

RESEARCH

Open Access



# SAFER: sub-hypergraph attention-based neural network for predicting effective responses to dose combinations

Yi-Ching Tang<sup>1</sup>, Rongbin Li<sup>2</sup>, Jing Tang<sup>3,4</sup>, W. Jim Zheng<sup>2</sup> and Xiaoqian Jiang<sup>1\*</sup>

\*Correspondence:  
Xiaoqian.Jiang@uth.tmc.edu

<sup>1</sup> Department of Health Data Science and Artificial Intelligence, McWilliams School of Biomedical Informatics, The University of Texas Health Science Center at Houston, 7000 Fannin Street, Houston, TX, USA

<sup>2</sup> Department of Bioinformatics and Systems Medicine, McWilliams School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, TX, USA

<sup>3</sup> Research Program in Systems Oncology, Faculty of Medicine, University of Helsinki, Helsinki, Finland

<sup>4</sup> Department of Biochemistry and Developmental Biology, Faculty of Medicine, University of Helsinki, 00290 Helsinki, Finland

## Abstract

**Background:** The potential benefits of drug combination synergy in cancer medicine are significant, yet the risks must be carefully managed due to the possibility of increased toxicity. Although artificial intelligence applications have demonstrated notable success in predicting drug combination synergy, several key challenges persist: (1) Existing models often predict average synergy values across a restricted range of testing dosages, neglecting crucial dose amounts and the mechanisms of action of the drugs involved. (2) Many graph-based models rely on static protein–protein interactions, failing to adapt to dynamic and higher-order relationships. These limitations constrain the applicability of current methods.

**Results:** We introduce SAFER, a Sub-hypergraph Attention-based graph model, addressing these issues by incorporating complex relationships among biological knowledge networks and considering dosing effects on subject-specific networks. SAFER outperformed previous models on the benchmark and the independent test set. The analysis of subgraph attention weight for the lung cancer cell line highlighted JAK-STAT signaling pathway, PRDM12, ZNF781, and CDC5L that have been implicated in lung fibrosis.

**Conclusions:** SAFER presents an interpretable framework designed to identify drug-responsive signals. Tailored for comprehending dose effects on subject-specific molecular contexts, our model uniquely captures dose-level drug combination responses. This capability unlocks previously inaccessible avenues of investigation compared to earlier models. Furthermore, the SAFER framework can be leveraged by future inquiries to investigate molecular networks that uniquely characterize individual patients and can be applied to prioritize personalized effective treatment based on safe dose combinations.

**Keywords:** Hypergraph representation learning, Graph attention mechanisms, Drug combination prediction, Context-aware models, Dose–response drug combination data



## Background

Combination therapies are widely used therapeutic strategies in oncology, designed to enhance therapeutic outcomes and reduce the risk of acquired resistance [1]. By simultaneously targeting multiple pathways, these approaches offer the potential for more durable effects in cancer treatment. In certain cases, such combinations can provide greater benefits to cancer patients than single agents. However, finding effective drug combinations faces significant challenges, primarily due to the vast amount of potential drug and dosage combinations. Additionally, the limited or modest response to drugs complicates actionable drug combination discovery and clinical drug development. Notably, oncology drugs have the lowest clinical trial success rate (3.4%) among other types of treatment, which ranges from 15 to 33%, and the drug response rate in cancer patients is the lowest (25%) compared to patients with other diseases which had 30–80% of response rate [2]. This emphasizes the complexity of cancer development. Unfortunately, the complexity of highly variable drug responses can be attributed to several factors: (i) Individual variability. Cancer subtypes could respond differently to the specific targeted therapies [2]. (ii) Tissue specificity. Drugs may not always achieve their intended effects on tissue-specific pathways [3–5]. (iii) Inappropriate dosing. While a higher dosage may be effective for severe conditions, it could result in unexpected outcomes for patients with less severe conditions [6, 7]. Therefore, accurately predicting individual drug responses to combinatorial therapies is of critical importance for clinical drug discovery and development.

High-throughput drug screening, while efficiently screening large libraries of small molecules, encounters the challenge of combinatorial explosion. Furthermore, the efficiency of high-throughput screening techniques may unintentionally hinder the identification of meaningful synergistic interactions due to the inherent randomness in sample selection. In response to these challenges, advances in deep learning approaches have emerged as invaluable solutions. They help streamline drug development by accelerating the prioritization of drug candidates with large-scale training data, thus reducing both time and associated costs. As exemplified by DeepSynergy [8], the typical computational approach to predict drug synergistic effects is to learn the inter-correlation of drug molecule embeddings and baseline gene expression data through deep neural networks. Integration of heterogeneous data is crucial. High dimensional gene expression data were typically used to train a deep neural network, such as DeepSynergy, TranSynergy [9], and DeepDDs [10]. However, not all genes are predictive of drug response, and high-dimensional data are prone to overfitting [11, 12]. On the other hand, models have made attempts to extract drug embeddings from pair-wise similarity based on drug properties (i.e., CCSynergy [13]) or from a static protein–protein interaction (PPI) of drug-targeting genes (i.e., TranSynergy). DeepDDs employed graph attention networks to emphasize the sub-structures of drugs' molecular graphs as a drug feature representation. However, drug synergy is highly context-dependent, drugs may behave differently in different biological contexts, making it challenging to predict drug responses purely from similarity. Further, genes play different roles in different pathways, which indicate the nature of multiple relationships, and static PPI might not capture these nuances. Recently, hypergraph attention networks have gain increasing attention in text classification [14] and cancer subtype classification [15]. Different from a typical graph in which

an edge can only connect two nodes, hyperedges can connect two or more nodes, making it suitable for modelling complex functional networks, such as pathways. For example, in a pathway hypergraph, nodes represent genes and hyperedges connect genes that are in the same pathways. As a gene may participate in different pathways, there may be multiple edges connecting to this gene. Therefore, the use of hypergraph representation learning enables a model to learn a more expressive representation from a higher-order context.

In addition to multiple-relation modelling, dose amount is another critical factor for optimizing therapeutic outcomes and ensuring treatment safety. Nevertheless, current models are limited in their ability to determine personalized, safe, and effective treatment because they were designed to predict average synergistic effects within a narrow and inconsistent dose range. In this study, we present SAFER, a context-aware framework for drug combination prediction. Our contributions include: (1) We propose to use hypergraphs to represent complex relationships among drugs and among genes. This approach enables the model to comprehensively capture interactions within transcriptional regulatory networks associated with drug response. (2) We develop a multi-modal model that combines cell-gene interactions, drug-gene relationships, compound structures, and dosage data. This comprehensive framework enhances dose-level prediction accuracy by 20–30% compared to current models. (3) We demonstrated model's interpretability through subject-based attention weights, which uncovered distinct gene set importance signals for synergistic and antagonistic drug combination responses, including PRDM12, ZNF781, CDC5L-targeting genes, and the JAK-STAT signaling pathways (adjusted  $P$ -values  $< 0.05$ ).

## Materials and methods

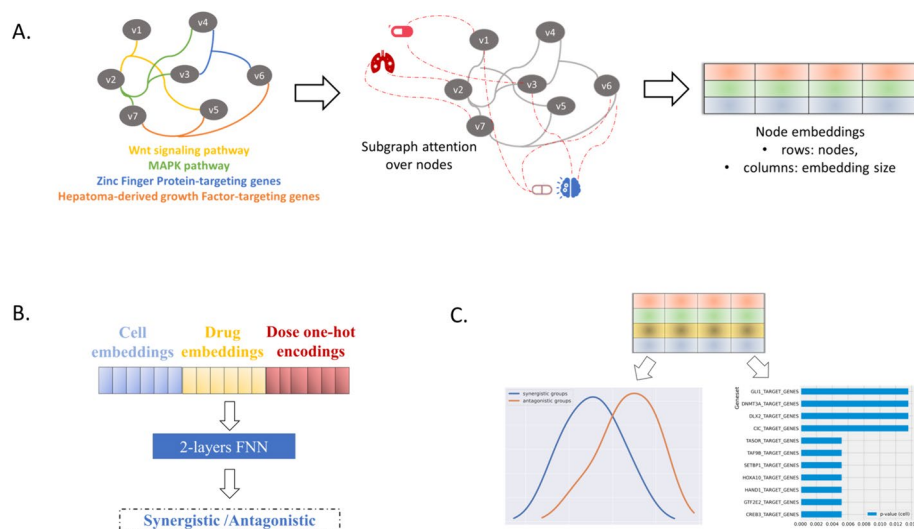
We first introduce SAFER, which stands for sub-hypergraph attention-based neural network for predicting effective responses to dose combinations. Then, we describe each component of the framework. Figure 1 shows the overview of the analysis pipeline. SAFER utilizes the attention-based hypergraph neural networks to learn the representations of cell line-gene, drug-gene, and drug-structure in the context of gene regulatory networks (Fig. 1A). It then uses two-layer feed-forward neural networks to learn the inter-correlation between these data representations along with dose combinations and synergistic effects at different dose combinations (Fig. 1B). Analysis of attention weights is designed to study the correlation between molecular mechanisms and responses to drug combinations (Fig. 1C).

### Sample collection

Table 1 describes the overview for the datasets used in this study. We provide detailed descriptions in the following.

### Drug combination data

We downloaded drug combination datasets from the DrugComb database (v1.5) [16]. Dose-level drug synergy measures were used in this study and were retrieved via the API (<https://api.drugcomb.org/response/>). The DrugComb database is the most comprehensive resource for studying drug synergism. It harmonized the degree of combination



**Fig. 1** A. Hypergraph representation learning: attention-based graph neural network was used to generate context-specific subject embeddings for downstream prediction tasks. B. Synergy prediction: A two-layer neural network takes in concatenation of three vectors, including cell line embedding, drug pair embedding, and one-hot encodings of drug-drug-cell line triplet, and output drug synergy label represented by 1 as synergistic or 0 as antagonistic. C. Attention weight analysis: comparison of attention weights between synergistic and antagonistic groups and highlight differentially weighted functional gene sets

**Table 1** Summarization of dataset collected and used in this paper

Category	Description	Source/data retrieval	Specifics/notes
Main dataset overview	Analysis of drug synergy measures	DrugComb database (v1.5)	Focus on non-additive drug combination effects; contains 225,608 data points, 104,819 combinations, 9078 pairs, 90 cancer cell lines
Cell line data	Genomics profiles of cancer cell lines	DepMap portal (22Q4)	Gene expression (log <sub>2</sub> TPM + 1), CRISPR screen gene essentiality scores, TF-gene regulatory network from GRAND
Drug data	Information on compound structures and associated genes	DrugComb database, DrugBank, TTD, STITCH	SMILE strings for compound identification, exclusion of non-human genes
Functional gene sets	Sets of genes related to canonical pathways and transcription factor-target genes	MSigDB (v2022)	Utilized C2 and C3 collections, gene size filtering based on one standard deviation from the mean
Benchmarks & validation	Reference datasets for comparison and validation	O'Neil et al. [21] (benchmark); Holbeck et al. [22] (validation)	Used to assess the performance of drug synergy predictions

additive, synergistic, or antagonistic by cell inhibition rate (i.e., percent inhibition or %inhibition) relative to the corresponding untreated control model, and presented with different synergy principles, including Bliss independence (BLISS) [17], Highest single agent (I) [18], Loewe additivity (LOEWE) [19], and Zero interaction potency (ZIP) [20] (more detailed description about the equations has been provided in the official website: <https://drugcomb.org/help/>). For each drug-drug-cell triplet, synergy scores were

averaged if it had experiment replicates. We removed non-cancer samples in the dataset (i.e., malaria and SARS-COV-2). The remaining samples that have available drug data and cell line data were included in this study. We focused on predicting drug combination effects deviating from additive, and therefore we excluded additive samples that had synergy scores within  $-10$  and  $10$  in all measures (i.e., BLISS, I, LOEWE, and ZIP). Additionally, samples with full agreement synergistic effects but had less ability to kill cancer cells (i.e., percent inhibition  $< 0$ ) than the untreated control were also excluded as they are not within the scope of our study. As a result, the datasets we used contain 225, 608 data points, consisting of 104,819 drug combinations and 9,078 drug pairs from 11 studies, and 90 cancer cell lines from 14 tissues. Finally, we used the dataset from the O'Neil study [21] as the benchmark for comparison purposes and used the Almanac dataset [22] for external validation.

### Cell line data

While baseline gene expression data remain widely used for their ability to predict drug response, gene essentiality scores and transcription factors have shown demonstrative potential [4, 9, 23]. Cancer cell lines used in this study were unified by DepMap ID, and their genomics profiles were retrieved from the DepMap portal (public 22Q4, <https://doi.org/https://doi.org/10.6084/m9.figshare.21637199.v2>), including the gene expression TPM values (i.e.,  $\log_2$  TPM + 1) of bulk RNA sequencing [24], gene knockout readouts derived from CRISPR screen (i.e., gene essentiality scores) [25–28]. In addition, we obtained the cell line-specific transcription factor (TF)-gene regulatory network from the GRAND [23] database and averaged across transcription factors to generate a cell line by gene matrix representing probabilities of being a TF-targeting gene in a cell line.

### Drug data

We leverage molecular structures and drug-targeting genes to learn property relationships and mechanism of actions. Compound structures (i.e., SMILE strings) were used and retrieved directly from the DrugComb database. In the dataset, drug names were coded using various labels, such as the company's serial number, the generic name, or chemical name. We unified drug names with unique SMILE strings to avoid data leakage. Drug-associated genes were collected from DrugBank [29], TTD [30], and STITCH [31] databases. Non-human genes were excluded, and gene symbols were assigned using NCBI's genes.

### Functional gene sets

Our previous findings revealed pathway-level signals have advantage to improve model performance and explainability [32]. Studies also showed tissue-specific drug response in transcription factors [4]. In this study, we used two different gene sets to learn biological contextual information of drug synergy. Functional gene sets were downloaded from the Molecular Signatures Database (MsigDB) [33], including C2 Canonical Pathways and C3 transcription factor-target genes (v2022). The C2 collection used in this study contains pathway gene sets from BioCarta [34], KEGG [35], PID [36], REACTOME [37], and WikiPathways [38]. The C3 collection is a subset of

GTRD [39] transcription factor (TF) binding genes. We retained gene sets using one standard deviation of the gene size. That is, gene size below or above one standard deviation of the mean was excluded.

### Sub-hypergraph representation learning

A hypergraph can be viewed as an incidence matrix,  $\mathcal{H} \in \{0, 1\}^{|\mathcal{V}| \times |\mathcal{E}|}$ , with nodes  $\mathcal{V}$  and hyperedges  $\mathcal{E}$ . It is different from a standard graph in that a hyperedge can connect to more than two nodes, making it suitable for learning complex interactions of molecular networks, such as gene-pathway relationships. The SHINE [15] framework represents gene-pathway relationship as a hypergraph in which nodes were genes and hyperedges were pathways, and then employs message passing neural networks to learn node representation. The main contribution of this learning process is that it allows for subgraph representation learning through both node and hyperedge representation. Specifically, the representation for a hyperedge  $p_j$  at layer  $k$ ,  $h_E^k(p_j)$ , is calculated from the nodes' representation at  $k - 1$  layer,  $h_V^{k-1}(g_i)$ , as in  $h_E^k(p_j) = \sigma \left( \sum_{g_i \in p_j} a_E(p_j, g_i) h_V^{k-1}(g_i) \right)$ , where  $\sigma$  is the ReLU non-linear function, and  $a_E(p_j, g_i)$  denotes the hyperedge attention which aggregates information over incident nodes and was computed as:

$$a_E(p_j, g_i) = \exp(c^T s(p_j, g_i)) / \left( \sum_{g_{i'} \in p_j} \exp(c^T s(p_j, g_{i'})) \right), \text{ where}$$

$c$  is a learnable context vector and  $s$  is the attention ready state. Likewise, the nodes' representation at layer  $k$  was obtained from the hyperedges' representation at layer  $k - 1$  as  $h_V^k(g_i) = \sigma \left( \sum_{p_j \ni g_i} a_V(g_i, p_j) h_E^{k-1}(p_j) \right)$ , where  $a_V(g_i, p_j)$  denotes the node attention which aggregates information over associated hyperedges and was computed as:

$$a_V(g_i, p_j) = \exp(c^T s(p_j, g_i)) / \left( \sum_{p_{j'} \ni g_i} \exp(c^T s(p_{j'}, g_i)) \right)$$

This dual attention allows the subsequent subgraph attention to learn nodes and hyperedges' representation simultaneously. Sub-hypergraph  $\mathcal{G}_j$  refers to a subset of nodes in a hypergraph and can be used to represent a subject. In other words, a subgraph representation,  $h(\mathcal{G}_j)$ , is characterized by its node representation in the hypergraph structure and was computed as  $h(\mathcal{G}_j) = \sigma \left( \sum_{g_i \in \mathcal{G}_j} a(\mathcal{G}_j, g_i) h_V^k(g_i) \right)$ , where  $a(\mathcal{G}_j, g_i)$  is the subgraph attention computed as:

$$a(\mathcal{G}_j, g_i) = \exp(F_{ji} b^T h_V^k(g_i)) / \left( \sum_{g_{i'} \in \mathcal{G}_j} \exp(F_{ji} b^T h_V^k(g_{i'})) \right), \text{ where } F_{ji} \text{ is the feature}$$

matrix in which each row represents a subject, and each column corresponds to node features. Finally, the output subgraph embedding,  $\mathcal{S}$ , is obtained through concatenation of all node representations:  $\mathcal{S} = [h(\mathcal{G}_1)^T | h(\mathcal{G}_2)^T | \dots | h(\mathcal{G}_v)^T]$ . In this study, we applied this subgraph attention mechanism to generate context-specific cell line-gene embedding, drug-gene embedding, and drug-structure embedding.



### Cancer-gene representation

We used the Python library, ssGSEApY (version 1.0.4) [40], to perform the Single-Sample Gene Set Enrichment Analysis (ssGSEA) using cancer cell lines' gene expression data, which outputs normalized enrichment score (NES) representing the level of enrichment of a gene set in a cell line. We represented the C3 TF-targeting genes as a hypergraph by using TFs as nodes and their targeting genes as hyperedge, and then applied the subgraph attention on the enrichment matrix (For SAFER-c2 model, the hypergraph is replaced by C2 pathway gene sets). By doing this, we can benefit from enrichment analysis to capture TF activities specific to cell lines and leverage the built-in dual attention mechanisms to learn the relationship across gene sets. To model dosage effects on cell line's transcriptional changes in the context of gene regulation, we injected dose combinations (e.g., the addition of drug dose a and drug dose b) into the node values by multiplying dose combination values with NES. Therefore, the subgraph embedding has the size of the number of subject and embedding dimensions, where subjects represent drug combinations (i.e., drugA-drugB-cell line) at different dose combinations (i.e., doseA-doseB).

### Drug-gene and drug-structure representation

Drug embeddings were extracted from the heterogeneous hypergraphs that consists of two types of nodes: molecular structures and drug-targeting genes. First, we created the structure hypergraph. To represent chemical structures in a hypergraph, we used the k-mer approach [41] which splits a SMILE string into sub-strings with a length of k as nodes. The choice of the k is determined by the ablation study. Second, we used hypergraph to model functional gene sets in which nodes are functional gene sets that drug-genes participated in (which is C2 pathways for SAFER-c2 model, and C3 TFs for SAFER-c3 model). A hypergraph can be represented as an incidence matrix where rows are nodes and columns are hyperedges. As we represent drugs as hyperedges, such that a cell value of one at node  $i$  for hyperedge drug  $j$  represents this drug has the substructure/TF-associated gene in  $i$ . To reflect dose effects, we injected the drug dose of each drug into the node values of its associated genes and sub-structures, respectively. We then used the same attention mechanism described above to generate subgraph embeddings for drug-gene and drug-structure multiple relationships.

### Dose combination representation

We also used one-hot encoding to represent drug doses in the dataset as we observed an increase in per-triplet AUPRC from 0.81 to 0.83. In addition, we encoded phenotypes of cancer cell lines, including tissue of origin, patient gender, and age in a one-hot vector as this information help to improve the overall AUPRC from 0.87 to 0.89. Therefore, we concatenated them all, including the dose combination vector.

### Synergy prediction

#### *Model architecture*

The resulting embeddings described above were concatenated and used as the input to the two-layer feed-forward neural (FNN) networks to generate synergy predictions.

We applied batch normalization to the input layer. The ELU activation function was used for the FNN layers. We binarized Loewe scores into synergistic samples ( $LOEWE > 10$ ) and antagonists ( $LOEWE < -10$ ) and used the BCELoss function combined with a sigmoid layer (i.e., the BCEWithLogitsLoss function in PyTorch library version 2.1.2 [42]). As for model training, we used group k fold to split dataset into ten folds to ensure that there is no overlapping drug-drug-cell line triplet in the training and the test folds. To reduce the impact of graph smoothing effects, we implemented regularization methods including L2 regularization, employed dropout rate, and limited the number of attention layers to two following previous approach [15]. We performed Bayesian optimization for hyperparameter tuning with the Optuna [43] package (version 3.1.0), and we tuned the hidden dimension of sub-hypergraph representation  $\in (100, 200, 300, 400, 500)$ , the dropout rate layer for the subgraph attention layer and for the feed-forward neural networks  $0.2 \leq dropoutrate \leq 0.8$ . All hyperparameters were tuned on the validation set in a ten-fold cross-validation manner to optimize the area under the precision–recall curve (AUPRC). The best hyperparameters were determined by averaging AUPRC on the validation set across all folds. We used a learning rate scheduler, so we did not tune the learning rate and weight decay. The final hyperparameters used in this study are as follows: hidden dimension = 200, dropout rates = 0.2, 0.3 for the subgraph layer and FNN layers, respectively.

### **Model evaluation**

The area under the receiver operating characteristic curve (AUROC) and the area under the precision–recall curve (AUPRC) were chosen to assess the overall performance of our binary classifier. To assess the model's capability of predicting dose-level drug combination response, we averaged AUROC and AUPRC by drug-drug-cell line groups which we called them per-triplet AUROC and per-triplet AUPRC, respectively. Model performance was evaluated internally and externally under the following scenarios:

- Cross-validation on the benchmark dataset

In this setting, we used the O'Neil dataset which is a screening of 490 drug pairs on 34 cell lines on a 5 by 5 dose-grid matrix. We split the data into 10 folds with a ratio of 80:10:10 for train, validation, and test partition. We then averaged performance on the test set across all folds for our models and the competing models, including the typical deep learning model, DeepSynergy, and three other state-of-the-art models, which are DeepDDs, TranSynergy, and CCSynergy. We reported the results of CCSynergy V as it shows a better performance among the other four versions. Hence, the CCSynergy mentioned in this study refers to CCSynergy V.

- Split the O'Neil dataset by pair-wise drug pair similarity

In this out-of-distribution (OOD) scenario, we created test sets distinct from the training and test data by dividing drug pairs based on their compound properties and overlapping targeting genes, respectively. We used two measures: Tanimoto similarity for comparing SMILE strings of compounds, and cosine similarity for target gene-based comparisons. This resulted in two data splits: one based on target similarity



(i.e., OOD-target) and another on compound similarity (OOD-structure). Test on the unseen data.

We used the dataset from the Almanac study that contains 4 by 4 and 4 by 6 dose-response matrices for 3811 drug pairs and 46 cell lines. We excluded the common drug-drug-cell line triplets seen in the O'Neil dataset, generating an independent test with 72,785 samples. We trained our model on the O'Neil data to obtain the best-performing model and then applied it to the Almanac data to generate predictions. We repeated this procedure ten times and reported mean and standard deviation.

- Evaluation on different dose areas

We obtained the half-maximal inhibitory concentration (i.e., IC<sub>50</sub>) for each drug from the DrugComb database and used that to dissect dose-response matrix into four quadrants for each drug-drug-cell line triplet in the dataset. We defined a higher dose if the tested concentration is above the IC<sub>50</sub> of a drug, a lower dose implies the drug is tested at a dose lower than its IC<sub>50</sub>. When graphically represented, these are the lower-left (or low dose area), upper-left (or low-high dose area), upper-right (or high dose area), and lower-right (or high-low dose area) quadrants. The overall AUPRC and AUROC for these areas is reported.

### **Model ablation analysis**

To find the best-performing model, we first compared different sizes of  $k$  for the compound structure hypergraph used for creating drug pair embeddings, and then we tested different cell line-gene embedding using single transcriptomics data against multi-omics. To construct the multi-omics model, we used DepMap's gene essentiality scores (GES) and GRAND's cell line-specific TF-gene regulatory networks (TF-GRNs) to obtain subgraph embeddings, respectively. Then, we concatenated them with the gene expression-based subgraph embeddings. The final comparison was conducted by comparing different molecular contexts including the C2 and the C3 collections, representing extracellular signaling pathways and intracellular regulatory networks, respectively.

### **Attention weight analysis**

#### **Statistical testing**

We obtained weighted subgraphs by multiplying attention weights by subgraph values (e.g., dosage effects on enrichment scores for cell lines), then we applied the Mann–Whitney U test to compare difference in weights between synergistic and antagonistic samples. The  $P$ -values were adjusted using the Benjamini–Hochberg method to control the false discovery rate at a level of 0.05.

#### **Sample pools**

We performed attention weight analysis for the top 20 drug-drug-cell line triplets our model performed well for. In this experiment, we sought to identify functional gene sets that were significantly differentially weighted between synergistic and antagonistic interactions. For each triplet, we first search such gene sets in a sample pool consisting of different dose combinations, and then search in another sample pool consisting of different cell lines that were also screened against the same pair of drugs. Gene sets with adjusted  $P$ -value smaller than 0.05 were reported.

## Results

We used hypergraphs to learn intricate relationships among cell line-gene, drug-gene, and drug-structure in the context of gene regulation. Dose effects on subject-specific networks captured by the subgraph attention mechanism allow us to study how individual molecular networks would be influenced by drug combinations and dose concentrations. The multi-modal model integrates different types of data to represent cell line data. Hence, we conducted the ablation study to find the best-performing modal.

Table 2 shows that increasing the size of k-mer can improve performance, among which using 9-mer to construct the chemical hypergraph obtained the best overall performance on the benchmark dataset with an average AUPRC of  $0.9 \pm 0.006$  and an average AUROC of  $0.773 \pm 0.009$ . While gene expression is beneficial for the prediction of drug response, studies have shown predictability of signals within GES [9] and TF-GRN [23]. Thus, we further constructed the multi-omics model incorporating gene expression, GES, and TF-GRN (see Methods), and compared it to the single transcriptomics model. The result showed that adding more features into the model slightly decreased the performance from 0.9 of AUPRC to 0.889 and 0.773 of AUROC to 0.751. We then continued with the single omics model that uses only gene expression as cell line representation for further comparison. The experiment of comparing different cellular contexts also showed that combining multiple cellular contexts did not benefit model performance. However, different gene sets contribute to model performance differently. Using the C3 TF-targeting gene set achieved better performance for all evaluation metrics, including AUPRC ( $0.9 \pm 0.06$  vs.  $0.889 \pm 0.005$ ), AUROC ( $0.773 \pm 0.009$  vs.  $0.751 \pm 0.008$ ), per-triplet AUPRC ( $0.825 \pm 0.229$  vs.  $0.808 \pm 0.238$ ), and per-triplet AUROC ( $0.697 \pm 0.362$  vs.  $0.668 \pm 0.370$ ). We continued with these two models. From now on, we term them as SAFER-C2 and SAFER-C3, respectively.

We went on to compare our models to other state-of-the-art models. All models were evaluated on the same data splits. As shown in Table 3, all models can perform well on the benchmark dataset with the average AUPRC above 0.8 and the average AUROC above 0.7. We observed that two deep neural networks, CCSynergy

**Table 2** Model ablation analysis under different settings

Experiment settings	AUPRC	AUROC	Per-triplet AUPRC	Per-triplet AUROC
Wo one-hot	$0.873 \pm 0.005$	$0.73 \pm 0.011$	$0.812 \pm 0.239$	$0.662 \pm 0.387$
With one-hot	<b><math>0.900 \pm 0.006</math></b>	<b><math>0.773 \pm 0.009</math></b>	<b><math>0.825 \pm 0.229</math></b>	<b><math>0.697 \pm 0.362</math></b>
3-mer	$0.888 \pm 0.005$	$0.750 \pm 0.08$	$0.811 \pm 0.236$	$0.671 \pm 0.369$
6-mer	$0.887 \pm 0.004$	$0.750 \pm 0.08$	$0.810 \pm 0.236$	$0.669 \pm 0.369$
9-mer	<b><math>0.900 \pm 0.006</math></b>	<b><math>0.773 \pm 0.009</math></b>	<b><math>0.825 \pm 0.229</math></b>	<b><math>0.697 \pm 0.362</math></b>
12-mer	$0.888 \pm 0.004$	$0.749 \pm 0.07$	$0.811 \pm 0.236$	$0.670 \pm 0.369$
Single transcriptomics	<b><math>0.900 \pm 0.006</math></b>	<b><math>0.773 \pm 0.009</math></b>	<b><math>0.825 \pm 0.229</math></b>	<b><math>0.697 \pm 0.362</math></b>
Multi-omics	$0.889 \pm 0.005$	$0.751 \pm 0.10$	$0.811 \pm 0.236$	$0.671 \pm 0.369$
C2	$0.889 \pm 0.005$	$0.751 \pm 0.008$	$0.808 \pm 0.238$	$0.668 \pm 0.370$
C3	<b><math>0.900 \pm 0.006</math></b>	<b><math>0.773 \pm 0.009</math></b>	<b><math>0.825 \pm 0.229</math></b>	<b><math>0.697 \pm 0.362</math></b>
C2 and C3	$0.887 \pm 0.005$	$0.749 \pm 0.009$	$0.808 \pm 0.237$	$0.668 \pm 0.369$

The best scores were bolded

Single transcriptomics: gene expression as cell line representation

Multi-omics: a concatenation of gene expression, cell line-specific TF-GRNs, and GES as cell line representation

**Table 3** Prediction performance on the O'Neil dataset

Model	AUPRC	AUROC	Per-triplet AUPRC	Per-triplet AUROC
SAFER-C2	0.889 ± 0.005	0.751 ± 0.008	0.808 ± 0.238	0.668 ± 0.370
SAFER-C3	0.900 ± 0.006	0.773 ± 0.009	<b>0.825 ± 0.229</b>	<b>0.697 ± 0.362</b>
CCSynergy	<b>0.943 ± 0.005</b>	<b>0.864 ± 0.009</b>	0.548 ± 0.216	0.500 ± 0.008
DeepDDs	0.884 ± 0.009	0.773 ± 0.011	0.564 ± 0.223	0.495 ± 0.136
DeepSynergy	0.935 ± 0.015	0.845 ± 0.015	0.548 ± 0.216	0.50 ± 0.0
TranSynergy	0.90 ± 0.006	0.771 ± 0.011	0.553 ± 0.217	0.50 ± 0.010
Baseline	0.73	0.5	0.56	0.5

The best scores were bolded

and DeepSynergy, performed better than the graph-based models, DeepDDs, TranSynergy, SAFER-C2, and SAFER-C3, achieving superior overall performance with  $0.943 \pm 0.005$ ,  $0.935 \pm 0.15$  of AUPRC, and  $0.864 \pm 0.009$ ,  $0.845 \pm 0.013$  of AUROC, which could be due to a larger number of hidden layer neurons between 2000 to 4096 that are able to capture more nuances in the data. However, such “deep” architecture is not suitable attention-based models because of graph smoothing problem [10]. Furthermore, given the subtle differences less than 0.05 in AUPRC, increasing the number of neurons in the feed-forward layers is not necessary for graph-based models. When testing on out-of-distribution (OOD) scenarios with substantial differences in drug pairs between training and test sets (average pairwise similarities of 0.13 for OOD-structure and 0.11 for OOD-target), all models showed reduced accuracies, as shown in Tables 4 and 5. However, SAFER-c2 and SAFER-c3 consistently outperformed DeepDDs, the graph attention network, suggesting that hypergraphs can better capture multiple relationships. Notably, deep neural networks generally produced

**Table 4** Prediction performance on OOD-target splits

Model	AUPRC	AUROC	Per-triplet AUPRC	Per-triplet AUROC
SAFER-C2	0.877 ± 0.06	0.732 ± 0.027	0.803 ± 0.239	0.657 ± 0.373
SAFER-C3	0.887 ± 0.022	0.750 ± 0.035	<b>0.817 ± 0.232</b>	<b>0.680 ± 0.368</b>
CCSynergy	0.921 ± 0.015	0.820 ± 0.019	0.548 ± 0.216	0.500 ± 0.0
DeepDDs	0.884 ± 0.026	0.772 ± 0.020	0.565 ± 0.223	0.500 ± 0.133
DeepSynergy	0.907 ± 0.033	0.789 ± 0.054	0.548 ± 0.216	0.499 ± 0.005
TranSynergy	<b>0.922 ± 0.011</b>	<b>0.820 ± 0.025</b>	0.552 ± 0.216	0.500 ± 0.0
Baseline	0.73	0.5	0.54	0.500 ± 0.0

The best scores were bolded

**Table 5** Prediction performance on OOD-smile splits

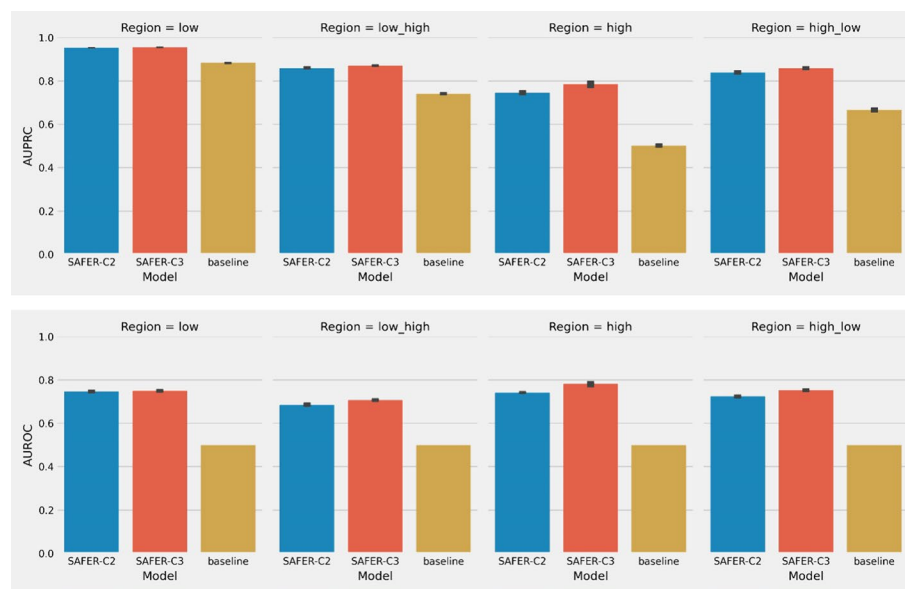
Model	AUPRC	AUROC	Per-triplet AUPRC	Per-triplet AUROC
SAFER-C2	0.880 ± 0.271	0.737 ± 0.031	0.802 ± 0.238	0.657 ± 0.370
SAFER-C3	0.891 ± 0.215	0.755 ± 0.024	<b>0.815 ± 0.233</b>	<b>0.681 ± 0.366</b>
CCSynergy	<b>0.921 ± 0.016</b>	<b>0.819 ± 0.023</b>	0.548 ± 0.216	0.500 ± 0.0
DeepDDs	0.883 ± 0.019	0.773 ± 0.014	0.567 ± 0.224	0.500 ± 0.133
DeepSynergy	0.914 ± 0.021	0.797 ± 0.034	0.548 ± 0.216	0.500 ± 0.0
TranSynergy	0.920 ± 0.012	0.815 ± 0.012	0.552 ± 0.216	0.500 ± 0.0
Baseline	0.73	0.5	0.54	0.500 ± 0.0

The best scores were bolded

better results compared to graph embedding methods, indicating that larger embedding sizes may be more effective.

On the comparison of per-triplet performance, SAFER surpassed all the other competing models, with a 20–30% of performance gain. In contrast, all previous state-of-the-art models showed poor performance no better than random guesses. This result suggests that our approach can capture response variations at the dose level, which can be confirmed by the result of the baseline model devised to predict the most prevalent label in the dataset. We can see that the baseline model can benefit from the unbalanced ratio between positive and negative samples (nearly 4 to 1) to achieve an overall performance above a random guess (AUPRC=0.73). However, it fails to distinguish dose-level variations as demonstrated by its' poor per-triplet performance that is equal to random guesses (per-triplet AUPRC=0.56, per-triplet AUROC=0.50). Such comparison demonstrated that dosing information is helpful for drug combination response classification.

It is of clinical interest to know if drug combinations can improve clinical outcomes without harm. Typically, higher dosages are more likely to cause unexpected outcomes than lower doses. This drives us to evaluate model performance at different dose levels. We used IC50 values to define low and high-dose areas and then averaged their AUPRC and AUROC scores (see Materials and Methods). The results show that our models perform well regardless of dose regions as demonstrated by their superior performance than the baseline model. Furthermore, SAFER-C2 and SAFER-C3 have equally good performance in all dose areas with a small standard deviation of 0.01 as displayed by the short error bars in Fig. 2. Averaging the AUPRCs of the two models, we obtained 0.953, 0.865, 0.765, and 0.855 for the low, low–high, high, and high–low dose areas, respectively. As for the mean AUROC, they were 0.748, 0.696, 0.765, and 0.738, respectively. Properly utilizing IC50 values is non-trivial as it is commonly used



**Fig. 2** Bar plots of AUPRC (upper panel) and AUROC (bottom panel) at four dose areas for models SAFER-C2 (blue) and SAFER-C3 (red). Error bars indicate standard deviation across all folds

to estimate drug response in vitro and in vivo. Furthermore, cross validation on the Almanac data that tested with 4 by 4 dose grid different from the O’Neil data which is 5 by 5 demonstrated SAFER’s applicability to different dose ranges (Average of AUPRC of two models = 0.775, AUROC = 0.777, average of per-triplet AUPRC = 0.8, per-triplet AUROC = 0.64). Together, in this experiment, we demonstrate that ranking drug combinations based on effective dosages is possible with our approach given its robust performance across different dose ranges.

The Almanac dataset contains cancer cell lines that were not included in the O’Neil’s study, including brain, haematopoietic and lymphoid, and kidney tissue despite the two datasets share similar age and sex distribution, providing a valuable independent data to investigate tissue-specific drug responses (see Additional file 3, Figure S1–3). To test whether our context-dependent model can benefit model generalizability across tissue types, we applied our models trained on the O’Neil dataset to generate predictions for the non-overlapping Almanac dataset. The baseline performance on this independent data is 0.498 and 0.5 for AUPRC and AUROC because the class ratio is one-to-one (35,648 positive samples and 37,277 negative samples). As shown in Table 6, most of the models had prediction performance close to the baseline, indicating the difficulties of this task. SAFER-C2 outperformed all models in all measures and all scores were above the baseline (AUPRC =  $0.540 \pm 0.004$ , AUROC =  $0.538 \pm 0.004$ , per-triplet AUPRC =  $0.677 \pm 0.314$ , per-triplet AUROC =  $0.549 \pm 0.430$ ), echoing previous findings that pathway information is more generalizable because of sharing pathways across cancer types [15, 32]. SAFER-C2 and SAFER-C3 showed better per-triplet performance standing at top of the other competing models, which once again shows the importance of dosing information to dose-level synergy predictions. While integrating different modalities could help improve accuracy, the overall performance may not be optimal for all competitors. Study has revealed the tissue-specificity nature of TF activities, which could provide insights into drug responses across various cancer types [4]. Hence, we reasonably hypothesize that this specificity might contribute to the observed low generalizability in our experiment. In other words, knowledge learned from the O’Neil dataset were not sufficient to capture the associations with drug responses in the brain, haematopoietic and lymphoid, and kidney tissues. This highlights the necessity of tailoring learning models to specific contexts.

**Table 6** Prediction performance on the non-overlapping Almanac dataset

Model	AUPRC	AUROC	Per-triplet AUPRC	Per-triplet AUROC
SAFER-C2	<b>0.540 ± 0.004</b>	<b>0.538 ± 0.004</b>	<b>0.677 ± 0.314</b>	<b>0.549 ± 0.430</b>
SAFER-C3	0.471 ± 0.014	0.452 ± 0.018	0.630 ± 0.240	0.560 ± 0.317
CCSynergy	0.512 ± 0.004	0.530 ± 0.004	0.500 ± 0.000	0.500 ± 0.142
DeepDDs	0.492 ± 0.007	0.504 ± 0.012	0.509 ± 0.140	0.499 ± 0.006
DeepSynergy	–	–	–	–
TranSynergy	0.321 ± 0.000	0.396 ± 0.000	0.625 ± 0.176	0.500 ± 0.000
Baseline	0.489	0.50	0.5	0.5

–Indicates processed data (i.e., features) not available

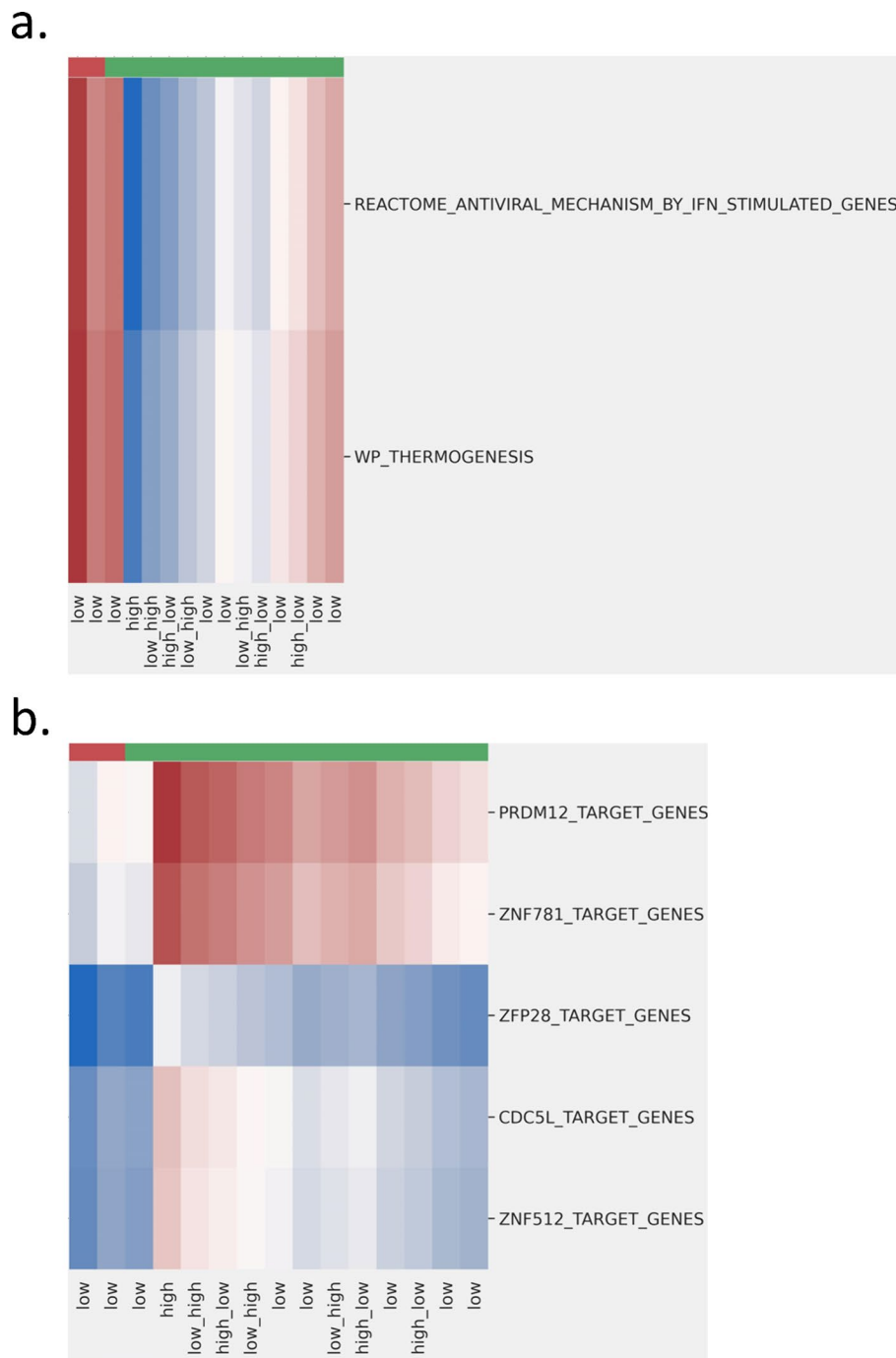
The best scores were bolded

Further exploration of activities of transcription factors within these tissues could offer valuable insight into their potential influence on drug responses.

SAFER has some model interpretation advantages that it can learn the importance of functional gene set through subject-level enrichment scores and subject-level attention weights. Through statistical testing, we identified gene sets that are differentially weighted between synergistic and antagonistic samples. Here, we highlight an example for the drug-drug-cell line triplet: 5-FU, Lapatinib, and MSTO, a lung cancer cell line that exhibits fibroblast morphology, and we provided the detailed list in the supplementary file (see Additional files 1–2, Table S1–2). Briefly, this triplet was tested with 16 different non-zero dose combinations, and the drug pair exhibited synergistic effects in this lung cancer cell line only at the low dose area. Our attention weight analysis identified several zinc finger proteins and two signaling pathways involved in immune response that might account for drug combination responses in this lung cancer cell line ( $n = 15$ , adjusted  $P$ -value = 0.03) (Additional file 1, Table S1). Notably, these functional gene sets were also found in the other lung cancer cell lines in the dataset ( $n = 14,218$ , adjusted  $P$ -value = 0.01) (Additional file 2, Table S2). The results of our data-driven analysis can be confirmed by literature findings presented below.

Patients with lung fibroblasts typically have difficulties to breath. This is because zinc deficiency can cause oxidative stress and because such zinc-dependent DNA damages will alter binding activity of zinc finger transcription factors [44, 45]. Studies have recognized the regulation roles of some zinc-finger proteins in oncogenesis [45, 46]. While different types of zinc finger proteins may have different contributions to diseases, our findings of the positive regulatory domain (PRDM12) that encodes for zinc-finger proteins, zinc-finger protein 781 (ZNF781), and cycle division cycle 5-like (CDC5L) echo previous research findings that suggest they could be biomarkers for lung adenocarcinoma. For example PRDM12 was found only upregulated in lung, ovary, and prostate, breast, colon, kidney, and liver cancers and is not expressed in normal tissues [47]. CDC5L may contribute to metastasis in lung cancer through regulating cell division [48]. In this example, we can see that drug response variations might be related to transcriptional changes resulted from changes in TF binding activities and/or subsequent post-translational modification. From SAFER-C2, we found REACTOME's antiviral mechanism by interferons (IFN) stimulated genes (i.e., JAK-STAT signaling pathway) is closely related to various cancers including non-small cell lung cancer [49–51]. The other pathway is the thermogenesis pathways of WikiPathways, a typical heat production process during immune reactions. Two pathways are related to immune response. It is worth noting that, the low doses, synergistic and the high doses, antagonistic responses were clustered in different groups by all the functional gene sets discussed above, as shown in the cluster heatmaps of Fig. 3. While it is known that adaptive immune system will develop more specific antibodies when carrying certain tumors for a long time, but little is known about if dose amount can influence the adaptive immunity. Further investigations are needed to find out whether increasing dose amount will induce a new immune response through heat production or enhance the existing defence mechanisms. Likewise, whether the amount of dose combination will influence TF-binding activity identified from SAFER-C3 needs to be elucidated.





**Fig. 3** Heatmaps of functional gene sets differentially weighted between synergistic and antagonistic samples from SAFER-C2 **(a)** and SAFER-C3 **(b)**. These figures utilize a diverging color palette, where higher weights are represented in red and lower weights in blue. Synergistic and antagonistic effects are annotated on the top of the heatmap, with red indicating synergistic and green indicating antagonistic interactions

## Discussions

We have presented the context-aware hypergraph-based prediction model. To the best of our knowledge, this study is the first work to employ attention-based hypergraph models for learning dose-level drug combination responses. It is known that previous state-of-the-art suffer from distinguishing dose-level synergistic interactions. On the other hand, the subgraph attention mechanisms were useful for capturing dosage effects on gene activities within molecular contexts. We injected dose information in every component, each of them provides distinct but complementary features that helps to improve prediction performance.

The internal and external validation indicates different molecular contexts have different contributions to predictive models: SAFER-C2 is more generalizable while SAEFR-C3 is more tissue-specific. Given their distinctive roles, integration analysis of signaling pathways and transcription factors could potentially provide a more comprehensive perspective of cellular communication [52]. However, our model performance dropped when using both C2 and C3 gene sets, indicating that current approach has limitation. We provided two future directions to address this: (1) the gene-set enrichment scores can be used to exclude irrelevant gene sets. (2) the attention-based hypergraph representation learning is based on static hypergraph structure, but different cancers or subtypes may develop unique pathways as disease progress. Effectively pruning hyperedges based on an individual's molecular characteristics could be an interesting direction. While we recognize that our approach may not match the performance of deep neural networks, expanding the embedding size would increase computational costs and potentially lead to a smoothing effect, diminishing the model's ability to differentiate between synergistic effects. Future research should not only focus on reducing the hypergraph size, as mentioned earlier but also explore hierarchical relationships among gene sets. Implementing hierarchical models could capture local and global contextual information, thereby preserving diversity in node embeddings without requiring overly large embedding dimensions. We expect our framework to be a useful platform for discovering mechanistic insights into dose-dependent drug combination responses.

Despite we carefully selected an independent data, the Almanac data is not patient tumor data. Although determination of drug doses is essential for clinical drug discovery, studies often release only drug responses either IC<sub>50</sub> values or clinical outcomes (e.g., Response Evaluation Criteria in Solid Tumors, or RECIST) when published their datasets. The lack of proper training data poses an obstacle for us to apply our models to tumor data. We suggest future investigators to also release dosing information for their studies, to increase the translational value of computational models by bridging the gap between immortal cancer cell lines and patient-derived tumors. Based on our previous success for monotherapies [32], we believe that SAFER-C2 has the transferrable potential and can generalize well to a tumor data through transfer learning.

## Conclusions

We have presented a multi-modal hypergraph-based model tailored for context-aware drug combination prediction and mechanistic interpretation of drug responses. We have demonstrated SAFER's capacity to effectively capture dose-level variations

in drug combination responses, surpassing the limitations of previous models. By integrating gene set knowledge and molecular contexts, SAFER offers a comprehensive understanding of the underlying mechanisms driving drug responses. Through a proof-of-concept application on established biological knowledge networks associated with context-specific drug response, SAFER showcases its adaptability to diverse biological gene sets of interest. Its innovative design enables the elucidation of the intricate relationship between dosage effects on molecular networks and drug synergistic interactions. SAFER provides an interpretable framework for exploring molecular context-dependent genetic changes and unlocking the potential for personalized medicine. By leveraging SAFER, future research can delve deeper into individual patient molecular profiles, ultimately enhancing our ability to tailor treatments and improve patient outcomes in diverse clinical settings.

#### Abbreviations

AUPRC	Area under precision–recall curve
AUROC	Area under receiver operating curve
IFN	Interferon
TF	Transcription factor
GES	Gene essentiality scores
TF-GRNs	TF-gene regulatory networks
PRDM12	Positive regulatory domain
ZNF781	Zinc-finger proteins, zinc-finger protein 781
CDC5L	Cycle division cycle 5-like
RECIST	Response evaluation criteria in solid tumors

#### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12859-024-05873-9>.

Additional file 1: Table S1. Differentially weighted functional gene sets per triplet.

Additional file 2: Table S2. Differentially weighted functional gene sets per drug pair.

Additional file 3: Comparison of data distribution between the O'Neil and the Almanac datasets. Figure S1. Bar plots of tissue distribution. Figure S2. Bar plots of age distribution. Figure S3. Bar plots of sex distribution.

Additional file 4.

Additional file 5.

Additional file 6.

#### Acknowledgements

Not applicable.

#### Author contributions

Y.T and X.J conceived of the project idea, developed the study and the computational framework. Y.T. carried out the implementation and statistical analysis and was a major contributor in writing the manuscript. R. L assisted with comparison of competing models. J.T, W.Z were involved in planning. X.J supervised the work. Y.T drafted the manuscript. All authors read and approved the final manuscript.

#### Funding

X.J is CPRIT Scholar in Cancer Research (RR180012), and he was supported in part by Christopher Sarofim Family Professorship, UT Stars award, UTHealth startup. This work is partially supported by the National Institute of Health (NIH) under award number R01AG066749, R01AG066749-03S1, R01LM013712, R01LM014520, R01AG082721, 5R56AG069880-02, U01AG079847, U24LM013755, U01CA274576, 1OT2OD032581-02–211, 1OT2OD032581-02–164, and the National Science Foundation (NSF) #2124789. R.L is supported by a training fellowship from the Gulf Coast Consortia, on the NLM Training Program in Biomedical Information & Data Science (T15LM007093). This work is also supported by the Academy of Finland under award number 317680.

#### Availability of data and materials

All datasets used in this study were openly available in public databases as described in our material and method sections. The processed datasets supporting the results of this article are available at <https://zenodo.org/records/10822806>. The open-source code can be found at <https://github.com/TangYiChing/SAFER/>.

## Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Competing interests

The authors declare no competing interests.

Received: 23 April 2024 Accepted: 17 July 2024

Published online: 30 July 2024

## References

- Fouquier J, Guedj M. Analysis of drug combinations: current methodological landscape. *Pharmacol Res Perspect*. 2015;3(3):e00149.
- Spear BB, Heath-Chiozzi M, Huff J. Clinical application of pharmacogenetics. *Trends Mol Med*. 2001;7(5):201–4.
- Cohen RL, Settleman J. From cancer genomics to precision oncology—tissue's still an issue. *Cell*. 2014;157(7):1509–14.
- Garcia-Alonso L, Iorio F, Matchan A, Fonseca N, Jaaks P, Peat G, et al. Transcription factor activities enhance markers of drug sensitivity in cancer. *Cancer Res*. 2018;78(3):769–80.
- Hao Y, Quinnes K, Realubit R, Karan C, Tatonetti NP. Tissue-specific analysis of pharmacological pathways. *CPT Pharmacomet Syst Pharmacol*. 2018;7(7):453–63.
- Bayat Mokhtari R, Homayouni T, Baluch N, Morgatskaya E, Kumar S, Das B. Combination therapy in combating cancer. *Oncotarget*. 2017;30:8.
- Tolcher AW, Mayer LD. Improving combination cancer therapy: the CombiPlex<sup>®</sup> development platform. *Future Oncol*. 2018;14(13):1317–32.
- Preuer K, Lewis RPI, Hochreiter S, Bender A, Bulusu KC, Klambauer G. DeepSynergy: predicting anti-cancer drug synergy with deep learning. *Bioinformatics*. 2018;34(9):1538–46.
- Liu Q, Xie L. TranSynergy: mechanism-driven interpretable deep neural network for the synergistic prediction and pathway deconvolution of drug combinations. *PLOS Comput Biol*. 2021;17(2):e1008653.
- Wang J, Liu X, Shen S, Deng L, Liu H. DeepDDS: deep graph neural network with attention mechanism to predict synergistic drug combinations. *Brief Bioinform*. 2022;23(1):bbab390.
- Adaptive Control Processes [Internet]. 2016 [cited 2023 Nov 13]. Available from: <https://press.princeton.edu/books/hardcover/9780691652214/adaptive-control-processes>
- Mirza B, Wang W, Wang J, Choi H, Chung NC, Ping P. Machine learning and integrative analysis of biomedical big data. *Genes*. 2019;10(2):87.
- Hosseini SR, Zhou X. CCSynergy: an integrative deep-learning framework enabling context-aware prediction of anti-cancer drug synergy. *Brief Bioinform*. 2023;24(1):bbac588.
- Ding K, Wang J, Li J, Li D, Liu H. Be more with less: hypergraph attention networks for inductive text classification. In: Proceedings of the 2020 conference on empirical methods in natural language processing (EMNLP) [Internet]. Online: Association for Computational Linguistics; 2020 [cited 2024 Jun 21]. p. 4927–36. Available from: <https://www.aclweb.org/anthology/2020.emnlp-main.399>
- Luo Y. SHINE: SubHypergraph inductive neural network. 2022 [cited 2023 Jun 14]; Available from: <https://arxiv.org/abs/2210.07309>
- Zheng S, Aldahdooh J, Shadbahr T, Wang Y, Aldahdooh D, Bao J, et al. DrugComb update: a more comprehensive drug sensitivity data repository and analysis portal. *Nucleic Acids Res*. 2021;49(W1):W174–84.
- Bliss CI. The toxicity of poisons applied jointly 1. *Ann Appl Biol*. 1939;26(3):585–615.
- Berenbaum MC. What is synergy? *Pharmacol Rev*. 1989;41(2):93–141.
- Loewe S. The problem of synergism and antagonism of combined drugs. *Arzneimittelforschung*. 1953;3(6):285–90.
- 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;25(13):504–13.
- O'Neil J, Benita Y, Feldman I, Chenard M, Roberts B, Liu Y, et al. An unbiased oncology compound screen to identify novel combination strategies. *Mol Cancer Ther*. 2016;15(6):1155–62.
- Holbeck SL, Camalier R, Crowell JA, Govindharajulu JP, Hollingshead M, Anderson LW, et al. The National Cancer Institute ALMANAC: a comprehensive screening resource for the detection of anticancer drug pairs with enhanced therapeutic activity. *Cancer Res*. 2017;77(13):3564–76.
- Ben Guebila M, Lopes-Ramos CM, Weighill D, Sonawane AR, Burkholz R, Shamsaei B, et al. GRAND: a database of gene regulatory network models across human conditions. *Nucleic Acids Res*. 2022;50(D1):D610–21.
- Ghandi M, Huang FW, Jané-Valbuena J, Kryukov GV, Lo CC, McDonald ER, et al. Next-generation characterization of the cancer cell line encyclopedia. *Nature*. 2019;569(7757):503–8.
- Dempster JM, Rossen J, Kazachkova M, Pan J, Kugener G, Root DE, et al. Extracting Biological Insights from the Project Achilles Genome-Scale CRISPR Screens in Cancer Cell Lines [Internet]. bioRxiv; 2019 [cited 2023 Nov 4]. p. 720243. Available from: <https://doi.org/10.1101/720243v1>
- Dempster JM, Boyle I, Vazquez F, Root D, Boehm JS, Hahn WC, et al. Chronos: a CRISPR cell population dynamics model [Internet]. bioRxiv; 2021 [cited 2023 Nov 4]. p. 2021.02.25.432728. Available from: <https://doi.org/10.1101/2021.02.25.432728v1>

27. Meyers RM, Bryan JG, McFarland JM, Weir BA, Sizemore AE, Xu H, 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.
28. Pacini C, Dempster JM, Boyle I, Gonçalves E, Najgebauer H, Karakoc E, et al. Integrated cross-study datasets of genetic dependencies in cancer. *Nat Commun.* 2021;12(1):1661.
29. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.* 2018;46(Database issue):D1074–82.
30. Zhou Y, Zhang Y, Zhao D, Yu X, Shen X, Zhou Y, et al. TTD: therapeutic target database describing target druggability information. *Nucleic Acids Res.* 2023;15:gkad751.
31. Szklarczyk D, Santos A, von Mering C, Jensen LJ, Bork P, Kuhn M. STITCH 5: augmenting protein–chemical interaction networks with tissue and affinity data. *Nucleic Acids Res.* 2016;44(Database issue):D380–4.
32. Tang YC, Powell RT, Gottlieb A. Molecular pathways enhance drug response prediction using transfer learning from cell lines to tumors and patient-derived xenografts. *Sci Rep.* 2022;12(1):16109.
33. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. *Bioinformatics.* 2011;27(12):1739–40.
34. Nishimura D. BioCarta. *Biotech Softw Internet Rep.* 2001;2(3):117–20.
35. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 2000;28(1):27–30.
36. Schaefer CF, Anthony K, Krupa S, Buchoff J, Day M, Hannay T, et al. PID: the pathway interaction database. *Nucleic Acids Res.* 2009;37:674–9.
37. Jassal B, Matthews L, Viteri G, Gong C, Lorente P, Fabregat A, et al. The reactome pathway knowledgebase. *Nucleic Acids Res.* 2020;48(D1):D498–503.
38. Martens M, Ammar A, Riutta A, Waagmeester A, Slenter DN, Hanspers K, et al. WikiPathways: connecting communities. *Nucleic Acids Res.* 2021;49(D1):D613–21.
39. Yevshin I, Sharipov R, Kolmykov S, Kondrakhin Y, Kolpakov F. GTRD: a database on gene transcription regulation-2019 update. *Nucleic Acids Res.* 2019;47(D1):D100–5.
40. Fang Z, Liu X, Peltz G. GSEApy: a comprehensive package for performing gene set enrichment analysis in Python. *Bioinforma Oxf Engl.* 2023;39(1):btac757.
41. Saifuddin KM, Bumgardner B, Tanvir F, Akbas E. HyGNN: drug-drug interaction prediction via hypergraph neural network. In: 2023 IEEE 39th international conference on data engineering (ICDE) [Internet]. 2023 [cited 2023 Oct 20]. p. 1503–16. Available from: <https://ieeexplore.ieee.org/abstract/document/10184559/authors#authors>
42. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. PyTorch: an imperative style, high-performance deep learning library. In: *Advances in neural information processing systems*; 2019. p. 8026–37.
43. Akiba T, Sano S, Yanase T, Ohta T, Koyama M. Optuna: A next-generation hyperparameter optimization framework [Internet]. arXiv; 2019 [cited 2023 Oct 20]. Available from: <http://arxiv.org/abs/1907.10902>
44. Ho E, Courtemanche C, Ames BN. Zinc deficiency induces oxidative DNA damage and increases p53 expression in human lung fibroblasts. *J Nutr.* 2003;133(8):2543–8.
45. Liu X, Ali MK, Dua K, Xu R. The role of zinc in the pathogenesis of lung disease. *Nutrients.* 2022;14(10):2115.
46. Cassandri M, Smirnov A, Novelli F, Pitolli C, Agostini M, Malewicz M, et al. Zinc-finger proteins in health and disease. *Cell Death Discov.* 2017;3:17071.
47. Rienzo M, Di Zazzo E, Casamassimi A, Gazerro P, Perini G, Bifulco M, et al. PRDM12 in health and diseases. *Int J Mol Sci.* 2021;22(21):12030.
48. Zhou RH, Zhang JT, Chen C, Xu ZH, Lv XB, Ye L, et al. Identification of CDC5L as bridge gene between chronic obstructive pulmonary disease and lung adenocarcinoma. *Epigenomics.* 2020;12(17):1515–29.
49. Hu X, Li J, Fu M, Zhao X, Wang W. The JAK/STAT signaling pathway: from bench to clinic. *Signal Transduct Target Ther.* 2021;6:402.
50. Wong KY, Cheung AHK, Chen B, Chan WN, Yu J, Lo KW, et al. Cancer-associated fibroblasts in nonsmall cell lung cancer: from molecular mechanisms to clinical implications. *Int J Cancer.* 2022;151(8):1195–215.
51. Yoshida GJ. Regulation of heterogeneous cancer-associated fibroblasts: the molecular pathology of activated signaling pathways. *J Exp Clin Cancer Res CR.* 2020;39(1):112.
52. Weidemüller P, Kholmatov M, Petsalaki E, Zaugg JB. Transcription factors: bridge between cell signaling and gene regulation. *Proteomics.* 2021;21(23–24):e2000034.

## Publisher's Note

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