MGTDR: A Multi-modal Graph Transformer Network for Cancer Drug Response Prediction

Chi Yan
Department of Information and Communication
Officers college of PAP
Chengdu,610213,China
yanchi8@qq.com

2024 7th International Conference on Artificial Intelligence and Big Data (ICAIBD) | 979-8-3503-8510-6/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICAIBD62003.2024.10604610

***Abstract*—Drug response prediction in cancer cell lines can guide researchers to design personalized treatments for different patients. However, accurately predicting drug response remains a challenging task. This study proposes MGTDR, a multi-modal graph transformer framework for drug response prediction. First, using an auto-encoder, MGTDR learns the latent features of cancer cell lines. Secondly, It employs graph convolutional neural networks (GCN) and multi-layer perceptrons (MLP) to understand features of drugs from the simplified molecular input line entry specification (SMILES) and molecular fingerprints of drugs. Thirdly, it utilizes miRNA expression, DNA methylation, and drug physicochemical properties to calculate cell line similarity and drug similarity. Subsequently, it constructs a heterogeneous network by combining cell line similarity and drug similarity. The cell line features and drug features calculated earlier are then employed as the features of the nodes in the network. Finally, it applies graph transformer networks and MLP to predict drug sensitivity. Extensive experiments on publicly available datasets demonstrated the effectiveness and efficiency of the proposed method in predicting drug response and its potential value in guiding personalized therapy.**

***Index Terms*—Drug response prediction, multi-omics fusion, drug structure, graph convolutional neural network**

I. Introduction

Cancer remains a global challenge and a principal obstacle to enhancing life expectancy worldwide. To confront this multifaceted menace, developing effective treatment strategies is paramount. In recent years, precision medicine has emerged as a transformative approach, unraveling the intricate genomic intricacies of cancer and pioneering tailored therapeutic interventions. Nevertheless, one formidable barrier persists—our inability to accurately forecast how an individual will respond to a specific drug.

The rapid progress in pharmacogenomic research has yielded a rich repository of genomic data and comprehensive drug sensitivity and resistance profiles across a spectrum of cancer cell lines. Noteworthy initiatives in this domain include the Cancer Cell Line Encyclopedia (CCLE) [1], Genomics of Drug Sensitivity in Cancer (GDSC) [2], and the US National Cancer Institute 60 human tumor cell line anticancer drug screen (NCI60) [3], among others. Benefiting from these multi-modal data, many computational methods based on these data have been developed to predict drug response.

Conventional statistical models and sophisticated machine learning approaches to drug response prediction often rely on empirical observations and statistical analyses of clinical data [4], [5]. A variety of computational methods have been developed for drug response prediction and the discovery of drug response biomarkers [6]–[8]. Nonetheless, these computational models continue to exhibit room for improvement in predictive performance and generalizability [9].

Recently, deep learning has provided a promising avenue for capturing biological data's inherent complexity and nonlinear relationships due to its ability to handle vast amounts of high-dimensional and noisy data. For example, Ma et al. [10] have developed a new TCRP method to train drug response predictors in cancer cell lines and optimize their performance in higher complex cancer model systems via few-shot learning. And many other deep learning methods like SCAD [11], Deepdsc [12], CDRscan [13], DeepCDR [14], MOLI [15], tCNNs [16] and DeepTTA [17] have been proposed. However, deep learning methods based on Euclidean manifolds are limited to learning non-Euclidean features in drug response prediction.

In response to the limitations of traditional deep learning, graph neural networks (GNN) based methods have been proposed. For example, Deng et al. [18] proposed a technique based on a heterogeneous network for predicting the association of miRNA drug sensitivity. Jin et al. [19] proposed a general hypergraph learning algorithm for multitask drug prediction in biomedical networks. And many other graph learning-based models like MAOFGCN [20] and GADRP [21] have been proposed.

Although the above methods have achieved promising results, there is still room for improvement. Firstly, these methods should have considered the multi-modal features of drugs or cell lines. Secondly, both of these methods should have considered the molecular structure information of drugs.

Considering the above limitations, we proposed a Multi-modal Graph Transformer network for cancer Drug Response prediction (MGTDR). First, using an auto-encoder, MGTDR learns the latent features of cancer cell lines. Secondly, It employs GCNs and MLP to understand features of drugs from the SMILES and molecular fingerprints of medicines. Thirdly, it utilizes miRNA expression, DNA methylation, and drug physicochemical properties to calculate cell line similarity and drug similarity. Subsequently, it constructs a heterogeneous network by combining cell line similarity and drug similarity.

The cell line features and drug features calculated earlier are then employed as the features of the nodes in the network. Finally, it applies graph transformer networks and a MLP to predict drug sensitivity. Extensive experiments on publicly available datasets demonstrated the effectiveness and efficiency of the proposed method in predicting drug response and its potential value in guiding personalized therapy. Importantly, our proposed method incorporates several novel features:

- MGTDR integrates the multi-omics features of cancer cell lines and multi-modal features of drugs into an end-to-end graph learning framework to boost the performance of drug response prediction.
- The drugs are presented as graphs where the nodes are atoms, and the edges are the bonding of atoms. MGTDR extracts the molecular structure information of drugs represented by graphs through graph convolutional networks.
- MGTDR employs the graph transformer framework to extract features of the constructed drug-cell network.

II. Materials and Methods

A. Datasets

Drug physicochemical properties and molecular fingerprint: For 1448 drugs, the physicochemical characteristics and molecular fingerprints were downloaded from the PubChem database [22].

Drug simplified molecular input line entry specification (SMILES): The SMILES string of 1448 drugs was downloaded from the PubChem database [22]. Then, by leveraging the RDKit toolkit [23], the SMILES string of each Drug can be transformed into a molecular graph that reflects interactions between atoms inside drugs [24].

Multi-omics data for cancer cell lines: The gene expression data, miRNA expression data, DNA copy number data, and DNA methylation data of cancer cell lines were downloaded from the CCLE database [1]. Subsequently, cell lines lacking any of these four features were excluded, resulting in a final dataset comprising 388 cell lines.

Cancer cell line–drug response data: The IC_{50} values corresponding to 1,448 drugs and 388 cancer cell lines were retrieved from the PRISM Repurposing database [25] to measure the responses between drugs and cancer cell lines.

B. Overview of framework

The overall network architecture of our proposed graph learning network is illustrated in Fig. 1 with four main modules: i) Extraction of cell line features, ii) Extraction of drug features, iii) Construction of drug-cell network, and iv) Predicting of drug response.

Extraction of cell line features: An autoencoder was used for dimensionality reduction to address the high dimensionality of cancer cell line multi-omics profiles. The input features comprised 23,316-dimensional DNA copy number data and 48,392-dimensional gene expression data. Subsequently, the autoencoder reduced these high-dimensional features to 400 dimensions, serving as the low-dimensional representations for DNA copy number and gene expression data.

Extraction of drug features: We employ the SMILES representation, which captures global structural information, and the molecular fingerprint, which captures local structural information, as drug features. For drug molecular fingerprints, we use a MLP to extract their features. To extract comprehensive structural details of drugs, the SMILES representation of a drug is transformed to an undirected graph $G = (V, E)$ with N nodes $v_i \in V$, edges $(v_i, v_j) \in E$, a feature matrix X of node feature vectors and an adjacency matrix $A \in R^{N \times N}$. Then, we use a graph convolutional network to extract the features from the graph of SMILES:

$$\mathbf{X}^{(l+1)} = \sigma\left(\tilde{\mathbf{D}}^{-\frac{1}{2}}\tilde{\mathbf{A}}\tilde{\mathbf{D}}^{-\frac{1}{2}}\mathbf{X}^{(l)}\mathbf{W}^{(l)}\right) \tag{1}$$

where l is the number of layers, $\tilde{D}$ is the diagonal degree matrix of the graph, $\tilde{A}$ is the graph adjacency matrix with added self loop, W is the trainable parameter matrix.

Construction of drug-cell network: We explore the similarity between cells or drugs to create both a cell line similarity network and a drug similarity network. In this study, we employ the Pearson correlation coefficient to quantify the similarity among cell lines or drugs:

$$\text{pearsonr}(X, Y) = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \tag{2}$$

where X and Y a represent the features of cell lines or drugs, and n is the number of cell lines or drugs. X_i and Y_i represent individual cell line or drug, $\bar{X}$ and $\bar{Y}$ represent the mean of X and Y, respectively.

Predicting of drug response: We use the graph transformer network to extract node features when constructing the Drug-Cell network. Given node features $X = \{x_1, x_2, ..., x_n\}$, the multi-head attention for each edge from node j to node i is calculated as following:

$$q_{c,i} = W_{c,q}x_i + b_{c,q} \tag{3}$$

$$k_{c,j} = W_{c,k}x_j + b_{c,k} \tag{4}$$

$$v_{c,j} = W_{c,v}h_j + b_{c,v} \tag{5}$$

$$\alpha_{c,ij} = \frac{\langle q_{c,i}, k_{c,j}\rangle}{\sum_{u \in \mathcal{N}(i)}\langle q_{c,i}, k_{c,u}\rangle} \tag{6}$$

For the c_{th} head attention, we firstly transform the source feature x_i and distant feature x_j into query vector $q_{c,i} \in R^d$, key vector $k_{c,j} \in R^d$ and value vector $v_{c,j}$ respectively. $W_{c,q}$, $W_{c,k}$, $b_{c,q}$, $b_{c,k}$ denote the trainable parameters, respectively. After getting the graph multi-head attention, we make a message aggregation from the distant node j to the source node i:

$$\hat{x}_i = \text{Concat}(\sum_{j \in \mathcal{N}(i)} \alpha_{c,ij}v_{c,j}) \tag{7}$$

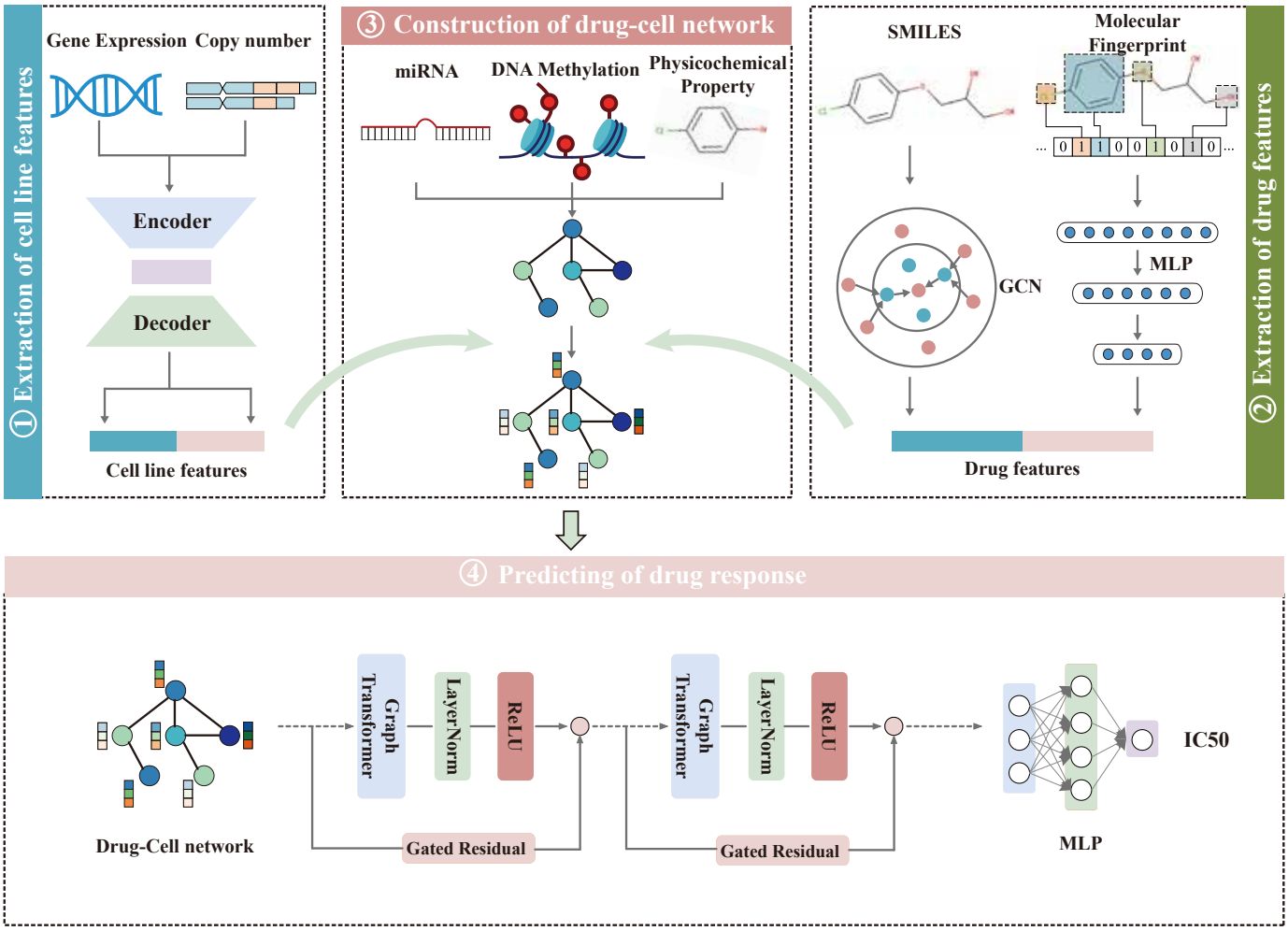

Fig. 1. **Overview of the MGTDR**. (1) Extraction of cell line features. The autoencoder learns the latent features of gene expression and copy number. (2) Extraction of drug features. The SMILES and molecular fingerprint features are extracted by a graph convolutional network and MLP, respectively. (3) A drug-cell network is constructed based on miRNA, DNA methylation, and physicochemical data. The features of each node are formed by integrating features extracted from (1) and (2). (4) The constructed drug-cell network is fed into a graph transformer and a multi-layer perceptron to predict the IC_{50} values.

Finally, the extracted features were fed into a multi-layer perceptron to predict the IC_{50} value.

C. Baselines

Aiming to evaluate the performance of our proposed method on cancer drug response prediction, we compared our method with eight state-of-the-art deep learning methods, including: MOLI [15], CDRscan [13], tCNNs [16], DeepDSC [12], DeepCDR [14], MAOFGCN [20], DeepTTA [17] and GADRP [21] based on the same dataset and metrics, and using the default or optimal performance parameters for each model. All baseline models shared the same input features and preprocessing methods.

D. Evaluation Metrics

We evaluate the performance of GADRP and baselines by four metrics: root mean squared error (RMSE), Pearson's correlation coefficient (PCC), Spearman's correlation coefficient (SCC), and R^2. The calculation of these metrics is as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \tilde{y}_i)^2} \tag{8}$$

$$PCC = \frac{\mathrm{cov}(Y, \tilde{Y})}{\sigma_Y \sigma_{\tilde{Y}}} \tag{9}$$

$$SCC = 1 - \frac{6\sum_{i=1}^{N} {d_i}^2}{N(N^2-1)} \tag{10}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \tilde{y}_i)^2}{\sum_{i=1}^{N}(y_i - \tilde{y}_i)^2} \tag{11}$$

III. Experiments

A. MGTDR outperforms existing methods in drug response prediction

To evaluate the performance of the proposed MGTDR in drug response prediction, we compared it with eight state-of-the-art methods introduced in baselines. As shown in Fig. 2,

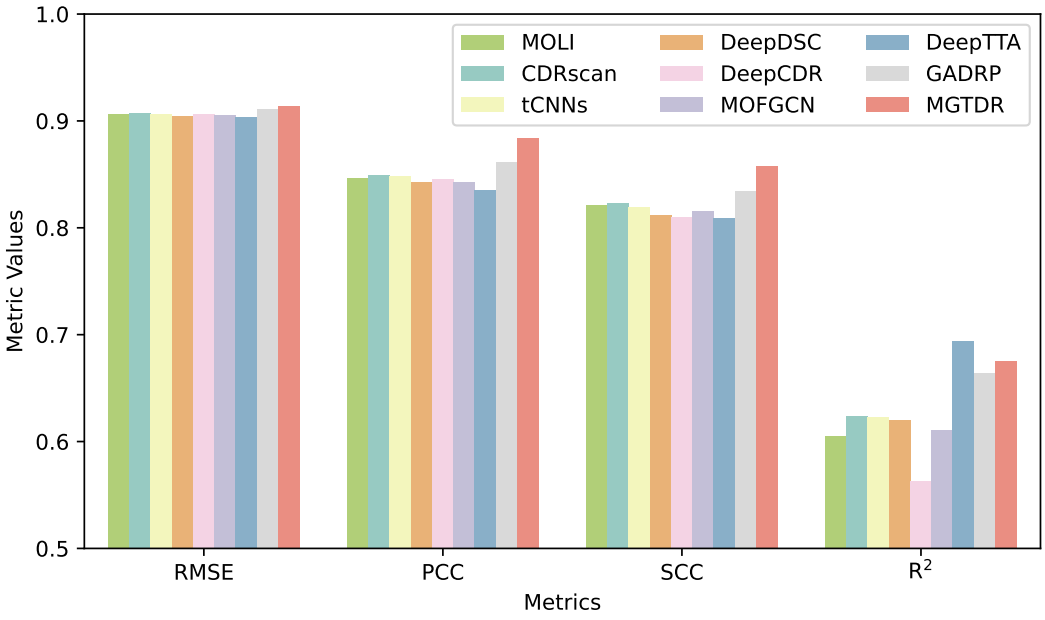

Fig. 2. **Performance comparison of MGTDR with other baselines.** The performances of our method and existing methods on drug response prediction. RMSE, PCC, SCC, and R^2 are evaluation metrics.

MGTDR achieved the best performance among all of the tested methods, with an RMSE of 0.0862, PCC of 0.8835, SCC of 0.8573 and R^2 of 0.6754. It should be noted that, for a more intuitive comparison with other evaluation metrics, the actual values of the R^2 when plotting the bar chart are represented as $1 - R^2$. Therefore, our method is more effective than the baselines in drug response prediction.

B. Exploration of the optimal architecture of our model

To investigate the optimal architecture of our model, we compared the results of different variants:

- **Graph Convolutional Network (GCN):** We use the graph convolutional network (GCN) to extract features from the drug-cell heterogeneous network.
- **Graph Attention Network (GAT):** We use the graph attention network (GAT) to extract features from the drug-cell heterogeneous network.
- **Graph Transformer Network (Trans):** We use the graph transformer network (Trans) to extract features from the drug-cell heterogeneous network.

These variants' performances and our model are shown in Fig. 3. These observations illustrate that the graph transformer network is superior to the graph convolutional network and the graph attention network on all four metrics.

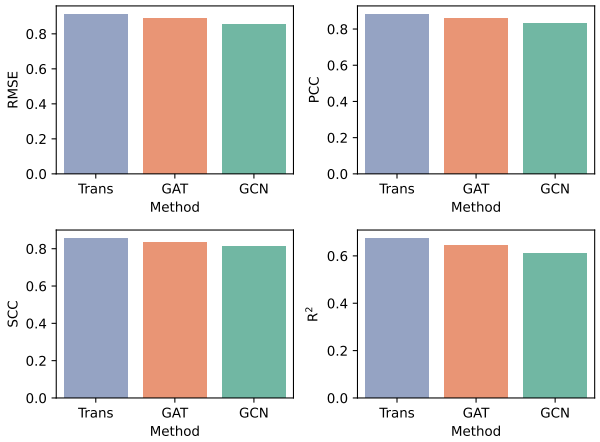

Fig. 3. **Exploration of the optimal architecture of MGTDR.** The RMSE, PCC, SCC and R^2 are displayed as histograms. Trans, GAT, and GCN denote that we use the graph transformer network, graph attention network, and graph convolutional network to extract features from the drug-cell network.

C. MGTDR benefits from multi-modal feature fusion

To validate the performance of MGTDR on a subset of data types, we conducted feature ablation experiments on drug features and cell line multi-omics data. The results can be observed from Fig. 4. Using all drug features and cell line features significantly improved the model's performance compared to using only part of the features. The result demonstrates the effectiveness of integrating multi-modal data.

IV. Conclusion

In this study, we develop MGTDR, a multi-modal graph transformer network for cancer drug response prediction. MGTDR integrates cell line and drug similarity networks to construct a heterogeneous network and combines multi-modal data into a single predictive model. Compared with other

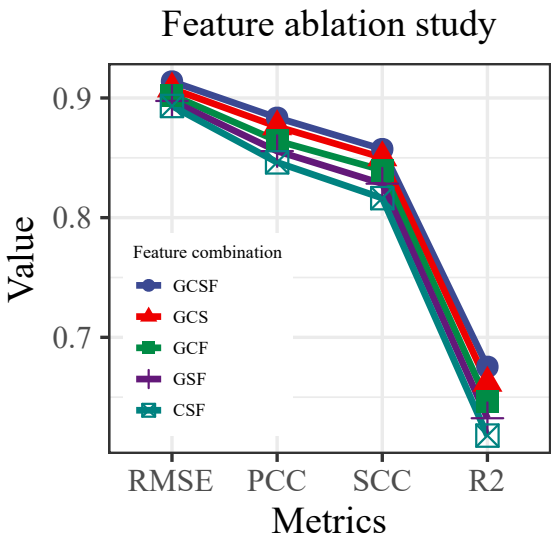

Fig. 4. **Results of feature ablation experiments.** GCSF: gene expression + copy number + SMILES + molecular fingerprint. GCS: gene expression + copy number + SMILES. GCF: gene expression + copy number + molecular fingerprint. GSF: gene expression + SMILES + molecular fingerprint. CSF: copy number + SMILES + molecular fingerprint.

state-of-the-art methods designed for drug response prediction, our model achieved superior performance regarding RMSE, PCC, SCC, and R^2. Although our MGTDR model achieves improved performance in predicting drug response, there is still room for improvement. Cancer drug treatments suffer from low efficacies caused by cancer heterogeneity. Fortunately, the revolutionary single-cell RNA-sequencing (scRNA-seq) technique offers an unparalleled opportunity to unveil the diverse gene expressions within cancer subpopulations in response to specific drugs. In the future, we plan to design computational methods to infer cancer drug responses at the single-cell level.

REFERENCES

[1] Jordi Barretina, Giordano Caponigro, Nicolas Stransky, Kavitha Venkatesan, Adam A Margolin, Sungjoon Kim, Christopher J Wilson, Joseph Lehár, Gregory V Kryukov, Dmitriy Sonkin, et al. The cancer cell line encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature*, 483(7391):603–607, 2012.

[2] Wanjuan Yang, Jorge Soares, Patricia Greninger, Elena J Edelman, Howard Lightfoot, Simon Forbes, Nidhi Bindal, Dave Beare, James A Smith, I Richard Thompson, et al. Genomics of drug sensitivity in cancer (gdsc): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic acids research*, 41(D1):D955–D961, 2012.

[3] Robert H Shoemaker. The nci60 human tumour cell line anticancer drug screen. *Nature Reviews Cancer*, 6(10):813–823, 2006.

[4] Jose Luis Perez-Gracia, Miguel F Sanmamed, Ana Bosch, Ana Patiño-Garcia, Kurt A Schalper, Victor Segura, Joaquim Bellmunt, Josep Tabernero, Christopher J Sweeney, Toni K Choueiri, et al. Strategies to design clinical studies to identify predictive biomarkers in cancer research. *Cancer Treatment Reviews*, 53:79–97, 2017.

[5] Muthu Dhandapani and Aaron Goldman. Preclinical cancer models and biomarkers for drug development: new technologies and emerging tools. *Journal of molecular biomarkers & diagnosis*, 8(5), 2017.

[6] Cai Huang, Roman Mezencev, John F McDonald, and Fredrik Vannberg. Open source machine-learning algorithms for the prediction of optimal cancer drug therapies. *PLoS One*, 12(10):e0186906, 2017.

[7] Stefan Naulaerts, Cuong C Dang, and Pedro J Ballester. Precision and recall oncology: combining multiple gene mutations for improved identification of drug-sensitive tumours. *Oncotarget*, 8(57):97025, 2017.

[8] Kaitlyn M Gayvert, Omar Aly, James Platt, Marcus W Bosenberg, David F Stern, and Olivier Elemento. A computational approach for identifying synergistic drug combinations. *PLoS computational biology*, 13(1):e1005308, 2017.

[9] Angeliki Kalamara, Luis Tobalina, and Julio Saez-Rodriguez. How to find the right drug for each patient? advances and challenges in pharmacogenomics. *Current opinion in systems biology*, 10:53–62, 2018.

[10] Jianzhu Ma, Samson H Fong, Yunan Luo, Christopher J Bakkenist, John Paul Shen, Soufiane Mourragui, Lodewyk FA Wessels, Marc Hafner, Roded Sharan, Jian Peng, et al. Few-shot learning creates predictive models of drug response that translate from high-throughput screens to individual patients. *Nature Cancer*, 2(2):233–244, 2021.

[11] Zetian Zheng, Junyi Chen, Xingjian Chen, Lei Huang, Weidun Xie, Qiuzhen Lin, Xiangtao Li, and Ka-Chun Wong. Enabling single-cell drug response annotations from bulk rna-seq using scad. *Advanced Science*, page 2204113, 2023.

[12] Min Li, Yake Wang, Ruiqing Zheng, Xinghua Shi, Yaohang Li, Fang-Xiang Wu, and Jianxin Wang. Deepdsc: a deep learning method to predict drug sensitivity of cancer cell lines. *IEEE/ACM transactions on computational biology and bioinformatics*, 18(2):575–582, 2019.

[13] Yoosup Chang, Hyejin Park, Hyun-Jin Yang, Seungju Lee, Kwee-Yum Lee, Tae Soon Kim, Jongsun Jung, and Jae-Min Shin. Cancer drug response profile scan (cdrscan): a deep learning model that predicts drug effectiveness from cancer genomic signature. *Scientific reports*, 8(1):8857, 2018.

[14] Qiao Liu, Zhiqiang Hu, Rui Jiang, and Mu Zhou. Deepcdr: a hybrid graph convolutional network for predicting cancer drug response. *Bioinformatics*, 36(Supplement_2):i911–i918, 2020.

[15] Hossein Sharifi-Noghabi, Olga Zolotareva, Colin C Collins, and Martin Ester. Moli: multi-omics late integration with deep neural networks for drug response prediction. *Bioinformatics*, 35(14):i501–i509, 2019.

[16] Pengfei Liu, Hongjian Li, Shuai Li, and Kwong-Sak Leung. Improving prediction of phenotypic drug response on cancer cell lines using deep convolutional network. *BMC bioinformatics*, 20(1):1–14, 2019.

[17] Likun Jiang, Changzhi Jiang, Xinyu Yu, Rao Fu, Shuting Jin, and Xiangrong Liu. Deeptta: a transformer-based model for predicting cancer drug response. *Briefings in bioinformatics*, 23(3):bbac100, 2022.

[18] Lei Deng, Ziyu Fan, Xiaojun Xiao, Hui Liu, and Jiaxuan Zhang. Dual-channel heterogeneous graph neural network for predicting micrornamediated drug sensitivity. *Journal of Chemical Information and Modeling*, 62(23):5929–5937, 2022.

[19] Shuting Jin, Yue Hong, Li Zeng, Yinghui Jiang, Yuan Lin, Leyi Wei, Zhuohang Yu, Xiangxiang Zeng, and Xiangrong Liu. A general hypergraph learning algorithm for drug multi-task predictions in micro-to-macro biomedical networks. *PLOS Computational Biology*, 19(11):e1011597, 2023.

[20] Wei Peng, Tielin Chen, and Wei Dai. Predicting drug response based on multi-omics fusion and graph convolution. *IEEE Journal of Biomedical and Health Informatics*, 26(3):1384–1393, 2021.

[21] Hong Wang, Chong Dai, Yuqi Wen, Xiaoqi Wang, Wenjuan Liu, Song He, Xiaochen Bo, and Shaoliang Peng. Gadrp: graph convolutional networks and autoencoders for cancer drug response prediction. *Briefings in Bioinformatics*, 24(1):bbac501, 2023.

[22] Yanli Wang, Jewen Xiao, Tugba O Suzek, Jian Zhang, Jiyao Wang, and Stephen H Bryant. Pubchem: a public information system for analyzing bioactivities of small molecules. *Nucleic acids research*, 37(suppl_2):W623–W633, 2009.

[23] Greg Landrum. Rdkit documentation. *Release*, 1(1-79):4, 2013.

[24] Tuan Nguyen, Giang TT Nguyen, Thin Nguyen, and Duc-Hau Le. Graph convolutional networks for drug response prediction. *IEEE/ACM transactions on computational biology and bioinformatics*, 19(1):146–154, 2021.

[25] Steven M Corsello, Rohith T Nagari, Ryan D Spangler, Jordan Rossen, Mustafa Kocak, Jordan G Bryan, Ranad Humeidi, David Peck, Xiaoyun Wu, Andrew A Tang, et al. Discovering the anticancer potential of non-oncology drugs by systematic viability profiling. *Nature cancer*, 1(2):235–248, 2020.