



Article

Circular RNA-Drug Association Prediction Based on Multi-Scale Convolutional Neural Networks and Adversarial Autoencoders

Yao Wang ¹, Xiujuan Lei ^{1,*} , Yuli Chen ¹ , Ling Guo ² and Fang-Xiang Wu ³

¹ School of Computer Science, Shaanxi Normal University, Xi'an 710119, China; wangyao_snnu.edu.cn (Y.W.); chen_yuli@snnu.edu.cn (Y.C.)

² College of Life Sciences, Shaanxi Normal University, Xi'an 710119, China

³ Division of Biomedical Engineering, Department of Mechanical Engineering and Department of Computer Science, University of Saskatchewan, 57 Campus Drive, Saskatoon, SK S7N 5A9, Canada

* Correspondence: xjlei@snnu.edu.cn

Abstract: The prediction of circular RNA (circRNA)-drug associations plays a crucial role in understanding disease mechanisms and identifying potential therapeutic targets. Traditional methods often struggle to cope with the complexity of heterogeneous networks and the high dimensionality of biological data. In this study, we propose a circRNA-drug association prediction method based on multi-scale convolutional neural networks (MSCNN) and adversarial autoencoders, named AAECDA. First, we construct a feature network by integrating circRNA sequence similarity, drug structure similarity, and known circRNA-drug associations. Then, unlike conventional convolutional neural networks, we employ MSCNN to extract hierarchical features from this integrated network. Subsequently, adversarial characteristics are introduced to further refine these features through an adversarial autoencoder, obtaining low-dimensional representations. Finally, the learned representations are fed into a deep neural network to predict novel circRNA-drug associations. Experiments show that AAECDA outperforms various baseline methods in predicting circRNA-drug associations. Additionally, case studies demonstrate that our model is applicable in practical related tasks.



Academic Editor: Sivanesan
Dakshanamurthy

Received: 10 January 2025

Revised: 6 February 2025

Accepted: 9 February 2025

Published: 11 February 2025

Citation: Wang, Y.; Lei, X.; Chen, Y.; Guo, L.; Wu, F.-X. Circular RNA-Drug Association Prediction Based on Multi-Scale Convolutional Neural Networks and Adversarial Autoencoders. *Int. J. Mol. Sci.* **2025**, *26*, 1509. <https://doi.org/10.3390/ijms26041509>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: circular RNA-drug association prediction; multi-scale convolutional neural network; adversarial autoencoder

1. Introduction

Circular RNAs (circRNAs) represent a class of non-coding RNA molecules characterized by a closed-loop structure, present across various organisms such as animals, plants, and microorganisms. These circRNAs are predominantly formed from precursor mRNA (pre-mRNA) through a process known as back-splicing, where the upstream splice donor connects with the downstream splice acceptor, resulting in a circular configuration. This distinctive circular formation makes circRNAs more resistant to degradation by conventional RNA exonucleases, thereby enhancing their stability within the cell [1].

CircRNAs have been demonstrated to be involved in numerous biological processes, with their functions encompassing miRNA sponging, transcription regulation, protein transport, and enhancement of protein-protein interactions [2]. As miRNA sponges, circRNAs can attach to miRNAs, preventing them from interacting with their intended gene targets, thereby influencing gene expression [3]. This capability is often described as the “miRNA sponge” effect. For instance, in bladder cancer, certain circRNAs have been observed to act as miRNA sponges. CircRNAs are also involved in regulating transcription.

Recent studies have revealed that circRNAs derived from the *insulin* gene interact with the RNA-binding protein TDP-43, playing a crucial role in the regulation of genes linked to insulin secretion [4]. CircRNAs also play a role in modulating protein function. For instance, *Circ-Amotl1* contributes to wound healing by interacting with *Dnmt3a* and *miR-17*, binding to *STAT3*, and *facilitating* its transport into the nucleus, where it upregulates the expression of fibronectin [5].

In the 1970s, circRNAs were first observed, but due to technological limitations and a lack of understanding of their function, they did not attract widespread attention. In 2012, Salzman et al. [6] first proposed that pre-mRNA could form circular RNA through back-splicing. These circRNAs were found in various *human* cell types. Jeck and colleagues identified over 25,000 specific types of RNA in human fibroblasts [7], thus uncovering the mechanism of circRNA formation. Since then, research on circRNAs has quickly gained significant attention.

Validating the relationships between circRNAs and drugs using conventional biomedical techniques can be both time-intensive and costly. Consequently, the demand for developing computational methods that can accurately and efficiently predict circRNA-drug associations is increasing. The Graph Attention Network (GAT) is a deep learning model designed for graph data, which merges the advantages of Graph Neural Networks (GNNs) with attention mechanisms. By incorporating attention mechanisms to handle graph-structured data, GAT enables the network to prioritize significant neighboring nodes, thereby enhancing the extraction of features from the graph. This model has been widely utilized in predicting circRNA-drug associations as well as in other association prediction tasks [8]. Deng et al. [9] proposed GATECDA, a computational framework utilizing a graph attention autoencoder, for predicting associations between circRNA and drug sensitivity. GATECDA employs a graph attention autoencoder (GATE) to produce compact representations of circRNAs and drugs, preserving critical information from sparse, high-dimensional features while seamlessly incorporating data from neighboring nodes. Additionally, Yang et al. [10] introduced an innovative method called MNGACDA to predict potential circRNA-drug sensitivity associations, aiming to facilitate deeper biomedical research. MNGACDA builds a multimodal network by combining multiple information sources related to circRNAs and drugs. It then utilizes node-level attention-based graph autoencoders to extract low-dimensional embeddings of circRNAs and drugs from this integrated network. In the final step, the model applies an inner-product decoder to compute association scores between circRNAs and drug sensitivity, using the generated embeddings as input. Similarly, Li et al. [11] developed DGATCCDA, a deep learning-driven computational method focused on detecting circRNA-drug sensitivity associations. DGATCCDA begins by constructing a multimodal network using the original feature data of circRNAs and drugs. Next, a DeepWalk-based graph attention network is employed to thoroughly capture feature information from the multimodal network, producing node embeddings. The multimodal network features are then merged via layer attention, followed by applying an inner-product technique to create an association matrix between circRNAs and drugs for prediction.

Moreover, other Graph Neural Network (GNN) approaches, such as Graph Convolutional Networks (GCNs) and autoencoders [12], have also proven effective in addressing the circRNA association prediction challenge [13–15]. Specifically, Lei et al. [16] proposed a method related to the denoising autoencoder and applied it to the prediction of microbe-disease associations. Moreover, the method proposed by Guo et al. [17], based on an Encoding–Decoding Framework Based on CNN for predicting circRNA–RBP associations has achieved remarkable results. Liu et al. [18] developed a computational framework named GraphCDD to predict associations between circRNAs and drug resistance. By em-

ploying multimodal GNNs, GraphCDD integrates various data types to generate effective representations of circRNAs, diseases, and drugs, thereby constructing a predictive model. Lu et al. [19] developed a new method, DHANMKF, to predict possible circRNA-drug sensitivity interactions, with the goal of advancing biomedical validation efforts. DHANMKF constructs a multimodal network by incorporating multiple information sources from circRNAs and drugs. The model then utilizes a dual-type, multi-relational heterogeneous graph to capture detailed intra-type and inter-type node features, which are subsequently enhanced through an attention-based encoder with a hierarchical design. The method integrates these intra-type and inter-type embeddings using a multi-core fusion technique. As the last step, the method employs Dual Laplacian Regularization Least Squares (DLapRLS) to predict potential associations between circRNAs and drug sensitivity within both the circRNA and drug spaces. Li et al. [20] proposed the MNCLCDA computational framework to identify potential circRNA-drug sensitivity associations, with the goal of contributing to medical research advancements. By utilizing drug structural information, circRNA gene sequences, and GIP kernel data, MNCLCDA computes the similarity between a given drug and circRNA. In order to mitigate noise within the similarity data, a preprocessing technique utilizing random walk with restart is implemented, allowing for the efficient extraction of key features from the similarity network. Finally, potential associations are predicted using kernel matrices in the respective feature spaces. The existing links between circRNAs and drug sensitivity are not fully understood, and numerous associations remain unidentified. Consequently, there is a pressing demand for more accurate computational approaches to predict trustworthy circRNA-drug sensitivity relationships.

In this study, we introduce a novel method named AAECDA for predicting circRNA-drug associations aimed at enhancing biomedical screening and validation processes. Our approach integrates circRNA-drug similarity with adversarial autoencoders. At the outset, we compute the combined similarity between circRNAs and drugs by integrating drug structural data with circRNA host gene sequence information, utilizing Gaussian kernel similarity for the calculation. Subsequently, adversarial autoencoders are utilized to learn low-dimensional latent representations from the feature network. Finally, a deep neural network is employed to predict potential circRNA-drug associations. Experimental results across four datasets indicate that AAECDA outperforms existing baseline methods. Additionally, case studies reveal that our model is practically applicable in real-world scenarios.

2. Results and Discussion

2.1. Experiment Settings and Evaluation Metrics

We applied five (ten) fold cross-validation to assess the model's effectiveness in predicting circRNA-drug associations. At the outset, the known circRNA-drug associations were considered as positive samples. To create a balanced dataset, a random selection of unobserved circRNA-drug associations was used as negative samples. Both the positive and negative samples were then split equally into five parts. During each round of validation, four parts were used as the training set, while the remaining part served as the test set. The model's overall performance was then evaluated by calculating the average performance across these five versions. Furthermore, we compared the model's performance with that of other baseline methods to comprehensively evaluate its strengths and weaknesses.

To assess the performance of the circRNA-drug association prediction model, we utilized seven key metrics: the area under the ROC curve (AUC), the area under the precision-recall curve (AUPR), *F1 score*, accuracy, specificity, precision, and recall. These metrics are mathematically defined in Equations (1)–(4). The AUC metric evaluates the model's capability to accurately distinguish between positive and negative samples. AUC

is particularly advantageous because it performs consistently well across various sample distributions, offering high robustness and stability. This helps to minimize performance variability caused by differences in test sets, thereby enabling more precise model evaluations. In binary classification tasks, precision and recall are commonly used to measure a model's effectiveness. The *F1 score*, which is the weighted average of precision and recall, provides a more holistic view of the model's overall performance. The ROC curve is constructed by plotting the true positive rate (TPR) against the false positive rate (FPR) across multiple thresholds, providing a comprehensive evaluation of the model's performance. Similarly, the precision-recall (PR) curve is plotted based on precision and recall values at different thresholds, illustrating the relationship between these two metrics in the prediction task. A higher AUC and AUPR indicate better model performance.

$$TPR = \frac{TP}{TP + FN}, FPR = \frac{FP}{TN + FP} \quad (1)$$

$$Precision = \frac{TP}{TP + FP}, Recall = \frac{TP}{TP + FN} \quad (2)$$

$$Specificity = \frac{TN}{TN + FP} \quad (3)$$

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

The ROC and PR curves for the 5-CV scenario are depicted in Figure 1. In this scenario, AAECDA achieves an average AUC of 0.9438 and an average AP of 0.9553. Table 1 presents additional performance metrics, where the average values for *F1 score*, accuracy, recall, specificity, and precision are 0.8305, 0.8480, 0.8534, 0.8785, and 0.9302, respectively. For the 10-CV scenario, as shown in Figure 2, AAECDA attains an average AUC of 0.9464 and an average AP of 0.9570. The corresponding metrics, listed in Table 2, show average values of 0.8410 for *F1 score*, 0.8407 for accuracy, 0.8437 for recall, 0.8881 for specificity, and 0.9383 for precision.

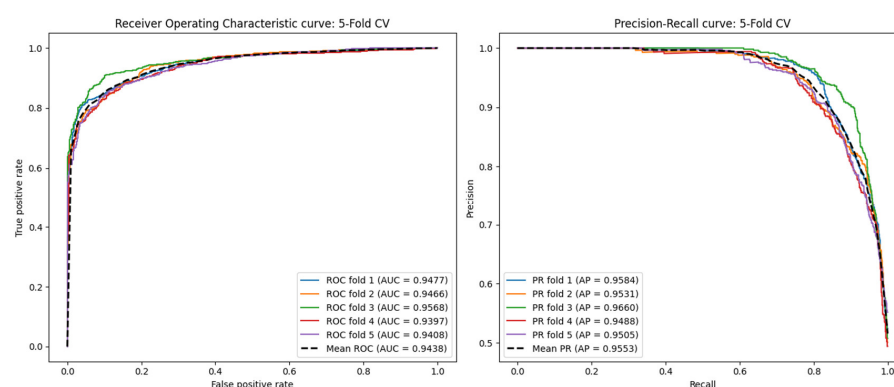


Figure 1. ROC curves and PR curves in fivefold cross-validation.

Table 1. Comparison with other methods based on 5-CV.

	GATECDA	MNGACDA	MKGCN	LAGCN	AAECDA
AUC	0.8846	0.9139	0.8664	0.8642	0.9438
AUPR	0.8928	0.9209	0.8662	0.8738	0.9553
F1_SCORE	0.8279	0.8489	0.8023	0.8084	0.8305
ACCURACY	0.8190	0.8310	0.7985	0.8019	0.8480
RECALL	0.8348	0.8498	0.8186	0.8265	0.8534
SPECIFICITY	0.8065	0.8343	0.7889	0.7539	0.8785
PRECISION	0.8273	0.8590	0.7857	0.7912	0.9302

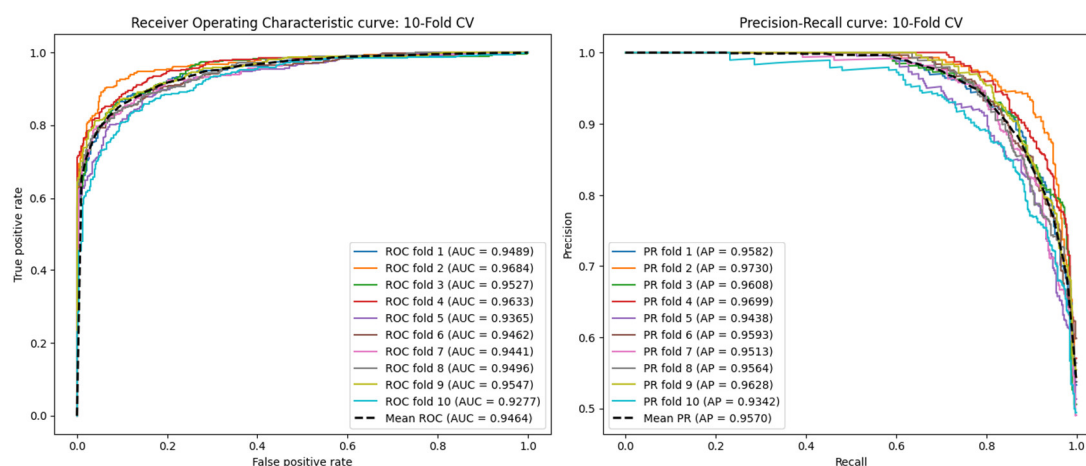


Figure 2. ROC curves and PR curves in tenfold cross-validation.

Table 2. Comparison with other methods based on 10-CV.

	GATECDA	MNGACDA	MKGCN	LAGCN	AAECDA
AUC	0.8918	0.9182	0.8683	0.8710	0.9464
AUPR	0.9015	0.9249	0.8769	0.8818	0.9570
F1_SCORE	0.8267	0.8373	0.8047	0.8133	0.8410
ACCURACY	0.8271	0.8427	0.8026	0.8076	0.8407
RECALL	0.8312	0.8536	0.8173	0.8312	0.8437
SPECIFICITY	0.8135	0.8323	0.7973	0.7486	0.8881
PRECISION	0.8225	0.8517	0.7937	0.7956	0.9383

2.2. Performance Comparison with Other Methods Under 5-CV and 10-CV Experiments

Based on available knowledge, existing computational approaches for predicting circRNA-drug interactions remain limited. To assess the predictive capability of AAECDA, we conducted a comparative analysis against five leading models in the field: GATECDA [9], MNGACDA [10], MNCLCDA [20], LAGCN [21], and MKGCN [22] under the same experimental settings. Additionally, the hyperparameters used in the experiments were set according to the recommendations in the authors' studies. Notably, LAGCN and MKGCN are established models commonly utilized in various bioinformatics prediction tasks, including drug-disease association prediction. Below is a brief overview of each model.

- **GATECDA:** This computational model employs a graph attention autoencoder and DNN to predict associations between circRNAs and drugs.
- **MNGACDA:** This model utilizes a node-level attention-based graph autoencoder to extract feature representations and employs an inner-product decoder to predict associations.
- **MNCLCDA:** This model applies a random walk with a restart method to preprocess the similarity network and capture features, followed by using a mixed-neighborhood graph convolutional network to acquire node neighborhood information.
- **MKGCN:** This method predicts microbe-drug associations by integrating multiple data sources and applying dual Laplacian regularized fewest squares on multiple kernel matrices.
- **LAGCN:** The model constructs a heterogeneous network, applies graph convolution to obtain the weights of each layer's embedding, and then predicts disease-drug associations.

In the 5-CV experiments, as shown in Table 1, the average AUC of AAECDA was 0.9438, which was 5.92% higher than GATECDA, 2.99% higher than MNGACDA, 7.74% higher than MKGCN, and 7.96% higher than LAGCN. The AUPR results are shown in Figure 2. The average AUPR score of AAECDA was 0.9553, which was 6.25% higher than GATECDA, 3.44% higher than MNGACDA, 8.91% higher than MKGCN, and 8.15% higher than LAGCN. Additionally, other performance metrics are shown in Table 1.

In the 10-CV experiments, as shown in Table 2, the average AUC of AAECDA was 0.9464, which was 5.46% higher than GATECDA, 2.82% higher than MNGACDA, 7.81% higher than MKGCN, and 7.54% higher than LAGCN. The AUPR results are shown in Figure 2. The average AUPR score of AAECDA was 0.9570, which was 5.55% higher than GATECDA, 3.21% higher than MNGACDA, 8.01% higher than MKGCN, and 7.52% higher than LAGCN. Additionally, other performance metrics are shown in Table 2.

Furthermore, as illustrated in Figure 3, the model demonstrates exceptional overall performance. In conclusion, the experiments outlined above confirm that AAECDA is a highly effective computational model for predicting circRNA-drug associations.

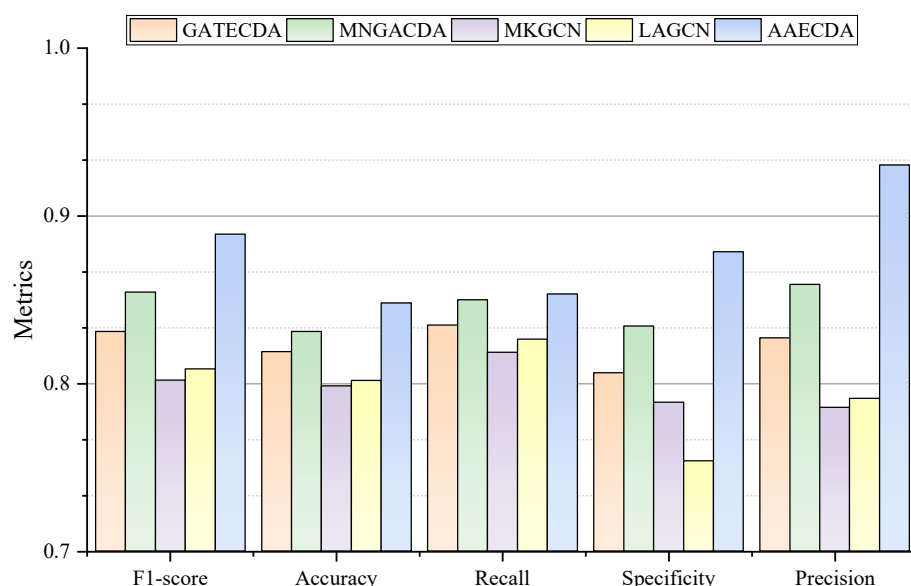


Figure 3. Performance comparison of AAECDA and baseline methods.

2.3. Parameter Sensitivity Analysis

The model's parameters can greatly affect its predictive performance. Therefore, we conducted experiments using 5-fold cross-validation (5-CV) to analyze the sensitivity of some key parameters, including (1) the training ratio $\frac{\lambda_1}{\lambda_2}$ of the encoder and discriminator, (2) the number of MSCNN convolutional layers c .

In the AAECDA, the training ratio between the encoder and discriminator is a critical parameter that determines the frequency at which the encoder and discriminator are trained during the adversarial process, directly affecting the model's balance and final performance. To investigate the impact of different training ratios on model performance, this study selected three training ratios for experimental analysis: 1:1, 1.2:1, and 0.8:1. A ratio of 1:1 means the encoder and discriminator are trained alternately, maintaining a balanced adversarial learning pace. A ratio of 1.2:1 means the encoder is trained 1.2 times before the discriminator is trained once, relatively increasing the training frequency of the encoder. A ratio of 0.8:1 means the encoder is trained 0.8 times before the discriminator is trained once, relatively increasing the training frequency of the discriminator. The experimental results, as shown in Figure 4, indicate that different training ratios have significant effects

on metrics such as AUC and AUPR. The specific performance under each training ratio is as follows:

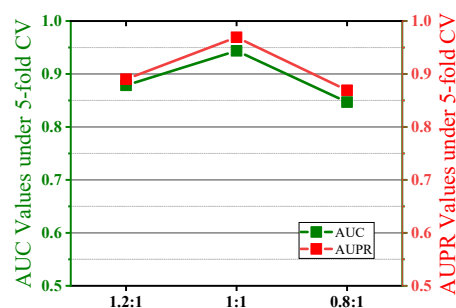


Figure 4. Effect of encoder and discriminator training ratio.

At a ratio of 1:1, the AUC reaches 0.9438, showing the best performance, indicating that the adversarial balance between the encoder and discriminator is ideal at this ratio. At a ratio of 1.2:1, the AUC is 0.8986, decreasing by approximately 4.8% compared to the baseline ratio, indicating a slight decline in performance. This suggests that when the encoder's training frequency increases, the discriminator's learning ability is weakened, disrupting the adversarial balance and affecting overall model performance. At a ratio of 0.8:1, the AUC is 0.8769, showing a further decrease of approximately 7.1% compared to the baseline ratio and a further decline compared to 1.2:1. This indicates that while the increased training frequency of the discriminator enhances its discriminative ability, it also suppresses the learning of the encoder, leading to a decrease in the quality of generated features and overall model performance. The experimental findings indicate that when the training ratio is set to 1:1, the training frequencies of the encoder and discriminator remain balanced, providing the most stable and superior adversarial balance during model training.

Next, the number of convolutional layers c in the multi-scale convolutional neural networks (MSCNN) affects the model's ability to capture features at different scales. In this study, we set the number of convolutional layers c to {2, 3, 4, 5} and observed its impact on model performance. As shown in Figure 5, when the number of convolutional layers $c = 3$, both the AUC and AUPR of the model reached their optimal values. When the number of layers is 2, the AUC value is 0.9145, which is approximately 3.1% lower than that of three layers. When the number of layers is 4, the AUC value is 0.9273, a decrease of about 1.8%, and when the number of layers is 5, the AUC value is 0.9021, a decrease of approximately 4.4%. This indicates that with a three-layer convolutional network structure, the model can balance feature diversity and computational complexity, achieving the best performance.

In addition, we used Xavier initialization for the model's parameters and the RMSprop optimizer during training. The learning rate was set to 0.0001, and weight decay was set to 0.0001.

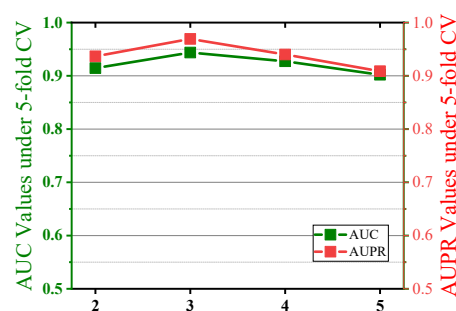


Figure 5. Effect of the number of convolution layers.

2.4. Ablation Experiments

To assess the contribution of each module in our proposed model to the final prediction accuracy and evaluate the impact of each module on the model's overall performance, we performed a series of ablation studies. The key components of the model worth investigating are the feature extraction for circRNAs and drugs, as well as the combined feature extraction of the two. The two main feature extraction parts are the MSCNN and the AAE.

To extract better features from circRNAs and drugs, one important part of the model is the MSCNN, which is designed to extract multi-scale features from the input data. To assess its contribution, we designed the AAECDA_{no}MSCN model, which retains the AAE module but omits the MSCNN module. This experiment helps to determine the role of the MSCNN in extracting useful features and enhancing the model's learning ability.

In the AAE, we designed two additional models: AAECDA_{no}encoder and AAECDA_{no}discriminator, to separately evaluate the impact of the encoder and discriminator modules on model performance.

1. AAECDA_{no}encoder removes the encoder, and the original input data are processed directly by the discriminator without encoding. This is equivalent to using the raw input data for adversarial training without compressed representations.
2. AAECDA_{no}discriminator degrades the AAE into a regular autoencoder, where the model structure contains only the encoder and decoder without adversarial training. This design evaluates the importance of the discriminator in guiding the AAE to learn the latent space distribution.

As shown in Figure 6, the AUC and AUPR values of AAECDA are significantly higher than those of AAECDA_{no}MSCN, AAECDA_{no}encoder, and AAECDA_{no}discriminator. Table 3 shows that its remaining performance metrics are also always superior to other ablation methods, indicating that each module in the AAECDA model plays an important role in enhancing predictive performance.

Specifically, the MSCNN provides the capability to capture multi-scale features during the feature extraction phase, which is crucial for the model's accuracy and generalization ability. The ablation experiment results demonstrate that when the MSCNN module is omitted (i.e., in the AAECDA_{no}MSCN model), the overall performance of the model significantly decreases, proving the critical role of the MSCNN in effectively extracting complex and multi-scale information from the input data.

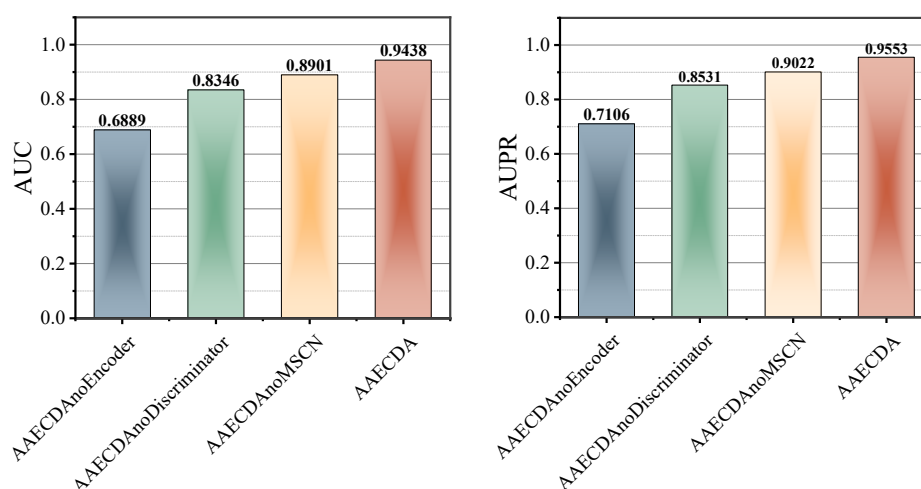


Figure 6. The results of AAECDA and its variants in the ablation study.

Table 3. Results of ablation experiments for AAECDA.

Variant	AUC	AUPR	F1_SCORE	ACCURACY	RECALL	SPECIFICITY	PRECISION
AAECDAnoEncoder	0.6889	0.7106	0.7205	0.7022	0.7864	0.6975	0.7764
AAECDAnoDiscriminator	0.8346	0.8531	0.8014	0.8012	0.8248	0.7967	0.8083
AAECDAnoMSCN	0.8901	0.9022	0.8095	0.8237	0.8533	0.7991	0.8238
AAECDA	0.9438	0.9553	0.8305	0.8480	0.8534	0.8785	0.9302

In the ablation study of the AAE component, both the AAECDA_no_encoder model, which removes the encoder, and the AAECDA_no_discriminator model, which removes the discriminator, exhibited performance degradation. The AAECDA_no_encoder model shows that without the encoder for data compression, the model's predictive ability is noticeably weakened, indicating the encoder's crucial role in reducing data redundancy and focusing on key features. Meanwhile, the AAECDA_no_discriminator model, which lacks the discriminator, experienced even more significant performance degradation, further illustrating the importance of the discriminator in adversarial training. By guiding the encoder to learn the latent space distribution, the discriminator enables the encoder to more effectively capture the structural information and differences in the data.

2.5. Performance Under the Blind Test Set

In previous studies, we employed fivefold and tenfold cross-validation to evaluate the performance of the model. Although this approach is effective for training and validating within the dataset, it cannot fully assess the model's predictive ability in real-world scenarios, where the associations between circRNAs and drugs are usually unknown. To address this issue, we introduced a blind test set to evaluate the model's performance on truly unseen data.

The blind test set was constructed by randomly sampling 20% of the total samples from the original dataset, ensuring that each sample had an equal probability of being selected. Importantly, the blind test set was completely isolated from the model training and hyperparameter optimization processes. The remaining 80% of the data was used for model selection and hyperparameter tuning through 5-fold cross-validation. After training, we evaluated the model's performance on the blind test set to assess its generalization ability to new circRNA-drug interactions. To ensure the robustness of our results, we repeated the blind test set evaluation five times with different random splits and reported the average performance metrics.

The ROC and PR curves for the blind test set are shown in Figure 7. Additionally, we compared our model with two state-of-the-art methods, GATECDA and MNGACDA, using the same blind test set division method. The experimental results, presented in Table 4, demonstrate that our model achieved an AUC of 0.8183 on the blind test set, outperforming both GATECDA (AUC = 0.7760) and MNGACDA (AUC = 0.7947). This indicates that our model has a strong generalization ability for predicting circRNA-drug interactions in real-world scenarios.

Table 4. Comparison with other methods on the blind test set based on 5-CV.

	GATECDA	MNGACDA	AAECDA
AUC	0.7760	0.7947	0.8183
AUPR	0.7821	0.8013	0.8342

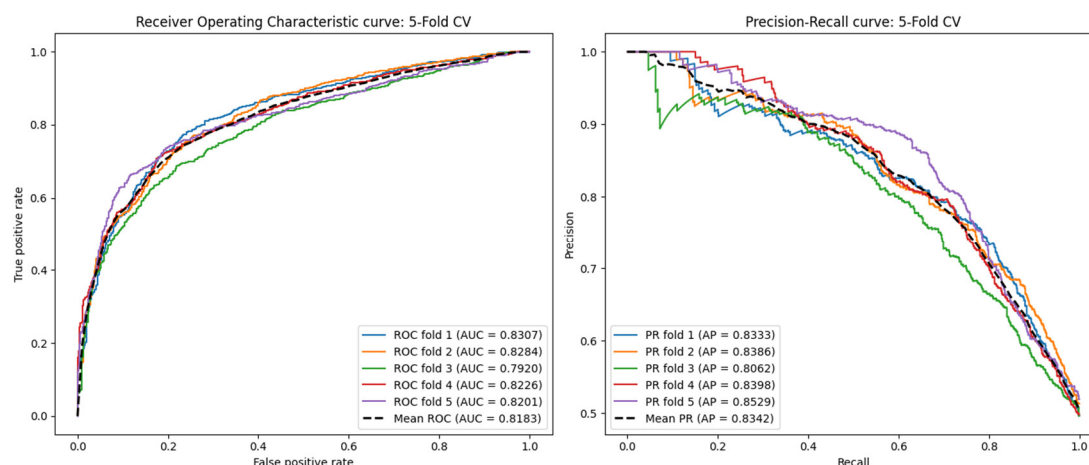


Figure 7. ROC curves and PR curves in fivefold cross-validation under the blind test set.

2.6. Case Studies

In our study, we employed the AAECDA-based model. Given that known associations were obtained from the GDSC database [23], we utilized the circRNA-drug sensitivity associations from GDSC as the training set and the circRNA-drug associations from the CTRP database [24] as the test set. To predict 10 potential circRNA-drug association combinations, we selected two specific drugs: Temozolomide and Cisplatin. The predicted scores were ranked from highest to lowest, where higher scores indicate stronger potential associations. These combinations may have therapeutic potential for certain diseases, and some have been supported by existing literature. The associations between circRNAs and drugs can be visually represented using Figure 8.

Temozolomide [25] is an alkylating antitumor drug that is primarily employed in the treatment of malignant brain tumors, including glioblastoma multiforme (GBM) and anaplastic astrocytoma. This oral chemotherapy drug boasts excellent bioavailability and tissue penetration, particularly its ability to cross the blood-brain barrier, making it highly effective in treating brain tumors. It is regarded as one of the standard treatments for gliomas. As indicated in Table 5, of the top 10 circRNAs predicted to be associated with Temozolomide, 8 have already been validated by the CTRP through literature evidence.

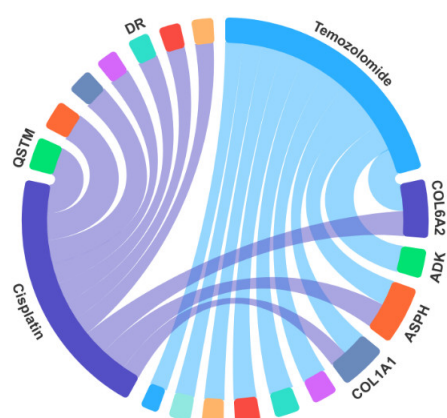


Figure 8. circRNAs association with Temozolomide and Cisplatin.

Cisplatin [26] is a well-established chemotherapy drug extensively used in clinical settings, classified under platinum-based anticancer agents. This drug is mainly utilized in treating a variety of solid tumors, such as those affecting the head and neck, lungs, bladder, ovaries, and testes. Cisplatin exerts its anticancer effects by binding to DNA and interfering

with normal cellular functions. The DNA damage it causes activates several cellular stress responses, including the activation of the *p53 gene*, which leads to apoptosis (programmed cell death). It is considered one of the cornerstone drugs for treating various cancers. As presented in Table 6, 9 of the top 10 predicted circRNAs associated with Cisplatin have been confirmed by the CTRP, with additional support from existing research literature.

Table 5. The Top 10 circRNAs associated with the drug Temozolomide.


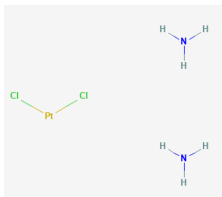
Drug	Rank	circRNA	Evidence
 Temozolomide	1	COL6A2	CTPR
	2	ADK	CTPR
	3	ASPH	CTPR
	4	COL1A1	CTPR
	5	EFEMP1	CTPR
	6	RPN1	NA
	7	MYH9	CTPR
	8	ADGRG1	CTPR
	9	COPG1	CTPR
	10	KATNB1	NA

Table 6. The Top 10 circRNAs associated with the drug Cisplatin.

Drug	Rank	circRNA	Evidence
 Cisplatin	1	SQSTM1	CTPR
	2	CALR	CTPR
	3	ASPH	CTPR
	4	COL6A2	CTPR
	5	LTBP1	CTPR
	6	VIM	CTPR
	7	WDR5	CTPR
	8	MYADM	CTPR
	9	POLR2A	NA
	10	COL1A1	CTPR

3. Materials and Methods

3.1. Dataset

The datasets and processing methods employed in this study are detailed as follows. We utilized a dataset originally proposed by Deng et al. [9], where circRNA-drug sensitivity associations were sourced from the circRic database [27], and drug sensitivity data were retrieved from the GDSC database [23]. This dataset encompasses 80,076 associations involving 404 circRNAs and 250 drugs. To establish the relationships between each circRNA and drug sensitivity pair, the Wilcoxon test was applied. Correlations with a false discovery rate of less than 0.05 were classified as significant associations. The benchmark dataset used in this study includes only the circRNA-drug sensitivity pairs with significant associations, comprising a total of 4134 associations, 218 drugs, and 271 circRNAs. From this, we constructed the circRNA-drug association matrix $A \in R^{271 \times 218}$. For elements in A , $A_{ij} = 1$ indicates that circRNA i is associated with the sensitivity of drug j , otherwise $A_{ij} = 0$, where i and j are the indices of circRNAs and drugs in A , respectively. Furthermore, in order to compute the similarity between circRNAs and drugs, we retrieved the circRNA host gene sequences from the NCBI Gene database [28], while the drug structural information was gathered from the PubChem database [29] provided by NCBI. This dual-source data acquisition allowed for a comprehensive analysis of both genetic and chemical features.

3.2. Construction of the Similarity Network

3.2.1. Sequence Similarity of Host Genes of circRNAs

Since circRNAs are formed from exons, introns, or a combination of both from their host genes, their sequence information is closely related to the host gene. If two circRNAs originate from similar or identical genomic regions, their host gene sequences will exhibit higher similarity. Therefore, by comparing the sequences of host genes, potential circRNA similarities can be identified. The similarity of circRNAs can be calculated using the sequence information of their host genes. The Levenshtein distance metric [30] is a tool for measuring the difference between two strings. By calculating the Levenshtein distance between two circRNA sequences, the degree of similarity between them can be quantified. We denote the similarity between circRNAs as $SS_C \in R^{271 \times 271}$.

$$SS_{C_{\text{leven}}}(c_i, c_j) = 1 - \frac{\text{trans}}{\text{len}(c_i) + \text{len}(c_j)} \quad (5)$$

where *trans* represents the minimum cost of transforming one circRNA into another, and *len*(·) denotes the length of the circRNA sequence.

3.2.2. Structural Similarity of Drugs

Obtaining the similarity between drugs by comparing their chemical structures is one of the key methods in drug design and screening. First, after acquiring the chemical structure information of drugs from the PubChem database, RDKit [31] is used to compute the topological fingerprints for each drug. Then, the Tanimoto method is employed to calculate their structural similarity [32]. As a result, we obtain the structural similarity matrix between drugs, denoted as $SS_D \in R^{218 \times 218}$.

3.2.3. Gaussian Interaction Profile Kernel Similarity of circRNAs and Drugs

In circRNA-drug interaction data, data sparsity is a prevalent challenge. To address this, GIP kernel similarity has been extensively utilized in prior research for similarity calculations [33–35]. By leveraging interaction profiles, GIP kernel similarity aids in overcoming the difficulties associated with sparse data, thereby improving the model's generalization capability. The GIP kernel similarity matrix for circRNA is calculated from the association matrix *A*, which is denoted as

$GS_C \in R^{271 \times 271}$, and the calculation is as follows:

$$GS_C(c_i, c_j) = \exp\left(-\eta_c \left\|IP(c_i) - IP(c_j)\right\|^2\right) \quad (6)$$

where $GS_C \in R^{M \times M}$, $IP(c_i)$ refers to the column corresponding to circRNA c_i within the circRNA-drug association matrix, while the parameter η_c is employed to control the kernel bandwidth. The parameter η_c is calculated as the average number of their associations. It is defined as follows:

$$\eta_c = \eta'_c / \left(\frac{1}{n_c} \sum_{k=1}^{n_c} IP(c_k)^2\right) \quad (7)$$

where $\eta_c = 1.0$, and η_c represents the number of circRNAs. Similarly, the similarity matrix for drugs is denoted as $GS_D \in R^{218 \times 218}$, and the GIP kernel similarity is calculated as follows:

$$GS_D(d_i, d_j) = \exp\left(-\eta_d \left\|IP(d_i) - IP(d_j)\right\|^2\right) \quad (8)$$

$IP(d_j)$ represents the row corresponding to drug d_j in the association matrix. The parameter η_d is the same as above:

$$\eta_d = \eta'_d / \left(\frac{1}{n_d} \sum_{k=1}^{n_d} \text{IP}(c_k)^2 \right) \quad (9)$$

where $\eta_d = 1.0$ is the number of drugs.

3.2.4. Similarity Fusion

As previously mentioned, we have separately calculated the similarities between circRNAs and drugs. To obtain more accurate similarities and integrate more biological information, we fuse the circRNA similarities with their Gaussian kernel similarities to construct a circRNA integrated similarity matrix, denoted as S_C . S_C is defined as follows:

$$S_{C_{ij}} = \begin{cases} \frac{(SS_{C_{ij}} + GS_{C_{ij}})}{2}, & \text{if } SS_C \neq 0 \\ GS_{C_{ij}}, & \text{otherwise} \end{cases} \quad (10)$$

Similarly, the integrated similarity matrix for drugs is calculated in the same manner as follows:

$$S_{D_{ij}} = \begin{cases} \frac{(SS_{D_{ij}} + GS_{D_{ij}})}{2}, & \text{if } SS_D \neq 0 \\ GS_{D_{ij}}, & \text{otherwise} \end{cases} \quad (11)$$

3.2.5. Multi-Scale Convolutional Neural Network

Based on the integrated similarities mentioned above, we input the fused features into a multi-scale convolutional neural network (MSCNN). In a traditional Convolutional Neural Network (CNN), convolutional layers extract features by applying fixed-size filters [36]. However, filters of a single scale may not capture all feature information. For instance, smaller filters are better at capturing detailed features, while larger filters are more suited to capturing global structures. Therefore, we use multiple convolutional kernels of different sizes to capture feature representations at different scales, which helps the model obtain better feature representations.

The core concept of a MSCNN involves utilizing multiple convolutional kernels of varying sizes within the same layer. The outputs from these different kernels are then concatenated to create a multi-scale feature representation. This method offers the advantage of capturing both fine-grained and coarse-grained feature information simultaneously, thereby enhancing the model's expressiveness and overall performance. Specifically, the network includes three parallel convolutional layers, each using convolutional kernels of sizes 3, 5, and 7 to perform convolution operations on the input features. Each convolutional layer extracts feature representations at different scales, providing richer and more diverse feature representations for subsequent analysis.

From the integrated similarity matrices of circRNAs and drugs, corresponding similarity vectors can be obtained, which serve as their respective feature representations. Each circRNA-drug pair in the dataset is denoted as (c_i, d_i) , where c_i represents the similarity vector of each circRNA, and d_i represents the similarity vector of each drug. The numbers of circRNAs and drugs in the dataset are denoted as c_l and d_l , respectively. Based on this, we learn better feature representations of vector c_i from the feature vectors E_{c_i} . Specifically, the feature representation generated by the convolution of vector E_{c_i} using the l -th filter is

$$s_{c_i}^l = \psi \left(W_{c_i}^l * E_{c_i} + b_{c_i}^l \right) \quad (12)$$

where $*$ represents the convolution operation, $W_{c_i}^l$ is the weight matrix, $b_{c_i}^l$ is the corresponding bias term, and $s_{c_i}^l$ denotes the features obtained through the l filter. Additionally, $\psi(\cdot)$ is the nonlinear activation function. $K \times n$ represents the size of the filter. Similarly,

the feature representation for drugs is denoted as $s_{d_i}^l$. Then, the obtained feature vectors are concatenated in the feature merging layer to form the integrated feature representation, which is expressed as:

$$h_{c_i} = \text{Concat}\left(s_{c_i}^1 \| s_{c_i}^2 \| \dots \| s_{c_i}^l\right) \quad (13)$$

where $\text{Concat}(\cdot)$ represents the concatenation operation for the features. Similarly, the integrated features for drugs are denoted as h_{d_i} .

Next, the fully connected layer further extracts and compresses the feature information of both through nonlinear transformations. The fully connected layer extracts and compresses the feature information through a nonlinear transformation, and the operation is represented as follows:

$$x_{c_i} = \phi\left(W_{c_i}x_{c_i} + b_{c_i}\right) \quad (14)$$

where W_{c_i} represents the weights of the fully connected layer, b_{c_i} is the corresponding bias matrix, and $\phi(\cdot)$ is the activation function. Similarly, the feature of the drug after passing through the fully connected layer is denoted as x_{d_i} .

Finally, the feature vectors x_{c_i} and x_{d_i} for circRNAs and drugs processed by the fully connected layers are concatenated to form the final integrated feature representation x , which is expressed as follows:

$$x = \text{Concat}(x_{c_i} \| x_{d_i}) \quad (15)$$

The integrated feature representations of circRNAs and drugs will be input into the subsequent Adversarial Autoencoder (AAE) for more accurate prediction of the associations between diseases and circRNAs.

3.3. AAECDA

In this study, we developed a circRNA-drug association prediction model based on an adversarial autoencoder, termed AAECDA. The model is primarily composed of two parts: a MSCNN and an AAE. First, the input to the model is the integrated feature information of circRNA and drugs, which is fed into the MSCNN for feature extraction. The extracted circRNA and drug similarity representations are then input into the AAE for further feature learning. The AAE incorporates a discriminator network, enabling the encoder to not only reconstruct the input features but also generate latent representations in the hidden space that align with the true data distribution. After training, the features output by the encoder are used as input to a deep neural network (DNN) for the final association prediction. The DNN applies multiple layers of nonlinear transformations to further refine and optimize the feature representations, ultimately generating the predicted circRNA-drug association results.

The proposed circRNA-drug prediction model, AAECDA, is illustrated in Figure 9 and consists of the following steps:

- (1) Construction of similarity networks, as well as the sensitivity association network.
- (2) Extraction of integrated circRNA and drug features using the MSCNN.
- (3) Extraction of the latent representations of circRNA and drugs using the AAE.
- (4) Inputting the extracted latent representations into the DNN to predict the circRNA-drug association score.

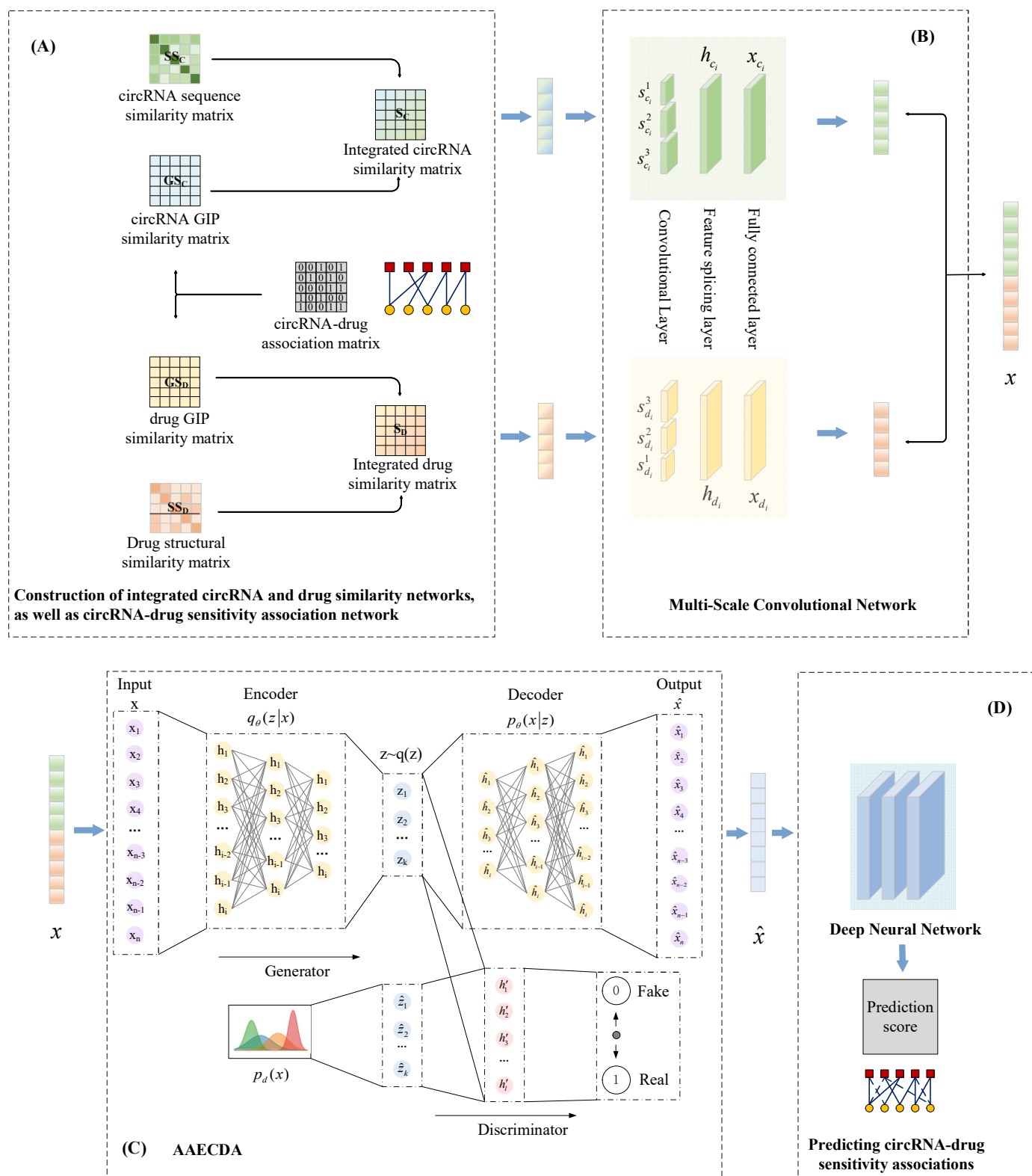


Figure 9. Overall architecture of AAECDA. (A) Construction of similarity network, (B) Details of extracting integrated features using MSCNN, (C) Details of extracting latent representations using AAE, (D) DNN predicting association scores.

3.3.1. Adversarial Autoencoder

The Adversarial Autoencoder (AAE) is a deep learning model that integrates the strengths of an Autoencoder (AE) with a Generative Adversarial Network (GAN) [37]. It combines the reconstruction abilities of an autoencoder with the adversarial training of a

GAN to effectively learn the latent representation of data. AAE incorporates an adversarial network, typically a discriminator, that aims to align the distribution of the latent space with a predefined prior distribution. Throughout the training, the discriminator and autoencoder are trained simultaneously with opposing objectives. The discriminator's task is to differentiate between genuine latent variables and those produced by the encoder, while the encoder works to fool the discriminator by making its generated latent variables closely resemble the prior distribution. As a generative autoencoder, AAE employs variational inference and GANs to enforce a specific prior distribution on the feature space. The main benefit of this method lies in its capacity to generate features that align with a predefined prior distribution while capturing the underlying data manifold, irrespective of the process's state.

A drawback of traditional AE-based feature extraction methods is that the distribution of the extracted features tends to be random and unregulated. To overcome this limitation, we propose a circRNA-drug association prediction method that utilizes the AAE.

A typical AE learns to generate an output identical to the input data x , with the encoder network and decoder network as its two components. The encoder extracts the feature vector z , and the decoder reconstructs the original vector. However, a limitation of AEs is that they randomly map input data into the feature space, resulting in uncontrolled feature distributions. To address this issue, AAE incorporates GAN. The generator (G) and discriminator (D) are the primary components of GANs, which are optimized through adversarial training [38]. The generator's objective is to learn the data distribution, while the discriminator's role is to evaluate the similarity between the generated data z and the real data x . Unlike the Variational Autoencoder (VAE), backpropagation through KL divergence requires an explicit functional form of the prior distribution. In contrast, AAEs only require sampling from the prior distribution, which enhances the model's flexibility.

We use AAE for feature extraction. The integrated feature representation x of circRNAs and drugs mentioned above is supplied to the encoder of the AAE to obtain low-dimensional latent feature representations. Part of the AAE's encoder is treated as the generator (G) to generate feature vectors analogous to real data, while the discriminator (D) is trained to discern between the distribution of the real data $p_d(x)$ and the distribution of the extracted feature vectors $q(z)$. Here, we define the posterior distribution of the hidden layer $q(z)$ as shown below:

$$q(z) = \int_x q_\phi(z|x) p_d(x) dx \quad (16)$$

In the equation, $q_\phi(z|x)$ represents the encoder network.

The training process of the AAE model involves two distinct stages: the reconstruction stage and the regularization stage. In the reconstruction phase, the encoder and decoder undergo training with the objective of minimizing the reconstruction error L . The reconstruction error L is specified as follows:

$$\mathcal{L}(\phi, \theta; x) = -\mathbb{E}_{q_\phi(z|x)} [\log(p_\theta(x|z))] \quad (17)$$

In the equation, $p_\theta(x|z)$ represents the decoder network. In the regularization stage, the adversarial network uses the cross-entropy loss function V to train the discriminator to distinguish between real experimental samples and generated samples, and update the parameters of the discriminator λ , as shown below:

$$-\mathcal{V}(\phi, \lambda; x, z) = -\log(d_\lambda(p(z))) - \log(1 - d_\lambda(q(z))) \quad (18)$$

In this equation, $p(z)$ represents the prior distribution and $q(z)$ represents the posterior distribution. Then, Training the generator and updates ϕ as follows:

$$-\mathcal{V}(\phi, \lambda; x, z) = \log(d_\lambda(q(z))) \quad (19)$$

Finally, the reconstruction loss and regularization loss are minimized through backpropagation.

After being decoded by the L-layer decoder, the reconstructed feature representation of circRNA and drugs is denoted as \hat{x} . Specifically, the generation process of \hat{x} is as follows:

$$\hat{x} = p_\theta(\hat{x}|z) \quad (20)$$

Through adversarial training, AAE approximates the distribution of the latent space to a predefined prior distribution (such as a standard normal distribution), making the latent space representation more regular and meaningful. This approach enhances the interpretability of the latent space and makes generating new samples from the latent space more stable. Compared to VAE, AAE does not require an exact specification of the prior distribution's functional form. Instead, adversarial training ensures that the latent distribution approximates a target distribution, increasing the model's flexibility and adaptability. Through adversarial training, AAE is able to learn more stable and generalized latent representations, which helps improve the model's performance when handling new data or performing transfer learning.

3.3.2. Association Prediction Based on Deep Neural Networks

After obtaining the latent feature representations using the previously described AAE, we employ the DNN to predict the associations between circRNAs and drugs. Each layer typically contains several neurons that are interconnected with the neurons in both the preceding and subsequent layers, forming a complex network structure. DNNs offer significant advantages, particularly in their ability to manage complex nonlinear relationships. By utilizing multiple hidden layers and nonlinear activation functions, DNNs are capable of capturing intricate patterns and features embedded within the data. Additionally, DNN models trained on large datasets exhibit strong generalization capabilities, performing well on previously unseen data and minimizing the risk of overfitting.

First, \hat{x} is input into the first fully connected layer to generate the hidden layer representation m_1 . Then, the hidden layer representation undergoes multiple transformations through fully connected layers, resulting in the final high-level feature representation m_n . After each fully connected layer, we introduce Dropout and Batch Normalization operations to prevent overfitting and accelerate model convergence.

Finally, the last hidden layer representation m_n is input into the output layer. The output layer passes through the Sigmoid activation function to generate the final predicted probability p :

$$p = \text{Sigmoid}(W_{\text{out}} \cdot m_n + b_{\text{out}}) \quad (21)$$

The Sigmoid activation function restricts the output value between $[0, 1]$, representing the predicted association probability.

For the training of the DNN, since our task is a binary classification problem of predicting circRNA-drug associations, we use the binary cross-entropy loss function. The formula for binary cross-entropy loss is:

$$L = -\frac{1}{N} \sum_{i=1}^N [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)] \quad (22)$$

where N is the number of samples, y_i is the true label of the i -th sample (taking values 0 or 1), and p_i is the predicted probability of the i -th sample. This loss function measures the difference between the predicted probabilities and the actual labels. Minimizing the binary

cross—entropy loss during the training process helps the DNN to make more accurate predictions.

In our code, we set the threshold for classification as 0.5. That is, when $p \geq 0.5$, the prediction result is positive (association exists); when $p < 0.5$, the prediction result is negative (no association exists).

Once the model has been trained, it can be utilized to identify and predict new associations between previously unknown circRNA and drug pairs within the dataset.

4. Conclusions

With the advancement of research on cancer and other diseases, increasing evidence shows that the expression of circRNAs in human cells can impact drug sensitivity, thereby playing a crucial role in treatment outcomes. As a result, predicting the associations between circRNAs and drug sensitivity not only facilitates the development of novel therapeutics but also aids in overcoming drug resistance in cells, ultimately improving treatment efficacy. There is an urgent need to develop a computational method for identifying potential associations. In this study, we introduce a novel computational framework, AAECDA. We first built a bimodal network based on known association data and quantified the similarities between drugs and circRNAs.

Next, we processed the network using a multi-scale convolutional neural network and further extracted and refined features through an adversarial autoencoder. Finally, DNN is used to learn low-dimensional features for prediction.

To assess the effectiveness of the AAECDA model, we conducted cross-validation across multiple datasets and compared its performance with various existing methods. The results indicate that AAECDA consistently outperformed other approaches in predicting circRNA-drug associations. In addition, we conducted a case study, and the results further validated the effectiveness of the AAECDA model in predicting novel associations. However, the experimentally validated associations available are still limited, which may affect the accuracy of the model's predictions to some extent. Moving forward, we aim to gather more circRNA-drug sensitivity data, along with additional bioinformatics data such as drug-disease and circRNA-disease associations, to improve the model's predictive performance by integrating multi-source data.

In conclusion, while current methods for predicting associations between drug sensitivity and circRNAs have certain limitations, this study offers valuable insights for advancing research and applications in this area. Future work will be necessary to enhance the accuracy and reliability of prediction models, paving the way for more effective approaches in this field.

Author Contributions: Conceptualization, Y.W. and X.L.; Methodology, Y.W.; Software, Y.W.; Validation, Y.W. and Y.C.; Formal analysis, Y.W.; Investigation, Y.W.; Resources, L.G. and F.-X.W.; Writing—original draft, Y.W.; Writing—review & editing, Y.W.; Visualization, Y.W. and Y.C.; Supervision, X.L.; Project administration, X.L.; Funding acquisition, X.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (62272288) and the Fundamental Research Funds for the Central Universities, Shaanxi Normal University (GK202302006).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The required data are available on GitHub <https://github.com/yjslzx/GATECDA> (accessed on 10 June 2024).

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Kumar, L.; Shamsuzzama; Haque, R.; Baghel, T.; Nazir, A. Circular RNAs: The Emerging Class of Non-coding RNAs and Their Potential Role in Human Neurodegenerative Diseases. *Mol. Neurobiol.* **2017**, *54*, 7224–7234. [CrossRef] [PubMed]
- Zhao, X.; Zhong, Y.; Wang, X.; Shen, J.; An, W. Advances in Circular RNA and Its Applications. *Int. J. Med. Sci.* **2022**, *19*, 975–985. [CrossRef] [PubMed]
- Syvänen, A.C. Accessing genetic variation: Genotyping single nucleotide polymorphisms. *Nat. Rev. Genet.* **2001**, *2*, 930–942. [CrossRef] [PubMed]
- Stoll, L.; Rodríguez-Trejo, A.; Guay, C.; Brozzi, F.; Bayazit, M.B.; Gattesco, S.; Menoud, V.; Sobel, J.; Marques, A.C.; Venø, M.T.; et al. A circular RNA generated from an intron of the insulin gene controls insulin secretion. *Nat. Commun.* **2020**, *11*, 5611. [CrossRef] [PubMed]
- Yang, Z.-G.; Awan, F.M.; Du, W.W.; Zeng, Y.; Lyu, J.; Wu, D.; Gupta, S.; Yang, W.; Yang, B.B. The Circular RNA Interacts with STAT3, Increasing Its Nuclear Translocation and Wound Repair by Modulating Dnmt3a and miR-17 Function. *Mol. Ther.* **2017**, *25*, 2062–2074. [CrossRef]
- Salzman, J.; Gawad, C.; Wang, P.L.; Lacayo, N.; Brown, P.O. Circular RNAs Are the Predominant Transcript Isoform from Hundreds of Human Genes in Diverse Cell Types. *PLoS ONE* **2012**, *7*, e30733. [CrossRef]
- Jeck, W.R.; Sorrentino, J.A.; Wang, K.; Slevin, M.K.; Burd, C.E.; Liu, J.; Marzluff, W.F.; Sharpless, N.E. Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA* **2013**, *19*, 426. [CrossRef]
- Guo, Y.; Lei, X.; Liu, L.; Pan, Y. circ2CBA: Prediction of circRNA-RBP binding sites combining deep learning and attention mechanism. *Front. Comput. Sci.* **2022**, *17*, 175904. Available online: <https://link.springer.com/article/10.1007/s11704-022-2151-0> (accessed on 14 September 2024). [CrossRef]
- Deng, L.; Liu, Z.; Qian, Y.; Zhang, J. Predicting circRNA-drug sensitivity associations via graph attention auto-encoder. *BMC Bioinform.* **2022**, *23*, 160. [CrossRef]
- Yang, B.; Chen, H. Predicting circRNA-drug sensitivity associations by learning multimodal networks using graph auto-encoders and attention mechanism. *Brief. Bioinform.* **2023**, *24*, bbac596. [CrossRef]
- Li, G.; Li, Y.; Liang, C.; Luo, J. DeepWalk-aware graph attention networks with CNN for circRNA-drug sensitivity association identification. *Brief. Funct. Genom.* **2023**, *23*, 418–428. [CrossRef] [PubMed]
- Yang, J.; Lei, X. Predicting circRNA-disease associations based on autoencoder and graph embedding—ScienceDirect. *Inf. Sci.* **2021**, *571*, 323–336. Available online: <https://www.sciencedirect.com/science/article/pii/S0020025521004084> (accessed on 14 September 2024). [CrossRef]
- Lei, X.; Tie, J.; Pan, Y. Inferring Metabolite-Disease Association Using Graph Convolutional Networks. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2022**, *19*, 688–698. [CrossRef] [PubMed]
- Fan, C.; Lei, X.; Pan, Y. Prioritizing CircRNA-disease Associations with Convolutional Neural Network Based on Multiple Similarity Feature Fusion. *Front. Genet.* **2020**, *11*, 540751. [CrossRef]
- Yang, J.; Lei, X.; Pan, Y. Predicting circRNA-disease Associations by Using Multi-Biomolecular Networks Based on Variational Graph Auto-Encoder with Attention Mechanism. *Chin. J. Electron.* **2024**, *33*, 1526–1537. [CrossRef]
- Lei, X.; Chen, Y.; Pan, Y. Multi-source Data with Laplacian Eigenmaps and Denoising Autoencoder for Predicting Microbe-disease Association via Convolutional Neural Network. *J. Comput. Sci. Technol.* **2024**. [CrossRef]
- Guo, Y.; Lei, X.; Pan, Y. An Encoding-Decoding Framework Based on CNN for circRNA-RBP Binding Sites Prediction. *Chin. J. Electron.* **2024**, *33*, 256–263. [CrossRef]
- 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**. Available online: <https://webofscience.clarivate.cn/wos/alldb/summary/f9308609-b5af-4a7c-bea0-26f1cc0a0375-01063c258d/relevance/1> (accessed on 6 September 2024).
- Lu, S.; Liang, Y.; Li, L.; Liao, S.; Zou, Y.; Yang, C.; Ouyang, D. Inferring circRNA-drug sensitivity associations via dual hierarchical attention networks and multiple kernel fusion. *BMC Genom.* **2023**, *24*, 796. [CrossRef]
- 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*, 291. [CrossRef]
- Yu, Z.; Huang, F.; Zhao, X.; Xiao, W.; Zhang, W. Predicting drug-disease associations through layer attention graph convolutional network. *Brief. Bioinform.* **2021**, *22*, bbaa243. [CrossRef]
- Yang, H.; Ding, Y.; Tang, J.; Guo, F. Inferring human microbe-drug associations via multiple kernel fusion on graph neural network. *Knowl.-Based Syst.* **2022**, *238*, 107888. [CrossRef]
- Yang, W.; Soares, J.; Greninger, P.; Edelman, E.J.; Lightfoot, H.; Forbes, S.; Bindal, N.; Beare, D.; Smith, J.A.; Thompson, I.R.; et al. Genomics of Drug Sensitivity in Cancer (GDSC): A resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res.* **2013**, *41*, D955–D961. [CrossRef] [PubMed]

24. Rees, M.G.; Seashore-Ludlow, B.; Cheah, J.H.; Adams, D.J.; Price, E.V.; Gill, S.; Javaid, S.; E Coletti, M.; Jones, V.L.; E Bodycombe, N.; et al. Correlating chemical sensitivity and basal gene expression reveals mechanism of action. *Nat. Chem. Biol.* **2016**, *12*, 109–116. [[CrossRef](#)]
25. Gonzalez, R.D.; Small, G.W.; Green, A.J.; Akhtari, F.S.; Havener, T.M.; Quintanilha, J.C.F.; Cipriani, A.B.; Reif, D.M.; McLeod, H.L.; Motsinger-Reif, A.A.; et al. RYK Gene Expression Associated with Drug Response Variation of Temozolomide and Clinical Outcomes in Glioma Patients. *Pharmaceuticals* **2023**, *16*, 726. [[CrossRef](#)] [[PubMed](#)]
26. El Charif, O.; Mapes, B.; Trendowski, M.R.; Wheeler, H.E.; Wing, C.; Dinh, P.C.; Frisina, R.D.; Feldman, D.R.; Hamilton, R.J.; Vaughn, D.J.; et al. Clinical and Genome-wide Analysis of Cisplatin-induced Tinnitus Implicates Novel Ototoxic Mechanisms. *Clin. Cancer Res.* **2019**, *25*, 4104–4116. [[CrossRef](#)]
27. Ruan, H.; Xiang, Y.; Ko, J.; Li, S.; Jing, Y.; Zhu, X.; Ye, Y.; Zhang, Z.; Mills, T.; Feng, J.; et al. Comprehensive characterization of circular RNAs in ~ 1000 human cancer cell lines. *Genome Med.* **2019**, *11*, 55. [[CrossRef](#)]
28. Rangwala, S.H.; Kuznetsov, A.; Ananiev, V.; Asztalos, A.; Borodin, E.; Evgeniev, V.; Joukov, V.; Lotov, V.; Pannu, R.; Rudnev, D.; et al. Accessing NCBI data using the NCBI Sequence Viewer and Genome Data Viewer (GDV). *Genome Res.* **2021**, *31*, 159–169. [[CrossRef](#)]
29. Wang, Y.; Bryant, S.H.; Cheng, T.; Wang, J.; Gindulyte, A.; Shoemaker, B.A.; Thiessen, P.A.; He, S.; Zhang, J. PubChem BioAssay: 2017 update. *Nucleic Acids Res.* **2017**, *45*, D955–D963. [[CrossRef](#)]
30. Saowapa, K.; Kaneko, H.; Fujiwara, E. Systematic binary deletion/insertion error correcting codes capable of correcting random bit errors. *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.* **2000**, *E83A*, 2699–2705.
31. Landrum, G. RDKit: Open-source cheminformatics from machine learning to chemical registration. *Abstr. Pap. Am. Chem. Soc.* **2019**, 258. [[CrossRef](#)]
32. Shen, L.; Liu, F.; Huang, L.; Liu, G.; Zhou, L.; Peng, L. VDA-RWLRLS: An anti-SARS-CoV-2 drug prioritizing framework combining an unbalanced bi-random walk and Laplacian regularized least squares. *Comput. Biol. Med.* **2022**, *140*, 105119. [[CrossRef](#)] [[PubMed](#)]
33. Niu, M.; Zou, Q.; Wang, C. GMNN2CD: Identification of circRNA-disease associations based on variational inference and graph Markov neural networks. *Bioinformatics* **2022**, *38*, 2246–2253. [[CrossRef](#)] [[PubMed](#)]
34. Li, Z.; Li, J.; Nie, R.; You, Z.-H.; Bao, W. A graph auto-encoder model for miRNA-disease associations prediction. *Brief. Bioinform.* **2021**, *22*, bbaa240. [[CrossRef](#)] [[PubMed](#)]
35. Zhang, Y.; Lei, X.; Fang, Z.; Pan, Y. CircRNA-Disease Associations Prediction Based on Metapath2vec++ and Matrix Factorization. *Big Data Min. Anal.* **2020**, *3*, 280–291. [[CrossRef](#)]
36. Xu, X.; Zhao, X.; Wei, M.; Li, Z. A comprehensive review of graph convolutional networks: Approaches and applications. *Electron. Res. Arch.* **2023**, *31*, 4185–4215. [[CrossRef](#)]
37. Jang, K.; Hong, S.; Kim, M.; Na, J.; Moon, I. Adversarial Autoencoder Based Feature Learning for Fault Detection in Industrial Processes. *IEEE Trans. Ind. Inform.* **2022**, *18*, 827–834. [[CrossRef](#)]
38. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial Networks. *Commun. ACM* **2020**, *63*, 139–144. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.