

## Methods

journal homepage: [www.elsevier.com/locate/ymeth](http://www.elsevier.com/locate/ymeth)



# Improving anti-cancer drug response prediction using multi-task learning on graph convolutional networks

Hancheng Liu<sup>a</sup>, Wei Peng<sup>a,b,\*</sup>, Wei Dai<sup>a,b</sup>, Jiangzhen Lin<sup>a</sup>, Xiaodong Fu<sup>a,b</sup>, Li Liu<sup>a,b</sup>, Lijun Liu<sup>a,b</sup>, Ning Yu<sup>c</sup>

<sup>a</sup> Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650050, China

<sup>b</sup> Computer Technology Application Key Lab of Yunnan Province, Kunming University of Science and Technology, Kunming 650050, China

<sup>c</sup> State University of New York, The College at Brockport, Department of Computing Sciences, 350 New Campus Drive, Brockport NY 14422

## ARTICLE INFO

### Keywords:

Anti-cancer drug response  
Graph convolutional neural network  
Multi-task learning

## ABSTRACT

Predicting the therapeutic effect of anti-cancer drugs on tumors based on the characteristics of tumors and patients is one of the important contents of precision oncology. Existing computational methods regard the drug response prediction problem as a classification or regression task. However, few of them consider leveraging the relationship between the two tasks. In this work, we propose a Multi-task Interaction Graph Convolutional Network (MTIGCN) for anti-cancer drug response prediction. MTIGCN first utilizes an graph convolutional network-based model to produce embeddings for both cell lines and drugs. After that, the model employs multi-task learning to predict anti-cancer drug response, which involves training the model on three different tasks simultaneously: the main task of the drug sensitive or resistant classification task and the two auxiliary tasks of regression prediction and similarity network reconstruction. By sharing parameters and optimizing the losses of different tasks simultaneously, MTIGCN enhances the feature representation and reduces overfitting. The results of the experiments on two in vitro datasets demonstrated that MTIGCN outperformed seven state-of-the-art baseline methods. Moreover, the well-trained model on the in vitro dataset GDSC exhibited good performance when applied to predict drug responses in in vivo datasets PDX and TCGA. The case study confirmed the model's ability to discover unknown drug responses in cell lines.

## 1. Introduction

Predicting the therapeutic effect of anti-cancer drugs on tumors based on the characteristics of tumors and patients is one of the important contents of precision oncology [1]. Anti-cancer drug response prediction can help clinicians choose the most suitable personalized treatment plan for patients, improve treatment outcomes, reduce adverse reactions and costs [2,3]. Because tumors are heterogeneous, their responses to drugs are influenced by many factors, including the genomic, transcriptomic, proteomic, metabolomic and other characteristics of the tumors. Additionally, the complexity of the tumor micro-environment, including the influence of the immune system and microbiome, adds another layer of complexity to drug response prediction [4]. Therefore, predicting the therapeutic effect of anti-cancer drugs on tumors is a complex and challenging task in precision medicine.

The emergence of high-throughput technologies has facilitated the

generation of large-scale anti-cancer drug response data, such as Genomics of Drug Sensitivity in Cancer(GDSC) [5], Cancer Cell Line Encyclopedia(CCLE) [6], The Cancer Therapeutics Response Portal (CTRP) [7], which provide a wealth of information for designing computational methods such as machine learning and deep learning to predict anti-cancer drug response [1,8]. The drug response prediction problem is usually regarded as a classification or regression task. The classification task divides the drug response into two categories: sensitive and resistant, while the regression task predicts the quantitative IC50 sensitivity score of the anti-cancer drug response. These methods usually establish the association between the input features of drugs and cell lines (such as chemical structure, gene expression, mutation, etc.) and the output of drug response (such as half-maximal inhibitory concentration IC50, sensitive/resistant classification, etc.), and evaluate the performance of the model through training and testing datasets.

In recent years, various types of methods have been proposed for anti-cancer drug response prediction problems, and we divide them into

\* Corresponding author at: Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650050, China.  
E-mail addresses: [weipeng1980@gmail.com](mailto:weipeng1980@gmail.com) (W. Peng), [daiwei@kust.edu.cn](mailto:daiwei@kust.edu.cn) (W. Dai), [ieall@kmust.edu.cn](mailto:ieall@kmust.edu.cn) (L. Liu), [nyu@brockport.edu](mailto:nyu@brockport.edu) (N. Yu).

three categories according to the time sequence: regression model, classification model, and link prediction model. Early regression models, such as ridge regression [9], LASSO [10] and elastic net [11], infer regression values such as IC50 concentration between cell line expression profile and drug response. These regression methods can quickly obtain accurate and continuous prediction results but are not very sensitive to nonlinear relationships. Classification models leverage deep learning models to extract cell line and drug features and then combine these features to implement predictions. Deep neural network(DNN) [12], convolutional neural network (CNN) [13], AutoEncoder(AE) [14], and attention mechanism [15] are popular deep learning models that have been employed to encode cell line genomic features (such as expression profile, mutation status and copy number variation). For drug encoding, DNN and CNN are adopted to process molecular fingerprints and descriptors. CNN [13] encodes SMILES strings to obtain drug features [13]. Recently, drugs usually are described as molecular structure graphs, where nodes are atoms and edges denote their join keys. Graph neural network (GNN) models run on the molecular graph to learn drug feature embeddings [16]. After obtaining cell line and drug features, popular classifiers, such as support vector machine (SVM), random forest, DNN or CNN, are employed to combine these features for IC50 regression prediction or classification [14,16–18]. However, classification models often ignore the complex relationship between drugs and cell lines. Previous work found that similar drugs exhibit similar responses to similar cell lines and vice versa. Hence, some studies build heterogeneous networks [19,20], where nodes include drugs and cell lines, and edges include the known reactions between drugs and cell lines. Then they regard drug response prediction as a link prediction problem. For example, Zhang et al. [21] construct a heterogeneous network containing cell lines, drugs and target genes and their connections and use an information flow-based algorithm to infer the potential response of drugs to cell lines. This method can use multi-source heterogeneous data to enhance prediction ability, but it also requires much prior knowledge to construct a reasonable network structure. Wang et al. [22] consider the similarity among cell lines, drugs and targets and employ the similarity regularized matrix factorization (SRMF) method to decompose the known drug-cell line association into drug features and cell line features and implement association reconstruction to finish predictions. Peng et al. [23] fuse cell line multi-omics data and drug chemical structure data, constructed a cell line-drug heterogeneous network, and update cell line and drug features with graph convolution operation that can both consider node network structures and node features in the feature learning process. Liu et al. [24] design a drug response prediction method based on GCN and contrastive learning framework to enhance the difference between positive and negative samples and improve the model's generalization ability.

Our previous work, NIHGCN, focused on predicting the continuous values of half-maximal inhibitory concentration (IC50) and classifying the cell-drug responses using a graph convolutional network (GCN) and neighborhood information (NI) layers [25]. NIHGCN achieved good performance in these tasks. It and other previous methods treated them separately without considering their high correlation. In reality, cell lines and drugs with higher response concentrations tend to exhibit higher resistance between them. Therefore, it is beneficial to leverage the relationship between these two tasks. To address this, we propose a Multi-task Interaction Graph Convolutional Network (MTIGCN) for anti-cancer drug response prediction. It builds upon the neighborhood interaction GCN and incorporates an auxiliary task A, which focuses on cell line and drug IC50 regression prediction, in addition to the main task of cell line and drug binary classification. This multi-task learning (MTL) approach improves the model's generalization performance by sharing feature representations between the related tasks [26]. Moreover, the learned embeddings should preserve the intrinsic structure of drug and cell line similarity. To achieve this, the model MTIGCN adds an auxiliary task B, which involves reconstructing the cell line/drug

similarity network. During the training process, the model parameters are learned by simultaneously minimizing classification loss, regression loss and similarity network reconstruction loss. Extensive experiments were conducted to evaluate the effectiveness of the MTIGCN model on GDSC and CCLE datasets. The results demonstrated that MTIGCN outperformed state-of-the-art algorithms. Furthermore, when trained on the GDSC dataset, MTIGCN successfully predicted PDX and TCGA samples, indicating its portability of transferring from in vitro cell lines to in vivo datasets.

## 2. Materials

### 2.1. In vitro datasets

We test our model on two vitro datasets from the GDSC (Genomics of Drug Sensitivity in Cancer) and CCLE (Cancer Cell Line Encyclopedia) databases. GDSC Database provides two tables, TableS4A<sup>1</sup> and TableS5C<sup>2</sup> to help us infer drug sensitivity and resistance status. TableS4A Contains logarithmized half-maximal inhibitory concentration (IC50) values for various cell line/drug combinations, including 990 cancer cell lines and 265 tested drugs. TableS5C records the sensitivity thresholds of 265 drugs. IC50 values in TableS4A are compared to sensitivity thresholds recorded in TableS5C to determine drug sensitivity or resistance status. CCLE database provides 11,670 records of cell line-drug trials. Each record reports experimental information, such as drug target, dose, log(IC50) and effective area. Similar to previous methods [27], drug response is determined by comparing the z-score normalized log(IC50) value with a predefined threshold ( $-0.8$  in this case).

Our method involves cell line gene expression features and drug substructure fingerprint features. Drug substructure fingerprints were obtained from PubChem [28] database. Gene expression data from GDSC and CCLE databases were preprocessed [14] and normalized using the RMA [29] method, log-transformed and aggregated to the level of genes. The compound ID (CID) of the target compound is determined using the compound name from the PubChem website. PubChemPy package is used to retrieve the drug substructure fingerprint features. Cell lines without histological data and drugs with the same compound ID (CID) are excluded from the analysis. After preprocessing, the GDSC database comprises 962 cell lines and 228 drugs. Similarly, the CCLE database comprises 436 cancer cell lines and 24 drugs (see Table 1).

### 2.2. In vivo datasets

A key challenge in drug response research is clinical efficacy, i.e., whether the research results can be translated into actual patients [9]. This work focused on transferring cancer drug response from cell lines (in vitro data) to two different types of in vivo data: patient-derived xenografts (PDX) and patient tumors from The Cancer Genome Atlas (TCGA) dataset.

**Table 1**  
Statistics of in vitro datasets.

Dataset	Number of drugs	Number of cell lines	Sensitive	Resistant	Total
GDSC	228	962	20,851	156,512	177,363
CCLE	24	436	1696	8768	10,464

<sup>1</sup> [https://www.cancerrxgene.org/gdsc1000/GDSC1000\\_WebResources//Data/suppData/TableS4A.xlsx](https://www.cancerrxgene.org/gdsc1000/GDSC1000_WebResources//Data/suppData/TableS4A.xlsx).

<sup>2</sup> [https://www.cancerrxgene.org/gdsc1000/GDSC1000\\_WebResources//Data/suppData/TableS5C.xlsx](https://www.cancerrxgene.org/gdsc1000/GDSC1000_WebResources//Data/suppData/TableS5C.xlsx).

We retrieved the gene expression profiles and drug response values from the supplementary files of Gao et al. [30] for the PDX dataset and Ding et al. [31] for the TCGA dataset. To ensure consistency and comparability, the gene expression values were preprocessed by converting them to transcripts per million (TPM) and log-transforming them, following the approach described in Sharifi-Noghabi et al. [14]. We retrieved 6 drugs shared by GDSC and PDX. PDX drug response was divided into two groups, sensitive (“CR” and “PR”) and resistant (“SD” and “PD”). We obtained 22 drugs shared by GDSC and TCGA. TCGA drug response was divided into two groups, sensitive (“complete remission” and “partial remission”) and resistant (“stable disease” and “progressive disease”). Table 2 summarizes the in vivo datasets used in the study. There are 191 response records of 6 drugs from the PDX dataset and 430 response records of 22 drugs from the TCGA dataset.

### 3. Method

#### 3.1. Overview

Fig. 1 illustrates the flowchart of predicting anti-cancer drug response using interactive graph convolutional network with multi-task learning (MTIGCN). The algorithm comprises three main parts: the main task (binary classification), auxiliary task A (IC50 regression) and auxiliary task B (similarity reconstruction).

The model takes as input feature matrices of cell lines, drugs, and responses. Then it utilizes a graph convolutional network-based model called NIHGCN to aggregate information from the node features. This step produces embeddings for both cell lines and drugs. After that, the model employs multi-task learning to predict anti-cancer drug response, which involves training the model on three different tasks simultaneously:

- Main Task (Binary Classification): The main task focuses on binary classification, aiming to predict whether a drug response will be sensitive or resistant.
- Auxiliary Task A (IC50 Regression): Auxiliary task A involves IC50 regression, which predicts the half-maximal inhibitory concentration (IC50) values for the drug response.
- Auxiliary Task B (Similarity Reconstruction): Auxiliary task B is concerned with similarity network reconstruction, aiming to reconstruct the similarity relationships between drugs.

#### 3.2. NIHGCN model learning cell line and drug features

The model takes as input feature matrices of cell lines, drugs, and responses. Then it utilizes NIHGCN [25] model to learn cell line and drug feature representations from the bipartite heterogeneous network.

Let  $G = (A \in \{0, 1\}^{m \times n}, C \in R^{m \times h}, D \in R^{n \times h})$  represent the bipartite heterogeneous network that captures the relationship between cell lines and drugs.  $A$  is the network adjacency matrix, where rows correspond to cell lines, and columns correspond to drugs. A value of 1 in matrix  $A$  indicates that the cell line is sensitive to the drug.  $C$  and  $D$  are the attribute matrices of cell line nodes and drug nodes in the network, obtained by Eqs. (1) and (2):

**Table 2**  
Statistics of in vivo datasets.

Dataset	Number of drugs	Number of tumors/ samples	Sensitive	Resistant	Total
PDX	6 (shared with GDSC)	118	24	167	191
TCGA	22(shared with GDSC)	403	201	229	430

$$C = X_c \bullet \theta_c \quad (1)$$

$$D = X_d \bullet \theta_d \quad (2)$$

where  $X_c \in R^{m \times h_c}$  is the z-score normalized cell line gene expression matrix over all cell lines [25],  $m$  is the number of cell lines,  $\theta_c \in R^{h_c \times h}$  represents the parameter set of cell line linear transformation.  $X_d \in R^{n \times h_d}$  is the drug molecular fingerprint matrix,  $n$  is the number of drugs,  $\theta_d \in R^{h_d \times h}$  represents the drug linear transformation parameter set.

After preparing the network and node attributes, NIHGCN takes the interaction module to learn feature representations for drugs and cell lines. The interaction module consists of a parallel graph convolution network (GCN) layer and a neighborhood interaction (NI) layer, aggregating features from neighbors at the node and element level. In the parallel GCN layer, NIHGCN implemented two parallel graph convolution operations on the bipartite heterogeneous network and independently aggregated node-wise features from neighbors for the cell lines and drugs. In the NI layer, NIGCN multiplies the elements of target node and its neighbor nodes to capture fine-grained neighbor features. Mathematically, Eqs. (3) and (4) produce the feature representations of cell lines and drugs.

$$H_c^{(k)} = \sigma((1 - \alpha)(SC + \mathcal{L}_c H_d^{(k-1)})W_c^{(k)} + \alpha(SC + \mathcal{L}_c H_d^{(k-1)}) \odot H_c^{(k-1)} W_c^{(k)}) \quad (3)$$

$$H_d^{(k)} = \sigma((1 - \alpha)(SD + \mathcal{L}_d H_c^{(k-1)})W_d^{(k)} + \alpha(SD + \mathcal{L}_d H_c^{(k-1)}) \odot H_d^{(k-1)} W_d^{(k)}) \quad (4)$$

where  $\mathcal{L}_c = D_c^{-\frac{1}{2}} A D_d^{-\frac{1}{2}}$  and  $\mathcal{L}_d = D_d^{-\frac{1}{2}} A^T D_c^{-\frac{1}{2}}$  are the Laplace transform of the adjacency matrix for cell line and drug, respectively.  $D_{c(ij)} = \sum_j A_{ij} + 1$  and  $D_{d(ij)} = \sum_i A_{ji} + 1$ . Considering the features of the nodes themselves, we introduce cell line self-features  $SC = (D_c^{-1} + I_m)H_c^{(k-1)}$  and drug self-features  $SD = (D_d^{-1} + I_n)H_d^{(k-1)}$ .  $\sigma$  is the ReLU activation function,  $\alpha$  is a hyperparameter to balance the contributions of GCN layer and NI layer.  $H_c^{(k)} \in R^{m \times f}$  and  $H_d^{(k)} \in R^{n \times f}$  are the final representations of cell lines and drugs obtained after  $k$  steps of embedding propagation.  $H_c^{(0)} \in R^{m \times h} = C$  and  $H_d^{(0)} \in R^{n \times h} = D$  are the initial attributes of cell lines and drugs.  $W_c^{(k)} \in R^{h \times f}$  and  $W_d^{(k)} \in R^{h \times f}$  are the weight parameters of cell line and drug aggregators. We use different weight matrices to aggregate features for cell lines and drugs in the GCN models.

#### 3.3. Main task: binary classification

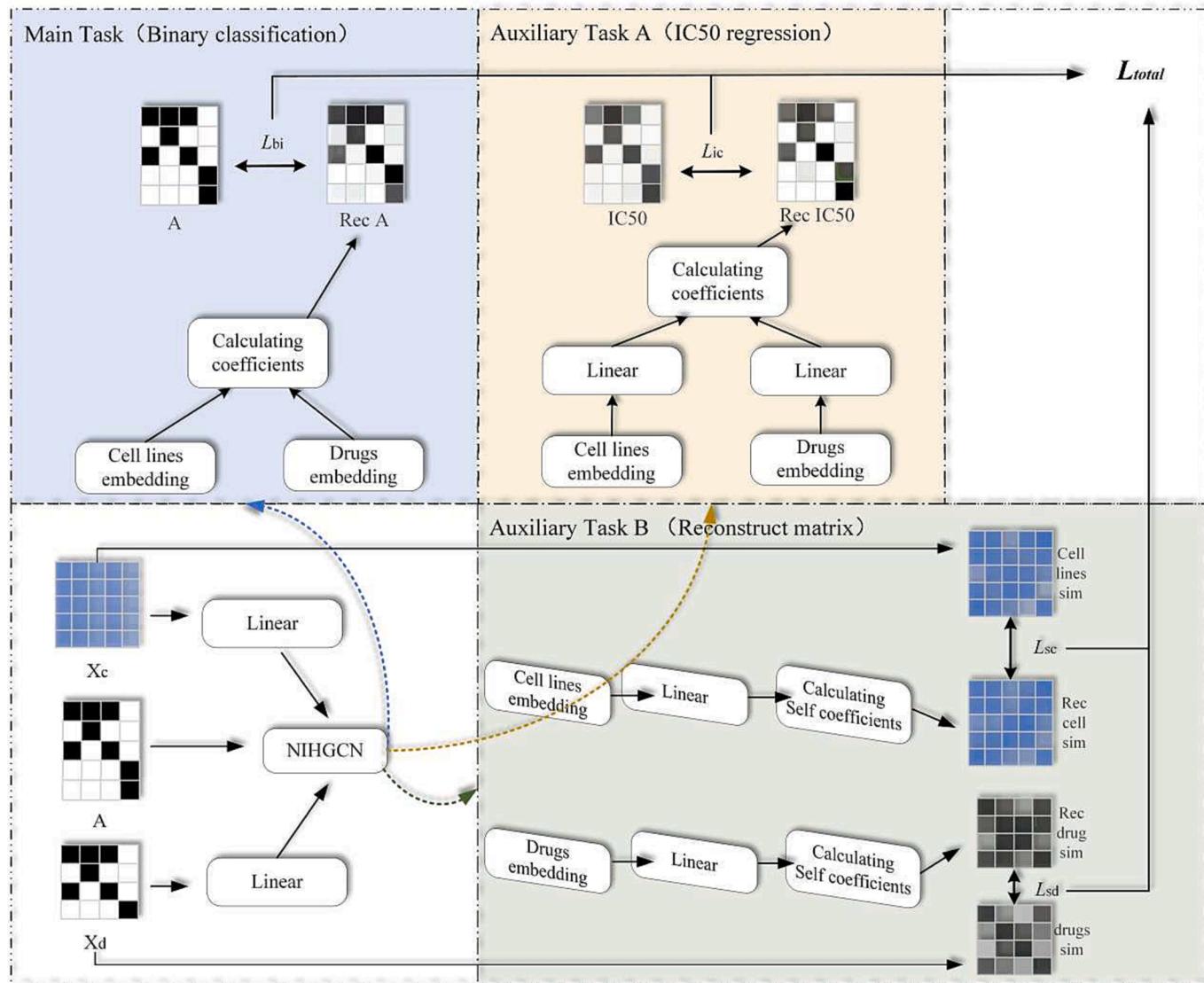
The main task of our method aims to predict whether a drug response will be sensitive or resistant. In this task, the normalized feature matrix  $X_c$  of cell lines, the drug feature matrix  $X_d$  and the binary response matrix  $A$  are input to the NIHGCN model to obtain the cell line and drug feature embeddings,  $H_c^{(k)}$  and  $H_d^{(k)}$ :

$$H_c^{(k)} = NIHGCN^{(k)}(X_c, X_d, A) \quad (5)$$

$$H_d^{(k)} = NIHGCN^{(k)}(X_d, X_c, A^T) \quad (6)$$

we calculate the reactions between drugs and cell lines using the linear correlation coefficient of cell line embeddings and drug embeddings, which is defined as follows:

$$\text{Corr}(h_i, h_j) = \frac{(h_i - \mu_i)(h_j - \mu_j)^T}{\sqrt{(h_i - \mu_i)(h_i - \mu_i)^T} \sqrt{(h_j - \mu_j)(h_j - \mu_j)^T}} \quad (7)$$



**Fig. 1.** Framework diagram of MTIGCN algorithm.

where  $h_i \in H_c^{(k)}$  and  $h_j \in H_d^{(k)}$  are the feature representation vectors of the i-th cell line and the j-th drug, respectively,  $\mu_i$  and  $\mu_j$  are the mean values of  $h_i$  and  $h_j$ , respectively. Finally, the cell line-drug response matrix is reconstructed as:

$$\hat{A} = \varphi\left(\text{Corr}\left(H_c^{(k)}, H_d^{(k)}\right)\right) \quad (8)$$

where  $\varphi$  is the sigmoid activation function. Since the main task focuses on binary classification, we use binary cross entropy as the drug response prediction loss:

$$\ell_b(A, \hat{A}) = -\frac{1}{m \times n} \sum_{ij} M_{ij} [A_{ij} \ln(\hat{A}_{ij}) + (1 - A_{ij}) \ln(1 - \hat{A}_{ij})] \quad (9)$$

where m and n represent the number of cell lines and drugs. The indicator matrix M has dimensions  $m \times n$ , where each element  $M_{ij}$  represents whether the association between the i-th cell line and the j-th drug is in the training set.

### 3.4. Auxiliary task (A): regression prediction

Previous studies have achieved good results on two specific tasks:

predicting cell line and drug binary response (sensitivity or resistance) and predicting cell line and drug response concentration (IC50 value). However, these models were trained solely on single task indicators, neglecting the potential benefits of jointly training these correlated tasks. Drug-sensitive cell lines usually have a lower response concentration. Hence, incorporating auxiliary task A, which involves predicting the drug response IC50 value, is a natural choice to complement the primary task of binary response prediction.

We use output of the model NIHGCN shared with the main task to get the embedding of cell line  $H_c^{(k)}$  and drug  $H_d^{(k)}$  as input of the IC50 regression task, and apply a linear transformation layer separately to get the final cell line and drug embedding representation,  $H_{rc}$  and  $H_{rd}$ :

$$H_{rc} = W_{rc} H_c^{(k)} + b_c \quad (10)$$

$$H_{rd} = W_{rd} H_d^{(k)} + b_d \quad (11)$$

we estimate the concentration (IC50 value) of drug response to cell lines by calculating the linear correlation coefficient between the cell line embedding  $H_{rc}$  and the drug embedding  $H_{rd}$ . Finally, the cell line-drug IC50 response matrix  $A_r$  is reconstructed as:

$$\hat{A}_r = \text{Corr}(H_{rc}, H_{rd}) \quad (12)$$

where the function  $\text{Corr}()$  represents the linear correlation coefficient. Since this task is a regression task, we use mean squared error as the auxiliary task A loss:

$$\ell_r(A_r, \hat{A}_r) = -\frac{1}{m \times n} \sum_{i,j} M_{ij} (A_{r(ij)} - \hat{A}_{r(ij)})^2 \quad (13)$$

where m and n represent the number of cell lines and drugs, respectively. M is an indicator matrix, which helps identify the specific cell line-drug associations used for training.

### 3.5. Auxiliary task (B): similarity reconstruction

The intrinsic structure of drug similarity and cell line similarity should be preserved in the embeddings learned from the network, so we added auxiliary task B to reconstruct the cell line/drug similarity network. Based on the gene expression profile of the cell lines, we measure the similarity between cell lines as follows:

$$S_c(i, j) = e^{-\frac{\|c_i - c_j\|}{2\epsilon^2}} \quad (14)$$

If there are m cell lines, then  $S_c \in R^{m \times m}$ ,  $C_i$  represents the gene expression of the i-th cell line,  $\epsilon$  represents the regularization constant. We take  $\epsilon = 3$  in the experiment. The molecular fingerprint is often used as a drug feature descriptor, which encodes the characteristics of small drug molecules as a binary bit string of length 920. We evaluate their similarity based on the Jaccard coefficient of the molecular fingerprints of drugs as follows.

$$S_d(i, j) = \frac{|d_i \cap d_j|}{|d_i \cup d_j|} \quad (15)$$

If there are n drugs, then  $S_d \in R^{n \times n}$ , D represents the molecular fingerprint feature of the drug. In the similarity reconstruction task, we obtain the cell line embedding  $H_c^{(k)}$  and drug embedding  $H_d^{(k)}$  through the model NIHGCN shared with the main and auxiliary tasks, and add a separate linear transformation layer to obtain the final embedding representation:

$$H_{sc} = W_{sc} H_c^{(k)} + b_{sc} \quad (16)$$

$$H_{sd} = W_{sd} H_d^{(k)} + b_{sd} \quad (17)$$

we estimate the cell line similarity or drug similarity by calculating the linear correlation coefficient between the cell line embedding  $H_{sc}$  or drug embedding  $H_{sd}$ . Finally, the cell line similarity matrix  $\hat{S}_c$  and drug similarity matrix  $\hat{S}_d$  are reconstructed as:

$$\hat{S}_c = \text{Corr}(H_{sc}, H_{sc}) \quad (18)$$

$$\hat{S}_d = \text{Corr}(H_{sd}, H_{sd}) \quad (19)$$

where  $\text{Corr}()$  represents the linear correlation coefficient. We use mean squared error to calculate the cell line similarity reconstruction loss  $\ell_c$  and the drug similarity reconstruction loss  $\ell_d$ :

$$\ell_c(S_c, \hat{S}_c) = -\frac{1}{m \times n} \sum_{i,j} M_{ij} (S_{c(ij)} - \hat{S}_{c(ij)})^2 \quad (20)$$

$$\ell_d(S_d, \hat{S}_d) = -\frac{1}{m \times n} \sum_{i,j} M_{ij} (S_{d(ij)} - \hat{S}_{d(ij)})^2 \quad (21)$$

where m and n represent the number of cell lines and drugs, respectively. M is an indicator matrix, whose elements identify whether the specific cell line-drug associations used for training.

### 3.6. Joint loss

We obtain the multi-task joint loss by weighting the losses of the above different tasks:

$$L_{total} = \theta_b \ell_b + \theta_r \ell_r + \theta_c \ell_c + \theta_d \ell_d \quad (22)$$

where  $\theta = \{\theta_b, \theta_r, \theta_c, \theta_d\}$  controls the weights of different tasks and balances their importance. In the experiment, we used the Adam optimizer to optimize the loss function. The weight parameters for balancing tasks were selected from [0.05, 0.1, 0.15..., 0.95]. We chose the parameters that produce the highest AUC prediction results under the cross-validation. Finally, we set the embedding size and the convolution layer output size to 1024. The interaction module balance parameter in NIHGCN was set to 0.25. the parameters  $\theta$  for balancing different tasks were set to 0.75, 0.25, 0.1, and 0.25, respectively. The learning rate was 0.001. The weight decay was 1e-5, and the epoch was 2000.

## 4. Experiment

### 4.1. Baseline

To assess how our model performs in drug response prediction, we compared our method with the following baselines:

- HNMDRP [21] constructs a heterogeneous network by combining various types of information, such as gene expression profiles, drug chemical structures, drug target interactions, and protein interactions. It uses an information diffusion algorithm on the network to calculate response scores for each cell line-drug pair.
- SRMF [22] extracts drug and cell line features through a regularized matrix factorization model. Then it uses these feature vectors to reconstruct the drug-cell line response matrix and predicts unknown drug responses.
- DeepDSC [17] utilizes stacked deep autoencoders to extract cell features from gene expression data. These extracted features are combined with drug chemical features to predict drug response.
- DeepCDR [16] uses GCNs to encode drug chemical structures and combines cell line features of multi-omics data to predict drug response.
- MOFGCN [23] uses the similarity between cell lines and drugs as input features, and performs graph convolution operations on a homogeneous graph to diffuse similarity information and reveal potential connections.
- GraphCDR [24] employs the GCN model and contrastive learning to predict drug response.
- NIHGCN [25] inputs cell line and drug feature vectors into a heterogeneous network and uses an interaction module with convolutional and interactive layers to learn node-level and element-level features of cell lines and drugs, respectively.

### 4.2. Experimental design

We performed tests under five different settings to evaluate the performance of our model and the baselines.

Test 1: In this test, we assessed the ability of each model to recover known cell line-drug associations. We randomly removed a certain number of associations from the training set and used them as the test set.

Test 2: This test aimed to evaluate each model's ability to predict new cell lines or drug responses. We zeroed out either the rows or columns of the cell line-drug association matrix and used the zeroed-out rows or columns as the test set.

Test 3: This test aimed to verify whether the model can use in vitro data to train the model to predict drug responses in vivo data. We used GDSC, an in vitro dataset, as the training set and PDX and TCGA, two in

vivo datasets, as the test set.

Test 4: This test conducted ablation studies on our model to verify the performance contribution of different components.

Test 5: The purpose of this test was to perform case studies on our model to examine its ability to discover unknown drug responses in cell lines.

All methods input gene expression of cell lines and molecular fingerprints of drugs for fair comparison. For evaluation, classification tasks utilized AUC and AUPRC metrics from the ROC curves and PR curves, respectively. Regression results were evaluated using three metrics: Pearson correlation coefficient (PCC), Spearman correlation coefficient (SCC) and root mean square error (RMSE).

### 4.3. Experimental results

#### 4.3.1. Random zeroing experiment (Test 1)

In the random zeroing experiment, we divided the known cell line-drug associations into five parts and selected 1/5 of the positive samples and an equal number of negative samples as test data. The remaining positive and negative samples were used for training. Our model can output a classification label value of 1 or 0 to indicate sensitivity or resistance but also outputs a quantitative value representing IC50 concentration, which allows us to evaluate the performance of our method on both classification and regression tasks.

**Table 3** shows the average AUC and AUPRC values of each method on GDSC and CCLE datasets when classifying whether a drug is sensitive or resistant to a cell line. The bold text in the table indicates the best results. We observed that Our model MTIGCN always outperforms all baselines on both datasets for the classification of drug sensitivity or resistance in cell lines. The improved performance of MTIGCN may be attributed to the introduction of auxiliary tasks, which allows for sharing features between related tasks. By leveraging information from the auxiliary task of quantitative IC50 concentration prediction, the model gains better generalization performance on the original sensitivity/resistance classification task.

**Fig. 2** illustrates the performance of our model compared to the baseline model on the regression task. We compare the accuracy of our model and the baseline model in predicting IC50. We observed that MTIGCN had the highest Pearson's correlation coefficient (PCC) and Spearman's correlation coefficient (SCC) and the lowest root-mean-square error (RMSE) on the GDSC dataset, showing good drug response prediction performance. The performance of MTIGCN was slightly lower than that of the NIHGCN method on the CCLE dataset due to the fact that the CCLE had a smaller amount of data and the MTIGCN has higher parameters and complexity than NIHGCN, which may cause

**Table 3**

Comparison of random zeroing cross-validation performance on GDSC and CCLE datasets.

Algorithm	GDSC		CCLE	
	AUC	AUPRC	AUC	AUPRC
HNMDRP	0.7258 ± 3 × $10^{-5}$	0.7198 ± 4 × $10^{-5}$	0.7104 ± 1 × $10^{-4}$	0.6956 ± 2 × $10^{-4}$
SRMF	0.6563 ± 2 × $10^{-4}$	0.6605 ± 5 × $10^{-5}$	0.7669 ± 4 × $10^{-5}$	0.7418 ± 2 × $10^{-5}$
DeepCDR	0.7849 ± 5 × $10^{-5}$	0.7827 ± 6 × $10^{-5}$	0.8289 ± 1 × $10^{-4}$	0.8185 ± 2 × $10^{-4}$
DeepDSC	0.8118 ± 4 × $10^{-4}$	0.8311 ± 1 × $10^{-4}$	0.8594 ± 1 × $10^{-4}$	0.8607 ± 1 × $10^{-4}$
MOFGCN	0.8684 ± 7 × $10^{-6}$	0.8730 ± 1 × $10^{-5}$	0.8608 ± 1 × $10^{-4}$	0.8589 ± 1 × $10^{-4}$
GraphCDR	0.8136 ± 4 × $10^{-5}$	0.8193 ± 3 × $10^{-5}$	0.8474 ± 2 × $10^{-4}$	0.8495 ± 2 × $10^{-4}$
NIHGCN	0.8760 ± 1 × $10^{-5}$	0.8803 ± 1 × $10^{-5}$	0.8806 ± 1 × $10^{-4}$	0.8803 ± 1 × $10^{-4}$
MTIGCN	<b>0.8870 ± 6 × <math>10^{-6}</math></b>	<b>0.8907 ± 7 × <math>10^{-6}</math></b>	<b>0.8810 ± 1 × <math>10^{-4}</math></b>	<b>0.8813 ± 2 × <math>10^{-4}</math></b>

MTIGCN to overfit on some tasks, thus reducing performance. However, MTIGCN still outperforms most other methods, which suggests that in multi-task learning, related tasks can be mutually reinforcing to improve generalization.

We compared the runtime cost of our model with that of all benchmark methods on the same computer with 4 CPU cores and 16 GB RAM. The results are shown in **Table 4**, where our model takes longer to train than MOFGCN, GraphCDR, and NIHGCN, which use GCN models, but not significantly so. Adding tasks may increase the computational complexity and the number of parameters, but it also improves the performance and generalization ability of the model. In addition, we verified the convergence of the models on the GDSC and CCLE datasets, as shown in **Fig. 3**, as the model training epochs increase, the training loss decreases, and the AUC on the test set increases and reaches a higher level of stable values, which suggests that our models can converge well and avoid overfitting.

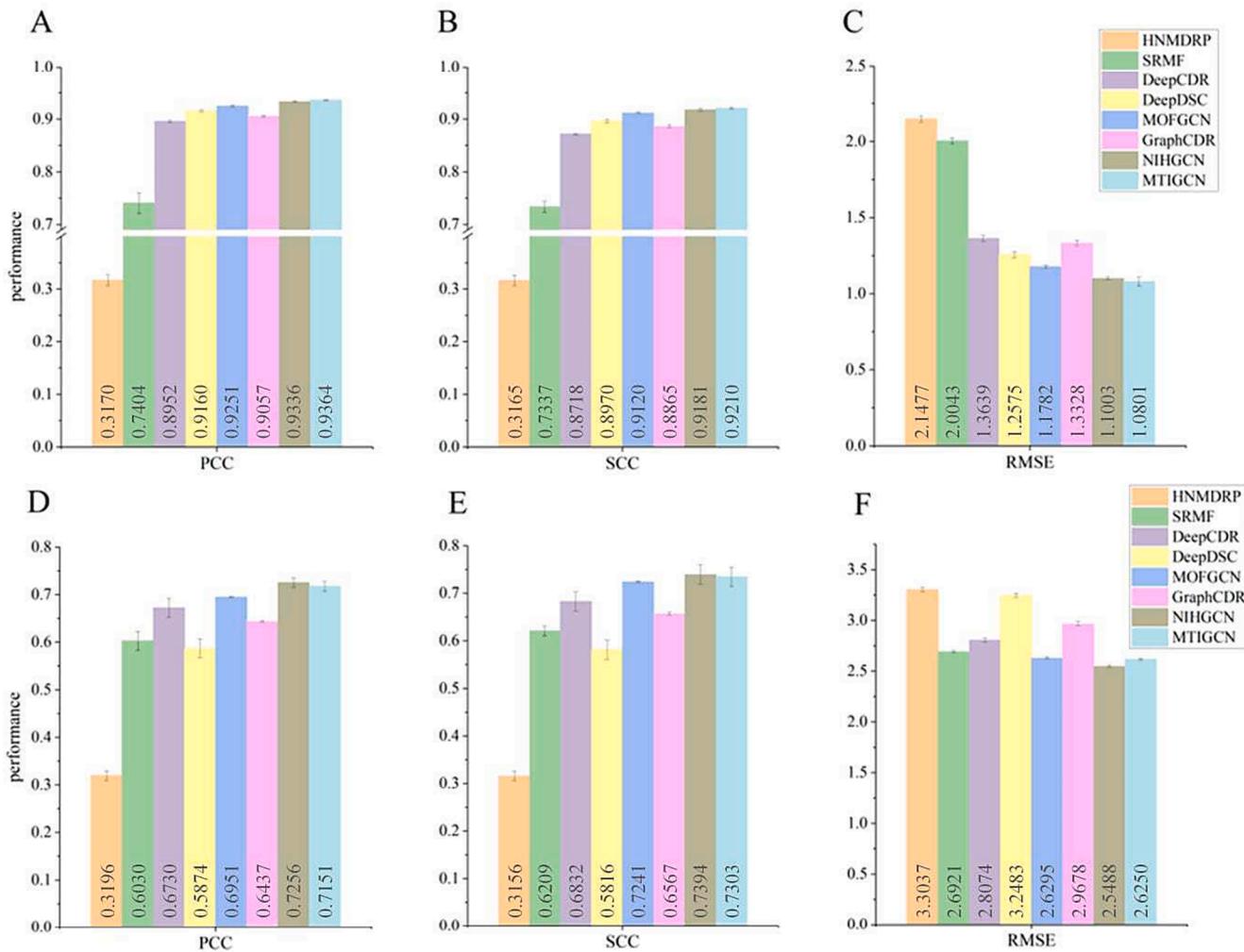
#### 4.3.2. New drug/new cell line prediction (Test 2)

To evaluate the algorithm's ability to predict new cell line or drug responses, we created a test set by removing either a row (cell line) or a column (drug) from the cell line-drug association matrix. This ensures that the test set contains cell lines or drugs that were not used during the training phase. We only used cell lines or drugs that have more than 10 positive samples to avoid extreme cases that are either too general or too specific in their responses. Therefore, we chose 658 out of 962 cell lines and 227 out of 228 drugs from GDSC dataset, and 26 out of 436 cell lines and 20 out of 24 drugs from CCLE database for experiments.

**Tables 5 and 6** show the comparison results of eight algorithms on GDSC and CCLE datasets. We observed that our model demonstrates superior predictive performance for new drug response prediction in GDSC and CCLE datasets compared to the baselines, whose AUC and AUPRC values were 0.93 % and 0.57 % higher than the second best method (NIHGCN) on GDSC dataset, and 1.06 % and 0.55 % higher than NIHGCN on CCLE dataset. For new cell line response prediction, our model MTIGCN performs the best among all comparing methods on GDSC dataset and has the second best performance on CCLE dataset, whose AUC and AUPRC values were slightly lower than NIHGCN on CCLE dataset. It may be the limited number of cell lines for testing on the CCLE dataset. Overall, our model consistently achieved the best overall performance, demonstrating that our multi-task learning framework can enhance the model's generalization and improve the prediction ability for new drugs or new cell lines.

#### 4.3.3. In vivo drug response prediction (Test 3)

This experiment aims to evaluate the transferability of the model across different datasets in drug response prediction. We trained our model and other baseline methods using GDSC in vitro dataset and then applied the models to predict drug responses of two in vivo datasets (i.e., PDX mouse xenograft and TCGA patient data). The samples in the GDSC, PDX and TCGA datasets contain different numbers of genes, so the model focuses only on the genes shared by GDSC and the other datasets (PDX and TCGA). For PDX samples, 18,942 shared genes were selected; for TCGA samples, 18,948 shared genes were used as input features. We applied the z-score normalization on both cell line (GDSC) and patient data (PDX and TCGA) to remove batch effects between the datasets. **Table 7** shows the performance comparison results between our model and baseline methods on two tasks: predicting drug-cell line responses on the PDX dataset (191 drug-cell line responses) and predicting drug-cell line responses on the TCGA dataset (430 drug-cell line responses). The multi-task learning model outperforms the baseline methods in both tasks. On the PDX dataset, the model improves 0.76 % and 4.11 % on AUC and AUPRC, respectively, compared to the best baseline method (NIHGCN). When predicting drug responses on the TCGA dataset, the model still maintains the best performance, with improvements of 3.41 % and 5.35 % on AUC and AUPRC, respectively. The results indicate that multi-task learning methods have good generalization performance and



**Fig. 2.** Random zero-cross validation results on GDSC and CCLE datasets in the regression task. (A)-(C) are Pearson correlation (PCC), Spearman correlation (SCC) and root mean square error (RMSE) on GDSC dataset, respectively. (D)-(F) are Pearson correlation (PCC), Spearman correlation (SCC) and root mean square error (RMSE) on CCLE dataset, respectively.

**Table 4**  
Comparison of running time of every method (seconds).

Dataset	HNMDRP	SRMF	DeepCDR	DeepDSC	MOFGCN	GraphCDR	NIHGCN	MTIGCN
GDSC	0.04 s	11.87 s	216.71 s	153.28 s	5.04 s	13.83 s	8.73 s	16.33 s
CCLE	0.01 s	2.56 s	11.60 s	11.19 s	4.18 s	3.48 s	3.47 s	5.45 s

transferability.

#### 4.3.4. Ablation study

MTIGCN is a multi-task learning model combining regression and similarity reconstruction tasks to improve drug response prediction. To understand the performance contribution of the multi-task learning aspect of MTIGCN, we designed four model variants for comparison under the random zeroing experiment.

**MTIGCN:** This is the complete MTIGCN model, including regression prediction and similarity reconstruction tasks. It combines these two tasks to improve drug response prediction.

**MTIGCN-A:** This variant involves training the model solely on the auxiliary task B (similarity reconstruction). The auxiliary task A (regression prediction) is excluded from the training process.

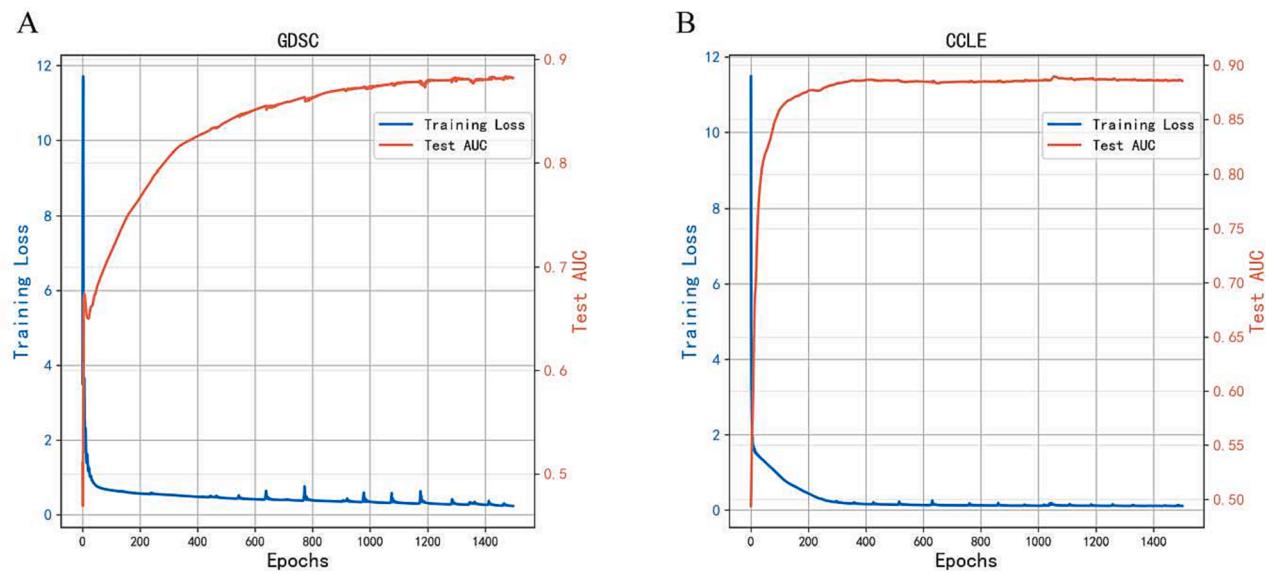
**MTIGCN-B:** This variant involves training the model solely on the auxiliary task A (regression prediction task). The auxiliary task B (similarity reconstruction task) is excluded from the training process.

**MTIGCN-AB:** This variant removes both auxiliary task A and auxiliary task B.

Based on the analysis provided in Table 8, we can draw the following conclusions:

Removing the auxiliary task A, which involves IC50 regression prediction, reduces in the overall performance of MTIGCN. This suggests that the IC50 regression task helps the model learn important feature representations that contribute to better predictions of drug response outcomes.

On both datasets (GDSC and CCLE), the MTIGCN model without the similarity reconstruction task (Task B) exhibits lower predictive performance in AUC and AUPRC values than the original MTIGCN model. The similarity reconstruction task plays a crucial role in preserving the relationships between the original features during the learning process. By maintaining the similarity information, the model can better understand the underlying structure of the data and make more accurate predictions.



**Fig. 3.** Training loss & Test AUC curves. (A) and (B) are the training loss and test AUC values for different training epochs under GDSC and CCLE datasets, respectively.

**Table 5**

Comparison of prediction performance of new cell line or new drug response on GDSC dataset.

Algorithm	New cell lines		New drugs	
	AUC	AUPRC	AUC	AUPRC
HNMDRP	–	–	$0.6951 \pm 1 \times 10^{-2}$	$0.6935 \pm 1 \times 10^{-2}$
SRMF	$0.5807 \pm 1 \times 10^{-2}$	$0.6153 \pm 1 \times 10^{-2}$	$0.6683 \pm 6 \times 10^{-3}$	$0.6757 \pm 6 \times 10^{-3}$
DeepCDR	$0.7526 \pm 8 \times 10^{-3}$	$0.7664 \pm 8 \times 10^{-3}$	$0.7605 \pm 9 \times 10^{-3}$	$0.7565 \pm 1 \times 10^{-2}$
DeepDSC	$0.7831 \pm 8 \times 10^{-3}$	$0.7994 \pm 8 \times 10^{-3}$	$0.7472 \pm 1 \times 10^{-2}$	$0.7514 \pm 1 \times 10^{-2}$
MOFGCN	$0.7190 \pm 5 \times 10^{-3}$	$0.7366 \pm 5 \times 10^{-3}$	$0.7601 \pm 7 \times 10^{-3}$	$0.7558 \pm 8 \times 10^{-3}$
GraphCDR	$0.7122 \pm 9 \times 10^{-3}$	$0.7061 \pm 9 \times 10^{-3}$	$0.7614 \pm 8 \times 10^{-3}$	$0.7501 \pm 9 \times 10^{-3}$
NIHGNCN	$0.8267 \pm 7 \times 10^{-3}$	$0.8346 \pm 8 \times 10^{-3}$	$0.7927 \pm 6 \times 10^{-3}$	$0.7877 \pm 6 \times 10^{-3}$
MTIGCN	$0.8289 \pm 7 \times 10^{-3}$	$0.8357 \pm 7 \times 10^{-3}$	$0.8020 \pm 6 \times 10^{-3}$	$0.7934 \pm 7 \times 10^{-3}$

**Table 6**

Comparison of prediction performance of new cell line or new drug response on CCLE dataset.

Algorithm	New cell lines		New drugs	
	AUC	AUPRC	AUC	AUPRC
HNMDRP	–	–	$0.6947 \pm 6 \times 10^{-3}$	$0.6871 \pm 6 \times 10^{-3}$
SRMF	$0.6138 \pm 8 \times 10^{-3}$	$0.6187 \pm 1 \times 10^{-2}$	$0.4873 \pm 9 \times 10^{-3}$	$0.5288 \pm 5 \times 10^{-3}$
DeepCDR	$0.8830 \pm 6 \times 10^{-3}$	$0.8913 \pm 6 \times 10^{-3}$	$0.7389 \pm 4 \times 10^{-3}$	$0.7300 \pm 5 \times 10^{-3}$
DeepDSC	$0.8935 \pm 4 \times 10^{-3}$	$0.9073 \pm 4 \times 10^{-2}$	$0.7315 \pm 4 \times 10^{-3}$	$0.7295 \pm 6 \times 10^{-3}$
MOFGCN	$0.8108 \pm 8 \times 10^{-3}$	$0.8137 \pm 8 \times 10^{-3}$	$0.7215 \pm 2 \times 10^{-3}$	$0.7113 \pm 2 \times 10^{-3}$
GraphCDR	$0.7613 \pm 1 \times 10^{-2}$	$0.7694 \pm 1 \times 10^{-2}$	$0.7506 \pm 4 \times 10^{-3}$	$0.7280 \pm 5 \times 10^{-3}$
NIHGNCN	$0.9084 \pm 3 \times 10^{-3}$	$0.9186 \pm 3 \times 10^{-3}$	$0.7620 \pm 5 \times 10^{-3}$	$0.7483 \pm 5 \times 10^{-3}$
MTIGCN	$0.9045 \pm 3 \times 10^{-3}$	$0.9158 \pm 3 \times 10^{-3}$	$0.7726 \pm 5 \times 10^{-3}$	$0.7538 \pm 7 \times 10^{-3}$

**Table 7**

Comparison of drug response prediction performance on PDX and TCGA datasets.

Algorithm	PDX		TCGA	
	AUC	AUPRC	AUC	AUPRC
HNMDRP	–	–	–	–
SRMF	$0.3816 \pm 2 \times 10^{-5}$	$0.1135 \pm 9 \times 10^{-6}$	$0.4617 \pm 1 \times 10^{-6}$	$0.4324 \pm 1 \times 10^{-5}$
DeepCDR	$0.6085 \pm 1 \times 10^{-3}$	$0.1987 \pm 1 \times 10^{-3}$	$0.6957 \pm 6 \times 10^{-3}$	$0.6519 \pm 6 \times 10^{-3}$
DeepDSC	$0.5956 \pm 2 \times 10^{-3}$	$0.1948 \pm 3 \times 10^{-3}$	$0.6589 \pm 3 \times 10^{-3}$	$0.6167 \pm 1 \times 10^{-3}$
MOFGCN	$0.5266 \pm 2 \times 10^{-3}$	$0.1654 \pm 1 \times 10^{-3}$	$0.5647 \pm 2 \times 10^{-3}$	$0.5204 \pm 3 \times 10^{-3}$
GraphCDR	$0.5719 \pm 1 \times 10^{-4}$	$0.1631 \pm 6 \times 10^{-4}$	$0.6722 \pm 1 \times 10^{-3}$	$0.6537 \pm 9 \times 10^{-4}$
NIHGNCN	$0.6200 \pm 5 \times 10^{-4}$	$0.2280 \pm 1 \times 10^{-3}$	$0.7118 \pm 1 \times 10^{-3}$	$0.6356 \pm 2 \times 10^{-3}$
MTIGCN	$0.6276 \pm 4 \times 10^{-4}$	$0.2691 \pm 1 \times 10^{-3}$	$0.7459 \pm 5 \times 10^{-4}$	$0.6891 \pm 1 \times 10^{-3}$

**Table 8**

Ablation study on GDSC and CCLE datasets.

Dataset	Methods	AUC	AUPRC
GDSC	MTIGCN	<b><math>0.8870 \pm 6 \times 10^{-6}</math></b>	<b><math>0.8907 \pm 7 \times 10^{-6}</math></b>
	MTIGCN-A	$0.8766 \pm 8 \times 10^{-6}$	$0.8810 \pm 1 \times 10^{-5}$
	MTIGCN-B	$0.8841 \pm 9 \times 10^{-6}$	$0.8886 \pm 1 \times 10^{-5}$
	MTIGCN-AB	$0.8760 \pm 1 \times 10^{-5}$	$0.8803 \pm 1 \times 10^{-5}$
	MTIGCN	<b><math>0.8810 \pm 1 \times 10^{-4}</math></b>	<b><math>0.8813 \pm 2 \times 10^{-4}</math></b>
	MTIGCN-A	$0.8782 \pm 1 \times 10^{-4}$	$0.8786 \pm 1 \times 10^{-4}$
	MTIGCN-B	$0.8745 \pm 1 \times 10^{-4}$	$0.8753 \pm 1 \times 10^{-4}$
	MTIGCN-AB	$0.8806 \pm 1 \times 10^{-4}$	$0.8803 \pm 1 \times 10^{-4}$

MTIGCN-A means removing auxiliary task A (IC50 regression task) from our original model.

MTIGCN-B means removing auxiliary task B (similarity reconstruction task) from our original model.

MTIGCN-AB means removing both auxiliary task A and auxiliary task B from our original model.

On the CCLE dataset, removing IC50 regression or removing similarity reconstruction task model performance decreased. We found that the data volume of CCLE is small, while the parameters and complexity

of MTIGCN increase after adding tasks, which may cause MTIGCN to overfit on some tasks, resulting in performance degradation.

The MTIGCN model benefits from the joint learning of regression prediction and similarity reconstruction tasks, leading to improved performance compared to its variants.

#### 4.3.5. Case study

It was found that about 20 % of the drug response data were missing in the existing dataset [24,25]. To address this data gap and verify whether MTIGCN can discover unknown drug responses in cell lines, we utilized all the known drug response data in the GDSC dataset to train the MTIGCN model to predict these missing responses using the trained model. Table 9 presents the top 10 cell lines the MTIGCN model predicted most sensitive to Dasatinib and GSK690693. After conducting a non-exhaustive literature search, we found that some cell lines had already been confirmed to be sensitive to the respective drugs in previous studies or clinical trials.

For Dasatinib, three cell lines (NCI-H292, JURL-MK1, and 786-0) were among the top predicted sensitive cell lines. These findings aligned with previous studies: Dasatinib was confirmed to inhibit the growth of NCI-H292 cells by inhibiting the autophosphorylation of ACK kinase in a previous study [32]. Obr et al. also found that Dasatinib affects the survival and death of CML cells JURL-MK1 by inhibiting BCR-ABL1 fusion kinase [33]. Moreover, in an experiment by Roseweir et al., Dasatinib reduced the activity and colony formation ability of 786-0 cells, induced cell death, and prevented cell migration and invasion [34].

For GSK690693, three cell lines (RCH-ACV, JEKO-1, and MOLT-16) were consistent with previous study observations or clinical trials. Levy et al. studied the response of pre-B cell RCH-ACV and T cell line MOLT-16 to GSK690693 and found that it effectively inhibits their proliferation [35]. Additionally, Liu et al. demonstrated that GSK690693 can effectively inhibit the proliferation of the MCL cell line JeKo-1 [36].

These case study results indicate that MTIGCN successfully predicted cell lines that were already known to be sensitive to Dasatinib and GSK690693, confirming its ability to discover unknown drug responses in cell lines. In line with previous experimental findings, the model's predictions further validate its potential utility in identifying novel drug sensitivities, which could be valuable for drug development and personalized medicine approaches.

## 5. Conclusion

In this study, we proposed a multi-task learning method called MTIGCN to predict anti-cancer drug response. MTIGCN utilizes the complementary information between the drug sensitive or resistant classification task and the IC50 regression prediction task. By sharing parameters and optimizing the losses of different tasks simultaneously, MTIGCN enhances the feature representation and reduces overfitting. In addition, MTIGCN incorporates a similarity reconstruction task, which preserves the intrinsic structure of drug similarity and cell line similarity in the embeddings learned by the network, and prevents the loss of similarity between the original features during the learning process.

- (1) The results of the experiments and evaluations on two in vitro datasets demonstrated that MTIGCN outperformed seven state-of-the-art baseline methods. The success of multi-task learning allowed the model to perform better generalization by sharing information between related tasks.
- (2) Moreover, the well-trained model on the in vitro dataset GDSC exhibited good performance when applied to predict drug responses in in vivo datasets PDX and TCGA, which indicates the potential of MTIGCN for personalized treatment in clinical applications.

**Table 9**  
Top 10 predicted sensitive cell lines for Dasatinib and GSK690693.

Drug	Rank	Cell line	PMID
Dasatinib	1	A204	N/A
	2	NCI-H292	20,190,765
	3	RCC-JF	N/A
	4	JURL-MK1	25,198,091
	5	HCC-44	N/A
	6	SW1710	N/A
	7	Hs-633 T	N/A
	8	786-0	26,984,511
	9	TT2609-CO2	N/A
	10	NCI-H2369	N/A
GSK690693	1	RCH-ACV	19,064,730
	2	JEKO-1	32,120,074
	3	KP-1 N	N/A
	4	GA-10	N/A
	5	LB647-SCLC	N/A
	6	SCC90	N/A
	7	CRO-AP2	N/A
	8	DOHH-2	N/A
	9	NCI-H929	N/A
	10	MOLT-16	19,064,730

(3) The case study confirmed the model's ability to discover unknown drug responses in cell lines, suggesting that MTIGCN could be a helpful reference for cancer pharmacology research.

here are some limitations and shortcomings in our work that need to be addressed in future research. For example, when designing a multi-task learning model, the correlation between the selected tasks is crucial. A good set of tasks can promote each other and improve the generalization ability, while a poor set of tasks, such as the CCLE dataset regression task, can instead reduce the overall performance of the model. In addition, when predicting *in vivo* responses, our approach does not account for batch effects between patients and cell lines. Future research will explore new methods, such as few-shot learning [37], aligning cell line and patient domains through loss function [38] and other powerful embedding models [39], to address the challenges between preclinical models and clinical applications.

## CRediT authorship contribution statement

**Hancheng Liu:** Data curation, Writing – original draft, Software.  
**Wei Peng:** Conceptualization, Methodology, Writing – review & editing.  
**Wei Dai:** Visualization, Investigation, Validation. **Jianzhen Lin:** .  
**Xiaodong Fu:** Supervision. **Li Liu:** Conceptualization. **Lijun Liu:** Conceptualization. **Ning Yu:** Writing – review & editing.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

The data and source code are available at: <https://github.com/weiba/MTIGCN>

## Acknowledgments

This work is supported in part by the National Natural Science Foundation of China under grant No.61972185, Natural Science Foundation of Yunnan Province of China (2019FA024), Yunnan Ten Thousand Talents Plan young.

## References

- [1] D. Baptista, P.G. Ferreira, M. Rocha, Deep learning for drug response prediction in cancer, *Brief. Bioinform.* 22 (2021) 360–379, <https://doi.org/10.1093/bib/bbz171>.
- [2] J. Marquart, E.Y. Chen, V. Prasad, Estimation of the percentage of US patients with cancer who benefit from genome-driven oncology, *JAMA Oncol.* 4 (2018) 1093, <https://doi.org/10.1001/jamaoncol.2018.1660>.
- [3] G. Adam, L. Rampášek, Z. Safikhani, P. Smirnov, B. Haibe-Kains, A. Goldenberg, Machine learning approaches to drug response prediction: challenges and recent progress, *Npj Precis. Oncol.* 4 (2020) 19, <https://doi.org/10.1038/s41698-020-0122-1>.
- [4] I. Vitale, E. Shema, S. Loi, L. Galluzzi, Intratumoral heterogeneity in cancer progression and response to immunotherapy, *Nat. Med.* 27 (2021) 212–224, <https://doi.org/10.1038/s41591-021-01233-9>.
- [5] F. Iorio, T.A. Knijnenburg, D.J. Vis, G.R. Bignell, M.P. Menden, M. Schubert, N. Aben, E. Gonçalves, S. Barthorpe, H. Lightfoot, T. Cokelaer, P. Greninger, E. Van Dyk, H. Chang, H. De Silva, H. Heyn, X. Deng, R.K. Egan, Q. Liu, T. Mironenko, X. Mitropoulos, L. Richardson, J. Wang, T. Zhang, S. Moran, S. Sayols, M. Soleimanl, D. Tamborero, N. Lopez-Bigas, P. Ross-Macdonald, M. Esteller, N. S. Gray, D.A. Haber, M.R. Stratton, C.H. Benes, L.F.A. Wessels, J. Saez-Rodriguez, U. McDermott, M.J. Garnett, A landscape of pharmacogenomic interactions in cancer, *Cell* 166 (2016) 740–754, <https://doi.org/10.1016/j.cell.2016.06.017>.
- [6] J. Barretina, G. Caponigro, N. Stransky, K. Venkatesan, A.A. Margolin, S. Kim, C. J. Wilson, J. Lehár, G.V. Kryukov, D. Sonkin, A. Reddy, M. Liu, L. Murray, M. F. Berger, J.E. Monahan, P. Morais, J. Meltzer, A. Korejwa, J. Jané-Vilbuena, F. A. Mapa, J. Thibault, E. Bric-Furlong, P. Raman, A. Shipway, I.H. Engels, J. Cheng, G.K. Yu, J. Yu, P. Aspasia, M. De Silva, K. Jagtap, M.D. Jones, L. Wang, C. Hatton, E. Palescandolo, S. Gupta, S. Mahan, C. Sougnez, R.C. Onofrio, T. Lifield, L. MacConaill, W. Winckler, M. Reich, N. Li, J.P. Mesirov, S.B. Gabriel, G. Getz, K. Ardlie, V. Chan, V.E. Myer, B.L. Weber, J. Porter, M. Warmuth, P. Finan, J. L. Harris, M. Meyerson, T.R. Golub, M.P. Morrissey, W.R. Sellers, R. Schlegel, L. A. Garraway, The cancer cell line encyclopedia enables predictive modelling of anticancer drug sensitivity, *Nature* 483 (2012) 603–607, <https://doi.org/10.1038/nature11003>.
- [7] B. Seashore-Ludlow, M.G. Rees, J.H. Cheah, M. Cokol, E.V. Price, M.E. Coletti, V. Jones, N.E. Bodycombe, C.K. Soule, J. Gould, B. Alexander, A. Li, P. Montgomery, M.J. Wawer, N. Kuru, J.D. Kotz, C.-S.-Y. Hon, B. Munoz, T. Liefeld, V. Dančík, J.A. Bittker, M. Palmer, J.E. Bradner, A.F. Shamji, P.A. Clemons, S. L. Schreiber, Harnessing connectivity in a large-scale small-molecule sensitivity dataset, *Cancer Discov.* 5 (2015) 1210–1223, <https://doi.org/10.1158/2159-8290.CD-15-0235>.
- [8] B. Shen, F. Feng, K. Li, P. Lin, L. Ma, H. Li, A systematic assessment of deep learning methods for drug response prediction: from in vitro to clinical applications, *Brief. Bioinform.* 24 (2023) bbac605, <https://doi.org/10.1093/bib/bbac605>.
- [9] P. Gelehrter, N.J. Cox, R.S. Huang, Clinical drug response can be predicted using baseline gene expression levels and in vitro drug sensitivity in cell lines, *Genome Biol.* 15 (2014) R47, <https://doi.org/10.1186/gb-2014-15-3-r47>.
- [10] R. Tibshirani, Regression shrinkage and selection via the lasso, *J. R. Stat. Soc. B. Methodol.* 58 (1996) 267–288, <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>.
- [11] H. Zou, T. Hastie, Regularization and variable selection via the elastic net, *J. R. Stat. Soc. Ser. B Stat Methodol.* 67 (2005) 301–320, <https://doi.org/10.1111/j.1467-9868.2005.00503.x>.
- [12] T. Sakellaropoulos, K. Vougas, S. Narang, F. Koinis, A. Kotsinas, A. Polyzos, T. J. Moss, S. Piha-Paul, H. Zhou, E. Kardala, E. Damianioudi, L.G. Alexopoulos, I. Aifantis, P.A. Townsend, M.I. Panayiotidis, P. Sfikakis, J. Bartek, R.C. Fitzgerald, D. Thanos, K.R. Mills Shaw, R. Petty, A. Tsirigos, V.G. Gorgoulis, A deep learning framework for predicting response to therapy in cancer, *Cell Rep.* 29 (2019) 3367–3373.e4, <https://doi.org/10.1016/j.celrep.2019.11.017>.
- [13] Y. Chang, H. Park, H.-J. Yang, S. Lee, K.-Y. Lee, T.S. Kim, J. Jung, J.-M. Shin, Cancer drug response profile scan (CDRscan): a deep learning model that predicts drug effectiveness from cancer genomic signature, *Sci. Rep.* 8 (2018) 8857, <https://doi.org/10.1038/s41598-018-27214-6>.
- [14] H. Sharifi-Noghabi, O. Zolotareva, C.C. Collins, M. Ester, MOLI: multi-omics late integration with deep neural networks for drug response prediction, *Bioinformatics* 35 (2019) i501–i509, <https://doi.org/10.1093/bioinformatics/btz318>.
- [15] M. Manica, A. Oskooei, J. Born, V. Subramanian, J. Sáez-Rodríguez, M. Rodríguez Martínez, Toward explainable anticancer compound sensitivity prediction via multimodal attention-based convolutional encoders, *Mol. Pharm.* 16 (2019) 4797–4806, <https://doi.org/10.1021/acs.molpharmaceut.9b00520>.
- [16] Q. Liu, Z. Hu, R. Jiang, M. Zhou, DeepCDR: a hybrid graph convolutional network for predicting cancer drug response, *Bioinformatics* 36 (2020) i911–i918, <https://doi.org/10.1093/bioinformatics/btaa822>.
- [17] M. Li, Y. Wang, R. Zheng, X. Shi, Y. Li, F.-X. Wu, J. Wang, DeepDSC: a deep learning method to predict drug sensitivity of cancer cell lines, *IEEE/ACM Trans. Comput. Biol. Bioinf.* 18 (2021) 575–582, <https://doi.org/10.1109/TCBB.2019.2919581>.
- [18] R. Su, X. Liu, L. Wei, Q. Zou, Deep-Resp-Forest: A deep forest model to predict anti-cancer drug response, *Methods* 166 (2019) 91–102, <https://doi.org/10.1016/jymeth.2019.02.009>.
- [19] Y. Zhang, X. Lei, Z. Fang, Y. Pan, CircRNA-disease associations prediction based on metapath2vec++ and matrix factorization, *Big Data Min. Anal.* 3 (2020) 280–291, <https://doi.org/10.26599/BDMA.2020.9020025>.
- [20] C. Fan, X. Lei, L. Guo, A. Zhang, Predicting the associations between microbes and diseases by integrating multiple data sources and path-based HeteSim scores, *Neurocomputing* 323 (2019) 76–85, <https://doi.org/10.1016/j.neucom.2018.09.054>.
- [21] F. Zhang, M. Wang, J. Xi, J. Yang, A. Li, A novel heterogeneous network-based method for drug response prediction in cancer cell lines, *Sci. Rep.* 8 (2018) 3355, <https://doi.org/10.1038/s41598-018-21622-4>.
- [22] L. Wang, X. Li, L. Zhang, Q. Gao, Improved anticancer drug response prediction in cell lines using matrix factorization with similarity regularization, *BMC Cancer* 17 (2017) 513, <https://doi.org/10.1186/s12885-017-3500-5>.
- [23] W. Peng, T. Chen, W. Dai, Predicting drug response based on multi-omics fusion and graph convolution, *IEEE J. Biomed. Health Inform.* 26 (2022) 1384–1393, <https://doi.org/10.1109/JBHI.2021.3102186>.
- [24] X. Liu, C. Song, F. Huang, H. Fu, W. Xiao, W. Zhang, GraphCDR: a graph neural network method with contrastive learning for cancer drug response prediction, *Brief. Bioinform.* 23 (2022) bbab457, <https://doi.org/10.1093/bib/bbab457>.
- [25] W. Peng, H. Liu, W. Dai, N. Yu, J. Wang, Predicting cancer drug response using parallel heterogeneous graph convolutional networks with neighborhood interactions, *Bioinformatics* 38 (2022) 4546–4553, <https://doi.org/10.1093/bioinformatics/btac574>.
- [26] W. Peng, Q. Tang, W. Dai, T. Chen, Improving cancer driver gene identification using multi-task learning on graph convolutional network, *Brief. Bioinform.* 23 (2022) bbab432, <https://doi.org/10.1093/bib/bbab432>.
- [27] Chemosensitivity prediction by transcriptional profiling, (n.d.). <https://doi.org/10.1073/pnas.191368598>.
- [28] E.E. Bolton, Y. Wang, P.A. Thiessen, S.H. Bryant, PubChem: Integrated Platform of Small Molecules and Biological Activities, in: *Annu. Rep. Comput. Chem.*, Elsevier, 2008: pp. 217–241. [https://doi.org/10.1016/S1574-1400\(08\)00012-1](https://doi.org/10.1016/S1574-1400(08)00012-1).
- [29] R.A. Irizarry, Summaries of Affymetrix GeneChip probe level data, *Nucleic Acids Res.* 31 (2003) 15e–e, <https://doi.org/10.1093/nar/gng015>.
- [30] H. Gao, J.M. Korn, S. Ferretti, J.E. Monahan, Y. Wang, M. Singh, C. Zhang, C. Schnell, G. Yang, Y. Zhang, O.A. Balbin, S. Barbe, H. Cai, F. Casey, S. Chatterjee, D.Y. Chiang, S. Chuai, S.M. Cogan, S.D. Collins, E. Dammasca, N. Ebel, M. Embry, J. Green, A. Kauffmann, C. Kowal, R.J. Leary, J. Lehar, Y. Liang, A. Loo, E. Lorenzana, E. Robert McDonald, M.E. McLaughlin, J. Merkin, R. Meyer, T. L. Naylor, M. Patrawan, A. Reddy, C. Röelli, D.A. Ruddy, F. Salangsang, F. Santacroce, A.P. Singh, Y. Tang, W. Tinetto, S. Tobler, R. Velazquez, K. Venkatesan, F. Von Arx, H.Q. Wang, Z. Wang, M. Wiesmann, D. Wyss, F. Xu, H. Bitter, P. Atadjia, E. Lees, F. Hofmann, E. Li, N. Keen, R. Cozens, M.R. Jensen, N. K. Pryer, J.A. Williams, W.R. Sellers, High-throughput screening using patient-derived tumor xenografts to predict clinical trial drug response, *Nat. Med.* 21 (2015) 1318–1325, <https://doi.org/10.1038/nm.3954>.
- [31] Z. Ding, S. Zu, J. Gu, Evaluating the molecule-based prediction of clinical drug responses in cancer, *Bioinformatics* 32 (2016) 2891–2895, <https://doi.org/10.1093/bioinformatics/btw344>.
- [32] J. Li, U. Rix, B. Fang, Y. Bai, A. Edwards, J. Colinge, K.L. Bennett, J. Gao, L. Song, S. Eschrich, G. Supererti-Furga, J. Koomen, E.B. Haura, A chemical and phosphoproteomic characterization of dasatinib action in lung cancer, *Nat. Chem. Biol.* 6 (2010) 291–299, <https://doi.org/10.1038/nchembio.332>.
- [33] A. Obr, P. Röselová, D. Grebenová, K. Kuželová, Real-time analysis of imatinib- and dasatinib-induced effects on chronic myelogenous leukemia cell interaction with fibronectin, *PLoS One* 9 (2014) e107367.
- [34] A.K. Roseweir, T. Qayyum, Z. Lim, R. Hammond, A.I. MacDonald, S. Fraser, G. M. Oades, M. Aitchison, R.J. Jones, J. Edwards, Nuclear expression of Lyn, a Src family kinase member, is associated with poor prognosis in renal cancer patients, *BMC Cancer* 16 (2016) 229, <https://doi.org/10.1186/s12885-016-2254-9>.
- [35] D.S. Levy, J.A. Kahana, R. Kumar, AKT inhibitor, GSK690693, induces growth inhibition and apoptosis in acute lymphoblastic leukemia cell lines, *Blood* 113 (2009) 1723–1729, <https://doi.org/10.1182/blood-2008-02-137737>.
- [36] Y. Liu, Z. Zhang, F. Ran, K. Guo, X. Chen, G. Zhao, Extensive investigation of benzylidene N-containing substituents on the pyrrolopyrimidine skeleton as Akt inhibitors with potent anticancer activity, *Bioorganic Chem.* 97 (2020) 103671, <https://doi.org/10.1016/j.bioorg.2020.103671>.
- [37] J. Ma, S.H. Fong, Y. Luo, C.J. Bakkenist, J.P. Shen, S. Mourragui, L.F.A. Wessels, M. Hafner, R. Sharari, J. Peng, T. Ideker, Few-shot learning creates predictive models of drug response that translate from high-throughput screens to individual patients, *Nat. Cancer.* 2 (2021) 233–244, <https://doi.org/10.1038/s43018-020-00169-2>.
- [38] H. Sharifi-Noghabi, P.A. Harjandi, O. Zolotareva, C.C. Collins, M. Ester, Out-of-distribution generalization from labelled and unlabelled gene expression data for drug response prediction, *Nat. Mach. Intell.* 3 (2021) 962–972, <https://doi.org/10.1038/s42256-021-00408-w>.
- [39] M. Chen, Y. Jiang, X. Lel, Y. Pan, C. Ji, W. Jiang, Dug-target Interactions Prediction based on Signed Heterogeneous Graph Neural Networks, *Chin. J. Electron.* 33 (1) (2024) 1–13.