Interdisciplinary Sciences: Computational Life Sciences (2023) 15:316–330
https://doi.org/10.1007/s12539-023-00558-y

ORIGINAL RESEARCH ARTICLE

Predicting Drug Synergy and Discovering New Drug Combinations Based on a Graph Autoencoder and Convolutional Neural Network

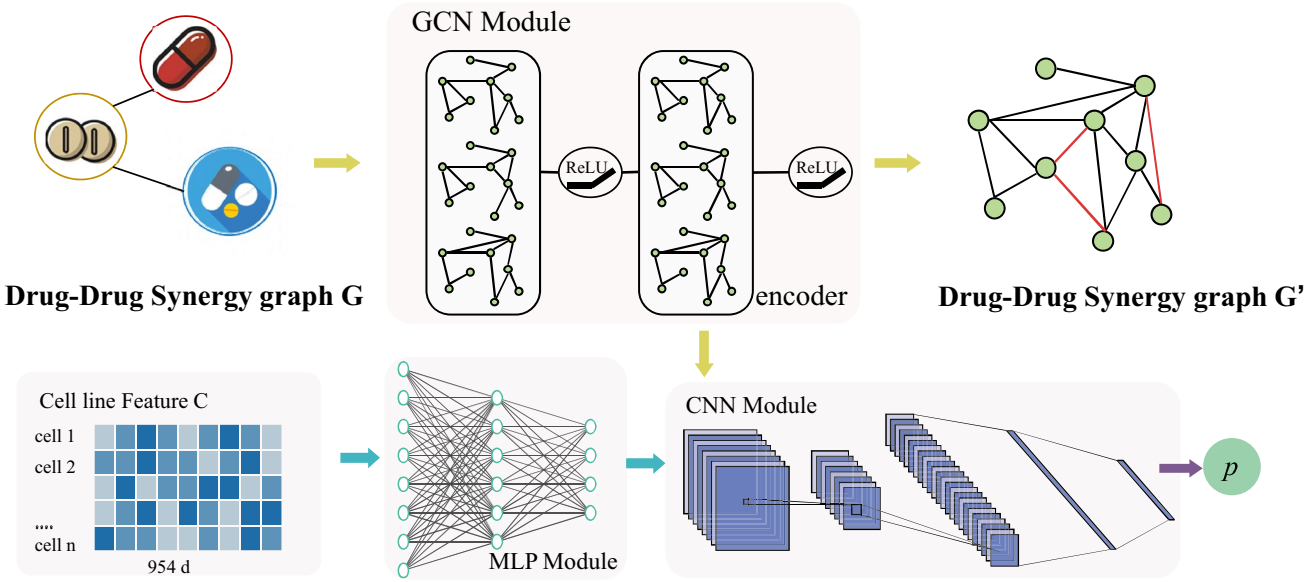
Huijun Li[1] · Lin Zou[1] · Jamal A. H. Kowah[2] · Dongqiong He[2] · Lisheng Wang[1] · Mingqing Yuan[1] · Xu Liu[1]

Received: 1 June 2022 / Revised: 23 February 2023 / Accepted: 23 February 2023 / Published online: 21 March 2023

Abstract

Drug synergy is a crucial component in drug reuse since it solves the problem of sluggish drug development and the absence of corresponding drugs for several diseases. Predicting drug synergistic relationships can screen drug combinations in advance and reduce the waste of laboratory resources. In this research, we proposed a model that utilizes graph autoencoder and convolutional neural networks to predict drug synergy (GAECDS). Our methods include a graph convolutional neural network as an encoder to encode drug features and use a matrix factorization method as a decoder. Multilayer perceptron (MLP) was applied to process cell line features and combine them with drug features. Furthermore, the latent vectors generated during the encoding process are being used to predict drug synergistic scores using a convolutional neural network. By measuring prediction performance using AUC, AUPR, and F1 score, GAECDS superior to other state-of-the-art models. In addition, four pairs of the predicted top 10 drug combinations were found to work well enough for evaluation. The case study shows that the GAECDS approach is useful for identifying potential drug synergy.

Graphical Abstract



Keywords Drug synergy · Deep learning · Graph autoencoder · Graph convolutional neural network

✉ Xu Liu
wendaoliuxu@163.com

Extended author information available on the last page of the article

1 Introduction

Recently, research on drug synergy has grown significantly and demonstrated considerable potential. Reduced drug resistance, improved treatment options, and decreased dosage and toxicity are the main aims of drug pairing or the administration of multiple drugs. Compared with new drug development, the drug synergy method saves considerable manpower and material resources. Traditional drug synergy mainly relies on experience to screen and analyze drugs one by one, which is time-consuming. Computational approaches for drug combination screening have shown significant promise as science and technology have developed. The application of computational methods for screening and then experimental verification greatly shortens the time for drug combination discovery. Traditional computer methods include high-throughput screening [1], molecular docking, and conventional machine learning (ML) methods like support vector machines and random forests. [2, 3] However, researchers are currently concentrating on deep learning (DL) techniques as these approaches are ineffective when dealing with vast amounts of data. Deep learning is a novel research direction of ML that is utilized neural networks to represent and process data [4]. Due to their network characteristics, deep learning methods are superior for handling larger quantities and multidimensional data. Deep learning methods have been frequently employed in machine translation [5–7], computer vision [8–10], speech recognition [11–13], and text sentiment classification [14–17].

Recently, deep learning has made significant advancements in the prediction of drug synergy. For example, DeepSynergy applies deep neural networks to process drug combinations and cell line information to obtain synergistic scores of drug combinations on cell lines [18]. DeepSynergy was the first deep-learning model for drug synergy prediction, and it was the most outstanding model at that time. After that, more deep synergy models were proposed [19–22]. DTF combined tensor factorization and deep neural networks to predict the synergy of drug combinations [19]. Kim et al. [21] proposed a multi-task deep neural network-based drug combination prediction model that achieved drug synergy prediction for some unstudied tissues by using a transfer learning approach from data-rich to data-poor tissues.

The models previously recommended are primarily dependent on the structural data of drugs, making it simple to ignore the relationship between drug interactions. A graph neural network (GNN) can update its vector using neighboring relations, which should be an excellent application for drug synergy prediction. Numerous notable models, including graph convolution networks (GCNs) [23], graph sample and aggregate (GraphSAGE) [24], and graph attention networks (GATs) [25], have been proposed with the development of graph neural networks. Furthermore, graph neural networks are increasingly being used to predict drug synergy. DeepDDS converts drugs to graphs and is processed by GCN and GAT methods, combined with cell line information to predict drug synergy, which achieved better results compared with other methods [26]. Jin et al. [27] parameterized the drug-target interaction (DTI) to a directional message passing neural network (DMPNN) [28], with bonds as edges and atoms as nodes, combining other structural information to predict drug synergy. Effective drug combinations against COVID-19 were predicted using the model and validated experimentally.

In this article, we proposed GAECDS (Graph Auto Encoder with Convolution neural network for Drug Synergy prediction), using a graph autoencoder combined with a convolutional neural network to predict the synergy of drug combinations. The GAECDS model contains two modules, the GAE module for encoding vectors and the CNN module for predicting synergistic scores. The GAE module applied GCN methods to process and encode the information of drug combinations that used drugs as nodes and the synergistic relation as edges and then decoded them by the matrix factorization method. Furthermore, the latent vectors generated by the encoder during the experiment are input into the CNN to predict synergistic scores. In comparison to support vector machine (SVM), random forest, gradient boost machine (GBM), and extreme gradient boosting (XGBoost), and the previously proposed state-of-the-art deep learning methods, DeepSynergy and DeepDDS, our method significantly outperformed other competitive methods. In addition, GAECDS was employed to predict novel drug combinations, and the experimental verification shows that the model is indeed available, indicating that the GAECDS model will serve as a practical tool in drug synergy prediction.

2 Preface

Before introducing our method, we give the definitions related to our method as follows:

Definition Graph. According to graph theory, a graph is made up of several user-defined points and the line connecting two points. This particular type of graph is typically used to explain a specific relationship between two objects, and the line shows this kind of relationship between the corresponding two objects. The formula for a graph is $\boldsymbol{G=(V, E)}$, where $\boldsymbol{V}$ stands for the set of nodes and $\boldsymbol{E}$ for the set of edges.

Example In this study, we construct a graph of drug synergy, in which drugs are nodes and drug synergy relationships are edges.

3 Data and Methods

3.1 Datasets

Drug combination data were downloaded from DrugComb [29], and we extracted the breast, kidney, lung, and liver data as our dataset. Cancer Cell Line Encyclopedia (CCLE) [30] provided the gene expression information for cancer cell lines, and Transcripts Per Million (TPM), a measurement of gene expression based on the genome-wide read count matrix, was used to normalize the data.

After removing and merging data (the process step was shown in the Online Resource), we obtained 5693 combinations, including 197 drugs and 12 cell lines, and Table 1 shows the data distribution in each cell line. The drug representation data of Simplified Molecular Input Line Entry Specification (SMILES) [31] in this study are obtained through PubChem [32] and converted to Extended Connectivity Fingerprints (ECFP) [33] by RDKit [34]. RDKit counts substructures within the radius by setting a radius from a specific atom and performs operations on the substructures to produce a list of features to form a molecular fingerprint. Different molecular fingerprints and their lengths can be obtained by adjusting the radius and bits. In this study, the radius of the ECFP is set to 6, and the number of bits is 300.

Table 1 The distribution of data in each cell line

Cell line	Drug combinations	Synergistic combination	Antagonistic combinations	Single drug
A549	46	26	20	27
ACHN	1	0	1	2
BT549	5000	1111	3889	119
HCC1187	3	1	2	5
HS578T	76	16	60	60
HUH7	16	5	9	9
KPL1	314	137	177	53
MCF7	26	5	21	32
MDAMB468	175	35	140	21
UO31	3	0	3	6
X786O	1	0	1	2
ZR751	32	8	24	33

3.2 Method Overview

In this part, we describe the GAECDS model's architecture, as depicted in Fig. 1. Module A is the GAE module, which inputs the drug feature and adjacency matrix into the GCN layers to encode the feature. Then, the decoder was applied to decode the feature and reconstruct the drug synergy graph. Module B is the MLP module, which inputs the cell line feature into the network to extract the feature for further prediction. Module C is the CNN module, which combines the feature of the cell line processed by the MLP network and the drug feature processed by the GCN network and inputs the combined feature to the CNN network to predict drug synergistic scores.

3.3 DDS Graph and Feature Matrix

3.3.1 DDS graph

A drug-drug synergy (DDS) graph was constructed with synergistic relations as edges and drugs as nodes. The DDS graph is denoted by $\boldsymbol{G}=(\boldsymbol{V}, \boldsymbol{E})$, where $\boldsymbol{V}$ is the set of drugs, $\boldsymbol{E}$ is the set of synergistic relations, $\boldsymbol{V}$ is denoted by a drug feature matrix $\boldsymbol{F}$, and $\boldsymbol{E}$ is denoted by an adjacency matrix $\mathbf{A}$. The adjacency matrix $\mathbf{A} \in \mathrm{R}^{N*N}$, where N stands for the unduplicated drug's number. If there is a synergic relation between drug i and drug j, $\mathbf{A}(i, j)=1$; otherwise, if drugs i and j have no synergic relation or unknown relation, $\mathbf{A}(i, j)=0$. Moreover, since the DDS graph is undirected, that is, the synergistic relationship between drugs i and j is the same as that between drugs j and i, $\mathbf{A}(i, j)=\mathbf{A}(j, i)$.

3.3.2 Drug Feature Matrix

Drug feature matrix $\boldsymbol{F} \in \mathrm{R}^{N*M}$, where N represents the drug's number and M represents the drug feature's dimension. In this paper, drug features are represented by molecular fingerprints, which are abstract representations of molecules that convert or encode molecules into a series of bits, with one bit on the molecular fingerprint corresponding to a molecular fragment. The typical process is to extract the structural features of the molecule and then hash them to generate bit vectors. We employed the ECFP to represent drug features. It is a circular fingerprint and transforms the drug structure into a string of bits. We obtain the ECFP by RDKit; its radius is 6, and the number of bits is 300. The drug features are arranged in the order of PubChem CID from small to large to form the drug feature matrix $\boldsymbol{F}$.

3.3.3 Cell line Feature Matrix

The cell line data were obtained from Wang et al. [26], and the process was as follows. First, significant genes were

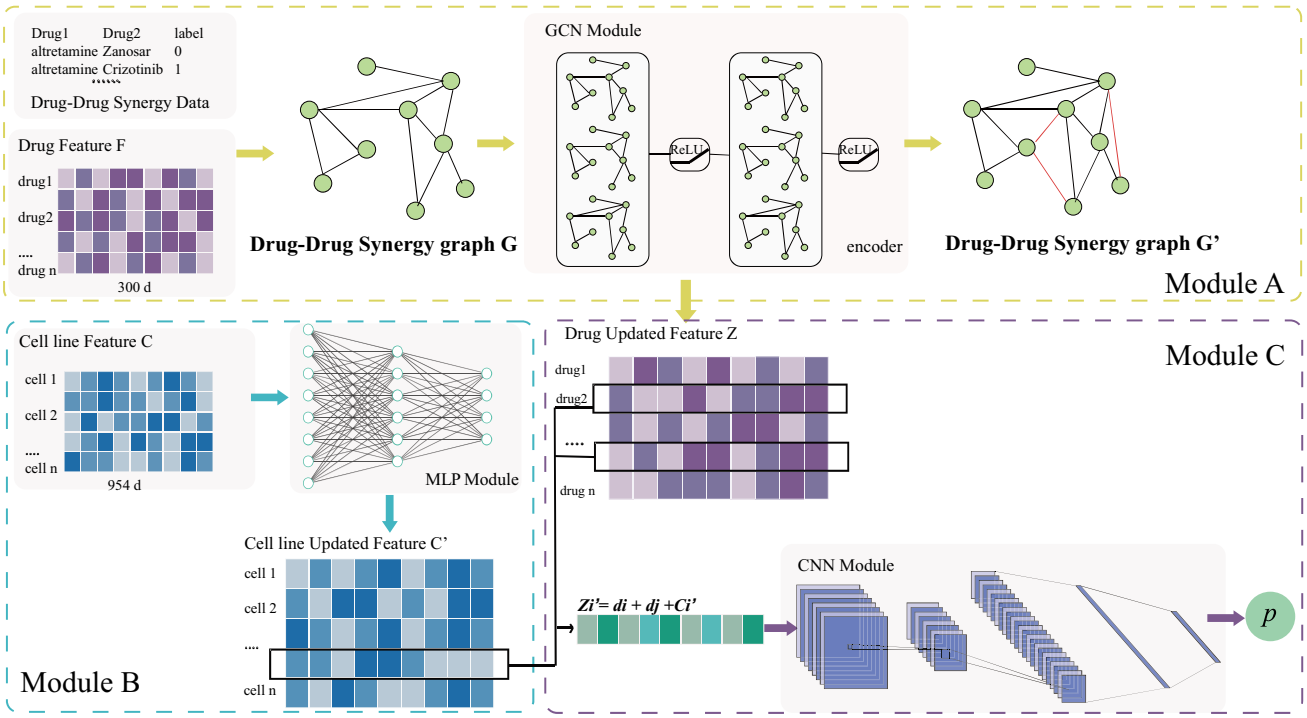


Fig. 1 The overview of GAECDS modules. Module **A**: The graph convolutional neural network was used as the encoder to obtain the latent features, and then decode them to obtain the reconstructed drug synergy graph and the new relationships. Module **B**: Generating the feature representation of cell line by MLP module. Module **C**: Utilizing the CNN module to predict drug synergy by combining drug features and cell line features

collected from the LINCS project [35], and the cross-data of the CCLE dataset and significant genes were obtained [30]. Then, the gene annotation information was downloaded from CCLE and GENCODE [36], obtaining cell line data of 954 dimensions by removing redundant data and noncoding RNA.

The cell line feature matrix $\boldsymbol{C} \in \mathrm{R}^{P*Q}$ was obtained according to the abovementioned methods, where P denotes the cell line's number and Q denotes the cell line feature's dimensions, which are 12 and 954, respectively, in this paper.

We employed the multilayer perceptron (MLP) module to extract the cell line feature data, which contained a three-layer fully connected layer, and the relevant settings of MLP are shown in the Experiment setting section. After the MLP process, we obtain an updated matrix $\boldsymbol{C}^{'} \in \mathrm{R}^{P*Q^{'}}$, where P denotes the cell line's number, $Q^{'}$ denotes the updated cell line feature's dimensions, and $\boldsymbol{C}^{'}{}_{i}$ is the ith cell line feature.

3.4 Reconstruction of the DDS Graph

The adjacency matrix **A** and the drug feature matrix $\boldsymbol{F}$ are the inputs of multilayer GCN [23]. After being calculated and updated by GCN layers, drug feature matrix $\boldsymbol{F}$ is converted to matrix $\boldsymbol{Z}$, in which $\boldsymbol{F} \in \mathrm{R}^{N*M}$ and $\boldsymbol{Z} \in \mathrm{R}^{N*L}$, L stand for the updated latent feature's dimensions, and N stands drug's number. The iteration process can be defined as follows:

$$\boldsymbol{h}^{l+1} = f(\boldsymbol{h}^{l}, \boldsymbol{A}) \tag{1}$$

$$f = \delta \tilde{\boldsymbol{D}}^{-\frac{1}{2}} \tilde{\boldsymbol{A}} \tilde{\boldsymbol{D}}^{-\frac{1}{2}} \boldsymbol{h}^{l} W_{G}^{l} + b_{G}^{l} \tag{2}$$

where $\tilde{\boldsymbol{A}} = \boldsymbol{A} + \boldsymbol{I}$ is the adjacency matrix of the undirected graph with added self-connections, and $\boldsymbol{I}$ is the identity matrix. $\tilde{\boldsymbol{D}} = \sum_{j} \tilde{\boldsymbol{A}}_{i,j}$ is the degree matrix, and the sum of each row of the adjacency matrix $\boldsymbol{A}$ is the degree of each node; h^{l} is the input of the lth layer, and as the premier input, $\boldsymbol{h}^{\boldsymbol{0}}$ is the drug feature matrix $\boldsymbol{F}$, W_{G}^{l} and b_{G}^{l} are the weight matrix and bias matrix of the lth layer, respectively. $\delta(\cdot)$ is the activation function.

According to Kipf et al. [37], the GAE model applied GCN layers as the encoder and factorization as the decoder. For a nonprobabilistic variant of the VGAE model, Kipf et al. [37] calculate embedding Z and the reconstructed adjacency matrix $\hat{A}$ as follows: $\hat{A} = \sigma(Z Z^{T})$; where $Z = \mathrm{GCN}(X; A)$, and $\sigma(\cdot)$ is the activation function. In our method, we used the GAE method to encode and decode the DDS graph. In the process of reconstructing

the DDS graph, two layers of GCN networks are applied as the encoder, and the decoder adopts the method of Kipf et al. [37]. The process of the decoder is as follows:

$$A' = \delta\left(ZZ^T\right) \tag{3}$$

where A' is the reconstructed adjacency matrix, $\delta(\cdot)$ is the activation function, Z is the drug feature matrix after being calculated and updated by GCN layers, and matrix Z^T is the transpose of matrix Z. The reconstructed adjacency matrix $A' \in R^{N*N}$, N stands for the unduplicated drug's number, and the reconstructed adjacency matrix A' can be employed to express the reconstructed graph G'. The reconstructed graph G' contains the original synergy as well as the new synergy.

3.5 Predicting Drug-Drug Synergy

By reconstructing the DDS graph, we obtained the updated drug feature matrix Z, in which $Z \in R^{N*L}$, N represents the drug's number, L represents the updated latent feature's dimension, and Z_i is the updated feature of drug i (d_i), as shown in Fig. 2. We employed the latent feature for drug synergy prediction; similarly, latent vectors can be used to predict other properties. Before predicting the synergistic scores of drugs, we convert drug feature matrix Z to drug combination feature matrix Z', in which matrix Z' is the set that combines the feature of drug i (d_i) and drug j (d_j) from drug combinations and the cell line, and Z_i' is represented as $Z_i' = d_i + d_j + C'_i$. The '+' indicates that the vectors are added together instead of connected or otherwise. After obtaining the drug combination feature matrix Z', the matrix Z' is input into the prediction modules to obtain the synergistic scores of drug combinations.

In this study, we applied the CNN module for drug synergy prediction. The CNN module contains a two-layer CNN layer and a fully connected layer [38, 39], and the predicted score is obtained after using the sigmoid function. The process of the convolution layer is as follows:

$$X^{(k+1)} = \delta\left(X^k W_C^k + b_C^k\right) \tag{4}$$

where X^k is the input feature of the kth layer. For the prediction of drug i and drug j, the initial input feature is the drug combination feature Z_i'. $\delta(\cdot)$ is the activation function, and W_C^k and b_C^k are the weight parameter matrix and bias parameter matrix of the kth layer, respectively.

3.6 Experiment Setting

To evaluate the drug synergy model, we used fivefold cross-validation. The data are specifically separated into five pieces, one of the five sections of the data is chosen every time as the test set, and the remaining data as the train set. The performance of the model is assessed utilizing a five-fold cross-validation average result. In our method, the drug combination data take Loewe scores as the basis to obtain the label, and scores greater than 0 are labeled 1. Otherwise, the label is 0. We set the dimension of each layer in the GAE module to {300, 256, 128}, the dimensions of each layer in the CNN module to {128, 64, 32, 1}, the dimensions of each layer in the MLP of the cell line to {954, 512, 256, 128}, the learning rate of the GAE module, CNN module, and MLP module to {0.01, 0.001, 0.0001,0.00001}, and the activation function to {ReLU, sigmoid, tanh}.

After training, we used some metrics to evaluate the models' performance. True-positive (TP) and true-negative (TN) refer to the amount of correctly detected positive and negative samples, respectively. False-positive (FP) and false-negative (FN) indicate the number of positive and negative samples that were mistakenly recognized, respectively. The true positive rate (TPR) and false-positive rate (FPR) calculated under various thresholds were combined

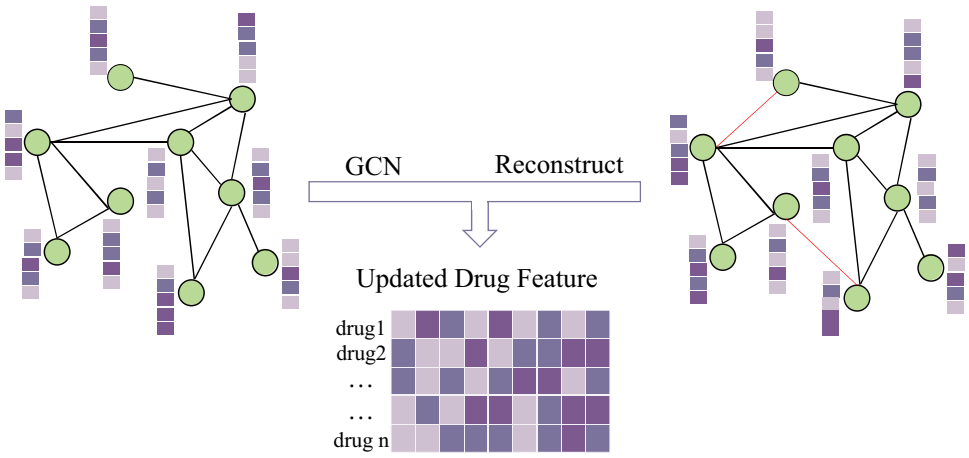


Fig. 2 The process of updating drug features

to form the receiver operating characteristic curve (ROC). These rates are derived as follows:

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}} \tag{5}$$

Another often used evaluation metric is the area under the precision-recall (AUPR) curve, where recall is the percentage of true positive samples accurately detected and precision is the percentage of predicted true positive samples in predicted positive samples. This is shown as follows:

$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{FN+TP}} \tag{7}$$

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{FP+TP}} \tag{8}$$

The accuracy rate is the percentage of predictions in both the positive and negative examples that were correct, and it is represented as follows:

$$\mathrm{Accuracy} = \frac{\mathrm{TP+TN}}{\mathrm{TP+TN+FP+FN}} \tag{9}$$

Moreover, the F1 score is also a useful metric that is used to comprehensively evaluate recall and precision metrics, defined as follows:

$$F1 = 2\frac{\mathrm{Precision} * \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}} = \frac{2\mathrm{TP}^2}{2\mathrm{TP} + \mathrm{FN} + \mathrm{FP}} \tag{10}$$

To evaluate the GAECDS and other models in this work, we employed the accuracy, AUC-ROC, AUPR, precision, recall, and F1 score.

3.7 Loss Function

Both the GAE module and CNN module employed a cross-entropy function to calculate the loss.

In GAE module, we obtained the updated adjacency matrix $\boldsymbol{A}'$ after reconstructing the DDS graph. We calculated the loss of prediction values and real values to adjust the parameters of the network. The loss function is defined as follows:

$$\mathrm{L1} = -\frac{1}{\mathrm{N}}\left(\sum_{\mathrm{i,j}}\left(\boldsymbol{A}_{i,j}\log \boldsymbol{A}_{i,j}^{2''} + \left(1-\boldsymbol{A}_{i,j}\right)\log\left(1-\boldsymbol{A}_{i,j}^{2''}\right)\right)\right) \tag{11}$$

where matrix $\boldsymbol{A}''$ is the matrix $\boldsymbol{A}'$ that only retains the synergy relationship.

In the CNN module, we obtained prediction scores p and calculated the loss between prediction scores p and real values y to adjust the parameters of the network. The loss function of the CNN module is defined as follows:

$$\mathrm{L2} = -\frac{1}{\mathrm{N}}\left(\sum\nolimits_{\mathrm{i,j}}\left(y\log p + (1-y)\log(1-p)\right)\right) \tag{12}$$

In addition, we appeal to adaptive moment estimation (Adam) as an optimizer to minimize both losses [40].

3.8 Cell Experiment

The Cell Bank of the Chinese Academy of Sciences (Shanghai, China) provided the human hepatocellular carcinoma cells (HepG2) and human cervical cancer cells (HeLa) used in the investigations. At a constant temperature of 37 °C and 5% CO_2, the cell lines were grown in complete DMEM (Dulbecco's Modified Eagle Medium) with 10% heat-inactivated fetal bovine serum (FBS, Four Seasons Green) and 1% antibiotics (streptomycin and penicillin combination). Cell passaging was done after the cells had grown to a confluence of about 80 to 95 percent. The drugs used were obtained from MCE.

After 48 h of drug treatment, both alone and in combination, cell proliferation was measured using the MTT method. Trypsin-digested cells in the logarithmic growth phase were inoculated in 96-well plates with roughly 5000 cells and 150 µL of the complete medium in each well, and incubated overnight at 37 °C in a 5% CO2 cell culture incubator. After the cells were plastered, a series of drug combination concentrations were configured and treated with drugs for 48 h. The 96-well plate was removed, 15 µL MTT solution (5 mg/mL) was added to each well, and continuing the incubation for another 4 h. The supernatant in the wells was aspirated, and 150 µL of DMSO was added to each well and shaken for 10 min at room temperature to fully dissolve the crystalline methylzan products. By using an enzyme marker, the 490 nm optical density (OD) values of each well were evaluated. Cell viability is calculated as follows: [(OD of the experimental group—OD of the blank group) / (OD of the control group—OD of the blank group)] × 100%.

4 Results and Analysis

4.1 Comparison with Other Methods

We examined a variety of superior methods, including several ML mosels and DL models, to assess the performance of GAECDS. DeepSynergy and DeepDDS are the DL models for comparison. DeepSynergy was based on a fully connected layer of conical structures, with drug combinations and cell lines information as input, to predict the synergistic scores of drug combinations on cell lines [18]. DeepDDS employs two types of GNNs, GATs and GCNs to obtain features of drugs, and the features of cancer cells are processed by the MLP model and connect drug combination features

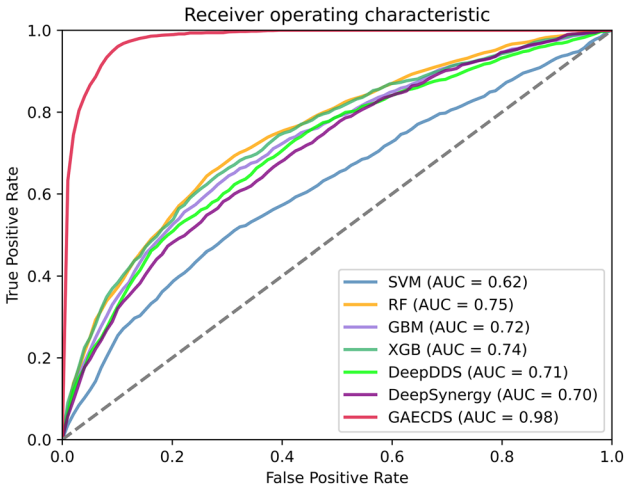


Fig. 3 The ROC curves of all methods

and cell line features into fully connected layers for drug synergy prediction [26]. Figure 3 shows the AUC values of DeepSynergy, DeepDDS, support vector machine (SVM), random forest (RF), gradient boost machine (GBM), extreme gradient boost (XGB), and our method GAECDS.

We used fivefold cross-validation to obtain the AUC values, as shown in Fig. 3. The AUC value of GAECDS, DeepDDS, and DeepSynergy are 0.98, 0.71, and 0.70, respectively. The AUC values of SVM, RF, GBM, and XGB are 0.62, 0.75, 0.72, and 0.74, respectively. Figure 3 shows that GAECDS has superior performance, and RF is the second best. Table 2 shows the results of the evaluation. The model's performance is further indicated by the AUPR value, which for the GAECDS model is 0.93, higher than that of the models [DeepDDS (0.44), DeepSynergy (0.43), SVM (0.62), RF (0.53), GBM (0.48), XGB (0.53)]. Furthermore, as shown in Table 2, the accuracy, F1 score, recall, and precision of the GAECDS model were 0.87, 0.77, 0.83, and 0.78, respectively, superior to other models. We found that GAECDS, when compared with other models, has optimal performance and is a promising model.

The GAECDS model is superior for several reasons. First, compared to directly converting drugs into graphs to extract features, the GAECDS model uses drugs as nodes and synergistic relationships as edges, taking the relationships between different drugs into account, making feature extraction more comprehensive. Second, the drug synergy graph is reconstructed using the GAE module, the latent vectors generated under this process are used again for prediction, and the prediction performance is well-improved in this cycle of reconstruction prediction. Furthermore, when reconstructing the DDS graph, 197 drugs were combined in pairs, and there were 19,306 interaction relationships. Except for the known 5,693 pairs of synergistic combinations, there were 13,613 pairs of nonsynergistic or unknown relationships. When reconstructing the DDS graph, to avoid the influence of negative examples on the accuracy rate, we mainly calculate the accuracy of correctly identifying 5693 pairs of synergistic combinations. We set the threshold for labels to be identified as positive to be 0.9, and the accuracy of the reconstructed map is 0.99. Moreover, some new synergistic relationships were generated in the reconstructed DDS graph $\boldsymbol{G}'$.

4.2 Model Stability Testing

To test the stability of the model and avoid data leakage from two modules in the model from having an impact on the results, we tested the GAE module (Fig. 1 Module A) and the CNN module (Fig. 1 Module C) of GAECDS separately. As shown in Fig. 4, first, we employed part of the data for training and the rest for validation, and the training set to validation set ratio is 8:2. During training, as shown in the blue route in Fig. 4, only the labels of the training data are used to calculate the loss, and the model adjusts the weights in light of the training set's data, at which time the labels of test data has not been exposed to the model while all vectors are updated. Second, as shown in the yellow route in Fig. 4, for the CNN module, we use the data not exposed to labels in the GAE test as the validation set, the training set in the GAE module is still used as the training set of the CNN module, and the CNN module is trained to make predictions on the validation set directly. Since the training and validation sets are completely separated, resulting in some combinations of single drugs in the training and testing sets, respectively, not participating in the experiments, the final number of training sets in this module is 3377, and the

Table 2 The results of GAECDS and comparison models

Method	GAECDS	DeepDDS	DeepSynergy	SVM	RF	GBM	XGB
Accuracy	0.87	0.76	0.77	0.76	0.78	0.78	0.79
AUC	0.98	0.71	0.70	0.62	0.75	0.72	0.74
AUPR	0.93	0.44	0.43	0.62	0.53	0.48	0.53
Recall	0.83	0.33	0.13	0.01	0.21	0.16	0.34
Precision	0.78	0.50	0.58	0.41	0.62	0.60	0.58
F1 score	0.77	0.40	0.21	0.03	0.31	0.25	0.43

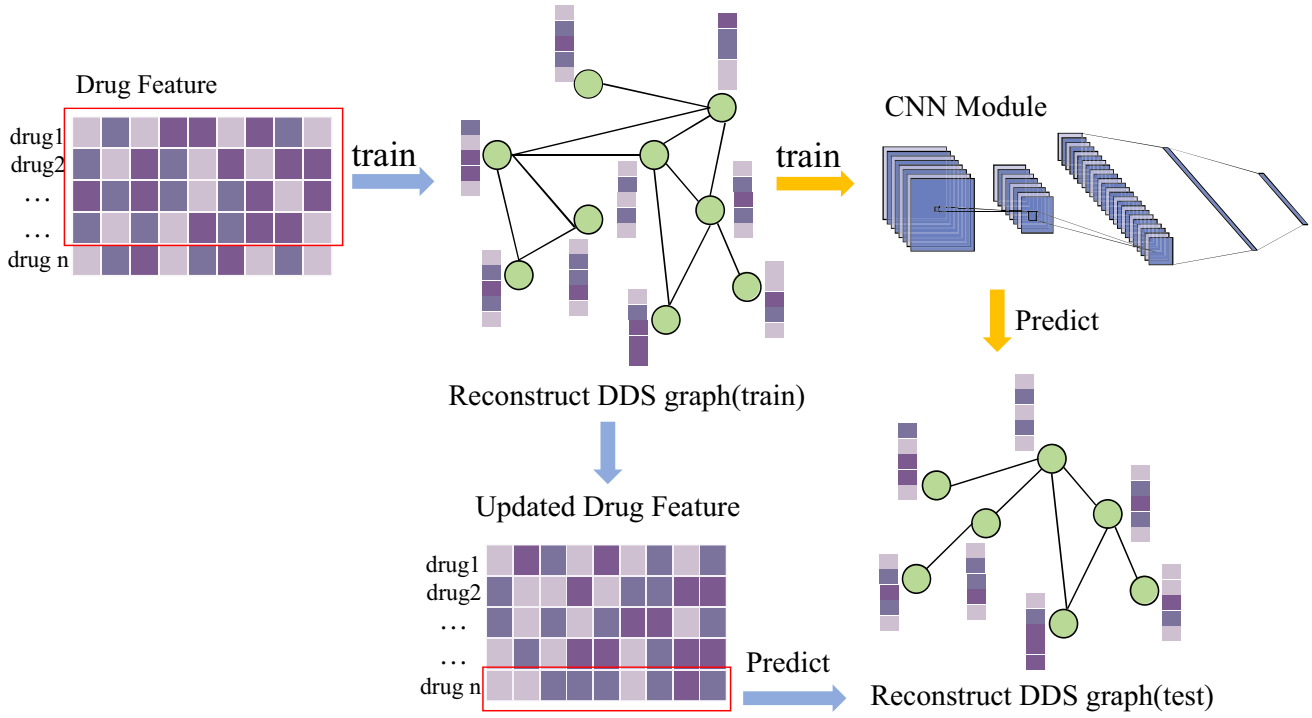


Fig. 4 The process of verifying the stability of GAECDS

Table 3 Results of the validation of GAE stability

	Accuracy	AUC	AUPR	Recall	Precision	F1 score
GAE_test	0.99	0.99	0.99	0.99	0.99	0.99
CNN_test	0.68	0.71	0.47	0.68	0.48	0.56

number of validation sets is 279. The test results of the GAE module and CNN module are shown in Table 3.

As shown in Table 3, the accuracy and AUC value of the GAE test were both 0.99, and since we calculate the accuracy and AUC value only for the synergistic combination, it shows that the model can accurately predict synergistic combinations without touching the label. The accuracy and AUC values in the CNN test were 0.68 and 0.71, respectively. The best results of accuracy and AUC value for the classification task in the independent test set of DeepDDS in a previous study were 0.64 and 0.67, respectively [26]. The CNN test's validation set and training set have exactly no overlap, including the overlap of drug combinations and the overlap of single drugs, we believe that this result indicates that GAECDS has a strong generalization ability on completely new combinations.

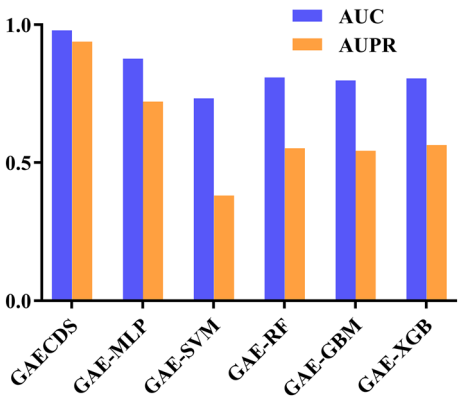


Fig. 5 AUC and AUPR values with different prediction structures

4.3 Ablation Study

For the GAECDS model, different network structures used in the prediction part have different influences on the model. In this section, we compared several different structures, such as the SVM, RF, GBM, XGB, and MLP models, with the current CNN structure to evaluate the best structure, and Fig. 5 shows the test results.

Table 4 Comparison of results after applying GAE to process data

Method	GAECDS	GAE-MLP	GAE-SVM	GAE-RF	GAE-GBM	GAE-XGB
Accuracy	0.87	0.77	0.77	0.80	0.80	0.80
AUC	0.98	0.88	0.73	0.81	0.80	0.81
AUPR	0.93	0.72	0.38	0.55	0.54	0.56
Recall	0.83	0.50	0.08	0.32	0.33	0.38
Precision	0.78	0.75	0.58	0.65	0.63	0.62
F1 score	0.77	0.60	0.14	0.42	0.43	0.47

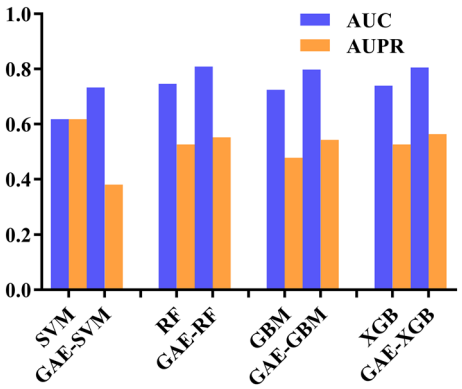


Fig. 6 The machine learning method's AUC and AUPR values before and after applying the GAE module

Figure 5 shows that the GAECDS model with the CNN structure has the best results, with higher AUC and AUPR values than the model with other structures, and the model with the MLP structure provided the second-best performance. We employed the GAE module as a data processing module and replaced the prediction module to evaluate the performance of each structure, as shown in Table 2. Using the GAE module for data processing before prediction can considerably improve model performance.

Tables 2 and 4 show that the AUC values of SVM, RF, GBM, and XGB before using GAE to process the data are 0.62 0.75, 0.72, and 0.74, respectively. In contrast, the AUC values after using the GAE module to process the data are 0.73, 0.81, 0.80, and 0.81, respectively, indicating a considerable increase in the AUC values.

Figure 6 demonstrates the comparison results of AUC and AUPR of the machine learning method before and after applying GAE. It is observed that the effect of data processing with GAE is significantly better than unprocessed data, indicating that the GAE module can truly improve the model's performance.

Moreover, some drug pairs may be deficient in cell line features; as a result, we evaluated GAECDS' performance without considering cell line features. In this section, the data for testing GAECDS with and without cell lines were both 5693 groups, and the difference between the two sets of data was the presence or absence of cell line data. The AUC values of the GAECDS models with and without cell line information were 0.99 and 0.99, respectively, as shown in Table 5. According to Table 5, when compared to the model utilizing cell line features, the GAECDS model without cell line features was not significantly different except for the recall and F1 score. The recall of the model with and without cell line features is 0.79 and 0.88, respectively. The recall without cell line data was higher because when the model data were tuned to additional cell line features, more parameters were used to adjust the model and the model fit data more easily, but there was an overfitting problem. When we trained the model, the overfitting problem was unavoidable, and when the data volume was high, the overfitting problems could be mitigated. Therefore, we believe that the high data volume of the model with cell line data will perform better

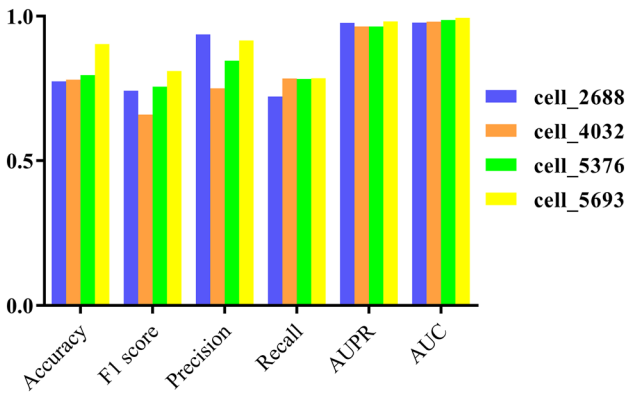


Fig. 7 Results of GAECDS with cell line data under multiple datasets

Table 5 Results of models with and without cell lines

Metric	Accuracy	AUC	AUPR	Recall	Precision	F1 score
GAECDS_cell	0.90	0.99	0.98	0.79	0.92	0.81
GAECDS_nocell	0.92	0.99	0.98	0.88	0.91	0.88

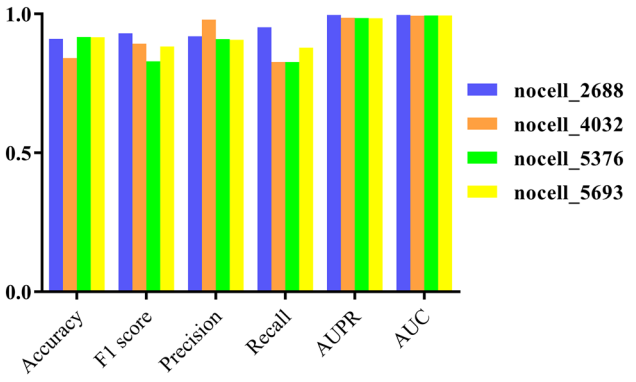


Fig. 8 Results of GAECDS without cell line data under multiple datasets

Table 6 Hyperparameter adjustment of GAECDS

Variable name	Variables			
Learning rate(GCN)	0.01	0.001	0.0001	0.00001
Learning rate(CNN)	0.01	0.001	0.0001	0.00001
Learning rate(MLP)	0.01	0.001	0.0001	0.00001
Activation function(GCN)	ReLU	tanh	sigmoid	
Activation function(CNN)	ReLU	tanh	sigmoid	
Activation function(MLP)	ReLU	tanh	sigmoid	
Vector dimensions(latent)	32	64	128	256

than the model without cell line data because it has more tunable parameters.

To further verify our conjecture, we performed validation of the GAECDS under multiple datasets, the results with and without cell lines data are shown in Fig. 7 and Fig. 8, respectively. As shown in Figs. 7 and 8, with increasing data, the overall effectiveness of the model with cell line data increases, while the overall effectiveness of the model without cell line data remains relatively constant, this shows that the model with cell line data performs better when data volume increases. Furthermore, the addition of cell line data makes the model interpretable and gives the model better prospects.

4.4 Parameter Adjustment

This section looks into the effects of various parameters. All parameters that were tested are shown in Table 6.

Figure 9 shows the results of learning rate adjustment, the best effect of the GAE module, CNN module, and MLP model when learning rates were 0.00001, 0.00001, and 0.01, respectively. The activation function in neural networks is used to increase the network's nonlinearity, and different activation

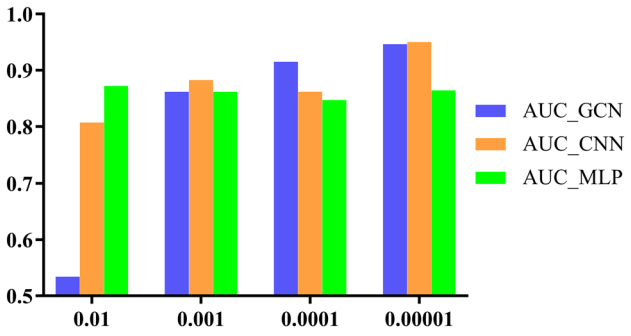


Fig. 9 Average AUC value for different learning rates

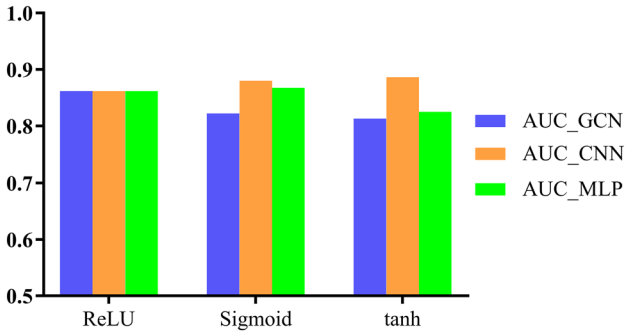


Fig. 10 Average AUC value for different activation function

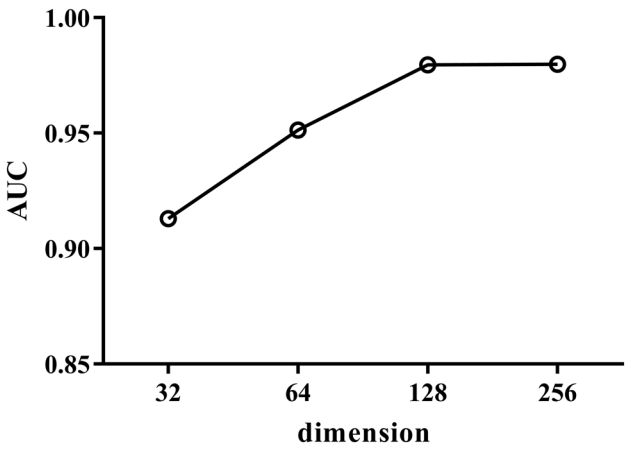


Fig. 11 Average AUC value for the different dimensions of latent vector

functions lead to different effects on the network. The applied activation functions are shown as follows: Fig. 10

$$\mathrm{ReLU} = \max(0, x) \tag{13}$$

$$\tanh = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \tag{14}$$

Table 7 The influence of the different proportions of negative samples on GAECDS

Ratio	Positive samples size	Negative samples size	AUC
1:1	1344	1344	0.99
1:2	1344	2688	0.99
1:3	1344	4032	0.99

$$\text{sigmoid} = \frac{1}{1+e^{-x}} \tag{15}$$

Figure 11 shows the adjustment results of the activation function. The tanh function makes the CNN module work the best, the ReLU function makes the GCN module work the best, and the sigmoid function makes the MLP module work the best. To further analyze the influence of the dimensions of the latent vectors on the model, we tested latent vectors of different dimensions, and the obtained results are shown in Fig. 11, indicating that the latent vector dimension of 128 is optimal for the model's performance.

4.5 Analysis of Negative Sampling

Our data were imbalanced, with a disproportionately large number of negative samples compared to tiny percentages of positive samples. Most datasets have a comparatively high number of negative samples. We train the model using various data ratios in this section to assess the effects of negative examples on the model, as shown in Table 7. Table 7 demonstrates that there was no discernible difference in the AUC values for various negative data proportions, indicating that the model effect is relatively stable and not affected by the distribution of model data.

4.6 Case Studies

Case studies are performed to confirm the capabilities of GAECDS. GAECDS was applied to predict the drugs that we are currently focusing on. The cell line data were not used in this module due to missing cell line data.

The dataset has three sources:

1. The first source was derived from calculations in a publication [41], which suggested that drugs sharing a single target may have synergistic effects, and the probability of a large number of drugs sharing a single target was calculated and experimentally validated.
2. The second source of data came from drug combinations approved by the US Food and Drug Administration (FDA) [42], which included skin, cardiovascular, tumor, and other categories.
3. Third, the drug combinations were confirmed by laboratory experiments and reported in relevant publications [43–45].

Based on the results from these sources, we obtained 1390 drug combinations, including 679 single drugs in the dataset. The dataset contains 1321 sets of data for training and testing and 69 sets of data for prediction. For 713 pairs of drugs determined to be synergistic, the label is 1, and for 608 pairs of drugs determined to be nonsynergistic, the label is 0. The labels of 69 pairs of drugs with unknown relationships were set to 0 when constructing the DDS graph.

The GAECDS model was used to make multiple predictions and take the top 10 of the overlapping parts of the predictions for experimental verification. Table 8 summarizes the prediction results.

Based on the combinations shown in Table 8, we have experimentally verified the combinations in the table, and the results and analysis are as follows.

Analysis was performed using the ChoueTalalay method and CalcuSyn software (version 2, Biosoft, Cambridge, UK). The combination index (CI) was used to quantify interactions and determine whether they had antagonistic (CI > 1), additive (CI = 1), or synergistic (CI < 1) effects. Graphing data Combination Index values were obtained with GraphPad Prism 7. The relationship between dose and inhibition rate is represented by a heatmap, which was plotted using Python. Draw a heatmap with the dose-inhibition rate; the text in the heatmap box indicates the inhibition rate, and the unit of the dose is μM. The heatmaps of the dose and inhibition rate and combination index plot for several drugs are shown in Figs. 12, 13, 14, and 15. As shown in these figures, the inhibition rates of the drug combinations were higher than those of single drugs, and most synergistic combinations had a CI value of less than 1, indicating that the drug combinations were indeed effective.

Table 8 The top 10 repeated parts of the prediction result

Drug1	PubChem id	Drug2	PubChem id	Evidence
Propranolol	4946	Mitoxantrone	4212	–
Methylprednisolone	6741	Mitoxantrone	4212	Verified
Flumethasone	16,490	Mitoxantrone	4212	Verified
Valsartan	60,846	Mitoxantrone	4212	–
Enzalutamide	15,951,529	Mitoxantrone	4212	–
Capivasertib	25,227,436	Mitoxantrone	4212	Verified
Flumethasone	16,490	Daunomycin	30,323	–
Flumethasone	16,490	Idarubicin	42,890	–
Flumethasone	16,490	Fludarabine	657,237	–
Mitoxantrone	4212	Trametinib	11,707,110	Verified

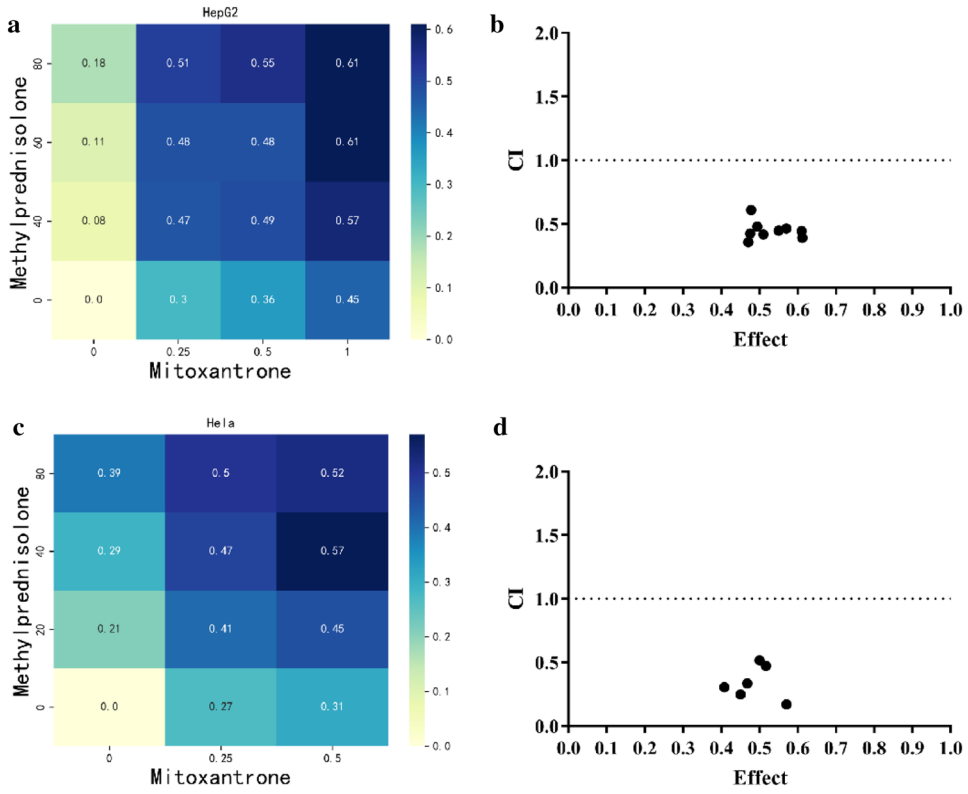


Fig. 12 The heatmaps of inhibition rate and combination index plot for Mitoxantrone and Methylprednisolone. **a** Heatmap of dose-inhibition rate of Mitoxantrone and Methylprednisolone in HepG2. **b** Combination index plot of Mitoxantrone and Methylprednisolone in HepG2. **c** Heatmap of the dose-inhibition rate of Mitoxantrone and Methylprednisolone in HeLa. **d** Combination index plot of Mitoxantrone and Methylprednisolone in HeLa

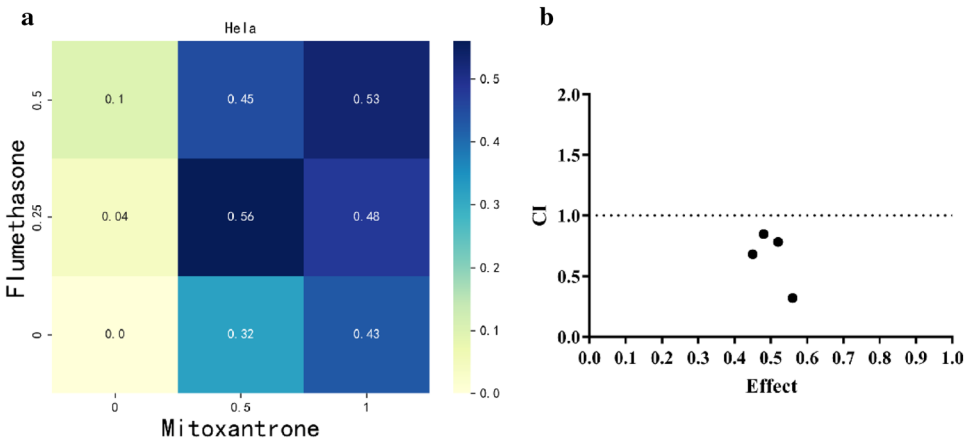


Fig. 13 The heatmaps of inhibition rate and combination index plot for Mitoxantrone and Flumethasone. **a** Heatmap of the dose-inhibition rate of Mitoxantrone and Flumethasone in HeLa. **b** Combination index plot of Mitoxantrone and Flumethasone in HeLa

5 Conclusion

In conclusion, we introduced GAECDS, a novel approach to drug synergy prediction that combines the GAE module for data processing with the CNN module to predict the synergistic scores of drug combinations. GAECDS outperforms machine learning methods (SVM, RandomForest, GBM, XGBoost) and deep learning methods (DeepDDS, DeepSynergy). Furthermore, we tested several variants of the model using GAE with various structures, and the obtained results show that the effect of data processed by GAE is significantly improved, indicating that utilizing GAE as a data processing module is promising. For data test effect models, we first tested the effect of the model with and without cell line features, and the AUC values were 0.99 and 0.99, respectively. Second, we tested the effect of different proportions of positive and negative data on the model, and the AUC values for each ratio (1:1, 1:2, 1:3) were 0.99, 0.99, and 0.99, respectively, demonstrating that the GAECDS has good stability. Moreover, we applied GAECDS to predict drug synergy in a novel dataset and verified by experiments that there are four drug combinations (Mitoxantrone and Methylprednisolone,

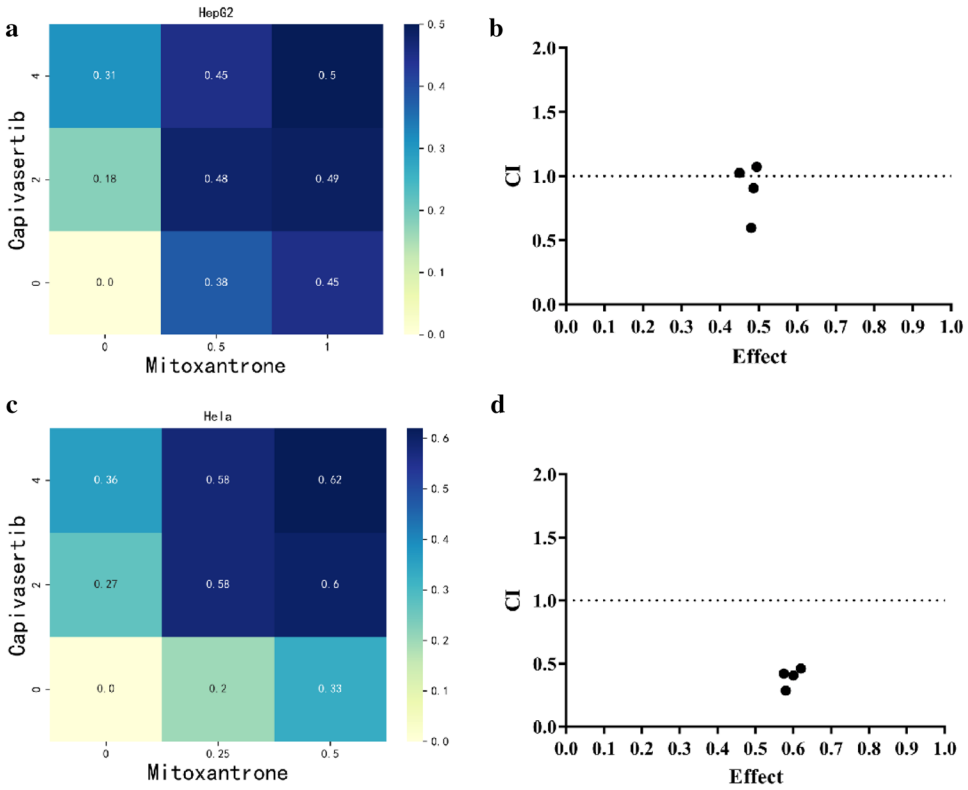


Fig. 14 The heatmaps of inhibition rate and combination index plot for Mitoxantrone and Capivasertib. **a** Heatmap of the dose-inhibition rate of Mitoxantrone and Capivasertib in HepG2. **b** Combination index plot of Mitoxantrone and Capivasertib in HepG2. **c** Heatmap of the dose-inhibition rate of Mitoxantrone and Capivasertib in HeLa. **d** Combination index plot of Mitoxantrone and Capivasertib in HeLa

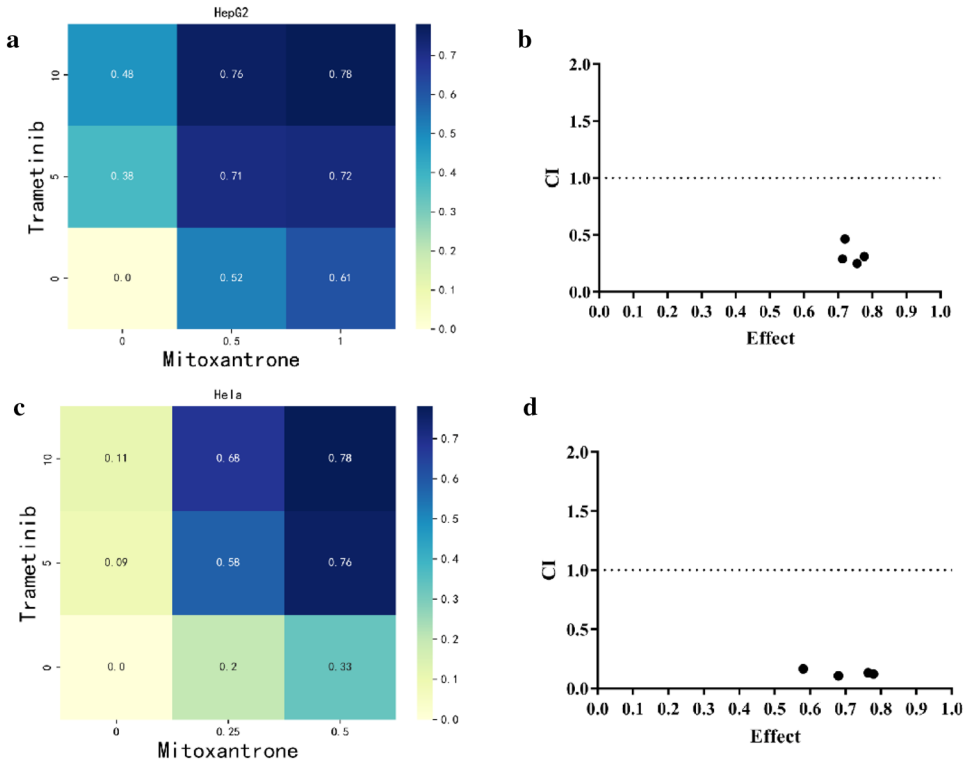


Fig. 15 The heatmaps of inhibition rate and combination index plot for Mitoxantrone and Trametinib. **a** Heatmap of the dose-inhibition rate of Mitoxantrone and Trametinib in HepG2. **b** Combination index plot of Mitoxantrone and Trametinib in HepG2. **c** Heatmap of the dose-inhibition rate of Mitoxantrone and Trametinib in HeLa. **d** Combination index plot of Mitoxantrone and Trametinib in HeLa

Mitoxantrone and Flumethasone, Mitoxantrone and Capivasertib, Mitoxantrone and Trametinib) that have been experimentally validated to have a synergistic effect with CI values less than 1, suggesting that GAECDS could be an effective method for drug combination prediction. Furthermore, our GAECDS model with and without cell line data performed better when we compared it with other methods, indicating that GAECDS is a promising model.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12539-023-00558-y.

Acknowledgements This work was financially supported by the National Natural Science Foundation of China (No. 22078073), the Guangxi Innovation-Driven Development Special Fund Project (GUANGXI AA18242040), and the Guangxi key research and development program (GUANGXI AB18221121). The work was supported by the Guangxi Key Laboratory of Traditional Chinese Medicine Quality Standards (Guangxi Institute of Traditional Medical and Pharmaceutical Sciences) (guizhongzhongkai201703) and the foundation of the Key Laboratory of Trusted Software (No. kx201703).

Data availability https://github.com/junelyemm/GAECDS.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests and Funding or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Yan X, Yang Y, Chen Z, Yin Z, Deng Z, Qiu T, Tang K, Cao Z (2020) H-RACS: a handy tool to rank anti-cancer synergistic drugs. Aging-US 12(21):21504–21517. https://doi.org/10.18632/aging.103925
2. Cuvitoglu A, Zhou JX, Huang S, Isik Z (2019) Predicting drug synergy for precision medicine using network biology and machine learning. J Bioinform Comput Biol 17(2):1950012. https://doi.org/10.1142/S0219720019500124
3. Wildenhain J, Spitzer M, Dolma S, Jarvik N, White R, Roy M, Griffiths E, Bellows DS, Wright GD, Tyers M (2015) Prediction of synergism from chemical-genetic interactions by machine learning. Cell Syst 1(6):383–395. https://doi.org/10.1016/j.cels.2015.12.003
4. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444. https://doi.org/10.1038/nature14539
5. Ali MNY, Rahman ML, Chaki J, Dey N, Santosh KC (2021) Machine translation using deep learning for universal networking language based on their structure. Int J Mach Learn Cybern 12(8):2365–2376. https://doi.org/10.1007/s13042-021-01317-5
6. Popel M, Tomkova M, Tomek J, Kaiser L, Uszkoreit J, Bojar O and Zabokrtsky Z (2020) Transforming machine translation: a deep learning system reaches news translation quality comparable to human professionals. Nat Commun 11(1): 4381 https://www.ncbi.nlm.nih.gov/pubmed/32873773
7. Xu T, Chen W, Zhou J, Dai J, Li Y and Zhao Y (2020) Neural machine translation of chemical nomenclature between English and Chinese. J Cheminform 12 (1): 50. https://www.ncbi.nlm.nih.gov/pubmed/33431023
8. Esteva A, Chou K, Yeung S, Naik N, Madani A, Mottaghi A, Liu Y, Topol E, Dean J and Socher R (2021) Deep learning-enabled medical computer vision. NPJ Digit Med 4 (1): 5. https://www.ncbi.nlm.nih.gov/pubmed/33420381
9. Venkateswara H, Chakraborty S, Panchanathan S (2017) Deep-learning systems for domain adaptation in computer vision: learning transferable feature representations. IEEE Signal Process Mag 34(6):117–129. https://doi.org/10.1109/msp.2017.2740460
10. Voulodimos A, Doulamis N, Doulamis A and Protopapadakis E (2018) Deep Learning for Computer Vision: A Brief Review. Comput Intell Neurosci 2018: 7068349. https://www.ncbi.nlm.nih.gov/pubmed/29487619
11. Zhang ZX, Geiger J, Pohjalainen J, Mousa AED, Jin WY, Schuller B (2018) Deep learning for environmentally robust speech recognition: an overview of recent developments. ACM Trans Intell Syst Technol 9(5):1–28. https://doi.org/10.1145/3178115
12. Purwins H, Li B, Virtanen T, Schluter J, Chang S-Y, Sainath T (2019) Deep learning for audio signal processing. IEEE J Selected Topics Signal Processing 13(2):206–219. https://doi.org/10.1109/jstsp.2019.2908700
13. Zhang Z, Geiger J, Pohjalainen J, Mousa AE-D, Jin W, Schuller B (2018) Deep learning for environmentally robust speech recognition. ACM Trans Intel Sys Tech 9(5):1–28. https://doi.org/10.1145/3178115
14. Onan A (2022) Bidirectional convolutional recurrent neural network architecture with group-wise enhancement mechanism for text sentiment classification. J King Saud Univ-Com 34(5):2098–2117. https://doi.org/10.1016/j.jksuci.2022.02.025
15. Onan A, Korukoğlu S (2016) A feature selection model based on genetic rank aggregation for text sentiment classification. J Inform Sci 43(1):25–38. https://doi.org/10.1177/0165551515613226
16. Onan A, Korukoğlu S, Bulut H (2016) Ensemble of keyword extraction methods and classifiers in text classification. Expert Syst Appl 57:232–247. https://doi.org/10.1016/j.eswa.2016.03.045
17. Onan A, Tocoglu MA (2021) A Term weighted neural language model and stacked bidirectional LSTM based framework for sarcasm identification. IEEE Access 9:7701–7722. https://doi.org/10.1109/access.2021.3049734
18. Kristina P, Richard L, Sepp H, Andreas B, Krishna CB, Günter K (2018) DeepSynergy: predicting anti-cancer drug synergy with deep learning. Bioinformatics (Oxford, England) 34(9):1538–1546. https://doi.org/10.1093/bioinformatics/btx806
19. Zexuan S, Shujun H, Peiran J, Pingzhao H (2020) DTF: deep tensor factorization for predicting anticancer drug synergy. Bioinformatics (Oxford, England) 36(16):4483–4489. https://doi.org/10.1093/bioinformatics/btaa287
20. Kuenzi BM, Park J, Fong SH, Sanchez KS, Lee J, Kreisberg JF, Ma J, Ideker T (2020) Predicting drug response and synergy using a deep learning model of human cancer cells. Cancer Cell 38(5):672-684.e6. https://doi.org/10.1016/j.ccell.2020.09.014
21. Kim Y, Zheng S, Tang J, Jim Zheng W, Li Z, Jiang X (2021) Anticancer drug synergy prediction in understudied tissues using transfer learning. J Am Med Inform Assoc 28(1):42–51. https://doi.org/10.1093/jamia/ocaa212
22. Liu Q, Xie L (2021) TranSynergy: Mechanism-driven interpretable deep neural network for the synergistic prediction and pathway deconvolution of drug combinations. PLoS Comput Biol 17(2):e1008653. https://doi.org/10.1371/journal.pcbi.1008653
23. Kipf TN and Welling M (2016) Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907. https://arxiv.org/abs/1609.02907
24. Hamilton WL, Ying R and Leskovec J (2017) Inductive representation learning on large graphs. arXiv preprint arXiv:1706.02216. https://arxiv.org/abs/1706.02216
25. Veličković P, Cucurull G, Casanova A, Romero A, Lio P and Bengio Y (2017) Graph attention networks. arXiv preprint arXiv:1710.10903. https://arxiv.org/abs/1710.10903

26. Wang J, Liu X, Shen S, Deng L and Liu H (2021) DeepDDS: deep graph neural network with attention mechanism to predict synergistic drug combinations. Brief Bioinform 23 (1). https://www.ncbi.nlm.nih.gov/pubmed/34571537
27. Wengong J, Jonathan MS, Richard TE, Zina I, Alexey VZ, James JC, Tommi SJ, Regina B (2021) Deep learning identifies synergistic drug combinations for treating COVID-19. P Natl Acad Sci 118(39):e2105070118. https://doi.org/10.1073/pnas.2105070118
28. Yang K, Swanson K, Jin W, Coley C, Eiden P, Gao H, Guzman-Perez A, Hopper T, Kelley B, Mathea M, Palmer A, Settels V, Jaakkola T, Jensen K, Barzilay R (2019) Analyzing learned molecular representations for property prediction. J Chem Inf Model 59(8):3370–3388. https://doi.org/10.1021/acs.jcim.9b00237
29. Zheng S, Aldahdooh J, Shadbahr T, Wang Y, Aldahdooh D, Bao J, Wang W and Tang J (2021) DrugComb update: a more comprehensive drug sensitivity data repository and analysis portal. Nucleic Acids Res 49(W1): W174-W184. https://www.ncbi.nlm.nih.gov/pubmed/34060634
30. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, Wilson CJ, Lehár J, Kryukov GV, Sonkin D, Reddy A, Liu M, Murray L, Berger MF, Monahan JE, Morais P, Meltzer J, Korejwa A, Jané-Valbuena J, Mapa FA, Thibault J, Bric-Furlong E, Raman P, Shipway A, Engels IH, Cheng J, Yu GK, Yu J, Aspesi P, de Silva M, Jagtap K, Jones MD, Wang L, Hatton C, Palescandolo E, Gupta S, Mahan S, Sougnez C, Onofrio RC, Liefeld T, MacConaill L, Winckler W, Reich M, Li N, Mesirov JP, Gabriel SB, Getz G, Ardlie K, Chan V, Myer VE, Weber BL, Porter J, Warmuth M, Finan P, Harris JL, Meyerson M, Golub TR, Morrissey MP, Sellers WR, Schlegel R, Garraway LA (2012) The cancer cell line encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature 483(7391):603–607. https://doi.org/10.1038/nature11003
31. Weininger D (1988) SMILES, a chemical language and information system 1 introduction to methodology and encoding rules. J Chem Inf Comput Sci 28(1):31–36. https://doi.org/10.1021/ci00057a005
32. Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, Li Q, Shoemaker BA, Thiessen PA, Yu B, Zaslavsky L, Zhang J, Bolton EE (2020) PubChem in 2021: new data content and improved web interfaces. Nucleic Acids Res 49(D1):D1388–D1395. https://doi.org/10.1093/nar/gkaa971
33. Rogers D, Hahn M (2010) Extended-connectivity fingerprints. J Chem Inf Model 50(5):742–754. https://doi.org/10.1021/ci100050t
34. Landrum. (2010) "RDKit: Open-source cheminformatics. Release 2014.03.1." from https://www.rdkit.org
35. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, Bindal N, Beare D, Smith JA, Thompson IR, Ramaswamy S, Futreal PA, Haber DA, Stratton MR, Benes C, McDermott U and Garnett MJ (2013) Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Res 41(Database issue): https://doi.org/10.1093/nar/gks1111
36. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A, Knowles DG, Lagarde J, Veeravalli L, Ruan X, Ruan Y, Lassmann T, Carninci P, Brown JB, Lipovich L, Gonzalez JM, Thomas M, Davis CA, Shiekhattar R, Gingeras TR, Hubbard TJ, Notredame C, Harrow J, Guigó R (2012) The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res 22(9):1775–1789. https://doi.org/10.1101/gr.132159.111
37. Kipf TN and Welling M (2016) Variational Graph Auto-Encoders. arXiv preprint arXiv:1611.07308. https://arxiv.org/abs/1611.07308v1
38. LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. P IEEE 86(11):2278–2324. https://doi.org/10.1109/5.726791
39. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. Proceed IEEE Conference Computer Vision Pattern Recognition. https://doi.org/10.1109/CVPR.2015.7298594
40. Kingma DP and Ba JL (2014) Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980 [v4]. https://doi.org/10.48550/arXiv.1412.6980
41. Campillos M, Kuhn M, Gavin AC, Jensen LJ and Bork P (2008) Drug target identification using side-effect similarity. Science 321(5886): 263–266. https://www.ncbi.nlm.nih.gov/pubmed/18621671
42. Das P, Delost MD, Qureshi MH, Smith DT, Njardarson JT (2019) A Survey of the structures of US FDA approved combination drugs. J Med Chem 62(9):4265–4311. https://doi.org/10.1021/acs.jmedchem.8b01610
43. Kano Y, Suzuki K, Akutsu M, Suda K (1992) Effects of mitoxantrone in combination with other anticancer agents on a human leukemia cell line. Leukemia 6(5):440–445. https://doi.org/10.1002/hon.2900100314
44. Manuela R, De Michele S, Paola B, Monica A, Alice B, Shoeb A, Ornella A, Eleonora F, Armando C, Roberta C, Marco V (2021) A Phase I dose escalation study of oxaliplatin, cisplatin and doxorubicin applied as PIPAC in patients with peritoneal carcinomatosis. Cancers 13(5):1060. https://doi.org/10.3390/cancers13051060
45. Kano Y, Akutsu M, Tsunoda S, Suzuki K, Ichikawa A, Furukawa Y, Bai L, Kon K (2000) In vitro cytotoxic effects of fludarabine (2-F-ara-A) in combination with commonly used antileukemic agents by isobologram analysis. Leukemia. https://doi.org/10.1038/sj.leu.2401684

Authors and Affiliations

Huijun Li[1] · Lin Zou[1] · Jamal A. H. Kowah[2] · Dongqiong He[2] · Lisheng Wang[1] · Mingqing Yuan[1] · Xu Liu[1]

[1] School of Medicine, Guangxi University, Nanning 530004, China

[2] School of Chemistry and Chemical Engineering, Guangxi University, Nanning 530004, China