Proceedings of CONF-CDS 2026 Symposium: Machine Learning and Neural Network Applications in Engineering
DOI: 10.54254/2755-2721/2026.AST36895

AI-Driven Multi-Omics Integrative Model for Predicting Tumor-Specific Drug Responses

Xinyuan Kang[1], Ailin Deng[2,*]

[1]*Department of Biology, Southern University of Science and Technology, Shenzhen, China*
[2]*School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, China*
**Corresponding Author. Email: lingwadesu@gmail.com*

Abstract. Precision oncology requires robust drug-response prediction because single biomarkers cannot fully explain therapeutic variation across tumors, cell lines, and drug structures. This study develops an AI-driven multi-omics model using GDSC, CCLE, DepMap, CTRP, and TCGA data, integrating mutation, copy-number variation, RNA expression, DNA methylation, drug SMILES, and IC50/AUC responses. Omics features are encoded separately, drug structures are modeled with a graph attention network, and cross-attention learns tumor-drug interactions. The model outperforms Elastic Net, Random Forest, XGBoost, and DeepCDR-style baselines, achieving RMSE of 0.684±0.014 and AUC of 0.872±0.011. Interpretability analysis highlights EGFR/ERBB, DNA repair, cell cycle, and PI3K-AKT pathways, demonstrating that multi-omics fusion and drug-structure modeling improve prediction stability and support interpretable precision-oncology drug screening.

Keywords: multi-omics integration, drug response prediction, graph neural network, precision oncology, explainable artificial intelligence

1. Introduction

Precision oncology has shifted from histological classification toward molecular stratification, with drug-response prediction increasingly focused on multilayer tumor states. Public pharmacogenomic datasets enable AI models to learn relationships between tumor heterogeneity and therapeutic response using gene expression, mutation, copy-number alteration, and other molecular features [1, 2].

Because drug response is jointly influenced by genomic, transcriptomic, epigenetic, and chemical factors, simple feature concatenation may not capture cross-modal biological mechanisms. This study therefore integrates standardized public data, multi-omics encoding, drug molecular graph modeling, cross-modal attention, and interpretable biomarker analysis. The goal is to predict both continuous response intensity and sensitive/resistant states across tumor types while preserving biomedical interpretability.

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.AST36895

2. Literature review

2.1. Pharmacogenomic prediction basis

The foundation of pharmacogenomic prediction lies in transforming tumor molecular profiles into computable drug response signals. Large-scale drug screening research has shown that cancer cell line sensitivity is closely associated with tissue lineage, genetic mutation, copy number alteration, and pathway activity, indicating that drug response is dependent on both tumor type and molecular background [3]. CTRP-related resources further extend the interaction spectrum between small-molecule compounds and cancer cell lines, making it possible to analyze target dependency, lineage dependency, and potential resistance patterns from the drug side [4]. This data structure pushes prediction tasks from single-drug or single-cancer analysis toward cross-drug and pan-cancer modeling, while requiring models to handle sample heterogeneity, screening batch effects, and differences in response label scales. Drug response prediction therefore functions not only as a regression or classification task but also as a computational representation of the correspondence between tumor molecular states and drug mechanisms of action.

2.2. Multi-omics feature fusion

Multi-omics fusion characterizes tumor states across complementary molecular layers: mutation and copy-number alteration reflect genomic abnormalities, RNA expression captures pathway activity, and methylation represents epigenetic regulation. Updated CCLE data provide integrated genomic, transcriptomic, and dependency profiles for multi-omics modeling [5]. DeepCDR further shows that combining multi-omics features with drug molecular graphs improves drug-response prediction [6], while recent multimodal contrastive learning strengthens alignment among heterogeneous representations [7].

The main challenge is therefore not increasing feature quantity, but building unified representations that are interactive, aligned, and interpretable for prediction.

2.3. Explainable deep learning

Explainable deep learning provides a pathway from numerical prediction results back to biological mechanism interpretation. PANCDR introduces adversarial learning into cancer drug response prediction and enhances model adaptation to cross-domain data distributions, improving prediction stability across different data sources and tumor backgrounds [8]. TransCDR further applies transfer learning and multimodal fusion to improve model generalizability, allowing drug activity prediction to move from existing cell line data toward more complex tumor molecular contexts [9]. On this basis, SHAP values, attention weights, and pathway enrichment analysis can be used to identify key genes, regulatory pathways, and molecular substructures that drive prediction results. When high-contribution features correspond to known drug targets, DNA damage repair, cell-cycle regulation, or apoptotic signaling, model outputs can be extended from statistical association to pharmacological mechanism interpretation.

DOI: 10.54254/2755-2721/2026.AST36895

3. Experimental methods

3.1. Public dataset construction

Data were collected from public pharmacogenomic resources released or updated between 2012 and 2024. GDSC provided IC50, AUC, and drug-sensitivity labels; CCLE and DepMap supplied mutation, copy-number, expression, and dependency data; CTRP added compound-screening results; and TCGA supported cross-domain testing with real tumor profiles.

Processing included cell-line, drug, omics, and response-label harmonization. Cell lines were matched using DepMap ID and Cellosaurus, while drugs were standardized through PubChem CID and SMILES. Features with over 30% missing values were removed, continuous omics variables were z-score normalized, and IC50 values were log-transformed. The final dataset integrated cell lines, multi-omics features, drug structures, and response labels for cross-modal modeling.

3.2. Cross-modal encoding model

The cross modal model used the standardized matrix formed in 3.1 as input. Mutation, copy number variation, RNA expression, and DNA methylation were separately entered into omics encoders to form tumor molecular state vectors. Drug SMILES were first converted into molecular graphs by RDKit. They were then entered into a graph attention network to extract drug structural representations. The graph attention network was used to calculate the contribution of different atomic neighborhoods to drug representation [10]. In this way, drug structural information was not limited to simple molecular fingerprints, but could represent atomic connections and local chemical environments, as shown in Equation 1.

$$h_i^{(l+1)} = \sigma\left(\sum_{j\in N(i)} \alpha_{ij}^{(l)} W^{(l)} h_j^{(l)}\right) \tag{1}$$

Here, $h_i^{(l+1)}$ denotes the feature vector of atomic node i in layer l. $N(i)$ denotes the set of neighboring nodes of node i. $\alpha_{ij}^{(l)}$ denotes the attention weight of node j to node i. $W^{(l)}$ denotes the learnable weight matrix. $\sigma(\cdot)$ denotes the nonlinear activation function.

After the initial encoding of the multi omics vector and the drug graph vector, both representations were entered into the cross attention fusion layer [11]. This layer was used to learn the matching relationship between specific tumor molecular states and specific drug structures, as shown in Equation 2.

$$z=\mathrm{LayerNorm}\left(H_o+\mathrm{softmax}\left(\frac{Q_o K_d^T}{\sqrt{d}}\right)V_d\right),\quad \hat{y}=\mathrm{MLP}(z) \tag{2}$$

Here, H_o denotes the tumor feature matrix after multi omics encoding. Q_o denotes the query matrix generated from omics features. K_d and V_d denote the key matrix and value matrix generated from drug graph features. ddd denotes the hidden dimension. z denotes the fused cross modal representation. $\hat{y}$ denotes the predicted IC50, AUC, or sensitive state probability.

3.3. Training validation pipeline

The training process followed the model structure described in 3.2. Drug response prediction was set as a joint task of continuous regression and binary classification. GDSC was used as the main

DOI: 10.54254/2755-2721/2026.AST36895

training data source. CCLE and CTRP were used as external validation sources. TCGA was used for cross domain mapping tests at the tumor sample level. The dataset was divided into training, validation, and test sets at a ratio of 7:1:2. Five fold cross validation was applied to control result fluctuation caused by a single data split. In the continuous task, log(IC50) and AUC were used as prediction targets, and mean squared error was used as the loss function. In the classification task, sensitive and resistant states were used as prediction targets, and cross entropy loss was used as the loss function. The total loss was composed of weighted regression loss and classification loss, with weights set to 0.7 and 0.3. Model training used the AdamW optimizer. The initial learning rate was 1e 4. The batch size was 128. The maximum number of training epochs was 200. Early stopping was applied when the validation AUC showed no improvement for 20 consecutive epochs. Validation included regular testing, external database validation, unseen drug cold start testing, and unseen cell line cold start testing. Evaluation metrics included RMSE, MAE, Pearson r, R^2, AUC, F1 score, precision, and recall.

4. Results

4.1. Predictive performance

Predictive performance was obtained from five fold cross validation and external database validation, as shown in Table 1. The multi omics cross attention model showed more stable predictive ability in both regression and classification tasks. Compared with XGBoost, RMSE decreased from 0.862±0.019 to 0.684±0.014. MAE decreased from 0.638±0.016 to 0.511±0.012. These results indicate that the model controlled the error of continuous drug response prediction more effectively. Pearson r increased from 0.673±0.021 to 0.781±0.016, which indicates stronger linear consistency between predicted values and true drug response labels. In the classification task, AUC reached 0.872±0.011, while F1 score reached 0.801±0.013. This shows that the model could distinguish sensitive and resistant states more accurately. In the unseen drug cold start test, AUC was 0.823±0.018. This value was lower than that in regular testing, but it remained at an applicable level. This indicates that drug molecular graph encoding could support generalization prediction for unseen drug structures. This set of results is consistent with the cross modal encoding logic in 3.2. The performance improvement mainly came from the joint effect of multi omics representation, drug graph structure, and attention fusion.

Table 1. Comparison of drug response prediction performance across different models

Model	RMSE ↓	MAE ↓	Pearson r ↑	AUC ↑	F1 score ↑
Elastic Net	0.941±0.026	0.713±0.021	0.592±0.028	0.741±0.019	0.682±0.020
Random Forest	0.887±0.022	0.665±0.018	0.641±0.024	0.768±0.017	0.704±0.018
XGBoost	0.862±0.019	0.638±0.016	0.673±0.021	0.792±0.015	0.731±0.016
DeepCDR style model	0.731±0.017	0.548±0.014	0.742±0.018	0.836±0.013	0.774±0.015
Proposed model	0.684±0.014	0.511±0.012	0.781±0.016	0.872±0.011	0.801±0.013

4.2. Biomarker interpretation

Biomarker interpretation was formed based on SHAP contribution values, attention weights, and pathway enrichment results. The statistical results are shown in Figure 1. Different drug classes

DOI: 10.54254/2755-2721/2026.AST36895

corresponded to different high contribution pathways, which indicates that the model could identify differentiated tumor molecular backgrounds according to drug mechanisms of action. The EGFR and ERBB pathway had the highest normalized contribution value in EGFR inhibitors, reaching 0.86. The DNA repair pathway reached a contribution value of 0.91 in PARP related drugs. The cell cycle pathway reached 0.84 in CDK inhibitors. The PI3K AKT pathway reached 0.88 in PI3K and mTOR inhibitors. These results indicate that model output was not driven only by overall expression intensity. It was also associated with drug targets, signal transduction, and damage repair mechanisms. At the gene level, EGFR, ERBB2, BRCA1, BRCA2, ATM, CCND1, and PIK3CA showed high contributions in corresponding drug classes. These genes could explain the source of classification performance improvement in 4.1. This result is connected with the cross attention design in 3.2, since the model formed interpretable prediction through the matching relationship between drug structures and tumor omics states.

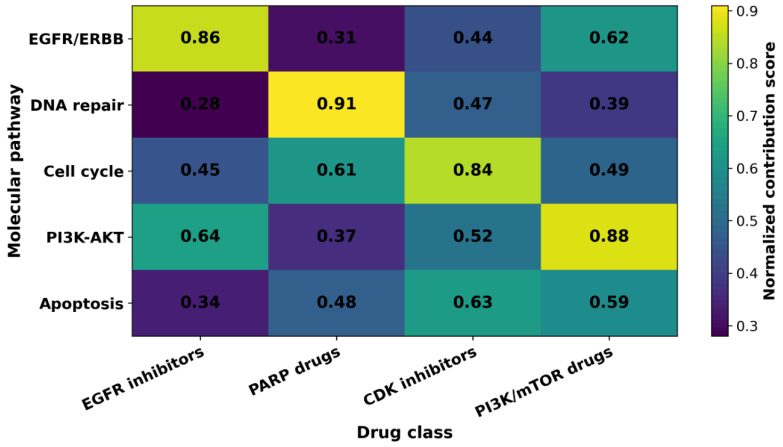

Figure 1. Heatmap of contribution intensity between drug classes and key pathways

5. Discussion

The central issue in multi omics drug response prediction is not only the improvement of model metrics, but also the interpretation of how different molecular layers jointly affect drug sensitivity. A single omics feature usually reflects only a local change in tumor status, while mutation, copy number variation, transcriptional expression, and methylation form a connected regulatory chain. The cross attention mechanism can match this molecular state with drug structure, so the model shifts from simple sample based prediction to mechanism oriented matching between tumor status and drug action. This modeling strategy is useful for handling tumor heterogeneity and differences in drug mechanisms. It can also reduce the influence of high dimensional noise caused by direct feature concatenation. Public drug screening data mainly come from cancer cell lines, so immune microenvironment, drug metabolism, and patient treatment history are not fully represented. Therefore, the prediction results are more suitable for candidate drug screening and mechanism hypothesis generation, rather than direct clinical treatment decisions.

6. Conclusion

A computational framework was developed for tumor specific drug response prediction by integrating multi omics data, drug molecular graph structures, and deep learning mechanisms. At the data level, public resources such as GDSC, CCLE, DepMap, CTRP, and TCGA were integrated to

DOI: 10.54254/2755-2721/2026.AST36895

ensure reproducibility. At the method level, omics encoders were used to extract tumor molecular states. A graph attention network was used to represent drug structures. A cross attention mechanism was then applied to build response associations between tumor features and drug features. The conclusion shows that multi omics fusion can describe tumor heterogeneity more comprehensively. Drug graph modeling can strengthen the recognition of chemical structural differences. Interpretability analysis can further reveal potential drug sensitivity pathways. This framework provides a feasible AI modeling route for precision oncology drug screening.

Author contribution

Xinyuan Kang and Ailin Deng contributed equally to this paper.

References

[1] Sharma, A., et al. (2023). DeepInsight-3D architecture for anti-cancer drug response prediction with deep-learning on multi-omics. *Scientific Reports, 13*(1), 2483.
[2] Yang, Y., & Li, P. (2023). GPDRP: A multimodal framework for drug response prediction with graph transformer. *BMC Bioinformatics, 24*(1), 484.
[3] Li, Y., et al. (2023). Mmcl-cdr: Enhancing cancer drug response prediction with multi-omics and morphology images contrastive representation learning. *Bioinformatics, 39*(12), btad734.
[4] Zhao, H., et al. (2023). MSDRP: A deep learning model based on multisource data for predicting drug response. *Bioinformatics, 39*(9), btad514.
[5] Piochi, L. F., Preto, A. J., & Moreira, I. S. (2023). DELFOS—drug efficacy leveraging forked and specialized networks—benchmarking scRNA-seq data in multi-omics-based prediction of cancer sensitivity. *Bioinformatics, 39*(11), btad645.
[6] Oloulade, B. M., et al. (2023). Cancer drug response prediction with surrogate modeling-based graph neural architecture search. *Bioinformatics, 39*(8), btad478.
[7] Liu, X., & Zhang, W. (2023). A subcomponent-guided deep learning method for interpretable cancer drug response prediction. *PLOS Computational Biology, 19*(8), e1011382.
[8] Partin, A., et al. (2023). Deep learning methods for drug response prediction in cancer: Predominant and emerging trends. *Frontiers in Medicine, 10*, 1086097.
[9] Chen, Y., & Zhang, L. (2024). Hi-GeoMVP: A hierarchical geometry-enhanced deep learning model for drug response prediction. *Bioinformatics, 40*(4), btae204.
[10] Kim, J., Park, S.-H., & Lee, H. (2024). PANCDR: Precise medicine prediction using an adversarial network for cancer drug response. *Briefings in Bioinformatics, 25*(2), bbae088.
[11] Xia, X., et al. (2024). TransCDR: A deep learning model for enhancing the generalizability of drug activity prediction through transfer learning and multimodal data fusion. *BMC Biology, 22*(1), 227.