivity association results and promote the discovery of potential mechanisms of action of circRNA drugs in treating diseases. In

this study, MiGNN2 CDS uses an attention mechanism to identify the top 3 metapath instances in a given circRNA–drug pair. Since the metapaths in MiGNN2 CDS have biomedical importance, the metapath instances can be used to explain the model prediction results to discover the key groups and potential therapeutic phenotypes of circRNAs. We selected the circRNA–drug association pair “ASPH–Trametinib” with the highest prediction probability in the dataset as an example, output the metapath instance of the association pair and its attention coefficient, and used the metapath instance corresponding to the Top 3 attention coefficient as a case to explain the potential mechanism of action presented by the circRNA–drug sensitivity association pair. The overall workflow and results are shown in Fig. 4.

The results showed that MiGNN2 CDS first successfully predicted the association information of the CDSA; then, MiGNN2 CDS generated a meta-path instance of “circRNA⇌circRNA⇌drug⇌drug” the attention coefficient of the meta-path instance, which was “ASPH–CUX1–trametinib–nilotinib”. The chemical structure similarity of the drugs and the annotation of the circRNAs were checked, and the explanation of this meta-path instance was as follows. In the meta-path of “ASPH–CUX1–trametinib–nilotinib”, both nilotinib and trametinib are small-molecule tyrosine kinase inhibitors used for cancer treatment. In terms of molecular structure, both nilotinib and trametinib contain a pyrid pyrimidine structure, which is a common feature of tyrosine



**Fig. 4** A typical example of practical CDSA predictions and model interpretability of MiGNN2 CDS

kinase inhibitors. There is a strong association between the two, and there is also evidence to prove the association between Trametinib and CUX1 [37]. Therefore, the overall results suggest that this meta-pathway may explain the CDSA association between CUX1 and nilotinib, and all parts of the meta-pathway can be demonstrated by chemical structures or references.

## Discussion

The proposed MiGNN2 CDS framework demonstrates significant advancements in CDSA prediction by MIL with a heterogeneous GNN. Our model outperforms existing methods, highlighting its ability to effectively capture contextual information and long-term dependencies in circRNA-drug interactions. The pseudometa-path generator and bidirectional translation embedding method contribute to robust feature learning, while the multiscale interpretable joint predictor enhances model transparency—a crucial aspect for biomedical applications.

However, several challenges remain. First, the computational complexity of MiGNN2 CDS is higher than that of traditional methods, which may limit its scalability in large-scale applications. Future work could explore model compression or efficient training strategies to mitigate this issue. Second, the meta-path generation process, while effective, may not fully capture all biologically relevant regulatory pathways. Refining meta-path definitions using domain knowledge or dynamic learning mechanisms could improve reliability. Finally, dataset imbalance remains a critical issue in association prediction tasks. Incorporating weighted networks or adversarial learning (e.g., GANs) could help address bias and improve generalization.

Despite these limitations, MiGNN2 CDS represents a meaningful step toward interpretable and context-aware CDSA prediction, offering both methodological innovation and practical utility for precision medicine.

## Conclusions

In this study, we proposed MiGNN2 CDS, a novel heterogeneous graph network framework for circRNA-drug sensitivity association (CDSA) prediction. By integrating multi-instance learning (MIL) with a pseudometa-path generator, bidirectional translation embedding, and a multiscale interpretable predictor, our model achieves state-of-the-art performance while providing biologically meaningful explanations for predictions. Ablation studies confirm the contributions of key components, and case analyses demonstrate the model's real-world applicability in identifying potential drug-circRNA interactions. However, challenges such as computational complexity, meta-path optimization, and dataset imbalance require further investigation. Future work will focus on efficiency improvements, dynamic meta-path learning, and imbalance-aware training

strategies to enhance robustness. Overall, MiGNN2 CDS advances the field by offering the first interpretable CDSA prediction model with strong practical potential, paving the way for more reliable and explainable computational approaches in drug discovery and precision oncology.

## Methods

### Dataset

In this study, we used CDSA datasets provided in the literature [23] as our benchmark dataset. The dataset was downloaded from the GDSC database [38], which collects drug sensitivity data that are significantly related to circRNA expression, as identified via the Wilcoxon test. The GDSC database systematically characterized the circRNA expression profiles of 935 cancer cell lines in 22 cancer lineages from the Cancer Cell Line Encyclopedia and analysed circRNA biogenesis regulators, the impact of circRNAs on drug response, and the relationships between circRNAs, mRNAs, proteins, and mutations. We downloaded 80,076 associations involving 404 circRNAs and 250 drugs from the database. Then, we defined the associations with a false discovery rate less than 0.05 as significant associations, which are the CDSA data in the final dataset S (S is defined as Eq. (1)). Dataset S includes 4134 associations involving 271 circRNAs and 218 drugs.

$$s_{ij} = \begin{cases} 1, & \text{circRNAs are associated with drug sensitivity} \\ 0, & \text{circRNA is not associated with drug sensitivity} \end{cases} \quad (1)$$

where  $i$  and  $j$  are indices of circRNAs and drugs, respectively.

### CircRNA sequence similarity

In this section, we calculate the sequence similarity between circRNAs on the basis of the Levenshtein distance [39]. First, we downloaded the circRNA sequence data from NCBI. Then, for sequences A and B, the Levenshtein distance calculates the number of deletion, insertion, and substitution operations required to transform A into B; that is, the calculation formula is as follows:

$$CL(A, B) = 1 - \frac{t}{len(A) + len(B)} \quad (2)$$

where  $t$  is the minimum cost of conversion between A and B, and  $len(A)$  and  $len(B)$  are the sizes of A and B, respectively.

### Structural similarity of drugs

We know that the structure of a drug affects its function, and the fingerprint of a compound is necessary for considering the similarity of compounds via a computer. Therefore, the SMILES structural data of the drug were downloaded from PubChem [40]. Then, the data are



converted into RDKit molecular objects, and the topological fingerprint of each drug is calculated via the Tanimoto method to express the structural similarity between drugs. Ultimately, the drug structure similarity matrix  $T$  was obtained. A compound topological fingerprint is a method of encoding compound structure information into a binary vector. Therefore, for Compounds  $C$  and  $D$ , the similarity  $T$  calculation formula is shown in Eq. (3).

$$T(C, D) = \frac{|C \cap D|}{|C \cup D|} \quad (3)$$

where  $|C \cap D|$  is the number of elements in the intersection of  $C$  and  $D$  and where  $|C \cup D|$  represents the number of elements in the union of  $C$  and  $D$ .

#### Gaussian kernel function similarity features between CircRNAs and drugs

After the above steps, we obtain the sequence similarity matrix  $CL$  of the circRNA. However,  $CL$  has sparse features and cannot contain enough information. To obtain a more comprehensive disease similarity, we use Gaussian interaction profile kernel similarity (GIPKS) to supplement the features [41]. The calculation formula of the GIPKS matrix  $GC$  of a circRNA is shown in Eq. (4).

$$GC(c(i), c(j)) = \exp(-\theta_c ||V(c(i)) - V(c(j))||^2) \quad (4)$$

$$\theta_c = \frac{1}{n} \sum_{i=1}^n ||V(c(i))||^2 \quad (5)$$

$\theta_c$  and  $n$  have the same meanings as  $\theta_d$  and  $m$ .

The GIPKS matrix  $GD$  of drugs is calculated via Eq. (6).

$$GD(d(i), d(j)) = \exp(-\theta_d ||V(d(i)) - V(d(j))||^2) \quad (6)$$

Where  $\theta_d$  is the bandwidth of GIPKS.

$$\theta_d = \frac{1}{m} \sum_{i=1}^m ||V(d(i))||^2 \quad (7)$$

where  $m$  is the number of rows of  $M$ .

#### Characteristic fusion of CircRNAs and drugs

As mentioned above, we calculated the similarity features  $CL$ ,  $GC$ ,  $T$ , and  $GD$  of circRNAs and drugs. Here, we fuse the above features. The comprehensive similarity matrix fusion formula of a circRNA is as follows:

$$CSim(c(i), c(j)) = \begin{cases} CL(c(i), c(j)) + GC(d(i), d(j)) & \text{if } CL(c_i, c_j) \neq 0 \\ GD(d(i), d(j)) & \text{otherwise} \end{cases} \quad (8)$$

Similarly, the formula for calculating the comprehensive similarity matrix of drugs is Eq. (9).

$$DSim(d(i), d(j)) = \begin{cases} T(d(i), d(j)) + GD(d(i), d(j)) & \text{if } T(d(i), d(j)) \neq 0 \\ GD(d(i), d(j)) & \text{otherwise} \end{cases} \quad (9)$$

$$FV(c(i), d(j)) = [CSim(c(i)), DSim(d(j))] \quad (10)$$

#### Heterogeneous graph neural networks based on MIL

The CDSA prediction problem can be understood as predicting the probability of whether there is an association between the circRNA ( $c$ ) and drug ( $d$ ) nodes. Assume an association graph  $G = (V, E)$ , where  $V$  is a node set,  $E$  is an edge set.

#### The heterogeneous graph node embedding extraction module

The first step involved constructing a heterogeneous circRNA-drug network. When extracting heterogeneous graph node embedding features, we added a feature conversion layer to project multiple node features into the same dimension. The vector conversion process is shown in Eq. (11).

$$H^{(\text{Linear})} = \begin{bmatrix} W_c H_c^{(0)} + b_c \\ W_d H_d^{(0)} + b_d \end{bmatrix} \quad (11)$$

where  $H_c^{(0)}$  and  $H_d^{(0)}$  are the initial features of the circRNA and drug, respectively.  $W_c$ ,  $W_d$ ,  $b_c$ , and  $b_d$  is a learnable parameter matrix.

We then designed heterogeneous GNN to learn deep feature representations of circRNAs and drugs. After performing graph convolution, the heterogeneous graph features of the  $l$ th layer can be obtained. We then used an aggregator to learn the embeddings of neighbouring node, achieving graph learning and node embedding updates. For the circRNA/drug graph, the final embedded node in the heterogeneous network layer of the  $l$ th layer is Eq. (12).

$$H_c^{(l)} = \sum_{j \in N_i} H_j^{(l)} \quad (12)$$

In addition, to alleviate the oversmoothing problem of graph networks, we used a residual structure as the graph

node embedding extractor as the final node embedding  $H(\text{node})$ .

$$H^{(\text{node})} = \begin{bmatrix} W_r \text{Conact}(H_c^0, H_c^1, \dots, H_c^{(l)}) \\ W_d \text{Conact}(H_d^0, H_d^1, \dots, H_d^{(l)}) \end{bmatrix} \quad (13)$$

### Metapath-level instance embedding projector

For pseudo-metapath generation, to generate appropriate instances, we used the network topology to obtain pseudo-metapaths as  $c_i - c_i$  instances. First, we define a metapath of the form “circRNA  $\Leftrightarrow$  circRNA  $\Leftrightarrow$  drug  $\Leftrightarrow$  drug”. The intermediate circRNA node  $c_n$  and drug node  $d_m$  sample survival from the adjacent node sets  $N_{c_i}$  and  $N_{d_j}$ ; that is, the metapath form is expressed as “ $c_i - c_n - d_m - d_j$ ”. The process is expressed in Eq. (14).

$$I = [c_i c_n d_m d_j], \forall c_n \in N_{c_i} \text{ and } d_m \in N_{d_j} \quad (14)$$

Moreover, the node embedding vector learned by the node encoding module is used to represent the instance representation  $H_{c_i-d_j}^{(In)}$ .

$$H_{c_i-d_j}^{(In)} = \begin{bmatrix} H_{c_i}^{(Node)} & H_{c_n}^{(Node)} & H_{d_m}^{(Node)} & H_{d_j}^{(Node)} \\ H_{r_i}^{(Node)} & \vdots & \vdots & H_{d_j}^{(Node)} \\ H_{c_i}^{(Node)} & H_{c_n}^{(Node)} & H_{d_m}^{(Node)} & H_{d_j}^{(Node)} \end{bmatrix} \quad (15)$$

Then, we proposed a bidirectional translating embedding approach (BiTrans) to aggregate and update the representations of meta-path instances. For a given undirected meta-path “ $c_i \rightarrow c_n \rightarrow d_m \rightarrow d_j$ ”, we divided it into two directed meta-path “ $c_i \rightarrow c_n \rightarrow d_m \rightarrow d_j$ ” and “ $d_j \rightarrow d_m \rightarrow c_n \rightarrow c_i$ ”. And used a “ $c_i \rightarrow d_j$ ” and “ $d_j \rightarrow c_i$ ” to acquire aggregated embeddings with different message passing directions. The two directed instance representations  $H_{c_i \rightarrow c_n \rightarrow d_m \rightarrow d_j}$  and  $H_{d_j \rightarrow d_m \rightarrow c_n \rightarrow c_i}$  are obtained via Eq. (16) and Eq. (17).

$$H_{c_i \rightarrow c_n \rightarrow d_m \rightarrow d_j} = \text{Mean}(\text{Mean}(\text{Linear}(\text{Mean}(H_{c_i}, H_{c_n})), H_{d_k}), H_{d_j}) \quad (16)$$

$$H_{d_j \rightarrow d_m \rightarrow c_n \rightarrow c_i} = \text{Mean}(\text{Mean}(\text{Linear}(\text{Mean}(H_{d_j}, H_{d_m})), H_{c_n}), H_{c_i}) \quad (17)$$

Finally, we concatenate the representations of the two directed instances as the final metapath instance representation  $H_{c_i-d_j}^{(In)}$  (as Eq. (18)).

$$H_{c_i-d_j}^{(In)} = \begin{bmatrix} \text{Concat}(H_{c_i \rightarrow c_n \rightarrow d_m \rightarrow d_j}, H_{d_j \rightarrow d_m \rightarrow c_n \rightarrow c_i}) \\ \text{Concat}(H_{c_i \rightarrow c_n \rightarrow d_m \rightarrow d_j}, H_{d_j \rightarrow d_m \rightarrow c_n \rightarrow c_i}) \end{bmatrix} \quad (18)$$

### Multiscale interpretable joint predictor

After obtaining the instance embedding of  $c_i - d_j$ , we design a multiscale interpretable predictor that uses an attention mechanism to obtain a “package”-level representation and attention coefficient  $a_{c_i-d_j} H_{c_i-d_j}^{(Bag)}$  +

$$a_{c_i-d_j} H_{c_i-d_j}^{(Bag)} = \text{Attention}(H_{c_i-d_j}^{(In)}, W^{Attn}, q^{Attn}) \quad (19)$$

where  $W^{Attn}$  and  $q^{Attn}$  is the trainable parameter matrix.

The attention mechanism can aggregate metapath instance embeddings and dynamically adjust the importance weights. Through the importance of weights, we can find the metapath that is most important for model explanation. If there are k metapaths, the weight of the Mth metapath can be expressed as Eq. (20).

$$w_m^{(In)} = \frac{1}{M} \sum_{k \leq M} q^T \cdot \text{Linear}(H_k^{(In)}) \quad (20)$$

where  $q$  is the attention vector.

The final package that embeds  $H_{c_i-d_j}^{(Bag)}$  is Eq. (21).

$$H_{c_i-d_j}^{(Bag)} = \sum_{k \leq M} a_m^{(In)} H_m^{(In)} \quad (21)$$

Among them, in the package predictor, we use a multi-layer perceptron (MLP) [42] to output association probabilities  $a_{c_i-d_j}^{(Bag)}$ .

$$a_{c_i-d_j}^{(Bag)} = \text{MLP}(H_{c_i-d_j}^{(Bag)}) \quad (22)$$

In the instance predictor, we use the inner product with a linear layer to predict the association probability of all metapath instances; that is, the association probability matrix  $a_{c_i-d_j}^{(In)}$  is Eq. (23).

$$a_{c_i-d_j}^{(In)} = \begin{bmatrix} a_1 \\ \vdots \\ a_M \end{bmatrix} = \begin{bmatrix} \sigma(H_{c_i \rightarrow c_n \rightarrow d_m \rightarrow d_j} W H_{d_j \rightarrow d_m \rightarrow c_n \rightarrow c_i}^T) \\ \vdots \\ \sigma(H_{c_i \rightarrow c_n \rightarrow d_m \rightarrow d_j} W H_{d_j \rightarrow d_m \rightarrow c_n \rightarrow c_i}^T) \end{bmatrix} \quad (23)$$

Therefore, we use TOP K filtering [43] to force the instance predictor to focus on the reliable instances with the highest K attention coefficients. Therefore, the predicted association probability  $a_{c_i-d_j}^{(In)}$  is Eq. (24).

$$a_{c_i-d_j}^{(In)} = \text{Mean}(\text{Max}_{topk}([\begin{smallmatrix} a_1^{(In)} \cdot a_1 \\ \vdots \\ a_m^{(In)} \cdot a_m \end{smallmatrix}])) \quad (24)$$

For a given circRNA-drug pair  $c_i - d_j$ , the output of its association probability is Eq. (25).

$$a_{c_i-d_j} = \text{Mean}(a_{c_i-d_j}^{(Bag)}, a_{c_i-d_j}^{(In)}) \quad (25)$$

Before averaging, we use a feature conversion layer to project the latent space of the head node type to the latent space of the tail node type. Therefore, the calculation of projector  $p_{c-d}$  and projector  $p_{d-c}$  can be expressed as Eqs. (26) and (27).

$$\begin{aligned} \text{Projector}_{c-d} & \left( H_{c_i \rightarrow c_n \rightarrow d_m \rightarrow d_j}^{(In)} \right) \\ &= \text{Mean}(\text{Mean}(\text{Linear}(\text{Mean}(H_{c_i}, H_{c_n})), H_{d_k}), H_{d_j}) \end{aligned} \quad (26)$$

$$\begin{aligned} \text{Projector}_{c-d} & \left( H_{c_i \rightarrow c_n \rightarrow d_m \rightarrow d_j}^{(In)} \right) \\ &= \text{Mean}(\text{Mean}(\text{Linear}(\text{Mean}(H_{d_i}, H_{d_m}))), H_{c_n}), H_{c_j}) \end{aligned} \quad (27)$$

## Evaluation metrics

To fully evaluate the prediction performance of the MiGNN2 CDS, we conducted FFCV on the dataset S. We used seven common evaluation indicators: AUPR, AUC, accuracy, recall, precision, specificity, and F1-score. In addition, we plotted the ROC curve to intuitively show the performance of our model. The larger the AUC and AUPR values are, the better the prediction performance of the model [44].

## Abbreviations

circRNA	Circular RNA
TNBC	Triple-negative breast cancer
NSCLC	Non-small cell lung cancer
MIL	Multi-instance learning
DL	Deep learning
GAN	Graph attention network
GCN	Graph convolutional neural network
CDSA	CircRNA-drug sensitivity association
RF	Random forest
LR	Learning rate
ROC	Receiver operating characteristic
AUPR	Precision-recall curve
AUC	Area under the ROC curve
FFCV	Fivefold cross-validation
ELM	Extreme learning machine
SVM	Support vector machine
ITSV	Independent test set validation
GIPKS	Gaussian interaction profile kernel similarity

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12915-025-02223-w>.

Additional file 1: Table S1. Performance comparison of different DL frameworks. Table S2. Performance comparison with state-of-the-art predictors.

## Acknowledgements

We thank the participants for partaking in this study. The authors would like to thank three anonymous reviewers, whose constructive comments are very helpful for strengthening the presentation of this paper.

## Authors' contributions

Q.Z. and M.N. conceived and designed the experiment. M.N. and Y.C. performed the experiment. M.N. and C.W. analyzed the results. M.N. and C.W. wrote and revised the manuscript. M.N., Q.Z. and X.L. reviewed and edited the manuscript. Q.Z. and X.L. provided funding and resources and project administration. All authors provided feedback on the manuscript. All authors read and approved the final manuscript.

## Funding

The National Natural Science Foundation of China (62231013, 62473268, 62371347, 62302341); the Special Science Foundation of Quzhou (No. 2024D001). Shenzhen Science and Technology Program (Grant No. RCB20231211090800004).

## Data availability

All code and data generated or analyzed during this study are included in this published article, its additional file, and publicly available repositories. Which are available in the Zenodo repository (<https://zenodo.org/records/15104136>) and github repository (<https://github.com/nmt315320/MiGNN2CDS.git>).

## Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Competing interests

The authors declare no competing interests.

Received: 10 December 2024 Accepted: 25 April 2025

Published online: 14 May 2025

## References

- Chen L-L. The expanding regulatory mechanisms and cellular functions of circular RNAs. *Nat Rev Mol Cell Biol.* 2020;21(8):475–90.
- Fu P, Cai Z, Zhang Z, Meng X, Peng Y. An updated database of virus circular RNAs provides new insights into the biogenesis mechanism of the molecule. *Emerg Microbes Infect.* 2023;12(2): 2261558.
- Wang X, Chen T, Li C, Li W, Zhou X, Li Y, et al. CircRNA-CREIT inhibits stress granule assembly and overcomes doxorubicin resistance in TNBC by destabilizing PKR. *J Hematol Oncol.* 2022;15(1):122.
- Sun Y, Ma J, Lin J, Sun D, Song P, Shi L, et al. Circular RNA circ\_ASAP2 regulates drug sensitivity and functional behaviors of cisplatin-resistant gastric cancer cells by the miR-330-3p/NTSE axis. *Anticancer Drugs.* 2021;32(9):950–61.
- Cai Z, Lu C, He J, Liu L, Zou Y, Zhang Z, et al. Identification and characterization of circRNAs encoded by MERS-CoV, SARS-CoV-1 and SARS-CoV-2. *Brief Bioinform.* 2021;22(2):1297–308.
- Liang Y, Song X, Li Y, Su P, Han D, Ma T, et al. circKDM4C suppresses tumor progression and attenuates doxorubicin resistance by regulating miR-548p/PBLD axis in breast cancer. *Oncogene.* 2019;38(42):6850–66.

7. Tong S. Circular RNA. SMARCA5 may serve as a tumor suppressor in non-small cell lung cancer. *J Clin Lab Anal.* 2020;34(5):e23195.
8. Hua L, Huang L, Zhang X, Feng H, Shen B. Knockdown of circular RNA CEP128 suppresses proliferation and improves cytotoxic efficacy of temozolomide in glioma cells by regulating miR-145-5p. *NeuroReport.* 2019;30(18):1231–8.
9. Meng Q, Li S, Liu Y, Zhang S, Jin J, Zhang Y, et al. Circular RNA circSCAF11 accelerates the glioma tumorigenesis through the miR-421/SP1/VEGFA axis. *Molecular Therapy-Nucleic Acids.* 2019;17:669–77.
10. Yi Q, Feng J, Liao Y, Sun W. Circular RNAs in chemotherapy resistance of lung cancer and their potential therapeutic application. *IUBMB Life.* 2023;75(3):225–37.
11. Li J, Ma M, Yang X, Zhang M, Luo J, Zhou H, et al. Circular HER2 RNA positive triple negative breast cancer is sensitive to Pertuzumab. *Mol Cancer.* 2020;19:1–18.
12. Pan Z, Zheng J, Zhang J, Lin J, Lai J, Lyu Z, et al. A novel protein encoded by exosomal CircATG4B induces oxaliplatin resistance in colorectal cancer by promoting autophagy. *Advanced Science.* 2022;9(35): 2204513.
13. Huang B, Ren J, Ma Q, Yang F, Pan X, Zhang Y, et al. A novel peptide PDHK1-241aa encoded by circPDHK1 promotes ccRCC progression via interacting with PPP1CA to inhibit AKT dephosphorylation and activate the AKT-mTOR signaling pathway. *Mol Cancer.* 2024;23(1):34.
14. Xu T, Wang M, Jiang L, Ma L, Wan L, Chen Q, et al. CircRNAs in anticancer drug resistance: recent advances and future potential. *Mol Cancer.* 2020;19:1–20.
15. Mahabady MK, Mirzaei S, Saebfar H, Gholami MH, Zabolian A, Hushmandi K, et al. Noncoding RNAs and their therapeutics in paclitaxel chemotherapy: mechanisms of initiation, progression, and drug sensitivity. *J Cell Physiol.* 2022;237(5):2309–44.
16. Wang Y, Zhai Y, Ding Y, Zou Q. SBSM-Pro: support bio-sequence machine for proteins. *Sci China Inf Sci.* 2024;67(11): 212106.
17. Yang B, Chen H. Predicting circRNA-drug sensitivity associations by learning multimodal networks using graph auto-encoders and attention mechanism. *Briefings in Bioinformatics.* 2023;24(1):bbac596.
18. Lu S, Liang Y, Li L, Liao S, Zou Y, Yang C, et al. Inferring circRNA-drug sensitivity associations via dual hierarchical attention networks and multiple kernel fusion. *BMC Genomics.* 2023;24(1):796.
19. Li G, Zeng F, Luo J, Liang C, Xiao Q. MNCLCDA: predicting circRNA-drug sensitivity associations by using mixed neighbourhood information and contrastive learning. *BMC Med Inform Decis Mak.* 2023;23(1):291.
20. Li G, Li Y, Liang C, Luo J. DeepWalk-aware graph attention networks with CNN for circRNA–drug sensitivity association identification. *Brief Funct Genomics.* 2024;23(4):418–28.
21. Luo Y, Deng L. DPMGCDA: Deciphering circRNA–Drug Sensitivity Associations with Dual Perspective Learning and Path-Masked Graph Autoencoder. *J Chem Inform Model.* 2024;64(10):4359–72.
22. Liu Z, Dai Q, Yu X, Duan X, Wang C. Predicting circRNA-drug resistance associations based on a multimodal graph representation learning framework. *IEEE J Biomed Health Inform.* 2023;29(3):1838–1848.
23. Deng L, Liu Z, Qian Y, Zhang J. Predicting circRNA-drug sensitivity associations via graph attention auto-encoder. *BMC Bioinformatics.* 2022;23(1):160.
24. Loeffler CML, El Nahhas OS, Muti HS, Carrero ZI, Seibel T, van Treeck M, et al. Prediction of homologous recombination deficiency from routine histology with attention-based multiple instance learning in nine different tumor types. *BMC Biol.* 2024;22(1):225.
25. Moulin TC, Dey S, Dashi G, Li L, Sridhar V, Safa T, et al. A simple high-throughput method for automated detection of *Drosophila melanogaster* light-dependent behaviours. *BMC Biol.* 2022;20(1):283.
26. Niu M, Wang C, Zhang Z, Zou Q. A computational model of circRNA-associated diseases based on a graph neural network: prediction and case studies for follow-up experimental validation. *BMC Biol.* 2024;22:24.
27. Niu M, Zou Q, Wang C. GMNN2CD: identification of circRNA–disease associations based on variational inference and graph Markov neural networks. *Bioinformatics.* 2022;38(8):2246–53.
28. Cui X, Qu X, Li D, Yang Y, Li Y, Zhang X. Mkgcn: multi-modal knowledge graph convolutional network for music recommender systems. *Electronics.* 2023;12(12): 2688.
29. Ai N, Yuan H, Liang Y, Lu S, Ouyang D, Lai QH, et al. Multi-view multiattention graph learning with stack deep matrix factorization for circRNA-drug sensitivity association identification. *IEEE J Biomed Health Inform.* 2024;28:7670–82.
30. Lou Z, Cheng Z, Li H, Teng Z, Liu Y, Tian Z. Predicting miRNA–disease associations via learning multimodal networks and fusing mixed neighborhood information. *Briefings in Bioinformatics.* 2022;23(5):bbac159.
31. Yu Z, Huang F, Zhao X, Xiao W, Zhang W. Predicting drug–disease associations through layer attention graph convolutional network. *Briefings in bioinformatics.* 2021;22(4):bbac243.
32. Lan W, Wu X, Chen Q, Peng W, Wang J, Chen YP. GANLDA: graph attention network for lncRNA-disease associations prediction. *Neurocomputing.* 2022;469:384–93.
33. Ma Z, Kuang Z, Deng L. CRPGCN: predicting circRNA-disease associations using graph convolutional network based on heterogeneous network. *BMC Bioinformatics.* 2021;22:1–23.
34. Rees MG, Seashore-Ludlow B, Cheah JH, Adams DJ, Price EV, Gill S, et al. Correlating chemical sensitivity and basal gene expression reveals mechanism of action. *Nat Chem Biol.* 2016;12(2):109–16.
35. Di Martino V, Verhoeven DW, Verhoeven F, Aubin F, Avouac J, Vuitton L, et al. Busting the myth of methotrexate chronic hepatotoxicity. *Nat Rev Rheumatol.* 2023;19(2):96–110.
36. Musa S, Amara N, Selawi A, Wang J, Marchini C, Agbarya A, et al. Overcoming Chemoresistance in Cancer: The Promise of Crizotinib. *Cancers.* 2024;16(13): 2479.
37. Molho-Pessach V, Hartshtark S, Merims S, Lotem M, Caplan N, Alfassi H, et al. Giant congenital melanocytic naevus with a novel CUX1–BRAF fusion mutation treated with trametinib. *Br J Dermatol.* 2022;187(6):1052–4.
38. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res.* 2012;41(D1):D955–61.
39. Lu P, Wu J, Zhang W. Identifying circRNA-disease association based on relational graph attention network and hypergraph attention network. *Anal Biochem.* 2024;694: 115628.
40. Wang Y, Xiao J, Suzek TO, Zhang J, Wang J, Bryant SH. PubChem: a public information system for analyzing bioactivities of small molecules. *Nucleic Acids Res.* 2009;37(suppl\_2):W623–33.
41. Wu Q, Deng Z, Zhang W, Pan X, Choi K-S, Zuo Y, et al. MLNGCF: circRNA–disease associations prediction with multilayer attention neural graph-based collaborative filtering. *Bioinformatics.* 2023;39(8):btad499.
42. Liang J, Zhao X, Li M, Zhang Z, Wang W, Liu H, et al. MMMLP: Multi-modal multilayer perceptron for sequential recommendations. *Proceedings of the ACM Web Conference 2023*;2023:1109–1117. <https://doi.org/10.1145/3543507.3583378>.
43. Shi Q, Xu Y, Qi J, Li W, Yang T, Xu Y, et al. Cuckoo counter: Adaptive structure of counters for accurate frequency and top-k estimation. *IEEE/ACM Trans Networking.* 2023;31(4):1854–69.
44. Ao CY, Jiao SH, Wang YS, Yu L, Zou Q. Biological sequence classification: a review on data and general methods. *Research.* 2022;2022:0011.

# Publisher's Note

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