GPDRP: a multimodal framework for drug response prediction with graph transformer

Yingke Yang[1] and Peiluan Li[1,2]*

*Correspondence:
15038522015@163.com

[1] School of Mathematics
and Statistics, Henan University
of Science and Technology,
Luoyang 471000, China
[2] Longmen Laboratory,
Luoyang 471003, China

Abstract

Background: In the field of computational personalized medicine, drug response prediction (DRP) is a critical issue. However, existing studies often characterize drugs as strings, a representation that does not align with the natural description of molecules. Additionally, they ignore gene pathway-specific combinatorial implication.

Results: In this study, we propose drug Graph and gene Pathway based Drug response prediction method (GPDRP), a new multimodal deep learning model for predicting drug responses based on drug molecular graphs and gene pathway activity. In GPDRP, drugs are represented by molecular graphs, while cell lines are described by gene pathway activity scores. The model separately learns these two types of data using Graph Neural Networks (GNN) with Graph Transformers and deep neural networks. Predictions are subsequently made through fully connected layers.

Conclusions: Our results indicate that Graph Transformer-based model delivers superior performance. We apply GPDRP on hundreds of cancer cell lines' bulk RNA-sequencing data, and it outperforms some recently published models. Furthermore, the generalizability and applicability of GPDRP are demonstrated through its predictions on unknown drug-cell line pairs and xenografts. This underscores the interpretability achieved by incorporating gene pathways.

Keywords: Drug response prediction (DRP), Multimodal deep learning model, Graph transformer, Drug molecular graphs, Pathway activity scores

Background

Cancer, a highly complex disease, is caused by the interaction of various carcinogenic factors. It significantly impacts global human health and poses a threat to human life. Individuals with the disease exhibit heterogeneity in both genetic and phenotypic aspects, primarily due to the tumor microenvironment's clonal diversity of cancer cells and non-malignant cells with changed phenotypes. This heterogeneity leads to partial or non-responsiveness of certain patients to therapeutic strategies such as chemotherapy, targeted therapy, and immunotherapy during the cancer treatment process [1]. In other words, even when implementing the same therapeutic strategies for patients of the same cancer type, there are still variations in treatment responses, making responses to cancer treatment generally unpredictable. Additionally, it is important to note that

not all cancers and anticancer drugs are strongly associated with targetable genetic biomarkers. Therefore, relying solely on the relationship between drug targets or mutation status may be insufficient to predict the efficacy of specific targeted therapies [2, 3]. And implementing targeted therapies without taking drug resistance into account may reduce patient survival rates. Drug resistance may show up as the activation of alternative signaling pathways promoting tumor growth or clonal expansion under the selective pressure induced by treatment [4]. Therefore, drug response prediction (DRP) is critically important in cancer therapy and has become a significant topic in personalized medicine research. Accurate prediction of treatment response assists in designing more effective treatment plans for patients and provides valuable insights for the development of novel disease-inhibiting drugs.

With the rapid development of high-throughput genomics technologies, large-scale pharmacogenomics databases have gradually accumulated. The Cancer Cell Line Encyclopedia (CCLE) [5] provides a platform for systematic study of cell lines. The Genomics of Drug Sensitivity in Cancer (GDSC) [6] is one of the largest public databases, covering information regarding the sensitivity of cancer cells to drugs and related molecular markers. The Cancer Therapeutics Response Portal (CTRPv2) [7] provides extensive data on drug sensitivity. These high-throughput screening research resources collectively form a vast knowledge base [1]. Based on these abundant data resources, numerous researchers have established various DRP models to predict the response of anticancer drugs.

Menden et al. compared anticancer drug sensitivity prediction models constructed using different methods by utilizing two large-scale drug genomics datasets, and demonstrated that genomics can validate the response of specific drugs as an explanatory variable [8]. Ammad-Ud-Din et al. employed a novel nuclear norm-based Bayesian matrix factorization approach that combined drug chemical structure features and genomic characteristics for DRP [9]. Zhang et al. introduced an integrated model to predict drug response in a specified cell line and demonstrated its superiority over the elastic net model [10]. Wang et al. predicted drug response by utilizing the chemical structure of drugs and gene expression profiles, employing a similarity regularized matrix factorization method [11]. Chang et al. proposed the CDRscan, which utilizes cell lines' genomic mutations and molecular fingerprints of drugs for predicting drug efficacy [12]; Sakellaropoulos et al. constructed a model named Precily based on gene expression data for DRP and demonstrated its superior performance over Elastic Net and Random Forest models [13]; Choi et al. introduced an innovative deep neural network model named RefDNN for better drug resistance prediction and biomarker identification related to drug response [14].

Despite significant progress in DRP research, there are some issues worth considering. For instance, most studies represent drugs as strings, which is an unnatural way of representing molecules and may result in the loss of structural information [15]. Additionally, the pathway-specific combinatorial implication (or gene sets) of genes are disregarded, and gene expression levels are treated as independent variables, which may overly emphasize machine learning techniques [16, 17].

To address these issues, we propose GPDRP (Graph and Pathway based Drug response prediction), a novel multimodal deep learning architecture, that can predict

drug responses on cell lines by modeling drugs as molecular graphs. In addition, Graph Transformer was combined with Graph Isomorphism Network (GIN) to improve the capacity for more precise DRP. We compared GPDRP with two recently published works: Precily [18], which represents drug moleculars using simplified molecular-input line-entry system (SMILES) strings, and GraTransDRP [19], encoding cell lines' genomic and epigenomic characteristics through one-hot encoding. Our approach performs better considering root mean square error (RMSE) and Pearson correlation coefficient (PCCs), according to experimental results. Also, by applying GPDRP to 15,094 drug-cell line pairs lacking response values and xenograft datasets, we demonstrated the potential of the model to predict unknown drug-cell line pairs, as well as the applicability of the model and interpretability using gene pathway scores. The primary contributions of GPDRP include:

1. We integrate the drug molecular graph with gene pathway activity score, leveraging the strengths of both types of data to enhance the predictive power of our model.
2. We introduce GPDRP, a novel multimodal framework for DRP, which leverages Graph Convolutional Networks in conjunction with Graph Transformer and deep neural networks. The performance of GPDRP is demonstrated using the CCLE/ GDSC dataset, and it outperforms two recently published models, Precily and GraTransDRP.
3. GPDRP demonstrates the potential to predict unknown drug-cell line pairs. It was utilized to predict the pairs that were missing from the GDSC, and some published works were located and discussed that supported our predictions.
4. GPDRP exhibits excellent applicability. We applied it to predict the LNCaP xenograft dataset and provided explanations based on gene activity pathway scores.

Results

Performance comparison on the CCLE/GDSC dataset

To assess GPDRP's prediction accuracy, we trained the model using the CCLE/GDSC dataset and employed the same data splitting strategy as in Precily [18]. We separated the dataset according to the cell lines, making sure that the test, validation, and training sets did not share any cell lines. Of the total drug-cell line pairs (80,056), we randomly selected 90% (72,156) for the dataset, with 80% of cell lines allocated to the training set and 10% to the validation set for hyperparameter tuning. The remaining 10% (7900) of the pairs were designated for the testing set. The test results revealed a PCCs value of 0.8833 and a RMSE value of 0.0321 in the best model as shown in Fig. 1.

We then compared GPDRP with some recently published models. For methods relying on the identical dataset, PCCs and RMSE were computed. The performance is shown in Fig. 1B and Table 1. Evidently, our model GPDRP outperforms Precily and GraTransDRP for almost all graph convolutional networks. Among three GNN models: Graph Convolutional Networks (GCN), Graph Attention Networks (GAT), and GIN, the GIN model performed the best, achieving a PCCs of 0.8827. This illustrates GIN's potential for graph representation and lends credence to the idea that GIN is one of the most potent GCN models [20]. Therefore, we considered combining GIN with the Graph Transformer, resulting in the best PCCs of 0.8833 and the best RMSE of 0.0321.
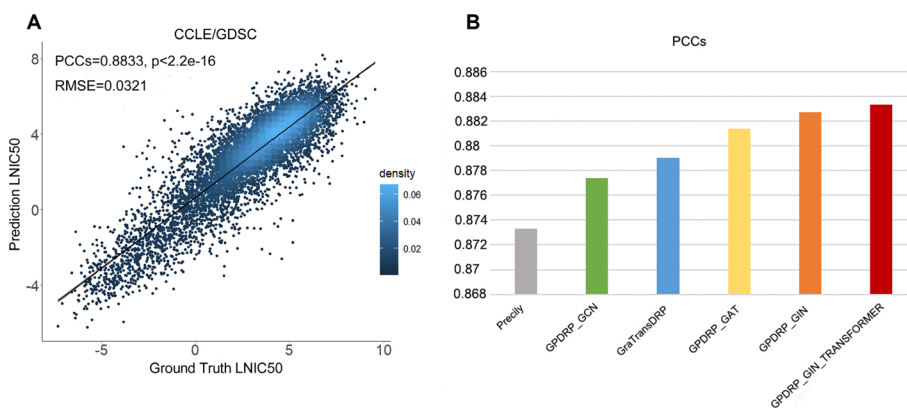
Fig. 1 Performance comparison. **A** Scatter plot demonstrating the performance of GPDRP across all drug-cell line pairs in the CCLE/GDSC test data. P-value was calculated using a two-sided t-test. **B** Barplot shows the Pearson's correlation coefficients (PCCs) for different models

Table 1 The performance comparison of PCCs and RMSE on the GDSC/CCLE dataset (the best performance is in bold)

Model	PCCs	RMSE
Precily [18]	0.8733	1.3773
GraTransDRP [19]	0.8790	0.0333
GPDRP_GCN	0.8774	0.0325
GPDRP_GAT	0.8814	0.0322
GPDRP_GIN	0.8827	0.0323
GPDRP_GIN_TRANSFORMER	**0.8833**	**0.0321**

Prediction of responses for unknown drug-cell line pairs

In this part, we used the optimal model, GPDRP_GIN_TRANSFORMER, to predict the response for the processed 15,094 drug-cell line pairs lacking response values (see Additional file 1: Table S3). All the prediction results are provided in Additional file 1: Table S4. The predicted LN IC50 values for the unknown response pairs grouped by drug are displayed in Fig. 2 using a box plot. Drugs are sorted by the median of their distributions, with each drug's box representing the numerical distribution of values associated with its corresponding cell lines. The figure displays six drugs with the highest values and six drugs with the lowest medians. As the true values for these unknown response pairs are unavailable, the accuracy of our prediction is determined by works as follows.

The LN IC50 is the logarithm of the concentration IC50 at which a drug inhibits biological activity. A smaller value indicates greater sensitivity of the cell lines to the drug, indicating its effectiveness. Our predictions identified the top six most effective drugs as Bortezomib, Daporinad, Vinblastine, Vinorelbine, Paclitaxel, and Vincristine. It is noteworthy that Bortezomib, Vinblastine, Paclitaxel, and Vincristine were also identified as potentially effective drugs in the pioneering model proposed by Liu et al. [21].

Our analysis identified Bortezomib as the most potent drug. Bortezomib has demonstrated extensive antitumor activity and has been shown to enhance the efficacy
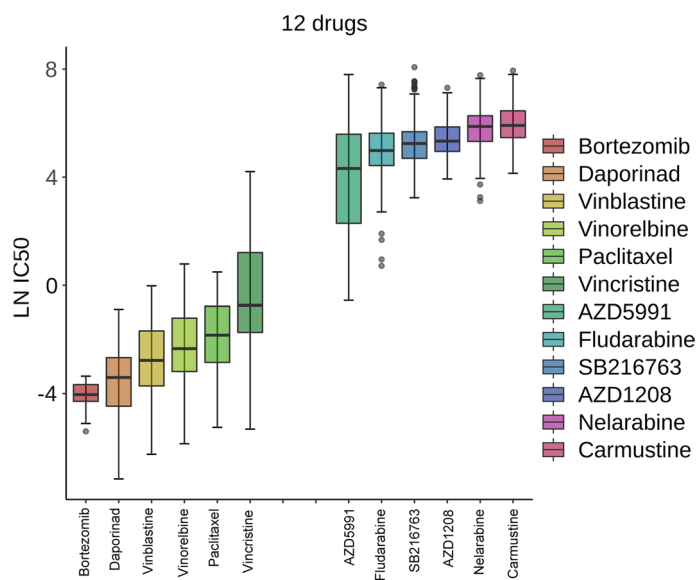
Fig. 2 Box plot of predicted LN IC50 values for unknown response pairs. The drugs are arranged based on the median of their predicted LN IC50 values for cell lines. The horizontal axis denotes the drug names, and the vertical axis denotes their LN IC50 values with cell lines. The top 6 drugs with the lowest median LN IC50 values indicate that they may be the most effective drugs, while 6 drugs with the highest median LN IC50 values suggest that they may be the most ineffective drugs

of various chemotherapeutic drugs [22]. Its capacity to sensitize cell lines to numerous other drugs was noted in a study by Friedman et al. [23]. Bortezomib, the initial proteasome inhibitor authorized for the treatment of malignant diseases, is approved for addressing multiple myeloma and mantle cell lymphoma. It has shown positive clinical outcomes as a standalone treatment or as part of a combination therapy, enhancing the effects of chemotherapy/radiation or overcoming drug resistance [24]. Notably, in our predictions, the DOHH2 cell line exhibited the highest sensitivity to Bortezomib among all the unknown drug-cell line combinations. DOHH2, a human non-Hodgkin lymphoma cell line, is frequently used in lymphoma research, and there is evidence supporting Bortezomib's potential in treating non-Hodgkin lymphoma [25].

Daporinad, a potential small molecule compound, exhibits anti-tumor and antiangiogenic properties. It binds to and inhibits nicotinamide phosphoribosyltransferase (NMPRTase), thereby suppressing the biosynthesis of nicotinamide adenine dinucleotide (NAD+) from nicotinamide (vitamin B3). This activity has the potential to exhaust energy reserves in metabolically active tumor cells and trigger apoptosis. Furthermore, Daporinad may hinder the production of vascular endothelial growth factor (VEGF) in tumor cells, thereby inhibiting tumor angiogenesis. Daporinad has been clinically tested for treating melanoma, cutaneous T-cell lymphoma, and B-cell chronic lymphocytic leukemia [26].

Vinblastine is employed in treating various cancers, including breast cancer, testicular cancer, lymphoma, neuroblastoma, Hodgkin's and non-Hodgkin's lymphoma, as well as fungal infections, histiocytosis, and Kaposi sarcoma [27]. Research by Brugie'res et al. suggested that vinblastine might be effective in treating relapsed anaplastic large cell

lymphoma, leading to durable remissions [28]. Vinorelbine, another vinca alkaloid drug, is frequently employed in cancer therapy, encompassing non-small cell lung cancer and breast cancer [29]. Vincristine has maintained a steady role in cancer therapy research, being an integral part of anti-cancer treatment [30].

Paclitaxel, often referred to as an 'anti-cancer superstar,' is a naturally occurring secondary metabolite extracted and purified from the bark of the yew tree, Taxus brevifolia. It has been clinically validated to possess superior anti-tumor properties and is widely used in treating malignancies such as breast cancer, ovarian cancer, and gastric cancer. It is one of the most frequently used chemotherapeutic drugs in clinical practice [31].

The six least effective drugs identified in our study are AZD5991, Fludarabine, SB216763, AZD1208, Nelarabine, and Carmustine. A literature review revealed that these drugs are typically used in combination therapies. For example, Nelarabine is used to treat relapsed or refractory T-cell acute lymphoblastic leukemia (T-ALL) and T-cell lymphoblastic lymphoma (T-LBL) following the failure of at least two previous treatment regimens [32]. Fludarabine can have significant side effects, and careful monitoring of hematologic and non-hematologic toxicities is recommended when used as an anti-cancer drug.

In conclusion, our method has shown exceptional performance in predicting drug responses for unknown drug-cell line pairs, thereby confirming the accuracy and practicality of GPDRP. This allows us to better understand the effects of drugs on specific cell lines, offering robust support for drug development and the creation of personalized treatment strategies.

Predictions in LNCaP xenografts

Patient-Derived Xenografts (PDXs) are widely used in vivo tumor models to investigate therapeutic responses and forecast drug responses in cancer patients sharing analogous traits. In our study, we applied the GPDRP method to analyze the GSE211856 dataset, which was obtained from the NCBI GEO database (www.ncbi.nlm.nih.gov/geo/). This dataset comprises bulk RNA-seq data from an extensively annotated study on the progression of prostate cancer, focusing on the responsiveness and development of resistance to AR-targeted therapies. Androgens are required for the establishment and early growth of LNCaP xenograft tumors in male mice (pre-castration group, PRE-CX). Castration reduces androgen receptor (AR) activity and tumor growth (post-castration group, POST-CX). This initial sensitivity to castration, however, consistently progresses to castration resistance (castration-resistant prostate cancer, CRPC). Further treatment of CRPC with the AR targeting drug enzalutamide (ENZ) produces an initial therapeutic reaction (ENZ Sensitive, ENZS), but resistance develops over time (ENZ Resistant, ENZR). The dataset includes a total of 54 samples, encompassing multiple biological replicates for each condition and treatment group, as summarized in Table 2.

To predict drug responses, we used the GPDRP_GIN_TRANSFORMER model trained on the CCLE/GDSC dataset. By applying this model to the 54 samples (see Additional file 1: Table S5), we obtained the predicted sensitivity of 173 drugs on LNCaP xenograft tumor samples, as depicted in Fig. 3. As our response values are continuous and Z-score normalized with a mean of 0 and a standard deviation of 1, we employed Euclidean distance for clustering analysis to enable comparison on a consistent scale. Figure 4 reveals

Table 2 GSE211856 dataset overview

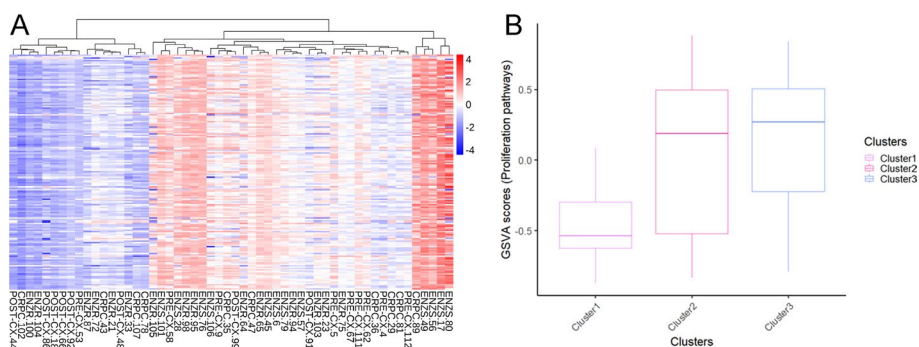| Sample	Sample Size	Group
PRE-CX	9	Condition
POST-CX	8	Condition
CRPC	10	Condition
ENZS	12	treatment
ENZR	15	treatment

Fig. 3 Predictions and analysis in LNCaP xenografts. **A** Heatmap represents the predicted LN IC50 values of 173 drugs across the 54 samples, where lower LN IC50 values are indicated by bluer color bars, indicating greater sensitivity of the predicted samples to the drugs. The samples were grouped based on the Euclidean distance. **B** Boxplots showing the distribution of GSVA scores of proliferation-related pathways (n = 12) across three clusters (n = 12, n = 25 and n = 17 samples from cluster 1, cluster 2 and cluster 3, respectively)
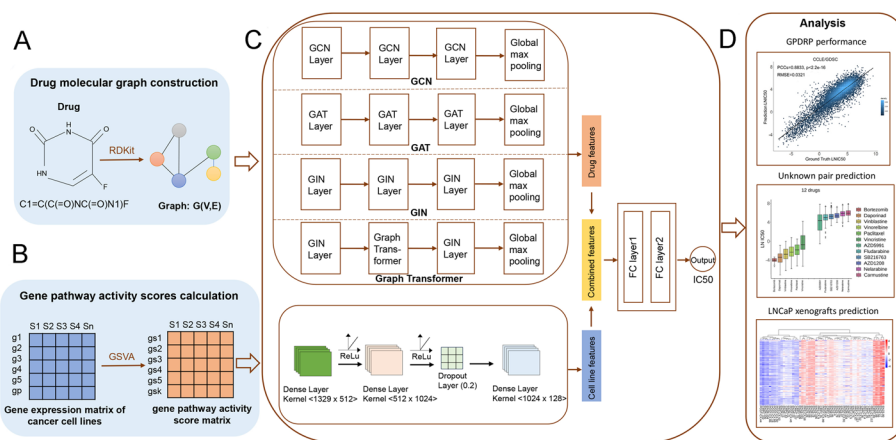

Fig. 4 Illustration of the predictive analysis workflow of GPDRP. **A** Drug molecular graph construction. The structure information of drugs was collected from PubChem and we represented drugs as molecular graphs using RDKit. **B** Gene pathway activity scores calculation. For the cancer cell lines obtained from CCLE, we computed pathway activity scores for canonical pathways using GSVA. **C** Two subnetworks for learning drug features and cell line features respectively. GPDRP took molecular graphs of drugs and gene pathway activity scores of cell lines as inputs to the drug subnetworks and cell line subnetworks, respectively. The two representations are then concatenated and put through two FC layers to predict the response. **D** Results and downstream analysis of this work. Including performance comparison, prediction of unknown drug-cell line response and predictions in LNCaP xenografts

the division of the 54 samples into three main clusters. We summarized the samples with the highest predicted values as Cluster 1, shown as the most red-colored region in Fig. 3A. This cluster exhibits the strongest drug resistance, indicating the lowest drug sensitivity, and is mainly composed of tumor samples treated with ENZ (a total of 12 samples, with 7 ENZS and 3 ENZR). Conversely, we summarized the samples with the lowest predicted values as Cluster 3, shown as the most blue-colored region in the figure, which demonstrates the highest sensitivity to the 173 drugs. Notably, ENZR samples are distributed across all three clusters, suggesting heterogeneity in treatment outcomes and implying that ENZ resistance may involve different underlying mechanisms, potentially involving interactions with stromal components in the tumor microenvironment [18].

To further elucidate the clustering results, we focused on pathway activity scores related to cell proliferation, as shown in Fig. 3B. The pathways related to cell proliferation that we utilized are provided in Additional file 1: Table S6. Box plots were utilized to illustrate the variances in pathway activity scores among the three clusters. Cluster 1 exhibited the lowest pathway activity scores in cell proliferation-related pathways, which may account for the lowest sensitivity to drug responses in this cluster. Conversely, Cluster 3 displayed the highest pathway activity scores, indicating a higher proliferation index, thereby explaining the increased sensitivity to drug responses in this cluster. Therefore, the use of gene activity scores makes the model results more interpretable.

Discussion

Accurate prediction of drug response in cancer cells is pivotal for personalized oncology. This work introduces GPDRP, a multimodal deep learning framework leveraging the Graph Transformer architecture to forecast the response to cancer treatment, utilizing information from both drug molecular graphs and gene pathway activity. We employed four GNN variants: GCN, GAT, GIN and Graph Transformer with the combination of GIN, used for learning drug features. Subsequently, the drug-cell line pairs were used to predict LN IC50 values. Notably, our model combines drug molecule graphs with gene pathway activity scores, outperforming some recently published methods in terms of performance comparisons based on RMSE and PCCs.

The experimental results indicate that GPDRP outperforms in terms of RMSE and PCCs. Through performance comparison, we believe that representing drugs using graphical structures may preserve the essence of their chemical structures, making it more appropriate than using strings. In this experiment, GPDRP_GIN_TRANS-FORMER demonstrated superior performance, possibly due to the addition of a Graph Transformer layer. Firstly, the multi-layer feature extraction capabilities of GIN and Graph Transformer complement each other. GIN excels in capturing local neighborhood features, while the Graph Transformer layer effectively captures long-range dependencies and global relationships among nodes through its self-attention mechanism. The combination of these two layers enables the model to learn more comprehensive and informative graph structure features. Secondly, the integration of local and global information enhances the model's representational power. GIN's neighborhood aggregation process may overlook long-distance relationships, which can be effectively addressed by the Graph Transformer layer's ability to capture global dependencies. By incorporating the Graph Transformer layer after the GIN layers, the model achieves a better

fusion of local and global information. This integration leverages the strengths of both models, allowing for comprehensive feature extraction, effective combinations of local and global information, and improved generalization abilities on graph data. Furthermore, when predicting responses for drug-cell line pairs with unknown responses in the CCLE/GDSC dataset, we identified Bortezomib, Daporinad, Paclitaxel, and vinca alkaloids as drugs with the lowest response values, highlighting their anti-tumor properties. Conversely, drugs with the highest response values exhibited lower sensitivity to cancer, illustrating the model's potential to learn from data and predict responses for new drug-cell line pairs. We further demonstrate the applicability of GPDRP in LNCaP xenografts and its interpretability using gene pathway activity scores.

One limitation of GPDRP is the interpretability of the model. We employ GNN to learn the latent features of drug molecular graphs. While Nguyen et al. [15] demonstrated that GNN can assign significance to clearly defined chemical features automatically without prior knowledge, the majority of the learned latent variables still defy explanation using available descriptors (specific details provided in Additional file 1: Supplementary Materials C). Furthermore, our study solely focuses on cell lines, and when it comes to data splitting based on drug compounds, the model falls short of achieving the anticipated outcomes (specific details provided in Additional file 1: Supplementary Materials D). This may be attributed to the vast chemical space of drug compounds. In the future, we will place a particular emphasis on researching model interpretability and give greater attention to drug-based research to enhance the model's interpretability and improve its effectiveness in predicting drug responses. Additionally, RGCN and RGAT may enhance the predictive capabilities of the model, and we will explore their use to achieve better predictive performance.

Conclusions

In this paper, we propose a multimodal deep learning model, GPDRP, which enables more accurate prediction of drug responses. By employing drug molecular graphs as the representation of drugs and leveraging GNN with Graph Transformer for feature extraction, this approach may better preserve the structural information of drug molecules, enhancing the model's understanding and predictive capability of drug features. Furthermore, through the incorporation of gene pathway activity scores, GPDRP provides valuable interpretability. The introduction of this model holds significant implications, offering a precise tool for personalized medicine and cancer treatment, and driving advancements in cancer research.

Methods

We propose a multimodal deep learning architecture, called GPDRP for DRP. The DRP problem is formulated as a regression task, wherein a drug-cell line pair serves as the input and a continuous measurement of the response value LN IC50 of that pair serves as the output. Molecular graphs are used to represent drugs, which allows the model to directly capture atom-to-atom bonds. GPDRP is trained using the Pytorch [33]. Figure 4 illustrates the proposed framework.

Data acquisition

For comparison purposes, we followed the same procedure as in Precily [18], obtaining 550 CCLE cell lines' bulk RNA-seq gene expression profiles that overlap with the GDSC2 dataset of the GDSC database. The relevant response data was extracted from the GDSC2 dataset. We collected information on drug responses for 173 compounds, and their SMILES notations were retrieved using PubChemPy [34]. Specific data processing is provided in Additional file 1: Supplementary Materials A.

Drug molecular graph construction

For drug features, we perceive drug compounds as graphs depicting the interactions among atoms. Firstly, 173 molecular compounds' chemical structure data was obtained in terms of a Canonical SMILES using PubChemPy (see Additional file 1: Table S1). Then using the open-source cheminformatics program RDKit [35], we translated the Canonical SMILES into the corresponding molecular graphs and extracted atomic features. We employed a collection of atomic attributes adapted from DeepChem [36] to characterize a node in the molecular graph. Each node is represented as a multidimensional binary feature vector conveying five distinct pieces of information: the atomic symbol, the number of neighboring atoms, the number of neighboring hydrogen atoms, the implicit valence of the atom, and whether the atom is part of an aromatic structure. The presence of a bond between a pair of atoms triggers the establishment of an edge. Consequently, an indirect binary graph, comprising nodes endowed with associated attributes, is constructed for each input Canonical SMILES.

Gene pathway activity scores calculation

For cell lines features, we used pathway activity scores (see Additional file 1: Table S2). Based on the gene expression matrix, we computed Gene Set Variation Analysis (GSVA) scores using the GSVA [37] R software package, utilizing 1329 gene sets from the Molecular Signatures Database (MSigDB) [38] make up the c2 canonical pathway collection (MSigDB.CP.v.6.1, see Additional file 1: Supplementary), with min.sz set to 5. By calculating GSVA scores, we transformed the gene expression matrix into a GSVA score matrix comprising 1329 pathway activity scores and 550 cell lines. The resulting GSVA score matrix served as the cell line feature matrix. To enhance the convergence and stability of the model, each feature is normalized to the [0,1] range using min−max scaling. For the kth cell line on the ith pathway, the normalization is performed as follows:

$$\hat{x}_{ik} = \frac{x_{ik} - \min(x_i)}{\max(x_i) - min(x_i)},$$

where x_{ik} represents the kth cell line's pathway activity score on the ith pathway, while $min(x_i)$ and $max(x_i)$ respectively denote the minimum and maximum values of pathway i across all cell lines.

Processing of the response variable

After processing the drug and cell line data, we obtained 95,150 drug-cell line pairs. There were 15,094 pairs for which corresponding response values LN IC50 are not

available in the GDSC database, and 80,056 pairs have LN IC50. Therefore, we used these 80,056 pairs along with their corresponding response values for model training and testing. Additional file 1: Supplementary Materials B contained the dataset's summary statistics. In addition, we scaled the drug response values LN IC50 values within the range of (0,1) to facilitate the training. For a given LN IC50 value x, the actual value is $y = e^x$, and the subsequent function is employed to normalize y: the actual

$$\hat{y} = \frac{1}{1 + y^{-0.1}},$$

in order to distribute the result more evenly on (0, 1),the parameter value of -0.1 is typically selected when y is very small ($< 10^{-3}$) [8].

Two subnetworks for drugs and cell lines

Conceptually, GPDRP can be viewed as a multimodal deep learning model comprising two subnetworks designed for processing drug and cell line features.

For drug features, graph convolutional networks may be well-fitting for DRP because a graph is used to represent the drug's molecular structure. In light of the widespread utilization of Graph Convolutional Networks (GCN) in the context of drug response prediction [19, 39, 40], we investigated four graph convolutional models, including Graph Convolutional Networks (GCN) [41], Graph Attention Networks (GAT) [42], Graph Isomorphism Network (GIN) [43] and Graph Transformer with the combination of GIN, all of which we described as follows. Following the GNN, a fully connected layer (FC layer) was additionally utilized to transform the outcome into 128 dimensions.

For cell line features, we used pathway activity scores and employed deep neural networks (DNN) with three hidden layers to learn features. The DNN architecture consisted of an input layer succeeded by three dense layers with sizes of 512, 1024, and 128, respectively, using Rectified Linear Unit (ReLU) as the activation function. The architecture incorporated a dropout layer with a rate set to 0.2 after the second dense layer to prevent overfitting. Then the output was flattened to a 128-dimensional vector. Subsequently, the 256-dimensional vector, encompassing both drug and cell line features, traversed two FC layers to predict drug response, with 1024 and 128 nodes respectively. The LN IC50 was used to quantify GPDRP output and indicated how well a medication inhibited the growth of a particular cancer cell line. A high level of drug efficacy was indicated by small IC50 values, which suggested that the drug was sensitive to the corresponding cancer cell line [28]. The hyper-parameters utilized in our experiments are listed in Table 3. They were chosen on the basis of prior research experience rather than tuned.

Graph convolutional networks (GCN)

Predicting a continuous value that represents the LN IC50 of drug sensitivity in cell lines is our main goal in this work. We employ GCN to learn about each drug graph representation. Formally, $G = (V, E)$ denotes the graph of a given drug, where V is the set of $N \in R$ nodes, each characterized as a C-dimensional vector, and E represents the set of edges, which is denoted by an adjacency matrix $A \in R^{N \times N}$. A node feature matrix $X \in R^{N \times C}$ and an adjacency matrix A are inputs to the multi-layer GCN. Then

Table 3 Hyper-parameters for different graph neural network variants used in our experiments

Hyper-parameters	Setting
Activation	ReLu
Optimizer	Adam
Learning rate	0.0001
Dropout	0.2
GCN Layers	3
GAT Layers	3
GIN Layers	3
GIN_TRANSFORMER Layers	3
DNN Layers	3

it generates a node-level output $Z \in R^{N \times F}$ with F denoting the quantity of features each node output. A normalized form is employed to express the propagation rule as follows:

$$H^{(l+1)} = \sigma(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^{(l)} W^{(l)}),$$

where $\tilde{A} = A + I_N$, $\tilde{D}$ is the graph diagonal degree matrix. And σ is an activation function, $H^{(l)} \in R^{N \times C}$ is the l - th layer's activation matrix, $H^{(0)} = X$, W is learnable parameters.

Three consecutive GCN layers are used in our GCN-based model, and the ReLU function is applied after each layer. After the last GCN layer, a global max pooling layer is incorporated to capture the representation vector of the entire graph, which is then combined with the representation of the cell line to predict the response value.

Graph attention networks (GAT)

The GAT is constructed through the layering of a graph attention layer. It introduces an attention-based structure to acquire latent node representations within a graph, employing a self-attention mechanism. The GAT layer uses a weight matrix W to apply a linear transformation to each node in a set of graph nodes that it receives as input. And the attention coefficients between node i and its first-order neighbors j are computed in the graph as

$$\alpha(W_{x_i}, W_{x_j}).$$

Subsequently, these attention coefficients undergo normalization through a softmax function and are employed to calculate the output features for the nodes as

$$\sigma \left(\sum_{j \in N(i)} \alpha_{ij} W_{x_j} \right),$$

where $\sigma(\cdot)$ is a non-linear activation function and α_{ij} are the normalized attention coefficients.

Our GAT-based model consists of three GAT layers activated by a ReLU function, followed by a global max pooling layer to obtain the graph representation vector.

Multi-head-attentions are used for the first GAT layer, with the number of heads set to ten. The second and third GAT's output features are limited to 128.

Graph Isomorphism network (GIN)

The GIN is a recent approach believed to attain optimal discriminative capability within graph neural networks. It employs a multi-layer perceptron (MLP) model for updating the node features as

$$MLP\left((1+\mu)x_i + \sum_{j\in N(i)} x_i\right),$$

where μ is either a learnable parameter or a fixed scalar, x is the node feature vector, and $N(i)$ is the set of nodes neighbor to i.

Three GIN layers are stacked in our GIN-based model to build architecture, with a batch normalization layer added after each layer. A global max pooling layer is added for aggregating a graph representation vector, similar to previous architectures.

Graph transformer

GCN and GAT are designed to learn on homogeneous graphs. GIN updates node representations by utilizing only the features of local neighboring nodes, which may result in insufficient capture of global information. In contrast, Transformer can facilitate better feature learning for more generalized drug graphs. With its self-attention mechanism, Transformer can simultaneously consider the information from all nodes in the graph, enabling a more effective integration of global information.

Drug graph $G = (V, E)$ has a set of node type T^v, and a set of edge type T^e. There is an adjacency tensor $A \in R^{N \times N \times K}$, where $K = |T^e|$ and feature matrix $X \in R^{N \times F}$. A meta-path is defined to predict new connections among nodes as

$$A_P = A_{t1} \cdots A_{tp},$$

where A_{ti} is an adjacency matrix for the ith edge type of meta-path. For A_{ti}, a soft adjacency matrix Q using 1×1 convolution is

$$Q = F(A, W_\phi) = \phi(A, soft\max(W_\phi)),$$

where ϕ is a convolution layer and $W_\phi \in R^{1 \times 1 \times K}$. Combining with GCN, node representations are constructed as

$$Z = \left\|_{i=1}^{C}\sigma(\tilde{D}_i^{-1}\tilde{A}_i^{(l)}XW).$$

Z is a function of neighborhood connectivity. Extracting features from the graph poses challenges in determining the node positions because of the inherent characteristics of the graph, the Graph Transformer utilizes Laplacian eigenvectors to address this concern as.

$$\Delta = I - D^{-1/2}AD^{-1/2} = U^T \Lambda U,$$

where U and Λ are eigenvectors and eigenvalues, respectively.

In our model, we combined one Graph Transformer layer with two GIN layers to improve feature extraction and prediction accuracy.

Performance evaluation

Two metrics were utilized to assess the performance of the models: Root Mean Squared Error (RMSE) and Pearson Correlation Coefficient (PCCs). RMSE is calculated as the square root of the mean squared error, representing the average squared difference between the actual and predicted responses. PCCs endeavors to gauge the presence of a linear correlation between two variables. Given n samples, O is the actual response value, and Y is the predicted response value. The actual response walue of ith sample is o_i, and ith sample's predicted response value is y_i. RMSE is calculated as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i}^{n} (o_i - y_i)^2}.$$

The PCCs of o_i and y_i is defined as follows:

$$PCCs = \frac{\sum_{i}^{n} (o_i - y_i)^2}{\sigma_O \sigma_Y}.$$

where σ_O and σ_Y are the standard deviations of ground-truth O and predicted value Y, respectively.

Abbreviations
GPDRP   Graph and gene pathway based drug response prediction method
GNN   Graph neural networks
DRP   Drug response prediction
CCLE   The Cancer Cell Line Encyclopedia
GDSC   The Genomics of Drug Sensitivity in Cancer
CTRPv2   The cancer therapeutics response portal
GIN   Graph isomorphism network
SMILES   Simplified molecular-input line-entry system
PCCs   Pearson correlation coefficient
RMSE   Root mean square error
GCN   Graph convolutional networks
GAT   Graph attention networks

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12859-023-05618-0.

> **Additional file 1**. Supplementary materials and tables.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

References

1. Feng F, Shen B, Mou X, Li Y, Li H. Large-scale pharmacogenomic studies and drug response prediction for personalized cancer medicine. J Genet Genom. 2021;48(7):540–51.
2. Adam G, Rampášek L, Safikhani Z, Smirnov P, Haibe-Kains B, Goldenberg A. Machine learning approaches to drug response prediction: challenges and recent progress. NPJ Precis Oncol. 2020;4:19.
3. Maeda H, Khatami M. Analyses of repeated failures in cancer therapy for solid tumors: poor tumor-selective drug delivery, low therapeutic efficacy and unsustainable costs. Clin Transl Med. 2018;7(1):11.
4. Lopez JS, Banerji U. Combine and conquer: challenges for targeted therapy combinations in early phase trials. Nat Rev Clin Oncol. 2017;14(1):57–66.
5. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim SG, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature. 2012;483(7391):603–7.
6. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Res. 2013;41(Database issue):D955–61.
7. Seashore-Ludlow B, Rees MG, Cheah JH, Cokol M, Price EV, Coletti ME, et al. Harnessing connectivity in a large-scale small-molecule sensitivity dataset. Cancer Discov. 2015;5(11):1210–23.
8. Menden MP, Iorio F, Garnett M, McDermott U, Benes CH, Ballester PJ, Saez-Rodriguez J. Machine learning prediction of cancer cell sensitivity to drugs based on genomic and chemical properties. PLoS ONE. 2013;8(4):e61318.
9. Ammad-ud-din M, Georgii E, Gönen M, Laitinen T, Kallioniemi O, Wennerberg K, et al. Integrative and personalized QSAR analysis in cancer by kernelized Bayesian matrix factorization. J Chem Inf Model. 2014;54(8):2347–59.
10. Zhang N, Wang H, Fang Y, Wang J, Zheng X, Liu XS. Predicting anticancer drug responses using a dual-layer integrated cell line-drug network model. PLoS Comput Biol. 2015;11(9):e1004498.
11. Wang L, Li X, Zhang L, Gao Q. Improved anticancer drug response prediction in cell lines using matrix factorization with similarity regularization. BMC Cancer. 2017;17(1):513.
12. Chang Y, Park H, Yang HJ, Lee S, Lee KY, Kim TS, et al. Cancer Drug Response Profile scan (CDRscan): a deep learning model that predicts drug effectiveness from cancer genomic signature. Sci Rep. 2018;8(1):8857.
13. Sakellaropoulos T, Vougas K, Narang S, Koinis F, Kotsinas A, Polyzos A, et al. A deep learning framework for predicting response to therapy in cancer. Cell Rep. 2019;29(11):3367-3373.e4.
14. Choi J, Park S, Ahn J. RefDNN: a reference drug based neural network for more accurate prediction of anticancer drug resistance. Sci Rep. 2020;10(1):1861.
15. Nguyen T, Le H, Quinn TP, Nguyen T, Le TD, Venkatesh S. GraphDTA: predicting drug-target binding affinity with graph neural networks. Bioinformatics. 2021;37(8):1140–7.
16. Ein-Dor L, Zuk O, Domany E. Thousands of samples are needed to generate a robust gene list for predicting outcome in cancer. Proc Natl Acad Sci USA. 2006;103(15):5923–8.
17. Khatri P, Sirota M, Butte AJ. Ten years of pathway analysis: current approaches and outstanding challenges. PLoS Comput Biol. 2012;8(2):e1002375.
18. Chawla S, Rockstroh A, Lehman M, Ratther E, Jain A, Anand A, et al. Gene expression based inference of cancer drug sensitivity. Nat Commun. 2022;13(1):5680.
19. Chu T, Nguyen TT, Hai BD, Nguyen QH, Nguyen T. Graph transformer for drug response prediction. IEEE/ACM Trans Comput Biol Bioinform. 2023;20(2):1065–72.
20. Liu Q, Hu Z, Jiang R, Zhou M. DeepCDR: a hybrid graph convolutional network for predicting cancer drug response. Bioinformatics. 2020;36(Supplement_2):i911–8.
21. Liu P, Li H, Li S, Leung KS. Improving prediction of phenotypic drug response on cancer cell lines using deep convolutional network. BMC Bioinform. 2019;20(1):1–4.
22. Ross JS, Schenkein DP, Pietrusko R, Rolfe M, Linette GP, Stec J, et al. Targeted therapies for cancer 2004. Am J Clin Pathol. 2004;122(4):598–609.

23. Friedman AA, Amzallag A, Pruteanu-Malinici I, Baniya S, Cooper ZA, Piris A, et al. Landscape of targeted anticancer drug synergies in melanoma identifies a novel BRAF-VEGFR/PDGFR combination treatment. PLoS ONE. 2015;10(10):e0140310.
24. Chen D, Frezza M, Schmitt S, Kanwar J, Dou PQ. Bortezomib as the first proteasome inhibitor anticancer drug: current status and future perspectives. Curr Cancer Drug Targets. 2011;11(3):239–53.
25. Smith MR, Jin F, Joshi I. Bortezomib sensitizes non–Hodgkin's lymphoma cells to apoptosis induced by antibodies to tumor necrosis factor–related apoptosis-inducing ligand (TRAIL) receptors TRAIL-R1 and TRAIL-R2. Clin Cancer Res. 2007;13(18):5528s-s5534.
26. PubChem [Internet]. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; 2004-. PubChem Compound Summary for CID 6914657, Daporinad; [cited 2023 July 17]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/Daporinad
27. Wishart DS, Knox C, Guo AC, Shrivastava S, Hassanali M, Stothard P, Chang Z, Woolsey J. DrugBank: a comprehensive resource for in silico drug discovery and exploration. Nucleic Acids Res. 2006;34(suppl_1):D668-72.
28. Brugières L, Pacquement H, Le Deley MC, Leverger G, Lutz P, Paillard C, et al. Single-drug vinblastine as salvage treatment for refractory or relapsed anaplastic large-cell lymphoma: a report from the French Society of Pediatric Oncology. J Clin Oncol. 2009;27(30):5056–61.
29. Xu B, Sun T, Wang S, Lin Y. Metronomic therapy in advanced breast cancer and NSCLC: vinorelbine as a paradigm of recent progress. Expert Rev Anticancer Ther. 2021;21(1):71–9.
30. Škubník J, Pavlíčková VS, Ruml T, Rimpelová S. Vincristine in combination therapy of cancer: emerging trends in clinics. Biology. 2022;10(9):849.
31. Zhou X, Zhu H, Liu L, Lin J, Tang K. A review: recent advances and future prospects of taxol-producing endophytic fungi. Appl Microbiol Biotechnol. 2010;86:1707–17.
32. Gandhi V, Keating MJ, Bate G, Kirkpatrick P. Nelarabine. Nat Rev Drug Discovery. 2006;5(1):17–9.
33. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. Pytorch: An imperative style, high-performance deep learning library. Advances in neural information processing systems 2019; 32.
34. Swain, M. PubChemPy: A way to interact with PubChem in Python. (2014).
35. Landrum G. RDKit: Open-source cheminformatics. 2006. Google Scholar 2006.
36. Ramsundar B, Eastman P, Walters P, Pande V. Deep learning for the life sciences: applying deep learning to genomics, microscopy, drug discovery, and more. O'Reilly Media, Inc; 2019
37. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinform. 2013;14:1–5.
38. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci. 2005;102(43):15545–50.
39. Huang Z, Zhang P, Deng L. DeepCoVDR: deep transfer learning with graph transformer and cross-attention for predicting COVID-19 drug response. Bioinformatics. 2023;39(39 Suppl 1):i475–83.
40. Kim S, Bae S, Piao Y, Jo K. Graph convolutional network for drug response prediction using gene expression data. Mathematics. 2021;9(7):772.
41. Kipf TN, Welling M. Semi-supervised classification with graph convolutional networks. arXiv:1609.02907. 2016 Sep 9.
42. Velickovic P, Cucurull G, Casanova A, Romero A, Lio P, Bengio Y. Graph Attention Networks Stat. 2017;1050(20):10–48550.
43. Xu K, Hu W, Leskovec J, Jegelka S. How powerful are graph neural networks?. arXiv:1810.00826. 2018 Oct 1.

Publisher's Note

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