GNNSynergy: A multi-view graph neural network for predicting anti-cancer drug synergy

This report gives supplementary information to the manuscript "GNNSynergy: A multi-view graph neural network for predicting anti-cancer drug synergy". It provides more detailed information about dataset, parameter setting in both baseline methods and GNNSynergy, and additional experimental results.

1. DATASET

Table S1 shows the cancer cell lines included in the DrugComb database. There are 81 cell lines originated from 11 different tissues.

Table S1. The 81 cell lines tested in the DrugComb Database, covering 11 different tissue types. Amount indicated the amount of cell lines in the tissue.

Tissue (# cell lines)	Number of drug pairs	Cell Line
Large Intestine (10)	35297	SW-620; HT29; HCT116; HCT-15; KM12; COLO 205; HCC-2998; LOVO; SW837; RKO
Breast (8)	29421	T-47D; MDA-MB-231; MCF7; BT-549; MDA-MB-468; HS 578T; OCUBM; MDAMB436;
Lymphoid (8)	25234	K-562; CCRF-CEM; SR; MOLT-4; RPMI-8226; L-1236; HDLM-2; L-428;
Kidney (8)	37363	ACHN; SN12C; TK-10; A498; 786-0; CAKI-1; UO-31; RXF 393;
Lung (13)	39934	NCIH23; NCI-H226; A549; EKVX; NCI-H322M; NCI-H522; HOP-92; HOP-62; A427; SKMES1; NCIH2122; NCIH520; NCIH1650
Ovary (11)	31186	SK-OV-3; OVCAR3; OVCAR-5; IGROV1; OVCAR-8; OVCAR-4; A2780; PA1; ES2; UWB1289; OV90;
Skin (11)	33389	UACC62; UACC-257; SK-MEL-28; SK-MEL-5; M14; LOX IMVI; SK-MEL-2; A375; RPMI7951; A2058; HT144
Brain (6)	24883	SF-268; U251; SF-539; SF-295; SNB-75; T98G
Prostate (3)	9872	PC-3; DU-145; VCAP
Bone (2)	1878	A-673; TC-71
Soft (1)	80	RD

Fig. S1 shows the distribution of true synergy scores and predicted synergy scores based on the processed DrugComb dataset. We can observe that the true score distribution presents a normal distribution, and there are few synergistic drug pairs and antagonistic drug pairs. In addition, the score distribution predicted by GNNSynergy is basically consistent with the real score distribution.

Besides, we performed the t-SNE analysis to visualize the cell lines in 2D space to reflect relationships between them (Fig. S2). Originally, each cell line has 972 features for its gene

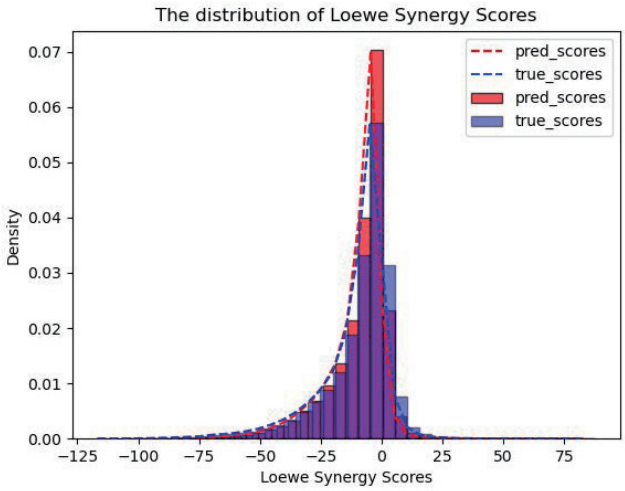


Fig. S1. The distribution of true synergy scores and predicted synergy scores based on the processed DrugComb dataset.

expression profile, and here we applied t-SNE to map the 972-dimensional vectors into the 2D space. It shows that cell lines in the same tissues are closer to each other. Given a specific cell line, the information of other cell lines in the same tissue would be useful for drug synergy prediction.

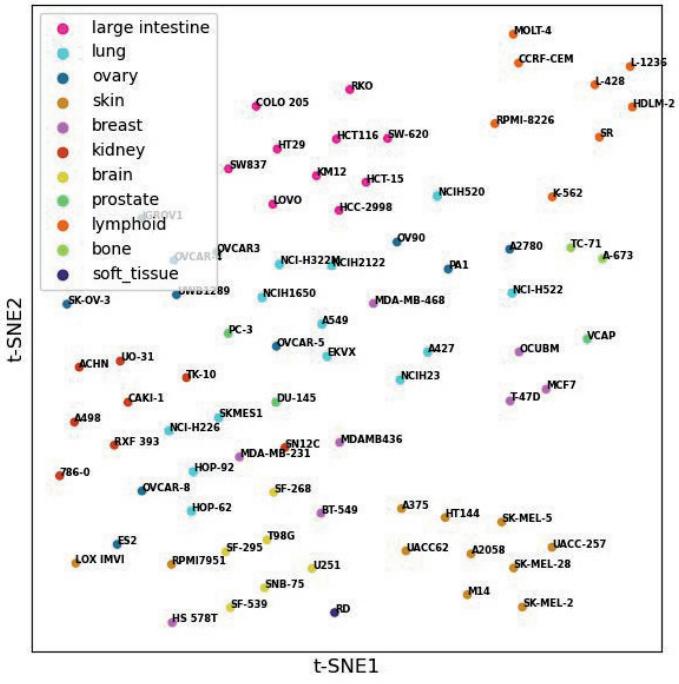


Fig. S2. Visualization of different cell lines with t-SNE analysis. Different colors indicate different tissues. It shows that cell lines in the same tissues are closer to each other.

2. HYPERPARAMETERS SETTING AND ANALYSIS

A. Hyper-parameter Setting

The hyper-parameters of all the methods were optimized by validation set, using grid search. Table S2 shows the search range for different hyper-parameters in baseline methods, including Elastic Net (EN), Random Forest (RF), Gradient Boosting Machine (GBM), TreeCombo (XGBoost) and comboLTR, respectively. For the deep learning methods, such as DeepSynergy, TranSynergy and MatchMaker, we used the default parameters used in their original papers.

Table S2. Hyperparameter space considered for Baselines

Model	Hyperparament	Values consided
EN	Constant α	0.001; 0.01; 0.1; 1; 10; 100
	L1 ratio	0.2; 0.4; 0.6; 0.8
RF	numbers of estimators (decision trees)	10; 100; 500; 1000
	maximum tree depth	4; 6; 8; 10; 12
GBM	numbers of estimators (decision trees)	10; 100; 500; 1000
	maximum tree depth	4; 6; 8; 10; 12
	learning rate	0.05; 0.10; 0.15
XGBoost	maximum tree depth	4; 6; 8; 10; 12
	learning rate	0.05; 0.10; 0.15
comboLTR	rank_uv	20; 32; 64; 128; 256
	repeats	20; 70; 120; 170; 220; 270
	ranks	20; 40; 60; 80; 100

For our GNNSynergy, we used 1-layer GCN with output feature dimensionality $d = 256$ and dropout rate $\rho = 0.7$. We used 4-layer MLPs (i.e., 1 input layer, 2 hidden layers and 1 output layer) in both single-view and multi-view encoders. The input layer of MLPs in single-view encoder has $3d = 768$ neurons, and it has $3d \times k$ neurons in multi-view encoder with k sub-view cell lines. For MLPs in single-view encoder, their 2 hidden layers and 1 output layer had 256, 512 and 256 neurons, respectively, and the dropout rates for 2 hidden layers were set as 0.5 and 0.2. For MLPs in multi-view encoder, their 2 hidden layers and 1 output layer had 1280, 512 and 256 neurons, respectively, and the dropout rates for 2 hidden layers were set as 0.5 and 0. Note that we pre-trained the single-view encoder for each cell line with a learning rate of 1e-3, while the learning rate to optimize the overall GNNSynergy model was set as 1e-5. Lastly, we set the maximum number of training epochs to 2000, and we considered the early-stop mechanism that the optimization will stop if the validation loss does not decline within the recent 300 epochs. Table S3 summarizes the parameter settings in our GNNSynergy.

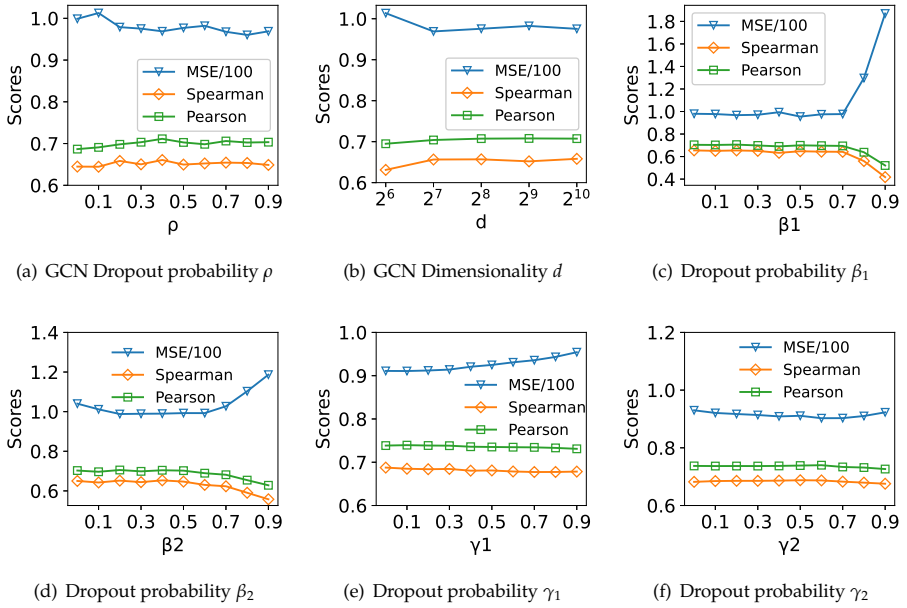
B. Parameter Sensitivity Analysis

We present the sensitivity analysis for the parameters in our GNNSynergy, including the GCN dimensionality d, GCN dropout probability ρ, the dropout probability β_1, β_2 of MLP in Single-View, and the dropout probability γ_1, γ_2 of MLP in Multi-View, as shown in Fig. S3.

From the Fig. S3 (a), the performance of GNNSynergy is relatively stable when ρ is set as different values. Thus, we recommend to set ρ in the range $[0.5, 0.9]$. In Fig. S3 (b), we could observe that the medium values for d e.g., $d = 2^7$ or $d = 2^8$ are more favorable. In our study, we used $d = 2^8 = 256$. Fig. S3 (c) and Fig. S3 (d) show the performance of GNNSynergy using different values for β_1 and β_2 (the dropout probabilities in MLPs in single-view encoder), respectively. We change one of them and fix the other to its default value. As for β_1, we can observe that the MSE of GNNSynergy starts to increase when β_1 is larger than 0.7. Similarly, the performance get worse when $\beta_2 > 0.5$. In this paper, we set β_1 as 0.2 and β_2 as 0.5 in our experiments. In Fig. S3 (e) and Fig. S3 (f), we can observe that GNNSynergy achieves the best performances when $\gamma_1 = 0, \gamma_2 = 0.5$.

Table S3. Hyper-parameter settings for GNNSynergy

Module	Hyperparameter	Value
Global	learning rate in single-view (η_1)	1e-3
	learning rate in multi-view (η_2)	1e-5
	max epoch	2000
	early-stop epoch	300
GCN	dimensionality (d)	256
	dropout probability (ρ)	0.7
MLP in Single-View	dimensionality	[256, 512, 256]
	dropout probability (β)	[0.2,0.5]
MLP in Multi-View	dimensionality	[1280, 512, 256]
	dropout probability (γ)	[0.0, 0.5]



(a) GCN Dropout probability ρ (b) GCN Dimensionality d (c) Dropout probability β_1

(d) Dropout probability β_2 (e) Dropout probability γ_1 (f) Dropout probability γ_2

Fig. S3. Parameter sensitivity analysis for GNNSynergy.

3. RESULTS

A. Investigation of prediction performance among all specific cell lines

We further investigated the performance of GNNSynergy among different cell lines. The spearman correlation coefficients and pearson correlation coefficients for GNNSynergy, MatchMaker and DeepSynergy are shown in Fig. S4. Obviously, GNNSynergy performs better than the MatchMaker in most of cell lines. Besides, Fig. S5 shows the cell line-specific MSE predicted by GNNSynergy.

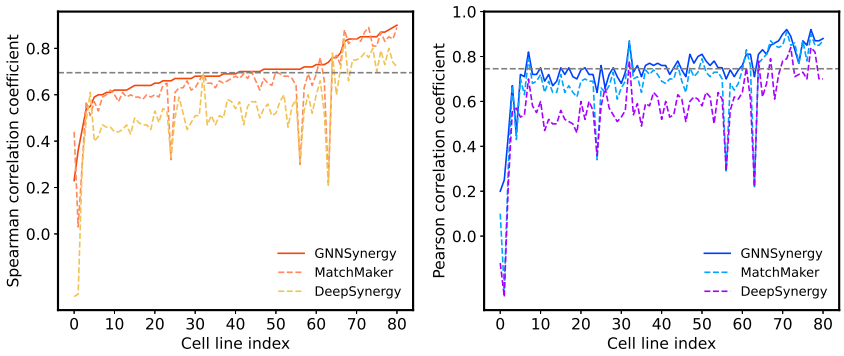


Fig. S4. The cell line-specific predicted Spearman correlation coefficient and Pearson correlation coefficient comparison among GNNSynergy, DeepSynergy and MatchMaker. Here, MatchMaker is the best baseline as shown in our main manuscript. The gray horizontal dotted line represents the average score of GNNSynergy.

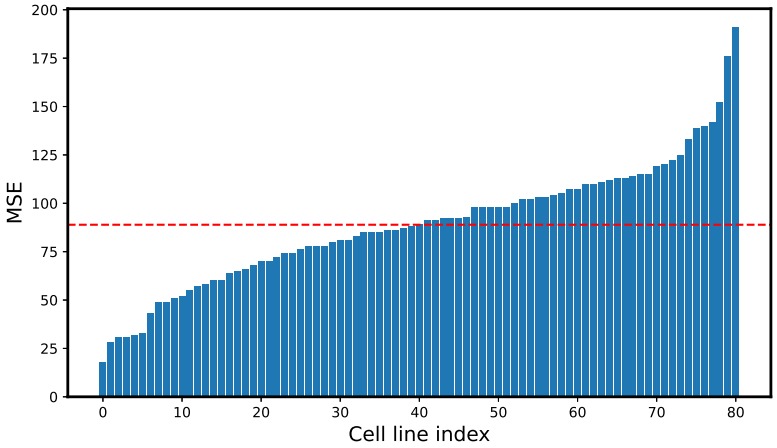


Fig. S5. The MSE scores predicted by GNNSynergy among all cell lines. The red horizontal dotted line represents the average score of GNNSynergy.

B. Model Ablation Study

In addition to the attention weights automatically learned in GNNSynergy, we also tried a simple way to derive the weighted concatenation from different cell lines, i.e., we used fixed equal weights for all the sub-view cell lines. Table S4 shows the comparison between 'Equal Weights' and 'Attention Weights', demonstrating that the attention weights can help to effectively integrate multiple cell lines for learning drug embeddings and predicting drug synergy scores.

Recall that we build three DDS graphs in each cell line to describe different drug synergistic effects. Here, we derive three additional variants by using different number of DDS graphs and

Table S4. The weights for different sub-view cell lines in large intestine, when the main view cell line is SW-620.

Strategy	Equal Weights	Attention Weights
sub-view 1	0.1111	0.1236
sub-view 2	0.1111	0.1558
sub-view 3	0.1111	0.0953
sub-view 4	0.1111	0.1083
sub-view 5	0.1111	0.0417
sub-view 6	0.1111	0.1699
sub-view 7	0.1111	0.0519
sub-view 8	0.1111	0.0643
sub-view 9	0.1111	0.1893

compare their performances. The variant 'One Graph' refers to using a single DDS graph to describe all the drug combination pairs (i.e., we do not differentiate synergistic or antagonistic effects in this variant). The variant 'Two Graphs (t=0)' considers two DDS graphs, namely, synergy graph where edges have positive Loewe scores and antagonism graph with negative edges (i.e., threshold t is set as 0). The variant 'Two Graphs (t=10)' considers two DDS graphs, namely, synergy graph where edges have Loewe scores greater than 10 and antagonism graph with negative Loewe scores less than -10 (i.e., threshold t is set as 10). The variant 'Three Graphs' is our GNNSynergy with threshold t set as 10. As shown in Fig. S6, GNNSynergy with three DDS graphs achieves the best performance.

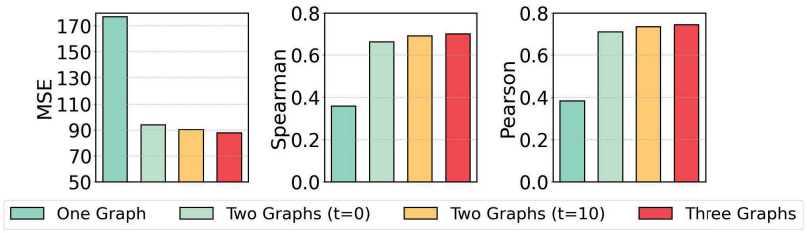


Fig. S6. The performance comparison of three variants to divide the DDS graph.

C. Case studies for novel drug combination prediction

As mentioned in our main manuscript, we found that 19 out of these 4,050 predicted drug pairs are considered as synergistic pairs in DrugCombDB. These 19 predicted pairs are listed in Table S5.

Table S5. 19 predicted drug combination pairs with literature support.

#	Rank	Drug 1	Drug 2	Cell line
1	302	AURANOFIN	TRAMETINIB	NCI-H226
2	497	AURANOFIN	PD325901	NCI-H226
3	834	AURANOFIN	CISATRACURIUM BESYLATE	NCIH23
4	992	CISATRACURIUM BESYLATE	AURANOFIN	UWB1289
5	1207	AURANOFIN	CARFILZOMIB (PR-171)	SF-268
6	2192	AURANOFIN	TRAMETINIB	HOP-92
7	2356	CISATRACURIUM BESYLATE	AURANOFIN	LOVO
8	3219	CARFILZOMIB (PR-171)	AURANOFIN	LOVO
9	3454	ZALCITABINE	AURANOFIN	SR
10	3700	NILOTINIB	SORAFENIB	T98G
11	3704	NILOTINIB	VEMURAFENIB	T98G
12	3714	IMATINIB	NILOTINIB	T98G
13	3716	NILOTINIB	GEFITINIB	T98G
14	3722	SORAFENIB	VEMURAFENIB	T98G
15	3729	IMATINIB	SORAFENIB	T98G
16	3742	ERLOTINIB HYDROCHLORIDE	NILOTINIB	T98G
17	3907	ISONIAZID	BORTEZOMIB	HDLM-2
18	3909	LEFLUNOMIDE	BORTEZOMIB	HDLM-2
19	3925	PROCARBAZINE	BORTEZOMIB	HDLM-2