frontiers | Frontiers in Pharmacology

TYPE Original Research
PUBLISHED 28 April 2025
DOI 10.3389/fphar.2025.1543966

Check for updates

OPEN ACCESS

EDITED BY
Michael Heinrich,
University College London, United Kingdom

REVIEWED BY
Bo-Wei Zhao,
Chinese Academy of Sciences (CAS), China
Zeheng Wang,
Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia

*CORRESPONDENCE
Huixia Zhou,
huixia99999@163.com
Mengnan Yu,
mmnnyu@163.com

†These authors have contributed equally to this work and share first authorship

RECEIVED 12 December 2024
ACCEPTED 15 April 2025
PUBLISHED 28 April 2025

CITATION
He F, Duan L, Xing G, Chang X, Zhou H and Yu M (2025) AMFGNN: an adaptive multi-view fusion graph neural network model for drug prediction.
*Front. Pharmacol.* 16:1543966.
doi: 10.3389/fphar.2025.1543966

COPYRIGHT
© 2025 He, Duan, Xing, Chang, Zhou and Yu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# AMFGNN: an adaptive multi-view fusion graph neural network model for drug prediction

Fang He[1,2,3,4†], Lian Duan[1,3,4,5†], Guodong Xing[1,3,4,5†], Xiaojing Chang[1,3,4,5], Huixia Zhou[1,3,4,5*] and Mengnan Yu[1,3,4,5*]

[1]Faculty of Pediatrics, The Chinese PLA General Hospital, Beijing, China, [2]Department of Child Growth and Development Clinic, The Seventh Medical Center of PLA General Hospital, Beijing, China, [3]National Engineering Laboratory for Birth Defects Prevention and Control of Key Technology, Beijing, China, [4]Beijing Key Laboratory of Pediatric Organ Failure, Beijing, China, [5]Department of Pediatric Surgery, The Seventh Medical Center of PLA General Hospital, Beijing, China

**Introduction:** Drug development is a complex and lengthy process, and drug-disease association prediction aims to significantly improve research efficiency and success rates by precisely identifying potential associations. However, existing methods for drug-disease association prediction still face limitations in feature representation, feature integration, and generalization capabilities.

**Methods:** To address these challenges, we propose a novel model named AMFGNN (Adaptive Multi-View Fusion Graph Neural Network). This model leverages an adaptive graph neural network and a graph attention network to extract drug features and disease features, respectively. These features are then used as the initial representations of nodes in the drug-disease association network to enable efficient information fusion. Additionally, the model incorporates a contrastive learning mechanism, which enhances the similarity and differentiation between drugs and diseases through cross-view contrastive learning, thereby improving the accuracy of association prediction. Furthermore, a Kolmogorov-Arnold network is employed to perform weighted fusion of various final features, optimizing prediction performance.

**Results:** AMFGNN demonstrates a significant advantage in predictive performance, achieving an average AUC value of 0.9453, which reflects the model's high accuracy in prediction.

**Discussion:** Cross-validation results across multiple datasets indicate that AMFGNN outperforms seven advanced drug-disease association prediction methods. Additionally, case studies on Hepatoblastoma, asthma and Alzheimer's disease further confirm the model's effectiveness and potential value in real-world applications.

KEYWORDS

**drug prediction, drug-disease association prediction, graph attention network, contrastive learning, Kolmogorov-Arnold network**

## 1 Introduction

Since the outbreak of the COVID-19 pandemic, the global healthcare system has faced unprecedented challenges (Vaz et al., 2023; Meng et al., 2024), making the need for safe and effective treatment strategies more urgent than ever. Drug repositioning has attracted widespread attention because of its ability to rapidly identify new therapeutic options,

effectively reducing both the cost and timeline of new drug development (Pushpakom et al., 2019; Tang et al., 2023). Recent breakthroughs in high-throughput screening technologies and continuous improvements in computational methods have significantly enhanced the efficiency and accuracy of computer-aided drug repositioning in identifying potential drug-disease associations (Singh et al., 2024; Zeng et al., 2024).

In the field of drug-disease association (DDA) prediction, research methods have gradually evolved from traditional machine learning models to deep learning techniques (Lavecchia, 2015). Traditional machine learning-based DDA prediction identifies potential drug-disease associations through data modeling and feature extraction, and the application of diverse algorithms has further expanded the depth of research in this area. For example, Gao et al. (2022a) combined similarity fusion technology with Laplacian regularization algorithms to accurately predict new indications for drugs and diseases; Yang et al. (2021) developed the MKDGRLS model, employing a multi-kernel approach and Laplacian regularization to handle complex interactions and optimizing model parameters through alternating least squares; additionally, Zhang et al. (2020) proposed the Bayesian inductive matrix completion (DRIMC) method, which integrates features from multiple data sources for analysis in latent space, effectively predicting new drug applications; Niu et al. (2024) introduced the SRR-DDI model, which utilizes a self-attention mechanism to represent drug substructures finely and incorporates drug similarity features, significantly enhancing the stability and performance of drug interaction predictions.

Compared to traditional methods, drug-disease association (DDA) prediction methods based on graph neural networks (GNNs) have made significant breakthroughs in recent years (Zhang et al., 2021; Meng et al., 2024; Zeng et al., 2024). Their unique advantage lies in the ability of GNNs to effectively process multimodal data and model complex network structures, a characteristic that has also gained attention in other fields (Liu et al., 2023). For example, Gao et al. (2022b) proposed the CTST model, which constructs a heterogeneous network of drugs and diseases, using graph convolutional autoencoders to encode shared and unique features of nodes. The model also integrates features through an attention mechanism, significantly improving prediction accuracy. Zhao et al. (2022a) developed the HINGRL model, which leverages a heterogeneous information network encompassing drug-disease and protein-protein interactions, enhances feature recognition through graph representation learning, and combines a random forest algorithm for precise drug indication prediction. Yang et al. (2024) designed the GCNGAT model, which integrates graph convolutional networks with graph attention networks, particularly suitable for drug repositioning. This model analyzes drug-disease associations by constructing heterogeneous graphs and extracts key interaction features in multi-disease contexts. Additionally, Liu et al. (2024) proposed the AMDGT framework, which uses a dual-graph transformer technique to integrate similarity data and complex biochemical information, deeply merging drug and disease features to efficiently predict potential drug associations. Wang et al. (2025) proposed an automatic collaborative learning framework that integrates neighbor interaction metrics with the message-passing mechanism of Graph Neural Networks to enhance prediction accuracy.

Although machine learning and deep learning methods have made significant progress in drug-disease association (DDA) prediction, existing approaches still face challenges in handling complex, multidimensional data and efficiently integrating information from multiple perspectives. To address this issue, we propose an Adaptive Multi-view Fusion Graph Neural Network (AMFGNN) model. The model first constructs drug-drug similarity networks and disease-disease similarity networks, using graph attention networks to extract drug and disease features, which are then used as initial features for the downstream drug-disease association network nodes. To further improve the accuracy of drug-disease association prediction, the model incorporates a contrastive learning mechanism that enhances the similarity and dissimilarity between drugs and diseases through cross-view contrastive learning. The final features are predicted using a Kolmogorov-Arnold Networks (KAN), further improving the prediction accuracy of drug-disease associations. By integrating multi-source heterogeneous information and achieving adaptive weighted fusion, the model flexibly handles complex drug and disease data, dynamically adjusting the weights of different information sources, thereby improving the accuracy, stability, and generalization of predictions. This approach establishes deeper connections between multidimensional data and multi-view information, providing stronger support for drug discovery and personalized medicine.

## 2 Materials and methods

In this section, we first describe the benchmark datasets used in the proposed model. Next, we introduce the AMFGNN model framework, which consists of three main components. As shown in Figure 1, the framework includes: (i) construct similarity network, (ii) feature extraction and fusion module, and (iii) prediction module.

### 2.1 Dataset

To comprehensively evaluate the performance of the proposed model, we adopted three benchmark datasets: Gottlieb et al. (2011), Luo et al. (2016), and Liang et al. (2017), which are widely used in drug repositioning research. Fdataset is a validated dataset containing 1,632 known drug-disease associations, involving 468 drugs and 298 diseases, providing a reliable reference standard for drug repositioning studies. Cdataset includes 663 drugs, 409 diseases, and 2,352 drug-disease interaction pairs, with the data first appearing in Luo et al.'s study. LRSSL consists of 3,051 validated drug-disease associations, involving 763 drugs and 681 diseases, and has been widely applied in drug repositioning research.

### 2.2 Graph attention networks

In this study, Graph Attention Network (GAT) (Veličković et al., 2017) is introduced for drug-disease prediction tasks. Graph Neural Networks (GNNs) (Wu et al., 2020) learn

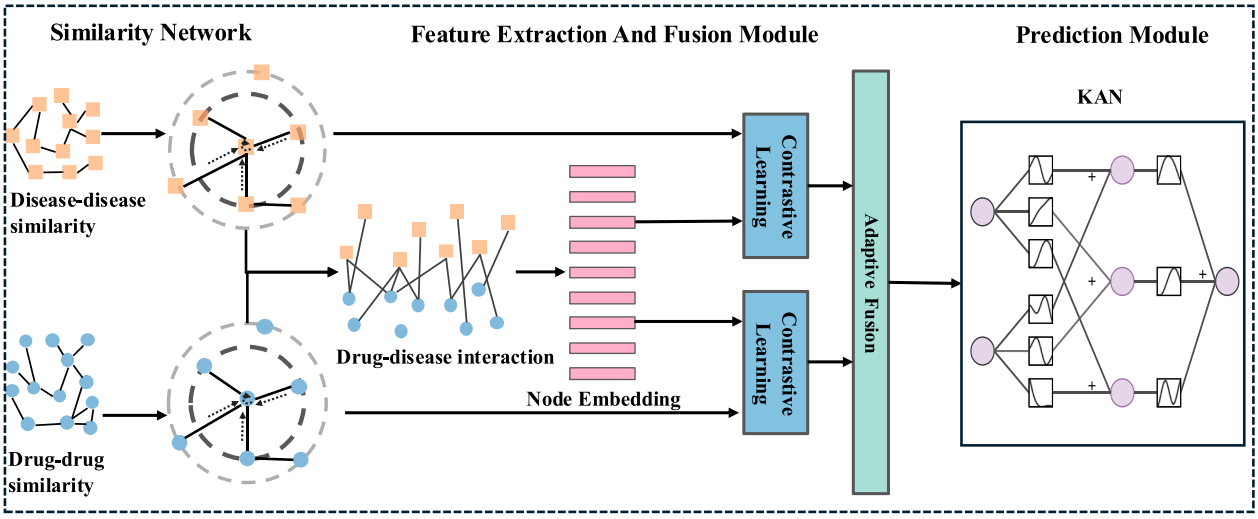


FIGURE 1
Illustration of the AMFGNN. The process begins with the construction of a similarity network, followed by feature extraction and fusion using the Graph Attention Network (GAT) and Adaptive Fusion Layer with contrastive learning. Finally, the prediction module employs Kolmogorov-Arnold Networks (KAN) for the final prediction.

interactions between nodes and edges in a graph to perform tasks such as node classification, graph classification, and link prediction. GAT extends this by assigning different weights to each node and its neighbors using an attention mechanism. These weighted features are aggregated to learn the node's embedding representation. We use a three-layer twin GAT network to extract features for both drugs and diseases. Taking the drug similarity network as an example, the drug similarity graph consists of a feature matrix $F_a$ and an adjacency matrix $A_a$. To construct the adjacency matrix $A_a$, we use a k-nearest neighbors (KNN) algorithm to identify the $K$ most similar drugs to a drug $a_i$, and establish connections. For drugs that do not belong to the $K$ nearest neighbors, no connection is made.

In GAT, the attention mechanism computes the importance of a drug $a_j$ for its neighboring drug $a_i$.Specifically, the attention coefficient between drug $a_i$ and its neighbor $a_j$ is calculated as follows Equation 1:

$$\alpha = \frac{exp\left(\text{LeakyReLU}\left(Wh_i \| Wh_j\right)\right)}{\sum_{k \in N_e(a_i)} exp\left(\text{LeakyReLU}\left(Wh_i \| Wh_k\right)\right)} \tag{1}$$

Here,$W$ is the learnable weight matrix, and $h_i$ and $h_j$ represent the feature vectors of drugs $a_i$ and $a_j$, respectively. The LeakyReLU function serves as the activation function, and the attention coefficient $\alpha_{ij}$ represents the importance of neighboring drug node $a_j$ to the central drug node $a_i$. A higher value of $\alpha_{ij}$ indicates a greater contribution from the features of node $a_j$ to updating the representation of node $a_i$. Through this mechanism, GAT effectively identifies and emphasizes connections in the drug similarity network that carry higher predictive significance, thus enhancing the quality of drug embedding representations and ultimately improving the accuracy of drug-disease association predictions.

After computing the attention coefficients, the feature vector of drug $a_i$ is updated by aggregating the features of all its neighboring drugs weighted by the attention coefficients. The updated feature vector $\widehat{h_i}$ is calculated as follows Equation 2:

$$\widehat{h_i} = \sigma\left(\sum_{j \in Ne(a_i)} \alpha W h_j\right) \tag{2}$$

where $\sigma$ is the activation function, and LeakyReLU is used. To capture the complex relationships between nodes, GAT introduces a multi-head attention mechanism that learns multiple sets of attention weights. The multi-head attention calculation is expressed as Equation 3:

$$\widehat{h_i} = \|_{k=1}^{K} \sigma\left(\sum_{j \in N_e(a_i)} \alpha_{ij}^k W^k h_j\right) \tag{3}$$

Here, $K$ is the number of attention heads, and $\alpha_{ij}^k$ represents the weight of the $k$-th attention head between drug $a_i$ and drug $a_j$.$W^k$ is the weight matrix for the $k$-th attention head.

The multi-head attention mechanism allows GAT to capture diverse relationships between drugs from multiple perspectives, alleviating information bottlenecks and enhancing the model's generalization ability. To integrate information from multiple heads, GAT averages the outputs of the different heads in the final layer, providing a more comprehensive and stable embedding representation Equation 4:

$$\widehat{h_i} = \frac{1}{K} \sum_{k=1}^{K} \widehat{h_i}^k \tag{4}$$

Through this multi-head attention mechanism, GAT can more effectively capture the complex dependencies between drugs and generate accurate drug embeddings. In this study, both the drug similarity and disease similarity networks employ a three-layer GAT structure to improve the accuracy of drug-disease prediction and the model's expressive power.

## 2.3 Adaptive fusion

The core idea of the Graph Attention Network (GAT) is to update node features based on the importance weights of neighboring nodes. However, as the depth of the GAT network increases, an issue of over-smoothing may arise, where the features of all nodes in the graph become too similar, severely affecting the model's prediction accuracy. To address this issue, we introduce residual connections to ensure that GAT can adaptively retain the original features while updating node features. The specific operation is as follows Equations 5-8:

$$R\left(H_{as}^{0}\right) = \mathrm{Elu}\left(W_a H_{as}^{0} + b_a\right) \tag{5}$$

$$R\left(H_{ds}^{0}\right) = \mathrm{Elu}\left(W_d H_{ds}^{0} + b_d\right) \tag{6}$$

$$R_{ad} = \left(R\left(H_{as}^{0}\right); R\left(H_{ds}^{0}\right)\right) \tag{7}$$

$$\begin{aligned} R\left(H_{ad}^{0}\right) &= \mathrm{Elu}\left(W_{ad} H_{ad}^{0} + b_{ad}\right) = \mathrm{Elu}\left(W_{ad}\left(H_{as}^{0}; H_{ds}^{0}\right) + b_{ad}\right) \\ &= \left(R_a; R_d\right) \end{aligned} \tag{8}$$

Here, $\mathrm{R}(\cdot)$ denotes the residual connection operation,$H_{as}^{0}$ denotes the initial feature representation of the drug similarity network (i.e., the pre-processed embedding derived from the original drug feature matrix $F_a$), while $H_{ds}^{0}$ represents the initial feature representation of the disease similarity network (i.e., the pre-processed embedding derived from the original disease feature matrix $F_d$). By introducing residual connections, the model adaptively preserves initial feature information, effectively addressing the issue of over-smoothing that typically occurs as the depth of the network increases. $H_{ad}^{0}$ corresponds to the feature of the drug-disease association view. The weight matrix is represented by $W_k$, where $k \in \{a, d, ad\}$, $b_k$ is the bias term, and $R_k$ denotes the residual result. The activation function used is the Elu function.

We fuse the residual connections of different views with the features from the previous layer, as shown in the following Equations 9-11:

$$H_a^{(1)} = \epsilon_1 H_{as}^{l} + (1 - \epsilon_1) R_a \tag{9}$$

$$H_d^{(1)} = \epsilon_2 H_{ds}^{l} + (1 - \epsilon_2) R_d \tag{10}$$

$$H_{ad} = \epsilon_3 H_{ad}^{l} + (1 - \epsilon_3) R_{ad} = \left(H_a^{(2)}; H_d^{(2)}\right) \tag{11}$$

Here, $H_a^{(q)}$ denotes the feature vector of the $q$-th view of the drug, $H_{as}^{l}$ is the feature representation of the drug similarity modality in the final layer of GAT, and $\epsilon_i$ is an adaptive variable learned during training, used to control the fusion weight between residual and original features.

To effectively integrate multi-view features of drugs and diseases, we adopt an adaptive feature fusion strategy Equations 12, 13:

$$H_a = \left[\eta_1 H_a^{(1)} + (1 - \eta_1) H_a^{(2)}; H_a^{(1)}; H_a^{(2)}\right] \tag{12}$$

$$H_d = \left[\eta_2 H_d^{(1)} + (1 - \eta_2) H_d^{(2)}; H_d^{(1)}; H_d^{(2)}\right] \tag{13}$$

Here, $H_a$ is the drug embedding after multi-view fusion, $H_d$ is the disease embedding after multi-modal fusion, and $\eta_i$ is the adaptive variable learned during model training. Finally, we input the fused drug embedding $M_i$ and disease embedding $D_j$ into KAN to compute the predicted potential connection $\hat{S}_{ij}$ Equation 14.

$$\hat{S}_{ij} = M_i^T \cdot D_j \tag{14}$$

During training, we use the cross-entropy loss function to minimize the error between the model's predictions and the true labels Equation 15:

$$\mathcal{L}_{CE} = -\sum_{(i,j)\epsilon x^+ \cup x^-} \left[S_{ij} \ln \hat{S}_{ij} + \left(1 - S_{ij}\right) \ln\left(1 - \hat{S}_{ij}\right)\right] \tag{15}$$

Here, $x^+$ and $x^-$ represent the positive and negative sample sets in the dataset, respectively, and $S_{ij}$ is the true association score between drug $a_i$ and disease $d_j$. To further improve the model's prediction accuracy and constrain the parameter updates across modalities, we introduce multi-view contrastive learning as a regularization term. This method reduces the distance between features of the same sample from different views, while increasing the distance between features of different samples from different views, thereby enhancing the model's feature representation ability. The objective function is defined as follows Equation 16:

$$dis\left(x, x^+\right) \ll d\left(x, x^-\right) \tag{16}$$

For each drug sample $a_i$, the contrastive learning loss function is defined as Equation 17:

$$\mathcal{L}_{MC} = \frac{1}{N} \sum_{j=1}^{N} \left(dis\left(m_i^1, m_i^2\right) - dis\left(m_i^1, m_j^2\right)\right) \tag{17}$$

Here, $N$ is the total number of drug samples,$m_i^p$ represents the embedding of drug sample $a_i$ in the $p$-th modality, and d (u, v) is the distance function calculated using cosine similarity Equation 18:

$$dis\left(u, v\right) = -\frac{u^T v}{\|u\| \cdot \|v\|} \tag{18}$$

Similarly, for each disease sample $d_j$, the loss function is defined as Equation 19:

$$\mathcal{L}_{DC} = \frac{1}{N} \sum_{k=1}^{N} \left(dis\left(d_j^1, d_j^2\right) - dis\left(d_j^1, d_k^2\right)\right) \tag{19}$$

Finally, the overall loss function of the model is defined as Equation 20:

$$\mathcal{L} = \mathcal{L}_{CE} + \omega_{MC} \mathcal{L}_{MC} + \omega_{DC} \mathcal{L}_{DC} \tag{20}$$

## 2.4 Kolmogorov-Arnold Networks

In the drug repositioning task, to improve the parameter efficiency of the model, we modified the traditional multilayer perceptron (MLP) structure by replacing the final MLP module with Kolmogorov-Arnold Networks (KAN). KAN introduces learnable activation functions, replacing traditional linear weight matrices, which significantly enhance the network's expressive power while maintaining or even improving model performance. The traditional MLP captures complex mappings through linear transformations and fixed nonlinear activation functions, which often limits model flexibility and increases parameter redundancy. Mathematically, an MLP is expressed as Equation 21:

$$MLP(Z) = (W_{K-1} \circ \sigma \circ W_{K-2} \circ \sigma \ldots \circ W_1 \circ \sigma \circ W_0)Z \quad (21)$$

Where $Z$ is the input vector,$W_k$ is the weight matrix, and $\sigma$ is the activation function. Although effective in learning complex functions, the fixed linear transformations in MLPs may restrict the adaptability of the network. KAN, in contrast, utilizes learnable nonlinear activation functions instead of fixed linear weight matrices, providing greater flexibility in capturing complex relationships between input features. Specifically, each connection in KAN is modeled by a combination of a parametric basis function and B-spline functions Equation 22:

$$f(x) = f(x_1, \ldots, x_n) = \sum_{q=1}^{2n+1} \Phi_q \left( \sum_{p=1}^{n} \phi_{q,p}(x_p) \right) \quad (22)$$

Each layer's $\Phi_i$ consists of a set of learnable activation functions, represented as Equation 23:

$$\Phi = \{\phi_{q,p}\}, \quad p = 1, 2, \ldots, n_{in}, \quad q = 1, 2, \ldots, n_{out} \quad (23)$$

We assume that a KAN can be expressed as $[n_0, n_1, \ldots, n_L]$ $n_i$ represents the number of neurons in the $i$-th layer. We use $(l, i)$ to denote the $i$-th neuron in the $l$-th layer and $X_{(l,i)}$ to represent the activation value of the neuron $(l, i)$. Between the $l$-th and $(l+1)$-th layers, there are $n_l \times n_{l+1}$ activation functions: the activation function connecting $(l, i)$ and $(l+1, i)$ is represented as Equation 24:

$$\phi_{l,j,i}, \quad l = 0, 1, \ldots, L-1, \quad i = 1, 2, \ldots, n_l, \quad j = 1, 2, \ldots, n_{l+1} \quad (24)$$

The pre-activation value of $\phi_{l,j,i}$ is $x_{(l,i)}$, and the post-activation value of $\phi_{l,j,i}$ is represented as Equation 25:

$$\hat{x}_{l,j,i} = \phi_{l,j,i}(x_{l,i}) \quad (25)$$

The activation value of neuron $x_{(l+1,i)}$ is the sum of all incoming post-activation values Equation 26:

$$x_{l+1,j} = \sum_{i=1}^{n_l} \hat{x}_{l,j,i} = \sum_{i=1}^{n_l} \phi_{l,j,i}(x_{l,i}), \quad j = 1, \ldots, n_{l+1} \quad (26)$$

The activation function $\phi$ is composed of a weighted sum of a basis function $b(x)$ and a B-spline function $spline(x)$ Equations 27-29:

$$\phi(x) = w_b b(x) + w_s \text{spline}(x) \quad (27)$$

$$b(x) = \text{SiLU}(x) = \frac{x}{1 + e^{-x}} \quad (28)$$

$$\text{spline}(x) = \sum_i c_i B_i(x) \quad (29)$$

Here, $w_b$ and $w_s$ represent the weights of the basis function and B-spline function, respectively. $c_i$ are the trainable parameters in the B-spline function, and $B_i(x)$ is the B-spline basis function defined on a grid. By using adaptive activation functions, KAN significantly enhances the representational power of neural networks, allowing the model to learn smoother and more complex transformations without increasing model complexity excessively. This flexible representation is particularly beneficial for integrating diverse features from multi-view data, such as drug similarity, disease similarity, and their interactions, leading to improved predictive performance for drug-disease associations.

# 3 Results and discussion

## 3.1 Parameter settings

We perform 10-fold cross-validation to evaluate the performance of AMFGNN. In the 10-fold cross-validation, all known and unknown drug-disease associations are randomly divided into 10 subsets of approximately equal size. Each subset is used as the test set in turn, while the remaining nine subsets serve as the training set.

We set the feature embedding size to 128 to achieve the best prediction performance in drug-drug similarity, disease-disease similarity, and drug-disease association graphs, and we set the dropout rate to 0.2 to optimize the training process of the network layers. For the graph model selection, we choose to use a 3-layer GAT instead of GCN (Graph Convolution Network) because GAT outperforms GCN in terms of AUC. The 3-layer structure helps prevent both information redundancy and over-smoothing issues. For optimization, we employed the Adam optimizer with a learning rate of 0.001 and weight decay of 0.001, training the model for 300 epochs.

## 3.2 Model evaluation and cross-validation

To rigorously evaluate the performance of our proposed model, we employed a standard 10-fold cross-validation approach. Specifically, drug-disease associations from each benchmark dataset were randomly divided into ten subsets of approximately equal size. In each fold, nine subsets were combined to form the training set, while the remaining subset served as the test set for evaluating the model's performance. This procedure was repeated ten times, with each subset serving as the test set exactly once. We reported the model's performance using the average and standard deviation across these ten evaluations. Furthermore, this cross-validation procedure was independently conducted on all three benchmark datasets (Fdataset, Cdataset, and LRSSL) to ensure the reliability and generalizability of our results.

## 3.3 Baseline methods

To evaluate the performance of AMFGNN, we performed 10-fold cross-validation on three public datasets: Fdataset, Cdataset, and LRSSL. The models compared in this study include LBMFF, SCPMFDD, SCPMF, MKGCN, and MNGACDA.

- LBMFF is a model for drug-disease relationship prediction that combines latent bilinear matrix factorization and focal loss. The model captures latent associations between drugs and diseases through matrix factorization techniques and introduces focal loss to address the class imbalance problem, enhancing the model's ability to handle sparse data and hard-to-predict instances (Kang et al., 2023).
- SCPMFDD is a semi-supervised learning model for drug-disease prediction. The model combines collaborative projection matrix factorization and semi-supervised learning strategies to enhance prediction performance by

TABLE 1 Performance comparison of different methods across datasets using AUC.

| Datasets | LBMFF | SCPMFDD | SCPMF | MKGCN | MNGACDA | DDAGDL | RGLDR | AMFGNN |
|---|---|---|---|---|---|---|---|---|
| Fdataset | 0.7953 ± 0.035 | 0.7740 ± 0.001 | 0.8957 ± 0.001 | 0.8870 ± 0.001 | 0.8179 ± 0.005 | 0.9266 ± 0.001 | 0.9311 ± 0.005 | 0.9328 ± 0.014 |
| Cdataset | 0.9069 ± 0.001 | 0.7937 ± 0.001 | 0.9117 ± 0.002 | 0.9109 ± 0.001 | 0.8406 ± 0.005 | 0.9256 ± 0.001 | 0.9359 ± 0.003 | 0.9443 ± 0.012 |
| LRSSL | 0.9139 ± 0.002 | 0.7668 ± 0.001 | 0.8977 ± 0.001 | 0.8596 ± 0.001 | 0.7936 ± 0.002 | 0.8990 ± 0.001 | 0.8032 ± 0.005 | 0.9588 ± 0.006 |
| Avg | 0.8720 | 0.7782 | 0.9017 | 0.8858 | 0.8173 | 0.9171 | 0.8901 | 0.9453 |

leveraging known drug-disease relationships and unlabeled data (Li et al., 2022).

- SCPMF is a semi-supervised learning model for drug-disease prediction. It learns the latent relationships between drugs and diseases through collaborative matrix factorization, while also utilizing semi-supervised learning to enhance the model's learning capability by incorporating unlabeled data. (Meng et al., 2021).
- MKGCN is a model for complex drug-disease prediction tasks. By introducing multiple kernel functions (Multi-Kernel), it integrates different types of graph structure features. MKGCN uses Graph Convolutional Networks (GCN) to process drug-disease graph data and applies kernel functions to weight different features, thereby more accurately capturing the complex relationships between drugs and diseases and improving the model's ability to model and predict (Cui et al., 2023).
- MNGACDA is a graph neural network model for drug-disease prediction. The model combines multi-node graph attention mechanisms and dual attention mechanisms, effectively processing the drug and disease relationship graph through Graph Convolutional Networks (Yang and Chen, 2023).
- DDAGDL (Zhao et al., 2022b) is a model for drug–disease prediction that applies geometric deep learning over heterogeneous information networks. It integrates biological information into the network structure and uses an attention mechanism to learn effective representations of drugs and diseases, enabling improved performance on non-Euclidean biomedical data.
- RGLDR (Zhao et al., 2025) combines regulation-aware graph representation learning with meta-path-based connectivity patterns to capture diverse regulatory mechanisms in heterogeneous biological networks. It enhances drug and disease embeddings using a multi-view attention mechanism and predicts drug-disease associations with an XGBoost classifier. Experimental results demonstrate its superior performance over state-of-the-art methods on benchmark datasets.

According to the results shown in Table 1, AMFGNN achieves the highest AUC in all three datasets. AMFGNN demonstrates outstanding performance across different datasets. On the F dataset, the AUC value of AMFGNN is 0.9328, significantly higher than other models such as LBMFF (0.7953), SCPMFDD (0.7740), and others. Similarly, on the C dataset and LRSSL dataset, AMFGNN also shows higher prediction accuracy, with AUC values of 0.9443 and 0.9588, respectively.

The results indicate that the AUC (Area Under the Curve) of the AMFGNN model outperform those of other models. This demonstrates that the AMFGNN model effectively improves the accuracy and stability of drug-disease association prediction by integrating multi-source heterogeneous information and dynamically adjusting the weights of different information sources. In addition, based on the results of Recall (Table 2) and F1-score (Table 3), AMFGNN also performs excellently in these metrics. In terms of Recall, AMFGNN effectively captures positive samples and reduces false negatives, indicating its high recall ability. A high Recall value means the model can identify more positive samples, which is especially important for drug-disease association prediction, as missing positive samples could lead to the omission of crucial information. Regarding F1-score, AMFGNN demonstrates a good balance between Precision and Recall, indicating its advantages in both accuracy and recall ability. A high F1-score means the model reduces false positives while effectively capturing more positive samples, avoiding the performance imbalance that may arise from optimizing a single metric. Overall, AMFGNN shows outstanding performance in improving prediction accuracy, stability, and comprehensiveness, further confirming its effectiveness in drug-disease association prediction, especially on complex and imbalanced datasets.

As shown in Figure 2, the AUC (Area Under the Curve) values of the AMFGNN model outperform those of other models. This demonstrates that the AMFGNN model effectively improves the accuracy and stability of drug-disease association prediction by integrating multi-source heterogeneous information and dynamically adjusting the weights of different information sources.

## 3.4 Ablation studies

To thoroughly investigate the effectiveness of different components in our proposed AMFGNN model, we conduct comprehensive ablation studies. Specifically, we evaluate four variants of our model:

- Full Model: The complete AMFGNN architecture with all components, including the Kolmogorov-Arnold Networks and contrastive learning loss.
- AMFGNN w/o KAN: A variant without the Kolmogorov-Arnold Network, while maintaining the contrastive learning mechanism.
- AMFGNN w/o CL: A variant that removes the contrastive learning loss while retaining the Kolmogorov-Arnold Networks, utilizing only the main task loss function for optimization.

TABLE 2 Recall comparison of various models across multiple datasets.

| Datasets | LBMFF | SCPMFDD | SCPMF | MKGCN | MNGACDA | DDAGDL | RGLDR | AMFGNN |
|---|---|---|---|---|---|---|---|---|
| Fdataset | 0.7328 | 0.0255 | 0.4128 | 0.0098 | 0.0801 | 0.4560 | 0.8814 | 0.9152 |
| Cdataset | 0.7006 | 0.0217 | 0.4767 | 0.0644 | 0.1695 | 0.4830 | 0.8724 | 0.9443 |
| LRSSL | 0.7165 | 0.0151 | 0.3812 | 0.0233 | 0.0737 | 0.4328 | 0.6975 | 0.9371 |
| Avg | 0.7166 | 0.0208 | 0.4236 | 0.0325 | 0.1078 | 0.4573 | 0.8171 | 0.9322 |

TABLE 3 F1 -Score performance of different approaches across datasets.

| Datasets | LBMFF | SCPMFDD | SCPMF | MKGCN | MNGACDA | DDAGDL | RGLDR | AMFGNN |
|---|---|---|---|---|---|---|---|---|
| Fdataset | 0.0577 | 0.1174 | 0.4024 | 0.0195 | 0.1475 | 0.6007 | 0.8615 | 0.8699 |
| Cdataset | 0.2343 | 0.0857 | 0.4556 | 0.1210 | 0.2875 | 0.6013 | 0.8806 | 0.8890 |
| LRSSL | 0.2304 | 0.0249 | 0.3548 | 0.0452 | 0.1352 | 0.5538 | 0.7194 | 0.9040 |
| Avg | 0.1741 | 0.0760 | 0.4043 | 0.0619 | 0.1901 | 0.5853 | 0.8205 | 0.8876 |

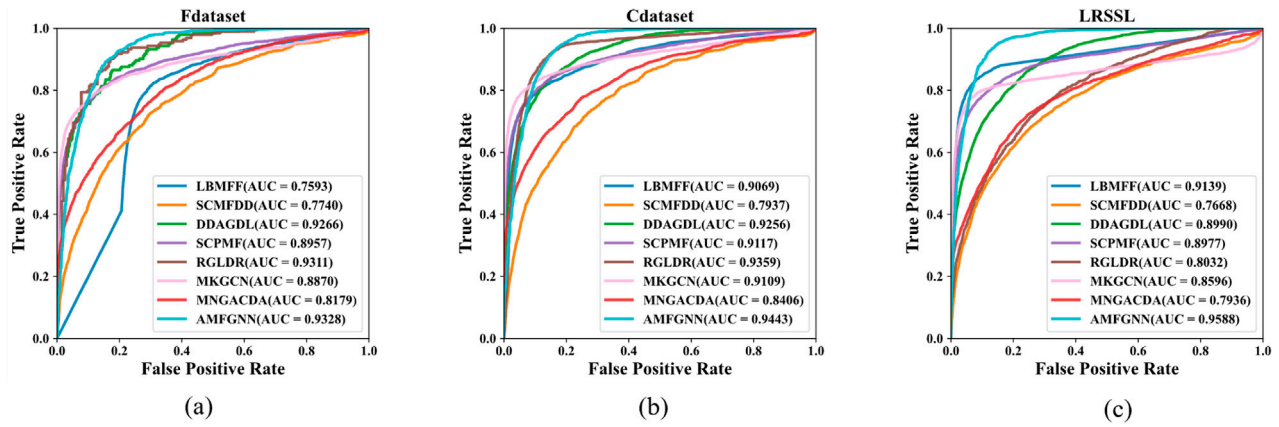


FIGURE 2
The ability of various methods to predict potential drugs for novel diseases is evaluated on public datasets using AUROC serving as the metric: **(a)** Fdataset **(b)** Cdataset **(c)** LRSSL.

- AMFGNN w/o KAN & CL: The baseline variant that removes both the Kolmogorov-Arnold Network and contrastive learning loss, maintaining only the basic MLP and main task loss.

As shown in Figure 3,the experimental results demonstrate several key findings: The full AMFGNN model achieves the best performance across all metrics, validating the effectiveness of our proposed architecture. Removing the Kolmogorov-Arnold Networks (w/o KAN) leads to a performance decrease of AUC and AUPR, highlighting the importance of KAN layer in our model. The absence of contrastive learning (w/o CL) also results in a drop, indicating that the contrastive learning mechanism plays a crucial role in learning more discriminative feature representations. The baseline variant (w/o KAN & CL) shows the most significant performance degradation, confirming that both components contribute substantially to the model's effectiveness.

These ablation studies clearly demonstrate that each component in our proposed AMFGNN makes meaningful contributions to the overall performance, and their combination leads to the optimal results.

## 3.5 Case study

To evaluate the practical application value of the model, we conducted a case study. Specifically, the model was trained on the Fdataset to predict potential drugs associated with asthma and hepatoblastoma. The predicted drugs were ranked in descending order based on their probability scores, and the top ten candidates were selected for further analysis. To assess the reliability of the model's predictions, comprehensive validation was performed using multiple authoritative data sources, including DrugCentral, Avram et al. (2021) CTD, and ClinicalTrials databases.

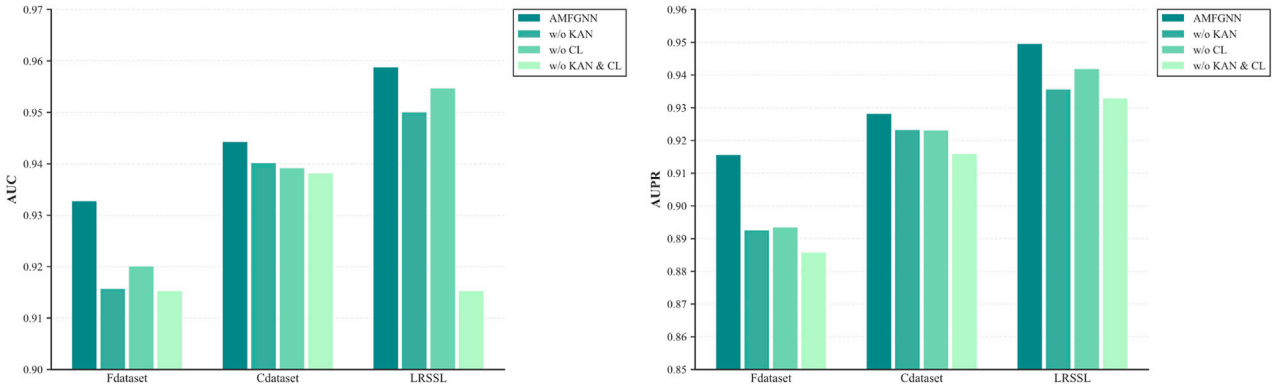


FIGURE 3
Results of ablation experiments on three datasets.

TABLE 4 TOP 5 candidate drugs for Hepatoblastoma predicted by AMFGNN.

| Rank | DrugBank IDs | Candidate drugs | Evidences |
|---|---|---|---|
| 1 | DB00397 | Phenylpropanolamine | DrugCentral |
| 2 | DB00313 | Valproic acid | DrugCentral, CTD |
| 3 | DB00215 | Citalopram | DrugCentral, CTD |
| 4 | DB00448 | Lansoprazole | DrugCentral, CTD |
| 5 | DB00167 | Isoleucine | Unconfirmed |

Hepatoblastoma is a malignant liver tumor primarily occurring in children, typically diagnosed during infancy or early childhood (Sharma et al., 2017). Table 4 highlights five potential therapeutic agents for hepatoblastoma predicted by the AMFGNN model, four of which have been validated by authoritative databases. Furosemide, widely used in both adults and children to manage hypertension and edema caused by liver dysfunction, was identified by the AMFGNN model as a promising candidate for hepatoblastoma treatment. This prediction has been corroborated by the DrugCentral database. Furthermore, the model predicted citalopram as another potential therapeutic agent for hepatoblastoma, with this conclusion supported by both the DrugCentral and CTD databases. These findings provide new perspectives and directions for drug development targeting hepatoblastoma.

Asthma is a chronic inflammatory disease influenced by both genetic and environmental factors (Toskala and Kennedy, 2015), making it a complex hereditary condition. Table 5 lists the top five potential asthma treatments predicted based on the F dataset, of which four are verified through reliable databases or clinical trials, further supporting the accuracy and practicality of the model's predictions. AMFGNN predicts flunisolide as a potential drug for treating asthma, a conclusion supported by both DrugBank and ClinicalTrials.gov. Additionally, studies show that uncontrolled asthma is often associated with gastroesophageal reflux disease (GERD) (Harding, 2003). As a proton pump inhibitor, esomeprazole is widely used to treat GERD, and AMFGNN also predicts that esomeprazole might have therapeutic effects on asthma, a prediction that is verified by ClinicalTrials.gov.

Alzheimer's Disease (AD) is a neurodegenerative disorder characterized primarily by progressive cognitive decline and memory impairmentAbubakar et al. (2022). Table 6 lists five potential drug candidates for the treatment of Alzheimer's Disease as predicted by the AMFGNN model, four of which have already been validated by authoritative pharmaceutical databases. Memantine, an approved N-methyl-D-aspartate (NMDA) receptor antagonist, was identified by AMFGNN as an effective therapeutic agent for Alzheimer's Disease. Additionally, Methylphenidate and Levothyroxine were also predicted as promising therapeutic candidates, with supporting evidence from clinical trials documented in the ClinicalTrials.gov database, suggesting their potential clinical application in the future.

TABLE 5 TOP 5 candidate drugs for asthma predicted by AMFGNN.

| Rank | DrugBank IDs | Candidate drugs | Evidences |
|---|---|---|---|
| 1 | DB00180 | Flunisolide | DB, ClinicalTrials.gov |
| 2 | DB00736 | Esomeprazole | ClinicalTrials.gov |
| 3 | DB00182 | Amphetamine | DrugCentral, ClinicalTrials.gov |
| 4 | DB00181 | Baclofen | Unconfirmed |
| 5 | DB00695 | Furosemide | ClinicalTrials.gov |

TABLE 6 TOP5 candidate drugs for Alzheimer's Disease predicted by AMFGNN.

| Rank | DrugBank IDs | Candidate drugs | Evidences |
|---|---|---|---|
| 1 | DB01043 | Memantine | ClinicalTrials.gov, CTD, DrugCentral |
| 2 | DB00642 | Methylphenidate | ClinicalTrials.gov |
| 3 | DB00636 | Clofibrate | Unconfirmed |
| 4 | DB00413 | Pramipexole | ClinicalTrials.gov, DrugCentral |
| 5 | DB00337 | Levothyroxine | ClinicalTrials.gov |

TABLE 7 Molecular binding energies (kcal/mol) between the top 5 candidate drugs for asthma predicted by AMFGNN and 5 target proteins.

| Drug | 1KTJ | 3FY1 | 4P0I | 5BOW | 7XXW |
|---|---|---|---|---|---|
| Flunisolide | −7.4 | −9.6 | −9.6 | −7.5 | −6.8 |
| Esomeprazole | −8.0 | −8.5 | −6.6 | −7.0 | −5.9 |
| Amphetamine | −5.7 | −5.5 | −4.8 | −5.1 | −4.4 |
| Baclofen | −6.0 | −6.6 | −5.7 | −6.4 | −4.3 |
| Furosemide | −6.2 | −7.4 | −6.2 | −6.8 | −5.9 |

Moreover, we selected five asthma-related target proteins and conducted molecular docking simulations to assess their binding abilities with five candidate drugs using AutoDock Vina (Trott and Olson, 2010). The interactions between the ligands and target proteins were further analyzed using Discovery Studio (DS) visualization software. Regarding the relevance of baclofen to asthma, we chose acidic mammalian chitinase (AMCase, PDB code: 3FY1) as the target protein and found that baclofen has a binding energy of −6.6 kcal/mol with AMCase (Table 7). Figure 4 shows the van der Waals interactions between baclofen and several specific amino acid residues (ALA:183, MET:385, MET:210, TYR:267, GLU:140, GLY:98, PHE:58). In addition, other types of molecular interactions are observed. For example, the oxygen atom forms conventional hydrogen bonds with residues TYR:212, ASP:213, TRP:99, and NA:1, while the nitrogen atom exhibits carbon-hydrogen bond interactions with residue ASP:138. Furthermore, $\pi-\sigma$ interactions are observed between the small molecule and residues TYR:27 and TRP:360, and covalent bonds form between the oxygen atom and the functional group of residue NA:1.

## 4 Conclusion

This paper presents the Adaptive Multi-View Fusion Graph Neural Network (AMFGNN), a novel model designed for drug-disease association prediction. AMFGNN integrates multiple data sources, including drug similarity, disease similarity, and drug-disease interactions, using an adaptive multi-view feature fusion strategy. The

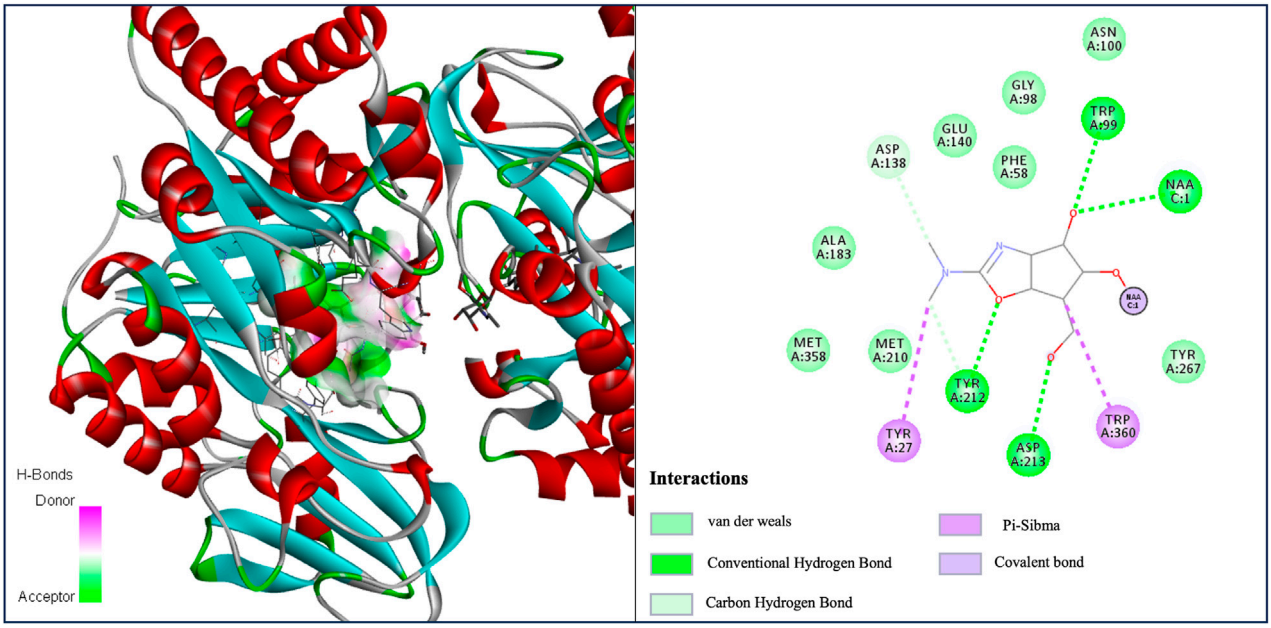


FIGURE 4
Docking and interactions of Baclofen (DrugBank ID: DB00181) with AMCase (PDB code: 3FY1).

model combines Graph Attention Networks (GAT) with contrastive learning to improve the accuracy of drug-disease predictions by effectively capturing relationships between nodes in the graph. Additionally, replacing traditional multilayer perceptron (MLP) layers with Kolmogorov-Arnold Networks (KAN) enhances the model's flexibility, expressive capability, and overall predictive performance.

Through 10-fold cross-validation on three benchmark datasets (F dataset, C dataset, and LRSSL), AMFGNN outperforms existing models, achieving high area under the curve (AUC) scores of 0.9328, 0.9443, and 0.9588, respectively. These results demonstrate that AMFGNN significantly improves drug-disease prediction accuracy, making it a valuable tool in drug repositioning and personalized medicine.

## Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding authors.

## Author contributions

FH: Writing – original draft, Writing – review and editing. LD: Writing – original draft, Writing – review and editing. GX: Writing – original draft, Writing – review and editing. XC: Investigation, Writing – original draft. HZ: Supervision, Writing – review and editing. MY: Supervision, Writing – review and editing.

## Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by the Innovation Talent Fund of Senior Department of Pediatrics, The Seventh Medical Center of PLA General Hospital (Grant No. QZX-04-EKCLJH-5).

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

Abubakar, M. B., Sanusi, K. O., Ugusman, A., Mohamed, W., Kamal, H., Ibrahim, N. H., et al. (2022). Alzheimer's disease: an update and insights into pathophysiology. *Front. aging Neurosci.* 14, 742408. doi:10.3389/fnagi.2022.742408

Avram, S., Bologa, C. G., Holmes, J., Bocci, G., Wilson, T. B., Nguyen, D.-T., et al. (2021). Drugcentral 2021 supports drug discovery and repositioning. *Nucleic acids Res.* 49, D1160–D1169. doi:10.1093/nar/gkaa997

Cui, X., Qu, X., Li, D., Yang, Y., Li, Y., and Zhang, X. (2023). Mkgcn: multi-modal knowledge graph convolutional network for music recommender systems. *Electronics* 12, 2688. doi:10.3390/electronics12122688

Gao, C.-Q., Zhou, Y.-K., Xin, X.-H., Min, H., and Du, P.-F. (2022a). Dda-skf: predicting drug–disease associations using similarity kernel fusion. *Front. Pharmacol.* 12, 784171. doi:10.3389/fphar.2021.784171

Gao, L., Cui, H., Zhang, T., Sheng, N., and Xuan, P. (2022b). Prediction of drug–disease associations by integrating common topologies of heterogeneous networks and specific topologies of subnets. *Briefings Bioinforma.* 23, bbab467. doi:10.1093/bib/bbab467

Gottlieb, A., Stein, G. Y., Ruppin, E., and Sharan, R. (2011). Predict: a method for inferring novel drug indications with application to personalized medicine. *Mol. Syst. Biol.* 7, 496. doi:10.1038/msb.2011.26

Harding, S. M. (2003). Acid reflux and asthma. *Curr. Opin. Pulm. Med.* 9, 42–45. doi:10.1097/00063198-200301000-00007

Kang, H., Hou, L., Gu, Y., Lu, X., Li, J., and Li, Q. (2023). Drug–disease association prediction with literature based multi-feature fusion. *Front. Pharmacol.* 14, 1205144. doi:10.3389/fphar.2023.1205144

Lavecchia, A. (2015). Machine-learning approaches in drug discovery: methods and applications. *Drug Discov. today* 20, 318–331. doi:10.1016/j.drudis.2014.10.012

Li, X., Zhang, Q., Zuo, Z., Yan, R., Zheng, C., and Zhang, F. (2022). "Predicting drug-disease associations by self-topological generalized matrix factorization with neighborhood constraints," in *International conference on intelligent computing* (Springer), 138–149.

Liang, X., Zhang, P., Yan, L., Fu, Y., Peng, F., Qu, L., et al. (2017). Lrssl: predict and interpret drug–disease associations based on data integration using sparse subspace learning. *Bioinformatics* 33, 1187–1196. doi:10.1093/bioinformatics/btw770

Liu, J., Guan, S., Zou, Q., Wu, H., Tiwari, P., and Ding, Y. (2024). Amdgt: attention aware multi-modal fusion using a dual graph transformer for drug–disease associations prediction. *Knowledge-Based Syst.* 284, 111329. doi:10.1016/j.knosys.2023.111329

Liu, W., Tang, T., Lu, X., Fu, X., Yang, Y., and Peng, L. (2023). Mpclcda: predicting circrna–disease associations by using automatically selected meta-path and contrastive learning. *Briefings Bioinforma.* 24, bbad227. doi:10.1093/bib/bbad227

Luo, H., Wang, J., Li, M., Luo, J., Peng, X., Wu, F.-X., et al. (2016). Drug repositioning based on comprehensive similarity measures and bi-random walk algorithm. *Bioinformatics* 32, 2664–2671. doi:10.1093/bioinformatics/btw228

Meng, Y., Jin, M., Tang, X., and Xu, J. (2021). Drug repositioning based on similarity constrained probabilistic matrix factorization: covid-19 as a case study. *Appl. soft Comput.* 103, 107135. doi:10.1016/j.asoc.2021.107135

Meng, Y., Wang, Y., Xu, J., Lu, C., Tang, X., Peng, T., et al. (2024). Drug repositioning based on weighted local information augmented graph neural network. *Briefings Bioinforma.* 25, bbad431. doi:10.1093/bib/bbad431

Niu, D., Xu, L., Pan, S., Xia, L., and Li, Z. (2024). Srr-ddi: a drug–drug interaction prediction model with substructure refined representation learning based on self-attention mechanism. *Knowledge-Based Syst.* 285, 111337. doi:10.1016/j.knosys.2023.111337

Pushpakom, S., Iorio, F., Eyers, P. A., Escott, K. J., Hopper, S., Wells, A., et al. (2019). Drug repurposing: progress, challenges and recommendations. *Nat. Rev. Drug Discov.* 18, 41–58. doi:10.1038/nrd.2018.168

Sharma, D., Subbarao, G., and Saxena, R. (2017). Hepatoblastoma. *Seminars diagnostic pathology Elsevier* 34, 192–200. doi:10.1053/j.semdp.2016.12.015

Singh, S., Gupta, H., Sharma, P., and Sahi, S. (2024). Advances in artificial intelligence (ai)-assisted approaches in drug screening. *Artif. Intell. Chem.* 2, 100039. doi:10.1016/j.aichem.2023.100039

Tang, X., Zhou, C., Lu, C., Meng, Y., Xu, J., Hu, X., et al. (2023). Enhancing drug repositioning through local interactive learning with bilinear attention networks. *IEEE J. Biomed. Health Inf.* 29, 1644–1655. doi:10.1109/JBHI.2023.3335275

Toskala, E., and Kennedy, D. W. (2015). Asthma risk factors. *Int. forum allergy and rhinology Wiley Online Libr.* 5, S11–S16. doi:10.1002/alr.21557

Trott, O., and Olson, A. J. (2010). Autodock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* 31, 455–461. doi:10.1002/jcc.21334

Vaz, E. S., Vassiliades, S. V., Giarolla, J., Polli, M. C., and Parise-Filho, R. (2023). Drug repositioning in the covid-19 pandemic: fundamentals, synthetic routes, and overview of clinical studies. *Eur. J. Clin. Pharmacol.* 79, 723–751. doi:10.1007/s00228-023-03486-4

Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., and Bengio, Y. (2017). Graph attention networks. *arXiv Prepr. arXiv:1710.10903*. doi:10.48550/arXiv.1710.10903

Wang, Y., Meng, Y., Zhou, C., Tang, X., Zeng, P., Pan, C., et al. (2025). Automatic collaborative learning for drug repositioning. *Eng. Appl. Artif. Intell.* 139, 109653. doi:10.1016/j.engappai.2024.109653

Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., and Philip, S. Y. (2020). A comprehensive survey on graph neural networks. *IEEE Trans. neural Netw. Learn. Syst.* 32, 4–24. doi:10.1109/tnnls.2020.2978386

Yang, B., and Chen, H. (2023). Predicting circrna-drug sensitivity associations by learning multimodal networks using graph auto-encoders and attention mechanism. *Briefings Bioinforma.* 24, bbac596. doi:10.1093/bib/bbac596

Yang, H., Ding, Y., Tang, J., and Guo, F. (2021). Drug–disease associations prediction via multiple kernel-based dual graph regularized least squares. *Appl. Soft Comput.* 112, 107811. doi:10.1016/j.asoc.2021.107811

Yang, R., Fu, Y., Zhang, Q., and Zhang, L. (2024). Gcngat: drug–disease association prediction based on graph convolution neural network and graph attention network. *Artif. Intell. Med.* 150, 102805. doi:10.1016/j.artmed.2024.102805

Zeng, P., Zhang, B., Liu, A., Meng, Y., Tang, X., Yang, J., et al. (2024). Drug repositioning based on tripartite cross-network embedding and graph convolutional network. *Expert Syst. Appl.* 252, 124152. doi:10.1016/j.eswa.2024.124152

Zhang, W., Xu, H., Li, X., Gao, Q., and Wang, L. (2020). Drimc: an improved drug repositioning approach using bayesian inductive matrix completion. *Bioinformatics* 36, 2839–2847. doi:10.1093/bioinformatics/btaa062

Zhang, X.-M., Liang, L., Liu, L., and Tang, M.-J. (2021). Graph neural networks and their current applications in bioinformatics. *Front. Genet.* 12, 690049. doi:10.3389/fgene.2021.690049

Zhao, B.-W., Hu, L., You, Z.-H., Wang, L., and Su, X.-R. (2022a). Hingrl: predicting drug–disease associations with graph representation learning on heterogeneous information networks. *Briefings Bioinforma.* 23, bbab515. doi:10.1093/bib/bbab515

Zhao, B.-W., Su, X.-R., Hu, P.-W., Ma, Y.-P., Zhou, X., and Hu, L. (2022b). A geometric deep learning framework for drug repositioning over heterogeneous information networks. *Briefings Bioinforma.* 23, bbac384. doi:10.1093/bib/bbac384

Zhao, B.-W., Su, X.-R., Yang, Y., Li, D.-X., Li, G.-D., Hu, P.-W., et al. (2025). Regulation-aware graph learning for drug repositioning over heterogeneous biological network. *Inf. Sci.* 686, 121360. doi:10.1016/j.ins.2024.121360