



可解释药物协同作用预测的多模态相互关注网络

SynergyX: a multi-modality mutual attention network for interpretable drug synergy prediction

Yue Guo , Haitao Hu, Wenbo Chen, Hao Yin, Jian Wu, Chang-Yu Hsieh, Qiaojun He and Ji Cao 

Corresponding authors: Jian Wu, Second Affiliated Hospital School of Medicine, School of Public Health, Zhejiang University, Hangzhou 310058, Zhejiang, China. Tel.: +86-571-86713379; E-mail: wujian2000@zju.edu.cn; Chang-Yu Hsieh, College of Pharmaceutical Sciences, Zhejiang University, Hangzhou 310058, Zhejiang, China. E-mail: kimhsieh@zju.edu.cn; Qiaojun He, College of Pharmaceutical Sciences, Zhejiang University, Hangzhou 310058, Zhejiang, China. Tel.: +86-571-88208400; E-mail: qiaojunhe@zju.edu.cn; Ji Cao, College of Pharmaceutical Sciences, Zhejiang University, Hangzhou 310058, Zhejiang, China. Tel.: +86-571-88208401; E-mail: caoji88@zju.edu.cn

Abstract

SynergyX: 多模态相互注意网络提高抗肿瘤药物协同作用预测

Discovering effective anti-tumor drug combinations is crucial for advancing cancer therapy. Taking full account of **intricate** biological interactions is highly important in accurately predicting drug synergy. However, the extremely limited prior knowledge poses great challenges in developing current computational methods. To address this, we introduce SynergyX, **a multi-modality mutual attention network to improve anti-tumor drug synergy prediction**. It dynamically captures cross-modal interactions, allowing for the modeling of complex biological networks and drug interactions. A convolution-augmented attention structure is adopted to integrate multi-omic data in this framework effectively. Compared with other state-of-the-art models, SynergyX demonstrates superior predictive accuracy in both the General Test and Blind Test and cross-dataset validation. By exhaustively screening combinations of approved drugs, SynergyX reveals its ability to identify promising drug combination candidates for potential lung cancer treatment. Another notable advantage lies in its multidimensional interpretability. Taking Sorafenib and Vorinostat as an example, SynergyX serves as a powerful tool for uncovering drug-gene interactions and deciphering cell selectivity mechanisms. In summary, SynergyX provides an illuminating and interpretable framework, poised to **catalyze** the expedition of drug synergy discovery and deepen our comprehension of rational combination therapy.

Keywords: drug synergy prediction; combination therapy; mutual attention; multi-modality; transformer

预测药物协同作用

INTRODUCTION

Drug synergy therapy has emerged as an effective strategy in cancer treatment, as it can enhance therapeutic efficacy and reduce toxic side effects. It is also regarded as an effective strategy to overcome drug tolerance [1]. **At present, the discovery of drug synergy mainly relies on experimental screening and clinical experience, resulting in low efficiency and high costs.** In addition, due to the multiple heterogeneity of cancer and the uncertainty of drug synergistic mechanisms, relying solely on human expertise falls short of fully understanding the interactions between drugs and intricate biological networks. Therefore, it is essential to develop an efficient strategy for discovering drug synergies [2].

With the emergence of some public datasets of drug combinations [3], researchers have grown interested in exploring

computational methods for **predicting drug synergy**. Machine learning (ML)-based approaches, especially deep learning (DL) methods, have gained increasing popularity in the field of drug combination prediction owing to their remarkable predictive performance and expansive prediction capabilities [4]. **Generally, there are two approaches to building DL models for drug synergy prediction.** 一般来说,有两种方法来建立药物协同作用预测的深度学习模型。

The first approach category involves using two branches to **extract features from drugs and cells** separately, followed by feature concatenation for synergy prediction. DeepSynergy [2] and MatchMaker [5] were the very early DL algorithms for predicting drug synergy. They consider the physicochemical properties of drugs and gene expression data as input features, using layers of fully connected neural networks for prediction. Subsequent research attempted to introduce additional input

Yue Guo is a Ph.D. student at College of Pharmaceutical Sciences, Zhejiang University, focusing on deep learning for biological network construction and drug target discovery.

Haitao Hu is a master's student at Zhejiang University, focusing on artificial intelligence for single-cell data analysis.

Wenbo Chen is a master's graduate from Zhejiang University, specializing in bioinformatics and deep learning for biomedicine.

Hao Yin is a master's candidate at Polytechnic Institute, Zhejiang University, conducting research in artificial intelligence for drug discovery.

Jian Wu is a professor at the School of Public Health, Zhejiang University, with research interests in artificial intelligence for clinical diagnosis and treatment.

Chang-Yu Hsieh is a professor at the College of Pharmaceutical Sciences, Zhejiang University, specializing in the application of computational technology in drug development.

Qiaojun He is a professor at the College of Pharmaceutical Sciences, Zhejiang University, with research interests in oncology pharmacology and innovative drug development.

Ji Cao is a professor at the College of Pharmaceutical Sciences, Zhejiang University, specializing in drug discovery and novel target discovery.

Received: September 24, 2023. Revised: January 11, 2024

© The Author(s) 2024. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

features and improve the feature extractors. For example, Deep-DDS innovatively **utilizes graph neural networks for drug feature extraction** [6]. EC-DFR was the first to introduce drug-induced gene expression profiles [7]. MGAE-DC constructed multi-channel graph autoencoders, taking into account synergistic, additive and antagonistic combinations, to obtain drug embeddings that are more suitable for drug synergy tasks [8]. These studies demonstrate the **tremendous** potential of DL methods in drug synergy prediction. However, a limitation is their insufficient emphasis on the interactions between cell lines and drug pairs, which is crucial for the effects of drug combinations.

Other approaches often **construct graphs to capture drug, cell and gene associations**. These methods fully leverage drug-target interactions (DTIs), drug-drug interactions (DDIs) and protein-protein interactions (PPIs) to characterize biological networks and drug-cell interactions. In the early work, Jiang P et al. constructed cell line-specific graphs by integrating drug-drug combinations, DTI and PPI to predict potential drug synergy [9]. Subsequently, GraphSynergy adapted a GCN-based network to capture relationships between drug-targeted proteins and cell line-associated proteins [10]. NEWMIN, on the other hand, built a multiplex drug-drug similarity network [11]. These methods aim to leverage biological prior knowledge such as DTI and DDI to enhance drug synergy prediction. However, a significant challenge lies in obtaining sufficient features in the real world. This may also restrict the model's applicability to new drugs and those with limited research.

A potential solution is to encode the cross-modal interactions *in silico* rather than relying solely on prior knowledge. In a previous study, DTSyn introduced a two-branch Transformer encoder framework to extract biological associations among molecules, proteins and cell lines [12]. It pioneered using a model-based approach for interaction extraction, but certain limitations may exist. Specifically, using pooled chemical features obtained from GCN blocks could cause the subsequent Transformer encoder to overlook the local information of molecules. Additionally, directly concatenating drug and gene representations may not be optimal for handling multi-modal data and cross-modal interactions. A similar work, DFFNDDS, innovatively attempted to utilize two networks (a multi-head attention mechanism and a highway network) to extract and combine input features [13]. However, it also directly concatenated different modal data of drugs and cell lines as the input for its fusion encoder. Subsequently, Deep-TraSynergy also focused on multi-modality, attempting to integrate a variety of prior knowledge such as protein-protein, drug-protein and cell-protein interactions, and then employed a simple matrix multiplication for multi-modal fusion [14]. DGSSynADR made a novel attempt using Multi-Layer Perceptron (MLP)-based bilinear predictors, which fused the feature representations of the two drugs to enhance model performance [15]. These studies made preliminary attempts to employ specialized yet simplistic modules for achieving multi-modal interactions. However, prior research [16] has demonstrated that focusing algorithm design on multi-modal interaction steps rather than feature extraction steps can yield better modal fusion and predictive outcomes.

Based on the above considerations, this study aims to incorporate the mutual-attention mechanism [17] to introduce multidimensional cross-modal interactions in the drug synergy prediction process. Herein, we propose a multi-modality mutual attention network named SynergyX, specifically focused on extracting cross-modal biological knowledge. It features a 'sandwich'-like block consisting of multiple mutual-attention

modules and self-attention modules, allowing the extraction of drug-cell and drug-drug interactions. To ensure that the attention modules explicitly focus on drug substructures, we employ substructure encoding as the initial representation of drugs. Furthermore, by utilizing both CNN to extract local features and attention modules to capture context features, we achieve the effective integration of multi-omic features for cell lines. Extensive experiments demonstrate that SynergyX outperforms other state-of-the-art methods, including ML and DL methods. Additionally, we highlight the substantial potential of SynergyX in unraveling the underlying mechanisms of drug synergy.

MATERIALS AND METHODS

Synergy dataset

Drug combination data were collected from a comprehensive database, DrugComb (v1.5) [3, 18], which includes 739 964 combination experiments. **It provides a novel drug synergy measurement called the S score [19], which defines the level of drug synergy by measuring the disparity between the dose-response curves of a drug combination and its single drugs.** Malyutina et al. demonstrated that the S score exhibits high accuracy (AUC>0.99) in predicting the most synergistic and antagonistic drug combinations when compared with all four existing synergy scores (HSA [20], Bliss [21], Loewe [22] and ZIP [23]) [19]. We chose the S score as the quantitative measure for the interactions of a drug pair in a specific cell line.

We selected all available data from the **DrugComb database** for further data cleansing. **First**, we removed entries with unclear or missing information regarding drugs and cell lines. **Next**, we identified a significant data imbalance issue within the dataset, where 2157 drugs (85.9%) appeared fewer than 10 times, contributing only 1.37% (4587 items) of the dataset. To address this data imbalance and enhance the dataset's quality, **we eliminated drugs that appeared fewer than 10 times**. Furthermore, we employed the 3σ principle [24] to identify outliers within the dataset, and then used the mean ± 3 standard deviations as a threshold and assigned boundary values to scores outside this range. Ultimately, **we obtained a dataset of 330 917 drug combinations involving 354 drugs and 167 cell lines (Table 1)**. It is worth noting that despite DrugComb being one of the largest drug combination databases available, the labeled data are still significantly smaller compared with the prediction space. In our case, the combination space of 354 drugs across 167 cell lines amounts to approximately 21 million possibilities, whereas the available data only cover a mere 1.56% of the total possibilities.

Additionally, four independent datasets were used for further evaluation of the model: the O'Neil dataset [25], the ALMANAC dataset [26], the Oncology Screen dataset [25] and the DrugCombDB dataset [27] (Table 1). All datasets underwent a data processing workflow consistent with the DrugComb dataset. **In the O'Neil dataset**, 23 062 samples involving 39 cell lines and 38 unique drugs were included. The Loewe scores were retained as the synergy score. **As for the ALMANAC dataset**, a total of 296 503 drug combination data were retained, covering 118 drugs and 60 cell lines. The synergy scores **in the ALMANAC dataset** were measured by ComboScore, which quantifies synergy as the cumulative discrepancy between observed and expected growth fractions across all dose combinations [26]. The remaining two datasets, the Oncology Screen dataset and a shortened version of the DrugCombDB dataset, align with the benchmark datasets utilized in prior study [14]. We obtained 4176 drug synergy records for **the Oncology Screen dataset**, involving 21 unique drugs and

Table 1: Details of the datasets after data cleansing. 细胞系

Dataset	Drugs	Cell lines	Combinations	Synergy score
DrugComb	354	170	330 917	S score
O'Neil	38	39	23 062	Loewe score
ALMANAC	118	118	296 503	ComboScore
Oncology Screen	21	29	4176	ZIP score
DrugCombDB	600	68	60 932	ZIP score

使用四个独立的数据集对模型进行进一步评估

29 cell lines. For the DrugCombDB dataset, we acquired 60 792 records involving 600 drugs and 68 cell lines. In both datasets, the ZIP score was used as the synergy score.

Drug substructure encoding

For each drug, its SMILES were obtained from the ChEMBL database [28], and after that, the RDKit library [29] was used to convert it into canonical SMILES. We further transformed the drug's SMILES format into the Explainable Substructure Partition Fingerprint (ESPF) [30] to enable the subsequent attention module to capture interpretable pharmacophore information. ESPF forms a sequential encoding of the drug by decomposing it into a set of discrete, moderate-sized substructures. Huang et al. identified around 2700 valuable substructures to form a substructure dictionary [30]. In the end, drugs were represented as fixed-length vectors of size 165, matching the dimensions of cell features. If the number of substructures in a drug is less than 165, the remaining positions will be filled with zeros.

Following the aforementioned steps, each drug is represented as a one-hot vector, denoting the substructure index for the i th substructure of the drug sequence. In line with the approaches of previous studies [31], we further transform these features. Initially, we employ a learnable dictionary lookup matrix $W_c \in \mathbb{R}^{k \times l}$, where k represents the maximum number of substructures and l represents the vector dimension. This allows us to transform each drug's substructure index into a vector of length l (where $l=128$), thereby obtaining the initial features of each drug $D_i \in \mathbb{R}^{165 \times 128}$. Subsequently, to capture the positional information of the substructures for a drug, a positional representation was computed using a lookup dictionary $W_{pos} \in \mathbb{R}^{k \times l}$:

$$P_i = W_{pos} I_i, \quad (1)$$

where $I_i \in \mathbb{R}^l$ is a one-hot vector where the i th position is set to 1. Therefore, we generated the final drug substructure encoding by summing the initial features and positional representations:

$$X_d = D_i + P_i \quad (2)$$

After this feature transformation, we obtained the final drug substructure encoding $X_d \in \mathbb{R}^{165 \times 128}$, which serves as the input for the subsequent Cross-Modal Fusion Encoder.

Unified gene set

In order to identify a subset of potentially significant genes that contribute to drug synergy, we selected marker genes from various sets and integrated them. The Library of Integrated Network-Based Cellular Signatures (LINCS) project has published a landmark gene set comprising 978 genes, which is known to capture 81% of the information from the entire transcriptome [32]. These genes are included in our final gene set. Besides, we chose the top 1000 genes with the highest variation in gene expression for

cancer cell lines in the CCLE project [33, 34] and the top 1500 varied genes from tumor samples in TCGA [35]. These genes can effectively capture the heterogeneity of different samples.

Furthermore, based on the PPI network contained in the STRING database [36, 37], we filtered out interactions with a combined score higher than 0.7 and then identified the top 1000 proteins that exhibit the most interactions with other proteins. Genes encoding these proteins are commonly regarded as being more crucial and functionally significant in biological networks. In addition, 1000 drug-targeted genes associated with the highest number of drugs were also included, derived from the STITCH database [38]. Finally, we combined the five sets of selected genes mentioned above, which are categorized as L1000_gene, ccle_gene, tcga_gene, ppi_gene and dti_gene, to form the final set consisting of 4079 genes (Figure 4), after removing genes with missing features.

细胞系的多基因组整合

Multi-Omic integration for cell lines

SynergyX utilized six types of omic data downloaded from the DepMap database [39, 40]: gene expression (exp), gene mutations (mut), gene copy number (cn), gene methylation (met), gene effect (eff) and gene dependency probability (dep). The gene mutation data were processed as a binary variable, where 0 represented a normal gene and 1 represented a mutated gene. For each cell line, we aimed to preserve the omic features for the mentioned 4079 genes. While we collected complete gene features of exp, mut and cn across all 167 cell lines, there were instances of missing data for the remaining features. Specifically, for eff and dep, data were collected for 3456 genes across 138 cell lines, and for met, data were collected for 2279 genes across 143 cell lines. In case any omic feature was missing, it was imputed with the average value of that gene across the remaining cell lines. We applied tanh normalization, similar to DeepSynergy [2], for preprocessing the raw data of different omic types. For a specific cell line, the six omic features were integrated in a fixed gene order, resulting in an input feature $X_c^0 \in \mathbb{R}^{4079 \times 6}$.

Inspired by the convolution-augmented Transformer structure (Conformer) [41], we employ a 1D convolutional neural network (CNN) to preliminarily integrate raw multi-omic data of cell lines before the attention modules, as illustrated in (Figure 1A, B). Given the relatively long sequence of cell line features, we introduce subsampling using a maxpooling layer between the convolutional layers. CNNs excel at extracting local features, and when combined with cascaded attention modules that are good at modeling long-range global context, they lead to a more adequate representation of cell line features. After the CNN module, we obtained integrated multi-omic features of cell lines $X_c \in \mathbb{R}^{165 \times 128}$.

采用一维卷积神经网络 (CNN) 对细胞系的原始多组学数据进行初步整合

Cross-Modal Fusion Encoder 跨模态融合编码器

The core module of SynergyX is the Cross-Modal Fusion Encoder, which employs a variety of attention modules to achieve feature updating and fusion, while also extracting underlying drug-cell 跨模态融合编码器由多个注意力模块组成!

能够捕获整个转录组81%的信息

将药物的smile
转为
ESPE

将ESPE
放入子
结构注
意力模
块

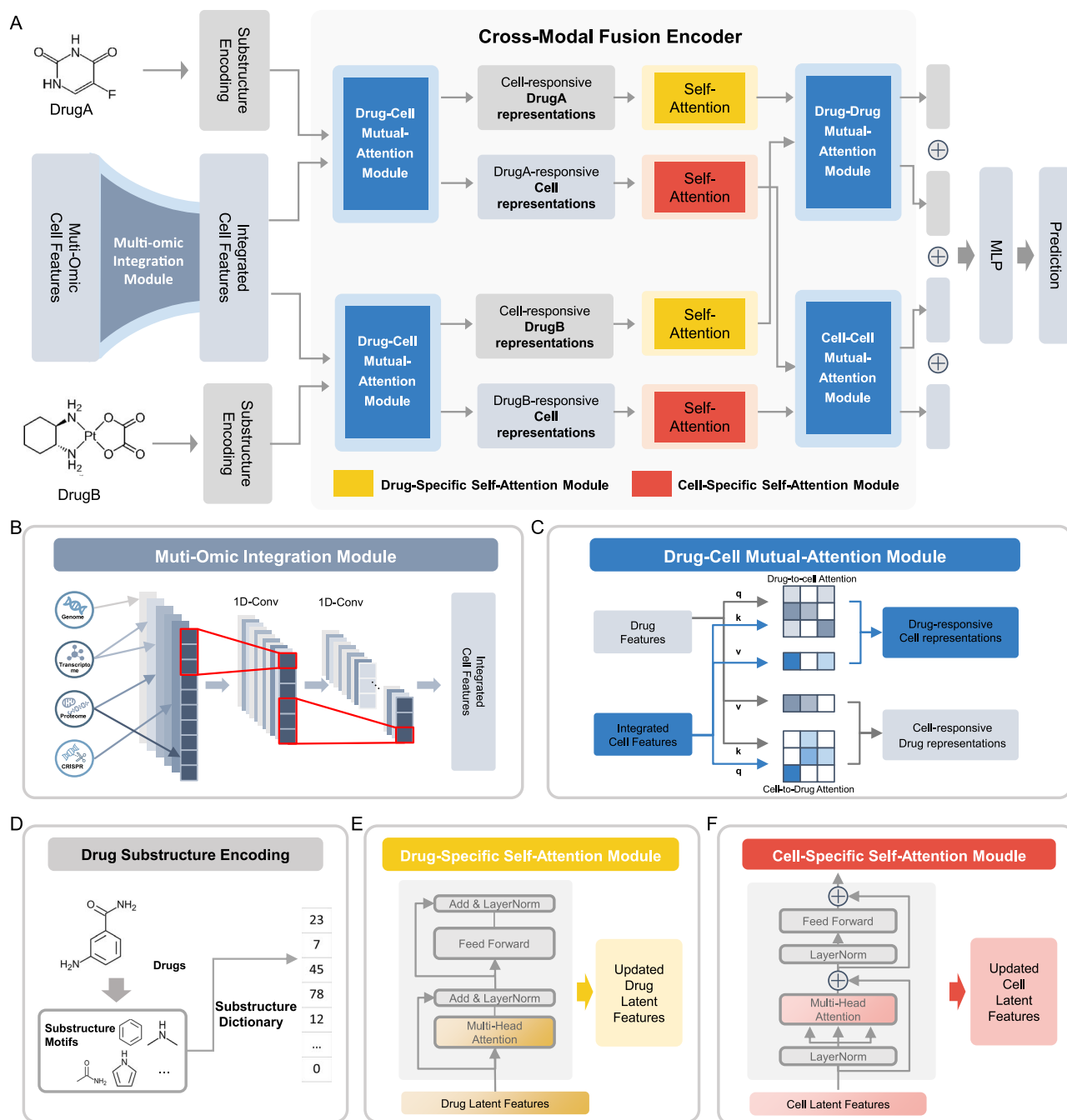


Figure 1. The overview of SynergyX. (A) The architecture of SynergyX consists of three main components: the Multi-Omic Integration Module, the Cross-Modal Fusion Encoder and the Prediction Module. SynergyX takes the features of a drug pair and a cell line as input and outputs a synergy score. (B) The Multi-Omic Integration Module employs a 1D convolutional network to integrate multi-omics data of cell lines. (C) The DCMA module utilizes drug features and integrated cell features to calculate cell-to-drug and drug-to-cell attention, which are used to update the features of drugs and cell lines. (D) Steps for generating drug substructure encoding. The drug molecules are decomposed into a series of appropriately sized substructures, which are then mapped to a predefined substructure dictionary to obtain the index of each substructure, serving as the initial representation of the drug. (E) and (F) illustrate the structure of the drug-specific self-attention module and cell-specific self-attention module.

and drug-drug interactions. We experimented with different combinations of mutual-attention and self-attention modules and ultimately settled on a three-layer ‘sandwich’ structure in our SynergyX. The outer layers are mutual-attention modules, while the middle layer consists of self-attention modules.

Cross-Modal Mutual-Attention

The Cross-Modal Fusion Encoder begins with a Drug-Cell Mutual-Attention (DCMA) module for cross-modal associated features

extraction (Figure 1A). The DCMA module consists of multi-head mutual-attention networks and feed-forward neural networks (Figure 1C). The core step involves calculating the weights of drug-to-cell and cell-to-drug interactions. The drug-to-cell attention matrix is used to obtain drug-responsive cell latent representations. In contrast, the cell-to-drug attention matrix is used to obtain cell-responsive drug latent representations.

Taking the generation of cell-responsive drug representations as an example, the following computations are performed. First,

the Query matrix is generated based on the drug's input features X_d . Then, the Key and Value matrix are generated based on the cell line's integrated multi-omic features X_c .

$$Q_i = X_d W_i^Q \quad (3)$$

$$K_i = X_c W_i^K \quad (4)$$

$$V_i = X_c W_i^V, \quad (5)$$

where Q_i , K_i and V_i are the Query, Key and Value of the i -th head in the multi-head attention; W_i^Q , W_i^K and W_i^V are learnable parameters.

Then, the cell-to-drug single-head attention is calculated as follows:

$$\text{Attention}_{c2d}(Q_i, K_i, V_i) = \text{Softmax}\left(\frac{Q_i K_i^T}{\sqrt{d_k}}\right) V_i, \quad (6)$$

where K_i^T represents the transpose of the matrix K_i , d_k denotes the dimension of K_i divided by the number of heads and Softmax denotes the normalization operation.

Finally, the outputs of each single-head attention are concatenated to obtain the updated drug features:

$$\text{MHAttn}_{c2d} = \text{Concat}(\text{head}_1, \dots, \text{head}_h) W^O, \quad (7)$$

where

$$\text{head}_i = \text{Attention}_{c2d}(Q_i, K_i, V_i),$$

where MHAttn_{c2d} represents the updated drug features obtained after performing multi-head attention between drugs and cell lines, h denotes the number of heads, W^O is a learnable parameter matrix and Concat represents the concatenation operation.

After the DCMA module, we obtain four fused features: the cell-responsive drugA's primary features, the cell-responsive drugB's primary features, the drugA-responsive cell's primary features and the drugB-responsive cell's primary features.

Self-attention for specific features updation

The second layer of the Cross-Modal Fusion Encoder is made up of two specific self-attention modules, namely the Drug-specific Self-Attention (DSSA) Module and the Cell-specific Self-Attention (CSSA) Module, which are designed to capture internal context within different entities (Figure 1A).

The DSSA Module takes as input the cell-responsive drugA's and cell-responsive drugB's primary features, which are utilized to update the hidden features of both drugs further. This module employs a classic Transformer encoder [42] to learn the interdependencies among various substructures within each drug (Figure 1E). Specifically, the DSSA Module begins by generating query, key and value matrices from the latent features of the drugs. These matrices are then fed into a multi-head self-attention layer to calculate the self-attention within each drug. Following the multi-head self-attention layer, we apply a feed-forward neural network with two-layer linear transformations and ReLU activation functions. Subsequently, the results pass through a residual connection layer and layer normalization to obtain the updated hidden features of the drugs.

Similarly, the CSSA Module captures the correlations between different genes in cells and updates the cell line features. The

input of CSSA is the drugA-responsive cell's primary features and the drugB-responsive cell's primary features, respectively. An architecture similar to Vision Transformer (ViT) [43] is adopted in the CSSA module (Figure 1E). The key difference from the DSSA module lies in the order of connections between different layers. Specifically, layer normalization is placed before the multi-head self-attention layer and before the feed-forward neural network to better stabilize the forward input distribution after passing through the CNN. The rest of the processes in the CSSA module are similar to those in the DSSA module.

Unimodal Mutual-Attention

In the third layer, we employ two Unimodal Mutual-Attention modules to extract coarser-grained fusion information. Among them, the drug-drug Mutual-Attention Module, named DDMA, integrates the latent features of drugA and drugB. The Cell-Cell Mutual-Attention Module, named CCMA, is utilized to combine the cell latent features that are responsive to drugA and drugB, respectively. Both DDMA and CCMA consist of a multi-head mutual-attention layer and a feed-forward neural network.

Similar to the process described in the cross-modal mutual-attention section, the DDMA initially uses drugA and drugB to generate their respective Key, Query and Value matrices. It then computes attention weights for drugA-to-drugB and drugB-to-drugA, utilizing these weights to update the features of drugB and drugA, respectively. This process captures the intricate interplays and potential synergies between the drugs. Consistent with the DSSA Module, the DDMA adopts the Transformer architecture as its foundation but replaces self-attention with mutual-attention. In contrast, the CCMA module utilizes the Vision Transformer (ViT) architecture as its backbone, following the design principles of the CSSA module. The CCMA module follows a computation process analogous to that of the DDMA. This crucial step effectively fuses the cell latent features responsive to drugA and drugB, enabling the module to reflect the combined impact of the two drugs on cellular gene networks.

Prediction module

After the aforementioned Cross-Modal Fusion Encoder, we obtain four final features: drugA's final features, drugB's final features, drugA-responsive cell's final features and drugB-responsive cell's final features. These features are flattened and concatenated, then passed into the Prediction module to output predicted synergy scores. The Prediction module is composed of three fully connected layers with ReLU activation functions applied between the layers. Additionally, we also introduced Dropout layers to prevent overfitting.

Experimental setups

Hyperparameters

Given the substantial number of adjustable hyperparameters in our model, exhaustively exploring all parameter combinations is challenging. Following thorough investigation, we determined initial values and defined the scope for adjustment of select parameters. Preliminary experiments were then undertaken to evaluate the model's responsiveness to various parameter alterations. Ultimately, we focused on tuning several key parameters identified through this screening process, as detailed in Supplementary Table S1. For the Cross-Modal Fusion Encoder, we varied the number of layers in the attention module and how different attention layers were connected. Note that all attention modules

shared the same hyperparameters, including the number of attention heads and hidden size. We also experimented with different learning rates and batch sizes, among others.

Comparison methods

To assess the performance of SynergyX, we compared it with six representative state-of-the-art DL methods and two outstanding ML methods, Random Forest (RF) and XGBoost. All methods were trained and evaluated on the same dataset used for SynergyX. The six DL methods mentioned are DeepSynergy [2], MatchMaker [5], DeepDDS [6], DTSyn [12], MGAE-DC [8] and DFFNDDS [13]. For each method, we attempted to preserve consistent input features, model architectures and optimal training parameters mentioned in their respective studies. Specifically, DeepDDS utilizes GCN or GAT for drug feature extraction, and we designated these two models as DeepDDS-GCN and DeepDDS-GAT, respectively. Furthermore, for models initially designed for classification tasks, such as DeepDDS and DFFNDDS, we made minor adjustments to their prediction modules and loss functions. We also optimized their training parameters within a reasonable scope to better suit them for regression tasks. For the RF and XGBoost, a grid search-like method was used to find the optimal training parameters. All considered parameters and their corresponding results are listed in Supplementary Tables S2 and S3. Notably, the cuML package was employed to leverage GPU acceleration to train the RF model.

Model evaluation

SynergyX and all comparison methods were assessed in two experimental settings: (1) General test for rediscovering known drug synergy, and (2) Leave-drug combination-out test for discovering unknown drug pairs [2]. In the General test, the whole dataset was partitioned into train, validation and test sets with an 8:1:1 ratio. For the Leave-drug combination-out test, we employed a stratified sampling approach based on drug pairs, ensuring that the test set did not include any drug pairs present in the train set. The final train, validation and test sets comprised approximately 80%, 10% and 10% of all drug pairs, respectively. All experiments were repeated five times with different random seeds. We used Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), R-squared (R^2) and Spearman correlation (Spearman) as evaluation metrics for the regression prediction task. Additionally, we further evaluated SynergyX on four independent datasets to demonstrate its robust prediction capability across different datasets. The dataset splitting strategy and experimental settings were consistent with those described above.

RESULTS

Model performance

Table 2 and Table 3 present the comparison results between SynergyX and nine other state-of-the-art models in the General Test and Leave-drug combination Test, respectively. Among the evaluated methods, SynergyX demonstrates the best performance across most metrics in both tests. In the General Test, SynergyX achieved MSE, RMSE, MAE, R^2 , Pearson and Spearman of, respectively, 79.55, 8.92, 6.14, 0.85, 0.92 and 0.85, surpassing the second-best model by 2 to 10% across all metrics. Notably, its performance significantly outperforms other Transformer-based models, DFFNDDS and DTSyn. In the Leave-drug combination-out Test, all models exhibit varying degrees of performance decline compared with the General Test, but SynergyX maintains its superior performance (Table 3). When measured by the R^2 metric, SynergyX's performance decreases by approximately 2%.

This indicates that current models have not yet achieved the dual objectives of high predictive accuracy and robust generative capacity simultaneously. On the other hand, we attribute this to the significant gap between labeled data and the prediction space. We further assessed SynergyX's cross-dataset prediction capabilities using four additional benchmark datasets: O'Neil, ALMANAC, Oncology Screen and DrugcombDB. Combining the results from all datasets, SynergyX demonstrates a significant performance improvement ($P < 0.05$) in three out of five datasets compared with the second-best model (Figure 2B, Supplementary Table S4). While both DeepSynergy and MGAE-DC appear to exhibit strong predictive performance across various datasets, it is important to emphasize that these two models are based on MLPs and are unable to provide additional interpretative information beyond the prediction results.

Furthermore, SynergyX demonstrates a relatively balanced predictive performance for the majority of cancer types (11 out of 16) in the test dataset, characterized by an RMSE metric lower than the average and a Pearson correlation coefficient (PCC) greater than 0.75, as illustrated in Figure 2C. For cancer types with poorer prediction outcomes, possible reasons might include insufficient data for specific cancer types or an imbalanced distribution between train and test sets. Additionally, significant variations may exist among different cell lines within the same cancer type.

These results indicate that the architecture of SynergyX is robust and exhibits good generalization. One of its advantages lies in capturing the impact of different drugs on cell lines, which is achieved by the Cross-Modal Fusion Encoder. Using the Uniform Manifold Approximation and Projection (UMAP) method, we visualized the cell embedding of KATO III (a Gastric Cancer cell line) after passing through the Cross-Modal Fusion Encoder. Figure 2C shows the cell embeddings when AZD4527 (an investigational FGFR inhibitor) is combined with 23 other drugs or applied individually to KATO III. Additional benefits can be ascribed to SynergyX's successful integration of omic data and its encoding of interactions among diverse entities, which will be elaborated upon in the following section.

Model ablation study

SynergyX comprises six sub-modules: CNN (the Multi-Omic Integration Module), DCMA, DSSA, CSSA, DDMA and CCMA. Different components of SynergyX were removed to evaluate their contribution to the overall performance. The variant models are labeled SynergyX-B, SynergyX-C, SynergyX-D, SynergyX-E, SynergyX-F and SynergyX-G (Table 4). Five random experiments were conducted for each model to investigate whether the absence of certain modules significantly affects the model's performance. As shown in Table 4, the overall best performance is exhibited by the complete SynergyX model, which demonstrates the importance of each component in contributing to the model.

For SynergyX-B, we removed the CNN module for integrating multi-omic data and employed a strategy similar to the Vision Transformer (ViT) model [43] to encode cell lines (Figure 3). Specifically, we divided the 4079 genes into 50 patches, each containing 81 genes (29 genes were randomly discarded to align the features). The flattened six omic data of the 81 genes formed the feature vector for each patch. All patches were concatenated and served as the initial feature input for the DCMA module. The results show that SynergyX-B exhibits slightly lower performance across most of the metrics.

We removed different Mutual-Attention modules in SynergyX-C, SynergyX-D and SynergyX-E. All models exhibited a noticeable

Table 2: Performance comparison of SynergyX and baseline methods in General Test.

General Test								
Method	MSE	P-value	95% Confidence Interval	RMSE	MAE	R ²	Pearson	Spearman
SynergyX	79.55±0.86		[78.48, 80.63]	8.92±0.07	6.14±0.05	0.85±0.00	0.92±0.00	0.85±0.00
MAGE-DC	81.20±1.00	3.20×10^{-2}	[79.96, 82.44]	9.01±0.06	6.29±0.06	0.84±0.00	0.92±0.00	0.85±0.01
DeepSynergy	88.08±1.30	8.00×10^{-3}	[86.47, 89.69]	9.38±0.07	6.57±0.04	0.83±0.00	0.91±0.00	0.84±0.00
DeepDDS-GCN	109.26±1.17	8.00×10^{-3}	[107.80, 110.71]	10.45±0.06	7.49±0.05	0.79±0.00	0.89±0.00	0.79±0.01
XGBoost	146.73±2.10	8.00×10^{-3}	[144.12, 149.34]	12.11±0.09	8.74±0.06	0.72±0.01	0.85±0.00	0.74±0.01
DFFNDDS	168.19±7.31	8.00×10^{-3}	[159.12, 177.26]	12.97±0.28	9.31±0.19	0.68±0.02	0.90±0.01	0.80±0.02
DeepDDS-GAT	171.41±13.45	8.00×10^{-3}	[154.70, 188.11]	13.08±0.52	9.31±0.38	0.67±0.03	0.82±0.02	0.71±0.02
DTSyn	229.95±17.41	8.00×10^{-3}	[208.34, 251.57]	15.16±0.57	10.91±0.37	0.56±0.03	0.75±0.02	0.61±0.02
Random Forest	222.38±2.48	8.00×10^{-3}	[265.87, 269.74]	14.91±0.08	10.38±0.05	0.57±0.00	0.77±0.00	0.69±0.00
MatchMaker	267.81±1.56	8.00×10^{-3}	[219.30, 225.45]	16.36±0.05	10.93±0.03	0.48±0.00	0.70±0.00	0.61±0.00

For each metric, we report the mean ± standard deviation values. Statistical significance is evaluated using a Wilcoxon rank-sum test, with P<0.05 indicating significance. The optimal performance over all competitive methods is highlighted in bold.

Table 3: Performance comparison of SynergyX and baseline methods in Leave-drug combination-out Test.

Leave-drug combination-out Test								
Method	MSE	P-value	95% Confidence Interval	RMSE	MAE	R ²	Pearson	Spearman
SynergyX	90.43±5.16		[85.54, 94.69]	9.51±0.27	6.59±0.17	0.83±0.01	0.91±0.00	0.84±0.01
MAGE-DC	93.77±4.26	9.50×10^{-2}	[88.48, 99.06]	9.68±0.22	6.83±0.15	0.82±0.01	0.91±0.00	0.82±0.00
DeepSynergy	103.92±6.81	8.00×10^{-3}	[95.47, 112.37]	10.19±0.33	7.22±0.23	0.80±0.01	0.90±0.01	0.81±0.01
DeepDDS-GCN	119.32±5.22	8.00×10^{-3}	[112.84, 125.81]	10.92±0.24	7.85±0.17	0.77±0.01	0.88±0.00	0.77±0.01
XGBoost	169.16±4.36	8.00×10^{-3}	[163.74, 174.58]	13.01±0.17	9.55±0.13	0.68±0.01	0.82±0.00	0.69±0.01
DFFNDDS	178.65±4.56	8.00×10^{-3}	[172.99, 184.32]	13.37±0.17	9.62±0.12	0.66±0.01	0.89±0.00	0.79±0.01
DeepDDS-GAT	181.97±7.71	8.00×10^{-3}	[172.40, 191.54]	13.49±0.29	9.76±0.19	0.65±0.01	0.81±0.01	0.68±0.01
DTSyn	234.22±14.4	8.00×10^{-3}	[216.34, 252.10]	15.30±0.46	11.04±0.33	0.55±0.03	0.75±0.02	0.61±0.02
Random Forest	224.32±6.66	8.00×10^{-3}	[216.05, 232.59]	14.98±0.22	10.63±0.16	0.57±0.01	0.78±0.01	0.68±0.01
MatchMaker	274.39±4.98	8.00×10^{-3}	[268.21, 280.57]	16.56±0.15	11.11±0.16	0.48±0.01	0.70±0.01	0.61±0.01

For each metric, we report the mean ± standard deviation values. Statistical significance is evaluated using a Wilcoxon rank-sum test, with P<0.05 indicating significance. The optimal performance over all competitive methods is highlighted in bold.

decrease in performance, emphasizing the effectiveness of Mutual-Attention mechanisms in providing interactive information and improving prediction performance. Among them, SynergyX-C, which omits the DCMA module, performed the poorest among all tested variants, underscoring the vital role of the DCMA module. It is reasonable to note that removing all mutual attention modules (SynergyX-E) had a relatively more minor impact compared with solely removing the DCMA module (SynergyX-C). While the function of the CCMA module is to fuse the cell representations responsive to drugA and drugB, this step becomes redundant and even introduces unnecessary complexity without the DCMA module. Furthermore, to compare with more simplified variant models, we designed SynergyX-F and SynergyX-G. SynergyX-F retains only the self-attention encoders, while SynergyX-G uses CNN to extract cell line features and self-attention to extract drug features. However, both of them perform worse than SynergyX-E, demonstrating that the composition of SynergyX is not redundant.

Taking into account SynergyX, SynergyX-B, SynergyX-E, SynergyX-F and SynergyX-G, we can observe that regardless of the presence of mutual-attention modules, the convolution-augmented attention architecture—combing CNN and attention—tends to yield superior outcomes compared with using either of them alone (SynergyX versus SynergyX-B and SynergyX-E versus SynergyX-F). This aligns with our hypothesis that CNN features capture local perceptual characteristics, while self-attention can effectively capture long-range semantic information. Combining both can achieve better cell line representations.

In summary, the ablation experiments demonstrate the importance of different components in SynergyX. Notably, all variants still maintain superior performance compared with baseline models. Additionally, the architecture of SynergyX is just one possible configuration, and for different tasks, diverse combinations of self-attention and cross-attention modules can be explored to achieve optimal results.

Muti-omic data study

Two subsequent questions we aim to figure out are: (1) Is using multi-omic data always more effective than using single omic data? (2) Are specific types of omic data more suitable for drug synergy prediction? Therefore, we explored the impact of different combinations of omic data. A total of six types of omic were initially applied in SynergyX. To simplify the experiments, we randomly selected six combinations for the cases involving 2, 3, 4 and 5 kinds of omic data. Figure 4(A, B) and Supplementary Table S5 display the different omic combinations and experimental results.

In single-omic experiments, model performance varies across different omic types. The cn data exhibit the lowest MSE value of 82.17, while exp, mut and met perform relatively well. However, using eff or dep leads to significantly poorer results. This is because we lacked sufficient eff and dep features, which could serve as supplementary features but are supposed to be ineffective under single-omic experiments. Moreover, we observed that SynergyX consistently outperforms other models when using only exp, mut, cn or met data, underscoring the superiority of the SynergyX architecture. As the number of omic data types

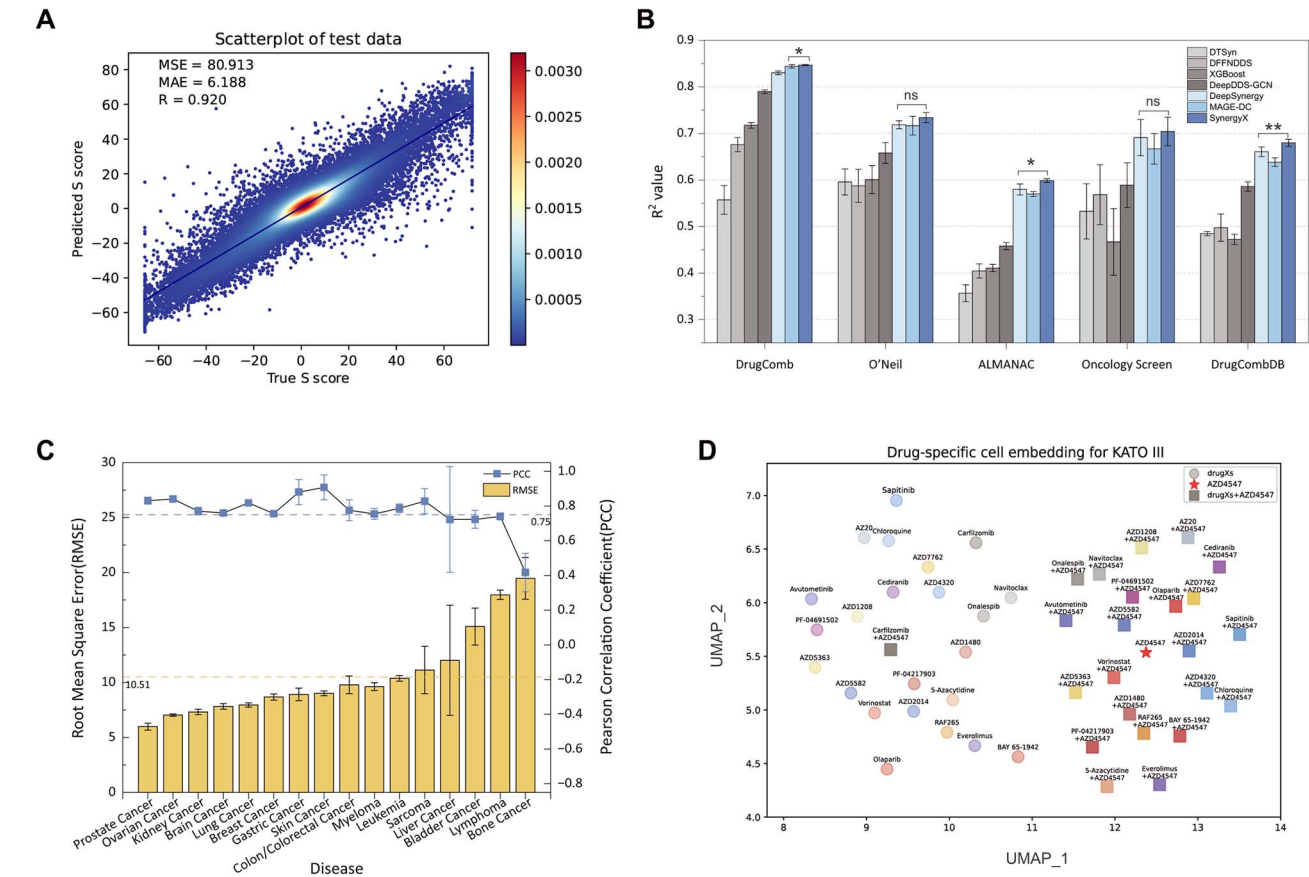


Figure 2. Performance of SynergyX. (A) Scatter density plot shows a significant positive correlation between Predicted and True S score in the test dataset. The density of the dots is color-coded, with lighter colors indicating areas of higher density. (B) Performance of SynergyX across five datasets compared with other models. Error bars represent the standard deviation from five random experiments. Significance calculated using the Wilcoxon rank-sum test is denoted as follows: *, **, *** for $P < 0.05$, $P < 0.01$, $P < 0.001$, respectively; 'ns' denotes no significance. (C) SynergyX's Predictive Results for different Cancer Types. The bar graph displays the mean \pm standard deviation (SD) for the RMSE metric for each cancer type, while the line graph shows the PCC metric, both calculated from five random experiments. The blue dashed line represents a PCC threshold of 0.75, and the yellow dashed line indicates the average RMSE across all cancer types. (D) Visualization of drug-specific cell embedding of KATO III cell line by UMAP. Specific drugs or drug pairs are labeled near each point.

Table 4: Results of model ablation experiment.

Model	Description	MSE	Confidence Interval	P-value	RMSE	MAE	R ²
SynergyX	—	79.555\pm0.865	[78.48, 80.63]		8.919\pm0.048	6.138\pm0.035	0.847\pm0.001
SynergyX-B	without CNN	81.612 \pm 1.526	[79.72, 83.51]	3.20×10^{-2}	9.012 \pm 0.08	6.214 \pm 0.098	0.843 \pm 0.004
SynergyX-C	without DCMA	85.715 \pm 1.797	[83.48, 87.95]	8.00×10^{-3}	9.258 \pm 0.097	6.469 \pm 0.069	0.835 \pm 0.004
SynergyX-D	without DDMA,CCMA	81.248 \pm 1.007	[80.00, 82.50]	3.60×10^{-2}	9.014 \pm 0.056	6.209 \pm 0.042	0.843 \pm 0.001
SynergyX-E	without DCMA,DDMA,CCMA	81.565 \pm 0.822	[80.54, 82.59]	3.60×10^{-2}	9.031 \pm 0.045	6.271 \pm 0.051	0.843 \pm 0.001
SynergyX-F	without CNN,DCMA,DDMA,CCMA	82.554 \pm 1.02	[81.29, 83.82]	8.00×10^{-3}	9.086 \pm 0.056	6.331 \pm 0.055	0.841 \pm 0.003
SynergyX-G	without CSSA,DCMA,DDMA,CCMA	82.876 \pm 1.896	[80.52, 85.23]	1.60×10^{-2}	9.103 \pm 0.104	6.353 \pm 0.094	0.84 \pm 0.003

The bold values represent the optimal performance over all models. Statistical significance was evaluated using a Wilcoxon rank-sum test, with $P < 0.05$ indicating significance.

increased, the model's predictive capacity improved to a certain degree. The optimal outcomes were obtained when all six omic data were employed.

Interestingly, SynergyX appears to be insensitive to the specific omic types. As depicted in (Figure 4A, B), when the number of omic types exceeds 2, using any omic type appears to yield similar predictions. One possible explanation is that different omic types provide complementary rather than independent information. While individual omic datasets provide limited information, including two or three omic types is adequate to obtain comprehensive features and remain cost-effective. Specifically, the combination

of exp, mut and cn performs similarly to using all six omic and surpasses most combinations involving more omic.

Comparison of different gene set

We selected three gene sets of varying sizes to observe the impact of different gene sets on drug combination prediction (Figure 4C). The smallest gene set, Geneset 1, consists of 963 genes obtained from the landmark gene published by the LINCS project [32] after feature selection. Geneset 2 is the Unified gene set used in SynergyX, containing 4079 genes. For Geneset 3, we selected all genes from the CCLE dataset that possess exp, mut and cn

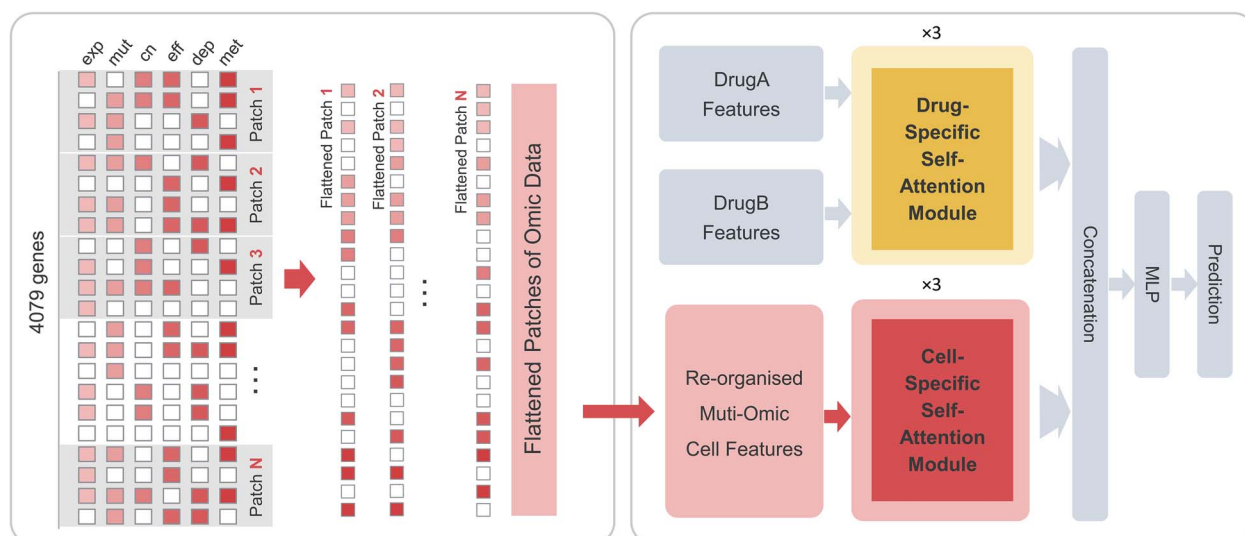


Figure 3. The framework of SynergyX-F. SynergyX-F is a simplified variant of SynergyX, consisting solely of three layers Drug-specific Self-Attention (DSSA) Module and Cell-specific Self-Attention (CSSA) Module. With the CNN module removed, we divided the feature genes of each cell line into 50 patches, each containing 81 genes. The flattened omic data formed the feature vector for each patch.

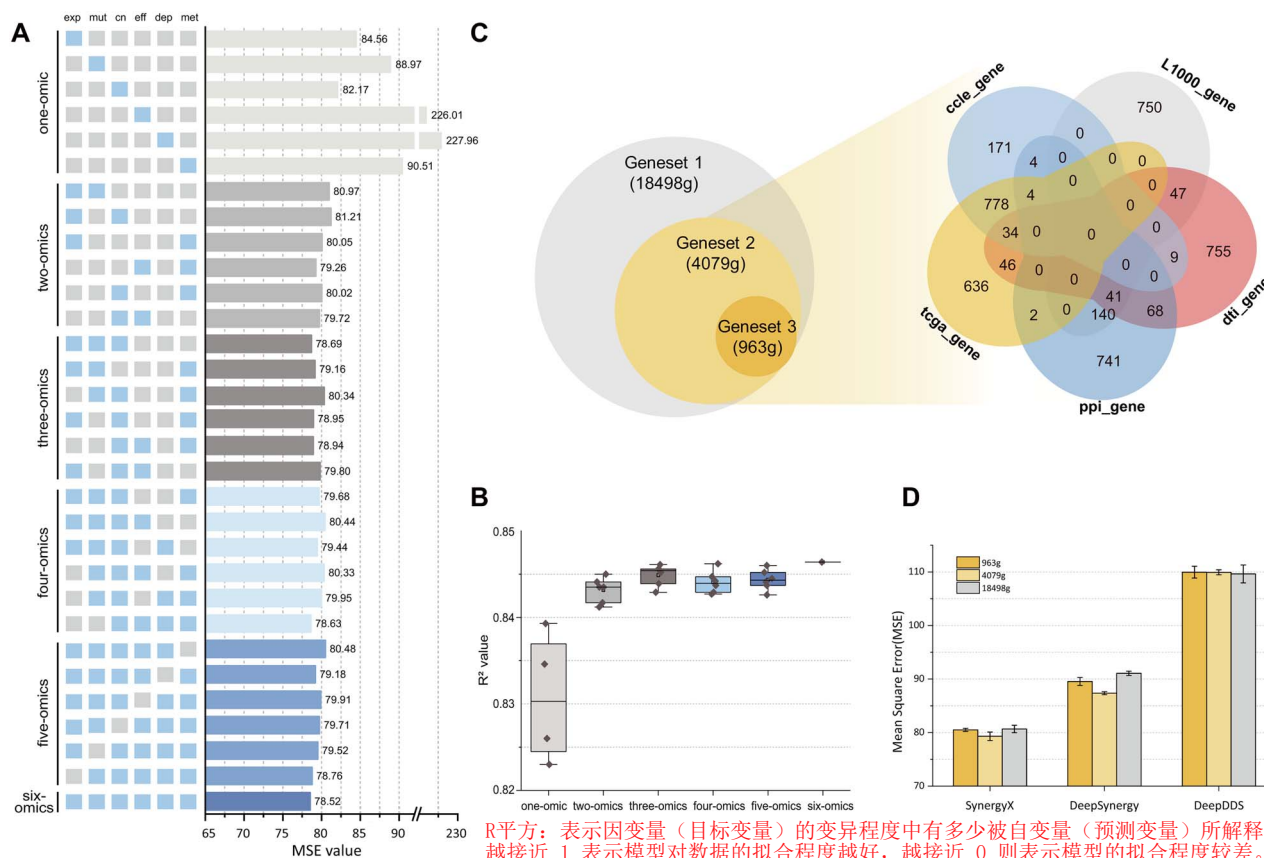


Figure 4. Results of multi-omic study and gene set comparison. (A) MSE values for experiments of different combinations of omic data. The light blue squares represent the included omic types, while the gray squares indicate absence. (B) Boxplot showing R^2 values grouped by the number of omic types. We excluded the two worst-performing one-omic experiments for clarity. (C) Three gene sets for the comparison experiment. Geneset 2 (Right) is a union of five gene sets. The numbers indicate gene count in each set. The Venn diagram on the right was generated using omicshare. (D) Grouped bar chart displaying the MSE values of the SynergyX, DeepSynergy and DeepDDS-GCN models for different gene sets in the DrugComb dataset. The error bars represent the standard deviation from five random experiments.

features, totaling 18 498 genes, representing the entire genome. We conducted comparative experiments with different gene sets in the DrugComb dataset. Specifically, we utilized these three gene

sets as input features for cells to train SynergyX, DeepSynergy and DeepDDS-GCN, thereby assessing the sensitivity of different models to different gene sets. For each gene set, we adjusted the

Table 5: Top 10 predicted synergistic drug combinations in NCI-H460 cell line.

DrugA	DrugB	Cell line	Predicted score	Supporting publications
Bortezomib	Vinblastine	NCI-H460	43.83	NA
Bortezomib	Palbociclib	NCI-H460	41.03	PMID:31120355
Actinomycin	Toremifene	NCI-H460	39.40	NA
Erlotinib	Vinblastine	NCI-H460	38.86	PMID:31262905
Vincristine	Bortezomib	NCI-H460	37.22	PMID:20582937, PMID:25738670
Lomustine	Regorafenib	NCI-H460	35.61	NA
Vorinostat	Vinblastine	NCI-H460	35.42	PMID:26156322
Bortezomib	Temozolomide	NCI-H460	35.05	PMID:34689859, PMID:29722661, PMID:21850464
Bortezomib	Erlotinib	NCI-H460	35.00	PMID:33804833, PMID:19633475
Temozolomide	Vinorelbine	NCI-H460	33.99	PMID:22890892, PMID:17987262

model’s training parameters accordingly. Importantly, in the gene set comparison experiments, we employed exp, mut and cn as the cell line inputs for SynergyX.

Using MSE for evaluation, we discovered that both SynergyX and DeepSynergy performed best with Geneset 2. In contrast, DeepDDS-GCN showed no significant performance variation across different gene sets, as shown in Figure 4D. Larger gene sets offer a more comprehensive view of the intracellular gene network but can introduce redundant features and increase computational complexity. Therefore, it is advisable to choose genes based on specific algorithm goals. For instance, the LINCS’s landmark gene set may suffice for representative cell line features. However, considering different gene sets becomes essential when exploring specific latent biological information. Our approach of seeking suitable gene sets based on biological knowledge proves to be a reliable method.

Novel drug synergy prediction

Subsequently, we applied SynergyX to predict new drug synergies on three cancer cell lines: NCI-H460 (Lung Cancer), T-47D (Breast Cancer) and HCT116 (Colon Cancer). We selected 85 clinically approved drugs, resulting in 7140 potential drug combinations. For each cell line, we removed existing drug pairs from the dataset. Table 5 displays the top 10 drug pairs with the highest predicted synergy scores for NCI-H460. To validate these predictions, we conducted literature searches and found that over half of the results had supporting evidence, confirming SynergyX’s reliability in predicting new drug synergies.

Specifically, among the Top 10 cases in Table 5, five involve combinations with Bortezomib and other drugs. Bortezomib, a proteasome inhibitor, has demonstrated anti-lung cancer activity by affecting various signaling pathways [44]. The combination of Bortezomib and Vinblastine (VBL) had the highest predicted synergy score. Prior research has shown a synergistic effect of proteasome inhibitors and VBL in lung cancer cell lines [45]. Given similar mechanisms, Bortezomib and VBL may represent a potential drug synergy. Bortezomib and the CDK 4/6 inhibitor Palbociclib were also predicted to have a strong synergistic effect, supported by a prior clinical phase I trial [46]. Furthermore, we found supporting evidence for the combinations of Bortezomib with Vincristine [47, 48], Temozolomide [49–51] and Erlotinib [52, 53].

The above evidence indicates that SynergyX has identified promising drug combination candidates for potential lung cancer treatment. We emphasize that combinations with high predicted scores yet not reported warrant further validation of their potential activity. Additionally, the top drug synergy predictions for

the other two cell lines are provided in Supplementary Tables S6 and S7.

Interpretability of SynergyX

The Cross-Modal Fusion Encoder of SynergyX is able to extract cross-modal attention information between different entities. The DCMA module primarily focuses on substructure-gene interactions between individual drugs and specific cell lines, while the two mutual-attention modules output information at both the chemical and omic levels that closely align with the prediction results.

One common strategy for anti-tumor combination therapy involves the combination of chemotherapy drugs and targeted drugs [54]. Chemotherapy drugs have a broad cell-killing effect, while targeted therapies specifically inhibit molecules that are activated in tumor cells. For instance, the synergistic effects of combining Gemcitabine [55] and Dasatinib [56] have been extensively demonstrated [57, 58]. Gemcitabine is a widely used chemotherapy drug, while Dasatinib is an effective dual Abl/Src kinase inhibitor that is used in chronic myeloid leukemia and other solid tumors [56].

Another strategy involves dual blockade of oncogenic signals, where two drugs target different signaling pathways or upstream and downstream molecules of the same signaling pathway [54]. Vorinostat is the first FDA-approved histone deacetylase (HDAC) inhibitor [59]. Sorafenib is a multi-kinase inhibitor that simultaneously targets tumor cell growth and tumor angiogenesis [60]. HDAC, being a proangiogenic factor, directly regulates the expression of vascular endothelial growth factor (VEGF). The dual inhibition of HDAC and VEGFR has been considered as a promising anticancer strategy [61, 62]. Several studies have already explored the combination of Sorafenib and Vorinostat [63, 64].

DCMA Module Reveals Drug-Cell Interactions

We first selected two synergistic cases, Gemcitabine and Dasatinib in A549 (Lung Cancer) and Vorinostat and Sorafenib in CAKI-1 (Renal Cancer), to explore the interpretability information provided by the DCMA module (Figure 5, Supplementary Figure S1). We focused on the top gene-substructure attention scores to investigate the most emphasized features of the model.

Based on observations of gene-substructure attention scores across different ranges, such as top 10%, top 20% and top 30%, among others, our initial findings revealed a strong correlation between the drug substructures highlighted by DCMA and the known pharmacophores of the drugs. For instance, Vorinostat exhibits greater attention toward its hydroxamic acid group and a chain of methylene groups (Figure 5A, Supplementary Figure S3).

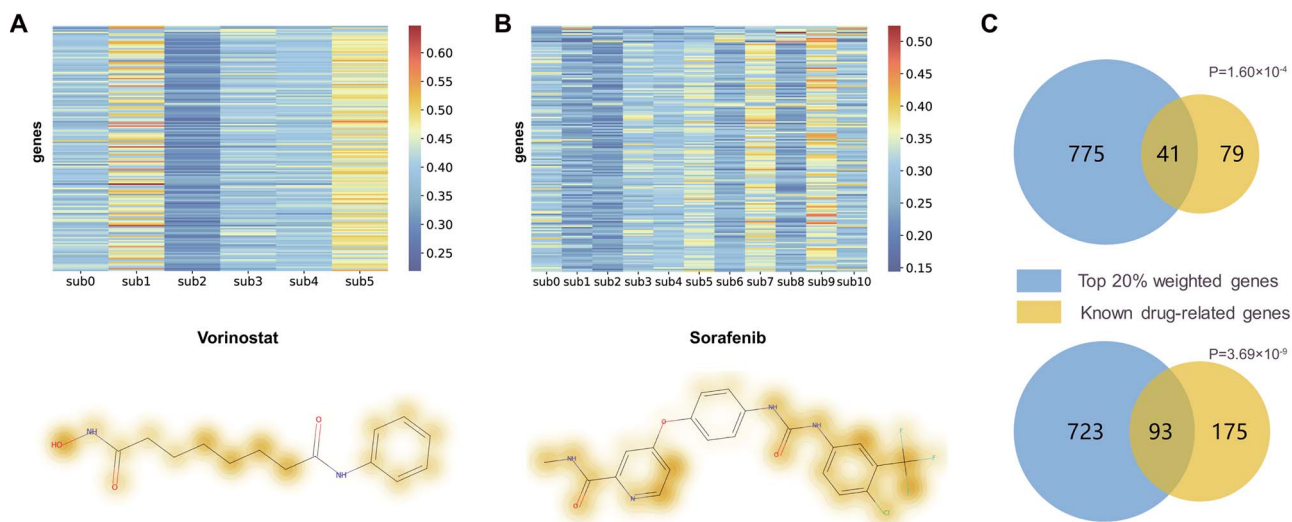


Figure 5. Interpretability of the DCMA module for the synergy of Vorinostat and Sorafenib in CAKI-1 cell line. (A) and (B) Heatmaps showing the Cell-to-Drug attention scores output by the DCMA module and the visualization of the average top 30% attention on each substructure. The deeper colors indicate higher importance. (A) represents Vorinostat, and (B) represents Sorafenib. (C) Venn diagram displaying the intersection of the top 20% genes with the highest Drug-to-Cell attention (in yellow) and the Known drug-related genes in the STITCH database [38] (in blue), separately for Vorinostat (top) and Sorafenib (bottom). P-value is evaluated by Hypergeometric test, with $P < 0.05$ indicating significance.

This aligns with previous reports, where the hydroxamic acid group with intensive polarity facilitates efficient chelation of the zinc ion within the active center of HDAC, and the number of methylene groups directly impacts its activity [65]. The DCMA module also revealed three critical pharmacophores of sorafenib (Figure 5B, Supplementary Figure S4), which is consistent with previous reports [66].

Next, we collected drug-protein interactions from the STITCH database [38] to analyze the critical genes captured by the DCMA module. The coarse-grained cell line attention scores from the DCMA were linearly mapped to the 4079 input genes. We observed a significant overlap between high-weighted genes and known drug-related genes. For instance, we identified 176 DTIs of Gemcitabine from the STITCH dataset, and 84 of these genes (47.7%, $P=1.60 \times 10^{-4}$) were among the top 20% of genes ranked by attention scores (Figure 5C). Similarly, for Dasatinib, Vorinostat and Sorafenib, the intersection of genes was 63/212 (29.7%, $P=3.69 \times 10^{-9}$), 41/120 (34.1%, $P=2.30 \times 10^{-17}$) and 93/268 (34.7%, $P=3.44 \times 10^{-4}$), respectively (Figure 5C, Supplementary Figure S1C). It is reasonable to speculate that non-overlapping genes may serve as promising DTIs to be explored in future research. Further, we extended our analysis to encompass all drugs in the DrugComb dataset to examine the overlap between top-ranked genes and known drug targeted genes. After excluding drugs whose DTIs were not among the 4079 genes in the STITCH database, 299 drugs remained for further analysis. As a result, we found that 186 out of these 299 drugs (62.2%, $P < 0.05$) significantly enriched an average of 42.6% of the DTI genes (Supplementary Figure S5).

These findings highlight the remarkable ability of the DCMA module to capture essential drug pharmacophores while simultaneously enriching drug-specific target genes. It enables the extraction and transmission of interaction information into the drug synergy prediction process and offers an approach to uncover molecular-level mechanisms explicitly.

CCMA Module Infers Cell Selectivity of Drug Combinations

Considerable variations in drug synergy effects are observed across different cell lines when exposed to the same drug pair

(Figure 6A). We attempted to identify the molecular-level factors influencing cell line selectivity using the output of the CCMA module. For each drug pair, we selected cell lines with synergy scores greater than 10 as synergistic cell lines and cell lines with scores below -10 as antagonistic cell lines, ensuring an equal number of both types of cell lines. We extracted the gene attention matrix from the CCMA and then averaged the top 20% of weighted genes for synergistic and antagonistic cell lines. Subsequently, pathway enrichment analysis was performed. Taking Sorafenib and Vorinostat as an example (Figure 6B), in the synergistic cell lines, there is a notable enrichment of pathways related to cell proliferation, such as the G2-M checkpoint pathway, E2F Targets pathway and MYC Targets V1 pathway (Figure 6C). Both Sorafenib and Vorinostat are known to exert their anticancer effects by inhibiting tumor cell proliferation and growth. Besides, Oishi *et al* [67] found a positive correlation between G2M checkpoint score and the area under the curve (AUC) of the Sorafenib single-agent sensitivity curve.

We mainly focused on the differences in enriched pathways between antagonistic and synergistic cell lines, which could indicate the mechanisms of drug resistance in specific cellular contexts (Figure 6C, D). Encouragingly, we observed that some of these differential pathways align with previously reported resistance mechanisms. For instance, alterations in the DNA repair machinery and the downregulation of key factors in the IFN signaling pathway have been identified as resistance mechanisms to Sorafenib [68–70]. In addition, the activation of the IL-6/JAK/STAT3 pathway has been reported as a resistance mechanism to both Sorafenib and HDAC inhibitors [71, 72]. Remarkably, the pathways mentioned above were enriched in the top-weighted genes of antagonistic cell lines rather than in the synergistic cell lines.

The volcano plot in Figure 6E displays genes with significant changes in attention scores between antagonistic and synergistic cell lines. We observed that Rac GTPase Activating Protein 1 (RACGAP1) garnered notable attention. RACGAP1 is a cytokinesis-regulatory protein that has been reported to serve as an oncogenic enhancer and predictive biomarker for poor prognosis in various malignancies [73–75]. Our findings suggest that the RACGAP1

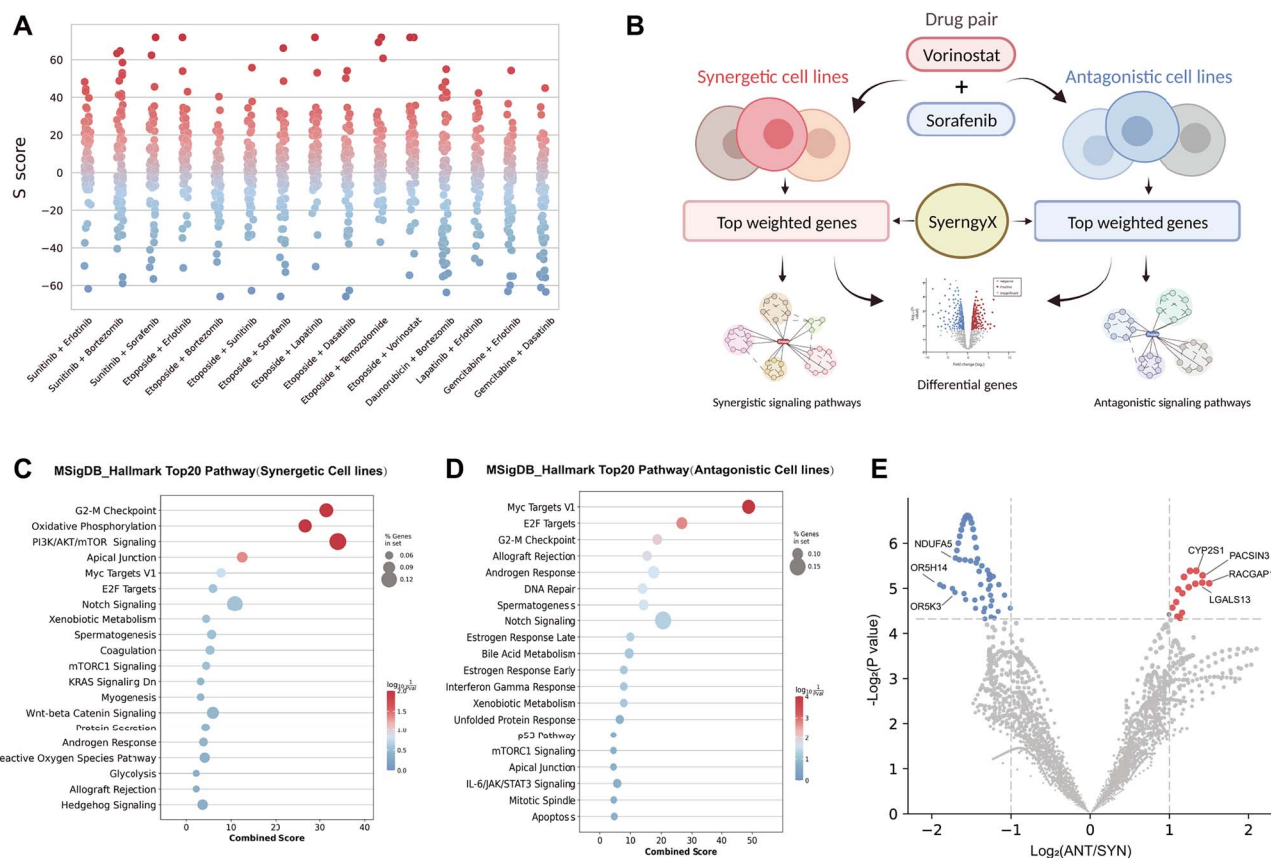


Figure 6. Investigating cross-cell line synergy selectivity for drug combinations using SynergyX. (A) Grouped scatter plot showing the distribution of S scores across different cell lines for various drug combinations. The color gradient from red to blue represents the transition from synergy to antagonism. (B) Overview of using SynergyX to analyze the cross-cell line synergy selectivity of Vorinostat and Sorafenib. Created by BioRender.com. (C) and (D) Pathway enrichment plots of the top 20% genes with the highest attention score output by the CCMA module, separately for synergistic cell lines (C) and antagonistic cell lines (D). (E) Volcano plot presenting genes with significant changes in attention scores between antagonistic and synergistic cell lines. Red dots represent genes with significantly increased attention scores, while blue dots represent genes with significantly decreased attention scores. P-values are evaluated using an independent sample t-test.

gene may potentially serve as a biomarker for resistance to the combination of Sorafenib and Vorinostat.

Regarding the combination of Gemcitabine and Dasatinib, we found specific enrichment of the oxidative phosphorylation (OXPHOS) pathway and the Wnt/ β -catenin Signaling pathway (Supplementary Figure S2). Recent studies have indicated that OXPHOS can drive tumor resistance [76, 77]. Our findings may suggest that Gemcitabine and Dasatinib are more inclined toward antagonism in tumors with high OXPHOS activity. Additionally, the activation of the Wnt/ β -catenin pathway has been reported as one of the mechanisms of resistance to Gemcitabine in pancreatic cancer [78].

Overall, our findings underscore the effectiveness of SynergyX in accurately simulating and revealing the diverse effects of drug combinations within distinct cellular contexts, providing valuable insights into drug synergy and resistance mechanisms.

DISCUSSION

The development of accurate drug synergy prediction models becomes imperative. Drug combination is a complex biological process involving numerous unknown and intricate interactions [54]. Previous studies have emphasized the significance of incorporating biological knowledge like DTI and drug-induced gene expression profiles [6, 7, 79], but limited prior biological knowledge has restricted the development of broader, more applicable

models. In this study, we made a pioneering attempt to introduce the mutual attention mechanism to encode cross-modal interactions between drugs and cells, thereby embedding interaction knowledge *in silico*. Furthermore, our findings indicate the importance of designing more specialized multi-modal fusion modules rather than solely focusing on single-feature extraction steps. This discovery aligns with conclusions drawn in text-image multi-modal research [16]. Additionally, by adopting a convolution-augmented attention structure, we devised a novel multi-omic feature extraction method capable of simultaneously capturing local features and context awareness, enabling the comprehensive encoding of intracellular gene networks.

The results of extensive evaluations comparing SynergyX with eight other state-of-the-art models showcase its superior performance in predicting drug synergy across different datasets. When further applied to predict novel drug synergy, SynergyX yields predictions that align with previously reported experimental results. One noteworthy aspect of SynergyX is its ability to handle omic data effectively. Our findings indicate that the quantity of omic data, rather than the specific omic category, exerts a larger influence. Using two or three types of omic data is sufficient to obtain stable results. Another key strength of SynergyX lies in its interpretability. The attention scores generated by the multiple mutual-attention modules offer valuable insights into the intricate dynamics of drug interactions in diverse cellular contexts.

One area for improvement lies in the inability to capture attention at the single-gene resolution in cell lines due to memory constraints, which reduces the precision of model interpretability. Moreover, it is essential to recognize the constraints of SynergyX concerning its generative capabilities. Existing models may struggle to handle highly heterogeneous primary tumor samples or reflect the complex intracorporal process of drugs. For instance, consistent with preclinical experimental results, we successfully predicted a strong synergistic effect between Bortezomib and Erlotinib in NCI-H460 cell line (Table 5). However, a phase 2 study of erlotinib in combination with bortezomib in non-small cell lung cancer was discontinued due to insufficient activity [53]. To address these limitations, one promising approach is to utilize large-scale single-cell pre-trained models [80], which effectively capture the complex gene connections and offer a universally applicable cell representation, thus bridging the gap between cancer cells and clinical samples. Furthermore, incorporating strategies like transfer learning to boost the model's generalization capacity or adopting few-shot learning and active learning techniques to extend the prediction capabilities based on limited experimental data can be advantageous. When combined with these strategies, SynergyX may hold the potential to address clinical challenges. Furthermore, developing disease-specific models that are better aligned with clinical needs and unique therapeutic mechanisms will be a key direction. Customized models that consider disease-specific contexts can lead to more accurate and clinically relevant predictions.

In summary, SynergyX provides a comprehensive framework for drug synergy prediction, achieving enhanced predictive performance and interpretability. The capacity of SynergyX to decipher the intricate landscape of drug interactions and identify context-specific responses may lead to more effective discovery of drug synergy and advance toward rational combination therapy.

Key Points

- SynergyX presents an innovative and scalable architecture, introducing the mutual attention mechanism to capture cross-modal interactions and mimic complex biological responses, thereby enhancing drug synergy prediction.
- We highlight that the convolution-augmented attention structure is a better approach to integrate multi-omic data and characterize cellular gene networks for specific cell lines.
- SynergyX offers interpretability in deciphering intricate drug-cell interactions and context-specific responses of drug combinations, advancing rational combination therapy.

ACKNOWLEDGEMENTS

We would like to express our gratitude to Dr. Ruiwei Feng, Jian Gao and Qing Ye for their valuable advice and assistance provided to the project.

AUTHOR CONTRIBUTIONS STATEMENT

Y.G. conceived the research idea, designed the algorithms, conducted experiments and drafted the manuscript. H.H. performed some experiments and revised the initial draft. W.C. and H.Y.

were responsible for data collection and preprocessing. J.W., Q.H. and J.C. provided supervision and acquired funding. J.C. and C.H. contributed to reviewing and editing the manuscript.

FUNDING

This work was supported by the Key Projects of Hangzhou Agricultural and Social Development Program [20200416A36 to Q.H.]; National Natural Science Foundation of China [81872885 to J.C.]; Zhejiang Provincial Natural Science Foundation [LR22H310002 to J.C.]; the Fundamental Research Funds for the Central Universities [226-2023-00059] and Zhejiang Key R&D Program of China [2023C03053].

DATA AVAILABILITY

All the data used in the article were obtained from online databases. Drug combination data were sourced from DrugComb (v1.5) [18]. Genomic data were downloaded from the DepMap database [40]. Drugs' chemical information was obtained from the ChEMBL database [28]. The source code of SynergyX and processed data used in this study can be accessed at <https://doi.org/10.5281/zenodo.8289295> and <https://github.com/GSanShui/SynergyX>.

REFERENCES

1. Rees MG, Brenan L, Bajrami B, et al. Systematic identification of biomarker-driven drug combinations to overcome resistance. *Nat Chem Biol* 2022;**18**(6):615–24.
2. Preuer K, Lewis RPI, Hochreiter S, et al. DeepSynergy: predicting anti-cancer drug synergy with deep learning. *Bioinformatics* 2018;**34**(9):1538–46.
3. Zagidullin B, Aldahdooh J, Zheng S, et al. DrugComb: an integrative cancer drug combination data portal. *Nucleic Acids Res* 2019;**47**(W1):W43–51.
4. Wu L, Wen Y, Leng D, et al. Machine learning methods, databases and tools for drug combination prediction. *Brief Bioinform* 2022;**23**(1):bbab355.
5. Kuru HI, Tastan O, Cicek AE. Matchmaker: a deep learning framework for drug synergy prediction. *IEEE/ACM Trans Comput Biol Bioinform* 2021;**19**(4):2334–44.
6. Wang J, Liu X, Shen S, et al. DeepDDS: deep graph neural network with attention mechanism to predict synergistic drug combinations. *Brief Bioinform* 2022;**23**(1):bbab390.
7. Lin W, Wu L, Zhang Y, et al. An enhanced cascade-based deep forest model for drug combination prediction. *Brief Bioinform* 2022;**23**(2):bbab562.
8. Zhang P, Tu S. Mgae-dc: predicting the synergistic effects of drug combinations through multi-channel graph autoencoders. *PLoS Comput Biol* 2023;**19**(3):e1010951.
9. Jiang P, Huang S, Fu Z, et al. Deep graph embedding for prioritizing synergistic anticancer drug combinations. *Comput Struct Biotechnol J* 2020;**18**:427–38.
10. Yang J, Xu Z, Wu WKK, et al. Graphsynergy: a network-inspired deep learning model for anticancer drug combination prediction. *J Am Med Inform Assoc* 2021;**28**(11):2336–45.
11. Yu L, Xia M, An Q. A network embedding framework based on integrating multiplex network for drug combination prediction. *Brief Bioinform* 2022;**23**(1):bbab364.
12. Hu J, Gao J, Fang X, et al. Dtsyn: a dual-transformer-based neural network to predict synergistic drug combinations. *Brief Bioinform* 2022;**23**(5):bbac302.

13. Xu M, Zhao X, Wang J, et al. Dfndds: prediction of synergistic drug combinations with dual feature fusion networks. *J Chem* 2023;**15**(1):1–12.
14. Rafiei F, Zeraati H, Abbasi K, et al. Deeptrasynergy: drug combinations using multimodal deep learning with transformers. *Bioinformatics* 2023;**39**(10):btad607.
15. Gan Y, Huang X, Guo W, et al. Predicting synergistic anticancer drug combination based on low-rank global attention mechanism and bilinear predictor. *Bioinformatics* 2023;**39**(10):btad607.
16. Kim W, Son B, Kim I. Vilt: Vision-and-language transformer without convolution or region supervision. In: Meila M, Zhang T (eds). *International Conference on Machine Learning*. New York: PMLR, 2021, 5583–94.
17. Zheng X, Wang B, Du X, et al. Mutual attention inception network for remote sensing visual question answering. *IEEE Trans Geosci Remote Sens* 2021;**60**:1–14.
18. Zheng S, Aldahdooh J, Shadbahr T, et al. Drugcomb update: a more comprehensive drug sensitivity data repository and analysis portal. *Nucleic Acids Res* 2021;**49**(W1):W174–84.
19. Malyutina A, Majumder MM, Wang W, et al. Drug combination sensitivity scoring facilitates the discovery of synergistic and efficacious drug combinations in cancer. *PLoS Comput Biol* 2019;**15**(5):e1006752.
20. Borisy AA, Elliott PJ, Hurst NW, et al. Systematic discovery of multicomponent therapeutics. *Proc Natl Acad Sci* 2003;**100**(13):7977–82.
21. Bliss CI. The toxicity of poisons applied jointly 1. *Annals of applied biology* 1939;**26**(3):585–615.
22. Loewe S. The problem of synergism and antagonism of combined drugs. *Arzneimittelforschung* 1953;**3**:285–90.
23. Yadav B, Wennerberg K, Aittokallio T, Tang J. Searching for drug synergy in complex dose–response landscapes using an interaction potency model. *Comput Struct Biotechnol J* 2015;**13**:504–13.
24. Pukelsheim F. The three sigma rule. *The American Statistician* 1994;**48**(2):88–91.
25. O'Neil J, Benita Y, Feldman I, et al. An unbiased oncology compound screen to identify novel combination strategies. *Mol Cancer Ther* 2016;**15**(6):1155–62.
26. Holbeck SL, Camalier R, Crowell JA, et al. The national cancer institute almanac: a comprehensive screening resource for the detection of anticancer drug pairs with enhanced therapeutic activity. *NCI Thesaurus* 2017;**77**(13):3564–76.
27. Liu H, Zhang W, Zou B, et al. Drugcombd: a comprehensive database of drug combinations toward the discovery of combinatorial therapy. *Nucleic Acids Res* 2020;**48**(D1):D871–81.
28. Gaulton A, Bellis LJ, Bento AP, et al. ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res* 2012;**40**(D1):D1100–7.
29. Landrum G, et al. Rdkit: a software suite for cheminformatics, computational chemistry, and predictive modeling. *Greg Landrum* 2013;**8**:31.
30. Huang K, Xiao C, Glass L, et al. Explainable substructure partition fingerprint for protein, drug, and more. In: *NeurIPS Learning Meaningful Representation of Life Workshop*, Vancouver, BC, Canada, 2019.
31. Jiang L, Jiang C, Yu X, et al. DeepTTA: a transformer-based model for predicting cancer drug response. *Brief Bioinform* 2022;**23**(3):Bbac100.
32. Subramanian A, Narayan R, Corsello SM, et al. A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. *Cell* 2017;**171**(6):1437–1452.e17.
33. Barretina J, Caponigro G, Stransky N, et al. The cancer cell line encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* 2012;**483**(7391):603–7.
34. Ghandi M, Huang FW, Jané-Valbuena J, et al. Next-generation characterization of the cancer cell line encyclopedia. *Nature* 2019;**569**(7757):503–8.
35. Hutter C, Zenklusen JC. The cancer genome atlas: creating lasting value beyond its data. *Cell* 2018;**173**(2):283–5.
36. Szklarczyk D, Gable AL, Nastou KC, et al. The string database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res* 2021;**49**(D1):D605–12.
37. Szklarczyk D, Kirsch R, Koutrouli M, et al. The string database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res* 2023;**51**(D1):D638–46.
38. Szklarczyk D, Santos A, Von Mering C, et al. Stitch 5: augmenting protein–chemical interaction networks with tissue and affinity data. *Nucleic Acids Res* 2016;**44**(D1):D380–4.
39. Tsherniak A, Vazquez F, Montgomery PG, et al. Defining a cancer dependency map. *Cell* 2017;**170**(3):564–576.e16.
40. Meyers RM, Bryan JG, McFarland JM, et al. Computational correction of copy number effect improves specificity of crispr-cas9 essentiality screens in cancer cells. *Nat Genet* 2017;**49**(12):1779–84.
41. Gulati A, Qin J, Chiu CC, et al. Conformer: convolution-augmented transformer for speech recognition. *arXiv preprint arXiv:2005.08100*. 2020.
42. Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. *Advances in neural information processing systems* 2017;**30**:1–11.
43. Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16x16 words: transformers for image recognition. In: *Proceedings of the 9th International Conference on Learning Representations (ICLR)*, Virtual Event, Austria. 2021. OpenReview.net.
44. Davies AM, Lara PN, Jr, Mack PC, Gandara DR. Incorporating bortezomib into the treatment of lung cancer. *Clin Cancer Res* 2007;**13**(15):4647s–51s.
45. Sooman L, Gullbo J, Bergqvist M, et al. Synergistic effects of combining proteasome inhibitors with chemotherapeutic drugs in lung cancer cells. *BMC Res Notes* 2017;**10**(1):1–9.
46. Martin P, DiLiberto M, Mason CE, et al. The combination of palbociclib plus bortezomib is safe and active in patients with previously treated mantle cell lymphoma: final results of a phase i trial. *Blood* 2013;**122**(21):4393.
47. Messinger Y, Gaynon P, Raetz E, et al. Phase i study of bortezomib combined with chemotherapy in children with relapsed childhood acute lymphoblastic leukemia (all): a report from the therapeutic advances in childhood leukemia (tacl) consortium. *Pediatr Blood Cancer* 2010;**55**(2):254–9.
48. Robak T, Huang H, Jin J, et al. Bortezomib-based therapy for newly diagnosed mantle-cell lymphoma. *New England Journal of Medicine* 2015;**372**(10):944–53.
49. Pak O, Zaitsev S, Shevchenko V, et al. Effectiveness of bortezomib and temozolomide for eradication of recurrent human glioblastoma cells, resistant to radiation. *Prog Brain Res* 2021;**266**:195–209.
50. Kong XT, Nguyen NT, Choi YJ, et al. Phase 2 study of bortezomib combined with temozolomide and regional radiation therapy for upfront treatment of patients with newly diagnosed glioblastoma multiforme: safety and efficacy assessment. *International journal of radiation oncology* biology**. *Phys Ther* 2018;**100**(5):1195–203.

51. Portnow J, Frankel P, Koehler S, et al. A phase i study of bortezomib and temozolomide in patients with advanced solid tumors. *Cancer Chemother Pharmacol* 2012;**69**:505–14.
52. Saafan H, Alahdab A, Michelet R, et al. Constitutive cell proliferation regulating inhibitor of protein phosphatase 2a (cip2a) mediates drug resistance to erlotinib in an egfr activating mutated nsccl cell line. *Cell* 2021;**10**(4):716.
53. Lynch TJ, Fenton D, Hirsh V, et al. A randomized phase 2 study of erlotinib alone and in combination with bortezomib in previously treated advanced non-small cell lung cancer. *J Thorac Oncol* 2009;**4**(8):1002–9.
54. Jin H, Wang L, Bernards R. Rational combinations of targeted cancer therapies: background, advances and challenges. *Nat Rev Drug Discov* 2023;**22**(3):213–34.
55. Toschi L, Finocchiaro G, Bartolini S, et al. Role of gemcitabine in cancer therapy. *Future Medicine* 2005;**1**:7–17.
56. Kantarjian H, Jabbour E, Grimley J, Kirkpatrick P. Dasatinib. *Nat Rev Drug Discov* 2006;**5**(9):717–8.
57. Secord AA, Teoh D, Jia J, et al. Dasatinib (bms-35482) interacts synergistically with docetaxel, gemcitabine, topotecan, and doxorubicin in ovarian cancer cells with high src pathway activation and protein expression. *International journal of gynecologic. Cancer* 2014;**24**(2):218–25.
58. Ma L, Wei J, Su GH, Lin J. Dasatinib can enhance paclitaxel and gemcitabine inhibitory activity in human pancreatic cancer cells. *Cancer Biol Ther* 2019;**20**(6):855–65.
59. Iwamoto M, Friedman EJ, Sandhu P, et al. Clinical pharmacology profile of vorinostat, a histone deacetylase inhibitor. *Cancer Chemother Pharmacol* 2013;**72**:493–508.
60. Abdelgalil AA, Alkahtani HM, Al-Jenoobi FI. Sorafenib. *Profiles of drug substances, excipients, and related methodology*. 2019;**44**: 239–66.
61. Aggarwal R, Thomas S, Pawlowska N, et al. Inhibiting histone deacetylase as a means to reverse resistance to angiogenesis inhibitors: phase i study of abexinostat plus pazopanib in advanced solid tumor malignancies. *J Clin Oncol* 2017;**35**(11): 1231–9.
62. Xue X, Zhang Y, Liao Y, et al. Design, synthesis and biological evaluation of dual hdac and vegfr inhibitors as multitargeted anticancer agents. *Invest New Drugs* 2022;**40**(1): 10–20.
63. Hsu FT, Liu YC, Chiang I, et al. Sorafenib increases efficacy of vorinostat against human hepatocellular carcinoma through transduction inhibition of vorinostat-induced ERK/NF- κ B signaling. *Int J Oncol* 2014;**45**(1):177–88.
64. Jeannot V, Busser B, Vanwonterghem L, et al. Synergistic activity of vorinostat combined with gefitinib but not with sorafenib in mutant kras human non-small cell lung cancers and hepatocarcinoma. *Onco Targets Ther* 2016;**9**:6843–55.
65. Marks PA, Breslow R. Dimethyl sulfoxide to vorinostat: development of this histone deacetylase inhibitor as an anticancer drug. *Nat Biotechnol* 2007;**25**(1):84–90.
66. Wilhelm S, Carter C, Lynch M, et al. Discovery and development of sorafenib: a multikinase inhibitor for treating cancer. *Nat Rev Drug Discov* 2006;**5**(10):835–44.
67. Oshi M, Patel A, Le L, et al. G2m checkpoint pathway alone is associated with drug response and survival among cell proliferation-related pathways in pancreatic cancer. *Am J Cancer Res* 2021;**11**(6):3070–84.
68. Yadav A, Kumar B, Teknos TN, Kumar P. Sorafenib enhances the antitumor effects of chemoradiation treatment by down-regulating ercc-1 and xrcc-1 dna repair proteins. *Mol Cancer Ther* 2011;**10**(7):1241–51.
69. Cabral LKD, Tiribelli C, Sukowati CH. Sorafenib resistance in hepatocellular carcinoma: the relevance of genetic heterogeneity. *Cancer* 2020;**12**(6):1576.
70. Bhatt R, Zhang L, Schor-Bardach R, et al. Effects of sorafenib administration on the ifn signaling pathway and on the expression of ifn-inducible genes in rcc tumor cells. *Cancer Res* 2008;**68**(9_Supplement):1128–8.
71. Xu Z, Yang F, Wei D, et al. Long noncoding rna-srlr elicits intrinsic sorafenib resistance via evoking il-6/stat3 axis in renal cell carcinoma. *Oncogene* 2017;**36**(14):1965–77.
72. Zeng H, Qu J, Jin N, et al. Feedback activation of leukemia inhibitory factor receptor limits response to histone deacetylase inhibitors in breast cancer. *Cancer Cell* 2016;**30**(3):459–73.
73. Yang XM, Cao XY, He P, et al. Overexpression of rac gtpase activating protein 1 contributes to proliferation of cancer cells by reducing hippo signaling to promote cytokinesis. *Gastroenterology* 2018;**155**(4):1233–1249.e22.
74. Imaoka H, Toiyama Y, Saigusa S, et al. Racgap1 expression, increasing tumor malignant potential, as a predictive biomarker for lymph node metastasis and poor prognosis in colorectal cancer. *Carcinogenesis* 2015;**36**(3):346–54.
75. Pliarchopoulou K, Kalogeras K, Kronenwett R, et al. Prognostic significance of racgap1 mrna expression in high-risk early breast cancer: a study in primary tumors of breast cancer patients participating in a randomized hellenic cooperative oncology group trial. *Cancer Chemother Pharmacol* 2013;**71**:245–55.
76. Zhao Z, Mei Y, Wang Z, He W. The effect of oxidative phosphorylation on cancer drug resistance. *Cancer* 2022;**15**(1):62.
77. Ashton TM, McKenna WG, Kunz-Schughart LA, Higgins GS. Oxidative phosphorylation as an emerging target in cancer therapy. *Clin Cancer Res* 2018;**24**(11):2482–90.
78. Zhou C, Yi C, Yi Y, et al. Lncrna pvt1 promotes gemcitabine resistance of pancreatic cancer via activating wnt/ β -catenin and autophagy pathway through modulating the mir-619-5p/pygo2 and mir-619-5p/atg14 axes. *Mol Cancer* 2020;**19**:1–24.
79. Wang X, Zhu H, Jiang Y, et al. PRODeepSyn: predicting anticancer synergistic drug combinations by embedding cell lines with protein–protein interaction network. *Brief Bioinform* 2022;**23**(2):bbab587.
80. Theodoris CV, Xiao L, Chopra A, et al. Transfer learning enables predictions in network biology. *Nature* 2023;**618**(7965):616–24.