4-15-2023

SYNDEEP: a deep learning approach for the prediction of cancer drugs synergy.

Anna Torkamannia
Tabriz University

Yadollah Omidi
Nova Southeastern University, yomidi@nova.edu

Reza Ferdousi
Tabriz University

scientific reports

OPEN

SYNDEEP: a deep learning approach for the prediction of cancer drugs synergy

Anna Torkamannia[1], Yadollah Omidi[2] & Reza Ferdousi[1✉]

Drug combinations can be the prime strategy for increasing the initial treatment options in cancer therapy. However, identifying the combinations through experimental approaches is very laborious and costly. Notably, in vitro and/or in vivo examination of all the possible combinations might not be plausible. This study presented a novel computational approach to predicting synergistic drug combinations. Specifically, the deep neural network-based binary classification was utilized to develop the model. Various physicochemical, genomic, protein–protein interaction and protein-metabolite interaction information were used to predict the synergy effects of the combinations of different drugs. The performance of the constructed model was compared with shallow neural network (SNN), k-nearest neighbors (KNN), random forest (RF), support vector machines (SVMs), and gradient boosting classifiers (GBC). Based on our findings, the proposed deep neural network model was found to be capable of predicting synergistic drug combinations with high accuracy. The prediction accuracy and AUC metrics for this model were 92.21% and 97.32% in tenfold cross-validation. According to the results, the integration of different types of physicochemical and genomics features leads to more accurate prediction of synergy in cancer drugs.

Cancer is one of the most detrimental diseases with high mortality worldwide and is considered a challenging barrier in terms of increasing life expectancy[1]. The currently used treatments fail to satisfactorily cure the disease, in large part due to the emergence of drug resistance, and severe side effects[2,3]. In cancer therapy, the foremost target is usually specified in eradicating malignant cells using anticancer cytotoxic drugs through the induction of apoptosis and cell death in the diseased cells and tissues. However, cancer cells can develop escape mechanisms, initiate bypasses in the networks, and emerge alternative pathways for further proliferation and recurrence[4,5]. Remarkably, combinational pharmacotherapy might be the prime strategy that can intensify the therapeutic impacts of anticancer drugs and overcome the drug resistance mechanisms of cancer cells[6,7]. Such an approach can impose synergistic effects and potentially reduce the dose of monotherapy and drug resistance and avoid toxicity while the efficacy of the drug increases[8,9]. Hence, the discovery of the drug combination with synergistic effects is of eminent necessity in treating cancer.

Combinational pharmacotherapy should be provided and designed by in vitro and in vivo experiments based on the US Food and Drug Administration (FDA), the European Medicines Agency, and the World Health Organization guidelines[10]. It should be noted that predicting the possible drug combinations with synergistic effects solely via in vitro and/or in vivo experimentation is an extremely laborious task with no/trivial outcomes[11]. Besides, predicting the synergistic drug combination with clinical experiments is inefficient, time-consuming, and cost-intensive[7,12]. Therefore, in silico approaches can be reliable tools to facilitate and prioritize identifying synergistic drug candidates for experimental strategies[13].

The use of the in silico approaches as powerful tools has empowered the opportunity of exploring the wide variety of synergistic gaps with the diversity of chemical structures of drugs and genomic data from cancer cell lines[9,14]. Such approaches have paved the way for clinical trial experiments by providing accurate and adequate predictions[15]. Accordingly, over the last years, predicting drug combinations with synergistic effects has been increased by computational methods and has provided satisfactory outcomes. In this line, *Jiang* et al. developed a computational model to predict synergistic drug combinations on 39 cell lines using a graph convolutional network[16]. They used the multimodal network of the drug-drug synergy network, drug-target interaction network, and protein–protein interaction network. As a result of their analysis, the value of AUC was higher than 80%. The AuDNNsynergy method was proposed by *Zhang* et al.[17], which identified the synergistic effects based

[1]Department of Health Information Technology, School of Management and Medical Informatics, Tabriz University of Medical Sciences, Tabriz 51656/65811, Iran. [2]Department of Pharmaceutical Sciences, College of Pharmacy, Nova Southeastern University, Fort Lauderdale, FL 33328, USA. ✉email: ferdousi.r@gmail.com

1

on chemical structure and genomic data. As a result of this approach, the omics data enhance prediction accuracy, and the value of Pearson correlation obtained 0.74. DeepSynergy was developed to detect the drug combinations effect based on the deep neural network[18]. The method has utilized the chemical descriptors of drug pairs and gene expression profiles. DeepSynergy demonstrated the best performance compared to other state-of-art methods. *Yang* and colleagues proposed a model based on the functional similarity of target proteins to prioritize and stratify synergistic drug combinations[19]. This approach capitalized on the protein targets to predict synergy effects on breast cancer cell lines and experimentally validated the BRAF/insulin receptor combination in 48 colorectal cancer cell lines.

In the current study, a novel deep neural network model was proposed as an SYNDEEP to predict the synergistic drug combinations based on cancer cell lines and drug information. The model utilized the different features of physicochemical, genomic, protein–protein interaction, and protein-metabolite interaction information. Then the feature network was constructed based on eight various feature groups. The vector of features was developed according to the structure of the feature network. Finally, the feature vectors are fed into the deep neural network to achieve synergistic prediction results.

Results

This section summarizes the results of the synergy prediction model on the NCI-ALMANAC dataset. The 74 drugs were utilized in this study. The final experiment dataset consists of 22,228 drug pairs combinations of 74 unique drugs against 60 cell lines. Also, the final features for one drug consisted of 777 features, and total number of the features for a pairs of drugs were 1614 with similarity measurement and cell lines. In addition, the diverse hyperparameters of the algorithms were implemented to adjust the optimal state. As a result, the final optimization parameters are as follows: SNN: hidden layer = 1, dropout rate = 0.8, epochs = 300, learning rate = 10^{-2}, KNN: n = 4, metric = 'Minkowski', p = 2, SVM: degree = 3, kernel = 'linear', cache size = 200; RF: max depth = 4, n estimators = 10; GBCs: max depth = 3, n_estimators = 100.

Performance of deep neural network with different feature groups. SYNDEEP was implemented with different feature group to evaluate performance. Furthermore, to confirm the deep neural network is robust to training, we implemented tenfold cross-validation (CV) for all six feature group groups. Table 1 summarizes the performance results of six groups. The results highlighted that the four groups achieved accuracy over 90%, and the two remaining groups achieved accuracy of 89%. Best outperforming score was achieved by DC_networkII with the highest accuracy of 92.21%, and the second highest accuracy was achieved by DC_networkI with 92.16%. By including the protein–protein interaction and protein-metabolite interaction similarity score on DC_networkI to generate DC_networkII, the accuracy score slightly increased.

It can be seen from the results that by adding gene mutation and gene expression information to the DT_CL group, accuracy slightly dropped. However, accuracy increased significantly when combined with other feature groups. To further investigate the effect of gene mutation and gene expression on synergy prediction, we left out their information. Substantial differences were not observed in the forecast's performance by excluding the information (*ACCU.* = 92.21%). In drug investigations, gene mutations and gene expression are highly predictive[20]. In this sense, adding their information seems crucial to predicting the synergy effect.

The DC_networkII feature set group was selected as a complete feature set group. To better investigate the model's prediction ability, we obtained the values of sensitivity, specificity, precision, F_{-score}, Matthews correlation coefficient (MCC), AUC, and Cohen's kappa. The considered evaluation metrics comprehensively reflect the model performance in the values of F_{-score}, MCC, and AUC were obtained 92.15%, 84.41% and 97.32%, respectively. In addition, the Kappa value represents the model's inter-rater reliability with a score of 84.37%.

Figure 1 shows the area under the receiver operating characteristic curve (ROC-AUC) and the accuracy in evaluating the performance of each fold. As can be seen from the graphs, the curves obtained by ten folds had the most covering coordinate space. The result of the tenfold cross-validation indicates that SYNDEEP provided a more reliable performance on the entire feature set.

Performance comparison of the various machine learning methods. We reported the primary evaluation criteria for SYNDEEP, SNN, KNN, SVM, RF, and GBCs in Table 2. We observed from the table that SYNDEEP achieved the highest results in the entire evaluation criteria compared to other methods. The three models with the highest performance based on accuracy were (i) SYNDEEP, (ii) the GBCs, and (iii) the KNN. Accordingly, SYNDEEP accuracy was 92.21%, which is approximately 3.8% and 4.23% higher than the GBCs

Group name	ACCU. (%)	Sen. (%)	Spec. (%)	Prec. (%)	F_{score} (%)	MCC. (%)	AUC (%)	Kappa (%)
DT_CL	90.08	89.86	90.29	90.23	90.05	80.16	96.14	80.16
DCGM	89.78	87.84	91.37	91.38	89.57	79.63	95.80	79.57
DCGME	89.63	87.03	92.22	91.78	89.34	79.36	95.94	79.26
DCGMEM	90.21	89.32	91.10	90.92	90.12	80.44	95.96	80.43
DC_networkI	92.16	91.31	93.00	92.88	92.09	84.33	97.14	84.32
DC_networkII	92.21	91.53	92.91	92.80	92.15	84.41	97.32	84.37

Table 1. The overall results of the performance on six groups of features.
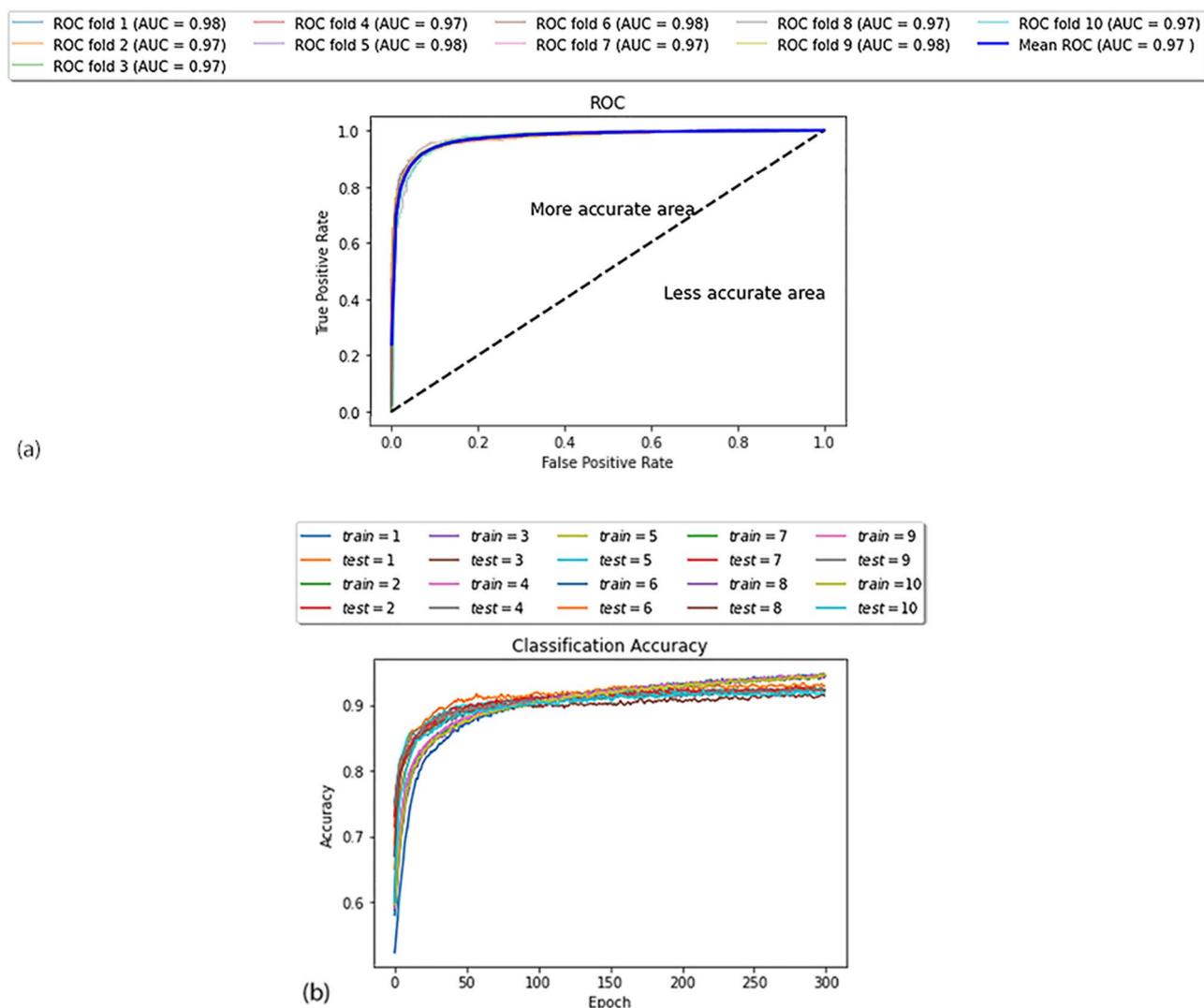
Figure 1. The ROC and accuracy diagram of tenfold-cross validation. (**a**) The ROC curves and (**b**) accuracy diagram were generated on the entire feature set.

Method	ACCU. (%)	Sen. (%)	Spec. (%)	Prec. (%)	F_{score} (%)	MCC. (%)	AUC (%)	Kappa (%)
Gradient boosting classifiers	88.41	88.73	88.10	87.97	88.35	76.83	0.94	76.83
SYNDEEP	92.21	91.53	92.91	92.80	92.15	84.41	0.97	84.37
Random forest	74.85	80.15	71.15	66.00	72.39	50.49	0.84	49.69
Support vector machines	84.23	84.42	84.04	83.92	84.17	68.46	0.90	68.46
Shallow neural network	86.61	86.48	86.74	86.71	86.59	73.23	0.93	73.23
k-nearest neighbor	87.98	84.35	92.36	93	88.47	76.39	0.88	75.99

Table 2. The overall results of the different state-of-art methods.

model, and the KNN, respectively. Among the models examined, the RF achieved poor accuracy (74.85%). Also, SYNDEEP achieved remarkable results in Matthews Correlation Coefficient (*MCC.* = 84.41%), representing the correlation between predictions and labels. Notably, SYNDEEP obtained a high Kappa score (*Kappa.* = 84.37%) while the RF had the lowest score (*Kappa.* = 49.69%).

The ROC curve was used as another evaluation measure. The ROC curve plots the true-positive rate (TPR) versus the false-positive rate (FPR). The area under the ROC curve (AUC) was calculated to reflect predictive accuracy. Figure 2 shows the visualization of the area under the ROC curve for SYNDEEP and the other classifiers to evaluate the performance of binary predictions. The AUCs for SYNDEEP, GBCs, and SNN were 0.97, 0.94 and 0.93, respectively. This shows SYNDEEP's potential for synergy prediction in drug combinations.
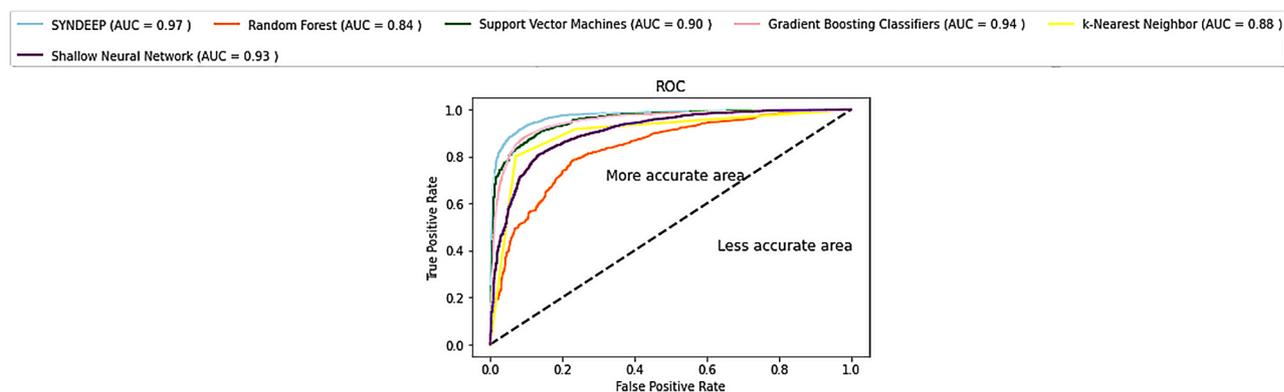
Figure 2. ROC curve comparison of different machine learning methods.

We conducted McNemar's Test to assess classifiers' performance. The null hypothesis of this test states that the probability of the synergistic effect being correctly identified is equal to the probability of the non-synergy effect being correctly identified. Also, the probability of the synergy effect being incorrectly classified is equal to the probability of the non-synergy effect being incorrectly classified. The p-value is calculated, and a p-value < 0.05 is considered significant, thus rejecting the null hypothesis. Hence, we obtained $X^2 = 154.0$, with a p-value of 0.0, which is below the set significance threshold (p-value = 0.05, degree of freedom = 1) and leads to rejection of the null hypothesis; we can conclude that SYNDEEP's performance is reliable. The values of McNemar's Test for GBC, RF, SVM, SNN, and KNN were achieved at 248.0, 363.0, 344.0, 295.0, and 154.0, respectively. Furthermore, RF and KNN obtained p-value < 0.05, while the comparison between GBC, SVM, and SNN gave a non-significant p-value > 0.05, so the null hypothesis was accepted. Generally, SYNDEEP achieved substantial performance compared with other methods to predict drug synergy.

Discussion

Cancer remains the primary cause of morbidity and mortality worldwide, despite the pharmaceutical and clinical research in cancer treatment[21]. Furthermore, the mono-therapeutic strategies are inefficient in cancer treatment because the targets are single proteins or pathways, and drug resistance occurs in these strategies[22,23]. Accordingly, combining pharmacotherapy with synergistic effects is a promising approach[6,7]. However, it is infeasible to identify and examine the possible combinations of anticancer drugs[24]. Therefore, in silico approaches are beneficial to overcome the limitations.

Hence, we have developed a novel deep neural network model as an SYNDEEP that accurately predicts the synergistic effect of drug combinations for cancer cell lines. In this study, a data-driven approach to predict the synergy effects of drug combinations.

Previous studies used a small subset of genomic and drug-related data to predict synergy[16–19,25,26]. Some studies, such as Jiang's model[16], have considered drug-drug synergy, drug-target interaction, and protein–protein interaction networks based on 39 cell lines. Moreover, the DeepSynergy[18] model has been defined based on chemical descriptors and genomic features on 39 cell lines. AuDNNsynergy[17] is another drug synergy model that uses multi-omic and chemical data for synergy prediction. The most striking result among previous studies emerged from the studies that utilized genomic and drug information.

To the best of our knowledge, there is no study in the literature that utilized the comprehensive feature set of genomic, physicochemical, and drug information to predict synergy. Utilizing genomic, PPI, PMI, and physicochemical data is important for overcoming drug resistance. The genomic data would be practical for predicting the drug combination synergy[27–29]. In this study, a novel comprehensive feature set of genomic, physicochemical, and drug information (i.e. drug-target, protein–protein interaction, protein-metabolite interaction, gene mutation, gene expression, differential methylation, chemical structure, and cell lines) was constructed. Therefore, we used the comprehensive feature set of various data types to optimize the model's performance.

The modification in depth and width of model architecture is the pivotal factor in deep neural networks to enhance performance[30]. Generally, applying this property of deep neural networks leads to developing highly impactful architectures for diverse tasks[30,31]. Several previous studies have utilized network-based and machine-learning methods, such as random forests and extremely randomized trees[14,25,32–34]. In contrast, rare studies used deep neural networks that focused on a small range of features[18,35–40]. Thus, we have used comprehensive biological features with a deep neural network. This strategy obtained satisfactory results, as shown in Table 2.

The Kappa value of SYNDEEP was 84.37%. The Kappa value demonstrated that classes have independent distributions. Few related studies in the literature have calculated the kappa coefficient[16–18], while most studies only reported the accuracy/ AUC value[14,34,37,38]. The kappa values of DeepSynergy, AuDNNsynergy, and Jiang's model were 0.51, 0.51, and 0.584, respectively. The kappa value in previous studies was just over 0.50%[16–18], which may represent fair to a reasonable agreement beyond chance.

Previous studies used different datasets to predict the drug combination effects based on the deep neural network[17,39–41]. The NCI-ALMANAC and Merck datasets are the two large-scale pan-cancer datasets employed in most drug synergy prediction studies. For example, among these studies, the performance of the DeepSynergy, AuDNNsynergy, TranSynergy[39], and Jiang's models has been evaluated on the Merck dataset. The nature

of these studies in terms of recruited information (e.g. multi-omics, phenotypic and biophysical features) was different. For example, some studies utilized drug pathways information, gene expression information, and/or microRNA information. Several previous studies reported the accuracy value, some of which stated the MSE value[17,18,39–41]. The accuracy values of DeepSynergy, AuDNNsynergy, and Jiang's models were 0.92%, 0.93% and 0.919%, respectively.

In this study, SYNDEEP's performance has been tested on the NCI-ALMANAC dataset. Moreover, among the previous studies, the SYNPRED[40] and *Xia* et al.[35] utilized the NCI-ALMANAC dataset based on a deep neural network. The accuracy value of SYNPRED was 0.85%, while Xia's model[35] reported the R^2, Pearson correlation, and Spearman correlation values.

In this study, the accuracy value of the novel deep neural network was 92.21%. While the AuDNNsynergy has reported a value score of 0.93% but the kappa value was 0.51%. However, the results of the current study highlighted the strength of the presented SYNDEEP, which was the large volume of features and network of features. Hence, the proposed model achieved a high score in kappa and accuracy values *(Kappa.* = 84.37%, *ACCU.* = 92.21%).

Evidence of in silico methods represented that combining the biological, chemical, and phenotypic properties had a substantial role in the modeling and prediction procedures[42,43]. According to the performance result of feature sets, it was evident that combining the genomic and physicochemical features had significant effects on improving the performance.

The network-based approaches evaluate the interactions among the various agents. Agents in the prediction of drug effects have different natures. Several studies have used network-based analysis to predict the drug combinations effect[19,25,26,34,37].

Previous studies prove that integrating the genomic data, drug targeting networks, chemical structure, cell lines, genomic profiling data, gene mutation, gene expression, and protein interaction based on the network approach has a significant role in synergy prediction[19,25,26,34,37].

Therefore, we have investigated the effect of drug interactions among the diverse feature agents (i.e. drug-target, protein–protein interaction, metabolite-protein interaction, gene expressions, gene mutations, chemical structures, cell lines, and differential methylation). However, the previous studies investigated and analyzed the network of interaction while not examining the network by deep neural network[19,25,26,34,37].

Here, we have used a network of features comprised of different relationships based on the nature of the information for the first time. As shown in Table 2, the features network has achieved high performance and improved the deep neural network model.

Table 2 has proved that deep neural networks had superior outcomes compared to other state-of-art methods by the network of features.

Deep neural networks have significant capabilities in complex dimensional spaces by various structures[44,45]. Therefore, deep neural network algorithms have been proven to be a more capable system and have a better accuracy rate than other algorithms in classifying synergistic effects in drug combinations.

Conclusion

The current study proposes SYNDEEP for the prediction of synergy in drugs combination. It is well known that the critical factor in predicting drugs effect is extracting practical features, so the main superiority of this study was the network of extensive features. SYNDEEP integrated the various type of physicochemical and genomics information to generate the comprehensive features set for successfully and robustly predicting drug combinations' effects. SYNDEEP obtained 92.21% prediction accuracy utilizing the tenfold cross-validation in the NCI-ALMANAC dataset. In the experiment, we compared SYNDEEP, SNN, KNN, SVM, RF, and GBC methods that SYNDEEP achieved superior performance. The result indicated that the deep neural network is the competent learning technique to determine synergy. For better and more accurate predictions in the future, updating and complementing the information of cell lines and targets of drugs are needed. Although the *in silico* approaches provide substantial insights into in vivo and in vitro experiments and accelerate the procedures, well-designed experimental investigations are required to prove the data resulting from *in silico* computational analyses.

Materials and methods

The first step of the proposed deep neural network was data preparation which consisted of data acquisition, features extraction, and network of features construction. The next step was prediction model construction which comprised the synergy prediction and model evaluation. Figure 3 illustrates a schematic representation of how the model uses the distinct features to predict synergistic drug combinations. The network of features is recruited as input for a deep neural network to predict the synergistic effect. The network has been constructed from the sequence of physicochemical, genomic, protein–protein interaction, and protein-metabolite interaction data. The data has been produced from different databases based on drug-pairs information. The following figure illustrates the overall steps of the study.

Data acquisition

In this study, the NCI-ALMANAC dataset was used[46]. This dataset is the most well-known anticancer drug combinations effect dataset, covering the combinations of drug pairs against the NCI-60 cell lines with different concentrations. The drugs in the dataset contain FDA-approved drugs in oncology. The NCI-60 panel comprises 60 human tumor cell lines derived from nine various tumor types. The NCI-ALMANAC introduced the ComboScore to quantify the combination benefit of pairs of drugs which initially modified the version of the Bliss independence score. The ComboScore defines the combination activity by three classes of positive, negative, and
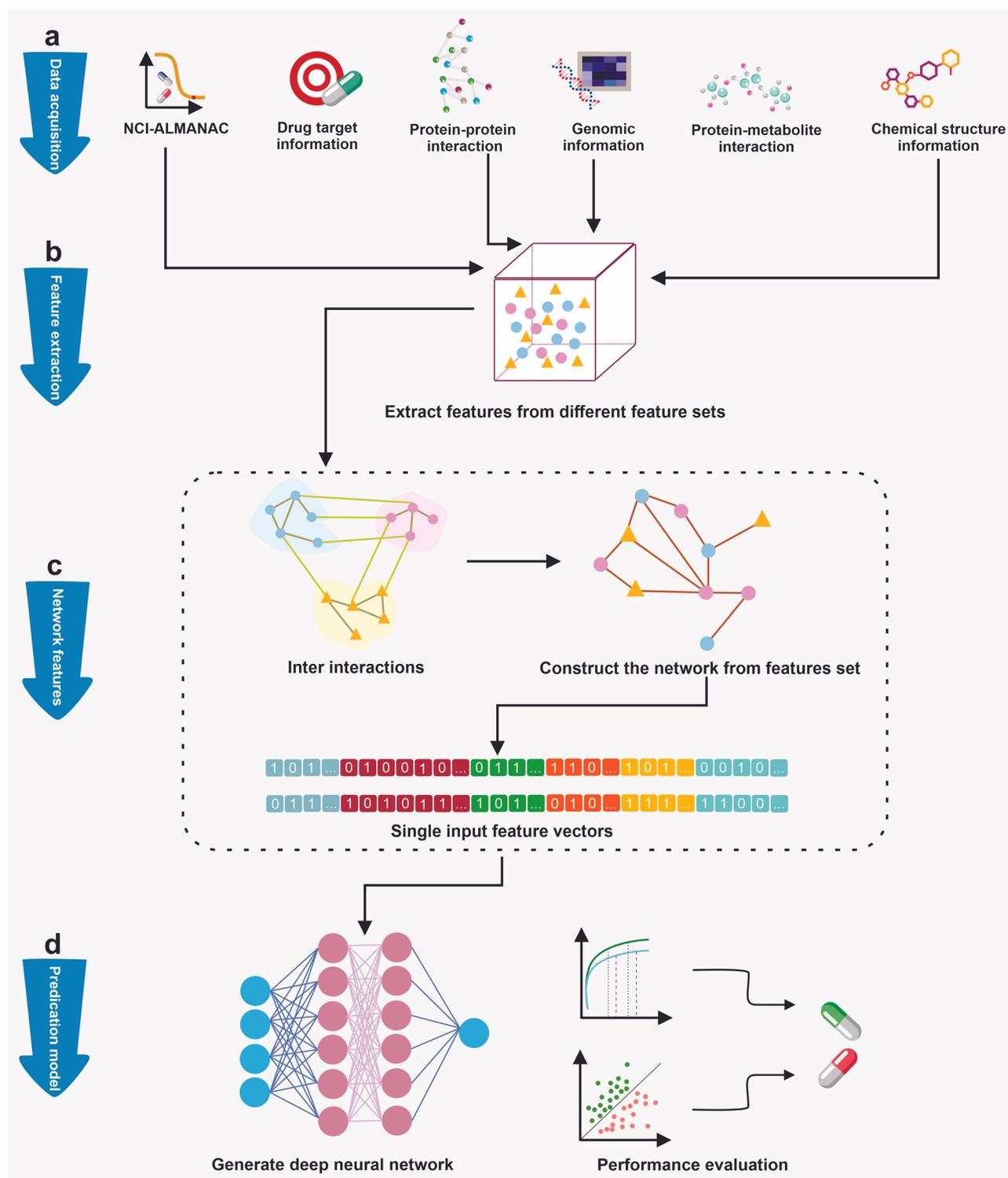
Figure 3. Overall steps to construct drug synergy prediction model. (**a**) eight feature sets were used to predict the synergy effect, drug-target, protein–protein interaction, protein-metabolite interaction, genomic features (gene expression, mutations, and differential methylation), chemical structure, and cell lines. (**b**) Features from different sets were extracted to generate the network. (**c**) The heterogeneous network was created from various features. (**d**) The prediction model was constructed to predict drug combinations' effects.

zero values. Hence, the combination of drug pairs with positive values represents to be synergistic, whereas the negative values indicate antagonist combinations and zero values correspond to additive combinations.

The ComboScore has defined the score range from 1 to 200 as synergistic, and − 1 to − 228 has specified the antagonistic effect. The score of drug combinations was converted to 1 and 0 to reduce the computational complexity. One represents the synergistic drug-pairs combination, and 0 illustrates the antagonistic drug-pairs combination. By sorting the dataset, the number of antagonist combination effects was more than the synergistic combinations. To avoid the occurrence of an imbalanced dataset, we selected an equal number of drug pairs based on the number of synergistic combinations.

In the present study, different feature groups have been extracted for each drug in the dataset. The extracted feature groups are as follows:

Extraction of the drug-target interactions. The total indexed drugs and related information in DrugBank(version 5.1.8.) was downloaded[47]. In the next step, the FDA-approved anticancer drugs were extracted from the XML file of DrugBank. Hence, the list of drug-target interactions (DT) for each anticancer drug based on protein targets was elicited. For each drug, the pairs of (Di, DTj) were created, where D_i was observed in the approved drug list, and DT_j was extracted as a protein target.

Extraction of the protein–protein interactions. The critical resource of protein–protein interaction data is the STRING (Search Tool for the Retrieval of Interacting Genes/Proteins)[48] database which is a reliable tool for providing the properties of proteins. As the first step, the sequences of proteins and accessions were retrieved from STRING (version 11.5). Then, protein–protein interaction (PPI) was extracted based on every protein related to DTs. Next, the pair of DTs and PPI was created as (DT_i and PPI_j), where $DT_i \in$ {protein target reported for $drug_i$} and $PPI_j \in$ {PPI stated for protein $target_j$}.

Extraction of the genomic features (gene expression, mutations, and differential methylation). The Catalogue Of Somatic Mutations In Cancer (COSMIC)[49] database represents worthwhile information on genomic sequences and microarray expression data. The COSMIC database provides extensive data for individual genes/ cell lines. The COSMIC (version 95-24) has provided the individual files for gene expression (GE), mutation (GM), and differential methylation (DM), which all relevant files downloaded. However, the vast majority of information was not relevant to predicting synergy, which might result in a bias. Hence, we filtered the data based on the cell lines and protein targets. As a result, we had the pairs of (D_i, GE_j), (D_i, GM_k), and (D_i, DM_n), where D_i was considered in the list of drugs, GE_j, GM_k and DM_n were extracted as gene expression, mutation, and differential methylation.

Extraction of the protein-metabolite interaction. The Human Metabolome Database (HMDB)[50] was utilized to construct the protein-metabolite interaction(PMI) data in this study. The metabolite and protein data lists were downloaded from HMDB version 4.0. The list of the corresponding metabolites for every protein in DTs was extracted. Then, the pair of (DT_i, M_j) was considered, where $DT_i \in$ {protein target reported for $drug_i$} and $M_i \in$ {metabolite reported for $protein_j$}.

Extraction of the chemical structure. We used the Morgan fingerprint counts, total polar surface area, molecular weight, logP, aliphatic and aromatic rings, H-bond donors, and acceptors as chemical structures (CF). The Morgan fingerprint counts are designed to represent the number of times a particular substructure detects in the molecule. To integrate the features into the deep neural network model, we used the RDKit library. As a result, chemical features for each drug were obtained.

Dimension reduction based on the similarity measure. In total, there were 19,888 features for one drug except for cell lines. The entire features for drug pairs have consisted of 39,776 features with cell lines. These extensive features are considered the high dimensionality problem in constructing the synergy prediction model. The PPI and PMI comprised the high-dimension features among the other feature groups. Hence, the Russell-Rao similarity measure[51] on PPI and PMI on pairs of drugs was used to overcome this issue. Therefore, the similarity values of PPI and PMI were included in the vector instead of the total features. The Russell-Rao similarity measure is defined as follows:

$$S_{Russell-Rao} = \frac{x}{d}$$

where x denotes the number of features where the values of vector one and vector two are one, which means positive matches, d denotes the total sum of the length of vectors.

Feature network construction. The network feature in SYNDEEP (Fig. 3) was a heterogeneous network consisting of eight components: DT, PPI, PMI, GM, GE, DM, CF, and cell line (CL). In the previous section, it was described how to build each group of features.

In this study, different types of features were utilized; hence we constructed an undirected network U = (V, E), where V is a set of N nodes, in which V = {$V^{DT} \cup V^{PPI} \cup V^{PMI} \cup V^{GM} \cup V^{GE} \cup V^{DM} \cup V^{CF} \cup V^{CL}$} is composed of cell lines, and seven sets of feature components, and E is a set of M edges such as protein–protein interaction and drug target. These N nodes have node feature vectors $a_1, a_2, a_3, \ldots, a_N \in \mathbb{R}^d$ where d is the dimension of the feature vector. As for the edges, for example, (v_i, v_j) represents the link between node v_i and v_j. The total set of feature vectors has been combined to generate a feature matrix. As shown in Fig. 4, was defined feature matrix F as:

$$
F = \begin{cases}
F^{D-DT} \in \mathbb{R}^{N_{V^D} \times N_{V^{DT}}} \\
\quad if\ v_i \in V^D, v_j \in V^{DT}. \\
F^{DT-PPI} \in \mathbb{R}^{N_{V^{DT}} \times N_{V^{PPI}}} \\
\quad if\ v_i \in V^{DT}. v_j \in V^{PPI}. \\
F^{DT-PMI} \in \mathbb{R}^{N_{V^{DT}} \times N_{V^{PMI}}} \\
\quad if\ v_i \in V^{DT}. v_j \in V^{PMI}. \\
F^{DT-GM} \in \mathbb{R}^{N_{V^{DT}} \times N_{V^{GM}}} \\
\quad if\ v_i \in V^{DT}. v_j \in V^{GM}. \\
F^{DT-GE} \in \mathbb{R}^{N_{V^{DT}} \times N_{V^{GE}}} \\
\quad if\ v_i \in V^{DT}. v_j \in V^{GE}. \\
F^{DT-DM} \in \mathbb{R}^{N_{V^{DT}} \times N_{V^{DM}}} \\
\quad if\ v_i \in V^{DT}. v_j \in V^{DM}. \\
F^{D-CF} \in \mathbb{R}^{N_{V^D} \times N_{V^{CF}}} \\
\quad if\ v_i \in V^D. v_j \in V^{CF}. \\
F^{D-CL} \in \mathbb{R}^{N_{V^D} \times N_{V^{CL}}} \\
\quad if\ v_i \in V^D. v_j \in V^{CL}.
\end{cases}
$$

where D is a set of drugs that are related to cancer drugs. Hence, $F_{ij} \in F^{D-DT} = 1$ if a D node v_i and a DT node v_j are related according to the drug and drug target, $F_{ij} = 0$ otherwise, which means between drug and drug target there is no interaction. $F^{DT-PPI} \in [0, n]$ is indicated the similarity values between pairs of PPI nodes, in which n represents the non-zero values. $F^{DT-PMI} \in [0, n]$ is represented the non-zero values of similarity between pairs of PMI nodes. $F_{ij} \in F^{DT-GM} = 1$ indicates DT i is related according to the drug target and gene mutation with j-th GM, $F_{ij} = 0$ otherwise. $F_{ij} \in F^{DT-GE} = 1$ represents the i-th DT is associated with a j-th GE, $F_{ij} = 0$ otherwise. $F_{ij} \in F^{DT-DM} = 1$ if the interaction between a DT and DM nodes has been observed, $F_{ij} = 0$ otherwise. In this study, to represent the chemical features of drugs $F^{D-CF} \in [0, n]$ is utilized. where D − CF is the chemical features of drugs, and n indicates the non-zero values of features. $F_{ij} \in F^{D-CL} = 1$ if a D node v_i and a CL node v_j are related according to the NCI-ALMANAC data, $F_{ij} = 0$ otherwise.
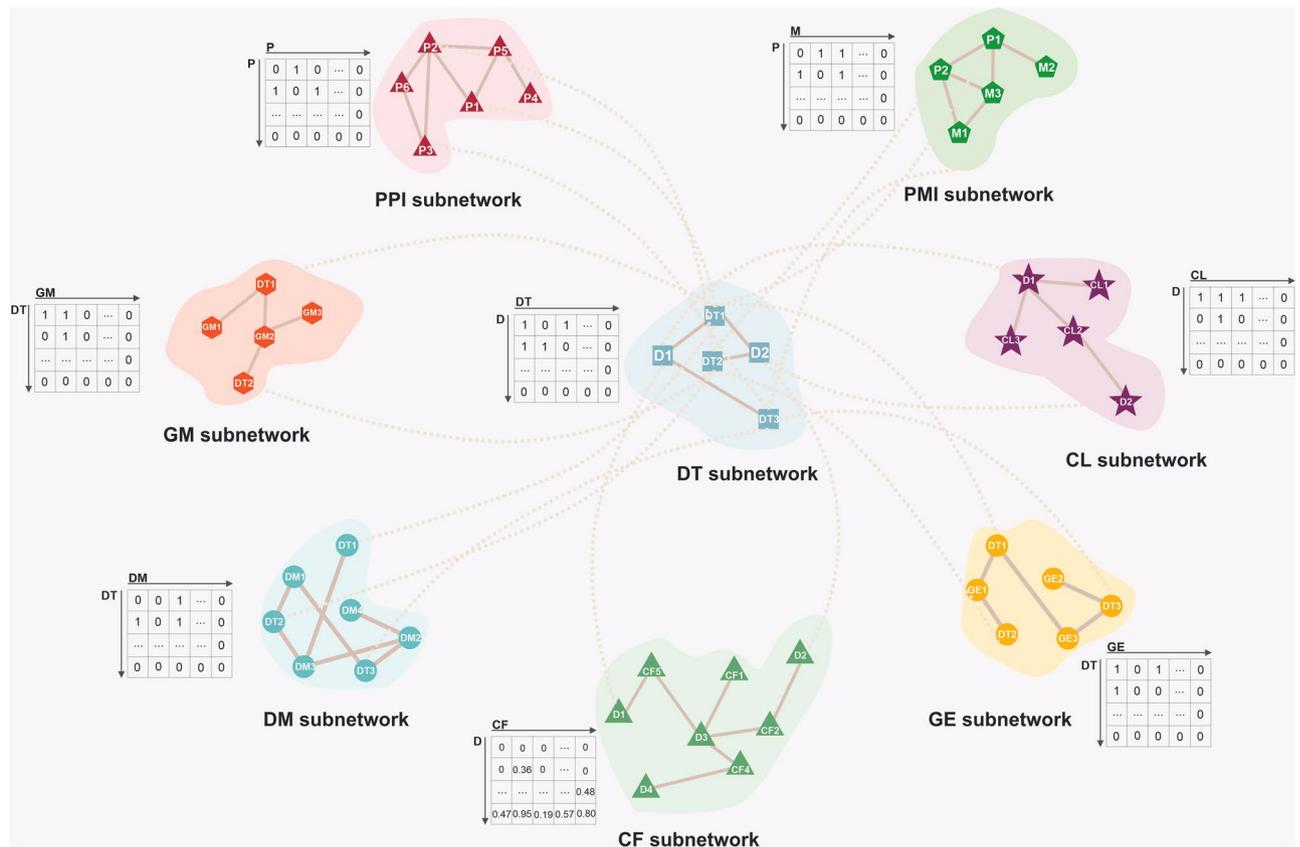
Figure 4. Construction of the heterogeneous network. The heterogeneous feature matrix was generated from drug-target, protein–protein interaction, protein-metabolite interaction, gene mutation, gene expression, differential methylation, chemical structures, and cell lines. The matrix values comprised zero, and non-zero values.

To utilize the topological structural relations among the cell lines, and seven sets of features embedded in the network U, we formulated the multi-structure of topology as single input feature vectors (Fig. 5). As an input feature vector, the vector was divided into eight different groups: seven sets of features, and cell line related features. As mentioned above, the value of each group consists of two sets: (i) zero, and (ii) non-zero values.

Feature groups construction. The various features group was established to assess the benefits of features on classifying the drug synergy, along with the network of features. Six groups of features were constructed. Where, the first group consisted of drug targets and cell lines, and the last group was the network of features comprised of total feature groups. Table 3 details the information on groups.

Construction of deep neural network model. The Multi-layer perceptron (MLP) was selected to build the model for synergy prediction. The shape of the MLP model was conic. The different hyperparameter settings were considered (i.e. number of layers, number of neurons, and the learning rate). The six diverse different layers (3, 4, 5, 6, 7) were tested. Various numbers of neurons (1024, 512, 256, 128, 64, 32) and learning rates (10^{-1}, 10^{-2}, 10^{-5}, 10^{-8}, 10^{-75}) were examined. After comparing the outcomes of different deep networks, the best result was observed from the five-layer network for binary classification of the drug synergy. The selected model had an input layer, 3 hidden layers, and an output layer (Fig. 6). The two activation functions have been used in the model: the ReLu and the sigmoid functions.

$$ReLu(x) = \begin{cases} x \; if \; x \geq 0 \\ 0 \; otherwise \end{cases}$$
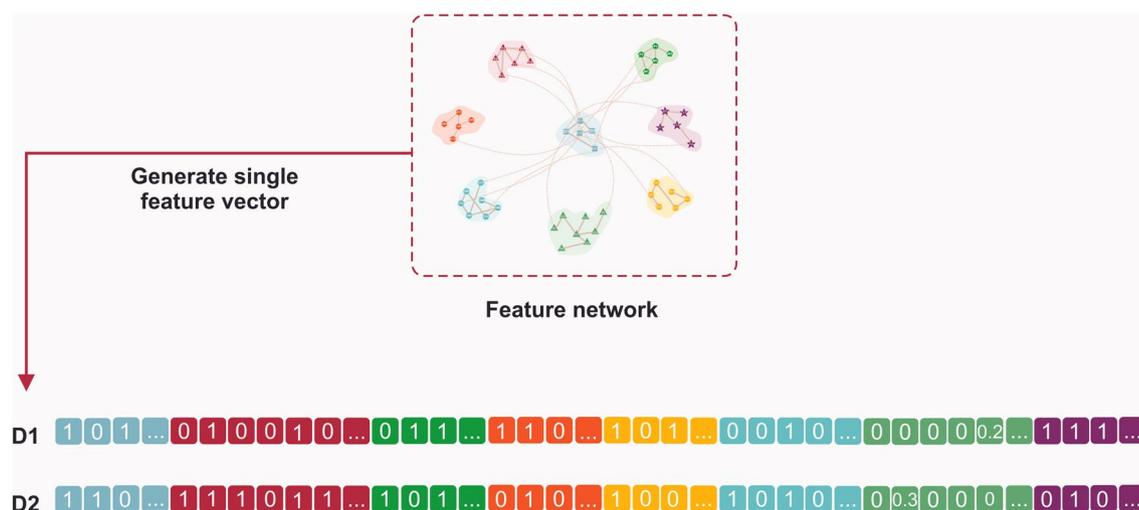
Figure 5. Single input feature vectors. Single input feature vectors derived from multi-topological structural relations among the drug-target, protein–protein interaction, protein-metabolite interaction, gene mutation, gene expression, differential methylation, chemical structures, and cell lines.

Group	Group name	Features	Number of features
Group 1	DT_CL	Drug-target, Cell-lines	393
Group 2	DCGM	Drug-target, Cell-lines, Gene-mutation	488
Group 3	DCGME	Drug-target, Cell-lines, Gene-mutation, Gene-Expression	752
Group 4	DCGMEM	Drug-target, Cell-lines, Gene-mutation, Gene-Expression, Differential-methylation	1086
Group 5	DC_networkI	Drug-target, Cell-lines, Gene-mutation, Gene-Expression, Differential-methylation, Chemical-features	1612
Group 6	DC_networkII	Drug-target, Cell-lines, Gene-mutation, Gene-Expression, Differential-methylation, Chemical-features, Protein–Protein Interaction, Protein-metabolite Interaction	1614

Table 3. Feature group description. The name of each group has been determined based on their feature groups. *DT_CL* drug targets and cell lines, *DCGM* drug targets, cell lines, and gene mutation, *DCGME* drug targets, cell lines, gene mutation, and gene expression, *DCGMEM* drug targets, cell lines, gene mutation, gene expression, and differential methylation, *DC_networkI* drug targets, cell lines, gene mutation, gene expression, differential methylation, and chemical features, *DC_networkII* drug targets, cell lines, gene mutation, gene expression, differential methylation, chemical features, protein–protein interaction, protein metabolite interaction.
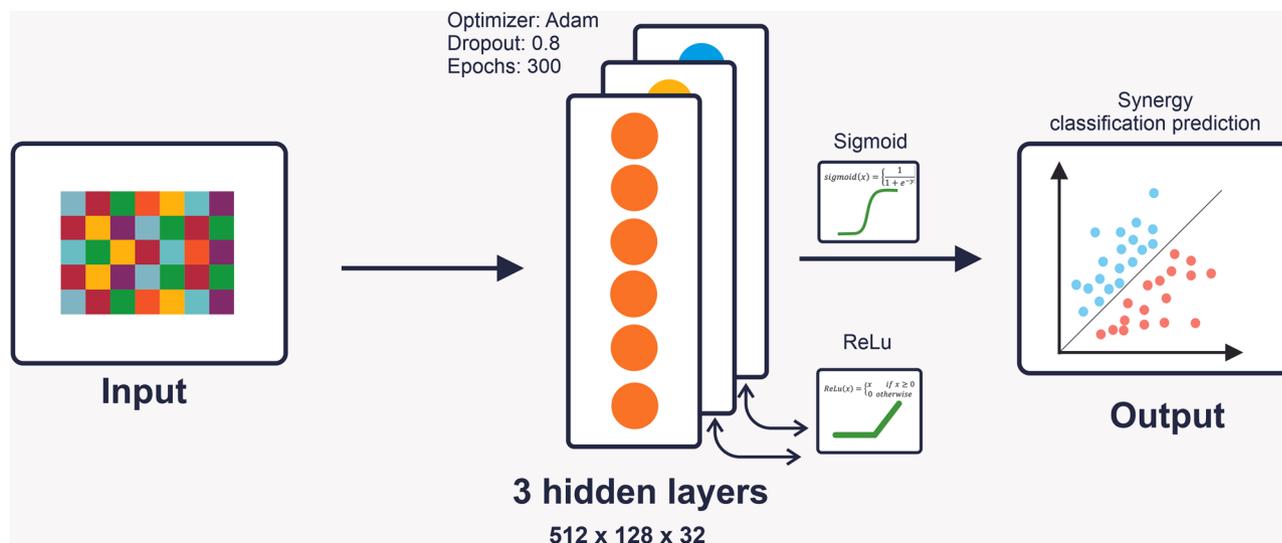
Figure 6. Schematic illustration of the proposed method (SYNDEEP). The model takes the feature vectors as input. The input layer is connected to 3 hidden layers. The number of nodes in each of the three hidden layers was 512, 128, and 32, respectively. Between each of the three hidden layers, there is a ReLu activation function. For the last layer, we adopt a sigmoid activation function. The learning rate was 10^{-2}.

The ReLu activation function is termed Rectified Linear Unit or rectifier, which is one of the common activation functions for deep learning. The ReLu function is used in hidden layers to detect the patterns, and the performance of the model increase by overcoming the gradient vanishing. The activation function of the last layer was the sigmoid function:

$$sigmoid(x) = \left\{ \frac{1}{1 + e^{-y}} \right.$$

To optimize, we used the binary cross-entropy function as a loss function.

$$L = -\frac{1}{N} \sum_{i=1}^{N} y_i \cdot \log(s_i) + \left(1 - y_i\right) \cdot \log(1 - s_i)$$

where y_i is the actual synergy label of each drug pair, s_i is the predicted synergy score of each combination, and N is the number of drug combinations.

Applying other machine learning methods. To evaluate the performance of the predictive model, we compared the deep neural network model with different machine learning algorithms. Shallow neural network (SNN), k-nearest neighbor (KNN), Support vector machines (SVMs), random forest (RF), and gradient boosting classifiers (GBCs) were the algorithms that were recruited in this step. All the methods were examined by different hyperparameters. The performance of considered algorithms is largely dependent on hyperparameters. Therefore, the various hyperparameters of the algorithms were adjusted to the optimal state.

Evaluation criteria of presented models. The tenfold cross-validation and some popular evaluation criteria were utilized in the proposed experiments to evaluate the models' performance. The tenfold cross-validation involves randomly splitting the whole dataset into ten independent subsets equally sized. Each time one fold is used for testing, and nine remaining folds are used for the training, which this procedure is repeated iteratively. This process is performed ten times to ensure each fold is tested at once. To evaluate the performance of the prediction model, widely used evaluation criteria, including accuracy (Accu.), sensitivity (Sen.), specificity (Spec.), precision (Prec.), F-Score (F_{score}), Matthews Correlation Coefficient (MCC.), and Cohen's kappa coefficient (Kappa) were calculated as follow.

$$Accu. = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Sen. = \frac{TP}{TP + FN}$$

10

$$Spec. = \frac{TN}{TN + FP}$$

$$Prec. = \frac{TP}{TP + FP}$$

$$F_{score} = 2 \times \frac{Sen. \times Prec.}{Sen. + Prec.}$$

$$MCC. = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

$$Kappa. = \frac{Accu. - P_c}{1 - P_c}$$

where TP, FP, TN, and FN denote true positive, false positive, true negative, and false negative, respectively. The current study used Cohen's kappa to evaluate the agreement grade between observed accuracy and expected accuracy and the performance of the prediction model. Where P_c is the probability value of agreements expected by chance.

In addition, we performed McNemar's Test to compare SYNDEEP results with proposed state-of-the-art computational models. McNemar's Test is defined as follows:

$$X^2 = \frac{(|B - C| - 1)^2}{B + C}$$

where B is the number of drug combinations that were detected correctly as non-synergy and incorrectly detected as synergy effects, while C is the number of drug combinations that were detected correctly as synergy and incorrectly as non-synergy effects.

Computational equipment.
In this study, the deep neural network model is implemented in python language using Python 3.7 version. To implement the deep learning methods and machine learning algorithms, we used the Keras and Scikit-learn libraries, respectively. The software environment to develop the experiments was Google colab.

Data availability
Dataset: NCI-ALMANAC Data Resource (https://dtp.cancer.gov/ncialmanac). Drug-target Interactions Data: Drug-target interactions (https://go.drugbank.com/releases/latest). Protein–Protein Interactions Data: PPI (https://string-db.org/). Genomic Features Data: gene expression, mutations, and differential methylation (https://cancer.sanger.ac.uk/cosmic). Protein-Metabolite Interaction Data: protein-metabolite interaction (https://www.mhmdb.co.uk/). Scripts: The source code of SYNDEEP is available in (https://github.com/annatorkamannia/SYNDEEP).

References
1. Nagai, H. & Kim, Y. H. Cancer prevention from the perspective of global cancer burden patterns. *J. Thorac. Dis.* **9**, 448–451. https://doi.org/10.21037/jtd.2017.02.75 (2017).
2. Alamzadeh, Z. *et al.* Ultrastructural and optical characteristics of cancer cells treated by a nanotechnology based chemo-photo-thermal therapy method. *J. Photochem. Photobiol. B* **192**, 19–25. https://doi.org/10.1016/j.jphotobiol.2019.01.005 (2019).
3. Piscitello, A. & Edwards, D. K. Estimating the screening-eligible population size, ages 45–74, at average risk to develop colorectal cancer in the United States. *Cancer Prev. Res.* **13**, 443–448. https://doi.org/10.1158/1940-6207.Capr-19-0527 (2020).
4. Wang, X., Zhang, H. & Chen, X. Drug resistance and combating drug resistance in cancer. *Cancer Drug Resist.* **2**, 141–160. https://doi.org/10.20517/cdr.2019.10 (2019).
5. Hassan, M., Watari, H., AbuAlmaaty, A., Ohba, Y. & Sakuragi, N. Apoptosis and molecular targeting therapy in cancer. *Biomed. Res. Int.* **2014**, 150845. https://doi.org/10.1155/2014/150845 (2014).
6. Jardim, D. L., De Melo Gagliato, D., Nikanjam, M., Barkauskas, D. A. & Kurzrock, R. Efficacy and safety of anticancer drug combinations: A meta-analysis of randomized trials with a focus on immunotherapeutics and gene-targeted compounds. *Oncoimmunology* **9**, 1710052. https://doi.org/10.1080/2162402x.2019.1710052 (2020).
7. Falzone, L., Salomone, S. & Libra, M. Evolution of cancer pharmacological treatments at the turn of the third millennium. *Front. Pharmacol.* **9**, 1300. https://doi.org/10.3389/fphar.2018.01300 (2018).
8. Ayoub, N. M. Editorial: Novel combination therapies for the treatment of solid cancers. *Front. Oncol.* **11**, 708943. https://doi.org/10.3389/fonc.2021.708943 (2021).
9. Torkamannia, A., Omidi, Y. & Ferdousi, R. A review of machine learning approaches for drug synergy prediction in cancer. *Brief. Bioinform.* https://doi.org/10.1093/bib/bbac075 (2022).
10. Foucquier, J. & Guedj, M. Analysis of drug combinations: Current methodological landscape. *Pharmacol. Res. Perspect.* **3**, e00149. https://doi.org/10.1002/prp2.149 (2015).
11. Jaaks, P. *et al.* Effective drug combinations in breast, colon and pancreatic cancer cells. *Nature* **603**, 166–173. https://doi.org/10.1038/s41586-022-04437-2 (2022).

12. Jamali, A. A. *et al.* DrugMiner: Comparative analysis of machine learning algorithms for prediction of potential druggable proteins. *Drug Discov. Today* **21**, 718–724. https://doi.org/10.1016/j.drudis.2016.01.007 (2016).
13. Brogi, S., Ramalho, T. C., Kuca, K., Medina-Franco, J. L. & Valko, M. Editorial: In silico methods for drug design and discovery. *Front. Chem.* **8**, 612. https://doi.org/10.3389/fchem.2020.00612 (2020).
14. Celebi, R., Bear Don't Walk, O., Movva, R., Alpsoy, S. & Dumontier, M. In-silico prediction of synergistic anti-cancer drug combinations using multi-omics data. *Sci. Rep.* **9**, 8949. https://doi.org/10.1038/s41598-019-45236-6 (2019).
15. Ferdousi, R., Safdari, R. & Omidi, Y. Computational prediction of drug-drug interactions based on drugs functional similarities. *J. Biomed. Inform.* **70**, 54–64. https://doi.org/10.1016/j.jbi.2017.04.021 (2017).
16. Jiang, P. *et al.* Deep graph embedding for prioritizing synergistic anticancer drug combinations. *Comput. Struct. Biotechnol. J.* **18**, 427–438. https://doi.org/10.1016/j.csbj.2020.02.006 (2020).
17. Zhang, T., Zhang, L., Payne, P. R. O. & Li, F. Synergistic drug combination prediction by integrating multiomics data in deep learning models. *Methods Mol. Biol.* **2194**, 223–238. https://doi.org/10.1007/978-1-0716-0849-4_12 (2021).
18. Preuer, K. *et al.* DeepSynergy: Predicting anti-cancer drug synergy with deep learning. *Bioinformatics* **34**, 1538–1546. https://doi.org/10.1093/bioinformatics/btx806 (2018).
19. Yang, M. *et al.* Stratification and prediction of drug synergy based on target functional similarity. *NPJ Syst. Biol. Appl.* **6**, 16. https://doi.org/10.1038/s41540-020-0136-x (2020).
20. Iorio, F. *et al.* A landscape of pharmacogenomic interactions in cancer. *Cell* **166**, 740–754. https://doi.org/10.1016/j.cell.2016.06.017 (2016).
21. Sung, H. *et al.* Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **71**, 209–249. https://doi.org/10.3322/caac.21660 (2021).
22. Karami Fath, M. *et al.* Anti-cancer peptide-based therapeutic strategies in solid tumors. *Cell. Mol. Biol. Lett.* **27**, 33. https://doi.org/10.1186/s11658-022-00332-w (2022).
23. Housman, G. *et al.* Drug resistance in cancer: An overview. *Cancers (Basel)* **6**, 1769–1792. https://doi.org/10.3390/cancers6031769 (2014).
24. Fan, K., Cheng, L. & Li, L. Artificial intelligence and machine learning methods in predicting anti-cancer drug combination effects. *Brief. Bioinform.* https://doi.org/10.1093/bib/bbab271 (2021).
25. Huang, L. *et al.* DrugComboRanker: Drug combination discovery based on target network analysis. *Bioinformatics* **30**, i228-236. https://doi.org/10.1093/bioinformatics/btu278 (2014).
26. Bansal, M. *et al.* A community computational challenge to predict the activity of pairs of compounds. *Nat. Biotechnol.* **32**, 1213–1222. https://doi.org/10.1038/nbt.3052 (2014).
27. Ghosh, A. & Saha, S. Survey of drug resistance associated gene mutations in Mycobacterium tuberculosis, ESKAPE and other bacterial species. *Sci. Rep.* **10**, 8957. https://doi.org/10.1038/s41598-020-65766-8 (2020).
28. Shen, Y. & Yan, Z. Systematic prediction of drug resistance caused by transporter genes in cancer cells. *Sci. Rep.* **11**, 7400. https://doi.org/10.1038/s41598-021-86921-9 (2021).
29. Kang, H. C. *et al.* Identification of genes with differential expression in acquired drug-resistant gastric cancer cells using high-density oligonucleotide microarrays. *Clin. Cancer Res.* **10**, 272–284. https://doi.org/10.1158/1078-0432.ccr-1025-3 (2004).
30. Sarker, I. H. Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions. *SN Comput. Sci.* **2**, 420. https://doi.org/10.1007/s42979-021-00815-1 (2021).
31. Pacal, I., Karaboga, D., Basturk, A., Akay, B. & Nalbantoglu, U. A comprehensive review of deep learning in colon cancer. *Comput. Biol. Med.* **126**, 104003. https://doi.org/10.1016/j.compbiomed.2020.104003 (2020).
32. Huang, L. *et al.* Driver network as a biomarker: Systematic integration and network modeling of multi-omics data to derive driver signaling pathways for drug combination prediction. *Bioinformatics* **35**, 3709–3717. https://doi.org/10.1093/bioinformatics/btz109 (2019).
33. Sidorov, P., Naulaerts, S., Ariey-Bonnet, J., Pasquier, E. & Ballester, P. J. Predicting synergism of cancer drug combinations using NCI-ALMANAC data. *Front. Chem.* **7**, 509. https://doi.org/10.3389/fchem.2019.00509 (2019).
34. Li, X. *et al.* Prediction of synergistic anti-cancer drug combinations based on drug target network and drug induced gene expression profiles. *Artif. Intell. Med.* **83**, 35–43. https://doi.org/10.1016/j.artmed.2017.05.008 (2017).
35. Xia, F. *et al.* Predicting tumor cell line response to drug pairs with deep learning. *BMC Bioinform.* **19**, 486. https://doi.org/10.1186/s12859-018-2509-3 (2018).
36. Zhang, H., Feng, J., Zeng, A., Payne, P. & Li, F. Predicting tumor cell response to synergistic drug combinations using a novel simplified deep learning model. *AMIA Annu. Symp. Proc.* **2020**, 1364–1372 (2020).
37. Sun, Y. *et al.* Combining genomic and network characteristics for extended capability in predicting synergistic drugs for cancer. *Nat. Commun.* **6**, 1–10 (2015).
38. Li, H., Li, T., Quang, D. & Guan, Y. Network propagation predicts drug synergy in cancers. *Can. Res.* **78**, 5446–5457 (2018).
39. Liu, Q. & Xie, L. TranSynergy: Mechanism-driven interpretable deep neural network for the synergistic prediction and pathway deconvolution of drug combinations. *PLoS Comput. Biol.* **17**, e1008653 (2021).
40. Preto, A. J., Matos-Filipe, P., Mourão, J. & Moreira, I. S. SynPred: prediction of drug combination effects in cancer using full-agreement synergy metrics and deep learning. *GigaScience* **11**, giac087 (2022).
41. Kuenzi, B. M. *et al.* Predicting drug response and synergy using a deep learning model of human cancer cells. *Cancer Cell* **38**, 672-684.e676 (2020).
42. Jamal, S., Goyal, S., Shanker, A. & Grover, A. Predicting neurological Adverse Drug Reactions based on biological, chemical and phenotypic properties of drugs using machine learning models. *Sci. Rep.* **7**, 872. https://doi.org/10.1038/s41598-017-00908-z (2017).
43. Ye, Z., Chen, F., Zeng, J., Gao, J. & Zhang, M. Q. ScaffComb: A phenotype-based framework for drug combination virtual screening in large-scale chemical datasets. *Adv. Sci.* **8**, e2102092. https://doi.org/10.1002/advs.202102092 (2021).
44. Kraus, M., Feuerriegel, S. & Oztekin, A. Deep learning in business analytics and operations research: Models, applications and managerial implications. *Eur. J. Oper. Res.* **281**, 628–641 (2020).
45. Min, S., Lee, B. & Yoon, S. Deep learning in bioinformatics. *Brief. Bioinform.* **18**, 851–869 (2017).
46. Holbeck, S. L. *et al.* The National Cancer Institute ALMANAC: A comprehensive screening resource for the detection of anticancer drug pairs with enhanced therapeutic activity. *Can. Res.* **77**, 3564–3576 (2017).
47. Wishart, D. S. *et al.* DrugBank: A comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res.* **34**, D668–D672 (2006).
48. Mering, C. V. *et al.* STRING: A database of predicted functional associations between proteins. *Nucleic Acids Res.* **31**, 258–261 (2003).
49. Forbes, S. A. *et al.* COSMIC: Exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Res.* **43**, D805–D811 (2015).
50. Wishart, D. S. *et al.* HMDB: The human metabolome database. *Nucleic Acids Res.* **35**, D521–D526 (2007).
51. Willett, P. *Similarity-Based Approaches to Virtual Screening* (Portland Press Ltd., 2003).

Acknowledgements

Author contributions

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to R.F.

Reprints and permissions information is available at www.nature.com/reprints.

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