MHGCN: a Multi-channel Hybrid Graph Convolutional Neural Network for Cancer Drug Response Prediction

Peisheng Yang[1], Changxiang He[1*], Ping Zhang[1], Xiaofei Qin[2], Qingqian Zhang[1] and Die Li[1]

***Abstract*—Due to the heterogeneity of cancer cells, personalized treatment plans for cancer patients remain a continuous concern. High-throughput drug screening technology has led to the development of deep learning models that generate personalized therapies. However, most existing models fail to account for the topological relationships between cell line-drug pair (CDP) nodes, thereby ignoring their intrinsic connections. This paper proposes a multi-channel hybrid graph convolutional neural network (MHGCN) for predicting cancer drug response (CDR). First, we define CDPs by integrating gene expression and drug molecular fingerprints. These CDPs are refined through denoising autoencoders to eliminate noise. Second, we compute pairwise cosine similarities among CDPs to build a similarity network, while simultaneously establishing a heterogeneous response graph connecting cell lines and drugs. Third, MHGCN processes the CDP network via graph convolutional layers and generates the response matrix through linear projection. Concurrently, a heterogeneous graph convolutional neural network learns the response heterogeneous network. Following data augmentation, we derive feature embeddings for cell lines and drugs, then compute their similarity matrix. Finally, CDR predictions are generated through weighted matrix fusion of these components. To the best of our knowledge, MHGCN represents the first framework explicitly incorporating CDP topology into CDR prediction. Experiments demonstrate MHGCN's statistically significant improvements over state-of-the-art methods.**

***Index Terms*—cancer drug response prediction, cell line-drug pair, graph convolutional neural network**

I. Introduction

CANCER remains a leading cause of global mortality, posing a significant health challenge due to its profound impact on human life [1]. Developing effective anti-cancer therapies is therefore an urgent priority. However, tumor heterogeneity—driven by genomic variations and cellular diversity—leads to divergent therapeutic responses among patients with the same cancer type [2-4]. Personalized treatment is consequently essential to improve patient survival and longevity. Traditional biological methods for identifying individualized therapies are often clinically impractical, as they are time-consuming, costly, and may delay optimal intervention. Recent advances in genomics and biological sciences have enabled researchers to screen thousands of cancer cell lines and compile drug sensitivity data in public repositories such as the Cancer Cell Line Encyclopedia (CCLE) [5] and Genomics of Drug Sensitivity in Cancer (GDSC) [6]. These databases integrate genomic profiles (e.g., gene expression) with pharmacogenomic data, providing vital resources for predicting personalized drug responses.

Building on these resources, recent advances in high-throughput drug screening have accelerated the development of machine learning algorithms for pharmacological sensitivity prediction. These approaches are broadly categorized as classical machine learning or deep neural networks. Classical methods utilize network science principles, constructing bipartite graphs with cell lines and drugs as nodes and employing algorithms such as random walk [7-8], restartable random walk [9], or information flow [10-11] for link prediction. Classification-based approaches first extract molecular descriptors, then apply predictive models including support vector machines (SVM) [12-13], random forests [14-15], logistic regression [16], or ensemble methods [17]. Deep learning techniques, conversely, harness hierarchical feature learning through architectures such as autoencoders, stacked autoencoders [18], variational autoencoders (VAE) [19], convolutional neural networks (CNN) [20-21], recurrent neural networks (RNN) [22], and multilayer perceptrons (MLP) [23-24] to learn discriminative representations of cellular and pharmacological features. Notable examples include DeepDSC [25], which employs stacked denoising autoencoders for unsupervised pre-training of gene expression profiles to predict drug sensitivity; CCL-ASP [26], introducing collaborative contrastive learning with adaptive self-paced sampling for drug-target interaction prediction; and MIDTI [27], integrating multi-view similarity networks with deep interactive attention mechanisms. However, despite demonstrating superior accuracy over classical methods, these approaches often neglect inherent topological relationships between cell lines and drugs—essential for capturing systemic interaction patterns.

Recent advances in graph representation learning have firmly established Graph Neural Networks (GNNs) as powerful tools for CDR prediction, largely due to their capacity to model complex biological interaction systems. Several notable graph-based approaches have emerged: Liu et al. [21] propose DeepCDR, a hybrid graph convolutional network that learns drug chemical structures through uniform graph convolution, fuses multi-omics data with drug characteristics via 1D-CNN, and predicts IC50 values; Peng et al. [28] develop MOFGCN, constructing cell line-drug heterogeneous networks from multi-omics and chemical structure data to enable similarity-based prediction via graph-derived node em-

*Correspondence: changxiang-he@163.com
[1]College of Science, University of Shanghai for Science and Technology, Shanghai 200093, China
[2]School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China

beddings; Hostallero et al. [29] introduce BiG-DRP, which builds response heterogeneous networks from log(IC50) values and employs heterogeneous graph convolution for feature extraction; Peng et al. [32] propose NIHGCN with neighborhood interactions; Liu et al. [33] present a self-supervised GNN method with contrastive learning; and AMDGT [34] employs dual-graph transformers to model inter-modal dependencies. Despite these advances in extracting topology information between cell lines and drugs, current graph-based methods have largely neglected interactions between individual Cell-Drug Pairs (CDPs). CDP-centric modeling offers the distinct advantage of simultaneously optimizing both node attributes (cellular/drug features) and relationship patterns. This represents a critical unexplored dimension for significantly enhancing prediction robustness, especially for rare cell-drug combinations.

To address these limitations, we introduce MHGCN (Multi-channel Hybrid Graph Convolutional Network), a novel framework that explicitly models pairwise Cell-Drug Pair (CDP) interactions to jointly learn cell line and drug representations. Our approach constructs a CDP network using denoised features and cosine similarity metrics, processing it through graph convolutional layers to extract topological features. Concurrently, we build a heterogeneous cell line-drug graph using IC50 values, employing heterogeneous GCNs to extract embedded features. Linear projection layers then map CDP features to pharmacological response matrices, enabling the derivation of interaction patterns from topological relationships, while similarity matrices capture cross-modal dynamics in the heterogeneous graph space. By integrating CDP network topology with heterogeneous network information, MHGCN generates comprehensive predictions of cell line-drug interactions for robust drug efficacy assessment in personalized oncology. Experimental validation on CCLE and GDSC demonstrates statistically significant superiority over state-of-the-art baselines.

Our research has made the following contributions:

- MHGCN introduces CDP and fuses the topological characteristics of CDP nodes with heterogeneous information of cell lines and drugs for the first time.
- We use a data enhancement layer to make cell line and drug features more prominent, thereby improving the prediction performance of CDR.
- Comparative experiments with competitive models on open databases demonstrate significant improvements in the performance of our model.

II. Materials and Methods

A. Materials

This study leverages two authoritative pharmacogenomic databases for comprehensive analysis. The Cancer Cell Line Encyclopedia (CCLE; https://depmap.org/portal/) provides multi-omics profiles including gene expression, mutations, and copy number variations across 1,072 human cancer cell lines, serving as our primary source for baseline molecular characterization. Complementing this, the Genomics of Drug Sensitivity in Cancer (GDSC; https://www.cancerrxgene.org/) contains dose-response data for 265 anticancer compounds tested on 809 cell lines, supplying quantitative drug sensitivity measurements represented by IC50 values (half-maximal inhibitory concentration). Following NCI-60 screening standards [5-6], these IC50 values quantify drug potency, with lower values ($< 1\ \mu m$) indicating higher sensitivity.

For quantitative analysis, we obtained log-transformed IC50 values from GDSC/CCLE and binarized the continuous responses following established protocols [28,35], using:

$$A_{ij} = \begin{cases} 1 & \sigma(response_j)_{ij} \leq \alpha \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where σ represents the z-score normalization, $response_j$ is the logarithm of IC50 for j-th drug with other cells, and α is a predefined threshold. We consider the drug and cancer cell line have an interactive relationship if $A_{ij} = 1$. Conversely, if $A_{ij} = 0$, the cell line is deemed resistant to the drug. For cancer cell lines and drugs without IC50 values, we assume that there is no interaction between them. Gene expression data for cancer cell lines is obtained from two databases, namely GDSC and CCLE. The fingerprint characteristics of drugs are sourced from the PubChem database.

Following database screening, we excluded drugs lacking fingerprint characteristics and cell lines missing gene expression data. From CCLE, we obtained 436 cancer cell lines and 24 drugs, yielding 1,696 drug-sensitive interactions and 8,768 resistant samples. The GDSC dataset provided 527 cell lines and 100 drugs, comprising 10,996 interactions and 41,704 resistant samples. For missing IC50 values (absent in 4,058 GDSC cases but none in CCLE), we treated these drug-cell line pairs as resistant cases. While these missing-value samples participate in model training, they are masked during optimization to prevent influencing parameter updates. The filtered datasets described above were used in all experiments in Section III (EXPERIMENTS) except for Test 2.

B. Methods

Our proposed model architecture, illustrated in Figure 1, comprises four key components: (1) feature extraction, (2) network construction, (3) embedding learning, and (4) CDR prediction. The model processes three primary inputs: gene expression profiles from cancer cell lines, drug fingerprint characteristics, and the corresponding cell line-drug response matrix. In the feature extraction stage, we first construct cell line-drug pairs (CDPs) and employ a Denoising Autoencoder (DAE) [32] to derive low-dimensional CDP representations. The network construction phase then builds two complementary structures: a CDP similarity network generated through pairwise similarity computations, and a cell line-drug heterogeneous network derived from the response matrix. For embedding learning, we utilize distinct graph neural architectures: a homogeneous graph convolutional network processes the CDP network topology, while a heterogeneous graph convolutional network coupled with data enhancement techniques learns representations from the cell line-drug interactions. The final CDR prediction integrates these learned representations

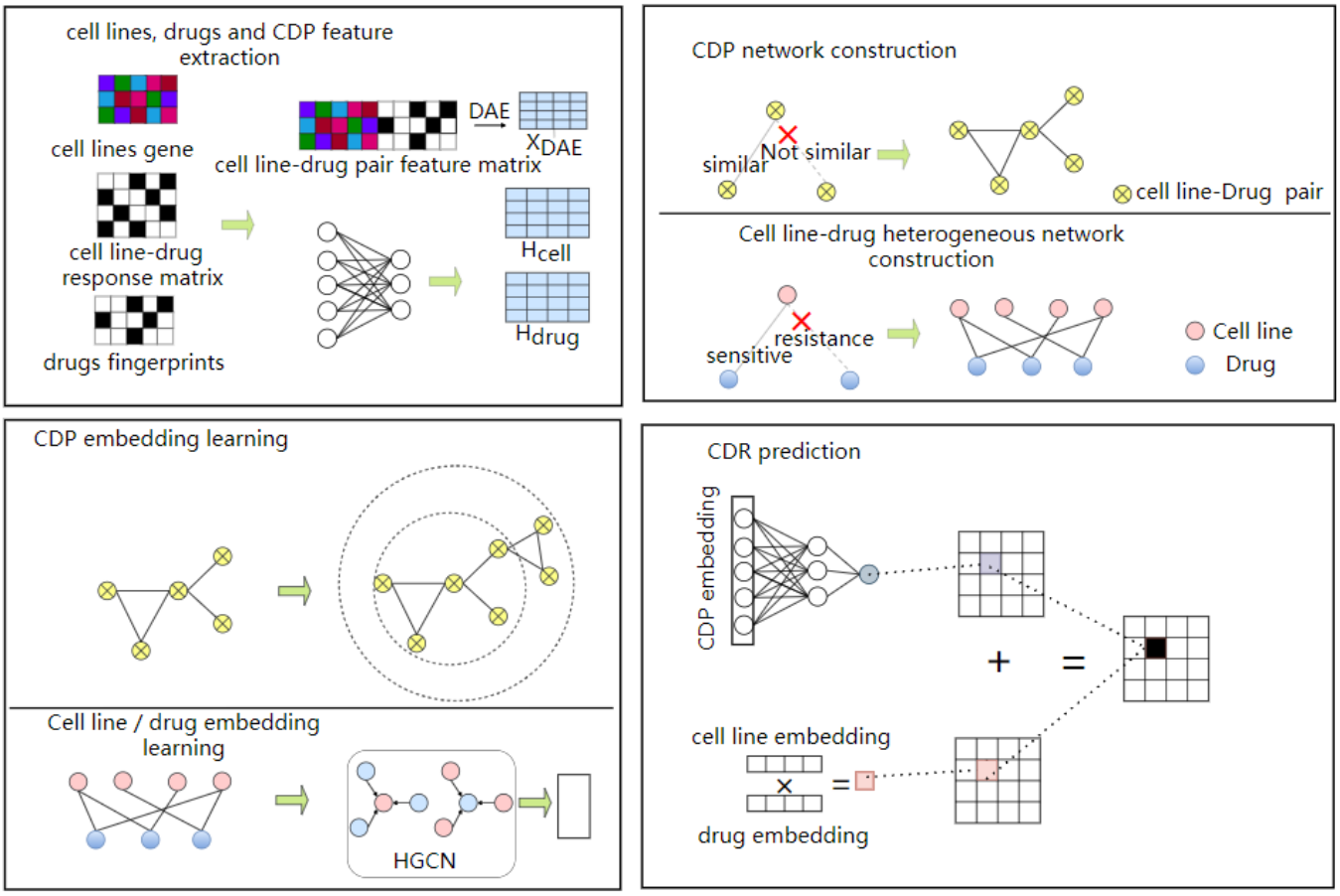

Fig. 1. Architecture of the Multi-channel Hybrid Graph Convolutional Network (MHGCN). The framework comprises four stages: (a) Feature Extraction: CDP node construction with denoising autoencoder (DAE) preprocessing. (b) Networks Construction: CDP similarity network via pairwise cosine similarity computation and cell line-drug heterogeneous network based on interaction information. (c) Embedding Learning: Graph convolutional networks (GCN) process the CDP topology while heterogeneous graph convolutional networks (HGCN) analyze interaction of drugs and cell lines. (d) CDR Prediction: Fusion of node-level regression (CDP network) and edge-level interaction prediction (heterogeneous network). Comparative experiments with competitive models on open databases demonstrate significant improvements in the performance of our model.

by fusing the response matrix with cell line-drug similarity matrices.

In this section, the above four parts will be described in detail.

1) Feature Extraction: We use the gene expression matrix $X_{cell} \in R^{N_{cell} \times d_{cell}}$ of the cancer cell lines and the fingerprint feature matrix $X_{drug} \in R^{N_{drug} \times d_{drug}}$ of the drug as the original input information of the model, where N_{cell} and N_{drug} represent the number of cancer cell lines and drugs, respectively, while d_{cell} and d_{drug} represent their initial feature dimensions. In order to capture the deep information between cancer cell lines and drugs to improve the accuracy of the model, we combine each cancer cell lines and drugs to form CDPs. We use CDP_k to represent the k-th CDP. Its feature can be expressed as $H^k_{CDP} = \{X_{cell}, X_{drug}\}$, where H^k_{CDP} represents gene expression feature of the i-th cancer cell line and the fingerprint feature of the j-th drug connected to each other. Then there is an initial feature matrix $H_{CDP} \in R^{N_{CDP} \times d_{CDP}}$ for all CDPs, where $N_{CDP} = N_{cell} \times N_{drug}$ and $d_{CDP} = d_{cell} + d_{drug}$.

To extract noise-robust low-dimensional representations of Cell-Drug Pairs (CDPs), we process them through a Denoising Autoencoder (DAE) model. As illustrated in Figure 2, the DAE's key functionality resides in its hidden layer representations. The model first intentionally corrupts input signals by adding controlled noise (Equation 2), then learns to reconstruct the original clean inputs through its hidden layer features.

$$\widetilde{X}_{CDP} = H_{CDP} + \alpha \cdot \epsilon \tag{2}$$

where α is the scaling factor used to control the intensity of noise and ϵ is a random vector that follows a normal distribution. $\widetilde{X}_{CDP}$ is the feature of CDPs after adding noise. We obtain the required low dimensional features of CDPs with added noise $X_{DAE} \in R^{N_{CDP} \times d_{DAE}}$ using the Auto Encoders (reference(3)). Finally, we use the Decoders (reference (4)) to reconstruct the original signal Y.

$$X_{DAE} = \sigma(W_1 \widetilde{X}_{CDP} + b_1) \tag{3}$$

$$Y = W_2 X_{DAE} + b_2 \tag{4}$$

where σ is the Softplus activation function, W_1 and W_2 are learnable weight matrices, b_1 and b_2 are learnable bias vectors.

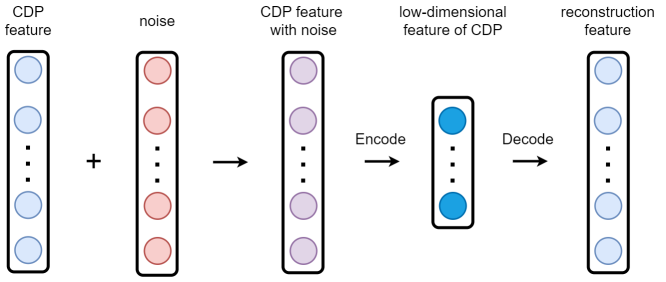

Fig. 2. Denoising Auto-Encoders(DAE) flowchart. The DAE first adds noise to the input CDP signal. Input the generated damaged signal into the encoder to obtain the CDP feature X_{DAE} with noise we need. Finally, the lossless signal is reconstructed through the decoder.

The loss function of the DAE is defined by (5) to be minimization formula. We obtain the low dimensional characteristics of CDPs with noise.

$$J = \frac{1}{2}\Sigma(Y - H_{CDP})^2. \tag{5}$$

To address dimensional inconsistency between the original cell line and drug input features, we employ a linear projection layer to align their feature dimensions. Furthermore, given the substantial variability in gene expression profiles across different cell lines - evident in both variance and standard deviation measures - we perform normalization preprocessing for each cell line using Equation (6).

$$\hat{x}^i_{cell} = \frac{x^i_{cell} - \mu_i}{\sigma_i} \tag{6}$$

where $\hat{x}^i_{cell}$ is the normalized feature vector of the i-th cell line, μ_i represents the average value of gene expression in the i-th cancer cell line, and σ_i represents the standard deviation based on expression in the i-th cancer cell line. In this way, we can obtain the preconditioned cancer cells gene expression matrix $\hat{X}_{cell}$. Then we apply (7) and (8) to linearly transform the features of cancer cell lines and drugs, converting the features of each cancer cell line and drug into the same dimensions.

$$H_{cell} = \hat{X}_{cell} \times \theta_{cell} \tag{7}$$

$$H_{drug} = X_{drug} \times \theta_{drug} \tag{8}$$

where $H_{cell} \in R^{N_{cell} \times d}$ and $H_{drug} \in R^{N_{drug} \times d}$ are the feature matrices of cancer cell lines and drugs with the same dimensions, respectively, θ_{cell} and θ_{drug} represent transition matrices for linear transformation of cancer cell lines and drugs, respectively.

2) Networks Construction: **CDP network construction.** Each row of X_{DAE} represents the low dimensional characteristics of each CDP with noise. For each CDP$_i$, we apply (9) to calculate its cosine similarity with any CDP$_j$,

$$C(X^i_{DAE}, X^j_{DAE}) = \frac{X^i_{DAE} \times X^j_{DAE}}{|| X^i_{DAE} || \times || X^j_{DAE} ||} \tag{9}$$

where X^i_{DAE} and X^j_{DAE} represent the DAE feature of CDP$_i$ and CDP$_j$, respectively, $|| \cdot ||$ represents the modulus of a vector. In this way, we can obtain the cosine similarity between CDP$_i$ and any CDP$_j$, and compare it with a specified threshold to get the similar neighbors of CDP$_i$.

$$A_{CDP_{ij}} = \begin{cases} 1 & C(X^i_{DAE}, X^j_{DAE}) \geq \alpha \quad or \quad i = j \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

If $A_{CDP_{ij}} = 1$, we define CDP$_j$ as the similar neighbor of CDP$_j$, otherwise, they are dissimilar. Besides, the diagonal elements of A_{CDP} are 1. Therefore, we can construct a CDP network $G_{CDP} = (A_{CDP}, H_{CDP})$, where A_{CDP} is the adjacency matrix and H_{CDP} represent the feature of CDPs.

Cell line-drug response heterogeneous network construction. A is the interaction matrix between the cell line and the drug. So we can construct a cell line drug response heterogeneous network $G_h = (A, H_{cell}, H_{drug})$, where A is the adjacency matrix of a heterogeneous network. If $A_{ij} = 1$, it means that the i-th cell has an interaction with the j-th drug, then there is an edge between them. If $A_{ij} = 0$, then the cell line and the drug are not connect to each other. H_{cell} and H_{drug} represent the feature of cell lines and drugs, respectively.

3) Embedding Learning: **CDPs embedding learning.** For $G_{CDP} = (A_{CDP}, H_{CDP})$, we use graph convolutional neural network to capture the embedding of CDP nodes. We apply (11) to obtain the k-th layer output of the graph convolutional neural network.

$$Z^k_{CDP} = \sigma(\widetilde{D}^{-\frac{1}{2}}_{CDP}\widetilde{A}_{CDP}\widetilde{D}^{-\frac{1}{2}}_{CDP}Z^{k-1}_{CDP}W^k) \tag{11}$$

where σ is the Relu activation function, Z^{k-1}_{CDP} is the output of the $k-1$ layer graph convolutional neural network, $\widetilde{A}_{CDP} = A_{CDP} + I$, $\widetilde{D}_{CDP}$ is the diagonal degree matrix of $\widetilde{A}_{CDP}$, W^k is learnable weight matrix of k-th layer of GCN, Z_{CDP} represents the output of the last layer of CPDs in GCN. The Graph Convolutional Network (GCN) effectively captures and models the complex relationships between cell lines and drugs by learning their topological associations through the CDP network structure. This architecture simultaneously extracts two critical information types: (1) the inherent topological patterns among Cell-Drug Pairs (CDPs), and (2) the latent connection features between cell lines and drugs themselves. Through this dual learning mechanism, GCN identifies and represents their correlated features in a unified embedding space.

Cell lines and drugs embedding learning. For heterogeneous network $G_h = (A, H_{cell}, H_{drug})$, considering the characteristics of heterogeneous graphs with multiple types of nodes, we use parallel heterogeneous graph convolutions to aggregate the feature of cell lines and drugs. We apply (12) and (13) to represent the outputs of the k-th layer heterogeneous graph convolution neural network of cell lines and drugs, respectively.

$$Z^k_{cell} = \sigma\{[(D^{-1}_{cell} + I_{cell})Z^{k-1}_{cell} + L_{cell}Z^{k-1}_{drug}]W^k_1\} \tag{12}$$

$$Z^k_{drug} = \sigma\{[(D^{-1}_{drug} + I_{drug})Z^{k-1}_{drug} + L_{drug}Z^{k-1}_{cell}]W^k_2\} \tag{13}$$

where σ is the ReLU activation function, Z_{cell}^{k-1} and Z_{drug}^{k-1} represent the $k-1$ layer heterogeneous convolutional neural network output of of the cell lines and drugs, respectively. When k = 0, $Z_{cell}^{0} = H_{cell}$ and $Z_{drug}^{0} = H_{drug}$. D_{cell} and D_{drug} are the degree matrices of cell lines and drugs, respectively, where $D_{cell(ij)} = \Sigma_j A_{ij} + 1$ and $D_{drug(ij)} = \Sigma_j A_{ji} + 1$. $L_{cell} = D_{cell}^{-\frac{1}{2}} A D_{drug}^{-\frac{1}{2}}$ and $L_{drug} = D_{drug}^{-\frac{1}{2}} A^T D_{cell}^{-\frac{1}{2}}$ are Laplacian matrices of cell lines and drugs. W_1^k and W_2^k are a learnable weight matrix of k-th layer HGCN. Z_{cell} and Z_{drug} represent the output of the last layer of cell lines and drugs in HGCN, respectively.

Finally, we perform a data enhancement layer on the output features of HGCN of cell lines and drugs (reference (14) and (15)). This operation can enhance the features of cell lines and drugs. Thereby it improves the accuracy of the model.

$$F_{cell} = \sigma[(Z_{cell} \odot Z_{cell})W_1] \tag{14}$$

$$F_{drug} = \sigma[(Z_{drug} \odot Z_{drug})W_2] \tag{15}$$

where σ is the ReLU activation function, F_{cell} and F_{drug} are the final feature matrix of cell lines and drugs, W_1 and W_2 are the learnable weight matrix, and $\odot$ is the Hadamard product.

In this module, we obtain the embedding vectors F_{cell}, F_{drug}, Z_{CDP} of all cell lines, drugs and CDPs.

4) CDR prediction: In order to achieve the regression of CDP network features, we perform the following linear transformation on the final embedding of CDP nodes, converting the feature vector of each CDP into a real number. So Z_{CDP} is converted into a column vector.

$$F_{CDP} = Z_{CDP} \times \theta \tag{16}$$

θ is a transition matrix of a linear transformation. $F_{CDP} \in R^{N_{CDP} \times 1}$. We obtain the response matrix $R \in R^{N_{cell} \times N_{drug}}$ for cells and drugs by dimensional transformation of F_{CDP}, where R_{ij} represents the reaction degree of the i-th cell and the j-th drugs.

At the same time, we calculate the linear correlation coefficient between the cells and the drugs to indicate whether there is an interaction between them. We apply (17) to calculate the linear correlation between the i-th cell and the j-th drug.

$$Corr(c_i, d_j) = \frac{(c_i - \mu_i)(d_j - \mu_j)^T}{\sqrt{(c_i - \mu_i)(c_i - \mu_i)^T}\sqrt{(d_j - \mu_j)(d_j - \mu_j)^T}} \tag{17}$$

where $c_i \in F_{cell}$ and $d_j \in F_{drug}$ are the final embedding of cell lines and drugs, respectively. μ_i and μ_j are the average values of c_i and d_j. In order to make the similarity between cell lines and drugs more apparent, we introduce a super parameter γ. Then, we apply (18) to obtain the linear correlation matrix M of the cell lines to the drugs, where M_{ij} indicates the degree of interaction of the i-th cell line to the j-th drug.

$$M = \gamma Corr(F_{cell}, F_{drug}). \tag{18}$$

We apply (19) to integrate the correlation matrix and response matrix of cell lines and drugs to predict whether there is interaction between cell lines and drugs.

$$\hat{A} = \sigma\{\alpha M + (1 - \alpha)R\} \tag{19}$$

Algorithm 1 MHGCN

Input: cell lines gene expression matrix, drug fingerprints matrix, cell line-drug response matrix
Output: CDR prediction matrix: $\hat{A}$

```
Obtaining CDP feature matrix H_CDP by constructing CDPs
X_CDP ← DAE(H_CDP)
Calculating H_cell via Eq.(6)
Calculating H_drug via Eq.(7)
for i to N_CDP do
  Calculating the top-k similar CDPs of CDP_i via Eq.(8)
end for
Constructing A_CDP
G_CDP ← (A_CDP, H_CDP)
G_h ← (A, H_cell, H_drug)
for i = 1 to epoch do
  Z^0_CDP = H_CDP, Z^0_cell = H_cell, Z^0_drug = H_drug
  for i = 1 to k do
    Calculating Z^i_CDP via Eq.(9)
  end for
  for i = 1 to t do
    Calculating Z^i_cell via Eq.(10)
    Calculating Z^i_drug via Eq.(11)
  end for
  calculating F_cell via Eq.(12)
  calculating F_drug via Eq.(13)
  constructing response matrix R via Eq.(14)
  constructing similarity matrix M via Eq.(16)
  calculating CDR precision matrix Â via Eq.(17)
  Calculate loss ℓ via Eq.(18)
  Updating learnable parameters by gradient descent and backpropagation
end for
return Â
```

where σ is the Sigmod activation function and α is a super parameter that adjusts the weight of M and R.

We train our model using the following formula as a loss function.

$$\ell = -\frac{1}{m \times n}\Sigma_{ij} K_{ij}[A_{ij} ln(\hat{A}_{ij}) + (1 - A_{ij}) ln(1 - \hat{A}_{ij})] \tag{20}$$

where m and n are the number of cell lines and drugs respectively, and K is an indicator matrix. When the association between cell line i and drug j is in the training set, $K_{ij} = 1$, otherwise $K_{ij} = 0$.

III. EXPERIMENTS

A. Parameter settings

Our MHGCN framework processes two complementary biological graph structures using specialized neural architectures: (1) a homogeneous Graph Convolutional Network (GCN) that models topological relationships between Cell-Drug Pairs (CDPs), and (2) a Heterogeneous GCN (HGCN)

TABLE I
THE AUC AND AUPRC RESULTS OF CDR PREDICTION ON CCLE.

Model	AUC	AUPRC
HNMDRP	$0.7149\pm2\times10^{-3}$	$0.7112\pm1\times10^{-4}$
tCNNs	$0.7479\pm5\times10^{-4}$	$0.7425\pm2\times10^{-5}$
DeepCDR	$0.8291\pm3\times10^{-4}$	$0.8172\pm4\times10^{-3}$
MOFGCN	$0.8602\pm1\times10^{-5}$	$0.8587\pm1\times10^{-4}$
GraphCDR	$0.8421\pm4\times10^{-4}$	$0.8488\pm2\times10^{-4}$
NIHGCN	$0.8756\pm3\times10^{-4}$	$0.8767\pm2\times10^{-4}$
MHGCN	$\mathbf{0.8814\pm1\times10^{-5}}$	$\mathbf{0.8822\pm1\times10^{-5}}$

* The font bold indicates the optimal model.

TABLE II
THE AUC AND AUPRC RESULTS OF CDR PREDICTION ON GDSC.

Model	AUC	AUPRC
HNMDRP	$0.7129\pm5\times10^{-4}$	$0.7182\pm4\times10^{-5}$
tCNNs	$0.7356\pm3\times10^{-5}$	$0.7445\pm2\times10^{-3}$
DeepCDR	$0.7669\pm3\times10^{-4}$	$0.7641\pm5\times10^{-5}$
MOFGCN	$0.8143\pm1\times10^{-5}$	$0.8145\pm1\times10^{-3}$
GraphCDR	$0.8017\pm2\times10^{-5}$	$0.8049\pm3\times10^{-4}$
NIHGCN	$0.8492\pm4\times10^{-5}$	$0.8501\pm4\times10^{-4}$
MHGCN	$\mathbf{0.8563\pm1\times10^{-4}}$	$\mathbf{0.8612\pm1\times10^{-4}}$

* The font bold indicates the optimal model.

that captures pharmacological interaction patterns between cell lines and drugs. We evaluate the framework on two established pharmacogenomic databases - the Cancer Cell Line Encyclopedia (CCLE) and Genomics of Drug Sensitivity in Cancer (GDSC) - with architecture dimensions scaled to each dataset's characteristics.

The model configuration adapts to dataset scale: CCLE implementations use 512-unit layers for feature projection and HGCN processing, while GDSC requires 1024 units to handle its expanded interaction space. Both datasets employ consistent GCN architectures ($1024 \rightarrow 64$ hidden units) for CDP network analysis, followed by a final linear projection to single-unit sensitivity predictions. Implemented in PyTorch 2.0.1, we optimize the model using Adam (learning rate=0.001) with dataset-adjusted batch sizes (CCLE:256, GDSC:512). The complete computational procedure is formalized in Algorithm 1.

B. Baselines

We benchmark MHGCN against six state-of-the-art approaches in pharmacological interaction prediction:

• **HNMDRP** employs network propagation to integrate multi-omics data from cell lines, drug molecular descriptors, and protein-protein interaction networks for comprehensive CDR prediction.

• **tCNNs** utilizes 1D convolutional neural networks with drug representations encoded as one-hot vectors from SMILES sequences and cell line features derived from somatic mutation profiles to predict pharmacological responses.

• **DeepCDR** implements a hybrid architecture that processes drug chemical structures through uniform graph convolutional networks (U-GCN) while fusing cell line multi-omics data via 1D-CNNs, enabling multimodal feature integration for regression prediction.

• **MOFGCN** constructs cell line-drug heterogeneous graphs using pharmacological similarity metrics, then applies graph neural networks to model interaction for linear correlation-based prediction.

• **GraphCDR** combines graph neural networks for molecular graph feature extraction with deep neural networks for multi-omics data fusion, enhanced by contrastive learning regularization to improve prediction robustness.

• **NIHGCN** constructs a heterogeneous network consisting of drugs and cell lines, using parallel GCN and neighborhood interaction layers to capture nodes features. Finally, drug response prediction is performed by calculating the linear correlation coeffcient between cell line and drug feature representation.

C. Evaluation metrics

We evaluate model performance using two established metrics for imbalanced pharmacological response classification: the Area Under the Receiver Operating Characteristic Curve (AUC-ROC) and the Area Under the Precision-Recall Curve (AUPRC).

The AUC-ROC metric quantifies a classifier's ability to discriminate between sensitive and resistant interactions, representing the probability that a randomly chosen sensitive sample receives a higher prediction score than a resistant one. Ranging from 0 to 1, values of 0.5 indicate random performance while 1.0 represents perfect separation. In clinical applications, AUC-ROC scores exceeding 0.75 are typically considered meaningful for treatment decision support.

AUPRC provides a more focused assessment of predictive performance by evaluating the precision-recall trade-off across classification thresholds. This metric is particularly valuable for pharmacological prediction tasks with class imbalance, as it specifically measures the model's ability to identify the minority class (sensitive responses). Precision reflects the proportion of correctly identified sensitive cases among all predicted positives, while recall indicates the fraction of actual sensitive cases successfully detected. AUPRC values range from 0 to 1, with higher scores indicating better performance, where 1.0 represents ideal classification with complete sensitive case identification and no false positives.

D. Experimental Results

1) Performances evaluated by 5-fold cross-validation: We implement rigorous experimental protocols to ensure fair model comparisons. All models process identical input features: cancer cell line gene expression profiles and drug molecular fingerprints from the filtered datasets described in Section II (Materials and Methods). Specifically, the CCLE dataset comprises 436 cell lines and 24 drugs, while GDSC

TABLE III
THE AUC AND AUPRC RESULTS OF NEW CELL LINES RESPONSE PREDICTION

Dataset	Model	AUC	AUPRC
CCLE	HNMDRP	-	-
	tCNNs	0.7596±3×10⁻4	0.7612±3×10⁻5
	DeepCDR	0.8768±3×10⁻4	0.8807±3×10⁻5
	MOFGCN	0.8172±3×10⁻4	0.8098±3×10⁻5
	GarphCDR	0.7756±3×10⁻4	0.7654±3×10⁻5
	NIHGCN	0.8830±3×10⁻4	0.8776±3×10⁻5
	MHGCN	**0.8889±1×10⁻5**	**0.8875±1×10⁻5**
GDSC	HNMDRP	-	-
	tCNNs	0.7546±3×10⁻4	0.7532±3×10⁻5
	DeepCDR	0.7536±3×10⁻4	0.7468±3×10⁻5
	MOFGCN	0.7242±3×10⁻4	0.7244±3×10⁻5
	GarphCDR	0.7145±3×10⁻4	0.7124±3×10⁻5
	NIHGCN	0.8035±4×10⁻4	0.8105±5×10⁻3
	MHGCN	**0.8112±3×10⁻5**	**0.8129±3×10⁻5**

* The font bold indicates the optimal model.

TABLE IV
THE AUC AND AUPRC RESULTS OF NEW DRUGS RESPONSE PREDICTION.

Dataset	Model	AUC	AUPRC
CCLE	HNMDRP	0.7012±3×10⁻4	0.6983±3×10⁻5
	tCNNs	0.7517±3×10⁻4	0.7551±3×10⁻5
	DeepCDR	0.7210±3×10⁻4	0.7289±3×10⁻5
	MOFGCN	0.7251±3×10⁻4	0.7213±3×10⁻5
	GarphCDR	0.7521±3×10⁻4	0.7425±3×10⁻5
	NIHGCN	0.7522±5×10⁻3	0.7275±4×10⁻4
	MHGCN	**0.7807±1×10⁻5**	**0.7690±1×10⁻5**
GDSC	HNMDRP	0.6252±3×10⁻4	0.6125±3×10⁻5
	tCNNs	0.6427±3×10⁻4	0.6592±3×10⁻5
	DeepCDR	0.6721±3×10⁻4	0.6787±3×10⁻5
	MOFGCN	0.6698±3×10⁻4	0.6812±5×10⁻5
	GarphCDR	0.6713±4×10⁻4	0.6661±2×10⁻5
	NIHGCN	0.7085±4×10⁻4	0.7057±5×10⁻5
	MHGCN	**0.7118±3×10⁻5**	**0.7132±3×10⁻5**

* The font bold indicates the optimal model.

contains 527 cell lines and 100 drugs. Our evaluation employs five repetitions of five-fold cross-validation, where each repetition randomly divides the complete dataset into five non-overlapping subsets. During validation, models train on four combined subsets (80% of data) and evaluate on the remaining subset (20%), guaranteeing each sample appears in a test set exactly once per repetition. To preserve class distribution consistency, we randomly allocate 20% of positive interaction pairs to the test set in each fold, using the remaining positive pairs and corresponding negative samples for training. This procedure yields five distinct training-test splits per repetition, with final performance metrics computed as the average across all 25 evaluations (5 repetitions × 5 folds).

Our comprehensive evaluation demonstrates MHGCN's superior performance against state-of-the-art methods on both CCLE and GDSC databases through rigorous five-fold cross-validation (Tables 1-2). MHGCN achieves outstanding predictive accuracy, with AUC/AUPRC scores of 0.8814/0.8822 (CCLE) and 0.8563/0.8612 (GDSC). These results represent significant improvements over existing approaches: a 0.1665 AUC and 0.171 AUPRC advantage over HNMDRP on CCLE (0.1434/0.1430 on GDSC), and 0.1335 AUC/0.1397 AUPRC gains versus tCNNs on CCLE (0.1207/0.1167 on GDSC).

This performance superiority originates from MHGCN's innovative dual-graph architecture, which synergistically combines: (1) heterogeneous graph convolution employing attention-based message passing to explicitly model pharmacological interactions, and (2) CDP similarity network utilizing multi-hop graph convolution to capture higher-order topological constraints. Unlike conventional approaches (e.g., DeepCDR, MOFGCN) that process these relationships independently, MHGCN's joint optimization framework enables complementary learning of both local pairwise interactions (through the CDP network) and global pharmacological contexts (via the heterogeneous graph), resulting in enhanced prediction robustness.

2) Predicting new cell line or new drug response: A persistent challenge in pharmacogenomic research stems from the inevitable emergence of untested drug-cell line combinations. This occurs both when new therapeutic compounds are developed without comprehensive screening against existing cancer models, and when tumor evolution generates novel cell line variants not represented in current databases. To rigorously assess our model's ability to predict responses for these previously unseen cases, we designed a controlled experiment that mimics real-world drug discovery scenarios through a masked prediction approach. Our evaluation protocol systematically withholds either complete rows (simulating predictions for novel cell lines) or columns (representing new drugs) from the pharmacological interaction matrix, using all remaining known interactions for model training. To ensure statistically meaningful results, we only evaluate cell lines or drugs with at least 10 documented interactions in the database. Specifically, a cell line i participates in the current evaluation round if $\Sigma_j A_{ij} > 10$, and similarly, a drug j is included if $\Sigma_i A_{ij} > 10$. This criterion yields 26 testable cell lines (from 463 total) and 20 drugs (from 24) in CCLE, while GDSC's more comprehensive annotation profile allows inclusion of all 527 cell lines and 100 drugs under the same threshold.

Notably, while HNMDRP's architecture fundamentally cannot predict novel entities due to its dependence on pre-existing interaction networks for feature propagation, MHGCN demonstrates consistent superiority across all evaluation scenarios. On CCLE, MHGCN achieves state-of-the-art performance in novel cell line prediction (AUC=0.8889, AUPRC=0.8875) and novel drug prediction (AUC=0.7807, AUPRC=0.7690), outperforming the strongest baseline NIHGCN by 0.0059/0.0099 and 0.0285/0.0415 respectively. This advantage extends to GDSC, where MHGCN obtains AUC/AUPRC scores of 0.8112/0.8129 for cell line cold-start (improvements of

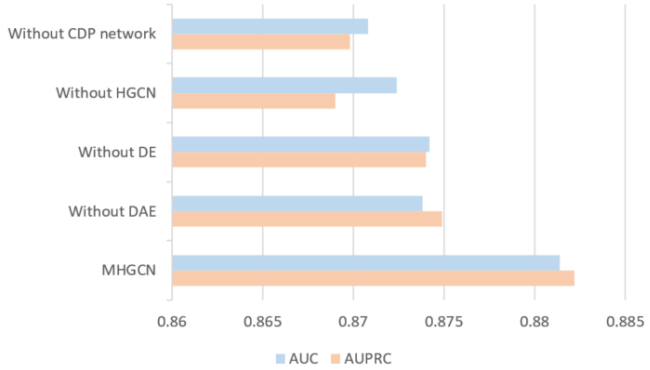

Fig. 3. The result of the ablation experiment. It is apparent that MHGCN performs best when none module is missing

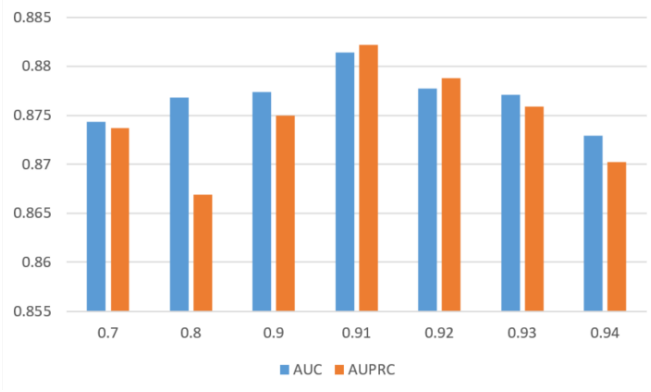

Fig. 4. The results of AUC and AUPRC on different number of edges of CDP network. It is apparent that MHGCN performs best when α is 0.91

0.0077/0.0024 over NIHGCN) and 0.7118/0.7132 for drug cold-start (gains of 0.0033/0.0075), demonstrating robust cross-dataset generalizability through its novel architecture that overcomes the limitations of existing approaches.

MHGCN demonstrates robust generalizability to novel pharmacological entities absent from training data through its innovative dual-graph representation learning architecture. The framework's effectiveness stems from its ability to jointly model cell line-drug interactions in the heterogeneous graph while capturing topological constraints between Cell-Drug Pairs (CDPs), generating transferable embeddings that reveal latent similarity relationships between known and novel entities. Specifically, the architecture employs message passing to propagate molecular features across pharmacologically similar nodes in the heterogeneous graph, complemented by CDP graph convolutions that preserve higher-order interaction patterns through multi-hop neighborhood aggregation. This dual mechanism enables simultaneous learning of intrinsic node attributes and contextual topological signatures, achieving accurate sensitivity predictions even for completely novel cell lines or drugs. Notably, MHGCN outperforms existing approaches in predicting new cell lines in the CCLE database, validating its superior generalization capability.

3) Ablation Experiment: To evaluate the contribution of each module in MHGCN, we conducted ablation studies by systematically removing individual components. Crucially, all variant models retain identical architectural structures and hyperparameter configurations to the original MHGCN, differing only in the exclusion of one specific module per variant.

• **Without CDP network**: We remove the CDP graph that is constructed by the CDP node and only use the construction to learn the features of the cell lines and drugs for CDR prediction.

• **Without HGCN**: We remove the HGCN used for training the cell line-drug response heterogeneous network and directly utilize entity features for data augmentation to CDR prediction.

• **Without Data Enhancement (DE)**: We remove the data enhancement layer after the HGCN layer and directly use the cell lines and drugs feature obtained from the HGCN layer to calculate the similarity matrix of the cell line and drug for CDR prediction.

• **Without DAE**: We remove the DAE layer and directly use the initial feature construction to calculate the cosine similarity of CDP nodes.

Five-repetition five-fold cross-validation on CCLE reveals consistent performance degradation across all ablated configurations (Fig.3), validating the necessity of each module in our architecture. Consistent findings are observed in GDSC under identical parameter settings, where all modules remain indispensable (detailed results in Appendix 1). Notably, ablating the CDP network module exhibits the most significant performance drop, demonstrating its critical role in extracting inter-CDP information to capture deeper drug and cell line characteristics. This further proves that the topological structure between CDPs is essential for CDR prediction. Comparative experiments show that the heterogeneous graph convolution neural network (HGCN) outperforms ordinary GCN in aggregating cell line-drug information. The data enhancement module and DAE-based feature denoising also contribute substantially to model robustness and accuracy.

4) Parameter Sensitivity Analysis: In our architecture, we construct two distinct graph structures, a cell line-drug interaction heterogeneous graph and a CDP similarity graph, which are processed through dedicated heterogeneous GCN (HGCN) and standard GCN modules respectively.

Number of Edges of CDP network The edge count in a CDP network directly corresponds to the number of neighbors for each CDP node. This parameter critically affects model performance: excessive neighbors may induce training overfitting, while insufficient neighbors can lead to inadequate feature learning and information loss. To investigate this balance, we systematically evaluate the impact of edge density by adjusting the threshold parameter α in Equation (10). As shown in Figure 4, model accuracy initially improves then declines as α increases, peaking at $\alpha = 0.91$. This optimal value emerges because lower thresholds prevent effective aggregation of topological information between CDPs via GCN, while higher thresholds ($\alpha > 0.91$) introduce excessive similarity by connecting too many identical neighbors, causing the GCN to aggregate overly redundant information. Our findings demonstrate that while CDP topological characteristics are essential for accurate CDR prediction, excessive aggregation

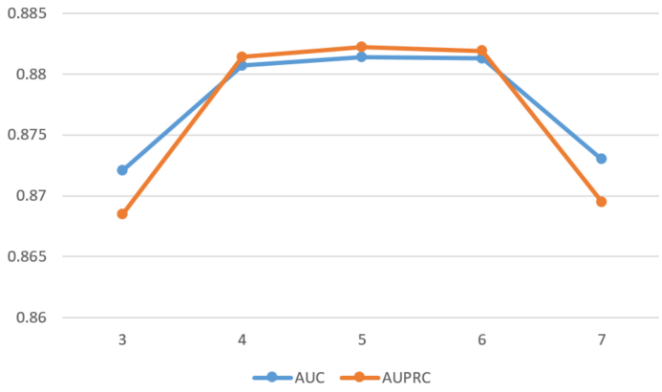

Fig. 5. AUC and AUPRC results on different number of GCN layers. It is apparent that MHGCN performs best when there are 5 layers of GCN.

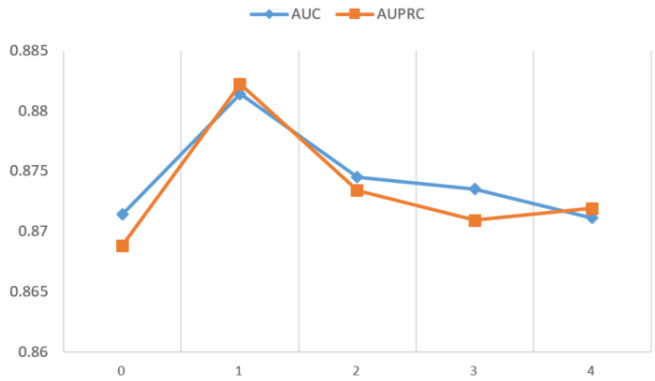

Fig. 6. AUC and AUPRC results on different number of HGCN layers. It is apparent that MHGCN performs best when there is 1 layer of HGCN.

of similar node information can compromise generalization by promoting test overfitting, ultimately degrading prediction performance.

Number of GCN Layers Graph Convolutional Networks (GCNs) extend traditional graph neural networks by introducing convolutional operations, where each additional layer aggregates information from one more hop of neighboring nodes. To determine the optimal message-passing depth for CDP network processing, we evaluate GCN architectures with 3 to 7 layers, as illustrated in Figure 5. Our results demonstrate peak performance at 5 layers, with both shallower and deeper configurations yielding inferior results. With fewer than 5 layers, the model cannot adequately aggregate neighborhood information from cell lines and drugs, failing to capture the complete topological structure essential for accurate CDR prediction. Conversely, architectures exceeding 5 layers exhibit gradually degrading performance, likely due to excessive aggregation of similar CDP information from distant neighbors. This oversmoothing phenomenon, where node representations become overly similar through multiple aggregation steps, can lead to model overfitting and reduced discriminative power. The observed 5-layer optimum suggests a careful balance is required between capturing sufficient neighborhood information and maintaining node distinctiveness in CDP networks.

Number of HGCN Layers The Heterogeneous Graph Convolutional Network (HGCN) extends conventional GCNs to handle heterogeneous networks by aggregating information from diverse neighbor types during each propagation step. To investigate the impact of architectural depth, we evaluate HGCN configurations ranging from 0 to 5 layers, where layer 0 represents replacement with a feedforward neural network (FNN). As shown in Figure 6, optimal performance occurs with a single HGCN layer. The FNN baseline (0 layers) fails to capture critical topological relationships between cell lines and drugs due to its inability to aggregate heterogeneous neighborhood information. This confirms the essential role of pharmacological network structure in CDR prediction. Performance degrades with additional layers (¿1) as excessive message passing leads to oversmoothing - the undesirable homogenization of node representations through repeated aggregation of similar information. Our findings demonstrate that a single, carefully designed HGCN layer optimally balances the incorporation of heterogeneous topological signals while avoiding representation degradation.

IV. Conclusion

We propose MHGCN (Multi-channel Hybrid Graph Convolutional Network), a novel framework for cell line-drug response (CDR) prediction. The model operates through the following key steps: First, it constructs Cell-Drug Pairs (CDPs) by concatenating cell line gene expression profiles with drug molecular fingerprints. These CDPs then undergo denoising via a Denoising Autoencoder (DAE). Second, the framework builds a CDP similarity network by computing cosine similarity between all CDP pairs. Concurrently, it processes cell line and drug features through linear projection into a unified feature space to construct a pharmacological interaction heterogeneous graph. The core innovation of MHGCN lies in its parallel processing of these two graph structures: (1) a standard Graph Convolutional Network (GCN) analyzes the CDP similarity network, while (2) a Heterogeneous GCN (HGCN) processes the interaction graph. The model generates two complementary outputs - a refined response matrix and a similarity matrix - whose fusion produces the final CDR predictions. Through comprehensive experiments, we demonstrate MHGCN's superior performance, particularly in:

- MHGCN demonstrates superior predictive accuracy compared to baseline approaches, achieving this by fusing the topological characteristics of CDP nodes with heterogeneous information of cell lines and drugs. The framework establishes a robust representation learning paradigm for precision oncology applications.
- New cell lines and new drug experiments show that our model can effectively predict CDR in extreme situations, indicating that MHGCN has strong robustness.
- The ablation experiment demonstrates the indispensability of various modules in HMGCN, and shows that our propose data augmentation module can highlight the node characteristics of cell lines and drugs, and improve the accuracy of the model.

Our model can be extended to other task scenarios, such as drug target prediction, protein-protein interaction prediction,

drug-drug interaction prediction and so on. Inspired by previous algorithms, we believe that fusing multi omics information from cell lines could improve the effectiveness of the model. Therefore, in the future, we will improve our model from these directions.

V. Funding

This work was supported by the "Artificial Intelligence Promotes Scientific Research Paradigm Reform and Empowers Discipline Leapfrogging Program". The funders have no role in study design, data collection, data analysis, data interpretation or writing of the manuscript.

Conflict of Interest: The authors declare no competing financial interest.

References

[1] S. Noorolyai, A. Baghbani, E. Asadi, M. B. Kojabad, A. Mogaddam, M. M. Baradaran and B. Behzad, "The role of microRNAs involved in PI3-kinase signaling pathway in colorectal cancer," *J. Cellular Physiol*, vol. 234, no. 5, pp. 5664-5673, May 2019.

[2] X. Sagaert, A. Vanstapel and S. Verbeek, "Tumor heterogeneity in colorectal cancer: What do we know so far?," *Pathobiology*, vol. 85, no. 1, pp. 72-84, Jan. 2018.

[3] M. A. Rubin, "Health: Make precision medicine work for cancer care," *Nature*, vol. 520, no. 7547, pp. 290-291, Apr. 2015.

[4] J. P. Lloyd, M. B. Soellner, S. D. Merajver and J. Z. Li, "Impact of between-tissue differences on pan-cancer predictions of drug sensitivity," *PLoS Comput. Biol.*, vol. 17, no. 2, Feb. 2021, doi: 10.1371/journal.pcbi.1008720.

[5] J. Barretina et al., "The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity," *Nature*, vol. 483, pp. 603-607, Mar. 2012.

[6] W. Yang et al., "Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells," *Nucleic Acids Res.*, vol. 41, pp. D955-D961, Jan. 2013.

[7] Z. Stanfield, M. Coşkun and M. Koyutürk, "Drug Response Prediction as a Link Prediction Problem," *Sci. Rep.*, vol. 7, Jan. 2017, doi: 10.1038/s41598-017-06631-9.

[8] T. Turki and Z. Wei, "A link prediction approach to cancer drug sensitivity prediction," *BMC Syst. Biol.*, vol. 11, no. 5, pp. 1-14, Jan. 2017.

[9] I. Lee and H. Nam, "Identification of drug-target interaction by a random walk with restart method on an interactome network," *BMC Bioinformatics*, vol. 19, no. 8, p. 208, Jan. 2018.

[10] N. Q. Zhang et al., "Predicting Anticancer Drug Responses Using a Dual-Layer Integrated Cell Line-Drug Network Model," *PLoS Comput. Biol.*, vol. 11, no. 9, e1004498, Sep. 2015.

[11] F. Zhang, M. Wang, J. Xi, J. Yang and A. Li, "A novel heterogeneous network-based method for drug response prediction in cancer cell lines," *Sci. Rep.*, vol. 8, no. 1, p. 3355, Feb. 2018.

[12] C. Huang, R. Mezencev, J. F. McDonald and F. Vannberg, "Open source machine-learning algorithms for the prediction of optimal cancer drug therapies," *PLoS One*, vol. 12, no. 10, e0186906, Oct. 2017.

[13] J. Yang, A. Li, Y. Li, X. Guo and M. Wang, "A novel approach for drug response prediction in cancer cell lines via network representation learning," *Bioinformatics*, vol. 35, no. 9, pp. 1527-1535, May 2019.

[14] A. P. Lind and P. C. Anderson, "Predicting drug activity against cancer cells by random forest models based on minimal genomic information and chemical properties," *PLoS One*, vol. 14, no. 7, e0219774, Jul. 2019.

[15] R. Su, X. Y. Liu, L. Y. Wei and Q. Zou, "Deep-Resp-Forest: A deep forest model to predict anti-cancer drug response," *Methods (San Diego, Calif.)*, vol. 166, pp. 91-102, Jan. 2019.

[16] Y. Liang, D. Zhou, L. Gao and Y. Zha, "Prediction of drug response in multilayer networks based on fusion of multiomics data," *Methods (San Diego, Calif.)*, vol. 192, pp. 85-92, May 2021.

[17] H. Gerdes et al., "Drug ranking using machine learning systematically predicts the efficacy of anti-cancer drugs," *Nature communications*, vol. 12, no. 1, p. 1850, Mar. 2021.

[18] P. F. Liu, H. J. Li, S. Li and K. S. Leung, "Improving prediction of phenotypic drug response on cancer cell lines using deep convolutional network," *BMC Bioinformatics*, vol. 20, no. 1, p. 408, Aug. 2019.

[19] L. Rampášek, D. Hidru, P. Smirnov, B. Haibe-Kains and A. Goldenberg, "Dr.VAE: Improving drug response prediction via modeling of drug perturbation effects," *Bioinformatics*, vol. 35, no. 19, pp. 3743-3751, Oct. 2019.

[20] P. F. Liu, H. J. Li, S. Li and K. S. Leung, "Improving prediction of phenotypic drug response on cancer cell lines using deep convolutional network," *BMC Bioinformatics*, vol. 20, no. 1, p. 408, Aug. 2019.

[21] Q. Liu, Z. Hu, R. Jiang and M. Zhou, "DeepCDR: A hybrid graph convolutional network for predicting cancer drug response," *IEEE J. Biomed. Health Inform.*, vol. 24, no. 1, pp. 367-374, Jan. 2020.

[22] Q. Li, J. Huang, H. Zhu and Q. Liu, "Prediction of cancer drug effectiveness based on multi-fusion deep learning model," *2020 10th Annual Computing and Communication Workshop and Conference (CCWC)*, Jan. 2020, pp. 330-335.

[23] M. Li, Y. Wang, R. Zheng, X. Shi, Y. Li, F. Wu and J. Wang, "DeepDSC: A deep learning method to predict drug sensitivity of cancer cell lines," *IEEE/ACM Trans. Comput. Biol. Bioinform.*, vol. 18, no. 2, pp. 575-582, Mar-Apr. 2021.

[24] J. Choi, S. Park and J. Ahn, "RefDNN: A reference drug based neural network for more accurate prediction of anticancer drug resistance," *Sci. Rep.*, vol. 10, no. 1, p. 1861, Feb. 2020.

[25] M. Li, Y. Wang, R. Zheng, X. Shi, Y. Li, F. Wu and J. Wang, "DeepDSC: A deep learning method to predict drug sensitivity of cancer cell lines," *IEEE/ACM Trans. Comput. Biol. Bioinform.*, vol. 18, no. 2, pp. 575-582, Mar-Apr. 2021.

[26] Z. Tian, Y. Yu, Fengming Ni and Q. Zou. Drug-target interaction prediction with collaborative contrastive learning and adaptive self-paced sampling strategy. BMC Biology. 2024, 22: 216.

[27] W. Song, . Xu, C. Han, Z. Tian and Quan Zou. Drug-target interactionpredictions with multi-view similarity network fusion strategy and deepinteractive attention mechanism. Bioinformatics.2024, 40(6): btae346.

[28] W. Peng, T. Chen and W. Dai, "Predicting drug response based on multi-omics fusion and graph convolution," *IEEE J. Biomed. Health Inform.*, vol. 26, no. 3, pp. 1384-1393, Mar. 2022.

[29] J. Shin, Y. Piao, D. Bang, S. Kim and K. Jo, "DRPreter: Interpretable anticancer drug response prediction using knowledge-guided graph neural networks and transformer," *Int. J. Mol. Sci.*, vol. 23, no. 22, p. 13919, Nov. 2022.

[30] W. Peng, H. Liu, W. Dai, N. Yu and J. Wang, "Predicting cancer drug response using parallel heterogeneous graph convolutional networks with neighborhood interactions," *Bioinformatics*, vol. 38, no. 19, pp. btac574, Oct. 2022.

[31] X. Liu, C. Song, F. Huang, H. Fu, W. Xiao and W. Zhang, "GraphCDR: A graph neural network method with contrastive learning for cancer drug response prediction," *Brief. Bioinform.*, vol. 23, no. 1, pp. bbab457, Jan. 2022.

[32] J. Choi, J.-K. Rhee and H. Chae, "Cell subtype classification via representation learning based on a denoising autoencoder for single-cell RNA sequencing," *IEEE Access*, vol. 9, pp. 14540-14548, Jan. 2021.

[33] A. Brahms, "Representation error for real numbers in binary computer arithmetic," IEEE Computer Group Repository, Paper R-67-85.

[34] J. Liu, S. Guan and Q. Zou, Hongjie Wu*,Prayag Tiwari, Yijie Ding. AMDGT: Attention aware multi-modal fusion using adual graph transformer for drug–disease associations prediction. Knowledge-BasedSystems. 2024, 284: 111359.

[35] J. Staunton, D. Slonim, H. Coller, P. Tamayo, M. Angelo, J. Park, U. Scherf, J. Lee, W. Reinhold, J. Weinstein, J. Mesirov, E. Lander, T. Golub.Chemosensitivity prediction by transcriptional profiling.[J].Proceedings of the National Academy of Sciences of the United States of America,2001,Vol.98(19): 10787-10792