Unlocking the therapeutic potential of drug combinations through synergy prediction using graph transformer networks

Waleed Alam [a], Hilal Tayara [b,*], Kil To Chong [a,c,*]

[a] *Department of Electronics and Information Engineering, Jeonbuk National University, Jeonju, 54896, South Korea*
[b] *School of International Engineering and Science, Jeonbuk National University, Jeonju, 54896, South Korea*
[c] *Advanced Electronics and Information Research Center, Jeonbuk National University, Jeonju, 54896, South Korea*

ARTICLE INFO

ABSTRACT

Drug combinations are frequently used to treat cancer to reduce side effects and increase efficacy. The experimental discovery of drug combination synergy is time-consuming and expensive for large datasets. Therefore, an efficient and reliable computational approach is required to investigate these drug combinations. Advancements in deep learning can handle large datasets with various biological problems. In this study, we developed a SynergyGTN model based on the Graph Transformer Network to predict the synergistic drug combinations against an untreated cancer cell line expression profile. We represent the drug via a graph, with each node and edge of the graph containing nine types of atomic feature vectors and four bonds features, respectively. The cell lines represent based on their gene expression profiles. The drug graph was passed through the GTN layers to extract a generalized feature map for each drug pairs. The drug pair extracted features and cell-line gene expression profiles were concatenated and subsequently subjected to processing through multiple densely connected layers. SynergyGTN outperformed the state-of-the-art methods, with a receiver operating characteristic area under the curve improvement of 5% on the 5-fold cross-validation. The accuracy of SynergyGTN was further verified through three types of cross-validation tests strategies namely leave-drug-out, leave-combination-out, and leave-tissue-out, resulting in improvement in accuracy of 8%, 1%, and 2%, respectively. The Astrazeneca Dream dataset was utilized as an independent dataset to validate and assess the generalizability of the proposed method, resulting in an improvement in balanced accuracy of 13%. In conclusion, SynergyGTN is a reliable and efficient computational approach for predicting drug combination synergy in cancer treatment. Finally, we developed a web server tool to facilitate the pharmaceutical industry and researchers, as available at: http://nsclbio.jbnu.ac.kr/tools/SynergyGTN/.

1. Introduction

Drug combination therapies have gained attention for treating patients with complex diseases, especially cancer, because of their efficacy [1,2]. Moreover, drug combinations have increased the potential for cancer treatment by concurrently targeting the abundant molecular mechanisms of cancer cells. Mono-therapy can cure several human diseases; however, it has many limitations, such as resistance or inefficiency [3,4]. Human diseases are caused by complex interactions between genomics and phenotypic factors. Usually, a single drug targets a single pathway or protein, which is not sufficient for a complex disease. The combination of drugs can overcome these shortcomings, including decreased unfavorable side effects [5] and increased efficacy [1,2]. Although drug combinations are mostly very useful, they have some side effects [6]. For example, the combination of panitumumab and bevacizumab decreases the progression-free survival of metastatic colorectal cancer patients because of increased toxicity [6]. Hence, accurate identification of the synergistic effect of drug combinations is necessary to treat different cancer types.

The exponential growth of drug combinations with respect to the increase in cancer types makes the study of drug synergy more challenging. These drug combinations were proposed through wet-lab experiments, which are time-consuming and costly [7]. Moreover, drug combination trials may cause side effects or harmful reactions in patients [8]. Therefore, preclinical strategies such as high-throughput screening (HTS) have been introduced for the identification and determination of drug combinations in different cancer cell line effects [5, 9,10]. Owing to advancements in HTS technology, data analysis and system-level management have applications in integrative cancer drug combination data portal and DrugComb database [11]. O'Neil et al.

(2016) performed HTS experiments to identify 23062 pairs of drug combinations from 583 drug combinations along with 39 different cancer types of cell-line profiles [5]. The target assessment methods based on druggability analysis using the Therapeutic Target Database (TTD), categorizing nine druggability characteristics across diverse targets, and highlights the potential of TTD and similar databases in advancing the discovery and validation of innovative drug targets [12].

In recent decades, several computational methods have been proposed for the prediction of synergistic drug combination effect scores. Initially, system biological methods, stochastic search algorithms, and mathematical and statistical methods were introduced for novel drug pair prediction. System biological methods mainly focus on the analysis of biological networks, which have limitations such as biological knowledge, making it impossible to obtain large amounts of data on drug combinations [13]. Stochastic search algorithms operate based on the possibility that drugs are simultaneously combined and predict their effects in a vast space [14]. Stochastic search algorithms are efficient for small datasets but owing to their high computational costs, they are not suitable for large-scale applications. Mathematical and statistical methods focus on the quality of the hypotheses behind the models [15,16]. Addressing the foremost challenge in cancer treatments, drug resistance, the Re-Sensitizing Drug Prediction (RSDP) introduces a novel computational strategy that accurately predicts personalized cancer drug combinations by reversing the resistance signature of drug A+B, utilizing a comprehensive approach integrating various biological features, and establishes a substantial resource of cell line—specific cancer drug resistance signatures, offering a promising avenue for guiding personalized medicine decisions in the realm of oncology [17]. Advancement in deep learning technologies helped the researcher using it for several problems such as image processing [18–21], natural language processing [22–24], speech recognition [25] and bioinformatics sequence analysis [26–29].

More recently, several machine learning models have been introduced for the prediction of the synergistic scores of drug combinations. First, Sidorov et al. [30] proposed traditional machine learning methods for drug combination effect prediction, which were based on classifiers such as random forest and extreme gradient boosting (XGboost). These methods utilizes physicochemical properties as features for drugs in trained models for every cell line. In both methods, the XGboost model performed slightly better than the random forest model. In 2018, a new model was published to introduce a deep learning method for predicting the synergy effect scores of drug pairs. This model applied the multiple layer perceptron (MLP) method with a chemical descriptor for drug representation and gene expression in the cell line. This was the first regression deep-learning model to predict the synergy of drug combinations [31]. The AuDNNsynergy model is the latest model for synergy score prediction of drug combinations, which also uses the MLP method along with the integration of multi-omics data from cancer cell lines [32]. However, the prediction of the synergistic score requires a more accurate computational method.

The most promising development in the deep learning is the graph based networks, which has gained significant attention and become a hot topic for current research of molecular graph data. Numerous graph neural networks have been proposed as a result of the rapid growth of graph machine learning [33]. Representative models among them are residual gated graph convolutional networks (RGG) [34], topology adaptive graph convolutional networks (TAG) [35], and heterogeneous graph neural network (HGN) [36]. Data can be represented in graphs in a variety of ways, and many kinds of data can be converted into graphs for predicting drug combination synergy effects and identification are fundamentally edge prediction problems. For the purpose of predicting drug–drug interaction, Cheng et al. [37] suggested an end-to-end deep learning method based on a graph attention network and several self-attention mechanisms. A multi-head self-attention mechanism and a graph attention network are used to enhance the feature extraction of drug and MLP to reduce the dimension of cell-line expression [38]. The structural characteristic information of drug and cell line expres-sion is only represented using one-dimensional data, and much more sophisticated characteristic information is lost in prediction. Motivated by these observations, Through the residual GNN, this network learns the intricate graph aspects. They combine these properties with the attention module to create a sophisticated cell-line vector for processing multilayer perceptron's. The majority of graph neural network-based models, however, only look at the connections between cell-line and drug, ignoring many of the connections between each set of drugs and cell-line. Based on these drawbacks, this work suggests a graph transformer-based method for forecasting synergy of drug combination that accounts for the connections between each collection of drug– drug combinations as well as the details of the nodes and the entire graph in order to forecast the synergy between drugs and targets. The proposed method overcome the shortcoming of the existing methods and achieved the ROC-AUC, PR-AUC, and ACC of 0.98, 0.97, and 0.93 on the 5-fold cross-validation, respectively. The data flow of the proposed method is shown in Fig. 1. Finally, we have established a GitHub repository to facilitate academia and the pharmaceutical industry in reproducing the code and accessing the results. The repository is available at https://github.com/waleed551/SynergyGTN.

2. Results

The efficiency of the proposed model was first evaluated on 5-fold cross-validation using standard evaluation metrics. To ensure the validity of the experimental results and eliminate the possibility of pseudo-random results, all models were trained five times under identical conditions, with the results averaged and standard deviations calculated. The ROC-AUC and PR-AUC were used to assess the performance of the proposed model. The ROC-AUC and PR-AUC values were found to be 0.98 and 0.97, which were improved 5% and 4%, respectively, higher than the previous best model, DeepDDS [38]. In the task of predicting drug combination synergy on cancer cell-lines, the SynergyGTN model outperformed all existing state-of-the-art methods.

Additionally, the superior performance of the SynergyGTN in identifying novel drug combination pairs in the cross-validation test datasets can be attributed to the accuracy of the drug feature vector extraction process performed by the graph transformer network layer. This layer is able to effectively learn from the nodes, edges, and edge attributes of each drug, composing a comprehensive feature representation of the drug. The feature distribution was computed using t-distributed stochastic neighbor embedding (t-SNE) and the captured feature distribution was visualized using the t-SNE implementation from the Scikit-Python library (https://scikit-learn.org). The results of the t-SNE visualization are displayed in Fig. 2, where the red and blue dots represent the Synergistic and Antagonistic drug pairs for cancer-affected cell lines, respectively. The figure clearly demonstrates the robustness of the SynergyGTN in accurately predicting the optimal drug pairs for the given cell lines.

2.1. Comparisons with existing models on K-fold cross validation

The proposed method, SynergyGTN, has been evaluated and compared to existing state-of-the-art methods using the same evaluation measures. Therefore, we employed a K-fold cross-validation strategy with a k value of 5, which aligns with the existing model's configuration for training and validating the model's performance. The evaluation of the SynergyGTN models performance was conducted based on the confusion matrix, as presented in 3a. In addition, the receiver operating characteristic curve and its corresponding area under the curve (ROC-AUC) values were analyzed and the results, along with their standard deviation, were illustrated in Fig. 3b. The results of the comparison are presented in Table 1 and Fig. 4. SynergyGTN achieved an ROC-AUC, PR-AUC, and ACC of 0.98, 0.97, and 0.93, respectively. The results indicate that SynergyGTN improved the ROC-AUC, PR-AUC and ACC by 5%, 4% and 8% respectively when compared to the existing best
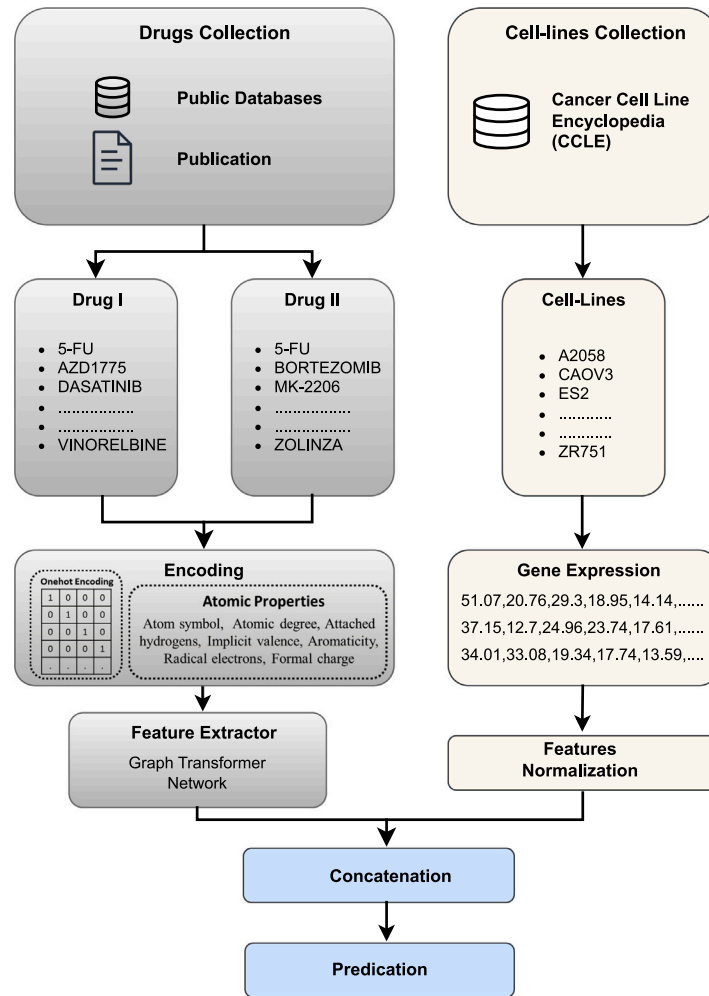
Fig. 1. The data-flow of the proposed method.

Table 1

Performance comparison of the proposed model with state-of-the-art model on 5-fold cross-validation.

Methods	ROC-AUC	PR-AUC	ACC	BACC	PREC	TPR	KAPPA
SynergyGTN	**0.98**	**0.97**	**0.93**	**0.93**	**0.91**	**0.92**	**0.87**
DeepDDS-GAT	0.93	0.93	0.85	0.85	0.85	0.85	0.71
DeepDDS-GCN	0.93	0.92	0.85	0.85	0.85	0.84	0.70
XGBoost	0.92	0.92	0.83	0.83	0.84	0.84	0.68
Random Forest	0.86	0.85	0.77	0.77	0.78	0.74	0.55
GBM	0.85	0.85	0.76	0.76	0.77	0.74	0.53
Adaboost	0.83	0.83	0.74	0.74	0.74	0.72	0.48
MLP	0.65	0.63	0.56	0.56	0.54	0.53	0.12
SVM	0.58	0.56	0.54	0.54	0.54	0.51	0.08
AuDNNsynergy	0.91	0.63	0.93	NA	0.72	NA	0.51
TranSynergy	0.90	0.89	0.83	0.83	0.84	0.80	0.64
DTF	0.89	0.88	0.81	0.81	0.82	0.77	0.63
DeepSynergy	0.88	0.87	0.80	0.80	0.81	0.75	0.59

method on 5-fold cross validation. Furthermore, we tested the proposed method by switching the positions of Drug A and Drug B as input and found the same results, providing evidence that the position of Drug A and Drug B as input has no influence on the proposed methods probability prediction.

2.2. Comparisons on three cross-validation with existing models

To ensure a comprehensive evaluation of the proposed SynergyGTN model and a fair comparison with existing methods, three

cross-validation strategies were employed. The results of each strategy are presented in Table 2 Fig. 5. The first strategy, leave-drug-combination cross-validation, demonstrated a 1% improvement in the ROC-AUC, PR-AUC, and ACC for the SynergyGTN model compared to existing methods. The second strategy, leave-drug-out cross-validation, revealed poor performance of existing methods in predicting synergistic effects. In contrast, the SynergyGTN model achieved the improvement of 7% and 8% in ROC-AUC and ACC, respectively. The third strategy, leave-tissue-out validation, showed that the SynergyGTN model achieved an average ACC of 0.76, which represents a 2% improvement compared to all other competitive existing models. Furthermore, the proposed model achieved superior results for individual tissues, including breast, colon, lung, melanoma, ovarian, and prostate, with ROC-AUC values of 0.857, 0.868, 0.857, 0.843, 0.834, and 0.812, respectively. These results demonstrate the generalization ability and superiority of the SynergyGTN model in all three cross-validation strategies. The ROC-AUC and PR-AUC curves results for each test case and their folds along the standard deviation are shown in Supplementary Figure S(1–6).

2.3. Comparison of independent dataset (ASTRAZENECA)

The proposed model has been evaluated on an independent dataset, which was not seen during the training process, in order to assess its generalizability. The ASTRAZENECA datasets was imbalanced, the primary focus was on the balanced accuracy (BACC). The SynergyGTN, demonstrated an improvement of 13% in BACC on the ASTRAZENECA
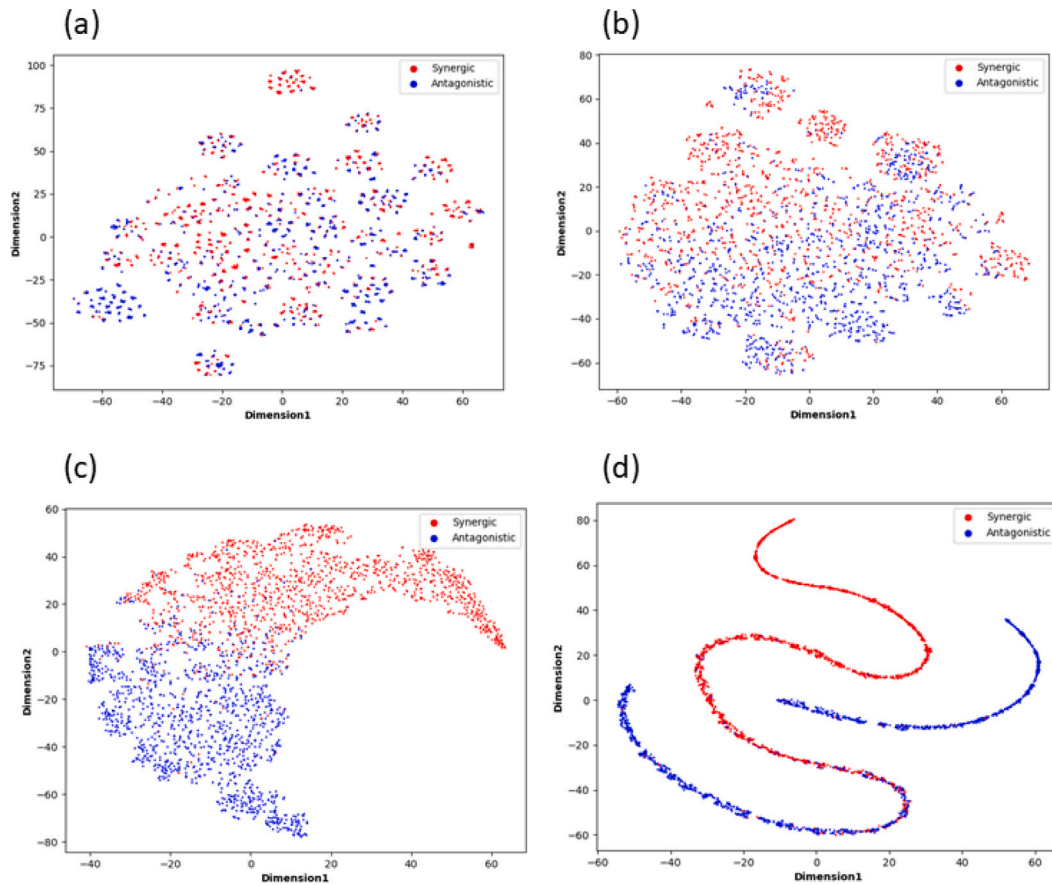
Fig. 2. The captured feature distribution of the different layer of SynergyGTN for 5-fold cross-validation. **(a)** Showcases the concatenated feature representation of both drug1 and drug2, as well as the reduced cell-line feature, captured by the GTN. **(b)** Demonstrates the feature distribution after the application of the dropout layer. **(c)** Depicts the third linear layer, which effectively differentiates between the synergistic and antagonistic classes, though with some overlap in the features. **(d)** Displays the results of the final layer, which exhibits clear clustering of the synergistic and antagonistic classes with minimal overlap.
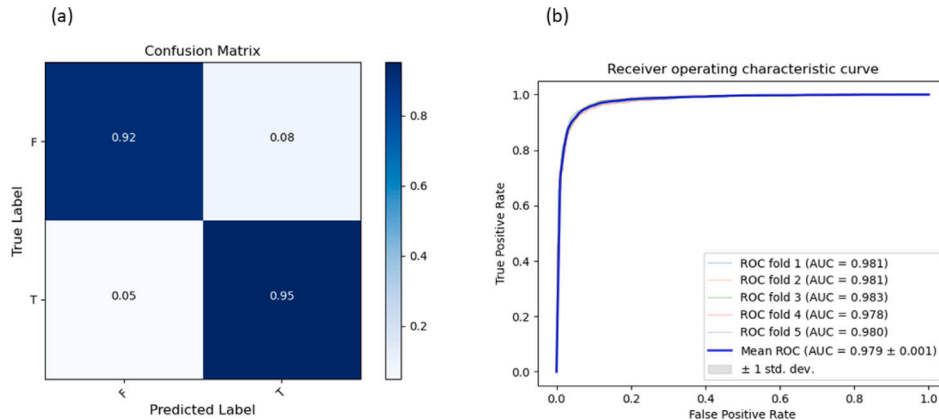

Fig. 3. The evaluation of proposed model on 5-fold cross-validation. (a) Confusion matrix. (b) The auROC curves of the 5-fold on testing data, the mean auROC is 0.979 and their standard deviation of 0.001.

dataset as compared with existing state-of-art-models. This significant improvement in BACC, along with outperforming all other evaluation parameters, serves as evidence of the methods generalizability. The detailed results of the ASTRAZENECA dataset are presented in Table 3 and Fig. 6.

2.4. Evidence for the prediction of novel drug combination

In our study, the combination of Dasatinib and AZD1775 demonstrated a high predication probability of 0.99% in the HCT116 cell line.

Dasatinib is a tyrosine kinase inhibitor that targets multiple kinases, including Src, Abl, and various receptor tyrosine kinases (RTKs), while AZD1775 is a small molecule inhibitor of the cell cycle checkpoint kinase 1 (CHK1) [58]. Preclinical studies have shown that combining these two drugs can lead to synergistic antitumor activity in various cancer cell lines, while having minimal toxicity in normal cells. A study demonstrated that the combination of Dasatinib and AZD1775 induced DNA damage and cell death in lung cancer cell lines and showed promising activity in vivo in a mouse model of lung cancer [59]. Karpel-Massler et al. showed that the combination of these drugs had
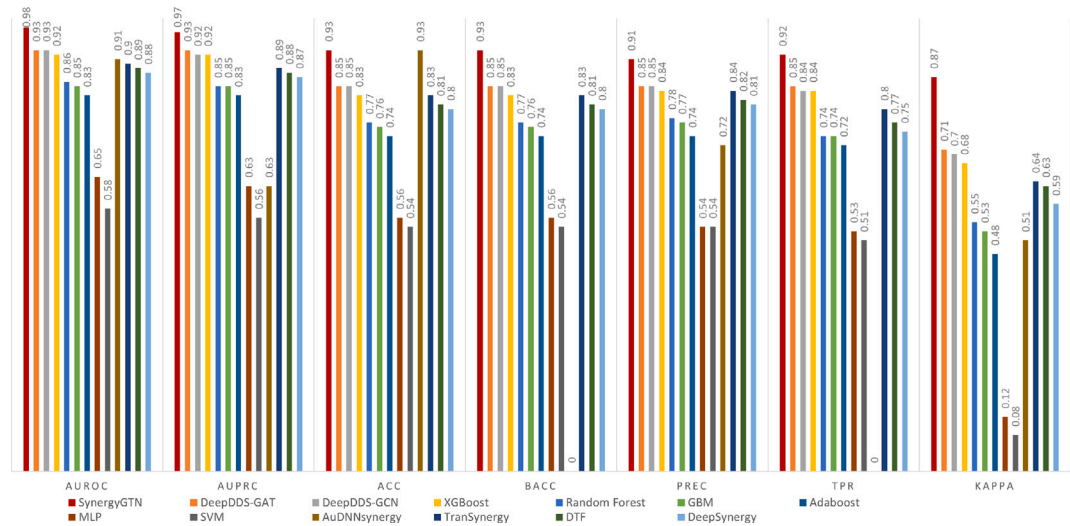
Fig. 4. The performance of proposed model compared with the existing published models on 5-fold cross-validation.

Table 2

Performance comparison of the proposed models with state-of-the-art model on three different cross-validation strategies.

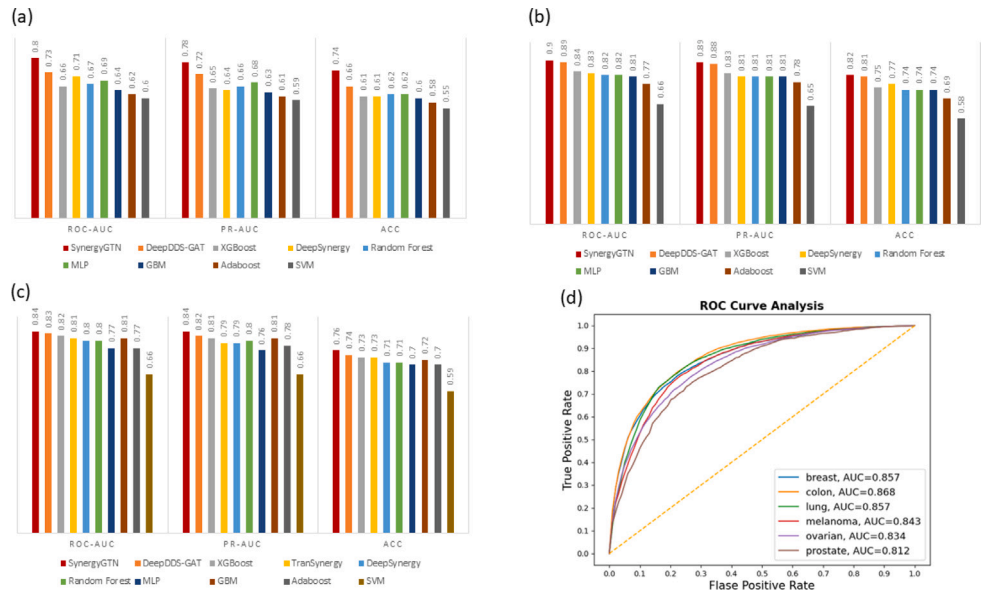| Methods	Leave-drug-out			Leave-combination-out			Leave-cell-line-out		
	ROC-AUC	PR-AUC	ACC	ROC-AUC	PR-AUC	ACC	ROC-AUC	PR-AUC	ACC
SynergyGTN	**0.80**	**0.78**	**0.74**	**0.90**	**0.89**	**0.82**	**0.84**	**0.84**	**0.76**
DeepDDS-GAT	0.73	0.72	0.66	0.89	0.88	0.81	0.83	0.82	0.74
XGBoost	0.66	0.65	0.61	0.84	0.83	0.75	0.82	0.81	0.73
TranSynergy	NA	NA	NA	NA	NA	NA	0.81	0.79	0.73
DeepSynergy	0.71	0.64	0.61	0.83	0.81	0.77	0.80	0.79	0.71
Random Forest	0.67	0.62	0.62	0.82	0.81	0.74	0.80	0.80	0.71
MLP	0.69	0.68	0.62	0.82	0.81	0.74	0.77	0.76	0.70
GBM	0.64	0.63	0.60	0.81	0.81	0.74	0.81	0.81	0.72
Adaboost	0.62	0.61	0.58	0.77	0.78	0.69	0.77	0.78	0.70
SVM	0.60	0.59	0.55	0.66	0.65	0.58	0.66	0.66	0.59

Fig. 5. The comparison of the proposed model and previous established models on three type of cross validation test. **(a)** The leave-drug-out cross-validation results. **(b)** The leave-combination-out cross-validation results. **(c)** The leave-tissue-out cross-validation results. **(d)** The ROC-AUC curves for leave-tissue-out.

synergistic effects in breast cancer cell lines and increased apoptosis, while sparing normal cells [60]. Moreover, a phase I clinical trial by Morel et al. evaluated the safety and efficacy of the Dasatinib and AZD1775 combination in patients with advanced solid tumors [61].

The results showed promising activity, with some patients achieving a partial response and others experiencing stable disease. These preclinical and clinical studies provide a strong rationale for further investigating the Dasatinib and AZD1775 combination in larger clinical

Table 3

Performance comparison of the proposed model with state-of-the-art models on the independent data-set (ASTRAZENECA).

Methods	ROC-AUC	PR-AUC	ACC	BACC	PREC	TPR	KAPPA
SynergyGTN	0.81	0.89	0.80	0.75	0.85	0.87	0.52
DeepDDS-GAT	0.66	0.82	0.64	0.62	0.80	0.67	0.21
DeepDDS-GCN	0.67	0.83	0.60	0.63	0.83	0.56	0.21
DeepSynergy	0.55	0.71	0.47	0.53	0.75	0.39	0.04
Random Forest	0.53	0.76	0.50	0.54	0.75	0.49	0.06
MLP	0.53	0.74	0.53	0.53	0.74	0.53	0.05

Table 4

Validation and biological verification of prediction of proposed method.

| DrugI	DrugII	Publish year	Predicated probability	Reference
5-FU	AZD1775	2018	0.99	[39]
5-FU	BEZ-235	2015	0.99	[40]
5-FU	BEZ-235	2015	0.98	[40]
5-FU	CYCLOPHOSPHAMIDE	1991	0.97	[41]
5-FU	DASATINIB	2018	0.97	[42]
5-FU	ERLOTINIB	2007	0.96	[43]
5-FU	L778123(Tipifarnib)	2017	0.98	[44]
5-FU	lapatinib	2017	0.96	[45]
5-FU	Mitomycine	2019	0.99	[46]
5-FU	MK-2206	2016	0.99	[47]
5-FU	MK-4541	2023	0.99	[48]
5-FU	MK-5108	2021	0.99	[49]
5-FU	MK-8669	2021	0.99	[50]
5-FU	MK-8776	2014	0.99	[51]
5-FU	MK-003	2021	0.98	[50]
5-FU	Paclitaxel	2022	0.99	[52]
5-FU	SUNITINIB	2012	0.96	[53]
5-FU	TEMOZOLOMIDE	2019	0.98	[54]
5-FU	TOPOTECAN	2003	0.95	[55]
5-FU	VINBLASTINE	1997	0.87	[56]
5-FU	ZOLINZA	2019	0.99	[57]
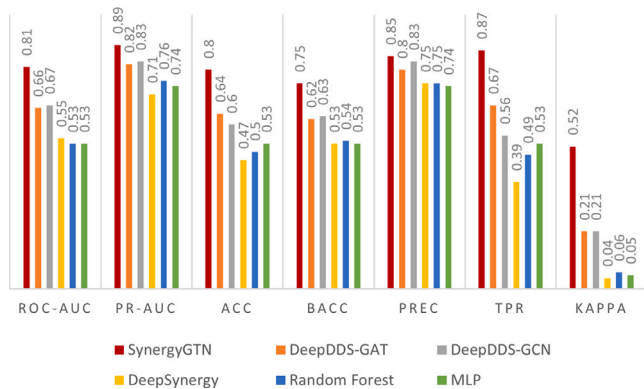
Fig. 6. The comparison of the proposed method with existing state-of-the-art models on independent dataset (ASTRAZENECA).

trials, to evaluate its potential as a treatment option for cancer patients. In addition, the proposed model prediction of other novel drugs combinations for HCT116 are shown in Fig. 7a.

In the A375 human melanoma cell line, the combination of MK-8669 and METFORMIN had the highest prediction probability of 0.99%. MK-8669, a potent and selective mTOR inhibitor, has been shown to inhibit the proliferation of various tumor cell lines and xenografts [62]. METFORMIN, a commonly prescribed drug for type II diabetes, has been demonstrated to possess strong anti-cancer properties. It activates adenosine monophosphate-activated protein kinase (AMPK), which in turn inhibits the mTOR signaling pathway [63, 64]. Previous studies have shown that the combination treatment with mTOR inhibitors and METFORMIN can synergistically inhibit the growth of pancreatic cancer in vitro and in vivo [65]. Thus, novel predications of several drug combination are shown in Fig. 7b for cell

A375 for further investigations. The remaining cell-lines heatmap of predicated probability score are shown in Supplementary Figure S7.

In order to validate the findings of the heatmap analysis in Fig. 7b, we selected 5-FU along with its positively predicted combinations involving 19 approved drugs. The presentation of these combinations, as well as 5-FU, is detailed in Table 4. Notably, 18 out of the 19 drug pairings with 5-FU are novel with respect to the proposed methodology, as they are not included in the benchmark dataset. Comprehensive validation and verification of these predictions were performed through a complete review of relevant literature sources, which are appropriately referenced in the table. As a result, we assert that the predictions made by the proposed method exhibit significant efficacy and accuracy within the context of cancer treatment.

2.5. Web-server for prediction of drug combination synergy

Finally, A web server, based on the Graph Transformer Network, has been developed to predict the synergy of drug combinations against the gene expression profiles of cancer cell lines. The design of the web server is user-friendly and easy to navigate, making it accessible for researchers in wet-lab and the pharmaceutical industry to determine the optimal drug combination. The web-server interface is demonstrated in Figs. 10 and 11. The following steps should be followed to use the SynergyGTN web server:

(1) The first two tabs allow the user to input the SMILES of drug1 and drug2, respectively.

(2) A scroll button is available to select the untreated cancer cell line by name or all cell-lines for the given drug pair.

(3) The third tab enables the user to input a choice of cell-line expression that is not available in the scroll list.

Note: The cell-line expression feature vector, consisting of 957 features, can be created using the "Cell line representation" section to create your own cell-line expression for the prediction.

3. Materials and methods

3.1. Benchmark dataset

The benchmark dataset plays a key role in the development of an efficient and accurate machine learning model. Therefore, we utilized a valid benchmark dataset containing 23062 samples, introduced by Merck research laboratories using an oncology screening technique [5]. These data samples were derived from seven types of human cancers by testing 38 verified drugs, along with 39 untreated cell lines. The drug samples contained 14 experimental and 24 drugs approved by the United States Food and Drug Administration. We utilized the open-source chemical informatics software RDKit [66] to collect primary structured molecules from a drug Simplified Molecular Input Line Entry System (SMILES) Using the Combenefit tool, the synergy score for each drug pair was determined [67]. The final unique drug combinations were calculated as the average of replicating drug pairings. The CCLE [68], a separate initiative that works to describe genomes, messenger RNA expression, and anticancer treatment dosage responses across cancer cell lines, was used to gather the gene expression data for cancer cell lines. Transcripts Per Million is used to normalize the expression data based on the genome-wide read counts matrix. We collected the normalized gene expression profiles of required cancer cell-line from CCLE [68]. The drug pair-cell line triplets were classified using a threshold of 10 for label balancing and deleting noisy data. Positive triplets had synergistic values more than 10, whereas negative triplets had scores lower than 0 [38]. Finally, using 31 cell lines and 36 drugs, we were able to create 12415 distinct triplets for the proposed computational model. Furthermore, we used k-fold cross-validation to simultaneously split the dataset into training and testing samples. According to the recent literature on computational models,
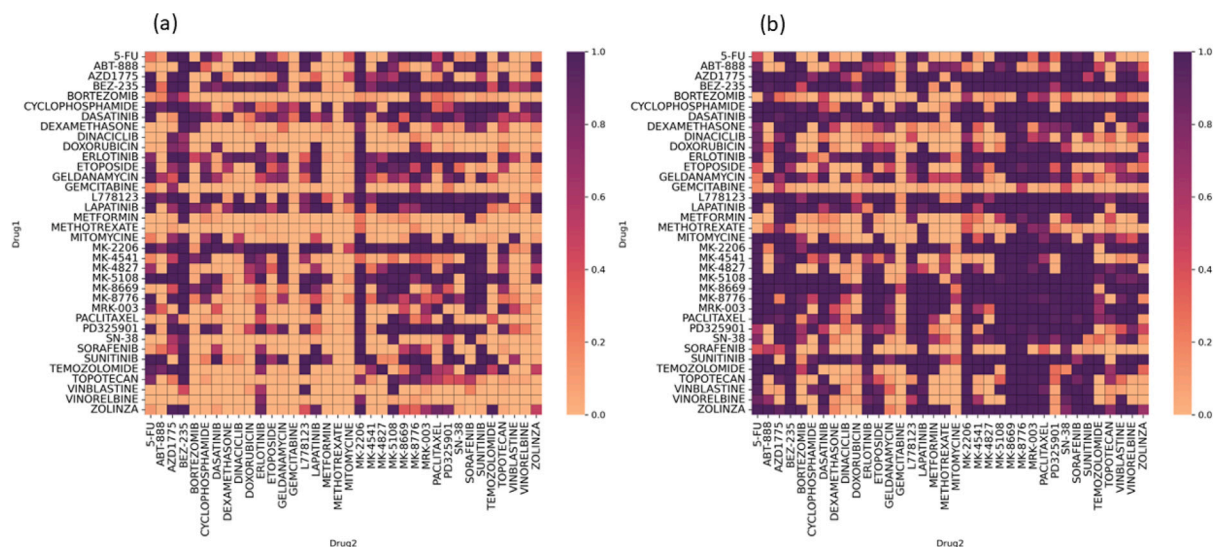
Fig. 7. The heat-maps has been adopted to represent the drug pair probability scores for two selected cell-lines HCT116 and A375, respectively. **(a)** The heat-map of HCT116 show drug AZD1775 pair have high probability score predicted by proposed model. **(b)** The heat-map of A375 show drug ABZ-235 and MK-8669 pair have high probability score predicted by proposed model.
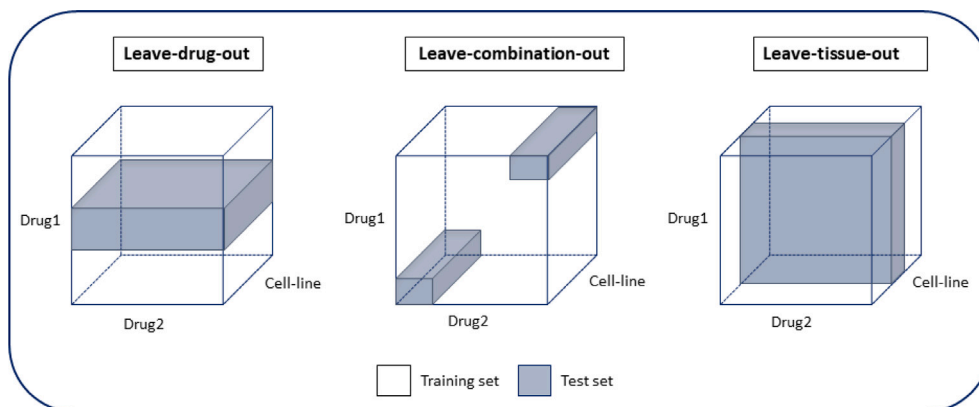

Fig. 8. The dataset was divided into various subsets to train and test the proposed method to verify the generalization ability. In leave-drug-out case, we choose one drug from the whole dataset for testing and remaining for the training. In drug-combination-out case, the selection of drug pair from the dataset for testing and the remaining samples has considered for training. In leave-tissue-out case, the triplets of a tissue has been selected for testing and the remaining data has been used for training.

evaluation of the model using k-fold cross-validation the outcome of k-fold combinations can be considered as different training and testing datasets. In this study, we set the k value to 5 to create a 5-fold validation of the dataset, similar to a previous study for fair evaluation and comparison; each fold contained 9932 training samples and 2483 testing samples.

3.2. Methods

3.2.1. Drug representation

Drug representations have played a key role in the development of efficient and reliable computational models. In this study, the drug is represented in a SMILES, which was developed to represent molecules that are readable by a computer [69]. SMILES contains much information on drug descriptors, such as the number of valence electrons and heavy atoms, which were used as features for synergy, affinity, and toxicity effect prediction. SMILES is also used directly as a string-to-feature by natural language processing and a convolutional neural network. In this study, the SMILES representation are converted into the molecules with the help RDKIT tool [70], check if the molecule is valid then generate graph object which contained the node and edge features. Avoiding the ambiguity in the graph representation

by using a unique number for each atom in a molecule and then traversed the molecular graph. The node and edge features are the composition of nine types of atomic features and four types of bond features, respectively. The one-hot encoding has been used to encode the node and bond features in categorical variables for the model training. The detail information of one-hot encoding bitwise is given in Supplementary Table S1. For instance, degree was encoded with an 8-bit one-hot vector, while hybridization was encoded with a 7-bit one-hot vector. The one-hot encoded feature vectors of molecules have been converted into graph representation by utilizing the Pytorch geometric library [71]. The resulting feature vectors for training are represented by a 92×10, where the 92 features correspond to the nodes and the 10 features correspond to the edge bonds.

3.2.2. Cell line representation

The cell-line gene expression profiles have been extracted and eliminate the redundant data and the transcripts of noncoding RNA using the gene annotation details from the CCLE [68] and the GENCODE annotation database [72]. We chose the significant genes in accordance with the Library of Integrated Network-Based Cellular Signatures (LINCS) project [73] in order to address the dimension imbalance between the feature vectors of drugs and cell lines. Based on the Connectivity Map

data, the LINCS project offers a collection of around 1000 carefully chosen genes known as the "landmark gene set" that can capture 80% of the information [74]. The genes that crossed over between the landmark set and the CCLE gene expression profiles were chosen for further study. The final features map contained 954 genes to feed into the model as each cell-line expression.

3.2.3. Data splitting

We adopt the standard KFold cross-validation from scikit-learn library (version 1.3.0) to split the data into 5-fold for training and testing. In this approach, the training data set consisted of four folds, whereas the testing dataset consisted of one-fold. The proposed model hyper-parameters were tuned to the 5-Fold cross validation and achieved the best performance. In addition, the several data splitting methodologies have been employed to evaluate the proposed model generalized ability of the prediction in multiple scenarios shown in Fig. 8. In the first scenario, we select one drug and their combination from the training data and evaluate the model performance on the selected combinations. In the second scenario, we choose drug pair for testing from the dataset and the remaining were used for the training to check the model ability of predication. In the third scenario, we keep one cell-line out from the training set for testing the model and evaluate the predication ability.

3.2.4. Graph transformer network

In 2017, Google introduced the transformer model, which is still frequently used today. The self-attention technique in this architecture enabled quick parallelism for machine translation workloads at initially. The graph transformer network can address the transformer model's main flaw, which is RNNs' sluggish training speed. To maintain the features of the graph, Dwivedi et al. [75] extended the transformer model to graphs. In particular, the multi-head attention of each edge from j to i is determined as follows given the node feature $H^{(l)} = H_1^{(l)}, H_2^{(l)}, \ldots, H_n^{(l)}$.

$$q_{c,i}^{(l)} = W_{c,q}^{(l)} + b_{c,q}^{(l)} \tag{1}$$

$$k_{c,i}^{(l)} = W_{c,k}^{(l)} + b_{c,k}^{(l)} \tag{2}$$

$$e_{c,i,j} = W_{c,e} e_{i,j} + b_{c,e} \tag{3}$$

$$\alpha_{c,i,j} = \frac{(q_{c,i}^{(l)}, k_{c,i}^{(l)} + e_{c,i,j})}{\sum_{u \in N_{(i)}} (q_{c,i}^{(l)}, k_{c,u}^{(l)}) + e_{c,i,u}} \tag{4}$$

The exponential scale dot-product function and d, the hidden size of each head, are both included in Formula (4). For the C head attention, first encode the edge features $e_{i,j}$ and add them to the key vector as additional information in each layer by converting them into $q_{c,i}^{(l)} \in \mathbb{R}^d$ and $k_{c,i}^{(l)} \in \mathbb{R}^d$ respectively, using distinct trainable parameters $W_{c,q}^{(l)}, W_{c,k}^{(l)}, b_{c,q}^{(l)}, b_{c,k}^{(l)}$. After getting the graph's multi-head attention, message aggregation is carried out for the following distances:

$$v_{c,i}^{(l)} = W_{c,v}^{(l)} + b_{c,v}^{(l)} \tag{5}$$

$$\hat{h}_i^{(l+1)} = \|_{c=1}^{C} \left[\sum_{j \in \aleph}^{n} \alpha_{c,ij}^{(l)} (v_{c,i}^{(l)} + e_{c,i,j}) \right] \tag{6}$$

According the equation C represent the number of multi-headed attentions while ∥ is the connection between attentions, and v_c is applied instead of the distance feature $h_j, j \in \mathbb{R}^d$ for the weighted sum. Furthermore, Shi et al. [76] use a gated residual connection between layers to prevent the model from becoming too smooth, a multi-headed attention matrix as the transfer matrix for message passing rather than the original normalized adjacency matrix, and finally, apply graph transformer on the final output layer to apply averaging to the multi-headed output and remove the non-linear transformation.

3.2.5. Network setup

Fig. 9 depicts the SynergyGTN deep learning framework designed for predicting synergistic drug combinations against untreated cancer cell-line expressions. In each pairwise drug combination, the initial step involves inputting the molecular graphs of the two drugs and the gene expression profiles of a cancer cell line treated with these drugs into the input layer. In the subsequent phase, each drug undergoes processing through two Graph Transformer Network (GTN) layers to extract feature embedding vectors. Simultaneously, the cell line expression is fed into the Multilayer Perceptron (MLP) for normalization and the importation of feature vectors. Both GTN layers were fine-tuned to identify optimal hyper-parameters, with filters, heads, and edge-dims set at 128, 20, and 10, respectively. Although the MLP comprises three dense layers, the node sizes were configured as 256, 128, and 64, respectively. Both the GTN and MLP layers of the network utilized the ReLU activation function. Dropout layers were incorporated to regulate network overfitting, with a dropout value set at 0.05. The resultant embedding vectors are then concatenated to create the ultimate feature representation for each drug pair and cell line. This representation undergoes propagation through fully connected layers, enabling binary classification to distinguish between synergistic and antagonistic drug combinations.

3.2.6. Hyperparameter settings

We set the input dimensions for the cell line expression, and chemical atomic vector in SynergyGTN to be 954, and 92, respectively. The ideal SynergyGTN parameters were tuned using a Optuna method [77]. The range of hyper-parameters are enlisted in Supplementary Table S2. The same hyper-parameters were tuned, which were used by comparative state-of-the-art models.

3.2.7. Performance measures

In order to evaluate and compare the SynergyGTN model with state-of-the-art existing models, which were utilized for various bioinformatics tools [31,38,78,79]. These evaluate metrics are area under the receiver operating characteristics (ROC-AUC), area under the precision recall (PR-AUC), accuracy (ACC), balanced accuracy (BACC), precision score (PREC), true positive Rate (TPR) and cohen's kappa co-efficient (KAPPA). The mathematical formulation of evaluation metrics is given in the following:

$$\text{ACC} = \frac{(\text{tp} + \text{tn})}{(\text{tp} + \text{tn} + \text{fp} + \text{fn})} \tag{7}$$

$$\text{BACC} = \frac{1}{2} \left(\frac{\text{tp}}{\text{tp} + \text{fn}} + \frac{\text{tn}}{\text{tn} + \text{fp}} \right) \tag{8}$$

$$\text{PREC} = \frac{\text{tp}}{\text{tp} + \text{fp}} \tag{9}$$

$$\text{TPR} = \frac{\text{tp}}{\text{tp} + \text{fn}} \tag{10}$$

$$\text{KAPPA} = \frac{\text{p}_o - \text{p}_e}{1 - \text{p}_e} \tag{11}$$

The ROC-AUC curve have defined the probability of false positive rate (FPR) on the horizontal axis and the probability of the true positive rate (TPR) on the vertical axis.

4. Conclusion and discussion

The combination of drugs has become a promising approach for treating complex diseases, particularly cancer. In this study, we introduced a Graph Transformer Network (GTN) based computational model for predicting accurate and reliable synergistic drug combinations for untreated cancer cell lines. The proposed SynergyGTN architecture consists of two consecutive graph transformer layers for extracting a generalized feature vector for each drug and several dense layers employed to reducing and extracting important the cell-line expression
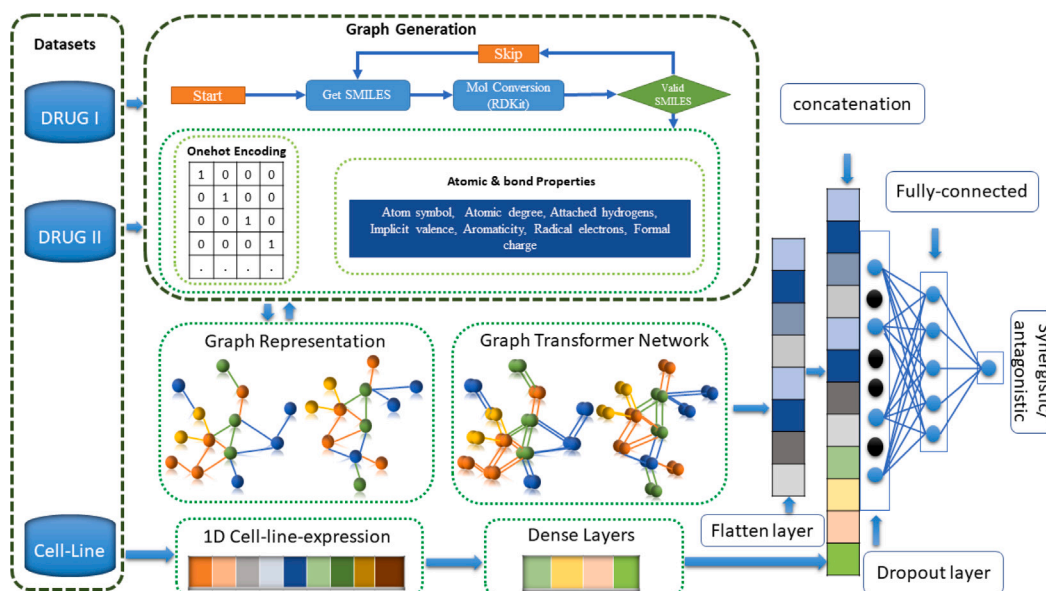
Fig. 9. The data-flow and architecture of the proposed model.

Fig. 10. The implemented web-server menu interface for inputs the desire drugs pair and cell-line, the backend has followed by proposed model for the prediction.

data points. The benchmark dataset used for training and testing was introduced and validated through oncology screening techniques by Merck Research Laboratories. Feature extraction plays a crucial role in enhancing the ability and effectiveness of the AI model learning. To this end, the RDKit tool was utilized to generate molecular structure graphs based on SMILES representations. Each graph contained nine types of atomic features in each node and four types of edges, allowing the GTN to determine the intra-molecular properties of a given chemical structure at the atomic level. The generalized learned features are then used by several fully connected layers to classify the synergistic effects.
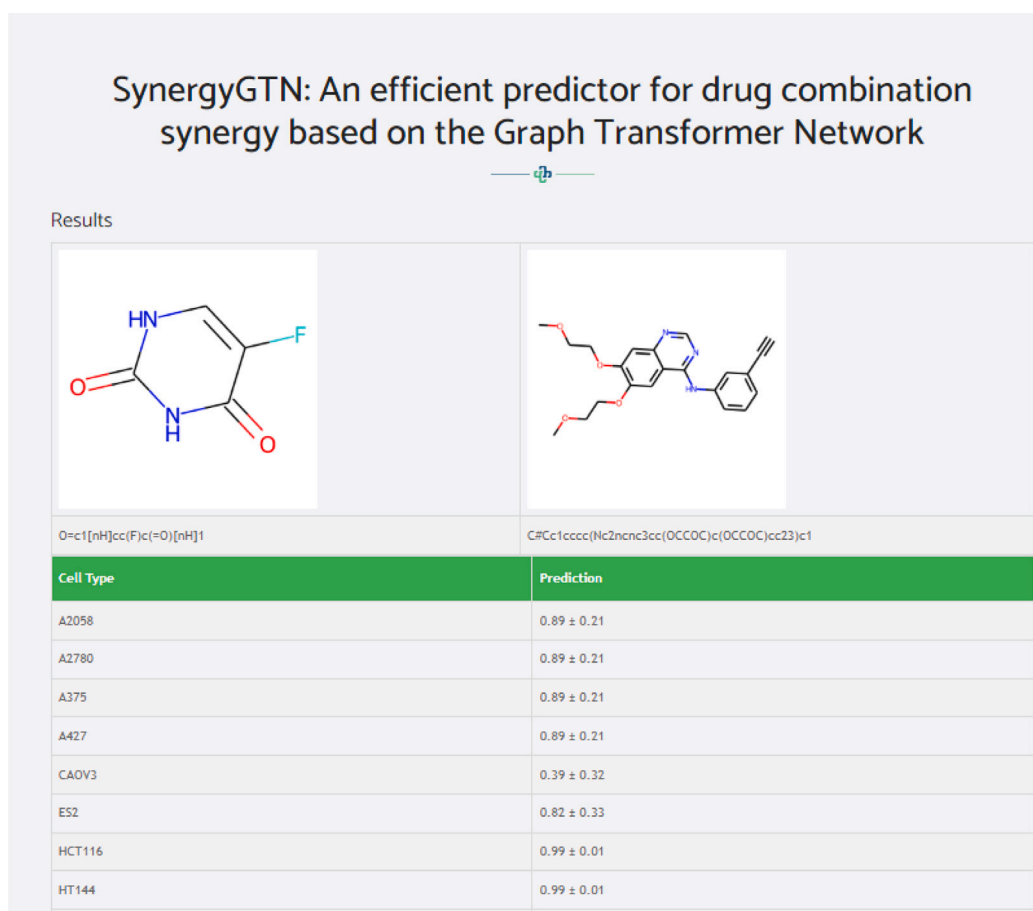
Fig. 11. The results interface of web-server illustrate the molecular structure of drugs and their synergistic probability for each cell-line.

The performance of the SynergyGTN model was found to surpass that of competitive state-of-the-art models. To validate the model's generalizability, a comprehensive set of test cases was applied. The results indicated a significant improvement in ROC-AUC on the leave-drug out and independent test datasets, respectively. Nevertheless, this work does have certain limitations, particularly in the area of dataset expansion. We are firmly dedicated to addressing this by incorporating additional datasets in future iterations, aiming to enhance the model's predictive sensitivity. Several publicly available datasets contain samples of protein and drug combination interactions, facilitating the prediction of synergy values. In subsequent developments, our goal is to design a more generalized deep learning framework capable of handling various input types for predicting synergistic drug combinations specific to a given target. Finally, to facilitate further investigation and research, we have introduced a user-friendly web server accessible to both the pharmaceutical industry and academic community.

CRediT authorship contribution statement

Waleed Alam: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Data curation, Conceptualization. **Hilal Tayara:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Investigation, Funding acquisition, Conceptualization. **Kil To Chong:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization.

Declaration of competing interest

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

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.compbiomed.2024.108007.

References

[1] J. Jia, F. Zhu, X. Ma, Z.W. Cao, Y.X. Li, Y.Z. Chen, Mechanisms of drug combinations: interaction and network perspectives, Nature Rev. Drug Discov. 8 (2) (2009) 111–128.

[2] P. Csermely, T. Korcsmáros, H.J. Kiss, G. London, R. Nussinov, Structure and dynamics of molecular networks: A novel paradigm of drug discovery: A comprehensive review, Pharmacol. Ther. 138 (3) (2013) 333–408.

[3] A.A. Borisy, P.J. Elliott, N.W. Hurst, M.S. Lee, J. Lehár, E.R. Price, G. Serbedzija, G.R. Zimmermann, M.A. Foley, B.R. Stockwell, et al., Systematic discovery of multicomponent therapeutics, Proc. Natl. Acad. Sci. 100 (13) (2003) 7977–7982.

[4] J. Lehár, A.S. Krueger, W. Avery, A.M. Heilbut, L.M. Johansen, E.R. Price, R.J. Rickles, G.F. Short Iii, J.E. Staunton, X. Jin, et al., Synergistic drug combinations tend to improve therapeutically relevant selectivity, Nature Biotechnol. 27 (7) (2009) 659–666.

[5] J. O'Neil, Y. Benita, I. Feldman, M. Chenard, B. Roberts, Y. Liu, J. Li, A. Kral, S. Lejnine, A. Loboda, et al., An unbiased oncology compound screen to identify novel combination strategies, Mol. Cancer Ther. 15 (6) (2016) 1155–1162.

[6] C. Guignabert, C. Phan, A. Seferian, A. Huertas, L. Tu, R. Thuillet, C. Sattler, M. Le Hiress, Y. Tamura, E.-M. Jutant, et al., Dasatinib induces lung vascular toxicity and predisposes to pulmonary hypertension, J. Clin. Invest. 126 (9) (2016) 3207–3218.

[7] K. Pang, Y.-W. Wan, W.T. Choi, L.A. Donehower, J. Sun, D. Pant, Z. Liu, Combinatorial therapy discovery using mixed integer linear programming, Bioinformatics 30 (10) (2014) 1456–1463.

[8] D. Day, L.L. Siu, Approaches to modernize the combination drug development paradigm, Genome Med. 8 (1) (2016) 1–14.

[9] L. He, E. Kulesskiy, J. Saarela, L. Turunen, K. Wennerberg, T. Aittokallio, J. Tang, Methods for high-throughput drug combination screening and synergy scoring, in: Cancer Systems Biology, Springer, 2018, pp. 351–398.

[10] M.P. Menden, D. Wang, Y. Guan, M.J. Mason, B. Szalai, K.C. Bulusu, T. Yu, J. Kang, M. Jeon, R. Wolfinger, et al., A cancer pharmacogenomic screen powering crowd-sourced advancement of drug combination prediction, 2018, 200451, BioRxiv.

[11] B. Zagidullin, J. Aldahdooh, S. Zheng, W. Wang, Y. Wang, J. Saad, A. Malyutina, M. Jafari, Z. Tanoli, A. Pessia, et al., DrugComb: An integrative cancer drug combination data portal, Nucleic Acids Res. 47 (W1) (2019) W43–W51.

[12] Y. Zhou, Y. Zhang, D. Zhao, X. Yu, X. Shen, Y. Zhou, S. Wang, Y. Qiu, Y. Chen, F. Zhu, TTD: Therapeutic target database describing target druggability information, Nucleic Acids Res. (2023) gkad751.

[13] J. Tang, K. Wennerberg, T. Aittokallio, What is synergy? The Saariselkä agreement revisited, Front. Pharmacol. 6 (2015) 181.

[14] R.G. Zinner, B.L. Barrett, E. Popova, P. Damien, A.Y. Volgin, J.G. Gelovani, R. Lotan, H.T. Tran, C. Pisano, G.B. Mills, et al., Algorithmic guided screening of drug combinations of arbitrary size for activity against cancer cells, Mol. Cancer Ther. 8 (3) (2009) 521–532.

[15] J.-H. Lee, D.G. Kim, T.J. Bae, K. Rho, J.-T. Kim, J.-J. Lee, Y. Jang, B.C. Kim, K.M. Park, S. Kim, CDA: Combinatorial Drug Discovery Using Transcriptional Response Modules, Public Library of Science San Francisco, USA, 2012.

[16] M. Bansal, J. Yang, C. Karan, M.P. Menden, J.C. Costello, H. Tang, G. Xiao, Y. Li, J. Allen, R. Zhong, et al., A community computational challenge to predict the activity of pairs of compounds, Nature Biotechnol. 32 (12) (2014) 1213–1222.

[17] X. Wang, L. Yang, C. Yu, X. Ling, C. Guo, R. Chen, D. Li, Z. Liu, An integrated computational strategy to predict personalized cancer drug combinations by reversing drug resistance signatures, Comput. Biol. Med. 163 (2023) 107230.

[18] A. Khan, H. Kim, L. Chua, PMED-net: Pyramid based multi-scale encoder-decoder network for medical image segmentation, IEEE Access 9 (2021) 55988–55998.

[19] T. Ilyas, Z.I. Mannan, A. Khan, S. Azam, H. Kim, F. De Boer, TSFD-net: Tissue specific feature distillation network for nuclei segmentation and classification, Neural Netw. 151 (2022) 1–15.

[20] M.U. Rehman, S. Cho, J. Kim, K.T. Chong, BrainSeg-net: Brain tumor MR image segmentation via enhanced encoder–decoder network, Diagnostics 11 (2) (2021) 169.

[21] R.R. Irshad, S. Hussain, S.S. Sohail, A.S. Zamani, D.Ø. Madsen, A.A. Alattab, A.A.A. Ahmed, K.A.A. Norain, O.A.S. Alsaiari, A novel IoT-enabled healthcare monitoring framework and improved grey wolf optimization algorithm-based deep convolution neural network model for early diagnosis of lung cancer, Sensors 23 (6) (2023) 2932.

[22] M. Sundermeyer, T. Alkhouli, J. Wuebker, H. Ney, Translation modeling with bidirectional recurrent neural networks, in: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP, 2014, pp. 14–25.

[23] S. Abimannan, E.-S.M. El-Alfy, Y.-S. Chang, S. Hussain, S. Shukla, D. Satheesh, Ensemble multifeatured deep learning models and applications: A survey, IEEE Access (2023).

[24] S. Abimannan, E.-S.M. El-Alfy, S. Hussain, Y.-S. Chang, S. Shukla, D. Satheesh, J.G. Breslin, Towards federated learning and multi-access edge computing for air quality monitoring: Literature review and assessment, Sustainability 15 (18) (2023) 13951.

[25] A. Graves, J. Schmidhuber, Framewise phoneme classification with bidirectional LSTM and other neural network architectures, Neural Netw. 18 (5–6) (2005) 602–610.

[26] M.U. Rehman, H. Tayara, K.T. Chong, DCNN-4mC: Densely connected neural network based N4-methylcytosine site prediction in multiple species, Comput. Struct. Biotechnol. J. 19 (2021) 6009–6019.

[27] A. Siraj, D.Y. Lim, H. Tayara, K.T. Chong, Ubicomb: A hybrid deep learning model for predicting plant-specific protein ubiquitylation sites, Genes 12 (5) (2021) 717.

[28] Z. Abbas, H. Tayara, K. Chong, ZayyuNet a unified deep learning model for the identification of epigenetic modifications using raw genomic sequences, IEEE/ACM Trans. Comput. Biol. Bioinform. (2021).

[29] S.D. Ali, W. Alam, H. Tayara, K. Chong, Identification of functional piRNAs using a convolutional neural network, IEEE/ACM Trans. Comput. Biol. Bioinform. (2020).

[30] P. Sidorov, S. Naulaerts, J. Ariey-Bonnet, E. Pasquier, P.J. Ballester, Predicting synergism of cancer drug combinations using NCI-ALMANAC data, Front. Chem. (2019) 509.

[31] K. Preuer, R.P. Lewis, S. Hochreiter, A. Bender, K.C. Bulusu, G. Klambauer, Deep-Synergy: predicting anti-cancer drug synergy with deep learning, Bioinformatics 34 (9) (2018) 1538–1546.

[32] T. Zhang, L. Zhang, P.R. Payne, F. Li, Synergistic drug combination prediction by integrating multiomics data in deep learning models, in: Translational Bioinformatics for Therapeutic Development, Springer, 2021, pp. 223–238.

[33] N.A. Asif, Y. Sarker, R.K. Chakrabortty, M.J. Ryan, M.H. Ahamed, D.K. Saha, F.R. Badal, S.K. Das, M.F. Ali, S.I. Moyeen, et al., Graph neural network: A comprehensive review on non-euclidean space, IEEE Access 9 (2021) 60588–60606.

[34] X. Bresson, T. Laurent, Residual gated graph convnets, 2018, 2017, URL https://Openreview.Net/Forum.

[35] J. Du, S. Zhang, G. Wu, J.M. Moura, S. Kar, Topology adaptive graph convolutional networks, 2017, arXiv preprint arXiv:1710.10370.

[36] C. Zhang, D. Song, C. Huang, A. Swami, N.V. Chawla, Heterogeneous graph neural network, in: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2019, pp. 793–803.

[37] Z. Cheng, C. Yan, F.-X. Wu, J. Wang, Drug-target interaction prediction using multi-head self-attention and graph attention network, IEEE/ACM Trans. Comput. Biol. Bioinform. 19 (4) (2021) 2208–2218.

[38] J. Wang, X. Liu, S. Shen, L. Deng, H. Liu, DeepDDS: deep graph neural network with attention mechanism to predict synergistic drug combinations, Brief. Bioinform. 23 (1) (2022) bbab390.

[39] E. Méndez, C.P. Rodriguez, M.C. Kao, S. Raju, A. Diab, R.A. Harbison, E.Q. Konnick, G.M. Mugundu, R. Santana-Davila, R. Martins, et al., A phase I clinical trial of AZD1775 in combination with neoadjuvant weekly docetaxel and cisplatin before definitive therapy in head and neck squamous cell carcinoma, Clin. Cancer Res. 24 (12) (2018) 2740–2748.

[40] H. Wang, L. Zhang, X. Yang, Y. Jin, S. Pei, D. Zhang, H. Zhang, B. Zhou, Y. Zhang, D. Lin, PUMA mediates the combinational therapy of 5-FU and NVP-BEZ235 in colon cancer, Oncotarget 6 (16) (2015) 14385.

[41] C. Zamagni, A. Martoni, L. Ercolino, M. Baroni, S. Tanneberger, F. Pannuti, 5-fluorouracil, epirubicin and cyclophosphamide (FEC combination) in advanced breast cancer, J. Chemother. 3 (2) (1991) 126–129.

[42] Y. Fu, G. Yang, P. Xue, L. Guo, Y. Yin, Z. Ye, S. Peng, Y. Qin, Q. Duan, F. Zhu, Dasatinib reduces 5-fu-triggered apoptosis in colon carcinoma by directly modulating Src-dependent caspase-9 phosphorylation, Cell Death Discov. 4 (1) (2018) 61.

[43] A.-R. Hanauske, J. Cassidy, J. Sastre, C. Bolling, R.J. Jones, A. Rakhit, S. Fettner, U. Brennscheidt, A. Feyereislova, E. Diaz-Rubio, Phase 1b dose escalation study of erlotinib in combination with infusional 5-fluorouracil, leucovorin, and oxaliplatin in patients with advanced solid tumors, Clin. Cancer Res. 13 (2) (2007) 523–531.

[44] M. Gilardi, Z. Wang, M. Proietto, A. Chillà, J.L. Calleja-Valera, Y. Goto, M. Vanoni, M.R. Janes, Z. Mikulski, A. Gualberto, et al., Tipifarnib as a precision therapy for HRAS-mutant head and neck squamous cell carcinomas, Mol. Cancer Ther. 19 (9) (2020) 1784–1796.

[45] G. Shepard, E.R. Arrowsmith, P. Murphy, J.H. Barton Jr., J.D. Peyton, M. Mainwaring, L. Blakely, N.A. Maun, J.C. Bendell, A phase II study with lead-in safety cohort of 5-fluorouracil, oxaliplatin, and lapatinib in combination with radiation therapy as neoadjuvant treatment for patients with localized HER2-positive esophagogastric adenocarcinomas, Oncologist 22 (10) (2017) 1152–e98.

[46] A. Saint, L. Evesque, A.T. Falk, G. Cavaglione, L. Montagne, K. Benezery, E. Francois, Mitomycin and 5-fluorouracil for second-line treatment of metastatic squamous cell carcinomas of the anal canal, Cancer Med. 8 (16) (2019) 6853–6859.

[47] P. Jin, C.C. Wong, S. Mei, X. He, Y. Qian, L. Sun, MK-2206 co-treatment with 5-fluorouracil or doxorubicin enhances chemosensitivity and apoptosis in gastric cancer by attenuation of Akt phosphorylation, Oncotargets Therapy (2016) 4387–4396.

[48] A. Mafi, A. Rezaee, N. Hedayati, S.D. Hogan, R.J. Reiter, M.-H. Aarabi, Z. Asemi, Melatonin and 5-fluorouracil combination chemotherapy: opportunities and efficacy in cancer therapy, Cell Commun. Signal. 21 (1) (2023) 33.

[49] R. Du, C. Huang, K. Liu, X. Li, Z. Dong, Targeting AURKA in cancer: molecular mechanisms and opportunities for cancer therapy, Mol. Cancer 20 (2021) 1–27.

[50] G.-R. Chang, C.-Y. Kuo, M.-Y. Tsai, W.-L. Lin, T.-C. Lin, H.-J. Liao, C.-H. Chen, Y.-C. Wang, Anti-cancer effects of zotarolimus combined with 5-fluorouracil treatment in HCT-116 colorectal cancer-bearing BALB/c nude mice, Molecules 26 (15) (2021) 4683.

[51] E. Martino-Echarri, B.R. Henderson, M.G. Brocardo, Targeting the DNA replication checkpoint by pharmacologic inhibition of Chk1 kinase: A strategy to sensitize APC mutant colon cancer cells to 5-fluorouracil chemotherapy, Oncotarget 5 (20) (2014) 9889.

[52] D.N. Kumar, A. Chaudhuri, D. Dehari, A. Shekher, S.C. Gupta, S. Majumdar, S. Krishnamurthy, S. Singh, D. Kumar, A.K. Agrawal, Combination therapy comprising paclitaxel and 5-fluorouracil by using folic acid functionalized bovine milk exosomes improves the therapeutic efficacy against breast cancer, Life 12 (8) (2022) 1143.

[53] M. Miyake, S. Anai, K. Fujimoto, S. Ohnishi, M. Kuwada, Y. Nakai, T. Inoue, A. Tomioka, N. Tanaka, Y. Hirao, 5-fluorouracil enhances the antitumor effect of sorafenib and sunitinib in a xenograft model of human renal cell carcinoma, Oncol. Lett. 3 (6) (2012) 1195–1202.

[54] L. de Mestier, T. Walter, H. Brixi, C. Evrard, J.-L. Legoux, P. de Boissieu, O. Hentic, J. Cros, P. Hammel, D. Tougeron, et al., Comparison of temozolomide-capecitabine to 5-fluorouracile-dacarbazine in 247 patients with advanced digestive neuroendocrine tumors using propensity score analyses, Neuroendocrinology 108 (4) (2019) 343–353.

[55] R. Nagourney, B. Sommers, S. Harper, S. Radecki, S. Evans, Ex vivo analysis of topotecan: advancing the application of laboratory-based clinical therapeutics, Br. J. Cancer 89 (9) (2003) 1789–1795.

[56] F. Nole, F. De Braud, M. Aapro, I. Minchella, M. De Pas, M. Zampino, S. Monti, G. Andreoni, A. Goldhirsch, Phase I–II study of vinorelbine in combination with 5-fluorouracil and folinic acid as first-line chemotherapy in metastatic breast cancer: A regimen with a low subjective toxic burden, Ann. Oncol. 8 (9) (1997) 865–870.

[57] G. Piro, M.S. Roca, F. Bruzzese, C. Carbone, F. Iannelli, A. Leone, M.G. Volpe, A. Budillon, E. Di Gennaro, Vorinostat potentiates 5-fluorouracil/cisplatin combination by inhibiting chemotherapy-induced EGFR nuclear translocation and increasing cisplatin uptake, Mol. Cancer Ther. 18 (8) (2019) 1405–1417.

[58] S. Vakili-Samiani, A.T. Jalil, W.K. Abdelbasset, A.V. Yumashev, V. Karpisheh, P. Jalali, S. Adibfar, M. Ahmadi, A.A.H. Feizi, F. Jadidi-Niaragh, Targeting Wee1 kinase as a therapeutic approach in hematological malignancies, DNA Repair 107 (2021) 103203.

[59] Y. Oku, N. Nishiya, T. Tazawa, T. Kobayashi, N. Umezawa, Y. Sugawara, Y. Uehara, Augmentation of the therapeutic efficacy of WEE 1 kinase inhibitor AZD 1775 by inhibiting the YAP–E2F1–DNA damage response pathway axis, FEBS Open Bio. 8 (6) (2018) 1001–1012.

[60] G. Karpel-Massler, M.-A. Westhoff, S. Zhou, L. Nonnenmacher, A. Dwucet, R.E. Kast, M.G. Bachem, C.R. Wirtz, K.-M. Debatin, M.-E. Halatsch, Combined inhibition of HER1/EGFR and RAC1 results in a synergistic antiproliferative effect on established and primary cultured human glioblastoma cells, Mol. Cancer Ther. 12 (9) (2013) 1783–1795.

[61] D. Morel, G. Almouzni, J.-C. Soria, S. Postel-Vinay, Targeting chromatin defects in selected solid tumors based on oncogene addiction, synthetic lethality and epigenetic antagonism, Ann. Oncol. 28 (2) (2017) 254–269.

[62] V.M. Rivera, R.M. Squillace, D. Miller, L. Berk, S.D. Wardwell, Y. Ning, R. Pollock, N.I. Narasimhan, J.D. Iuliucci, F. Wang, et al., Ridaforolimus (AP23573; MK-8669), a potent mTOR inhibitor, has broad antitumor activity and can be optimally administered using intermittent dosing regimens, Mol. Cancer Ther. 10 (6) (2011) 1059–1071.

[63] B.J. Quinn, H. Kitagawa, R.M. Memmott, J.J. Gills, P.A. Dennis, Repositioning metformin for cancer prevention and treatment, Trends Endocrinol. Metabol. 24 (9) (2013) 469–480.

[64] A. Mohammed, N.B. Janakiram, M. Brewer, R.L. Ritchie, A. Marya, S. Lightfoot, V.E. Steele, C.V. Rao, Antidiabetic drug metformin prevents progression of pancreatic cancer by targeting in part cancer stem cells and mTOR signaling, Transl. Oncol. 6 (6) (2013) 649–IN7.

[65] J.-W. Zhang, F. Zhao, Q. Sun, Metformin synergizes with rapamycin to inhibit the growth of pancreatic cancer in vitro and in vivo, Oncol. Lett. 15 (2) (2018) 1811–1816.

[66] B. Ramsundar, P. Eastman, P. Walters, V. Pande, Deep Learning for the Life Sciences: Applying Deep Learning to Genomics, Microscopy, Drug Discovery, and More, O'Reilly Media, 2019.

[67] G.Y. Di Veroli, C. Fornari, D. Wang, S. Mollard, J.L. Bramhall, F.M. Richards, D.I. Jodrell, Combenefit: an interactive platform for the analysis and visualization of drug combinations, Bioinformatics 32 (18) (2016) 2866–2868.

[68] J. Barretina, G. Caponigro, N. Stransky, K. Venkatesan, A.A. Margolin, S. Kim, C.J. Wilson, J. Lehár, G.V. Kryukov, D. Sonkin, et al., The cancer cell line encyclopedia enables predictive modelling of anticancer drug sensitivity, Nature 483 (7391) (2012) 603–607.

[69] D. Weininger, SMILES, A chemical language and information system. 1. Introduction to methodology and encoding rules, J. Chem. Inf. Comput. Sci. 28 (1) (1988) 31–36.

[70] G. Landrum, et al., RDKit: Open-source cheminformatics. 2006, Google Scholar (2006).

[71] Z. Wu, B. Ramsundar, E.N. Feinberg, J. Gomes, C. Geniesse, A.S. Pappu, K. Leswing, V. Pande, MoleculeNet: A benchmark for molecular machine learning, Chem. Sci. 9 (2) (2018) 513–530.

[72] T. Derrien, R. Johnson, G. Bussotti, A. Tanzer, S. Djebali, H. Tilgner, G. Guernec, D. Martin, A. Merkel, D.G. Knowles, et al., The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression, Genome Res. 22 (9) (2012) 1775–1789.

[73] W. Yang, J. Soares, P. Greninger, E.J. Edelman, H. Lightfoot, S. Forbes, N. Bindal, D. Beare, J.A. Smith, I.R. Thompson, et al., Genomics of drug sensitivity in cancer (GDSC): A resource for therapeutic biomarker discovery in cancer cells, Nucleic Acids Res. 41 (D1) (2012) D955–D961.

[74] L. Cheng, L. Li, Systematic quality control analysis of LINCS data, CPT: Pharmacomet. Syst. Pharmacol. 5 (11) (2016) 588–598.

[75] V.P. Dwivedi, X. Bresson, A generalization of transformer networks to graphs, 2020, arXiv preprint arXiv:2012.09699.

[76] Y. Shi, Z. Huang, S. Feng, H. Zhong, W. Wang, Y. Sun, Masked label prediction: Unified message passing model for semi-supervised classification, 2020, arXiv preprint arXiv:2009.03509.

[77] T. Akiba, S. Sano, T. Yanase, T. Ohta, M. Koyama, Optuna: A next-generation hyperparameter optimization framework, in: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2019, pp. 2623–2631.

[78] F. Khan, S. Hussain, S. Basak, M. Moustafa, P. Corcoran, A review of benchmark datasets and training loss functions in neural depth estimation, IEEE Access 9 (2021) 148479–148503.

[79] F. Khan, S. Hussain, S. Basak, J. Lemley, P. Corcoran, An efficient encoder–decoder model for portrait depth estimation from single images trained on pixel-accurate synthetic data, Neural Netw. 142 (2021) 479–491.