MAF-CDR: multi-omics data integration for cancer-drug response prediction model

Sizhe Zhang[1,a]
[1]School of Software and Engineering,
Xinjiang University,
Urumqi, Xinjiang, China
[a]e-mail: 1845953587@qq.com

Xuecong Tian[2,b]
[2]School of Information Science and Engineering,
Xinjiang University,
Urumqi, Xinjiang, China
[b]e-mail: 599502625@qq.com

Ying Su[3,c]
[3]School of Information Science and Engineering,
Xinjiang University,
Urumqi, Xinjiang, China
[c]e-mail: 2550501601@qq.com

Wanhua Huang[4,d]
[4]School of Information Science and Engineering,
Xinjiang University,
Urumqi, Xinjiang, China
[d]e-mail: 199999@stu.xju.edu.cn

Xiaoyi Lv[5,e]
[5]School of Software and Engineering,
Xinjiang University,
Urumqi, Xinjiang, China
[e]e-mail: xjuwawj01@163.com

Keao Li[6,f*]
[6]New Cicon Pharmaceutical Co LTD,
Urumqi, Xinjiang, China,
[*]Corresponding author: [f]260687990@qq.com

***Abstract*—Achieving precise application and rapid discovery of anticancer drugs has been an important topic in medicine and pharmacology. Traditional anticancer drug discovery relies on in vivo experiments and in vitro drug screening, methods that are important in the discovery of new drugs, yet require more time and resources. Although there are some traditional methods based on machine learning to predict drug-cancer cell lineage response (CDR), these methods are usually based on a single source of information, limiting the interpretation of the panoramic view of cancer cells. Using multi-omics information to comprehensively respond to cancer cells and anticancer drug responses remains a great challenge. In this experiment, we propose MAF-CDR, a CDR prediction model based on the alignment and fusion of multi-omics information. The model fuses multi-omics information in a unified framework and achieves cross-modal information representation through adversarial training to achieve aligned fusion of different modalities. Meanwhile, in order to better capture the correlation between drugs and cancer cell lines, MAF-CDR constructed a graph neural network encoder based on contrast learning to achieve accurate CDR prediction. The experimental results showed that the AUC value of MAF-CDR was higher than 0.8, which was better than other current baseline methods. This indicates that the MAF-CDR method can efficiently promote the development and innovation in the field of drug discovery, and proves the great potential value of deep learning methods in guiding the rapid discovery and precise application of anti-cancer drugs.**

Keywords-Multimodal Fusion, Graph neural network, Contrastive learning, Drug discovery

I. INTRODUCTION

Cancer is a major public health problem worldwide [1], and drug discovery plays a key role in precision cancer therapy. With the development of gene sequencing technology, cancer genomic information and cancer drug susceptibility information provides valuable information resources for precise drug discovery and provide researchers with new research directions [2][3]. These important data provide strong support for the development of cancer cell-drug response mechanism (CDR) prediction models and significantly shorten the cycle time required for drug discovery [2][4].

However, although previous cancer-drug prediction models have made great progress, they are usually based on a single source of information and provide only a limited view of cancer cell characteristics. Therefore, in order to overcome the limitation of a single information source, this experiment developed Multi-modal Alignment Fusion for Cell-Drug Response (MAF-CDR). The multi-modal alignment fusion (MAF) module was used to integrate different genomic information to reveal the characteristics and mechanisms of cancer in a more comprehensive way [5]. Meanwhile, a graph neural network (GNN) based on contrast learning was used as an encoder to improve the accuracy of drug-cancer cell response prediction while effectively solving the imbalance between cancer cell lines and drug labels.

II. METHOD

A. Dataset

For the processing of the dataset, this experiment was based on the processing method of Liu et at. [4] Specifically, for the multi-omics data of cancer cells, we downloaded genomic mutation, transcriptome, and DNA methylation information related to various cancer cell lines from Cancer Cell Line Encyclopedia (CCLE) [6] and DeMap website. For drug

2023 4th International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI) | 979-8-3503-4311-3/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICHCI58871.2023.10277928

molecular data, the IC50 of drug and cancer cell line response was collected from Genomics of Drug Sensitivity in Cancer (GDSC) [7] as a response indicator, and the SMILES sequence of each drug was obtained from PubChem [8].

The labels (sensitive or insensitive) were set for the drug IC50 according to the cancer cell screening concentration thresholds provided by GDSC. The final experiment collected a dataset containing 561 cancer cell lines and 222 drugs, which had a total of 1375 sensitive responses and 5932 insensitive responses.

B. Multimodal Alignment Fusion Module

The general framework of the multimodal alignment fusion module (MAF) is shown in Figure 1. In order to extract multimodal information, MAF first obtains each modality invariant representation for the multimodal data of cancer cell lines by the shared encoder E^S:

$$F_m^S = E^S(f_m; \theta^s), m \in \{G, E, T\} \quad (1)$$

where f_m , $m \in \{G, E, T\}$ denote the multi-omics features of cancer cell lines, respectively. The shared encoder E^S maps the multi-source data to a common space by optimizing the parameter θ^s to learn modality invariant representation.

Meanwhile, two discriminators, D_1 and D_2, were used to align and reduce the modal gap. Specifically, the discriminator uses the genomics information as the main modality and the rest information as the source modality. The alignment of the two modalities is achieved by minimizing the alignment loss L_{alig}, which is implemented as follows:

$$\begin{aligned} L_{alig} = &-\left[\log\left(D_1\left(F_G^S\right)\right) + \log\left(D_2\left(F_T^S\right)\right)\right] \\ &-\left[\log\left(1 - D_1\left(F_G^S\right)\right) + \log\left(D_1\left(F_E^S\right)\right)\right] \\ &-\left[\log\left(1 - D_2\left(F_T^S\right)\right) + \log\left(1 - D_2\left(F_E^S\right)\right)\right] \end{aligned} \quad (2)$$

where $D_1(\cdot)$ and $D_2(\cdot)$ denote the discriminator's identification values for different modal information. By training the shared encoder E^S to minimize L_{alig}, the two source modalities are aligned to the main modal to confuse the discriminator, thus achieving multimodal alignment while learning modal invariant representation and alleviating the modal gap problem.

After achieving multimodal alignment, in order to unify the cross-modal representation, the experiments first adopt the gated network structure as the attention mechanism to obtain the weight factors for the invariant representation of each modality. The gated network is capable of adaptively adjusting the importance weights of the inputs according to the characteristics and contextual information of the input data, thus achieving a unified representation of the cross-modal states. The structure is implemented as follows:

$$\alpha_m = w_m * \left(F_G^S \oplus F_E^S \oplus F_T^S\right), m \in \{G, E, T\} \quad (3)$$

where w_m denotes the learnable weight matrix and $\oplus$ denotes the concatenation operator. the obtained weight factors are multiplied with the invariant features of each modality and combined with the low-rank multimodal fusion (LMF) method of Yu et al. [9] to make full use of the aligned inter-modal relationship features to obtain a unified cross-modal representation. The implementation is as follows.

$$F_{fu}^S = \mathrm{LMF}[\mathrm{ReLU}(\alpha_m * F_m^S)], m \in \{G, E, T\} \quad (4)$$

where ReLU denotes the activation function. In addition, three independent encoders E^P were designed to obtain each histological modality-specific representation F_m^P:

$$F_m^p = E^p\left(f_m; \theta_m^p\right), m \in \{G, E, T\} \quad (5)$$

The modality-specific representation encoder E^P maps the multi-omics data into the private space of each modality by optimizing the parameters θ_m^p under different modal spaces to learn the modality-invariant representation. In order to ensure that mode-invariant and mode-specific representations can learn different features between modes separately, this experiment introduces the orthogonality constrained loss L_{orth} to achieve feature separation in public and private spaces.

$$L_{orth} = \sum_{m \in \{G,E,T\}} \left\| H_m^p H_m^{s\,T} \right\|_F^2 \quad (6)$$

where H_m^p and H_m^s denote the matrices, whose rows represent the hidden vectors F_m^p and H_m^s for modality m, $\|\cdot\|$ is the squared Frobenius norm. To ensure that mode-specific and mode-invariant features can jointly provide multimodal fusion information to the molecule, we developed the encoder D_{rec} to perform fusion reconstruction of the learned molecular features by reconstructing to ensure that the details of the respective modalities can be captured. The specific implementation is as follows:

$$F_m = D_{rec}\left(F_{fu}^S \oplus H_m^p\right), m \in \{G, E, T\} \quad (7)$$

However, although multiple loss constraints have been aligned for different modalities, the reconstruction fusion process inevitably creates encoders that learn different information from the original features, and encourage the reconstructed cross-modal feature vectors to be similar to the original features at the same time, we use the reconstruction loss L_{rec} for the constraints, which is implemented as follows:

$$L_{rec} = \frac{1}{2} \sum_{m \in \{G,E,T\}} \|F_m - f_m\|_2^2 \quad (8)$$

Where $\|\cdot\|_2^2$ is denotes squared L_2-norm. Finally, the obtained cross-modal representation F_{fu}^S and the mode-specific representation F_m^P are spliced to obtain the final fusion feature F_{Cell}:

$$F_{Cell} = F_{fu}^{S} \oplus F_{G}^{P} \oplus F_{E}^{P} \oplus F_{T}^{P} \quad (9)$$

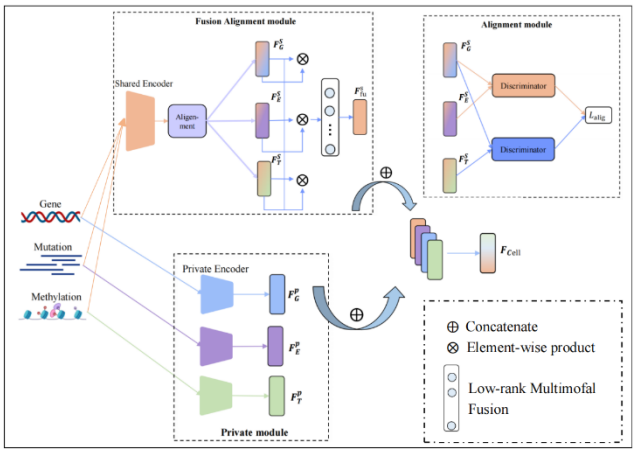

Figure 1. Illustration of the proposed MAF framework

C. Drug Node Representation Module

In this experiment, the SMILES sequences in the dataset were compiled into a molecular graph using the RDKit package, with atoms and chemical bonds as nodes and edges of the graph, respectively. The node features were processed similarly as in Zhu et al. [10], and each atomic node feature in the graph was of 75 dimensions. Meanwhile, to extract the structural features of the molecular graph more fully, the graph convolution algorithm (GCN) was used to extract the embedded representation of molecular nodes in this experiment, and finally the drug molecule embedding representation F_{Drug} was obtained.

D. CDR Prediction Module

The graph encoder framework based on contrast learning is shown in Figure 2. In this experiment, the drug-cancer cell line reaction is represented as an undirected graph G = (V, E). Where V denotes the set of nodes, which includes cancer cell nodes and drug nodes, and cancer cell fusion feature F_{Cell} and drug embedding representation F_{Drug} are used as the node features of G. E denotes the set of edges, and 1375 sensitive responses of drug-cancer cell line are used as the edges of G.

The potential relationship $\Omega(G) \rightarrow H_G$ between graph G nodes is learned using the encoder Ω based on GCN algorithm, and H_G denotes the node features obtained after G optimization. This experiment further obtains the final embedding H_{Drug} of drug nodes and the final embedding H_{Cell} of cancer cell lines from H_G for the final prediction. To perform drug-cancer cell line response prediction, this experiment predicts the response probability by inner product and Sigmoid function as follows:

$$\hat{P} = \text{Sigmoid}\left(H_{Drug} H_{cell}^{T}\right) \quad (10)$$

Task loss L_{task} for cancer cell line-drug response prediction measured by cross-entropy loss function:

$$L_{task} = -\frac{1}{N} \sum_{(c,d) \in N} P \log \hat{P} + (1 - P) \log\left(1 - \hat{P}\right) \quad (11)$$

where N is the number of data sets, c and d represent cancer cell line nodes and drug nodes, respectively, and P represents the true label.

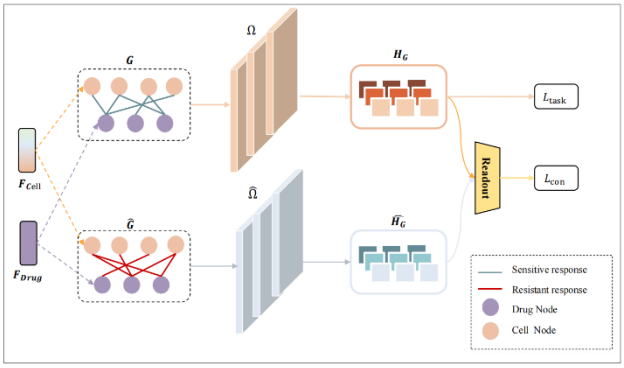

Figure 2. Contrast learning-based graph encoder framework

E. Comparative Learning Module

To construct the contrast learning graph $\hat{G}$, this experiment continued to represent the drug-cancer cell line insensitive responses as $\hat{G} = (V, \hat{E})$, and the 5932 insensitive responses of the drug-cancer cell line were used as the edges of $\hat{G}$. Similar to the treatment of G, the nodal features $\widehat{H_G}$ of the comparison graph $\hat{G}$ were obtained in this experiment using the encoder $\hat{\Omega}$. To maximize the drug-cancer cell response relationship, the graph-level embeddings s and $\hat{s}$ between the two graphs were further obtained from H_G and $\widehat{H_G}$ by reading out the function D_{con}. This experiment increases the sensitivity of the model to the response by expanding the two graph-level embedding losses L_{con} by:

$$L_{con} = -\frac{1}{2|V|}\left(\sum_{V} \log U(H_G, s) + \sum_{V} \log\left(1 - U\left(\widehat{H_G}, s\right)\right)\right) - \frac{1}{2|V|}\left(\sum_{V} \log U\left(\widehat{H_G}, \hat{s}\right) + \sum_{V} \log\left(1 - U(H_G, \hat{s})\right)\right) \quad (12)$$

where V is the feature dimension of the node features H_G and $\widehat{H_G}$, and $U(\cdot,\cdot)$ is a comparison discriminator constructed from a simple bilinear function for estimating the similarity between node-level embeddings and graph-level embeddings. Ultimately, in order to achieve both the alignment fusion task of the model and the CDR prediction task based on contrast learning, this experiment combines all of the above loss functions to optimize the following objective function:

$$L_{task} = L_{task} + \alpha L_{alig} + \beta L_{orth} + \gamma L_{rec} + \delta L_{Con} \quad (13)$$

where α, β, γ and δ are hyperparameters that balance the contributions of different loss functions, respectively. By simultaneously optimizing this objective function, the MAF-

CDR framework was successfully constructed to achieve accurate prediction of drug-cancer cell line responses.

III. RESULT

In order to evaluate the performance of the model, the area under curve (AUC) and the area under the precision-recall (AUPR) were used as evaluation metrics in this experiment, and the data set was divided into training set and independent test set in the ratio of 8:2, while a five-fold cross-validation was performed to ensure the generalization of the model.

In order to compare the performance of the models and evaluate the advantages of MAF-CDR, the current deep-learning models with excellent performance in predicting drug-cancer cell line responses were selected as the baseline models in this experiment [4]. In order to ensure the reliability of the results, the same data segmentation method was used for each type of baseline model algorithm. As shown in Figure 3. the MAF-CDR proposed in this paper has a better performance than the baseline model with AUC and AUPR of 0.826 and 0.496, respectively. Compared with other baselines, the AUC scores are about 2% and 5% higher than the next best baseline models, NRL2DRP and tCNNs, and the AUPR scores are about 7% and 15% higher, which fully demonstrates the excellent prediction performance of the proposed model.

The above experimental results show that MAF-CDR outperforms current advanced predictive CDR methods, and this superiority can be attributed to the multimodal alignment fusion and comparative learning strategy adopted by MAF-CDR. By integrating multi-omics information, MAF-CDR can capture the relevant features of cancer cell lines more comprehensively and thus improve the prediction performance. Meanwhile, the application of contrast learning enables the model to effectively deal with label imbalance, further improving the prediction accuracy. Taken together, MAF-CDR shows excellent performance in the inductive learning process, which provides powerful support for precision drug discovery and treatment.

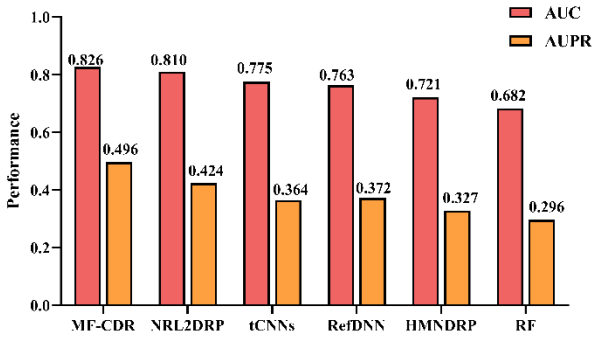

Figure 3. Comparison results between MF-CDR and other baseline methods

IV. CONCLUSION

In this paper, we propose a drug-cancer cell line response prediction model based on aligned fusion of multimodal information, namely MAF-CDR, which aligns and fuses multimodal genomic information of cancer cell lines to obtain a comprehensive cancer cell line landscape. In order to better capture the correlation between drugs and cancer cell lines, this experiment proposes to use a comparative learning approach as part of the graph optimizer to achieve accurate CDR prediction while solving the label imbalance problem. The final experimental results show that the proposed MAF-CDR is superior to other advanced drug-cancer cell line prediction models and has great potential value in guiding the rapid discovery and precise application of anti-cancer drugs.

In the future, we will continue to improve the MAF-CDR framework, such as (1) fusing more multimodal data of drug molecules to obtain more comprehensive molecular characterization; (2) deploying the constructed MAF-CDR model to the Web site for researchers to facilitate cancer drug discovery.

ACKNOWLEDGMENT

Our work is supported by the Major scientific and technological special projects in Xinjiang Uygur Autonomous Region (2022A03016-4).

REFERENCES

[1] Ferlay J, Colombet M, Soerjomataram I et al. Cancer statistics for the year 2020: An overview, International journal of cancer 2021;149:778-789.

[2] Roelofs C, Hollande F, Redvers R et al. Breast tumour organoids: promising models for the genomic and functional characterisation of breast cancer, Biochemical Society Transactions 2019;47:109-117.

[3] Ben-David U, Beroukhim R, Golub TR. Genomic evolution of cancer models: perils and opportunities, Nature Reviews Cancer 2019;19:97-109.

[4] Liu X, Song C, Huang F et al. GraphCDR: a graph neural network method with contrastive learning for cancer drug response prediction, Briefings in Bioinformatics 2022;23:bbab457.

[5] Karim MR, Islam T, Lange C et al. Adversary-aware multimodal neural networks for cancer susceptibility prediction from multiomics data, IEEE Access 2022;10:54386-54409.

[6] Nusinow DP, Szpyt J, Ghandi M et al. Quantitative proteomics of the cancer cell line encyclopedia, Cell 2020;180:387-402. e316.

[7] Yang W, Soares J, Greninger P et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells, Nucleic Acids Research 2012;41:D955-D961.

[8] Kim S, Chen J, Cheng T et al. PubChem 2023 update, Nucleic Acids Research 2022;51:D1373-D1380.

[9] Yu W, Xu H, Yuan Z et al. Learning modality-specific representations with self-supervised multi-task learning for multimodal sentiment analysis. In: Proceedings of the AAAI conference on artificial intelligence. 2021, p. 10790-10797.

[10] Zhu W, Zhang Y, Zhao D et al. HiGNN: A Hierarchical Informative Graph Neural Network for Molecular Property Prediction Equipped with Feature-Wise Attention, Journal of Chemical Information and Modeling 2022.